

**UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE CIÊNCIA DA INFORMAÇÃO**

Italo Rodrigues Castro

**PROPOSTA DE MODELO DE COMPATIBILIDADE DE LINGUAGEM ENTRE
SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO: A CLASSIFICAÇÃO
DECIMAL UNIVERSAL E O AGROTERMOS**

Belo Horizonte
2021

Italo Rodrigues Castro

PROPOSTA DE MODELO DE COMPATIBILIDADE ENTRE SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO: A CLASSIFICAÇÃO DECIMAL UNIVERSAL E O AGROTERMOS

Tese apresentada ao Programa de Pós-Graduação em Gestão e Organização do Conhecimento da Escola de Ciência da Informação da Universidade Federal de Minas Gerais, como requisito parcial para a obtenção do título de Doutor em Gestão e Organização do Conhecimento.

Área de concentração: Ciência da Informação

Linha de Pesquisa: Arquitetura e Organização do Conhecimento

Orientadora: Prof^a. Dr^a. Gercina Ângela de Lima

Coorientador: Dr. Ivo Pierozzi Junior

Belo Horizonte
2021

C355p

Castro, Italo Rodrigues.

Proposta de modelo de compatibilidade entre sistemas de organização do conhecimento [recurso eletrônico] : a classificação decimal universal e o agrotermos / Italo Rodrigues Castro. - 2021.

1 recurso online (312 f. : il., color.) : pdf.

Orientadora: Gercina Ângela de Lima

Coorientador: Ivo Pierozzi Junior

Tese (Doutorado) – Universidade Federal de Minas Gerais, Escola de Ciência da Informação.

Referências: f. 251-277.

Anexos: f. 278-312.

Exigências do sistema: Adobe Acrobat Reader.

1. Ciência da informação – Teses. 2. Organização da informação – Teses. 3. Sistemas de recuperação da informação – Teses. 4. Tesouros – Teses. 5. Classificação decimal universal – Teses. I. Título. II. Lima, Gercina Ângela Borém de Oliveira. III. Pierozzi Junior, Ivo. IV. Universidade Federal de Minas Gerais, Escola de Ciência da Informação.

CDU: 025.4

Ficha catalográfica: Rosimeire Silva Campos de Lima CRB:6/3145

Biblioteca Profª Etelvina Lima, Escola de Ciência da Informação da UFMG.



UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE CIÊNCIA DA INFORMAÇÃO - ECI
PROGRAMA DE PÓS-GRADUAÇÃO EM GESTÃO E ORGANIZAÇÃO DO CONHECIMENTO - PPG-GOC

FOLHA DE APROVAÇÃO

Proposta de Modelo de compatibilidade entre sistemas de organização do conhecimento: a Classificação Decimal Universal e o AgroTermos

ITALO RODRIGUES CASTRO

Tese submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em GESTÃO E ORGANIZAÇÃO DO CONHECIMENTO, como requisito para obtenção do grau de Doutor em GESTÃO E ORGANIZAÇÃO DO CONHECIMENTO, área de concentração CIÊNCIA DA INFORMAÇÃO, linha de pesquisa Arquitetura e Organização do Conhecimento.

Aprovada em 27 de agosto de 2021, por videoconferência, pela banca constituída pelos membros:

Prof(a). Gercina Ângela de Lima (Orientadora)
ECI/UFMG

Dr(a). Ivo Pierozzi Júnior (Coorientador)
Embrapa

Prof(a). Benildes Coura Moreira dos Santos Maculan
ECI/UFMG

Prof(a). Cíntia de Azevedo Lourenço
ECI/UFMG

Prof(a). Lucinéia Souza Maia
UFOP

Prof(a). Maria Luiza de Almeida Campos
UFF e UFBA

Prof(a). Marcel Ferrante Silva
UFG

Belo Horizonte, 27 de agosto de 2021.



Documento assinado eletronicamente por **Gercina Angela de Lima, Professora do Magistério Superior**, em 02/09/2021, às 14:07, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Cintia de Azevedo Lourenco, Professora do Magistério Superior**, em 02/09/2021, às 16:30, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Benildes Coura Moreira dos Santos Maculan, Professora do Magistério Superior**, em 02/09/2021, às 16:36, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Maria Luiza de Almeida Campos, Usuário Externo**, em 02/09/2021, às 17:29, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Ivo Pierozzi Junior, Usuário Externo**, em 06/09/2021, às 14:38, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Marcel Ferrante Silva, Usuário Externo**, em 08/09/2021, às 14:19, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Lucinéia Souza Maia, Usuário Externo**, em 21/09/2021, às 15:12, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0927871** e o código CRC **26F292FD**.



UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE CIÊNCIA DA INFORMAÇÃO - ECI
PROGRAMA DE PÓS-GRADUAÇÃO EM GESTÃO E ORGANIZAÇÃO DO CONHECIMENTO - PPG-GOC

ATA DA DEFESA DE TESE DO ALUNO

ITALO RODRIGUES CASTRO

Realizou-se, no dia 27 de agosto de 2021, às 09:00 horas, por videoconferência, da Universidade Federal de Minas Gerais, a defesa de tese, intitulada *Proposta de Modelo de compatibilidade entre sistemas de organização do conhecimento: a Classificação Decimal Universal e o AgroTermos*, apresentada por ITALO RODRIGUES CASTRO, número de registro 2016719901, graduado no curso de CIÊNCIA DA COMPUTAÇÃO, como requisito parcial para a obtenção do grau de Doutor em GESTÃO E ORGANIZAÇÃO DO CONHECIMENTO, à seguinte Comissão Examinadora: Prof(a). Gercina Ângela de Lima - ECI/UFMG (Orientadora), Dr(a). Ivo Pierozzi Júnior - Embrapa (Coorientador), Prof(a). Benildes Coura Moreira dos Santos Maculan - ECI/UFMG, Prof(a). Cíntia de Azevedo Lourenço - ECI/UFMG, Prof(a). Lucinéia Souza Maia - UFOP, Prof(a). Maria Luiza de Almeida Campos - UFF e UFBA, Prof(a). Marcel Ferrante Silva - UFG.

A Comissão considerou a tese:

Aprovada

Reprovada

Finalizados os trabalhos, lavrei a presente ata que, lida e aprovada, vai assinada por mim e pelos membros da Comissão.

Belo Horizonte, 27 de agosto de 2021.

Assinatura dos membros da banca examinadora:



Documento assinado eletronicamente por **Gercina Angela de Lima, Professora do Magistério Superior**, em 02/09/2021, às 14:07, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Cintia de Azevedo Lourenco, Professora do Magistério Superior**, em 02/09/2021, às 16:30, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Benildes Coura Moreira dos Santos Maculan, Professora do Magistério Superior**, em 02/09/2021, às 16:36, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Maria Luiza de Almeida Campos, Usuário Externo**, em 02/09/2021, às 17:29, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Ivo Pierozzi Junior, Usuário Externo**, em 06/09/2021, às 14:36, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Marcel Ferrante Silva, Usuário Externo**, em 08/09/2021, às 14:19, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Lucinéia Souza Maia, Usuário Externo**, em 21/09/2021, às 15:12, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0927835** e o código CRC **C7B6708A**.

À minha família por ser meu alicerce
e meus amigos pelo apoio.

AGRADECIMENTOS

À Deus por sempre estar ao meu lado, me guiando para que eu consiga vencer várias etapas da minha vida, principalmente com saúde, me dando força e encorajamento para enfrentar cada obstáculo, fazendo com que eu chegasse até o final desta etapa.

À minha orientadora Prof^a Dr^a Gercina Lima, que além de amiga, é para mim um exemplo a ser seguido. Nunca vou esquecer a excelente formação que me deu como pesquisador e como pessoa durante o doutoramento. Agradeço de coração todas as nossas conversas, orientações de várias horas seguidas, sempre me colocando no caminho certo, além de sua postura profissional admirável! Sou eternamente grato pela paciência, pelo cuidado, cada palavra, conselhos, gestos de carinho e motivação, o que tornou todo o processo mais leve, sempre transparecendo seu lado humano independente da situação. Vou levar isso comigo, e não só para a carreira acadêmica, mas também para a vida. Muito obrigado!

Ao meu coorientador Dr. Ivo Pierozzi Júnior, que me ajudou muito por compartilhar seus conhecimentos durante a minha pesquisa, sempre disposto a esclarecer minhas dúvidas, com várias sugestões de melhorias. Muito obrigado pela amizade e parceria, foi um grande aprendizado!

À minha família, pelo amor incondicional que sempre me deu, que mesmo de longe, sempre me apoiaram e me deram a base essencial para que eu conseguisse conquistar várias etapas da minha vida, e agora mais esta vitória que compartilhamos e comemoramos juntos!

Aos meus amigos, pela força e pela compreensão nos meus momentos de ausência, e mesmo assim, sempre ficaram do meu lado na torcida e me apoiando nos momentos que mais precisei. Muito obrigado a todos!

À Escola de Ciência da Informação e ao programa de Pós Graduação em Gestão e Organização do Conhecimento da Universidade Federal de Minas Gerais (UFMG), que me acolheu, me dando a oportunidade de estudar e desenvolver minha pesquisa de doutorado. Agradeço a todos os professores pela dedicação, em especial às professoras Dr^a Marlene Oliveira, Dr^a Dalgiza Andrade, Dr^a Renata Baracho, Dr^a Beatriz Cendon, Dr^a Célia da Consolação Dias, Dr^a Elisângela Cristina Aganette e aos professores Dr. Marcello Peixoto Bax e Dr. Maurício Barcellos Almeida com quem eu tive o prazer de conhecer e alguns fazer disciplinas na pós-graduação. Muito obrigado a todos por me acompanhar durante o curso, pelo incentivo, pelo conhecimento compartilhado e experiências que adquiri.

À coordenação e às secretárias pela presteza e gentileza para ajudar na resolução de qualquer tipo de problema ou situação em relação ao curso.

Aos amigos da ECI/UFMG, que trabalham ou que fizeram aulas comigo, espero não perder o contato com vocês, nunca esquecerei nossas conversas, trabalhos, muitos estudos e discussões, o que fizeram muita diferença no meu aprendizado, tanto durante as disciplinas, palestras, seminários, quanto na parte da escrita. Desejo muito sucesso a cada um de vocês!

À Empresa Brasileira de Pesquisa Agropecuária (Embrapa) pela parceria formada com o Grupo de Pesquisa MHTX, no qual faço parte, que me forneceu os dados necessários para aplicar em minha pesquisa.

Aos amigos do Grupo de Pesquisa MHTX, em especial: Webert Araújo, Graciane Bruzinga, Patrícia Lopes, Elaine Diamantino, Helder Firmino, Tereza Cristina, Mayara Silva, Danielle Rioga, Rosy Oliveira, Cleide Vieira, Rainer Finelli, Celsiane Araújo, Heber Tormentino, Reuber Fonseca e Júnio Lopes, que sempre estiveram dispostos a ajudar tanto no lado acadêmico, quanto no pessoal.

À banca de qualificação e de tese, pela disponibilidade e pelas valiosas contribuições que enaltecem os resultados do meu trabalho final.

À Profª Drª Benildes Maculan, que enxergou meu potencial e me apoiou desde o início do doutorado, sempre com palavras amigas e motivadoras, me ajudando nas revisões finais tanto na defesa de qualificação quanto na de tese, com uma leitura bem detalhada, enxergando os pontos a serem melhorados, o que resultou em melhorias bastantes significativas, e pelo grande aprendizado que tive a partir de sua tese e com os resultados que utilizei.

À Profª Drª Lucinéia Maia, que me ajudou com muita presteza e dedicação com seu conhecimento nas revisões finais, que engrandeceu minha pesquisa na parte textual e principalmente a partir dos resultados de sua pesquisa de doutorado que usei em minha pesquisa.

À Profª Drª Cíntia Azevedo Lourenço pela ajuda tanto na defesa de qualificação, quanto na de tese, me incentivando a incrementar os resultados da minha pesquisa, agregando mais valor, o que enriqueceu ainda mais os resultados.

À Profª. Drª Maria Luiza Campos, o Prof. Dr. Marcel Ferrante e a Profª Drª Cristiane Mendes Netto que contribuíram com minha pesquisa com seus vastos conhecimentos interdisciplinares, tanto em relação a área da Ciência da Informação, quanto da Ciência da Computação.

Por fim, à todos que de alguma forma colaboraram para a realização da minha pesquisa, não somente em relação ao desenvolvimento, mas também com palavras e dicas que me motivaram a querer fazer sempre o melhor. Os nomes que não mencionei aqui, não são menos importantes que os demais. Meus sinceros agradecimentos a todos!

“O modo como você reúne, administra e usa a informação determina se vencerá ou perderá”.
(Bill Gates)

RESUMO

Os sistemas de organização do conhecimento são estruturados com o propósito de facilitar os processos de organização e recuperação da informação em ambientes físicos e digitais. A Classificação Decimal Universal adentra-se nesse contexto, pois apresenta uma estrutura semifacetada, com um alto grau de flexibilidade e especificidade de assuntos pertencentes a qualquer domínio do conhecimento, tendo em vista ampliar o meio de acesso ao recurso informacional. O AgroTermos é um tesouro utilizado pela Empresa Brasileira de Pesquisa Agropecuária, que possui um conjunto de termos que se referem aos conceitos e suas relações semânticas para a representação de assuntos do domínio da agricultura. O objetivo desta pesquisa é propor um modelo de compatibilidade de linguagem entre a Classificação Decimal Universal e o tesouro AgroTermos com o intuito de melhorar a recuperação da informação, em um domínio específico, entre diferentes sistemas de informação. Este estudo tem importância por valorizar um sistema que contribui para a organização de bibliotecas e que pode ser estendido para uso no ambiente da web semântica, sobretudo com o auxílio de outros sistemas de organização do conhecimento. No contexto digital, precisam ser considerados processos específicos, tais como: compatibilidade de linguagens, interoperabilidade entre vocabulários controlados, bem como em sistemas de informação em formatos do tipo Simple Knowledge Organization System, o que torna a temática deste estudo ainda mais relevante. Como metodologia, o estudo é exploratório, descritivo, aplicado e de abordagem qualitativa e quantitativa, e os procedimentos técnicos são: revisão sistemática da literatura e estudo de caso, tendo como referência empírica ou experimental o tesouro AgroTermos no contexto da Empresa Brasileira de Pesquisa Agropecuária. Para conduzir o estudo de caso, na primeira etapa, procedeu-se à compatibilidade entre os termos das classes da Classificação Decimal Universal e os termos que representam os conceitos relacionados com a temática “Intensificação Agropecuária” do domínio da agricultura, processo em que os termos foram extraídos do tesouro AgroTermos, gerando um modelo conceitual; na segunda etapa, esse modelo foi representado no padrão Simple Knowledge Organization System, por meio do software *TemaTres*, inserindo as classes e os termos da Classificação Decimal Universal e os descritores do tesouro AgroTermos com seus respectivos relacionamentos semânticos. Pelos resultados, percebeu-se que as linguagens de ambos os vocabulários são correspondentes e isso representa o primeiro passo para que haja a interoperabilidade entre os sistemas, no contexto da web semântica. Por fim, concluiu-se que a Classificação Decimal Universal acoplada a um tesouro, como o AgroTermos, contribui para melhorias na recuperação da informação, visto que esses sistemas são compatíveis, estabelecendo relações semânticas capazes de direcionar as buscas, com flexibilidade e mantendo a lógica entre as relações. As melhorias na recuperação também se devem ao fato de que ambos os sistemas garantem a interoperabilidade, pois mantêm uma padronização de formato suficiente para ser interpretada por sistemas de informação, contribuindo para a recuperação, por meio do intercâmbio de informações e recursos no contexto da web semântica.

Palavras-chave: Sistemas de Organização do Conhecimento. Classificação Decimal Universal. Tesouros. Compatibilidade de Linguagem. Recuperação da Informação.

ABSTRACT

Knowledge organization systems are structured with the purpose to facilitate information organization and retrieval processes in physical and digital environments. The Universal Decimal Classification is inserted into that context, because it presents a semi-faceted structure, with a high degree of flexibility and specificity of subjects belonging to any domain of knowledge, expanding the means of access to the informational resource. AgroTermos is a thesaurus used by the Brazilian Agricultural Research Corporation, which has a set of terms that refer to concepts and their semantic relationships for the representation of subjects in the domain of agriculture. The objective of this research is to propose a language compatibility model between the Universal Decimal Classification and the AgroTermos thesaurus with intention to improve information retrieval, in a specific domain, between different information systems. This study is important to value a system that greatly contributes to the organization of libraries and that can be extended for use in the semantic web environment, above all with the help of other knowledge organization systems. In the digital context, specific processes need to be considered, such as: language compatibility, interoperability between controlled vocabularies, as well as information systems in Simple Knowledge Organization System-type formats, which makes the subject of this study even more relevant. As methodology the study is exploratory, descriptive, applied with a qualitative and quantitative approach, and the technical procedures are: systematic literature review and case study, having as empirical or experimental reference the AgroTermos thesaurus in the context of the Brazilian Agricultural Research Corporation. To conduct the case study, in the first stage, proceeded the compatibility between the terms of the Universal Decimal Classification classes and the terms that represent the concepts related to the topical "Agricultural Intensification" of domain of agriculture, process in which the terms were extracted of thesaurus AgroTermos, generating a conceptual model; in the second stage, this model was represented in the Simple Knowledge Organization System standard, through TemaTres software, inserting the classes and the terms of the Universal Decimal Classification and the descriptors of the AgroTermos thesaurus with their respective semantic relationships. For the results, it is perceived that the languages of both vocabularies are matching and this represents the first step towards the interoperability between the systems, in the context of the semantic web. Finally, it is concluded that the Universal Decimal Classification coupled with a thesaurus, such as the AgroTermos, contributes to improvements in the information retrieval, since the these systems are compatible, establishing semantic relationships capable of directing the searches, with flexibility and keeping the logic between the relationships. Improvements in retrieval are also due to the fact that both systems ensure the interoperability, because they maintain a standardization of a format sufficient to be interpreting for information systems, contributing to retrieval through the exchange of information and resources in the context of the semantic web.

Keywords: Knowledge Organization Systems. Universal Decimal Classification. Thesaurus. Language Compatibility. Information Retrieval.

LISTA DE FIGURAS

FIGURA 1 - Etapas do processo de Revisão Sistemática da Literatura	31
FIGURA 2 - Principais estruturas e funções dos SOC tradicionais.....	65
FIGURA 3 - Classes vinculadas ao universo da gricultura agrupadas por semelhanças e diferenças	75
FIGURA 4 - Macroestrutura da Classificação Decimal Universal e síntese das notações....	91
FIGURA 5 - Sintaxe permitida na Classificação Decimal Universal para combinação de assuntos	93
FIGURA 6 - Exemplo de notação complexa como um assunto simples	97
FIGURA 7 - Exemplo de notação complexa na recuperação	98
FIGURA 8 - Triângulo semântico do conceito	102
FIGURA 9 - Métodos para a extração de termos	105
FIGURA 10 - Tipos de mapeamentos previstos em tesouros.....	107
FIGURA 11 - Níveis de um tesouro baseado em conceitos	108
FIGURA 12 - Níveis conceituais e ajuste de níveis lexicais multilíngues.....	109
FIGURA 13 - Interface principal de navegação do <i>TemaTres</i>	115
FIGURA 14 - Pesquisa simples e avançada no <i>TemaTres</i>	117
FIGURA 15 - Esquema ilustrativo dos tipos de relações semânticas	121
FIGURA 16 - Relações semânticas (gênero - espécie) no contexto da CDU	126
FIGURA 17 - Relações partitivas (todo-parte) no âmbito dos sistemas de classificação....	126
FIGURA 18 - Tesouro baseado em termos e tesouro baseado em conceitos	128
FIGURA 19 - Exemplo de classe de um recurso do tipo conceitual	150
FIGURA 20 - Exemplo de recursos com relações hierárquicas no domínio das espécies animais	151
FIGURA 21 - Exemplos de hierarquias transitivas e intransitivas com níveis de especificidade	153
FIGURA 22 - Exemplo de recursos com relações associativas.....	154
FIGURA 23 - Esquema de representação de um conceito no vocabulário SKOS	155
FIGURA 24 - Delineamento do universo, amostra e recorte temático	160
FIGURA 25 - Fluxograma da pesquisa - Procedimentos metodológicos	162
FIGURA 26 - Estrutura classificatória da CDU a ser compatibilizada	167
FIGURA 27 - Categorias dos termos utilizados no AgroTermos	169
FIGURA 28 - Estrutura gerada a partir da compatibilidade entre as classes da CDU, AgroTermos e relações semânticas	203
FIGURA 29 - Tela do <i>TemaTres</i> - Geração das relações e de links.....	227
FIGURA 30 - Código e descrição dos termos exportados no formato SKOS	233
FIGURA 31 - Termo “Intensificação Agropecuária” e suas relações no <i>Tematres</i>	234
FIGURA 32 - URI do termo “Intensificação” no tesouro AgroTermos e Agrovoc	235
FIGURA 33 - Termo “Intensificação Agropecuária” do AgroTermos no SKOS	236
FIGURA 34 - Termo “Mudança de cobertura da terra” do AgroTermos e suas relações no <i>Tematres</i>	238
FIGURA 35 - Termo “Mudança de cobertura da terra” do AgroTermos no SKOS	239
FIGURA 36 - Conversão do termo “Pousio” do AgroTermos e suas relações no <i>Tematres</i>	241
FIGURA 37 - Termo “Pousio” do AgroTermos no SKOS.....	242

LISTA DE QUADROS

QUADRO 1 - Expressões utilizadas para busca na literatura.....	34
QUADRO 2 - Principais dados levantados no processo de análise.....	38
QUADRO 3 - Comparação entre os sistemas de classificação bibliográfica	76
QUADRO 4 - Divisões principais da Classificação Decimal Universal	87
QUADRO 5 - Detalhamento dos símbolos das Tabelas Auxiliares Comuns e Especiais	92
QUADRO 6 - Exemplos de flexibilidade em tabelas auxiliares comuns	97
QUADRO 7 - Principais abreviaturas utilizadas no AgroTermos	110
QUADRO 8 - Exemplo de hierarquia estabelecida no AgroTermos	111
QUADRO 9 - Principais definições para relações hierárquicas, equivalentes e associativas	122
QUADRO 10 - Relações de equivalência no AgroTermos	129
QUADRO 11 - Relações de hierarquia no AgroTermos	130
QUADRO 12- Relações associativas no AgroTermos.....	131
QUADRO 13 - Matriz de compatibilidade conceitual de Dahlberg.....	144
QUADRO 14 - Estudos que sustentaram a constituição do modelo baseado em dois instrumentos de organização do conhecimento	165
QUADRO 15 - Insumos para a análise de dados da compatibilização entre o vocabulário da CDU e AgroTermos.....	166
QUADRO 16 - Compatibilidade da classe 061::63 com o termo Embrapa	171
QUADRO 17 - Compatibilidade da classe 631.15 com os termos Intensificação agropecuária e Manejo da cultura.....	172
QUADRO 18 - Compatibilidade da classe 631.15:314.8.061 com o termo Densidade demográfica	174
QUADRO 19 - Compatibilidade da classe 631.15:314.8.062 com o termo Pressão populacional.....	175
QUADRO 20 - Compatibilidade da classe 631.15:314.82 com o termo Crescimento populacional.....	176
QUADRO 21 - Compatibilidade da classe 631.15:519.246.8 com o termo Análise de Séries Temporais.....	177
QUADRO 22 - Compatibilidade da classe 631.115 com os termos Posse da terra e Escassez da terra	178
QUADRO 23 - Compatibilidade da classe 631.151.1 com o termo Agricultura extensiva...	179
QUADRO 24 - Compatibilidade da classe 631.151.2 com o termo Agricultura intensiva...	180
QUADRO 25 - Compatibilidade da classe 631.153.7 com o termo Mudança Agrícola	182
QUADRO 26 - Compatibilidade da classe 631.173 com os termos Meio ambiente e Ambiente físico	183
QUADRO 27 - Compatibilidade da classe 631.23 com o termo Cultura	184
QUADRO 28 - Compatibilidade da classe 631.4 com o termo Solo	186
QUADRO 29 - Compatibilidade da classe 631.544.7 com os termos Cobertura da terra e Mudança de cobertura da terra	187
QUADRO 30 - Compatibilidade da classe 631.547 com o termo Ciclo da cultura	188
QUADRO 31 - Compatibilidade da classe 631.57 com o termo Produto agropecuário	189
QUADRO 32 - Compatibilidade da classe 631.58 com o termo Cultura anual	190
QUADRO 33 - Compatibilidade da classe 631.15:502 com os termos Meio ambiente e Ambiente físico	191
QUADRO 34 - Compatibilidade da classe 631.581 com os termos Pousio e Período de pousio	192
QUADRO 35 - Compatibilidade da classe 631.86 com o termo Biomassa.....	193
QUADRO 36 - Compatibilidade da classe 631.874 com o termo Adubo verde	195
QUADRO 37 - Compatibilidade da classe 632.95 com o termo Pesticida	196
QUADRO 38 - Compatibilidade da classe 636 com o termo Pecuária	197
QUADRO 39 - Estrutura textual para descrição e análise dos resultados	205

QUADRO 40 - Quantidade de classes e subclasses da CDU relacionadas à temática do domínio da agricultura.....	207
QUADRO 41 - Alguns assuntos expressos nas classes da CDU - recorte temático	207
QUADRO 42 - Exemplos de conceitos pertencentes ao AgroTermos (Intensificação Agropecuária)	211
QUADRO 43 - Classes da CDU e termos do AgroTermos por categorias	216
QUADRO 44 - Tipos de relações semânticas por classe da CDU e termos do AgroTermos	219
QUADRO 45 - Detalhamento dos elementos do modelo proposto.....	228

LISTA DE GRÁFICOS

GRÁFICO 1 - Total de publicações recuperadas na primeira etapa.....	35
GRÁFICO 2 - Total de publicações recuperadas na segunda etapa.....	36
GRÁFICO 3 - Total de publicações recuperadas na terceira etapa.....	36
GRÁFICO 4 - Número de publicações dos autores mais produtivos.....	42
GRÁFICO 5 - Instituições dos autores com maior produção científica.....	43
GRÁFICO 6 - Anos com maior número de publicações.....	44
GRÁFICO 7 - Veículos de publicação.....	45
GRÁFICO 8 - Número de ocorrências das principais palavras-chave.....	46
GRÁFICO 9 - Principais domínios da seleção bibliográfica.....	46
GRÁFICO 10 - Categorias dos trabalhos sobre a temática.....	48
GRÁFICO 11 - Quantidade de classes da CDU utilizadas no modelo e quantidade de classes utilizadas no recorte.....	209
GRÁFICO 12 - Categorias de classes da CDU utilizadas no modelo.....	210
GRÁFICO 13 - Ocorrências entre as classes e os termos da CDU e AgroTermos.....	214
GRÁFICO 14 - Classes e termos com mais de duas ocorrências conceituais.....	216
GRÁFICO 15 - Quantidade de classes da CDU e termos do AgroTermos por categorias.....	218
GRÁFICO 16 - Quantidade de relações semânticas por tipo entre as classes da CDU e os termos do AgroTermos.....	222

LISTA DE ABREVIATURAS

AA	-	<i>Art & Architecture Thesaurus</i>
AGROVOC	-	<i>Agricultural Information Management Standards</i>
BC2	-	<i>Bliss Bibliographic Classification</i>
BCI	-	Biblioteconomia e Ciência da Informação
BT	-	<i>Broader Term</i>
BTG	-	<i>Broader Term Generic</i>
BTP	-	<i>Broader Term Partitive</i>
CAB	-	<i>Cab Thesaurus</i>
CC	-	<i>Cólon Classification</i>
CDD	-	Classificação Decimal de Dewey
CDU	-	Classificação Decimal Universal
CRG	-	<i>Classification Research Group</i>
EIA	-	Embrapa Informática Agropecuária
EMBRAPA	-	Empresa Brasileira de Pesquisa Agropecuária
INTAGRO	-	Intensificação Agropecuária em Polos de Produção de Soja e Cana de Açúcar: Territorialidade, Sustentabilidade e Competitividade
ISKO	-	<i>International Society for Knowledge Organization</i>
ISO	-	<i>International Organization for Standardization</i>
ISTA	-	<i>Information Science Technology Abstracts</i>
KOS	-	<i>Knowledge Organization System</i>
LCC	-	<i>Library of Congress Classification</i>
LISA	-	<i>Library and Information Science Abstracts</i>
LISTA	-	<i>Library and Information Science Technology Abstracts</i>
MHTX	-	Modelagem Conceitual Hipertextual
MRF	-	<i>Master Reference File</i>
NE	-	Nota de Escopo
NT	-	<i>Narrower Term</i>
NTG	-	<i>Narrower Term Generic</i>
OPAC	-	<i>On-line Public Access Catalogue</i>
PMEST	-	<i>Personality, Matter, Energy, Space, Time</i>
RI	-	Recuperação da informação
RSL	-	Revisão Sistemática da Literatura
RT	-	<i>Related Term</i>
SE	-	Sumário Expandido
SGBD	-	Sistema de Gerenciamento de Banco de dados
SKOS	-	<i>Simple Knowledge Organization System</i>
SOC	-	Sistemas de Organização do Conhecimento
SRI	-	Sistema de Recuperação da Informação
TA	-	Termo Associado
TAF	-	Teoria da Análise Facetada
TE	-	Termo Específico
THESAGRO	-	<i>Thesaurus Agrícola Nacional</i>
TICs	-	Tecnologias da informação e comunicação
TR	-	Termo Relacionado
UDC	-	<i>Universal Decimal Classification</i>
UFMG	-	Universidade Federal de Minas Gerais
UP	-	Usado Para
URI	-	<i>Uniform Resource Identifier</i>
URL	-	<i>Uniform Resource Locator</i>
USE	-	<i>Use</i>
XML	-	<i>Extensible Markup Language</i>

WWW - *World Wide Web*
W3C - *World Wide Web Consortium*

SUMÁRIO

1 INTRODUÇÃO	21
1.1 Justificativa e problema	23
1.2 Pressupostos.....	26
1.3 Objetivos	27
1.3.1 Objetivo geral	27
1.3.2 Objetivos específicos.....	27
1.4 Estrutura da Tese	27
2 REVISÃO DA LITERATURA.....	29
2.1 Revisão Sistemática da Literatura: definição e caracterização	29
2.2 Procedimentos iniciais da RSL: descrição das etapas	30
2.3 Resultados da aplicação da RSL	33
2.3.1 Análise quantitativa da RSL	42
2.3.2 Análise qualitativa dos resultados da aplicação da RSL	48
2.3.3 Interpretação dos resultados e conclusões da RSL	57
3 REFERENCIAL TEÓRICO E METODOLÓGICO	62
3.1 Sistemas de Organização do Conhecimento	62
3.1.1 Sistema de Classificação Bibliográfica	68
3.1.1.1 Classificação Facetada	77
3.1.1.2 Classificação Decimal Universal	80
3.1.2 Tesouros	99
3.1.2.1 Software <i>TemaTres</i>	114
3.2 Relações Semânticas na Biblioteconomia e Ciência da Informação	118
3.2.1 Classificação das relações semânticas	120
3.2.2 Relações semânticas em Sistemas de Organização do Conhecimento	125
3.3 Recuperação da informação	132
3.4 Compatibilidade e Convertibilidade de Linguagem	136
3.5 Simple Knowledge Organization System.....	145
4 CARACTERIZAÇÃO E DESCRIÇÃO DA METODOLOGIA	156
4.1 Caracterização da pesquisa	156

4.2 Delimitação do universo da pesquisa	159
4.3 Procedimentos metodológicos	160
4.4 Delimitação do ambiente da pesquisa.....	168
4.4.1 A Embrapa	169
4.4.2 O recorte temático e o AgroTermos	170
5 PROPOSTA DE UM MODELO PARA A RECUPERAÇÃO DA INFORMAÇÃO BASEADO NA CDU.....	173
5.1 Compatibilidade de linguagem entre os termos da CDU e AgroTermos	174
6 ANÁLISE E DESCRIÇÃO DOS RESULTADOS.....	205
6.1 Mapeamento do recorte temático.....	206
6.2 Compatibilidade de linguagem entre os SOC.....	213
6.3 Representação do modelo proposto na linguagem SKOS	226
7 CONSIDERAÇÕES FINAIS	244
REFERÊNCIAS	251
ANEXO A - ESTRUTURA CLASSIFICATÓRIA DA CDU (CLASSE 631 - AGRICULTURA EM GERAL)	278
ANEXO B - CONCEITOS UTILIZADOS NA AMOSTRA DO AGROTERMOS	289
ANEXO C - ESTRUTURA SEMÂNTICA DA AMOSTRA DO AGROTERMOS	296
ANEXO D - MODELO DE RELAÇÕES SEMÂNTICAS PARA OS INSTRUMENTOS DE ORGANIZAÇÃO NO CONTEXTO DA BIBLIOTECONOMIA E CIÊNCIA DA INFORMAÇÃO	308
ANEXO E - ESQUEMA DE NEVILLE PARA NÍVEIS DE CORRESPONDÊNCIAS ENTRE TESAUROS	309

1 INTRODUÇÃO

“A ciência consiste em substituir o saber que parecia seguro por uma teoria, ou seja, por algo problemático”.

(José Ortega y Gasset)

Com o avanço das tecnologias, é possível solucionar problemas causados pela falta de organização da informação em diversos ambientes informacionais. Saracevic (1996) verificou que os resultados da recuperação da informação (RI) devem ser coerentes com a expressão utilizada no momento da busca pelo usuário, de modo que esses resultados sejam relevantes e pertinentes em diversas áreas do conhecimento¹. Segundo o autor, pode-se dizer que a informação² organizada é um dos principais insumos para sustentar o desenvolvimento do Sistema de Recuperação da Informação (SRI), tendo em vista proporcionar resultados relevantes da informação recuperada.

Saracevic (1999) considera a RI a vertente tecnológica da Ciência da Informação (CI) resultado da relação que essa área estabelece com a área da Ciência da Computação (CC). A CI trabalha com a construção de Sistemas de Organização do Conhecimento (SOC), com o fim de representar as informações que circulam no contexto social, institucional ou individual, facilitando a comunicação entre os usuários e considerando as necessidades de informação de diferentes grupos de usuários. No campo da CC, o foco está na construção de tecnologias de informação para o processamento de um volume grande de dados, permitindo melhorar o armazenamento e recuperação da informação. O autor relata também, que o principal objetivo da RI é fornecer aos usuários o acesso ao conteúdo informacional dos documentos de um domínio do conhecimento em um SRI.

A partir do desenvolvimento das tecnologias da informação e comunicação (TICs), tornaram-se necessários mecanismos de organização e representação da informação, com o intuito de facilitar a RI a partir do conteúdo dos documentos, caracterizados como objetos informacionais. Esses objetos podem estar indexados no formato de texto, imagem, mapas, vídeos ou qualquer outro, desde que possibilitem o suporte à informação, cujo objetivo é registrar e comunicar a informação (SARACEVIC, 1999). Entretanto, em ambientes digitais, os mecanismos de representação devem possuir um nível de especificidade e exaustividade de estrutura semântica capaz de auxiliar o usuário a recuperar documentos relevantes, alinhados às estratégias de busca.

¹ Conhecimento é adquirido da informação processada e transformada de acordo com a necessidade de informação dos usuários e,

² Informação é o conhecimento registrado dos dados que são organizados ou processados, seja de forma impressa ou digital (SEMIDÃO, 2007).

A relevância tem sido um dos grandes desafios da RI, tanto em ambientes tradicionais, digitais ou na web, tornando-se um dos principais campos estudados na área da CI. Sabe-se que o processo de RI visa atender as necessidades dos usuários, com mais agilidade, o que demanda um trabalho colaborativo. Contudo, quando realizado no ambiente digital, esse processo necessita ser mais sofisticado, permitindo que os conceitos que representam o assunto de um documento correspondam aos termos utilizados pelos usuários durante a busca em um SRI.

Nesse contexto, a aplicação dos SOC, mesmo os tradicionais, para melhorar a RI on-line, tem sido estudada desde o surgimento da web. Todavia, o potencial do uso de estruturas de conhecimento na interface do usuário ainda não tem sido muito difundido. Para além das funções de representação, o uso dos SOC como instrumentos que garantem suporte à navegação por assunto tem tido bastante repercussão na literatura e na experiência profissional. Por isso, observa-se o aumento em relação do debate acerca da visualização do acesso por meio de apresentações hierárquicas dos assuntos, permitindo, assim, uma recuperação mais interativa das informações. Os SOC, quando incorporados aos serviços da web, podem funcionar como mapas semânticos que orientam os usuários nas buscas.

Com o aumento da padronização dos objetos representados, por meio do uso de padrões de metadados, a busca por assunto se tornou um elemento crucial na eficácia de um SRI. Assim, valorizou-se o uso dos SOC, que podem ser de vários tipos, sendo analisados, no âmbito desta pesquisa, dois SOC, os sistemas de classificação bibliográfica, mais especificamente a Classificação Decimal Universal (CDU) e os tesouros.

Adentrando-se à delimitação do tema da pesquisa, o objeto de pesquisa é a estrutura hierárquica da Classificação Decimal Universal (CDU) e os descritores de um tesouro, com o objetivo de trazer melhorias na recuperação da informação. Nesta tese, a CDU é utilizada como um recurso de classificação pré-coordenado e pós-coordenado, sendo acoplada a um tesouro, para que seja avaliado o potencial dos dois SOC, com o objetivo de organizar o conhecimento produzido em um domínio específico.

Apesar de a CDU ser um instrumento que possui uma estrutura classificatória com um nível elevado de especificidade de assuntos, ela não é totalmente efetiva na RI, ou seja, não resolve todos os problemas oriundos da falta de relevância durante a busca pelos usuários. Assim, durante esta pesquisa, foram propostas melhorias para as atividades de representação e recuperação no domínio da agricultura. Isso possibilitou evidenciar as relações semânticas entre os termos relacionados na CDU e os descritores do AgroTermos³ durante a compatibilidade de linguagem entre os dois vocabulários, permitindo a interoperabilidade semântica e o aumento da relevância na recuperação. As atividades de

³Disponível em: <https://sistemas.sede.embrapa.br/AgroTermos/#>

pesquisa foram facilitadas mediante o convênio/parceria entre a Universidade Federal de Minas Gerais (UFMG), o Grupo Modelagem Conceitual para Organização e Representação da Informação Hipertextual (MHTX)⁴ e a Empresa Brasileira de Pesquisa Agropecuária (Embrapa), em específico a Embrapa Informática Agropecuária (EIA), sediada em Campinas, no Estado de São Paulo.

1.1 Justificativa e problema

Para Creswell (2010), o trabalho de investigação possibilita benefícios, a partir dos resultados obtidos com a pesquisa. O autor explicita que os resultados serão importantes para o público, cujas justificativas estão associadas a contribuições de ordem teórica ou prática.

No campo da organização do conhecimento, percebe-se a falta de soluções eficientes para a realização dos processos de representação, armazenamento, busca e acesso a informações relevantes para satisfazer uma necessidade de informação do usuário. De acordo com Lancaster (2004), para uma recuperação eficiente em ambientes digitais, o primeiro passo é organizar e representar o conhecimento. A organização promove o tratamento do registro do conhecimento, seja de forma impressa ou digital, com o objetivo de prover a comunicação efetiva entre o sistema de informação e o usuário.

Atualmente, são necessários grandes esforços na organização e representação do conhecimento e da informação para melhorar a precisão⁵ e revocação⁶ na recuperação da informação. A razão disso é que os dados, informações e documentos são armazenados em bases de dados locais ou na web, na maioria das vezes de forma desordenada. Logo, os sistemas de informação possuem um esquema relacional entre as entidades (categorias) e os conceitos para padronizar a informação que será extraída dos termos armazenados em banco de dados. No contexto desta pesquisa, conforme relatado por Frâncu (2003a, 2003b), as relações semânticas da CDU e dos tesouro permitem o relacionamento entre os termos do tesouro com os da CDU que representam os conceitos em cada SOC, trazendo resultados relevantes para a RI, no contexto de um determinado domínio do conhecimento, sobretudo no ambiente digital.

⁴ Grupo de pesquisa coordenado pela Profa. Dra. Gercina Ângela de Lima, orientadora do autor desta tese. Registrado desde 2004 no diretório dos grupos de pesquisa do Brasil no CNPQ na página <http://dgp.cnpq.br/dgp/espelhogrupo/19689>. O site do grupo está hospedado na ECI/UFMG disponível em: <https://mhtx.eci.ufmg.br/>.

⁵ Precisão é a relação entre o número de documentos recuperados que são considerados relevantes em uma busca e o número total de documentos recuperados, ou seja, este índice está relacionado com a especificidade (qualidade) na RI, e a

⁶ Revocação é a relação entre o número de documentos relevantes que são recuperados sobre um tema em relação ao total de documentos existentes na base de dados sobre o mesmo tema, ou seja este índice está relacionado com a exaustividade (quantidade) na RI (LANCASTER, 1986).

Para que o usuário obtenha a informação desejada em determinado domínio, a preocupação com a falta de compatibilidade de linguagem não é recente. Em linhas gerais, os profissionais que atuam na organização de documentos, tais como os profissionais da informação, realizam diferentes esforços, com o uso de vários métodos para reduzir conflitos causados pela linguagem natural, tais como: ambiguidade⁷, homônimos⁸, polissemia⁹, sinonímia¹⁰ e termos que não representam o conteúdo dos documentos. Entretanto, em ambientes digitais, é necessário que a realização da busca ou navegação não seja apenas pela sintaxe expressa com o uso de termos isolados. Nessa perspectiva, é necessário considerar, também, o significado semântico por meio de uma notação que expresse, de maneira única, a informação contida nos documentos, considerando a limitação terminológica existente para a RI, mesmo com o uso de tesouros (SARACEVIC, 1996).

A falta de flexibilidade, conectividade e interatividade entre os assuntos são fatores preocupantes para a RI. Essas falhas são causadas pela ausência de representação de assuntos e do relacionamento entre eles para fins de recuperação, a partir da busca textual e navegacional. Sparck Jones (1970) aponta a necessidade de um grande esforço intelectual durante a classificação de novos termos para a representação temática de um domínio. Desse modo, todos os requisitos necessários para representar um domínio devem ser inseridos no sistema de classificação, devendo ser substituídas, também, as classificações que foram representadas de modo ineficiente durante os procedimentos de agrupamento (*clustering*), ou as classificações que ainda não foram representadas. Por essa razão, o processo de indexação é relevante para representar os assuntos contidos nos documentos e suas possíveis relações com outros assuntos.

Para que a utilização dos sistemas de classificação seja eficiente, torna-se necessário estabelecer quais os conteúdos serão indexados, utilizando padrões de metadados ou criando metadados para atender os requisitos do domínio, melhorando a recuperação da informação em ambientes digitais. Essa questão pode ser um problema extremamente complexo e que vai se evoluindo, o que justifica o desenvolvimento de pesquisas, em especial no contexto de um domínio específico.

É importante destacar que diversos autores apontam relevância significativa no uso de outros sistemas de classificação mais comuns, tais como: a CDD, *Library of Congress*

⁷ Diz respeito a um efeito linguístico, cujos termos ou palavras possuem significados distintos, ocorrendo diversas interpretações (BRASCHER, 2002).

⁸ São termos ou palavras que apresentam igualdade ou semelhança na pronúncia com mesma grafia, porém com significados distintos (HJØRLAND, 2007a).

⁹ Refere-se a um único signo linguístico (termo ou expressão) que podem ter vários sentidos, modificando-se, conforme o contexto. Ex.: arquivo no sentido de: móvel, instituição, conjunto de documentos (BRASCHER, 2002).

¹⁰ É relacionado a termos ou palavras que possuem o mesmo significado, e se diferem apenas pelo signo linguístico (MURPHY, 2003).

Classification (LCC), *Cólon Classification* (CC) e a *Bliss Bibliographic Classification* (BC2), visando à recuperação da informação. A maioria dos estudos publicados nessa abordagem justifica a escolha desses sistemas, considerando a estrutura hierárquica e enumerativa dessas classificações. De fato, tais estruturas facilitam a decomposição dos assuntos por meio de termos (vocabulário controlado) e da notação que representa um índice por assuntos (COATES, 1988; CHAN, 2000; MATVEYEVA, 2002; HJØRLAND; NISSEN PEDERSEN, 2005; LERTMAHAKIAT; KAMPEERAPAAPPAT; MINGKHWAN, 2008; GEDAM; PARADKAR, 2013; AHN; LIN; KHOO, 2014; FENG; CHIN, 2015).

Para além das funções de representação, os SOC, tais como os sistemas de classificação, tesouros, entre outros, caracterizam-se como instrumentos que fornecem suporte à navegação por assunto e permitem a visualização do acesso por meio de apresentações hierárquicas de diversas áreas, possibilitando, assim, uma recuperação mais interativa das informações. A utilização dos SOC tem sido notória e, acredita-se que auxiliam os usuários na RI, funcionando como links semânticos para obter informações pertinentes durante a busca em um SRI.

O valor da classificação na RI também foi reconhecido como significativo no ambiente digital e na web (VICKERY, 1980; SOERGEL, 1999; HODGE, 2000; CURRIER; WAKE, 2001; SIMÕES; FREITAS; RODRÍGUEZ-BRAVO, 2018). Por ter uma estrutura hierárquica que pode ser expandida no momento da pesquisa, a classificação em um SRI permite uma seleção contextualizada, auxiliando na desambiguação dos termos de busca, o que torna os resultados mais relevantes. Outro aspecto importante é a possibilidade de fornecer acesso a assuntos em ambientes multilíngues, cujos sistemas de classificação bibliográfica são benéficos e muito utilizados. Isso decorre do uso dos termos de indexação, que formam as notações, não somente por serem independentes do idioma, mas também porque esses esquemas foram traduzidos para vários idiomas. Portanto, não é estranho que os sistemas de classificação disponíveis e conhecidos passaram a ser utilizados no suporte à descoberta de recursos na internet, desde o início da década de 1990.

No âmbito da Biblioteconomia e Ciência da Informação (BCI), as classificações facetadas são frequentemente discutidas em pesquisas recentes. Dentre as diversas abordagens de estudo, destacam-se com forte predominância estudos que versam sobre a contribuição desse sistema classificatório para organização do conhecimento em ambientes digitais (MACULAN, 2011; PONTES, 2013; SILVA, 2013; FRICKÉ, 2017; GREEN, 2017; SLAVIC; DAVIES, 2017; ULLAH; KHUSRO; ULLAH, 2017; GNOLI, 2019; ZENG, 2019).

Algumas características das classificações facetadas podem ser observadas na CDU, como, por exemplo, o potencial analítico-sintético com o uso dos processos de análise e síntese para melhorar a precisão e revocação na recuperação da informação. Por meio da classificação em facetas, torna-se possível a representação de assuntos compostos ou

complexos por meio da combinação de números de classes da CDU em diversas facetas pelo processo de síntese. Assim, esse tipo de classificação auxilia na exibição de assuntos, pois os assuntos serão mais facilmente localizados, pois estarão organizados na estrutura hierárquica, interligados de forma multidirecional por meio de remissivas e associados por uma área ou domínio do conhecimento. A partir do agrupamento em facetas, o processo classificatório pode ser usado para navegação linear, pré-sequenciada e inflexível de estruturas classificatórias que exibem o assunto em uma extensa cadeia de classes e subclasses (hierarquia), para que seja possível recuperar a informação para responder a uma necessidade de informação dos usuários.

Diante do exposto, o problema que norteia este estudo é: Qual é o potencial da estrutura classificatória da CDU compatibilizada com um tesauro para a representação de assuntos em um domínio específico para melhorar a RI?

Levando-se em consideração a justificativa e a relevância do problema exposto, são listados, a seguir, os pressupostos, o objetivo geral e os objetivos específicos que foram necessários para alcançar os resultados e que sustentam a questão proposta.

1.2 Pressupostos

Os pressupostos que justificam o desenvolvimento desta pesquisa são variados, sobretudo por constituir um tema muito comum no cotidiano dos profissionais que organizam a informação para serem posteriormente recuperadas. Assim, com o problema formulado e a relevância do tema, tanto em termos práticos quanto teóricos, são definidos os seguintes pressupostos:

- 1) A estrutura classificatória da CDU em conjunto com a do tesauro, possibilita a ocorrência de vários tipos de relações semânticas, entre os termos de um determinado assunto.
- 2) A utilização da CDU e do tesauro de modo pós-coordenado, a partir da combinação dos números de classes e descritores, permitem a padronização terminológica (vocabulário controlado) para resolver problemas de ambiguidade da linguagem natural.
- 3) A compatibilidade de vocabulários controlados melhora a interoperabilidade semântica.

- 4) A utilização de padrões de organização do conhecimento no âmbito da web semântica, possibilita a interoperabilidade entre diferentes sistemas de informação.

1.3 Objetivos

Diante das justificativas, o problema e pressupostos apresentados nesta pesquisa, foi possível a formulação dos objetivos propostos. Portanto, a seguir, apresentam-se os objetivos que conduziram o desenvolvimento desta tese.

1.3.1 Objetivo geral

- Propor um modelo de compatibilidade de linguagem entre a Classificação Decimal Universal e o tesouro AgroTermos com o intuito de melhorar a recuperação da informação, em um domínio específico, entre diferentes sistemas de informação.

1.3.2 Objetivos específicos

- 1) Analisar a compatibilidade dos termos da estrutura classificatória da CDU com os descritores do tesouro AgroTermos, na representação do assunto “Intensificação Agropecuária” do domínio da agricultura.
- 2) Verificar o potencial da recuperação com o uso da expansão dos termos e seus relacionamentos semânticos mapeados nesses dois SOC.
- 3) Representar o modelo de compatibilidade proposto, utilizando o padrão Simple Knowledge Organization System, no âmbito da web semântica.

1.4 Estrutura da Tese

O conteúdo desta pesquisa foi estruturado refletindo as etapas que foram sendo desenvolvidas à medida que os procedimentos eram aplicados e interpretados. Para facilitar a organização dos conteúdos contidos nesses capítulos, fez-se necessário agrupar esses conteúdos em sete blocos maiores, os quais constituíram os capítulos da pesquisa.

O capítulo 1 apresenta a introdução do tema. Para tanto, são elencados aspectos necessários para delimitar o tema, tais como a descrição da justificativa, o problema da pesquisa, os pressupostos, os objetivos e, por fim, o modo com que a pesquisa foi organizada (estrutura).

O capítulo 2 reservou-se à construção da revisão da literatura, mencionando os trabalhos já publicados e relacionados com a temática desta pesquisa. Esse procedimento é necessário, pois mapeia as diferentes pesquisas relacionadas ao tema da tese, ou seja, identifica e seleciona pesquisas que utilizam a CDU e o tesauro para a RI. Com essa finalidade, foi realizada uma revisão sistemática, verificando os resultados obtidos em cada trabalho, em correspondência com a proposta desta pesquisa.

No capítulo 3 é exibido o referencial teórico e metodológico, apresentando informações gerais para o entendimento do domínio investigado e associado à área da Ciência da Informação. Portanto, são explorados conceitos e características dos campos inter-relacionados para o entendimento teórico-metodológico utilizados durante toda a pesquisa.

Os procedimentos metodológicos são detalhados no capítulo 4. Nessa parte do trabalho, apresenta-se a metodologia que auxiliou no desenvolvimento da pesquisa, descrevendo como o modelo de compatibilidade foi desenvolvido e os procedimentos necessários para utilizar a CDU e o AgroTermos, haja vista facilitar o processo de recuperação.

No capítulo 5, é proposto um modelo a partir da CDU e do tesauro AgroTermos com a aplicação do estudo de caso a partir dos termos que representam os conceitos do assunto “Intensificação Agropecuária” do domínio da agricultura e suas relações semânticas com a validação e interpretação dos dados utilizados.

O capítulo 6 é destinado a apresentação da análise e descrição dos resultados obtidos no estudo de caso, os quais foram comparados com a literatura, de modo a possibilitar o atendimento ou não aos objetivos e pressupostos da tese. É nesta parte, também, que são evidenciados alguns apontamentos acerca do potencial da CDU e do AgroTermos para melhorar a recuperação.

Por fim, no capítulo de número 7, são formuladas as considerações finais, as quais apontam os resultados do estudo e o que eles representam para a evolução do conhecimento até então produzido sobre a temática. Além das conclusões, o capítulo também apresenta as limitações encontradas no percurso da pesquisa, as contribuições reveladas com os procedimentos e as indicações de estudos futuros.

2 REVISÃO DA LITERATURA

“Descobrir consiste em olhar para o que todo mundo está vendo e pensar uma coisa diferente”.
(Roger Von Oech)

Neste capítulo, apresenta-se a Revisão Sistemática da Literatura (RSL), cujo objetivo foi mapear e identificar os estudos relevantes sobre a utilização da estrutura classificatória da CDU e de um tesouro com o propósito de explorar os estudos que apresentam melhorias na RI em ambientes digitais.

2.1 Revisão Sistemática da Literatura: definição e caracterização

As RSL conforme relatado por Higgins *et al.* (2019), procuram agrupar todas as evidências que se encaixam nos critérios de elegibilidade (inclusão ou exclusão) de estudos predefinidos, a fim de abordar uma questão de pesquisa específica. Para isso, utilizam-se métodos explícitos e sistemáticos que possuem orientação metodológica para a preparação e execução dos procedimentos exigidos na revisão. Assim, são realizadas diferentes atividades, como: avaliação crítica, síntese dos resultados e a diminuição de conflitos de definições relacionadas às investigações de uma determinada temática.

A revisão de literatura narrativa se difere da revisão sistemática, pois responde a uma pergunta específica. Os autores caracterizam a revisão sistemática como uma forma de abranger a literatura de modo não tendencioso, no que tange à seleção, coleta, síntese e interpretação dos dados. Nessa perspectiva, a revisão sistemática segue uma metodologia explícita e sistemática para buscar e selecionar pesquisas, além de avaliar a relevância e a validade das informações encontradas. As publicações são selecionadas com o objetivo de minimizar o caminho a ser percorrido, com descobertas mais confiáveis, haja vista responder questões relevantes para formular conclusões e tomar decisões. Com isso, buscam-se selecionar estudos de qualidade, fazer análises estatísticas (meta-análise) dos dados após a coleta, além de identificar a escolha dos métodos e os resultados obtidos pelos pesquisadores, durante o percurso da pesquisa (TORLONI *et al.*, 2009).

Os procedimentos para a busca sistemática foram retirados das orientações metodológicas registradas no manual da *Cochrane Collaboration*, organização fundada em 1993, e que se dedica à valorização e empreendimentos em prol desse método de pesquisa, contando com um grande número de colaboradores em diversos países. Os editores-chefes do manual da Cochrane, Higgins e Green (2011), contam com uma grande equipe de autores e editores, composta por diversos perfis profissionais, vinculados, principalmente, às áreas da Saúde e da Assistência Social. Nessa equipe, há pesquisadores, advogados e outros

profissionais, cuja missão é promover a tomada de decisões com base em evidências, produzindo revisões sistemáticas relevantes, acessíveis e de alta qualidade.

A RSL proposta por essa instituição permite, antes da publicação dos resultados, reduzir o impacto dos vieses nos estudos selecionados. Desse modo, é possível verificar a transparência dos métodos e processos, o que diminui a ocorrência de trabalhos duplicados, como também a escolha de trabalhos revisados por pares. O manual da Cochrane baseia-se na experiência de mais de cem colaboradores, desenvolvedores de métodos e editores que formam um grupo com estes profissionais no mundo inteiro. Este grupo que realiza a revisão por pares, é responsável por conduzir pesquisas para embasar a metodologia da Cochrane (CUMPSTON, 2019).

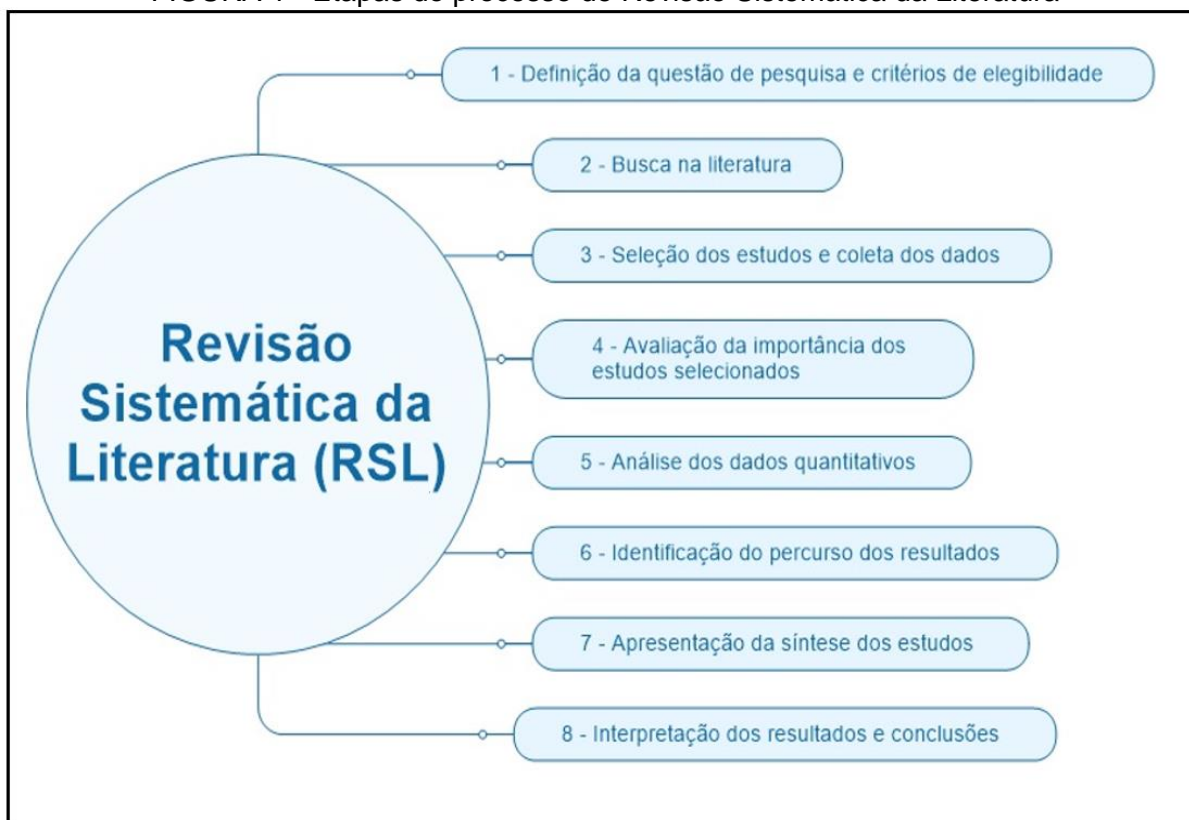
De-La-Torre-Ugarte *et al.* (2011) apontam que a revisão sistemática é caracterizada por um conjunto de objetivos claramente definidos com critérios de elegibilidade de estudos predefinidos, além de ter uma metodologia explícita e reproduzível. Na RSL, a busca deve ser de forma sistemática para identificar todos os estudos que atendam aos critérios de elegibilidade e seja possível verificar a validade dos resultados dos estudos, por exemplo, mediante a avaliação do caminho percorrido na pesquisa bem como uma apresentação sistemática da síntese das características e resultados dos estudos incluídos.

2.2 Procedimentos iniciais da RSL: descrição das etapas

Nesta pesquisa, foram utilizadas as etapas definidas por Higgins e Green (2011), a partir das recomendações do manual da *Cochrane Collaboration*¹¹. A versão utilizada foi definida em oito etapas: 1) Definição da questão de pesquisa e critérios de elegibilidade; 2) Busca na literatura; 3) Seleção dos estudos e coleta dos dados; 4) Avaliação da importância dos estudos selecionados; 5) Análise dos dados quantitativos; 6) Identificação do percurso dos resultados; 7) Apresentação da síntese dos estudos; e 8) Interpretação dos resultados e conclusões. O desenho metodológico das etapas citadas é ilustrado na Figura 1.

¹¹As revisões produzidas por essa equipe são disponibilizadas no site da instituição: www.cochrane.org. A versão do manual da *Cochrane Collaboration* utilizada nesta tese foi a versão 5.1.0, disponível em: <https://handbook-5-1.cochrane.org/>.

FIGURA 1 - Etapas do processo de Revisão Sistemática da Literatura



FONTE: Elaborada pelo autor (2020).

Para operacionalizar todo o estudo, realizou-se a pesquisa bibliográfica de dezembro de 2018 até o mês de agosto de 2019, como recorte temporal desta revisão.

Na primeira etapa, foi definida a questão de pesquisa e os critérios de elegibilidade que determinaram a inclusão e exclusão dos estudos na revisão.

A segunda etapa nomeada “Busca na literatura”, requer um nível mais profundo de investigação, com estratégias objetivas e que sejam reprodutíveis, tendo em vista localizar os estudos selecionados. Então, sugere-se a realização de uma busca por assunto livre, em ordem alfabética de títulos, pelo título de periódicos, ou em fontes de dados para recuperar a literatura sobre o tema, como, por exemplo: (a) bases de dados bibliográficas, (b) revistas e outras fontes não bibliográficas de banco de dados e (c) estudos não publicados e em andamento (HIGGINS; GREEN, 2011).

Para esses autores, a maioria das pesquisas é realizada mediante a consulta às bases de periódicos relevantes, ou seja, bases relacionadas à área do conhecimento que está sendo investigado, e que tenham relação com o tema da pesquisa, motivo pelo qual esse tipo de fonte de informação foi escolhido para nortear a RSL desta tese. Sampaio e Mancini (2007) declaram que uma busca eficaz envolve não só uma estratégia onde são incluídos os termos adequados, mas também, a escolha da base de dados, que pode ser bibliografias gerais ou especializadas, desde que contenham estudos focados ao tema da pesquisa. As autoras esclarecem que, a partir da seleção de material confiável, espera-se mapear boa parte da

produção científica que versa sobre o assunto, o que possibilitará resultados mais íntegros e consistentes, fortalecendo as tomadas de decisão, conforme delimitado na questão norteadora. Na segunda fase, também são definidos os termos a serem utilizados nas expressões de busca, o tempo entre os estudos ou outras fontes de informação, a fim de encontrar estudos relacionados com a temática proposta (HIGGINS; GREEN, 2011).

Após a realização das buscas, na terceira etapa da revisão, acontece a seleção dos estudos e coleta dos dados, onde são avaliados a elegibilidade dos estudos e a extração dos dados das publicações. Para Higgins e Green (2011), os dados podem ser relatados em diferentes formatos, mas geralmente podem ser convertidos em variáveis pertinentes à pesquisa, a partir dos materiais selecionados. Os resultados de uma revisão sistemática dependem, criticamente, de decisões relacionadas aos estudos incluídos e quais serão os dados desses estudos que serão apresentados e analisados. Em seguida, os métodos utilizados para essas decisões devem ser transparentes e escolhidos para minimizar possíveis erros e vieses (HIGGINS; GREEN, 2011).

O processo de seleção dos estudos é geralmente realizado com a utilização de um software de gerenciamento de referências para inclusão e remoção de registros duplicados. Higgins e Green (2011) examinam os títulos e resumos para remover as publicações irrelevantes, selecionam o texto completo das publicações que realmente possuem um potencial relevante e verificam a conformidade com os estudos e critérios de elegibilidade para prosseguir com a coleta dos dados (HIGGINS; GREEN, 2011).

A avaliação da importância dos estudos selecionados, é executada na quarta etapa, observando os vieses ou percursos das publicações que foram incluídas em consonância com a temática. Higgins e Green (2011) consideram que as diferenças encontradas podem ajudar a explicar a variação nos resultados dos estudos, representando a falta de veracidade nos resultados ou inferências dos estudos incluídos em uma revisão sistemática. Assim, pode ser necessária a utilização de ferramentas para a investigação e análises estatísticas, de modo que o método de avaliação seja mais eficaz. Assim, todas as publicações são analisadas de forma superficial, a princípio, para ver se correspondem aos mesmos assuntos (HIGGINS; GREEN, 2011).

Na quinta etapa é realizada a extração dos dados das publicações para a combinação das análises dentro do universo de referências selecionado. Nessa etapa, são realizadas análises quantitativas para identificar informações que sejam estatisticamente significativas para um conjunto de estudos e não para estudos individuais. A princípio, são criadas categorias previamente estabelecidas, com alinhamento dos estudos a cada uma dessas categorias temáticas. Deeks *et al.* (2019) alertam que os estudos semelhantes ao tema de pesquisa devem ser agrupados em categorias, desde que esses agrupamentos não induzem a erros durante a análise. Com os dados retirados das publicações, foram produzidas

comparações estatísticas para auxiliar na escolha de soluções quando encontradas opiniões divergentes entre os estudos ou a criação de novas hipóteses (DEEKS *et al.*, 2019).

O percurso dos resultados obtidos, na sexta etapa, são identificados os vieses das pesquisas de acordo com a divulgação dos resultados das pesquisas, sendo influenciada pela natureza e direção dos resultados obtidos. De acordo com Higgins e Green (2011), os resultados 'positivos', são mais significativos, indicando maior probabilidade de serem publicados com maior rapidez. Outro fator importante é quando as publicações estão no idioma inglês e foram publicadas em periódicos de alto fator de impacto, o que leva maior probabilidade das publicações obterem maior número de citações por outros autores. No entanto, a contribuição feita em revisões sistemáticas de estudos com resultados 'negativos', mesmo sendo menos significativos, são tão importantes quanto aos estudos com resultados mais significativos.

Higgins e Green (2011) apresentam na sétima etapa, a análise qualitativa dos estudos com a síntese dos resultados obtidos na revisão por meio da avaliação mais detalhada, com apresentação dos conceitos vinculados à temática, bem como a descrição dos objetivos, das metodologias aplicadas para o desenvolvimento e os resultados de cada estudo. A análise qualitativa reconhece o tema de forma mais ampla, dificilmente a partir de uma questão específica bem definida, e não segue um processo rígido durante seu desenvolvimento, como acontece na análise quantitativa e por fim, a busca das fontes não é predeterminada e específica, sendo frequentemente menos abrangente.

Por fim, na oitava etapa, Higgins e Green (2011) recomendam encerrar as revisões com a análise dos resultados e conclusões da revisão, onde é apresentada a análise dos resultados da revisão de literatura e sua conclusão. Na revisão da literatura, buscou-se apresentar um mapeamento da produção científica sobre as metodologias de RI em ambientes digitais, que fazem uso da CDU associada a um tesouro, sendo que o resultado foi alcançado a partir das análises nos subcapítulos 2.3.1 e 2.3.2.

2.3 Resultados da aplicação da RSL

Na 1ª etapa, a questão definida buscou responder: “Quais pesquisas foram realizadas sobre a utilização da CDU e de um tesouro, com o propósito de melhorar a RI em ambientes digitais?”. Quanto aos critérios de elegibilidade, foram definidos os de inclusão dos estudos, assim estabelecidos: 1) publicações científicas (artigos de periódicos e de conferências, teses e dissertações); 2) relacionadas à questão de pesquisa; e 3) sem recorte temporal, por acreditar que a temática tenha sido pouco abordada. Os critérios de exclusão utilizados foram: 1) tipologia documental (monografias, editoriais, prefácios, sumários, entrevistas, notícias, revisões, tutoriais, workshops, painéis, pôsteres e publicações pagas);

2) estudos que estejam em desacordo com a questão de pesquisa; 3) trabalhos que não estejam escritos nos idiomas português, espanhol e inglês; e 4) publicações duplicadas.

A definição das expressões de busca é o resultado da 2ª etapa, as quais foram utilizadas nas buscas às bases vinculadas à área do conhecimento *Ciências Sociais Aplicadas* e subárea *Ciência da Informação*, a partir da listagem descrita no Portal de Periódicos da Capes, que é uma plataforma de acesso a diversos periódicos e bases de dados assinadas pela CAPES para uso das Universidades em diversas áreas do saber. Dessas bases, delimitando a área e subárea mencionadas, obteve-se o total de 88 bases. Em seguida, a partir desse quantitativo, foi selecionada a base de dados de conteúdo multidisciplinar, que é o Google Acadêmico, e cinco bases especializadas: *Library and Information Science Abstracts* (LISA), *Information Science & Technology Abstracts* (ISTA), *Library, Information Science & Technology Abstracts with Full Text* (LISTA), *Scopus* e *Web of Science*. Essas bases foram escolhidas, considerando a relevância e a relação que possuem com a área e com a subárea do conhecimento delimitadas para esta revisão. Após essas escolhas, partiu-se para a elaboração das estratégias de busca, as quais, junto aos termos escolhidos, foram utilizados símbolos, como: parênteses, operadores booleanos (*AND* e *OR*), e aspas dupla (“ ”), haja vista garantir uma busca mais refinada.

A princípio, as expressões foram aplicadas em todas as bases de dados bibliográficas selecionadas sem a utilização de critérios de inclusão ou exclusão. As buscas realizadas com o termo “tesauro” no plural e no singular, visto que não há padronização terminológica do termo nos estudos publicados, ora com uso no plural, ora no singular ou dos dois modos. Estes termos foram incluídos após a leitura preliminar de alguns estudos, que utilizavam a CDU, e que também usava um tesauro para auxiliar na RI, motivo pelo qual o tesauro foi incluído nas expressões de busca. Esses procedimentos manifestaram-se como um pré-teste, ou seja, aferir uma visão genérica do universo, como também, explorar todas as alternativas/ocorrências possíveis.

Após a realização dos testes iniciais, mesmo sem aplicação dos critérios de inclusão e exclusão, e utilizando as expressões E1, E2 e E3, foi retornado pelo sistema um total de 10.716 publicações nos idiomas inglês, espanhol e português. O Quadro 1 apresenta essas estratégias.

QUADRO 1 - Expressões utilizadas para busca na literatura

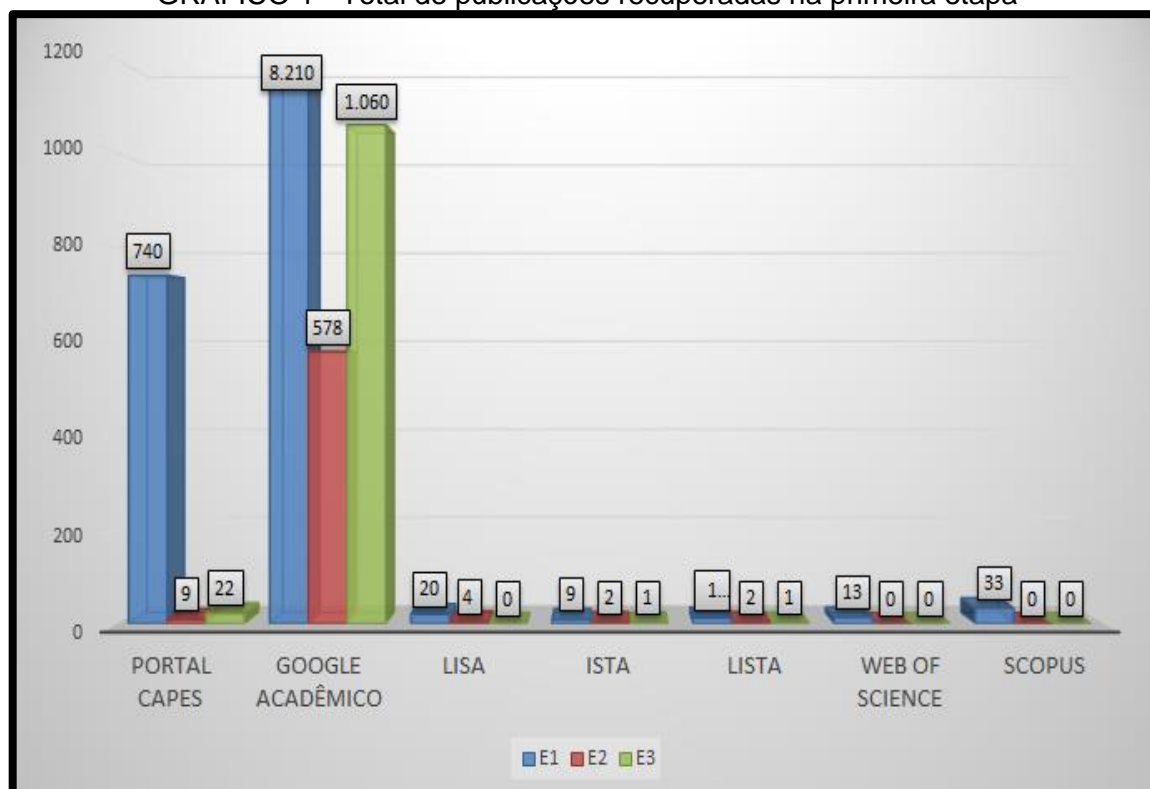
Sigla	Expressões de busca
E1	(“ <i>classification systems</i> ” AND “ <i>information retrieval</i> ” AND (“ <i>thesaurus</i> ” OR “ <i>thesauri</i> ”))
E2	(“ <i>sistemas de clasificación</i> ” AND “ <i>recuperación de información</i> ” AND (“ <i>tesauro</i> ” OR “ <i>tesauros</i> ”))
E3	(“ <i>sistemas de classificação</i> ” AND “ <i>recuperação da informação</i> ” AND (“ <i>tesauro</i> ” OR “ <i>tesauros</i> ”))
E4	(“ <i>udc</i> ” AND “ <i>information retrieval</i> ” AND (“ <i>thesaurus</i> ” OR “ <i>thesauri</i> ”))

E5	("cdu" AND "recuperación de información" AND ("tesauro" OR "tesauros"))
E6	("cdu" AND "recuperação da informação" AND ("tesauro" OR "tesauros"))
E7	("universal decimal classification" AND "information retrieval" AND ("thesaurus" OR "Thesauri"))
E8	("clasificación decimal universal" AND "recuperación de información" AND ("tesauro" OR "tesauros"))
E9	("classificação decimal universal" AND "recuperação da informação" AND ("tesauro" OR "tesauros"))

FONTE: Elaborado pelo autor (2020).

As expressões E1, E2 e E3 retornaram mais resultados pelo Portal Capes e na base de dados do Google Acadêmico, sendo que na primeira por se tratar de um ambiente de acesso a vários periódicos e a segunda por se tratar de uma base multidisciplinar. As expressões E2 e E3 não retornaram nenhum resultado nas bases *Web of Science* e *Scopus*, e a expressão E3 não retornou nenhum resultado na base LISA, conforme apresentado no Gráfico 1.

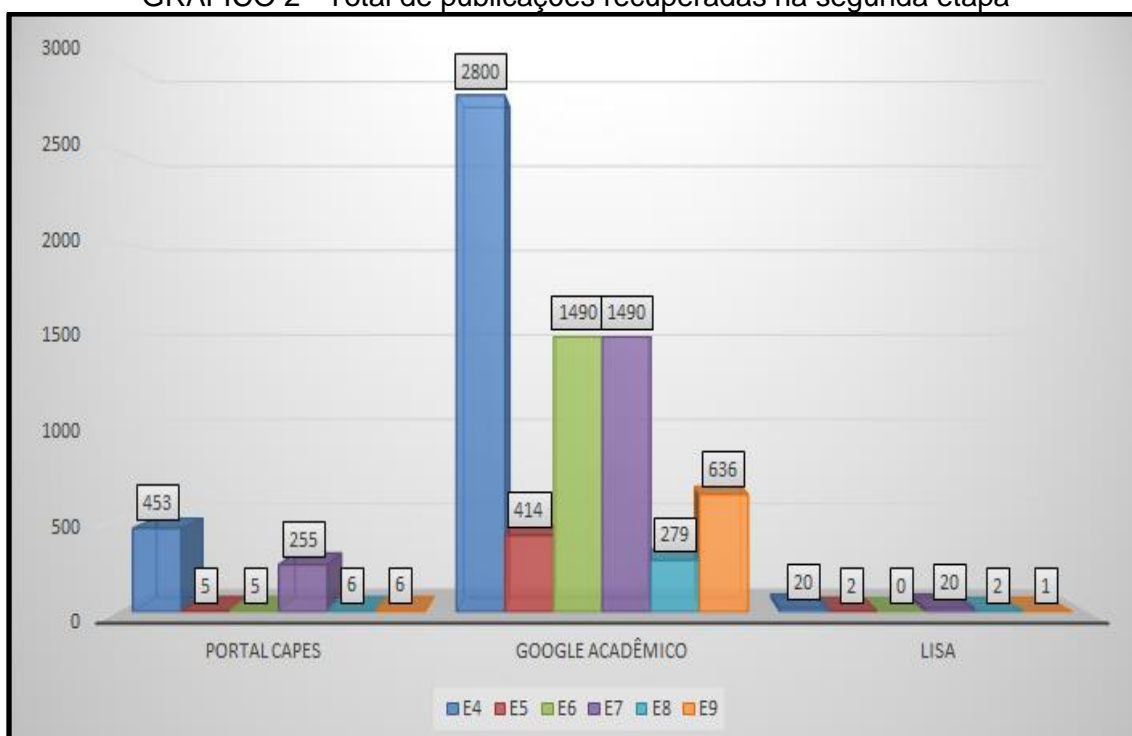
GRÁFICO 1 - Total de publicações recuperadas na primeira etapa



FONTE: Dados da pesquisa (2020).

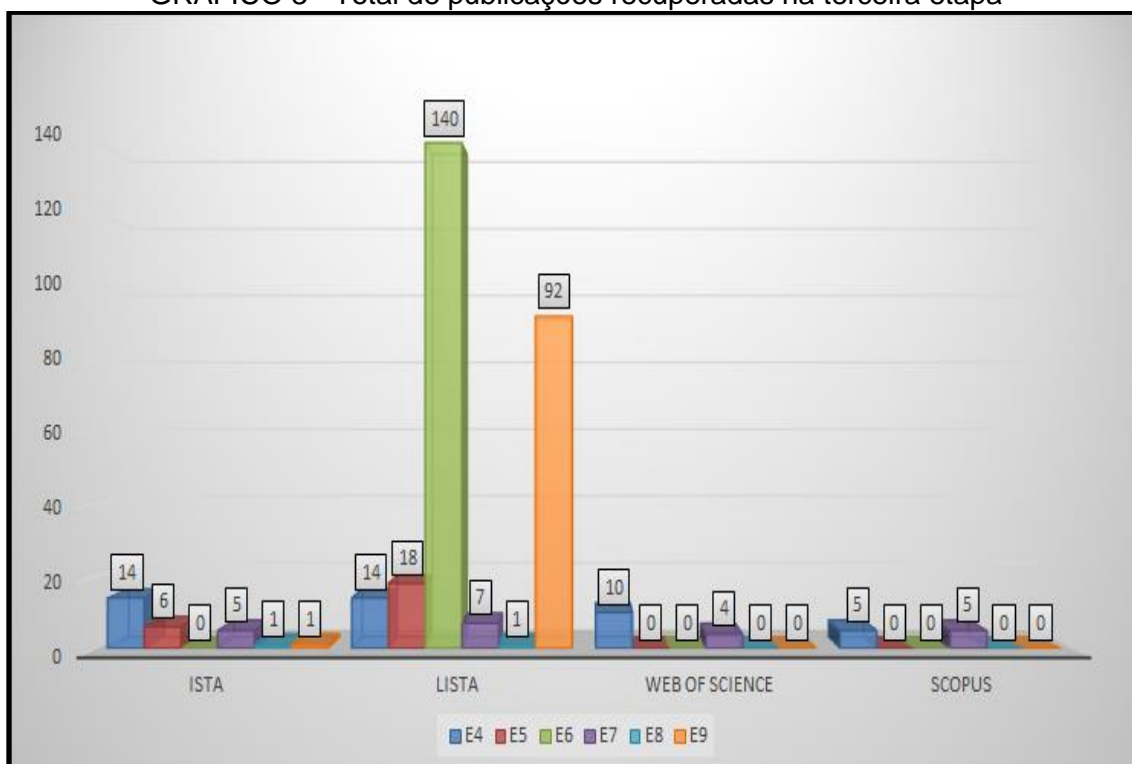
A fim de concluir a fase de busca na literatura, com o uso das demais expressões, obteve-se o total de 8.207 publicações nos três idiomas citados. Assim, verificou-se que as expressões com maior revocação durante os testes em cada base de dados foram: E4, E6 e E7, nas bases Google Acadêmico e a E6 na base LISA e ISTA. As expressões E5, E6, E8 e E9 não trouxeram resultados nas bases *Web of Science* e *Scopus*. Os gráficos 2 e 3 descrevem esses resultados.

GRÁFICO 2 - Total de publicações recuperadas na segunda etapa



FONTE: Dados da pesquisa (2020).

GRÁFICO 3 - Total de publicações recuperadas na terceira etapa



FONTE: Dados da pesquisa (2020).

Na 3ª etapa, ao selecionar estudos e coletar dados, foi realizada uma análise de conteúdo no título, resumo, palavras-chave e no texto completo, a fim de selecionar os documentos relevantes e pertinentes capazes de atender a questão de pesquisa desta revisão

da literatura. Para o gerenciamento e compartilhamento das referências bibliográficas, utilizou-se o software *Mendeley*. Assim, com a exclusão das publicações de acordo com os critérios de elegibilidade e análise de cada uma, o software gerou uma lista com 210 publicações.

Por fim, no que tange ao software para geração das referências, destaca-se que ele permite que elas sejam identificadas, recuperadas e classificadas, referenciando o texto original das publicações científicas ou acadêmicas. Ademais, o software analisa o conjunto de documentos, descartando os irrelevantes, de modo que sejam mantidos apenas aqueles que irão compor a amostra da revisão e que serão avaliados na próxima etapa.

A 4ª etapa avaliou os estudos selecionados, a partir das publicações que foram analisadas com leituras dinâmicas, tendo em vista identificar a relação do assunto do documento com o tema abordado nesta revisão. Nesse processo, foram levantados os seguintes dados das publicações selecionadas: autor e coautor, título da publicação, periódico, número de citação, idioma e ano da publicação. Para facilitar as análises dos dados, foram estabelecidas as seguintes categorias: autores que mais publicaram sobre a temática, idiomas das publicações, tipos de documentos (considerando os canais ou meios de comunicação dos trabalhos), ano de publicação, trabalho mais citado, periódico de maior impacto, e categorias temáticas abordadas nas publicações.

Os assuntos das publicações retirados dos títulos, palavras-chave e resumos foram: análise de assunto, análise facetada, banco de dados, catálogo on-line, catálogo web, CDU, classificação analítico-sintético, classificação bibliográfica, classificação facetada, combinação de descritores, compatibilidade de linguagens, especificidade, *gateways* de assunto, índice de assunto, interoperabilidade, linguagem de indexação, modelos de classificação, navegação, OPAC, organização da informação, padrões de vocabulário, relação hierárquica, relações ou fenômenos semânticos, sistema de biblioteca, sistema de recuperação da informação, sistemas de classificação, SOC, tesouro e web semântica.

Mediante o levantamento dos dados das 210 publicações, no final da 5ª etapa, foi possível excluir trabalhos que não correspondiam ao problema de pesquisa. Com efeito, após análises de relevância dos 210 estudos, chegou-se em 44 publicações referentes à utilização da CDU, vinculada com um tesouro, haja vista melhorar a RI.

As informações referentes às publicações foram retiradas do Google Acadêmico (número de citação) e o restante foi retirado de cada publicação, durante o processo de análise da revisão. O Quadro 2 apresenta, em ordem cronológica, as principais categorias extraídas das publicações após a coleta dos dados, o que possibilitou apresentar a amostra da pesquisa, conforme recomendações de Higgins e Green (2011).

QUADRO 2 - Principais dados levantados no processo de análise

Nº	Autoria e ano	Título	Tipos de documentos	Nº Citação	Idioma
1	Freeman (1964)	Computers and Classification Systems. Journal of Documentation	Artigo de periódico Journal of Documentation	18	Inglês
2	Mills (1970)	Progress in Documentation -Library Classification	Artigo de periódico Journal of Documentation	15	Inglês
3	Dahlberg (1971)	Possibilities for a new universal decimal Classification	Artigo de periódico Journal of Documentation	13	Inglês
4	Wall (1973)	Indexing language structure for automated retrieval	Artigo de periódico Information Storage and Retrieval	3	Inglês
5	Hindsoti (1979)	Reflections on the utilization of the Universal Decimal Classification	Artigo de periódico Information Storage and Retrieval	2	Inglês
6	Wall (1980)	Intelligent indexing and retrieval: A man-machine partnership	Artigo de periódico Information Processing and Management	21	Inglês
7	Fiałkowski, Muraszkiwicz e Tołwińska (1981)	Methodology of world clearinghouse activity	Artigo de periódico Information Storage and Retrieval	3	Inglês
8	Buxton (1990)	Computer searching of UDC numbers	Artigo de periódico Journal of Documentation	28	Inglês
9	McIlwaine (1997)	The Universal Decimal Classification: Some Factors Concerning Its Origins, Development, and Influence	Artigo de periódico Journal of the American Society for Information Science	61	Inglês
10	Chan e Zeng (2002)	Ensuring interoperability among subject vocabularies and knowledge organization schemes: a methodological analysis	Artigo de periódico IFLA Journal	52	Inglês
11	Frâncu (2003a)	The impact of specificity on the retrieval power of a UDC-based multilingual thesaurus	Artigo de periódico Cataloging & classification	4	Inglês
12	Frâncu (2003b)	Multilingual access to information using an intermediate language	Tese	5	Inglês
13	Chudamani (2004)	Classification Model for Libraries in the Digital Environment	Artigo de periódico International Caliber	0	Inglês
14	Schallier (2004)	What a subject search interface can do	Artigo de periódico Extensions & Corrections to the UDC	2	Inglês

15	Slavic (2004a)	UDC implementation: from library shelves to a structured indexing language	Artigo de periódico International Cataloging & Bibliographic Control	14	Inglês
16	Slavic (2005)	Classification management and use in a networked environment: the case of the Universal Decimal Classification	Tese	2	Inglês
17	Slavic (2006a)	Interface to classification: some objectives and options Item type Preprint Interface to Classification: Some Objectives and Options	Artigo de periódico Extensions and Corrections to the UDC	15	Inglês
18	Slavic (2006b)	The level of exploitation of Universal Decimal Classification in library OPACs: A pilot Study	Artigo de periódico Vjesnik bibliotekara Hrvatske	4	Inglês
19	Slavic (2006c)	UDC in subject gateways: experiment or opportunity?	Artigo de periódico Knowledge Organization	15	Inglês
20	Frâncu (2007)	Does convenience trump accuracy? The avatars of the UDC in Romania	Artigo de periódico Extensions and Corrections to the UDC	0	Inglês
21	Gnoli (2007)	Progress in synthetic classification: towards unique definition of concepts	Artigo de periódico Extensions and corrections to the UDC	9	Inglês
22	Pika (2007)	Universal Decimal Classification at the ETH-Bibliothek Zürich-a Swiss perspective	Artigo de periódico UDC Consortium	3	Inglês
23	Slavic, Cordeiro e Riesthus (2007)	Enhancement of UDC data for use and sharing in a networked environment Item type Conference Paper Enhancement of UDC data for use and sharing in a networked environment	Artigo de evento The 31st Annual Conference of the German Classification Society on Data Analysis, Machine Learning, and Applications	2	Inglês
24	La Barre (2007)	The heritage of early FC in document reference retrieval systems, 1920-1969	Artigo de periódico Library History	13	Inglês
25	Colillas e Nicolau (2007)	Aplicación de un nuevo sistema de indización en una colección de recursos especializados en ciencias de la educación	Artigo de periódico El profesional de la información	6	Espanhol

26	Broughton e Slavic (2007)	Building a faceted classification for the humanities: principles and procedures	Artigo de periódico Journal of Documentation	51	Inglês
27	Hjørland (2007b)	Arguments for 'the bibliographical paradigm': Some thoughts inspired by the new English edition of the UDC	Artigo de evento The Sixth International Conference on Conceptions of Library and Information Science-"Featuring the Future"	26	Inglês
28	Colillas e Nicolau (2008)	Improving subject searching in databases through a combination of descriptors and UDC	Artigo de periódico Zadar: Bobcatsss	4	Inglês
29	Kaosar (2008)	Merits and Demerits of using Universal Decimal Classification on Internet	Dissertação	1	Inglês
30	Slavic (2008)	Faceted classification: management and use	Artigo de periódico Axiomathes	42	Inglês
31	Slavic, Cordeiro e Riesthuis (2008)	Maintenance of the Universal Decimal Classification: overview of the past and preparations for the future	Artigo de periódico International Cataloguing and Bibliographic Control Journal	14	Inglês
32	Dahlberg (2008)	The Information Coding Classification (ICC): A modern, theory-based fully-faceted, universal system of knowledge fields	Artigo de periódico Axiomathes	24	Inglês
33	Balíková (2009)	The role of UDC classification in the Czech Subject Authority File	Artigo de periódico Extensions & Corrections to the UDC	0	Inglês
34	Colillas (2009)	Tratamiento de los Metadatos de Contenido en la Web Semántica o cuando la Lógica se hace Evidente	Artigo de evento ISKO	3	Espanhol
35	Hajdu-Barát (2009)	Integration of a thesaurus and Universal Decimal Classification (UDC) to improve subject access: the Hungarian experience	Artigo de periódico Extensions & Corrections to the UDC	0	Inglês
36	San Segundo (2009)	Crossroads: Multiple Directions to Usability Using MARC classification format for UDC and mappings to other KO systems for an enriched authority file	Artigo de periódico Extensions & Corrections to the UDC	4	Inglês

37	Broughton (2010)	Concepts and terms in the faceted classification: the case of UDC	Artigo de periódico Knowledge Organization	21	Inglês
38	Frâncu e Sabo (2010)	Implementation of a UDC-Based Multilingual Thesaurus in a Library Catalogue: The Case of BiblioPhil	Artigo de periódico Knowledge Organization	7	Inglês
39	Doorn e Polman (2010)	From classification to thesaurus ... and back? Subject indexing tools at the library of the Afrika-Studiecentrum Leiden	Artigo de periódico Knowledge Organization	3	Inglês
40	McIlwaine (2010)	Classification at a Crossroads: Multiple Directions to Usability	Artigo de periódico International Cataloguing and Bibliographic Control: Quarterly Bulletin of the IFLA UBCIM Programme	0	Inglês
41	Colillas (2011)	UDC on the Internet: Theory and project in evolution for use of indexing and retrieval systems	Artigo de periódico IFLA Journal	3	Inglês
42	Vukadin e Slavic (2014)	Challenges of Facet analysis and Concept Placement in Universal Classifications: the Example of Architecture in UDC	Artigo de periódico Knowledge Organization	0	Inglês
43	Lee (2017)	Numbers, instruments and hands: the impact of faceted analytical theory on classifying music ensembles	Artigo de periódico Knowledge Organization	4	Inglês
44	Slavic e Davies (2017)	Facet Analysis in UDC: Questions of Structure, Functionality and Data Formality	Artigo de periódico Knowledge Organization	0	Inglês

FONTE: Dados da pesquisa (2020).

Além dos dados descritos no Quadro 2, foram analisados os veículos de publicação (artigos de periódicos, anais de congresso, teses e dissertações), número de ocorrências de palavras-chave, afiliação, país da publicação e categorização dos estudos por semelhança. A amostra analisada foi composta de acordo com os critérios de elegibilidade, cujos trabalhos foram lidos, na íntegra, onde o resultado da seleção dos estudos desta revisão demonstrou uma busca exaustiva nas bases de dados. Essas atividades possibilitaram realizar o mapeamento da literatura, com estudos publicados de 1964 a 2017.

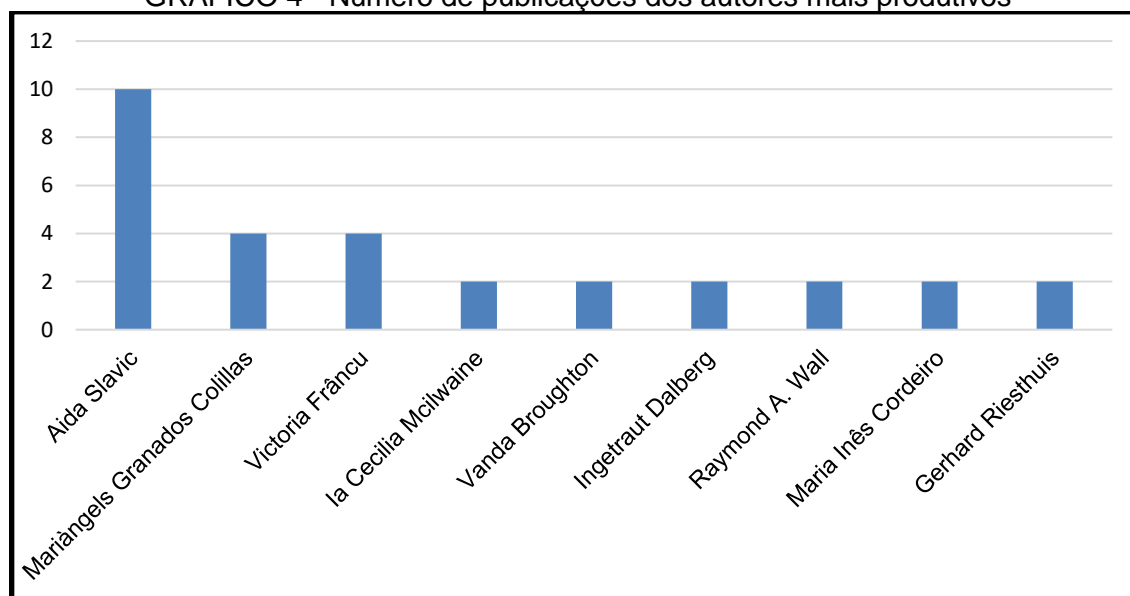
2.3.1 Análise quantitativa da RSL

Logo após a coleta, foram realizadas análises quantitativas e qualitativas do *corpus* da pesquisa por meio do uso de planilhas eletrônicas utilizando o software *Microsoft Excel*. Mediante o uso desses instrumentos, foi possível cruzar as informações, filtrá-las e apresentar os dados em gráficos. Além dos dados quantitativos constantes no quadro anterior, foram analisadas outras categorias, as quais serão apresentadas a seguir.

Com o objetivo de identificar os autores que mais publicaram sobre o tema desta pesquisa, foi realizada a contagem dos autores e coautores, sendo atribuída uma unidade para cada um deles em cada trabalho publicado. Observou-se um total de 38 autores, sendo que 29 autores tiveram apenas uma publicação no período de 1964 a 2017, o que corresponde a 76,31% do total de autores.

O Gráfico 4 apresenta quais foram os autores que tiveram mais de um estudo publicado no período de 1964 a 2017. Os nove autores com maior número de publicações representam 23,67% do total de autores selecionados, sendo que seis autores tiveram apenas duas publicações; dois autores publicaram quatro; e um autor publicou dez, o que caracteriza 15,78%, 5,26% e 2,63%, respectivamente. O total de publicações realizadas pelos autores que menos publicaram sobre o tema foram 29 trabalhos, o que correspondem a 65,9% do total das 44 publicações selecionadas.

GRÁFICO 4 - Número de publicações dos autores mais produtivos

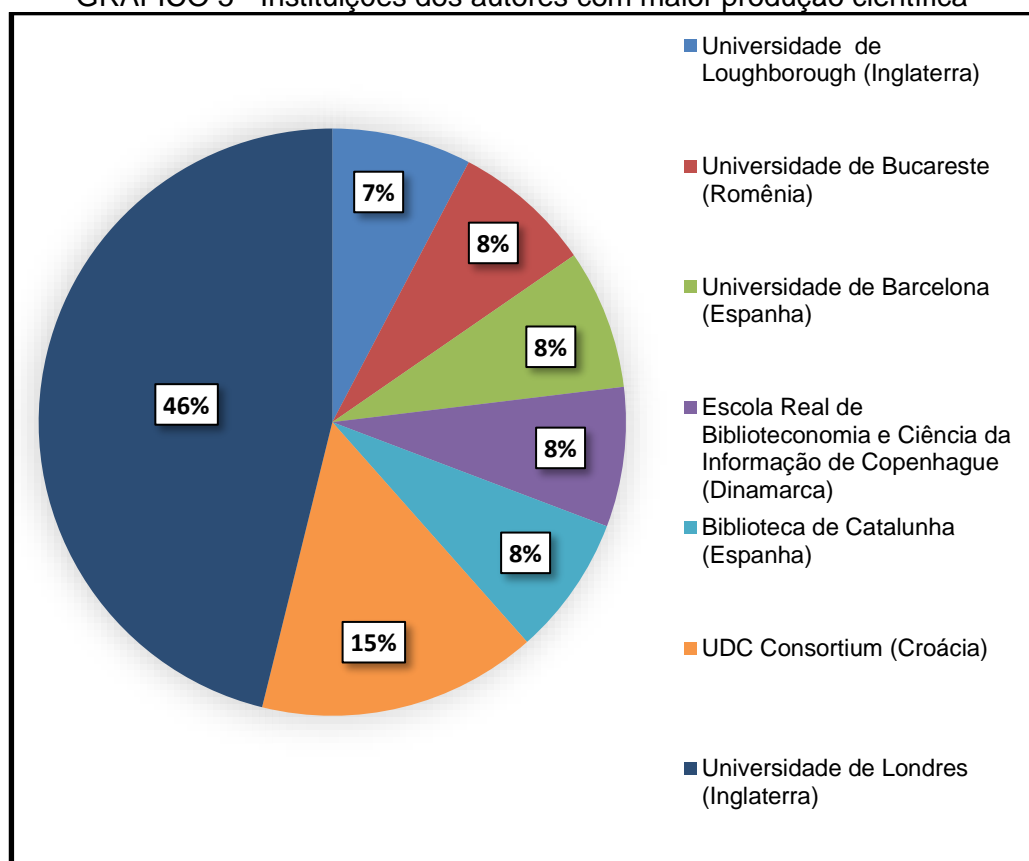


FONTE: Dados da pesquisa (2020).

No Gráfico 5, são apresentadas as instituições vinculadas aos autores que mais publicaram sobre a temática. Analisou-se um total de 26 publicações, cujos autores mais produtivos tiveram 2, 4 e 12 publicações. Os autores mais produtivos que tiveram 12 publicações, o que corresponde a 45% do total de publicações e estão vinculados à

Universidade de Londres (Inglaterra); quatro estão vinculados a *UDC Consortium* (Croácia), o que categoriza 15%. Por fim, os últimos dez autores estão vinculados a cinco instituições: Universidade de Loughborough (Inglaterra), Universidade de Bucareste (Romênia), Universidade de Barcelona (Espanha), Escola Real de Biblioteconomia e Ciência da Informação de Copenhague (Dinamarca) e a Biblioteca de Catalunha (Espanha), o que equivale ao percentual de 8% em cada instituição, totalizando 40% das instituições que tiveram dois autores vinculados a cada uma. O país que mais publicou sobre o tema foi a Inglaterra com 14 publicações; os demais países apresentam uma quantidade menor de trabalhos, distribuídos na seguinte proporção: Croácia (quatro trabalhos), Espanha (quatro), Romênia (dois) e a Dinamarca (dos estudos).

GRÁFICO 5 - Instituições dos autores com maior produção científica



FONTE: Dados da pesquisa (2020).

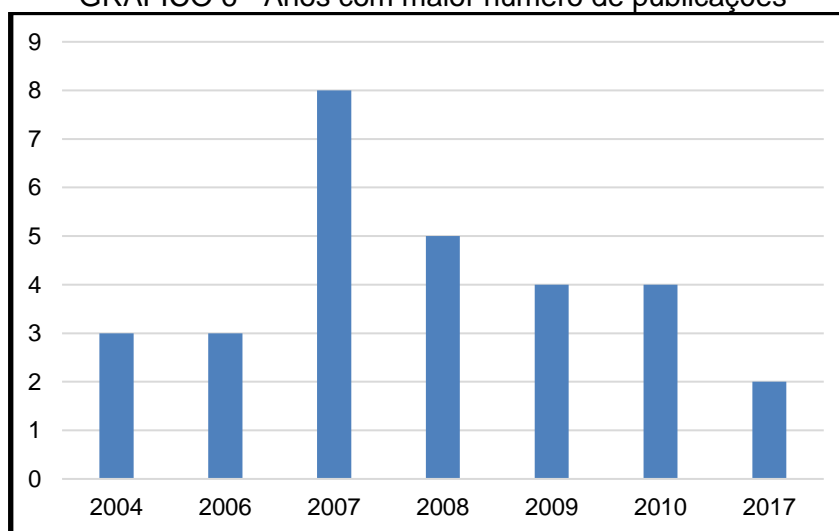
O idioma predominante foi o inglês com um total de 42 publicações e com apenas dois trabalhos em espanhol. O inglês representou quase o total de publicações do período, caracterizando 95,45%, o espanhol 4,54% e não tiveram trabalhos em português.

A avaliação do grau de relevância das publicações foi realizada com base nas 13 publicações que tiveram mais que 15 citações no Google Acadêmico, o que indica 29,54% do total das publicações selecionadas. No trabalho de Mcilwaine (1997), destacaram-se 61 citações, sendo a publicação mais citada sobre o tema desta pesquisa. Outro ponto relevante dessa análise foram os trabalhos mais citados da amostra defendidos pela autora Aida Slavic

(2006a, 2006c, 2008), os quais receberam 15, 15 e 42 citações, respectivamente; e o trabalho de Broughton e Slavic (2007), que recebeu 51 citações. Por meio dessa análise, observou-se que Aida Slavic é uma das principais autoras a estudar e pesquisar sobre a temática em análise.

O Gráfico 6 mostra o total de publicações por ano no período de 2004 a 2017, cujo critério de elegibilidade foram os anos que tiveram mais de três publicações e a comparação com o último ano de publicação que foi em 2017. Foi constatado um aumento de 11,36% dos anos de 2004 e 2006 para 2007, uma redução de 6,81% para 2008 e de 9,09% para os anos de 2009 e 2010. O ano que obteve o maior número de publicações foi o ano de 2007 com 8 trabalhos, o que corresponde a 27,58% do total das publicações. No último ano da amostra, houve uma redução de 4,54% em relação ao ano de 2010 do total da produção científica sobre o tema no período. A diminuição publicações em 2017 em relação aos anos anteriores precisa ser melhor investigado uma vez que pode trazer prejuízos à temática.

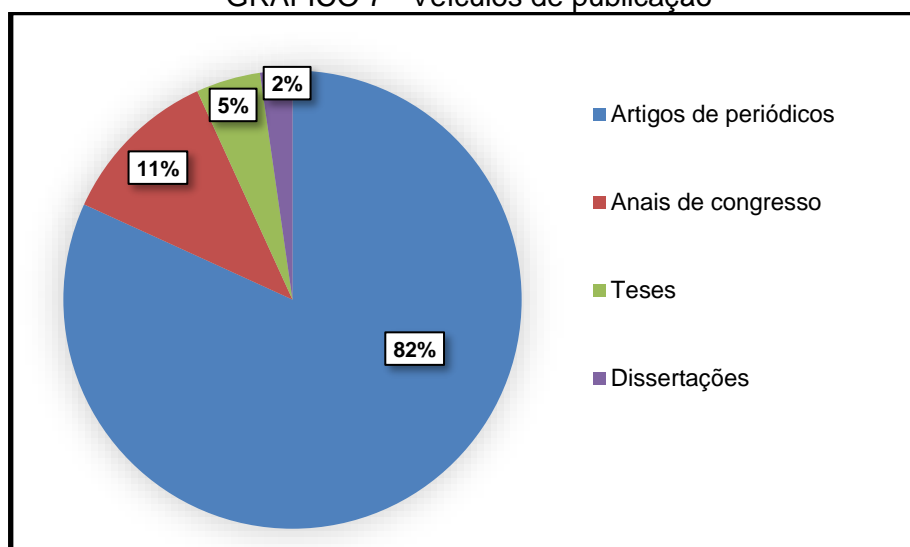
GRÁFICO 6 - Anos com maior número de publicações



FONTE: Dados da pesquisa (2020).

Quanto aos meios de publicação (artigos de periódicos, anais de congresso, teses e dissertações), os resultados revelam um total de 36, 5, 2 e 1, respectivamente. O Gráfico 7 apresenta as respectivas porcentagens, para esses resultados, considerando o total das 44 publicações selecionadas. Verificou-se que a maioria dos trabalhos foram publicados em artigos de periódicos e anais de congresso com um total respectivo de 36 e cinco trabalhos, o que representa 82% e 11% do total da produção científica analisada.

GRÁFICO 7 - Veículos de publicação



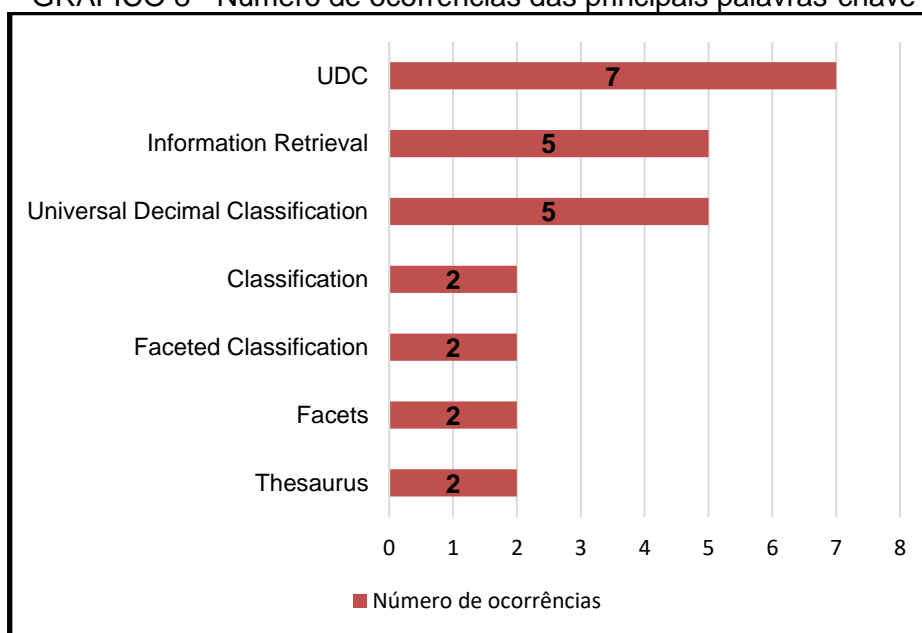
FONTE: Dados da pesquisa (2020).

Considerando que os periódicos foram os meios de publicação mais recorrentes, notou-se, também, que o periódico mais produtivo em relação ao tema foi o “*Extensions & Corrections to the UDC*”, com sete publicações, o que corresponde a 15,55% do total das publicações. O principal motivo pelo maior número de publicações deve-se ao envolvimento da autora Aida Slavic com o *UDC Consortium*, que publica pesquisas relacionadas a estrutura da CDU e suas respectivas alterações ou manutenções. A autora é responsável pelo desenvolvimento e manutenção da CDU, e também por ser a editora-chefe da publicação oficial desse periódico. Outros dois periódicos que mais publicaram sobre a temática foram o *Journal of Documentation e Knowledge Organization*, com cinco e oito publicações, o que corresponde a 11,36% e 18,18% do total de publicações.

Para verificar quais foram as principais palavras-chave inseridas nas publicações, foi contabilizado o número de ocorrências de cada palavra-chave, sem considerar as publicações que não apareceram as palavras-chave no próprio texto. O Gráfico 8 mostra as sete palavras-chave que ocorreram mais de uma vez nas publicações, com o total de 25 trabalhos, sendo que as palavras-chave que apareceram apenas uma vez tiveram no total 53 ocorrências. Foram destacadas as seguintes palavras-chave que tiveram o maior número de ocorrências: “*UDC*”, “*Information retrieval*”, “*Universal Decimal Classification*”, “*Classification*”, “*Faceted Classification*”, “*Facets*” e “*Thesaurus*”.

Das palavras mencionadas, a primeira obteve sete ocorrências, a segunda e terceira cinco e da quarta até a sétima, duas ocorrências, o que corresponde a 11,66%, 8,33% e 3,33%, respectivamente. Outra análise confirmou que a palavra-chave “Tesauro” apareceu tanto no plural (“*Thesauri*”) quanto no singular (“*Thesaurus*”) com uma ocorrência de cada uma. Sendo assim, definiu-se utilizar o termo no singular durante a análise quantitativa.

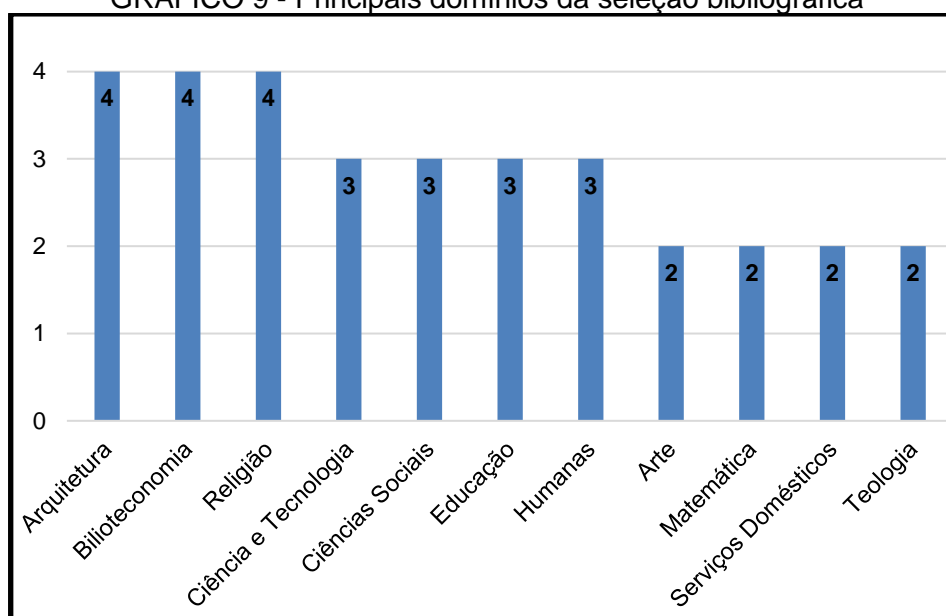
GRÁFICO 8 - Número de ocorrências das principais palavras-chave



FONTE: Dados da pesquisa (2020).

Foram levantados durante esta revisão da literatura, quais as principais áreas do conhecimento em que houveram experimentos realizados em cada trabalho selecionado. O Gráfico 9 apresenta quais foram esses domínios, sendo que mais estudados foram: Arquitetura, Biblioteconomia, Religião, Ciência e Tecnologia, Ciências Sociais, Educação e Humanas.

GRÁFICO 9 - Principais domínios da seleção bibliográfica



FONTE: Dados da pesquisa (2020).

Considerando o domínio de conhecimento delimitado para aplicação metodológica desta tese, qual seja, a agricultura, pelas informações apresentadas no Gráfico 9, é possível verificar que não há, dentre os trabalhos selecionados, estudo aplicado somente para esse

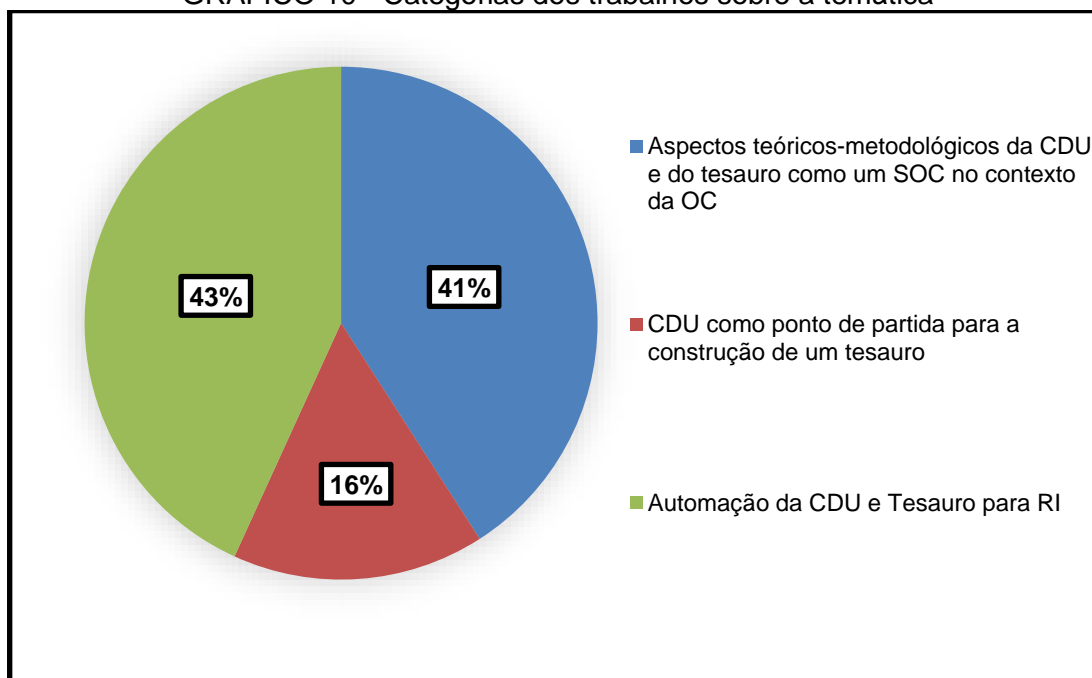
domínio. Dentre os 44 trabalhos, 32 tiveram dois ou mais estudos, tanto em instituições internacionais ou nacionais. Observou-se no momento, que nenhum autor com vínculo em uma instituição brasileira estudou o tema. Por esse motivo, pode-se comprovar que as técnicas e abordagens desta temática ainda sejam incipientes no Brasil em relação às publicações internacionais, o que torna esta pesquisa inédita para aplicação de um assunto específico do domínio da agricultura.

Após a análise dos 44 trabalhos sobre a utilização da CDU com o auxílio de um tesouro para a RI, foi possível definir a categorização dos temas abordados em cada trabalho. Fez-se necessário agrupar os temas em três categorias, mencionando a quantidade de trabalhos pertencentes à categoria, bem como as autorias e os respectivos anos da publicação. A distribuição foi definida, assim:

- **Aspectos teórico-metodológicos da CDU e do tesouro como SOC no contexto da organização do conhecimento:** esse tema foi abordado do ponto de vista teórico, sem aplicação na prática, havendo 18 trabalhos, o que corresponde a 41% da amostra com essa abordagem, a saber: (MILLS,1970; DAHLBERG, 1971; FRÂNCU, 2003a; FRÂNCU, 2003b; BROUGHTON; SLAVIC, 2007; GNOLI, 2007; HJØRLAND, 2007a; LA BARRE, 2007; PIKA, 2007; SLAVIC; CORDEIRO; RIESTHUIS, 2007; COLLILAS; NICOLAU, 2008; KAOSAR, 2008; HAJDU-BARÁT, 2009; BROUGHTON, 2010; MCILWAINE, 2010; VUKADIN; SLAVIC, 2014; LEE, 2017; SLAVIC; DAVIES, 2017);
- **CDU como ponto de partida para a construção de um tesouro:** os trabalhos abordaram as notações da CDU sem modificar sua estrutura original, inserindo termos do tesouro quando não existirem na CDU e uma notação para fazer a inserção das relações semânticas do tesouro com os termos da CDU. Sete trabalhos relataram esse aspecto (16%): (WALL, 1973; WALL, 1980; CHAN; ZENG, 2002; SLAVIC, 2006b; SAN SEGUNDO, 2009; DOORN; POLMAN, 2010; FRÂNCU; SABO, 2010);
 - **Automação da CDU vinculada a um tesouro para RI:** foi mencionado que se pode utilizar a estrutura original da CDU, com tesouros de diferentes domínios, cujo objetivo seja representar o conteúdo dos documentos. Nessa categoria, foram alinhados 19 trabalhos (43%): (FREEMAN, 1964; HINDSOTI, 1979; FIAŁKOWSKI; MURASZKIEWICZ; TOŁWIŃSKA, 1981; BUXTON, 1990; MCILWAINE, 1997; CHUDAMANI, 2004; SCHALLIER, 2004; SLAVIC, 2004a; SLAVIC, 2005; SLAVIC, 2006a; SLAVIC, 2006c; COLILLAS; NICOLAU, 2007; FRÂNCU, 2007; DAHLBERG, 2008; SLAVIC; CORDEIRO; RIESTHUIS, 2008; SLAVIC, 2008; BALÍKOVÁ, 2009; COLILLAS, 2009; COLILLAS, 2011).

O Gráfico 10 exhibe o enfoque das categorias temáticas que mais foram estudadas no âmbito do uso da CDU, com utilização de tesouro, para facilitar a RI em ambientes digitais.

GRÁFICO 10 - Categorias dos trabalhos sobre a temática



FONTE: Dados da pesquisa (2020).

No intuito de detalhar o conteúdo de cada um dos 44 trabalhos analisados na amostra desta revisão, procedeu-se à identificação e descrição do objetivo, método, resultado e conclusão explanados em cada um dos estudos. Esse procedimento foi facilitado com o uso de uma ficha de resumo, com preenchimento dos campos para cada elemento a ser extraído da publicação. Portanto, no subcapítulo seguinte, realiza-se a apresentação dos resultados desses aspectos, o que possibilita compreender melhor os detalhes de cada uma das pesquisas que versam sobre a temática desta tese.

Após finalizar a etapa de análise dos dados quantitativos, na 6ª etapa ampliaram-se as *strings* de buscas inserindo em três idiomas, português, inglês e espanhol no Google Acadêmico e nas cinco bases de dados bibliográficas mencionadas na segunda etapa, aumentando as fontes de busca com o objetivo de aumentar o número de publicações a serem incluídas de forma exaustiva, com o objetivo de diminuir a diferença entre os percursos ou métodos de pesquisa realizados para obter os resultados que mais se aproximam da temática.

2.3.2 Análise qualitativa dos resultados da aplicação da RSL

Na 7ª etapa, conforme as recomendações mencionadas por Higgins e Green (2011), foram selecionados 44 estudos que cumpriram os critérios estabelecidos como objeto de estudo desta pesquisa, sendo agrupados de acordo com a natureza e semelhança de cada um. A síntese dos resultados dos estudos obtidos na revisão da literatura é apresentada em ordem cronológica, iniciando-se da publicação mais antiga até a mais recente.

O primeiro trabalho selecionado na literatura foi desenvolvido por Freeman (1964), intitulado “*Computers and Classification Systems*”, que apresentou a estrutura da CDU, porém com a ordem de armazenamento diferente da estrutura original. A metodologia mostrou que os caracteres dos dados poderiam ser substituídos por outros caracteres ou ordenados de outro modo. Nesse sistema, a busca pode ser realizada apenas por assunto, utilizando uma terminologia exata em linguagem natural e por índices de assuntos, sendo por ordem alfabética ou catálogos de documentos em um SRI. Caso algum assunto mais específico não for reconhecido, o sistema seleciona o assunto mais genérico para não trazer resultados vazios. Os resultados da pesquisa apontaram que a notação classificatória da CDU pode ser uma ferramenta bastante eficiente em sistemas de informação, principalmente para mostrar as relações hierárquicas utilizando um tesouro (FREEMAN, 1964). Esse aspecto também pode ser observado na pesquisa de Hindsoti (1979).

Mills (1970) apresentou o artigo “*Progress in Documentation - Library Classification*”, propondo uma classificação moderna para biblioteca, baseada em índices, com uso da CDU associada a um tesouro, considerado como um sistema pós-coordenado que auxilia na RI. Essa metodologia baseou-se na criação de um índice na CDU para os termos de um tesouro, levando em consideração os sinônimos, termos gerais e específicos e as relações entre eles, cujo assunto representado deve estar em um único lugar e ser atribuída uma única notação. Todas as combinações necessárias são possíveis, respeitando a estrutura original da CDU com a descrição das regras de combinação separadas. Os resultados mostraram a consistência interna e a eficácia da RI, evidenciando que a CDU e o tesouro devem estar explicitamente conectados em uma base única, de acordo com a definição dos conceitos relacionados (MILLS, 1970). Os mesmos procedimentos metodológicos foram aplicados, anos depois, por Wall (1973) e Wall (1980).

No artigo “*Possibilities for a new universal decimal Classification*”, Dahlberg (1971) realizou a realocação de assuntos das classes principais da CDU, de maneira que todos os conceitos e objetos existentes possam ser integrados em um sistema. Os procedimentos metodológicos utilizados foram por meio de um sistema estruturado logicamente para facilitar a utilização e leitura pelo computador. Além disso, o sistema possibilita subdividir os assuntos, prevenindo omissões de assuntos sem prejudicar a notação. A partir do princípio da análise facetada existente na CDU, foi possível incluir, na nova classificação, um maior número de categorias para contemplar os assuntos e vocabulários existentes. Desse modo, a CDU pôde ser integrada a um sistema de recuperação, permitindo que seja utilizada pelos usuários com um tesouro dos assuntos em um domínio. Os resultados demonstraram que um sistema de classificação pode fazer a união entre a CDU e um tesouro, bem como pode fornecer aos bibliotecários e documentalistas um sistema de classificação que possua várias categorias de assuntos (DAHLBERG, 1971). Dahlberg (2008) e Hjørland (2007b) também concluíram que a

metodologia utilizada por Dahlberg (1971) fornece a classificação de assuntos das disciplinas específicas desenvolvidas por meio da utilização de tesouros e sistemas de classificação, conectando todas as disciplinas do conhecimento humano com um vocabulário controlado.

No trabalho de Fiałkowski, Muraszkievicz e Tołwińska (1981), intitulado “*Methodology of world clearinghouse activity*”, estabeleceram-se os princípios e a metodologia da atividade de *Clearing house* de Varsóvia, que é um serviço bibliográfico especializado para coletar, processar e disseminar informações sobre tesouros, sistemas de classificação, descritores, termos e listas de cabeçalho de assunto. Na recuperação, é fornecido o número e termos da CDU, que estão armazenadas em uma base de dados. Como resultado, foi necessário automatizar a atividade da *Clearing house*, em particular a RI no sistema *Thesauri e CLAssification* (TEKLA). Tal sistema processa a bibliografia e descritores de tesouros, esquemas de classificação, palavras-chave, categoria de assuntos dos documentos, idioma e ano de publicação, sendo esses elementos armazenados e processados pelo sistema (FIAŁKOWSKI; MURASZKIEWICZ; TOŁWIŃSKA, 1981).

O estudo de Buxton (1990) intitulado “*Computer searching of UDC numbers*” desenvolveu um SRI por meio de catálogos on-line, no projeto chamado *Automatic Direct Access to Information with the UDC On-line* (Audacious), para o sistema de informação da biblioteca ETH (Ética). O autor desenvolveu um método, cujo sistema utiliza apenas os números da CDU, com uma estrutura pouco aproveitada demandando grande esforço do classificador. Por outro lado, a vantagem do sistema é que a notação baseia-se em uma estrutura hierárquica. Com isso, os dados são mantidos no SRI utilizando um tesouro on-line, para buscar informações além da estrutura embutida na notação, de modo a controlar a sinonímia e garantir os termos relacionados. A princípio, utilizou-se um arquivo de sinônimos que foi vinculado aos números da CDU, associado com os relacionamentos semânticos entre descritores que podem ser ajustados sem levar em conta as relações embutidas na estrutura ou notação da CDU. Nos resultados, verificou-se que a biblioteca utilizou descritores adicionais e alterar as relações entre eles, sendo inviável incluir todas as possíveis combinações de termos entre os números das tabelas principais e das tabelas auxiliares no arquivo de autoridade do registro bibliográfico (BUXTON, 1990).

Na literatura, um estudo semelhante aos resultados de Fiałkowski, Muraszkievicz e Tołwińska (1981) é o desenvolvido por McIlwaine (1997), intitulado “*The Universal Decimal Classification: Some Factors Concerning Its Origins, Development, and Influence*”. Nos procedimentos metodológicos utilizados na citada pesquisa, recorreu-se ao uso dos números e termos da CDU, em vez de utilizar um sistema de recuperação. Os resultados revelaram que as pesquisas por termos mostram que o mesmo termo pode estar em mais de um número de classificação da CDU, pois representam diferentes aspectos sobre um tema. Portanto, quanto menor o número no sistema de classificação, menor a quantidade de locais que estará

os termos da CDU. Nesse sentido, com o advento da automação, é possível desenvolver um esquema eficiente na RI, que não seja apenas a simples ordenação de documentos nas estantes.

No trabalho de Chan e Zeng (2002), com o título “*Ensuring interoperability among subject vocabularies and knowledge organization schemes: a methodological analysis*”, foi apresentada a interoperabilidade de diferentes Sistemas de Organização do Conhecimento, como vocabulários controlados, sistemas de classificação e tesouros. Verificou-se que todos os instrumentos podem ser interoperáveis como se estivessem todos dentro de um único sistema. Desse modo, criou-se um esquema único de classificação baseado na CDU, mapeando termos existentes aplicado a uma estrutura de tesouro adicional para cada assunto específico. Ademais, foi possível criar mapas mentais do usuário que utilizam termos representativos para otimizar a recuperação. Os resultados alcançados foram a necessidade de interoperabilidade entre vocabulários multilíngues na organização do conhecimento em sistemas, sendo necessário construir uma correspondência (*matching*) entre os vocabulários envolvidos.

Esse método também foi proposto por Schallier (2004), com o uso da CDU em um *On-line Public Access Catalog* (OPAC), para a recuperação dos assuntos dos títulos da Biblioteca do Congresso e da *Medicine Subject Headings* (MeSH). Ademais, o trabalho de San Segundo (2009) também utilizou o mesmo método que foi chamado de metatesouro da Biblioteca Nacional de Medicina (UMLS) da *Unified Medical Language System*, nos Estados Unidos, com a criação de vocabulários de cabeçalhos de assunto, descritores de tesouros e sistemas de classificação interoperáveis com técnicas baseada em unidades lexicais.

No ano de 2003, foi apresentada a tese “*Multilingual access to information using an intermediate language*” (FRÂNCU, 2003a) e publicado o artigo “*The impact of specificity on the retrieval power of a UDC-based multilingual thesaurus*” (FRÂNCU, 2003b). Os estudos objetivaram explorar os impactos da especificidade no potencial de recuperação de um tesouro multilíngue baseado na CDU com indicadores de desempenho, tais como: revocação, precisão e relevância. A metodologia adotada visou atender os requisitos de uma ferramenta de RI altamente eficiente, capaz de satisfazer as necessidades do usuário, com alto desempenho de pesquisa e as expectativas do indexador, respeitando as regras de classificação. Foi utilizado a compatibilidade de linguagem com a CDU, o primeiro tesouro com a notação e descrição no tesouro LTHES, que é um tesouro multilíngue interdisciplinar nos idiomas: inglês, francês e romeno, baseado na versão resumida da CDU com baixo nível de especificidade. O segundo tesouro utilizado foi o PTHES (*Pocket Edition*), que também é um tesouro multilíngue nesses idiomas, que foi compatibilizado com a CDU, porém como este tesouro possui um nível maior de detalhes, obteve resultados melhores na RI em relação ao primeiro. A diferença entre os dois é que o tesouro LTHES permite a busca de informações

apenas para assuntos amplos, enquanto o PTHES permite pesquisas mais refinadas, resultando em mais precisão e, portanto, maior relevância dos resultados durante a busca. Em suma, pelos resultados desses estudos, constatou-se que a utilização de tesouros com maior nível de detalhamento, permitindo aos usuários ampliar os resultados obtidos em suas pesquisas e com o tesouro com baixo nível de especificidade para restringir os resultados durante a busca, conforme as necessidades de informação dos usuários. Por se tratar de três SOC e com diferentes idiomas, utilizou-se uma linguagem intermediária, permitindo a compatibilidade entre a CDU e os dois tesouros, para que o usuário tivesse a oportunidade de ampliar e restringir a busca de acordo com a especificidade de cada tesouro, suprimindo suas necessidades, funcionando como dispositivo de desambiguação. O mesmo padrão de pesquisa também foi descrito no trabalho de Frâncu (2007) e Frâncu e Sabo (2010). Os trabalhos de Kaosar (2008), Mcilwaine (2010), Pika (2007), Slavic, Cordeiro e Riesthuis (2007), basearam-se somente em aspectos multilíngue, sem a utilização de linguagem intermediária.

Slavic (2004a) publicou o artigo intitulado “*UDC implementation: from library shelves to a structured indexing language*”, mostrando que algumas bibliotecas estão construindo, gradualmente, sistemas com tesouros ou títulos de assuntos baseados nos dados existentes na CDU. Os métodos utilizados foram aplicados em usuários fora do domínio da biblioteca que gerenciam sistemas de classificação com dados disponíveis em arquivos de autoridade de assunto da CDU, disponíveis para compartilhamento. Os resultados mostraram que as ferramentas de organização do conhecimento devem ser tratadas como dados independentes e separados do próprio registro dos metadados. Isso se deve por tratar de termos disponíveis para vários idiomas e serem pré-coordenados, visto que, geralmente, perdem os relacionamentos de sintaxe se não forem codificados. Por isso, a CDU é recomendada como um sistema padrão de organização do conhecimento para o uso em vários padrões de metadados, que permitem a codificação do esquema no qual o termo também pode ser usado com o uso de *Uniform Resource Identifiers* (URI). Nesse caso, a URI deve ser usada quando os dados de classificação são disponibilizados na web e sejam acessíveis por diferentes aplicativos, possibilitando um grande esforço para garantir a permanência desses identificadores e permitir que os dados dos assuntos forneçam formas eficazes para expressar novos conceitos, garantindo a integridade e a consistência na recuperação (SLAVIC, 2004a). A mesma metodologia pode ser observada em Balíková (2009).

A tese de Slavic (2005), com título “*Classification management and use in a networked environment: the case of the Universal Decimal Classification*”, e o artigo “*Faceted classification: management and use*” (SLAVIC, 2008), descreveram as preocupações existentes no design, construção e uso da classificação. Percebeu-se que as qualidades do

conteúdo ou lógicas da estrutura do sistema dependem das características estruturais e funcionais para que a Classificação Facetada seja utilizada de modo eficiente na recuperação. Nos procedimentos metodológicos, definiram-se alguns requisitos funcionais comuns às classificações analítico-sintéticas, como, por exemplo, a utilização do padrão do vocabulário. Os resultados demonstraram que nenhum desses padrões atende aos requisitos típicos de linguagens de indexação pré-coordenada e não possuem solução para expressar relacionamentos sintagmáticos. Esses padrões não resolvem o problema das relações entre assuntos de uma classificação analítico-sintética, o que exige um grande esforço intelectual dos classificionistas e grupos de trabalho dentro de padrões específicos. O formato do banco de dados da CDU proposto é uma sugestão para um modelo de dados mais robusto, haja vista facilitar e contribuir para o uso de um sistema baseado na CDU em ambiente web (SLAVIC, 2005, 2008). Esse mesmo método foi utilizado por Chudamani (2004), Slavic (2006a), Gnoli (2007), La Barre (2007) e Broughton (2010). Nesses estudos, destacaram-se os sistemas que fazem parte do projeto *Faceted Analytico-Synthetic Theory* (FAST): o *Thesauro Facet* e o Tesouro da Unesco, que discute o uso da análise facetada na construção de vocabulários controlados. Pelos resultados desses estudos, percebe-se que a CDU continua sendo um sistema de classificação altamente utilizada pela University College London no desenvolvimento do projeto FATHUM (*Facet Analytic Theory – Humanities*).

Slavic (2006b) utilizou a CDU em uma interface de portais de assunto, verificando o impacto do uso da classificação na web. O método utilizado foi o mesmo do sistema *German Harvest Automated Retrieval and Directory* (GERHARD), por meio de um banco de dados com classificação automática e um serviço integrado de busca e navegação em um diretório de assunto. Os metadados gerados e o índice de documentos foram processados em um banco de dados relacional (ORACLE), e cada um dos termos de uma classe da CDU foi hierarquicamente exibido e representado com uma legenda de classe curta, levando em conta a precisão na recuperação. Os resultados verificaram que as bibliotecas reconhecem e usam as classificações principalmente com o propósito de organização das coleções. Quando usada no ambiente web, a classificação possibilita o controle de vocabulário e as relações semânticas entre assuntos relevantes para a descoberta de recursos, utilizando ou não sua notação (SLAVIC, 2006b). Os mesmos aspectos foram considerados no estudo de Slavic (2006c).

Broughton e Slavic (2007) também exploraram o potencial da CDU com a Teoria da Análise Facetada (TAF) associada a um tesouro, o que permitiu o desenvolvimento futuro de tesouros. A abordagem adotada para a CDU foi mais integrada à estrutura conceitual, utilizando tabelas auxiliares, além das “facetadas” de tempo, lugar, forma e suas subfacetadas tais como língua, que abrangem as tabelas, geralmente, aplicadas a pessoas, materiais, propriedades, processos e operações, as quais, ao mesmo tempo, removem exemplos das

classes principais. A metodologia utilizada foi a partir das combinações de facetas utilizando símbolos adicionais que irão separar e indicar a natureza do relacionamento durante a indexação para facilitar a pesquisa utilizando uma interface de navegação entre as facetas. O aspecto analítico-sintético da CDU forneceu uma estrutura facetada para cada assunto utilizando o PMEST em diferentes níveis, sendo uma notação mista, utilizando dígitos e números de classes da CDU para conectar os assuntos. A notação da CDU com dois pontos e outros símbolos permitiu a combinação dos números da classe. Os resultados mostram a especificação da estrutura dos dados, sua semântica, estrutura de nomeação, listas de termos controlados, revelando como esses elementos são interpretados e relacionados em uma ferramenta de banco de dados, sendo possível a representação de múltiplas hierarquias das classes da CDU, algo considerado como uma vantagem adicional na recuperação.

Colillas e Nicolau (2008) apresentaram um sistema que permite a organização dos conceitos baseada nos relacionamentos das disciplinas e não no alfabeto, com maior flexibilidade para combinação de assuntos. Os métodos escolhidos pelos autores foram baseados em um conjunto de termos equivalentes, dispensando o índice temático da CDU, com uso de links para que as classes da CDU fossem equivalentes com as notações e descritores, haja vista representar o conteúdo dos documentos. Para isso, foi necessário utilizar o tesouro com diferentes descritores para a indexação que se relacionam sem gerar incompatibilidade de vocabulário (COLILLAS; NICOLAU, 2008). Hajdu-Barát (2009) utilizou a mesma técnica, comparando as experiências de dois projetos húngaros, que visavam criar um sistema complexo para combinar a CDU e tesouro em um projeto de construção de tesouros na Biblioteca Nacional Húngara (Széchényi).

Slavic, Cordeiro e Riesthuis (2008) descreveram uma visão geral da história e revisão da CDU. O propósito do estudo foi entender a finalidade de direcionar a construção de futuros sistemas de classificação. O principal objetivo da política de manutenção e desenvolvimento da CDU, ao longo dos anos, é continuar o trabalho iniciado em 1993, de modo a atualizar o sistema. A metodologia possibilitou melhorar o controle sobre as mudanças, sua consistência em todo o sistema e a garantia de que os usuários sejam totalmente suportados no rastreamento e automação dessas mudanças no ambiente de uso. Os resultados mostraram que as tabelas da CDU podem conter assuntos simples, códigos do sistema com números compostos e o índice na forma de um tesouro. Assim sendo, foram despendidos esforços para fornecer o sistema com um índice alfabético e mapeamentos para outros sistemas de assuntos, a fim de melhorar a indexação e a recuperação (SLAVIC; CORDEIRO; RIESTHUIS, 2008).

Colillas (2009), por meio do estudo "*Tratamiento de los Metadatos de Contenido em la Web Semántica o cuando la Lógica se hace Evidente*", propôs uma lógica nos metadados, baseada na CDU, como ferramenta contextualizada nos cabeçalhos de assuntos,

concebidos como descritores que possuem um padrão lógico na atribuição para representar o conteúdo dos documentos. A proposta dessa metodologia foi permitir sucessivas seleções em três níveis de hierarquia com as classes e subclasses por meio de um menu, contendo as notações da CDU e a expressão alfabética. A navegação pode ser entre as classes principais da CDU ou pela tradução alfabética das notações, como *clusters* para agrupar todos os recursos relevantes usando os operadores booleanos. A busca pós-coordenada está mais associada à recuperação da linguagem natural, cujas linguagens controladas são mais associadas às linguagens pré-coordenadas, possibilitando maior flexibilidade na sintaxe em relação à busca pós-coordenada. Os resultados mostraram que a RI implicaria a existência de um tesouro multilíngue ou uma adaptação específica dos cabeçalhos de assunto da Biblioteca do Congresso Americano. Por fim, a gestão da CDU pode ser conduzida com um software que auxilie o funcionamento automático da recuperação a partir do filtro de todos os dados relativos aos índices incluídos no sistema. Esse processo pode ser aplicado no ambiente tradicional de bibliotecas e também nos motores de busca da web (COLILLAS, 2009).

Na pesquisa de Doorn e Polman (2010), intitulada “*From classification to thesaurus... and back? Subject indexing tools at the library of the Afrika-Studiecentrum Leiden*”, também foi estabelecida a concordância entre os códigos da CDU com um tesouro, atribuindo os códigos desse sistema com os títulos do catálogo com descritores do tesouro. Os procedimentos metodológicos demonstraram como um tesouro de estudos africanos foi construído visando à indexação e recuperação de assunto na biblioteca, em conjunto com um sistema baseado em termos, considerando uma alternativa mais amigável para os códigos da CDU. Os resultados demonstraram que, para cada descritor, foi atribuído um código da CDU, utilizando a edição mais recente do *Master Reference File* (MRF). Isso possibilitou substituir vários códigos antigos que eram utilizados para conectar a um tesouro em diferentes versões linguísticas do MRF da CDU (DOORN; POLMAN, 2010).

Segundo Colillas (2011), no trabalho intitulado “*UDC on the Internet: Theory and project in evolution for use of indexing and retrieval systems*”, há uma ordem lógica para os termos de indexação, representando o conteúdo principal do documento, com uso dos descritores da CDU de forma universal, tendo em vista garantir maior precisão na RI. A metodologia utilizou um sistema de indexação e recuperação baseado na CDU no projeto *Faceted Application of Subject Terminology* (FAST). Essa ferramenta possibilitou a criação de novas associações, baseadas na indexação e na classificação, que aparecem nos documentos pela adaptação do citado projeto em conjunto com a CDU. Os resultados geraram vários documentos reais na base de dados do catálogo da Biblioteca do Congresso indexados de acordo com a versão on-line da CDU no sistema proposto, para serem recuperados de forma multidisciplinar na base de dados (COLILLAS, 2011).

Em 2014, Vukadin e Slavic (2014) mostraram a adequação da CDU no artigo intitulado “*Challenges of Facet analysis and Concept Placement in Universal Classifications: the Example of Architecture in UDC*” para a indexação e recuperação de assuntos sobre a área da Arquitetura. O estudo mencionou que todos os conceitos fundamentais do domínio estão presentes no esquema e na estrutura hierárquica, cujas relações são bem representadas e apoiadas por princípios sintéticos que permitem a descrição do assunto de modo detalhado e com menos dependência, como acontece em uma estrutura hierárquica rígida. O método utilizado para fins de melhorar a recuperação foi os princípios da análise facetada na CDU, com o tesouro *Art & Architecture Thesaurus On-line* (AAT), em 2013. Esse instrumento constitui uma linguagem de indexação alfabética, preocupada com aspectos linguísticos do controle de vocabulário da CDU, característicos da organização por facetas. Os resultados apontaram a definição única de conceitos e a representação notacional, as quais permitem uma organização e uma recuperação mais previsíveis na CDU, com recurso analítico-sintético dos conceitos listados sobre o domínio da Arquitetura (VUKADIN; SLAVIC, 2014).

O ano das publicações mais recentes desta revisão foi 2017, com o estudo de Lee (2017) intitulado “*Numbers, Instruments and Hands: The Impact of Faceted Analytical Theory on Classifying Music Ensembles*”, que também utilizou a análise facetada para realizar uma análise aprofundada na área musical. A metodologia utilizou um número significativo de tipos de informações relacionadas à classificação na área Música. Foram considerados a flexibilidade e o caráter universal da CDU, incluindo, apenas, um pequeno número de subfacetas e seus relacionamentos, de acordo com a ordem de citação e modelos para a construção de um tesouro, haja vista facilitar a recuperação. Os resultados apresentaram os aspectos principais da Música, por meio da análise das facetas e das relações, sendo utilizadas para análises de outros assuntos. Assim, a utilização da análise facetada e das relações forneceu uma nova maneira de entender e reestruturar a classificação de conjuntos instrumentais, como também apresentou um novo caminho e um potencial para utilizar a análise facetada em outros assuntos complexos (LEE, 2017).

Slavic e Davies (2017) utilizaram a mesma técnica, por meio de requisitos de sistemas de informação para a modelagem de dados, com o objetivo de criar um projeto de bancos de dados e a interface de sistemas de classificação bibliográfica em ambientes on-line. A utilização da CDU possibilitou a garantia de um progresso apropriado, com as classes (entidades) exibidas de modo correto. As autoras evidenciaram que, dependendo dos requisitos, devem ser utilizadas ferramentas adicionais, como, por exemplo, os tesouros para gerenciar as concordâncias entre as estruturas antigas e as novas (SLAVIC; DAVIES, 2017).

2.3.3 Interpretação dos resultados e conclusões da RSL

Na etapa 8, após os resultados das análises verificou-se que as publicações selecionadas na amostra permitiram conhecer os pesquisadores de diversos países, não sendo identificado nenhum estudo publicado no Brasil, cujo idioma predominante foi o inglês, com um pequeno número de artigos em espanhol. A amostra contemplou um total de 44 trabalhos selecionados do total de 210, sendo 42 em inglês e dois em espanhol, constituindo o *corpus* da pesquisa.

Dentre os estudos analisados, percebeu-se que a autora com maior número de publicações é Aida Slavic. Os anos que obtiveram o maior número de publicações foi o ano de 2007 e 2008, com predominância em artigos de periódicos, no que tange aos tipos de documentos ou meios de publicação, sendo o principal periódico o “*Extensions & Corrections to the UDC*”. A diminuição do número de publicações de 2008 a 2017 deve ser mais investigada, como também análises dos trabalhos publicados em outros idiomas, haja vista verificar novas possibilidades de uso da CDU relacionada a tesouro, tendo em vista a recuperação em ambientes digitais não necessariamente somente na web. A instituição dos autores que mais publicaram sobre o tema foi a Universidade de Londres. A palavra-chave mais utilizada nos veículos de publicação foi “UDC”. O trabalho mais citado foi o de McIlwaine (1997), e os principais domínios estudados foram Arquitetura, Biblioteconomia e Religião.

Slavic (2005) observou que a busca por soluções está nos assuntos contidos em disciplinas do conhecimento, sendo desenvolvidos tesouros e sistemas de classificações bibliográficas para organização desses assuntos, conectando todo o conhecimento humano com um vocabulário controlado e terminologia utilizada em diversos domínios. De acordo com Frâncu (2003a, 2003b), o motivo está nas relações semânticas entre os termos, permitindo uma estrutura classificatória facetada de acordo com as regras de cada domínio em sua estrutura, com a utilização da CDU e de um tesouro, visando melhorar a consistência, reuso e a compatibilidade de linguagem. Para isso, deve-se realizar a compatibilização dos termos e as relações semânticas dos dois SOC com a classificação da CDU original (FREEMAN, 1964; MILLS, 1970).

No que tange ao conteúdo das 44 publicações, pelos resultados, foi possível constatar a potencialidade do uso da CDU que, geralmente, é utilizada para organização dos documentos em uma biblioteca ou centros de documentação. Os estudos reforçaram que, se a CDU for vinculada a um tesouro, essa é uma proposta viável, pois promoverá melhorias para aumentar o potencial na RI. Por outro lado, como reforça Colillas e Nicolau (2008), a recuperação pode ser mais relevante quando combinada ao alto grau de especificidade e flexibilidade da CDU, cujos termos, notações e relações semânticas potencializam a recuperação, sobretudo em tesouros de domínios específicos.

Verificou-se, também, a contribuição do tesauro ao fornecer termos padronizados que contemplam os assuntos de um domínio, em conformidade com os termos utilizados pelos usuários nas atividades de busca. Nessas atividades, a recuperação será mais efetiva se a informação registrada nas bases de dados levarem em consideração as necessidades de informação dos usuários. Portanto, com a junção dos dois SOC, é possível criar relacionamentos entre os termos da CDU com os descritores do tesauro, por meio de relações hierárquicas, associativas e de equivalência, sem alterar a estrutura original da CDU. Após a identificação das relações entre os dois SOC, é permitido um alto grau de flexibilidade em uma estrutura classificatória facetada (COLILLAS, 2009; LEE, 2017).

A capacidade de especificidade de representação de assuntos é possibilitada por meio de símbolos e tabelas indicando relações entre os assuntos. Durante o percurso da pesquisa, observou-se que é necessário incluir os descritores de um tesauro a um sistema de classificação bibliográfica para que se tenha maior grau de especificidade na RI. Desse modo, de acordo com o estudo de Slavic (2005), pode-se obter uma estrutura navegacional mais ampla e flexível, com a terminologia capaz de abranger um conjunto de termos que representem o domínio proposto, aumentando a precisão na RI. A CDU pode ser uma ferramenta bastante eficiente para a organização e representação do conhecimento, com o objetivo de melhorar a RI (SLAVIC, 2005).

A partir da análise das pesquisas de Chudamani (2004), Slavic (2006a), Gnoli (2007), La Barre (2007) e Broughton (2010), foi possível constatar que a característica mais específica da CDU é a sua extensa função analítico-sintética, cujos elementos da classificação podem ser livremente combinados em vários níveis. Durante o processo de análise da CDU, os conceitos são representados por meio de assuntos simples e, na etapa de síntese, as categorias são combinadas mediante o uso do vocabulário que é organizado de modo hierárquico e estruturado, de acordo com as categorias de assuntos. Por esse motivo, o vocabulário deve ser escolhido de maneira pragmática e as regras sintáticas utilizadas de modo flexível, com a inserção de símbolos que permitem a combinação de assuntos de diferentes categorias ou de diferentes disciplinas de acordo com os princípios da classificação facetada.

A partir desta revisão, foi possível verificar os princípios da classificação facetada de Ranganathan com os processos de análise e síntese de conceitos para serem organizados em categorias mutuamente exclusivas, aplicando-se o princípio de divisão única em cada nível de subdivisão, atribuindo para cada assunto uma notação. Essa notação é inserida de modo separado da estrutura de classificação em banco de dados, cuja única finalidade é a ordenação das classes, e seu sistema notacional não expressa os termos por meio de uma sintaxe ou hierarquia ordenada semanticamente, ficando atribuída esta tarefa para o classificador (SLAVIC, 2005).

Esta pesquisa possibilitou a compreensão da capacidade da CDU pelo potencial desse sistema representar conceitos, por meio de unidades de conhecimento com as possíveis combinações de assuntos. Com essas características, e seguindo as sugestões da Classificação Facetada de Ranganathan, é possível construir sistemas de classificação reproduzíveis, inserindo a terminologia necessária para indexar e classificar os termos a partir dos conceitos. Isso permitirá a navegação por assuntos, por meio dos termos e facetas, colaborando para a organização e representação da informação (SLAVIC, 2005, 2008).

Pelo estudo de Slavic (2005, 2008) verificou que o modelo de dados baseado na estrutura da CDU associada a um tesauro em um SRI pode ser simplificado, tratando a combinação de assuntos das classes principais, ou pelo relacionamento com as tabelas auxiliares, como se fosse um assunto simples. Chan e Zeng (2002) relatam que a falta de revocação na RI pode ser evitada fornecendo combinações predeterminadas, estabelecendo links semânticos entre os termos já existentes no tesauro com os da CDU para que sejam interoperáveis.

Durante a revisão da literatura, percebeu-se que a utilização do vocabulário (termos) da CDU, quando vinculado aos descritores de um tesauro, foi organizada de modo relacional em uma estrutura hierárquica (taxonomia). Essas regras de sintaxe utilizadas nessa organização possibilitaram a reutilização por meio de diferentes símbolos de expressões de assuntos, que podem ser combinadas e adequadas para uma indexação detalhada, de fácil expansão e aplicação em qualquer área do conhecimento. Por esse motivo, alguns autores como Slavic (2005) e Colillas (2009) estão preocupados com o reconhecimento dos elementos presentes na estrutura pré-coordenada da CDU, para depois inserir os descritores de um tesauro, haja vista melhorar a precisão e revocação na RI. Logo, ao inserir novos termos de classificação presentes no tesauro, aumenta-se o potencial da estrutura da CDU já existente. Assim, esse aumento não será somente pela união de assuntos compostos ou complexos e de sua notação existente, mas também por meio das relações já existentes de gênero-espécie e todo-parte, com a adição da relação de equivalência e associativa entre os descritores do tesauro (SLAVIC, 2005; COLLILAS, 2009).

Slavic (2004a, 2006b) afirma que as informações sobre as regras para a criação de um sistema baseado na estrutura da CDU e de um tesauro devem conter uma estrutura semântica dos elementos, por meio de um esquema de metadados em banco de dados. Essa estrutura promove um importante papel no desenvolvimento de sistemas em ambientes on-line, sendo utilizada para diversos propósitos, com o intuito de melhorar a RI.

Na revisão da literatura, concluiu-se que, para relacionar os termos do esquema de classificação da CDU, com os descritores do tesauro, existe uma grande complexidade por problemas de ambiguidade da linguagem natural, como, por exemplo: homônimos, polissemia, sinonímia, dentre outros, devendo haver um alto grau de esforço intelectual para

criar um novo instrumento de organização do conhecimento, gerando um tesauro com as notações e termos da CDU. A identificação e a inclusão de termos do tesauro na CDU, serão a partir do perfil do usuário de um domínio específico, por meio da inclusão de relações semânticas entre os termos para a combinação de assuntos, como, por exemplo, <<é um tipo de>>, <<é sinônimo de>>. Para isso, é necessário haver uma compatibilidade da linguagem entre os termos da CDU e descritores do tesauro, para que a descrição de um índice de classificação seja adequada para não comprometer o resultado da RI.

Algumas pesquisas confirmaram que a construção de um SRI que vise melhorar os processos de indexação e recuperação, somente será possível por meio do mapeamento dos códigos da CDU e seus termos associados com cada descritor de um tesauro para representar um assunto (CHAN; ZENG, 2002, FRÂNCU, 2003a, 2003b; FRÂNCU; SABO, 2010). Logo, com a terminologia e notação da CDU e suas relações semânticas com os descritores do tesauro, torna-se possível utilizá-la em diferentes línguas por meio de um tesauro multilíngue de um domínio. A revisão da literatura constatou que não existem pesquisas para utilização da CDU vinculada aos descritores e relações semânticas específicas de um tesauro para um contexto específico no âmbito do domínio da agricultura, que será realizado por meio de um estudo de caso nesta tese.

Ficou evidenciado que não existem pesquisas para utilização dos termos específicos de uma classe da CDU que, por sua vez, estão vinculados aos descritores e relações semânticas de um tesauro para um domínio específico. Para isso, torna-se necessário extrair os descritores que representam um domínio para identificar com precisão o conceito referente ao assunto de um documento. Outrossim, o tesauro possui os descritores que contemplam os assuntos de um domínio em conformidade com os termos utilizados pelos usuários em um SRI, visando à recuperação, de acordo com a necessidade de informação, considerando a informação registrada na base de dados a ser pesquisada.

Nesse sentido, podem-se criar relacionamentos semânticos entre os conceitos da CDU e do tesauro, preservando os relacionamentos já existentes nos dois SOC e unificar as duas linguagens. Com a notação da CDU é possível unificar vários SOC independente do idioma, porém para o mapeamento entre os vocabulários multilíngue é necessário levar em conta a tradução para os termos específicos de um domínio. Outro ponto a ser considerado é a experiência do indexador, entre outros fatores durante o processo de indexação para que se tenha resultados relevantes na RI. Além do mais, com a criação de vários tipos de relacionamentos entre a CDU e o tesauro, os dois SOC irão ser capazes de utilizarem as relações hierárquicas (gênero-espécie e todo-parte), de equivalência e associativas, deixando de ser uma estrutura somente hierárquica para ser uma estrutura multidimensional, permitindo a relação entre os termos que denotam conceitos em qualquer nível da hierarquia da CDU.

Dentro do percurso desta tese, neste capítulo foram descritas as etapas utilizadas na revisão sistemática da literatura e os procedimentos realizados em cada etapa, incluindo as análises a partir dos resultados obtidas nesta revisão. No próximo capítulo, será apresentado o referencial teórico-metodológico que evidencia os principais fundamentos utilizados na elaboração desta tese.

3 REFERENCIAL TEÓRICO E METODOLÓGICO

“A alegria que se tem em pensar e aprender faz-nos pensar e aprender ainda mais”.

(Aristóteles)

O estudo proposto foi fundamentado em conceitos, métodos e teorias disponíveis na literatura científica da BCI. Assim, a fundamentação da pesquisa explora os seguintes temas: Sistemas de Organização do Conhecimento (SOC); Sistema de Classificação Bibliográfica, em particular a Classificação Decimal Universal (CDU); Tesouros; Relações Semânticas, Recuperação da Informação; Compatibilidade e Convertibilidade de Linguagem; e Simple Knowledge Organization (SKOS).

Considerando essas temáticas, nos próximos subcapítulos, são abordados os assuntos associados ao problema desta pesquisa.

3.1 Sistemas de Organização do Conhecimento

O termo “Sistemas de Organização do Conhecimento” é a expressão em português traduzida do inglês “Knowledge Organization Systems”, que designa todos os tipos de esquemas para organizar informações e promover a gestão do conhecimento (HODGE, 2000). Como exemplos de SOC, citam-se: *gazetteers*, bancos de dados lexicais, taxonomias, tesouros e ontologias (HODGE, 2000).

Esses sistemas modelam a estrutura semântica subjacente de um domínio capaz de auxiliar a navegação e tradução por meio de rótulos, definições, digitação, relacionamentos e propriedades para conceitos (HILL *et al.*, 2004; KOCH; TUDHOPE, 2004; ZENG, 2008). Incorporados como serviços da web, tais sistemas facilitam a busca e recuperação de recursos, agindo como mapas semânticos, o que possibilita uma orientação comum para indexadores e futuros usuários ou máquinas (KOCH; TUDHOPE, 2003, 2004).

Na organização da informação, os SOC visam à padronização terminológica para auxiliar na representação descritiva (física) e temática (assunto) de um recurso informacional (SVENONIUS, 2004). Para Guinchat e Menou (1994), a representação da informação é realizada por meio dos processos de catalogação, indexação, classificação, elaboração de resumos, gestão de registros bibliográficos e criação de bases de dados bibliográficas, de modo padronizado, para que a informação contida em um recurso informacional possa ser recuperada.

Os autores Prasad e Madalli (2009) apontam que os SOC conectam as informações entre diferentes recursos em diversos domínios, correlacionando-as, independente se elas estão no próprio SOC, armazenadas em bases de dados ou em arquivos. As informações dos recursos são conectadas por meio de índices ou notações da

classificação, de modo que sejam recuperadas, bem como acessadas em outros recursos informacionais, para que o usuário consiga recuperar documentos relevantes (PRASAD; MADALLI, 2009). Desse modo, a representação da informação é constituída por um conjunto de atributos capazes de descrever o conteúdo informacional de um determinado recurso, que pode estar no formato impresso ou digital, traduzido em termos ou palavras que representam esse conteúdo (PRASAD; MADALLI, 2009).

Analisando o contexto da representação da informação, é demonstrado no estudo de Brascher e Café (2008) que o processo de representar um recurso é realizado por meio de elementos descritivos dos dados ou metadados. Os metadados são propriedades que representam uma entidade ou objeto do mundo real, conceituados como “dados sobre dados” ou “definição dos dados”, que descrevem a semântica dos dados referentes aos recursos informacionais (documentos), permitindo o relacionamento entre seus elementos e a representação do conhecimento mediante o uso de diferentes tipos de SOC (BRASCHER; CAFÉ, 2008).

Soergel (2014) considera que a utilização de padrões ou esquemas de metadados em um SOC são o agrupamento de metadados com características similares, sendo considerados elementos-chave para a organização e representação da informação em ambientes digitais. Para tanto, segundo esse autor, existem vários padrões de metadados utilizados para diversas finalidades, como a *Extensible Markup Language* (XML), que é uma linguagem de marcação (rótulo ou etiquetagem) para o intercâmbio de dados entre sistemas de informação, já o *Dublin Core* (DC) é estruturado para padronizar os dados referentes ao conteúdo de um recurso informacional e seus relacionamentos semânticos, por exemplo, os dados autor, ano, editora, título, dentre outros.

Em relação à padronização dos dados, Lima (2020) ressalta a necessidade da utilização de metadados entre diferentes tipos de SOC, com o propósito de compartilhar e disseminar a informação, ou seja, permitir a interoperabilidade. Durante o processo de interoperabilidade, devem-se levar em conta as características peculiares das coleções, dos sistemas de informação, dos vocabulários, estabelecendo relações semânticas entre diferentes tipos de SOC (LIMA, 2020). A autora reforça que, no ambiente físico, a interoperabilidade deve ser capaz de integrar diversos sistemas de informação, aplicativos (softwares), repositórios digitais ou bases de dados de conhecimento, além de permitir que a informação esteja disponível de modo consistente para suprir a necessidade de informação dos usuários em ambiente on-line.

Do ponto de vista de Zeng (2019), durante o processo de interoperabilidade entre SOC, podem ocorrer diversos problemas durante o intercâmbio de informações que por sua vez, se manifestam em diversos níveis: sistêmico, sintático, estrutural e semântico. Em relação à interoperabilidade entre sistemas, observam-se problemas de incompatibilidade

durante a troca de dados pela rede, entre computadores, softwares e serviços na web (ZENG, 2019). Já a interoperabilidade entre diferentes vocabulários, no nível sintático, para essa autora, é possível haver problemas na troca de dados que não podem ser tratados em relação ao formato ou padrão de dados durante a codificação, decodificação e representação dos dados; já na camada estrutural ou arquitetural, os problemas mais comuns referem-se a modelos de dados, estrutura de dados e esquema de dados, ou seja, em relação à estrutura dos dados e metadados (propriedades dos dados) (ZENG, 2019). Por fim, na camada semântica, são apresentados problemas relativos ao entendimento ou significado correto da linguagem, terminologia e/ou metadados utilizados, conforme afirmação de Zeng (2019), confirmando também que, sem a interoperabilidade em todos os três níveis citados acima, os metadados não podem ser compartilhados de modo eficiente.

No ambiente digital ou na web, com a existência de diversas bases de dados formadas por diferentes tipos de dados, torna-se necessário que esses dados sejam interoperáveis (WEISS, 2019). Para isso, é necessária a utilização de estratégias, tais como o reuso de vocabulário controlado, utilizado com frequência na construção de SOC, por exemplo em sistemas de classificação, ontologias, dentre outros, tendo em vista representar conceitos já existentes em diversos estudos (CARDILLO *et al.*, 2014).

Além da necessidade de que os dados entre diferentes tipos de SOC sejam interoperáveis, Soergel (1999) considera o uso dos SOC de acordo com a base de conhecimento que se deseja organizar, tendo em vista três objetivos principais: indexar de maneira automática; traduzir e recuperar informações; e possibilitar o processamento da linguagem natural (PLN) em diversos idiomas.

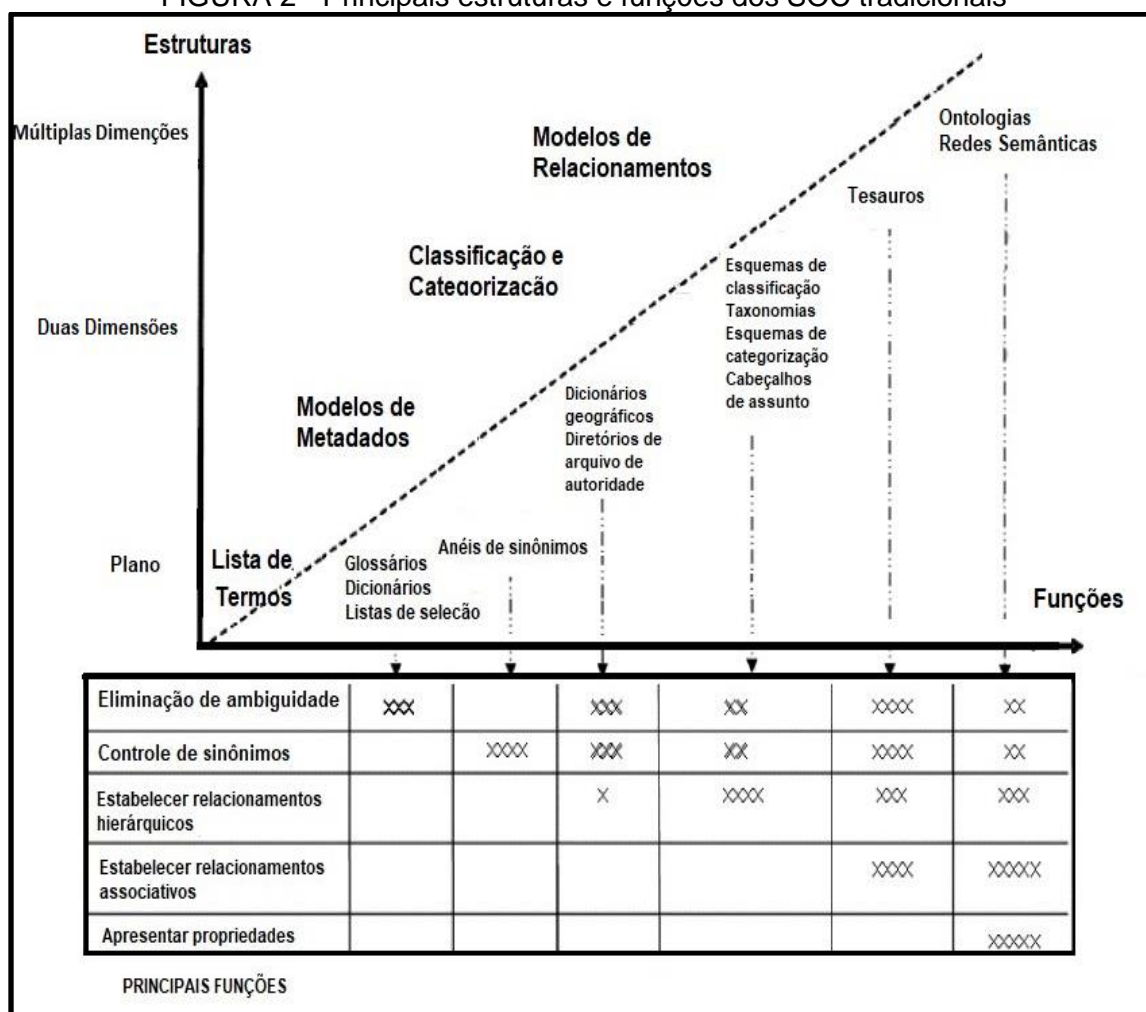
Outros autores, entre eles Souza, Tudhope e Almeida (2012), também apresentam a mesma ideia acerca do objetivo de um SOC, reforçando que esses sistemas ajudam na organização da informação, com o objetivo de garantir e melhorar a recuperação da informação. Nessa perspectiva, os autores acreditam no potencial da área da BCI, visto que é missão dessa área “[...] organizar e facilitar a recuperação da riqueza de informação que surge do conhecimento produzido em todos os outros campos, e isso envolve a criação de políticas epistemológicas e substitutas ontológicas [...]” (SOUZA; TUDHOPE; ALMEIDA, 2012, p. 179). Nas palavras desses autores:

[...] muitas das teorias, processos e instrumentos dependem de produtos de representação, modelados através de abstrações sucessivas sobre as características relevantes de um mundo ou domínio, ou as informações coletadas e processadas sobre elas, registrado em sistemas de informação e documentos. As representações, conhecidas como Sistemas de Organização do Conhecimento (SOC), variam enormemente em formato e exibição, mas compartilham a característica geral de auxiliar a organização do conhecimento, visando promover a recuperação da informação (SOUZA; TUDHOPE; ALMEIDA, 2012, p.179).

Na opinião de Mayr *et al.* (2016), os SOC promovem a semântica em um ambiente controlado e documentado. Além disso, os SOC são utilizados no suporte ao conhecimento para informação, mapas conceituais, ferramentas de comunicação que possibilitam um framework conceitual para a modelagem conceitual de um domínio (MAYR *et al.*, 2016). Campos e Gomes (2008) compartilham da ideia, afirmando que, para modelar um domínio utilizando um SOC, deve-se antes analisar o método utilizado em sua construção para que o SOC utilizado possa auxiliar a organização do conhecimento de maneira eficiente. Ademais, as autoras salientam que os SOC verificam como estão definidos os objetos da representação, a relação entre eles, de acordo com a semântica entre os conceitos de um determinado domínio e as formas de representação gráfica que um modelo pode adotar.

Desse modo, os SOC se apresentam em diferentes estruturas, com a finalidade de facilitar a representação da informação de um recurso informacional em um SRI, com o objetivo de representar o conhecimento de um domínio específico (LIMA, 2020). A síntese dos principais SOC existentes é exibida na Figura 2, a qual indica as principais funções e a complexidade de cada sistema, de acordo com a estrutura adotada.

FIGURA 2 - Principais estruturas e funções dos SOC tradicionais



FONTE: Adaptada de Zeng (2008, p. 161).

Mediante a leitura da Figura 2, percebe-se que a estrutura dos SOC possui desde instrumentos com baixa complexidade até os mais complexos. A complexidade está relacionada com o nível de controle do vocabulário. Os instrumentos mais simples são baseados em uma estrutura sintática mais próximo da linguagem natural, ou seja, possuem apenas o nível de controle de sinônimos, com base na terminologia, tais como, as listas de termos, os dicionários, os glossários e os anéis de sinônimos (ZENG, 2008). Conforme expresso nessa Figura e percorrido pela autora citada, os instrumentos mais complexos foram construídos com um nível de controle maior do vocabulário, que é realizada por meio de relacionamentos semânticos, hierárquicos, associativos, entre outras propriedades (características). Estes instrumentos são conhecidos como arquivos de autoridade de assuntos, sistemas de classificação, taxonomias, tesouros, redes semânticas e ontologias (ZENG, 2008).

Nesse contexto, Zeng (2008) mostra diferentes tipos de SOC que foram estruturados e implementados com as seguintes funções: eliminação de ambiguidade, garantia do controle de sinônimos, estabelecimento dos relacionamentos hierárquicos entre os conceitos, formação de associações e a apresentação das propriedades. Além do mais, os SOC podem ser utilizados como relações complexas a partir de regras e axiomas, visando ao processamento por máquina, cujas ontologias e as redes semânticas possuem todas as funções mencionadas e variados tipos de relações semânticas (ZENG, 2008).

Ainda sobre as funções dos SOC, Svenonius (2000) ressalta os problemas oriundos dos diversos modos utilizados para descrever um determinado assunto, causados pela ambiguidade da linguagem natural, sobretudo com o uso dos homógrafos e polissemias, o que revela a necessidade de se estabelecer alternativas de desambiguação. A autora recomenda algumas e possíveis alternativas, a saber: a especificação do domínio, a inserção de qualificadores entre parênteses, notas de escopo, exibições hierárquicas, dentre outras. Salienta-se que esse tipo de informação é de interesse não só das unidades de informação, no que tange à organização de seus recursos, mas também para os desenvolvedores de sistemas de informação que poderão utilizar algum tipo de SOC, por exemplo, taxonomias, sistemas de classificação ou ontologias, tendo em vista garantir a organização do conhecimento em um contexto específico (SVENONIUS, 2000).

Para Broughton (2006), com a utilização dos SOC, é possível obter uma recuperação mais efetiva e consistente, pois os instrumentos desses sistemas, como os cabeçalhos de assuntos, taxonomias, sistemas de classificação, dentre outros, possibilitam organizar o conhecimento, a partir de relações hierárquicas; já os tesouros, redes semânticas e ontologias, além de utilizarem as relações semânticas já mencionadas, também utilizam as relações associativas e incluem as propriedades (características) do conhecimento a ser representado. Com a utilização dos SOC no ambiente digital ou na web, facilita-se a utilização

de ferramentas de busca para recolher ou expandir consultas, entre outras possibilidades (BROUGHTON, 2006). Em relação aos sistemas de classificação, é possível fazer um mapeamento interdisciplinar por meio de notações, a serem representadas por um código único, servindo como base de conhecimento multilíngue para o processamento da linguagem natural (PLN) (BROUGHTON, 2006).

De acordo com Lima e Maculan (2017), a estrutura da maioria dos SOC, tais como os tesouros, as taxonomias e os sistemas de classificação, não possui um grau de formalização semântica capaz de alcançar um nível descritivo e consistente como acontece nas ontologias. Para as autoras, as ontologias possuem conceitos bem definidos que representam as características de um domínio e seus relacionamentos, possibilitando uma representação formal e lógica, o que torna esse instrumento capaz de fazer inferências do domínio que está sendo modelado, haja vista expandir o conhecimento. Nesse sentido, os SOC possuem estruturas com características específicas que dependem de sua aplicação e do contexto e, para cada tipo de SOC, é necessário um conjunto de regras e metodologias diferentes para sua construção e/ou atualização, com o objetivo de garantir o controle terminológico e semântico (LIMA; MACULAN, 2017).

As aplicações dos SOC possibilitam o controle da terminologia de um domínio, mediante o uso de um vocabulário controlado, que é uma linguagem composta por termos que podem ser organizados em uma estrutura de relações semânticas ou notacional (HODGE, 2000). Isso quer dizer, segundo a autora, e utilizando outras palavras, que os vocabulários controlados permitem a tradução para uma linguagem de indexação, com consistência, gerando relacionamentos que garantam a busca e recuperação, de modo específico, eficaz e rápido.

Além disso, os SOC são comumente utilizados para diversos propósitos, por exemplo: nos refinamentos de relações semânticas propondo uma reengenharia de tesouros (MACULAN, 2015); na avaliação de ontologias do ponto de vista de um domínio específico, mostrando os pontos que precisam ser melhorados para ampliar a abrangência, de modo que seja suficiente para representar a visão de um especialista ou usuário do domínio (ARAÚJO, 2016); na visualização de ontologias de um domínio, analisando a necessidade de informação dos usuários e questões de usabilidade, sendo apontados os níveis de eficácia e eficiência que atendam, de modo satisfatório, uma comunidade de usuários específica (NETTO, 2017), no reúso de vocabulários controlados para construção de outros vocabulários, a serem utilizados como recursos disponibilizados, em particular, no contexto da web semântica (FIRMINO, 2019); dentre muitas outras aplicações para diversos domínios.

Considerando essas características mencionadas para os SOC, segundo Fujita, Santos e Alves (2018), os termos SOC, linguagem documentária e linguagem de indexação, por possuírem aplicações bem próximas, são considerados sinônimos, pois são instrumentos

de representação do conhecimento com a finalidade de garantir a recuperação. A linguagem documentária, de origem francesa, foi adotada, a princípio, por Gardin, ao tratar de um conjunto de termos que possuem ou não regras sintáticas, com o fim de representar conteúdos de documentos para classificação e recuperação. Por fim, o termo linguagem de indexação, de origem inglesa, relaciona-se com o processo de indexar, o qual utiliza ferramentas de representação da informação, com fins de recuperação, sendo fundamentado por autores como Foskett, Vickery, Slype, dentre outros (FUGITA *et al.*, 2018).

Nesse sentido, os SOC denotam uma nova terminologia a nível conceitual dos termos “linguagem documentária” e “linguagem de indexação” em comparação com os instrumentos tradicionais de organização do conhecimento, tais como os sistemas de classificação e os tesouros (FUGITA *et al.*, 2018). Desse modo, em concordância com os autores citados, os SOC abrangem os aspectos desses dois termos, com foco no desenvolvimento de tecnologias de organização do conhecimento, sendo bastante utilizados em estudos da web semântica, devido ao grande volume de informações no ambiente digital, visando à recuperação da informação.

Diante do exposto, os SOC são aplicados na padronização, organização e classificação do conteúdo dos documentos a partir da necessidade do usuário. Com efeito, tem-se a aproximação da linguagem dos SOC com a dos usuários, para que as informações sejam recuperadas com precisão e consistência em qualquer tipo de aplicação ou domínio (ZENG, 2008).

O próximo subcapítulo discorre sobre a Classificação Bibliográfica, apresentando as definições, características, funções, tipos e estruturas desse tipo de classificação. Em seguida, detalha-se o sistema CDU como instrumento para auxiliar na RI, por ser o insumo ou objeto desta pesquisa.

3.1.1 Sistema de Classificação Bibliográfica

O ato de classificar pode ser definido como a descrição dos vários aspectos do mundo real, agrupados de acordo com as semelhanças e as respectivas relações semânticas entre eles (PIEIDADE, 1977). Segundo Ranganathan (1968), classificar consiste em traduzir e/ou codificar ou decodificar os assuntos dos documentos da linguagem natural para a linguagem artificial (vocabulário controlado). Este processo pode ser realizado com o uso dos sistemas de classificação bibliográfica, caracterizados como sistemas de conceitos que possibilitam uma efetiva recuperação dos documentos. Por esse motivo, existe a necessidade do tratamento da informação, que tem como principal propósito tornar documentos ou quaisquer recursos informacionais organizados, por conseguinte, passíveis de serem recuperados.

Desse modo, é preciso prever as tecnologias e os processos necessários para facilitar a organização e o acesso às informações pelos usuários (RANGANATHAN, 1968).

Piedade (1977) enfatiza que classificar é a ação de reunir em classes os objetos ou elementos existentes no mundo, com base em critérios estabelecidos, que permitam identificar as semelhanças e as diferenças entre esses objetos. Essa ação é, segundo a autora, tão antiga quanto a própria humanidade, progredindo do ponto de vista filosófico ou científico para o bibliográfico. Barbosa (1969) lembra que as classificações filosóficas foram criadas por filósofos, idealizadores desse tipo de classificação; já as classificações bibliográficas surgiram com o intuito de organizar documentos, desde sua disposição física até a sua recuperação. Nessa direção, Dahlberg (1976) ressalta que, no mundo antigo, a organização dos documentos em estantes, catálogos, entre outras possibilidades, era feita de acordo com as ideias preconcebidas, devido à ausência de um sistema de classificação bibliográfica capaz de acompanhar a evolução do conhecimento, com o objetivo de torná-lo acessível, do ponto de vista da organização e recuperação.

Os autores Nunes e Tálamo (2009, p. 34) definem classificar como uma prática relevante para a organização da informação, visto que possibilita “[...] a transparência e o compartilhamento de informações, as quais são caminhos para tomadas de decisões, para preservação da memória técnica e administrativa das organizações contemporâneas e também para o exercício da cidadania [...]”. Ademais, defendem que a classificação é reconhecida como matricial, pois dela manifestam-se outras funções ligadas à organização do conhecimento na sociedade. Assim, “[...] desde o momento que seleciona o documento para guardar, o classificador trabalha com uma série de funções que discute a dimensão do conhecimento a qual permite a prioridade de procedimentos” (NUNES; TÁLAMO, 2009, p. 34). Para Chan e Salaba (2015), a preocupação em classificar e organizar todo e qualquer material identifica a importância da ação humana na geração dos registros do conhecimento, sendo esses gerados ao longo dos diferentes períodos da história da humanidade.

Hjørland e Nissen Pedersen (2005) perceberam que, para classificar um documento, é necessário um processo abstrato ou empírico, com o intuito de indicar uma compreensão positivista. Essa compreensão se sustenta em conceitos que são representados de acordo com os aspectos cognitivos, por meio da ciência e do conhecimento humano ou filosófico, considerando o que se percebe do mundo. Portanto, a classificação caracteriza-se como um "processo mental", estudado principalmente por profissionais da computação e da informação, sendo executado de modo consciente e/ou inconsciente, por qualquer ser humano (HJØRLAND; NIESSEN PEDERSEN, 2005). Nesse contexto, a classificação remete a uma compreensão pragmática, com o agrupamento de termos para a representação de conceitos que são alinhados em classes ou categorias, cuja finalidade é auxiliar o usuário a

organizar e entender, de maneira lógica e fácil, como catalogar e localizar os recursos informacionais (VICKERY, 1990).

O termo “classificação” é uma derivação da palavra latina "*classis*", que significa o agrupamento de alguma coisa, e constitui um processo responsável pela organização e representação de recursos informacionais utilizados nas atividades de organização do conhecimento, tendo como resultado a geração de classes distribuídas em uma sequência lógica e útil (SIMÕES, 2010). No âmbito da Biblioteconomia e Ciência da Informação (BCI), as classificações bibliográficas representam um mapeamento completo de um domínio ou disciplina do conhecimento, cujo classificador atribui a um recurso informacional um único lugar no sistema de classificação (PIEDE, 1977). Nos centros de informação, segundo essa autora, o item classificado recebe uma localização física nas estantes, de acordo com os conceitos e suas relações que, por sua vez, são acessadas por meio de símbolos linguísticos (termos).

Nesse contexto, as classificações bibliográficas representam os assuntos a partir do processo de sistematização dos conceitos e da categorização, realizada durante a análise de assunto de um documento, com o propósito de tornar a informação acessível aos usuários (LIMA, 2007). No entendimento de Simões (2010), um sistema de classificação permite a representação temática do conhecimento, gerando um conjunto de classes sistematizadas de modo coordenado e subordinado, o que amplia os campos de investigação dos pesquisadores interessados na temática, sobretudo com o aumento do número de recursos no formato digital.

Dahlberg (1993) relata que o conhecimento, por si só, não pode ser representado, a menos que seja por unidades de conhecimento e suas possíveis combinações de assuntos. Portanto, conforme relatado por Dahlberg (1993), o agrupamento de termos não necessariamente caracteriza uma linguagem específica, pois esses termos podem ter diferentes significados, fato esse que justifica a organização do conhecimento, de modo a estabelecer linguagens de indexação como, por exemplo, as classificações bibliográficas. Essas linguagens precisam ser elaboradas com foco no usuário, pois o profissional da informação trabalha com diversas áreas do conhecimento e existem usuários que necessitam de informações sobre qualquer disciplina do conhecimento (DAHLBERG, 1993).

Para uma recuperação mais específica, Dodebei (2002) observa que os itens classificados devem possuir as seguintes qualidades: serem completos, irreduzíveis e mutuamente exclusivos, considerando a organização conceitual do ponto de vista da intensão ou extensão. Esclarece a autora que a intensão, no âmbito conceitual das classificações, corresponde ao número de características que o objeto possui, e a extensão representa as instâncias ou espécies de uma classe de conceitos. Logo, quanto maior a intensão, menor é a extensão, ou seja, quanto maior o número de características do objeto, menor será o número de itens retornados, aumentando, assim, a eficiência e a precisão na recuperação (DODEBEI,

2002). Partindo dessas características, depreende-se que, como o conhecimento está em constante desenvolvimento, as classificações devem ser flexíveis e mutáveis, ou seja, capazes de acompanhar as mudanças, a partir de estruturas classificatórias com relacionamentos hierárquicos e com o conhecimento multidimensional (DAHLBERG, 1978b).

Uma das características da classificação bibliográfica apresentada por Vickery (1980) é a criação das classes ou categorias, que se apresentam no formato de árvore invertida, cujas folhas, que se encontram no mesmo ramo, possuem características mais semelhantes do que outras que estão em ramos diferentes. A mesma concepção é adotada por Aristóteles, nas classificações filosóficas, ao dividir o conhecimento em dez categorias principais, processo também realizado por Ranganathan, na classificação bibliográfica. Assim, nas classificações, as categorias e subcategorias de conceitos foram organizadas em hierarquias com propriedades inerentes que auxiliam no acesso e na recuperação (VICKERY, 1980). Logo, pode-se dizer que a classificação bibliográfica possui uma lista de classes com suas principais subdivisões de assuntos em um sistema de classificação, visando à organização dos itens em um catálogo bibliográfico, coleção ou no índice de classificação de uma área do conhecimento (VICKERY, 1980).

Para Barbosa (1972), o sistema de classificação possui um índice que é uma lista alfabética de todas as disciplinas ou áreas com a notação de cada classe de assuntos. Nesse contexto, um sistema de classificação possui tabelas com notações adicionais, que podem ser usadas no número que identifica a classe para representar assuntos com maior grau de especificidade (BARBOSA, 1972). Portanto, destaca a autora que o índice pode ser relativo, quando inclui tópicos de assuntos gerais em ordem alfabética, ou seja, com uma lista de todos os aspectos do assunto; ou o índice pode ser específico, contendo uma lista de assuntos própria, também em ordem alfabética, mas sem a lista de tópicos relacionados com os assuntos gerais.

Caso um tópico de assunto esteja relacionado em duas ou mais classes, será considerado o número que leva em cada tabela e sua localização nas estantes e a notação do leitor para a disposição nas estantes (KOCH, *et al.*, 1997). Dahlberg (1993) caracteriza os índices como um arranjo lógico de todos os assuntos que estão no sistema de classificação, sendo que esse arranjo geralmente é hierárquico, capaz de mostrar a relação de assuntos específicos com o assunto geral. Por fim, os índices possuem uma notação ou símbolo de classificação mostrado em cada assunto (DAHLBERG, 1993).

Concebem-se os sistemas de classificação bibliográfica para cumprir a função pela qual foram criados, ou seja, serem utilizados para fins de sistematização do conhecimento, tornando mais fácil a localização da informação registrada, a qual é agrupada em classes (DAHLBERG, 1976). As classes, de acordo com Piedade (1977), precisam ser consistentes e exaustivas, de modo a atender o princípio da divisão. Nesse princípio, a autora

indica que cada elemento de uma classe pode ter apenas uma característica, ou seja, precisa seguir a ordem da natureza entre os seres ou objetos, agrupando-os em classes de acordo com as propriedades de cada um.

Na visão de Simões e Freitas (2013), a partir da classificação bibliográfica, o usuário pode encontrar alguma informação incorporada nos objetos com mais rapidez e agilidade, estejam os objetos no formato impresso ou digital. Sendo assim, a classificação auxilia a organização das informações, de modo coerente e consistente, e isso revela que os sistemas de classificação bibliográfica existem para contrariarem o fenômeno da dispersão do conhecimento, visto que organizam as entidades físicas e abstratas em grandes categorias ou classes (SIMÕES; FREITAS, 2013).

As classificações bibliográficas são comumente utilizadas para organizar documentos como livros, teses, dissertações, dentre outros, os quais constituirão os acervos das bibliotecas, dos arquivos e demais unidades de informação, e estão ordenados em referências bibliográficas ou em fichas catalográficas (DAHLBERG, 1976). Considera-se, segundo essa autora, que a função dos sistemas de classificação vai muito além da organização de documentos nas estantes, uma vez que possibilita o registro da informação produzida em uma determinada área do conhecimento. Entretanto, no decorrer dos tempos, com o surgimento de novos campos ou áreas do conhecimento, as classificações podem ficar obsoletas, caso não sejam atualizadas pelos classificionistas (DAHLBERG, 1976).

Quanto aos classificionistas, Piedade (1977) explica que eles elaboram os sistemas de classificação, para que os classificadores utilizem esses instrumentos, sem que ocorra insatisfação junto à comunidade científica, visando à qualidade da classificação dos recursos informacionais. Contudo, os classificadores, em suas práticas de trabalho, podem classificar um assunto de diversos modos, recebendo influência de alguns fatores, como a experiência do classificador e dos usuários que irão utilizar os objetos classificados, fato esse que indica que a classificação pode ser passível de falhas (PIECADE, 1977).

Para evitar falhas no processo classificatório, é importante compreender que existem diversos tipos de classificação bibliográfica (TRISTÃO *et al.*, 2004). Dentre as tipologias, esses autores citam três, a saber:

- ✓ **especializadas e gerais:** as classificações são especializadas quando objetivam classificar um assunto em particular como, por exemplo, o sistema de classificação da *United Classification for the Construction Industry* (Uniclass), utilizado para a indústria de construção; as classificações são gerais quando abrangem um domínio do conhecimento como, por exemplo, a Classificação Decimal de Dewey (CDD);
- ✓ **analíticas e documentais:** as analíticas também chamadas de classificações científicas ou taxonômicas pretendem sistematizar fenômenos físicos com base na explicação e no entendimento como, por exemplo, a classificação do reino animal; as

classificações documentais são usadas para classificar documentos ou tipos de informação, haja vista facilitar a recuperação como, por exemplo, a Classificação Decimal Universal (CDU);

- ✓ **enumerativas e facetadas:** constiuem os tipos de classificação mais conhecidos e utilizados na literatura. As enumerativas são sistematizadas em sistemas numéricos, ao passo que as facetadas estruturam-se em facetas e focos, especificando os assuntos dos documentos (TRISTÃO *et al.*, 2004).

Na opinião de Vizine-Goetz (1996), as classificações enumerativas ou decimais, tais como a CDD e a CDU, são reconhecidas por possuírem uma estrutura hierárquica com números arábicos, que representam as categorias predefinidas de assuntos do universo do conhecimento. Assim, o classificador tem que definir o assunto e representá-lo a partir do número de classe ou notação relacionada, sem precisar criar nenhum número de classe (VIZINE-GOETZ, 1996). Entretanto, segundo esse autor, essas classificações permitem ser atualizadas e, com frequência, precisam ser adaptadas e/ou modificadas.

Desse modo, em um sistema de classificação enumerativo, os assuntos simples ou gerais são subdivididos em assuntos mais específicos, gerando um grande número de assuntos compostos com duas ou mais facetas da mesma classe principal (geral); e complexos, com duas ou mais facetas de diversas classes principais (VIZINE-GOETZ, 1996). Um ponto relevante apontado por Langridge (1977) é que os usuários podem encontrar as classes coordenadas e subordinadas com mais facilidade, e criar um mapa do assunto. Por outro lado, esse esquema é mais complexo para acomodar novos assuntos e podem ser necessárias revisões frequentes (LANGRIDGE, 1977).

Existem diversos sistemas de classificação que permitem classificar o conhecimento nas unidades de informação (SOUZA, 2012). Dentre eles, o autor detalha as características de quatro tipos, que são:

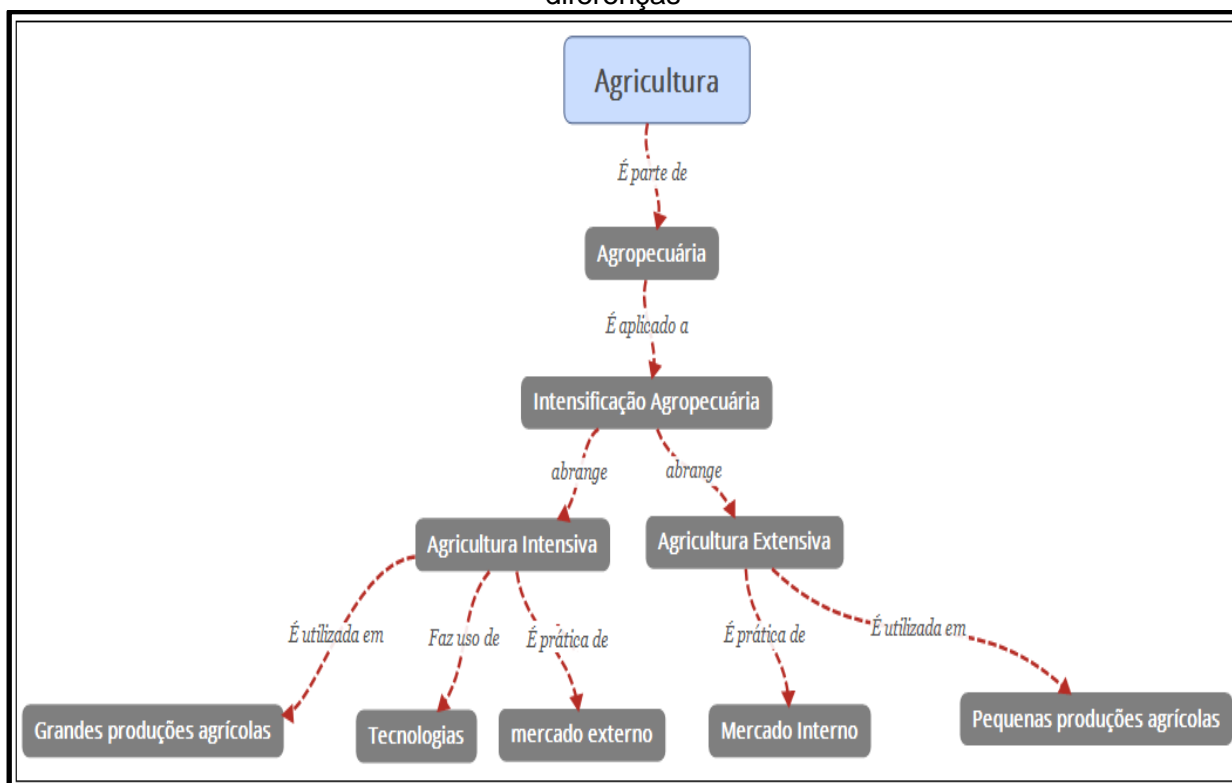
- ✓ **Classificação Decimal de Dewey (CDD):** esse sistema foi elaborado em 1876, com características de um sistema de classificação enumerativo. Nesse sistema, há possibilidade do mesmo assunto estar em vários locais do sistema de classificação de acordo com o contexto, além de o sistema ser complementado pelas tabelas auxiliares comuns que possuem características de assuntos. A CDD possui uma estrutura rígida que divide o conhecimento humano em dez classes, cuja notação é baseada em números arábicos e em sequência decimal;
- ✓ **Classificação da Biblioteca do Congresso, do inglês *Library of Congress Classification* (LCC):** sistema desenvolvido no ano de 1897, que representa os assuntos do acervo e, com isso, possibilita o acesso à coleção. Esta classificação possui como principal característica a flexibilidade dos assuntos, ou seja, permite a expansão das classes. O sistema LCC divide o conhecimento humano em vinte

classes, com a notação representada por letras maiúsculas e números com divisão geográfica, cronológica e de forma;

- ✓ **Classificação Decimal Universal (CDU):** surgiu no ano de 1905, com uma estrutura enumerativa, bastante flexível e semifacetada. É caracterizado como uma classificação analítico-sintética pelo uso dos princípios da Classificação Facetada. As tabelas principais são usadas para a combinação de assuntos simples, indicados por uma classe principal e sem subdivisões; pela combinação de assuntos compostos que contêm duas ou mais características da mesma classe principal, permitindo subdivisões; e pela combinação de assuntos complexos que contêm duas ou mais características de duas ou mais classes principais.
- ✓ **Classificação Facetada, Dois Pontos ou *Cólon Classification* (CC):** com sua primeira edição publicada em 1933, por Ranganathan, esse sistema é totalmente facetado, formado por classes de assuntos básicos, o que permite relacionar diversos assuntos entre a tabela principal, tabelas auxiliares comuns e tabelas auxiliares especiais. Similar à CDU, no que tange a ser considerada uma classificação analítico-sintética a CC, cuja principal característica é a subdivisão dos assuntos em facetas e foco, no plural *foci*. Como exemplo dos elementos dessa divisão, apresenta-se: na classe “Animais” e na faceta “Aquáticos”, podem ser estabelecidos vários *foci* como, por exemplo, “Peixe”, “Tubarão” e “Baleia” (SOUZA, 2012).

Em correlação com as definições apresentadas por Langridge (1977) e Piedade (1977), Vickery (1980) menciona que a estrutura de uma classificação bibliográfica deve pautar-se no agrupamento por aproximações e por distanciamentos quanto aos assuntos dos documentos. Isso quer dizer que, ao estruturar um sistema de classificação, deve-se “[...] reunir coisas ou ideias que sejam semelhantes entre si, e separar as que apresentam diferenças” (VICKERY, 1980, p. 23). Entende-se que a estrutura apresentada por Vickery (1980) pode ser exemplificada no domínio da agricultura. Nesse domínio, as classes mais distantes como, por exemplo, “Grandes Produções Agrícolas” e “Pequenas Produções Agrícolas”, possuem menor grau de semelhança e os relacionamentos correspondem às características nas quais as classes encontram-se agrupadas, conforme exposto na Figura 3.

FIGURA 3 - Classes vinculadas ao universo da agricultura agrupadas por semelhanças e diferenças



FONTE: Elaborada pelo autor (2020).

Pela Figura 3, infere-se que o domínio da agricultura se divide em grandes e pequenas classes; as últimas possuem um nível mais elevado de especificidade, ou seja, são mais específicas. A análise à figura indica que as divisões vão surgindo, conforme o grau de conhecimento dos classificionistas e do contexto, e considerando o desenvolvimento dos assuntos de um domínio, ao longo do tempo. Portanto, infere-se que a prática de classificar está muito impregnada ao nível de conhecimento dos classificionistas e dos classificadores, envolvendo conhecimentos individuais, experiências, questões culturais, regionais, entre outras, conforme relatado por Piedade (1977). Por essa razão, os modelos tradicionais, por possuírem estrutura rígida, podem apresentar dificuldades ou incompletudes na tentativa de organizar, com exatidão, o grande volume de informações surgidas nas últimas décadas do século XX, sobretudo com o advento das tecnologias digitais (PIEADADE, 1977).

Com base na opinião de Dahlberg (2008), as limitações presentes na estrutura dos sistemas de classificação tradicionais explicam o surgimento de novas edições de classificações com segmentações distintas. Na maioria dos casos, essas segmentações não possibilitam a união entre os termos com as diversas áreas do conhecimento, sendo incapazes de manter uma hierarquia de classes, as quais, no momento da classificação do conteúdo dos recursos, deveriam correlacionar áreas distintas (DAHLBERG, 2008).

Considerando a existência de diversos tipos de sistemas de classificação bibliográfica, é importante identificar as semelhanças e diferenças entre suas características,

como realizado no estudo de Ullah, Khusro e Ullah (2017). Segundo esses autores, no que tange à propriedade relacionada à complexidade estrutural, essa característica está relacionada com as classificações que possuem notações e regras complexas, tais como a Classificação Facetada e a CDU, cuja complexidade estrutural da primeira é maior que a da segunda. Ao contrário disso, a CDD e a LCC apresentam complexidade estrutural mais simples, comparadas às demais, conforme evidenciado no estudo de Satija e Martínez-Ávila (2015).

Quanto à brevidade notacional dos sistemas de classificação, Chatterjee *et al.* (2016) consideram a estrutura da CDD e da LCC com brevidade notacional, ou seja, apresentam notações curtas. As classificações que utilizam notações longas e complexas são a CC e a CDU, sendo elas mais flexíveis, por conseguinte, não possuem estrutura predefinida ou rígidas, onde a classificação não depende tanto do classificador, como acontece na estrutura da CDD e da LCC (CHATTERJEE *et al.*, 2016).

De acordo com as características dos diferentes sistemas de classificação citados, Tennis (2011) descreveu que a complexidade de regras determina o nível de dificuldade na aplicação delas em artefatos do conhecimento. A Classificação Facetada, a CDU e a LCC apresentam um conjunto de regras, cujos processos de desenvolvimento e o entendimento delas são mais complexos, o que não é percebido na CDD, visto que constitui um sistema relativamente simples, por possuir assuntos mais abrangentes e completos (TENNIS, 2011).

A CDD e a LCC possuem baixa complexidade de busca mediante o grande valor mnemônico que possuem (notação e termos), em comparação com a CC e a CDU (TENNIS, 2011). Para Ullah, Khusro e Ullah (2017), a CDD, CDU e a LCC são consideradas classificações enumerativas e hierárquicas, baseadas no princípio de dividir o universo do conhecimento em classes definidas. A CDU possui um número maior de conceitos com diversos aspectos predefinidos, afirmando que ela pode ser considerada o único sistema de classificação que possui uma estrutura facetada. A estrutura classificatória da CC não é enumerativa, por isso, não permite a interoperabilidade, por não possuir códigos ou símbolos para que facilite o link entre os conceitos e nem possui a busca facetada, por possuir apenas as classes básicas de assuntos em poucas categorias predefinidas, diferente dos demais sistemas de classificação bibliográfica (RANGANATHAN, 1968).

Em suma, percebe-se que há semelhanças e diferenças entre os diversos sistemas classificatórios. Com o fim de apresentar um esclarecimento mais nítido desses sistemas, apresenta-se o Quadro 3, o qual compara as características das classificações, inclusive da CDU.

QUADRO 3 - Comparação entre os sistemas de classificação bibliográfica

MÉTRICAS\ CLASSIFICAÇÃO	CLASSIFICAÇÃO FACETADA	CLASSIFICAÇÃO DECIMAL UNIVERSAL	CLASSIFICAÇÃO DECIMAL DE DEWEY	CLASSIFICAÇÃO DO CONGRESSO
Complexidade estrutural	√	√	X	X

Brevidade notacional	X	X	√	√
Estrutura predefinida	√	X	√	√
Complexidade de regras	√	√	X	√
Complexidade de busca	√	√	X	X
Interoperabilidade	X	√	√	√
Estrutura facetada	X	√	X	X
Busca facetada	X	√	√	√

FONTE: Adaptado de Ullah, Khusro e Ullah (2017, p. 54).

LEGENDA: √: indica que contém o item correspondente.

X: indica que não contém o item correspondente.

No contexto desta tese, a classificação mais associada à temática é a Classificação Facetada. O motivo deve-se pelo fato da CDU já possuir os princípios da classificação facetada em sua estrutura antes mesmo de Ranganathan propor a CC e a sua teoria, sendo apresentados nos próximos subcapítulos uma breve descrição dos princípios da classificação facetada e logo após, será apresentada a CDU de forma mais detalhada.

3.1.1.1 Classificação Facetada

A classificação facetada surgiu no século XX e foi proposta pelo indiano Shiyali Ramamritam Ranganathan. Os estudos desenvolvidos por esse autor possibilitaram identificar a importância das classificações bibliográficas, cujos argumentos utilizados foram suficientes para ampliar as discussões e os estudos sobre a temática, surgindo, portanto, a Teoria da Análise Facetada (TAF) (BARBOSA, 1972).

Os princípios da Teoria da Análise Facetada (TAF) de Ranganathan surgiu como base para a criação da Classificação Facetada ou Classificação de dois pontos (*Cólon Classification*), permitindo organizar e representar o conhecimento científico em qualquer domínio, por meio do agrupamento de termos que possuem características comuns em diversas categorias (facetas) de assuntos, que poderão dar origem a outras subcategorias (subfacetas) (BARBOSA, 1972).

Nesse sentido, Ranganathan (1967) determinou cinco categorias fundamentais para facilitar a combinação de assuntos, que são: personalidade (P), matéria (M), energia (E), espaço (S) e tempo (T). As categorias criadas por Ranganathan ficaram conhecidas, na literatura, pela sigla PMEST, para representar qualquer área do conhecimento (RANGANATHAN, 1967). Ranganathan observou que cada categoria representa uma característica de um domínio do conhecimento: a personalidade é a característica que representa o assunto; a matéria é o material físico que o assunto pode ser composto; a energia constitui a ação que ocorre referente ao assunto; o espaço é o componente geográfico da localização de um assunto; e o tempo representa a relação temporal associada com cada assunto (RANGANATHAN, 1967).

Silva e Neves (2010) exemplificam os termos representativos e as categorias definidas por Ranganathan, a partir da ordem de citação entre as facetas. Os autores citam um exemplo associado à busca e à recuperação de um livro, cujo assunto principal é “Classificação realizada pelo bibliotecário na biblioteca central hoje”. Nas palavras dos autores, a descrição do assunto procurado pode ser assim:

[P] Personalidade (*personality*) é representada pelo termo “bibliotecário”. O [M] representa Matéria (*matter*), como, por exemplo, “livro”. Já o [E] representa Energia (*energy*), ação de Personalidade, por exemplo, “Classificação”. Quanto ao [S], Espaço (*space*), lugar: “biblioteca central”. E, por fim, o Tempo (*time*) [T] como, por exemplo, o termo “hoje” (SILVA; NEVES, 2010, p. 7).

Após a formulação da Teoria da Análise Facetada (TAF), surgiu a proposta do *Classification Research Group* (CRG), em 1952, na tentativa de ampliar as categorias estabelecidas por Ranganathan. Assim, o grupo sugeriu a inserção de 13 categorias, a saber: entidade, tipo, parte, propriedade, material, processo, operação, agente, paciente, produto, subproduto, espaço e tempo (BROUGHTON, 2006). De acordo com a autora, além das categorias apresentadas por Ranganathan e pelo CRG, novas categorias podem ser incluídas, conforme a especificidade do domínio que está sendo modelado, por meio da identificação dos conceitos.

Nesse sentido a Classificação Facetada é mais flexível, ou seja, possui a capacidade de representar assuntos em múltiplas dimensões, listando as várias categorias (facetas) de cada disciplina ou classe principal, além de fornecer um conjunto de regras para construir números de classe por meio da análise de facetas (RANGANATHAN, 1967). Em síntese, segundo o autor, os esquemas de classificação facetados consolidam-se como uma alternativa para as deficiências encontradas nos sistemas de classificação tradicionais e enumerativos.

Ranganathan (1967) afirma que sua classificação recorre a processos de análise e de síntese, combinando o método hierárquico com o analítico-sintético. Na abordagem analítico-sintética, o assunto de um determinado documento é dividido em subdivisões de assuntos e, em seguida, o sistema de classificação é usado para encontrar notações para cada subdivisão, que serão combinadas de acordo com um conjunto de regras para a construção do número de classificação (RANGANATHAN, 1967). Ainda profere que, após as subdivisões, os assuntos são representados por meio de elementos simples pelo processo de análise; logo, os assuntos podem ser combinados ou decompostos pelo processo de síntese, utilizando elementos que pertencem a mais de uma categoria e que possuem uma notação.

Quanto ao processo de análise, Ranganathan (1967) aponta que é realizado pelo classificador para indicar o assunto do recurso informacional; a síntese, por sua vez, é usada para relacionar o assunto desse recurso com a notação correspondente. Contudo, uma

notação representa o conjunto de termos de um sistema de classificação por meio de letras, números, sinais ou símbolos matemáticos, sendo exibida no recurso informacional ou no catálogo da biblioteca (VICKERY, 1990). Para construir uma notação, o classificador deve seguir um conjunto de regras, cuja notação serve como um link entre o índice do recurso e a classe do sistema de classificação, e/ou entre o catálogo da biblioteca e as estantes (VICKERY, 1990).

Nos sistemas de classificação analítico-sintética são previstos os problemas que ocorrem em sistemas de classificação enumerativos, pois, ao fornecer várias tabelas, notações específicas e regras, evita-se a utilização de uma lista com muitas classes e, portanto, será um sistema de classificação menor em tamanho (LUCAS; CORRÊA; EGGERT-STEINDEL, 2016). Além disso, conforme Lucas, Corrêa e Eggert-Steindel (2016), esse método classificatório fornece flexibilidade aos usuários, que podem construir notações específicas e o classificador não fica restrito à disponibilidade de um assunto específico; porém, o trabalho dos classificadores será mais complexo, pois precisarão construir as notações, ao invés de apenas selecionar uma notação em uma lista (LUCAS; CORRÊA; EGGERT-STEINDEL, 2016).

De acordo com Tristão *et al.* (2004), ao utilizar a Classificação Facetada, é possível fragmentar um assunto complexo pela síntese de mais de uma faceta, que indica um conceito diferente em relação às outras categorias ou facetas. Nesse sentido, a classificação analítico-sintética segue os princípios e técnicas que permitem sintetizar vários aspectos de um determinado assunto, mediante os múltiplos conceitos que são relacionados, tendo em vista organizar e facilitar a recuperação. Assim, é possível descrever e relacionar os assuntos de modo infinito, por meio da utilização de um símbolo de dois pontos (:), de modo a possibilitar inúmeras dimensões ao conhecimento, mesma abordagem utilizada na construção de tesouros e ontologias (TRISTÃO *et al.*, 2004).

Segundo Barbosa (1972), os termos de uma Classificação Facetada são organizados nas subfacetas com características mutuamente exclusivas, não permitindo a sobreposição, ou seja, as facetas estabelecidas determinam a ordem de citação nas quais serão apresentadas no sistema de classificação, ou seja, de forma hierárquica. Com isso, os elementos devem ser ordenados, seguindo a ordem de arquivamento na base de dados, de modo que seja possível inserir o assunto geral antes do específico (BARBOSA, 1972). Embora tenha característica, também, de um sistema enumerativo, Vickery (1980) menciona que o sistema facetado prevê o processo de síntese para construir a notação da classe; logo, os termos que representam o assunto de uma classe possuem uma notação, que é baseada no agrupamento de duas ou mais classes. Enfim, o sistema de classificação pode receber uma notação, que deverá ser flexível para permitir a inclusão de novas classes, e cada índice de

classificação deverá conter todos os termos e suas respectivas notações agrupadas (VICKERY, 1980).

No estudo de Gomes e Campos (2019), constatou-se que o método de faceta utilizada na classificação de Ranganathan é bastante utilizado em portais a partir dos formulários de entrada de dados, por exemplo, cada categoria pode oferecer ao usuário a taxonomia apropriada para preencher dados relativos a filmes. Desse modo, a categoria duração, que também é uma das facetas da taxonomia, pode oferecer ao usuário uma lista com os vários tipos de filmes, como, por exemplo, curta, média e longa ou o inverso, que seria longa, média e curta, ou seja, a que seja mais usada (GOMES; CAMPOS, 2019). O método também permite, segundo as autoras, a inclusão de metadados ou de outras facetas, por exemplo: gênero, terror, romance, ação, entre outras, além de ser possível adicionar a categoria origem, tal como nacional ou estrangeiro, dentre outras possibilidades cabíveis para representação dos dados dos objetos.

3.1.1.2 Classificação Decimal Universal

A CDU surgiu no final do século XIX, no ano de 1895, em Bruxelas, durante a Conferência Internacional de Bibliografia, quando os belgas Paul Otlet e Henry La Fountain receberam a tarefa de organizar o Repertório Bibliográfico Universal (PIEIDADE, 1977). Com essa missão, Piedade (1977) menciona que os autores utilizaram como base a CDD, sobretudo por ela facilitar a organização dos documentos nas bibliotecas, seja de maneira impressa ou digital. Otlet e La Fontaine constataram que a CDU pode se caracterizar como uma taxonomia do conhecimento humano e que poderia ser expressa em uma linguagem internacional (PIEIDADE, 1977).

Pelo fato de os esquemas de classificação existentes não serem adequados a contemplar uma abrangência universal, foi criada a primeira edição da CDU em francês entre os anos de 1902 a 1907 intitulada “Manual do Repertório Bibliográfico Universal”, em francês “*Manuel du Répertoire Bibliographique Universel*”, com 33.000 subdivisões de classes de assuntos (SOUZA, 2004). Posteriormente, entre os anos de 1927 a 1933, foi publicada a segunda edição também no idioma francês, conhecida como *Classification Décimale Universelle* e contendo 70.000 subdivisões (SOUZA, 2012).

Segundo Andrade, Bruna e Sales (2011), a terceira edição foi publicada de 1934 a 1951, em alemão, incluindo 140.000 subdivisões. Outros idiomas que possuem edições completas da CDU são: espanhol japonês, sendo que a publicação da tradução em inglês foi iniciada em 1943, intitulada "Classificação Decimal Universal" e foi designada como a quarta edição internacional. Em 1961, a *British Standards Institution* (BSI) divulgou a terceira

edição revisada no idioma inglês e, logo após, a edição abreviada da CDU foi publicada também em outros idiomas (ANDRADE; BRUNA; SALES, 2011).

Desde o ano de 1949, as propostas de alteração das edições da CDU, de acordo com Slavic (2004b), são publicadas no periódico "Extensões e correções do UDC", a cada ano, em conjunto com as últimas alterações aprovadas pelo consórcio da CDU, do inglês, *UDC Consortium* (UDCC). O UDCC é uma instituição de pesquisa sem fins lucrativos que está sob os cuidados da lei holandesa, sendo responsável pela gestão e manutenção da CDU, iniciada após a dissolução da Federação Internacional de Informação e Documentação (FID), no ano 2000 (SLAVIC, 2004b). O trabalho do comitê do UDCC é auxiliado, direta ou indiretamente, pelos Comitês Nacionais, cujos assuntos especiais em cada país ou internacionalmente possuem um procedimento descentralizado de revisão da CDU. Assim, caso os usuários da CDU queiram sugerir alterações, elas precisam ser enviadas ao órgão responsável de seus respectivos países (SLAVIC, 2004b).

As mudanças ou atualizações da CDU são comunicadas aos seus usuários por meio do periódico responsável por sua publicação, que assumiu essa responsabilidade a partir do final de 1991. No entanto, os programas de pesquisa realizados nos Estados Unidos e em outros países como Grã-Bretanha, Alemanha, Dinamarca e Suíça, nos anos 1960, ajudaram a CDU a ser utilizada como uma linguagem de indexação para o controle informatizado e de processamento de informações em diversas áreas do conhecimento (UDC CONSORTIUM, 2020).

O consórcio da CDU foi formado pela FID, *Business Standards Company* (BSI) e quatro outras editoras como membros fundadores, com o intuito de reestruturar as finanças, gerir e fazer a manutenção do esquema e garantir a sua existência, cujo cargo de editora-chefe está sob responsabilidade de Aida Slavic, em 2021 (SOUZA, 2012). Em suma, o UDCC é responsável pela manutenção da CDU, permitindo que o sistema seja reconhecido como uma classificação totalmente facetada, e sejam removidas as inconsistências que permeiam a classificação e dificultam o fornecimento de uma estrutura sintética. Além disso, o UDCC tem o objetivo de eliminar conceitos compostos e enumerar apenas assuntos simples, que constituem parte da síntese (UDC CONSORTIUM, 2020).

Nesse contexto, o UDCC publicou edições completas da CDU nos idiomas inglês, francês e alemão, com cerca de 220.000 entradas válidas, e esses idiomas continuaram sendo o meio para disseminação da CDU de 1949 até 1992. Após esse período, foi definido o inglês como a língua oficial da UDCC, considerando as mudanças nas extensões das classes até o nível de edição completa, a fim de serem inseridas no banco de dados *Master Reference File* (MRF), no futuro (SOUZA, 2012).

A versão em inglês da CDU¹² é uma edição padrão, completa e mais atualizada, com cerca de 72.000 classes estendidas e com mais de 12.000 registros de dados históricos dos números cancelados da CDU. O serviço da CDU online possui vários recursos para facilitar a construção de novos números na CDU, além da pesquisa, navegação, análise, validação e arquivamento. Como também são exibidos os números cancelados com o objetivo de ajudar os usuários a rastrear estes números, para verificar as substituições, exclusões ou redirecionamentos (remissivas) dessas classes ou subclasses, pelo fato dessas classes que foram canceladas possam ainda existirem coleções de centros de informações e bibliotecas (UDC CONSORTIUM, 2020). No Brasil, a primeira edição foi publicada em 1976 na língua portuguesa, pelo Instituto Brasileiro de Informação e Tecnologia (IBICT); já a segunda e a terceira edição foram publicadas nos anos de 1987 e 1997 respectivamente (BUXTON; MCILWAINE, 1998).

Ademais, o banco de dados *Master Rerence File* (MRF), a princípio, foi projetado usando o pacote de software *Computerized Documentation System - Integrated Set of System* (CDS/ISIS) do projeto Sistema de Informação das Nações Unidas em Ciência e Tecnologia (UNISIST) em conjunto com a Organização das Nações Unidas para Educação, Ciência e Cultura (UNESCO), a fim de ser usado em um sistema de informação científica mundial, que ofereça suporte a padrões internacionais para o formato de intercâmbio de dados bibliográficos (ISO 22709, 2011). No início, o MRF possuía cerca de 60.000 classes, escolhidas como uma seleção gerenciável e de modo representativo das informações e conhecimentos registrados, aumentando para 68.000 classes, no ano de 2010 (ISO 22709, 2011). Após o ano de 2010, de acordo com o consórcio da CDU (*UDC Consortion*), esse sistema de classificação tornou-se um dos sistemas de classificação mais usados em bibliotecas, bibliografias, centros de documentação e informações; também foi utilizada em torno de 130 países e publicada em mais de 40 idiomas (SOUZA, 2012).

Mcilwaine (1997) conceitua a CDU como um sistema classificatório de abrangência universal, pois contempla todo o campo do conhecimento, além de ser uma ferramenta de classificação de assuntos multilíngue para organizar todos os tipos de conhecimento registrado na biblioteca. Para Buxton e Mcilwaine (1998), a CDU é um instrumento de indexação e recuperação da informação do conteúdo documental, que visa a organizar, de modo semântico, os assuntos pertencentes a qualquer domínio do conhecimento, tendo em vista ampliar o meio de acesso ao recurso informacional (BUXTON; MCILWAINE, 1998).

Na visão de Lancaster e Warner (1993), a CDU é um instrumento de representação do conhecimento utilizado para organizar acervos bibliográficos em um

¹² Disponível em: <https://udc-hub.com/en/login.php>.

catálogo. Tal instrumento de classificação é conhecido por ser independente do contexto e da linguagem, pois consegue resolver, de maneira implícita, os problemas de sinonímia e elimina as ambiguidades. Entretanto, um conceito pode ter vários significados, a partir do contexto a que se referem e os usuários podem não ter ciência dessa diferenciação, cabendo à CDU especificar o significado do termo (LANCASTER; WARNER, 1993). Em qualquer caso, a CDU pode ajudar a resolver problemas de homônimos e polissemias da linguagem natural, mostrando a posição ou contexto do termo na estrutura da organização do conhecimento, bem como a inclusão dos termos e suas variações que denotam um ou mais conceitos. Portanto, a CDU auxilia o usuário na escolha do termo que corresponda ao assunto referenciado, sendo que os demais assuntos relacionados e as estratégias de desambiguação precisam estar incluídos no sistema de classificação (LANCASTER; WARNER, 1993).

Segundo Souza (2012), a CDU conta com a cooperação internacional que hospeda e faz as atualizações e revisões necessárias, com o propósito de acrescentar informações na base de dados única da CDU, o *Master Reference File* (MRF), a qual possui diversas versões e idiomas. Além disso, esse sistema contém a versão definitiva e autorizada da CDU, mantida pelo consórcio do sistema, o qual gerencia esse banco de dados, de modo a preservar o conteúdo da CDU, priorizando as necessidades das revisões e acompanhando as mudanças, ao longo do tempo. Além disso, o conteúdo do sistema é salvo de maneira duplicada na base de dados, para facilitar as manutenções ou atualizações regulares durante o ano e, assim, cria-se um novo arquivo-mestre a cada ano (SOUZA, 2012). Com efeito, a nova versão do MRF contendo as alterações em cada ano é usada para fornecer a saída, para distribuir o conteúdo aos membros do consórcio e detentores da licença e para inserção nos aplicativos de software que estiverem aprovados para utilizar o MRF da versão completa da CDU (SOUZA, 2012).

Além da versão completa da CDU, foi criada a versão abreviada da CDU com o título *UDC Summary* (UDCS), que oferece uma seleção de cerca de 2.600 classes de todo o esquema, o que corresponde a mais de 70.000 entradas e está disponível em 57 idiomas (UDC CONSORTIUM, 2020). Importante salientar que esse site foi utilizado neste estudo como ambiente de consulta, haja vista identificar as subclasses da classe 631 (Agricultura, em geral), o recorte temático da CDU¹³ abreviada foi investigada no idioma português, bem como identificar outras classes.

A CDU é um instrumento que se caracteriza pelo arranjo de livros, documentos, entre outros recursos informacionais em bibliotecas ou centros de informação, aproximando de uma linguagem de indexação, cuja finalidade é organizar itens em índices de catálogos, servindo como ponto de acesso entre o catálogo e o item na estante. Miranda (2005) verificou

¹³ Disponível em: <http://www.udcsummary.info/php/index.php?lang=pt&pr=Y>.

que a CDU pode gerar códigos que proporcionam a organização do conhecimento em um SRI, indicando que esse tipo de classificação é considerado abrangente, com uma linguagem universal, que serve de padrão ao integrar os assuntos correlatos para atividades de classificação em unidades de informação (MIRANDA, 2005).

Desse modo, a CDU possui uma estrutura menos rígida, mais ampla para incluir facetas ou aspectos particulares de um tópico de assunto, sendo caracterizada, também, como uma classificação semifacetada (PIEDADE, 1977). Por esse motivo, é considerada mais flexível e com capacidade analítico-sintética, haja vista representar um assunto e ordenar os conceitos de uma área de conhecimento em facetas (classes ou categorias), cujos termos que representam um assunto de um recurso informacional podem estar em mais de uma faceta (PIEDADE, 1977).

Vukadin e Slavic (2014) caracterizaram a CDU como um sistema de classificação desenvolvido, sobremaneira, com o objetivo de indexar e organizar o conteúdo de grandes acervos bibliográficos, incluindo não apenas livros, mas também todo o tipo de recurso informacional que congrega facetas e hierarquias. Pelo potencial de relacionamentos que possui, a CDU aproxima-se da análise facetada, sustentada pelo princípio da análise e da síntese, que é capaz de permitir a combinação de facetas, de modo a representar qualquer tipo de assunto de um recurso (VUKADIN; SLAVIC, 2014).

Com esse raciocínio, Kaosar (2008) refere-se à CDU como a combinação da análise facetada com um esquema hierárquico e enumerativo; logo, nessa perspectiva, acredita-se que a CDU pode ter um bom desempenho na recuperação, no comparativo com outros esquemas como, por exemplo, a CDD e a Classificação Bibliográfica de Bliss (BC2)¹⁴. Isso porque, diferente desses sistemas, a CDU estabelece a classificação a partir do uso de facetas e esquemas enumerativos, além dos relacionamentos semânticos entre conceitos. A natureza analítico-sintética da CDU, de acordo com a autora, é baseada na organização relacional do vocabulário em facetas, as quais possuem significados próprios e hierárquicos de conceitos mutuamente exclusivos. A partir de sua estrutura semifacetada, é possível o reúso de conceitos em diferentes combinações para expressar assuntos com um nível maior de detalhes ou assuntos complexos (KAOSAR, 2008).

Outra característica importante da CDU é que ela possui relacionamentos semânticos, por exemplo, gênero-espécie, todo-parte, dentre outros, e esses são

¹⁴ Constitui um sistema elaborado até minúcias não previstas por outros classificadores, tal como o excesso de tabelas auxiliares a serem usadas em determinadas classes. Dá liberdade ao classificador para colocar certos assuntos alternadamente em classes diferentes. As regras são inadequadas para uso consistente das tabelas auxiliares para assuntos compostos. Como o sistema é rico em opções, é necessário que o usuário decida cuidadosamente qual opção utilizar em cada caso. É um dos sistemas menos adotados e particularmente preferidos pelas bibliotecas de Educação (UNIVERSIDADE DE SÃO PAULO, 2015).

coordenados por subdivisões hierárquicas de assuntos (SIMÕES, 2010). Nesse sentido, a autora destaca que é preciso utilizar classes mais específicas e símbolos distintos para diferentes relacionamentos, além do conjunto de regras para ordenar os assuntos dentro dos relacionamentos.

Essas relações são sistematizadas à medida que os conceitos são introduzidos, eliminados, reintegrados ou editados, e as relações sintáticas podem ser observadas na CDU por meio das tabelas auxiliares comuns e especiais, conforme esclarecido por Mcilwaine (1997). Logo, quando as relações se estabelecem dentro de uma mesma classe ou subclasse, elas representam assuntos compostos, já quando estão em classes ou subclasses diferentes referem-se a assuntos complexos. Por fim, as relações semânticas da CDU são responsáveis pelo significado do vocabulário usado para representar os assuntos dos documentos (MCILWAINE, 1997).

Quando o vocabulário dos assuntos está conectado a partir das relações semânticas no sistema de classificação, os usuários dos sistemas de informação, além de utilizarem a linguagem natural, podem usar as notações, que são expressões já definidas para um assunto ou combinação deles, ao passo que o vocabulário é formado por termos que são usados durante a busca (DAHLBERG, 1971). Contudo, para que os usuários realizem a busca por expressão notacional, o sistema de informação deve estar preparado com a combinação de elementos notacionais ou números de classes dos assuntos de modo pré-estabelecido, podendo utilizar as tabelas principais e auxiliares da CDU. Por essa razão, deve-se preservar a ordem de citação das classes do sistema para que o formato original de seus dados não seja modificado, de maneira que esses dados sejam reconhecidos e recuperados pela máquina (DAHLBERG, 1971).

Para Svenonius (1983), um número de classe da CDU é pré-combinado durante o processo de indexação e a construção da notação é facilitada por dispositivos notacionais e pelo conjunto de regras que determina a combinação desses números de classes, com respeito à ordem de citação. Na maioria dos casos, os números de classes da CDU são preservados em seu formato original, para que os dados de seus elementos estejam estruturados a ponto de serem reconhecidos em qualquer sistema de informação, bases de dados, idiomas, visando à interoperabilidade e à recuperação da informação de modo eficiente (SVENONIUS, 1983).

Salienta-se que a interoperabilidade interfere no potencial da CDU, considerando o modo de estruturação do sistema, que aproveita recursos e permite mais agilidade nas práticas de representação. Assim, na CDU, para a representação das categorias (entidades) e dos termos, é utilizado um código único (FRÂNCU, 2003a). Portanto, essas características do sistema trazem consistência, eficácia e reconhecimento a ponto da CDU ser utilizada por diversas instituições de informação (ANDRADE; BRUNA; SALES, 2011).

O sistema da CDU pode ser utilizado de forma pós-coordenada no momento da busca, sendo tanto por meio de notações, quanto de forma textual, com o uso de operadores booleanos para restringir a pesquisa, e com uso das classes de uma ou mais disciplinas do conhecimento, conforme as relações estabelecidas de maneira interdisciplinar, multidisciplinar e não linear (SVENONIUS, 1983). Portanto, a CDU facilita a manipulação do vocabulário, independente do uso de termos utilizados em facetas mutuamente exclusivas. Nesse contexto, a busca pós-coordenada da CDU pode ser feita de forma mais abrangente, utilizando o operador “?”, por exemplo: 8?: essa representação significa que será considerada a busca por qualquer país da América do Sul (SVENONIUS, 1983).

A partir do uso da CDU em sistemas de biblioteca e na web, é evidente que, mesmo que o conteúdo intelectual de uma classificação possa ser útil, esse sistema ainda é pouco usado na recuperação da informação, sendo necessário que seu conteúdo seja legível por máquina e distribuído em algum tipo de formato padronizado para ser processado por sistemas de informação (SAN SEGUNDO, 2009). Mesmo com o uso do formato *Machine Readable Cataloging* (MARC), isso não garante o emprego efetivo da CDU na web, já que a classificação começa a ser compartilhada fora de um domínio específico. Por esse motivo, devem ser utilizados padrões para implementação e compartilhamento de SOC a partir de um domínio (SAN SEGUNDO, 2009).

Cox (1992) enfatiza a importância da CDU tanto na indexação e classificação quanto na recuperação, estando o sistema baseado em facetas (categorias), que garantem a representação em uma linguagem e exibem o resultado esperado por usuários de sistemas de informação. Portanto, o sistema agregará documentos que lidam com o mesmo campo de assunto, porém em diferentes níveis de especificidade, visto que a classificação em mera ordem alfabética poderia resultar em uma dispersão do campo de assunto, sem levar em conta todos os conjuntos de termos que representam um assunto (COX, 1992).

No processo de navegação por assunto, de acordo com Vickery (1990), uma classe da CDU totalmente expandida, muitas vezes, pode ser muito detalhada, abrangente e difícil para os usuários obterem uma visão geral de uma área de assunto e a direção a qual se quer navegar. Os pontos de entrada para navegação são muito importantes e devem ser sugeridos e evidenciados pelo sistema de classificação. Assim, a navegação se inicia por seleção de modo aleatório (*top / down*), ou por uma pesquisa no índice de assuntos, que deve ser de forma rápida e flexível, principalmente quando se trata de estruturas de sistemas de classificação robustos, como é o caso da CDU (VICKERY, 1990). Portanto, é necessário que se possa “reduzir” e “expandir” os assuntos de uma hierarquia, com o objetivo de direcionar o usuário no caminho correto, durante a busca, em virtude dos relacionamentos semânticos que conectam os assuntos (VICKERY, 1990).

Por ser bastante utilizada em diversas instituições de informação, no Brasil e no exterior, infere-se que a CDU possui muitas funções importantes. Segundo Vickery (1990), a CDU tem como função a organização de documentos, com o propósito de auxiliar os usuários de biblioteca nas buscas por bibliografias universais e especializadas (VICKERY, 1990). Em linhas gerais, a CDU permite classificar os documentos, conforme a evolução do conhecimento humano, partindo dos assuntos gerais para os específicos. Nota-se que, a partir da CDU, é possível recuperar documentos por meio de diferentes áreas de assuntos, usando a combinação de termos em relação a uma determinada classe ou área do conhecimento, como também pelos próprios termos do dicionário baseado na CDU. Nesse contexto, a CDU possui a função de organizar o conhecimento de forma lógica e sistemática em catálogos, processo realizado por meio de notações ordenadas de modo hierárquico, facilitando o acesso ao conteúdo dos documentos. Além disso, a CDU possibilita a organização física dos documentos nas estantes, permitindo aos usuários conhecer as obras da biblioteca sobre os mais variados assuntos, mesmo sobre os que não são de seu conhecimento prévio (VICKERY, 1990).

Com esse objetivo, a CDU é essencialmente um sistema prático que codifica, numericamente, as informações de qualquer recurso; uma vez codificado e arquivado corretamente, o recurso pode ser facilmente encontrado durante a busca (SLAVIC, 2005). Portanto, é considerada uma classificação científica com sinais de conexão e relações de forma universal. Outra função da CDU é aprimorar o processamento e intercâmbio de dados bibliográficos referentes ao título, autor, editora, dentre outros detalhes de um objeto, dados esses que, no ambiente digital, são armazenados nas bases de dados, por meio de padrões de metadados (SLAVIC, 2005).

Piedade (1977) descreve a CDU como sendo uma classificação que possui uma estrutura decimal, assim como é a CDD, cujas classes são representadas por códigos numéricos e as notações são geradas com códigos alfanuméricos. Por ser decimal, segundo a autora, é possível dividir o conhecimento a ser representado em dez grandes classes básicas de 0 a 9, denominadas de classes principais e descritas por um um número (um dígito), o que possibilita representar todos os assuntos dentro dos domínios do conhecimento humano. Uma observação é quanto à classe 4, que está vaga desde 1960, quando foi agregada à classe 8, tendo em vista permitir futuras expansões de classes ou disciplinas (PIEIDADE, 1977). As dez classes da CDU são apresentadas no Quadro 4.

QUADRO 4 - Divisões principais da Classificação Decimal Universal

Classe	Descrição
0	Generalidades. Informação. Organização.
1	Filosofia. Psicologia.
2	Religião. Teologia.

Classe	Descrição
3	Ciências Sociais. Economia. Direito. Política. Assistência Social. Educação.
4	Classe vaga.
5	Matemática e Ciências Naturais.
6	Ciências Aplicadas. Medicina. Tecnologia.
7	Arte. Belas-Artes. Recreação. Diversões. Desportos.
8	Linguagem. Linguística. Literatura.
9	Geografia. Biografia. História.

FONTE: Elaborado pelo autor (2020).

Em geral, a estruturada CDU é semifacetada com um índice relativo em ordem alfabética, produzido pela permutação de termos e seus sinônimos para auxiliar na busca dos assuntos e suas subdivisões nas tabelas principais e auxiliares, sendo que as auxiliares permitem a combinação de assuntos básicos para que se tornem assuntos complexos. Para isso, são utilizados os símbolos das tabelas auxiliares para conectar, iniciar ou finalizar os números simples da CDU. Assim, a CDU permite analisar o entendimento de expressões complexas por meio de seu sistema notacional, cuja sintaxe da síntese dos números da CDU pode ser aplicada em toda a estrutura do sistema (BUXTON; MCILWAINE, 1998).

Nesse sentido, a estrutura analítico-sintética da CDU permite, a partir da criação das classes principais do sistema, o estabelecimento de expressões mais complexas, de modo semelhante à descrição das palavras em linguagem natural na criação de frases complexas, tendo como base o sentido semântico (FRÂNCU, 2003a, 2003b). Desse modo, as práticas de organização do conhecimento, mediante a utilização da CDU, possibilitam um aumento significativo na quantidade de termos e relações semânticas, com a finalidade de permitir a representação de um domínio com maior relevância na recuperação (VICKERY, 1990). Com efeito, a recuperação está associada à estruturação do sistema, ou seja, os relacionamentos firmados entre as classes e a estrutura flexível do sistema possibilitam melhorias no processo de recuperação (SOUZA, 2012), o que justifica a necessidade de compreender, em detalhes, a estrutura da CDU, incluindo suas classes, subclasses e tabelas auxiliares.

A maioria dos termos do índice da CDU não indica o contexto em que eles aparecem, por isso a necessidade de qualificadores (notas explicativas) para identificar as entradas, retiradas principalmente da descrição da classe principal (RIESTHUIS, 2003). Isso torna imprescindível que os usuários consultem as tabelas da notação para compreenderem plenamente os diferentes contextos dos termos indexados, e o índice serve apenas como um guia para a localização da notação. Nesse sentido, o índice fornece o acesso mais rápido aos

números da CDU a partir das pesquisas por assuntos, cujos termos estão em linguagem natural e não em um vocabulário controlado (RIESTHUIS, 2003).

Piedade (1977) explica que a estrutura semifacetada da CDU é pelo fato de não conter todas as facetas (categorias e subcategorias), necessárias para representar todo o conhecimento humano, visto que ele está em constante evolução. De modo similar, Barbosa (1972) afirma que esse sistema classificatório utiliza as disciplinas do conhecimento como categorias do pensamento, além da associação das classes e subclasses do próprio sistema com o uso de remissivas (ver também).

A representação de remissivas, que são usadas para orientar o classificador das notações que representam conceitos relacionados ou próximos, mas com outro ponto de vista. Assim, essas remissivas são indicadas pelo símbolo "→", que equivale a "ver também", com o objetivo de apontar para uma ou mais classes associadas a uma determinada classe da CDU, onde são referenciados os conceitos associados semânticamente em diferentes hierarquias (VICKERY, 1990).

É importante salientar que a utilização da estrutura da CDU pode ser compartilhada entre diversos grupos de assuntos, representando qualquer domínio do conhecimento (SOUZA, 2012). Para isso, segundo Slavic (2005), utilizam-se as combinações entre as tabelas presentes na estrutura da CDU, que são:

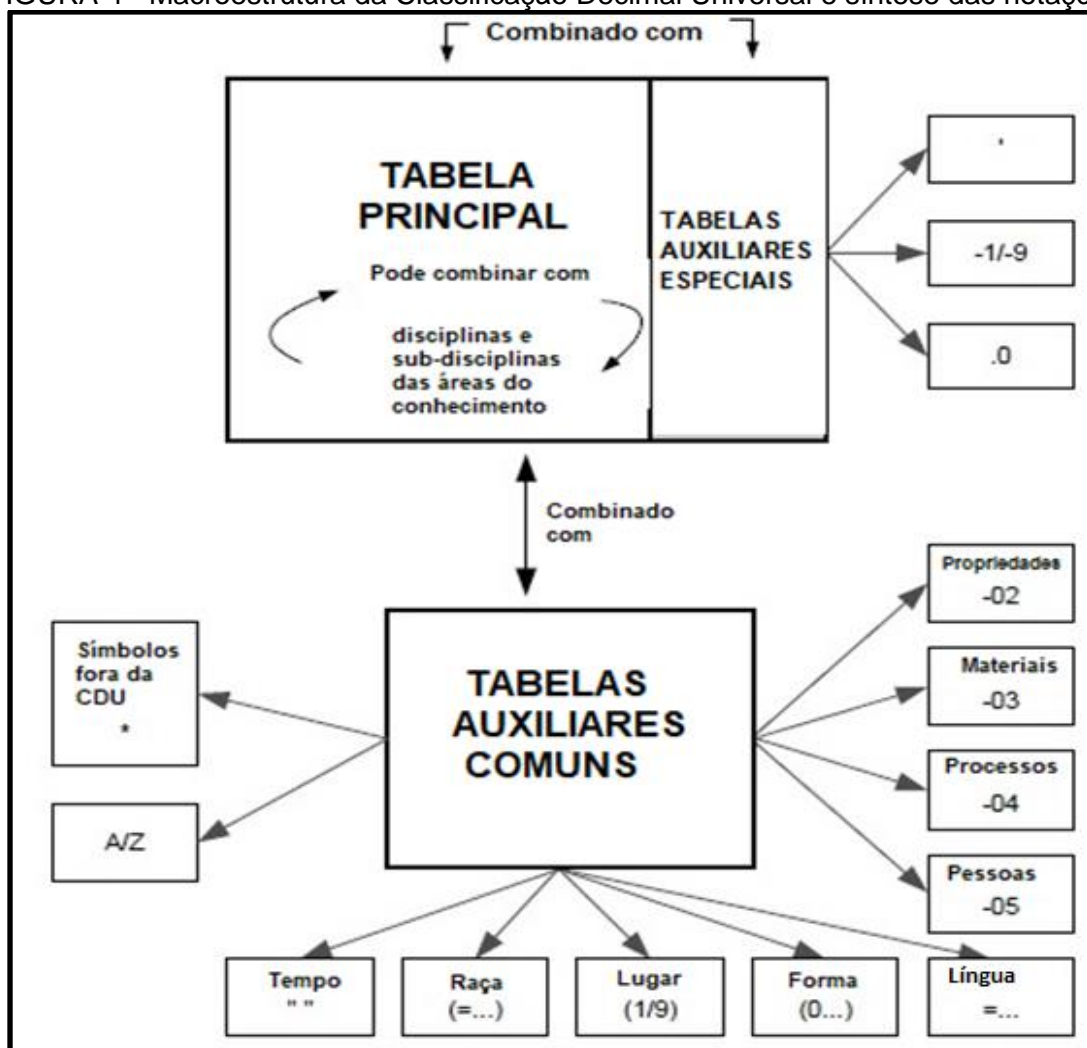
- **Tabelas Principais:** contêm códigos para classificação de disciplinas, subdisciplinas ou campos do conhecimento específicos;
- **Tabelas Auxiliares Comuns:** possuem informações gerais sobre os conceitos, que denotam características comuns a todos, aplicados às tabelas principais e fornecem meios de expressar as interações entre os assuntos de forma hierárquica. Essas tabelas são responsáveis pela eliminação de repetição de conceitos existentes nas tabelas de propriedades, materiais, processos, pessoas, língua, lugar, raça¹⁵, forma e tempo. As subdivisões auxiliares comuns são divididas em independentes e dependentes. As tabelas auxiliares independentes constituem das Tabelas 1ce 1g e as dependentes são encontrados nas Tabelas de 1(i) a 1k. Por fim, as tabelas auxiliares comuns também permitem o uso de símbolos ou sinais que não estão na CDU (" * ") e as letras do alfabeto ("A/Z");
- **Tabelas Auxiliares Especiais:** A maioria dos conceitos das tabelas ou subdivisões auxiliares especiais podem ser incluídos dentro de uma classe ou subclasse de uma tabela principal, levando em conta sua natureza e necessidade para a representação de assuntos. Por este motivo, em alguns casos, pode existir mais de uma subdivisão

¹⁵O agrupamento "raça" é derivado de tabelas auxiliares comuns de língua e de lugar. Assim, cada notação de linguagem pode ser usada para construir uma notação que expresse o agrupamento de raça (SLAVIC, 2005).

auxiliar especial em uma determinada classe principal e ainda possuir indicações de remissivas. Isso proporciona maior nível de especificidade para a representação de assuntos, os quais podem conter vários conceitos que são representados por termos. As classes das tabelas auxiliares especiais são agrupadas de forma hierárquica, assim como as tabelas auxiliares comuns e principais. Portanto, as subdivisões auxiliares especiais auxiliam na definição de conceitos com características de acordo com a localização que estão inseridas no sistema, que na maioria das vezes é de forma enumerativa, com um alto nível de síntese para representar assuntos compostos ou complexos. Assim, é possível combinar assuntos das tabelas auxiliares especiais com a tabela principal, onde o procedimento para ordenação é inserir o número de classe da tabela principal precedida a um ou mais números de classes das subdivisões auxiliares especiais (BUXTON; MCILWAIN, 1998; SLAVIC, 2005). As classes 631.3.02 a 631.3.077 são exemplos de classes da CDU que são subdivisões auxiliares especiais e estão subordinadas a classe principal 631.3 - Máquinas, implementos e equipamentos agrícolas (CLASSIFICAÇÃO DECIMAL UNIVERSAL, 1997, p. 887). Além da forma enumerativa das subdivisões auxiliares especiais na CDU, é possível obter um maior nível de especificidade de assuntos mediante o uso dos símbolos (“ ’ ”, “-1/-9” e “.0”) (SLAVIC, 2005).

O uso dos símbolos mencionados nas tabelas auxiliares especiais podem ser melhor compreendidos a partir da análise à Figura 4.

FIGURA 4 - Macroestrutura da Classificação Decimal Universal e síntese das notações



FONTE: Adaptada de Slavic (2005, p. 452).

A partir da Figura 4, percebe-se que Slavic (2005) analisa a CDU como um sistema de classificação bibliográfica que está fundamentado em uma lógica flexível, capaz de pluralizar as relações, sobretudo pelo fato das notações estarem localizadas em diferentes tabelas. Ademais, as notações possuem origens diferentes, ou seja, um indicador de faceta, estabelecido por uma notação com uma sintaxe, contendo símbolos das tabelas auxiliares comuns e especiais para mostrar o princípio da divisão da classificação facetada que se estabelece entre um conjunto de termos de uma mesma classe ou entre classes diferentes (SLAVIC, 2005).

Salienta-se que cada um dos símbolos das tabelas auxiliares comuns especiais possui um significado, que precisa ser compreendido pelo classificador e pelo usuário, haja vista possibilitar a localização de um item procurado no acervo (SOUZA, 2012). O Quadro 5 descreve os significados de cada um dos símbolos dessas tabelas.

QUADRO 5 - Detalhamento dos símbolos das Tabelas Auxiliares Comuns e Especiais

Símbolos	Descrição	Tabelas da CDU
+	Significa adição (e). Permite a associação semântica (por coordenação) ou não de assuntos separados (não consecutivos) com a tabela principal. É utilizado para representar assuntos compostos quando não existe um assunto simples. Ex.: 631+633.	Tabela 1a
/	Símbolos de extensão consecutivos (até), permitindo que se encontrem conceitos consecutivos na hierarquia de assuntos da tabela principal e auxiliares comuns. Ex.: 631/632.	Tabela 1a
:	Relação simples para representar vários assuntos coordenados ou subordinados da tabela principal e auxiliares comuns e especiais. Ex.: 061:63 - Organizações do ramo da agricultura.	Tabela 1b
::	Relação composta (por ordenação), para estabelecer a ordem que os conceitos serão apresentados, evitando dúvidas na representação da notação de assuntos, quando existe uma relação de subordinação entre conceitos para não alterar o sentido. Ex.: 78::63 – Música sobre agricultura e não o contrário, pois não teria sentido.	Tabela 1b
[]	Sub-agrupamento para uma combinação complexa de uma notação da CDU, esclarecendo as relações entre seus componentes, que, geralmente, são ligados pelo sinal de (+, :) da tabela principal. Ex.: [030+050]:63 – Obras e periódicos da agricultura.	Tabela 1b
*	Notações localmente externas à CDU, como, por exemplo, um número extraído da classificação de Dewey. Nessa ocorrência, é preciso inserir uma nota de escopo, explicando essa situação.	Tabela 1h
A/Z	Especificação alfabética.	Tabela 1h
()	Indica lugar, raça e forma, sendo que raça e forma devem ter os símbolos de “=” e “precedência de números”, respectivamente. Ex.: (1/9), (=...), (0...).	Tabelas 1d, 1e e 1f
“ ”	Este símbolo “ ” expressa o tempo cronológico e deve ter, no mínimo, dois dígitos.	Tabela 1g
-	O hífen é utilizado quando está associado aos auxiliares comuns -02 (Propriedades), -03 (Materiais), -04 (Processos) e -05(Pessoas).	Tabela 1k
=	O sinal de igual é usado para os auxiliares comuns de língua e os auxiliares de pessoas, etnias e nacionalidades devem ser precedidos de parênteses. Ex.: =134.3 – português.	Tabelas 1c e 1f
.	O ponto pode representar auxiliares especiais, por exemplo: .1/.9, ou a especificidade de um assunto de uma tabela principal. Para a tabela principal, o sistema separa os assuntos em até três algarismos para facilitar o entendimento, com exceção das tabelas auxiliares especiais que não utilizam esse princípio.	Tabela Principal
’	O apóstrofo tem a função de síntese, integrando assuntos compostos, por meio de notações complexas. Ele possui elementos da tabela principal com subdivisões paralelas. Ex.: 81’276.2 - Estudo de gírias; 811.111’276.2 – Gírias da língua inglesa.	Tabelas Auxiliares Especiais

FONTE: Elaborado pelo autor (2020). Adaptado de Simões (2010).

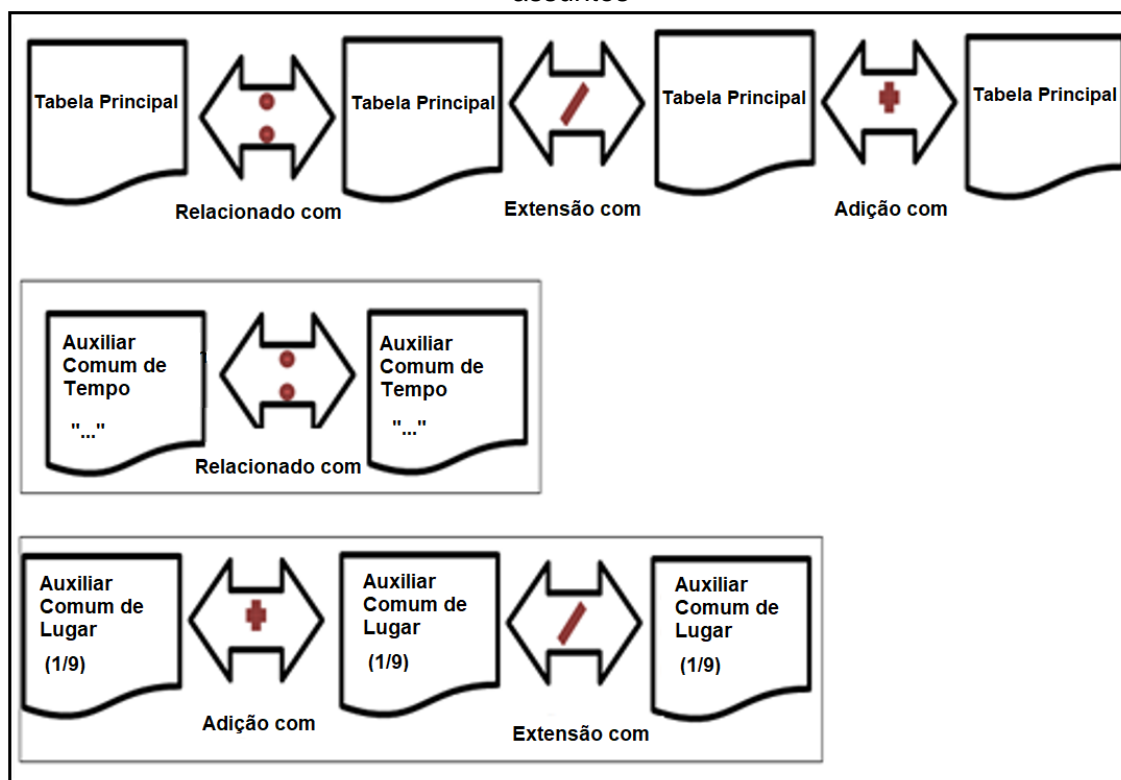
O Quadro 5 especifica, de modo mais detalhado em relação à Figura 4, os símbolos utilizados para relacionar a tabela principal com as tabelas auxiliares comuns e

especiais. Por outro lado, pela análise ao referido Quadro, constata-se que, quando duas ou mais notações são combinadas e/ou detalhadas, é possível determinar onde um elemento termina e o outro começa, conforme descrito por Simões (2010).

Em relação à representação de assuntos de recursos informacionais, Simões (2010) relata que a CDU faz a seleção de combinações úteis de assuntos, sem limitação de conceitos, cuja hierarquia de conceitos pode ser compactada para reduzir a notação. Nesse sentido, a CDU é composta por expressões (símbolos e signos linguísticos auxiliares) tendo em vista acrescentar características à notação do sistema. Ademais, a CDU possui uma linguagem normalizada, com termos uniformes e consistentes, cujo propósito é garantir resultados pertinentes e precisos na recuperação, ou seja, a precisão na recuperação depende da precisão e profundidade da indexação (SIMÕES, 2010).

Slavic (2005) descreve a sintaxe da CDU permitida para que ocorram as combinações entre uma ou mais notações da tabela principal e auxiliares comuns do mesmo tipo como, por exemplo, tabelas auxiliares comuns de tempo, raça, lugar, forma, língua, propriedades, entre outras. A Figura 5 exemplifica a sintaxe permitida, ao utilizar elementos da tabela principal e das tabelas auxiliares comuns, em combinação com os símbolos de relação, extensão e adição, a partir do exemplo elaborado por Slavic (2005).

FIGURA 5 - Sintaxe permitida na Classificação Decimal Universal para combinação de assuntos



FONTE: Elaborada pelo autor (2020). Adaptada de Slavic (2005).

No contexto da Figura 5, os símbolos das relações (“:”, “/”, “+”) podem ser utilizados entre tabelas ou classes principais e auxiliares, embora, na Figura mencionada,

utilizaram-se, especificamente, as auxiliares comuns de tempo e lugar. Nesse sentido, o símbolo de relação (“:”) representa uma sintonia entre um assunto com outro, podendo ser de modo coordenado ou subordinado¹⁶. Em seguida, o símbolo de extensão (“/”) denota uma série de assuntos consecutivos da tabela principal e auxiliares comuns (SLAVIC, 2005). Por fim, o símbolo de adição (“+”) indica assuntos consecutivos para associar assuntos com a tabela principal, formando um assunto composto quando não existe um assunto simples (SLAVIC, 2005). A autora observou que, na utilização da CDU, é possível fazer distinção entre três tipos de síntese no sistema, que são: inclusão direta, combinação com sinais e flexibilidade na ordem dos elementos. Os dois primeiros tipos de síntese de notação são aplicáveis em todo o nível do sistema de classificação. Entretanto, o uso da síntese de flexibilidade na ordem dos elementos não é muito recorrente entre os usuários ou desenvolvedores da CDU, que estão pesquisando uma maneira de substituir essa síntese pelas duas primeiras apresentadas (VICKERY, 1990).

Quanto à navegação por assuntos, a estrutura da CDU estimula a intuição do usuário, pois sua estrutura flexível permite decompor a análise de um determinado assunto em várias dimensões, representando assuntos em diversas categorias, desde os mais simples até os mais complexos (MIRANDA, 2005). Piedade (1977) justifica essa flexibilidade ao mencionar que a estrutura do sistema se torna expressiva, hierárquica, mnemônica e enumerativa. Por esse motivo, o assunto dos documentos é representado por notações que permitem armazenar e navegar por diversos tipos de assuntos, de modo a gerar uma estrutura de conceitos multidimensional. Com isso, cada assunto pode ser representado por notações em diversos aspectos, o que identifica o potencial analítico-sintético da estrutura semântica da CDU (PIEADADE, 1977).

Souza (2012) indica o processo de síntese como sendo dois elementos notacionais misturados, capazes de criar uma nova notação, ao passo que os elementos, quando separados, não são facilmente distinguíveis. Nessa perspectiva, esse tipo de síntese pode ser baseado em uma regra de subdivisão paralela, que é indicada para tal situação, e descreve a maneira pela qual a notação é criada (SOUZA, 2012). A notação de “origem” (821.111) foi sendo alterada e adicionada a um número definido para criar uma notação de “destino”, derivada da notação de origem, por meio da subdivisão paralela, conforme expresso neste exemplo, extraído do estudo de Souza (2012):

=111 Língua Inglesa (Tabela Auxiliar Comum de Língua)

(= 111) Ingleses (População) (Tabela Auxiliar Comum de Lugar)

811.111 Linguística - Língua Inglesa (Tabela Principal)

¹⁶ Subordinado significa que na estrutura classificatória existe uma ou várias classes subordinadas, ou seja, mais específicas em relação a classes superordenadas, que correspondem a classes de assuntos mais genéricos em relação a cadeia em questão.

821.111 Literatura Inglesa (Tabela Principal)

Em linhas gerais, o uso dessas tabelas e dos respectivos sinais/símbolos possibilitará melhorias quanto à revocação na recuperação da informação, desde que a estrutura semântica da CDU esteja preparada para uma pesquisa mais ampla, ou seja, que expresse a hierarquia, caso se obtenha poucos resultados em uma pesquisa específica, o que torna a recuperação mais precisa (RIESTHUIS *et al.*, 1998). Por outro lado, pode-se utilizar o truncamento da pesquisa durante a navegação por assuntos, de modo que comece em uma classe mais ampla, a abranger vários números de classes no mesmo nível de divisão (RIESTHUIS *et al.*, 1998).

Vickery (1990) sugere que cada notação de uma classe seja descrita usando uma legenda em formato de texto, o qual precisa ser associado aos termos de pesquisa adicionais em cada classe, por meio de um índice alfabético de assunto. Por esse motivo, a CDU oferece um dicionário, apresentado no formato de classificação, rede semântica ou lista alfabética complementada pelo sistema de classificação com o número da CDU correspondente (VICKERY, 1990).

Quanto à inclusão direta, o sistema notacional da CDU facilita a construção e a análise de notações simples, que utiliza uma classe específica ou complexa, pelo uso de combinações de classes, expressando uma combinação de conceitos ou atributos, o que evidencia a importância da classificação analítico-sintética na construção de sistemas de classificação (SOUZA, 2012). No que tange à inclusão direta de uma notação à outra, essa regra é válida em todos os casos, cujos indicadores de faceta são notações a serem conectadas, tendo em vista preservar a integridade de cada elemento (MOMM; LESSA, 2009). Esse tipo de síntese permeia todo o sistema e obedece às seguintes determinações:

- pode ocorrer entre uma notação da tabela principal com uma das tabelas auxiliares comuns. Exemplo: **631.111.3 + “322”**, que corresponde à adição do assunto da tabela principal **631.111.3** (Qualidade da Terra) e da tabela auxiliar comum de tempo **“322”** (Verão), o que possibilita a representação do assunto: “Qualidade da Terra no Verão”;
- Ao criar qualquer notação de uma tabela principal com uma ou mais tabelas auxiliares comuns, não é necessário um sinal para conectar os números e símbolos, formando assuntos ilimitados e mais específicos, por meio da combinação de assuntos ou inserção de novas notações. Exemplo: **63(81-022):631.151.2**, que corresponde ao assunto: **63** (Agricultura) da tabela principal; **(81)**, que significa Brasil na tabela auxiliar comum de lugar e **(-022)** (Grande Propriedades), da tabela auxiliar comum de propriedades; **(“:”)** relacionado com **631.151.2** (Sistema Agrícola Intensivo), o que corresponde ao assunto: “Agricultura no Brasil em Grandes Propriedades de Sistema Agrícola Intensivo” (MOON; LESSA, 2009).

O outro tipo de síntese conforme Momm e Lessa (2009) é a combinação de sinais, que serve para relacionar assuntos, de modo que essa regra é válida sempre que duas notações, quando unidas, resultam em uma distinção entre os elementos. Nesse tipo de síntese, prevalecem as seguintes determinações:

- Para formar notações entre as tabelas principais, considera-se o seguinte exemplo: entre tabelas principais (disciplinas) **63:004**, essa notação é traduzida como a relação de assuntos referentes às disciplinas de Agricultura (**63**) e de Ciência da Computação (**004**);
- No caso das tabelas de auxiliares comuns da mesma classe, tem-se o exemplo: para as tabelas auxiliares comuns de lugar (**81:474**), essa notação pode ser entendida como: (**81**) refere-se ao Brasil e (**474**) aos Estados Unidos (SLAVIC, 2005).

Por fim, o último tipo de síntese faz referência à flexibilidade na ordem dos mesmos elementos da notação com uma notação da tabela principal, sendo adicionadas, também, duas ou mais tabelas auxiliares comuns e com apenas tabelas auxiliares comuns. Essa regra se manifesta quando a ordem de citação da notação da tabela principal não sofre modificação na notação, alterando, apenas, a ordem de citação das tabelas auxiliares comuns, haja vista não modificar o assunto (SLAVIC, 2005).

Svenonius (1983) relata que, por exemplo, considerando a notação com as tabelas 63 (tabela principal - Agricultura), auxiliar de tempo (século 19) e auxiliar de lugar ((81) - Brasil), escrita como **63"18"(81)**, os números de classes pré-coordenados permitem o agrupamento e a exibição da agricultura por século e país, mas também permite uma pesquisa pós-coordenada, por exemplo, **63 AND (81)** ou Agricultura no Brasil. No entanto, no exemplo anterior, pode-se escolher a ordem de assuntos relevantes para uma determinada coleção, por exemplo, declarando um auxiliar de lugar seguido por um auxiliar de tempo, **63 (81) "18"**, e não o contrário: **63 "18" (81)**, cujo exemplo mencionado inicia-se por país e não por século. Exemplo: **631(81)"19"** e **631"19"(81)**, onde está a tabela principal; a notação **631** (Agricultura em geral); tabelas auxiliares comuns de lugar e tempo; Brasil, que possui a notação (**81**) e século XX, que possui a notação "**19**", sendo que essas duas notações correspondem ao mesmo significado: "Agricultura em Geral do Brasil no Século XX" (SVENONIUS, 1983).

Por outro lado, a ordem de arquivamento da CDU é obrigatória e padronizada, sem flexibilização, ou seja, é a partir da classificação dos conceitos mais genéricos para os mais específicos, do mesmo modo acontece com a ordem de citação, ou seja, a ordem em que os elementos da notação são combinados para formar uma notação composta. Entretanto, nas tabelas auxiliares especiais, a ordem de arquivamento é: =... Língua, (0...) Forma, (1/9) Lugar, (=...) Raça e "... Tempo e pode ser flexível. O Quadro 6 exemplifica a flexibilidade na ordem de citação da notação utilizada, apenas, nas tabelas auxiliares comuns.

QUADRO 6 - Exemplos de flexibilidade em tabelas auxiliares comuns

Número	Tabelas Auxiliares Comuns	Notação	Descrição (assunto)
1	Lugar	(729.828)	Granada (País da Região do Caribe)
2	Forma e Lugar	(086.8)(729.828)	Gravações de Vídeo em Granada
3	Lugar e Tempo	(729.828)“18”	Granada no Século XIX
4	Lugar, Tempo e Forma	(729.828)“18”(086.8)	Gravações de Vídeo de Granada no Século XIX

FONTE: Elaborado pelo autor (2020). Adaptado de Slavic (2005).

Outro aspecto a ser analisado na CDU diz respeito à escolha das notações compostas do sistema. Para esclarecer esse caso, apresenta-se a Figura 6, adaptada do estudo de Slavic (2005). Essa Figura mostra a construção de notações complexas, apresentando um assunto simples na recuperação da informação.

FIGURA 6 - Exemplo de notação complexa como um assunto simples

<u>Notação</u>	<u>Descrição</u>
631.8"19"(81)	Fertilizantes do Brasil no século XX

FONTE: Elaborada pelo autor (2020). Adaptada de Slavic (2005).

Diante da Figura 6, pode-se perceber que a notação **631.8"19"(81)** representa um assunto complexo. Esse assunto, conforme exemplificado por Slavic (2005), foi indexado como assunto simples na tabela principal, considerando a descrição da Figura: “Fertilizantes do Brasil no Século XX”.

A Figura 7 apresenta a utilização de combinação de assuntos com um exemplo de uma estrutura complexa de notação da CDU na recuperação. A notação foi dividida, utilizando uma tabela principal, pela notação **631.8**, que corresponde ao assunto “Fertilizantes” na CDU; em seguida, a notação foi combinada com uma tabela auxiliar comum de tempo, com a notação **“19”**, que significa “Século XX”; por fim, combinou-se a notação com um assunto da tabela auxiliar comum de lugar, **“(81)”**, nesse caso, “Brasil”. Cada notação representa um assunto simples separadamente, que já estava classificado em cada tabela

mencionada e que, no momento da recuperação, foi formada uma notação complexa (SLAVIC, 2005).

FIGURA 7 - Exemplo de notação complexa na recuperação



FONTE: Elaborada pelo autor (2020). Adaptada de Slavic (2005).

A Figura 7 mostra os blocos de cada notação já classificados e suas respectivas descrições separadamente, conforme explicado. Por fim, outro ponto a ser mencionado a respeito da CDU é a derivação de uma notação a partir de outra notação (derivação paralela) a partir da subdivisão paralela, sendo que a derivação e a subdivisão paralela são possíveis, apenas, nos casos previstos (SLAVIC, 2005).

Uma das principais aplicações da CDU é durante a classificação de itens informacionais, que é realizada de forma escalonável, sendo utilizada como uma taxonomia simples ou para a construção de números complexos que permite precisão no momento de representar e recuperar (RIESTHUIS, 1999). Para ser utilizada em potencial, qualquer que seja a complexidade de sua implementação, a classificação deve ser correspondente com a linguagem natural e, para isso, deve-se criar um índice alfabético de assunto para o controle de termos e vinculação às classes da CDU (RIESTHUIS, 1999).

Para Broughton (2010), a CDU tem sido bastante utilizada no cotidiano dos bibliotecários, além de ser relatada na literatura da área da BCI, com menção, também, em diversas outras áreas. A CDU permite o aumento da flexibilidade na classificação de assuntos, o que facilita a recuperação, sobretudo por possuir uma estrutura pré-coordenada que garante a criação de notações detalhadas, que são construídas de modo hierárquico com um número limitado de conceitos (BROUGHTON, 2010). Entretanto, segundo a autora, a estrutura coordenada do sistema tem a função de formar classes (renques) a partir de uma única característica de divisão de assuntos, dentro de uma categoria, permitindo a classificação de um número ilimitado de assuntos complexos.

De acordo com Soergel (1999), a CDU pode ter várias aplicações na recuperação da informação, tais como: fornecer mapas semânticos; melhorar a comunicação e

aprendizagem; indicar uma base conceitual para a realização de pesquisas; revelar classificação para ações; apoiar a recuperação de informações; fornecer uma base conceitual para sistemas baseados em conhecimento; sugerir a base conceitual para a definição de elementos de dados e hierarquias de objetos em sistemas de software; criar o mapeamento entre disciplinas, línguas e culturas; e servir como base para o processamento de linguagem natural, de modo a estabelecer uma forte relação entre a classificação e a máquina.

Para que se tenha uma implementação adequada da CDU em um SRI, Soergel (2009) relata que as soluções devem contemplar aplicações mais amplas de esquemas de classificação on-line. Assim, a tendência crescente de compartilhamento de SOC na web pode influenciar de maneira significativa o modo como usamos os sistemas de classificação na descoberta de recursos e como fonte de relações de conceitos de assuntos complexos, haja vista explorar a informação e o conhecimento no futuro.

Considerando que a CDU facilita o estabelecimento de vários relacionamentos, por conseguinte, garante melhorias na recuperação (SOUZA, 2012), porém, no ambiente digital, o uso desse sistema demanda da combinação com outros instrumentos de organização, haja vista potencializar a recuperação (SAN SEGUNDO, 2009), é importante discorrer sobre outros instrumentos de organização. Portanto, no próximo subcapítulo, será discutido sobre os tesouros, instrumentos de organização do conhecimento utilizados, nesta tese, com o propósito de auxiliar a CDU no processo de recuperação.

3.1.2 Tesouros

A palavra "*thesaurus*", no sentido etimológico, vem do grego "thesauros" e do latim que significa "estoque de tesouros", ou do inglês como "*treasure store*", o que designa de forma léxica "tesouro de palavras", que surgiu a partir da publicação de *Thesaurus of English Words and Phrases* de Peter Mark Roget, em Londres, no ano de 1852 (JARMASZ; SZPAKOWICZ, 2012). No final dos anos 1950, esse conceito ficou conhecido como dicionário de sinônimos, que representava uma lista de descritores (termos) relacionados em uma relação de equivalência. Além dessa relação, o tesouro também possui relacionamentos hierárquicos e associativos, sendo considerado um sistema de conceitos (SIMÕES, 2007). Nesse sentido, os tesouros se diferem dos dicionários tradicionais, visto que nesses parte-se de um termo para saber os significados ou conceitos que ele pode deter, ao passo que naqueles parte-se de um conceito para chegar aos termos que melhor o representem, como em um sistema de classificação, que relaciona as ideias e lista os termos e/ou tópicos de forma organizada (GOMES; CAMPOS, 2004).

Com o crescimento do volume de informações específicas produzidas em diversas áreas do conhecimento, houve a necessidade da criação de instrumentos capazes de auxiliar

nos processos de armazenamento (indexação) e recuperação, principalmente no contexto digital (BRASCHER, 2009). Nesse sentido, os tesouros são instrumentos de controle terminológico, caracterizados como vocabulários controlados, organizados em ordem alfabética e utilizados por indexadores, pesquisadores e usuários de sistemas de informação possuem o objetivo de realizar a tradução da linguagem dos documentos (linguagem natural) para uma linguagem controlada (BRASCHER, 2009).

Para Campos (2001), a definição de tesouros a partir da vertente norte-americana, mais especificamente dos Estados Unidos, é uma lista alfabética de descritores que foram criados a partir de cabeçalhos de assunto. Esta definição se deve pelo fato dos tesouros não atender de forma adequada a representação de assuntos compostos, ou seja, com mais de um descritor durante o processo de indexação. Já em relação a representação de assuntos simples, esta definição atende de forma adequada, já que é necessário apenas um único descritor como ponto de acesso em ordem alfabética ao assunto dos documentos. Por outro lado, tendo em vista que a maioria dos assuntos são representados como assuntos compostos em relação aos assuntos simples, onde os descritores são ordenados alfabeticamente, fez com que Lancaster (1986) considerasse o sistema unitermo como sendo o responsável pela origem dos tesouros.

O sistema Unitermo foi criado por Mortimer Taube em 1951 e ficou conhecido também como “sistema coordenado”, por permitir a combinação de assuntos compostos na recuperação. Desse modo, para cada documento ou conjunto deles é atribuído um número em ordem crescente, de acordo com a entrada no arquivo de documentos. Por este motivo, o sistema Unitermo não possui nenhum controle do vocabulário, e logo depois passou a ser chamado de “sistema pós-coordenado” (LANCASTER, 1986).

De acordo com Foskett (1973), nos sistemas pós-coordenados, os assuntos compostos são representados pelos cabeçalhos de assuntos compostos e não são acessados de forma direta, somente a partir de índices ou remissivas como ponto de acesso a informação. Este método traz prejuízos na recuperação, pois não levam em consideração todos os aspectos de um documento. Por este motivo, o tesouro da vertente Européia, foi definido como sendo uma parte sistemática, ou seja, por meio de categorias de assuntos baseada na teoria da classificação facetada de Ranganthan. Já a outra parte, refere-se como um índice alfabético de assuntos, para que os assuntos sejam organizados de forma apropriada para facilitar a recuperação (LANCASTER, 1986).

Aitchison (1970,1972) consolida a vertente Européia com o *Thesaurofacet*, conhecido como “tesouro em facetas”, que contém a integração de um sistema de classificação facetado com um tesouro organizado em ordem alfabética. Dessa forma, foi possível criar subdivisões de todo o conhecimento humano em disciplinas, possibilitando a criação de categorias de assuntos e suas relações semânticas para a construção de tesouros.

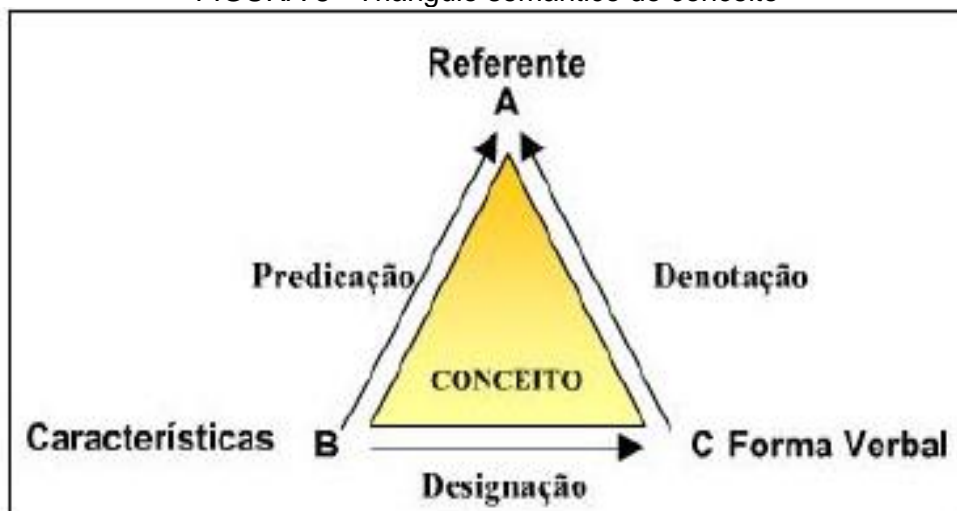
Em 1970, o grupo *Classification Research Group* (CRG) contribuiu para mudanças na construção de tesouros indexados de forma alfabética, a partir da Teoria da Análise Facetada de Ranganathan (AITCHISON, 1970, 1972). Após as contribuições do grupo CRG, o *Theraurofacet* se consolidou com as normas de construção de tesouros ISO 25964 (2011) e ISO 25964 (2013).

Após a definição de tesouros na vertente norte-americana e Européia, existem diversas definições de tesouros por parte de alguns autores, como Cervantes (2009) que descreve esse instrumento como um conjunto de descritores de um determinado domínio do conhecimento. Esse conjunto de descritores são conhecidos por diferentes indivíduos e comunidades científicas, os quais concordam na sua utilização, haja vista referir-se aos conceitos e suas relações semânticas, de maneira uniformizada (CERVANTES, 2009).

Na opinião de Tristão *et al.* (2004), o tesouro é um vocabulário controlado de descritores relacionados de modo genérico e específico, sobre determinada área do conhecimento. Portanto, os tesouros são mecanismos que formam uma estrutura de conceitos, criando uma lista de termos relacionados entre si. Tais instrumentos são desenvolvidos para atender várias áreas do conhecimento, e constituem uma seleção de termos (geral e específico) baseados em análise de conceitos. O termo geral possui o conceito com maior abrangência e os termos específicos expressam os conceitos menores (TRISTÃO *et al.*, 2004). Com efeito, os tesouros são constituídos por um conjunto de conceitos que são representados por termos em linguagem natural e utilizam regras sintáticas e semânticas com o objetivo de representar um domínio específico, tendo em vista facilitar a recuperação (CINTRA *et al.*, 2002).

Para melhor entender a definição de conceito, elemento principal representado nos tesouros, Dahlberg (1978a) explicita a teoria do conceito por meio do triângulo semântico. Segundo a autora, o conceito é materializado no termo, como sendo uma unidade indivisível, indicando o objeto (referente), que possui características (propriedades e relações), designado por uma forma verbal, conforme ilustrado na Figura 8.

FIGURA 8 - Triângulo semântico do conceito



FONTE: Dahlberg (1978a, p.149).

Conforme descrito na Figura 8, considerando a teoria proposta por Dahlberg (1978a), que segue a lógica analítica Aristotélica para classificar as coisas do mundo e logo depois, surgiu a classificação analítico-sintética de Ranganathan. Desse modo, percebe-se que as características de um objeto (referente) em um determinado domínio fazem parte do conceito, onde o conceito é apresentado como uma unidade do conhecimento, ou seja, é uma representação mental mais objetiva desse objeto. Assim, é estabelecida uma relação entre o conceito do referente e sua forma verbal, por meio de um signo linguístico ou termo para representar o referente (DAHLBERG, 1978a).

Gomes, Campos e Guimarães (2010) apontam que a abordagem metodológica mais apropriada para estabelecer o conceito é a onomasiológica, cujo conceito é o ponto de partida. Nessa metodologia, parte-se do objeto (referente) e através da atribuição de suas propriedades em um dado contexto, chega-se a forma mais apropriada de definição e denominação. Já na abordagem semasiológica, o ponto de partida é o termo (palavra), ou seja, inicia-se com o levantamento de unidades léxicas específicas que denotam um conceito, para que se possa identificar o conceito que a unidade representa em cada área do conhecimento (GOMES; CAMPOS; GUIMARÃES, 2010).

Nesse sentido, os tesouros foram criados para organizar os assuntos tratados nos documentos (GOMES, 1996). Além disso, segundo a autora, os tesouros orientam os procedimentos de busca em uma base de dados, haja vista padronizar o uso da terminologia, a qual não auxilia, apenas, a indexação das informações, mas também a recuperação. Logo, os tesouros são instrumentos de organização da informação utilizados em sistemas de buscas tradicionais, na indexação manual e automática e na recuperação da informação (GOMES, 1996).

Segundo Sales e Café (2009), o tesouro pode ser utilizado para recuperar informação em diversos catálogos on-lines, sobretudo em sistemas de informação de áreas

especializadas. Dessa forma, o tesouro auxilia na representação e recuperação de informações referentes aos dados que indicam as obras existentes no catálogo das bibliotecas, tais como: autor, título e o assunto de cada uma delas. Esse catálogo de assuntos pode conter uma lista de obras, ordenadas de modo alfabético, incluindo remissivas (ver também) e o cruzamento de referências para relacionar as diversas obras sobre um tema. Assim, o catálogo de assunto serve para auxiliar o usuário a encontrar um recurso informacional, sem a necessidade de conhecer o título, o autor, a editora, entre outros aspectos não conhecidos (SALES; CAFÉ, 2009).

De qualquer modo, o usuário precisa ter o conhecimento sobre um determinado tema no momento de realizar a busca, de modo que fique menos complexo fazer o arranjo e o cruzamento de informações. Isso possibilitará ao sistema de informação recuperar um número satisfatório de resultados relevantes, alinhados às necessidades dos usuários (SALES; CAFÉ, 2009). Em síntese, de acordo com Tristão *et al.* (2004), pode-se dizer que os tesouros são uma lista alfabética de termos que representa um conceito e cada conceito é seguido por outros que se relacionam.

Nesse sentido, Currás (1995) relata que os descritores do tesouro constituem uma linguagem padronizada, podendo representar conceitos que estão em uma ou mais disciplinas do conhecimento, ou seja, os tesouros podem ser do tipo mono ou multidisciplinares. Além dessa classificação, a literatura tem apresentado diversas divisões para esses instrumentos, considerando critérios distintos. Assim, os tesouros podem ser classificados quanto ao nível de especificidade, caracterizados como macrotesauros, compostos por termos que representam conceitos genéricos e específicos, e como microtesauros, que utilizam termos que correspondem a um assunto mais específico (CURRÁS, 1995).

Por serem utilizados para controle terminológico, os tesouros assumem a característica de padronização, ou seja, a informação pode ser acessada, com foco nos termos que estão relacionados de modo prévio e normalizado. Para garantir esse controle terminológico, são utilizadas metodologias de compatibilidade da linguagem, independente dela possuir um ou vários idiomas (monolíngue ou multilíngue) (CURRÁS, 1995). Esse fato pode ser observado em diversos tipos de SOC como, por exemplo, os sistemas de classificação, que são orientados conforme o desenvolvimento do conhecimento. Por caracterizarem-se como instrumentos que padronizam a linguagem, os tesouros manifestam-se como linguagem pós-coordenada. Neste tipo de linguagem, os termos são combinados no momento da recuperação, seja pelo indexador ou pelo usuário do sistema, ao contrário das linguagens pré-coordenadas como, por exemplo, os sistemas de classificação bibliográfica e os cabeçalhos de assuntos (CURRÁS, 1995).

Campos e Gomes (2003) mostram que os sistemas de classificação servem de base para a construção dos tesouros, por meio das relações semânticas de hierarquia, que

são utilizadas em serviços tradicionais de resumo como, por exemplo, o *Chemical Abstracts*, o *Common wealth Bureau of Agriculture* e o *Biological Abstracts*. As autoras revelam que, nesses serviços, os instrumentos de controle ampliam as possibilidades de se utilizar outros campos de metadados, tais como o título, nome dos autores, editora, ano, entre outros, tendo em vista a expansão da busca pelos usuários.

Partindo dessa característica, os conceitos devem estar de acordo com os descritores indexados, o que corresponde aos termos utilizados pelos usuários durante a busca, sendo que a principal representação é a seleção do termo preferido (descritores), utilizado na indexação, que identifica, com mais precisão, a necessidade de informação dos usuários (PASTOR-SÁNCHEZ; MARTÍNEZ MÉNDEZ; RODRÍGUEZ-MUÑOZ, 2009). Os termos não preferidos (não descritores), também conhecidos como termos alternativos, são relacionados com os termos preferidos por meio de termos gerais, específicos e relacionados, e a nota de escopo é utilizada para definir um termo (PASTOR-SÁNCHEZ; MARTÍNEZ MÉNDEZ; RODRÍGUEZ-MUÑOZ, 2009).

Para atingir esse objetivo, deve-se levar em conta a relação entre os termos que denotam o conceito de forma padronizada, haja vista evitar inconsistências (ISO 25964, 2011, 2013). Portanto, a expressão de busca deve ser correspondente para a mesma expressão utilizada, levando em conta os metadados que representam o conteúdo de um documento, por meio do intercâmbio entre os conceitos e entre os vocabulários utilizados, identificando as relações entre os conceitos, visando à precisão na recuperação (ISO 25964, 2011, 2013).

Para Hjørland (2016), o tesauro caracteriza-se como uma ferramenta semântica flexível (vocabulário controlado), que possui uma estrutura relacional e ampliada, visto que apresenta um número maior de relações semânticas do que nos sistemas de classificação. Além disso, utilizam códigos ou notações para representar o conteúdo informacional dos documentos. O autor descreve os tesauros como conjuntos de relações semânticas significativas entre os descritores, tais como: associativas, hierárquicas e de equivalência, relações essas que serão detalhadas no subcapítulo 3.2.

Semelhante aos sistemas de classificação, o objetivo dessas relações é garantir uma lógica, levando o indexador a representar um determinado domínio que, traduzido em uma linguagem artificial, seja capaz de eliminar a ambiguidade da linguagem natural, por conseguinte, o descritor representado possa ser utilizado na recuperação, gerando resultados relevantes às expectativas dos usuários (HJØRLAND, 2016). Nesse sentido, a representação dos conceitos de um domínio por meio de termos possibilita um complexo cruzamento de referências que permite a opção de termos gerais ou específicos durante o processo de indexação (HJØRLAND, 2016).

Quanto às funções, Cervantes (2009) menciona que os tesauros possuem diferentes abordagens para ajudar a estruturar, classificar, modelar e/ou representar conceitos

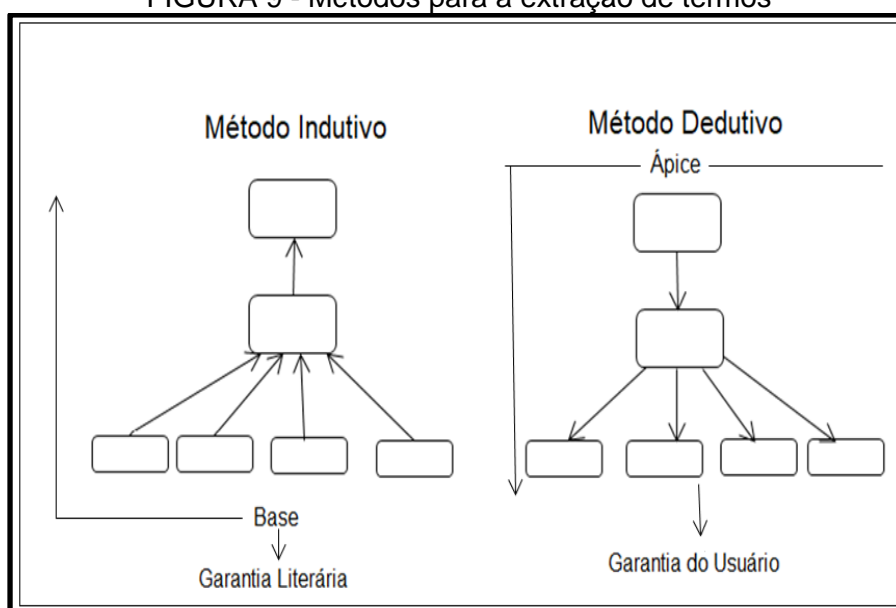
e relações de uma área específica. Além disso, o tesauro garante o controle dos termos, contemplando, também, a inclusão de novos termos, desde que esses expressem um conceito. Desse modo, cada termo pode estar em diferentes níveis de especificidade, dependendo do domínio em que está inserido (CERVANTES, 2009).

Com frequência, os tesouros são usados para agrupar termos em conjuntos de tópicos, que são baseados em conceitos e em relações semânticas. Os termos e os relacionamentos do tesauro são estabelecidos com lógica, objetividade e precisão, cujos elementos estruturais podem estar implícitos em uma Classificação Facetada, haja vista organizar a informação para ser recuperada, posteriormente (HODGE, 2000). Para o autor, essas relações semânticas podem ser estabelecidas, considerando as propriedades ou características entre os conceitos, que permitem identificar categorias ou classes, a partir das diferenças e semelhanças. Os descritores do *thesaurofacet* (tesauro com facetas) são identificados como descritores para indexação, classificação e recuperação, desde que representem um domínio do conhecimento (HODGE, 2000).

Na área da BCI, utiliza-se a Classificação Facetada pelo método dedutivo (*top-down*) para classificar o conhecimento dentro de um domínio, partindo da dedução das classes de conceitos inseridas, a partir do conceito maior ou de classes mais gerais até as mais específicas (GREEN, 2008). Isso possibilita a evidenciação do processo de divisão lógica, de acordo com a experiência dos usuários que são especialistas do domínio (GREEN, 2008).

A Figura 9 apresenta o método dedutivo, que classifica o conhecimento por meio de regras da lógica de predicados. Com efeito, a aplicação do método possibilita o alcance de uma conclusão, por meio das classes mais abrangentes.

FIGURA 9 - Métodos para a extração de termos



FONTE: Adaptada de Dodebei (2002, p. 66).

Pela análise à Figura 9, considerando os apontamentos de Dodebei (2002), percebe-se que, quando se utilizam os métodos, tanto dedutivo quanto indutivo, para classificar o conhecimento, se as premissas são verdadeiras e as leis aplicadas estiverem corretas, pode-se dizer que a conclusão é verdadeira. Já no método indutivo, são utilizadas categorias (facetadas) para atingir níveis mais específicos na hierarquia dos termos ou por agrupamento (*clusters*) desses termos, até chegar à representação da estrutura desejada de conceitos e termos. Com o método indutivo, é possível inferir conceitos gerais, com base na garantia literária (teses, dissertações, livros, artigos de periódicos, dentre outros tipos), por meio de conceitos mais específicos, permitindo o mapeamento dos termos (descritores) que são inseridos depois de serem validados, contribuindo, assim, para a organização e representação do conhecimento (DODEBEI, 2002).

Para Fujita e Moreira (2021), a relação entre os termos (descritores) e facetadas acontece por meio da combinação dos assuntos, de modo sistemático, ou seja, estabeleceu-se a relação hierárquica e facetada, a partir dos sistemas de classificação bibliográfica. Os autores salientam que essas relações podem ocorrer, também, de forma alfabética como nos tesouros convencionais. Por essa razão, a categorização dos assuntos surgiu por meio da utilização da Classificação Facetada, sendo possível, assim, definir as relações semânticas entre os termos dos tesouros (FUJITA; MOREIRA, 2021).

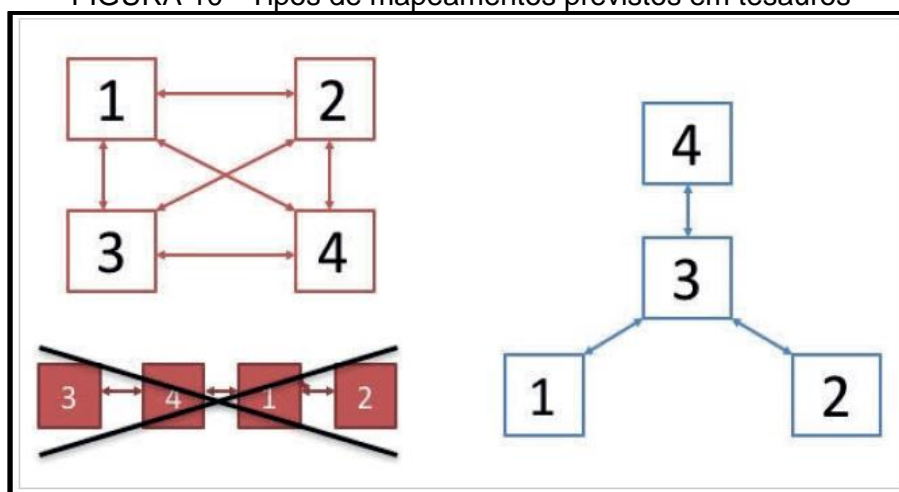
Os descritores de um tesouro, segundo Soergel (1995), são utilizados para representar um determinado domínio, cuja função é eliminar a ambiguidade da linguagem natural, por conseguinte, facilitar a recuperação. Assim, o tesouro manifesta-se como uma lista de termos autorizados, com a função de conduzir o usuário de um conceito para outro, de modo intuitivo (SOERGEL, 1995). Foskett (1973) salienta que as principais funções dos tesouros são: o controle de sinônimos e quase sinônimos, a distinção de homógrafos e a facilidade da busca por meio dos descritores relacionados e das referências cruzadas (estrutura sintética). Ademais, o autor menciona que o uso desse instrumento melhora a consistência da indexação e garante o controle da terminologia em função do conteúdo ou significado entre as relações dos conceitos por meio de comparações, sejam elas lógicas, formais, abstratas ou semânticas.

Além dessas funções, os tesouros facilitam a escolha dos termos que representem o significado dos conceitos que os usuários procuram, e são compostos por um conjunto de termos alternativos, cuja informação contida no documento precisa estar de acordo com o significado desejado pelos usuários (HJØRLAND, 2016). Nesse sentido, Foskett (1973) destaca que os tesouros podem ser utilizados em diferentes contextos, oferecendo termos representativos para a padronização da linguagem, e devem ser atualizados de acordo com a evolução ou atualização dos conceitos de um domínio para que não aconteçam problemas em relação à linguagem natural.

García-Marco (2017) verificou que os problemas que ocorrem na estrutura dos tesouros relativos à dispersão da linguagem natural podem ser de natureza léxica, sintática e semântica. Como problemas léxicos, citam-se, por exemplo, ambiguidades oriundas da ocorrência de sinonímia e polissemia; já os problemas sintáticos surgem pela formação e estruturação das frases, ou seja, pela sua expressão verbal, cujo conceito é representado por uma linguagem verbal ou terminológica, como também pelas características (signo, símbolo, forma, imagem som, dentre outras). Por fim, a semântica oferece sentido, significado ou denotação para a estruturação do instrumento (GARCÍA-MARCO, 2017).

Nesse contexto, os mapeamentos entre os termos (vocabulário) devem ser formulados de modo objetivo, cujos conceitos e as estruturas relacionais sejam equivalentes, por exemplo, entre dois tesouros. Recomenda-se, também, a existência de um link direto, possibilitando que todos os vocabulários tenham uma ligação e possam ser mapeados em ambas as direções, o que fortalecerá a estrutura do sistema dos vocabulários (GARCÍA-MARCO, 2017). O autor esclarece que um mapeamento seletivo precisa estar previsto entre partes dos vocabulários, sem obter um mapeamento total, mas uma aproximação, sendo essa uma técnica bastante utilizada em muitos projetos de construção de instrumentos terminológicos, conforme ilustrado na Figura 10.

FIGURA 10 - Tipos de mapeamentos previstos em tesouros



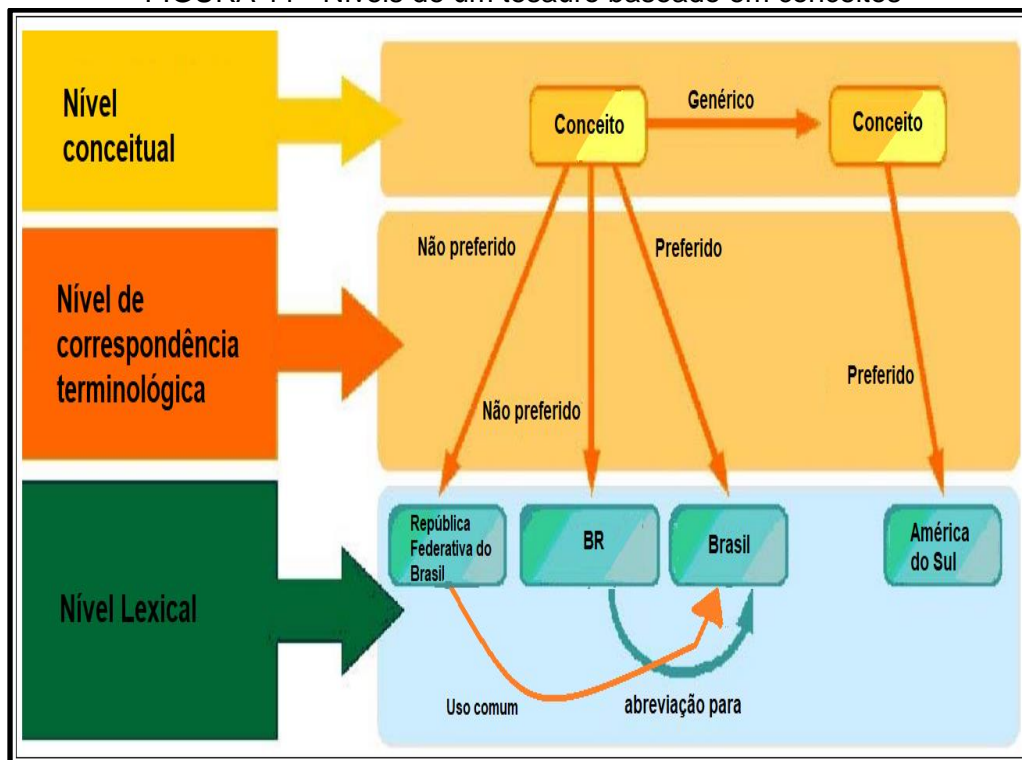
FONTE: García-Marco (2017, p. 306).

A Figura 10 indica que, nos dois elementos não riscados, todos os termos foram mapeados em ambas as direções. Além disso, esses termos possuem uma relação com os outros, ao contrário do elemento riscado, cujos termos não estão todos relacionados entre si (GARCÍA-MARCO, 2017).

A estrutura de um tesouro baseado em termos é, portanto, arquitetada em um único nível terminológico-léxico, enquanto um tesouro baseado em conceito tem uma estrutura de três níveis (PASTOR-SÁNCHEZ; MARTÍNEZ MÉNDEZ; RODRÍGUEZ-MUÑOZ, 2009). Os autores caracterizam cada um desses níveis, a saber: 1 - nível conceitual: nesse

nível, os conceitos são identificados e suas inter-relações estabelecidas; 2 - nível de correspondência terminológica: os termos são associados (preferido ou não preferido) aos seus respectivos conceitos; e 3 - nível lexical: as relações lexicais são definidas para interconectar termos. A fim de visualizar com mais clareza os níveis do tesauro baseado em conceitos, apresenta-se a Figura 11.

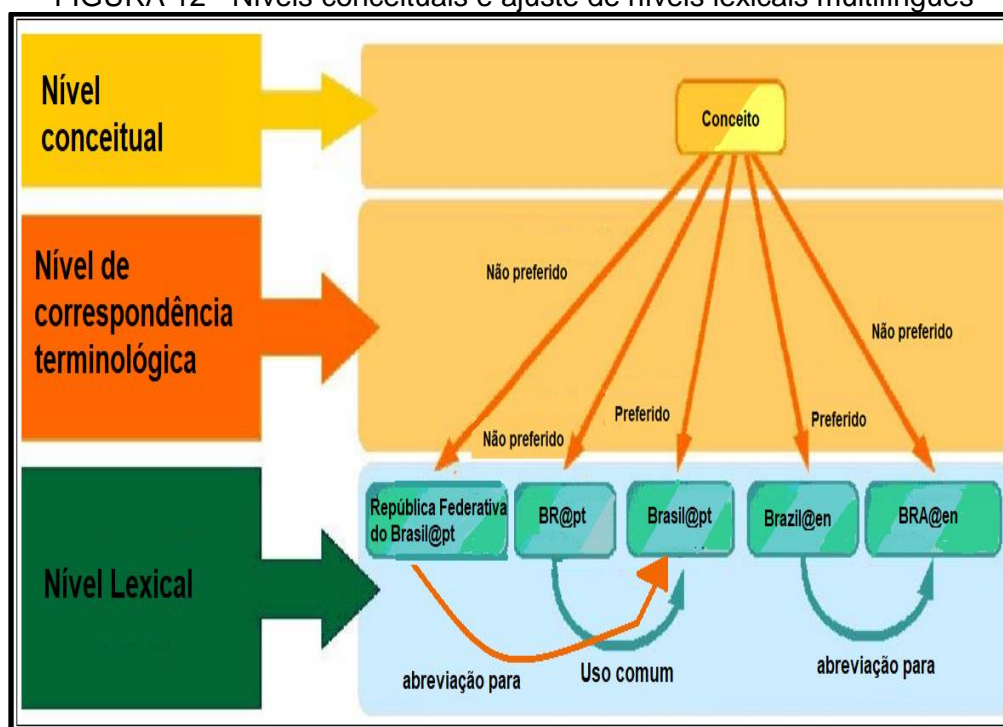
FIGURA 11 - Níveis de um tesauro baseado em conceitos



FONTE: Adaptada de Pastor-Sánchez, Martínez Méndez e Rodríguez-Muñoz (2009, p. 4).

O detalhamento da Figura 11 revela que, no nível lexical, os termos “**República Federativa do Brasil**” e “**BR**” são descritores não preferidos de um conceito. O primeiro possui a relação “**Uso Comum**” com o descritor preferido desse conceito que é “**Brasil**” e o segundo possui a relação “**abreviação para**” também com o descritor preferido. O descritor preferido “**América do Sul**” está em um nível superior, sendo mais genérico que os demais, que estão em um nível mais específico (PASTOR-SÁNCHEZ; MARTÍNEZ MÉNDEZ; RODRÍGUEZ-MUÑOZ, 2009). Portanto, segundo esses autores, deve-se associar um conceito a diferentes descritores preferidos para cada idioma, cujo conceito expresso no tesauro permitirá estabelecer relações lexicais adaptadas à realidade terminológica em cada língua. A variação terminológica associada ao idioma é explicitada na Figura 12.

FIGURA 12 - Níveis conceituais e ajuste de níveis lexicais multilíngues



FONTE: Adaptada de Pastor-Sánchez, Martínez Méndez e Rodríguez-Muñoz (2009, p. 5).

A partir da Figura 12, com base nas considerações de Pastor-Sánchez, Martínez Méndez e Rodríguez-Muñoz (2009), percebe-se que, no final dos descritores, é inserido o símbolo “@”, seguido da abreviatura da língua. Nessa figura, são utilizados os descritores “**República Federativa do Brasil**”, “**BR**” e “**Brasil**”, com a abreviatura “**pt**” (língua portuguesa). Por fim, são mostrados os descritores “**Brazil**” e “**BRA**”, com a abreviatura “**en**” (língua inglesa).

De acordo com Maculan (2015), os tesouros são estruturados utilizando abreviaturas (símbolos), necessárias para indicar o tipo de relação estabelecida entre os termos. Nos tesouros brasileiros, por exemplo, na maioria das vezes, apresentam-se os significados equivalentes também em inglês, sendo mencionado o significado das relações entre termos ou descritores dos tesouros que são utilizados de forma geral, ou seja, para qualquer domínio (MACULAN, 2015). No Quadro 7, são explicitados os tipos de relações semânticas utilizadas no AgroTermos¹⁷, a abreviatura e a explicação de cada abreviação, em português e em inglês. Reforça-se, segundo Maculan (2015), que os tesouros utilizam as abreviaturas para representar as relações de hierarquia, equivalência e associativa entre os termos que evidenciam um vocabulário específico.

¹⁷Esse tesouro caracteriza-se como um vocabulário controlado, reunindo inicialmente cerca de 240 mil termos pertinentes ao domínio do conhecimento agropecuário. Modelado conceitual e computacionalmente, o instrumento viabiliza diversos serviços web para aplicações em processos de gestão da informação e do conhecimento, como indexação documental, recuperação de informações e interoperabilidade entre sistemas de informações (EMBRAPA, 2018, p. 37).

QUADRO 7 - Principais abreviaturas utilizadas no AgroTermos

Símbolo em português	Símbolo em inglês	Tipo de relacionamento	Significado das relações
USE (Descritor ou termo Preferido)	USE	Relação de Equivalência	Representa um conceito utilizando termos preferidos, o que indica remissivas para o termo preferido.
UP (Usado Para)	UF (<i>Used for</i>)	Relação de Equivalência	Usado para termos não preferidos, com uso no controle de sinônimos, o que representa uma remissiva desses termos.
NE (Nota de Escopo)	SN (<i>Scope Note</i>)	Não se aplica	Define o termo e orientações sobre como utilizá-lo na indexação.
TG (Termo Genérico)	BT (<i>Broader Term</i>)	Relação Hierárquica	Indica um conceito mais abrangente ou superordenado representado por um termo.
TGG (Termo Genérico da Relação Genérica)	BTG (<i>Broader Term Generic</i>)	Relação Hierárquica	Utilizado para indicar uma classe mais ampla que pertence a um conceito específico ou subordinado.
TEG (Termo Específico Genérico)	NTG (<i>Narrower Term Generic</i>)	Relação Hierárquica	Indica uma classe específica que representa um conceito mais amplo.
TE (Termo Específico)	NT (<i>Narrower Term</i>)	Relação Hierárquica	Representa um conceito mais específico de acordo com um conceito mais abrangente.
TR (Termo Relacionado)	RT (<i>Related Term</i>)	Relação Associativa	Representa relações associativas (não hierárquicas) entre conceitos, possibilitando uma maior abrangência ou limitação durante a busca.
TGP (Termo Genérico Partitivo)	BTP (<i>Broader Term Partitive</i>)	Relação Hierárquica	É um termo que representa um termo, descrito de modo genérico por um conceito em relação a suas partes.
TEP (Termo Específico Partitivo)	NTP (<i>Narrower Term Partitive</i>)	Relação Hierárquica	Representa o termo que representa uma parte inerentemente em relação ao termo que representa o conceito como um todo.

FONTE: Elaborado pelo autor (2020). Adaptado de Maculan (2015).

As abreviaturas mencionadas no Quadro 7 indicam os relacionamentos estabelecidos entre os termos. Tais relacionamentos “[...] são elaborados para possibilitar a comunicação entre a informação e o usuário, pelo compartilhamento de um mesmo vocabulário [...]” (MACULAN, 2015, p. 140). No entendimento de Oliveira (2018), os relacionamentos instituídos nos tesouros podem ser melhorados com acréscimos de novos

termos, muitos deles oriundos de outros SOC, possibilitando que um instrumento contribua para melhoria e adaptação de outro instrumento. Isso evidencia, no entender de Dias, Moreira e Alves (2020), que esse instrumento é dinâmico, ou seja, está em constante adaptação, sobretudo se utilizado para a padronização de termos no ambiente digital.

Nesse contexto, entende-se que os tesouros precisam ser embasados em regras objetivas e claras, no sentido de que seja possível distinguir e desambiguar o entendimento de quem faz uso desses instrumentos. Portanto, deve-se especificar o significado do termo, ou seja, contextualizá-lo, a fim de evitar problemas da linguagem natural, tais como homônimos, polissemia, homógrafos (HJØRLAND, 2007a). Ao criar relacionamentos mais objetivos e lógicos, condizentes com o domínio de conhecimento, cujo conceito é utilizado pelos especialistas, tais relacionamentos ficam mais delimitados e, portanto, capazes de orientar os usuários no processo de representação ou nas buscas realizadas nos sistemas de recuperação (HJØRLAND, 2007a).

No domínio da agricultura, por exemplo, conforme estruturado no AgroTermos, é possível notar que as relações tradicionais desse instrumento, como as relações de equivalência (USE e UP), sendo que a notação UP é usada a partir do termo de entrada ou pontos de acesso (termo não-preferido) para o termo preferido. Sendo assim, a sigla Usado Para (UP) deve mostrar os sinônimos e quase-sinônimos em relação ao termo preferido, por exemplo: Pássaros (termo preferido) USE Aves (termo não-preferido), e Aves UP Pássaros. As relações hierárquicas (TG e TE) e associativas (TR) podem gerar interpretações dúbias para o usuário (MACULAN, 2015). Sobre isso, a autora descreve uma hierarquia referente ao campo da Agricultura Intensiva, conforme apresentado no Quadro 8.

QUADRO 8 - Exemplo de hierarquia estabelecida no AgroTermos

AGRICULTURA INTENSIVA TG CRIAÇÃO INTENSIVA TE SISTEMA DE EXPLORAÇÃO AGRÍCOLA TE TIPO DE AGRICULTURA TR UTILIZAÇÃO INTENSIVA DA TERRA

FONTE: Adaptado de Maculan (2015, p. 24).

Pela análise à estrutura apresentada no Quadro 8, nota-se a presença de uma relação hierárquica entre os termos “Agricultura Intensiva” e “Sistema de Exploração Agrícola”. Assim, como pontuou Maculan (2015), essa relação não possibilita, de modo objetivo, a compreensão do leitor que não é especialista do domínio. Isso porque não fica claro, dentre as relações hierárquicas, qual o tipo de relação estabelecido entre esses dois termos, ou seja, se é uma relação do tipo gênero-espécie, partitiva ou outros tipos (MACULAN, 2015).

No tesouro, uma lista de termos só pode ser compilada por meio da indexação de conceitos, construindo uma linguagem de indexação de assunto pós-coordenada. Portanto, o

instrumento classificatório que mais se aproxima é o esquema da Classificação Facetada e os respectivos índices, devido à coleção de termos organizada que o sistema apresenta, não se limitando a uma mera ordem alfabética, como em um dicionário (GOMES; CAMPOS, 2004). De acordo com Cervantes (2009) existem normas para o desenvolvimento de tesouros monolíngues, ou seja, os que são elaborados para um idioma, e para os tesouros multilíngues, formulados em dois ou mais idiomas. Ambos seguem uma norma padrão de relacionamentos semânticos, a fim de identificar os descritores e regras para atribuir os relacionamentos entre os termos (ANSI/NISO Z39.19 (2005); CERVANTES, 2009).

A construção dos tesouros segue as normas internacionais da *International Organization for Standardization* (ISO) 25964 (2011) e ISO 25964 (2013), as quais apontam, com detalhes, a construção e os tipos de relacionamentos semânticos, além de indicarem o uso do tesouro com outros vocabulários controlados como, por exemplo, em conjunto com os sistemas de classificação bibliográfica, permitindo, assim, o intercâmbio de informações, com o propósito de garantir a recuperação da informação. Para que as informações sejam compartilhadas entre dois ou mais vocabulários controlados, a expressão de busca deve ser a mesma nos vocabulários que serão integrados, o que permitirá a interoperabilidade semântica levando em conta o idioma de cada um desses vocabulários (ISO 25964, 2011, 2013).

Fujita *et al.* (2020) aponta que a construção dos tesouros deve basear-se no controle terminológico, seguir as normas internacionais e especificar as tecnologias a serem utilizadas, visando à interoperabilidade semântica a partir da evolução do conhecimento. Os autores referem-se à importância dos tesouros, em especial, como contribuinte para a recuperação da informação em ambientes digitais, visto que esses instrumentos podem ser utilizados na web semântica e na construção de outros SOC (ANSI/NISO Z39.19, 2005).

Nesse contexto, em um SRI, o controle terminológico e o detalhamento sobre o tesouro são inseridos previamente, de modo organizado e de acordo com o contexto de utilização, seguindo as relações semânticas necessárias para a recuperação (ISO 25964, 2011, 2013). Desse modo, tendo em vista o escopo de aplicação dos tesouros utilizando as normas internacionais, é essencial considerar em um tesouro mono ou multilíngue, que um conceito deve ser associado a diferentes rótulos preferidos para cada idioma. Outra vantagem de representar um conceito a partir das estruturas de tesouros, é a possibilidade de estabelecer relações lexicais adaptadas à realidade terminológica de cada língua (ISO 25964, 2011, 2013).

Pastor-Sánchez, Martínez Méndez e Rodríguez-Muñoz (2009) consideram que os tesouros são elementos-chave para o desenvolvimento de um modelo correspondente aos padrões *Resource Description Framework* (RDF) e à *Extensible Markup Language* (XML), e podem ser utilizados na web semântica para a construção de ontologias. A tecnologia que é

requerida para a construção de ontologias mais estruturadas e a necessidade de padronização podem contribuir para a construção de tesouros mais sofisticados (PASTOR-SÁNCHEZ; MARTÍNEZ MÉNDEZ; RODRÍGUEZ-MUÑOZ, 2009). Segundo esses autores, para explicitar os conceitos e as relações semânticas dos tesouros, é preciso estabelecer um modelo de análise de categorias de assuntos e da Classificação Facetada, contribuindo, assim, para a construção de ontologias mais consistentes, tanto em termos semânticos quanto estruturais. Os autores citados afirmam que, em tesouros multilíngues, com aplicação dos tesouros expressos em RDF (web semântica), é imprescindível considerar os mecanismos que permitam expressar os termos que representam os conceitos em diversas línguas sem qualquer tipo de ambiguidade.

Diante do exposto, nota-se que o processo de elaboração dos tesouros deve romper qualquer tipo de ambiguidade ou informações que despertam uma interpretação dúbia ou equivocada. Isso porque, quando o objeto é analisado em uma situação empírica, a natureza do objeto observado pode variar de um contexto a outro (MACULAN, 2015). Portanto, “[...] enquanto um objeto pode significar ‘um fato’ em determinado ambiente, em outro, ele poderá ser interpretado como um ‘juízo de valor’. Logo, em cada contexto serão observados diversos tipos de propriedades, fenômenos e relações” (MACULAN, 2015, p. 141).

Com esse pensamento, a construção dos tesouros precisa obedecer a diversos critérios, mesmo com a aplicação de softwares especializados para geração desses instrumentos. Assim, é importante considerar os mecanismos para a criação e o controle de atualização de termos, a possibilidade de reutilização de outros tesouros e a elaboração de relatórios (CAMPOS *et al.*, 2006). Destaca-se que, o aspecto da consistência deve ser bem trabalhado, de modo que não haja a entrada duplicada de termos, para que os termos não sejam inseridos de qualquer forma, sem lógica entre os níveis hierárquicos, e também não ocorram relações duplicadas entre dois termos (CAMPOS *et al.*, 2006).

É nesse contexto de inconsistências nas relações estabelecidas em um tesouro, que se torna importante estabelecerem metodologias para adaptações dos tesouros utilizados em um domínio, sobretudo quando se trata da informação representada em ambientes digitais (CAMPOS *et al.*, 2006). Assim, segundo as autoras, é importante conhecer estratégias de elaboração de tesouros, tais como o uso dos softwares, visto que eles apresentam facilidades de armazenamento, manipulação e apresentação dos termos definidos e as relações estabelecidas. As autoras também evidenciam que os softwares utilizados para construção dos tesouros precisam ser concebidos a partir de uma visão integrada, cujos “[...] aspectos operacionais venham se juntar os aspectos metodológicos criando uma ferramenta mais inteligente” (CAMPOS *et al.*, 2006, p. 80).

3.1.2.1 Software *TemaTres*

O *TemaTres* é um software especializado para construção e gestão de vocabulários controlados, com capacidade de explorar dados em quase todos os esquemas de metadados disponíveis para a codificação dos vocabulários, mantendo um modelo de dados autônomo em relação a esses esquemas (FERREYRA, 2019). Manifesta-se como um software livre, com recursos para incrementar diversos vocabulários controlados. Este software permite enriquecer a estrutura do vocabulário, considerando a exploração de diferentes elementos e recursos aplicáveis aos instrumentos de organização e representação da informação. Por isso, a ferramenta é muito utilizada por profissionais que lidam com o tratamento de recursos informacionais, podendo ser utilizada no ambiente profissional e acadêmico, permitindo qualquer que seja o tamanho do vocabulário controlado a ser construído (GALDINO; SABBAG, 2017).

Desenvolvido no ano de 2004, uma importante contribuição do *TemaTres* é que possibilita a construção de tesouros para um domínio específico, utilizando padrões de metadados, haja vista permitir a interoperabilidade entre SOC e demais sistemas de informação, além de controle e reúso de terminologias e gerar recursos para facilitar a navegação na web (GALDINO; SABBAG, 2017).

O *TemaTres* foi escolhido para complementar os resultados desta pesquisa, por ele ser de fácil acesso e uso, além de possibilitar a construção de mecanismos que facilitam a navegação em ambiente digital. Sobre isso, Fujita *et al.* (2017, p. 69) salientam que o sistema “[...] permite desenvolver estruturas de navegação web e funcionar como complemento de um sistema de gerenciamento de bibliotecas, para a gestão da linguagem [...]”. Além dessa contribuição, outras facilidades de uso dizem respeito à aquisição, cujo software pode ser baixado por meio do acesso a sua página web. Quanto à interface de uso do *TemaTres*, ela é multilíngue e se encontra na versão 3.2, disponibilizada em vários idiomas, tais como: espanhol, alemão, francês, inglês, italiano e português (FUJITA *et al.*, 2017).

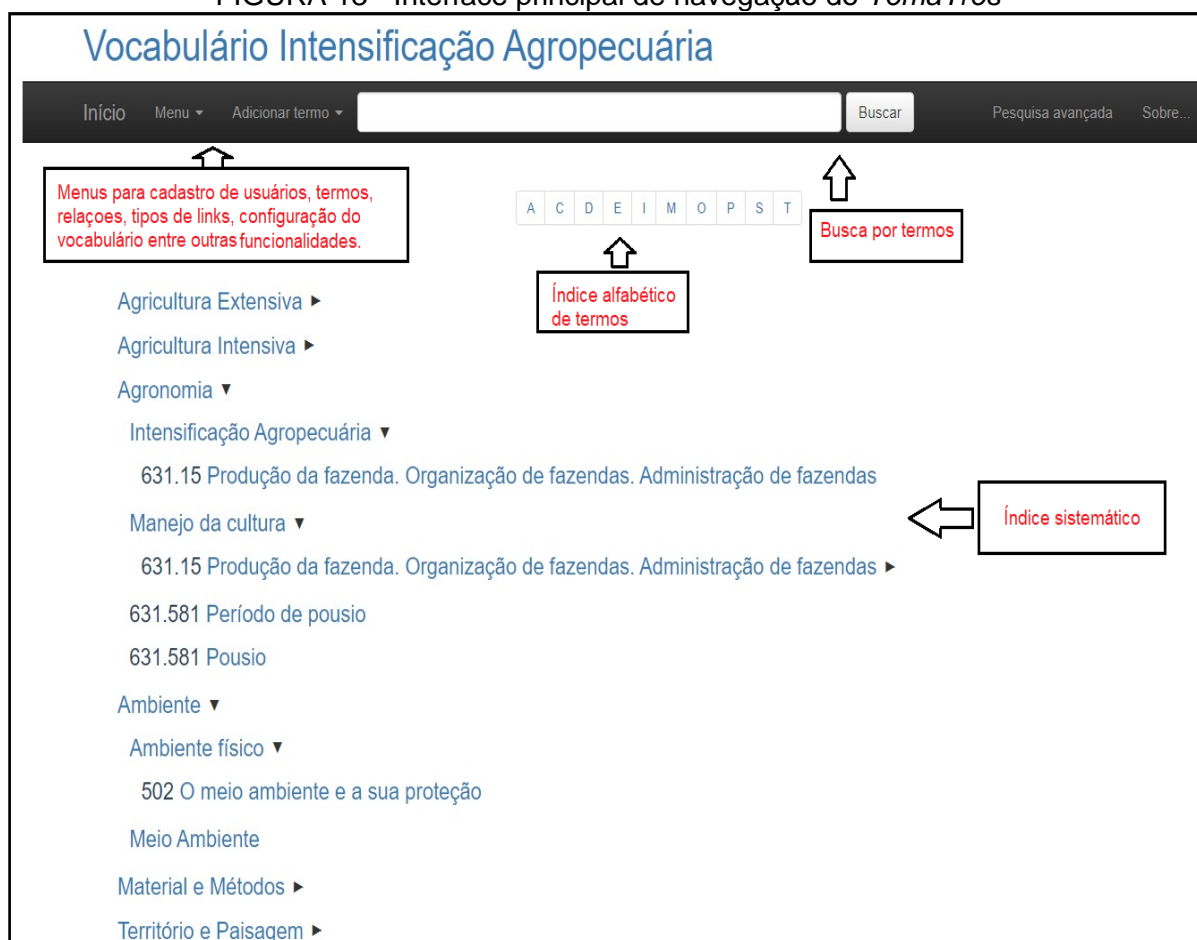
A ferramenta foi desenvolvida com uma estrutura simples e completa (CAMPOS *et al.*, 2006). Ademais, o software possui fácil entendimento pelo usuário, contendo a aplicação desenvolvida em PHP, normalmente utiliza-se um servidor Apache e banco de dados MySQL, no caso desta pesquisa utilizou-se o software Xampp¹⁸ que contém os dois servidores que foram instalados localmente. O *TemaTres* pode ser facilmente baixado e hospedado em sistema Linux, acessível pela web por meio de um navegador de internet com suporte a HTML5, CSS e JavaScript (COSTA; CASTRO; GOMES, 2019).

¹⁸ Disponível em: <https://www.apachefriends.org/download.html>.

O software apresenta quatro características básicas, que permitem: as edições dos usuários, o gerenciamento, a garantia da qualidade e a edição do vocabulário. Com essas características, torna-se possível estabelecer os relacionamentos e garantir que a estrutura gerada seja compatível com outros sistemas, ou seja, o software é vantajoso por proporcionar a eliminação de problemas de ambiguidade da linguagem natural e a interoperabilidade. Para isso, o *TemaTres* contém recursos para geração de listas em formatos de metadados para como SKOS/SKOS-Core, XML/BS8723-5 (tesauros), XML/DC (*Dublin Core*), *Metadata Authority Description Schema* (MADS), XTM (Mapa de tópicos), dentre outros formatos (SANTOS *et al.*, 2021).

Segundo Gonzales-Aguilar, Ramírez-Posada e Ferreyra (2012, p. 319), o *TemaTres* caracteriza-se como um aplicativo para uso na web, e, dentre as suas funcionalidades, destaca-se a oferta “[...] de acesso contínuo ao conteúdo informacional de documentos sobre temas específicos para estudo, pesquisa e tomada de decisão [...]”. Na interface do aplicativo, a disposição dos elementos (menus, botões, links) é de fácil entendimento, o que facilita a navegação. A Figura 27 apresenta a interface principal de navegação da ferramenta.

FIGURA 13 - Interface principal de navegação do *TemaTres*



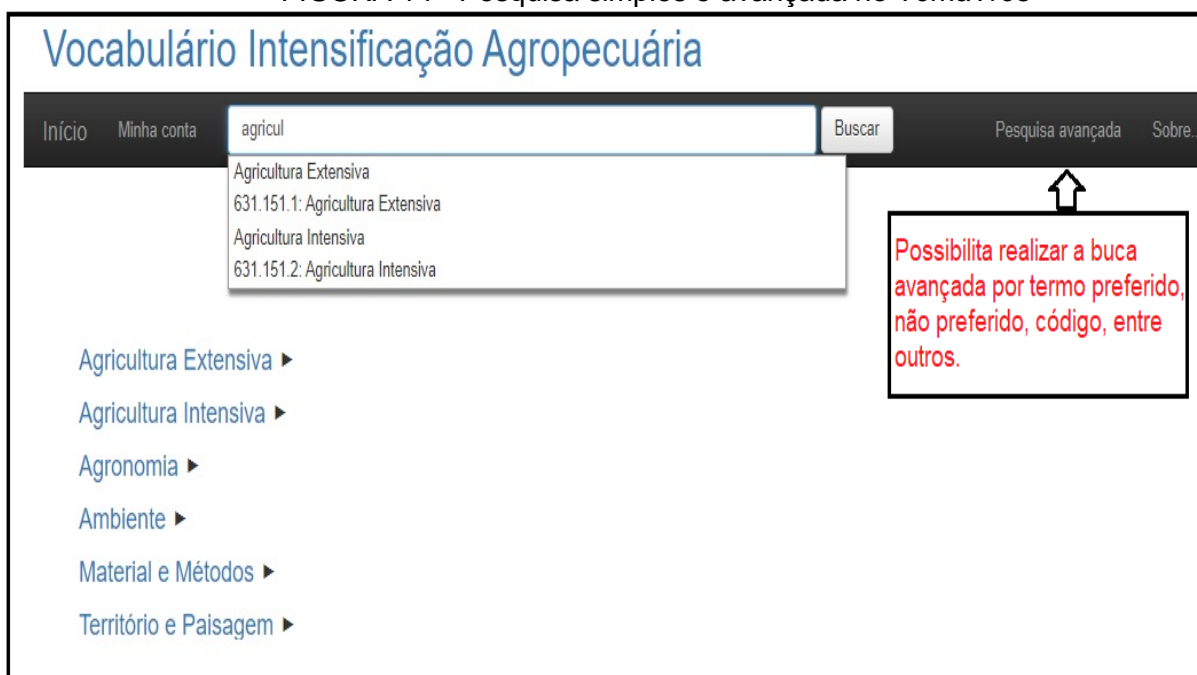
FONTE: Dados da pesquisa (2021).

Diante da Figura 13, observa-se os elementos da interface, onde são apresentados os principais a seguir. A navegação pode ser por meio do índice alfabético ou sistemático, ou seja, que ocorre pela hierarquia das classes ou categorias, permitindo a navegação coordenada por links estabelecidos pelas regras das relações semânticas estabelecidas entre os termos mais genéricos até os mais específicos que estão presentes em cada categoria. Dentre as principais funcionalidades do menu, estão evidenciadas a configuração geral do sistema, onde são realizadas a gestão de dados dos usuários e do vocabulário. Ainda de acordo com a Figura 13, ficou evidenciado que os termos “Intensificação Agropecuária” e “Manejo da cultura” estão subordinados a mesma classe da CDU (631.15) no *TemaTres*, conforme o modelo proposto no subcapítulo 5.1, assim como os termos “Intensificação Agropecuária” e “Pousio”, “Período de Pousio” foram compatibilizados com as classes 631.15 e 631.581 da CDU respectivamente.

A opção de importação de vocabulários é possível por meio de arquivos no formato TXT ou PDF, com o texto no formato de lista hierárquica (etiquetada), alfabética (tabulada), no formato SKOS, entre outros. A exportação é realizada pelas listas no formato hierárquico (etiquetado) e alfabético (tabulado), SKOS, *Dublin Core* (DC), *Metadata Authority Description Schema* (MADS), *Extensible Markup Language* (XML), *JavaScript Object Notation* (JSON), XTM (Mapa de tópicos), dentre outros. Nesta pesquisa, será detalhado somente o padrão de metadados na linguagem SKOS, por causa das limitações de escopo e tempo para o desenvolvimento da pesquisa. Após a exportação do vocabulário, é possível importá-lo em outros sistemas que tenham o mesmo propósito do *TemaTres*, sem haver qualquer perda de informação.

Para configurar o vocabulário é necessário inserir e editar dados como título, autor, idioma, data de criação, coautores, editora, email para contato, palavras-chave, URI, entre outros. Neste menu também é possível fazer a gestão dos tipos de nota (nota de escopo, nota catalográfica, entre outras), das relações semânticas (TG, TE, TR, entre outros), os tipos de links necessários, a serem mostrados a seguir. Os links estabelecem relações de correspondência (*matching*) entre outros sistemas, por meio da linguagem SKOS (<*skos:exactMatch*>, <*skos:broadMatch*> e <*skos:narrowMatch*>), utilizando um *Uniform Resource Identifier* (URI), conforme relatado por Sánchez *et al.* (2012).

No menu adicionar termo, é possível inserir um ou vários termos, colocando um termo por linha no campo “Novo termo”, onde cada termo inserido será um “termo candidato” na construção do vocabulário. Esta interface também possui um campo de busca que permite o usuário pesquisar pelos descritores/termos da CDU e do AgroTermos, seja por termos preferidos ou não preferidos, com a opção de autocomplete, que permite aparecer uma lista de opções dos termos à medida que o usuário vai inserindo o nome do termo no campo de pesquisa, conforme demonstrado na Figura 14.

FIGURA 14 - Pesquisa simples e avançada no *TemaTres*

FONTE: Dados da pesquisa (2021).

Ao observar a Figura 14, a opção de pesquisa simples durante a busca, caso o usuário selecionar um termo não preferido em cada SOC, o vocabulário controlado faz com que o usuário navegue até chegar no termo preferido e sua notação correspondente na CDU, de acordo com as relações semânticas estabelecidas no modelo proposto (subcapítulo 5.1).

Para cada categoria de termos, quando o usuário for sendo guiado durante a navegação por meio de links de navegação, até o termo preferido, torna-se possível obter o padrão de metadados com os termos dos dois SOC e seus relacionamentos semânticos nos diversos padrões já mencionados pelo *TemaTres*, que nesta pesquisa foi usada a linguagem SKOS. Esta linguagem foi escolhida, pois contém informações da estrutura de cada termo em diversos SOC, que no caso desta pesquisa, trata-se das unidades conceituais da CDU e do AgroTermos na linguagem SKOS/SKOS-Core. O termo preferido é o mesmo que está definido pelo tesouro, por se tratar de um descritor mais comum de ser utilizado pelos usuários durante a busca para suprir suas necessidades de informação e por estar relacionado com todos os outros termos não preferidos (alternativos). Ainda com relação a Figura 14, a opção de pesquisa avançada permite realizar consultas por termos preferidos, não preferidos, por código entre outros, com a opção da descrição dos termos ser de forma exata ou aproximada.

A aplicação dos vocabulários controlados no ambiente digital no *TemaTres*, são gerenciados, tendo em vista as representações formais do conhecimento. Esse software, segundo Gonzales-Aguilar, Ramírez-Posada e Ferreyra (2012), permite inserir e editar cada termo do vocabulário controlado, gerenciar e definir as relações semânticas entre os termos e possui níveis de hierarquia com número ilimitado de termos, podendo atribuir código único

para cada termo.

No contexto desta pesquisa, o tesauro foi utilizado para auxiliar a CDU e facilitar a delimitação, padronização, mapeamento de assuntos do domínio da agricultura e por fim, foi utilizado o software *TemaTres* para a representação e recuperação do modelo proposto.

Nesse sentido, no próximo subcapítulo, serão apresentadas as relações semânticas entre conceitos usados no contexto da BCI, com o intuito de minimizar as inconsistências existentes entre os termos e significados entre os conceitos nesse domínio do conhecimento.

3.2 Relações Semânticas na Biblioteconomia e Ciência da Informação

As relações semânticas são definidas como associações entre conceitos que explicitam os relacionamentos entre eles, ou seja, estabelecem um significado semântico, por meio de links entre os conceitos/entidades que estão sendo relacionados (KHOO e NA, 2006). Para os autores, as relações semânticas promovem a comunicação entre o usuário, desde a concepção inicial desse sujeito sobre um determinado assunto, até o conhecimento mais profundo. Esse processo baseia-se em teorias e experimentos realizados por pesquisadores em estudos específicos, com valiosa contribuição para a construção de instrumentos voltados à organização do conhecimento (KHOO e NA, 2006).

Para Hjørland (2007a), as relações semânticas são discutidas não somente na área da Biblioteconomia e Ciência da Informação, como também na Ciência da Computação, na Linguística, na Terminologia, entre outras áreas do conhecimento, constituindo, portanto, um campo interdisciplinar. Em suma, ao observar o mundo real e listar os conceitos existentes em um domínio ou sobre objetos, é possível representar o entendimento do mundo a partir das conexões entre os conceitos. Assim, os conceitos são conectados por meio das relações semânticas, de modo a estabelecer significados objetivos na produção do conhecimento e no estabelecimento das relações dentro de um domínio do conhecimento (DAHLBERG, 1978a; HJØRLAND, 2007a).

Na concepção de Dahlberg (1978a), as relações agrupam os conceitos por suas características. Dessa maneira, profere a autora que, quando as características são as mesmas entre conceitos diferentes, infere-se que possuem relações semânticas entre esses conceitos, por exemplo, os conceitos “preparar o solo”, “arar o solo” e “semear e colher” possuem uma característica em comum: a “atividade agrícola”. Com efeito, depreende-se que as relações semânticas constituem os relacionamentos essenciais que evitam problemas de incompatibilidade de linguagem entre os conceitos estabelecidos (DAHLBERG, 1978a). Essa definição básica das relações semânticas também foi apresentada em diversos estudos

publicados, tais como: Dahlberg (1981, 1983), Frâncu (2003a, 2003b), Slavic (2005), Brascher (2014), entre outros.

Hjørland (2007b) descreve dois tipos de relações: as semânticas, estabelecidas entre conceitos, e as lexicais, construídas entre as palavras, sendo que, nesta tese, serão abordadas, tão somente, as relações semânticas. Nesse sentido, é perceptível a necessidade da organização dos conceitos, de modo a facilitar o estabelecimento das relações semânticas entre eles, por meio de termos ou signos linguísticos (HJØRLAND, 2007b).

A partir da definição atribuída às relações semânticas, entende-se que, sem os conceitos, essas relações não existiriam (KHOO; NA, 2006). Portanto, os conceitos e as relações semânticas são peças fundamentais para representar o conhecimento sob uma única ótica ou linha de raciocínio, com possibilidades de se fazer inferências sobre um determinado assunto.

Uma das características das relações semânticas entre os conceitos é que elas podem ser identificadas por meio de um conhecimento prévio sobre o assunto, pela análise à literatura (garantia literária), pelos relacionamentos entre os conceitos do mundo real, entre outras possibilidades (KHOO; NA, 2006). Nesse sentido, a fim de classificar os conceitos e estabelecer as relações semânticas entre os SOC, não é necessário que os conceitos estejam incluídos no SOC, porém, sem as relações semânticas, não seria possível a criação do sistema (HJØRLAND, 2007a).

Nesse contexto, Brascher (2014) relata que as relações entre conceitos não representam as relações entre os termos para expressar os conceitos. É importante ressaltar a relevância de estudos e análises semânticas para a representação ou modelagem conceitual de diversos domínios do conhecimento, a partir das relações entre os conceitos para a construção do SOC. Esses sistemas são instrumentos que organizam, gerenciam e recuperam informações, cuja base dos principais SOC são os termos e as relações entre eles, o que garante ao sistema a capacidade de definir e compartilhar conceitos de áreas específicas do conhecimento (BRASCHER, 2014).

É importante ressaltar que as relações semânticas devem ser elaboradas de acordo com as necessidades de informação dos usuários, com o fim de tornar mais acessível a comunicação entre eles e a informação (GARCÍA-MARCO, 2017). No processo de construção dessas relações, segundo García-Marco (2017), são requeridas três propriedades básicas: a reflexiva, transitiva e a simétrica. A primeira ocorre quando a relação é entre o mesmo conceito; na segunda, os termos representantes dos conceitos são expressos por um relacionamento transitivo e explícito, por exemplo, se A está relacionado com B e B está relacionado com Z, isso implica que A é relacionado com Z, isso é, $A=B$ e $B=Z$ e, então, $A=Z$; por sua vez, a terceira propriedade indica uma relação de equivalência entre os conceitos, como: $A=B$ e $B=A$ (GARCÍA-MARCO, 2017).

Para Dahlberg (1978a), as características das relações semânticas possibilitam a categorização delas por semelhanças ou pontos comuns. Um exemplo de categorias por pontos comuns são as relações do tipo lógicas que utilizam a lógica proposicional, tais como:

- **Identidade:** a característica que identifica um conceito em relação a outro, ou seja, são equivalentes, sendo que as características do conceito A são iguais as do conceito B;
- **Implicação:** ocorre se uma característica do conceito de origem é verdadeira; implica-se, então, que essa mesma característica também será verdadeira no conceito de destino, mas não o contrário;
- **Intersecção:** refere-se às características comuns entre os dois conceitos, ou seja, alguma características de A também existe em B;
- **Disjunção:** manifestam-se quando as características comuns não existem, ou seja, as características de A são diferentes das características de B, isso é, são mutuamente exclusivas e;
- **Negação:** é estabelecida quando a característica entre os conceitos é de contradição, isso é, ocorre quando a característica de A é a negação em B, entre outras possibilidades (DAHLBERG, 1978a).

Além das características das relações semânticas lógicas entre os conceitos e/ou entidades, Khoo e Na (2006) apresentam outros tipos. Os autores citam três tipologias, a saber: (a) Abstração: apresenta ideias ou pensamentos; experiências em relação aos conceitos; (b) Conceitualização: indica as características a partir da análise cognitiva sobre os conceitos de um domínio específico; e (c) Distinção: ocorre quando existem diferentes significados entre os conceitos, a partir do agrupamento dos conceitos ou termos para definir a característica da relação entre eles. Essas tipologias de relações são estabelecidas de modo mais genérico, considerando as características essenciais das relações (KHOO; NA, 2006). Todavia, as relações podem ser classificadas a partir de características mais específicas, que possibilitam a formação de estruturas bem definidas, conforme apresentado no subcapítulo seguinte.

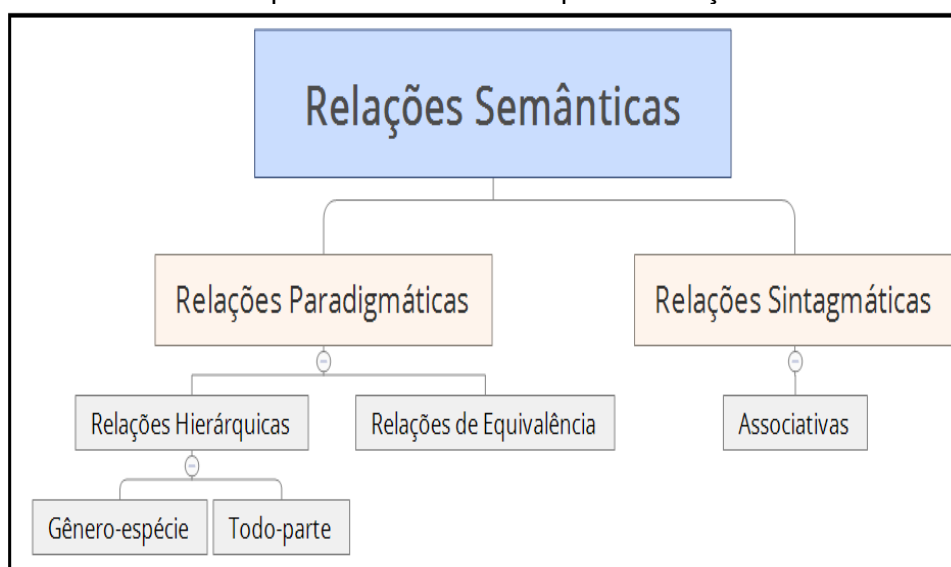
3.2.1 Classificação das relações semânticas

As relações semânticas são divididas em diferentes tipos e maneiras, por exemplo, verificando como os usuários de uma determinada área indicam os termos que expressam os conceitos, considerando que, um mesmo termo pode ser usado de diferentes maneiras, dependendo da área de atuação do usuário ou contexto onde está inserido (KHOO; NA, 2006). Assim, para garantir consistência na representação, as relações semânticas precisam ser estabelecidas a partir de diferentes aspectos, como: a garantia literária por meio

da análise documental, a garantia do usuário (especialistas de domínio e usuários finais), garantia da realidade que se pretende representar, entre outros (HJØRLAND, 2007a). Nesse sentido, é preciso definir o significado ou tipo de relação que pode ocorrer entre os conceitos envolvidos e qual a terminologia utilizada pelos linguistas, fortalecendo, assim, o uso das relações semânticas entre os SOC e não somente entre os termos (BROUGHTON *et al.*, 2005).

Em relação aos conceitos e seus relacionamentos entre SOC, Maniez (1988) e Svenonius (2000) apontam a existência de duas categorias: relações paradigmáticas e sintagmáticas. O primeiro tipo indica informações de conceitos com relacionamentos hierárquicos (gênero-espécie e todo-parte), estabelecidos previamente pela veracidade de relações entre os descritores. Desse modo, essas relações permitem organizar conceitos da mesma natureza, ou seja, de uma mesma categoria ou classe gramatical, contemplando, na maioria das vezes, substantivos, verbos, adjetivos, advérbios, entre outros. As relações sintagmáticas, por sua vez, são estabelecidas de forma ocasional e representadas pelas relações associativas, que ocorrem com palavras em uma mesma frase e com a definição de regras gramaticais e de sintaxe dentro de um domínio específico, verificada a veracidade da relação. Nas relações sintagmáticas, uma sentença genérica ou permanente contém alguma informação que é comumente conhecida, por exemplo, “a manga da camisa é azul” (MANIEZ, 1988; SVENONIUS, 2000). A Figura 13 apresenta um esquema ilustrativo para essa classificação específica das relações semânticas.

FIGURA 15 - Esquema ilustrativo dos tipos de relações semânticas



FONTE: Elaborada pelo autor (2021).

Na Figura 15, considera-se que, nas relações paradigmáticas, uma expressão que indica um conceito de forma ocasional ocorre desde que contenha informações sobre um contexto específico, como na estrutura: “o abacaxi é uma fruta” (STOCK, 2010). Entretanto,

Khoo e Na (2006) observaram que a diferença entre relações paradigmáticas e sintagmáticas não é precisa, pois as relações paradigmáticas podem ser expressas no nível sintático. Para os autores, conforme mostra a figura 15, as relações paradigmáticas representam as relações hierárquicas (Gênero-espécie), partitivas (Todo-parte) e de equivalência, capazes de representar uma possibilidade finita entre os termos, contidos do geral para o específico, e utilizados em uma abordagem empírica (KHOO; NA, 2006).

A relação hierárquica baseia-se na relação de gênero-espécie; a relação lateral (horizontal), nos renques; e a relação partitiva (todo e suas partes) baseia-se na relação de oposição (contrária ou contraditória). Por esse motivo, essas relações tornam-se limitadas em relação às ontologias, que apresentam as propriedades entre os conceitos no padrão da web semântica (ZENG, 2019). Khoo e Na (2006) relatam que é possível realizar relações semânticas entre diferentes conceitos de diversas categorias, de modo a possibilitar infinitas combinações de acordo com as necessidades dos usuários (KHOO; NA, 2006).

Na área da BCI, as relações semânticas propostas por Maia (2018) é abordada em profundidade, com a proposta de uma taxonomia com essas três tipologias (relações hierárquicas, equivalentes e associativas), motivo pelo qual essas relações também foram escolhidas. A seguir, foram detalhadas apenas as relações que foram utilizadas no mapeamento dos conceitos utilizados no modelo proposto nesta pesquisa, tendo em vista embasar a estruturação do produto final deste estudo (modelo proposto para a recuperação da informação com uso entre dois SOC: CDU e tesouro). A estrutura taxonômica proposta pela autora possui um nível de especificidade adequado para estabelecer relações semânticas entre conceitos que são comumente utilizados na área da BCI. Por esta razão, a padronização da terminologia tem a função de diminuir a ambiguidade da linguagem natural (homógrafos e polissemia), a partir do nível macro com as relações hierárquicas (Hiperônimo-Homônimo, Holônimo e Merônimo), de equivalência e associativas, até atingir os níveis específicos (MAIA, 2018).

Com o intuito de diferenciar cada uma das três relações abordadas por Maia (2018), apresenta-se o Quadro 9. Nesse Quadro, são apresentadas breves definições de cada um desses três tipos de estruturas.

QUADRO 9 - Principais definições para relações hierárquicas, equivalentes e associativas

Tipo de relação	Definições
Relações Hierárquicas	Constituem as tipologias mais utilizadas, organizadas do geral para o específico (gênero-espécie), formando uma estrutura que indica os níveis de superordenação (hiperonímia), para termos genéricos, e subordinação (hiponímia), para termos específicos formando classes/categorias. As relações utilizadas em SOC, para formular as classes, recorrem a dois métodos, que são o método indutivo (<i>bottom-up</i>) e o método dedutivo (<i>top-down</i>) (SOERGEL, 1974).
Relações de Equivalência	São definidas entre um ou mais termos (descritores) que possuem o mesmo significado para representar um conceito, ou seja, quando dois ou mais conceitos são correspondentes entre diferentes vocabulários,

	apresentando significados próximos em um determinado contexto. Os termos que são considerados sinônimos estabelecem uma relação de sinonímia entre os conceitos semelhantes, e a diferença entre os tipos de sinonímia são a abrangência de equivalência entre os conceitos (DODEBEI, 2002).
Relações Associativas	Referem-se àquelas que não ocorrem de forma hierárquica e nem são equivalentes, podendo ocorrer entre termos na mesma hierarquia ou em hierarquias diferentes. Esse tipo de relação pode ocorrer entre um termo com outros termos que são relacionados ou associados com o termo de origem, e são representados como remissivas (veja também) ou termo relacionado (TR) (KHOO; NA, 2006).

FONTE: Elaborado pelo autor (2021). Adaptado de Soergel (1974), Dodebei (2002) e Khoo e Na (2006).

Para Broughton *et al.* (2005) a relação de hierarquia Hiperônimo-hipônimo (gênero-espécie) ocorre quando existe uma relação firmada entre classes e subclasses. Em linhas gerais, uma classe genérica (superordenada) é constituída das mesmas características das classes subordinadas (específicas) a ela; já em relação à espécie, ela possui, além das mesmas características da classe genérica, uma ou mais características específicas que as diferenciam. Esse tipo de relação apresenta uma classe central que constitui o todo, de onde se subdividem diversas subclasses, vinculadas à classe central, considerada como um ponto de convergência, o que permite verificar se um objeto é subordinado (espécie) a outro mais superordenado (gênero) (BROUGHTON *et al.*, 2005). A relação Hierárquica - Hiperônimo-Hipônimo - Instância, conforme relata Stock (2010), é descrita como sendo uma relação entre um objeto que possui um nome em uma classe, por exemplo, UFMG é uma Universidade.

Conforme Chaffin e Herrmann (1984), a inclusão de classe acontece pelas relações “é um” ou “é um tipo de”, e o significado de uma classe está incluído em outra, por exemplo, uma classe de ipê-amarelo está incluída na classe de árvores, ou seja, o ipê-amarelo é uma árvore, ou é um tipo de árvore. Nesta pesquisa, foram utilizadas as classificações de *Inclusão de classe*, tais como: (1) funcionalmente subordinado: são objetos que se caracterizam por suas funções, por exemplo: trator é um veículo para atividades agrícolas; (2) atividade subordinada: envolve atividades entre os objetos, por exemplo: cultivo de plantas e hortaliças é uma atividade da agricultura, ou cultivo de plantas e hortaliças é um tipo de atividade da agricultura; (3) Taxonômica: é quando dois objetos possuem a relação de espécie-indivíduo, onde a espécie é representada por um conjunto de indivíduos com características ou semelhanças entre si. Esta relação é estabelecida quando não pode ser classificada em nenhum outro tipo de relação de inclusão de classe (CHAFFIN; HERRMANN, 1984).

Na relação Holônimo-Merônimo (todo-parte), é entre o todo e suas partes, ou seja, com a união das características das partes, que se forma o todo (CINTRA *et al.*, 2002). As relações de objeto estruturado que, nesta pesquisa, utilizou-se a relação (1) componente-complexo, conforme Stock (2010), tem a função de propriedade ou função do objeto, cujo objeto pode ser dividido em partes ou componentes, por exemplo, as folhas são partes da

árvore; e (2) Segmento-evento, que de acordo com o autor, o evento se difere de objeto como foi descrito anteriormente em (1), pois o evento que representa o todo, pode ter segmentos (partes) que podem ocorrer em diversos períodos de tempo. Um exemplo dessa relação é o evento “aquecimento global” que é causado por emissões de gases poluentes como o (CO₂ – dióxido de carbono) pela queima de combustíveis fósseis, desmatamento, mudanças no uso da terra, atividades industriais, entre outros fatores. Já a relação de instância é estabelecida entre categorias ou classes, como sendo um objeto dessa categoria, por exemplo, “João”, “Maria” é uma instância da classe “Pessoa” que são objetos representados por esta classe. (ANSI/NISO Z39.19, 2005).

No que se refere às relações de equivalência, nesta pesquisa, foram utilizadas classificações do tipo:

- **Sinônimo absoluto:** essa relação ocorre quando os termos possuem significados iguais em relação ao mesmo conceito em sua totalidade, ou seja, de forma abrangente e/ou quando possuem o mesmo termo em diferentes vocabulários, por exemplo: reinado, império ou monarquia;
- **Quase sinônimo:** esse tipo manifesta-se quando os termos compartilham significados próximos ou similares, mas podem estabelecer algumas diferenças por combinação ou sobreposição, por exemplo: “quase” e “por pouco”, ao apresentarem significados próximos, “Biologia” e “Informática”, que correspondem a estudos da área da Bioinformática (sobreposição);
- **Sinônimo próximo:** diz respeito aos termos que não possuem o mesmo sentido, ou seja, de forma idêntica, mas possui significados próximos, por exemplo, veículo e automóvel;
- **Sinônimo parcial:** nesse caso, os termos compartilham vários sentidos, iguais em muitos significados, mas se diferem em algum aspecto, por exemplo: língua no sentido de idioma e no sentido de um órgão do corpo humano (KHOO; NA, 2006).

Quanto às relações associativas, no contexto desta pesquisa, foram utilizadas as relações do tipo causal, ou seja, quando um sujeito (agente) causa uma ação ou efeito. Esse tipo de relação se manifesta em duas situações: diante de um agente-ação e diante de um agente-objeto. O primeiro caso ocorre quando um agente pratica uma ação, por exemplo: Fernando cortou a grama do jardim; já na relação do tipo agente-objeto, é estabelecido o relacionamento entre o agente e o instrumento utilizado, como também, uma relação com o que o agente produziu ou fabricou, por exemplo: o jardineiro utiliza uma máquina de aparar grama. Existem relações semânticas associativas que não representam causa, por exemplo a relação Disciplina-fenômeno e Processo-instrumento. A primeira relação indica um objeto que é estudado em uma área ou disciplina, por exemplo, os planetas são fenômenos que são

estudados na disciplina “Astronomia” e a segunda relação relaciona um processo com um instrumento, por exemplo, medição de tempo e relógio (ANSI/NISO Z39.19, 2005).

É importante salientar que esses tipos de relações constituem alguns exemplos, sendo possível estabelecer muitas outras modalidades, desde que seja estabelecida a lógica para firmar as relações. O Anexo D apresenta uma estrutura da taxonomia de relações semânticas no contexto da BCI, conforme a proposta de Maia (2018).

Diante do exposto, Maia (2018) relata a importância de se conhecer as relações semânticas utilizadas em cada domínio ou área do conhecimento, devido às diferenças terminológicas existentes, as quais podem ser utilizadas nos instrumentos de organização do conhecimento. A autora confirma a utilização das relações semânticas também no contexto digital, que pode ser realizada por inferências entre as relações e entre os conceitos, ação essa conduzida por pessoas e máquinas, o que permitirá, portanto, consistência na inteligibilidade para a representação do conhecimento (MAIA, 2018).

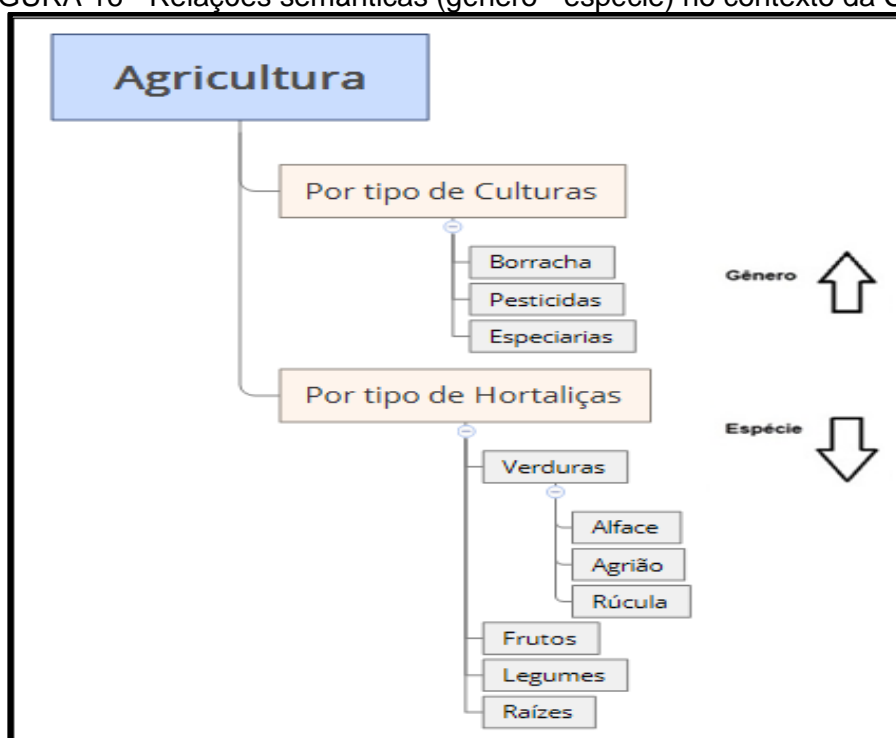
No próximo subcapítulo, são descritas as relações semânticas nos SOC, com foco na estrutura da CDU e do tesauro, além de expor reflexões dessas relações no domínio da agricultura. Essa apresentação fornecerá insumos para o desenvolvimento do modelo proposto nesta pesquisa, de modo que as relações estabelecidas sejam compatíveis, haja vista garantir melhorias à recuperação da informação.

3.2.2 Relações semânticas em Sistemas de Organização do Conhecimento

A CDU é composta por relações semânticas entre os conceitos que expressam o conteúdo dos itens informacionais existentes nos acervos das bibliotecas ou unidades de informação. O modo com que os conceitos estão estabelecidos possibilita que os recursos relacionados com a mesma temática fiquem fisicamente próximos (CINTRA *et al.*, 2002). Ademais, o sistema classificatório possui mecanismos ou sinais que facilitam a conexão e/ou combinação entre os conceitos, tais como “:” ou “+”. Essas conexões não precisam ser transportadas para a estrutura do SOC em questão (SLAVIC, 2005).

Com base nas propriedades estruturais e no amplo uso dos sistemas de classificação, tanto em bibliotecas convencionais quanto digitais, eles podem ser estruturados em categorias (facetadas) e subcategorias (subfacetadas) predeterminadas (SANTOS; MAZINI, 2011). Por essa razão, segundo os autores, a CDU utiliza as relações semânticas do tipo gênero-espécie e todo-parte, cujos conceitos são estabelecidos de acordo com as categorias/classes. A fim de exemplificar a relação gênero-espécie, a partir do modo com que é estruturada na CDU, apresenta-se a Figura 16.

FIGURA 16 - Relações semânticas (gênero - espécie) no contexto da CDU

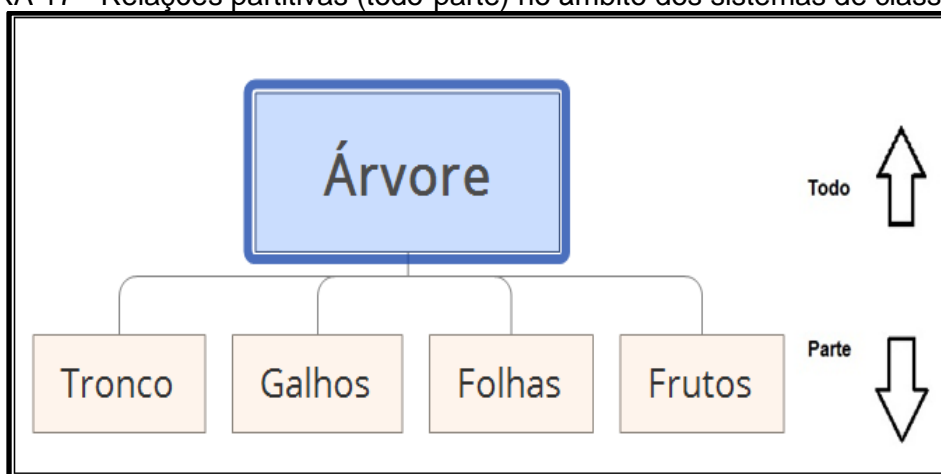


FONTE: Elaborada pelo autor (2020).

A Figura 16 expõe a relação gênero-espécie estabelecida em uma classe de assunto pertencente ao domínio da agricultura. Nota-se que as relações entre as subclasses possuem características específicas, ou seja, a “classe filha” (subclasse), representada na Figura por “tipos de Hortaliças”, possui uma ou mais características em relação à classe “tipos de Culturas” (classe principal) (VICKERY, 1980).

Outra relação possível de ser estabelecida entre as classes em um sistema de classificação é a relação partitiva (todo-parte). Para ilustrar esse tipo de relação, apresenta-se a Figura 17, a qual exemplifica a relação entre as partes de uma árvore.

FIGURA 17 - Relações partitivas (todo-parte) no âmbito dos sistemas de classificação



FONTE: Elaborada pelo autor (2020).

A Figura 17 mostra que as partes (raízes, tronco, galhos, folhas, flores e frutos) fazem parte do todo, que é a árvore. Dessa maneira, árvore possui divisões que são as partes, sendo a própria árvore o todo (CINTRA *et al.*, 2002).

Em relação aos tipos de relações semânticas dos tesouros, na visão de Weiss (2014), prevalecem: 1 - relações de equivalência (entre os termos preferidos e não preferidos); 2- relações hierárquicas (entre termos gerais e específicos); e 3 - relações associativas (ligações entre os termos que não são equivalentes e nem hierárquicos, mas que ainda assim são termos semanticamente ou conceitualmente associados a um ponto tal que se considera necessário explicitar a ligação entre eles).

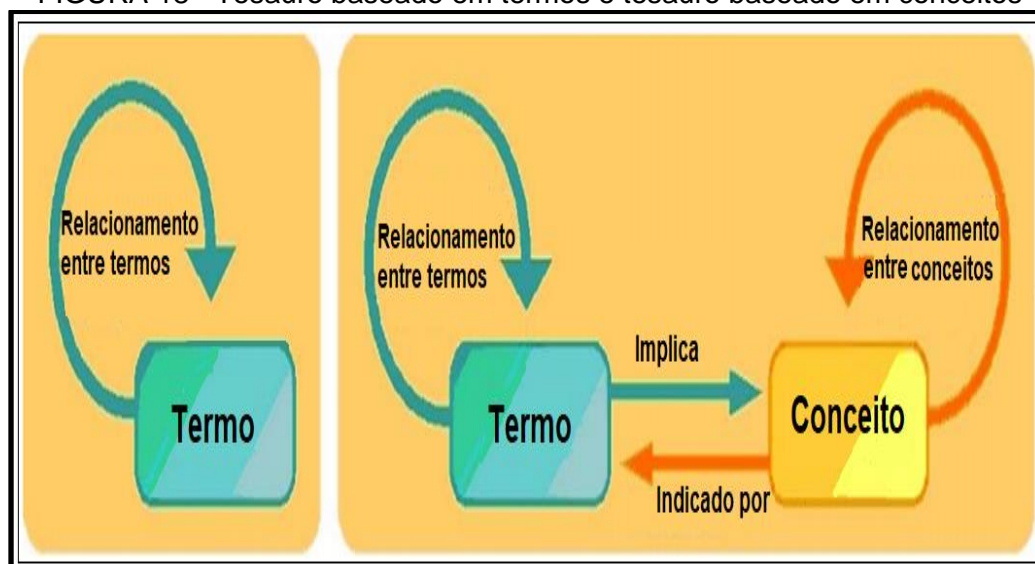
Na abordagem de construção dos tesouros, as relações associativas são, na maioria das vezes, aplicadas e detalhadas após certo nível de maturidade para representar e utilizar os termos. A princípio, esses termos podem ser escolhidos por meio das relações de hierarquia. Um exemplo dessa possibilidade é o sistema de informação da Biblioteca Nacional de Medicina, da Universidade *Oxford University Press*, que utiliza a *Unified Medical Language* (UMLS). Nesse sistema, definiram-se mais de 40 relacionamentos, os quais foram iniciados com as relações de hierarquia e de equivalência, e finalizados com as relações associativas (FRÂNCU, 2003a, 2003b).

As relações associativas existentes nos tesouros possibilitam ligar diferentes termos, a partir de suas semelhanças, o que favorece a expansão da notação que pode ser atribuída ao item informacional (CINTRA *et al.*, 2002). Nos tesouros, as relações associativas são importantes, visto que elas facilitam o resgate de termos correlatos, por conseguinte, ampliam as possibilidades de relacionamento entre outros termos correspondentes em um domínio do conhecimento (CAMPOS *et al.*, 2006).

Existem duas abordagens consideradas o núcleo central dos tesouros. A primeira caracteriza o tesouro como um conjunto de termos interconectados com diferentes tipos de relacionamento semântico. A representação clássica proposta pela ISO, pelos padrões da *American National Standards Institute* (ANSI) e pela *National Information Standards Organization* (NISO) descreve um tesouro como um conjunto de termos relacionados, com prevalência de relações hierárquicas (termo genérico e termo partitivo), associativas (termos relacionados) e de equivalência (termo equivalente, termo preferido e não preferido). Por sua vez, a segunda alternativa concebe o tesouro como um conjunto de conceitos firmados por relações semânticas semelhantes às de um dicionário de sinônimos baseado em termos. Porém, os conceitos estão atrelados aos termos no nível lexical, na forma de rótulos lexicais, ou seja, de termos (palavras) que podem ter mais de um significado, podendo ou não conter relações estabelecidas entre eles (PASTOR-SÁNCHEZ; MARTÍNEZ MÉNDEZ; RODRÍGUEZ-MUÑOZ, 2009).

Portanto, na concepção dos autores citados, entende-se que os tesouros podem ser estruturados por relações entre seus conceitos e entre seus termos. A Figura 16 apresenta a comparação entre os tesouros que são sustentados pelos relacionamentos entre os termos e os que são concebidos nos relacionamentos entre conceitos.

FIGURA 18 - Tesouro baseado em termos e tesouro baseado em conceitos



FONTE: Adaptada de Pastor-Sánchez, Martínez Méndez e Rodríguez-Muñoz (2009, p. 3).

As indicações da Figura 18 revelam que, no tesouro baseado em termos, as relações semânticas são estabelecidas entre os próprios termos. Já no tesouro baseado em conceitos, os relacionamentos semânticos são estabelecidos entre os conceitos, cujo termo designa o conceito, que é atribuído por uma unidade conceitual, sem importar com quantos termos sejam constituídos para identificar um conceito ou unidade conceitual. No tesouro baseado em conceitos, as relações entre os conceitos, na maioria das vezes, incluem aqueles que desenvolvem a hierarquia e as estruturas associadas, as quais denotam as relações entre os termos associados a um único conceito (abreviações, uso comum, técnicas, idioma, dentre outros) (PASTOR-SÁNCHEZ; MARTÍNEZ MÉNDEZ; RODRÍGUEZ-MUÑOZ, 2009).

Nesse contexto de estruturas semânticas, devem-se ficar claras as relações semânticas entre a CDU e o tesouro, de modo semelhante aos demais SOC. Isso precisa ser considerado na elaboração de instrumentos de organização em quaisquer domínios, incluindo a agricultura (MACULAN, 2015). Assim, é importante apresentar as características, definições e exemplos de como essas relações são firmadas quando aplicadas no domínio da agricultura, de modo a delimitar o escopo de aplicação metodológica desta tese.

Maculan (2015), ao refletir sobre o uso dos tesouros no domínio da agricultura, destacou que as relações de equivalência estão presentes no contexto de um discurso ou campo de conhecimento, com variações decorridas de modificações de propriedades ou criação de novos conceitos. Nos tesouros, “[...] essa ocorrência reflete o aspecto dinâmico da terminologia de diferentes domínios, e é necessário que as equivalências estejam

representadas no tesauro que modela tal campo de conhecimento [...]” (MACULAN, 2015, p. 145).

A relação de equivalência é destacada na orientação para desenvolver mecanismos de recuperação no domínio da agricultura, conforme salientado por Maculan (2015). Ressalta-se que, em uma abordagem pragmática, as relações de equivalência, nos tesauros, manifestam-se como a distinção do que foi considerado igual ou equivalente para fins de revocação na recuperação (WEISS; BRÄSCHER, 2017). As autoras entendem que esse tipo de relação diz respeito a uma medida que minimiza as implicações éticas de se escolher uma forma de expressão como termo preferido, além de proporcionar maior flexibilidade ao tesauro.

Considerando a importância de as relações de equivalência estarem claras nos tesauros, como mencionado por Weiss e Bräscher (2017), nota-se a necessidade de constantes análises e adequações nos tesauros, sobretudo em domínios específicos. No caso da agricultura, Maculan (2015), ao estudar o THESAGRO, que é um dos tesauros incorporados no AgroTermos, percebeu a existência dessa relação, embora a relação não se apresente de modo consistente. O Quadro 10 exemplifica a relação de equivalência para o termo “Mandioca”.

QUADRO 10 - Relações de equivalência no AgroTermos

MANDIOCA UF AIPIM UF MACAXEIRA RT MANHOT ESCULENTA RT MANHOT UTILÍSSIMA	MANIHOT ESCULENTA RT MANDIOCA
	MANIHOT UTILÍSSIMA RT MANDIOCA

FONTE: Adaptado de Maculan (2015, p. 202).

A análise ao Quadro 10 identifica uma relação de equivalência, visto que o descritor “Mandioca” é o preferido, e “Aipim” e “Macaxeira” são descritores não preferidos. No entanto, nota-se falta de clareza, ou melhor, ausência de detalhamento nos termos equivalente. Sobre isso, Maculan (2015, p. 203) mencionou que “[...] a falta de especificação do tipo de equivalência que há entre esses descritores não permite inferir que eles são variações regionais para um mesmo conceito [...]”. Essa constatação evidencia a necessidade de adaptações constantes nos instrumentos terminológicos, de modo a possibilitar um uso desambiguizado e íntegro, sobretudo para usuários não especialistas de um domínio. Além disso, um termo pode haver um ou vários nomes científicos, no exemplo do Quadro 10, Manihot Esculenta e Manihot Utilissima são exemplos para o termo mandioca com relação associativa e não de equivalência, já que se tratam do mesmo significado (MACULAN, 2015).

Em relação as relações hierárquicas, elas também se manifestam nos SOC utilizados no domínio da agricultura. Isso porque há dependência ou subordinação que termos mais amplos exercem sobre termos mais específicos (CINTRA *et al.*, 2002). Corroborar essa ideia Maculan (2015), ao mencionar que esse tipo de relação se fundamenta nos “[...]”

princípios de subordinação, que é o procedimento de subdivisão de um assunto (objeto) em subclasses (termos mais específicos), e de superordenação, que é a identificação das classes de nível mais alto (termos mais gerais) [...]” (MACULAN, 2015, p. 150).

No domínio da agricultura, observa-se a presença das relações hierárquicas. Consultando o THESAGRO, Maculan (2015) identificou vários desses relacionamentos, embora constatasse, também, algumas inconsistências. O Quadro 11 apresenta um exemplo de relação hierárquica, considerando as subdivisões para o conceito “Agricultura”.

QUADRO 11 - Relações de hierarquia no AgroTermos

AGRICULTURA NT AGRICULTURA DE PRECISÃO NT AGRICULTURA EM TEMPO PARCIAL NT AGRICULTURA FAMILIAR NT AGRICULTURA ALTERNATIVA NT AGRICULTURA SUSTENTÁVEL NT AGRICULTURA DE SUBSISTÊNCIA NT AGRIMENSURA NT AGRONOMIA NT FITOTECNIA NT PISCICULTURA CONSORCIADA NT SILVICULTURA NT ZOOTECNIA

FONTE: Maculan (2015, p. 199).

A estrutura apresentada no Quadro 11 indica uma relação de hierarquia para o conceito agricultura, cujo termo “agricultura” está em uma situação de coordenação em relação aos demais termos (*NT*: termos específicos), evidenciando que esses termos são subdivisões da classe superordenada. Pela análise a essa relação, Maculan (2015) aferiu que não está expresso no instrumento quais seriam as possíveis tipologias de hierarquia, ou seja, o termo geral e os específicos exercem relação de gênero-espécie, de todo-parte, dentre outras.

Semelhante às explicações de Maculan (2015), entende-se que, para identificar essas tipologias, o profissional indexador, no momento da representação, ou o usuário do sistema, no momento da busca, possuíse conhecimento prévio ou recorrente a fontes de consulta para garantir maior compreensão acerca dos conceitos expressos pelos termos. Portanto, assim como nas relações de equivalência, nota-se que as relações hierárquicas também precisam ser adaptadas, considerando, sobretudo, as necessidades dos usuários do sistema de informação, cujo conceito foi representado (MACULAN, 2015).

No estabelecimento das relações associativas em SOC da agricultura, é preciso considerar os benefícios gerados no que tange à expansão do assunto, indicando sua contextualização com outros aspectos do domínio, sendo essa uma atividade complexa (SLAVIC, 2005). O estudo de Maculan (2015) confirmou essa dificuldade de precisão das relações associativas. Como exemplo, a autora expõe a associação de um termo a partir de seu nome científico (conhecido por especialistas) e de seu nome popular (normalmente,

conhecido pelo usuário leigo). O Quadro 12 ilustra a existência da relação associativa no tesauro de domínio da agricultura.

QUADRO 12- Relações associativas no AgroTermos

<p>Nome científico: <i>Manihot esculenta</i> Autor: Crantz Nome popular: Mandioca, Macaxeira, Aipim Família: (Cronquist): <i>Euphorbiaceae</i> Gênero: <i>Maniho</i></p>	<p>Nome científico: <i>Manihot utilíssima</i> Autor: Pohl Nome popular: Mandioca, Macaxeira, Aipim Família: (APG2): <i>Euphorbiaceae</i> Gênero: <i>Maniho</i></p>
---	---

FONTE: Adaptado de Maculan (2015, p. 203).

No estudo de Maculan (2015), a autora percebeu que, mesmo o exemplo exposto no Quadro 12 evidencie a relação associativa, possibilitando benefícios para indexadores e usuários, em especial por ampliar as buscas no momento da recuperação, ele apresenta inconsistências, por não representar a equivalência entre um termo científico e outro popular. Nos dizeres da autora:

[...] percebe-se o uso de descritor preferencial tanto para a nomenclatura científica de um organismo vivo quanto para o seu nome popular, pois os descritores “*Manihot Esculenta*” e “*Manihot Stilissima*” são nomes científicos para “Mandioca”. Porém os descritores estão em uma relação associativa (RT), e, não, em uma relação de equivalência com o descritor preferido “Mandioca” (MACULAN, 2015, p. 203).

Além das inconsistências apontadas por Maculan (2015), a situação pode se tornar mais complexa quando se trata da representação e recuperação da informação no ambiente digital. Assim, considerando a aplicação das relações semânticas nos SOC do domínio da agricultura em ambientes digitais, é preciso verificar as condições do ambiente digital, tal como recomendado por Slavic (2005). Segundo essa autora, no processo de navegação, uma hierarquia de classificação pode ser pouco detalhada, com um escopo bastante abrangente. Isso poderá gerar dificuldades para os usuários “[...] obterem uma visão geral de uma área de assunto e a direção na qual navegar. É por isso que é importante ser capaz de 'reduzir' e 'expandir' uma hierarquia [...]” (SLAVIC, 2005, p. 116, tradução nossa).

A autora supracitada reforça que as relações em cadeias, na maioria das vezes, apresentam estruturas incompletas, sendo difícil, também, a incorporação de termos, quando esses instrumentos e estruturas são projetados para o ambiente digital. Segundo a autora, evidencia-se, com frequência, um dos problemas na automação da criação do índice da cadeia que é a adição de termos que estão faltando nas programações. “[...] Além disso, a extração automática de um índice de cadeia requer os dados “corrigidos” e processáveis por máquina na classe hierárquica [...]” (SLAVIC, 2005, p. 130-131, tradução nossa).

Ao comparar o uso das relações associativas em tesouros combinados com sistema de classificação para facilitar a recuperação em ambientes digitais, Slavic (2005) destacou que esses relacionamentos podem ligar diferentes aspectos de um assunto e isso pode ser implementado na recuperação. O propósito dessa combinação, segundo a autora, é

“[...] apoiar a exploração e expansão do assunto através da navegação lateral. Uma referência pode estar vinculada a várias etapas de associação” (SLAVIC, 2005, p. 115, tradução nossa).

Entende-se que, em qualquer tipo de ambiente, seja digital ou na web, um conjunto de relações pode ser considerado completo, na criação de um SOC, considerando que os conceitos precisam estar conectados e organizados de modo lógico, e sejam estabelecidas as relações semânticas capazes de gerar um sistema de conceitos. No contexto desta tese, é utilizada para a representação do conhecimento a estrutura classificatória da CDU, para melhorar a recuperação da informação no ambiente digital. Desse modo, para a representação de um modelo no ambiente digital, é necessária a utilização de padrões de metadados e da linguagem de modelagem Simple Knowledge Organization System (SKOS) no âmbito da web semântica. Além disso, são necessários estudos e análises sobre o domínio específico aplicado para a modelagem conceitual com foco na necessidade de informação dos usuários (ZENG, 2008). Portanto, a utilização de SOC para a representação de um domínio do conhecimento e suas relações semânticas auxilia na recuperação da informação, temática na qual será abordada no próximo subcapítulo.

3.3 Recuperação da informação

Com o aumento exponencial do volume de informação disponível atualmente, é necessária, cada vez mais, a utilização de métodos avançados de recuperação da informação, de modo a atender, com mais qualidade e exatidão, às necessidades dos usuários (MONTEIRO *et al.*, 2017). Saracevic (1996) descreve a recuperação da informação como um processo, que pretende resolver algum problema do usuário. Esse problema pode ser sanado, a partir dos resultados obtidos por meio da formulação de uma pesquisa sobre um determinado assunto. A recuperação ou a busca por informação estão relacionadas com o conhecimento prévio que o usuário possui sobre esse assunto. Em suma, necessita-se saber o que o usuário precisa conhecer para propor soluções ao seu problema, causado pela falta de informação sobre o assunto. Portanto, para garantir a recuperação, o usuário expressa sua necessidade de informação mediante uma estratégia de busca (SARACEVIC, 1996).

O processo de recuperação da informação foi considerado por Saracevic (1999) como um dos campos mais importantes da Ciência da Informação, devido às relações multi e interdisciplinares dessa área com outras áreas do conhecimento. O autor destaca a recuperação como um processo de comunicação entre o emissor (usuário) e o receptor (sistema), processo esse ocorrido em um SRI. Assim, o usuário emite uma pergunta para o sistema, que disponibiliza as informações pertinentes ou relevantes que possam atender uma necessidade de informação (SARACEVIC, 1999). A necessidade de informação, conforme descrito por Le Coadic (2004), constitui um desejo físico do ser humano, que busca solucionar

os problemas que ocorrem a todo o tempo em atividades do cotidiano. Em síntese, o processo de recuperação consolida-se por meio da entrada de dados, no sistema, seguido pelo processamento e finalizado com a saída de documentos que contenham informações relevantes para sanar as necessidades dos usuários (SARACEVIC, 1999).

Dessa maneira, a recuperação da informação é o conjunto de processos, métodos e procedimentos que viabilizam recuperar a informação registrada em um catálogo, um índice ou bases de dados (LANCASTER, 2004). Para isso, segundo o autor, são utilizados catálogos on-lines, arquivos ou bases de dados bibliográficas, que armazenam informações sobre o conteúdo de livros, periódicos, artigos, entre outros objetos, e que esse conteúdo precisa ser localizado pelo usuário, no momento das buscas.

Ao analisar sistemas de informação utilizados no ambiente web, Saracevic (1996) verificou a necessidade da organização da informação, por meio da descrição dos itens, atividade ocorrida na catalogação descritiva, na indexação (análise de assunto e tradução) e na busca de assuntos durante da recuperação (necessidade de informação e estratégias de busca). Em linhas gerais, as descrições dos objetos informacionais contemplam os processos de produção, processamento, armazenamento e disponibilização das informações que forem relevantes para os usuários (SARACEVIC, 1996). De acordo com Araújo Júnior e Tarapanoff (2006), é importante destacar que a organização da informação, no ambiente digital, usa metadados para representar e recuperar documentos armazenados, normalmente, em bases de dados. Esses documentos são representados por meio do tratamento descritivo (características extrínsecas) e temático (características intrínsecas ou de conteúdo), de modo a tornar esses recursos passíveis de recuperação, em face de um processo de busca pelo usuário.

Lancaster (2004) entende que a recuperação é caracterizada como o resultado do tratamento dos documentos, gerando os conceitos, capazes de descrever a informação contida nos documentos e que será recuperada pelos usuários de um SRI. De fato, a construção de estratégias de busca em um SRI só é possível graças aos mecanismos de representação e especificação, contidos nas listas de termos (citações dos assuntos), e das necessidades de informação dos usuários, expressas, no sistema, por meio dos mecanismos de busca ou motores de busca que possibilitam a recuperação em páginas da web (LANCASTER, 2004).

Singhal (2009) verificou a importância da necessidade do armazenamento e da recuperação de informações textuais após a invenção dos computadores, o que permitiu o processamento de grandes quantidades de informação para satisfazer uma necessidade individual ou coletiva. Segundo o autor, para satisfazer diferentes necessidades de informação, é necessário algum tipo de estratégia de busca (*query*) a ser formulada na interface das bases de dados.

Durante o processo de recuperação, deseja-se saber qual é a informação que está contida em um objeto informacional, e essa informação pode estar representada em um conjunto de documentos (SINGHAL, 2009). Esses documentos, geralmente, são textos que precisam estar estruturados, de modo semântico, possibilitando, assim, satisfazer uma necessidade. Tal necessidade pode ser expressa mediante os mecanismos disponibilizados nas bases de dados de coleções de documentos, com o propósito de garantir a recuperação, permitindo aos usuários identificar, navegar ou realizar a pesquisa, a partir de diferentes atributos (título, assunto, dentre outros) referentes aos documentos (SINGHAL, 2009).

Segundo Baeza-Yates e Ribeiro-Neto (1999), um dos principais objetivos da recuperação é conseguir informar ao usuário quais são os documentos relevantes e quais os irrelevantes, evitando a recuperação de documentos inúteis. Os autores consideram que esse processo garante a interoperabilidade semântica, a qual objetiva mostrar quais informações atendem a uma necessidade de informação, a partir de um conjunto de documentos (*corpus*) existente em um SRI.

Nesse contexto, o resultado obtido da recuperação deve apresentar desde os assuntos mais relevantes até os menos relevantes por meio de um ranque (*ranking*). A relevância representa algo subjetivo, que depende da necessidade de informação do usuário, expressa em linguagem natural e traduzida mediante o uso de termos de indexação (vocabulário controlado). No momento da iteração com o SRI, o usuário expressa sua necessidade, mediante o uso de uma expressão (*query*) de busca que indique os assuntos dos documentos inseridos na base do sistema. Salienta-se que o processamento da consulta e a ordenação dos documentos são realizados de acordo com o grau de relevância (BAEZA-YATES; RIBEIRO-NETO, 1999).

Durante a busca, os usuários podem formular uma expressão em linguagem natural, controlada ou por termos. Outra estratégia de busca é por navegação (*browsing*), cujo usuário navega por páginas selecionando links que sejam pertinentes à necessidade. Nesse sentido, quanto maior a especificidade durante a busca, maior será a precisão e qualidade dos itens recuperados, diminuindo a revocação, sendo que este último está relacionado com a quantidade de itens recuperados (exaustividade). Por fim, o resultado da pesquisa é apresentado para o usuário, permitindo uma retroalimentação, a partir da resposta ou conclusão do usuário no SRI, ou seja, o usuário de um sistema pode dar um *feedback* acerca dos documentos recuperados, informando quais foram mais úteis (SOUZA, 2006).

Os documentos requerem diferentes tipos de representação da informação, visando à recuperação. A partir disso, a informação contida nos documentos é desvinculada de sua forma registrada, tornando-se, portanto, organizada e representada (SÁNCHEZ; PANIAGUA, 2009). Nas bases dados, a representação pode ser por meio de arquivos sequenciais ou invertidos, o que possibilita ampliar os modos de relacionamento, por

consequente, elimina-se o modo linear de tratar o documento, nos ambientes digitais (SÁNCHEZ; PANIAGUA, 2009).

A ordem de saída dos documentos depende da ordem de citação de cada assunto de um documento, sendo caracterizada por meio de uma representação descritiva. Para isso, realiza-se a extração de termos representativos, retirados do documento, durante o processo de catalogação dos assuntos (BAEZA-YATES; RIBEIRO-NETO, 1999). Entretanto, os termos podem ser atribuídos pela representação temática, considerando o assunto tratado no documento, processo esse denominado de análise de assunto. Essa análise permite formar uma expressão linguística, descrita por meio de um conjunto de termos (descritores), retirados de uma linguagem de indexação ou vocabulário controlado (BAEZA-YATES; RIBEIRO-NETO, 1999).

Durante a busca, são realizadas consultas aos termos que são expressas em linguagem natural ou artificial (vocabulário controlado) e foram utilizadas pelo indexador, no momento da descrição (SÁNCHEZ; PANIAGUA, 2009). Para esses autores, devem-se recuperar apenas os documentos que contêm, exatamente, as palavras que o usuário informou durante a consulta. Por esse motivo, os documentos devem estar indexados de maneira consistente, visando à qualidade do conteúdo intrínseco e extrínseco, de modo que sejam recuperados com mais precisão, eliminando, sobretudo, problemas de ambiguidade da linguagem natural. Assim, o processo de recuperação possibilita a localização de documentos físicos e digitais, tais como: textos, imagens, sons, vídeos, páginas web, entre outros objetos obtidos após a recuperação, que pode ser realizada tanto nos ambientes convencionais quanto nos digitais (SÁNCHEZ; PANIAGUA, 2009).

Saracevic (1996) evidencia a importância da representação dos assuntos dos documentos em bases de dados, tornando-se necessária a elaboração de catálogos que representem esses objetos informacionais que serão pesquisados, no momento da busca. Por esse motivo, para a recuperação, utilizam-se os metadados, que representam os termos, descrevendo as características dos objetos informacionais, os quais são padronizados de modo a facilitar o acesso e o uso da informação pelos usuários (SARACEVIC, 1996). Nesse sentido, os usuários do SRI estão interessados em recuperar a informação que está contida nos assuntos dos documentos, traduzidos em expressões de busca. No âmbito da Ciência da Computação, são construídos sistemas de informação para recuperar os dados em banco de dados, o que representa a diferença entre a recuperação da informação em um SRI e em um Sistema de Gerenciamento de Banco de Dados (SGBD) (BAEZA-YATES; RIBEIRO-NETO, 1999).

Os SRI tradicionais representam o conteúdo de um documento (visão lógica), ou seja, por meio de um texto completo ou índice de termos, torna-se possível representar conteúdos para recuperar. Nesse sentido, para sanar o problema causado pela falta de

informação dos usuários, os documentos devem ser representados de acordo com as necessidades desses sujeitos, cujos elementos da representação (índices, catálogos, estratégias de busca, dentre outros) são armazenados por computadores em bases de dados (BAEZA-YATES; RIBEIRO-NETO, 1999). De acordo com os autores, os SRI estabelecem uma interface entre a coleção de recursos de informação em meio impresso ou digital e os usuários desses recursos. Os sistemas possibilitam gerenciar as atividades realizadas pelos profissionais e usuários, como representação, armazenamento, organização e acesso aos itens informacionais. Desse modo, para a representação adequada de documentos, é necessário criar sistemas de classificação ou indexação eficazes, que garantam resultados mais relevantes no momento da recuperação (BAEZA-YATES; RIBEIRO-NETO, 1999).

Outras aplicações significativas na recuperação da informação dizem respeito ao uso dos SOC que deram origem aos sistemas pós-coordenados, como os sistemas de classificação facetados (TRISTÃO *et al.*, 2004). Com o uso desses sistemas de organização, estimulou-se a construção de tesauros, que foram desenvolvidos baseados nas relações semânticas entre conceitos, fornecidos no momento da recuperação pelo usuário (TRISTÃO *et al.*, 2004).

Diante do exposto, para uma recuperação mais eficiente com o uso dos SOC, faz-se necessário estabelecer a interoperabilidade entre os dados e informações que serão recuperadas (WEISS, 2019). Portanto, é importante discorrer sobre a compatibilidade de linguagem entre os SOC, assunto refletido no próximo subcapítulo.

3.4 Compatibilidade e Convertibilidade de Linguagem

Desde 1960, a preocupação com os problemas relacionados à compatibilidade e à convertibilidade entre linguagens de indexação (vocabulário controlado), no contexto das bibliotecas e centros de informação, tem sido relatada na literatura científica. Alguns dos principais teóricos a realizarem estudos sobre essa temática foram: Newman (1965), Wall e Barnes (1969), Neville (1970), Soergel (1982), Dahlberg (1981, 1983, 1996), Lancaster (1986), Svenonius (1983), Maniez (1997), entre outros.

Os estudos sobre essa temática tiveram um alto índice na década de 1970 e 1990, conforme Maniez (1997), que observou a necessidade do intercâmbio de dados indexados em diversas bases de dados, e que essas possuem diferentes estruturas e vocabulários entre os diversos sistemas de informação em ambiente web. Na década de 1980, o acesso às bases de dados era restrito. Nas primeiras décadas do século XXI, a compatibilidade ainda é altamente necessária pelo fato do crescente número de bases de dados, que indexam recursos informacionais com diferentes terminologias, além de buscas que são realizadas simultaneamente em vários repositórios digitais com a finalidade de se obter uma

compatibilidade universal (ZHANG, 2006). Na literatura, o termo “compatibilidade”, segundo essa autora, pode ser relacionado a outros termos no sentido de conversão como: integração, harmonização, reconciliação e concordância (DAHLBERG, 1981, 1983, 1996; MANIEZ, 1997; SVENONIUS, 2000).

A compatibilidade, na visão de Newman (1965), é aplicada para cada termo de uma linguagem (vocabulário), que possui um termo correspondente com significado semelhante em outra linguagem, por esta razão, o termo pode ser convertido de uma linguagem para outra, sem alteração no significado. Para Maniez (1997), não é necessário converter a estratégia de busca, pois o usuário normalmente não muda as estratégias de busca para cada base de dados ou repositórios de um SRI, por esta razão, durante o processo de busca é comum ocorrer problemas de incompatibilidade (MANIEZ, 1997).

A compatibilidade de linguagem entre diferentes tipos de SOC é realizada para que dois ou mais sistemas possam se comunicar ou trocar informação, sem que haja qualquer problema de incompatibilidade (DOERR, 2001). Na visão de Soergel (1982), existem situações em que a compatibilização entre linguagens de indexação não é suficiente para garantir a eficiência no âmbito da recuperação, sem que haja qualquer perda de informação.

Nesse sentido, durante o processo de compatibilidade de vocabulário controlado, a interoperabilidade semântica entre diferentes tipos de SOC possui grande relevância no intercâmbio entre os vocabulários, com o fim de garantir a semântica entre os vocabulários (CHAN; ZENG, 2002, 2006). Na literatura da área da Ciência da Informação, é evidenciado que a compatibilidade de linguagem é realizada por meio de relacionamentos semânticos entre os conceitos de diferentes SOC, com vistas à recuperação, possibilitando a criação de sistemas de conceitos de um determinado domínio do conhecimento (CAMPOS; BARBOSA, 2020).

Svenonius (1983, p. 2, tradução nossa) define convertibilidade como sendo um termo aplicado a vocabulários controlados, ou seja, "a capacidade de uma linguagem de indexação ser convertida em outra", ou ainda, o mapeamento entre termos que representam conceitos que devem ser selecionados de maneira adequada para a formulação de uma estratégia de busca que, por sua vez, depende da compatibilidade. O problema de convertibilidade de linguagem ocorre quando a estratégia de busca ou a terminologia são diferentes entre duas ou mais linguagens, levando em consideração a subjetividade da indexação dos termos sobre um recurso (SVENONIUS, 1983).

Nesse sentido, Lancaster (1986) esclarece que a compatibilidade para ocorrer, requer que as linguagens de indexação sejam compatíveis, ou seja, utilizem os mesmos conjuntos de termos de forma que estejam integrados em bases de dados de modo cooperativo na indexação, contendo o mesmo vocabulário e a mesma semântica em todas as bases de dados, sem prejuízos durante a recuperação. Isso possibilitará o alcance de

resultados aproximados com o uso de descritores equivalentes, garantindo, assim, troca de informações em diferentes repositórios (LANCASTER, 1986). Dahlberg (1981, 1983, 1996) caracteriza a compatibilidade com a necessidade de intercâmbio das informações que, na maioria das vezes, acontece de forma simultânea, e precisa ser disseminada e compartilhada entre diversos SRI. Por essa razão, os termos que denotam os conceitos ou a combinação deles devem ser representados em qualquer linguagem de informação e vice-versa, bem como sua tradução para qualquer outro idioma, visando à recuperação (DAHLBERG, 1981, 1983, 1996).

Entretanto, pode ocorrer várias traduções quando a busca é por diferentes idiomas, de acordo com a pesquisa do usuário, o que torna a pesquisa morosa e com alto custo, e isso acontece mesmo com o uso de dois vocabulários no mesmo idioma. Além disso, mesmo em bases de dados diferentes, a linguagem deve ser compreendida de modo consistente, tendo em vista a interoperabilidade entre diferentes tipos de sistemas ordenados, atualmente denominados de SOC (DAHLBERG, 1981, 1983, 1996). Segundo Zeng e Chan (2004), a compatibilidade oportuniza fazer a busca cruzada em diversas bases de dados em dois ou mais sistemas, utilizando, por exemplo, um catálogo on-line com diversos vocabulários que foram indexados de modo separado, sendo possível a troca de dados e informações sem qualquer tipo de intervenção humana.

Maniez (1997) sinaliza que é necessário levar em consideração a estrutura das linguagens de indexação, que são determinadas principalmente pelas relações hierárquicas e, sobretudo, pelas regras utilizadas durante a indexação dos conceitos, atentando, também, para a subjetividade na escolha dos descritores, pois isso depende do conhecimento individual, das experiências, questões culturais, regionais, entre outras particularidades dos classificadores. Entretanto, o autor sinaliza que cada base de dados possui sua própria linguagem de indexação. Uma vez que não existe padronização ou compatibilidade entre esses vocabulários, e que são específicos de um determinado assunto, o mesmo conceito pode ser designado por diferentes termos em diferentes linguagens (MANIEZ, 1997).

Segundo Campos e Barbosa (2020), o primeiro passo a ser feito durante o processo de compatibilidade de linguagem entre SOC é o mapeamento entre os termos para verificar as possibilidades de serem compatíveis no nível conceitual. O mapeamento é a realização da correspondência (*matching*) entre os termos que serão compatibilizados em cada SOC, possibilitando a conexão dos termos que representam cada conceito, podendo estar em dois ou mais SOC, mesmo que os termos não sejam equivalentes. Nesse sentido, é necessário que os vocabulários entre os SOC sejam interoperáveis, ou seja, uma expressão de busca formulada pelo usuário deve ser convertida para outra expressão que tenha a mesma correspondência (*matching*) entre um ou mais vocabulários em cada SOC, sem levar em conta o idioma (CAMPOS, BARBOSA, 2020; ISO 25964, 2013). Além disso, a partir das

regras estabelecidas durante o processo de compatibilidade, possibilita a construção de algoritmos a partir da interoperabilidade semântica. A automação é possível somente a partir do mapeamento dos vocabulários dos SOC a serem compatibilizados, o que possibilita o intercâmbio de informações em bases de dados heterogêneas e diferentes domínios do conhecimento e idiomas (CHAN; ZENG, 2002, 2006).

Nesse contexto, Dahlberg (1983) aponta que, ao final da busca nas bases de dados, são exibidos os resultados e o tipo escolhido pelos usuários termo exato, sinônimo ou termo relacionado (quase sinônimo), termo geral para o específico ou vice-versa, sinônimo parcial, sem correspondência. Segundo Dahlberg (1981), estes tipos correspondem aos três padrões de compatibilidade conceitual entre elementos dos SOC, a saber: coincidência conceitual, correspondência conceitual e correlação conceitual. O primeiro elemento ocorre entre dois termos que se combinam em todas as características, ou seja, são equivalentes; a correspondência conceitual manifesta-se quando dois termos combinam grande parte de suas características, ou seja, são semelhantes. Por fim, a correlação conceitual acontece quando dois termos não são semelhantes, mas são relacionados em diferentes níveis estruturais ou em linguagens diferentes (DAHLBERG, 1981).

A partir desse princípio, a norma ISO 25964 é responsável por questões relacionadas à compatibilidade entre tesouros com outros vocabulários controlados, tais como sistemas de classificação, taxonomias, ontologias, entre outros. A norma descreve as principais recomendações para que seja realizada a definição e a manutenção de mapeamentos entre diversos tipos de SOC (ISO 25964, 2013).

Quando se trata de uma linguagem que vai ser compatibilizada com outra, a norma ISO 25964 estabelece alguns modelos, que podem ser seguidos, de modo a facilitar o mapeamento entre as linguagens, que por sua vez, não possuem somente relações de equivalência, podendo possuir também relações hierárquicas e associativas entre os termos, com o mesmo idioma e notação ou não. Os modelos de mapeamento mais utilizados são o modelo de ligação direta e o modelo de estrutura central. O modelo de ligação direta ocorre quando duas ou mais linguagens não possuem a mesma estrutura e idioma, tais como de taxonomias, sistemas de classificação bibliográfica, ontologias, entre outros. Dessa forma, os termos de um determinado SOC são mapeados de acordo com os termos de outros, podendo ser utilizada a base de qualquer um deles para recuperar a informação de um recurso indexado. Desse modo, os usuários, por meio de estratégias de busca em um SRI, podem recuperar informações sobre qualquer assunto indexado em qualquer uma das bases de dados envolvidas neste sistema de recuperação (ISO 25964, 2013).

No que tange ao modelo de estrutura central, ele ocorre quando é escolhido um vocabulário principal, cujos termos desse vocabulário serão mapeados para os outros termos, e a mesma coisa acontece ao contrário, ou seja, cada termo dos vocabulários secundários

deve ser mapeado com o vocabulário principal. Assim, a recuperação de qualquer recurso indexado nos vocabulários secundários poderá ocorrer também pelo vocabulário central, convertendo os termos indexados nesse vocabulário com os termos de um ou mais vocabulários secundários. A conversão também pode ocorrer de acordo com as expressões de busca nas bases dos vocabulários secundários; caso um vocabulário não possua todos os termos dos outros, o que pode acarretar em problemas de incompatibilidade, onde os resultados podem não ser satisfatórios (ISO 25964, 2013).

De acordo com Svenonius (1983), além dos tipos de relações semânticas conceituais já mencionadas, para se obter a compatibilidade entre linguagens é preciso realizar o mapeamento unidirecional ou bidirecional ou a tradução direta do vocabulário de uma linguagem de indexação para outra, como também pode ser construída uma linguagem de comutação ou intermediária (SVENONIUS, 1983). Em relação ao mapeamento, ele pode ocorrer de forma unidirecional, ou seja, entre duas linguagens de indexação, por exemplo, LA para LB ($LA \rightarrow LB$), onde todos os termos de LA devem ter termos correspondentes, ou que mais se aproximam de LB, não o contrário (RIESTHUIS, 1996). Caso seja necessário que todos os termos de LB estiverem em LA, é necessário que o mapeamento seja bidirecional, mapeando também os termos de LB em LA ($LB \rightarrow LA$), cujos termos da linguagem B terão termos correspondentes ou que mais se aproximam da linguagem A, ou seja, pode ser representado pela expressão booleana "OR" ($LA \text{ OR } LB$) (RIESTHUIS, 1996).

Em relação à linguagem de comutação ou intermediária, trata-se de uma linguagem que garante a compatibilidade semântica entre linguagens de indexação, independente da complexidade do vocabulário a ser compatibilizado. Sendo assim, é possível recuperar o conteúdo de um documento, por meio de termos de um vocabulário para outro, sem que haja perda significativa durante a recuperação. A linguagem intermediária é utilizada quando existem vários tipos de vocabulários envolvidos e/ou mesmo se forem apenas dois vocabulários envolvidos com idiomas diferentes. Desse forma, a linguagem intermediária é conectada por meio de uma única tabela de compatibilização conceitual, sendo que cada sistema de informação inclui notações e descritores aos documentos indexados com seus próprios vocabulários e com as notações da linguagem intermediária, haja vista possibilitar o intercâmbio ou incorporar informações de outras bases de dados (HORSNELL, 1975).

Para isso, deve-se verificar o nível de especificidade dos conceitos em cada linguagem, o nível de pré-coordenação que os descritores foram indexados e as relações entre os conceitos que pertencem a mais de um conceito (relações polihierárquicas) como, por exemplo, converter descritores de um tesouro para sistemas de classificação, considerando que os assuntos já estão inter-relacionados, ou seja, um conceito está subordinado a dois ou mais conceitos em mais de uma classe. Além disso, para o autor, existem problemas da linguagem natural (sinonímia, homônimos, polissemia e homógrafos),

das variações gramaticais e da inclusão de conceitos que existem em uma linguagem e não em outras. Desse modo, torna-se imprescindível elucidar as relações entre os conceitos durante o estabelecimento das relações de equivalência entre descritores das linguagens de indexação, como também precisa existir a compatibilidade considerando o uso da linguagem natural nos sistemas (SOERGEL, 1982).

Segundo Lancaster (1986), o mapeamento entre vocabulários exige um grande esforço intelectual e tempo. O autor apresenta um método de compatibilização de vocabulário chamado de “integração de vocabulários”, cujos termos equivalentes eram definidos durante as estratégias de busca dos usuários em ambiente on-line. Mesmo com alguns estudos para automatizar esse processo, nem sempre era possível converter, de modo automático, todos os vocabulários, mesmo utilizando termos equivalentes, que poderiam ser criados e aprimorados durante a expressão de busca em bases de dados em ambiente digital. O sistema que mais se destacou foi criado na década de 1970, chamado de *Vocabulary Switching System* (VSS), desenvolvido no *Batelle Columbus Laboratories*, sendo que o próprio usuário escolhia, no momento da busca, os tipos de correspondência (equivalência exata, sinônimos e termos relacionados ou adjacentes), o número de termos a serem exibidos, entre outras escolhas (LANCASTER, 1986).

A compatibilidade de linguagem como, por exemplo, converter o vocabulário da CDU para o de um tesouro, a comparação dos termos entre os vocabulários dos dois SOC estão em diversos níveis das estruturas hierárquicas, auxiliando na análise conceitual. Além disso, deve-se levar em consideração na compatibilidade de linguagem o fato de possuir inúmeras combinações de descritores para os termos que representam cada conceito.

Maniez (1997) destaca que existem problemas quanto à falta de projetos de compatibilidade com apoio de instituições nacionais e internacionais por questões de viabilidade, à dificuldade de consenso entre linguistas para tradução de vocabulários específicos de um domínio e idioma. Outro ponto a ser considerado é a escassez de desenvolvedores para automação do processo de compatibilização com custo baixo, baseado em poucas evidências de demonstração de eficácia e poucas justificativas pelo alto custo gerado em ambiente on-line (MANIEZ, 1997). Outros problemas que afetam a compatibilidade da linguagem referem-se à insuficiência de projetos e cientistas da informação qualificados para a realização do mapeamento das relações semânticas entre os termos em diversos vocabulários (MANIEZ, 1997).

Além do mapeamento entre as linguagens, Lancaster (1986) observou que também existem técnicas e algoritmos para o processamento pela máquina, embora essas técnicas sejam úteis apenas para uma parte do processo de compatibilidade de linguagem, sendo necessário estabelecer o mapeamento entre duas linguagens (unidirecional ou

bidirecional), com a utilização de linguagens intermediárias e tesouros (mais especificamente microtesouros, macrotesouros e o tesouro universal).

No contexto da compatibilização utilizando tesouros, manifestam-se os macrotesouros e os microtesouros (MANIEZ, 1997). Os macrotesouros possuem a representação de forma genérica dos termos de várias áreas do conhecimento, cujos termos se relacionam com outros tesouros ou vocabulários controlados de domínios específicos, com o intuito de unificar os termos utilizados em várias áreas do conhecimento (MANIEZ, 1997). Um exemplo de macrotesouros, de acordo com Lancaster (1986), é o *Broad system of Ordering* (BSO), com a finalidade de interconectar vários sistemas de informação, como também o *World Scientific Information Programme* (UNISIST) da UNESCO. O tesouro UNISIST adotou princípios da classificação facetada para o reconhecimento de relações semânticas entre os termos que representam os conceitos. Ademais, o tesouro da UNESCO foi desenvolvido para tratar assuntos referentes à informação científica e de tecnologia governamental, entre outros assuntos. Por essa razão, esse tesouro é conhecido como metatesouro, devido a sua equivalência com outros tesouros (LANCASTER, 1986).

Já os microtesouros são caracterizados por um conjunto de termos mais específicos e que têm relacionamento com tesouros de assuntos mais genéricos. Por fim, o tesouro universal, também conhecido como *Universal Source Thesaurus* (UST), serviu para a construção de outros tesouros com propósitos específicos, tendo em vista aumentar o grau de compatibilidade e especificidade para facilitar a construção de novos tesouros, de modo que a serem incorporados em um tesouro universal, que poderá ter mapeamento unidirecional ou bidirecional (SOERGEL, 1972).

Apesar de existirem vários métodos de compatibilização, Campos (2005) considera os métodos mais utilizados são: o método de reconciliação de tesouros Neville (1970), que será explicado a seguir, e sua aplicação no (Anexo E). Outro método de compatibilidade de linguagem, é a matriz de compatibilidade conceitual proposta por Dahlberg (1981), que será utilizada de forma adaptada nesta pesquisa.

O método de Neville, parte do princípio da indexação de conceitos com os descritores que denotam esses conceitos, como acontece nos tesouros, cujos termos são indexados e os descritores são usados apenas como rótulos ou etiquetas escolhidas para representar esses conceitos por meio de relações semânticas entre eles (NEVILLE, 1970). Nesse sentido, é possível utilizar diferentes descritores independente do idioma, tendo em vista representar o conceito, baseado na construção de uma linguagem intermediária, que exerce a função de um tesouro multilíngue (NEVILLE, 1970).

Esse método é baseado na construção de um sistema de conceitos com notações, sendo possível estabelecer a equivalência entre os conceitos utilizando descritores de diversas linguagens de indexação denominado método de reconciliação de tesouros

(NEVILLE, 1970). A notação permite recuperar itens de diversas bases de dados de forma automática, mesmo com diferentes idiomas como, por exemplo: a notação de um sistema de classificação quando compatibilizada com os termos de um tesauro deixa os dois SOC estruturados para serem indexados em diferentes bases de dados (NEVILLE, 1970).

É importante ressaltar que a estrutura original do tesauro não pode ser modificada, ou seja, nenhum descritor original é modificado ou excluído, podendo ser inseridos novos descritores em cada tesauro. Assim, um tesauro de um domínio pode indexar os mesmos conceitos utilizando descritores diferentes (NEVILLE, 1970). O (Anexo E) apresenta os níveis de reconciliação (compatibilidade) de tesauros, considerando o nível terminológico e conceitual representado pelo item tipo de compatibilidade, descrição de cada tipo com exemplos. Além disso, é importante ressaltar que podem ocorrer incompatibilidades ou reconciliação entre descritores do mesmo domínio, de acordo com o tipo de compatibilidade entre os vocabulários dos tesauros envolvidos. As incompatibilidades podem ocorrer também pois, o autor não considerou o tratamento de homonímia como por exemplo, “tanques” (veículos de combate)” e “tanques” (containers)”, apesar de ser menos comum quando de tesauros do mesmo domínio, sem utilizar a garantia literária somente a garantia conceitual dos tesauros envolvidos.

O método de Neville (1970) esclarece vários níveis de compatibilidade, porém o método de Dahlberg consiste em uma matriz de compatibilidade conceitual que se baseia no método analítico-sintético para estabelecer a compatibilidade entre diferentes linguagens de indexação como, por exemplo, em sistemas de classificação e tesauros por meio do método analítico-sintético (DAHLBERG, 1981).

A compatibilidade de duas ou mais linguagens, segundo Dahlberg (1981), pode ser realizada por meio da comparação verbal entre os vocabulários de diversos sistemas de informação que possuem o mesmo idioma, resultando em um grande número de termos com coincidência verbal ou equivalentes. Caso os sistemas de informação a serem compatibilizados estiverem em idiomas diferentes, os vocabulários dos sistemas devem ser traduzidos para a linguagem do SOC que foi escolhido como padrão de comparação (DAHLBERG, 1981).

O QUADRO 13 apresenta a matriz de compatibilidade conceitual proposta por Dahlberg (1981), em que as duas primeiras colunas representam o código dos conceitos e a linguagem do SOC principal apresentada de forma hierárquica. Logo após, as próximas três colunas indicam o nível de equivalência entre termos dos outros sistemas, com utilização de alguns símbolos, tais como o símbolo o ‘<’ para representar conceitos mais genéricos em relação à linguagem do sistema mestre (principal); o ‘>’ para conceitos mais específicos; e a coluna ‘cc’ representa o número de coincidências entre os conceitos das diferentes linguagens que serão compatibilizadas, possibilitando o cálculo do grau de compatibilidade conceitual.

Devem ser inseridos o número de coincidências conceituais na coluna 'cc', somente quando for a partir de dois ou mais o número total de coincidências entre os SOC (DAHLBERG, 1981).

QUADRO 13 - Matriz de compatibilidade conceitual de Dahlberg

Nº	Nome do Conceito	CDD	Classificação de BLISS	Tesouro universal (UNT) (Microtesouro)	Número de Coincidências conceituais
65	Bem-estar social	361 Problemas Sociais e Bem estar - Social	Q Bem - estar Social	R85/99 Bem - estar Social	3
65.1	Filosofia do bem-estar social	361.01 Filosofia e Teoria do bem-estar social	QAE Administração social e administração do bem-estar social.	R86 Filosofia do bem-estar social	3
65.2	Administração do bem-estar social	—	QAG Administração social, Administração do bem-estar social	R90 Bem-estar social	2
65.3	Política de bem-estar	>361.25 Ação dentro da estrutura social estabelecida (política)	QAG Política (Bem estar-social)	R87 Política Bem-estar	3
65.4	Planejamento do bem-estar social	>361.25 Ação dentro da estrutura social estabelecida (planejamento)	QAG P Planejamento para bem-estar, planejamento do bem-estar social	R88 Planejamento do bem-estar social	3
65.5	Economia do bem-estar social	—	QAT M/Z Gerenciamento do bem-estar social	R89 Economia do bem-estar social	2

FONTE: Elaborado pelo autor (2021). Adaptado de Dahlberg (1981, p. 89, tradução nossa).

Ainda com relação ao Quadro 13, nas situações de não correspondência e falta de hierarquias entre os termos que indicam os conceitos, as decisões devem ser tomadas no estabelecimento da matriz de compatibilidade de acordo com os métodos de classificação utilizando vários tipos de SOC, por meio do uso de sinais mencionados acima. Este quadro mostra a utilização destes sinais ou a falta de compatibilidade a partir da seleção de um SOC mestre (principal) como referência do ponto de partida, com os demais sistemas envolvidos, representados no Quadro 13 pela Classificação Decimal de Dewey (CDD), Classificação de Bliss (BC2) e o Tesouro Universal *Source Thesaurus* (UNT). Desse modo, Dahlberg (1981) indica como vocabulário principal o mais detalhado ou podem-se usar, também, todos os termos dos outros SOC envolvidos na comparação e na criação fora das hierarquias que estão sendo criadas, ou seja, uma hierarquia principal para a comparação dos conceitos relevantes em cada um de seus níveis.

Segundo Campos (2005), a matriz de compatibilidade é construída a partir do mapeamento semântico e estrutural das linguagens de indexação a serem compatibilizadas,

diferente do método de Neville. A autora menciona que, apesar dos níveis de compatibilidade apontados no método de Neville, ainda é necessário um conhecimento mais profundo do potencial das linguagens de indexação, pois tal método não expõe uma metodologia de análise em relação às linguagens de indexação. Por este motivo, optou-se por utilizar o método de Dahlberg (1981), que além de ser menos complexo, ele é um método de compatibilidade de linguagem que utiliza diversos tipos de SOC, sendo mais adequado à proposta para a construção do modelo proposto nesta tese que utiliza a compatibilidade entre a CDU e o tesauro AgroTermos.

Diante do exposto, infere-se que a compatibilização de vocabulários ainda é um assunto que precisa ser mais pesquisado em diversas áreas do conhecimento, tanto no âmbito da Ciência da Informação quanto na Ciência da Computação, sobretudo, em relação à compatibilização em diversos tipos de SOC. Além disso, a compatibilidade entre SOC permitirá que um ou mais sistemas troquem informações sem nenhum esforço, independente do contexto em que foram inseridos, mantendo a troca de dados por meio de relações estabelecidas entre outros SOC, que exigem padrões e formatos no âmbito da web semântica (ZENG; CHAN, 2004). Nesta tese, o foco é a CDU compatibilizada com o tesauro AgroTermos, tendo em vista garantir a representação semântica de conteúdos das duas estruturas, que são heterogêneas e serão analisadas no mesmo idioma.

No próximo subcapítulo, será detalhado a linguagem Simple Knowledge Organization System (SKOS) que permite a interoperabilidade entre SOC no âmbito da web semântica.

3.5 Simple Knowledge Organization System

O Simple Knowledge Organization System (SKOS) é uma linguagem de publicação de diversos tipos de SOC de forma padronizada, tais como listas de cabeçalhos de assuntos, taxonomias, sistemas de classificação, tesouros, entre outros tipos de vocabulário controlado, publicados no formato RDF no âmbito da web semântica (MARCONDES, 2021). O SKOS começou seu desenvolvimento em 2009 pelos membros de uma comunidade internacional conhecida como *World Wide Web Consortium (W3C)*, com o propósito de possibilitar que os SOC sejam interpretados por máquinas, tornando-se interoperáveis. A maior parte do conteúdo de documentos disponibilizados nos ambientes digitais ou na web é projetada para a leitura por pessoas e não por computadores, visto que as máquinas não conseguem manipular esse conteúdo de forma significativa (SOUZA; ALVARENGA, 2004).

Em geral, os computadores podem analisar adequadamente as páginas da web quanto ao layout e rotinas de processamento (hipertextos, transferência de arquivos, e-mails,

acesso remoto e links). A linguagem SKOS facilita essa análise, ao resolver problemas de interoperabilidade entre sistemas, considerando a semântica das estruturas textuais, pelo fato de que os computadores não possuem uma maneira confiável de processar o sentido das estruturas, como na frase: “Esta é a cidade de Belo Horizonte, onde está localizada a UFMG” (MARCONDES, 2021).

De acordo com Berners-lee, Hendler e Lassila (2001), a web semântica, também conhecida como web dos dados, é responsável pela estrutura do conteúdo significativo das páginas web, ao criar um ambiente capaz de desenvolver instrumentos computacionais que possuem sentido lógico e semântico para os dados que estão registrados na web. No contexto da W3C, a web semântica conecta os dados presentes em recursos na web ao invés de conectar uma página da web a outra (BERNERS-LEE; HENDLER; LASSILA, 2001).

Pastor-Sánchez, Martínez Méndez e Rodríguez-Muñoz (2009) observaram que os padrões da web semântica surgiram pela ausência de modelos adequados para a representação de um domínio em diferentes níveis de abstração para um formato legível por máquina. Estes níveis de abstração são por meio do processamento de linguagem natural (PLN), fazendo com que a máquina entenda o significado de um texto ou seja, traduzir um conjunto de palavras em linguagem natural para a linguagem de máquina. Segundo Isaac e Summers (2009) para atribuir a semântica a um conjunto de palavras em diferentes vocabulários, por meio do SKOS, é necessário fazer a análise léxica, sintática, semântica. Desse modo, é necessário que a máquina entenda o conjunto de palavras em um determinado idioma (análise léxica); quais são as regras para que estas palavras se liguem corretamente em uma frase (análise sintática) e por fim, verificar se estas palavras estão estruturadas corretamente, estabelecendo um sentido em dado contexto (análise semântica), de modo a garantir a interoperabilidade semântica em diferentes vocabulários (ISAAC; SUMMERS, 2009).

Desse modo, o SKOS utiliza tecnologias de linguagens de marcação (rótulos) para estruturar desde o nível léxico ao semântico os dados que estão disponibilizados na web, utilizando linguagens tais como a linguagem *Extensible Markup Language*(XML), a *Hyper Text Markup Language* (HTML), entre outras. Além disso, também são utilizadas linguagens de programação para o desenvolvimento da web (Java, PHP) e os sistemas de gerenciamento de banco de dados, como o MySQL e o Oracle (PASTOR-SÁNCHEZ; MARTÍNEZ MÉNDEZ; RODRÍGUEZ-MUÑOZ, 2009).

Para representar a informação de um documento em termos semânticos e os significados contextuais, recorre-se ao uso de notações ou rótulos, também conhecida como *tags*, as quais descrevem o conteúdo de um documento, utilizando o XML (SOUZA; ALVARENGA, 2004). A linguagem HTML é uma linguagem de marcação de hipertexto usada para descrever o layout e as ações, por exemplo, menus, links para outras páginas, entre

outras funcionalidades, que são interpretadas por navegadores (*browsers*) entre computadores/smartphones e servidores de aplicação web, para prover a comunicação, por meio do *HyperText Transfer Protocol* (HTTP); de modo diferente, da linguagem XML. Percebe-se que o SKOS constitui um modelo de dados, representado de maneira padronizada, e faz uso dos SOC para a organização e representação da estrutura básica do conteúdo informacional de um documento. A ideia geral inerente a esses padrões é a disponibilidade e a troca de vocabulários controlados na web, podendo contribuir para a descoberta de novos recursos informacionais a partir da padronização (SOUZA; ALVARENGA, 2004).

Os recursos informacionais disponíveis em ambiente digital ou na web são referenciados de forma unívoca, por meio do *Uniform Resource Identifier* (URI), que interliga um conjunto de dados relacionados (*Linked Data*), conhecido como dados ligados entre si, cujo objetivo é aproximar os dados e torná-los úteis (PASTOR-SÁNCHEZ; MARTÍNEZ MÉNDEZ; RODRÍGUEZ-MUÑOZ, 2009). O *Linked Data* foi disseminado em 2009, a partir da publicação de autoridade de nomes (cabeçalhos de assuntos) pela Biblioteca do Congresso, sendo esse o marco inicial do uso do *Linked Data* no SKOS. As relações semânticas do *Linked Data* na linguagem SKOS serviram para guiar a navegação, por exemplo, com a relação *skos:related* sendo usada para diferentes remissivas (ver também), o usuário é direcionado de forma consistente por meio de links no ambiente digital. A Biblioteca do Congresso liderou o desenvolvimento de vocabulários controlados no formato *Metadata Authority Description Schema in RDF* (MADS/RDF), que são esquemas de descrição de autoridade de nomes e assuntos, por meio de metadados em *Resource Description Framework* (RDF) formalmente mapeado para SKOS. Assim, o RDF fornece um modelo de dados com semântica com sintaxe padronizada escrita em XML para descrever documentos na web (SUMMERS *et al.*, 2008).

O MADS/RDF é baseado no formato de autoridade de catalogação legível por máquina, o *Machine Readable Cataloging* (MARC21), formato que identifica os elementos de autoridade que contém os termos (metadados) em relação as classes dos recursos informacionais e do registro catalográfico de forma clara, para que seja possível a leitura e interpretação dos dados em meio digital. Para o meio digital, a descrição dos elementos do registro é realizada por meio do *Resource Description Access* (RDA), atualização do padrão *Anglo American Cataloguing Rule* (AACR2), com uso do modelo conceitual *Functional Requirements for Bibliographic Records* (FRBR) e *Functional Requirements for Authority Data* (FRAD) (SUMMERS *et al.*, 2008; FUSCO, 2010). Dessa forma, é possível a representação descritiva de diversos recursos informacionais, tanto em meio físico quanto digital. Para inserir manualmente as informações contidas em uma ficha catalográfica no computador, será necessário um processo moroso com custo elevado, o que despertou a necessidade de

realizar essa tarefa de modo automatizado, de modo a interpretar com agilidade as informações encontradas nos registros bibliográficos (SUMMERS *et al.*, 2008; FUSCO, 2010).

No contexto do SKOS, apresenta-se, como exemplo, um assunto complexo “*madsrdf: ComplexSubject* (“Política - Brasil –século XX)” a partir de um registro de autoridade no padrão MARC como um tipo complexo “(*madsrdf: ComplexType*)” ou pela concatenação de assuntos simples “(*madsrdf: SimpleType*)”, por exemplo, “(A Política no Brasil no século XX)” no SKOS. O padrão MADS/RDF mostra uma maneira de estender a linguagem SKOS de modo coordenado aos cabeçalhos de assuntos (SUMMERS *et al.*, 2008; ISAAC *et al.*, 2011). Assim, um conceito e as suas partes constituintes podem ser representados como uma expressão coordenada e com os elementos simples pertencentes a um conjunto de tipos definidos, tais como o tópico de assunto, localização, gênero, relação temporal, entre outros. Portanto, o MADS/RDF é destinado a manter os dados (metadados) de arquivos de autoridade no formato MARC sobre a descrição dos objetos informacionais, utilizando o *Linked Data*. O SKOS de modo geral, é considerado um modelo de dados que identifica os conceitos, assim como o *Linked Data* com o objetivo definir vocabulários que expressam conceitos a partir de termos de representá-los independente do idioma (SUMMERS *et al.*, 2008; ISAAC *et al.*, 2011).

Nesse sentido, o *Linked Data*, no contexto da web semântica, refere-se a um conjunto de boas práticas recomendadas pela W3C para a publicação e o compartilhamento de dados estruturados e que são reconhecidos pelo padrão de endereços, conhecido como *Uniform Resource Identifier* (URL) com informações sobre sistemas disponibilizados na web. O endereço de rede do servidor do aplicativo e sua base de dados, são acessados pelo protocolo TCP/IP, que é responsável pelo transporte de informações pela Internet, onde fica armazenado o sistema e sua base de dados para transferência de informações. Esses princípios foram estabelecidos por Tim Berners-Lee, no ano de 2006, ao fazer uso de inferências, integrações e resolução de ambiguidades da linguagem natural, com o objetivo de fornecer o contexto semântico a qualquer recurso da web para pessoas e máquinas, permitindo a interoperabilidade entre sistemas (BERNERS-LEE *et al.*, 2006).

Contudo, o LOD é integrado com o SKOS para representar conceitos entre diversos conjuntos de dados (*datasets*), visando à interoperabilidade. Assim, as informações disponíveis na web são publicadas com licença para uso de forma aberta e com formatos que são comuns a qualquer usuário, tais como *Portable Document Format* (PDF), *Extensible Style Language* (XLS), *Comma-separated values* (CSV), RDF, dentre outras. Dessa forma, é possível que ocorra o intercâmbio dessas informações por qualquer pessoa, visando a interoperabilidade e reuso de vocabulários entre diferentes sistemas de informação (JAIN *et al.*, 2010).

Baker *et al.* (2013) verificaram que, no SKOS, os conceitos são representados por triplas (recurso, propriedade e valor) no padrão RDF, conforme recomendações da W3C. Os elementos da tripla são representados por um único URI e um prefixo, cujo recurso indica qual será o conceito a ser representado e a propriedade representará o relacionamento semântico entre o recurso e o valor. Por fim, o valor é a instância do objeto informacional, representado por um literal relacionado com o recurso (BAKER *et al.*, 2013). Os autores relataram que no SKOS pode ser observado no tesouro multilíngüe da área de agricultura e áreas relacionadas (AGROVOC), com o objetivo de formalizar ou estruturar, semanticamente, os conceitos e termos de um tesouro, haja vista permitir que sejam utilizados em ambiente web.

De acordo com Rizzo e Troncy (2011), a ideia geral desse compartilhamento de informações e da linguagem SKOS usados na web semântica é a disponibilidade e a troca de vários tipos de SOC em um ambiente digital. Desse modo, é possível contribuir para a descoberta de recursos informacionais por meio de inferências, resolvendo ambiguidades com o objetivo de fornecer um contexto semântico para o processamento de texto com vocabulários expressos de forma padronizada. Nesse sentido, a utilização de dados abertos vinculados, como o LOD, torna disponíveis as informações sobre o conteúdo de documentos classificados entre diversas bases de dados e disciplinas do conhecimento, para que sejam compartilhadas entre vários grupos de usuários ou instituições (RIZZO; TRONCY, 2011).

A partir desse princípio, quanto mais dados forem publicados nesse padrão, maior será a chance e a incidência de serem conectados de forma significativa e útil, por exemplo, conectar coleções de dados bibliográficos de qualquer natureza em RDF, tendo em vista o acesso pela comunidade acadêmica. A fim de estruturar vocabulários em RDF, é utilizado o *RDF Schema* (RDFS) que possui um conjunto de classes e propriedades como uma extensão do vocabulário básico do RDF usado para descrever recursos informacionais (SANTOS; MOREIRA, 2018).

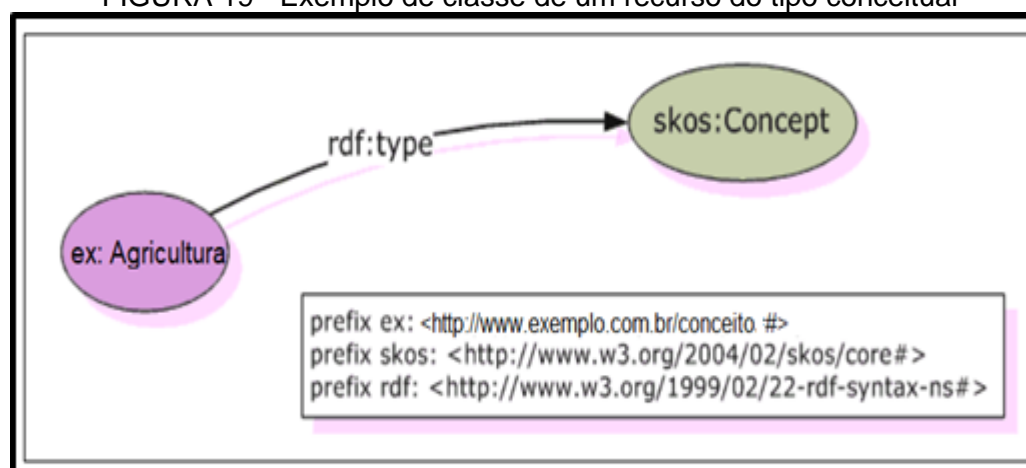
O SKOS permite a representação de um tesouro, o que facilita utilizar a linguagem OWL para expressar a semântica (significado) que eram usadas em outras linguagens tais como XML, RDF/RDFS. A linguagem OWL fornece uma representação de ontologias mais ampla com relação à semântica entre os conceitos e seus relacionamentos em um domínio em comparação com o padrão no formato RDF/RDFS, usados para representar a estrutura e as relações semânticas respectivamente (BAKER *et al.*, 2013). Zeng, Panzer e Salaba (2009) consideram a segunda versão da linguagem OWL melhor em alguns aspectos, pois atende a alguns requisitos específicos, por exemplo, os sistemas de classificação podem ocorrer de maneira enumerativa e sintética, cuja a OWL2, tem diversas expressões e axiomas que descrevem diferentes tipos de classes e propriedades, referentes a um objeto específico e tabelas auxiliares que podem ser representadas no SKOS.

Hartig, Bizer e Freytag (2009) verificaram que, para realizar consultas (*queries*) em bases de dados em formato RDF, utiliza-se a linguagem *Simple Protocol and RDF Query Language* (SPARQL). Desse modo, as aplicações que precisam acessar os dados em LOD devem utilizar serviços, chamados de SPARQL *endpoints*, que fornecem pontos de acesso aos dados, permitindo realizar consultas em bases de dados em RDF. Desse modo, os autores observaram que o RDF permite a combinação de SKOS e propriedades de outros esquemas, tais como os sistemas de classificação que possuem, em sua maioria, uma ordenação significativa a abranger hierarquias de classes sem a necessidade de modelar somente dentro do SKOS. Isso ocorre porque os sistemas de classificação têm diferentes princípios de ordenação, embora seja possível ordenar de modo individual as coleções em formato SKOS (HARTIG; BIZER; FREYTAG, 2009).

Atualmente, a utilização da estrutura da linguagem SKOS, na visão de Marcondes (2021), possui várias propriedades de etiquetagem ou rótulos (notações) que representam um conceito das relações semânticas em um domínio. Para o autor, as estruturas de dados que utilizam o RDF contêm as triplas a partir da definição de prefixos que identificam um recurso, capazes de representar um conceito. Com esse pensamento, os autores exemplificam que a URI <<http://www.exemplo.com.br/conceito#>> possui as seguintes representações: o recurso: <*skos:concept*>, a propriedade: <*rdf:type*> e o valor <“ex: Agricultura”>.

Miles e Brickley (2005) esclarecem que a representação no SKOS por meio das triplas, no padrão RDF que possibilita a formação de uma notação no sistema de classificação. A Figura 19 ilustra essa representação para um recurso da modalidade conceitual.

FIGURA 19 - Exemplo de classe de um recurso do tipo conceitual



FONTE: Adaptada de Miles e Brickley (2005).

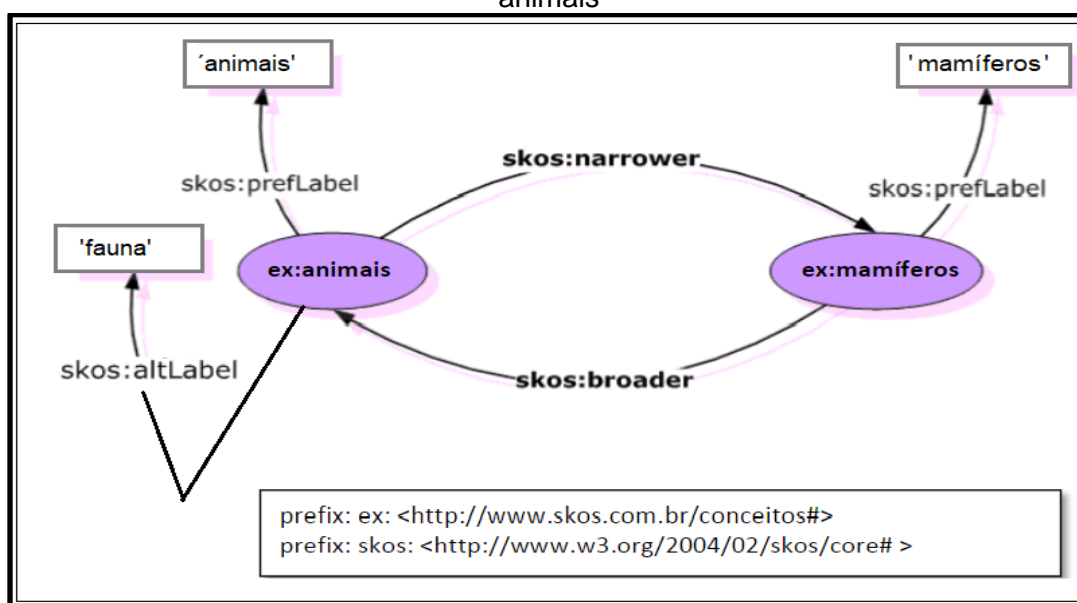
Pelo que consta na Figura 19, entende-se que a classe <*skos:Concept*> permite o reconhecimento de que recurso “Agricultura” é um recurso conceitual, ou seja, o próprio recurso é um conceito que é identificado pela URI:<<http://www.exemplo.com.br/conceito#>>. Este recurso pode estar em um ou vários idiomas, usando o *rdf:type* para identificar que este recurso identificado por esta URI é um conceito, por meio da *tag* <*skos:concept*>, conforme

pontuado por Miles e Brickley (2005). Os elementos `<skos:ConceptScheme>`, que é utilizado para mapear um conjunto de conceitos que são agregados para formar entidades e o `<skos:inScheme>` é usado para vincular outros conceitos a este mapeamento de conceitos. O SKOS utiliza o rótulo `<skos:hasTopConcept>` informa dentro de uma hierarquia de conceitos, quais são os conceitos mais genéricos e o rótulo `<skos:Collection>` para definir coleções de conceitos (ISAAC; SUMMERS, 2009).

Sánchez *et al.* (2012) evidencia a interoperabilidade entre SOC, a partir das propriedades de mapeamento entre dois ou mais conceitos no SKOS por meio das propriedades de *matching* (correspondência) entre diferentes SOC, tais como: `<skos:exactMatch>`, `<skos:closeMatch>`, `<skos:broadMatch>`, `<skos:narrowMatch>` e `<skos:relatedMatch>`. Os autores afirmam que para mapear conceitos equivalentes em diferentes SOC, pode ser representado no SKOS por meio das propriedades: `<skos:exactMatch>`, e `<skos:closeMatch>`. A propriedade `skos:exactMatch` permite que a equivalência entre dois conceitos possam ser disseminados para qualquer tipo de SOC, ou seja, possui um alto grau de semelhança com propriedade transitiva. Já a propriedade `skos:closeMatch` é intransitiva, ou seja, os conceitos são similares somente nos SOC que estão sendo mapeados. As propriedades `<skos:broadMatch>`, `<skos:narrowMatch>` e `<skos:relatedMatch>` são utilizados para conceitos em diferentes SOC, semelhante as propriedades que são utilizadas no mesmo vocabulário `<skos:broader>`, `<skos:narrower>` e `<skos:related>` que serão descritos a seguir (SÁNCHEZ, MÉNDEZ, MUÑOZ, 2012).

Um novo exemplo dessa representação é descrito por Miles e Brickley (2005), dessa vez, no domínio das espécies animais, com foco nos animais mamíferos. A Figura 20 expõe as relações hierárquicas desse assunto ou objeto no domínio das espécies animais.

FIGURA 20 - Exemplo de recursos com relações hierárquicas no domínio das espécies animais



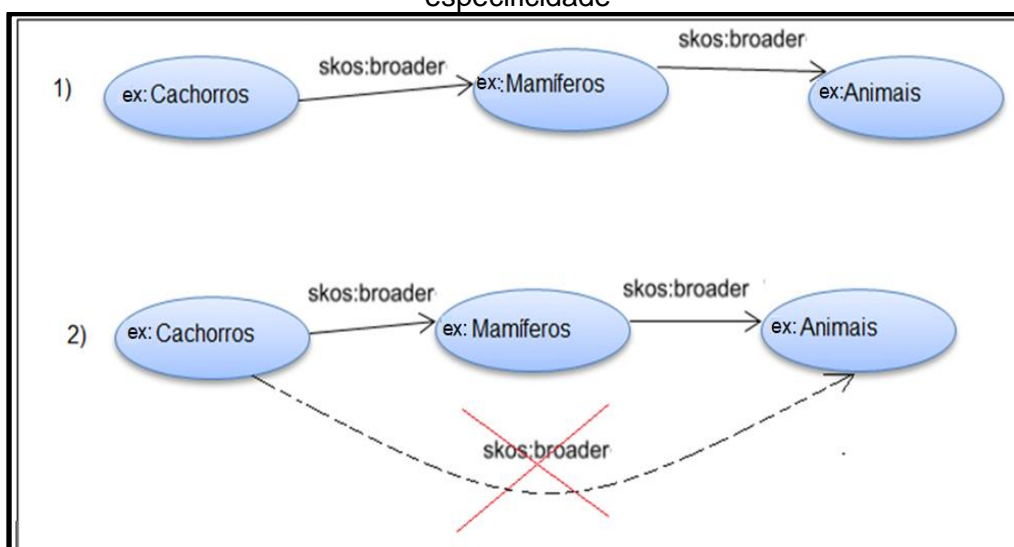
FONTE: Adaptada de Miles e Brickley (2005).

Da Figura 20, depreende-se que os recursos animais e mamíferos termos preferidos dos conceitos “Animais” e “Mamíferos”; as propriedades são: `<skos:broader>` e `<skos:narrower>`. Sánchez *et al.* (2012) relata que para representar um conceito que possui um significado mais amplo do que outro, utiliza-se a propriedade `<skos:broader>`. Já para representar o contrário, ou seja, um significado mais específico, utilizou-se a propriedade `<skos:narrower>`, o que significa que há, nesses termos que representam conceitos, a formação de relações de hierarquia. As propriedades identificadas pelas expressões `<skos:prefLabel>`, `<skos:altLabel>` e `<skos:hiddenLabel>`, representam relações de equivalência entre os conceitos com o intuito de representar termos preferidos; não preferidos (alternativos) e para ocultar as variações terminológicas, siglas que podem ser antecipadas, erros ortográficos em termos, respectivamente, sendo que esta última propriedade é utilizada como forma de melhorar internamente a representação, visando melhorar a recuperação (SANCHEZ *et al.* 2012).

Além disso, os termos podem ter uma *tag* (etiqueta) que indica a variação dos idiomas para os termos preferidos e não preferidos no SKOS, por exemplo, com `@en` para o idioma inglês, como foi utilizada na Figura 12 do capítulo 3.1.2. Nesse sentido, conforme a Figura 20 pode-se utilizar o conceito “Animais” com a propriedade `<skos:prefLabel>` que está associado ao termo, representado pelo retângulo com o literal ‘animais’, para indicar que o termo preferido é “animais”. Para mostrar os termos não preferidos ou alternativos, pode-se exemplificar o conceito “Animais” com a propriedade `<skos:altLabel>` em relação a “fauna” (SÁNCHEZ, MÉNDEZ, MUÑOZ, 2012).

Em uma estrutura hierárquica, as propriedades para representar as hierarquias em um SOC são representadas na Figura 21, sendo definidas como transitivas entre os termos “Mamíferos” e “Animais” (ISAAC; SUMMERS, 2009). Os autores representam, por meio de exemplificação, as relações transitivas e intransitivas, com foco no universo do conceito “Animais”, dessa vez, com novos níveis de especificidade, considerando a espécie “Mamíferos” do tipo “Cachorro”. A Figura 21 ilustra essas relações apontadas pelos autores.

FIGURA 21 - Exemplos de hierarquias transitivas e intransitivas com níveis de especificidade



FONTE: Adaptada de Isaac e Summers (2009).

Na Figura 21, as setas sólidas expressam a propriedade *<skos:broader>*, por exemplo, o conceito “Cachorros” em relação a “Mamíferos”. Assim, para a propriedade *<skos:narrower>* do conceito “Mamíferos” em relação a “Cachorros”, o conceito “Mamíferos” é mais geral do que “Cachorros” que, por sua vez, é mais específico do que “Mamíferos”. Do mesmo modo, seria com o conceito “Mamíferos” com a propriedade *<skos:broader>* em relação a “Animais”, e/ou o conceito “Mamíferos” em relação à propriedade *<skos:narrower>* do conceito “Animais” em relação a “Mamíferos”. Por outro lado, a semântica representada não suporta inferir diretamente a relação existente entre o conceito “Cachorros” e “Animais”, ou seja, não se pode dizer que cachorro é um animal, sem a inclusão das características do conceito “Mamíferos”, ou seja de forma direta conforme mostra a parte inferior da Figura 20 (ALLEMANG; HENDLER, 2011).

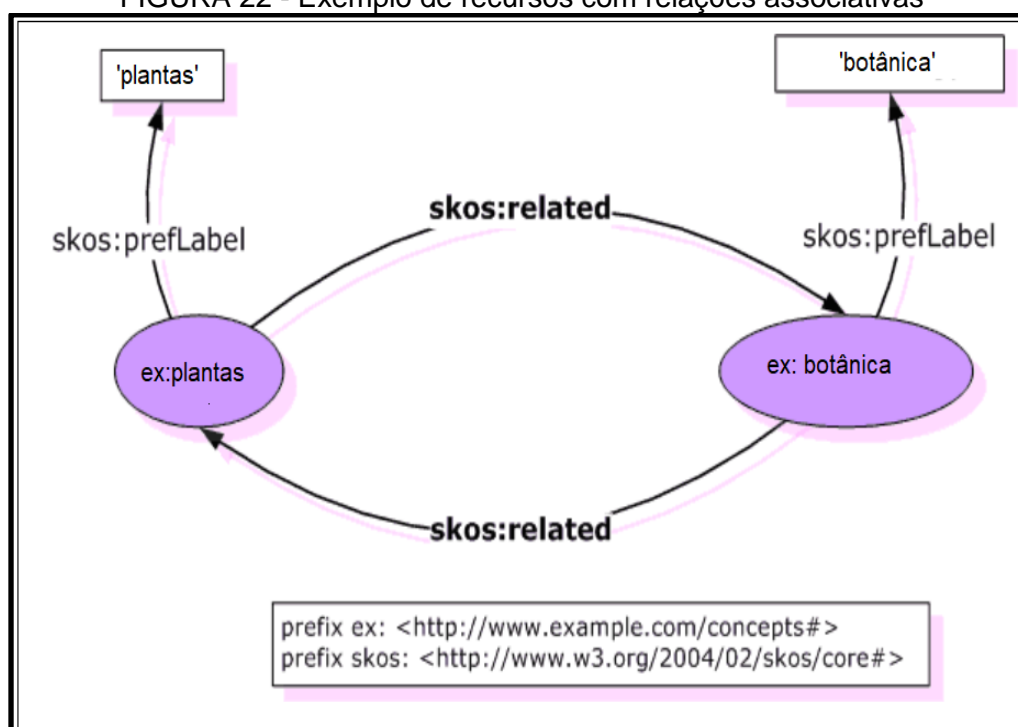
Desse modo, o conceito “Animais” possui a propriedade *<skos:broader>* em relação ao conceito “Mamíferos” e “Mamíferos” com a propriedade *<skos:broader>* em relação ao conceito “Cachorros”. Então, é definido como propriedade intransitiva, quando é informado que o conceito “Animais” possui propriedade *<skos:broader>* em relação a “Cachorros”; quando ocorre essa situação, é comum diferenciar as propriedades de representação utilizando setas pontilhadas (ISAAC; SUMMERS, 2009).

No contexto da Figura 21, considerando a reflexão proposta por Isaac e Summers (2009), entende-se que as setas sólidas representam declarações de conceitos afirmativos e, caso seja necessário inferir conceitos, deve-se utilizar a expressão *<skos:broaderTransitive>* e/ou *<skos:narrowerTransitive>*. Já as setas pontilhadas representam a declaração de inferência no modelo de dados no SKOS. Para representar a relação transitiva, devem ser inseridas as propriedades *<skos:broaderTransitive>* e/ou *<skos:narrowerTransitive>*, por

exemplo, o conceito “Cachorros” com a propriedade *<skos:broaderTransitive>* em relação a “Mamíferos”, com setas pontilhadas, também, pois, de modo similar, trata-se de outras propriedades (ISAAC; SUMMERS, 2009).

Em relação aos relacionamentos associativos (não hierárquicos), utilizam-se a expressão *<skos:related>*, visto que nenhum conceito é mais geral ou mais específico que o outro. A Figura 22 indica uma relação associativa estabelecida entre o conceito “Botânica” com a propriedade *<skos:related>* em relação a “Plantas” e com os termos preferidos Botânica e Plantas (MILES; BRICKLEY, 2005).

FIGURA 22 - Exemplo de recursos com relações associativas



FONTE: Adaptada de Miles e Brickley (2005).

Conforme ilustrado na Figura 22, considerando as explicações de Miles e Brickley (2005), a partir da relação descrita, a propriedade *<skos:related>* pode estabelecer relações associativas em diversos tipos de exemplos em uma ou mais disciplinas do conhecimento. Isso será possível, mediante a troca da disciplina “Botânica” por outra (MILES; BRICKLEY, 2005). Para Isaac e Summers (2009), a definição das notas de escopo no SKOS utiliza-se a propriedade *<skos:scopeNote>*, que são notas de escopo possuem uma breve descrição para um conceito para fins de documentação, podendo ainda esta documentação ser mais específica, utilizando as propriedades *<skos:definition>*, *<skos:example>* e *<skos:historyNote>*, que correspondem a definições formais, exemplos e notas históricas (mudanças no significado) para um conceito (ISAAC; SUMMERS, 2009).

Marcondes (2021) também exemplifica os tipos de relacionamentos hierárquico (termo genérico, termo específico), associativo (termo relacionado) ou terminológico (Usado, Usado para) entre conceitos na linguagem SKOS (MARCONDES, 2021). A Figura 23

apresenta um conceito representado em SKOS no formato RDF, conforme discorrido por Marcondes (2021).

FIGURA 23 - Esquema de representação de um conceito no vocabulário SKOS

```
<skos:ConceptScheme rdf: about="http://ClassificacaoFrutas/>
  <skos:Concept rdf: about = "http://ClassificacaoFrutas/
  Laranja">
    <skos:definition>Frutas cítricas ricas em vitamina C</
    skos:definition>
    <skos:prefLabel xml: lang = "pt">Laranja@pt</
    skos:prefLabel>
    <skos:broader rdf: resource = "http: //ClassificacaoFrutas/
    Fruta"/>
    <skos:narrower rdf: resource = "http: //
    ClassificacaoFrutas/Laranja_Lima"/>
    <skos:related rdf: resource = "http: //ClassificacaoFrutas/
    Tangerina"/>
  </skos: Concept>.
```

FONTE: Marcondes (2021, p. 73).

A partir do exemplo exposto na Figura 23, depreende-se que o vocabulário SKOS “[...] permite também especificar o mapeamento e compatibilidade entre conceitos de um SOC em outro [...]” (MARCONDES, 2021, p. 73). No exemplo apresentado, foi acrescentada uma linha “<skos:exactMatch rdf:resource = “http://aims.fao.org/aos/agrovoc/c_25492”/>”, que especifica o conceito “Laranja”, a partir da hipotética classificação de frutas equivalente ao conceito “Orange jasmine” do Tesouro AGROVOC¹⁹ (MARCONDES, 2021). Portanto, constata-se, segundo o autor, que uma das principais aplicações do SKOS é garantir o intercâmbio entre os sistemas, possibilitando uma representação mais consistente e abrangente.

Neste capítulo, apresentaram-se os fundamentos teórico-metodológicos que norteiam esta pesquisa. No próximo capítulo, serão mencionadas a caracterização e a descrição da metodologia proposta para esta tese.

¹⁹Disponível em: <http://aims.fao.org/est-registry/vocabularies/agrovoc>.

4 CARACTERIZAÇÃO E DESCRIÇÃO DA METODOLOGIA

“O cientista não é o homem que fornece as verdadeiras respostas; é quem faz as verdadeiras perguntas”.

(Claude Lévi-Strauss)

Descrever os procedimentos a serem seguidos na realização de uma pesquisa científica varia de acordo com as peculiaridades de cada estudo, área ou pesquisador, sendo necessário, em qualquer caso, apresentar os métodos e técnicas que serão utilizados. Neste capítulo, são apresentadas a caracterização e a descrição da pesquisa, contemplando a delimitação do universo investigado. Além disso, são apontados os procedimentos metodológicos para atingir o objetivo geral e os objetivos específicos, conforme descrito nos subcapítulos da introdução.

4.1 Caracterização da pesquisa

Seguindo os parâmetros estabelecidos por Gil (2010), nesta pesquisa, quanto aos objetivos, houve procedimentos de natureza **(1) exploratória** e **(2) descritiva**. A pesquisa exploratória deve-se ao aprimoramento de ideias ou descobertas realizadas com o objetivo de compreender melhor um fenômeno ou objeto de estudo, sobretudo pelo fato de existirem poucos estudos sobre a temática (GIL, 2010). Isso ocorre durante a pesquisa bibliográfica, cujo tema detalhado nesta tese é o uso da CDU com o auxílio de um tesouro para trazer melhorias nas taxas de exaustividade e especificidade (revocação e precisão) durante a recuperação da informação. A pesquisa descritiva, por sua vez, baseia-se em técnicas de coleta e análise de dados a partir da descoberta e interpretação entre as variáveis e suas relações sobre o fenômeno ou tema, com o objetivo de confrontar os dados com os pressupostos (GIL, 2010). Os termos do tesouro AgroTermos²⁰ foram inseridos na estrutura da CDU com o objetivo de contribuir na garantia de melhorias na recuperação.

O estudo foi de natureza **(3) aplicada**, ou seja, gerar conhecimento na prática, a partir da resolução de problemas específicos que ocorrem na realidade. Assim, o estudo propõe um modelo de compatibilidade de linguagem entre a CDU e o AgroTermos para o domínio da agricultura, para melhorar a recuperação. Quanto à forma de abordagem do problema, a pesquisa é **(4) quantitativa**, pois foi definida a quantidade de dados, de modo a facilitar as análises. Além disso, a pesquisa também tem aspecto **(5) qualitativo**, uma vez que a amostra de análise foi delimitada por meio de termos que denotam conceitos do mundo real, validados por um especialista do domínio da agricultura e, nesse processo, não foram

²⁰ Disponível em: <https://sistemas.sede.embrapa.br/AgroTermos/#>.

utilizados dados quantificáveis. Portanto, trata-se de uma abordagem focada no processo, haja vista procurar perspectivas e significados para a pesquisa. Nesse sentido, de acordo com Creswell (2010), o pesquisador procura credibilidade baseada em coerência, percepção, utilidade instrumental e integridade, mediante um processo de verificação da veracidade dos dados.

Segundo Gil (2010), o delineamento da pesquisa se refere ao seu planejamento geral, sendo definidos os procedimentos técnicos. Logo, nesta tese, recorreu-se à aplicação de uma **(6) pesquisa bibliográfica**, que levantou os resultados alcançados em estudos realizados por diversos autores que investigam a temática. Esses estudos foram publicados em livros e periódicos, cujos resultados forneceram insumos para a fundamentação teórica e metodológica, a revisão sistemática da literatura e para a escolha da metodologia a ser utilizada. Destaca-se que, durante a revisão de literatura, foram aplicados critérios de inclusão na seleção das fontes e foi utilizado o método indutivo, durante a análise realizada nos textos selecionados. Além disso, a síntese dos estudos para gerar resultados foi conduzida pela técnica da análise de conteúdo, com a distribuição dos trabalhos em categorias temáticas. Tal procedimento visou à identificação dos resultados com base nas mensagens dos textos, conforme sugerido por Bardin (2011).

Na visão de Gil (2010), a pesquisa bibliográfica é elaborada a partir de materiais já publicados, tais como: livros científicos ou técnicos, artigos publicados em periódicos ou anais de congresso, entre outros. Desse modo, foram escolhidas somente fontes bibliográficas e documentais confiáveis para não haver informações incorretas ou desatualizadas, de modo a não prejudicar a validade da pesquisa e, em todas as fontes analisadas, aplicou-se a análise de conteúdo, por meio da definição de categorias, construídas a partir das semelhanças entre os conteúdos dos textos.

Nesta tese, quanto aos procedimentos técnicos, foi realizado um estudo empírico aplicado à estrutura classificatória da CDU para melhorar a recuperação da informação, em específico na classe 631 - Agricultura em geral (Anexo A), com base nos trinta termos definidos conceitualmente por Maculan (2015) sobre a “Intensificação Agropecuária” (Anexo B). Também foram utilizados os termos sobre a temática, considerando a estruturação elaborada por Maculan (2015), cujos termos foram retirados do tesouro AgroTermos da Embrapa (domínio da agricultura) (Anexo C) para facilitar o estabelecimento das relações semânticas entre os dois SOC.

Em linhas gerais, na estrutura desses termos, percebem-se os fenômenos produzidos pela compatibilização de linguagem entre os conceitos nos dois instrumentos (CDU e AgroTermos), produzindo conhecimento a partir de dados reais. A escolha pela temática (Intensificação Agropecuária) e pelo domínio (Agricultura) pode ser justificada, por se tratar de uma estrutura semântica já consolidada, isto é, já está validada por um

especialista do domínio. Ressalta-se que essa escolha possibilita benefícios recíprocos, reforçando a relação teoria/prática, que se manifesta no contexto da universidade/sociedade, considerando-se os benefícios que podem surgir a partir dos resultados da pesquisa, com comum acordo, portanto, entre o autor desta tese e a Embrapa.

A escolha desse escopo foi bem recebida pelo autor desta tese, visto constituir-se de um problema que muito permeia o universo da organização da informação, área vinculada a área de pesquisa de doutorado do pesquisador. Além disso, o recorte temático escolhido constitui uma subárea da agricultura, que possui grupos de pesquisadores trabalhando com o tema. Com efeito, esta pesquisa poderá contribuir, dentre outros benefícios, para o fortalecimento e o desenvolvimento dos estudos realizados pela Embrapa ou por outras instituições. A contribuição especial volta-se à área da organização da informação, visto que serão reveladas estratégias de representação, para profissionais indexadores, por exemplo, como também, indicações que melhorem as estratégias de recuperação adotadas pelos usuários. Sendo assim, foram fornecidos insumos para o desenvolvimento da proposta de um modelo de utilização da CDU compatibilizada com o tesauro, acreditando que esse é um processo que promoverá melhorias na recuperação da informação, sobretudo em documentos disponibilizados na web ou em qualquer ambiente digital.

Ainda quanto aos procedimentos adotados, esta pesquisa se caracteriza como um **(7) estudo de caso**, conduzido por uma investigação profunda e exaustiva, a partir da análise de estudos similares, que trabalharam abordagens similares. Isso possibilita ao pesquisador identificar o problema de forma geral e observar as teorias e os procedimentos adotados para produzir conhecimento de modo descritivo, exploratório e pragmático, considerando os apontamentos de Gerring e Mcdermott (2007).

Assim, o objeto estudado com a finalidade de apresentar os aspectos de um grupo ou população (universo de pesquisa) relacionados a esse objeto, prevendo resultado sem algum contexto da realidade (amostra), o que possibilita uma visão genérica e específica do objeto de estudo (YIN, 2010). Além disso, o autor menciona que o estudo de caso facilita a compreensão do método utilizado na pesquisa e sua aplicação, seja em organizações, situações sociais, políticas, dentre outras circunstâncias.

Ventura (2007) relata que o estudo de caso permite realizar novas descobertas de acordo com a integridade do objeto a partir da definição, planejamento, técnicas e aplicação dos procedimentos metodológicos, desde a coleta até a análise dos dados, gerando os resultados da investigação. Com esse pensamento, o estudo de caso aplicado nesta tese foi norteado a partir da análise de uma metodologia que utiliza a compatibilidade de linguagem entre a estrutura classificatória da CDU no domínio da agricultura e os termos do tesauro Agrotermos com o propósito de trazer melhorias na recuperação. Portanto, foi validada uma

proposta de procedimentos metodológicos no âmbito da temática delimitada. Logo, o estudo de caso permitiu verificar a relevância do modelo proposto.

4.2 Delimitação do universo da pesquisa

A delimitação do universo da pesquisa, conforme relata Marconi e Lakatos (2017), consiste em explicitar quais serão os elementos constituintes da pesquisa, sejam pessoas, coisas, fenômenos, dentre outros. Nesse processo, as autoras afirmam que são enumeradas as características comuns desses elementos, como sexo, faixa etária, organização a que pertencem, comunidade onde vivem e outros aspectos.

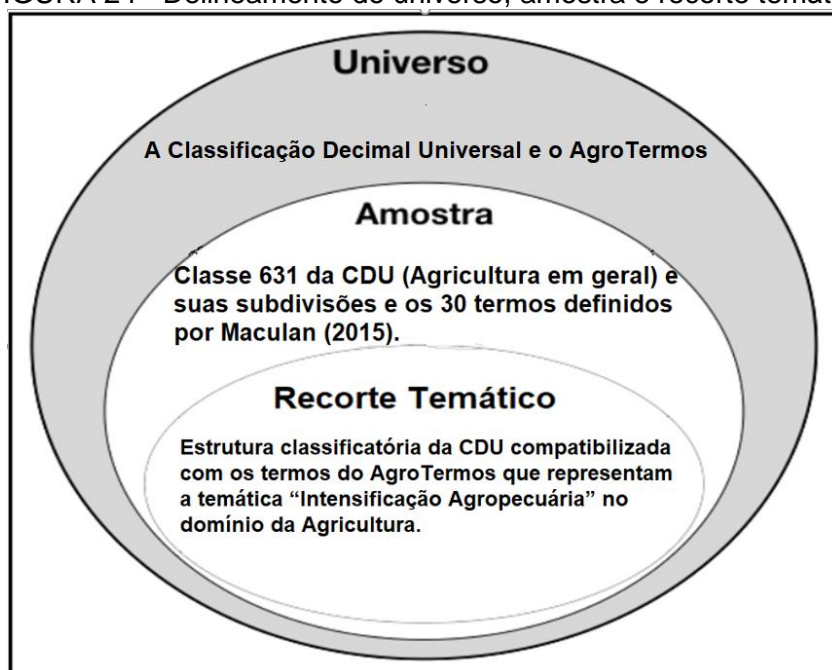
O universo da pesquisa selecionado para este estudo é: a Classificação Decimal Universal (CDU) e o tesouro AgroTermos. O objeto de estudo desta pesquisa é a estrutura classificatória da CDU e os termos do tesouro AgroTermos. A escolha da CDU se deve pelo fato desse sistema de classificação ser menos rígido e mais flexível, permitindo maior especificidade dos assuntos, algo com poucas possibilidades de ocorrer em sistemas mais engessados, como a Classificação Decimal de Dewey (CDD). Nesse sentido, a flexibilidade da CDU está associada à sua estrutura semifacetada, que possui princípios e regras para a combinação de elementos, com a utilização de notações e auxiliares relacionais, capazes de abranger assuntos compostos ou complexos de diversas áreas do conhecimento e aumentar a possibilidade de recuperar as informações representadas (CLASSIFICAÇÃO DECIMAL UNIVERSAL, 1997). Nesse sentido, a amplitude do universo requer a sua fragmentação, ou seja, é delimitada a amostra de pesquisa, capaz de representar o universo. No presente estudo, a seleção da amostra foi escolhida de forma não probabilística e intencional. Portanto, a amostra, no que diz respeito à CDU, foi a classe 631 (Agricultura em geral) e suas subdivisões (Anexo A). Já em relação ao tesouro AgroTermos, a amostra foi formada pelos trinta termos e seus relacionamentos semânticos, conforme apresentado no Anexo B.

A estrutura da classe 631 da CDU foi retirada do repositório da *UDC Consortium*, que possui uma edição completa, com serviço pago, mas pode ser consultada dentro de um período de teste gratuito nos idiomas inglês, espanhol, francês, alemão, norueguês, entre outros. Nesta tese, a CDU foi analisada no idioma inglês, disponível no link (<https://udc-hub.com/en/login.php>), cujo arquivo pode ser usado por estudantes e professores de Biblioteconomia, profissionais da área e pesquisadores. A escolha da classe 631 e suas subdivisões na CDU se justificam pelo fato dos assuntos se aproximarem mais dos assuntos do recorte temático, onde estão os termos representativos da estrutura do tesouro “Intensificação Agropecuária” no idioma português (Brasil). A classe 631 e as subdivisões utilizadas da CDU foram traduzidas para o português (Brasil) utilizando como base na versão impressa da CDU, não para tradução idêntica, mas para diminuir falhas na tradução e

umentar a compreensão dos conceitos de forma objetiva (CLASSIFICAÇÃO DECIMAL UNIVERSAL, 1997).

O recorte temático são as classes da CDU que foram compatibilizadas com os termos do tesauro AgroTermos que representam o assunto “Intensificação Agropecuária” dentro do domínio da agricultura. Foi utilizado o tesauro AgroTermos no ambiente da Embrapa para analisar os relacionamentos semânticos entre os termos da estrutura classificatória da CDU e os trinta termos representativos do recorte temático utilizado na tese de Maculan (2015). A autora propôs, de maneira estruturada, a representação dos conceitos com seus respectivos termos representativos do domínio da agricultura, cuja fundamentação teórica foi baseada em Boserup (1965) e validada pelo especialista de domínio. Ancorando-se na tese da autora citada, este estudo utilizou as relações semânticas incluídas entre os conceitos, a partir dos termos definidos por Maculan (2015), o que facilitou o entendimento da temática. A Figura 24 apresenta o universo, a amostra e o recorte temático utilizados nesta tese.

FIGURA 24 - Delineamento do universo, amostra e recorte temático



FONTE: Elaborada pelo autor (2021).

Após apresentar as características iniciais desta pesquisa, incluindo-se a sua ambientação, é importante detalhar os procedimentos realizados, considerando, sobretudo, a distribuição dessas atividades em etapas. O subcapítulo seguinte destina-se à apresentação desse detalhamento metodológico.

4.3 Procedimentos metodológicos

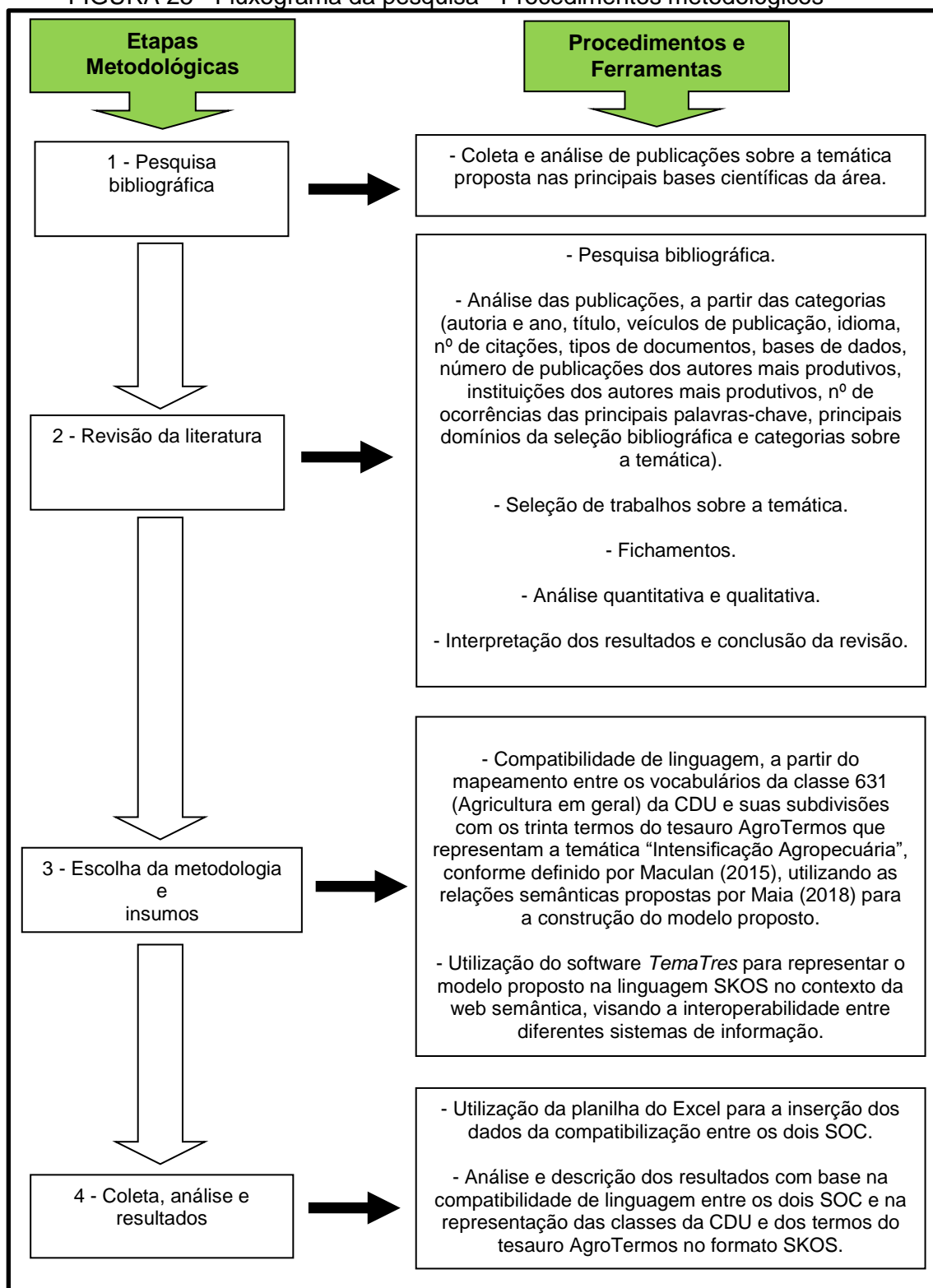
Neste subcapítulo, são apresentadas as principais etapas desta pesquisa, detalhando os procedimentos metodológicos utilizados. As atividades foram organizadas em

quatro etapas de acordo com o delineamento da pesquisa, cujos procedimentos foram distribuídos do seguinte modo:

- **Primeira etapa - pesquisa bibliográfica:** a pesquisa bibliográfica foi realizada durante todas as etapas do desenvolvimento desta tese, com o objetivo de se obter o conhecimento necessário acerca dos fundamentos teórico-metodológicos e da caracterização e descrição da metodologia utilizada no estudo de caso desta pesquisa.
- **Segunda etapa - revisão da literatura:** a apresentação da metodologia utilizada na revisão sistemática da literatura está descrita no capítulo 2, que expõe o detalhamento para selecionar e analisar os estudos relacionados com o tema desta pesquisa. No que tange ao processo de seleção dos materiais, foram utilizados, apenas, estudos que tange ao processo de seleção dos materiais, foram utilizados, apenas, estudos que focavam no “Uso da CDU com o auxílio de um tesouro para melhorar a recuperação da informação”. Dentre os principais estudos selecionados, citam-se os seguintes autores: Aida Slavic, Gerhard Riesthuis, la Cecilia Mcilwaine, Ingetraut Dahlberg, Maràngels Granados Colillas, Maria Inês Cordeiro, Raymond Wall, Vanda Broughton e Victoria Frâncu. A partir do mapeamento desses estudos, eles foram sintetizados e agrupados em categorias por semelhanças, conforme Bardin (2011), cujo estado da arte foi discutido de modo crítico. Os resultados dessa análise permearam o universo do problema, com estudos relevantes e focados, selecionados a partir de buscas em pesquisas e publicações na área, a fim de detectar trabalhos semelhantes ou idênticos ao que se pretende desenvolver, conforme explorado nos capítulos anteriores.
- **Terceira etapa - escolha da metodologia e insumos:** a partir dos fundamentos teóricos elucidados com as etapas anteriores, foi possível traçar a metodologia, delimitando-se o escopo e o campo de coleta e a análise dos dados para a proposta de um modelo de compatibilidade entre dois SOC. Reforça-se que tal modelo foi baseado na classe 631 da CDU e nos termos que representam os conceitos da temática “Intensificação Agropecuário” do AgroTermos. A proposta desse modelo foi a partir dos trabalhos de Dahlberg (1981,1983), Frâncu (2003a, 2003b), Slavic (2005), Maculan (2015) e Maia (2018). A etapa foi dividida entre insumos metodológicos (atividades) e tecnológicos (ferramentas).
- **Quarta etapa – coleta e análise dos dados e resultados:** constitui a fase de aplicação e discussão dos instrumentos metodológicos. Para tanto, apresentam-se em detalhes e de modo crítico a coleta e a análise dos dados no campo da investigação, ou seja, a análise realizada na amostra de pesquisa (domínio da agricultura e recorte temático da Intensificação Agropecuária), contextualizando o que os dados revelam para a geração dos resultados.

No intuito de explicitar o detalhamento da metodologia, apresenta-se o fluxograma da pesquisa (Figura 25). Tal fluxograma, além de destacar as quatro etapas da pesquisa, expõe os instrumentos de coleta e análise dos dados que possibilitaram a aferição dos resultados e das conclusões.

FIGURA 25 - Fluxograma da pesquisa - Procedimentos metodológicos



FONTE: Elaborada pelo autor (2020).

Conforme observado na Figura 25, nota-se que a pesquisa se desenvolveu de modo integrado, ou seja, mesmo sendo delimitadas quatro etapas, elas são complementares e dependentes umas das outras. Destaca-se que, na metodologia, manifesta-se o ambiente *in loco* da pesquisa, pois, como se trata de um estudo de caso, o trabalho foi desenvolvido no ambiente da Embrapa e será detalhado a seguir. Na terceira etapa da metodologia, além das especificações apresentadas na Figura 25, foram coletados dados que caracterizam essa instituição, de modo a descrevê-la, além da descrição dos instrumentos de controle terminológico utilizados por essa instituição, considerando o domínio da agricultura.

A técnica de coleta de dados foi a partir da estrutura classificatória da CDU, mais especificamente na Classe 631 (Agricultura em geral) e suas subdivisões, onde os dados foram retirados do site da instituição do Consórcio da CDU. Os termos e seus respectivos conceitos em relação à temática foram retirados da taxonomia proposta Maculan (2015), a partir do tesouro AgroTermos, com os relacionamentos semânticos entre os termos que representam cada conceito. Os relacionamentos semânticos entre os termos que representam os trinta termos da amostra, auxiliaram na compreensão dos conceitos representados pelos termos da CDU e do AgroTermos.

Nesse sentido, os termos que denotam os conceitos representados nos dois SOC, serão utilizados por usuários em um SRI, onde a linguagem utilizada durante a recuperação deve ser a mesma utilizada na indexação sobre o conteúdo dos recursos informacionais em relação ao recorte temático, levando em conta que o processo de representação que é permeado por subjetividade. Esta técnica de coleta de dados supracitada, serviu para atingir o primeiro objetivo específico, que foi analisar a compatibilidade dos termos da estrutura classificatória da CDU com os descritores do tesouro AgroTermos, na representação do assunto “Intensificação Agropecuária” do domínio da agricultura. Este objetivo vai permitir resolver o problema de entendimento dos conceitos do assunto “Intensificação Agropecuária” do domínio da agricultura a ser mapeado nos dois SOC, que é o ponto de partida para a compatibilidade de linguagem entre SOC.

O mapeamento dos termos da estrutura classificatória da CDU com seus descritores foi a partir da Classe 631 e suas subdivisões, que foram compatibilizados com os termos em linguagem natural no idioma português (Brasil), que representam os conceitos do recorte temático no tesouro AgroTermos no campo da “Intensificação Agropecuária”. Na CDU, já é atribuída a notação com seus respectivos descritores de acordo com sua estrutura classificatória original. Nessa perspectiva, o processo de inserção da notação foi de acordo com a ordem de citação da CDU, e a representação semântica de cada termo referente à temática proposta foi de acordo com os conceitos presentes na estrutura hierárquica desse sistema e das relações semânticas entre os trinta termos que representam o recorte temático no AgroTermos. A princípio foram inseridos os termos que representam a temática no tesouro

AgroTermos na base da Embrapa, o que possibilitou recuperar as relações semânticas entre os termos no domínio da agricultura.

Para realizar a compatibilidade entre as linguagens da CDU e do tesauro AgroTermos, foi a partir da pesquisa bibliográfica das relações semânticas no contexto da BCI e da análise do recorte temático nos dois SOC. Nesse sentido, os relacionamentos semânticos extraídos do AgroTermos facilitaram na definição das relações semânticas entre os termos nos dois SOC pelo autor desta tese, de acordo com a pesquisa de Maia (2018). Em seguida, com esforço intelectual, foi possível estabelecer as relações semânticas entre as classes da CDU em relação aos trinta termos definidos por Maculan (2015), que representam de maneira conceitual o recorte temático no domínio da agricultura.

Por fim, as relações semânticas definidas no contexto da BCI entre as classes da CDU e o termos do recorte temático no AgroTermos foram validadas pelo especialista do domínio, verificando as compatibilidades e incompatibilidades de cada termo e com seus respectivos conceitos, o que torna a proposta do modelo a ser construído adequado. Assim, foi possível agregar à CDU os termos comumente utilizados para representar os conceitos dessa temática com o auxílio da taxonomia “Intensificação Agropecuária”.

Para a realização do mapeamento, utilizou-se a matriz de compatibilidade conceitual de Dahlberg (1981) onde foi realizada a análise conceitual, a partir da estrutura classificatória da CDU e do AgroTermos. Para realizar este procedimento, foi utilizado o método de ligação direta entre os dois SOC, de forma unidirecional, sem a utilização de linguagem intermediária, por se tratar de um mapeamento entre duas linguagens com diferentes estruturas, mas no mesmo idioma e por fim, com o uso das relações semânticas proposta por Maia (2018) (HORSNELL, 1975; ISO 25964, 2013; MANIEZ, 1997). Além disso, foram considerados apenas as classes da CDU que tiveram duas ou mais coincidências conceituais (CC) na compatibilização, por meio de relações semânticas entre a CDU os termos no AgroTermos. Desse modo, justifica-se a compatibilização entre os dois SOC, ou seja, devem possuir mais de uma coincidência conceitual a partir das relações semânticas entre os termos de modo a serem representados, no modelo proposto. Os elementos oriundos desse mapeamento permitiram a construção do modelo proposto, com a inclusão desses dados em uma planilha do Excel, para depois serem usados como técnica para a análise dos dados, permitindo alcançar o segundo objetivo específico, que foi verificar o potencial da recuperação com o uso da expansão dos termos e seus relacionamentos semânticos mapeados nesses dois SOC.

Outra técnica de análise de dados foi a partir da aplicação do modelo proposto, utilizando o software *TemaTres* para ser representado na linguagem SKOS, por meio da sintetização do conteúdo expresso nos diversos documentos analisados sobre a estrutura classificatória da CDU e o AgroTermos. Esta técnica permitiu o cumprimento do terceiro e

último objetivo específico, que a representação do modelo de compatibilidade proposto, utilizando o padrão Simple Knowledge Organization System, no âmbito da web semântica. No percurso da terceira etapa, verificou-se a necessidade de um arcabouço teórico e metodológico para verificar a estrutura semântica dos SOC utilizados, levando em consideração o tipo de relação semântica entre os termos que representam a temática. Ressalta-se que as relações semânticas, no ambiente digital, foram estabelecidas por meio de uma correspondência (*matching*) entre os termos (metadados) que denotam os conceitos, possibilitando a representação, visando a melhoria na relevância (revocação e precisão) na recuperação da informação dos assuntos vinculados à temática.

Considerando a amplitude da terceira etapa, convencionou subdividi-la em duas fases: 1) insumos metodológicos e 2) insumos tecnológicos. Na primeira fase, foram descritos de forma resumida os estudos que serviram como base para a criação de um modelo de utilização da CDU com o auxílio de um tesouro. Nesse percurso, considerou-se a compatibilidade de linguagem entre os conceitos dos dois SOC por meio de relações semânticas. Com efeito, os procedimentos metodológicos utilizados para a proposta de um modelo baseado nos dois SOC utilizados nesta pesquisa foram sustentados por sete estudos, conforme detalhado no Quadro 14.

QUADRO 14 - Estudos que sustentaram a constituição do modelo baseado em dois instrumentos de organização do conhecimento

Autor	Ano	Evidências
Dahlberg	1981 e 1983	Nesses estudos, foi definida uma abordagem para compatibilizar vocabulários a partir de uma matriz de compatibilização semântica, organizada pelo mapeamento conceitual, que era chamado por Dahlberg como Sistemas Ordenados, que atualmente são conhecidos como SOC. Durante a compatibilização, os conceitos denotados por termos são analisados em qualquer nível de hierarquia por diferentes SOC, com vistas a permitir o intercâmbio de dados, mesmo que as linguagens de cada SOC estejam em diferentes bases de dados.
Frâncu	2003a e 2003b	Esses estudos sugerem uma estrutura lógica para a compatibilização da CDU baseada em dois tesouros multilíngues com a utilização de linguagem intermediária. A CDU foi escolhida por se tratar de um instrumento de controle terminológico por ser mais amigável e mais fácil de usar durante a busca. Com isso, a consistência e controle de notações são realizados conectando os termos relacionados dentro de uma ou em várias notações da CDU. Segundo a autora, não se deve utilizar uma notação aleatória baseada em notações simples ou de forma linear; deve-se utilizar a flexibilidade da CDU, combinando os assuntos presentes nas diversas classes do sistema. Contudo, uma vez que a lista de descritores for estabelecida e as características de relacionamento adicionadas à estrutura do tesouro, isso aumentará a eficácia no uso desse instrumento durante os processos de indexação e recuperação. Foi utilizado a compatibilização de linguagem com a CDU o primeiro tesouro com a notação e descrição no tesouro LTHES, que é um tesouro multilíngue interdisciplinar nos idiomas: inglês, francês e romeno, baseado na versão resumida da CDU com baixo nível de especificidade. O segundo tesouro utilizado foi o PTHES (<i>Pocket Edition</i>), que também é um tesouro multilíngue nesses idiomas, que foi compatibilizado com a CDU, porém como este tesouro possui um nível maior de detalhes, obteve resultados melhores na RI em relação ao primeiro. A diferença entre os dois é que o tesouro LTHES permite a busca de informações apenas para assuntos amplos, enquanto o PTHES permite

		pesquisas mais refinadas, resultando em mais precisão e, portanto, maior relevância dos resultados durante a busca. Em suma, pelos resultados desses estudos, constatou-se que a utilização de tesouros com maior nível de detalhamento, permite aos usuários ampliar os resultados obtidos em suas pesquisas e com o tesouro com baixo nível de especificidade, irá restringir os resultados durante a busca, conforme suas necessidades de informação dos usuários.
Slavic	2005	Esse estudo investigou a presença de sistemas de classificação, especificamente a CDU e o seu papel na organização do conhecimento na web. A autora descreveu as tendências e o desenvolvimento de SOC em ambiente digital focados em questões específicas da CDU, seu desempenho e suas funções para a recuperação. Para recuperar de modo eficiente, a classificação precisa desambiguar e contextualizar as consultas para melhorar a precisão e a recuperação, levando em conta a busca e navegação de assuntos na CDU, em ambiente on-line. O conjunto de termos pode ser utilizado para navegação de forma alfabética e sistemática. Na maioria das vezes, é exibido em ordem alfabética, mas dentro de cada entrada pode ser possível inserir uma hierarquia com novos termos. Assim, o conjunto de termos na forma de um tesouro pode ser explorado, posteriormente, de modo a facilitar a busca e a navegação de assuntos da CDU, que deve ser armazenada e distribuída em formato de banco de dados <i>Master Reference File</i> (UDC MRF) para seu uso em ambiente online.
Maculan	2015	Propôs a descrição dos relacionamentos entre os termos que representam a temática "Intensificação Agropecuária" a partir do tesouro AgroTermos da Embrapa. O modelo proposto nessa tese foi adaptado de acordo com a CDU (Classe 631 - Agricultura em geral) e os trinta termos do tesouro AgroTermos que representam a temática representada nesta pesquisa.
Maia	2018	Determinou uma metodologia de uso das relações semânticas para relacionar dois conceitos em um determinado domínio para a área da Biblioteconomia e Ciência da Informação. A autora propõe um modelo de extração de relações semânticas para a representação do conhecimento de documentos acadêmicos no contexto do idioma português brasileiro, a partir de vários estudos de diversos autores, com o intuito de diminuir problemas de inconsistência entre os termos estabelecidos. No contexto dessa pesquisa, foi gerado um novo modelo a partir das relações semânticas propostas pela autora entre os termos da CDU e do AgroTermos, inserindo a notação da CDU original, sendo validadas por um especialista do domínio.

FONTE: Elaborado pelo autor (2021).

A partir do produto gerado em cada estudo descrito no Quadro 14, foram fornecidos insumos metodológicos suficientes para embasar o resultado final desta tese. O Quadro 15 apresenta o diferencial apresentado neste estudo com a junção de metodologias, tais como a compatibilidade de linguagem proposta por Dahlberg (1981) para alcançar a melhoria na recuperação da informação. Foi utilizado o vocabulário da CDU com o vocabulário do tesouro AgroTermos, apresentando o número de coincidências conceituais. Por fim, foram definidas as relações semânticas pelo autor desta tese para a construção do modelo de compatibilidade proposto nos dois SOC, a partir do estudo de MAIA (2018) e dos termos propostos por Maculan (2015).

QUADRO 15 - Insumos para a análise de dados da compatibilização entre o vocabulário da CDU e AgroTermos

Nº	Estrutura classificatória CDU	AgroTermos Maculan (2015)	Relações semânticas entre o AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos Maia (2018)
----	-------------------------------	---------------------------	--	--

FONTE: Elaborado pelo autor (2021). Adaptado de Dahlberg (1981), Maculan (2015) e Maia (2018).

Nesse contexto, o Quadro 15 apresenta os insumos que permitiram compatibilizar os dois SOC, as quatro primeiras colunas correspondem a matriz de compatibilidade conceitual de Dahlberg (1981), que foi adaptada com o número de conceitos a serem compatibilizados do vocabulário da CDU com o vocabulário do AgroTermos. Foi excluída a coluna nome do conceito, por ser representado pelo termo na coluna “AgroTermos”, e por fim, o número de coincidências conceituais foi excluída para o contexto desta tese e será apresentada como análise e descrição dos resultados no capítulo 6. A quarta coluna corresponde as relações semânticas usadas no mapeamento entre os termos do AgroTermos para cada um dos termos do recorte temático proposto por Maculan (2015) extraídas da base da Embrapa. Por fim, a quinta coluna corresponde as relações semânticas entre os dois SOC que foram estabelecidas pelo autor desta tese, a partir da proposta de Maia (2018). A partir da compatibilização dos dois SOC, torna-se possível a construção de um modelo de representação dos conceitos específicos do assunto “Intensificação Agropecuária” no domínio da agricultura.

A segunda fase da terceira etapa da metodologia foi constituída pelo uso dos insumos tecnológicos, recorrendo às seguintes ferramentas: *EdrawMind*²¹, *AgroTermos* e *TemaTres*²². Esses softwares são utilizados para a construção, visualização e extração dos termos da CDU e do tesouro no domínio da agricultura, explicitando as relações semânticas entre os termos sobre a temática “Intensificação Agropecuária”, além de serem utilizados com a finalidade de aplicar o modelo proposto para a recuperação da informação, no ambiente digital.

Dentre as três ferramentas citadas, a princípio e com o objetivo de garantir a visualização da estrutura do modelo proposto, foi utilizada a ferramenta *EdrawMind*. Tal software gerencia as informações dos elementos do modelo e mostra, de modo visual, desde as informações simples até as mais complexas.

A segunda ferramenta utilizada foi o *AgroTermos*, que permite consultar e navegar em uma base de dados na procura de diferentes terminologias (vocabulário controlado) e suas relações semânticas de cada termo, que por sua vez, possui um *Uniform Resource Identifier* (URI) disponibilizada pela Embrapa. Além desse procedimento realizado com o *AgroTermos*, na terceira e quarta etapas da metodologia, utilizaram-se os aportes tecnológicos (planilha do Excel, base de dados do *AgroTermos* e os software *EdrawMind* e *TemaTres*). Estas etapas, facilitaram a coleta e análise dos dados (termos) e seus relacionamentos semânticos, os quais foram retirados da correspondência (*matching*) entre os termos da CDU e do *AgroTermos*.

²¹Disponível em: <https://www.edrawsoft.com/>.

²² Disponível em: <https://sourceforge.net/projects/tematres/>

Após a compatibilização entre os dois SOC, é possível representar um conceito no formato SKOS para o intercâmbio de dados entre os dois sistemas no âmbito da web semântica. Para isso, foi utilizada a terceira ferramenta de software, o *TemaTres*, que é um software construído para a web com o código aberto, cujo objetivo é gerenciar a construção de vocabulários controlados. Desse modo, possibilita estabelecer as relações semânticas entre os termos de diversos instrumentos como: glossários, taxonomias, sistemas de classificação, tesouros, ontologias e vocabulários multilíngues (SOBRINHO; FERREIRA, 2019). Para a construção do vocabulário controlado, é necessária a existência de um padrão de metadados para a recuperação da informação no ambiente digital, onde será aplicado no *TemaTres* o modelo proposto baseado na CDU e no AgroTermos. O *TemaTres* possui uma interface intuitiva para coletar, analisar, publicar e reutilizar dados, cuja a taxonomia do modelo proposto pode ser incluída nas categorias correspondentes às classes da CDU com os descritores e sua notação original compatibilizadas com os descritores do AgroTermos (GALDINO, 2017).

Outra característica do *TemaTres* é que permite a conversão de vocabulários em linguagem natural para vocabulários controlados no âmbito da web semântica, por meio de padrões de metadados. Este software permite a interoperabilidade entre sistemas de informação, bem como o reúso de vocabulários já existentes que estejam disponíveis para serem reutilizados durante a construção de sistemas de organização do conhecimento (SOC), de modo a aproveitar o que já foi desenvolvido e ainda compartilhar o vocabulário controlado construído, permitindo ampliar o conhecimento que está sendo representado (GALDINO, 2017).

O detalhamento deste software será apresentado no subcapítulo 6.3. Esse processo, em linhas gerais, subsidiará a proposta para melhoria do instrumento terminológico da instituição estudada, a Embrapa, cujo contexto institucional é descrito no subcapítulo seguinte.

4.4 Delimitação do ambiente da pesquisa

Este subcapítulo expõe o ambiente de investigação que sustentou o estudo de caso, com a finalidade de concluir a apresentação da metodologia do estudo. Em específico, são apresentadas informações institucionais sobre a Embrapa, como também são detalhados os processos de constituição do recorte temático a ser analisado.

4.4.1 A Embrapa

A Embrapa é uma instituição de pesquisa agropecuária vinculada ao Ministério da Agricultura, Pecuária e Abastecimento (MAPA), cujo propósito é produzir conhecimento e inovação tecnológica a partir de pesquisas na área agropecuária, sobretudo no campo da Intensificação Agropecuária. Fundada em 26 de abril de 1973, as pesquisas desenvolvidas contemplam o contexto brasileiro e internacional e a administração do sistema de pesquisa inclui governos estrangeiros (PENTEADO *et al.*, 2014).

Atualmente, os centros de pesquisa (unidades) da Embrapa estão localizados em todas as regiões do Brasil, além de possuírem laboratórios e/ou escritórios em diversos países como Estados Unidos, países da Europa, China e Coréia do Sul, o que leva o Brasil a ser um dos maiores produtores e exportadores de alimentos no âmbito mundial. Além disso, a Embrapa compõe o Sistema Nacional de Pesquisa Agropecuária (SNPA), também constituído por instituições brasileiras de diferentes segmentos como públicas federais, estaduais, universidades e empresas privadas, que executam as pesquisas de modo cooperativo entre diferentes áreas geográficas e campos do conhecimento científico (EMBRAPA, 2020).

As pesquisas realizadas pela Embrapa para o setor agropecuário implementam o modelo de Pesquisa, Desenvolvimento e Inovação (PD&I), centrado na demanda de tecnologias, produtos, processos, serviços e metodologias. Esse modelo é orientado pela necessidade de se adotar como estratégia o enfoque de sistemas de gestão pela qualidade da pesquisa e inovação com a finalidade de beneficiar a comunidade em geral, principalmente os produtores agrícolas (PIEROZZI JUNIOR; OLIVEIRA; ALENCAR, 2013). Berger *et al.* (2015) descrevem que os programas de PD&I seguem as diretrizes de mudanças de tecnologias, comunicação empresarial e de desenvolvimento institucional, tendo em vista aprimorar a gestão e aumentar a obtenção de resultados mais eficientes entre os grandes desafios nacionais e internacionais de competitividade e sustentabilidade. Nesse sentido, as políticas de pesquisa agropecuária destacaram-se pelo desenvolvimento técnico-científico direcionado à Agricultura e Pecuária Brasileira, atuando em todas as regiões do país (BERGER *et al.*, 2015).

Um dos objetivos da Embrapa é ser uma fonte de informação sobre a Agricultura e suas subáreas, de modo a contribuir para a formulação e o aprimoramento de políticas públicas em áreas relacionadas às demandas da agricultura. O resultado desse propósito é viabilizar, para a sociedade, o acesso à informação e ao conhecimento produzido de forma on-line, levando em conta os desafios a serem enfrentados no futuro (EMBRAPA, 2020). Para atender esses objetivos, a Embrapa possui a unidade Embrapa Informática Agropecuária (EIA), que é uma das 43 unidades da empresa com enfoque no desenvolvimento de soluções de tecnologia da informação e comunicação (TIC) voltadas para o setor agropecuário. Essa

unidade foi criada em 1985 e se coloca a serviço da formulação de políticas públicas para inovação com foco na sustentabilidade da agricultura em diversos aspectos (EMBRAPA, 2020).

A EIA está instalada na cidade de Campinas, estado de São Paulo, com infraestrutura de alto desempenho, capaz de garantir soluções estratégicas tanto para a instituição quanto para as empresas parceiras. A EIA possui um Laboratório Multiusuário de Bioinformática (LMB) e uma Unidade Mista de Pesquisa em Genômica Aplicada a Mudanças Climáticas (UMiP GenClima) e centros avançados para análises de alta complexidade (EMBRAPA INFORMÁTICA, 2018). Nesse sentido, essa unidade aplica métodos, técnicas e ferramentas computacionais, organizados por uma estrutura voltada à Computação Científica e Automação, Sistemas de Informação, Bioinformática e Biologia Computacional, Modelagem Agroambiental e Geotecnologias. Em suma, a EIA estabelece parcerias técnicas com governos, instituições de pesquisa e ensino, empresas públicas e privadas, *startups*, cooperativas, extensão rural, dentre outras empresas, a fim de fomentar o uso e o desenvolvimento de tecnologias emergentes e de inovação com foco na sustentabilidade (EMBRAPA INFORMÁTICA, 2018).

Alinhado ao contexto desta pesquisa, a EIA desenvolveu o tesouro AgroTermos, com o intuito de promover a organização e a representação dos recursos informacionais disponíveis na EMBRAPA, sobretudo no âmbito da web semântica. Para isso, torna-se necessária a utilização de ferramentas e padrões que disponibilizem a informação com a finalidade de minimizar os problemas de ambiguidade da linguagem natural. Assim, com o uso da tecnologia no contexto da web semântica, é possível melhorar a gestão da informação, considerando os problemas gerados com o aumento dos volumes de dados em diversas áreas do conhecimento, inclusive na área da agricultura (EMBRAPA INFORMÁTICA, 2018). Com o uso de recursos tecnológicos avançados, será possível otimizar a utilização de insumos, tendo em vista aumentar o rendimento. Além disso, possibilita o aumento da eficiência e torna os processos de trabalho mais eficientes, tais como o Melhoramento Genético a partir da Biotecnologia e da Bioinformática, estudos sobre Mudanças Climáticas, Geotecnologias para precisão da agricultura e os Sistemas para o Monitoramento e Tomada de Decisão (EMBRAPA INFORMÁTICA, 2018).

4.4.2 O recorte temático e o AgroTermos

A temática da agricultura foi abordada no projeto “Intensificação Agropecuária em Polos de Produção de Soja e Cana-de-Açúcar: Territorialidade, Sustentabilidade e Competitividade” (INTAGRO), projeto esse gerenciado pela Embrapa, instituído no ano de 2008 e finalizado em 2011. Desse modo, o objetivo desse projeto foi garantir a organização e

representação dos processos associados às atividades da Agricultura e da Pecuária (OLIVEIRA *et al.*, 2015). Pelo entendimento de Pierozzi Junior, Oliveira e Souza (2010), o campo da Intensificação Agropecuária é permeado pelo aumento da produtividade agrícola por área territorial, o que acarreta redução no período de pousio e multissafras. Além do mais, fortalece os processos de irrigação, fertilização, uso de animais para carga, maquinários, animais ou vegetais modificados geneticamente, entre outras atividades. Os autores apontam que o campo constitui-se um termo traduzido para o inglês como “*Agricultural Intensification*” e foi retirado do estudo de Boserup (1965), ao descrevê-lo como um processo agrossocioeconômico complexo que se contrapõe com a ideia malthusiana de que o aumento da população leva à escassez de alimentos.

Segundo Boserup (1965), o aumento populacional induz a intensificação dos processos agropecuários, como a modernização ou expansão agrícola sobre áreas ainda não utilizadas. Sendo assim, as áreas que possuem cobertura vegetal ainda nativa e os processos de exploração dessas terras, são necessários para suprir as necessidades do crescimento populacional. A partir da delimitação temática (Intensificação Agropecuária) e do domínio (Agricultura), a taxonomia com os termos sobre o recorte temático foi sugerida pelo Dr. Ivo Pierozzi Júnior, pesquisador em Agroterminologias e Agrossemântica, na Embrapa. Isso se deve pelo fato da amostra ser considerada consistente e representativa para o contexto do mundo real, uma vez que foi sugerida por especialista na temática.

Esta taxonomia foi proposta na tese de Maculan (2015) e validada pelo especialista do domínio com base em trinta termos do recorte temático com aproximadamente seiscentos relacionamentos semânticos entre os termos que denotam cada conceito de forma multidimensional. Além disso, a taxonomia contém remissivas para cada termo representativo de cada conceito, por meio das relações semânticas em relação à temática “Intensificação Agropecuária” com a estrutura proposta por Maculan (2015). A taxonomia proposta pela autora está estruturada em nove facetas, com o agrupamento semântico dos termos que correspondem aos conceitos do recorte “Intensificação Agropecuária”: 1 - Agricultura extensiva; 2 - Agricultura intensiva; 3 - material e métodos; 4 - ambiente; 5 - agronomia; 6 - território e paisagem; 7 - socioeconomia; 8 - espaço e tempo e 9 - instituições.

Em relação ao AgroTermos, ele possui cerca de 240 mil termos vinculados ao contexto agropecuário. Além disso, este tesouro possui uma modelagem conceitual e computacional baseada em SOC e no processamento de linguagem natural (PLN), que viabiliza diversos serviços na web. Assim, esses serviços podem ser aplicados à Gestão da Informação e do Conhecimento, como os processos de indexação, recuperação de informações e interoperabilidade entre sistemas de informação (PIEROZZI JUNIOR *et al.*, 2020).

Neste subcapítulo, apresentou-se a descrição de cada procedimento utilizado nas quatro etapas da metodologia, haja vista propor um modelo para a recuperação da informação com uso de dois instrumentos de controle terminológico: a CDU e o tesouro. Além disso, foi apresentada uma ferramenta de controle terminológico voltada ao âmbito da web semântica, o software *TemaTres*, que servirá de suporte para validar o modelo proposto, considerando a aplicação do modelo, tendo em vista facilitar a recuperação da informação na web e em outros ambientes digitais. A apresentação da proposta do modelo será detalhada no próximo subcapítulo.

5 PROPOSTA DE UM MODELO DE COMPATIBILIDADE DE LINGUAGEM BASEADO NA CDU E AGROTERMOS

"O segredo de um grande negócio consiste em saber algo que mais ninguém sabe".

(Aristóteles Onassis)

Neste capítulo, é apresentado um estudo de caso que visa a proposta de um modelo de compatibilidade de linguagem, tendo por base a classe 631 (Agricultura em geral) da CDU e o termos do AgroTermos. Nesse modelo, a classe 631 é compatibilizada com os termos do tesouro AgroTermos, conforme sugerido por Maculan (2015), com o intuito de representar o recorte temático "Intensificação Agropecuária" nos dois SOC, cujos termos utilizados estão no idioma português (Brasil).

A partir das recomendações de Yin (2010), foram definidas duas etapas a serem seguidas para o desenvolvimento do estudo de caso. A primeira etapa corresponde a execução do estudo de caso. A princípio, foram identificados os trinta termos do recorte temático e suas relações semânticas entre as classes da CDU com os termos do AgroTermos. As relações semânticas definidas no estudo de Maculan (2015) entre os termos que denotam cada conceito da amostra, auxiliaram na compreensão e definição das relações semânticas entre os termos estabelecidos no modelo proposto. Esse modelo foi construído a partir da classe 631 (Agricultura em geral) da CDU e de suas subdivisões e respectivas notações e descritores, sendo que cada classe foi compatibilizada com os termos para representar os conceitos sobre o recorte temático no AgroTermos.

Para isso, foi utilizado o método de compatibilidade conceitual de Dahlberg (1981) adaptado ao vocabulário de cada SOC, considerando a relação semântica de cada termo do recorte temático no AgroTermos e as relações semânticas propostas pelo autor desta tese, de acordo com a proposta de Maia (2018) para a área da BCI. Assim como Dahlberg (1981), durante o processo de compatibilização entre os dois SOC, foram consideradas as características de cada termo e seus relacionamentos semânticos que representam cada um dos trinta conceitos definidos por Maculan (2015). Por fim, foram apresentadas as categorias de cada termo para a compatibilidade de linguagem entre os termos nos dois SOC, com a validação pelo especialista do domínio. Após esses procedimentos, aplicou-se o modelo proposto no software *TemaTres*, tendo em vista a execução da segunda etapa.

O software foi hospedado somente para acesso local e inserção dos dados, de modo a representar o termo escolhido e expor as classes da CDU que tiveram mais de uma ocorrência na compatibilização, tendo como base as relações semânticas entre os termos no AgroTermos que representam essas classes. A partir do uso desse software, será possível verificar e analisar as relações semânticas no formato SKOS, o que permitirá a interpretação dos dados no ambiente digital no âmbito da web semântica. Além disso, será identificada uma

quantidade adequada de propriedades para serem representadas no SKOS com os descritores dos dois SOC e sua respectiva notação da CDU. Com isso, representa-se no SKOS o modelo proposto, inclusive com a *Uniform Resource Identifier* (URI) de cada termo disponibilizado pela Embrapa na web. Esse processo objetiva garantir melhorias na recuperação da informação nos sistemas de informação pertencentes ao domínio da Intensificação Agropecuária.

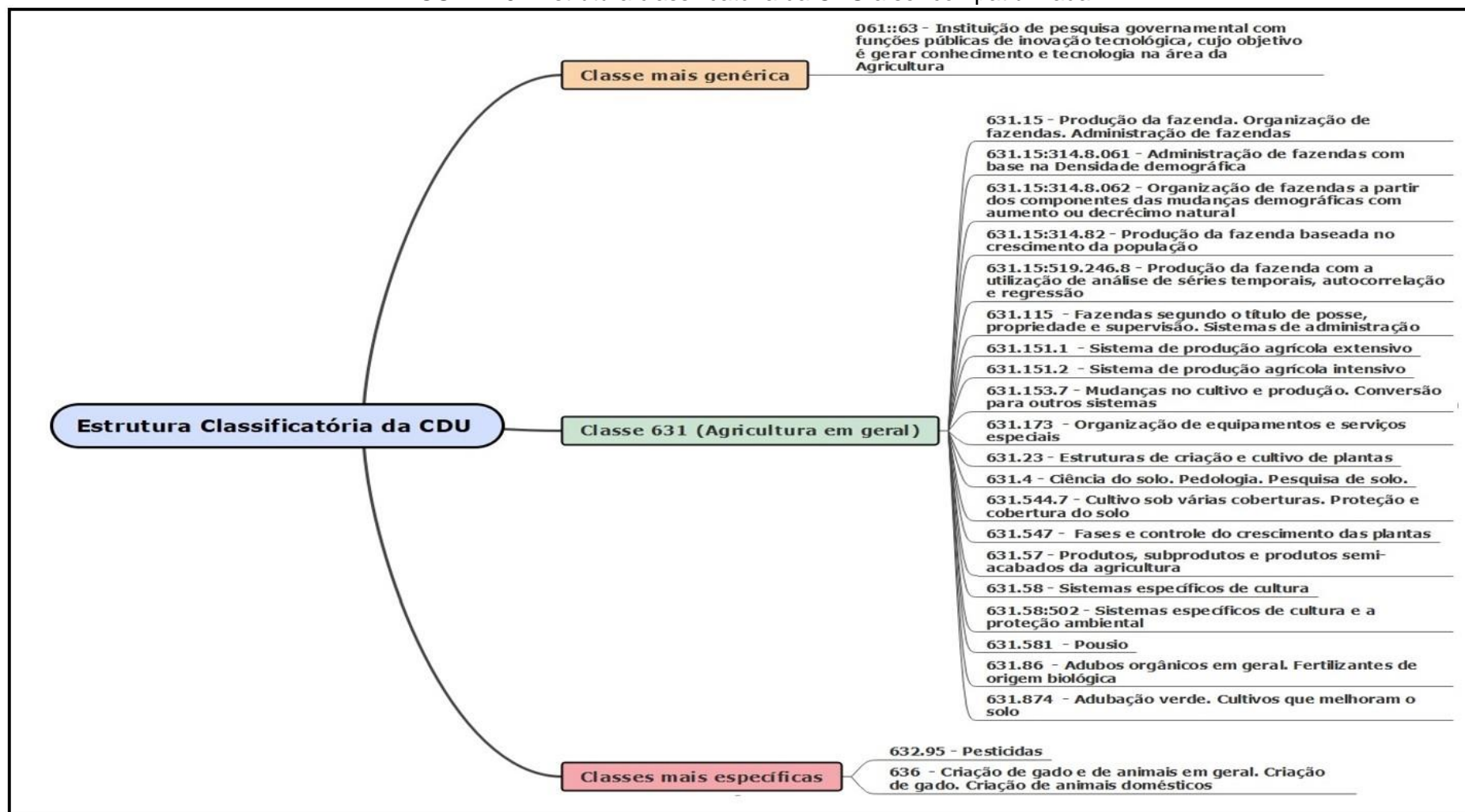
Na segunda etapa, refere-se a análise e descrição dos resultados do estudo de caso. Analisa-se o mapeamento e a compatibilidade de linguagem entre os conceitos nos dois SOC. A compatibilização entre os dois SOC representada no *TemaTres* em formato SKOS levará em consideração os termos que tiveram acima de duas ocorrências para cada classe da CDU, haja vista a grande quantidade de termos e relações semânticas presente no modelo proposto. Logo após, será analisada a relevância da representação dos termos e seus relacionamentos semânticos cujos resultados serão descritos a partir da compatibilidade de linguagem entre os dois SOC, visando à interoperabilidade entre sistemas de informação com bases de dados heterogêneas e o reúso de vocabulários controlados.

A seguir, é aplicado o método para a construção do modelo proposto, tendo como base a classe 631 e suas subdivisões que correspondem à temática “Intensificação Agropecuária” (primeira etapa do estudo de caso). O modelo foi definido seguindo a hierarquia da CDU, ou seja, da classe geral para as específicas, e considerou a notação decimal compatibilizada com os termos correspondentes à temática no *AgroTermos*. Os termos e seus respectivos conceitos utilizados nesse processo estão descritos no Anexo B.

5.1 Compatibilidade de linguagem entre os termos da CDU e *AgroTermos*

O modelo foi composto pelos termos (vocabulário) e notações da CDU, cuja maioria foi retirada da classe 631, a qual representa a amostra de análise desta tese. Em suma, durante a construção do modelo, foi necessária a utilização de outras classes da CDU, ou seja, que não pertenciam à classe 631. Isso se justifica, sobretudo pela estrutura da CDU possuir como base de classificação uma representação canônica, ou seja, são estabelecidas regras para a inclusão de classes e subclasses com a possibilidade de aderência a uma temática não canônica. Desse modo, o motivo da utilização de outras classes fora da amostra deve-se pelas inconsistências semânticas, pois a CDU é um sistema de classificação de assuntos, já o tesouro é um sistema de conceitos. Essas classes foram predefinidas, de modo a serem compatibilizadas com os termos relativos à temática no *AgroTermos*. A Figura 26 esquematiza a estrutura classificatória da CDU para a classe 631, que é complementada com classes gerais e específicas retiradas de outras classes do sistema classificatório.

FIGURA 26 - Estrutura classificatória da CDU a ser compatibilizada

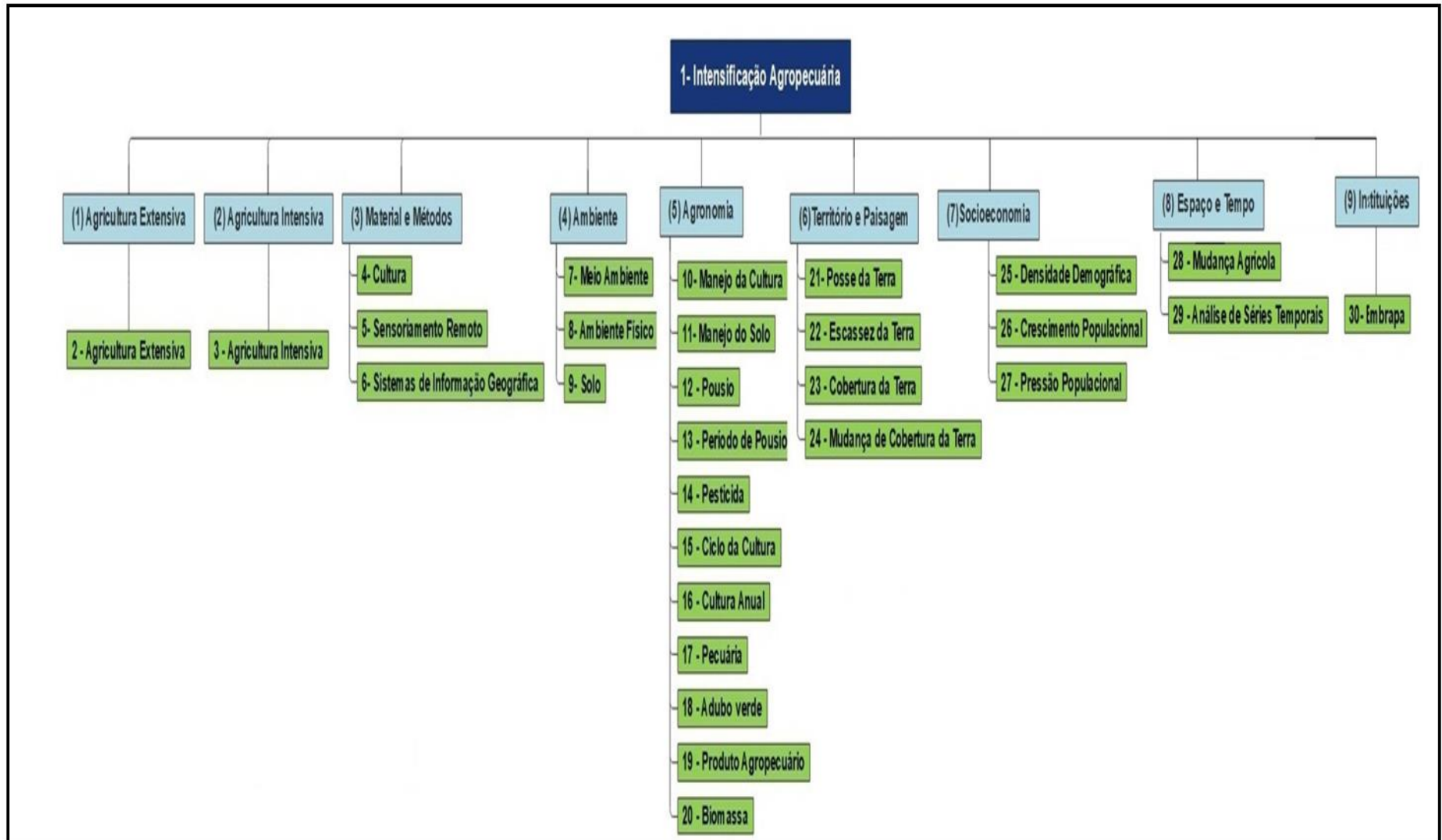


FONTE: Dados da pesquisa (2021).

A estrutura classificatória da CDU que compõe a amostra de análise deste estudo está representada na Figura 26. Observa-se que a classe referente a amostra representada pela classe 631, é formada por 15 subclasses capazes de possibilitar a classificação dos assuntos que permeiam o domínio da agricultura, conforme a estrutura classificatória da CDU. A primeira classe inserida no modelo foi relacionada à “Instituições de pesquisa” que se trata de uma classe mais genérica em relação a classe 631, duas subdivisões auxiliares especiais e três classes que foram combinadas com a classe 631.15 para a representação de assuntos associados à agricultura, tais como: Densidade demográfica, Componentes das mudanças demográficas, Crescimento Populacional, dentre outras. Por fim, a estrutura é complementada por duas classes específicas referentes a Pesticidas e Criação de Gados/Animais.

Após a identificação e o mapeamento das classes que representam o domínio da agricultura na CDU (Figura 26), apresenta-se o modelo proposto que possibilitou o agrupamento dos termos em nove categorias, conforme ilustrado na Figura 27.

FIGURA 27 - Categorias dos termos utilizados no AgroTermos



FONTE: Dados da pesquisa (2021).

O recorte do AgroTermos, mapeado e disposto na Figura 27, refere-se ao que foi proposto na pesquisa de Maculan (2015), considerando o campo da Intensificação Agropecuária. Por meio da identificação de 30 conceitos (Anexo B), presentes no tesouro, eles foram distribuídos em nove categorias. Além disso, foi necessário saber o significado dos termos e suas relações no AgroTermos, por meio do dicionário de termos técnicos da Agricultura Ditech online²³. Em relação às categorias do AgroTermos, a categoria com maior número de conceitos foi a de número 5 (Agronomia), seguida pela de número 6 (Território e Paisagem). A primeira categoria (Agricultura Extensiva), a segunda (Agricultura Intensiva) e a nona (Instituições) foram compostas, cada uma, por apenas um conceito. Cada uma dessas categorias e seus respectivos termos que representam os conceitos, foram compatibilizados com as classes da CDU, a partir do recorte delimitado e descrito na Figura 26.

A partir do processo de compatibilidade, foi possível comparar cada classe e termos da CDU com os conceitos representados pelos termos no AgroTermos e as respectivas relações semânticas existentes, conforme apontado nos quadros 16 ao 38. Assim, a primeira classe corresponde a combinação da classe: “<’ 061 - Organizações e outras formas de cooperação. Inclusive Sociedades. Associações. Congressos. Exposições. Empresas. Instituições de pesquisa. Funções Públicas. Organizações e cooperação governamentais”; com a classe “<’ 63 - “Agricultura. Ciências agrárias e técnicas relacionadas. Silvicultura. Explorações agrícolas. Exploração da vida selvagem”. O auxiliar relacional de “relação composta”, fixa a ordem de apresentação dos termos de acordo com suas características e definição do conceito, sendo indicada no sistema da CDU pelo símbolo “::”. Desse modo, foi possível a expansão na representação dos conceitos, combinando os termos da classe 061 e 63 da CDU, que, a partir dos conceitos, estabelecendo a ordem do assunto complexo, com os números de classe 061::63 da CDU e o auxiliar relacional composto. Assim, a partir de qualquer combinação entre os assuntos compostos ou complexos, não é necessário utilizar todos os descritores presentes nos dois SOC durante a “fusão” dos termos da classe 061 com os da classe 63. No entanto é necessário utilizar um tipo de auxiliar relacional (composto ou simples), sendo que este último será descrito adiante.

A classe 061::63 representa a classe mais genérica do domínio da agricultura para representar o termo “Embrapa” e seu respectivo conceito. A combinação das duas classes, permitiu verificar a flexibilidade da CDU para representar o conceito “Embrapa” com a utilização de um auxiliar relacional para incrementar a representação do conceito, aumentando o nível de especificidade do assunto.

²³ Disponível em: <https://www.ditech.com.br/dicionario/termos-tecnicos/Agricultura/>.

A relação semântica estabelecida durante o processo de compatibilidade entre a CDU e o AgroTermos de acordo com a proposta de Maia (2018) é de instância, como pode ser observado no Quadro 16.

QUADRO 16 - Compatibilidade da classe 061::63 com o termo Embrapa

Nº	Estrutura classificatória da CDU	AgroTermos - Maculan (2015)	Relações semânticas entre AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos - Maia (2018)
1	'<' 061::63 - Instituição de pesquisa governamental com funções públicas de inovação tecnológica, cujo objetivo é gerar conhecimento e tecnologia na área da agricultura	Embrapa	Pesquisa <Related Term> Institutos de Pesquisa	Hierárquica - Hiperônimo - Hipônimo - Instância

FONTE: Dados da pesquisa (2021).

O primeiro termo em análise, conforme apresentado no Quadro 16, corresponde à classe mais abrangente da CDU em relação ao recorte temático e compatibilizada com o termo “Embrapa” no AgroTermos. Esse termo foi inserido no número de classe '<' 061::63 da CDU, sendo que o sinal de '<', de acordo com a matriz de compatibilidade conceitual de Dahlberg (1981), corresponde a classes que são mais genéricas em um domínio, as quais, neste estudo, correspondem à classe 631 (Agricultura em geral). O termo “Embrapa” no tesauro representa uma instância em relação aos termos referente ao número de classe 061::63.

O termo “Pesquisa” nesse tesauro, representado pela sigla *RT*<Related Term> possui uma relação associativa em relação ao termo “Institutos de Pesquisa”. Os dados coletados do termo “Embrapa” e dos demais possibilitam constatar que, a partir do significado semântico dos termos no AgroTermos, há correspondências na CDU.

Ainda com base no Quadro 16, percebe-se que o número de classe 061::63 corresponde a uma classe mais genérica da CDU que representa o termo “Embrapa”, que é o nome de uma “Instituição de pesquisa governamental com funções públicas de inovação tecnológica, cujo objetivo é gerar conhecimento e tecnologia na área da agricultura”. Desse modo, o assunto composto neste exemplo, caso fosse invertido, seria representado por exemplo, “a agricultura e as instituições governamentais de pesquisa”, que significa que o assunto poderia ser “A agricultura em uma instituição de pesquisa do governo” e não que se trata de um assunto sobre “Uma instituição específica governamental realiza pesquisas na área da agricultura”.

Portanto, analisando a representação da classe 061::63, a relação semântica entre os dois vocabulários corresponde a uma relação Hierárquica - Hiperônimo - Hipônimo - Instância. Essa relação semântica entre os dois SOC se justifica devido ao fato do termo “Embrapa” ser um objeto (instância) do número de classe 061::63.

A categoria definida para o termo “Embrapa” corresponde à categoria de número 9, cujo sentido atribuído ao termo corresponde a instituições. Isso evidencia que a generalidade do termo, ou seja, dentre os diversos tipos de “Instituições” existentes, o significado do termo “Pesquisa” está associado a pesquisas em diversas áreas, realizadas por institutos de pesquisa, que por sua vez realiza vários tipos de pesquisa, tais como por área, produtos, tecnologias, entre outras possibilidades.

A próxima classe a ser analisada na estrutura classificatória da CDU é a de número 631.15. O Quadro 17 apresenta o resultado da compatibilidade entre os termos da classe mais geral, dentro do recorte temático da CDU correspondente a “631.15 - Produção da fazenda. Organização de fazendas. Administração de fazendas” e compatibilizado com os termos que denotam os conceitos “Intensificação Agropecuária” e “Manejo da cultura”, pertencentes ao AgroTermos.

QUADRO 17 - Compatibilidade da classe 631.15 com os termos Intensificação agropecuária e Manejo da cultura

Nº	Estrutura classificatória da CDU	AgroTermos Maculan - (2015)	Relações semânticas entre AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos - Maia (2018)
2	631.15 - Produção da fazenda. Organização de fazendas. Administração de fazendas	Intensificação Agropecuária	Exploração agrícola intensiva <Narrow Term> Intensificação	Hierárquica- Hiperônimo - Hipônimo - Inclusão de Classe - Atividade subordinada
		Manejo da Cultura	Sistema de Cultivo <Used For> Sistema de cultura	Hierárquica - Hiperônimo- Hipônimo - Inclusão de Classe - Funcionalmente subordinado

FONTE: Dados da pesquisa (2021).

Pelo que mostra o Quadro 17, a classe 631.15 é mais abrangente na CDU em relação às subclasses da classe 631, com a inclusão dos termos “Intensificação Agropecuária” e “Manejo da Cultura”. Isso aconteceu por não haver termo da CDU para uma representatividade específica ou de equivalência no tesauro analisado. A classe 631.15 possui o termo “Produção da fazenda. Organização de fazendas. Administração de fazendas”, manifestando-se, assim, como uma classe mais específica em relação à classe 061::63 de acordo com a estrutura classificatória da CDU, conforme o Quadro 17.

A partir da relação de hierarquia, o termo “Exploração agrícola intensiva” caracteriza-se como específico em relação ao termo “Intensificação”, pois se refere a sistemas agrícolas que envolvem intensificação de trabalho e/ou de insumos. Isso resulta em maior produção ou produtividade agrícola, sendo representado, no tesauro, pela sigla *NT*<Narrow

Term>. Em seguida, o termo “Sistema de cultivo” possui uma relação de equivalência com o termo “Sistema de cultura”, visto que faz menção aos procedimentos que são aplicados nos cultivos agrícolas, e está representado pela sigla *UF*<Used For>, no AgroTermos.

A classe 631.15 é mais genérica em relação ao termo “Intensificação Agropecuária”. O termo constitui um processo que visa o aumento da produtividade agropecuária e estabelece, entre os dois vocabulários, uma relação Hierárquica - Hiperônimo - Hipônimo - Inclusão de Classe - Atividade subordinada. Esse relacionamento pode ser explicado, considerando que a classe 631.15 possui termos com significados amplos, ou seja, abrangem qualquer tipo de atividade subordinada à produção, organização ou administração de fazenda; por isso o termo “Intensificação Agropecuária” ficou como sendo um termo específico da referida classe.

O termo “Manejo da cultura” faz alusão aos processos de cultivos agrícolas para melhorar as condições de plantio, incluindo os procedimentos de preparação do solo, fertilização, sistemas de cultivo, controle de pragas e doenças, entre outros aspectos. A compatibilidade desses dois vocabulários possibilitou a formação de uma relação Hierárquica-Hiperônimo-Hipônimo - Inclusão de Classe - Funcionalmente subordinado. Essa relação é bastante semelhante ao primeiro termo analisado e se diferencia pela classe 631.15 possuir termos mais amplos em relação a qualquer tipo de função relacionada à produção, organização ou administração de fazenda; com efeito, o termo “Manejo da cultura” também ficou como sendo um termo específico da referida classe.

O termo “Intensificação agropecuária” é mais abrangente em relação ao domínio do recorte temático, sendo similar ao termo “Exploração agrícola intensiva”, motivo pelo qual não foi incluído em nenhuma categoria. Já o termo “Manejo da cultura” foi inserido na categoria de número 5, no AgroTermos, cujo conceito atribuído ao termo corresponde à Agronomia. Constatou-se que esse termo é equivalente aos termos “Sistema de cultivo” e “Sistema de cultura”, que possuem significado relacionado a sistemas de controle de ambientes adequados para o cultivo de um determinado tipo de cultura agrícola.

Logo após, foi realizada a compatibilidade entre os termos da classe mais geral dentro do recorte temático da CDU, evidenciando a característica decimal da CDU em relação à subdivisão auxiliar especial 314.8.061. Portanto, a classe 631.15 foi combinada com a subdivisão auxiliar especial 314.8.061 - Densidade demográfica, para referir-se aos termos que representam o assunto “Administração de fazendas com base na Densidade demográfica”. A combinação de assuntos permitiu representar este assunto, por meio de um auxiliar relacional de relação simples na CDU, que é indicada pelo símbolo “:”, onde a ordem de apresentação dos termos das duas classes podem ser invertidos sem comprometer a definição do conceito. No caso de inverter os termos não afetaria no significado, por exemplo “A Densidade demográfica e a administração de fazendas”. Nos demais exemplos de

combinação de assuntos, foram utilizados apenas auxiliares relacionais simples, por isso não foram exemplificados. Neste exemplo, com a utilização do auxiliar relacional simples foi possível ampliar a representação de um assunto complexo, combinando a classe 631.15 e a subdivisão auxiliar especial 314.8.061 da CDU, sendo então definido pelo número de classe 631.15:314.8.061. Essa compatibilidade entre os dois vocabulários é apresentada no Quadro 18.

QUADRO 18 - Compatibilidade da classe 631.15:314.8.061 com o termo Densidade demográfica

Nº	Estrutura classificatória da CDU	AgroTermos - Maculan (2015)	Relações semânticas entre AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos - Maia (2018)
3	631.15:314.8.061 - Administração de fazendas com base na Densidade demográfica	Densidade demográfica	Densidade demográfica <Used For> Densidade populacional	Equivalência - Sinônimo parcial

FONTE: Dados da pesquisa (2021).

Conforme as informações descritas no Quadro 18, o termo “Densidade demográfica” possui uma representatividade parcial com o número de classe 631.15:314.8.061 da CDU, e estabelece uma relação de equivalência em relação ao termo “Densidade populacional” no AgroTermos.

Pelo Quadro 18, o termo “Densidade demográfica” é um termo equivalente no AgroTermos, que é representado pela sigla *UF*<Used For> em relação ao termo “Densidade populacional”. O termo “Densidade demográfica” é um termo equivalente a “Densidade populacional” pelo significado desse termo, que corresponde à média de indivíduos dividida pela superfície do território em que ocupam, ou seja, o número de pessoas ou outros seres vivos distribuídos por espaço habitado. Nesta análise, o termo “Densidade demográfica” pertencente ao número de classe 631.15:314.8.061 da CDU, manifesta-se como um termo que possui uma relação de equivalência - sinônimo parcial com o AgroTermos. Essa relação se justifica, sobretudo, pelos termos compartilharem o mesmo significado em relação a Densidade demográfica, mas se diferem no aspecto em relação a administração de fazendas.

A categoria do termo “Densidade demográfica” corresponde à categoria de número 7 no AgroTermos, cujo significado que foi atribuído ao termo corresponde à categoria Socioeconomia. Percebe-se que a especificidade do termo, ou seja, dentre os diversos tipos de “Densidade” existentes, o significado do termo “Densidade demográfica” está associado a aspectos econômicos, que impactam, por exemplo, na produtividade, no transporte, no uso da terra, dentre outros aspectos.

Em seguida, analisou-se classe 631.15 em relação a subdivisão auxiliar especial 314.8.062 “Componentes das mudanças demográficas. Inclusive aumento ou decréscimo

natural” da CDU, que foi representada pelo termo que indica o conceito “Pressão populacional” no AgroTermos.

Para a representação do assunto “Organização de fazendas a partir dos componentes de mudanças demográficas com aumento ou decréscimo natural.”, utilizou-se a classe 631.15 combinada com a subdivisão auxiliar especial 314.8.062 da CDU. A combinação de assuntos permitiu representar o conceito “Pressão populacional” do tesauro dentro do contexto da agricultura, utilizando para a combinação de assuntos um auxiliar relacional simples, onde foi representado o assunto complexo com o número de classe 631.15:314.8.062 da CDU.

A seguir, no Quadro 19, é apresentado o resultado dessa compatibilidade de linguagem.

QUADRO 19 - Compatibilidade da classe 631.15:314.8.062 com o termo Pressão populacional

Nº	Estrutura classificatória da CDU	AgroTermos - Maculan (2015)	Relações semânticas entre AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos - Maia (2018)
4	631.15:314.8.062 - Organização de fazendas a partir dos componentes das mudanças demográficas com aumento ou decréscimo natural.	Pressão populacional	Pressão Populacional <Related Term>Densidade demográfica	Hierárquica-Holônimo-Merônimo-Objeto estruturado-Segmento-evento

FONTE: Dados da pesquisa (2021).

De acordo com o Quadro 19, o número de classe 631.15:314.8.062 está no mesmo nível de hierarquia em relação ao número de classe 631.15:314.8.061 da CDU, ou seja são coordenadas, considerando o recorte temático na estrutura classificatória da CDU.

Nesse caso, o termo “Pressão populacional” possui uma relação associativa no AgroTermos em relação ao termo “Densidade demográfica”. Tal relação pode ser justificada pelo termo “Pressão populacional” caracterizar-se como um processo resultante da tensão entre as dinâmicas populacionais e as dinâmicas de uso e cobertura das terras. Dentre outras consequências, esse processo pode contribuir para os danos ambientais, tais como a poluição, perda de biodiversidade, efeitos esses representados pela sigla *RT* <Related Term> em relação ao termo “Densidade demográfica”.

Assim, nota-se que o termo da classe 631.15:314.8.062 possui uma relação hierárquica com o termo “Pressão populacional”, uma vez que, ambos se relacionam a partir da relação Hierárquica - Holônimo - Merônimo-Objeto estruturado-Segmento-evento. Essa relação se estabelece, sobretudo, pelo termo que representa o evento (todo), que é “Pressão populacional”, sendo que pode ter segmentos (partes) que ocorrem em períodos diferentes

para formar o todo, que neste caso, as partes podem ser por exemplo: as queimadas, inundações, entre outros fatores que causam a “Pressão populacional”.

Por fim, o Quadro 19 indica que o termo “Pressão populacional” pertence à categoria de número 7 no AgroTermos, cujo conceito atribuído ao termo corresponde à Socioeconomia, que está associado ao termo “Densidade demográfica”. A Densidade demográfica, nesse contexto, refere-se ao crescimento populacional e à degradação das terras e da paisagem, processo esse que afeta a produtividade agrícola.

A partir da classe 631.15 da CDU combinada com a classe 314.82 - Crescimento da população, verificou-se sua representatividade parcial pelo termo “Crescimento populacional” no AgroTermos. O processo de compatibilidade entre os dois vocabulários é apresentado no Quadro 20.

QUADRO 20 - Compatibilidade da classe 631.15:314.82 com o termo Crescimento populacional

Nº	Estrutura classificatória da CDU	AgroTermos - Maculan (2015)	Relações semânticas entre AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos - Maia (2018)
5	631.15:314.82 - Produção da fazenda baseada no crescimento da população	Crescimento populacional	População Humana <Related Term>Crescimento populacional	Equivalência - Sinônimo parcial

FONTE: Dados da pesquisa (2021).

Em relação ao Quadro 20, percebe-se que o termo “Produção da fazenda baseada no crescimento da população”, representada pelo número de classe 631.15:314.82 se apresenta no mesmo nível em relação à classe 631.15:314.8.062 da CDU, conforme o recorte temático na estrutura classificatória da CDU. Os termos da classe 631.15:314.82 possuem a mesma representatividade de equivalência no AgroTermos.

O resultado da compatibilidade indica que a classe 631.15:314.82 possui uma relação associativa no tesouro. Por esse motivo, o termo “População Humana” é um termo relacionado no AgroTermos, que é representado pela sigla *RT*<Related Term> em relação ao termo “Crescimento populacional”. Isso se justifica pelo significado atribuído ao termo “Crescimento populacional”, que está relacionado ao aumento do número de indivíduos em um determinado espaço. Já o termo “População humana” é definido como o conjunto de indivíduos da espécie humana que vive em uma mesma área.

Nesse sentido, o termo “Produção da fazenda baseada no crescimento da população” da classe 631.15:314.82 da CDU possui uma relação de equivalência - sinônimo parcial com o termo “Crescimento populacional” no AgroTermos. Assim, os termos estabelecem uma relação semântica, ou seja, eles possuem significados idênticos não em sua totalidade se diferenciando em alguns aspectos, sendo que ambos correspondem ao

contexto da agricultura, mas se diferem pelo termo “Produção de fazenda” da CDU, que pode aumentar a partir do crescimento do número de indivíduos.

No AgroTermos, a categoria do termo “Crescimento populacional” corresponde à categoria de número 7, cujo significado que foi atribuído ao termo corresponde à categoria Socioeconomia. Percebe-se que há relação de associação desse termo com o termo “População humana”, ou seja, o termo “População humana” é o tipo de população e o significado do termo “Crescimento populacional” está associado a aspectos relacionados com o motivo do crescimento da população. Diferentes aspectos podem levar ao crescimento da população, por exemplo, as taxas de mortalidade e natalidade que justificam o aumento ou diminuição da população em um determinado local.

Logo depois da classe 631.15:314.82 da CDU, verificou-se a classe 631.15 foi combinada com a classe 519.246.8 - Análise das séries temporais. Inclusive autocorrelação. Regressão. A classe 631.15:519.246.8 foi compatibilizada com o termo “Análise de séries temporais” existente no AgroTermos. A compatibilidade de linguagem entre os SOC é apresentada no Quadro 21.

QUADRO 21 - Compatibilidade da classe 631.15:519.246.8 com o termo Análise de Séries Temporais

Nº	Estrutura classificatória da CDU	AgroTermos - Maculan (2015)	Relações semânticas entre AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos - Maia (2018)
6	631.15:519.246.8 - Produção da fazenda com a utilização de análise das séries temporais, autocorrelação e regressão	Análise de séries temporais	Método Estatístico<Used For> Análise Estatística	Equivalência - Sinônimo parcial

FONTE: Dados da pesquisa (2021).

O Quadro 21 evidencia que a classe 631.15:519.246.8 com o termo “Produção da fazenda com a utilização de análise das séries temporais, autocorrelação e regressão.”, que refere-se a uma classe coordenada em relação à classe 631.15:314.82 da CDU. Por esse motivo, esse termo foi inserido na classe 631.15:519.246.8 da CDU, que possui uma representatividade parcial em relação ao do termo “Análise de séries temporais” no AgroTermos.

O termo “Método Estatístico” é um termo equivalente em relação ao termo “Análise Estatística”, sendo representado no tesauro pela sigla UF<Used For>. Assim, o termo “Método Estatístico” possui significado equivalente, ou seja, se refere a análise de informações estatísticas.

Os termos da classe 631.15:519.246.8 compatibilizada com o termo “Análise de séries temporais”, possui uma relação de equivalência - sinônimo parcial com o AgroTermos,

que se justifica, com base nos termos que possuem vários sentidos, sendo iguais em muitos significados, mas se diferenciam em algum aspecto. No caso analisado, a diferença está na inclusão dos termos autocorrelação e regressão. O primeiro termo mede o grau de correlação entre as observações em um determinado instante, a partir de uma série temporal, e o segundo é um método estatístico para a modelagem realizada em um conjunto de variáveis referentes a observações que acontecem de modo independente.

A categoria do termo “Análise de séries temporais” corresponde à categoria de número 8, no AgroTermos, cujo significado faz referência a Espaço e Tempo. Tal termo é representado, conceitualmente, por um conjunto de observações que são realizadas de maneira sequencial ao longo do tempo.

A próxima classe da hierarquia da CDU é a “631.115 - Fazendas segundo o título de posse, propriedade e supervisão. Sistemas de administração”, que está associada aos termos que indicam os conceitos “Posse da terra” e “Escassez da terra”, compatibilizados no AgroTermos. O Quadro 22 expõe os resultados dessa compatibilidade.

QUADRO 22 - Compatibilidade da classe 631.115 com os termos Posse da terra e Escassez da terra

Nº	Estrutura classificatória da CDU	AgroTermos - Maculan (2015)	Relações semânticas entre AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos - Maia (2018)
7	631.115 - Fazendas segundo o título de posse, propriedade e supervisão. Sistemas de administração	Posse da terra Escassez da terra	Posse da terra <Narrow Term> Propriedade da terra Planejamento da exploração agrícola <Related Term> Administração agrícola Posse da terra <Related Term> Economia	Equivalência - Sinônimo parcial Associativa - Causal - Agente-ação

FONTE: Dados da pesquisa (2021).

A classe 631.115 com o termo “Fazendas segundo o título de posse, propriedade e supervisão e sistemas de administração” descrita no Quadro 22 constitui uma classe mais específica em relação à classe 631.15 da CDU. O termo “Posse da terra” no tesouro, possui uma representatividade parcial em relação aos termos da classe 631.115 e o termo “Escassez da terra”, não possui um termo da CDU para uma representatividade específica no AgroTermos.

De acordo com as representações mencionadas no Quadro 22, identifica-se uma relação hierárquica entre o termo “Posse da terra” e o termo “Propriedade da terra”, manifestada mediante uma relação de propriedade ou posse entre o proprietário e a terra que lhe pertence, cujo proprietário detém os direitos sobre ela e pode atribuir-lhe o uso que

desejar. Essa relação é representada no tesauro pela sigla *NT* <Narrow Term>. Já o termo “Planejamento da exploração agrícola” possui uma relação associativa com o termo “Administração agrícola”, assim como o termo “Posse da terra”, que é um termo relacionado com o termo “Economia”. Isso se explica com base nos aspectos de planejamento e administração para o uso da terra, e pelo fato de que eles afetam questões econômicas, por causa da exploração e pela redução de quantidade de terra disponível para o uso. Tais relações estão definidas pela sigla *RT* <Related Term>.

A classe 631.115 da CDU formaliza uma relação de equivalência - sinônimo parcial com o termo “Posse da terra”, no tesauro analisado pois, os termos possuem vários sentidos, com vários significados iguais, mas que se diferenciam em algum aspecto. Isso se justifica, visto que nos dois SOC, os termos fazem referência aos direitos adquiridos, a partir de diferentes tipos de regimes de posse, e envolvem sistemas de administração tais como: arrendamento, desapropriação, usucapião e reforma agrária, mas se diferenciam pelo termo fazendas, que é uma propriedade rural que pode ter vários tipos de cultivos. O termo “Escassez da terra” estabelece uma relação Associativa - Causal - Agente - ação, que ocorre quando um sujeito (agente) causa uma ação ou efeito, ou seja, o sujeito pratica uma ação, por exemplo, a indisponibilidade para a posse ou qualquer tipo de uso agropecuário. Nesse processo, podem ser envolvidos diferentes fatores, tais como: crescimento populacional, zona urbana, políticas públicas, reforma agrária e cultura de alimentos; por conseguinte, tais fatores podem causar a redução do território (em quantidade) para a produção agrícola.

O termo “Posse da terra” e o termo “Escassez da terra” estão na categoria de número 6, no AgroTermos, e fazem alusão à propriedade e à redução da quantidade de terra em uma determinada área. Portanto, os termos correspondem a aspectos relacionados à posse e à redução do território disponível para o uso; por isso foram inseridos em Território e Paisagem.

Seguindo a hierarquia da CDU, a próxima classe é a “631.151.1 - Sistema de produção agrícola intensivo”. Verifica-se que essa classe estabelece relações com os termos designados ao conceito “Agricultura extensiva” ao ser compatibilizada com o AgroTermos. A compatibilidade entre os dois SOC é mostrada no Quadro 23.

QUADRO 23 - Compatibilidade da classe 631.151.1 com o termo Agricultura extensiva

Nº	Estrutura classificatória da CDU	AgroTermos - Maculan (2015)	Relações semânticas entre AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos - Maia (2018)
8	631.151.1 - Sistema de produção agrícola extensivo	Agricultura Extensiva	Agricultura extensiva <Used For> Produção extensiva	Equivalência - Sinônimo próximo

FONTE: Dados da pesquisa (2021).

O termo “Sistema de produção agrícola extensivo” da classe 631.151.1, de acordo com o Quadro 23, manifesta-se como mais específico em relação à classe 631.115 da CDU. Por sua vez, o termo “Agricultura extensiva” da classe 631.151.1 possui a mesma representatividade específica ou de equivalência no AgroTermos, de acordo com as demonstrações do Quadro 23.

Desse modo, o termo “Agricultura extensiva” é um termo equivalente no AgroTermos e está representado pela sigla *UP* <Used For> em relação ao termo “Produção extensiva”. Justifica-se essa ocorrência, tendo em vista que ambos os termos possuem o mesmo significado semântico, visto que o conceito deles representa um tipo de sistema de produção agrícola que utiliza pouco insumo tecnológico. Sendo assim, o conceito tem um alto grau de dependência quanto à incorporação de áreas naturais para exploração agrícola.

O termo “Sistema de produção agrícola extensivo” da classe da CDU, possui uma relação de equivalência - sinônimo próximo com o termo “Agricultura extensiva” no tesouro. Tal constatação deve-se ao fato de os termos da classe 631.151.1 não possuírem um sentido idêntico, porém possuem significados que se aproximam relação ao termo do tesouro. Logo, entende-se que o termo “sistema” presente no termo da classe da CDU, está relacionado com a organização da agricultura extensiva. Já a “Agricultura extensiva” pode ser realizada com ou sem um sistema. Em linhas gerais, a agricultura extensiva possui baixo investimento, por conseguinte, baixa produtividade, associada, por exemplo, a atividades pastoris, pastagens e a criações de peixes, sendo essa a principal fonte alimentar do agricultor. Além disso, no âmbito da baixa produtividade, são utilizadas instalações já existentes, sem se preocupar com técnicas para aumentar a área cultivada.

Por fim, quanto à categoria do termo “Agricultura extensiva”, corresponde à categoria de número 1, no AgroTermos. Essa constatação evidencia que o termo está relacionado a aspectos pertencentes ao contexto da extensificação da agricultura, devido ao uso de técnicas rudimentares e à baixa produtividade.

A classe “631.151.2 - Sistema de produção agrícola intensivo” é representada pelo termo que representa o conceito “Agricultura intensiva”, compatibilizada no AgroTermos. O resultado da compatibilidade entre as linguagens nos dois SOC é apresentado no Quadro 24.

QUADRO 24 - Compatibilidade da classe 631.151.2 com o termo Agricultura intensiva

Nº	Estrutura classificatória da CDU	AgroTermos - Maculan (2015)	Relações semânticas entre AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos - Maia (2018)
9	631.151.2 - Sistema de produção agrícola intensivo	Agricultura Intensiva	Agricultura intensiva <Used For> Produção intensiva	Equivalência - Sinônimo próximo

FONTE: Dados da pesquisa (2021).

O termo “Sistema de produção agrícola extensivo” da classe 631.151.2, de acordo com o Quadro 24, constitui uma classe coordenada em relação à classe 631.151.1 da CDU. Constatou-se, a partir do Quadro 24, que a classe 631.151.2 possui a mesma representatividade no AgroTermos; por esse motivo, o termo “Agricultura intensiva” está inserido na classe 631.151.1 da CDU, que possui uma relação de equivalência, no tesauro.

Nesse sentido, o termo “Agricultura intensiva” é um termo equivalente no AgroTermos e está representado pela sigla *UP* <Used For> em relação ao termo “Produção intensiva”. Isso ocorre porque os termos possuem significados iguais nos dois SOC em questão, ou seja, fazem referência a um conceito em comum. A Produção intensiva é conceituada como um tipo de sistema agrícola que envolve intensificação de trabalho e/ou de insumos, além do aumento da produção agrícola, podendo incluir, também, nesse processo, atividades realizadas em pequenos espaços, haja vista o cultivo de plantas ou a criação de animais.

A classe da CDU 631.151.2 que corresponde ao termo “Sistema de produção agrícola intensivo”, que estabelece uma relação de equivalência - sinônimo próximo com o termo “Agricultura intensiva”, no tesauro. O motivo dessa relação é comprovado pelo termo da classe 631.151.2, visto que não determina o mesmo sentido, no entanto, compartilha um ou mais significados em relação ao termo no tesauro. Isso se deve pelo termo “sistema” presente no termo da classe da CDU está relacionado com a organização da agricultura intensiva e o termo “Agricultura intensiva” pode ser realizada com ou sem um sistema. Portanto, os termos representados nos dois sistemas fazem referência ao aumento do cultivo, causando impacto no ecossistema, por exemplo, nos casos de ocorrer erosão ou despejamento de nutrientes químicos em reservatórios de água subterrânea, nascentes e rios. Para esse tipo de produção agrícola, utilizam-se máquinas, fertilizantes e pesticidas, de modo a reduzir a fertilidade e a capacidade de produção dos solos. Isso pode colocar em risco as fontes de água, quando se utilizam insumos e combustíveis que precisam dessas fontes para o funcionamento, mesmo que essas atividades sejam realizadas em espaços reduzidos para o cultivo ou para a criação de plantas ou animais.

No AgroTermos, a categoria 2 corresponde ao termo “Agricultura intensiva”, definida como a intensificação de trabalho relacionada à grande produção agrícola referente a insumos de cultivo e de pecuária. Além disso, esse é um processo que recorre ao uso de tecnologias de ponta para cultivos relacionados à Agropecuária.

A próxima classe da CDU, seguindo a estrutura classificatória do sistema, é a classe “631.153.7 - Mudanças no cultivo e produção. Conversão para outros sistemas” que está relacionada com o conceito indicado pelo termo “Mudança agrícola”, sendo compatibilizada, portanto, no AgroTermos. A compatibilidade do vocabulário nos dois SOC é apresentada no Quadro 25.

QUADRO 25 - Compatibilidade da classe 631.153.7 com o termo Mudança Agrícola

Nº	Estrutura classificatória da CDU	AgroTermos - Maculan (2015)	Relações semânticas entre AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos - Maia (2018)
10	631.153.7 - Mudanças no cultivo e produção. Conversão para outros sistemas	Mudança agrícola	Ordenamento do território <Related Term> Mudança no uso da terra	Equivalência - Sinônimo absoluto

FONTE: Dados da pesquisa (2021).

Pelo Quadro 25, nota-se que o termo “Mudanças no cultivo e produção. Conversão para outros sistemas” da classe 631.153.7 é mais específico em relação à classe 631.151.2 da CDU. O termo da classe 631.153.7 da CDU possui a mesma representatividade no AgroTermos, firmando uma relação associativa no tesouro, conforme dados registrados no Quadro 25.

Quanto ao relacionamento semântico no AgroTermos, o termo “Mudança no uso da Terra” é associado ao termo “Ordenamento do território”, pois engloba alterações diversas de cunho espaço-temporais relativas às atividades agrícolas e decorrentes da dinâmica de uso e cobertura das terras. O termo “Ordenamento do território” é caracterizado pela gestão das ocupações territoriais rurais para assegurar a preservação de recursos limitados em função da exploração antrópica, que é representada pela sigla *RT*<Narrow Term> em relação ao termo “Mudança no uso da terra”.

Portanto, o termo “Mudanças no cultivo e produção. Conversão para outros sistemas” da classe 631.153.7, possui uma relação de equivalência - sinônimo absoluto com o termo “Mudança agrícola”, no AgroTermos. Isso se explica, sobretudo, pela igualdade semântica dos termos, ou seja, eles possuem significados iguais em relação ao mesmo conceito em sua totalidade. Nesse contexto, os termos vinculam-se entre si, tendo em vista as semelhanças que os envolvem, que são: “mudanças agrícolas por alterações espaço-temporais das atividades desempenhadas” e “conversão de sistemas agrícolas” possuem o mesmo significado.

No AgroTermos, a categoria do termo “Mudança agrícola” corresponde à categoria de número 8, cujo significado atribuído ao termo se associa à categoria Espaço e Tempo. Percebe-se a equivalência desse termo, visto que o significado decorre dos aspectos espaço-temporais a envolver fatores múltiplos como a mecanização e as políticas de desenvolvimento rural.

A próxima classe da estrutura classificatória da CDU a ser analisada é a classe “631.173 Organização de equipamentos e serviços especiais”, que, no AgroTermos, foi associada aos termos “Sistema de Informação Geográfica” e “Sensoriamento remoto”. O Quadro 26 elucida as informações referentes a essa compatibilidade.

QUADRO 26 - Compatibilidade da classe 631.173 com os termos Meio ambiente e Ambiente físico

Nº	Estrutura classificatória da CDU	AgroTermos - Maculan (2015)	Relações semânticas entre AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos - Maia (2018)
11	631.173 - Organização de equipamentos e serviços especiais	Sistema de Informação Geográfica	Geoinformação <Narrow Term> geotecnologia Geoinformação <Narrow Term> sensoriamento remoto Geoinformação <Narrow Term> Geoprocessamento Geoinformação <Narrow Term> Georeferenciamento	Hierárquica - Hiperônimo - Hipônimo - Inclusão de Classe - Funcionalmente subordinado
		Sensoriamento remoto	Geoinformação <Narrow Term> Geoprocessamento Geoinformação <Narrow Term> Georeferenciamento Geoinformação <Narrow Term> Sensoriamento remoto	Hierárquica - Hiperônimo - Hipônimo - Inclusão de Classe - Funcionalmente subordinado

FONTE: Dados da pesquisa (2021).

Pela descrição do Quadro 26, ao mostrar a classe 631.173 referente ao termo “Organização de equipamentos e serviços especiais” da CDU, nota-se que essa é uma classe mais específica em relação à classe 631.153.7. O termo da classe 631.173 não possui representatividade específica ou equivalente no AgroTermos, conforme evidenciado no Quadro 26.

O termo “Sistema de informação geográfica” ou “Geoinformação” é um termo genérico em relação aos termos “Geotecnologia”, “Geoinformação”, “Geoprocessamento”, “Sensoriamento remoto” e “Georeferenciamento”, que representa tecnologias ou sistemas de informação para geoprocessamento, sensoriamento remoto e georeferenciamento, sendo representado pela sigla *NT*<Narrow Term>. Já o “Sensoriamento remoto” é um termo mais amplo em relação ao termo “Geoinformação”, visto que representa qualquer sistema de informação georeferenciado. Nesse caso, há correspondência no AgroTermos, mediante o estabelecimento de uma relação de hierarquia representada pela sigla *NT*<Narrow Term>.

A classe 631.173 da CDU possui uma relação de hierarquia em relação aos termos “Sistemas de informação geográfica” e “Sensoriamento remoto” do tipo Hierárquica -

Hiperônimo-Hipônimo - Inclusão de Classe - Funcionalmente subordinado. Essa ocorrência pode ser explicada, pois, os termos em questão constituem objetos que se caracterizam por suas funções. Nesse caso, as funções dos termos denotam tecnologias utilizadas para a obtenção de imagens e dados da superfície terrestre por meio da captação e do registro da energia refletida/emitida pela superfície estudada. Além disso, a integração do sensoriamento remoto com os Sistemas de Informação Geográfica (SIG) torna-se uma importante ferramenta para detectar a magnitude das mudanças geográficas e dos fenômenos agrícolas, de modo a colaborar para melhorias das perspectivas ambientais e agrícolas sustentáveis.

Os termos “Sistemas de informação geográfica” e “Sensoriamento remoto” foram inseridos na categoria de número 3, no AgroTermos, cujo conceito atribuído a esses termos corresponde a Material e Métodos. Nesse sentido, considera-se que os termos analisados estão representando, conceitualmente, métodos, técnicas e ferramentas diversos, tais como: mapas, radares, processamento de imagens, satélites de observação da terra, sensores, GPS, entre outros.

A classe seguinte da estrutura classificatória da CDU é “631.23 - Estruturas de criação e cultivo de plantas”, representada pelo termo que indica o conceito “Cultura”, no AgroTermos. No Quadro 27, apresenta-se o resultado do processo de compatibilidade de linguagem para essa classe e seus elementos.

QUADRO 27 - Compatibilidade da classe 631.23 com o termo Cultura

Nº	Estrutura classificatória da CDU	AgroTermos - Maculan (2015)	Relações semânticas entre AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos - Maia (2018)
12	631.23 - Estruturas de criação e cultivo de plantas	Cultura	produção vegetal <Used For> cultura	Hierárquica - Hiperônimo-Hipônimo - Inclusão de Classe - Atividade subordinada

FONTE: Dados da pesquisa (2021).

De acordo com o Quadro 27, a classe 631.23 correspondente ao termo “Estruturas de criação e cultivo de plantas” apresenta-se como uma classe mais específica em relação à classe 631.173 da CDU. As evidências do Quadro 27 indicam que o termo na classe 631.23 não possui uma representatividade com os termos do tesouro, visto que não foi identificada correspondência entre os termos.

Nessa compatibilização, o termo “Cultura” estabelece uma relação de equivalência e é representado pela sigla *UF* <Used For>, no tesouro, em relação ao termo “produção vegetal”. Esse termo faz alusão ao cultivo de qualquer tipo de planta, cultivo produzido pelo homem ou não, cujos produtos foram ou não colhidos em qualquer um dos

estágios de crescimento. Nesse sentido, as plantas que não foram cultivadas, mas cujos produtos são colhidos, não são consideradas culturas e a mesma situação se aplica ao contrário quando as plantas cultivadas não tiverem seus frutos colhidos.

Dessa maneira, o termo “Estruturas de criação e cultivo de plantas” da classe 631.23 possui uma relação hierárquica do tipo “Hierárquica - Hiperônimo - Hipônimo - Inclusão de Classe - Atividade subordinada” em relação ao termo “Cultura”. Essa relação se estabelece, sobretudo, com base nas atividades envolvidas entre os objetos. Portanto, a atividade “Cultura” está subordinada à estrutura de criação de plantas, tais como: cana de açúcar, café, soja, entre outras. Desse modo, para cada cultura, é utilizada uma estrutura para cultivar determinada espécie de planta, por exemplo: cultivo em céu aberto, sob telado e em estufa. O cultivo em céu aberto não possui nenhum tipo de proteção, pois, a planta permanece, durante o ciclo de desenvolvimento, totalmente a céu aberto, sendo mais utilizado em espécies menos exigentes e cujas condições ambientais são favoráveis. Já o cultivo sob telado faz o controle parcial de luminosidade por meio de cobertura com telas que proporcionam sombreamento de acordo com cada tipo de espécie. Por fim, o cultivo de plantas em estufa proporciona condições microclimáticas mais apropriadas, ou seja, permite controlar o ambiente interno de modo mais adequado, interferindo na luminosidade, ventilação, temperatura e na umidade.

No AgroTermos, o termo “Cultura” pertence à categoria de número 3, que se refere ao conceito cultura, atribuído à categoria Material e Métodos. Nesse sentido, Material e Métodos refere-se ao tipo de material utilizado (adubo, fertilizante, dentre outros) e o método aplicado para o cultivo diz respeito ao plantio direto por meio de sementes, técnicas de irrigação e agricultura de precisão com o uso de tecnologias, tais como SIG, sensoriamento remoto, entre outros.

Ao analisar os termos do tesouro, percebeu-se incompatibilidade deles com os vocabulários da CDU. O termo “Manejo do solo”, no AgroTermos, pode ser um manejo físico/mecânico ou manejo químico/biológico (relação polihierárquica). Diante dessa situação, considerando a estrutura classificatória da CDU, esse termo deveria estar em duas classes diferentes do sistema, o que não pode ocorrer, considerando a natureza analítico-sintética e hierárquica da CDU. Além disso, a organização do vocabulário da CDU segue o princípio da exclusão mútua para a organização em facetadas (classes/categorias), ou seja, cada termo que representa um determinado conceito deve estar em uma única categoria.

Para que o vocabulário da CDU e o vocabulário do AgroTermos sejam compatíveis em relação ao termo “Manejo do solo”, deve-se alterar a estrutura de um dos SOC. Para alterar a estrutura da CDU, sugere-se agrupar as classes “631.31 Máquinas e implementos para cultivo do solo. Máquinas e implementos auxiliares” e “631.8 Fertilizantes, adubos e enriquecimento do solo. Promoção do crescimento. Estimulantes vegetais”, de modo a

agrupar os conceitos de máquinas e adubos. Outra opção seria a criação de dois novos termos no tesouro AgroTermos: “Manejo físico do solo” e “Manejo químico do solo”. O primeiro químico do solo.

O termo “Manejo do solo” está inserido na categoria de número 5 do AgroTermos, cujo conceito atribuído ao termo corresponde à Agronomia, que se refere aos métodos, técnicas, ferramentas e tecnologias voltados ao preparo do solo, haja vista melhorar a produção de culturas. Tal melhoria deve-se a aspectos econômico-sociais, visando à conservação do ambiente, diminuição do gasto de energia e/ou recursos naturais para preservação das condições ambientais que são mutáveis. A partir desse resultado, decidiu-se não incluir o termo em análise no modelo proposto baseado na CDU e no AgroTermos (produto final desta tese). Isso porque a alteração da estrutura original ou reengenharia desses instrumentos está sob responsabilidade das instituições que os mantêm.

Em seguida, partiu-se para a análise à classe “631.4 - Ciência do solo. Pedologia. Pesquisa de solo” relativa ao termo “Solo” que denota o mesmo conceito ao ser compatibilizado no AgroTermos. A compatibilidade de linguagem para esses termos é apresentada no Quadro 28.

QUADRO 28 - Compatibilidade da classe 631.4 com o termo Solo

Nº	Estrutura classificatória da CDU	AgroTermos - Maculan (2015)	Relações semânticas entre AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos - Maia (2018)
13	631.4 - Ciência do solo. Pedologia. Pesquisa de solo	Solo	Solo <Related Term> ciência do solo	Associativa - Disciplina - fenômeno

FONTE: Dados da pesquisa (2021).

O Quadro 28 refere-se à classe 631.4 que contém o termo “Ciência do solo. Pedologia. Pesquisa de solo”, sendo uma classe mais específica em relação à classe 621.23 da CDU. De acordo com o Quadro 28, o termo “Solo” foi inserido na classe 631.4 da CDU, possibilitando uma relação associativa no AgroTermos.

O termo “Solo” está relacionado ao termo “Ciência do solo”, sendo representado no tesouro pela sigla *RT* <Related Term>. Assim, o termo “Solo” é conceitualmente designado como cobertura da terra, que recobre as superfícies não inundadas do planeta onde são cultivadas as plantações ou edificações de um modo geral.

O termo “Ciência do solo. Pedologia. Pesquisa de solo” da classe 631.4 representa uma relação Associativa-Disciplina- fenômeno com o termo “Solo”, no AgroTermos. Isso se justifica pelo fato do termo “Solo”, no AgroTermos, ser o componente ou parte do objeto estudado na disciplina representada pelos termos “Ciência do solo. Pedologia. Pesquisa do solo” da classe 631.4 da CDU. No contexto analisado, a diferença se concretiza pela classe do sistema classificatório envolver o estudo do solo como recurso natural da superfície

terrestre, incluindo a formação do solo, sua classificação, propriedades físicas, químicas, biológicas e de fertilização para o uso dos solos.

O termo “Solo” corresponde à categoria de número 4, que possui aspectos relacionados a Ambiente. Isso se justifica, pois, o conceito do termo diz respeito aos seres vivos que habitam a superfície local de um determinado ambiente, assim como os processos não naturais, de melhoramento e de manejo de solo.

A próxima classe é a “631.544.7 - Cultivo sob várias coberturas. Proteção e cobertura do solo” que está relacionada aos termos que designam os conceitos “Cobertura da terra e mudança de cobertura”, compatibilizados no AgroTermos. A compatibilidade entre os dois vocabulários para essa classe é apresentada no Quadro 29.

QUADRO 29 - Compatibilidade da classe 631.544.7 com os termos Cobertura da terra e Mudança de cobertura da terra

Nº	Estrutura classificatória da CDU	AgroTermos - Maculan (2015)	Relações semânticas entre AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos - Maia (2018)
14	631.544.7 - Cultivo sob várias coberturas. Proteção e cobertura do solo	Cobertura da terra	Cobertura do Solo <influences> Uso da Terra	Equivalência-Sinônimo próximo
		Mudança de cobertura da terra	Cobertura do Solo <influences> Uso da Terra	Associativa-Causal - Agente-objeto

FONTE: Dados da pesquisa (2021).

O termo “Cultivo sob várias coberturas. Proteção e cobertura do solo” da classe 631.544.7, de acordo com o Quadro 29, é mais específico em relação à classe 631.4. Percebe-se, pelo Quadro 29, que o termo da classe 631.544.7 não possui a mesma representatividade específica ou de equivalência no AgroTermos.

Pelas demonstrações do Quadro anterior, o termo “Cobertura da terra” possui uma relação associativa com o termo “Uso da Terra” no AgroTermos, sendo representado pela sigla <influences>. Portanto, o termo que representa o conceito “Cobertura da terra” influencia no uso da terra, visto que a cobertura da terra é um elemento da superfície terrestre correspondente à composição hidrogeológica, biológica, natural ou antrópica, que influencia a paisagem, o controle e o uso da terra. Já o termo “Mudança de cobertura da terra” também estabelece uma relação associativa <influences>, mas com propósitos diferentes, pois, se trata de um processo de modificação da cobertura do solo que é influenciado por diversos fatores, que podem ser de modo natural (chuva, vento, relâmpago, esgotamento de recursos naturais, dentre outros) ou não (a falta de vegetação ou desmatamento). Portanto, a mudança

em diferentes coberturas e proteção do solo afeta o cultivo. Já o termo “Classificação dos solos” é genérico em relação ao termo “Classificação”, uma vez que essa, além de possuir as características associadas à formação do solo e ao processo de modificação da cobertura do solo, também designa os níveis de classificação de acordo com o grupo, subgrupo, família, entre outras características referentes à classificação dos solos. O termo “Classificação dos solos” é representado pela sigla *NT* <Narrow Term>, no AgroTermos.

O termo “Cultivo sob várias coberturas. Proteção e cobertura do solo” da classe 631.544.7, possui uma relação de equivalência - sinônimo próximo com o termo “Cobertura da terra”, no tesouro. Isso se justifica, pois, o termo dessa classe não possui o mesmo sentido de forma idêntica, mas se conciliam em alguns significados que são próximos e se diferenciam em alguns aspectos, como por exemplo, “solo” e “terra” e/ou “cultivo” e “cobertura”. Portanto, esses termos, nos dois SOC, representam, em linhas gerais, os diversos tipos de cobertura geológica, física e biológica do solo. A relação hierárquica da classe 631.544.7 em relação ao termo que denota o conceito “Mudança de cobertura da terra” é do tipo “Associativa - Causal - Agente-objeto”, considerando a relação entre o agente e o instrumento utilizado, como também, a relação com o que o agente produziu ou fabricou. Dessa maneira, esse relacionamento é explicado pela mudança da cobertura da terra realizada pelo agente, tendo em vista fatores comerciais, demográficos ou referentes ao regime de propriedade do local.

Os termos “Cobertura da terra” e “Mudança de cobertura da terra” correspondem à categoria de número 6, o que significa que tais termos estão relacionados a aspectos de território e paisagem. Assim, a “Cobertura da terra” é um elemento da superfície do território tais como os atributos do solo e das diferentes classes de solo, ao passo que “Mudança de cobertura da terra” está relacionada aos processos de ocupação e da utilização da terra e de suas transformações.

A próxima classe da CDU é “631.547 - Fases e controle do crescimento das plantas”, que está relacionada ao conceito indicado pelo termo “Ciclo da cultura”, compatibilizado no AgroTermos. A compatibilidade do vocabulário nos dois SOC é apresentada no Quadro 30.

QUADRO 30 - Compatibilidade da classe 631.547 com o termo Ciclo da cultura

Nº	Estrutura classificatória da CDU	AgroTermos - Maculan (2015)	Relações semânticas entre AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos - Maia (2018)
15	631.547 - Fases e controle do crescimento das plantas	Ciclo da cultura	Estabelecimento de culturas agrícolas <Narrow Term> Plantio	Equivalência - Quase sinônimo

FONTE: Dados da pesquisa (2021).

Diante do Quadro 30, percebe-se que o termo “Fases e controle do crescimento das plantas” pertencente à classe 631.547 é mais específico em relação à classe 631.544.7. Além disso, esse termo possui a mesma representatividade no AgroTermos e foi inserido na classe 631.547 da CDU, constituindo uma relação de equivalência no tesouro.

Por meio da relação semântica no AgroTermos, o termo “Estabelecimento de culturas agrícolas” é específico em relação ao termo “Plantio”. Essa relação representada pela sigla *NT*<Narrow Term> pode ser explicada, pois, para cada cultura ou cultivo agrícola, é estabelecido um modo de plantio, por exemplo, com o cultivo do arroz, que é permeado por três grandes fases: vegetativa, reprodutiva e a maturação.

Portanto, o termo da classe 631.547 estabelece uma relação de equivalência - quase sinônimo em relação ao termo “Ciclo da cultura”, no AgroTermos. Esse tipo de relação manifesta-se quando os termos compartilham significados próximos ou similares, mas podem estabelecer algumas diferenças por combinação ou sobreposição. No exemplo do cultivo do arroz, conforme citado, as fases e o ciclo da cultura são iguais, mas o controle do crescimento é uma das etapas ou fases e não compreende todo o ciclo da cultura.

O termo “Ciclo da cultura” é relacionado à categoria de número 5, no AgroTermos, correspondendo à categoria Agronomia. Desse modo, entende-se que a relação é estabelecida, pois, o Ciclo da cultura compreende as fases do desenvolvimento de uma cultura, desde o plantio até a colheita.

Em seguida, analisou-se a classe da CDU “631.57 - Produtos, subprodutos e produtos semi-acabados da agricultura”, que foi representada pelo termo que indica o conceito “Produto agropecuário”, no AgroTermos. A seguir, no Quadro 31, é apresentado o resultado dessa compatibilidade de linguagem.

QUADRO 31 - Compatibilidade da classe 631.57 com o termo Produto agropecuário

Nº	Estrutura classificatória da CDU	AgroTermos - Maculan (2015)	Relações semânticas entre AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos - Maia (2018)
16	631.57 - Produtos, subprodutos e produtos semi-acabados da agricultura	Produto agropecuário	Produto <Used For> produto agrícola	Hierárquica - Hiperônimo- Hipônimo - inclusão de classe-taxonômica

FONTE: Dados da pesquisa (2021).

A classe 631.57 correspondente ao termo “Produtos, subprodutos e produtos semi-acabados da agricultura” e exposta no Quadro 31, é uma classe mais específica em relação à classe 631.547 da CDU. Consta-se que os termos na classe 631.57 possuem uma representatividade parcial com os termos do tesouro.

O termo “Produto agrícola” possui a mesma representatividade no AgroTermos em relação ao termo “Produto”. Por essa razão, o termo “Produto” ou “Produto agrícola”

estabelece uma relação de equivalência, representada pela sigla *UF* <Used For>. Entende-se que esse termo possui características de origem animal e vegetal e se manifesta, em linhas gerais, como produtos da agricultura.

Assim, nota-se que o termo “Produtos, subprodutos e produtos semi-acabados da agricultura” da classe 631.57 em relação ao termo “Produto agropecuário”, forma uma relação do tipo Hierárquica - Hiperônimo - Hipônimo - inclusão de classe-taxonômica. Tal relação se estabelece, sobretudo, pela relação entre uma espécie que é um conjunto de indivíduos que apresentam características comuns ou semelhantes de acordo com a natureza das subcategorias em relação a espécie. A relação taxonômica é devida aos termos da classe 631.57 ser um tipo de “Produto agropecuário”, que determina aspectos associados a produtos oriundos da Agricultura e da Pecuária, como: flores, plantas, carne, leite e seus derivados.

O termo “Produto agropecuário” pertence à categoria de número 5, no AgroTermos, cujo conceito atribuído ao termo corresponde à Agronomia. Salienta-se que esse termo está associado a práticas e princípios utilizados em produtos da Agropecuária que impactam na economia.

A próxima classe analisada da estrutura classificatória da CDU é a “631.58 - Sistemas específicos de cultura” representada pelo termo que indica o conceito “Cultura anual”, no AgroTermos. A seguir, no Quadro 32, é apresentado o resultado do processo de compatibilidade de linguagem.

QUADRO 32 - Compatibilidade da classe 631.58 com o termo Cultura anual

Nº	Estrutura classificatória da CDU	AgroTermos - Maculan (2015)	Relações semânticas entre AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos - Maia (2018)
17	631.58 - Sistemas específicos de cultura	Cultura anual	Estabelecimento de culturas agrícolas <Narrow term> plantio	Hierárquica - Hiperônimo - Hipônimo - Inclusão de Classe - Atividade subordinada

FONTE: Dados da pesquisa (2021).

A descrição apresentada no Quadro 32 indica que a classe 631.58, constituída pelo termo “Sistemas específicos de cultura”, é uma classe mais específica em relação à classe 631.57. Nota-se que o termo na classe 631.57 não possui uma representatividade específica ou equivalente com o termo do tesouro. Para essa compatibilidade, o termo “estabelecimento de culturas agrícolas” forma uma relação de hierarquia e é representado pela sigla *NT*<Narrow Term>, no tesouro, em relação ao termo “plantio”. A explicação para esse relacionamento é que, para o estabelecimento de uma cultura agrícola, é necessário definir o modo de plantio.

Assim, o termo “Sistemas específicos de cultura” da classe 631.57 possui uma relação hierárquica em relação ao termo “Cultura anual”, formulando uma relação do tipo

Hierárquica - Hiperônimo - Hipônimo - Inclusão de Classe - Atividade subordinada. Essa relação se estabelece, sobremaneira, por causa das atividades envolvidas entre os objetos. Portanto, a atividade “Cultura anual” está subordinada a um sistema específico de cultura, ou seja, um tipo de cultura, caracterizado pela replantação anual, visto que o ciclo produtivo encerra-se a cada ano.

O termo “Cultura anual” pertence à categoria de número 5, no AgroTermos, com referência a um tipo de cultura relacionado a aspectos agronômicos (categoria Agronomia). Esse tipo de cultura volta-se à conservação da qualidade e produtividade das plantações de modo cíclico, com periodicidade anual, o que afeta o período de pousio do solo, haja vista restaurar a fertilidade e, assim, garantir a realização de um novo cultivo.

A próxima classe da estrutura classificatória da CDU a ser analisada é a classe 631.58 em relação a classe 502 - O meio ambiente e a sua proteção, que é representada de forma conceitual por meio dos termos “Meio ambiente” e “Ambiente físico” no AgroTermos. A compatibilidade resultante dos dois vocabulários é apresentada no Quadro 33.

QUADRO 33 - Compatibilidade da classe 631.15:502 com os termos Meio ambiente e Ambiente físico

Nº	Estrutura classificatória da CDU	AgroTermos - Maculan (2015)	Relações semânticas entre AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos - Maia (2018)
18	631.58:502 - Sistemas específicos de cultura e a proteção ambiental	Meio ambiente	Proteção ambiental <Related Term> Meio ambiente	Equivalência - Sinônimo parcial
		Ambiente físico	Proteção ambiental <Related Term> Meio ambiente	Hierárquica - Holônimo - Merônimo – Objeto estruturado - Componente complexo

FONTE: Dados da pesquisa (2021).

O Quadro 33 apresenta a classe 631.58:502 com o termo “Sistemas específicos de cultura e a proteção ambiental” da CDU, manifestando-se como uma classe coordenada em relação à classe 631.58, tendo em vista o recorte temático na estrutura classificatória da CDU. Essa estrutura apresenta-se com representividade parcial em relação ao termo “Meio Ambiente” no AgroTermos e não possui nenhuma representatividade específica com o termo “Ambiente físico”, conforme representado no Quadro 33.

Pelas informações registradas no Quadro 33, nota-se que o termo “Meio Ambiente” é um termo associado em relação ao termo “Proteção ambiental”, pois representa todo o conjunto de práticas de proteção do ambiente físico para a sobrevivência do seres

vivos na biosfera. O termo “Meio Ambiente” é representado no tesauro pela sigla *RT<Narrow Term>* que possui uma relação associativa em relação ao termo “Proteção ambiental”.

A classe 631.58:502 da CDU possui uma relação de equivalência em relação ao termo “Meio Ambiente”, no AgroTermos. Esse termo representa um conjunto de condições de sobrevivência dos seres vivos na biosfera (conjunto no todo e em parte), incluindo o clima, solo, água e organismos, o que estabelece, portanto, uma relação de equivalência - sinônimo parcial. Tal relação significa que a classe 631.58:502 possui termos que compartilham vários sentidos, sendo iguais em muitos significados, mas diferentes em algum aspecto. No termo representado, a diferença está associada ao aspecto de proteção do meio ambiente.

Em relação ao Termo “Ambiente físico”, ele significa o conjunto de elementos e componentes naturais do meio ambiente que envolve, unicamente, os fenômenos e processos físicos e químicos, tais como: rochas, solo, chuva, nuvens, corpos d’água, atmosfera, luz, entre outros. A compatibilidade entre os dois vocabulários revela a existência de uma relação Hierárquica - Holônimo - Merônimo - Objeto estruturado - Componente complexo, constituindo-se em uma relação todo-parte, visto que possui aspectos relacionados à função de propriedade ou função do objeto. Nesse caso, o objeto pode ser dividido em partes ou componentes para formar o todo. Isso se justifica pelo fato do “Ambiente físico” ser composto pela parte terrestre, aquática e pelo ar atmosférico, pertencentes ao meio ambiente.

No AgroTermos, “Meio ambiente” e “Ambiente físico” foram inseridos na categoria de número 4, cujo conceito atribuído ao termo corresponde a “Ambiente”. Por sua vez, o conceito desses termos corresponde ao ambiente onde vivem os seres humanos e outros seres vivos, de modo geral.

A classe seguinte da CDU, considerando a estrutura classificatória analisada, é a classe “631.581 - Pousio” que está relacionada ao conceito denotado pelos termos “Pousio” e “Período de Pousio”, compatibilizados com o AgroTermos. A compatibilidade do vocabulário nos dois SOC é apresentada no Quadro 34.

QUADRO 34 - Compatibilidade da classe 631.581 com os termos Pousio e Período de pousio

Nº	Estrutura classificatória da CDU	AgroTermos - Maculan (2015)	Relações semânticas entre AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos - Maia (2018)
19	631.581 - Pousio	Pousio	Pousio<Used For> Cultivo pousio	Equivalência - Sinônimo absoluto
		Período de pousio	Pousio<Used For> Cultivo pousio	Equivalência - Quase sinônimo

FONTE: Dados da pesquisa (2021).

Pelas informações constantes no Quadro 34, nota-se que o termo “Pousio” da classe 631.581 constitui uma classe mais específica em relação à classe 631.57. Nesse sentido, os termos “Pousio” e “Período de pousio” possuem a mesma representatividade, no AgroTermos. Assim, esse termo foi inserido na classe 631.581 da CDU, cuja relação estabelecida no tesauro foi de equivalência.

O termo “Pousio” da classe 631.581 é um termo equivalente relacionado aos termos “Pousio” e “Cultivo de pousio”, no AgroTermos, pois, o termo “Pousio” é sinônimo de “Cultivo de Pousio” e se conceitua como o período de “descanso” da terra, ou seja, quando ela deixa de ser cultivada. Essa relação de equivalência é representada pela sigla *UF* <Used For>.

Em seguida, o termo “Pousio” da classe 631.581 estabelece uma relação do tipo semântica de equivalência - sinônimo absoluto em relação ao termo “Pousio”, no AgroTermos. Esse estabelecimento pode ser explicado, sobretudo, a partir da igualdade semântica dos termos, ou seja, eles possuem significados iguais em relação ao mesmo conceito que os vinculam entre si, o que os tornam semelhantes. Já o termo “Período de pousio” forma uma relação de equivalência - quase sinônimo, em relação ao termo “Pousio”. Tal relação acontece quando os termos compartilham significados próximos ou similares, podendo haver algumas diferenças por combinação ou sobreposição, visto que os termos apresentam significados próximos. Explica-se esse tipo de relacionamento, considerando que o período de pousio compartilha sentidos próximos de pousio, mas com finalidades diferentes neste contexto.

Percebe-se que os termos “Pousio” e “Período de pousio” correspondem à categoria de número 5, no AgroTermos, cujo significado que foi a eles atribuído corresponde à categoria Agronomia. Assim, a relação é equivalente, pois, o significado dos termos está associado a aspectos de interrupção do uso e das atividades agrárias no solo, por diferentes períodos de tempo, sem deixar de lado a sustentabilidade, a economia e a produção agrícola.

A próxima classe na hierarquia da CDU é a classe “631.86 - Adubos orgânicos em geral. Fertilizantes de origem biológica”, que está relacionada ao conceito indicado pelo termo “Biomassa”, sendo esse compatibilizado com o AgroTermos. A compatibilidade do vocabulário nos dois SOC é apresentada no Quadro 35.

QUADRO 35 - Compatibilidade da classe 631.86 com o termo Biomassa

Nº	Estrutura classificatória da CDU	AgroTermos - Maculan (2015)	Relações semânticas entre AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos - Maia (2018)
20	631.86 - Adubos orgânicos em geral. Fertilizantes de origem biológica	Biomassa	Biomassa <includes> Matéria Orgânica Recurso biológico <Used For> Recurso vivo	Equivalência - Quase sinônimo

FONTE: Dados da pesquisa (2021).

No Quadro 35, o termo “Adubos orgânicos em geral. Fertilizantes de origem biológica” da classe 631.86 corresponde a uma classe mais específica em relação à classe 631.581. Esse termo possui a mesma representatividade no AgroTermos e foi inserido na classe 631.547, constituindo uma relação de equivalência, no tesauro.

No AgroTermos, a relação semântica do termo “Biomassa” é do tipo associativa <includes> e está estabelecida com o termo matéria orgânica, ou seja, a biomassa é toda matéria orgânica, seja de origem animal ou vegetal, com a finalidade de produzir energia, por meio do vapor produzido pela combustão do material orgânico, cujo vapor rotaciona uma turbina conectada a um gerador de eletricidade. Entretanto, a biomassa pode ser utilizada para a fertilização dos solos, com o uso de resíduos florestais, resíduos agrícolas, excrementos de animais, entre outros. Além disso, o termo “Recurso biológico” é equivalente a “Recurso vivo”, no AgroTermos, e a equivalência é representada pela sigla *UF* <Used For>. Para explicar essa relação, considera-se que a Biomassa é uma substância ou recurso biológico e orgânico renovável, proveniente de organismos vivos de um determinado local.

Portanto, o termo da classe 631.547 possui uma relação de equivalência - quase sinônimo em relação ao termo “Ciclo da cultura”, no AgroTermos. Essa relação manifesta-se quando os termos compartilham significados próximos ou similares, mas podem estabelecer algumas diferenças por combinação ou sobreposição. No caso analisado, as fases e o ciclo da cultura são iguais, mas o controle do crescimento constitui uma das etapas ou fases e não compreende todo o ciclo da cultura.

Nota-se que o termo “Biomassa” pertence à categoria de número 5 com características que o aproxima da categoria Agronomia. Desse modo, a relação pode ser explicada, pois a biomassa não é um adubo a ser utilizado na produção agrícola, mas pode ser processada a ponto de virar adubo, pois o adubo é feito de biomassa. A biomassa também pode gerar eletricidade a ser usada na queima de cerâmicas e de carvão vegetal, na redução do ferro gusa, nos fornos siderúrgicos, na geração de combustível sustentável para fábricas, dentre outras aplicações.

A próxima classe da estrutura classificatória da CDU a ser analisada refere-se à classe “631.874 - Adubação verde. Cultivos que melhoram o solo” referente ao termo que representa o conceito “Adubo verde”, que foi compatibilizado no AgroTermos. O resultado do processo de compatibilidade entre os dois vocabulários é apresentado no Quadro 36.

QUADRO 36 - Compatibilidade da classe 631.874 com o termo Adubo verde

Nº	Estrutura classificatória da CDU	AgroTermos - Maculan (2015)	Relações semânticas entre AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos - Maia (2018)
21	631.874 - Adubação verde. Cultivos que melhoram o solo	Adubo verde	Adubo verde <Related Term> Adubo Fertilizante <Narrow Term> Adubo Orgânico	Associativa - Processo-instrumento

FONTE: Dados da pesquisa (2021).

No Quadro 36, é apresentado o termo “Adubação verde. Cultivos que melhoram o solo” da classe 631.874 da CDU, sendo essa mais específica em relação à classe 631.86. Nesse contexto, no AgroTermos, o referido termo não possui uma representatividade específica ou equivalente.

O termo “Adubo” é específico no AgroTermos, sendo representado pela sigla *NT* <Narrow Term> em relação ao termo “Adubo verde”. O termo “Adubo verde” é mais genérico pelo significado que possui, ou seja, é um adubo composto por plantas como leguminosas, gramíneas, crucíferas, entre outras, cujo objetivo é proteger e enriquecer o solo. Esse tipo de adubo é fragmentado, a fim de servir de cobertura, além de ser decomposto, logo após o plantio, seja em rotação, sucessão ou em consórcio de outros tipos de cultura. Já o termo “Fertilizante” é produzido em laboratórios com material sintético e é composto por produtos químicos que agem no solo, de modo a promover a fertilização. Assim, o termo “Fertilizante” estabelece uma relação de hierarquia representada por *NT* <Narrow Term> em relação ao termo “Adubo orgânico”. Essa relação ocorre, pois o adubo orgânico é constituído de resíduos de origem animal e vegetal, tais como folhas secas, gramas, restos vegetais, restos de alimentos e esterco animal. Esses resíduos se decompõem, tornando-se um insumo bastante utilizado na agricultura, sobretudo por quem prefere usar apenas produtos naturais.

Nesta análise, o termo “Adubação verde. Cultivos que melhoram o solo” da classe 631.874, estabelece uma relação Associativa - Processo-instrumento com o AgroTermos. Justifica-se tal relação, tendo em vista que os termos da classe 631.874 indicam os processos que melhoram o solo e o termo “Adubo verde” pode ser um dos instrumentos utilizados nesse processo.

O termo “Adubo verde” corresponde à categoria de número 5, ou seja, está associado à categoria Agronomia. Por isso, o termo faz alusão a aspectos que garantem a fertilidade do solo de modo natural, haja vista possibilitar a preparação da terra e o cultivo de vários tipos de cultura.

O Quadro 37 apresenta o resultado da compatibilidade entre os termos da classe da CDU “>’ - 632.95 Pesticidas”. Esses termos foram compatibilizados com o termo “Pesticida”, que representa o mesmo conceito presente no AgroTermos.

QUADRO 37 - Compatibilidade da classe 632.95 com o termo Pesticida

Nº	Estrutura classificatória da CDU	AgroTermos - Maculan (2015)	Relações semânticas entre AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos - Maia (2018)
22	'>' 632.95 - Pesticidas	Pesticida	Praguicida <Use> Pesticida	Equivalência - Sinônimo absoluto

FONTE: Dados da pesquisa (2021).

Pelo que está exposto no Quadro 37, afere-se que a classe “> 632.95 Pesticidas” é mais específica que a classe 631.874 apresentada no Quadro 36. Consta-se que aquela classe foi correspondida com o termo “Pesticida”, no AgroTermos, a partir do processo de compatibilidade. Esse termo foi inserido na classe da CDU, cujo sinal de '>', de acordo com a matriz de compatibilidade conceitual de Dahlberg (1981), corresponde a classes mais específicas em relação à hierarquia da CDU. O termo da classe 632.95 possui a mesma representatividade de equivalência ou similaridade em relação aos termos “Pesticida” e “Praguicidas”, no tesouro. No AgroTermos, a relação de equivalência entre o termo “Pesticida” e “Praguicida” é representada pela sigla *UF<Used For>*, a partir do significado semântico desses termos. Os termos significam substâncias químicas usadas na agricultura para a prevenção, eliminação e controle de pragas e doenças nas plantas e nos animais.

A classe 632.95 da CDU possui uma relação de equivalência - sinônimo absoluto entre o termo “Pesticida”, no AgroTermos. Essa relação ocorre quando os termos possuem significados iguais em relação ao mesmo conceito em sua totalidade, ou seja, os significados são abrangentes, e/ou quando possuem o mesmo termo em diferentes vocabulários. No caso analisado, os termos são iguais nos dois vocabulários.

O termo “Pesticida” corresponde à categoria de número 5, cujo significado corresponde à Agronomia. O termo “Pesticida” está associado a métodos e procedimentos que eliminam ou evitam pragas ou doenças na agropecuária, embora possam causar poluição do ar e/ou do solo.

Por fim, considerando a estrutura classificatória da CDU, após a classe mais específica, com o código 632.95, analisou-se a última classe mais específica da hierarquia, a classe “> 636 Criação de gado e de animais em geral. Criação de gado. Criação de animais domésticos”, referente ao termo que representa o conceito “Pecuária”, que foi compatibilizado com o mesmo termo presente no AgroTermos. Essa compatibilidade entre os dois vocabulários é apresentada no Quadro 38.

QUADRO 38 - Compatibilidade da classe 636 com o termo Pecuária

Nº	Estrutura classificatória da CDU	AgroTermos - Maculan (2015)	Relações semânticas entre AgroTermos em relação a cada termo (Embrapa)	Relações semânticas CDU e AgroTermos - Maia (2018)
23	'> 636 - Criação de gado e de animais em geral. Criação de gado. Criação de animais domésticos	Pecuária	Criação animal <Used For> Pecuária	Equivalência - Sinônimo parcial

FONTE: Dados da pesquisa (2021).

Nota-se que o termo “Criação de gado e de animais em geral. Criação de gado. Criação de animais domésticos” da classe 636 é mais específico em relação à classe 632.95. Nesse caso em específico, percebe-se que o termo da CDU possui a mesma representatividade no AgroTermos.

Portanto, o termo da classe 632.95 estabelece uma relação de equivalência em relação ao termo “Pecuária”, no tesouro. O termo “Pecuária” é um termo equivalente representado pela sigla *UF<Used For>* em relação ao termo “Criação animal”. O relacionamento pode ser explicado, pois, a criação de animais é parte da agricultura que envolve o cuidado e a gestão de diversos tipos de animais domesticáveis, incluindo atividades de produção de carne, leite e derivados do leite, além da confecção de produtos como couro e lã, cujos resultados dessa produção impactam na economia.

Dessa maneira, o termo “Pecuária” estabelece uma relação de equivalência - sinônimo parcial com o AgroTermos. Assim como em outros casos mencionados, a relação é justificada, sobremaneira, pois, os termos possuem vários sentidos, com vários significados iguais, mas que se diferenciam em algum aspecto. Nesta análise, o termo “Pecuária” do tesouro, compartilha o mesmo significado na CDU com a criação de gado, mas não compartilha com criação de animais em geral (Zoológico), bem como a criação de animais domésticos (cães e gatos).

O termo “Criação de gado e de animais em geral. Criação de gado. Criação de animais domésticos” corresponde à categoria de número 5, cujo significado que lhe foi atribuído equivale à categoria Agronomia, envolvendo aspectos associados à criação de animais para o abate visando à obtenção de carne, bem como em pecuária leiteira para a produção de leite. A pecuária de corte pode ser subdividida em intensiva, na qual o gado é criado em pequenos espaços e alimentado com ração específica, produzindo carne macia e de boa qualidade; e em extensiva, na qual o gado é criado solto e se alimenta de grama, produzindo uma carne mais dura, já que o gado desenvolve uma musculatura rígida.

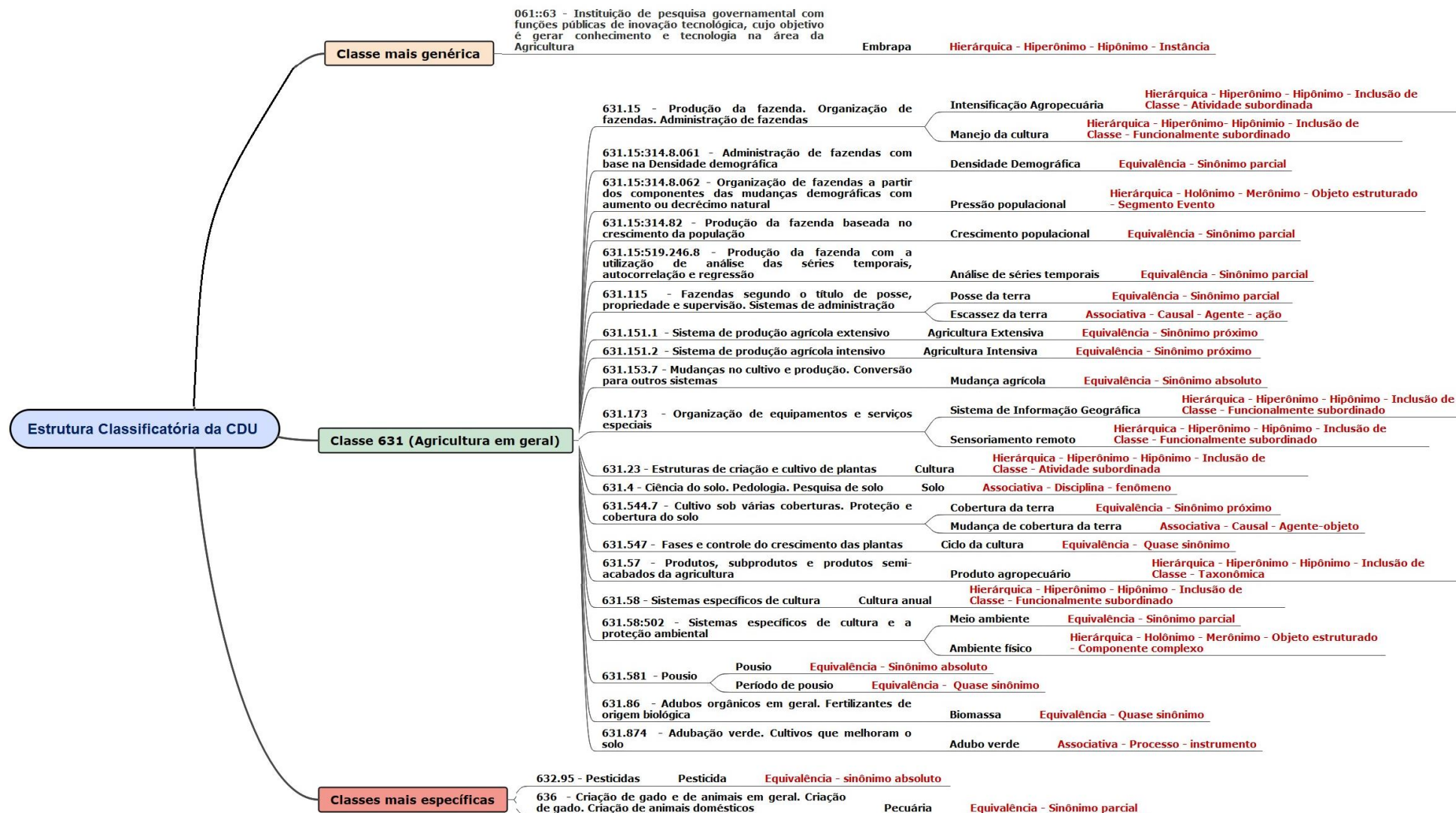
Diante do exposto, constatou-se que há compatibilidade de linguagem entre os dois SOC, não somente nos vocabulários em si, mas também, quando esses se encontram em diferentes bases de dados, considerando questões relacionadas ao reuso de termos ou

conceitos e à interoperabilidade dos SOC em questão. Isso possibilita a organização e tratamento da informação para que a recuperação seja consistente, mesmo que a informação, no ambiente digital, esteja em diferentes sistemas ou bases de dados.

Portanto, por causa do grande volume de dados disponibilizados na web ou em qualquer ambiente digital, é imprescindível a realização desses procedimentos, de modo que garanta uma recuperação mais precisa e consistente. Após a compatibilidade de linguagem, foi gerado o modelo de compatibilidade conceitual, a partir dos termos da estrutura classificatória da CDU e sua notação com os termos do AgroTermos, conforme exposto na Figura 28.

No processo de compatibilidade, cada classe da CDU foi comparada com os conceitos e categorias apresentados no modelo de Maculan (2015) e, nesse processo, verificou-se o tipo de relação semântica que poderia ser estabelecida entre classes e termos, cujos tipos de relações foram retirados da pesquisa de Maia (2018). Mediante esse processo, gerou-se uma nova estrutura, conforme apresentado na Figura 28.

FIGURA 28 - Estrutura gerada a partir da compatibilidade entre as classes da CDU, AgroTermos e relações semânticas



FONTE: Dados da pesquisa (2021).

A estrutura gerada com o processo de compatibilidade constitui o produto final desta tese (modelo proposto), cujos resultados serão analisados e descritos no capítulo seguinte. A princípio, conforme exposto na Figura 28, evidencia que houve compatibilidade entre os elementos analisados. Percebe-se, também, que as relações estabelecidas entre os elementos (classes da CDU e conceitos no tesouro) são variadas, contemplando relações do tipo hierárquica, equivalente e associativa. Esse resultado leva a pressupor que o número variado de relações pode condicionar novos modos de se recuperar a informação. Além disso, presume-se que a fusão dos dois SOC oportunizará mais flexibilidade ao sistema, aproximando-o de um sistema facetado, sem, contudo, perder a rigidez nos relacionamentos, ou seja, prevalece a característica básica da CDU: um sistema semifacetado.

Em síntese, este subcapítulo detalhou o processo de compatibilidade, evidenciando que as classes da CDU possuem correspondência com o AgroTermos e que diversas relações semânticas podem ser estabelecidas, de modo que sejam alcançadas melhorias na recuperação. O próximo subcapítulo apresenta a aplicação do modelo utilizando o software *TemaTres*, com a finalidade de representar os elementos do domínio analisado no formato Simple Knowledge Organization System (SKOS). Salienta-se que o propósito principal desse processo é garantir a interoperabilidade entre diferentes sistemas de informação.

6 ANÁLISE E DESCRIÇÃO DOS RESULTADOS

“Por vezes sentimos que aquilo que fazemos não é senão uma gota de água no mar. Mas o mar seria menor se lhe faltasse uma gota”.

(Madre Teresa de Calcutá)

Este capítulo reserva-se à análise e descrição dos resultados, a partir dos procedimentos metodológicos definidos e as possíveis correspondências com os objetivos formulados previamente para a investigação. O referido capítulo constitui, portanto, a segunda etapa do estudo de caso, e está dividido em três seções, com o propósito de facilitar o entendimento dos resultados, a partir do que eles indicam, por meio de inferências e proposições, devidamente alinhadas aos objetivos do estudo. O Quadro 39 resume o que será discorrido ao longo deste capítulo em alinhamento com os propósitos da pesquisa.

QUADRO 39 - Estrutura textual para descrição e análise dos resultados

Número do subcapítulo	Procedimento técnico	Objetivos da pesquisa	Principais atividades
6.1	Intercâmbio de informações sobre o assunto “Intensificação Agropecuária” do domínio da agricultura entre os vocabulários dos dois SOC.	<p>(Objetivo específico 1)</p> <p>- Analisar a compatibilidade dos termos da estrutura classificatória da CDU com os descritores do tesauro AgroTermos, na representação do assunto “Intensificação Agropecuária” do domínio da agricultura.</p>	<p>- Seleção das classes da CDU que correspondem ao assunto do domínio analisado.</p> <p>- Seleção de termos que representam o campo da “Intensificação Agropecuária” no AgroTermos.</p> <p>- Comparar os resultados com estudos similares.</p>
6.2	Criação do modelo de Compatibilidade de linguagem entre os SOC.	<p>(Objetivo específico 2)</p> <p>- Verificar o potencial da recuperação com o uso da expansão dos termos e seus relacionamentos semânticos mapeados nesses dois SOC.</p>	<p>- Relacionar as classes da CDU com os termos do AgroTermos, identificando os tipos de relações semânticas e a quantidade de ocorrências a partir do recorte mapeado.</p> <p>- Construir um modelo baseado na compatibilidade de linguagem entre a estrutura classificatória da CDU e dos termos do tesauro AgroTermos que represente a estrutura temática mapeada.</p>

			<ul style="list-style-type: none"> - Exposição do modelo proposto. - Comparar os resultados com estudos similares.
6.3	Representação do modelo de compatibilidade proposto entre os dois SOC na linguagem SKOS.	<p>(Objetivo específico 3)</p> <ul style="list-style-type: none"> - Representar o modelo de compatibilidade proposto, utilizando o padrão Simple Knowledge Organization System, no âmbito da web semântica. 	<ul style="list-style-type: none"> - Utilização do software <i>TemaTres</i> para representar a estrutura temática do modelo de compatibilidade proposto. - Representar o modelo na linguagem SKOS. - Comparar os resultados com estudos similares.

FONTE: Elaborado pelo autor (2021).

Pelo Quadro 39, nota-se que os três procedimentos técnicos são dependentes, ou seja, desenvolveram-se de modo sequencial. Isso quer dizer que, para a realização do processo de compatibilidade da linguagem, foi necessário, a princípio, delimitar o recorte temático no âmbito da CDU e do AgroTermos. De posse desse recorte, a compatibilização é realizada, tendo em vista demonstrar as possíveis correspondências, de modo a validar a proposta de uso de dois SOC para viabilizar a recuperação da informação. Com esses resultados, o modelo é apresentado de forma gráfica e aplicado no software *TemaTres*, que permitiu exibir o modelo proposto no formato SKOS para possibilitar a interoperabilidade entre diferentes sistemas de informação. As seções seguintes descrevem e analisam os resultados desses procedimentos técnicos.

6.1 Mapeamento do recorte temático

Com a finalidade de analisar a compatibilidade dos termos da estrutura classificatória da CDU com os descritores do tesauro AgroTermos, na representação do assunto “Intensificação Agropecuária” do domínio da agricultura (Objetivo específico 1), apresenta-se o recorte temático delimitado para este estudo. O levantamento preliminar realizado na CDU possibilitou a identificação da classe principal (631 – Agricultura em geral), por se tratar de uma classe mais genérica em relação à temática a ser representada. Esta classe principal, foi complementada por três outras classes que também representam aspectos associados ao assunto do domínio da agricultura, contendo uma classe mais

genérica e duas mais específicas em relação à classe principal. Isso foi possível por meio das combinações com subdivisões auxiliares especiais e outras classes, utilizando auxiliares relacionais compostos e simples para incrementar a representação e os níveis de especificidade na representação de assuntos nos dois SOC. Dessas classes, foram identificadas suas respectivas subclasses. O Quadro 40 sintetiza a quantidade de classes e subclasses, considerando a estrutura classificatória da CDU.

QUADRO 40 - Quantidade de classes e subclasses da CDU relacionadas à temática do domínio da agricultura

SOC	Classes	Quantidade de subclasses
CDU	061::63	0
	631	20
	632	1
	636	0

FONTE: Dados da pesquisa (2021).

LEGENDA: 0 Indica que não há item correspondente.

O Quadro 40 revela que o recorte delimitado para este estudo foi composto por 4 classes e 21 subclasses, ou seja, pela classe 061::63, 20 subclasses dentro da classe 631(recorte), a classe 632 e por fim, a classe 636, totalizando 23 classes da CDU. Esse recorte é específico de um domínio do conhecimento (Agricultura), e serve para agrupar diferentes itens informacionais vinculados a esse domínio, quanto a composição de um acervo. Tal agrupamento é estabelecido por meio das diferenças e semelhanças entre os assuntos do domínio, característica essa considerada como um dos principais princípios da classificação bibliográfica, como pontuado por Langridge (1977), Piedade (1977) e Vikery (1990).

Ao analisar os assuntos e a cobinações entre eles, utilizando as classes, subclasses e subdivisões especiais selecionadas para representar o recorte a ser investigado, percebeu-se que a estrutura classificatória é bastante representativa, sobretudo por contemplar assuntos superordenados, subordinados e coordenados. O Quadro 41 apresenta um resumo esquemático de alguns assuntos pertencentes ao domínio da agricultura e que foram selecionados para compor o recorte de análise deste estudo.

QUADRO 41 - Alguns assuntos expressos nas classes da CDU - recorte temático

Número da classe/subclasse	Assunto	Nível do assunto (Superordenado, subordinado e coordenado)
061::63	Instituição de pesquisa governamental com funções públicas de inovação tecnológica, cujo objetivo dessa instituição é gerar conhecimento e tecnologia na área da agricultura	Superordenado
631.15	Produção da fazenda...	Superordenado
631.15:314.8.061	Administração de fazendas com base na Densidade demográfica	Coordenado
631.15:502	Sistemas específicos de cultura e a proteção do meio ambiente	Coordenado

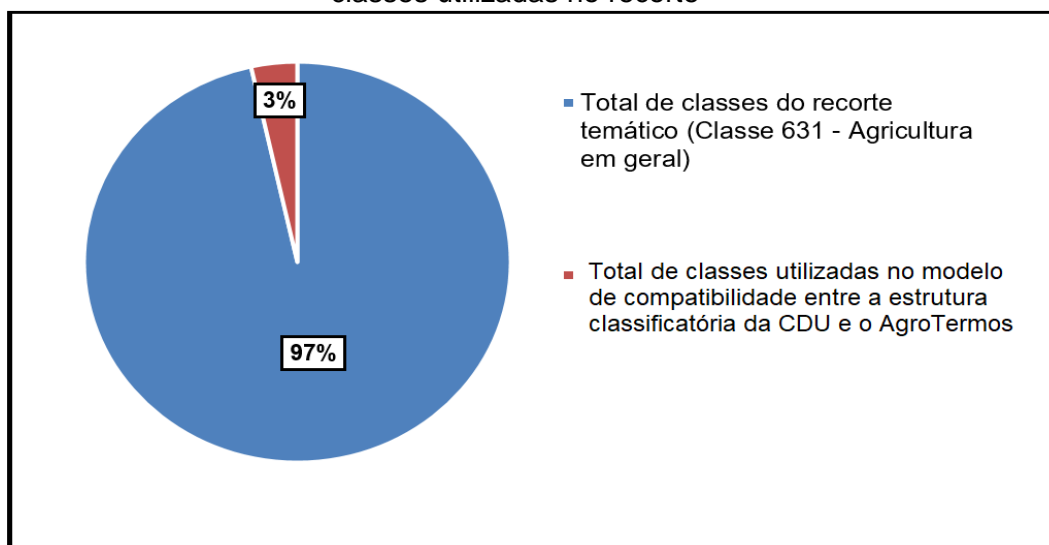
631.151.1	Sistema de produção agrícola extensivo	Subordinado
631.173	Organização de equipamentos e serviços especiais	Superordenado
631.58	Sistemas específicos de cultura	Subordinado
631.581	Pousio	Subordinado
631.86	Adubos orgânicos em geral. Fertilizantes de origem biológica	Subordinado
636	Criação de gado e de animais em geral ...	Subordinado

FONTE: Dados da pesquisa (2021).

A sequência das classes e suas divisões, conforme apresentado no Quadro 41, reforça a estrutura decimal, hierárquica e flexível da CDU. Quanto aos assuntos, parte-se de assuntos mais genéricos que vão sendo especificados (do geral para o específico) e novos assuntos podem ser inseridos, combinados com outras classes ou subdivisões auxiliares especiais, mantendo-se os níveis de exaustividade ou especificidade, cujas relações estabelecidas podem ser variadas. Ressalta-se que, no estabelecimento dessas relações, há um esforço intelectual no sentido de evitar inconsistências ou subjetividades, além de comprometer a lógica do sistema, acarretando problemas no ato de representar e de recuperar a informação, conforme argumentado por Spark Jones (1970). Assim, durante a construção da informação, aumenta a necessidade de que os SOC sejam mais flexíveis e dinâmicos, por depender de seres humanos, o que pode levar a ter sua qualidade comprometida durante a representação (WEISS; BRÄSCHER, 2017).

Identificadas as classes e subclasses e os assuntos gerais e específicos vinculados a elas, foi possível representar esse recorte com o total de 10 classes e subclasses, pois considera-se uma amostra adequada para representar as possíveis relações entre os níveis de assuntos (superordenados, subordinados e coordenados) entre as classes e subclasses do modelo. A partir desta análise, considerando o universo da classe geral 631, que é composta, conforme estrutura da CDU, por 636 subclasses, sendo que 20 classes foram utilizadas a partir da classe 631, uma classe mais genérica e duas mais específicas, totalizando 23 classes da estrutura classificatória da CDU para a construção do modelo proposto, o que corresponde a 3% do total de termos do recorte temático e 97% corresponde ao restante do número de classes e seus descritores referente à estrutura classificatória da CDU utilizada na construção do modelo de compatibilidade. Portanto, o Gráfico 11 ilustra o total de classes e subclasses utilizadas no recorte em relação a estrutura classificatória da CDU e o AgroTermos.

GRÁFICO 11 - Quantidade de classes da CDU utilizadas no modelo e quantidade de classes utilizadas no recorte



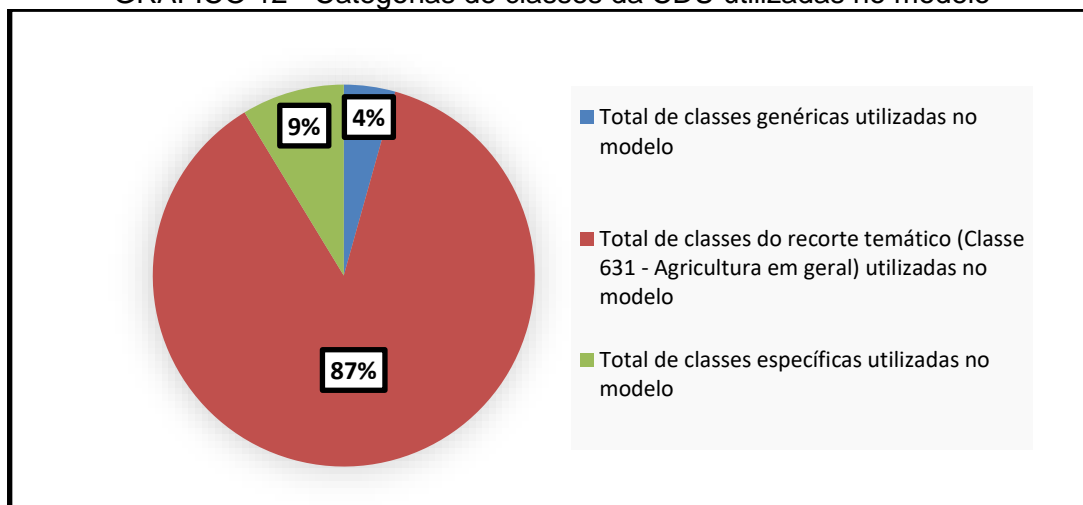
FONTE: Dados da pesquisa (2021).

De acordo com o Gráfico 11, a quantidade de classes utilizadas para o processo de compatibilidade que visou à constituição do modelo foi pequena no universo da CDU, embora essa quantidade foi suficiente para identificar a possibilidade de relacionamentos que pode ser estabelecida. Isso pode confirmar que a estrutura classificatória da CDU não é engessada, como outros sistemas de classificações, por exemplo, a CDD. Essa constatação vai ao encontro do que relatou Piedade (1997) e Souza (2012) acerca da flexibilidade do sistema, ou seja, a natureza das classes e subclasses é analítico-sintética, de modo a possibilitar a organização relacional do vocabulário em facetas. Assim, com uma quantidade pequena de classes, é possível gerar infinitos relacionamentos, permitindo, também, o reuso dos termos e com seus respectivos conceitos e suas combinações na representação de assuntos com um nível maior de detalhamento (KAOSAR, 2008). Para representar o recorte temático na CDU, foi necessário a representação de assuntos compostos e complexos, utilizado números de classes principais e subdivisões auxiliares especiais e nenhuma subdivisão auxiliar comum para a combinação de assuntos no modelo proposto. Portanto, para a representação de assuntos compostos ou complexos na CDU entre as tabelas principais e auxiliares especiais utilizou-se auxiliares relacionais compostos e simples para combinar os assuntos de cada tabela, ou classificar utilizando a estrutura da CDU para representar os assuntos compostos ou complexos de forma separada, e logo após, inserir no modelo como assunto simples, considerando sua flexibilidade.

Além do mapeamento da classe 631 da CDU e a identificação de classes passíveis de serem compatibilizadas e suficientes para constituir o produto final desta pesquisa, foram utilizadas, para o modelo, uma classe genérica, 20 classes vinculadas ao recorte (classe principal 631) e duas classes específicas. Esse resultado é apresentado no

Gráfico 12, com indicadores percentuais.

GRÁFICO 12 - Categorias de classes da CDU utilizadas no modelo



FONTE: Dados da pesquisa (2021).

O Gráfico 12 expõe o total de 23 classes utilizadas para a construção do modelo proposto: 20 classes pertencem ao recorte temático (classe 631), uma classe da CDU mais genérica em relação ao recorte e duas classes são mais específicas. O total de classes utilizadas para o modelo que pertence ao recorte temático representa 87%; o total de classes mais genéricas representa 4%; e, por fim, 9% representa as classes mais específicas em relação ao recorte.

Observa-se que a quantidade de classes utilizadas no modelo foi baixa, considerando o universo do recorte temático. Mesmo com um baixo percentual, evidenciou-se que as classes foram utilizadas em vários níveis da hierarquia do recorte, por conseguinte, foram possíveis estabelecer relações entre os assuntos dessas classes e os conceitos do tesauro. Esse resultado indica o aspecto relacional da CDU que, mesmo sendo um sistema decimal e hierárquico, possibilita a criação de novos relacionamentos, em uma teia infinita de possibilidades. Essa constatação vai ao encontro do que evidenciou Silva e Ganim (1994) acerca da flexibilidade do sistema, ou seja, sua estrutura está sucessível a receber novas divisões e subdivisões, num processo teoricamente infinito, que constitui o caráter hierárquico (enumerativo) do sistema.

Após a seleção das classes e a devida apresentação de suas características, partiu-se para a atividade de seleção dos termos que representam o assunto Intensificação Agropecuária no AgroTermos. Para isso, conforme detalhado nos procedimentos metodológicos, recorreu-se ao estudo de Maculan (2015), que mapeou 30 termos representativos do recorte temático da Intensificação Agropecuária, alinhados à teoria de Boserup (1965). No domínio da agricultura, a autora delimitou o campo da "Intensificação Agropecuária", considerando a estrutura do tesauro analisado. O resultado obtido pela autora foi utilizado para complementar o recorte temático deste estudo, cujo Quadro 42 sintetiza

alguns dos conceitos e seus respectivos significados, conforme o resultado do estudo de Maculan (2015).

QUADRO 42 - Exemplos de conceitos pertencentes ao AgroTermos (Intensificação Agropecuária)

Ordem (número do conceito)	Conceito	Definição
4	Cultura	Plantas cultivadas pelo homem ou não, cujos produtos foram, são ou serão colhidos em quaisquer de seus estágios de crescimento; é influenciada pelo calendário agrícola e pelo ciclo do tipo de cultura; também denominada “produção vegetal” e “planta cultivada”.
6	Sistemas de informação geográfica	Sistema de informação, de suporte e ferramenta de tecnologia geoespacial que reúne, armazena, organiza, analisa e exibe dados de fontes variadas a partir de uma abordagem de natureza cartográfica, altamente valorizada para a documentação de processos que se manifestam na superfície terrestre; usa GPS e dados espaciais; sigla SIG.
7	Meio ambiente	Entidade do ambiente que reúne um conjunto de elementos e processos físicos, químicos, biológicos e antrópicos e toda gama de manifestações decorrentes da combinação ou inter-relação desses componentes; abriga e rege a vida em todas as suas formas, afetando, por exemplo, a população e a economia e sendo afetado por elementos tais como biodegradação, catástrofes naturais, queimadas e políticas ambientais.
12	Pousio	Prática de manejo e sistema de cultivo que prevê a interrupção de atividades ou usos agrícolas, pecuários ou silviculturais do solo para descanso, por diferentes motivos (preparação, regeneração, fertilidade do solo), e por distintos períodos de tempo.
20	Biomassa	Substância ou recurso biológico e orgânico renovável, proveniente de organismos vivos de um determinado local, quantificada segundo seu peso, massa, volume ou área ocupada, e utilizada para a fertilização dos solos ou para a produção de energia.
24	Mudança de cobertura da terra	Processo de modificação da cobertura da terra, natural ou não, devido a fatores comerciais, demográficos ou referentes ao regime de propriedade do local.
25	Densidade demográfica	População humana ou de outros seres vivos, medida em um dado espaço habitado, geralmente expressa por um indicador que determina o número de habitantes por quilômetro quadrado; as características dessa população impactam, por exemplo, o transporte, o uso da terra e a economia.
28	Mudança agrícola	Desenvolvimento econômico das atividades agropecuárias que resulta das alterações espaço-temporais decorrentes da dinâmica de uso e cobertura das terras em função de sua exploração antrópica, que envolve fatores tais como a mecanização e as políticas de desenvolvimento rural.
29	Análise de séries temporais	Ferramenta de análise de informações estatísticas que tem por base padrões de dados temporais, e que auxilia nas previsões de padrões, de séries ao longo de um período de tempo e na determinação de ações estratégicas de prevenção às suas possíveis consequências.

FONTE: Maculan (2015, p. 329-336).

Semelhante à ordem de citação das classes da CDU, é possível estabelecer diferentes tipos de relacionamentos entre os conceitos expressos no Quadro 42. Há evidências de relações do tipo hierárquica (Meio ambiente - Pousio - Biomassa), relações equivalentes (Mudança de cobertura da terra - Mudança agrícola), dentre muitas outras. Portanto, comprova-se que os tesouros se compõem de uma base léxica, com descritores (termos preferidos) e não descritores (termos não preferidos)²⁴, que são definidos e descritos sob a denominação de nota de escopo ou nota de aplicação “[...] e são estruturados em relações hierárquicas (termos genéricos e específicos), não hierárquicas (associativas - termos relacionados) e de equivalência (sinônimos ou quase-sinônimos)” (BOCCATO; RAMALHO; FUJITA, 2008, p. 201).

Para Maculan (2015), os 30 conceitos são representativos, pois delimitam o campo conceitual da Intensificação Agropecuária, dentro do domínio da agricultura e foram validados pelo especialista do domínio. Nesse sentido, é possível que esse campo expresse boa parte do universo da agricultura, haja vista fornecer bases ou insumos para constituição de modelos de representação ou recuperação da informação. Na constituição desses, o processo é completo, mas contribui para a organização do conhecimento do domínio modelado.

Portanto, com o fim de mostrar, com mais clareza, os relacionamentos com o objetivo de estabelecer a interoperabilidade entre diferentes sistemas, o recorte temático a compor o modelo proposto nesta tese é oriundo de elementos vinculados a dois SOC: classes da CDU e termos que denotam os conceitos no AgroTermos, que se referem à amostra trabalhada por Maculan (2015). Embora esse recorte represente uma amostra do universo ou domínio da agricultura suficiente para representar esse modelo, isso possibilitará mais detalhamento e clareza acerca dos elementos que compõem os SOC. A delimitação do escopo ou recorte temático é uma constante nos estudos voltados aos instrumentos de organização do conhecimento, como: Kaosar (2008) - que investigou Clínica Geral (Medicina); Maculan (2015) - Intensificação Agropecuária (Agricultura); Lee (2017) - Instrumentos musicais (Música), dentre outros.

Por fim, considerando os resultados advindos do processo de mapeamento do recorte temático desta pesquisa, parte-se para as próximas análises, direcionadas ao processo de compatibilidade de linguagem entre os dois SOC estudados. Reforça-se que tal processo é formado pela correlação entre as classes e os termos e a identificação de

²⁴ Um conceito pode ser representado por vários termos, sendo que um desses termos será o termo preferido, que é representado pela sigla usado (USE) e os demais serão termos não preferidos, representado pela sigla usado para (UP), que são usados como pontos de acesso ao termo preferido no processo de indexação. Também são conhecidos como descritores e não descritores.

possíveis relações semânticas estabelecidas nessa correlação, conforme apresentado no subcapítulo seguinte.

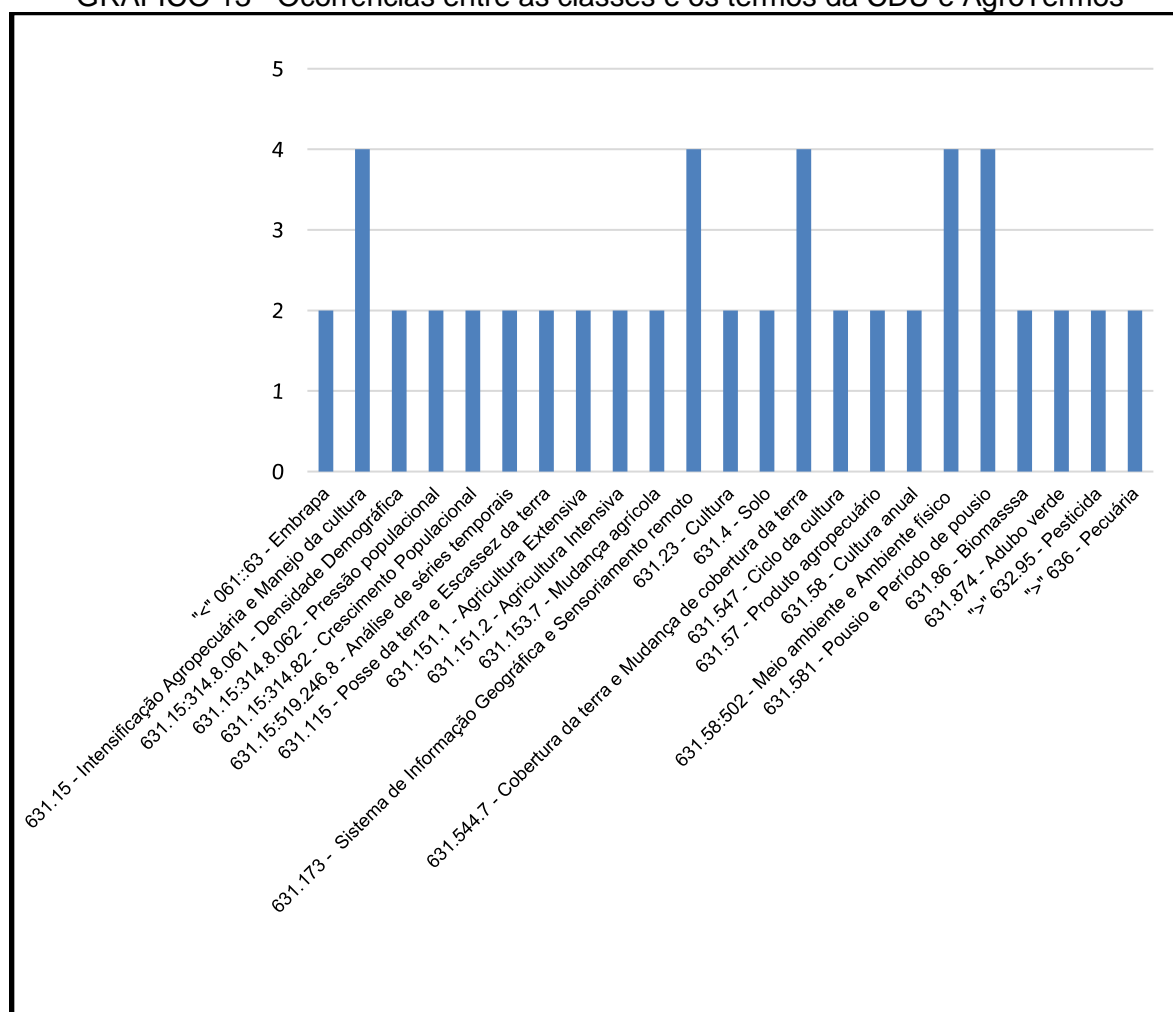
6.2 Compatibilidade de linguagem entre os SOC

Ao comparar as classes da CDU mapeadas com os termos do AgroTermos, foi possível apresentar a quantidade de ocorrências e as relações semânticas estabelecidas entre os dois SOC, propondo assim, um modelo de compatibilidade de linguagem da estrutura temática (Objetivo específico 2), concluindo, desse modo, o processo de compatibilidade, como foi orientado por Dahlberg (1981), de modo a atender o objetivo específico 2 desta tese.

Ressalta-se que, no processo de compatibilidade, recorreu-se ao uso da matriz proposta por Dahlberg (1981), com definição do conceito e as possíveis correspondências dele com as classes da CDU, momento pelo qual foi possível identificar o número de ocorrências conceituais (DAHLBERG, 1981). Tal processo foi complementado com o uso do método de ligação direta entre os dois SOC, de forma unidirecional, sem a utilização de linguagem intermediária, cujo efeito foi o estabelecimento das relações semânticas pelo autor desta tese, a partir do estudo de Maia (2018).

Pelos resultados, foi revelada uma variedade de ocorrências dos termos no contexto da agricultura; a grande maioria dos termos do AgroTermos foi correlacionada duas vezes com as classes da CDU, sendo que houve quatro ocorrências para seis termos. Assim, os termos “Meio ambiente e ambiente físico”, “Intensificação agropecuária e Manejo da cultura”, “Sensoramento remoto e sistema de informação geográfica”, “Cobertura da terra e mudança de cobertura” e “Pousio e período de pousio” obtiveram, cada um, quatro ocorrências, ao passo que os demais termos obtiveram duas ocorrências. Além disso, foi combinada a classe 061 com a classe 63, com o uso de um auxiliar relacional composto com o símbolo “::” e as demais combinações foram combinadas com um auxiliar relacional simples que possui o símbolo “:”, dentro do recorte temático (classe 631), sendo quatro combinações com a classe 631.15 e uma com a classe 631.58 da CDU, conforme detalhado no Gráfico 13.

GRÁFICO 13 - Ocorrências entre as classes e os termos da CDU e AgroTermos



FONTE: Dados da pesquisa (2021).

O Gráfico 13 descreve a quantidade de ocorrências entre as classes da CDU e os termos do AgroTermos. Constatou-se que, cada termo que representa o recorte temático foi compatibilizado com a CDU e com o AgroTermos. O primeiro termo compatibilizado foi "Embrapa" que, após a compatibilidade de linguagem com o vocabulário da CDU, teve duas ocorrências, ficando associado à classe mais geral da CDU no modelo proposto, a classe 061::63 - Instituição de pesquisa governamental com funções públicas de inovação tecnológica, cujo objetivo é gerar conhecimento e tecnologia na área da agricultura". Esse processo se repetiu para os demais termos, cada um sendo acoplado a uma classe da CDU, conforme detalhado no capítulo 5.

Percebeu-se que o único termo do tesouro que não foi possível compatibilizar foi o termo "Manejo do solo". Isso porque o conceito desse termo pode estar em duas classes da CDU que corresponde a máquinas e adubos, ou seja, a classe 631.31, que diz respeito a "Máquinas e implementos para cultivo do solo. Máquinas e implementos auxiliares; e a classe 631.8 - Fertilizantes, adubos e enriquecimento do solo. Promoção do crescimento. Estimulantes vegetais". Essa constatação confirma as afirmações de Lancaster e Warner

(1993) de que a estrutura classificatória da CDU possibilita o controle de ambiguidades, visto que um termo precisa ser direcionado ao real sentido que expressa, não podendo estar em mais de uma classe, desde que esteja delimitado por remissivas. Nesse sentido, o sistema, mesmo sendo flexível, podendo ser relacionado a outros vocabulários, conforme evidências deste estudo, possui recursos que auxiliam os problemas comuns da linguagem natural, como homônimos e polissemias (LANCASTER; WARNER, 1993).

A lógica estrutural do sistema precisa indicar o contexto do termo na estrutura da organização do conhecimento, além de possibilitar modificações na estrutura, caso seja necessário, tal como a inserção de termos e variações que venham a denotar mais de um conceito. A não correspondência do termo "Manejo do solo" com as classes da CDU considerando a variedade de conceitos atribuído ao termo é um indicativo de que o sistema tem suas classes de assunto organizadas a partir do sentido, ou seja, leva em consideração as relações semânticas, contribuindo, assim, para a navegação do usuário, na busca por informação. Como também relatou Khoo e Na (2006), as expressões geradas são complexas, semelhantes à linguagem natural; no entanto, a base de relacionamento entre termos e seus respectivos conceitos está no sentido semântico.

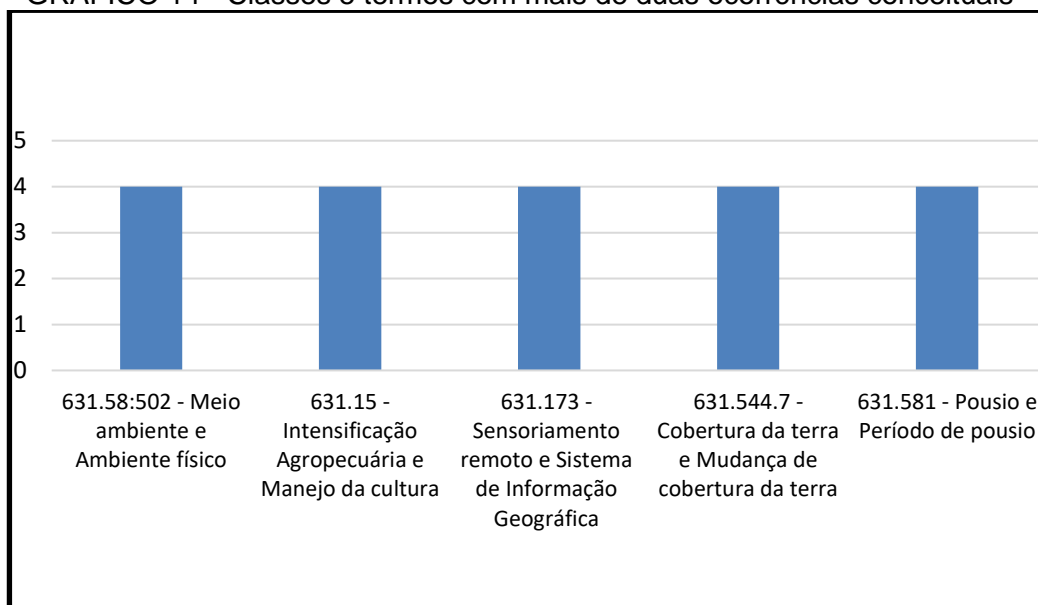
O estudo de Frâncu (2003a, 2003b), ao comparar dois tesouros com a CDU, também salientou a estrutura lógica do sistema classificatório, embora a autora não detalhou as relações semânticas que poderiam ser estabelecidas entre as classes da CDU e os termos nos dois tesouros multilíngues. De qualquer modo, a compatibilidade das linguagens, considerando a correspondência entre classes e termos, conforme os resultados apontados nesta pesquisa, ampliarão as possibilidades de conexão entre os elementos da CDU e do AgroTermos, impactando, assim, na recuperação. Portanto, o fato da maioria das classes da CDU estarem correspondidas com os termos do AgroTermos evidencia que o usuário terá mais possibilidade de refinar suas buscas, além de os termos por ele escolhidos permitirem o cruzamento entre as diferentes classes, fortalecendo, nesse sentido, a localização do item procurado.

A constatação de que a compatibilidade entre os dois SOC irá melhorar a recuperação foi presumida por Dahlberg (1981, 1983) e confirmada por Frâncu (2003a, 2003b), sendo reforçada com estes novos resultados. As classes da CDU e os termos do tesouro são compatíveis e podem ajudar no estabelecimento de outros termos, ligados por relações semânticas, o que ampliará os modos de representação e recuperação da informação. Sobre isso, Frâncu (2003a, 2003b) constatou que a utilização de tesouros com maior nível de detalhamento permite aos usuários expandir os resultados obtidos em suas pesquisas e com o tesouro com baixo nível de especificidade, irá restringir os resultados durante a busca, conforme as necessidades de informação dos usuários.

Outro resultado obtido a partir da identificação das ocorrências entre as classes e

os termos do tesauro diz respeito à prevalência de mais de duas ocorrências conceituais, resultado esse que, novamente, reforça a capacidade de ampliação dos níveis de relacionamento entre os termos. Assim, das 23 classes, 18 classes tiveram duas ocorrências entre os termos das classes da CDU e os termos do tesauro AgroTermos, o que representa 78% do total, e cinco classes obtiveram quatro ocorrências (ou coincidências conceituais), o que corresponde a 22%. Esse último resultado pode ser visualizado no Gráfico 14.

GRÁFICO 14 - Classes e termos com mais de duas ocorrências conceituais



FONTE: Dados da pesquisa (2021).

A evidência de que cinco termos podem se referir a mais de um conceito, conforme exposto no Gráfico 14, reforça a necessidade de controle terminológico para esses termos, como afirmado por Maia (2018). É nesse sentido que os SOC, sobretudo os tesouros, estabelecem relações de vários tipos, além das hierárquicas, sendo complementados por notações específicas que demonstram esses tipos, possibilitando um entendimento único e objetivo para o usuário, no momento da busca.

Foi possível correlacionar, também, durante o processo de compatibilidade, as classes da CDU com os descritores do tesauro, agrupados por categoria. Reforça-se que o estabelecimento das categorias foi proposto no estudo de Maculan (2015), cujo propósito dessa categorização é agrupar os termos que representam os conceitos do tesauro. O Quadro 43 descreve esse alinhamento com as classes da CDU por categorias.

QUADRO 43 - Classes da CDU e termos do AgroTermos por categorias

Categorias Maculan (2015)	CDU	AgroTermos
Categoria 9 - Instituições	"<" 061::63	Embrapa
Categoria 5 - Agronomia	631.15	Intensificação Agropecuária e Manejo da cultura
Categoria 7 - Socioeconomia	631.15:314.8.061	Densidade demográfica
Categoria 7 - Socioeconomia	631.15:314.8.062	Pressão populacional

Categoria 7 - Socioeconomia	631.15:314.82	Crescimento populacional
Categoria 8 - Espaço e Tempo	631.15:519.246.8	Análise de séries temporais
Categoria 6 - Território e paisagem	631.115	Posse da terra e Escassez da terra
Categoria 1 - Agricultura extensiva	631.151.1	Agricultura extensiva
Categoria 2 - Agricultura intensiva	631.151.2	Agricultura intensiva
Categoria 8 - Espaço e Tempo	631.153.7	Mudança agrícola
Categoria 3 - Material e Métodos	631.173	Sistema de Informação Geográfica e Sensoriamento remoto
Categoria 3 - Material e Métodos	631.23	Cultura
Categoria 4 - Ambiente	631.4	Solo
Categoria 6 - Território e Paisagem	631.544.7	Cobertura da terra e Mudança de cobertura da terra
Categoria 5 - Agronomia	631.547	Ciclo da cultura
Categoria 5 - Agronomia	631.57	Produto agropecuário
Categoria 5 - Agronomia	631.58	Cultura anual
Categoria 4 - Ambiente	631.58:502	Meio ambiente e Ambiente físico
Categoria 5 - Agronomia	631.581	Pousio e Período de pousio
Categoria 5 - Agronomia	631.86	Biomassa
Categoria 5 - Agronomia	631.874	Adubo verde
Categoria 5 - Agronomia	">" 632.95	Pesticida
Categoria 5 - Agronomia	">" 636	Pecuária

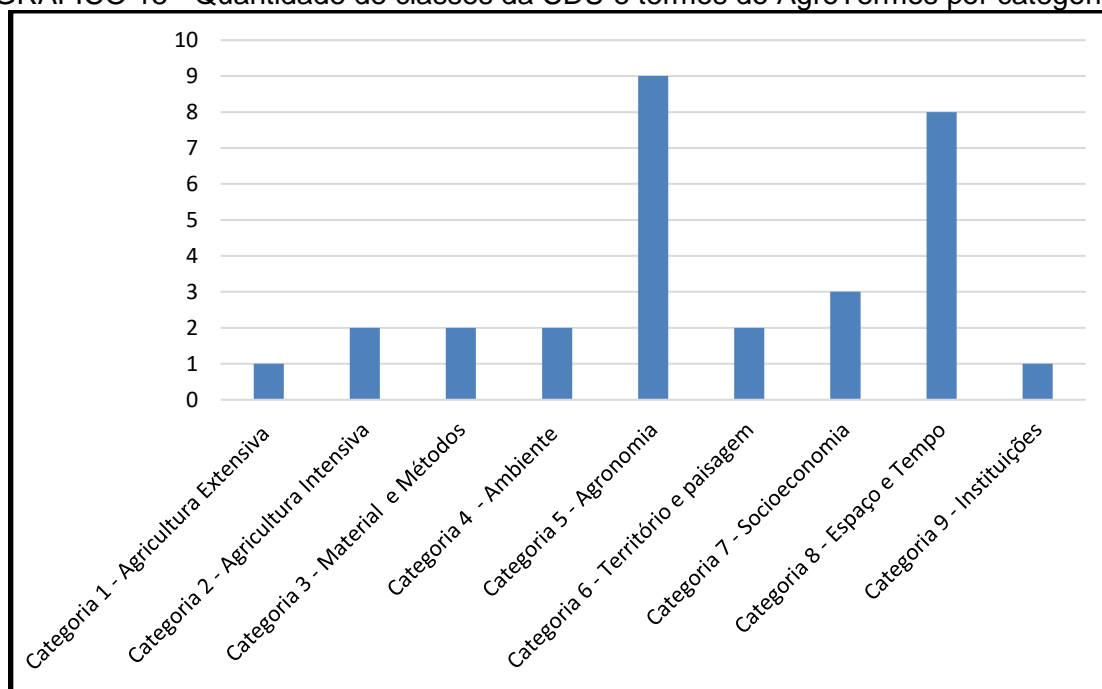
FONTE: Dados da pesquisa (2021).

Pelas demonstrações do Quadro 43, nota-se que, para a categoria 1 “Agricultura extensiva”, foi inserida apenas a compatibilidade da classe “631.151.1 - Sistema de produção agrícola extensivo” com o termo “Agricultura extensiva” do tesouro. Do mesmo modo, identifica-se que, na categoria 2, foi inserida a classe “631.151.2 - Sistema de produção agrícola intensivo” com o termo “Agricultura intensiva”. Em linhas gerais, os dados do Quadro 43 indicam que todas as classes foram correlacionadas com as categorias dos termos. Para o termo “Intensificação Agropecuária” do AgroTermos, como trata-se da temática representada, pelo estudo de Maculan (2015) não foi inserido em nenhuma categoria. Por este motivo, decidiu-se inserir este termo na categoria 5 “Agronomia”, na qual está inserido o termo Manejo da cultura, que estão compatibilizados com a classe 631.15 - Produção da Fazenda. Organização de fazendas. Administração de fazendas”.

O Gráfico 15 mostra a quantidade de classes da CDU e termos do AgroTermos agrupados por categoria do total . A categoria 5 “Agronomia” obteve a maior quantidade classes e termos compatibilizados nos dois instrumentos (9), seguida pela categoria 8

“Espaço e Tempo” (8). As demais categorias apresentaram uma quantidade reduzida de correspondências (entre 1 e 3).

GRÁFICO 15 - Quantidade de classes da CDU e termos do AgroTermos por categorias



FONTE: Dados da pesquisa (2021).

A leitura do Gráfico 15 revela que, à categoria 1 “Agricultura extensiva”, foi inserida uma classe da CDU: 631.151.1 e o termo “Agricultura extensiva” do AgroTermos, referente ao total 29 termos do AgroTermos que foram associados a classes da CDU. Na categoria 2 - “Agricultura intensiva”, foi atribuído o termo “Agricultura intensiva” do AgroTermos a essa categoria, nas duas categorias descritas, correspondem a categorias com menor número de classes e termos do associados, 3% do total de cada uma. Outras informações retiradas do Gráfico 15 dizem respeito às categorias com maior quantidade de classes e termos associados. Portanto, predomina a categoria 5 “Agronomia” (31%), seguida pela categoria 8 “Espaço e Tempo” (28%) e a categoria 7 “Socioeconomia” (11%).

Importante ressaltar que as categorias estabelecidas por Maculan (2015) possibilitaram sintetizar os conceitos que representam o campo da “Intensificação Agropecuária”, a partir de princípios de compatibilidade de linguagem postulados de por Dahlberg (1981) e pela Teoria da Classificação Facetada, ao considerarem que cada classe pode ser, também, caracterizada como um conceito no âmbito do domínio (MACULAN, 2015). A partir da análise às categorias, foi possível verificar que elas manifestam-se como facetas e, como tal, estão condicionadas a assumir diferentes direções dentro de um conjunto sistemático de itens. Isso evidencia, fortemente, o potencial das relações semânticas para expansão de um assunto, sem perder o significado do conceito, a partir do contexto que está sendo representado. Como afirmado pela autora,

analisando-se as características da área temática Intensificação Agropecuária, percebeu-se que ela é composta por processos inter-relacionados e, para entendê-la, é preciso levar em consideração diversas variáveis, tais como ambientais, agrônômicas, territoriais, socioeconômicas e os contextos institucionais específicos [...] (MACULAN, 2015, p. 206).

Constata-se, a partir desses resultados, que a fusão entre os dois SOC condicionará a existência de diversos tipos de relações. Isso evidencia que os dois sistemas, por serem compatíveis, “conversarão entre si”, sem perda de informação, o que os torna interoperáveis, como destacado no estudo de Doerr (2001). O efeito desse processo é garantir a semântica entre os vocabulários compatibilizados (CHAN; ZENG, 2002, 2006), o que indica a prevalência das relações ou o estabelecimento de novas relações, entre os termos e as classes.

Nesse contexto, evidencia-se a característica relacional da estrutura gerada com o processo de compatibilidade, sendo possível identificar relações semânticas nessa nova estrutura ou modelo gerado. Constatou-se a presença de boa parte das relações semânticas mencionadas no estudo de Maia (2018), considerando os três grandes grupos: hierárquica, equivalência e associativa. Nesses grupos, foram identificados alguns tipos de relações, como: sinônimo parcial, atividade subordinada, componente complexo, dentre outros. O Quadro 44 discrimina os grupos e os tipos de relações semânticas identificados no processo de compatibilidade de linguagem entre os dois SOC.

QUADRO 44 - Tipos de relações semânticas por classe da CDU e termos do AgroTermos

CDU	AgroTermos	Relações semânticas entre CDU e AgroTermos - Maia (2018)
"<" 061::63	Embrapa	Hierárquica - Hiperônimo - Hipônimo - Instância
631.15	Intensificação Agropecuária e Manejo da cultura	Hierárquica - Hiperônimo - Hipônimo - Inclusão de Classe - Atividade subordinada e Hierárquica - Hiperônimo - Hipônimo - Inclusão de Classe - Funcionalmente subordinado
631.15:314.8.061	Densidade demográfica	Equivalência - Sinônimo parcial
631.15:314.8.062	Pressão populacional	Hierárquica - Holônimo - Merônimo - Objeto estruturado - Segmento-evento
631.15:314.82	Crescimento populacional	Equivalência - Sinônimo parcial
631.15:519.246.8	Análise de séries temporais	Equivalência - Sinônimo parcial
631.115	Posse da terra e Escassez da terra	Equivalência - Sinônimo parcial e Associativa - Causal - Agentação
631.151.1	Agricultura extensiva	Equivalência - Sinônimo próximo

631.151.2	Agricultura intensiva	Equivalência - Sinônimo próximo
631.153.7	Mudança agrícola	Equivalência - Sinônimo absoluto
631.173	Sistema de Informação Geográfica e Sensoriamento remoto	Hierárquica - Hiperônimo- Hipônimo - Inclusão de Classe - Funcionalmente subordinado e Hierárquica - Hiperônimo- Hipônimo - Inclusão de Classe - Funcionalmente subordinado
631.23	Cultura	Hierárquica - Hiperônimo- Hipônimo - Inclusão de Classe - Atividade subordinada
631.4	Solo	Associativa - Disciplina - fenômeno
631.544.7	Cobertura da terra e Mudança de cobertura da terra	Equivalência - sinônimo próximo e Associativa- Causal - Agente-objeto
631.547	Ciclo da cultura	Equivalência - Quase sinônimo
631.57	Produto agropecuário	Hierárquica - Hiperônimo - Hipônimo - Inclusão de classe- taxonômica
631.58	Cultura anual	Hierárquica - Hiperônimo- Hipônimo - Inclusão de Classe - Atividade subordinada
631.58:502	Meio ambiente e Ambiente físico	Equivalência - Sinônimo parcial e Hierárquica - Holônimo - Merônimo - Objeto estruturado - Componente complexo
631.581	Pousio e Período de pousio	Equivalência - Sinônimo absoluto e Equivalência - Quase sinônimo
631.86	Biomassa	Equivalência - Quase sinônimo
631.874	Adubo verde	Associativa - Processo- instrumento
">" 632.95	Pesticida	Equivalência - Sinônimo absoluto
">" 636	Pecuária	Equivalência - Sinônimo parcial

FONTE: Dados da pesquisa (2021).

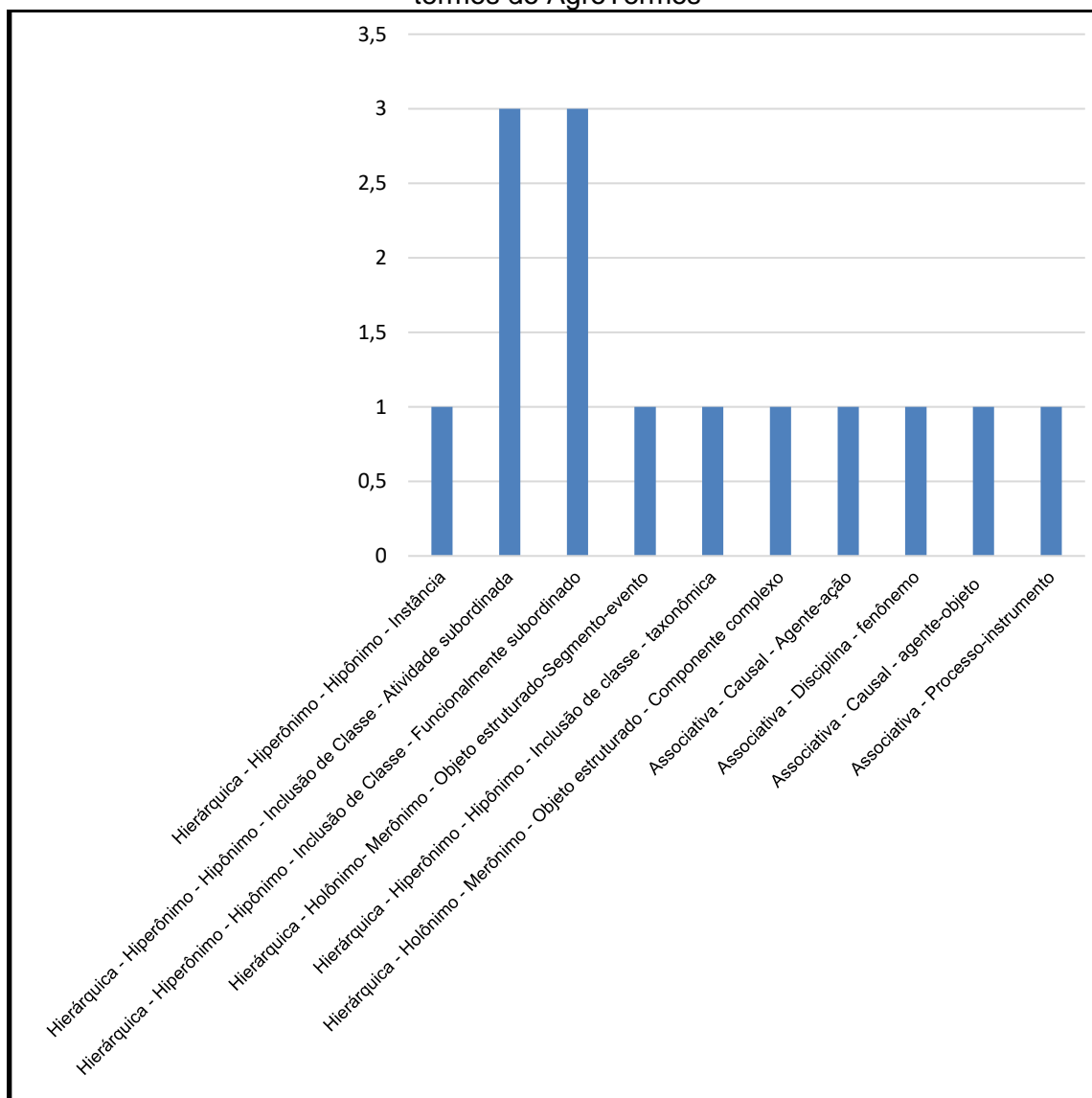
Nota-se, no Quadro 44, que, embora os três grupos principais de relações tenham sido identificados, nem todos os tipos da proposta de Maia (2018) - Anexo D, foram identificados nesta análise. Foi possível estabelecer relações em quatro classes da CDU (061, 631, 632 e 636), cujo total de relações firmadas foi de 29. A relação de equivalência (sinônimo parcial) predominou, com 6 ocorrências, seguida pelo restante das relações de equivalência (Sinônimo próximo, Sinônimo absoluto e Quase sinônimo) que obteve um total de três ocorrências cada, totalizando 15 ocorrências. Por fim, as relações Hierárquicas e Associativas

tiveram dez e quatro ocorrências respectivamente. Mesmo que nem todas as classes e termos tenham estabelecido algum tipo de relação, a quantidade encontrada indica certa sintonia entre os SOC. Esse resultado reforça a afirmação da autora, ao afirmar que “as relações semânticas são fundamentais para a compreensão da natureza da ligação entre dois conceitos em um domínio [...]”, cujos conceitos existentes em uma estrutura classificatória, quando alinhados, “[...] uma relação semântica entre eles pode existir de fato [...]” (MAIA, 2018, p. 8).

Ao estabelecer e ampliar o número de relações, é importante definir os níveis de exaustividade e especificidade entre os conceitos (FRÂNCU, 2003b). Embora esses níveis não foram abordados no modelo proposto, é importante salientar da importância deles, visto que esse critério interfere na recuperação. Quanto maior a especificidade, maior a probabilidade de o sistema atingir alta relevância. No contrário, um sistema que permite especificidade limitada tem probabilidade de atingir uma recuperação razoavelmente alta, mas com baixa relevância. A falta de especificidade não pode ser corrigida na fase de produção, cujo resultado é a necessidade imposta ao usuário de navegar de modo desnecessário para encontrar ou descobrir as informações que realmente são relevantes (FRÂNCU, 2003b).

Os dados constantes no Quadro 44 podem ser representados em indicadores percentuais, mediante uma formatação gráfica. Portanto, o Gráfico 16 ilustra essa apresentação, com destaque os três grupos de relações predominantes.

GRÁFICO 16 - Quantidade de relações semânticas por tipo entre as classes da CDU e os termos do AgroTermos



FONTE: Dados da pesquisa (2021).

Destaca-se, pelos dados expostos no Gráfico 16, que as relações de equivalência do tipo (sinônimo parcial, sinônimo próximo, sinônimo absoluto e quase sinônimo) abarcaram o maior número de ocorrências com o total de 6 (21%), 3 (11%), 3 (11%) e 3 (11%) respectivamente. Portanto, o número total de ocorrências das relações de equivalência resultaram em um total de 54%, o que evidencia a contribuição desse tipo de relação nos processos de representação e recuperação da informação, sobretudo com uso da CDU e tesauros. Isso faz inferir acerca do papel exercido por cada um dos SOC no que tange à aproximação entre o que está sendo representado, a partir da sua semântica, mesmo que a sintática (escrita) seja totalmente diferente. Portanto, aproximar os assuntos quando fazem referência a um mesmo conceito, ou seja, quando denotam o mesmo significado, constitui uma das funções da classificação bibliográfica, como salientou Piedade (1977) e Vickery (1980).

Além da função de aproximar os assuntos no contexto de um sistema classificatório, como pontuado por Piedade (1977), as relações de equivalência tentam eliminar a polissemia e a ambiguidade dos conceitos, tendo em vista garantir “[...] a monosssemia entre a forma do significante e a do significado” (CINTRA *et al.*, 2002, p. 71). Com efeito, reforça-se, como na pesquisa de Maia (2018, p. 62), acerca da importância desses tipos de relações, sobremaneira, por evitar a duplicidade de termos que expressem o mesmo conceito, visto que “[...] os significados ou sentidos (conceitos por trás das palavras) são os mesmos ou são considerados equivalentes em um dado contexto [...]”. Além disso, estas relações dependem da definição do termo preferido, onde os usuários são conduzidos, mesmo utilizando termos não preferidos a fazer a busca pelos termos preferidos, portanto não existe diferença entre os dois tipos. Dessa forma, a relação de equivalência permite recuperar tanto utilizando termos preferidos ou não preferidos a mesma quantidade de recursos, afetando diretamente na exaustividade e especificidade, ou seja, na melhoria das taxas de revocação e precisão durante a recuperação da informação (WEISS; BRÄSCHER, 2017).

Em síntese, os resultados oriundos do processo de compatibilidade revelam que há correspondências entre as classes da CDU e os termos do AgroTermos, com possibilidades de se estabelecer outras relações semânticas em relação as 29 relações estabelecidas entre os dois SOC. Os outros tipos de relações semânticas utilizadas tiveram menor abrangência, que foram as relações hierárquicas, que estabelecem correspondências mais abrangentes e específicas, de forma pragmática, ou seja, com conceitos predefinidos. As relações hierárquicas estabelecidas foram: Hiperônimo - Hipônimo – Instância; Hiperônimo - Hipônimo - Inclusão de Classe - Atividade subordinada; Hiperônimo - Hipônimo - Inclusão de Classe - Funcionalmente subordinado; Holônimo- Merônimo - Objeto estruturado-Segmento-evento; Hiperônimo - Hipônimo - Inclusão de classe - taxonômica e Holônimo - Merônimo - Objeto estruturado - Componente complexo, com a quantidade de ocorrências: 1 (3%), 3 (11%), 3 (11%), 1 (3%), 1 (3%) e 1 (3%), respectivamente, totalizando 34%.

Nas relações hierárquicas ao termos envolvidos devem-se referir para um todas as situações somente para algumas, por exemplo, todo funcionário trabalha em um departamento e todo departamento tem funcionários. No outro caso, todo professor já foi aluno, mas nem todo aluno é professor, neste último caso não seria utilizada uma relação de hierarquia, poderia ser utilizada uma relação associativa.

Nas relações hierárquicas (todo-parte), onde parte de um conceito se refere a um todo exclusivo, não podendo este todo ser representado por dois termos, por exemplo, toda roda pode fazer parte do conceito carro, bicicleta, moto, avião, entre outros, devendo também, se for o caso deve-se estabelecer uma relação associativa. Para as relações de instância, dois termos com nomes próprios são representados por substantivos que estão subordinados a uma classe genérica, por exemplo, os termos João e Maria estão subordinados a classe

pessoa. Por fim, quando um termo está associado a mais de uma classe, trata-se de um relacionamento polihierárquico, por exemplo Bioquímica está associado a Biologia e a Química (WEISS; BRÄSCHER, 2017). Este último caso, ocorreu durante o processo de compatibilidade, onde o termo “Manejo do solo”, que corresponde a máquinas e adubos do AgroTermos está relacionada a duas classes na CDU, mas não possui uma classe na CDU que represente estes dois significados desse termo.

E por fim, o último tipo de relação semântica foram as associativas (sintagmáticas), estabelecidas entre termos que estão em diferentes níveis de hierarquia do vocabulário nos dois SOC. As relações associativas, apesar de possuírem um maior diversidade de tipos de acordo com Maia (2018), na representação do modelo tiveram a menor representatividade entre todos os outros tipos de relacionamentos, sendo usadas apenas as relações associativas causal - agente - ação e causal, agente - objeto, Disciplina-fenômeno e Processo-instrumento, representando um total de 12%. Conforme argumentado por Weiss e Bräscher (2017), as relações associativas estão relacionadas a pares de termos, que não possuem relação de hierarquia, porém estão associados pelo seu significado semântico, seguindo o princípio da exclusão mútua, caso estejam na mesma hierarquia.

Dos 30 termos preferidos que representam a temática “Intensificação Agropecuária”, mapeados por Maculan (2015), diversos relacionamentos foram estabelecidos entre termos não preferidos, de modo a contribuir para a sistematização da estrutura ou modelo proposto. Quando se mapeia os conceitos, infere-se que os relacionamentos semânticos entre os termos não preferidos funcionam como pontos de acesso para cada termo que representa esses conceitos, ou seja, estes termos conduzem o usuário a chegar no termo preferido de cada assunto representado do recorte temático do domínio. A possibilidade de aproveitar esses relacionamentos e estabelecer outros é benéfica, desde que mantida a lógica do sistema, pois essa expansão auxiliará o usuário ter acesso a informação a partir de um número maior de termos durante a busca (MACULAN, 2015), como também permitirá a inserção de novos assuntos ou disciplinas que vão surgindo (FRÂNCO, 2003b).

Logo, no que tange à recuperação da informação,

[...] a possibilidade de refinamento das relações entre conceitos, através da sua explicitação na estrutura do tesauro, auxiliou a organização do conhecimento do domínio modelado, o que pode facilitar a sua exploração pelo usuário, uma vez que torna esses relacionamentos mais claros e específicos (MACULAN, 2015, p. 9).

Por sua vez, quanto à expansão dos assuntos e do sistema, independentemente das barreiras linguísticas, a atualização contínua dos termos é altamente necessária, visto que foram construídos em princípios fundamentais. A evolução dinâmica do conhecimento e os novos conceitos em evolução em cada disciplina ou domínio impõem um ajuste rápido e eficaz do conhecimento (SLAVIC, 2005).

Com a junção entre dois ou mais SOC é necessário desambiguar e contextualizar as consultas para promover melhorias na recuperação, levando em conta a busca e a navegação de assuntos na CDU em ambiente on-line (SLAVIC, 2005). O conjunto de termos e relações quando compatíveis, levando em conta a flexibilidade da CDU e a lógica das relações que são estabelecidas, conseqüentemente irá melhorar a recuperação. Isso reforça que o modelo de compatibilidade proposto entre a CDU e o AgroTermos com essa combinação, será viável para melhorar a recuperação no ambiente on-line.

Além do modelo se apresentar como consistente em termos de compatibilidade e de lógica nas relações, por contemplar as classes da CDU, que se apresentam como flexíveis, o modelo poderá ser melhorado, com inserção de novas classes, termos e relações. Slavic (2005) também sinalizou essa possibilidade, mesmo que não tenha investigado o potencial do sistema classificatório em junção com um tesouro em um domínio específico, apenas para fins de armazenamento de termos em bases de dados. Para a autora, o conjunto de termos na forma de um tesouro pode ser explorado, posteriormente, de modo a facilitar a busca e a navegação de assuntos da CDU, que deve ser armazenada e distribuída em formato de banco de dados *Master Reference File* (UDC MRF) para seu uso em ambiente on-line.

Nesse contexto, entende-se que os modelos de representação ou de recuperação da informação precisam ser estruturados com relações claras e lógicas, com o propósito de permitir mais semântica à estrutura, especificando o tipo de relação existente entre os termos e seus respectivos conceitos (FRÂNCO, 2003a, 2003b). De modo similar, os resultados da compatibilidade das classes com os termos do AgroTermos confirmam a possibilidade de haver ambigüidades, visto que a sintática de um termo é a mesma, mas equivale a diferentes significados. Logo, o uso das notações específicas (Frâncu e Sabo, 2010) faz-se necessário, sobretudo com a utilização de diferentes SOC para viabilizar a recuperação da informação, visto que os dois sistemas precisam ser interoperáveis, como constatado na pesquisa de Maculan (2015), acerca da recomendação sobre a construção dos tesouros.

As evidências do estudo de Maculan (2015) e de Maia (2018) se aproximam por constatar inconsistências quanto ao uso de termos que denotam os mesmos conceitos sem a presença de remissivas e/ou notações que promovam a desambiguação. Neste estudo, embora se utilize o mesmo recorte temático de Maculan, reforça-se a importância do uso de dois SOC para potencializar a recuperação. Para isso acontecer, os elementos desses instrumentos precisam ser compatíveis, por conseguinte, tornar os sistemas interoperáveis, haja vista contribuir para melhorar a recuperação em ambientes digitais, como foi recomendado por Dahlberg (1981,1983); Zeng e Chan (2004); Zeng (2019), entre outros e evidenciado com o processo de compatibilidade apresentado.

6.3 Representação do modelo proposto na linguagem SKOS

Com os resultados alcançados a partir do processo de mapeamento do recorte temático e da compatibilidade de linguagem, foi possível explorar, com mais detalhes, o produto final desta pesquisa, qual seja: o modelo de compatibilidade representado na linguagem SKOS com uso dos dois SOC, com vistas à interoperabilidade. Salienta-se que modelo, no contexto desta tese, é entendido a partir da definição mencionada por Lima (2004), ao conceituar modelos de representação da informação. Segundo a autora, modelo é definido como a representação abstrata e simplificada de um sistema real, de modo a possibilitar uma síntese esquemática dos processos de organizar, acessar e recuperar a informação, processos esses considerados complexos, o que justifica a formalização de um modelo.

Sendo assim, neste subcapítulo, o produto final da tese foi a representação do modelo proposto na linguagem SKOS, sobretudo em ambientes digitais, tendo como base o uso de dois SOC, visando a interoperabilidade entre diferentes sistemas. Esta representação do modelo é explicitada, cujos resultados são analisados e descritos a seguir (Objetivo específico 3).

Com esse fim, a princípio, o software é apresentado no subcapítulo 3.1.2.1, seguido pela exposição e detalhamento dos elementos do modelo. No final deste subcapítulo, é apresentada a aplicação do modelo proposto no *TemaTres*, com a exportação dos padrões de metadados no formato SKOS. Este formato de linguagem, se aproxima da estrutura relacional da CDU e do tesauro AgroTermos com seus relacionamentos de equivalência, hierárquicos e associativos, sendo considerado um modelo de dados para que os dados entre diferentes sistemas de informação sejam interoperáveis (CLARKE; ZENG, 2012).

Considerando os dados coletados com o mapeamento do recorte temático, a amostra de pesquisa foi inserida no *TemaTres*. A princípio, foram inseridos os termos das categorias oriundos do mapeamento. A atividade inicial contemplou o preenchimento dos campos para inserção dos termos das classe da CDU e do AgroTermos no sistema e suas relações semânticas e organizada a partir dos níveis hierárquicos do sistema classificatório. Para cada categoria, foram distribuídas as classes e subclasses da CDU para a aplicação do modelo proposto em ordem de citação correspondente à estrutura da CDU, ou seja, seguindo a hierarquia do sistema classificatório, que parte do geral para o específico. Confirma-se, portanto, que as classes da CDU são processadas pela máquina, cujos relacionamentos são validados pelo sistema de informação e podem ser utilizados para recuperação da informação, tanto em ambientes físicos quanto digitais, como foi revelado, também, nos achados de Slavic (2005).

As descobertas deste estudo estão em sintonia com boa parte das evidências encontradas por Slavic (2005), confirmando os pressupostos de que a CDU tem potencial

para uso em diferentes ambientes, particularmente portais de assuntos e catálogos de bibliotecas (SLAVIC, 2005). Se a conversão das classes não ocorrer de modo legível pela máquina, impede-se a implementação do sistema on-line e, portanto, aumenta o custo de uso (SLAVIC, 2005), além de acarretar falhas na recuperação pelo usuário.

Após a inserção das classes e subclasses nos campos do *TemaTres*, os termos foram inseridos, cujo sistema estabeleceu o relacionamento entre cada termo e sua classe correspondente, além de evidenciar possíveis relações. A inserção desses termos e a combinação com as classes, além de confirmarem a compatibilidade, também confirmam que a classificação tem função direcionada à recuperação de informações, por sua estrutura lógica, a qual pode ser usada favoravelmente na construção de ferramentas atualizadas para permitir acesso aprimorado à informação, como foi comprovado nos resultados do estudo de Frâncu (2003a, 2003b).

Quanto às relações semânticas formadas com a junção das classes com os termos, os resultados confirmaram a revelação de Frâncu (2003a, 2003b) e de Maia (2018), acerca da importância das relações semânticas e de links no fortalecimento da compatibilidade entre os vocabulários. Conforme já mencionado anteriormente, a Figura 29 apresenta o editor de relações e tipos de links no *TemaTres*.

FIGURA 29 - Tela do *TemaTres* - Geração das relações e de links

Editor de relações ▲

Tipo de relação	Legenda do tipo de relação:	Código do tipo de relação	Ordem	
TG/TE ▼	<input type="text"/>	<input type="text"/>	<input type="text"/>	Enviar
TR	Associativa - Causal - Agente - ação	AA	7	Eliminar relação
TR	Associativa - Causal - Agente - Objeto	AO	9	2
TG/TE	Hierárquica - (Hiperônimo - Hipônimo) - Inclusão de Classe - Atividade subordinada	AS	4	1
TG/TE	Hierárquica - (Holônimo - Merônimo - Objeto estruturado - Componente complexo	CC	5	Eliminar relação
TG/TE	Hierárquica - (Hiperônimo - Hipônimo) - Inclusão de Classe - Funcionalmente subordinado	FS	6	4
TG/TE	Associativa - Causal - Agente - objeto	AO	8	Eliminar relação
UP/USE	Equivalência - Sinônimo absoluto	SA	1	7
UP/USE	Equivalência - Sinônimo parcial	SP	2	4
UP/USE	Equivalência - Quase sinônimo	QS	3	1
9				

Editor de tipos de links ▲

Legenda do tipo de link:	Endereço do tipo de link	
<input type="text"/>	<input type="text"/>	Enviar
broadMatch	broadMatch	3
closeMatch	closeMatch	Excluir tipo de link
exactMatch	exactMatch	14
narrowMatch	narrowMatch	1
relatedMatch	relatedMatch	Excluir tipo de link
5		

FONTE: Dados da pesquisa (2021).

Nota-se, pela Figura 29, que as relações são estabelecidas, conforme os termos

do tesouro e que é possível atribuir links a elas, potencializando, assim, o uso dessas relações para navegação no ambiente digital. Constata-se, novamente, que a junção entre CDU e tesouro agregará valor para a recuperação da informação, sobretudo no contexto da web semântica. Slavic (2005) justificou essa constatação, afirmando que a vantagem da estrutura analítico-sintética e semi-facetada da CDU é que ela é menos rígida quando se trata de organizar o conhecimento baseado em disciplinas, uma vez que permite uma grande flexibilidade no relacionamento entre os assuntos de cada disciplina, por meio de relações interdisciplinares e multidisciplinares, o que facilita expandir os assuntos e incluir novos. O modelo baseado na CDU pode ser usado para apoiar a navegação por assunto, por meio de relações hierárquicas e associativas, na indexação pré-coordenada e pesquisa pós-coordenada, além de permitir a expansão da pesquisa pelo usuário (SLAVIC, 2005).

Apresentada a aplicação do modelo no *TemaTres* e as evidências da compatibilidade entre os dois SOC, como também o potencial para a recuperação da informação no ambiente digital, detalham-se os elementos do modelo gerado com esta pesquisa. O modelo gerado está constituído por quatro partes, esquematizadas no sentido vertical, a saber: categorias das classes, classes e subclasses, termos e relações. A representação gráfica do modelo permite constatar a ocorrência da compatibilidade, visto que as classes, seus respectivos termos e relações estão interligados e sequenciados de modo horizontal. O Quadro 45 detalha os elementos do modelo e as devidas correspondências entre eles, confirmando a ocorrência da compatibilidade.

QUADRO 45 - Detalhamento dos elementos do modelo proposto

Categorias	Número das classes	Assunto das classes	Termos	Relações
Genérica	061::63	Instituição de pesquisa governamental com funções públicas de inovação tecnológica, cujo objetivo é gerar conhecimento e tecnologia na área da agricultura	Embrapa	Hierárquica - Hiperônimo - Hipônimo - Instância
	631.15	Produção da fazenda. Organização de fazendas. Administração de fazendas	Intensificação Agropecuária	Hierárquica - Hiperônimo - Hipônimo - Inclusão de Classe - Atividade subordinada
			Manejo da Cultura	Hierárquica - Hiperônimo-Hipônimo - Inclusão de Classe - Funcionalmente subordinado

631.15:314.8.061	Administração de fazendas com base na Densidade demográfica	Densidade demográfica	Equivalência - Sinônimo parcial
631.15:314.8.062	Organização de fazendas a partir dos componentes das mudanças demográficas com aumento ou decréscimo natural.	Pressão populacional	Hierárquica - Holônimo-Merônimo - Objeto estruturado-Segmento-evento
631.15:314.82	Produção da fazenda baseada no crescimento da população	Crescimento populacional	Equivalência - Sinônimo parcial
631.15:519.246.8	Produção da fazenda com a utilização de análise das séries temporais, autocorrelação e regressão	Análise de séries temporais	Equivalência - Sinônimo parcial
631.115	Fazendas segundo o título de posse, propriedade e supervisão. Sistemas de administração	Posse da terra	Equivalência - Sinônimo parcial
		Escassez da terra	Associativa - Causal - Agente - ação
631.151.1	Sistema de produção agrícola extensivo	Agricultura Extensiva	Equivalência - Sinônimo próximo
631.151.2	Sistema de produção agrícola intensivo	Agricultura Intensiva	Equivalência - Sinônimo próximo
631.153.7	Mudanças no cultivo e produção. Conversão para outros sistemas	Mudança agrícola	Equivalência - Sinônimo absoluto
631.173	Organização de equipamentos e serviços especiais	Sistema de Informação Geográfica	Hierárquica - Hiperônimo - Hipônimo - Inclusão de Classe - Funcionalmente subordinado
		Sensoriamento remoto	Hierárquica - Hiperônimo - Hipônimo - Inclusão de Classe - Funcionalmente subordinado

	631.23	Estruturas de criação e cultivo de plantas	Cultura	Hierárquica - Hiperônimo-Hipônimo - Inclusão de Classe - Atividade subordinada
	631.4	Ciência do solo. Pedologia. Pesquisa de solo	Solo	Associativa - Disciplina - fenômeno
	631.544.7	Cultivo sob várias coberturas. Proteção e cobertura do solo	Cobertura da terra	Equivalência - sinônimo próximo
			Mudança de cobertura da terra	Associativa - Causal - Agente - objeto
	631.547	Fases e controle do crescimento das plantas	Ciclo da cultura	Equivalência - Quase sinônimo
	631.57	Produtos, subprodutos e produtos semi-acabados da agricultura	Produto agropecuário	Hierárquica - Hiperônimo - Hipônimo - Inclusão de classe - taxonômica
	631.58	Sistemas específicos de cultura	Cultura anual	Hierárquica - Hiperônimo - Hipônimo - Inclusão de Classe - Funcionalmente subordinado
	631.58:502	Sistemas específicos de cultura e a proteção ambiental	Meio Ambiente	Equivalência - Sinônimo parcial
			Ambiente físico	Hierárquica - Holônimo - Merônimo - Objeto estruturado - Componente complexo
	631.581	Pousio	Pousio	Equivalência - Sinônimo absoluto
			Período de pousio	Equivalência - Quase sinônimo
	631.86	Adubos orgânicos em geral. Fertilizantes de origem biológica	Biomassa	Equivalência - Quase sinônimo
	631.874	Adubação verde. Cultivos que melhoram o solo	Adubo verde	Associativa - Processo-instrumento
Específicas	632.95	Pesticidas	Pesticida	Equivalência - Sinônimo absoluto
	636	Criação de gado e de animais em geral. Criação de gado. Criação de animais domésticos	Pecuária	Equivalência - Sinônimo parcial

FONTE: Dados da pesquisa (2021).

Os elementos do modelo foram alinhados entre si, conforme exposto no Quadro 45, demonstrando o processo de compatibilidade. A compatibilidade entre dois SOC

promoverá contribuições para a recuperação da informação nos ambientes digitais, em especial a compatibilidade semântica, ou seja, a correspondência dos significados que os termos exprimem, sem ambiguidades, redundância ou quaisquer inconsistências (FRÂNCU, 2003b). A correspondência exata dos termos de um vocabulário conforme Neville (1970), é a função do processo de compatibilidade, cujo efeito mais importante é a garantia da interoperabilidade entre os sistemas. Sendo compatíveis e interoperáveis, é possível garantir o intercâmbio de informações, ou seja, a troca de dados pela rede, entre computadores, softwares e serviços da web, de acordo com Zeng (2019). Portanto, impacta na recuperação da informação pelo usuário (FRÂNCU, 2003a, 2003b).

A CDU e o AgroTermos, pelo que mostra o modelo ora exposto, juntam-se com o propósito de ampliarem os relacionamentos, sem gerar problemas de ambiguidades. Isso acarretará benefícios nas buscas pelo usuário, que terá condições de ampliar ou de restringi-las conforme suas necessidades, como mencionado, também, por Frâncu (2003b). Assim, o modelo de compatibilidade poderá proporcionar melhorias a recuperação, sobretudo quando se observa a predominância das relações de equivalência. Isso indica que a CDU possui, em grande parte, descritores equivalentes aos termos do tesouro, possibilitando maior precisão durante a recuperação. Ademais, o fato de que a grande maioria das relações semânticas ocorreu dentro do recorte temático da CDU (Classe 631), demonstrando a flexibilidade do sistema classificatório na combinação de assuntos compostos ou complexos. Foi necessário a utilização de poucas classes fora do recorte temático do domínio, para a representação do assunto “Intensificação Agropecuária” do domínio da agricultura. Essa constatação reforça que o sistema precisa ser armazenado e distribuído em formato de banco de dados UDC e *Master Reference File* (MRF) para seu uso em ambiente on-line, como salientou Slavic (2005).

Outra comprovação de que o modelo proposto é adequado à recuperação diz respeito à abrangência dos termos que representam o domínio analisado. A construção de modelos ou de instrumentos de representação e recuperação, tal como as taxonomias, por exemplo, ao compilarem diferentes tipos de relações semânticas, fortalecem a capacidade de representação de um domínio que foi evidenciado por Maia (2018), como também facilitam a busca, sobretudo em domínios especializados, como o da Intensificação Agropecuária, visto que podem ser refinadas as buscas pelos usuários (MACULAN, 2015). Logo, se houve compatibilidade da linguagem com um número representativo de termos do domínio da agricultura, espera-se que o usuário encontre os assuntos que busca com mais facilidade. Para resolver o problema do termo não compatibilizado (Manejo do solo), recomenda-se a modificação na estrutura dos vocabulários, de modo a torná-los compatíveis, como foi recomendado por Slavic (2005).

A compatibilidade da linguagem entre CDU e AgroTermos, sobremaneira, quando se trata de domínio específico com uso de vocabulário adaptado a esse domínio, facilita a

recuperação, como é o caso do AgroTermos. O uso dos termos em um mesmo idioma também facilitará a recuperação, conforme relata Maniez (1997). Além disso, torna-se adequada a descrição pelos indexadores, evitando, assim, erros de digitação, por conseguinte, problemas sintáticos, evidenciado por Frâncu (2003b) e de interoperabilidade (ZENG, 2019). Evidencia-se, também, que os termos da estrutura que denotam cada conceito podem se unir à estrutura da CDU por meio da ligação da notação com seus descritores para cada um dos trinta termos. Portanto, a estrutura formada será mais robusta, permitindo recuperar mais assuntos e com mais precisão e relevância, como foi salientado por Frâncu (2003b), ao acoplar mais de um tesouro à estrutura classificatória.

O modelo proposto e descrito no Quadro 45, contribui na construção e gestão dos ambientes digitais e colaborativos. Isso porque, como pontuado por Weiss (2019), os diversos bancos de dados precisam trocar informações, sem perdas ou inconsistências, ou seja, precisam ser interoperáveis, de modo a potencializar o compartilhamento de dados e informação por diferentes instituições. O benefício é no âmbito do usuário, no momento da recuperação, e no contexto das instituições, haja vista o reúso do vocabulário e a compatibilidade dos termos (CARDILLO *et al.*, 2014), poupando, desse modo, tempo e recursos.

Conforme mencionado por Slavic (2005) e em sintonia com o potencial do modelo ora proposto, entende-se que a criação e a gestão de um ambiente colaborativo e centralizado para o compartilhamento e intercâmbio de informações entre classificações e outros SOC facilitam o acesso e o processamento dos dados ou informações por diferentes sistemas distribuídos. Como resultado desse processo, reduz-se a duplicação de esforços, visando à interoperabilidade mais eficiente entre sistemas de informação, cujos dados, além de serem legíveis por máquina, estarão em uma linguagem padronizada (SLAVIC, 2005).

A linguagem padronizada e legível por máquina é outra contribuição do modelo proposto neste estudo, sobretudo por utilizar o formato SKOS. A inserção do modelo no *TemaTres* revelou essa ocorrência, cujos elementos do modelo, como classes, termos e relações foram reconhecidos pelo software, em uma linguagem controlada e entendida pela máquina. Esse procedimento foi necessário, pois, como discorrido na literatura, o formato SKOS possibilita que os SOC sejam interpretados pelas máquinas, garantindo, portanto, a interoperabilidade desses sistemas (SOUZA; ALVARENGA, 2004).

Considerando que esta pesquisa almeja contribuir para facilitar a recuperação da informação em ambientes digitais, torna-se necessário validar o modelo proposto no que tange à padronização no formato citado. Assim, ao adquirir uma linguagem legível pelo computador, o SOC terá a capacidade de analisar adequadamente as páginas da web quanto ao layout e rotinas de processamento (hipertextos, transferência de arquivos, e-mails, acesso remoto e links) (MARCONDES, 2021).

Para demonstrar os termos do modelos na interface do AgroTermos e no formato SKOS nesta tese, foi definido como recorte inicial para a representação no software *TemaTres* na linguagem SKOS, apenas as classes da CDU que tiveram mais de duas coincidências no AgroTermos, com o total de 10 termos. Como ainda tiveram muitos termos, para cada termo é representado por uma interface com o termo e seus relacionamentos semânticos com outros termos e outra que representa a linguagem SKOS, por este motivo, foi necessário adotar outro recorte para evitar muitas informações repetidas. Assim, foram selecionados três termos, seguindo o seguinte critério: termos que possuem diferentes tipos de relações de correspondência (*matching*) entre os SOC (*skos:exactMatch*, *skos:broadMatch* e *skos:narrowMatch*), com suas respectivas URI's disponibilizadas no AgroTermos pela Embrapa.

A Figura 30 apresenta a tabela “lc_tema” da base de dados do *TemaTres*, onde estão armazenados o código e a descrição dos termos que foram exportados no formato SKOS nos exemplos a seguir.

FIGURA 30 - Código e descrição dos termos exportados no formato SKOS

tema_id	code	tema
1	code_term	
26	NULL	Agronomia
27	NULL	Território e Paisagem
148	631.15	Intensificação Agropecuária
158	631.544.7	Mudança de cobertura da terra
159	NULL	Cultivo sob várias coberturas. Proteção e cobertur...
166	NULL	Produção da fazenda. Organização de fazendas. Admi...
170	631.58	Pousio

FONTE: Dados da pesquisa (2021).

Na Figura 30, de acordo com a base de dados do *TemaTres*, foi inserido o código dos termos no campo “tema_id”. O campo “code” armazena o código do termo na classe da CDU. O código nos descritores que correspondem às categorias do tesouro (“Agronomia” e “Território e Paisagem”) estão com valor nulo (NULL), por não possuírem um código específico no AgroTermos. O termo de código 148 está relacionado com o termo “Intensificação Agropecuária”, que possui o código 631.15, por ser um termo específico do descritor da CDU - “Produção da Fazenda. Organização de fazendas. Administração de fazendas”, com o código 631.15 da CDU e código 166 na base de dados. O descritor “Mudança de cobertura da terra” do AgroTermos, possui código 158 recebeu o código 631.544.7 da CDU, por se tratar de um termo com relação associativa de (Agente-objeto) de acordo com o descritor “Cultivo sob várias coberturas com o código 159. Proteção e cobertura do solo”, que por sua vez,

também não possui código devido ao fato de ser um termo relacionado, ou seja, que está subordinado a classe 631.544.7. Por fim, o descritor “Pousio” está no AgroTermos e na CDU, com o mesmo signo linguístico nos dois sistemas, com relação de equivalência, e portanto recebem o mesmo código da CDU (631.58), representado no *TemaTres* pelo código 170. As interfaces com as classes da CDU e os termos nos dois SOC e suas relações semânticas, bem como o código SKOS para representação dos descritores será detalhado logo a diante.

Para exemplificar, apresenta-se a inserção do termo “Intensificação Agropecuária”, que foi inserido na categoria 5 - “Agronomia” no *TemaTres* e sua respectiva notação da CDU e descritores e as relações semânticas atribuídas a este termo, conforme apresentado na Figura 31.

FIGURA 31 - Termo “Intensificação Agropecuária” e suas relações no *Tematres*

Vocabulário Intensificação Agropecuária

Início Menu Adicionar termo Buscar

Intensificação Agropecuária

Início → Agronomia → Intensificação Agropecuária

Termo Opções Adicionar Relações entre vocabulários Metadados

Intensificação Agropecuária

631.15

Termos genéricos

✘ TG↑ Agronomia

Termos específicos

✘ TEAS2 [código] Produção da fazenda. Organização de fazendas. Administração de fazendas

✘ BroadMatch: <https://sistemas.sede.embrapa.br/agrotermos/resources/43e7907a1789783315e08aad7f723848>

FONTE: Dados da pesquisa (2021).

LEGENDA: TEAS: indica que o descritor corresponde a uma relação hierárquica
Hiperônimo - Hipônimo - Inclusão de classe - Atividade Subordinada.

A Figura 31 apresenta a tela do *TemaTres* com a navegação do início do vocabulário, passando pela categoria do termo selecionado, até chegar no termo “Intensificação Agropecuária”, destacando a correspondência com a classe 631.15 da CDU, que possui como termo geral a categoria 5 (Agronomia), na qual está inserido. Esta classe da CDU e o termo “Intensificação Agropecuária” possui uma relação hierárquica (Hiperônimo-Hipônimo) - Inclusão de classe - Atividade Subordinada, indicada pela sigla “TEAS”, que foi escolhido para representar (Termo Específico - Atividade Subordinada). Os termos

específicos correspondentes a classe 631.15 são “Produção da fazenda. Organização de fazendas. Administração de fazendas”. E por fim, o tipo de link de correspondência com a classe 631.15 da CDU e seus descritores com o termo “Intensificação Agropecuária” com a relação (*skos:BroadMatch*) e sua respectiva URI²⁵, pois a classe 631.15 e seus descritores são mais genéricos que o termo do tesauro.

A partir da URI, é possível obter algumas informações sobre o vocabulário e sua correspondência com outros. A Figura 32 mostra a URI do termo “Intensificação”, disponibilizada no AgroTermos no ambiente da Embrapa, que foi compatibilizada com outro vocabulário, que no caso foi o tesauro AGROVOC, para fins de demonstração do formato de uma URI, bem como as informações disponibilizadas. Para os outros exemplos a seguir, não será demonstrado, pois a alteração está somente o nome do termo e no código da URI.

FIGURA 32 - URI do termo “Intensificação” no tesauro AgroTermos e Agrovoc

```
{"@context":{"skos":"http://www.w3.org/2004/02/skos/core#", "uri":"@id", "type":"@type", "lang":"@language", "value":"@value", "graph":"@graph", "prefLabel":"skos:prefLabel", "altLabel":"skos:altLabel", "inScheme":"skos:inScheme", "exactMatch":"skos:exactMatch"}, "graph":[{"uri":"https://sistemas.sede.embrapa.br/agrotermos/resources/43e7907a1789783315e08aad7f723848", "type":"skos:Term", "inScheme":{"uri":"http://aims.fao.org/aos/agrovoc"}, "prefLabel":{"lang":"pt-br", "value":"Intensificação"}}, {"dct:created":{"type":"http://www.w3c.org/2001/XMLSchema#dateTime", "value":"2017-03-09"}, "skos:notation":{"type":"https://sistemas.sede.embrapa.br/agrotermos/AgrotermosCode", "value":"43e7907a1789783315e08aad7f723848"}}]}
```

FONTE: Dados da pesquisa (2021).

Considerando a Figura 32, percebe-se que o termo “Intensificação” no tesauro AgroTermos está associado ao tesauro AGROVOC, destacando sua conversão para a linguagem de máquina. Para converter na linguagem SKOS, entre outros formatos já supracitados, deve-se clicar no link “metadados” e fazer a escolha. No caso da Figura 32, o termo “Intensificação” foi representado pela URI já mencionada na linguagem SKOS, o termo preferido, o não preferido, os tipos de correspondência (*skos:exactMatch*), entre outros. Isso comprova o uso de termos, por meio de URI, visando a interoperabilidade entre sistemas de informação para diferentes tipos de SOC.

A Figura 32 também indicou a URI do termo “Intensificação” no AGROVOC na língua português (Brasil), dados da criação do recurso²⁶, código do termo no tesauro AgroTermos (43e7907a1789783315e08aad7f723848), sua data de criação do vocabulário (2017-03-09) e notação (“ ”). Alguns links não estão funcionando, pois o vocabulário sempre deve estar sendo atualizado, além de outros campos que estão aparecendo com o valor vazio

²⁵<https://sistemas.sede.embrapa.br/agrotermos/resources/43e7907a1789783315e08aad7f723848>

²⁶ Disponível em: <http://www.w3c.org/2001/XMLSchema#dateTime>.

(" "), o que significa que não se aplica a este tipo de vocabulário ou não foi preenchido, como por exemplo, a "notação" que normalmente não aparece em tesouros.

A seguir, é apresentado o termo "Intensificação Agropecuária" e seus relacionamentos semânticos, com padrões de metadados que são agrupados no formato XML, RFS/RDFS, *Dublin Core* no formato SKOS, como está descrito a diante na Figura 33.

FIGURA 33 - Termo "Intensificação Agropecuária" do AgroTermos no SKOS

```

<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:skos="http://www.w3.org/2004/02/skos/core#" xmlns:map="http://www.w3c.rl.ac.uk/2003/11/21-skos-mapping#"
xmlns:dct="http://purl.org/dc/terms/" xmlns:dc="http://purl.org/dc/elements/1.1/">
  <skos:ConceptScheme rdf:about="http://localhost:8080/tematres/vocab/">
    <dc:title>Vocabulário Intensificação Agropecuária</dc:title>
    <dc:creator>Italo Rodrigues Castro</dc:creator>
    <dc:contributor/>
    <dc:publisher/>
    <dc:rights/>
    <dc:subject>Classificação Decimal Universal, Tesouros, Compatibilidade de linguagem, Relações Semânticas, Recuperação da Informaç
    </dc:subject>
    <dc:description>
      <![CDATA[ ]]>
    </dc:description>
    <dc:date>2021-02-25</dc:date>
    <dct:modified>2021-02-25 19:54:03</dct:modified>
    <dc:language>pt-BR</dc:language>
  </skos:ConceptScheme>
  <skos:Concept rdf:about="http://localhost:8080/tematres/vocab/xml.php?skosTema=148">
    <skos:prefLabel xml:lang="pt-BR">Intensificação Agropecuária</skos:prefLabel>
    <skos:notation>631.15</skos:notation>
    <skos:inScheme rdf:resource="http://localhost:8080/tematres/vocab/">
    <skos:broader rdf:resource="http://localhost:8080/tematres/vocab/xml.php?skosTema=26"/>
    <skos:narrower rdf:resource="http://localhost:8080/tematres/vocab/xml.php?skosTema=166"/>
    <skos:broadMatch>
      <skos:Concept rdf:about="https://sistemas.sede.embrapa.br/agrotermos/resources/43e7907a1789783315e08aad7f723848"/>
    </skos:broadMatch>
    <dct:created>2021-07-16 09:08:07</dct:created>
  </skos:Concept>
</rdf:RDF>

```

FONTE: Dados da pesquisa (2021).

Pela Figura 33, nota-se que o formato SKOS manifesta-se em várias linhas com uma linguagem baseada em XML namespace (Nome em XML para fornecer nomes de elementos e atributos para documentos em XML) e RDF/RDFS (Definição da tripla (recurso, propriedade, valor), que é a estrutura definida pela linguagem. O RDFS indica as relações semânticas na tripla), dispendo diferentes dados para caracterizar o termo "Intensificação Agropecuária". Isso quer dizer que em cada linha há representações dos metadados no formato *Dublin Core* (DC) referente ao termo, como título "Vocabulário Intensificação Agropecuária" (*dc:title*), criador "Italo Rodrigues Castro" (*dc:creator*), data de criação "2021-02-25" (*dc:date*), data e hora de modificação "2021-02-25 19:54:03" (*dct:modified*), assunto "Classificação Decimal Universal, Tesouros, Compatibilidade de Linguagem, Relações Semânticas, Recuperação da Informação" (*dc:subject*), idioma "pt-BR" (*dc:language*), dentre outros. O RDF constitui o formato de disponibilização dos dados, por URI's ou literal (MARCONDES, 2021), como apresentado na Figura 33.

O formato SKOS identifica também o conceito, o termo preferido "Intensificação

Agropecuária” (*skos:prefLabel*) de forma literal, número da classe “631.15” (*skos:notation*), que o termo do AgroTermos foi compatibilizado. Além de fornecer a URI do vocabulário que está sendo desenvolvido com o termo mais genérico, que pela URI tem o valor “Agronomia” (*skos:broader*) e o mais específico com o valor “Produção da fazenda. Organização de fazendas. Administração de fazendas” (*skos:narrower*). Além do tipo de correspondência para o conceito em outro sistema por meio da URI já mencionada anteriormente (*skos:broadMatch*), que possui informações do conceito disponibilizadas na web por outro sistema e por fim, a data de criação “2021-07-16 09:08:07” (*dct:created*). Isso reforça que os dois SOC, além de serem compatíveis, são interoperáveis, visto que são descritos em um padrão entendível pelo computador, o que evidencia o potencial dos SOC na recuperação da informação.

A constatação de que a CDU poderia ser compatibilizada com outros sistemas e que um formato padrão de descrição seria necessário foi levantada na tese de Slavic (2005). Segundo a autora, os formatos de descrição são imprescindíveis, visto que padronizam a linguagem, cuja uniformidade na representação possibilitará o processamento pela máquina. A portabilidade dos dados da CDU pode ser examinada e testada por meio de novos padrões para troca de vocabulário em rede (SLAVIC, 2005), como é demonstrado neste estudo. Ademais, as evidências ora expostas levam a entender que, além de serem padronizados pela máquina, a CDU e o AgroTermos são compatíveis. Como resultado desse processo, tem-se a garantia da interoperabilidade, logo, a promoção de melhorias na recuperação da informação em ambientes digitais.

Assim como o termo “Intensificação Agropecuária”, a seguir será apresentado o “Mudança de cobertura da terra”. A Figura 34 ilustra a descrição do termo “Mudança de cobertura da terra”, que está incluído na categoria 6 - “Território e paisagem” no *TemaTres* com sua respectiva notação da CDU, descritores e as relações semânticas atribuídas a este termo.

FIGURA 34 - Termo “Mudança de cobertura da terra” do AgroTermos e suas relações no *TemaTres*

The screenshot shows the 'Vocabulário Intensificação Agropecuária' interface. At the top, there is a navigation bar with 'Início', 'Menu', 'Adicionar termo', a search input field, and a 'Buscar' button. The main heading is 'Mudança de cobertura da terra' in green. Below it, a breadcrumb trail reads 'Início → Território e Paisagem → Mudança de cobertura da terra'. A secondary navigation bar includes 'Termo', 'Opções', 'Adicionar', 'Relações entre vocabulários', and 'Metadados'. The term 'Mudança de cobertura da terra' is listed with the class number '631.544.7'. Under 'Termos genéricos', there is a red 'x' icon, 'TG↑', and 'Território e Paisagem'. Under 'Termos relacionados', there is a red 'x' icon, 'TRAO⇌', and 'Cultivo sob várias coberturas. Proteção e cobertura do solo'. At the bottom, there is a red 'x' icon, 'NarrowMatch:', and the URI: <https://sistemas.sede.embrapa.br/agrotermos/resources/af88dfc5f6c51f44334d7b907cb8c044>.

FONTE: Dados da pesquisa (2021).

LEGENDA: TRAO: indica que o descritor corresponde a uma relação associativa - causal - agente - objeto.

A tela do *TemaTres* representado pela Figura 34 apresenta a navegação até o termo “Mudança de cobertura da terra”, que possui compatibilidade com a classe 631.544.7 da CDU, sendo inserido na categoria 6 (Território e Paisagem). Este termo representa uma relação Associativa - Causal - agente - objeto, com a sigla TEAO, que foi escolhido para representar (Termo relacionado - Agente - objeto) com a CDU. Os descritores da classe 631.544.7 são “Cultivo sob várias coberturas. Proteção e cobertura do solo” e o tipo de link de compatibilidade entre os vocabulários com a classe 631.544.7 da CDU e seus descritores com o termo “Mudança de cobertura da terra” com a relação (*skos:NarrowMatch*) e sua respectiva URI²⁷, pois a classe 631.544.7 e seus descritores são mais específicos que o termo representado no tesouro. A classe da CDU estabelece uma relação de associação com o termo “Mudança de cobertura da terra” do AgroTermos.

Adiante, é apresentado o termo “Mudança de cobertura da terra” e seus relacionamentos semânticos na linguagem SKOS, como está apresentado na Figura 35.

²⁷<https://sistemas.sede.embrapa.br/agrotermos/resources/af88dfc5f6c51f44334d7b907cb8c044>.

FIGURA 35 - Termo “Mudança de cobertura da terra” do AgroTermos no SKOS

```

<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:skos="http://www.w3.org/2004/02/skos/core#" xmlns:map="http://www.w3c.rl.ac.uk/2003/11/21-skos-mapping#"
xmlns:dct="http://purl.org/dc/terms/" xmlns:dc="http://purl.org/dc/elements/1.1/">
  <skos:ConceptScheme rdf:about="http://localhost:8080/tematres/vocab/">
    <dc:title>Vocabulário Intensificação Agropecuária</dc:title>
    <dc:creator>Italo Rodrigues Castro</dc:creator>
    <dc:contributor/>
    <dc:publisher/>
    <dc:rights/>
    <dc:subject>Classificação Decimal Universal, Tesouros, Compatibilidade de linguagem, Relações Semânticas, Recuperação da Informação.
    </dc:subject>
    <dc:description>
      <![CDATA[ ]]>
    </dc:description>
    <dc:date>2021-02-25</dc:date>
    <dct:modified>2021-02-25 19:54:03</dct:modified>
    <dc:language>pt-BR</dc:language>
  </skos:ConceptScheme>
  <skos:Concept rdf:about="http://localhost:8080/tematres/vocab/xml.php?skosTema=158">
    <skos:prefLabel xml:lang="pt-BR">Mudança de cobertura da terra</skos:prefLabel>
    <skos:notation>631.544.7</skos:notation>
    <skos:inScheme rdf:resource="http://localhost:8080/tematres/vocab/" />
    <skos:related rdf:resource="http://localhost:8080/tematres/vocab/xml.php?skosTema=159" />
    <skos:broader rdf:resource="http://localhost:8080/tematres/vocab/xml.php?skosTema=27" />
    <skos:narrowMatch>
      <skos:Concept rdf:about="https://sistemas.sede.embrapa.br/agrotermos/recursos/af88dfc5f6c51f44334d7b907cb8c044" />
    </skos:narrowMatch>
    <dct:created>2021-07-16 09:45:41</dct:created>
    <dct:modified>2021-07-16 09:52:27</dct:modified>
  </skos:Concept>
</rdf:RDF>

```

FONTE: Dados da pesquisa (2021).

A partir da Figura 35, percebe-se que o formato SKOS possui o detalhamento dos dados necessários para descrever o termo “Mudança de cobertura da terra”. O RDF constitui o formato de disponibilização dos dados, sendo representado por literais e/ou identificado por um recurso único (URI) (MARCONDES, 2021), como apresentado na Figura 34.

O formato SKOS identifica também o conceito, o termo preferido “Mudança de cobertura da terra” (*skos:prefLabel*), com sua respectiva classe na CDU “631.544.7” (*skos:notation*), que o termo do tesouro foi correspondido no momento da compatibilidade da linguagem. Além de fornecer a URI do vocabulário que está sendo desenvolvido com o termo mais genérico, que pela URI tem o valor “Território e Paisagem” (*skos:broader*), além do termo mais específico, representando pelo descritor “Cultivo sob várias coberturas. Proteção e cobertura do solo” (*skos:narrower*). Além do tipo de correspondência para o conceito em outro sistema por meio da URI, já descrita anteriormente (*skos:narrowMatch*) por meio do sistema da Embrapa por meio de um link, a data de criação “2021-07-16 09:08:07” (*dct:created*) e a data de modificação (*dct:modified*) 2021-07-16 09:52:27.

Assim como acontece com termos genéricos, os termos específicos apresentam a mesma formatação no SKOS, com indicação dos metadados, conforme disposto na Figura 35. Inere-se, portanto, que, independente do nível de representação, seja por exaustividade ou especificidade, a padronização do termo ocorre, tornando-se, portanto, legível pela

máquina. O requisito básico para a compatibilidade de dois vocabulários, como também, a padronização deles, é que suas respectivas configurações sejam, pelo menos, parcialmente compatível, se não totalmente compatível (em significado e estrutura). Essa afirmação foi declarada no estudo de Frâncu (2003a, 2003b) e é confirmada nas aplicações desta pesquisa, em conformidade com os resultados apresentados nas duas figuras anteriores.

Para Frâncu (2003a, 2003b), se esses requisitos não forem cumpridos, a recuperação da informação pode ser comprometida. Daí, a exigência de compatibilidade de linguagens de informação que podem coexistir e ser usadas alternativamente no acesso à informação pelos usuários, em ambientes físicos quanto digitais (FRÂNCU, 2003a, 2003b). A padronização no formato SKOS garante, também, o compartilhamento de informações, ou seja, os sistemas trocam informações (intercâmbio), o que facilita as atividades nos ambientes colaborativos (SLAVIC, 2005), sobretudo no contexto da web semântica, cujos recursos informacionais variados podem ser localizados e compartilhados por meio de inferências, eliminando-se ambiguidades, desde que os vocabulários tenham uma linguagem padronizada a ser processada pelas máquinas (RIZZO; TRONCY, 2011). Com efeito, os termos em formato SKOS, como ora demonstrado, servem para estimular o intercâmbio de dados e de gerenciamento de classificação entre diversos sistemas de bibliotecas (SLAVIC, 2005) e demais unidades que sistematizam as informações em acervos.

O modelo proposto nesta pesquisa é íntegro em termos de conversão para o formato SKOS. Isso porque todos os termos do AgroTermos foram testados, desde termos mais amplos a mais específicos. Com a finalidade de demonstrar a padronização do modelo no formato analisado, apresenta-se a normalização do termo “Pousio”, ou seja, um termo com um alto nível de especificidade em relação aos termos apresentados, como “Intensificação Agropecuária” e “Mudança de cobertura da terra”. A conversão desse termo para a linguagem da máquina é apresentada na Figura 36.

FIGURA 36 - Conversão do termo “Pousio” do AgroTermos e suas relações no *Tematres*

FONTE: Dados da pesquisa (2021).

LEGENDA: UPSA: indica que o descritor corresponde a uma relação de equivalência - Sinônimo Absoluto com o termo não preferido representado pela sigla UP (Usado para).

Como nas figuras anteriores, o termo do AgroTermos “Pousio” foi inserido no *TemaTres*, como está apresentado na Figura 36, que apresenta a navegação até o termo “Pousio”, compatibilizado com a classe 631.58 da CDU, incluído na categoria 5 (Agronomia). O termo “Pousio” representa uma relação Equivalência - Sinônimo absoluto, com a sigla UPSA, que possui significado (Usado para - Sinônimo - absoluto) em relação a classe da CDU. O descritor da classe 631.58 é “Pousio” e o tipo de link de compatibilidade entre os vocabulários com esta classe da CDU e seus descritores com o termo “Pousio” possui relação (*skos:ExactMatch*) e sua respectiva URI²⁸, pois a classe 631.58 e seus descritores são sinônimos em relação ao termo representado no tesouro. Esta classe da CDU possui uma relação de equivalência com o termo “Pousio” do AgroTermos, que além de possuírem o mesmo significado possuem o mesmo descritor nos dois SOC.

Por fim, o termo “Pousio” e seus relacionamentos semânticos são representados na linguagem SKOS, de acordo com a Figura 37.

²⁸<https://sistemas.sede.embrapa.br/agrotermos/resources/d3baaf75195ca2e887f6c07c66da8d4d>.

FIGURA 37 - Termo “Pousio” do AgroTermos no SKOS

```

<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:skos="http://www.w3.org/2004/02/skos/core#" xmlns:map="http://www.w3c.rl.ac.uk/2003/11/21-skos-mapping#"
xmlns:dct="http://purl.org/dc/terms/" xmlns:dc="http://purl.org/dc/elements/1.1/">
  <skos:ConceptScheme rdf:about="http://localhost:8080/tematres/vocab/">
    <dc:title>Vocabulário Intensificação Agropecuária</dc:title>
    <dc:creator>Italo Rodrigues Castro</dc:creator>
    <dc:contributor/>
    <dc:publisher/>
    <dc:rights/>
    <dc:subject>Classificação Decimal Universal, Tesouros, Compatibilidade de linguagem, Relações Semânticas, Recuperação da Informação.
    </dc:subject>
    <dc:description>
      <![CDATA[ ]]>
    </dc:description>
    <dc:date>2021-02-25</dc:date>
    <dct:modified>2021-02-25 19:54:03</dct:modified>
    <dc:language>pt-BR</dc:language>
  </skos:ConceptScheme>
  <skos:Concept rdf:about="http://localhost:8080/tematres/vocab/xml.php?skosTema=170">
    <skos:prefLabel xml:lang="pt-BR">Pousio</skos:prefLabel>
    <skos:notation>631.581</skos:notation>
    <skos:altLabel xml:lang="pt-BR">Pousio</skos:altLabel>
    <skos:inScheme rdf:resource="http://localhost:8080/tematres/vocab/">
    <skos:broader rdf:resource="http://localhost:8080/tematres/vocab/xml.php?skosTema=26"/>
    <skos:exactMatch>
      <skos:Concept rdf:about="https://sistemas.sede.embrapa.br/agrotermos/resources/d3baaf75195ca2e887f6c07c66da8d4d"/>
    </skos:exactMatch>
    <dct:created>2021-07-16 11:13:18</dct:created>
  </skos:Concept>
</rdf:RDF>

```

FONTE: Dados da pesquisa (2021).

Os dados referentes a Figura 37, mostram que o formato SKOS possui a formalização para a representação do termo “Pousio”. Este formato facilita que os dados estejam no formato legível por máquina, que pode ser apresentado meio de literais e/ou URI’s (MARCONDES, 2021).

O formato SKOS identifica também que quando o conceito está relacionado com o termo preferido e não preferido, utiliza-se o mesmo termo, que neste exemplo, foi o termo “Pousio” com os rótulos (*skos:prefLabel*) e (*skos:altLabel*). Isso se deve pois, o termo “Pousio” é idêntico nos dois SOC e está associado a classe “631.58” da CDU (*skos:notation*), que por sua vez, é a classe da CDU que houve a compatibilidade com o termo do tesouro. O label (*skos:inScheme*) possui a URI do vocabulário controlado que está sendo contruído. O termo genérico em relação ao termo “Pousio”, está relacionada com a categoria “Agronomia” (*skos:broader*). O tipo de correspondência para este termo está representado pela URI anteriormente apresentada (*skos:ExactMatch*), onde a Embrapa disponibilizou o link para acesso a este termo e a última informação é relativo a data de criação “2021-07-16 11:13:18” (*dct:created*).

Ao ser processado pela máquina e constatada a compatibilidade com as classes da CDU, o software formatou os dados que representam o termo (metadados) inserindo-os em uma estrutura padronizada (SKOS). Na descrição do formato, destaca-se a identificação da classe da CDU (631.58), dentre outros metadados, dispostos de modo idêntico em relação

ao termo do tesouro. Isso demonstra a funcionalidade do SKOS no que tange à uniformidade das descrições, tanto em termos sintáticos (estrutura), quanto em termos semânticos (significado), como salientou o estudo de Frâncu (2003a, 2003b). Essa uniformidade garantirá benefícios diversos (SUMMERS *et al.*, 2008), como melhorias na recuperação (SLAVIC, 2005) e compartilhamento de informação e recursos (RIZZO; TRONCY, 2011), por conseguinte, promoverá a interoperabilidade e o reúso (JAIN *et al.*, 2010).

Para garantir, com eficácia, a recuperação, não se trata, pois, dos termos serem compatíveis com as classes da CDU. É preciso descrever uma quantidade expressiva de dados referentes ao termo, visto que, quanto mais completa (mais dados descritos) for a descrição, maior será a chance e a incidência dos termos serem conectados de forma significativa e útil, por exemplo, conectar coleções de dados bibliográficos de qualquer natureza em RDF, tendo em vista o acesso pelos usuários (SANTOS; MOREIRA, 2018). Evidencia-se, então, que a CDU e o AgroTermos, ao serem acoplados, serão transportados para o SKOS, conforme apresentado nas figuras anteriores, cujo formato disponibilizará uma série de dados sobre o recurso informacional, tornando-se, assim, com mais chance de ser recuperado no momento das buscas.

Por fim, os dados descritos no formato SKOS, como apresentado nestas figuras, constituem propriedades de etiquetagem ou rótulos (notações) capazes de representar um conceito em um determinado domínio do conhecimento (MARCONDES, 2021). Esses rótulos de nós ou termos podem ser usados para mostrar a lógica de hierarquias necessárias em SOC, com o fim de agrupar termos mais específicos em categorias, embora não sejam descritores; logo, eles não devem ser usados como termos de indexação (FRÂNCU, 2003a, 2003b). Pelas evidências de outros estudos, os resultados ora apresentados, no que tange aos rótulos descritos no padrão SKOS, indicam que o modelo proposto está mais adequado à recuperação da informação (saída do sistema) e não ao processo de representação, realizado por profissionais da informação (entrada do sistema).

Em síntese, ao utilizar o *TemaTres* para validação do modelo proposto, confirma-se que o modelo é viável para a recuperação da informação, sobretudo em ambientes digitais, principalmente, pelos termos dos SOC poderem ser padronizados em uma linguagem ou formato capaz de ser interpretado pela máquina. O SKOS foi utilizado como requisito de validação do modelo, que poderá ser utilizado em ambientes digitais, em especial, no âmbito da web semântica, como demonstrado nestes resultados e complementado com as descobertas de estudos anteriores. Além da recuperação da informação, infere-se, também, que o modelo, por ser formatado em SKOS, garantirá o compartilhamento de informações e de recursos, apresentando-se, portanto, como de suma importância para os usuários da informação e as instituições prestadoras dos serviços de informação em diferentes domínios do conhecimento.

7 CONSIDERAÇÕES FINAIS

“O importante é não parar de questionar”.
(Albert Einstein)

Esta pesquisa, a partir da aplicação de uma RSL e de um estudo de caso no contexto da Embrapa, investigou a compatibilidade de dois SOC (a CDU e o tesouro AgroTermos) para o intercâmbio de informações entre diferentes sistemas de informação, ou seja, para torná-los interoperáveis. A partir dos insumos metodológicos e tecnológicos previamente definidos, foi possível, em linhas gerais, realizar a correspondência entre os dois sistemas, no sentido de inferir que a recuperação da informação em ambientes digitais pode ser facilitada.

Logo após, foi analisado a contribuição dos relacionamentos semânticos entre os termos que denotam os conceitos entre a estrutura classificatória da CDU com os descritores do tesouro AgroTermos, sendo considerado um instrumento de representação do conhecimento de modo pós-coordenado; a partir da combinação de números de classes e descritores (vocabulário controlado) do domínio da agricultura. Desse modo, os termos utilizados nesta pesquisa, são capazes de representar o assunto supracitado do domínio da agricultura de forma padronizada, com o intuito de eliminar a ambiguidade da linguagem natural. Ademais, torna-se necessário que o intercâmbio de informações entre os dois SOC sejam consistentes e coerentes, indicando que o objetivo geral, bem como o primeiro e o segundo pressupostos foram atendidos.

A princípio, por meio da RSL, foi revelado que há carências de experimentos que utilizem a correspondência da CDU e de um tesouro em domínios específicos do conhecimento. A maioria dos estudos publicados, principalmente no exterior, tanto os que aplicaram estudo de caso ou experimentos diversos, quanto os estudos voltados a tendências mais gerais, indicaram que a CDU é um sistema adequado para facilitar a organização do conhecimento em diferentes domínios, sobretudo no âmbito dos acervos de bibliotecas, e que a estrutura classificatória do sistema pode contribuir com a recuperação em ambientes digitais. Percebeu-se que alguns estudos utilizaram outro vocabulário controlado, que foi o tesouro, sendo utilizado para reforçar o potencial do sistema classificatório na recuperação da informação, com indicação da necessidade de se realizar novos estudos que investiguem a compatibilidade da linguagem e a formatação dos sistemas para sua leitura e interpretação automatizada.

A deficiência percorrida na literatura, acerca da escassez de estudos práticos que analisem a correspondência da CDU com um tesouro para auxiliar a recuperação da informação, e a recomendação de que é necessário verificar se os SOC são compatíveis e legíveis por máquina constituíram, pois, a principal motivação para o desenvolvimento dos

procedimentos práticos deste estudo. A constatação da literatura reforçou a necessidade de delimitação de um escopo temático, de modo que os procedimentos técnicos pudessem ser aplicados no prazo de desenvolvimento do estudo. Optou-se por analisar a classe 631 da CDU e outras classes que correspondem ao assunto “Intensificação Agropecuária” no campo da agricultura, com o uso de 30 termos mapeados na tese de Maculan (2015) acerca do vocabulário controlado da Embrapa.

Para facilitar a condução da pesquisa, convencionou-se dividi-la em etapas, *a priori*, definidas pela RSL e pelo estudo de caso. Assim, esses foram os dois grandes momentos da pesquisa (parte teórica e parte aplicada). A parte teórica forneceu os argumentos e conceitos sobre o tema estudado, além de ter identificado as aplicações da CDU quando acoplada a mais de um SOC e com vocabulários multilíngues, o que necessita de uma linguagem intermediária para a compatibilidade de linguagem para a representação de assuntos em diferentes sistemas e idiomas para a recuperação da informação em ambientes digitais. Mediante a identificação das aplicações, percebeu-se a importância de se desenvolver um modelo de compatibilidade de linguagem entre sistemas de organização do conhecimento, como um instrumento que pode ser utilizado por instituições que organizam informações de domínios especializados do conhecimento, como a Empresa Brasileira de Pesquisa Agropecuária (Embrapa).

Durante o estudo de caso, com os procedimentos técnicos adotados, foram realizadas diferentes atividades, resumidas do seguinte modo: formação do recorte temático a ser analisado; processo de compatibilidade e formação da estrutura esquemática do modelo; e conversão do modelo no padrão Simple Knowledge Organization System (SKOS). Essas atividades foram desenvolvidas de modo correspondente aos objetivos específicos previamente formulados, na tentativa de atendê-los, tendo em vista fortalecer a constituição dos resultados que iriam garantir a conclusão da pesquisa.

Assim, com o propósito de analisar as classes da CDU e descritores do AgroTermos que representam o assunto “Intensificação Agropecuária” do domínio da agricultura, foi realizado o mapeamento da estrutura classificatória da CDU com a estrutura do tesouro proposta por Maculan (2015), que permitiu a formação da amostra, constituído por 636 subclasses da estrutura classificatória da CDU e 30 termos do AgroTermos. Pelos resultados desse mapeamento, verificou-se a coerência do recorte, visto a existência de classes e termos que denotam conceitos genéricos quanto específicos. O resultado dessa aplicação permitiu constatar o alcance do primeiro objetivo específico e o terceiro pressuposto desta tese.

Em seguida, com a finalidade de apresentar a quantidade de ocorrências e as possíveis relações estabelecidas entre classes e termos, procedeu-se ao processo de compatibilidade da linguagem, cujos resultados revelaram que apenas um termo não foi compatível entre os demais. O termo em questão foi “Manejo do solo”, no AgroTermos, por ser um tipo de manejo físico/mecânico ou manejo químico/biológico, caracterizando-se como uma relação polihierárquica. Neste tipo de relação, considerando a estrutura classificatória da CDU, esse termo deveria estar em duas classes diferentes do sistema, por se tratar de dois tipos de manejo. Por este motivo, este exemplo não satisfaz o princípio da exclusão mútua para a organização do vocabulário da CDU, onde cada termo que representa um determinado conceito deve estar em uma única categoria. Para representar os 29 termos e seus respectivos conceitos, além de possuírem correspondência com a classe 631 da CDU e suas subdivisões; utilizou-se outras classes para a combinação de assuntos compostos e complexos dentro do domínio da agricultura, com o uso de auxiliares relacionais simples (:) e compostos (::) da CDU. Desse modo, estabeleceram-se diferentes termos e relacionamentos semânticos, com prevalência das relações do tipo equivalentes totalizando 23 classes utilizadas da CDU do total de 636 subclasses dentro da classe 631 (Agricultura em geral), o que representou o modelo proposto. Esse resultado reforça a característica flexível do sistema CDU que, embora mantenha uma lógica, com relacionamentos hierárquicos e associativos e rigidez em sua estrutura, possibilita a criação de infinitos relacionamentos, com inserção de novas classes e termos, ampliando o escopo de abrangência de um domínio do conhecimento. Isso se deve pela característica de análise e síntese desse sistema, que representou de forma exaustiva, o recorte temático do domínio da agricultura. Com isso, o modelo proposto pode melhorar interoperabilidade semântica entre os dois SOC e por conseguinte, a relevância na RI, de acordo com as necessidades de informação dos usuários em um SRI. Esta constatação identificada permitiu indicar o atendimento do segundo objetivo específico e do quarto e último pressuposto desta tese.

O alinhamento das classes da CDU com os descritores do AgroTermos e o estabelecimento das relações semânticas foram suficientes para confirmar que os dois SOC são compatíveis. Os resultados, ao serem confrontados com estudos correlatos, fortaleceram as evidências tendo em vista a confirmação de que as linguagens de ambos os vocabulários são correspondentes e isso representa o primeiro passo para que haja a interoperabilidade entre os sistemas, no contexto da web semântica. Se os sistemas são compatíveis e interoperáveis, torna-se possível refinar as buscas do usuário, que poderá navegar entre diferentes páginas da web, utilizando diferentes recursos, mas mantendo uma lógica na navegação. Ao aplicar o modelo de representação do domínio mapeado no software *TemaTres*, foi possível identificar a conversão dos termos do modelo e de seus relacionamentos para o formato SKOS no âmbito da web semântica.

Após o processo de compatibilização entre os termos a CDU e do Agrotermos, constatou-se que todos os termos puderam ser normalizados em um formato legível pela máquina. O formato SKOS é descrito no formato XML e RDF/RDFS, em diferentes campos, que indicam diferentes padrões de metadados referentes ao recurso informacional. Dentre esses metadados, destaca-se a classe da CDU, cujo termo do AgroTermos está vinculado. Todos os termos foram descritos no SKOS, com uma linguagem padronizada desse formato, desde os termos genéricos até os termos específicos. Os resultados alcançados demonstraram que o terceiro e último objetivo específico do estudo foi atingido, assim como o quarto e último pressuposto.

Destaca-se que o modelo de recuperação gerado possui uma estrutura formada por três grandes categorias: classe genérica, classe 631 e classes específicas. Para cada classe ou subclasse do sistema classificatório, houveram correspondências com um ou dois termos do AgroTermos, além do estabelecimento de uma ou duas relações semânticas. Reforça-se, também, que essa estrutura, ao ser formatada no SKOS, conforme evidências da literatura, possibilitará a correspondência entre os sistemas com os diversos recursos da web semântica. Tal fato facilitará as buscas pelo usuário, que poderá ampliar ou especificar os termos, mediante uma navegação direcionada, ou seja, há possibilidade de encontrar informações mais específicas e em maior quantidade, dentro do domínio de conhecimento procurado.

Além de garantir melhorias na recuperação, os resultados da conversão do modelo para o SKOS demonstraram que há possibilidade de as instituições serem beneficiadas, sobretudo na constituição dos ambientes colaborativos, visto que a linguagem padronizada enaltece o intercâmbio ou compartilhamento de informação e de recursos. Portanto, pelo compartilhamento, as instituições poupam tempo e recursos, no sentido de que a informação poderá ser reutilizada e intercambiada nos ambientes de colaboração. Essa possibilidade, para ser alcançada, demandará de sistemas informatizados que “conversem entre si”, ou seja, sistemas que sejam compatíveis, em termos de linguagem, e que sejam interoperáveis, em termos dos formatos que utilizam para leitura e interpretação por diferentes máquinas.

Por fim, concluiu-se que a CDU acoplada a um tesouro, como o AgroTermos, contribui para melhorias da recuperação da informação, visto que os SOC são compatíveis, estabelecendo relações semânticas capazes de direcionar as buscas, com flexibilidade e mantendo a lógica entre as relações. As melhorias na recuperação também se devem pelo fato de ambos serem interoperáveis, portanto mantêm uma padronização de formato suficiente para ser interpretada pelos sistemas informatizados. Com isso, possibilita o intercâmbio de informações e recursos, sobretudo no contexto da web semântica, indicando que todos os objetivos e pressupostos desta tese foram alcançados.

Espera-se que este estudo contribua para a formação de uma base teórica acerca da utilidade de mais de um SOC para a representação e recuperação da informação em ambientes digitais. Em termos práticos, a intenção é que o modelo proposto possa ser aplicado no contexto da Embrapa e em outras instituições que lidam com pesquisa ou organização da informação no domínio da agricultura. Há pretensão, também, que a pesquisa contribua para estimular o processo de compatibilidade e conversão de padrões entre diferentes SOC, investigando outros recortes temáticos, em diferentes campos do conhecimento, haja vista formalizar novos modelos que confirmem o potencial dos SOC na recuperação.

A literatura comprovou que há interesse de pesquisa com o uso da CDU, embora, conforme evidências deste estudo, seja necessária a aplicação de experimentos em outros recortes temáticos. Isso se deve pelo fato de validar novos modelos e, portanto, ampliar as comprovações acerca dos pontos fortes e fracos do sistema CDU para com a organização e recuperação da informação no ambiente digital. A literatura internacional destaca os estudos iniciados por Frâncu (2003a, 2003b) que aplicou procedimentos de compatibilidade e conversão de linguagens. Já outros trabalhos correlatos como o de Slavic (2005) dentre outros correlatos, não aplicaram estes procedimentos. Essas lacunas ainda não solucionadas ampliam o campo de investigação, sobretudo com as complexidades existentes no ambiente digital, no que tange à recuperação da informação. Logo, que os estudos se concentrem em abordagens interdisciplinares no campo da Ciência da Informação com a Tecnologia da Informação, tendo em vista aproveitar o potencial da CDU, visto que foi implementado para ser um sistema de caráter universal a ser aplicado em diferentes contextos.

Alguns problemas encontrados são dignos de serem mencionados, tais como: a necessidade de um recorte temático pequeno, por causa do tempo de desenvolvimento do estudo; a utilização, apenas, de um software para conversão do modelo em padrão de descrição informatizada; e a escassez de literatura produzida no Brasil. Esses problemas manifestaram-se como limitações do estudo, embora poderão constituir indicações para novas pesquisas, no sentido de alavancar os estudos sobre o potencial da CDU para a recuperação da informação, a compatibilidade entre diferentes SOC e a conversão desses sistemas em padrões que sejam legíveis pelas máquinas.

Finalmente, para pesquisas futuras, sugere-se um estudo comparativo entre as diferentes metodologias utilizadas pelos autores desta temática e a aplicação de estudos de caso em outros domínios, de modo a reconhecer o uso da CDU, no cotidiano de profissionais que lidam com a informação em ambientes digitais. Nesse sentido, a recomendação para estudos posteriores diz respeito ao uso da base de dados *Maste Reference File* (MRF) da CDU em formato de dados vinculados (*Linked Data*), de modo que seja possível incluir o maior número possível de dados em diferentes idiomas e mapeamentos para outros SOC.

Em relação aos dados oriundos desses mapeamentos e registros em bases de dados, faz-se necessário que sejam disponibilizados gratuitamente em várias linguagens, além do SKOS. Por exemplo: *Dublin Core (DC)*, *Metadata Authority Description Schema (MADS)*, *Extensible Markup Language (XML)*, *JavaScript Object Notation (JSON)*, XTM (Mapa de tópicos), dentre outros formatos no âmbito da web semântica, para facilitar a interoperabilidade. Para alcançar o objetivo sugerido, será necessário melhorar o conjunto de dados original em cada idioma e seus respectivos mapeamentos, a partir da contribuição voluntária de comunidades de usuários em todo o mundo para correções, manutenções ou evoluções desses estudos (UDC CONSORTIUM, 2020).

Além disso, é importante unificar os estudos, delegando a uma instituição a tarefa de disponibilizar de forma padronizada os conceitos e os relacionamentos entre os termos. Esses termos e seus respectivos conceitos devem representar os assuntos em cada domínio do conhecimento por meio de dados abertos (*Linked Open Data*), para que a informação possa ser compartilhada por diversas instituições nacionais ou internacionais e que esse conteúdo possa ser acessível para toda a comunidade da área de interesse.

Outra proposta de estudo em relação à temática é quanto ao uso do modelo de representação para a recuperação proposto nesta tese. Espera-se que ele possa ser aplicado em futuros experimentos, como por exemplo para a construção de ontologias com a utilização da linguagem OWL/OWL2. Este estudo comprovou que um SOC pode fornecer insumos para o outro, e quanto mais SOC estarem sendo compatibilizados melhor é a relevância do modelo que será proposto. Isso se deve pelo fato dos SOC se complementarem, o que garante maior controle do vocabulário e significativas melhorias na representação e recuperação da informação. Outro ponto importante, é no sentido de avaliar o impacto da recuperação promovida com o uso da CDU e do tesauro em diversas bases de dados, que contenham os dados de classificação e os termos referentes aos conceitos sobre os assuntos dos documentos registrados. Para isso, deve-se considerar a subjetividade do classificador durante o processo de classificação e por ser se tratar de um grande volume de informações (metadados) sobre uma coleção de documentos. Com isso, podem-se investigar diferentes abordagens de modelos de recuperação da informação, tais como: os métodos quantitativos (booleano, booleano estendido, vetorial, probabilístico e *fuzzy*) e os métodos dinâmicos no âmbito da Inteligência Artificial (sistemas especialistas, redes neurais e algoritmos genéticos), como o objetivo de aplicar diferentes algoritmos utilizados em algumas amostras para comparar e analisar a eficiência de cada um.

É importante ressaltar a realização de pesquisas em diferentes domínios do conhecimento, de modo a confrontar e comparar os resultados, avaliando a contribuição dos autores para ampliar o universo de termos presentes nos domínios, os quais estão em constantes mudanças. Além do mais, necessita-se verificar a interdisciplinaridade nos

diferentes domínios do conhecimento e a interoperabilidade. Nesse sentido, também deve-se atentar sobre a necessidade de realizar revisões sistemáticas da literatura em anos posteriores aos de 2017, pois esse foi o ano de publicação mais recente da revisão realizada nesta tese. Por esse motivo, é importante verificar, na atualidade, como está o nível de aplicação dos métodos propostos pelos autores estudados.

REFERÊNCIAS

AHN, Jae-wook; LIN, Xia; KHOO, Michael. Dewey Decimal Classification Based Concept Visualization for Information Retrieval. *In: CEUR WORKSHOP PROCEEDINGS*, 1., 2014, Dinamarca. **Proceedings** [...]. Dinamarca: [S.n.], 2014. p. 7-14.

AITCHISON, Jean. The thesurofacet: a multipurpose retrieval language tool. **Journal of documentation**, London, v.26, n. 3, p. 187-203, 1970.

AITCHISON, Jean. Thesurofacet: a new concept in subject retrieval schemes. **Subject retrieval in the seventies. Connecticut: Greenwood**, p. 72-98, 1972.

ALLEMANG, Dean; HENDLER, Jim. **Semantic Web for the Working Ontologist: Effective Modeling in RDFS and OWL**. 2nd. ed. [S.l.]: Morgan Kaufmann/Elsevier, 2011. Disponível em:
https://www.researchgate.net/publication/220692683_Semantic_web_for_the_working_ontologist_-_modeling_in_RDF_RDFS_and_OWL. Acesso em: 05 mai. 2021.

ANDRADE, Lucas Veras de; BRUNA, Dayane; SALES, Wesleyne Nunes de. Classificação: uma análise comparativa entre a Classificação Decimal Universal-CDU e a Classificação Decimal de Dewey-CDD. **Biblos**, Rio Grande, v. 25, n. 2, p. 31-42, jul./dez. 2011. Disponível em: <https://periodicos.furg.br/biblos/article/view/2088>. Acesso em: 6 jun. 2021.

ANSI/NISO Z39.19. **Guidelines for the construction, format and management of monolingual controlled vocabularies**. Bethesda: NISO, 2005. Disponível em:
<http://www.niso.org/publications/ansiniso-z3919-2005-r2010>. Acesso em: 5 maio 2020.

ARAÚJO JÚNIOR, Rogério Henrique de; TARAPANOFF, Kira. Precisão no processo de busca e recuperação da informação: uso da mineração de textos. **Ciência da Informação**, Brasília, v. 35, n. 3, p. 236-247, set./dez. 2006. Disponível em:
<http://revista.ibict.br/ciinf/article/view/1130/1278>. Acesso em: 6 jun. 2021.

ARAÚJO, Webert Junio. **Reúso de recursos da web semântica para a construção de vocabulários controlados no contexto da ciência da informação**. 2016. Dissertação (Mestrado em Ciência da Informação)– Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2016. Disponível em:
https://repositorio.ufmg.br/bitstream/1843/BUBD-AE3HRL/1/disserta_o_avaliao_de_ontologia_com_base_na_compara_o_a_um_corpus_um_estudo_da_ontoagrohidro_da_embrapa__webert_j_ni_1.pdf. Acesso em: 05 mar. 2020.

BAEZA-YATES, Ricardo; RIBEIRO-NETO, Berthier. **Modern Information Retrieval**. New York: ACM Press, 1999.

BAKER, Thomas *et al.* Key choices in the design of Simple Knowledge Organization System (SKOS). **Journal of Web Semantics**, [S.l.], v. 20, n. 1, p. 35-49, jan./jun. 2013. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1570826813000176>. Acesso em: 6 jun. 2021.

BALÍKOVÁ, Marie. The role of UDC classification in the Czech Subject Authority File Item type Meetings and Proceedings at a Crossroads: Multiple Directions to Usability The role of UDC classification in the Czech Subject Authority File. **Extensions & Corrections to the UDC**, [S.l.], p. 1-27, 2009. Disponível em: <http://hdl.handle.net/10150/198789>. Acesso em: 14 jul. 2018.

BARBOSA, Alice Príncipe. **Teoria e prática dos sistemas de classificação bibliográfica**. Rio de Janeiro: Instituto Brasileiro de Bibliografia e Documentação, 1969.

BARBOSA, Alice Príncipe. Classificações facetadas. **Ciência da Informação**, Brasília, v. 1, n. 2, p. 73-81, jul./dez. 1972. Disponível em: <http://revista.ibict.br/ciinf/article/view/10>. Acesso em: 6 jul. 2021.

BARDIN, Laurence. **Análise de conteúdo**. São Paulo: Edições 70, 2011.

BERGER, Julianna *et al.* EMBRAPA e disseminação da informação. *In*: WORKSHOP DE PESQUISA, TECNOLOGIA E INOVAÇÃO (PTI), 5.; SIMPÓSIO INTERNACIONAL DE INOVAÇÃO E TECNOLOGIA (SIINTEC), 1., 2015, Salvador. **Anais [...]**. Salvador: [S.n.], 2015. p. 25-28.

BERNERS-LEE, Tim; HENDLER, James; LASSILA, Ora. The semantic web. **Scientific american**, [S.l.], v. 284, n. 5, p. 34-43, maio 2001. Disponível em: https://www-sop.inria.fr/acacia/cours/essi2006/Scientific%20American_%20Feature%20Article_%20The%20Semantic%20Web_%20May%202001.pdf. Acesso em: 5 jun. 2021.

BERNERS-LEE, Tim *et al.* Tabulator: Exploring and analyzing linked data on the semantic web. *In*: INTERNATIONAL SEMANTIC WEB USER INTERACTION WORKSHOP, 3., 2006, [S.l.]. **Proceedings [...]**. [S.l.]: [S.n.], 2006. p. 159.

BOCCATO, Vera Regina Casari; RAMALHO, Rogério Aparecido Sá; FUJITA, Mariângela Spotti Lopes. A contribuição dos tesouros na construção de ontologias como instrumento de organização e recuperação da informação em ambientes digitais. **Ibersid**, [S. l.], v. 2, n. 1, p. 199-209, 2008. Disponível em: <https://www.ibersid.eu/ojs/index.php/ibersid/article/view/2235>. Acesso em: 14 jul. 2021.

BOSERUP, Ester. **The conditions of Agricultural growth: the economics of agrarian change under population pressure**. Chicago, IL: Aldine, 1965.

BRASCHER, Marisa. A ambiguidade na recuperação da informação. **DataGramZero - Revista de Ciência da Informação**, v.3, n.1, 2002. Disponível em: <https://ridi.ibict.br/handle/123456789/284>. Acesso em: 05 de mai. 2019.

BRASCHER, Marisa; CAFÉ, Lígia. Organização da informação ou organização do conhecimento?. *In*: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 9., 2008, São Paulo. **Anais [...]**. São Paulo: Ancib, 2008. p. 1-14. Disponível em: <http://repositorios.questoesemrede.uff.br/repositorios/handle/123456789/809>. Acesso em: 6 jul. 2021.

BRASCHER, Marisa. Tesouro, taxonomia e ontologia: uma evolução. *In*: CICLO DE CONFERÊNCIAS SOBRE ORGANIZAÇÃO DA INFORMAÇÃO LEGISLATIVA E JURÍDICA, 1., 2009, Brasília. **Anais [...]**. Brasília: [S.n.], 2009. Disponível em: <https://slideplayer.com.br/slide/5614662/>. Acesso em: 6 jun. 2021.

BRASCHER, Marisa. Semantic Relations in Knowledge Organization Systems. **Knowledge Organization**, [S. l.], v. 41, n. 2, p. 175-180, abr. 2014. Disponível em: <http://search.ebscohost.com/login.aspx?direct=true&db=iih&AN=95782279&lang=ptbr&site=ehost-live&authtype=ip,cookie,uid>. Acesso em: 5 mar. 2020.

BROUGHTON, Vanda *et al.* Knowledge organisation European Curriculum Reflections on Education in Library and information science (pp. 133-148). **Copenhagen: Royal School of Library and Information Science**, 2005. Disponível em: <<http://arizona.openrepository.com/arizona/handle/10150/105851>>. Acesso em: 20 abr. 2021.

BROUGHTON, Vanda. The need for a faceted classification as the basis of all methods of information retrieval. **Aslib Proceedings: New Information Perspective**, [S. l.], v. 58, n. 1, p. 49-72, jan./jun. 2006. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/00012530610648671/full/html>. Acesso em: 5 jun. 2021.

BROUGHTON, Vanda; SLAVIC, Aida. Building a faceted classification for the humanities: principles and procedures. **Journal of Documentation**, London, v. 63, n. 5, p. 727-754, nov./dez. 2007. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/00220410710827772/full/html>. Acesso em: 6 jun. 2021.

BROUGHTON, Vanda. Concepts and terms in the faceted classification: the case of UDC. **Knowledge Organization**, [S. l.], v. 37, n. 4, p. 270-279, out./nov. 2010. Disponível em: https://www.ergon-verlag.de/isko_ko/downloads/ko37_2010_4_c.pdf. Acesso em: 6 jun. 2021.

BUXTON, Andrew. Computer searching of UDC numbers. **Journal of Documentation**, London, v. 46, n. 3, p. 193-217, ago./set. 1990.

BUXTON, Andrew; MCILWAINE, Ia. **Guia para utilização da CDU**: um guia introdutório para o uso e aplicação da Classificação Decimal Universal. Brasília: Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), 1998.

CAMPOS, Maria Luiza Almeida. **Linguagem documentária**: teorias que fundamentam sua elaboração. Rio de Janeiro: Ed. UFF, 2001.

CAMPOS, Maria Luiza Almeida; GOMES, Hagar Espanha. Organização de domínio de conhecimento e os princípios ranganathianos. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 8, n. 2, p. 150-163, jul./dez. 2003. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/366/175>. Acesso em: 6 jun. 2021.

CAMPOS, Maria Luiza de Almeida. A problemática da compatibilização terminológica e a integração de ontologias: o papel das definições conceituais. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 6., 2005, Florianópolis. **Anais** [...]. Florianópolis: Universidade Federal de Santa Catarina, 2005. p. 1-12. Disponível em: <http://repositorios.questoesemrede.uff.br/repositorios/handle/123456789/39?show=full>. Acesso em: 6 jun. 2021.

CAMPOS, Maria Luiza de Almeida *et al.* Estudo comparativo de softwares de construção de tesouros. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 11, n. 1, p. 68-81, jan./abr. 2006. Disponível em: <https://www.scielo.br/j/pci/a/jNjm8LTYRz75b4tWkCyD3Rx/abstract/?lang=pt>. Acesso em: 6 jun. 2021.

CAMPOS, Maria Luiza de Almeida; GOMES, Hagar Espanha. Taxonomia e classificação: a categorização como princípio. **Datagramazero**: Revista de Ciência da Informação, v. 9, n. 4, p. 1-7, ago. 2008. Disponível em: <https://brapci.inf.br/index.php/res/v/6615>. Acesso em: 6 jun. 2021.

CAMPOS, Maria Luiza de Almeida; BARBOSA, Nilson Theobald. Interoperabilidade semântica: proposta metodológica para o mapeamento semântico entre SOCS em sistemas heterogêneos. **Informação & Sociedade: Estudos**, João Pessoa, v. 30, n. 4, p. 1-21, out./dez. 2020. Disponível em: <https://periodicos.ufpb.br/ojs2/index.php/ies/article/view/57046>. Acesso em: 6 jul. 2021.

CARDILLO, Elena *et al.* Towards the Reuse of Standardized Thesauri Into Ontologies. *In*: INTERNATIONAL CONFERENCE ON ONTOLOGY AND SEMANTIC WEB PATTERNS, 5., 2014, [S.l.]. **Proceedings** [...]. [S. l.]: [S. n.], 2014. p. 26-37. Disponível em: https://www.researchgate.net/publication/272794487_Towards_the_reuse_of_standardized_thesauri_into_ontologies. Acesso em: 6 jun. 2021.

CERVANTES, Brígida Maria Nogueira. **A construção de tesouros com a integração de procedimentos terminográficos**. 2009. 210f. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2009.

Disponível em: https://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/cervantes_bmn_do_mar.pdf. Acesso em: 20 jan. 2020.

CHAFFIN, Roger; HERRMANN, Douglas. The similarity and diversity of semantic relations. **Memory & Cognition**, [S.l.], v. 12, n. 2, p. 134-141, jan./dez. 1984.

CHAN, Lois Mai. **Exploiting LCSH, LCC, and DDC to Retrieve Networked Resources: Issues and Challenges**. Library of Congress, Washington, 2000. Disponível em: https://www.loc.gov/catdir/bibcontrol/chan_paper.html. Acesso em: 6 jun. 2021.

CHAN, Lois Mai; ZENG, Marcia Lei. Ensuring interoperability among subject vocabularies and knowledge organization schemes: a methodological analysis. **IFLA journal**, [S. l.], v. 28, n. 5-6, p. 323-327, dez. 2002. Disponível em: <https://archive.ifla.org/IV/ifla68/papers/008-122e.pdf>. Acesso em: 6 jun. 2021.

CHAN, Lois Mai; ZENG, Marcia Lei. Metadata interoperability and standardization—a study of methodology part I. **D-Lib magazine**, [S. l.], v. 12, n. 6, p. 1082-9873, dez. 2006. Disponível em: <https://www.dlib.org/dlib/june06/chan/06chan.html>. Acesso em: 6 jun. 2021.

CHAN, Lois Mai; SALABA, Athena. **Cataloging and classification**: an introduction. [S.l.]: Rowman & Littlefield, 2015.

CHATTERJEE, Amitabha *et al.* Universal Decimal Classification and Colon Classification: Their mutual impact. **Annals of Library and Information Studies (ALIS)**, [S. l.], v. 62, n. 4, p. 226-230, out./dez. 2016. Disponível em: <http://nopr.niscair.res.in/handle/123456789/33716>. Acesso em: 6 jun. 2021.

CHUDAMANI, K. S. Classification Model for Libraries in the Digital Environment. *In*: INTERNATIONAL CALIBER, 2., 2004, New Delhi. **Anais [...]**. New Delhi: [S. n.], 2004. p. 11-13. Disponível em: http://ir.inflibnet.ac.in/bitstream/1944/371/1/04cali_63.pdf. Acesso em: 5 out. 2018.

CINTRA, Ana Maria Marques *et al.* **Para entender as linguagens documentárias**. 2. ed. São Paulo: Polis, 2002.

CLARKE, Stella G. Dextre; ZENG, Marcia Lei. From ISO 2788 to ISO 25964: The evolution of thesaurus standards towards interoperability and data modelling. **Information Standards Quarterly (ISQ)**, v. 24, n. 1, 2012. Disponível em: <http://eprints.rclis.org/16818/>. Acesso em: 20 Jun. 2021.

CLASSIFICAÇÃO DECIMAL UNIVERSAL (CDU): edição-padrão internacional em língua portuguesa. Brasília: **Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)**, 1997.

COATES, Earl. The role of classification in information retrieval: action and thought in the contribution of Brian Vickery. **Journal of documentation**, London, v. 44, n. 3, p. 216-225, jul./ago. 1988.

COLILLAS, Mariàngels Granados; NICOLAU, Anna. Aplicación de un nuevo sistema de indización en una colección de recursos especializados en ciencias de la educación. **EI profesional de la información**, Barcelona, v. 16, n. 6, p. 627-635, nov. 2007.

COLILLAS, Mariàngels Granados; NICOLAU, Anna. Improving subject searching in databases through a combination of descriptors and UDC. **Zadar: Bobcatsss**, Barcelona, v. 1, n. 8, p. 1-29, ago. 2008. Disponível em: <http://edoc.huberlin.de/conferences/bobcatsss2008/granados-mariangels-151/PDF/granados.pdf>. Acesso em: 20 nov. 2018.

COLILLAS, Mariàngels Granados. Tratamiento de los Metadatos de Contenido en la Web Semántica o cuando la Lógica se hace Evidente. *In: NUEVAS PERSPECTIVAS PARA LA DIFUSIÓN Y ORGANIZACIÓN DEL CONOCIMIENTO*, 1., 2009, Barcelona. **Anais [...]**. Barcelona: [S. n.], 2009. Disponível em: <https://dialnet.unirioja.es/servlet/articulo?codigo=2925237>. Acesso em: 6 jun. 2021.

COLILLAS, Mariàngels Granados. UDC on the internet: theory and project in evolution for use of indexing and retrieval systems. **IFLA Journal**, [S. l.], v. 37, n. 4, p. 305-313, abr. 2011. Disponível em: http://eprints.rclis.org/18510/1/ifla-journal-37-4_305-313_2011.pdf. Acesso em: 6 jun. 2021.

COSTA, Lucas Rodrigues; CASTRO, Priscila de Paiva; GOMES, Rafael Fernandez. Instalação, configuração e manutenção. *In: SHINTAKU*, Milton (Org.). **Guia do usuário do TemaTres**. Brasília: Ibict, 2019. p. 13-16.

COX, Kevin. Information retrieval by browsing. *In: INTERNATIONAL CONFERENCE ON NEW INFORMATION TECHNOLOGY*, 5, 1992, Hongkong. **Proceedings [...]**. Hongkong: [S. n.], 1992. p. 1-12.

CRESWELL, Joh. **Projeto de pesquisa: métodos qualitativo, quantitativo e misto**. 3. ed. Porto Alegre: Artmed, 2010.

CUMPSTON, Miranda *et al.* Updated guidance for trusted systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions. **Cochrane Database Syst Rev**, [S. l.], v. 10, n. 3, p. 142, out. 2019. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/31643080/>. Acesso em: 6 jun. 2021.

CURRÁS, Emilia. **Tesauros**: linguagens terminológicas. Brasília, DF: IBICT, 1995.

CURRIER, Sarah; WAKE, Susannah. Negotiating subject access: resource discovery on the Web. **Library & information briefings**, [S. l.], n. 97, p. 1-14, jan./dez. 2001. Disponível em: <https://www.semanticscholar.org/paper/Negotiating-subject-access-%3A-Resource-discovery-on-Currier-Wake/480dfcb1547bb5ac475871054c4f0521d79da0a8>. Acesso em: 5 jun. 2021.

DAHLBERG, Ingetraut. Possibilities for a new universal decimal Classification. **Journal of Documentation**, London, v. 27, n. 1, p. 18-36, jan./jun. 1971.

DAHLBERG, Ingetraut. Teoria da classificação, ontem e hoje. *In*: CONFERÊNCIA BRASILEIRA DE CLASSIFICAÇÃO BIBLIOGRÁFICA, 2., 1976, Rio de Janeiro. **Anais [...]**. Rio de Janeiro: ABDF, 1976. p. 12-17.

DAHLBERG, Ingetraut. Teoria do conceito. **Ciência da Informação**, Brasília, v. 7, n. 2, p. 101-107, jul./dez. 1978a.

DAHLBERG, Ingetraut. Fundamentos teórico-conceituais da classificação. **Revista de Biblioteconomia de Brasília**, Brasília, v. 6, n. 1, p. 9-21, jan./jun. 1978b.

DAHLBERG, Ingetraut. Towards establishment of compatibility between indexing languages. **International Classification**, [S. l.], v. 8, n. 2, p. 88-91, jul./dez. 1981.

DAHLBERG, Ingetraut. Conceptual compatibility of ordering systems. **Ko knowledge organization**, London, v. 10, n. 1, p. 5-8, jan./jun. 1983.

DAHLBERG, Ingetraut. Knowledge organization: its scope and possibilities. **Knowledge Organization**, [S. l.], v. 20, n. 4, p. 211-222, abr. 1993.

DAHLBERG, Ingetraut. Compatibility and Integration of Order Systems 1960-1995: an annotated bibliography. *In*: COMPATIBILITY AND INTEGRATION OF ORDER SYSTEMS: RESEARCH SEMINAR OF THE TIP/ISKO MEETING, 1., 1995, Warsaw. **Proceedings [...]**. Warsaw: Wydawnictwo, 1996. p. 1-234.

DAHLBERG, Ingetraut. The Information Coding Classification (ICC): A modern, theory-based fully-faceted, universal system of knowledge fields. **Axiomathes**, [S. l.], v. 18, n. 2, p. 161-176, jan./dez. 2008. Disponível em: https://www.researchgate.net/publication/226779298_The_Information_Coding_Classification_ICC_A_Modern_Theory-Based_Fully-Faceted_Universal_System_of_Knowledge_Fields. Acesso em: 5 jun. 2021.

DEEKS, Jonathan *et al.* **Analysing data and undertaking meta-analyses**. Cochrane handbook for systematic reviews of interventions, [S. l.], p. 241-284, 2019. Disponível em: <https://training.cochrane.org/handbook/current/chapter-10>. Acesso em: 6 jun. 2021.

DE-LA-TORRE-UGARTE, Mônica Cecilia *et al.* Revisão sistemática: noções gerais. **Revista da Escola de Enfermagem da USP**, São Paulo, v. 45, n. 5, p. 1260-1266, out. 2011. Disponível em: <https://www.scielo.br/j/reeusp/a/CRjvBKKvRRGL7vGsZLQ8bQj/abstract/?lang=pt>. Acesso em: 6 jun. 2021.

DIAS, Danilo Camargo; MOREIRA, Walter; ALVES, Rachel Cristina Vesú. A representação temática de imagens digitais da NASA no Flickr: as contribuições dos sistemas de organização do conhecimento. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 18, n. 1, p. 1-14, jan./dez. 2020. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/8658560>. Acesso em: 3 jun. 2021.

DODEBEI, Vera. **Tesouro**: linguagem de representação da memória documentária. Rio de Janeiro: Intertexto, 2002.

DOERR, Martin. Semantic problems of thesaurus mapping. **Journal of Digital Information**, v. 1, n. 8, art. 52, Mar. 2001. Disponível em: <http://83.212.169.181/bbt/sites/default/files/Semantic-problems-of-thesaurus-mapping%282001%29.pdf/>. Acesso em: 25 dez. 2020.

DOORN, Marlene; POLMAN, Katrien. At a Crossroads: Multiple Directions to Usability From classification to thesaurus ... and back?. Subject indexing tools at the library of the Afrika-Studiecentrum Leiden [extended abstract]. **Knowledge Organization**, [S. l.], v. 37, n. 3, p. 203-208, mar. 2010. Disponível em: <http://www.ethnologue.com>. Acesso em: 29 fev. 2019.

EMBRAPA INFORMÁTICA AGROPECUÁRIA. **Relatório de Gestão 2015-2018**: pesquisa e inovação em tecnologia da informação e comunicação para a Agricultura, Campinas, 2018. Disponível em: <https://www.embrapa.br/busca-de-publicacoes/-/publicacao/1110042/relatorio-de-gestao-2015-2018-pesquisa-e-inovacao-em-tecnologia-da-informacao-e-comunicacao-para-a-Agricultura>. Acesso em: 5 jun. 2021.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA (EMBRAPA). **Relatório de Gestão**, [S. l.], 2018. Disponível em: <https://www.embrapa.br/documents/10180/1549626/Relat%C3%B3rio+de+Gest%C3%A3o+2018/600af295-9241-9094-1f76-d2e30b846417>. Acesso em: 3 jun. 2021.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA (EMBRAPA). **A Embrapa**, Brasília, 2020. Disponível em: <http://www.embrapa.br>. Acesso em: 5 jan. 2020.

FENG, Zhenxin; CHIN, Khan-Wu. A Novel Data Centric Information Retrieval Protocol for Queries in Delay Tolerant Networks. **Journal of Network and Systems Management**, Barcelona, v. 23, n. 4, p. 870-901, abr. 2015. Disponível em: https://www.researchgate.net/publication/271952080_A_Novel_Data_Centric_Information_Retrieval_Protocol_for_Queries_in_Delay_Tolerant_Networks. Acesso em: 6 jun. 2021.

FERREYRA, Diego. Prefácio. *In*: SHINTAKU, Milton (Org.). **Guia do usuário do TemaTres**. Brasília: Ibict, 2019. p. 51-74.

FIAŁKOWSKI, Konrad; MURASZKIEWICZ, Mieczysław; TOŁWIŃSKA, Alicja. Methodology of world clearinghouse activity. **Information Processing & Management**, Poland, v. 17, n. 1, p. 21-25, jan./jun. 1981.

FIRMINO, Helder Noel Monteiro. **Reúso de recursos da web semântica para a construção de vocabulários controlados no contexto da ciência da informação**. 2019. Tese (Doutorado em Gestão e Organização do Conhecimento) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2019. Disponível em: <https://repositorio.ufmg.br/handle/1843/32372> <http://aberto.univem.edu.br/handle/11077/722>. Acesso em: 10 mar. 2020.

FOSKETT, Antony Charles. **A abordagem temática da informação**. São Paulo: Polígono; Brasília: UnB, 1973.

FRÂNCU, Victoria. The impact of specificity on the retrieval power of a UDC-based multilingual thesaurus. **Cataloging & classification quarterly**, [S. l.], v. 37, n. 1-2, p. 49-64, jan./dez. 2003a. Disponível em: https://www.tandfonline.com/doi/abs/10.1300/J104v37n01_05. Acesso em: 7 jun. 2021.

FRÂNCU, Victoria. **Multilingual access to information using an intermediate language**. 203f. Tese (Doutorado) – Universiteit Antwerpen, Antwerpen, 2003b.

FRÂNCU, Victoria. Does convenience trump accuracy?. The avatars of the UDC in Romania. **Extensions and Corrections to the UDC**, [S. l.], v. 29, n. 1, p. 263-272, jan./jun. 2007. Disponível em: <https://repository.arizona.edu/handle/10150/106107>. Acesso em: 6 jun. 2021.

FRÂNCU, Victoria; SABO, Cosmin-Nicolae. Implementation of a UDC-Based Multilingual Thesaurus in a Library Catalogue: The Case of BiblioPhil. **Knowledge Organization**, [S. l.], v. 37, n. 3, p. 209-215, mar. 2010. Disponível em: https://www.researchgate.net/publication/277983264_Implementation_of_a_UDC-Based_Multilingual_Thesaurus_in_a_Library_Catalogue_The_Case_of_BiblioPhil. Acesso em: 6 jun. 2021.

FREEMAN, Robert. Computers and Classification Systems. **Journal of Documentation**, London, v. 20, n. 3, p. 137-145, mar. 1964.

FRICKÉ, Martin. Faceted Classification, Analysis and Search: Some Questions on their Interrelations. **Knowledge Organization**, [S. l.], v. 44, n. 6, p. 387-396, jun. 2017. Disponível em: https://www.ergon-verlag.de/isko_ko/downloads/ko_44_2017_6_a.pdf. Acesso em: 6 jun. 2021.

FUJITA, Mariângela Spotti Lopes *et al.* Avaliação das características do TemaTres e Multites para o controle de autoridades nas bibliotecas universitárias. **Scire**, [S. l.], v. 23, n. 2, p. 63-73, jul./dic. 2017. Disponível em: <https://dialnet.unirioja.es/servlet/articulo?codigo=6142605>. Acesso em: 15 jul. 2021.

FUJITA, Mariângela Spotti Lopes; SANTOS, Luciana Beatriz Piovezan dos; ALVES, Roberta Vesu. Linguagem de indexação e linguagem documentária são sistemas de organização do conhecimento?. uma análise bardiana da variação terminológica. **Scire**, [S. l.], v. 24, n. 2, p. 23-33, jul./dez. 2018. Disponível em: <https://docplayer.com.br/108846503-Linguagem-de-indexacao-e-linguagem-documentaria-sao-sistemas-de-organizacao-do-conhecimento-uma-analise-bardiana-da-variacao-terminologica.html>. Acesso em: 7 jun. 2021.

FUJITA, Mariângela Spotti Lopes *et al.* Evolução teórica sobre tesouro. **Scire**: representación y organización del conocimiento, [S. l.], v. 26, n. 1, p. 13-25, jan./jun. 2020. Disponível em: <https://dialnet.unirioja.es/ejemplar/551273>. Acesso em: 6 jun. 2021.

FUJITA, Mariângela Spotti Lopes; MOREIRA, Walter (Ed.). **Manual do planejamento, construção e manutenção do tesouro Unesp para bibliotecas**: do conceitual a práxis. Editora Oficina Universitária, 2021.

FUSCO, Elvis. **Modelos conceituais de dados como parte do processo da catalogação**: perspectiva de uso dos FRBR no desenvolvimento de catálogos bibliográficos digitais. 2010. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010. Disponível em: <http://aberto.univem.edu.br/handle/11077/722>. Acesso em: 5 jul. 2020.

GALDINO, Rejane; SABBAG, Deise Maria Antonio. TemaTres e a construção de tesouros: aspectos de aplicabilidade para o profissional da informação. **Revista Brasileira de Biblioteconomia e Documentação**, v. 13, p. 2566-2579, 2017. Disponível em: <https://rbbd.febab.org.br/rbbd/article/view/1040>. Acesso em: 05 mai. 2021.

GARCÍA-MARCO, Francisco-Javier. ISO 25964-2: interoperabilidad de los tesauros con otros sistemas de organización del conocimiento en contextos semánticos. **Anuario Think EPI**, [S. l.], v. 11, p. 305-309, jan./dez. 2017. Disponível em: https://www.researchgate.net/publication/316591173_ISO_25964-2_interoperabilidad_de_los_tesauros_con_otros_sistemas_de_organizacion_del_conocimiento_en_contextos_semanticos. Acesso em: 6 jun. 2021.

GEDAM, Pranali; PARADKAR, Ashwini. A study of web-based library classification schemes. **International journal of library and information science**, Índia, v. 5, n. 10, p. 386-393, maio 2013. Disponível em: https://academicjournals.org/article/article1381239290_Gedam%20and%20Paradkar.pdf. Acesso em: 6 jun. 2021.

GERRING, John; MCDERMOTT, Rose. An experimental template for case study research. **American Journal of Political Science**, [S. l.], v. 51, n. 3, p. 688-701, mar. 2007. Disponível em: <https://www.jstor.org/stable/4620092?seq=1>. Acesso em: 6 jun. 2021.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 5 ed. São Paulo: Atlas, 2010.

GNOLI, Claudio. Progress in synthetic classification: towards unique definition of concepts Item type Conference Paper. **Extensions and corrections to the UDC**, Haia, v. 29, n. 1, p. 4-5, jan./jun. 2007. Disponível em: <http://hdl.handle.net/10150/105614>. Acesso em: 12 jan. 2019.

GNOLI, Claudio. Notation. **Knowledge Organization**, [S. l.], v. 45, n. 8, p. 667-684, ago. 2019. Disponível em: <https://www.isko.org/cyclo/notation>. Acesso em: 6 jun. 2021.

GOMES, Hagar Espanha. Classificação, tesouro e terminologia: fundamentos comuns. **Palestra preparada para as Tertúlias do Departamento de Biblioteconomia da UNIRIO**, Rio de Janeiro, 1996.

GOMES, Hagar Espanha; CAMPOS, Maria Luiza de Almeida. Tesouro e normalização terminológica: o termo como base para intercâmbio de informações. **DatagramaZero**, Rio de Janeiro, v. 5, n. 1, p. 1-11, dez. 2004. Disponível em: <https://brapci.inf.br/index.php/res/v/5688>. Acesso em: 6 jun. 2021.

GOMES, Hagar Espanha; CAMPOS, Maria Luiza de Almeida; GUIMARÃES, Ludmila dos Santos. Organização da Informação e Terminologia: a abordagem onomasiológica. **DataGramZero: Revista de Ciência da Informação**, Rio de Janeiro, v. 11, n. 5, p. 1-12, maio 2010. Disponível em: <https://brapci.inf.br/index.php/res/v/7184>. Acesso em: 6 jun. 2021.

GOMES, Hagar Espanha; CAMPOS, Maria Luiza de Almeida. A organização do conhecimento na Web: contribuições de Shiyali Ramamrita Ranganathan e de Ingetraut Dahlberg. *In*: INSTITUTO DE ARTE E COMUNICAÇÃO SOCIAL, 1., 2019, Niterói. **Anais [...]**. Niterói: UFF, 2019.

GONZALES-AGUILAR, Audilio; RAMÍREZ-POSADA, María; FERREYRA, Diego. Tematres: software para gestionar tesauros. **El profesional de la información**, [S. l.], v. 21, n. 3, p. 319-325, maio/jul. 2012. Disponível em: <https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/epi.2012.may.14/17928>. Acesso em: 15 jul. 2021.

GREEN, Rebecca. Relationships in knowledge organization. **Knowledge Organization**, [S. l.], v. 35, n. 2/3, p. 150-159, fev./mar. 2008. Disponível em: https://www.researchgate.net/publication/286481942_Relationships_in_Knowledge_Organization. Acesso em: 6 jun. 2021.

GREEN, Rebecca. Facet Analysis and Semantic Frames. **Knowledge Organization**, [S. l.], v. 44, n. 6, p. 397-404, jun. 2017. Disponível em: https://www.ergon-verlag.de/isko_ko/downloads/ko_44_2017_6_b.pdf. Acesso em: 6 jun. 2021.

GUINCHAT, Claire; MENOUE, Michael. **Introdução geral às ciências e técnicas da informação e documentação**. Brasília: IBICT, 1994.

HAJDU-BARÁT, Ágnes. Integration of a thesaurus and Universal Decimal Classification (UDC) to improve subject access: the Hungarian experience. **Extensions & Corrections to the UDC**, [S. l.], v. 31, p. 197-207, jan./dez. 2009. Disponível em: https://repository.arizona.edu/bitstream/handle/10150/199891/Hadju-Barat_E%26C31_2009.pdf?sequence=1&isAllowed=y. Acesso em: 6 jun. 2021.

HARTIG, Olaf; BIZER, Christian; FREYTAG, Johann-Christoph. Executing SPARQL queries over the web of linked data. *In: INTERNATIONAL SEMANTIC WEB CONFERENCE*, 1., 2009, Berlin. **Proceedings** [...]. Berlin: Springer, 2009. p. 293-309. Disponível em: http://olafhartig.de/files/HartigEtAl_QueryTheWeb_ISWC09_Preprint.pdf. Acesso em: 6 jun. 2021.

HIGGINS, Julian; GREEN, Sally. **Cochrane Handbook of Systematic Reviews of Intervention**. London: The Cochrane Collaboration, 2011. Disponível em: <https://handbook-5-1.cochrane.org>. Acesso em: 10 set. 2019.

HIGGINS, Julian *et al.* (Ed.). **Cochrane handbook for systematic reviews of interventions**. [S. l.]: John Wiley & Sons, 2019.

HILL, Linda *et al.* Integration of knowledge organization systems into digital library architectures. **Data Analysis and Knowledge Discovery**, [S. l.], v. 20, n. 1, p. 4-8, jan./jun. 2004. Disponível em: <https://journals.lib.washington.edu/index.php/acro/article/view/13835>. Acesso em: 6 jun. 2021.

HINDSOTI, Richard. Reflections on the utilization of the Universal Decimal Classification. **Aslib Proceedings**, U.K., v. 31, n. 6, p. 305-311, jun. 1979.

HJØRLAND, Birger; NISSEN PEDERSEN, Karsten. A substantive theory of classification for information retrieval. **Journal of documentation**, London, v. 61, n. 5, p. 582-597, maio 2005. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/00220410510625804/full/html>. Acesso em: 5 jun. 2021.

HJØRLAND, Birger. Semantics and knowledge organization. **Annual review of information science and technology**, New York, v. 41, n. 1, p. 367-405, jan./dez. 2007a. Disponível em: https://www.researchgate.net/publication/220142011_Semantics_and_Knowledge_Organizat ion. Acesso em: 6 jun. 2021.

HJØRLAND, Birger. Arguments for 'The Bibliographical Paradigm.' Some Thoughts Inspired by the New English Edition of the UDC. **Information Research**, [S. l.], n. 4, v. 12, p. 1-17, abr. 2007b. Disponível em: <http://InformationR.net/ir/12-4/colis/colis06.html>. Acesso em: 10 set. 2019.

HJØRLAND, Birger. Does the traditional thesaurus have a place in modern information retrieval?. **Knowledge Organization**, [S. l.], v. 43, n. 3, p. 145-159, jan./dez. 2016. Disponível em: https://www.ergon-verlag.de/isko_ko/downloads/ko_43_2016_3_c.pdf. Acesso em: 6 jun. 2021.

HODGE, Gail. **Systems of knowledge organization for digital libraries**: beyond traditional authorities files. Washington, DC: Council on Library and Information Resources, 2000. Disponível em: <https://old.diglib.org/pubs/df090/df090.pdf>. Acesso em: 5 jan. 2020.

HORSNELL, Verina. The Intermediate Lexicon: an aid to international co-operation. **Aslib Proceedings**, [S. l.], v. 27, n. 2, p. 57-66, fev. 1975.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO) 22709. **Information and documentation**: Format for information exchange. Genebra: ISO, 2011.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO) 25964-1. **Information and documentation Thesauri and interoperability with other vocabularies**: Part 1: Thesauri for information retrieval. Genebra: ISO, 2011.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO) 25964-2. **Thesauri and interoperability with other vocabularies**: Part 2: Interoperability with other vocabularies. Genebra: ISO, 2013.

ISAAC, Antoine; SUMMERS, Ed. SKOS simple knowledge organization system primer. *In*: PRIMER, WORLD WIDE WEB CONSORTIUM (W3C), 1., 2009, [S. l.]. **Proceedings** [...]. [S. l.]: W3C, 2009. Disponível em: <https://www.w3.org/2004/02/skos/>. Acesso em: 6 jun. 2021.

ISAAC, Antoine *et al.* **Library Linked Data incubator group**: datasets, value vocabularies, and metadata element. World Wide Web Consortium, [S. l.], 2011. Disponível em: <https://www.w3.org/2005/Incubator/ld/XGR-ld-vocabdataset-20111025/>. Acesso em: 5 jun. 2021.

JARMASZ, Mario; SZPAKOWICZ, Stan. Roget's thesaurus: A lexical resource to treasure. **arXiv preprint arXiv:1204.0258**, 2012. Disponível em: <https://arxiv.org/abs/1204.0258>. Acesso em: 10 mai. 2020.

KAOSAR, Amina. **Merit & Demerit of using Universal Decimal Classification on the Internet**. 2008. Dissertação (Mestrado em Ciência da Informação) - Royal School of Library and Information Science, Universidade de Copenhagen, Copenhagen, 2008. Disponível em: <https://www.yumpu.com/en/document/view/19028676/merit-demerit-of-using-universal-decimal-classification-on-the->. Acesso em: 6 jun. 2021.

KHOO, Christopher; NA, Jin-Cheon. Semantic relations in information science. **Annual Review of Information Science and Technology**, [S. l.], v. 40, n. 1, p. 157-228, jan./jun. 2006. Disponível em: <https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/aris.1440400112>. Acesso em: 6 jun. 2021.

KOCH, Traugott *et al.* **The role of classification schemes in Internet resource description and discovery**, [S. l.], 1997.

KOCH, Traugott; TUDHOPE, Douglas. New applications of knowledge organization systems: introduction to a special issue. **Journal of Digital Information**, [S. l.], v. 4, n. 4, abr. 2003.

KOCH, Traugott; TUDHOPE, Douglas. **User-centred approaches to Networked Knowledge Organization Systems/Services (NKOS)**. [S. l.]: Background, 2004.

LANCASTER, Frederick Wilfrid. **Vocabulary control for information retrieval**. 2. ed. Arlington, Va.: Information Resources Press, 1986.

LANCASTER, Frederick Wilfrid; WARNER, Andrew Jackson. **Information Retrieval Today**. Arlington: Information Resources Press, 1993.

LANCASTER, Frederick Wilfrid. **Indexação e resumos: teoria e prática**. 2. ed. Brasília: Briquet de Lemos, 2004.

LANGRIDGE, Derek. **Classificação: abordagem para estudantes de biblioteconomia**. Rio de Janeiro: Interciência, 1977.

LA BARRE, Kathryn. The heritage of early FC in document reference retrieval systems, 1920-1969. **Library History**, Champaign, EUA, v. 23, n. 2, p. 129-149, jul./dez. 2007. Disponível em: <https://www.tandfonline.com/doi/abs/10.1179/174581607x205653>. Acesso em: 6 jun. 2021.

LEE, Deborah. Numbers, instruments and hands: the impact of faceted analytical theory on classifying music ensembles. **Knowledge Organization**, [S. l.], v. 44, n. 6, p. 405-415, jun. 2017. Disponível em: <https://openaccess.city.ac.uk/id/eprint/18645/>. Acesso em: 5 jun. 2021.

LERTMAHAKIAT, Wilaiporn; KAMPEERAPAAPPAT, Poorivat; MINGKHWAN, Anirach **DDC Multiple Relation Content Retrieval Framework**, Proceedings PGNet, U.K., 2008.

LE COADIC, Yves-François. **A ciência da informação**. Brasília: Briquet de Lemos, 2004.

LIMA, G. Â. B. O. **Mapa hipertextual (MHTX): um modelo para organização hipertextual de documentos**. 2004. 204 f. Tese (Doutorado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2004. Disponível em: <<http://www.bibliotecadigital.ufmg.br/dspace/handle/1843/LHLS-6BUPG9>>. Acesso em: 25 jun. 2021.

LIMA, Vânia Mara Alves. A informação documentária: codificação e decodificação. **Transinformação**, Campinas, v. 19, n. 2, p. 119-127, jul./dez. 2007. Disponível em: <https://www.scielo.br/j/tinf/a/cjDVFQPfgNgxqLP3sB8QX5d/abstract/?lang=pt>. Acesso em: 6 jun. 2021.

LIMA, Gercina Ângela de; MACULAN, Benildes Coura Moreira dos Santos. Estudo comparativo das estruturas semânticas em diferentes sistemas de organização do conhecimento. **Ciência da Informação**, Brasília, v. 46, n. 1, p. 60-72, jan./jun. 2017. Disponível em: <http://revista.ibict.br/ciinf/article/view/4014>. Acesso em: 6 jun. 2021.

LIMA, Gercina Ângela de. Organização e representação do conhecimento e da informação na web: teorias e técnicas. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 25, p. 57-97, fev. 2020. Disponível em: <https://periodicos.ufmg.br/index.php/pci/article/view/22283>. Acesso em: 31 maio 2021.

LUCAS, Elaine Rosangela de Oliveira; CORRÊA, Elisa Cristina Delfini; EGGERT-STEINDEL, Gisela. As contribuições de Ranganathan para a Biblioteconomia: reflexões e desafios. In: **As contribuições de Ranganathan para a Biblioteconomia: reflexões e desafios**. 2016. p. 222-222.

MACULAN, Benildes Coura Moreira dos Santos. **Taxonomia facetada navegacional: construção a partir de uma matriz categorial para trabalhos acadêmicos**. 2011. 195f. Dissertação (Mestrado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2011. Disponível em: <https://repositorio.ufmg.br/handle/1843/ECID-8LAN5K>. Acesso em: 31 maio 2021.

MACULAN, Benildes Coura Moreira dos Santos. **Estudo e aplicação de metodologia para reengenharia de tesouro: remodelagem do THESAGRO**. 2015. 339 f. Tese (Doutorado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2015. Disponível em: <http://www.bibliotecadigital.ufmg.br/dspace/handle/1843/BUBD9ZKMUV>. Acesso em: 13 mar. 2020.

MAIA, Lucinéia Souza. **Extração e explicitação de relações semânticas para a representação do conhecimento de documentos acadêmicos: um estudo de caso a partir de uma estrutura classificatória**. 2018. 263 f. Tese (Doutorado em Gestão e Organização do Conhecimento) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2018. Disponível em: <https://repositorio.ufmg.br/handle/1843/VAFA-BB5JHW>. Acesso em: 20 ago. 2020.

MANIEZ, Jacques. Relationships in thesauri: Some critical remarks. **KO KNOWLEDGE ORGANIZATION**, [S. l.], v. 15, n. 3, p. 133-138, jan./dez. 1988.

MANIEZ, Jacques. Database merging and the compatibility of indexing languages. **Knowledge Organization**, v. 24, n. 4. p. 213-224, 1997.

MARCONDES, Carlos Henrique. **Dados abertos interligados**: publicação, recuperação e integração de acervos de arquivos, bibliotecas e museus na web. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2021.

MARCONI, Marilia de Andrade; LAKATOS, Eva Maria. **Fundamentos de metodologia científica**. 8. ed. São Paulo: Atlas, 2017.

MATVEYEVA, Susan. A role for classification: the organization of resources on the Internet. **MLA Forum**, Vancouver, v. 1, n. 2, p. 1, jul./dez. 2002. Disponível em: <http://www.mlaforum.org/volumel/issue2/roleClassification.html/>. Acesso em: 5 jul. 2018.

MAYR, Philipp *et al.* Recent applications of Knowledge Organization Systems: introduction to a special issue. **International Journal on Digital Libraries**, London, v. 17, n. 1, p. 1-4, mar. 2016. Disponível em: <https://link.springer.com/article/10.1007/s00799-015-0167-x>. Acesso em: 5 jun. 2021.

MCILWAINE, Ia. The Universal Decimal Classification: Some Factors Concerning Its Origins, Development, and Influence. **Journal of the American Society for Information Science**, London, v. 48, n. 4, p. 331-339, abr. 1997.

MCILWAINE, Ia. Classification at a Crossroads: Multiple Directions to Usability. **International Cataloguing and Bibliographic Control**: Quarterly Bulletin of the IFLA UBCIM Programme, London, v. 39, n. 2, p. 41-42, jul./dez. 2010. Disponível em: <https://www.proquest.com/openview/7deb2f5435b182877a1ee7557816261d/1?pq-origsite=gscholar&cbl=60376>. Acesso em: 4 jun. 2021.

MILES, Alistair; BRICKLEY, Dan. **SKOS Core Guide**. [S. l.]: W3C, 2005. Disponível em: <http://www.w3.org/TR/2005/WD-wbp-skos-core-guide-20051102/>. Acesso em: 10 jul. 2020.

MILLS, Jack. Progress in Documentation: Library Classification. **Journal of documentation**, Califórnia, EUA, v. 26, n. 2, p. 120-160, jul./dez. 1970. Disponível em: <http://dx.doi.org/10.1108/eb026491>. Acesso em: 8 set. 2020.

MIRANDA, Marcos Luiz Cavalcanti de. **Organização e Representação do Conhecimento**: fundamentos teórico-metodológicos na busca e recuperação da informação em ambientes virtuais. 2005. 353f. Tese (Doutorado em Ciência da Informação) - Curso de Pós-Graduação em Ciência da Informação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005. Disponível em: <https://ridi.ibict.br/handle/123456789/690>. Acesso em: 6 jun. 2021.

MOMM, Christiane Fabíola; LESSA, Rafael Orivaldo. Sistema de classificação bibliográfica e a conceituação do turismo: uma visão da CDU. **Perspectivas em Ciência da Informação**, v. 14, n. 2, p. 141-154, 2009.

MONTEIRO, Silvana Drumond *et al.* Sistemas de recuperação da informação e o conceito de relevância nos mecanismos de busca: semântica e significação. **Encontros Bibli**: revista eletrônica de biblioteconomia e ciência da informação, Florianópolis, v. 22, n. 50, p. 161-175, jan./jun. 2017. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2017v22n50p161>. Acesso em: 6 jun. 2021.

MURPHY, M. Lynne. **Relações semânticas e o léxico: antonímia, sinonímia e outros paradigmas**. [S. l.], Cambridge University Press, 2003.

NETTO, Cristiane Mendes. **Proposta de modelo de requisitos para ferramentas de visualização de ontologia de domínio**. 2017. 163 f. Tese (Doutorado em Gestão e Organização do Conhecimento) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2017. Disponível em: <https://repositorio.ufmg.br/bitstream/1843/ECIP-AZHNS9/1/tesefinalcristianemendes.pdf>. Acesso em: 10 ago. 2020.

NEVILLE, Hugh. Feasibility study of a scheme for reconciling thesauri covering a common subject. **Journal of documentation**, [S. l.], v. 26, n. 4, p. 313-336, abr. 1970.

NEWMAN, Simon. **Information systems compatibility**. Washington: Spartan Books; London: Macmillan, 1965.

NUNES, Leiva; TÁLAMO, Maria de Fátima Gonçalves Moreira. Da filosofia da classificação à classificação bibliográfica. **RDBCI**: Revista Digital de Biblioteconomia e Ciência da Informação, Campinas, v. 7, n. 2, p. 30-48, jul./dez. 2009. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/1973>. Acesso em: 6 jun. 2021.

OLIVEIRA, Carolina. C. *et al.* A terminologia aplicada ao projeto INTAGRO na perspectiva da organização e difusão da informação e da representação do conhecimento. In: **Embrapa Informática Agropecuária-Artigo em anais de congresso (ALICE)**. In: CONGRESSO INTERINSTITUCIONAL DE INICIAÇÃO CIENTÍFICA, 6., 2012, Jaguariúna. Anais... Jaguariúna: Embrapa; ITAL, 2012., 2012.

OLIVEIRA, Elaine Diamantino. **Proposta de um modelo de hipertexto com abordagem semântica para a representação do conhecimento no domínio temático da intensificação agropecuária**. 2018. 233f. Dissertação (Mestrado em Gestão e Organização do Conhecimento) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2018. Disponível em: [https://repositorio.ufmg.br/handle/1843/ECIP-B2EJSW#:~:text=Abstract%3A,de%20Ester%20Boserup%20\(1965\)](https://repositorio.ufmg.br/handle/1843/ECIP-B2EJSW#:~:text=Abstract%3A,de%20Ester%20Boserup%20(1965)). Acesso em: 6 jun. 2021.

PASTOR-SÁNCHEZ, Juan-Antonio; MARTÍNEZ MÉNDEZ, Francisco Javier; RODRÍGUEZ-MUÑOZ, José Vicente. Advantages of Thesaurus Representation Using the Simple Knowledge Organization System (SKOS) Compared with Proposed Alternatives. **Information Research**: an international electronic journal, [S. l.], v. 14, n. 4, p. 4, abr. 2009. Disponível em: <http://informationr.net/ir/14-4/paper422.html>. Acesso em: 6 jun. 2021.

PENTEADO, Maria Isabel de Oliveira *et al.* A trajetória do planejamento da pesquisa na Embrapa. **Cadernos de Ciência & Tecnologia**, Brasília, v. 31, n. 1, p. 35-60, jan./abr. 2014. Disponível em: <https://seer.sct.embrapa.br/index.php/cct/article/view/19442/12567>. Acesso em: 6 jun. 2021.

PIEIDADE, Maria Antonieta Requião. **Introdução à teoria da classificação**. Rio de Janeiro: Interciência, 1977.

PIEROZZI JUNIOR, Ivo; OLIVEIRA, Leandro Henrique Mendonça de; SOUZA, Kleber Xavier Sampaio de. Construindo ontologias de domínio: o (re)conhecimento da intensificação agropecuária no Brasil. *In*: SEMINÁRIO DE PESQUISA EM ONTOLOGIA NO BRASIL, 3., 2010, Florianópolis. **Anais** [...]. Florianópolis: UFSC, 2010. p. 100-109.

PIEROZZI JUNIOR, Ivo; OLIVEIRA, Leandro Henrique Mendonça de; ALENCAR, Maria de Cleófas Faggion; ALMEIDA, Gladis Maria de Barcelos Almeida. **Sistema de organização do conhecimento sobre intensificação agropecuária**: agregando valor às redes de conhecimento da Embrapa. Embrapa Informática Agropecuária, Brasília, 2013. Disponível em: <https://ainfo.cnptia.embrapa.br/digital/bitstream/item/102345/1/isko2013.pdf>. Acesso em: 6 jun. 2021.

PIEROZZI JUNIOR, Ivo *et al.* **Engenharia da informação**: contribuições para a Agricultura digital. [S.l.]: Embrapa Informática Agropecuária, 2020.

PIKA, Jiri. Universal Decimal Classification at the ETH-Bibliothek Zürich-a Swiss perspective. **UDC Consortium**, London, v. 29, n. 1, p. 229–251, jan./jun. 2007. Disponível em: <http://hdl.handle.net/10150/105111>. Acesso em: 6 jun. 2021.

PONTES, Flávio Vieira. **Organização do conhecimento em bibliotecas digitais de teses e dissertações**: uma abordagem baseada na classificação facetada e taxonomias dinâmicas. 2013. 233f. Tese (Doutorado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2013. Disponível em: <http://repositorios.questoesemrede.uff.br/repositorios/handle/123456789/2049>. Acesso em: 6 jun. 2021.

PRASAD, A. R. D.; MADALLI, Devika P. Classificatory ontologies. **Extensions & Corrections to the UDC**, [S. l.], v. 31, n. 1, p. 223-232, jan./jun. 2009. Disponível em: <https://repository.arizona.edu/handle/10150/199910>. Acesso em: 6 jun. 2021.

RANGANATHAN, Shiyali Ramamrita. **Prolegomena to library classification**. Bombay: Asia Publishing House, 1967.

RANGANATHAN, Shiyali Ramamrita. Choice of scheme for classification. **Ranganathan Portal**, [S. l.], v. 5, n. 1, p. 1-69, jan./jun. 1968.

RIESTHUIS, Gerhard. Theory of Compatibility of Information Languages. *In*: COMPATIBILITY AND INTEGRATION OF ORDER SYSTEMS, 1., 1995, Warsawa. **Proceedings** [...]. Warsawa: Wydaw, 1996. p. 23-31.

RIESTHUIS, Gerhard. *et al.* Decomposition of UDC-numbers and the text of the UDC Master Reference. *In*: FIFTH INTERNATIONAL ISKO CONFERENCE, 1., 1998, Lille. **Proceedings** [...]. Lille: [S. n.], 1998. p. 25-29.

RIESTHUIS, Gerhard. Searching with words: re-use of subject indexing. **Extensions and corrections to the UDC**, [S. l.], v. 21, p. 24-32, jan./dez. 1999.

RIESTHUIS, Gerhard. Information Language and Multilingual Subject Access. *In*: IFLA SATELLITE MEETING HELD IN DUBLIN, 1., 2003, Ohio. **Proceedings** [...]. Ohio: [S.n.], 2003. p.11-18. Disponível em: <https://www.degruyter.com/document/doi/10.1515/9783110964912.11/html>. Acesso em: 5 jun. 2021.

RIZZO, Giuseppe; TRONCY, Raphaël. Nerd: evaluating named entity recognition tools in the web of data. *In*: WORKSHOP ON WEB SCALE KNOWLEDGE EXTRACTION (WEKEX'11), 1., 2011, [S. l.]. **Proceedings** [...]. [S. l.]: [S. n.], 2011. p. 1-16. Disponível em: https://www.researchgate.net/publication/228460023_NERD_Evaluating_Named_Entity_Recognition_Tools_in_the_Web_of_Data. Acesso em: 6 jun. 2021.

SALES, Rodrigo de; CAFÉ, Lígia. Diferenças entre tesouros e ontologias. **Perspectivas em Ciência da Informação**, v. 14, n. 1, p. 99-116, 2009.

SAMPAIO, Rosana Ferreira; MANCINI, Marisa Cota. Estudos de revisão sistemática: um guia para síntese criteriosa da evidência científica. **Revista Brasileira de Fisioterapia**, [S. l.], v. 11, n. 1, p. 83-89, fev. 2007. Disponível em: <https://www.scielo.br/j/rbfis/a/79nG9Vk3syHhnSgY7VsB6jG/?lang=pt>. Acesso em: 6 jun. 2021.

SÁNCHEZ, Montserrat Mateos; PANIAGUA, Carlos García-Figuerola. **Aplicación de técnicas de clustering en la recuperación de información web**. Gijón, Asturias: EdicionesTrea, 2009. Disponível em: <https://www.trea.es/books/aplicacion-de-tecnicas-de-clustering-en-la-recuperacion-de-la-informacion>. Acesso em: 6 jun. 2021.

SÁNCHEZ, Juan Antonio Pastor; MÉNDEZ, Francisco Xavier Martínez; MUÑOZ, José Vicente Rodríguez. SKOS application for interoperability of controlled vocabularies in the field of linked open data. **Profesional de la Información**. v.21, n. 3, p. 245–253, 2012. Disponível em: <https://1library.co/title/skos-application-interoperability-controlled-vocabularies-field-linked-open>. Acesso em: 05 mai. 2021.

SANTOS, Cibele Araújo Camargo Marques dos; MAZINI, Elizabeth Sardelli. Organização do conhecimento: das classificações e vocabulários controlados às taxonomias e ontologias na web. In: VALLS, Valéria Martin; VERGUEIRO, Waldomiro (Org.). **Tendências contemporâneas na gestão da informação**. São Paulo: Editora Sociologia e Política, 2011. p. 1-18.

SANTOS, José Carlos Francisco; MOREIRA, Walter. SKOS: uma análise sobre as abordagens e suas as aplicações na Ciência da Informação. **Informação & Informação**, v. 23, n. 3, p. 362-389, 2018.

SARACEVIC, Teko. Ciência da informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 1, n. 1, p. 41- 62, jan./jun. 1996. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/235>. Acesso em: 6 jun. 2021.

SARACEVIC, Teko. Information Science. **Journal of the american society for information science**, New Brunswick, New Jersey, v. 50, n. 12, p. 1051-1063, dez. 1999.

SATIJA, Mohinder; MARTÍNEZ-ÁVILA, Daniel. Features, Functions and Components of a Library Classification System in the LIS tradition for the e-Environment. **Journal of Information Science Theory and Practice**, [S. l.], v. 3, n. 4, p. 62-77, mar. 2015. Disponível em: https://www.researchgate.net/publication/288873086_Features_Functions_and_Components_of_a_Library_Classification_System_in_the_LIS_tradition_for_the_e-Environment. Acesso em: 5 jun. 2021.

SCHALLIER, Wouter. "What a subject search interface can do". **Extensions & Corrections to the UDC**, Bélgica, v. 26, n. 1, p.18-28, jan./jun. 2004. Disponível em: <https://repository.arizona.edu/handle/10150/105315>. Acesso em: 6 jun. 2021.

SAN SEGUNDO, Rosa. At a Crossroads: Multiple Directions to Usability Using MARC classification format for UDC and mappings to other KO systems for an enriched authority file. **Extensions & Corrections to the UDC**, Bélgica, v. 31, n. 1, p. 285–291, jan./jun. 2009. Disponível em: https://www.academia.edu/714857/Using_MARC_classification_format_for_UDC_and_mappings_to_other_KO_systems_for_an_enriched_authority_file. Acesso em: 6 jun. 2021.

SILVA, Odilson Pereira da; GANIM, Fátima. **Manual da CDU**. Brasília: Briquet de Lemos, 1994.

SILVA, Marcel Ferrante. **Proposta de modelo de colaboração para catálogo web facetado**. 2013. 269f. Tese (Doutorado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2013. Disponível em: <https://repositorio.ufmg.br/handle/1843/BUOS-9JLJJF>. Acesso em: 6 jun. 2021.

SILVA, Marcio Bezerra da; NEVES, Dulce Amélia de Brito. Estudo sobre o uso da teoria da classificação facetada em banco de dados. *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO*, 11., 2010, Rio de Janeiro. **Anais [...]**. Rio de Janeiro: ANCIB, 2010. p. 1-20. Disponível em: <http://repositorios.questoesemrede.uff.br/repositorios/handle/123456789/870>. Acesso em: 5 jun. 2021.

SEMIDÃO, Rafael Aparecido Moron. **Dados, informação e conhecimento enquanto elementos de compreensão do universo conceitual da ciência da informação: contribuições teóricas**. 2007. 199f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2007. Disponível em: [http:// https://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/semidao_ram_me_mar.pdf](http://https://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/semidao_ram_me_mar.pdf). Acesso em: 20 jun. 2019.

SIMÕES, Maria da Graça. **Relações conceptuais nos tesouros: norma internacional e sua aplicação prática**. Dissertação (Mestrado em Biblioteconomia), Facultad de Traducción y Documentación da Universidad de Salamanca, 2007. Disponível em: <https://eg.uc.pt/handle/10316/29011>. Acesso em: 05 dez. 2020.

SIMÕES, Maria da Graça. **A representação de etnia e a sua evolução na Classificação Decimal Universal**. 600f. Tese (Doutorado em Biblioteconomia e Documentação), Universidad de Salamanca, Salamanca, 2010. Disponível em: <https://estudogeral.uc.pt/handle/10316/20839>. Acesso em: 6 jun. 2021.

SIMÕES, Maria da Graça; FREITAS, Maria Cristina. A Classificação em arquivos e em bibliotecas à luz da Teoria da Classificação: pontos de convergência e de divergência. **PontodeAcesso**, Salvador, v. 7, n. 1, p. 81-115, jan./jun. 2013. Disponível em: <https://periodicos.ufba.br/index.php/revistaici/article/view/8050/0>. Acesso em: 6 jun. 2021.

SIMÕES, Maria da Graça; FREITAS, Maria Cristina; RODRÍGUEZ-BRAVO, Blanca. Theory of classification and classification in libraries and archives: Convergences and divergences. **Knowledge Organization**, [S. l.], v. 43, n. 7, p. 530-538, jul. 2018. Disponível em: <https://estudogeral.sib.uc.pt/handle/10316/36486>. Acesso em: 6 jun. 2021.

SINGHAL, Amit. Modern information retrieval: A brief overview. **Bulletin of the IEEE Computer Society Technical Committee on Data Engineering**, [S. l.], v. 1, p. 1-9, jan./jun. 2009. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.117.7676>. Acesso em: 6 jun. 2021.

SLAVIC, Aida. **UDC implementation: from library shelves to a structured indexing language.** International cataloguing and bibliographic control, London, 2004a. Disponível em: <https://www.semanticscholar.org/paper/UDC-implementation%3A-From-library-shelves-to-a-Slavic/5edf912d287294f6753198a4b1e9565bb09390dd>. Acesso em: 7 jun. 2021.

SLAVIC, Aida. UDC translations: a 2004 survey report and bibliography. **Extensions & Corrections to the UDC**, [S. l.], v. 26, p. 58-80, jan./dez. 2004b. Disponível em: https://www.researchgate.net/publication/277797344_UDC_Translations_a_2004_Survey_Report_and_Bibliography Acesso em: 10 mai. 2020.

SLAVIC, Aida. **Classification management and use in a networked environment: the case of the Universal Decimal Classification.** 2005. 463f. Tese (Doutorado em Filosofia), University of London, London, 2005. Disponível em: <https://discovery.ucl.ac.uk/id/eprint/1334914/>. Acesso em: 6 jun. 2021.

SLAVIC, Aida. Interface to classification: some objectives and options Item type Preprint Interface to Classification: Some Objectives and Options. **Extensions and Corrections to the UDC**, Haia, v. 28, p. 1-22, jan./dez. 2006a. Disponível em: <http://hdl.handle.net/10150/105459>. Acesso em: 5 out. 2018.

SLAVIC, Aida. The level of exploitation of Universal Decimal Classification in library OPACs: A pilot Study. **Vjesnik bibliotekara Hrvatske**, [S. l.], v. 49, n.3-4, p.155-182, jan./dez. 2006b. Disponível em: <https://repository.arizona.edu/handle/10150/105346>. Acesso em: 6 jun. 2021.

SLAVIC, Aida. UDC in subject gateways: experiment or opportunity?. **Knowledge Organization**, [S. l.], v. 2, n. 33, p. 67- 85, jul./dez. 2006c. Disponível em: <http://www.udcc.org>. Acesso em: 18 mai. 2019.

SLAVIC, Aida; CORDEIRO, Marialnes; RIESTHUIS, Gerhard. Enhancement of UDC data for use and sharing in a networked environment. *In: ANNUAL CONFERENCE OF THE GERMAN CLASSIFICATION SOCIETY ON DATA ANALYSIS, MACHINE LEARNING, AND APPLICATIONS*, 31., Freiburg, 2007. **Proceedings** [...]. Freiburg: [S.n.], 2007. Disponível em: <http://hdl.handle.net/10150/106330>. Acesso em: 8 set. 2020.

SLAVIC, Aida. Faceted classification: management and use. **Axiomathes**, [S. l.], v. 18, n. 2, p. 257, jul./dez. 2008. Disponível em: <https://arxiv.org/abs/1705.07047>. Acesso em: 6 jun. 2021.

SLAVIC, Aida; CORDEIRO, Maria Inês; RIESTHUIS, Gerhard. Maintenance of the Universal Decimal Classification: overview of the past and preparations for the future. **International Cataloguing and Bibliographic Control Journal International Cataloguing and Bibliographic Control**, London, v. 37, n. 2, p. 23-29, jul./dez. 2008. Disponível em: <https://repository.arizona.edu/handle/10150/105220>. Acesso em: 6 jun. 2021.

SLAVIC, Aida. Classification revisited: a web of knowledge. **arXiv preprint arXiv: 1705.07058**, London, v. 1, p. 1-17, maio 2017. Disponível em: <https://arxiv.org/abs/1705.07058>. Acesso em: 6 jun. 2021.

SLAVIC, Aida; DAVIES, Sylvie. Facet Analysis in UDC: Questions of Structure, Functionality and Data Formality. **Knowledge Organization**, [S. l.], v. 44, n. 6, p. 425-435, jan./dez. 2017. Disponível em: https://www.researchgate.net/publication/321985509_Facet_Analysis_in_UDC_Questions_of_Structure_Functionality_and_Data_Formality. Acesso em: 6 jun. 2021.

SOBRINHO, Heliomar Cavati; FERREIRA, Raquel Ellen Simões. TemaTres: uma análise acerca de suas funcionalidades. *In: CONGRESSO BRASILEIRO EM ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO*, 5., 2019, [S. l.]. **Anais [...]**. [S. l.]: [S. n.], 2019. p. 255-262. Disponível em: <https://brapci.inf.br/index.php/res/v/125256>. Acesso em: 6 jun. 2021.

SOERTEL, Dagobert. A universal source thesaurus as a classification generator. **Journal of the American Society for Information Science**, [S. l.], v. 23, n. 5, p. 229-305, set./out. 1972.

SOERTEL, Dagobert. **Indexing languages and thesauri**: construction and maintenance. Los Angeles: Melville, 1974.

SOERTEL, Dagobert. Compatibility of vocabularies. *In: THE CONTA CONFERENCE ON CONCEPTUAL AND TERMINOLOGICAL ANALYSIS IN THE SOCIAL SCIENCES*, 1., 1982, Frankfurt. **Proceedings [...]**. Frankfurt: [S. n.], 1982. p. 209-223.

SOERTEL, Dagobert. The art and architecture thesaurus (AAT): A critical appraisal. **Visual Resources**, [S. l.], v. 10, n. 4, p. 369-400, abr. 1995.

SOERTEL, Dagobert. The rise of ontologies or the reinvention of classification. **Journal of the American Society for Information Science**, London, v. 50, n. 12, p. 1119-1120, dez. 1999.

SOERTEL, Dagobert. Digital libraries and knowledge organization. *In: KRUK, Sebastian Ryszard (Org.). Semantic digital libraries*. Berlin: Springer, 2009. p. 9-39.

SOERTEL, Dagobert. **Knowledge Organization Systems**. Overview, Alexandria, 2014. Disponível em: <http://www.dsoertel.com/SoertelKOSOverview.pdf>. Acesso em: 20 ago. 2018.

SOUZA, Renato Rocha; ALVARENGA, Lúdia. A web semântica e suas contribuições para ciência da informação. **Ciência da Informação**, Brasília, v. 33, n. 1, p.132-141, jan./abr. 2004. Disponível em: <https://www.scielo.br/j/ci/a/sp3XpmZhXw384H5Fw9H89YL/abstract/?lang=pt>. Acesso em: 6 jun. 2021.

SOUZA, Renato Rocha. Sistemas de recuperação de informações e mecanismos de busca na web: panorama atual e tendências. **Perspectivas em ciência da informação**, Belo Horizonte, v. 11, n. 2, p. 161-173, jul./dez. 2006. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/320>. Acesso em: 5 jun. 2021.

SOUZA, Renato Rocha; TUDHOPE, Douglas; ALMEIDA, Maurício Barcellos. Towards a taxonomy of KOS: Dimensions for classifying Knowledge Organization Systems. **Knowledge Organization**, [S. l.], v. 39, n. 3, p. 179-192, mar. 2012. Disponível em: http://mba.eci.ufmg.br/downloads/Souza_Tudhope_Almeida_-_KOS_Taxonomy.Submitted.pdf. Acesso em: 6 jun. 2021.

SOUZA, Sebastião de. **CDU: como entender e utilizar a 2ª edição-padrão internacional em língua portuguesa**. 3. ed. Brasília: Thesaurus, 2012.

SPARCK JONES, Karen. Some thoughts on classification for retrieval. **Journal of Documentation**, London, v. 26, n. 2, p. 89-101, jul./dez. 1970.

STOCK, Wolfgang. Concepts and semantic relations in information science. **Journal of the American Society for Information Science and Technology**, [S. l.], v. 61, n. 10, p. 1951-1969, out. 2010. Disponível em: <http://onlinelibrary.wiley.com/doi/10.1002/asi.21382/full>. Acesso em: 9 jan. 2021.

SUMMERS, Ed *et al.* LCSH, SKOS and linked data. *In*: INTERNATIONAL CONFERENCE ON DUBLIN CORE AND METADATA APPLICATIONS, 1., 2008, [S.l.]. **Proceedings [...]**. [S. l.]: [S. n.], 2008. p. 25-33. Disponível em: <https://dcpapers.dublincore.org/pubs/article/viewFile/916/912>. Acesso em: 6 jun. 2021.

SVENONIUS, Elaine. Use of classification in on-line retrieval. **Library resources and technical services**, [S. l.], v. 27, n. 1, p. 76-80, jan./jun. 1983.

SVENONIUS, Elaine. **The intellectual foundation of information organization**. [S.l.]: MIT press, 2000.

SVENONIUS, Elaine. **The epistemological foundations of knowledge representations**, [S. l.], p. 572-587, 2004. Disponível em: https://sites.evergreen.edu/wp-content/uploads/sites/226/2016/08/svenonius-epistemological-roots-of-knowledge-organization.asp_.pdf. Acesso em: 6 jun. 2021

TENNIS, Joseph. Ranganathan's layers of classification theory and the FASDA model of classification. **Proceedings from North American Symposium on Knowledge Organization**, [S. l.], n. 1, p. 185-195, jan./jun. 2011. Disponível em: https://www.researchgate.net/publication/265035073_Ranganathan's_Layers_of_Classification_Theory_and_the_FASDA_Model_of_Classification. Acesso em: 6 jun. 2021.

TORLONI, Maria Regina *et al.* Prepregnancy BMI and the risk of gestational diabetes: a systematic review of the literature with meta-analysis. **Obesity reviews**, [S. l.], v. 10, n. 2, p. 194-203, jul./dez. 2009. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/19055539/>. Acesso em: 6 jun. 2021.

TRISTÃO, Ana Maria Delazar *et al.* Sistemas de classificação facetada e tesauros: instrumentos para organização do conhecimento. **Ciência da Informação**, Brasília, v. 33, n. 2, p. 161-171, maio/ago. 2004. Disponível em: <https://www.scielo.br/j/ci/a/kwqV8zpzRMbXJQH7HsTmjys/?lang=pt&format=pdf>. Acesso em: 6 jun. 2021.

UDC CONSORTIUM. **Universal Decimal Classification**, [S. l.], 2020. Disponível em: <http://www.udcc.org/index.php>. Acesso em: 10 nov. 2020.

ULLAH, Asim; KHUSRO, Shah; ULLAH, Irfan. Bibliographic Classification in the Digital Age: Current Trends and Future Directions. **Information Technology and Libraries**, [S. l.], v. 36, n. 3, p. 48-77, abr. 2017. Disponível em: <https://ejournals.bc.edu/index.php/ital/article/view/8930>. Acesso em: 6 jun. 2021.

UNIVERSIDADE DE SÃO PAULO. **Tesouro Organização do Conhecimento**, São Paulo, 2015. Disponível em: <http://www2.eca.usp.br/mapasconceituais/vocab/index.php>. Acesso em: 3 jun. 2021.

VENTURA, Magda Maria. O estudo de caso como modalidade de pesquisa. **Revista SoCERJ**, [S. l.], v. 20, n. 5, p. 383-386, maio 2007. Disponível em: http://sociedades.cardiol.br/socerj/revista/2007_05/a2007_v20_n05_art10.pdf. Acesso em: 6 jun. 2021.

VICKERY, Brian Campbell. **Classificação e indexação nas ciências**. Rio de Janeiro: BNG/Brasilart, 1980.

VICKERY, Brian Campbell. Classificatory principles in intelligent interfaces. Tools for knowledge organization and the human interface. *In*: INTERNATIONAL ISKO CONFERENCE, 1., 1990, Darmstadt. **Proceedings** [...]. Darmstadt: [S. n.], 1990. p. 14-20.

VIZINE-GOETZ, Diane. **Using library classification schemes for internet resources**. [S. l.]: OCLC, 1996.

VUKADIN, Ana; SLAVIC, Aida. Challenges of Facet analysis and Concept Placement in Universal Classifications: the Example of Architecture in UDC. **Knowledge Organization in the 21 st century**, [S. l.], v. 1, p. 236-243, jan./jun. 2014. Disponível em: <https://arxiv.org/ftp/arxiv/papers/1706/1706.07931.pdf>. Acesso em: 6 jun. 2021.

WALL, Eugene; BARNES, Julian M. **Intersystem compatibility and convertibility of subject vocabularies**. Philadelphia: Auerbach, 1969.

WALL, Raymond. Indexing language structure for automated retrieval. **Information Storage and Retrieval**, [S. l.], v. 9, n. 11, p. 607-617, nov. 1973.

WALL, Raymond. Intelligent indexing and retrieval: A man-machine partnership. **Information Processing & Management**, Loughborough, England, v. 16, n. 1, p. 33-90, jan./jun. 1980.

WEISS, Leila Cristina. **Relações semânticas em tesouros**: um estudo da abordagem pragmática. 2014. 168f. Dissertação (Mestrado em Ciência da Informação) - Centro de Ciências da Educação, Universidade Federal de Santa Catarina, Florianópolis, 2014. Disponível em: <https://repositorio.ufsc.br/handle/123456789/128984>. Acesso em: 05 fev. 2021.

WEISS, Leila Cristina; BRÄSCHER, Marisa. Relações semânticas em tesouros: contribuições da abordagem pragmática. **InCID: Revista de Ciência da Informação e Documentação**, Ribeirão Preto, v. 7, n. 2, p. 136-155, jul./dez. 2017.

WEISS, Leila Cristina. **Interoperabilidade semântica**: uma análise sob a perspectiva da abordagem ontológica de willard van orman quine. 2019. 299f. Tese (Doutorado em Ciência da Informação) - Centro de Ciências da Educação, Universidade Federal de Santa Catarina, Florianópolis, 2019. Disponível em: <https://repositorio.ufsc.br/handle/123456789/194269>. Acesso em: 10 abr. 2021.

YIN, Robert. **Estudo de caso**: planejamento e métodos. 4. ed. Porto Alegre: Bookman, 2010.

ZENG, Marcia Lei; CHAN, Lois Mai. Trends and issues in establishing interoperability among knowledge organization systems. **Journal of the American Society for information science and technology**, [S. l.], v. 55, n. 5, p. 377-395, maio 2004. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/asi.10387>. Acesso em: 7 jun. 2021.

ZENG, Marcia Lei. Knowledge Organization Systems (KOS). **Knowledge Organization**, [S. l.], v. 35, n. 2-3, p. 160-182, fev. 2008. Disponível em: <https://www.isko.org/cyclo/kos>. Acesso em: 6 jun. 2021.

ZENG, Marcia Lei; PANZER, Michael; SALABA, Athena. Expressing classification schemes with OWL 2 web ontology language. *In: INTERNATIONAL SOCIETY FOR KNOWLEDGE ORGANIZATION*, 1., 2009, [S. l.]. **Proceedings** [...]. [S. l.]: [S. n.], 2009. p. 356-362. Disponível em: https://www.ergon-verlag.de/isko_ko/downloads/aiko_vol_12_2010_51.pdf. Acesso em: 7 jun. 2021.

ZENG, Marcia Lei. Interoperability. **Knowledge Organization**, [S. l.], v. 46, n. 2, p. 122-146, fev. 2019. Disponível em: <https://www.isko.org/cyclo/interoperability>. Acesso em: 6 jun. 2021.

ZHANG, Xueying. Concept integration of document databases using different indexing languages. **Information Processing & Management**, [S. l.], v. 42, n. 1, p. 121-135, jan. 2006. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0306457304001347>. Acesso em: 6 jun. 2021.

ANEXO A - Estrutura Classificatória da CDU (Classe 631 - Agricultura em geral)

631 Agricultura em geral

631.1 Administração e organização de fazendas

631.1.016 Formação, crescimento, declínio de fazendas ou propriedades agrícolas. Fusões, conversões, etc.

631.1.017 Área e extensão da fazenda ou propriedade.

631.1.017.1 Fazendas, plantações, etc. estabelecimentos agrícolas de grande porte.

631.1.017.3 Fazendas pequenas. Sítios. Pequenas propriedades agrícolas, chácaras.

631.1.02 Condições especiais, acordos, etc.

631.1.027 Condições de comercialização e distribuição.

631.11 Fazendas, propriedades e sistemas agrícolas

631.111 Localização da fazenda, área e características do terreno (terreno).

631.111.1 Teoria sobre a localização das fazendas.

631.111.2 Situação e adequabilidade da propriedade agrícola (para plantios específicos, etc.).

631.111.3 Qualidade da terra.

631.111.4 Localização da fazenda em áreas especializadas.

631.112 Aquisição de fazendas. Compra, venda de fazendas.

631.115 Fazendas segundo o título de posse, propriedade e supervisão. Sistemas de administração.

631.115.1 Fazendas particulares.

631.115.11 Fazendas familiares.

631.115.13 Fazendas de parceria.

631.115.17 Fazendas de arrendamento. Fazendas arrendadas. Fazendas de parceria.

631.115.2 Fazendas com administrador. Fazendas corporativas.

631.115.6 Fazendas coletivas.

631.115.7 Fazendas de propriedade pública. Fazendas federais, comuns e municipais.

631.115.71 Fazendas estatais.

631.115.73 Fazendas comunitárias. Fazendas municipais.

631.115.75 Fazendas de empresa mistas.

631.115.8 Cooperativas agrícolas.

631.115.9 Fazendas de acordo com outros sistemas de administração.

631.117 Fazendas de ensino, pesquisa e demonstração.

631.117.2 Fazendas instrucionais. Fazendas escolares. Fazendas de treinamento.

631.117.4 Fazendas de pesquisa. Fazendas experimentais. Estações de pesquisa agrícola.

631.117.6 Fazendas de demonstração. Fazendas-piloto.

631.12 Equipamentos e materiais agrícolas em geral (seleção, aquisição, manutenção, etc.)

631.14 Lavoura e fazendas de acordo com produtos e produção

631.143 Fazendas sem gado.

631.145 Fábrica e Empresas agrícolas combinadas. Agricultura industrializada em geral.

631.147 Fazendas biotecnológicas, Ecológica (para tecnologia alternativa, intermediária e leve). Agricultura e fazendas orgânicas.

631.15 Produção da fazenda. Organização de fazendas. Administração de fazendas

631.151 Sistemas de produção e administração agrícola, da fazenda.

631.151.1 Sistema de produção agrícola extensivo.

631.151.2 Sistema de produção agrícola intensivo.

631.151.3 Sistema de produção agrícola especializado.

631.151.4 Sistema de produção agrícola diversificado.

631.151.6 Sistema de produção agrícola integrado.

631.151.61 Sistema de produção agrícola verticalmente integrado.

631.151.62 Sistema de produção agrícola horizontalmente integrado.

631.152 Princípios de administração da Agricultura.

631.153 Planejamento da produção agrícola.

631.153.3 Escolha do sistema de cultivo e criação.

631.153.4 Escolha do método de trabalho. Divisão de trabalho.

631.153.5 Supervisão de planos de produção.

631.153.7 Mudanças no cultivo e produção. Conversão para outros sistemas.

631.155 Controle de produção agrícola.

631.155.1 Metas e cotas de produção agrícola.

631.155.12 Objetivos de produção agrícola.

631.155.13 Quotas de produção agrícola.

631.155.2 Estoque de produtos.

631.155.3 Perdas de produtos.

631.155.4 Excedentes de produtos.

631.155.6 Subsídios.

- 631.157 Administração de empresas agrícolas, de fazendas.**
- 631.158 Administração de trabalho agrícola e pessoal.**
- 631.16 Finanças e avaliação econômica agrícolas**
- 631.162 Administração financeira da fazenda. Contabilidade, contabilidade, etc.**
- 631.164 Estimativa e avaliação agrícola.**
- 631.164.2 Princípios de estimativa.
 - 631.164.23 Investimento de capital de giro.
 - 631.164.24 Avaliação de culturas (incluindo o valor da terra).
 - 631.164.27 Avaliação de subprodutos e resíduos.
 - 631.164.28 Estimativa de dano.
 - 631.164.6 Estimativa e avaliação de empresas agrícolas inteiras.
- 631.165 Estimativa de colheitas e falhas de colheita.**
- 631.17 Tecnologia agrícola.**
- 631.171 Mecanização e automação da Agricultura e Agricultura.**
- 631.172 Uso de máquinas motorizadas.**
- 631.173 Organização de equipamentos e serviços especiais.**
- 631.173.2 Estações de tratores.
 - 631.173.4 Oficinas técnicas e de reparo.
 - 631.173.6 Serviços de uso conjunto de máquinas.
- 631.2 Prédios, estruturas e instalações agrícolas. Prédios para animais, produtos, maquinaria e equipamento.**
- 631.21 Casas de fazenda, sedes de estabelecimentos agrícolas, Incluse Conjuntos de prédios da fazenda.Eiras.Construções fora da sede. Construções agrícolas para múltiplas finalidades.**
- 631.22 Edifícios e estruturas para animais. Salas e acessórios.**
- 631.222 Estábulos. Edifícios similares para animais de casco.**
- 631.223 Galpões. Edifícios similares para animais de patas.**
- 631.224 Canis. Outros edifícios para animais de rapina.**
- 631.225 Casas e estruturas para coelhos e roedores.**
- 631.227 Casas e estruturas para aves e outros pássaros. Estruturas aviárias.**
- 631.229 Casas e estruturas para outros animais.**
- 631.23 Estruturas de criação e cultivo de plantas.**
- 631.234 Estufas.**
- 631.236 Estruturas de viveiro. Instalações para amadurecimento forçado.**
- 631.24 Edifícios para armazenamento, manuseio e processamento de produtos agrícolas.**
- 631.248 Poços e outras lojas para esterco e resíduos.**
- 631.25 Galpões e outros edifícios para máquinas e equipamentos agrícolas.**
- 631.27 Divisões de fronteira de terras agrícolas. Cercas.**
- 631.271 Sebes. Cercas vivas.**
- 631.272 Muros.**
- 631.273 Cercas. Grades. Porteiras, Roletas, Torniquetes, cercas elétricas.**
- 631.274 Trincheiras.Valas. Passagem sobre valas, inclusive mata-burro.**
- 631.3 Máquinas, implementos e equipamentos agrícolas.**
- 631.3.02 Componentes de máquinas e implementos agrícolas. Ferramentas. Partes da máquina. Acessórios.
 - 631.3.05 Máquinas e implementos com dispositivos para medição, calibrar, colher amostras, etc.
 - 631.3.06 Combinações de máquinas e implementos. Conjuntos de máquinas, etc. para uso em Sequência.
 - 631.3.07 Máquinas e implementos agrícolas de acordo com o tipo de acionamento e transmissão de energia.
 - 631.3.071 Implementos puxados por animais (por exemplo, puxados por cavalos, puxados por bois). Implementos manuais.
 - 631.3.072 Implementos rebocados para tratores de três rodas, quatro rodas e de esteira.
 - 631.3.074 Implementos rebocados para tratores de duas rodas.
 - 631.3.075 Equipamento montado em veículo ou embarcação.
 - 631.3.076 Implementos com força motriz própria. Máquinas e acessórios de autopropulsão,exceto tratores. Dispositivos motorizados.
 - 631.3.077 Implementos de tração a cabo (acionamento por corda ou corrente, acionado por guincho,etc.).
- 631.31 Máquinas e implementos para cultivo do solo. Máquinas e implementos auxiliares**
- 631.311 Máquinas e implementos gerais.**
- 631.311.3 Escavadoras de buracos de postes. Furadoras de buracos de plantas.
 - 631.311.5 Escavadoras de valas e fossos. Equipamento.
 - 631.311.7 Máquinas e implementos para arrancar raízes, tocos, touceiras, etc. Destocadores. Máquinas para recolher pedras.
 - 631.311.8 Dispositivos para resolver, misturar, regenerar e esterilizar o solo.

631.312 Arados.

- 631.312.02 Componentes de arados.
- 631.312.021 Corpos de arado. Inclusive Armaduras de arados. Reinas. Dentes. Aivecas. Dispositivos para abrir sulcos. Discos do arado.
- 631.312.024 Peças acessórias dos arados. Inclusive Segas de faca. Segas de disco. Segas raspadoras. Segas que enterram o esterco.
- 631.312.07 Arados segundo a força motriz.
- 631.312.2 Arados puxados por animais com relhas.
- 631.312.3 Arados mecânicos. Arados com ferramentas móveis. Inclusive arados rotativos. Máquinas escavadeiras.
- 631.312.4 Arado de trator com relhas.
- 631.312.5 Arados para subsolagem. Subsoladores. Arados para cavar em profundidade.
- 631.312.6 Arados especiais.
- 631.312.62 Arados modificados para resolver a terra, arrotear batatas, arrancar tubérculos, traçar sulcos, sachar.
- 631.312.63 Arados com relhas para abrir canais de irrigação e drenagem. Implementos para abrir canais. Deirrigação e drenagem.
- 631.312.68 Arados. Arados para limpar o terreno.
- 631.312.69 Arados mondedores ou desbastadores. Arados desmatadores. Arados destocadores.
- 631.312.8 Arados de disco.

631.313 Grades.**631.314 Rolos para tratamento do solo. Inclusive Compressores de torrões de terra. Escovadores.****631.315 Ferramentas manuais para tratamento do solo. Ferramentas para arrancar ervas daninhas.****Ferramentas de jardim.**

- 631.315.1 Enxadas. Enxadas estreitas e curtas.
- 631.315.2 Enxadas de puxar. Enxadações manuais. Grades manuais rotativas. Ferramentas de abrir Sulcos. Ganchos de cavar.
- 631.315.4 Ferramentas de limpeza de ervas daninhas. Sachos.
- 631.315.6 Ancinhos de solo. Ancinhos de jardim. Rastelos.
- 631.315.7 Pás. Espátulas de jardim. Forcados para escavar.

631.316 Capinadeiras. Escavadoras.

- 631.316.2 Capinadeiras (p. ex., capinadeiras rotativas).
- 631.316.4 Enxadas para máquinas (para trabalhar em fileiras e entre fileiras). Máquinas e Implementos paradesbastar e isolar plantas.
- 631.316.5 Aeradores de pastos. Cortadores de turfa.
- 631.316.6 Máquinas para arrancar ou cortar ervas daninhas. Máquinas mondadoras.

631.317 Trituradoras.**631.319 Máquinas e implementos para cultivo do solo para plantio e cuidado das plantas.**

- 631.319.2 Máquinas e implementos para fazer canais, regos, sulcos (para plantar, semear, esterocar). Implementos para abrir sulcos.
- 631.319.3 Máquinas e implementos para cavar e cobrir buracos de semeadura. Implementos para sachar (para plantio e cuidado das semeaduras em fileiras).
- 631.319.4 Trituradoras.

631.33 Máquinas e implementos para semear e plantar. Adubadeiras.

- 631.33.02 Componentes. Detalhes.
- 631.33.021 Recipientes para sementes e fertilizantes. Inclusive Caixas de sementes para regos de sementeiras. Tremonhas para distribuidores de fertilizantes.
- 631.33.022 Mecanismos alimentadores e distribuidores de sementes.
- 631.33.024 Dispositivos para introduzir as sementes no solo. Inclusive segas. Tubos de semente. Rodas de prensagem.

631.331 Máquinas para semear. Semeadeiras.

- 631.331.1 Lançadoras de sementes.
- 631.331.4 Rolos de anel para fazer sulcos.
- 631.331.5 Semeadeiras mecânicas. Furadeiras para semente. Inclusive Máquinas para semear em faixas. Máquinas para furar em filas. Máquinas para semear em aglomerados e grupos.
- 631.331.92 Equipamento para preparo da semente.
- 631.331.93 Máquinas para produzir fitas de semente.

631.332 Máquinas e equipamentos para plantar. Plantadeiras. Dispositivos para o preparo de plantinhas e mudas. Máquinas para fazer blocos de solo.

- 631.332.5 Máquinas para corte de mudas, plantar em vasos, etc.
- 631.332.7 Plantadeiras de batatas.
- 631.332.8 Máquinas e equipamento para plantios especiais.
- 631.332.81 Máquinas para plantar em sulcos.
- 631.332.87 Bastões plantadores.
- 631.333 Distribuidores de fertilizantes. Espalhadores de adubo. Máquinas e equipamentos para preparação e distribuição de fertilizantes.
- 631.333.4 Equipamento para espalhar esterco líquido. Tanques de esterco, bombas, borrifadores

de esterco. Tanques de lodo de esterco.

631.333.5 Máquinas para espalhar fertilizantes em pó e em grãos.

631.333.6 Espalhadores de adubo em terrenos agrícolas. Distribuidores. Máquinas para espalhar esterco animal, composto e turfa.

631.333.7 Distribuidores de sementes e fertilizantes combinados com unidades suplementares alimentadoras. Máquinas de plantar com unidades suplementares alimentadoras. Equipamento para abastecer as raízes das plantas com nutrientes.

631.333.92 Equipamento para picar esterco natural. Poços de armazenamento de esterco.

Instalações de tratamento para o esterco natural e líquido.

631.334 Máquinas e semear e plantar combinadas a distribuidoras de fertilizantes.

Plantadeiras combinadas.

631.335 Equipamento para propagação, multiplicação de plantas por mergulhia, exertia, etc. .

631.336 Equipamento para transplante. Inclusive para transplante de árvores.

631.338 Equipamento para cultivo de plantas em outros lugares, exceto no solo.

Inclusive equipamento para cultivo de plantas sem solo, p. ex., hidropônica.

631.34 Máquinas e equipamentos para cuidado e proteção das plantas

631.342 Ferramentas de corte. Inclusive tesouras. Podões. Facas de bota. Foices para cuidado das plantas. Machados. Machadinhas. Serrotes.

631.343 Equipamento de tratamento de solo para cuidar e proteger de plantas.

Inclusive Dispositivos de aquecimento do solo.

631.344 Dispositivos para cobrir e proteger as plantas.

631.344.1 Câmpanulas.

631.344.2 Armações contra o frio.

631.344.3 Canteiros quentes. Armações aquecidas. Canteiros para apressar o crescimento da planta.

631.344.4 Estufas frias (de vidro).

631.344.5 Estufas aquecidas. Estufas para apressar o crescimento das plantas.

631.344.7 Dispositivos de sombreamento. Esteiras.

631.344.8 Equipamento de aquecimento, ventilação e aguamento para estufas, locais onde se apressa o crescimento das plantas, etc.

631.345 Tutoros para plantas. Inclusive Varas, estacas. Arames e cordas. Dispositivos para orientar e amarrar as plantas.

631.346 Vasos, caixas, etc.

631.347 Dispositivos de regar, aguamento. Equipamento de irrigação.

631.347.1 Tubulações. Mangueiras. Veículos para o transporte de mangueiras.

Carretéis de enrolar mangueira.

631.347.2 Dispositivos de aguamento das raízes ou suprimentos. Regadores, irrigadores.

631.347.3 Equipamento de aspersão. Aspersores. Aspersores em jatos.

631.347.4 Aspersores giratórios.

631.347.8 Dispositivos para agrupamento de raízes ou suprimento de nutrientes.

Dispositivos de alimentação das raízes.

631.348 Máquinas e equipamentos para proteção de plantas. Inclusive Repelentes de pragas, espantelhos, etc. Aramadilhas. Controle químico de pragas e doenças. Fumigadores.

Geradores de fumaça. Aspersores de herbicidas e pesticidas. Equipamento para extinção de ervas daninhas. Lança chamas.

631.35 Equipamento de colheita. Colheitadeiras.

631.351 Ferramentas manuais para colheita. Ferramentas simples de colheita. Ganhos.

Forçados para feno e palha.

631.352 Máquinas ceifadeiras em geral. Segadeiras.

631.352.02 Componentes das ceifadeiras.

631.352.022 Mecanismos de corte.

631.352.027 Acessórios de ceifa e dispositivos de coleta.

631.352.92 Segadeiras para plantas aquáticas. Cortadores de caniço. Segadeiras de caniço. Cortadores de junco.

631.352.93 Cortadores de grama. Aparadores de grama.

631.352.94 Segadeiras para trabalho em valas, regos e encostas.

631.352.95 Máquinas de aparar sebes.

631.353 Máquinas para carregar feno. Máquinas para espalhar feno com dispositivo de amarrar.

Guinchos para feixes de feno.

631.353.2 Ancinhos de feno. Reviradores de feno. Varredores de feno. Ancinhos puxados a cavalo. Ancinhos tipo carretéis. Ancinhos e reviradores combinados.

Reviradores de ancinhos laterais eceifadores.

631.353.3 Máquinas para carregar feno. Máquinas para espalhar feno com dispositivos de amarrar. Guinchos para feixes de feno.

631.354 Máquinas para corte das colheitas de cereal. Enfeixadeiras. Colheitadeiras combinadas.

631.354.02 Componentes da máquina.

- 631.354.022 Mecanismos de corte. Dedos mecânicos. Barras de dedos mecânicos. Facas. Seções de facas.
- 631.354.024 Dispositivos para enfeixar. Enfeixadoras. Enfardadoras. Agulhas. Atadores de nós.
- 631.354.025 Transportadores de fichário. Transportadores de lona. Transportadores de corrente. Carregadores. Coletores de grãos, caixas.
- 631.354.026 Mecanismos de debulha.
- 631.354.028 Colheitadeiras (dedos). Divisores de colheita. Contadores de feixes. Dispositivos de proteção animal.
- 631.354.1 Fichários.
- 631.354.2 Ceifeiras debulhadoras.
- 631.354.3 Catadores. Cabeçalhos. Máquinas para recolher as cabeças de trevo, etc.
- 631.354.4 Ligantes de polia.

631.355 Colhedoras de milho.

631.356 Máquinas e implementos para a colheita de linhas.

- 631.356.2 Colhedores de raízes.
- 631.356.4 Colheitadeiras de batatas.

631.358 Ceifeiras para outras culturas

631.359 Equipamentos e máquinas para medir o rendimento das culturas

631.36 Máquinas e equipamentos para a preparação e armazenamento de culturas ou produtos.

Máquinas e equipamentos para a preparação de alimentos para animais.

631.361 Máquinas e equipamentos para separar ou remover partes de culturas

- 631.361.02 Componentes.
- 631.361.022 Aparelhos para debulhar, tambores, peneiras, etc.
- 631.361.025 Dispositivos de limpeza, agitadores, etc.
- 631.361.026 Aglutinantes de grãos. Ceifadores. Prensas de enfardamento
- 631.361.027 Transportadores de palha.
- 631.361.1 Ferramentas manuais para amarrar roldanas, cortar palha, debulhar
- 631.361.2 Máquinas de debulhar (debulhadores).
- 631.361.3 Joeiradoras.
- 631.361.4 Descascadoras de trevo (sementes). Cascas de trevo. Removedores de cascas, debulhadoras.
- 631.361.5 Máquinas para polir sementes ou grãos, etc.
- 631.361.7 Máquinas e equipamentos para a preparação de frutas.
- 631.361.72 Máquinas para separar e separar carne de grãos. Para remoção de sementes. Para apedrejar frutas. Para remoção de casca. Para descascar nozes, etc.
- 631.361.8 Máquinas e equipamentos para cortar legumes e frutas.
- 631.361.9 Máquinas para remover as coberturas de vegetais e frutas.

631.362 Máquinas e equipamentos para limpeza e triagem de produtos. Recipientes, peneiras, etc.

- 631.362.1 Dispositivos simples. Pás. Peneiras. Tanques.
- 631.362.3 Máquinas de limpeza e triagem.
- 631.362.5 Máquinas para remover a cobertura de sementes (casca, casca, remoção de resíduos, etc.). Máquinas para colheita e processamento de lúpulo.
- 631.362.6 Máquinas para lavar produtos. Máquinas centrífugas. Máquinas para remover resíduos químicos (por exemplo, pesticidas de frutas).

631.363 Máquinas e equipamentos para a preparação de alimentos para animais ou forragens.

- 631.363.1 Ensilagem, equipamento de fermentação. Silos (para forragem)..
- 631.363.2 Máquinas de trituração. Moedores, trituradores, espremedores (para forragens). Máquinas de cortar forragens. Máquinas para prensagem, compressão, cubagem, etc.
- 631.363.5 Equipamento para tratamento térmico de forragens.

631.364 Equipamento para empacotar, treliçar, amarrar, etc. Máquinas de embalagem.

631.365 Equipamento para secagem, arejamento, ventilação de produtos. Secadores.

Ventiladores para culturas.

631.367 Equipamento e arte para cobrir grampos, recipientes, etc., para o armazenamento de produtos agrícolas.

631.37 Máquinas e equipamentos para fornecer energia e tração. Equipamento de transporte agrícola.

631.371 Máquinas e equipamentos para fornecer energia motriz, acionamento.

631.372 Veículos motorizados.

631.373 Veículos não motorizados. Carrinhos de fazenda. trechos de um filme.

631.374 Equipamento de transporte. Equipamento de carga e descarga (exceto dispositivos estruturais)

631.4 Ciência do solo. Pedologia. Pesquisa de solo.

631.41 Propriedades químicas e físico-químicas do solo.

631.412 Propriedades físico-químicas do solo em geral.

631.413 Propriedades eletroquímicas do solo. Troca de íons.

- 631.413.2 Fixação.
- 631.413.3 Acúmulo de sal. Teor de sal.

- 631.413.4 Troca de cátions.
- 631.414 Propriedades coloidais do solo.**
 - 631.414.2 Estado coloidal. Endurecimento. Afrouxamento. Camadas impenetráveis.
 - 631.414.3 Absorção. Adsorção.
- 631.415 Reação do solo. Acidez e alcalinidade do solo. Concentração de íons hidrogênio no solo (valor do pH)**
 - 631.415.2 Acidez. Solos ácidos.
 - 631.415.3 Alcalinidade. Solos alcalinos.
 - 631.415.7 Plantas silvestres como indicadores de reação do solo.
- 631.416 Química inorgânica do solo. Natureza de vários componentes do solo.**
 - 631.416.1 Nitrogênio.
 - 631.416.2 Ácido fosfórico.
 - 631.416.3 Outros não metais e seus compostos.
 - 631.416.4 Potássio.
 - 631.416.5 Sódio.
 - 631.416.7 Cálcio.
 - 631.416.8 Outros metais e seus compostos.
- 631.417 Química orgânica do solo.**
 - 631.417.1 Teor de carbono.
 - 631.417.2 Conteúdo de húmus
 - 631.417.4 Proporção carbono: nitrogênio.
 - 631.417.7 Outros compostos orgânicos.
 - 631.417.8 Compostos organometálicos. Quelatos.
- 631.418 Solução do solo.**
- 631.42 Análise e experimentos de solo.**
- 631.421 Experimentos de campo e vaso. Métodos estatísticos. Lisimetria.**
- 631.422 Análise qualitativa do solo.**
- 631.423 Análise quantitativa do solo.**
 - 631.423.2 Teor de água. Teor de umidade.
 - 631.423.5 Sais dissolvidos.
 - 631.423.6 Composição do ar do solo.
 - 631.423.7 Determinação de bases trocáveis.
- 631.425 Análise física do solo**
- 631.427 Análise biológica do solo**
 - 631.427.1 Análise microbiológica.
 - 631.427.2 Análise microbiológica. Exame de microflora e microfauna.
 - 631.427.3 Métodos de análise usando plantas vivas.
 - 631.427.4 Análise por micróbios. Bioensaio.
- 631.43 Propriedades físicas e mecânicas do solo.**
- 631.431 Propriedades mecânicas do solo.**
 - 631.431.1 Densidade.
 - 631.431.2 Deformação. Cisalhamento.
 - 631.431.3 Calado de arado. Resistência ao arado.
 - 631.431.4 Coesão. Compressibilidade.
 - 631.431.5 Plasticidade. Viscosidade. Ponto pegajoso.
 - 631.431.6 Inchaço e encolhimento. Rachaduras.
 - 631.431.7 Compactação.
- 631.432 Água subterrânea. Hidrologia agrícola.**
 - 631.432.1 Lençol freático. Superfície freática.
 - 631.432.2 Umidade do solo.
 - 631.432.3 Permeabilidade. Lessivage. Lixiviação. Mobilidade dos constituintes do solo.
- 631.433 Atmosfera do solo.**
 - 631.433.1 Aeração do solo
 - 631.433.2 Condições de anaerobiose. Redução de solo.
 - 631.433.3 Respiração do solo.
 - 631.433.5 Composição da atmosfera do solo.
- 631.434 Estrutura do solo.**
- 631.435 Composição granulométrica do solo. Textura do solo.**
- 631.436 Propriedades térmicas do solo.**
 - 631.436.1 Transferência de calor
 - 631.436.2 Calor específico.
 - 631.436.5 Efeito do calor.
 - 631.436.6 Efeito do frio, geada.
- 631.437 Propriedades elétricas e magnéticas do solo.**
- 631.438 Radioatividade do solo.**
- 631.44 Classificação de solos.**
 - 631.44.06 Catenas. Associações de solo.

631.442 Classificação tradicional de solos (principalmente de acordo com textura e composição química).

- 631.442.1 Solo arenoso (solo arenoso).
- 631.442.2 Solo com cal (solo calcário).
- 631.442.3 Solo sedoso. Barro.
- 631.442.4 Solo argiloso (solo argiloso).
- 631.442.5 Solos orgânicos.
- 631.442.6 Solos pedregosos.

631.445 Classificação genética de solos.

- 631.445.1 Solos hidromórficos.
- 631.445.2 Solos podzólicos. Podzóis.
- 631.445.3 Solos marrons. Terras marrons
- 631.445.4 Solos chernozêmicos. Chernozem (terra negra)
- 631.445.5 Solos áridos. Solos do deserto.
- 631.445.6 Solos do Mediterrâneo.
- 631.445.7 Solos tropicais e subtropicais..
- 631.445.8 Rendzinas. Solos castanhos calcários.
- 631.445.9 Outros tipos de solo.

631.45 Fertilidade e toxicidade do solo.

- 631.452 Fertilidade do solo.
- 631.453 Toxicidade para o solo.
- 631.454 Exigência de controle do solo.
- 631.458 Empobrecimento do solo. Exaustão do solo.
- 631.459 Erosão.

631.46 Bacteriologia do solo. Biologia do solo.**631.461 Micróbios zimóticos do solo. Bactérias do solo. Microbiologia do solo.**

- 631.461.1 Bactérias amonizadoras. Bactérias ureolíticas. Proteólise. Amonificação.
- 631.461.2 Bactérias formadoras de nitrito.
- 631.461.3 Bactérias formadoras de nitrato. Nitrificação.
- 631.461.4 Desnitrificação.
- 631.461.5 Bactérias fixadoras de nitrogênio. Fixação de nitrogênio.
- 631.461.6 Decomposição de compostos não nitrogenados.
- 631.461.7 Conversão de substâncias inorgânicas, minerais.

631.462 Esterilização, desinfecção do solo.**631.463 Antibióticos. Antagonismo microbiano.****631.464 Vitaminas no solo.****631.465 Enzimas no solo.****631.466 Outra microflora.****631.467 Microfauna.****631.468 Macrofauna e mesofauna. Animais no solo.****631.47 Levantamento topográfico, classificação e utilização do solo.****631.472 Perfis de solo. Desenvolvimento do perfil do solo.**

- 631.472.5 Ninhada. Horizontes orgânicos.
- 631.472.6 Solo mineral em geral. Horizontes A-C.
- 631.472.7 Camadas do solo.
- 631.472.8 Micropedologia.

631.473 Levantamentos de solo.**631.48 Pedogênese. Formação do solo****631.482 Formação do solo por processos físicos e mecânicos.****631.483 Formação do solo por meios químicos.****631.484 Formação do solo por meios biológicos.****631.485 Solo transportado e deslocado.****631.5 Operações agrícolas.****631.51 Preparação e aradura do solo.**

- 631.51.02 Tempo de trabalho no solo.
- 631.51.021 Preparo do solo.
- 631.51.022 Trabalhos antes da semeadura ou plantio.
- 631.51.023 Trabalho durante o crescimento.

631.511 Preparação do solo sem dispositivos mecânicos.**631.512 Preparação do solo com ferramentas mecânicas.****631.513 Preparação do solo com explosivos.****631.52 Melhoria de linhagens de plantas.****631.522 Biologia da reprodução vegetal.****631.524 Variação de características.****631.526 Grupos e tipos de plantas cultivadas.**

- 631.526.1 Gênero.
- 631.526.2 Espécies e subespécies.

- 631.526.3 Variedade.
- 631.526.4 Linhagem genética.
- 631.527 Melhoria por seleção (seleção natural e artificial). Melhoramento seletivo de plantas.**
 - 631.527.2 Seleção de massa. Métodos rápidos de seleção.
 - 631.527.3 Seleção de linhagem.
 - 631.527.5 Híbridos, cruzamento endógeno e heterose.
 - 631.527.6 Seleção clonal.
 - 631.527.7 Seleção de espécies poliplóides
 - 631.527.8 Métodos de seleção.
 - 631.527.81 Castração.
 - 631.527.82 Polinização.
- 631.528 Melhoria por modificação e mutação (natural e induzida). Genética aplicada**
 - 631.528.1 Mutação de genes e cromossomos. Mutagênese.
 - 631.528.2 Poliploidização.
 - 631.528.6 Produção de mutações. Modificação genética.
- 631.529 Aclimação.**
- 631.53 Propagação vegetal. Propagação vegetativa.**
 - 631.53.01 Semente.
 - 631.53.02 Produção de sementes. Coleta, armazenamento e tratamento de sementes.
 - 631.53.027 Tratamento de sementes.
 - 631.53.027.2 Métodos químicos.
 - 631.53.027.3 Métodos físicos.
 - 631.53.027.32 Usando temperatura.
 - 631.53.027.33 Usando eletricidade.
 - 631.53.027.34 Usando radiação.
 - 631.53.03 Cultivo de plantas jovens, mudas.
 - 631.53.04 Sementeira. Plantio. Transplante.
 - 631.53.041 Semeadura direta, plantio.
 - 631.53.042 Semeadura à mão.
 - 631.53.048 Taxa de semeadura. Quantidade de sementes.
- 631.531 Sementes e esporos.**
- 631.532 Propagação por separação de parte de uma planta em crescimento.**
 - 631.532.1 Usando lâmpadas.
 - 631.532.2 Usando tubérculos. Divisão dos tubérculos em partes com um olho cada
 - 631.532.3 Divisão ou separação de aglomerados.
- 631.533 Destacamento de plantas separadas.**
 - 631.533.1 Ventosas do colarinho ou toco.
 - 631.533.2 Estolhos.
 - 631.533.3 Botões terminais e proliferativos.
- 631.534 Propagação por camadas.**
- 631.535 Propagação por corte. Estacas.**
- 631.538 Renovação de pomares e vinhedos, etc. Rejuvenescimento.**
- 631.54 Cuidados com as plantas.**
- 631.541 Enxertia.**
 - 631.541.1 Escolha do cavalo, do porta-exterco. Compatibilidade.
 - 631.541.3 Métodos de enxertia.
 - 631.541.4 Enxerto de acordo com a posição do enxerto no material.
 - 631.541.5 Enxerto de borbulha.
 - 631.541.6 Enxerto de encostia.
 - 631.541.7 Amarração e proteção do enxerto.
 - 631.541.8 Enxerto interno.
- 631.542 Desbaste. Poda.**
 - 631.542.1 Poda de inverno.
 - 631.542.2 Poda de verão.
 - 631.542.3 Corte, corte de plantas para fins específicos.
 - 631.542.32 Corte para modelar.
- 631.543 Plantio.**
 - 631.543.1 Perguntas gerais.
 - 631.543.2 Densidade de plantio. Espaçamento.
 - 631.543.3 Métodos de plantio.
 - 631.543.5 Remoção de plantas excedentes. Compensação. Desbaste.
 - 631.543.8 Tipos de plantio. Site em geral. Espaçamento (de linhas; de plantas). Plantando para sombra, como abrigo. Cuidado e proteção das plantações.
- 631.544 Proteção e abrigo de plantas. Cultivo coberto.**
 - 631.544.1 Cultivo sob cloches.
 - 631.544.2 Cultivo em viveiro. Cultivo em molduras quentes e frias.
 - 631.544.4 Cultivo de estufa (estufa).

631.544.7 Cultivo sob várias coberturas. Proteção e cobertura do solo.

631.546 Formas naturais e artificiais, formas.

631.547 Fases e controle do crescimento das plantas.

- 631.547.1 Germinação. Mudas.
- 631.547.2 Crescimento.
- 631.547.3 Desenvolvimento completo.
- 631.547.4 Floração. Florescimento.
- 631.547.5 Frutificação. Frutificação.
- 631.547.6 Amadurecimento.
- 631.547.66 Maturação.
- 631.547.67 Amadurecimento artificial.

631.55 Colheita.

- 631.55.03 Métodos especiais de colheita.
- 631.55.032 Colheita antes do vencimento. Colheita verde.
- 631.55.034 Colheita múltipla. Primeira, segunda colheita, etc.
- 631.55.035 Respiga.

631.552 Corte. Ceifa. Gadanha.

631.553 Produção de feno.

631.554 Colheita, colheita de cereais.

631.555 Colheita de milho.

631.558 Colheita de outras culturas.

631.559 Rendimento da colheita.

631.56 Preparo de produto agrícola. Tratamento após a colheita.

631.561 Debulha. Separação de frutas da casca. Casco, descasque, corte, etc.

631.562 Limpeza e triagem de produtos. Peneirar. Descamação. Lavagem, etc.

631.563 Preparação e armazenamento de produtos.

- 631.563.2 Secagem.
- 631.563.3 Redução de tamanho.
- 631.563.4 Mistura.
- 631.563.6 Fermentação.
- 631.563.8 Preparando e preservando para armazenamento.
- 631.563.9 Armazenamento.

631.564 Embalagem.

631.565 Transporte.

631.57 Produtos, subprodutos e produtos semi-acabados da Agricultura

631.571 Madeira. Casca de árvores.

631.572 Hastes. Fibras. Palha.

631.573 Raízes. Tubérculos. Bulbos.

631.574 Folhas. Folhagem.

631.575 Flores. Flores.

631.576 Frutas. Partes do fruto.

- 631.576.2 Frutos e nozes inteiras.
- 631.576.3 Sementes como produtos. Cereais. Nozes sem casca.
- 631.576.4 Cascas de frutas, casca, casca. Casca. Fibras. Farelo.

631.577 Sucos, sucos, gomas, essências etc.

631.58 Sistemas específicos de cultura.

631.581 Pousio.

631.582 Rotação de culturas.

- 631.582.1 Monocultura.
- 631.582.2 Cultivo itinerante. Rotação da terra.
- 631.582.3 Rotações de três estágios.
- 631.582.9 Outros sistemas.

631.583 Cultivo intensivo.

631.584 Cultura intercalada. Cultura de rodízio. Cultura exaustiva.

- 631.584.2 Sistema de cultura de rodízio.
- 631.584.4 Sistema de cobertura vegetal.
- 631.584.5 Cultura mista. Cultura intercalada.
- 631.584.9 Cultura exaustiva.

631.585 Sistema pastoral. Transumância

631.586 Agricultura a seco. Uso de níveis mais profundos do solo.

631.587 Cultivo em terras irrigadas.

631.588 Cultivo usando fenômenos especiais.

631.589 Cultivo em meios artificiais.

- 631.589.2 Hidropônica.
- 631.589.3 Culturas de areia. Outras mídias inertes.
- 631.589.5 Em meios plásticos.

631.6 Engenharia rural.

631.6.02 Conservação de solo e água. Controle de erosão.

631.6.03 Água: composição e propriedades.

631.61 Limpeza e utilização de tipos especiais de solos.**631.611 Recuperação de terras marginais em geral, pastagens, charnecas, etc.****631.612 Recuperação de terras pedregosas e rochosas. Recuperação de dunas.****631.613 Recuperação de terrenos inclinados.****631.614 Recuperação de florestas e matas virgens.****631.615 Recuperação de pântanos e charnecas.****631.616 Recuperação de terras abaixo do nível do mar. Formação e pôlderes.****631.617 Recuperação de desertos e terras áridas.****631.618 Recuperação de terras industriais e mineradas.****631.62 Drenagem.****631.67 Irrigação.****631.671 Necessidade de água.****631.672 Abastecimento de água.****631.674 Métodos de irrigação.**

631.674.1 Irrigação de superfície.

631.674.2 Irrigação por inundação.

631.674.3 Desvio de fluxo.

631.674.4 Irrigação subterrânea.

631.674.5 Irrigação por aspersão.

631.674.6 Irrigação por gotejamento. Irrigação por gotejamento..

631.675 Taxa e tempo de irrigação. Parâmetros de irrigação.**631.8 Fertilizantes, adubos e enriquecimento do solo. Promoção do crescimento.****Estimulantes vegetais.****631.8.022.3 Aditivos.****631.81 Generalidades. Nutrientes vegetais. Produção e uso de fertilizantes, adubos.**

631.81.031 Absorção, absorção de nutrientes

631.81.033 Conversão e distribuição na planta

631.81.036 Processos biodinâmicos

631.81.095.337 Requisito para certos elementos. Micronutrientes essenciais.

Vestígios

631.81.095.338 Falta de certos elementos. Deficiência mineral

631.811 Nutrição vegetal

631.811.98 Estimulantes vegetais (para promoção artificial do crescimento sem adubação).

Reguladores de crescimento de plantas, hormônios.

631.812 Produção, preparação, armazenamento, etc. de fertilizantes. Constituição e forma de fertilizantes.**631.813 Composição, reação (valor do pH), toxicidade e reações recíprocas de fertilizantes.****631.814 Princípios de perda de fertilizantes. Alterações no valor de gestão.****631.815 Aspectos econômicos.****631.816 Aplicação e distribuição de fertilizantes.**

631.816.1 Doses. Taxas de aplicação.

631.816.2 Tempo de aplicação (por exemplo, antes, depois e semeadura).

631.816.3 Método de aplicação.

631.82 Alterações minerais. Fertilizantes artificiais em geral.**631.821 Cal. Giz. Gesso.****631.822 Marga.****631.823 Polieletrólitos. Condicionadores de solo.****631.824 Magnésio.****631.825 Sulfato de ferro.****631.826 Lodo de inundações. Aluvião.****631.828 Outras alterações****631.829 Gases de combustão. Dióxido de carbono****631.83 Fertilizantes potássicos.****631.831 Cinza de madeira. Plante cinzas. Cinza turfa.****631.832 Cloretos.****631.833 Sulfatos.****631.834 Sais de potássio brutos.****631.835 Sylvinita.****631.836 Silicatos..****631.837 Carbonatos****631.84 Fertilizantes nitrogenados.****631.841 Fertilizantes de amônio. Amidas.****631.842 Fertilizantes nitrogenados. Salitros (nitratos de potássio e sódio), etc.****631.847 Fertilizantes nitrogenados vegetais. Inoculação bacteriana e micorrízica do solo.**

- 631.85 Fertilizantes fosfatados.
- 631.851 Fosfatos minerais. Fosfatos brutos.
- 631.852 Farinha de osso e produtos similares.
- 631.853 Escória básica. Escória de Thomas.
- 631.854 Guano (fezes de pássaros e morcegos). Guano de peixe.
- 631.855 Superfosfato.
- 631.857 Eletrofosfatos. Fosfatos condensados.
- 631.858 Silicofosfatos. Fosfatos da Renânia.
- 631.859 Outros fertilizantes fosfatados.
- 631.86 Adubos orgânicos em geral. Fertilizantes de origem biológica.
- 631.861 Adubos de origem animal em geral.
- 631.862 Esterco de pastos. Esterco animal. Esterco de curral (inclusive estéreo líquido).
- 631.863 Excremento humano, esgoto.
- 631.867 Outros fertilizantes de origem animal.
- 631.87 Outros adubos orgânicos.
- 631.871 Estrumes de origem vegetal em geral.
- 631.872 Detritos vegetais.
- 631.873 Matérias vegetais de origem aquática.
- 631.874 Adubação verde. Cultivos que melhoram o solo.
- 631.875 Adubação com plantas cortadas.
- 631.877 Adubação com bolos de sementes oleaginosas.
- 631.878 Turfa, linhite e carvão como fertilizantes
- 631.879 Outros resíduos como fertilizantes
 - 631.879.1 Lixo urbano, lixo como fertilizante.
 - 631.879.2 Esgoto como fertilizante.
 - 631.879.3 Resíduos industriais como fertilizantes.
 - 631.879.4 Compostagem.
- 631.89 Adubos compostos. Adubos mistos.
- 631.893 Misturas para fertilização puramente inorgânicas.
- 631.894 Misturas de adubação puramente orgânicas (adubos mistos).
- 631.895 Misturas de adubos orgânicos e inorgânicos.
- 631.92 Efeitos geofísicos do cultivo. Influência da Agricultura no clima, solo, cursos de água, erosão, deslizamentos de terra, etc.
- 631.95 Ecologia agrária.
- 631.96 Tipos de plantação. Avaliação do local.
- 631.961 Plantações em relação ao local.
- 631.962 Tipos de Formações vegetais e de matas.
 - 631.962.2 Densidade.
 - 631.962.3 Formações vegetais puras.
 - 631.962.4 Formações vegetais mistas.
 - 631.962.5 Formações vegetais de mesma idade.
 - 631.962.6 Formações vegetais de idadesdesiguais.
- 631.963 Origem e desenvolvimento de formações vegetais. Migração de espécies vegetais, etc. Invasão e naturalização de espécies vegetais.
 - 631.963.2 Formação de vegetais naturais em geral.
 - 631.963.3 Formação de vegetais artificiais em geral.

FONTE: Classificação Decimal Universal (1997).

ANEXO B - Conceitos utilizados na Amostra do AgroTermos

1. INTENSIFICAÇÃO AGROPECUÁRIA

Morfologia: Substantivo Feminino (s.f.)

Definição Inicial: Processo que visa o aumento da produtividade agropecuária, através de variadas práticas como, por exemplo, diminuição dos períodos de pousio, transição do uso das terras e fertilização.

Definição Especialista: Processo de aumento da produção ou da produtividade agropecuária numa mesma área, observado em microescalas, espaço-temporais e decorrente do uso intensificado de insumos, como trabalho, capital e tecnologia.

Definição Modelador: Processo agropecuário da produção ou da produtividade agropecuária em uma mesma área, decorrente do uso intensificado de insumos, como trabalho, capital e tecnologia, sem deixar de lado aspectos como a segurança alimentar, a sustentabilidade e a biodiversidade, sempre observado em relação às microescalas espaço-temporais e socioeconômicas.

2. AGRICULTURA EXTENSIVA

Morfologia: Substantivo Feminino (s.f.)

Definição Inicial: Agricultura caracterizada como tradicional na qual não há o uso de tecnologias, preparação do solo e seleção das sementes. Devido ao uso de técnicas rudimentares, possui baixo índice de produtividade.

Definição Especialista: Sistema de produção agrícola onde praticamente não se utilizam insumos tecnológicos e, portanto, altamente dependente da incorporação de áreas naturais para exploração agrícola.

Definição Modelador: Sistema de exploração agrícola onde praticamente não se utilizam insumos tecnológicos, altamente dependente da incorporação de áreas naturais para exploração agrícola e, em geral, com baixo investimento e baixa produtividade; favorece as atividades pastoris que têm nas pastagens sua principal fonte alimentar, e de outras criações extensivas (peixes, por exemplo) que utilizam instalações já existentes, sem preocupação com a técnica utilizada, e o aumento da extensificação (expansão da área cultivada).

3. AGRICULTURA INTENSIVA

Morfologia: Substantivo Feminino (s.f.)

Definição Inicial: Tipo de Agricultura, em larga escala, que pode causar grandes impactos negativos aos ecossistemas e capaz de aumentar a erosão, o despejamento de sedimentos e o escoamento de nutrientes e químicos agrícolas para os reservatórios de água subterrânea, nascentes e rios, sendo um dos responsáveis pelo excesso de fósforo e nitrogênio nas correntes de água. A expansão da Agricultura intensiva, por aumentar a produtividade no mercado, em espaços reduzidos, quando associado ao trabalho mecanizado, pode colocar os preços agrícolas de outros tipos de produções em declínio. Nas terras em que ocorre a Agricultura intensiva existe um aumento no cultivo e, por conta disso, um drástico uso de fertilizantes (nitratos) e pesticidas químicos que reduzem fertilidade e a capacidade produtiva dos solos, um grande uso de insumos e combustíveis, e uma alta demanda pela água que também pode colocar suas fontes em risco.

Definição Especialista: Refere-se a sistemas agrícolas que envolvem intensificação de trabalho e/ou de insumos, o que resulta em maior produção ou produtividade agrícolas, podendo agregar, ainda, referências a atividades realizadas em espaços reduzidos para cultivo de plantas ou criação de animais.

Definição Final: Sistema agrícola que envolve intensificação de trabalho e/ou de insumos, o que resulta em maior produção ou produtividade agrícolas, podendo agregar, ainda, referências a atividades realizadas em espaços reduzidos para cultivo de plantas ou criação de animais.

Definição Modelador: Sistema de exploração agrícola que envolve o uso intensivo da terra, com a intensificação sustentável do trabalho e/ou de insumos de cultivo e pecuária, o que resulta em maior produção ou produtividade agrícolas, podendo agregar, ainda, referências a atividades realizadas em espaços reduzidos e com tecnologias de ponta para o cultivo de plantas ou criação de animais.

MATERIAL E MÉTODOS

4. CULTURA

Morfologia: Substantivo Feminino (s.f.)

Definição Inicial: Cultivo de diferentes tipos de plantações como cana de açúcar, café, soja, entre outras.

Definição Especialista: Refere-se a qualquer planta, cultivada pelo homem ou não, cujos produtos foram, são ou serão colhidos em quaisquer de seus estágios de crescimento. Plantas que não foram

cultivadas, mas cujos produtos são colhidos não são consideradas culturas; o mesmo se aplica para plantas que foram cultivadas, mas que nunca terão seus produtos colhidos.

Definição Modelador: Plantas cultivadas pelo homem ou não, cujos produtos foram, são ou serão colhidos em quaisquer de seus estágios de crescimento; é influenciada pelo calendário agrícola e pelo ciclo do tipo de cultura; também denominada “produção vegetal” e “planta cultivada”.

5. SENSORIAMENTO REMOTO

Morfologia: Substantivo Masculino (s.m.)

Definição Inicial: Tecnologia de obtenção de imagens e dados da superfície terrestre através da captação e registro da energia refletida/emitida da superfície estudada. Pode ajudar a mapear variáveis climáticas, meteorológicas, mudanças de vegetação e até a qualidade da água de grandes reservatórios. Essas mudanças podem ser monitoradas em diferentes espaços e escalas de tempo. A integração do sensoriamento remoto com os Sistemas de Informação Geográfica (SIG) é uma importante ferramenta para detectar a magnitude das mudanças geográficas e dos fenômenos agrícolas. Desta forma, esses resultados podem, em futuras aplicações, colaborar para perspectivas ambientais e Agriculturas sustentáveis.

Definição Especialista: Tecnologia de obtenção de imagens e dados da superfície terrestre através da captação e registro da energia refletida/emitida da superfície estudada.

Definição Final: Tecnologia de obtenção de imagens e dados da superfície terrestre, através da captação e registro da energia refletida/emitida da superfície estudada.

Definição Modelador: Tecnologia composta por instrumentos para a obtenção de imagens e dados da superfície terrestre, por meio da captação e registro da energia refletida/emitida da superfície estudada; parte de disciplinas tais como Agrimensura e Geometria Geoespacial; usa métodos, técnicas e ferramentas tais como mapa, radar, processamento de imagens, satélites de observação da terra, sensores, GPS, entre outros.

6. SISTEMAS DE INFORMAÇÃO GEOGRÁFICA

Morfologia: Substantivo Masculino (s.m.)

Definição Inicial: Ferramentas computacionais que coletam, armazenam, transformam e exibem aspectos especiais a partir de uma abordagem de natureza cartográfica, altamente valorizada para documentação de processos de formação de paisagens agrícolas.

Definição Especialista: Ferramentas computacionais que reúnem, armazenam, organizam, analisam e exibem dados de fontes variadas a partir de uma abordagem de natureza cartográfica, altamente valorizada para documentação de processos que se manifestam na superfície terrestre. Para esse processamento, a geometria e os atributos dos dados devem estar georreferenciados, isto é, localizados na superfície terrestre e representados numa projeção cartográfica. A instância prática de um Sistema de Informação Geográfica, ou a sigla SIG, combina o programa computacional com equipamentos, dados, usuários e procedimentos, auxiliando planejamento e tomada de decisões.

Definição Modelador: Sistema de informação, de suporte e ferramenta de tecnologia geoespacial que reúne, armazena, organiza, analisa e exibe dados de fontes variadas a partir de uma abordagem de natureza cartográfica, altamente valorizada para a documentação de processos que se manifestam na superfície terrestre; usa GPS e dados espaciais; sigla SIG.

AMBIENTE

7. MEIO AMBIENTE

Morfologia: Substantivo Masculino (s.m.)

Definição Inicial: Conjunto de condições que envolvem e sustentam os seres vivos na biosfera, como um todo ou em parte desta, abrangendo elementos do clima, solo, água e de organismos.

Definição Especialista: Dimensão real do planeta que reúne elementos e processos físicos, químicos, biológicos e antrópicos e toda gama de manifestações decorrentes da combinação ou inter-relação desses componentes.

Definição Modelador: Entidade do ambiente que reúne um conjunto de elementos e processos físicos, químicos, biológicos e antrópicos e toda gama de manifestações decorrentes da combinação ou inter-relação desses componentes; abriga e rege a vida em todas as suas formas, afetando, por exemplo, a população e a economia e sendo afetado por elementos tais como biodegradação, catástrofes naturais, queimadas e políticas ambientais.

8. AMBIENTE FÍSICO

Morfologia: Substantivo Masculino (s.m.)

Definição Inicial: Espaço composto por elementos naturais como rios, lagos, oceanos, vegetação, formações rochosas, clima, etc.

Definição Especialista: Conjunto de elementos e componentes naturais do meio ambiente que envolvem unicamente os fenômenos e processos físicos e químicos, tais como rochas, solo, chuva, nuvens, corpos d'água, atmosfera, luz, etc.

Definição Modelador: Elemento do meio ambiente que engloba e envolve unicamente os fenômenos e processos naturais, físicos e químicos, tais como rochas, solo, chuva, nuvens, corpos d'água, atmosfera e luz; é afetado por fatores climáticos e sedimentação, e pode afetar a economia.

9. SOLO

Morfologia: Substantivo Masculino (s.m.)

Definição Inicial: Área que constitui a superfície sobre a qual são cultivados pastos, plantações ou estabelecidas edificações.

Definição Especialista: Camada que recobre as superfícies não inundadas do planeta e que se forma naturalmente pelo desgaste das rochas que a originam, envolvendo processos naturais de longo prazo tais como intemperização e interação com a biota que se estabelece localmente nessa camada.

Definição Modelador: Cobertura da terra e elemento do ambiente que recobre as superfícies não inundadas do planeta e que se forma naturalmente pelo desgaste das rochas que a originam, envolvendo vários componentes e processos naturais de longo prazo tais como intemperização e interação com a biota (seres vivos de um dado ambiente) que se estabelece localmente nessa superfície, assim como processos não naturais, de melhoramento e manejo de solo.

AGRONOMIA

10. MANEJO DA CULTURA

Morfologia: Substantivo Masculino (s.m.)

Definição Inicial: Prática que envolve a adoção de estratégias de controle para a promoção de um ambiente adequado para o desenvolvimento de uma cultura, de uma forma que seja impedido o crescimento de pragas e doenças. O desenvolvimento tecnológico nesta prática é capaz de aumentar a produtividade de uma área. A rotação de culturas é uma de suas principais alternativas de controle do manejo de cultura, visto que desta forma, as propriedades inerentes ao solo não podem ser alteradas.

Definição Especialista: Conjunto de procedimentos aplicados nos cultivos agrícolas que visam promover as melhores condições de plantio, crescimento das plantas e colheita, considerando os aspectos edáficos, hídricos e fitossanitários que envolvem as culturas. Esses procedimentos referem-se à preparação do solo, fertilização, sistemas de cultivo, métodos de controle de pragas e doenças, entre outros aspectos, sendo que várias deles podem ser mecanizados ou automatizados.

Definição Modelador: Manejo de plantas que envolve um conjunto de procedimentos tais como a preparação do solo, sistemas de cultivo, tratamentos de pré-colheita e métodos de controle de pragas e doenças, que podem ser mecanizados ou automatizados, que visam promover as melhores condições de plantio, crescimento das plantas e colheita, considerando os aspectos edáficos, hídricos e fitossanitários que envolvem as culturas; influencia o ambiente físico.

11. MANEJO DO SOLO

Morfologia: Substantivo Masculino (s.m.)

Definição Inicial: Conjunto de práticas que se destinam ao tratamento do solo a fim de que a produtividade seja maior, como por exemplo, o plantio direto.

Definição Especialista: Conjunto de práticas que se destinam à preparação, adequação ou tratamento do solo para garantia de sua sustentabilidade enquanto recurso natural, permitindo seu uso adequado a curto, médio e longo prazos para a agropecuária ou outros fins.

Definição Modelador: Prática de manejo que envolve métodos e procedimentos, mecanizados ou automatizados, empregados na preparação, adequação, recuperação e conservação do solo para garantia de sua sustentabilidade enquanto recurso natural; permite o uso adequado do solo a curto, médio e longo prazos.

12. POUISIO

Morfologia: Substantivo Masculino (s.m.)

Definição Inicial: Período de descanso em que a terra agrícola deixa de ser cultivada, visando recuperar e aumentar a fertilidade do solo. Nessa fase, que pode variar de um a vários anos, recupera-se a diversidade no solo e, dependendo do tempo do pouisio, restabelece-se também a diversidade de plantas e de animais nativos.

Definição Especialista: Período de descanso em que a terra agrícola deixa de ser cultivada.

Definição Modelador: Prática de manejo e sistema de cultivo que prevê a interrupção de atividades ou usos agrícolas, pecuários ou silviculturais do solo para descanso, por diferentes motivos (preparação, regeneração, fertilidade do solo), e por distintos períodos de tempo.

13. PERÍODO DE POUSIO

Morfologia: Substantivo Masculino (s.m.)

Definição Inicial: Tempo em que a área agrícola repousa a fim de restaurar a fertilidade do solo e a vegetação nativa. Não há um limite de tempo fixo, a duração do período pode variar conforme a intenção do agricultor e deve ocorrer entre os períodos de uso da área.

Definição Especialista: Tempo em que uma área agrícola permanece sem cultivos.

Definição Modelador: Pousio (ou interrupção) do uso e das atividades agrárias no solo, por determinados e diferentes períodos de tempo, para repouso da terra, sem deixar de lado elementos tais como a sustentabilidade e a economia.

14. PESTICIDA

Morfologia: Substantivo Masculino (s.m.)

Definição Inicial: Substância química usada na Agricultura para a prevenção, eliminação e controle de pragas.

Definição Especialista: Substância química usada na Agricultura para a prevenção, eliminação e controle de pragas e doenças das plantas e dos animais.

Definição Modelador: Substância química usada nas práticas de manejo para a prevenção, eliminação e controle de pragas e doenças das plantas e dos animais; o uso envolve métodos e procedimentos que pode causar poluição atmosférica e do solo, sendo necessário levar em conta a segurança alimentar.

15. CICLO DA CULTURA

Morfologia: Substantivo Masculino (s.m.)

Definição Inicial: Ciclo que compreende as fases do desenvolvimento de uma cultura. Por exemplo, no cultivo do arroz há três grandes fases: a fase vegetativa, a fase reprodutiva e a fase de maturação.

Definição Especialista: Ciclo que compreende as fases do desenvolvimento de uma cultura, do plantio à colheita.

Definição Modelador: Parte da Agricultura que compreende as fases do desenvolvimento de uma cultura, do plantio à colheita, cujas características afetam as práticas de manejo.

16. CULTURA ANUAL

Morfologia: Substantivo Feminino (s.f.)

Definição Inicial: Cultura que é caracterizada pela replantação anualmente, pois encerra seu ciclo produtivo em um ano ou menos. São exemplos desse tipo de cultura: a soja, o feijão, o trigo e a cevada.

Definição Especialista: Cultura que é caracterizada pela replantação anualmente, pois encerra seu ciclo produtivo em um ano ou menos. São exemplos desse tipo de cultura: a soja, o feijão, o trigo e a cevada.

Definição Modelador: Cultura cíclica e de sistema de cultivo anual, que tem uma única produção a cada período de um ano ou menos, e, após a colheita, é preciso novo plantio; afeta o período de pousio do solo.

17. PECUÁRIA

Morfologia: Substantivo Feminino (s.f.)

Definição Inicial: Atividade relacionada à criação de gado com o objetivo de fornecer principalmente carne, leite, couro e lã. Pode ser dividida em: pecuária de corte, que destina a criação para o abate visando à obtenção de carne; e em pecuária leiteira que conduz a criação para a produção de leite. A pecuária de corte pode ser subdividida em intensiva, na qual o gado é criado em pequenos espaços e alimentado com ração específica produzindo carne macia e de boa qualidade; e em extensiva, na qual o gado é criado solto e se alimenta de grama o que produz uma carne dura, já que o gado desenvolve uma musculatura rígida.

Definição Especialista: N/D

Definição Modelador: Parte da Agricultura que envolve a criação de diversos tipos de animais domesticáveis, e inclui atividades de produção de carne, leite (e seus derivados), e de produtos tais como couro e lã, cujos resultados impactam a economia.

18. ADUBO VERDE

Morfologia: Substantivo Masculino (s.m.)

Definição Inicial: Adubo composto por plantas como leguminosas, gramíneas, crucíferas, entre outras e que tem como objetivo proteger o solo. Após seu plantio em rotação, sucessão ou em consórcio com outras culturas, elas são fragmentadas e deixadas como cobertura até serem decompostas.

Definição Especialista: Adubo composto por plantas como leguminosas, gramíneas, crucíferas, entre outras e que tem como objetivo proteger e enriquecer o solo. Após seu plantio em rotação, sucessão ou em consórcio com outras culturas, elas são fragmentadas e deixadas como cobertura até serem decompostas.

Definição Modelador: Adubo e fertilizante orgânico que é resultado da fragmentação e da decomposição de plantas tais como leguminosas, gramíneas, crucíferas, que foram deixadas no solo como cobertura após seu plantio em rotação, sucessão ou em consórcio com outras culturas; tem como objetivo proteger e enriquecer o solo.

19. PRODUTO AGROPECUÁRIO

Morfologia: Substantivo Masculino (s.m.)

Definição Inicial: Produto proveniente da Agricultura e da pecuária, como por exemplo: flores, plantas, carne, leite e seus derivados.

Definição Especialista: Produto proveniente da Agricultura e da pecuária, como por exemplo: flores, plantas, carne, leite e seus derivados.

Definição Modelador: Produto em estado natural ou processado, proveniente da Agricultura e da pecuária, cuja produtividade impacta a economia.

20. BIOMASSA

Morfologia: Substantivo Feminino (s.f.)

Definição Inicial: Matéria orgânica proveniente de organismos vivos de um determinado local e é quantificado segundo seu peso, volume ou área ocupada. Pode ser utilizada para a fertilização dos solos ou para a produção de energia.

Definição Especialista: Matéria orgânica proveniente de organismos vivos de um determinado local, quantificada segundo seu peso, volume ou área ocupada. Pode ser utilizada para a fertilização dos solos ou para a produção de energia.

Definição Modelador: Substância ou recurso biológico e orgânico renovável, proveniente de organismos vivos de um determinado local, quantificada segundo seu peso, massa, volume ou área ocupada, e utilizada para a fertilização dos solos ou para a produção de energia.

TERRITÓRIO E PAISAGEM

21. POSSE DA TERRA

Morfologia: Substantivo Feminino (s.f.)

Definição Inicial: Relação de propriedade entre o titular e a terra, ou seja, o proprietário detém os direitos sobre a área que lhe pertence, podendo atribuir a ela o uso que desejar.

Definição Especialista: Relação de propriedade entre o proprietário e a terra que lhe pertence, detendo os direitos sobre ela e podendo atribuir a ela o uso que desejar.

Definição Modelador: Propriedade de extensão de terra cujos direitos foram adquiridos a partir de diferentes tipos de regimes de posse, e envolve fatores tais como arrendamento, desapropriação, usucapião e reforma agrária.

22. ESCASSEZ DE TERRA

Morfologia: Substantivo Feminino (s.f.)

Definição Inicial: Redução da quantidade de terra em uma determinada área por motivos como desflorestamento, degradação de recursos naturais, mudança climática, expansão populacional, urbanização, industrialização, entre outros. Uma das consequências da escassez de terra é a diminuição de áreas agrícolas destinadas à cultura de gênero alimentício, a qual interfere na segurança alimentar da população.

Definição Especialista: Redução da quantidade de terra disponível para posse e, portanto, para qualquer tipo de uso, incluindo para produção agrícola.

Definição Modelador: Território reduzido em sua quantidade, tornando-a indisponível para a posse ou qualquer tipo de uso agropecuário, cujo fenômeno envolve fatores tais como crescimento populacional, zona urbana, políticas públicas, reforma agrária e cultura de alimentos.

23. COBERTURA DA TERRA

Morfologia: Substantivo Feminino (s.f.)

Definição Inicial: Cobertura geológica, física e biológica do solo, como pastagens, culturas, entre outras.

Definição Especialista: Cobertura hidrogeológica e biológica, natural ou antrópica, das superfícies terrestres do planeta.

Definição Modelador: Elemento da superfície do território (ou da superfície terrestre) que corresponde à composição de uma cobertura hidrogeológica e biológica, natural ou antrópica; influencia a paisagem e o controle e uso da terra.

24. MUDANÇA DE COBERTURA DA TERRA

Morfologia: Substantivo Feminino (s.f.)

Definição Inicial: Processo de modificação da cobertura do solo devido a fatores comerciais, demográficos ou referentes ao regime de propriedade do local.

Definição Especialista: Idem à definição inicial.

Definição Modelador: Processo de modificação da cobertura da terra, natural ou não, devido a fatores comerciais, demográficos ou referentes ao regime de propriedade do local.

SOCIOECONOMIA

25. DENSIDADE DEMOGRÁFICA

Morfologia: Substantivo Feminino (s.f.)

Definição Inicial: Média de indivíduos divididos pela superfície do território em que ocupam, ou seja, o número de pessoas ou outros seres vivos por espaço habitado. A densidade populacional rural é primordial para a explicação dos padrões de uso das terras, visto que a pressão populacional associada à produtividade agrícola é maior.

Definição Especialista: Refere-se ao número de pessoas ou outros seres vivos por espaço habitado.

Definição Modelador: População humana ou de outros seres vivos, medida em um dado espaço habitado, geralmente expressa por um indicador que determina o número de habitantes por quilômetro quadrado; as características dessa população impactam, por exemplo, o transporte, o uso da terra e a economia.

26. CRESCIMENTO POPULACIONAL

Morfologia: Substantivo Masculino (s.m.)

Definição Inicial: Aumento do número de habitantes em um determinado espaço. Variáveis como a taxa de crescimento populacional, taxas de mortalidade e natalidade são propriedades primordiais para a explicação do aumento ou diminuição do número de habitantes de um determinado local. O crescimento populacional conduz a degradação por meio do aumento de áreas de cultivo e diminuição da rotatividade de culturas, acarretando no aumento da degradação. Porém, este não é um fator exclusivo de mudança da cobertura florestal.

Definição Especialista: Aumento do número de indivíduos em um determinado espaço.

Definição Modelador: Dinâmica populacional humana e de outros seres vivos que corresponde ao aumento do número de indivíduos em um determinado espaço, medido por valores tais como taxa de mortalidade e natalidade, e que envolve fatores como os de emigração, imigração e migração urbana e rural.

27. PRESSÃO POPULACIONAL

Morfologia: Substantivo Feminino (s.f.)

Definição Inicial: Processo resultado do envolvimento entre a população e o local ocupado que, dentre outras consequências, conduz ao desmatamento. Relaciona-se com a densidade populacional e o poder de compra da população local. O parcelamento urbano para fins de moradia e outras atividades reflete diretamente em uma contínua degradação das terras e da paisagem, assim afetando a produtividade agrícola. Porém, o crescimento da população, ligado à pressão populacional encoraja a Intensificação Agropecuária e sistemas pecuários mistos.

Definição Especialista: Processo resultante da tensão entre a dinâmicas populacionais e as dinâmicas de uso e cobertura das terras.

Definição Modelador: Condição resultante de fatores demográficos e de um processo de tensão entre a dinâmicas populacionais, eventos naturais e as dinâmicas de uso e cobertura das terras, e pode contribuir para os danos ambientais tais como o desmatamento, poluição, esgotamento de recursos e perda de biodiversidade.

ESPAÇO E TEMPO

28. MUDANÇA AGRÍCOLA

Morfologia: Substantivo Feminino (s.f.)

Definição Inicial: Fator de mudanças não lineares associadas às práticas agrícolas. Representa as mudanças, ao longo do tempo e do espaço, no uso e na cobertura das terras, em função de sua exploração antrópica e, em consequência, pode acarretar diminuição ou aumento na produção agropecuária e nas áreas de cultivo, uma vez que seu processo pode ou não ser associado ao controle de desenvolvimento planejado em um ambiente de produção.

Definição Especialista: Alterações espaço-temporais das atividades agrícolas decorrentes da dinâmica de uso e cobertura das terras em função de sua exploração antrópica.

Definição Modelador: Desenvolvimento econômico das atividades agropecuárias que resulta das alterações espaço-temporais decorrentes da dinâmica de uso e cobertura das terras em função de sua exploração antrópica, que envolve fatores tais como a mecanização e as políticas de desenvolvimento rural.

29. ANÁLISE DE SÉRIES TEMPORAIS

Morfologia: Substantivo Feminino (s.f.)

Definição Inicial: Ferramenta estatística que auxilia nas previsões de padrões ou de séries ao longo de um período de tempo, a partir da identificação de componentes e de sua análise quantitativa. A

identificação dos dados temporais e dos padrões de comportamento histórico possibilita não só a previsão de ocorrências futuras, como também a organização de ações preventivas em relação às suas consequências.

Definição Especialista: Ferramenta estatística que auxilia nas previsões de padrões ou de séries ao longo de um período de tempo, a partir da identificação de componentes e de sua análise quantitativa. A identificação dos dados temporais e dos padrões de comportamento histórico possibilita não só a previsão de ocorrências futuras, como também a organização de ações preventivas em relação às suas consequências.

Definição Modelador: Ferramenta de análise de informações estatísticas que tem por base padrões de dados temporais, e que auxilia nas previsões de padrões, de séries ao longo de um período de tempo e na determinação de ações estratégicas de prevenção às suas possíveis consequências.

INSTITUIÇÕES

30. EMBRAPA

Morfologia: Substantivo Feminino (s.f.).

Equivalência em inglês: *Embrapa*

Definição Inicial: Empresa de inovação tecnológica que tem por objetivo a geração de conhecimento e tecnologia para a agropecuária brasileira. A Empresa Brasileira de Pesquisa Agropecuária (Embrapa) nasceu em 26 de abril de 1973, em meio à intensificação agropecuária no Brasil. O crescimento acelerado da população e da renda *per capita* e a abertura para o mercado externo mostravam que, sem investimentos em ciências agrárias, o país não conseguiria reduzir o diferencial entre o crescimento da demanda e o da oferta de alimentos e fibras. Vinculada ao Ministério da Agricultura, Pecuária e Abastecimento (Mapa), a Embrapa assumiu um desafio: desenvolver, em conjunto com os parceiros do Sistema Nacional de Pesquisa Agropecuária (SNPA), um modelo de Agricultura e pecuária tropical genuinamente brasileiro, superando as barreiras que limitavam a produção de alimentos, fibras e energia. A empresa é responsável pela formulação de políticas agrícolas e por planejar, supervisionar, coordenar e controlar as atividades relacionadas à execução de pesquisa agropecuária no Brasil. A Embrapa conta hoje com 17 unidades centrais localizadas em Brasília; 46 Unidades Descentralizadas em todas as regiões do Brasil; quatro Laboratórios Virtuais no Exterior (Labex), localizados nos EUA, Europa, China e Coreia do Sul; e três Escritórios Internacionais, situados na América Latina e na África. Além de tirar o Brasil de uma condição de importador de alimentos básicos para a condição de um dos maiores produtores e exportadores mundiais, a Embrapa internacionalmente é uma das líderes na geração de conhecimento, tecnologia e inovação para a produção sustentável de alimentos, fibras e agroenergia.

Definição Especialista: Idem à definição inicial.

Definição Modelador: Instituição de pesquisa vinculada ao Ministério da Agricultura, Pecuária e Abastecimento (MAPA), criada em 26 de abril de 1973, formada por 17 unidades centrais localizadas em Brasília; 46 Unidades Descentralizadas em todas as regiões do Brasil; quatro Laboratórios Virtuais no Exterior (Labex), localizados nos EUA, Europa, China e Coreia do Sul; e três Escritórios Internacionais, situados na América Latina e na África; tem como objetivo a inovação tecnológica para a geração de conhecimento e tecnologia para a agropecuária brasileira, cujos resultados de pesquisa influenciam as políticas públicas.

FONTE: Maculan (2015, p. 329).

ANEXO C - Estrutura Semântica da Amostra do AgroTermos

1 INTENSIFICAÇÃO AGROPECUÁRIA

temTraduçãoEN (EQ-EN) AGRICULTURAL INTENSIFICATION
 temParte (TEP) AGRONOMIA
 temParte (TEP) INSTITUIÇÕES
 temParte (TEP) SOCIOECONOMIA
 termoEspecíficoGênero (TEG) AGRICULTURA EXTENSIVA
 termoEspecíficoGênero (TEG) AGRICULTURA INTENSIVA
 termoEspecíficoGênero (TEG) ESPAÇO E TEMPO
 termoEspecíficoGênero (TEG) MATERIAL E MÉTODOS
 termoEspecíficoGênero (TEG) AMBIENTE
 termoEspecíficoGênero (TEG) TERRITÓRIO E PAISAGEM

2 AGRICULTURA EXTENSIVA

temTraduçãoEN (EQ-EN) EXTENSIVE AGRICULTURE
 temSinonímiaAproximada (UP EQ-APR) FRONTEIRA AGRÍCOLA
 temSinônimo (UP EQ-SIN) PRODUÇÃO EXTENSIVA
 termoGenéricoGênero (TGG) SISTEMA DE EXPLORAÇÃO AGRÍCOLA
 termoEspecíficoGênero (TEG) ATIVIDADE PASTORIL
 termoEspecíficoGênero (TEG) CRIAÇÃO EXTENSIVA
 termoEspecíficoGênero (TEG) EXTENSIFICAÇÃO
 incluídoEm (TR) AGRICULTURA COM BAIXO INVESTIMENTO
 éPráticaPara (TR) GRANDE EXPLORAÇÃO AGRÍCOLA
 temObjetivoOuProcesso (TR) UTILIZAÇÃO EXTENSIVA DA TERRA
 inclui (TR) SISTEMA AGROPASTORIL
 influencia (TR) PRODUÇÃO

3 AGRICULTURA INTENSIVA

temTraduçãoEN (EQ-EN) INTENSIVE FARMING
 temSinonímiaAproximada (UP EQ-APR) AGRICULTURA INDUSTRIAL
 temSinônimo (UP EQ-SIN) EXPLORAÇÃO AGRÍCOLA INTENSIVA
 temSinônimo (UP EQ-SIN) PRODUÇÃO INTENSIVA
 termoGenéricoGênero (TGG) INTENSIFICAÇÃO AGROPECUÁRIA
 termoGenéricoGênero (TGG) SISTEMA DE EXPLORAÇÃO AGRÍCOLA
 termoEspecíficoGênero (TEG) CULTIVO INTENSIVO
 termoEspecíficoGênero (TEG) PECUÁRIA INTENSIVA
 termoEspecíficoGênero (TEG) INTENSIFICAÇÃO SUSTENTÁVEL
 termoEspecíficoGênero (TEG) INTENSIFICAÇÃO
 temObjetivoOuProcesso (TR) UTILIZAÇÃO INTENSIVA DA TERRA
 influencia (TR) PRODUÇÃO

MATERIAL E MÉTODOS

4 Cultura

temNotaDeEscopo (NE) NE-1 Todas as plantas úteis, cultivadas ou não.
 temTraduçãoEN (EQ-EN) CROP UP VEGETAL PRODUCTION
 temSinonímiaAproximada (UP EQ-APR) PLANTAS ECONÔMICAS
 temSinônimo (UP EQ-SIN) PLANTA CULTIVADA
 temSinônimo (UP EQ-SIN) PLANTAS ÚTEIS
 temSinônimoEspecífico (UP EQ-TE) PRODUÇÃO VEGETAL
 éParteDe (TGP) AGRONOMIA
 termoGenéricoGênero (TGG) MATERIAL E MÉTODOS
 termoGenéricoGênero (TGG) PLANTAS
 termoEspecíficoGênero (TEG) CULTURA ARÁVEL
 termoEspecíficoGênero (TEG) CULTURA AUTÓCTONE
 termoEspecíficoGênero (TEG) CULTURA CÍCLICA
 termoEspecíficoGênero (TEG) CULTURA CONSORCIADA
 termoEspecíficoGênero (TEG) CULTURA DE AÇÚCAR
 termoEspecíficoGênero (TEG) CULTURA DE ALIMENTO

termoEspecíficoGênero (TEG) CULTURA DE AMIDO
 termoEspecíficoGênero (TEG) CULTURA DE BORRACHA
 termoEspecíficoGênero (TEG) CULTURA DE CABAÇA
 termoEspecíficoGênero (TEG) CULTURA DE CAMPO
 termoEspecíficoGênero (TEG) CULTURA DE ESPECIARIA
 termoEspecíficoGênero (TEG) CULTURA DE FRUTA
 termoEspecíficoGênero (TEG) CULTURA DE GRÃO
 termoEspecíficoGênero (TEG) CULTURA DE INVERNO
 termoEspecíficoGênero (TEG) CULTURA DE OLEAGINOSA
 termoEspecíficoGênero (TEG) CULTURA DE ÓLEOS ESSENCIAL
 termoEspecíficoGênero (TEG) CULTURA DE PESTICIDA
 termoEspecíficoGênero (TEG) CULTURA DE PORTA-ENXERTO
 termoEspecíficoGênero (TEG) CULTURA DE PRIMAVERA
 termoEspecíficoGênero (TEG) CULTURA DE RAÍZ
 termoEspecíficoGênero (TEG) CULTURA DE SEMENTE
 termoEspecíficoGênero (TEG) CULTURA DE TERRAS ALTAS
 termoEspecíficoGênero (TEG) CULTURA EM ESTUFA
 termoEspecíficoGênero (TEG) CULTURA ENERGETICA
 termoEspecíficoGênero (TEG) CULTURA ESTIMULANTE
 termoEspecíficoGênero (TEG) CULTURA ESTRUTURAL
 termoEspecíficoGênero (TEG) CULTURA FIBRA
 termoEspecíficoGênero (TEG) CULTURA FORRAGEIRA
 termoEspecíficoGênero (TEG) CULTURA INDUSTRIAL
 termoEspecíficoGênero (TEG) CULTURA INTERCALAR
 termoEspecíficoGênero (TEG) CULTURA NUT
 termoEspecíficoGênero (TEG) CULTURA PARA ALIMENTAÇÃO ANIMAL
 termoEspecíficoGênero (TEG) CULTURA PARA ARMADILHA
 termoEspecíficoGênero (TEG) CULTURA PARA BEBIDA
 termoEspecíficoGênero (TEG) CULTURA PARA COMBUSTÍVEL
 termoEspecíficoGênero (TEG) CULTURA PARA SALADA
 termoEspecíficoGênero (TEG) CULTURA PERENE
 termoEspecíficoGênero (TEG) CULTURA TROPICAL
 termoEspecíficoGênero (TEG) ERVA CULINÁRIA
 termoEspecíficoGênero (TEG) HORTALIÇA
 termoEspecíficoGênero (TEG) IMPLANTACAO DE CULTURA
 termoEspecíficoGênero (TEG) LAVOURA
 termoEspecíficoGênero (TEG) PLANTA APÍCOLA
 termoEspecíficoGênero (TEG) PLANTA DE CERA
 termoEspecíficoGênero (TEG) PLANTA DE COBERTURA
 termoEspecíficoGênero (TEG) PLANTA DE CONTROLE DE EROSÃO
 termoEspecíficoGênero (TEG) PLANTA DE DROGAS
 termoEspecíficoGênero (TEG) PLANTA DE PROTEÇÃO
 termoEspecíficoGênero (TEG) PLANTA DE RECUPERAÇÃO DE SOLO
 termoEspecíficoGênero (TEG) PLANTA DE REVEGETAÇÃO
 termoEspecíficoGênero (TEG) PLANTA DE SOMBRA
 termoEspecíficoGênero (TEG) PLANTA PARA RUMINANTE
 termoEspecíficoGênero (TEG) PLANTA TAN
 termoEspecíficoGênero (TEG) PLANTA TINTUREIRA
 termoEspecíficoGênero (TEG) PLANTAÇÃO DE ÁRVORE
 termoEspecíficoGênero (TEG) PLANTA-MÃE
 termoEspecíficoGênero (TEG) PROTEAGINOSA
 éInfluenciadaPorOuDependeDe (TR) CALENDÁRIO AGRÍCOLA
 éInfluenciadaPorOuDependeDe (TR) TAXA DE LOTAÇÃO
 éInfluenciadaPorOuDependeDe (TR) CICLO DA CULTURA
 temBenefícioAPartirDe (TR) ÁRVORE QUEBRA-VENTO
 temBenefícioAPartirDe (TR) MELHORAMENTO VEGETAL
 temBenefícioAPartirDe (TR) PLANTA SELVAGEM
 temPrática (TR) DOMESTICAÇÃO

5 sensoriamento remoto

temTraduçãoEN (EQ-EN) REMOTE SENSING
 temSinonímiaAproximada (UP EQ-APR) IMAGEM DE SATELITE
 éParteDe (TGP) AGRIMENSURA
 éParteDe (TGP) GEOMETRIA GEOESPACIAL
 termoGenéricoGênero (TGG) MATERIAL E MÉTODOS
 fazUsoDe (TR) AEROFOTOGRAMETRIA
 fazUsoDe (TR) AVHRR
 fazUsoDe (TR) IMAGEM ESPECTRAL
 fazUsoDe (TR) IMAGEM LOCAL
 fazUsoDe (TR) IMAGEM DE SATÉLITE
 fazUsoDe (TR) LEVANTAMENTO AÉREO
 fazUsoDe (TR) MAPAS
 fazUsoDe (TR) PROCESSAMENTO DE IMAGEM
 fazUsoDe (TR) RADAR
 fazUsoDe (TR) SATÉLITE
 fazUsoDe (TR) SATÉLITE DE OBSERVAÇÃO DA TERRA
 fazUsoDe (TR) SENSOR
 fazUsoDe (TR) SENSOR REMOTO
 fazUsoDe (TR) SISTEMA DE POSICIONAMENTO GLOBAL
 usaProcesso (TR) FOTOGRAMETRIA
 usaProcesso (TR) TELEMETRIA

6 sistema de informação geográfica

temTraduçãoEN (EQ-EN) GEOGRAPHIC INFORMATION SYSTEM
 temAcrônimo (UP EQS-ACR) SIG
 temSinonímiaAproximada (UP EQ-APR) SISTEMA DE INFORMAÇÃO GEOESPACIAL
 termoGenéricoGênero (TGG) MATERIAL E MÉTODOS
 termoGenéricoGênero (TGG) SISTEMA DE INFORMAÇÃO
 termoGenéricoGênero (TGG) TECNOLOGIA GEOESPACIAL
 termoEspecíficoGênero (TEG) SISTEMA DE OBSERVAÇÃO DA TERRA
 éComponenteDe (TGP) SISTEMAS DE SUPORTE
 afeta (TR) AGRICULTURA DE PRECISÃO
 fazUsoDe (TR) CARTOGRAFIA
 fazUsoDe (TR) SISTEMA DE POSICIONAMENTO GLOBAL
 inclui (TR) DADO ESPACIAL

AMBIENTE**7 meio ambiente**

temTraduçãoEN (EQ-EN) ENVIRONMENT
 temSinonímiaAproximada (UP EQ-APR) NATUREZA
 temSinonímiaAproximada (UP EQ-APR) ZONAS ECOLÓGICAS
 termoGenéricoGênero (TGG) AMBIENTE
 termoGenéricoGênero (TGG) ENTIDADES
 termoEspecíficoGênero (TEG) AMBIENTE AQUÁTICO
 termoEspecíficoGênero (TEG) AMBIENTE BIÓTICO
 termoEspecíficoGênero (TEG) AMBIENTE ECONÔMICO
 termoEspecíficoGênero (TEG) AMBIENTE FÍSICO
 termoEspecíficoGênero (TEG) AMBIENTE NATURAL
 termoEspecíficoGênero (TEG) ECOSISTEMA
 termoEspecíficoGênero (TEG) FATOR AMBIENTAL
 termoEspecíficoGênero (TEG) HABITAT
 termoEspecíficoGênero (TEG) RECURSO NATURAL
 termoEspecíficoGênero (TEG) ZONAS RIBEIRINHAS
 temComponente (TEP) ECOSSISTEMA
 temComponente (TEP) FATOR AMBIENTAL
 afeta (TR) AMBIENTE SOCIOCULTURAL
 afeta (TR) AMBIENTE SOCIOECONÔMICO
 afeta (TR) INTERAÇÃO GENÉTICA

afeta (TR) SUSTENTABILIDADE
 afeta (TR) VEGETAÇÃO
 afeta (TR) ZONA URBANA
 afeta (TR) ZONEAMENTO ECOLÓGICO
 éAfetadoPor (TR) BIODEGRADAÇÃO
 éAfetadoPor (TR) CATÁSTROFE NATURAL
 éAfetadoPor (TR) EDUCAÇÃO AMBIENTAL
 éAfetadoPor (TR) FOGO
 éAfetadoPor (TR) FATOR CLIMÁTICO
 éAfetadoPor (TR) IMPACTO AMBIENTAL
 éAfetadoPor (TR) POLUIÇÃO AMBIENTAL
 éAfetadoPor (TR) SUBSTÂNCIA TÓXICA
 éAfetadoPor (TR) TEMPERATURA
 éInfluenciadoPor (TR) POLÍTICA AMBIENTAL
 éInfluenciadoPor (TR) AVALIAÇÃO DO IMPACTO AMBIENTAL
 éInfluenciadoPor (TR) AGRICULTURA SUSTENTÁVEL
 fazUsoDe (TR) ECOLOGIA
 influencia (TR) AEROBIOSE
 influencia (TR) AGRICULTURA SUSTENTÁVEL
 influencia (TR) ANAEROBIOSE
 influencia (TR) AVALIAÇÃO DO IMPACTO AMBIENTAL
 influencia (TR) BIOSSEGURANÇA
 influencia (TR) MICROCLIMA
 influencia (TR) POLÍTICA AMBIENTAL
 usaProcesso (TR) ZONEAMENTO ECOLÓGICO

8 ambiente físico

temNotaDeEscopo (NE) NE-2 Conjunto de recursos físicos e químicos, representado pelos fatores abióticos que compõem o meio ambiente.
 temTraduçãoEN (EQ-EN) PHYSICAL ENVIRONMENT
 termoGenéricoGênero (TGG) MEIO AMBIENTE
 termoEspecíficoGênero (TEG) ÁGUA
 termoEspecíficoGênero (TEG) AR
 termoEspecíficoGênero (TEG) SOLO
 éInfluenciadoPor (TR) CAMPO MAGNÉTICO
 éInfluenciadoPor (TR) FATOR CLIMÁTICO
 éInfluenciadoPor (TR) SEDIMENTAÇÃO
 incluídoEm (TR) BIOMA
 influencia (TR) PROCESSO QUÍMICO
 afeta (TR) SOCIOECONOMIA

9 solo

temNotaDeEscopo (NE) NE-3 Camada superior da terra, em que as plantas crescem; para os diferentes tipos de solo de uma região, país, etc., usar <TIPOS DE SOLO [soil types]>
 temTraduçãoEN (EQ-EN) SOIL
 éParteDe (TGP) CIÊNCIA DO SOLO
 termoGenéricoGênero (TGG) COBERTURA DA TERRA
 termoGenéricoGênero (TGG) AMBIENTE
 termoEspecíficoGênero (TEG) CONSERVAÇÃO DO SOLO
 termoEspecíficoGênero (TEG) FATOR MORFOLÓGICO DO SOLO
 termoEspecíficoGênero (TEG) FORMAÇÃO DO SOLO
 termoEspecíficoGênero (TEG) TIPOS DE SOLO
 éCompostoDe (TEP) MATÉRIA ORGÂNICA DO SOLO
 éCompostoDe (TEP) MINERAL DO SOLO
 temComponente (TEP) AR DO SOLO
 temComponente (TEP) ÁGUA DO SOLO
 temParte (TEP) RIZOSFERA
 temParte (TEP) SOLO SUPERIOR
 temParte (TEP) SUBSOLO
 éInfluenciadoPor (TR) BIOTA DO SOLO

éUtilizadoComo (TR) SUBSTRATO DE CULTURA
 influencia (TR) LIXIVIADO
 influencia (TR) MANEJO DO SOLO
 influencia (TR) PERDA DE SOLO
 influencia (TR) RESTO DE CULTURA
 usaProcesso (TR) MELHORAMENTO DO SOLO
 usaProcesso (TR) MANEJO DE SOLO

AGRONOMIA

10 manejo da cultura

temNotaDeEscopo (NE) NE-4 Decisões sobre práticas culturais, do plantio à colheita.
 temTraduçãoEN (EQ-EN) CROP MANAGEMENT
 temVariaçãoOrtográfica (UP EQS-ORT) MANEJO DE CULTURA
 termoGenéricoGênero (TGG) MANEJO
 éParteDe (TGP) AGRONOMIA
 éParteDe (TGP) CIÊNCIA VEGETAL
 termoEspecíficoGênero (TEG) MANEJO DE PRAGA
 termoEspecíficoGênero (TEG) MANEJO DE RESÍDUO DE CULTURA
 termoEspecíficoGênero (TEG) MANEJO INTEGRADO DA CULTURA
 termoEspecíficoGênero (TEG) VIVEIROS
 temComponente (TEP) CULTURA
 temComponente (TEP) ESCOLHA DE ESPÉCIES
 temComponente (TEP) OPERAÇÃO DE PRÉ-COLHEITA
 temComponente (TEP) PADRÕES DE CULTIVO
 temComponente (TEP) PROGRAMAÇÃO DE IRRIGAÇÃO
 temComponente (TEP) SISTEMAS DE CULTIVO
 temComponente (TEP) TRATAMENTO PRÉ-COLHEITA
 éInfluenciadoPor (TR) ESTÁGIO DA CULTURA
 fazUsoDe (TR) MANEJO DE FERTILIDADE
 fazUsoDe (TR) MÉTODO DE PLANTIO
 fazUsoDe (TR) MÉTODO DE SEMEADURA
 inclui (TR) MÉTODO DE CULTURA
 influencia (TR) FREQUÊNCIA DE COLHEITA
 utilizaValor (TR) CALENDÁRIO DE COLHEITA
 utilizaValor (TR) CUSTO DAS MUDAS
 utilizaValor (TR) CUSTO DE APLICAÇÃO
 utilizaValor (TR) CUSTO DE IRRIGAÇÃO
 utilizaValor (TR) DENSIDADES DE SEMEADURA
 utilizaValor (TR) ÉPOCA DE SEMEADURA
 utilizaValor (TR) ESPAÇAMENTO ENTRE MUDAS
 utilizaValor (TR) PERÍODO DE PLANTIO
 utilizaValor (TR) PERÍODO DE TRATAMENTO
 utilizaValor (TR) PROFUNDIDADE DE SEMEADURA
 utilizaValor (TR) TEMPO DE REPLANTAR
 éInfluenciado (TR) AMBIENTE FÍSICO

11 manejo do solo

temTraduçãoEN (EQ-EN) SOIL MANAGEMENT
 temSinonímiaAproximada (UP EQ-APR) PREPARO DO SOLO
 termoGenéricoGênero (TGG) MANEJO DE RECURSOS NATURAIS
 termoGenéricoGênero (TGG) PRÁTICA DE MANEJO
 éParteDe (TGP) AGRONOMIA
 éParteDe (TGP) CIÊNCIA DO SOLO
 termoEspecíficoGênero (TEG) CALAGEM
 termoEspecíficoGênero (TEG) COBERTURAS
 termoEspecíficoGênero (TEG) CONSERVAÇÃO DO SOLO
 termoEspecíficoGênero (TEG) CONTROLE DE SALINIDADE
 termoEspecíficoGênero (TEG) DRENAGEM
 termoEspecíficoGênero (TEG) FERTILIZANTES

termoEspecíficoGênero (TEG) FUMIGAÇÃO
 termoEspecíficoGênero (TEG) IRRIGAÇÃO
 termoEspecíficoGênero (TEG) PREPARAÇÃO LOCAL
 termoEspecíficoGênero (TEG) PROTEÇÃO DO SOLO
 termoEspecíficoGênero (TEG) TRATAMENTO DO SOLO
 afeta (TR) SUSTENTABILIDADE
 éProcessoPara (TR) SOLO
 fazUsoDe (TR) MÉTODO ALTERNATIVO
 fazUsoDe (TR) TERRAÇO
 inclui (TR) ESTERILIZAÇÃO
 inclui (TR) MELHORIA DO SOLO
 inclui (TR) POUSIO
 inclui (TR) PRÁTICA DE CULTURA VEGETAL
 influencia (TR) APLICAÇÃO DE FERTILIZANTES
 influencia (TR) AQUECIMENTO DO SOLO
 influencia (TR) CULTURA
 influencia (TR) MANEJO DA ÁGUA
 temObjetoDeAtividade (TR) EQUIPAMENTO DE MANEJO DO SOLO
 temPrática (TR) ESCARIFICAÇÃO DO SOLO
 temPrática (TR) RECUPERAÇÃO
 temPrática (TR) TERRAPLANAGEM

12 pousio

temNotaDeEscopo (NE) NE-5 Período de tempo em que a terra permanece não cultivada, por exemplo, entre uma safra e outra, em ciclos culturais anuais.
 temTraduçãoEN (EQ-EN) FALLOW
 temSinônimo (UP EQ-SIN) CULTIVO POUSIO
 termoGenéricoGênero (TGG) PRÁTICA DE MANEJO
 termoGenéricoGênero (TGG) SISTEMAS DE CULTIVO
 éParteDe (TGP) AGRONOMIA
 termoEspecíficoGênero (TEG) PERÍODO DE POUSIO
 incluídoEm (TR) PREPARAÇÃO DO SOLO
 influencia (TR) SISTEMAS DE POUSIO
 éAfetadoPor (TR) CICLO DA CULTURA
 éAfetadoPor (TR) CULTURA ANUAL

13 período de pousio

temTraduçãoEN (EQ-EN) FALLOW PERIOD
 termoGenéricoGênero (TGG) POUSIO
 termoEspecíficoGênero (TEG) POUSIO NU
 termoEspecíficoGênero (TEG) POUSIO LONGO
 termoEspecíficoGênero (TEG) POUSIO CURTO
 afeta (TR) SUSTENTABILIDADE
 afeta (TR) SOCIOECONOMIA

14 pesticida

temTraduçãoEN (EQ-EN) PESTICIDES
 temSinônimo (UP EQ-SIN) PRAGUICIDA
 éProdutoDe (TGG) INDÚSTRIA DE PESTICIDA
 termoGenéricoGênero (TGG) MANEJO DE PRAGA
 termoGenéricoGênero (TGG) SUBSTÂNCIA QUÍMICA
 termoEspecíficoGênero (TEG) ACARICIDA
 termoEspecíficoGênero (TEG) AVICIDA
 termoEspecíficoGênero (TEG) BIOPESTICIDA
 termoEspecíficoGênero (TEG) CARRAPATICIDA
 termoEspecíficoGênero (TEG) FUNGICIDA
 termoEspecíficoGênero (TEG) GERMICIDA
 termoEspecíficoGênero (TEG) HERBICIDA
 termoEspecíficoGênero (TEG) INSETICIDA
 termoEspecíficoGênero (TEG) MOLUSCICIDA

termoEspecíficoGênero (TEG) NEMATICIDA
 termoEspecíficoGênero (TEG) OVICIDA
 termoEspecíficoGênero (TEG) PESTICIDA ORGANOESTÂNICO
 termoEspecíficoGênero (TEG) PESTICIDA SINÉRGICO
 termoEspecíficoGênero (TEG) PESTICIDAS DE LIBERAÇÃO LENTA
 termoEspecíficoGênero (TEG) PESTICIDAS ORGANOCLORADO
 termoEspecíficoGênero (TEG) RATICIDA
 termoEspecíficoGênero (TEG) REPELENTE
 termoEspecíficoGênero (TEG) RODENTICIDA
 temComponente (TEP) BÓRAX
 temComponente (TEP) ADJUVANTES
 temComponente (TEP) INGREDIENTE ATIVO temComponente (TEP) REPELENTE
 temComponente (TEP) PRODUTOS QUÍMICOS
 afeta (TR) CONTROLE DE PRAGA
 causa (TR) POLUIÇÃO
 éControladoPor (TR) SEGURANÇA NO USO DE PESTICIDA
 éFeitoAPartirDe (TR) ADJUVANTE PESTICIDA
 éFeitoAPartirDe (TR) SUBSTÂNCIA TÓXICA
 éUsadoComo (TR) CONSERVANTE DE MADEIRA
 éUtilizadoComo (TR) APLICAÇÃO DE PESTICIDA
 influencia (TR) ENCHARCAMENTO DO SOLO
 seDesenvolveEm (TR) RESÍDUOS DE PESTICIDAS
 temMétodoDeControle (TR) CONTROLE QUÍMICO
 temPropriedade (TR) AÇÃO SISTÊMICA
 temPropriedade (TR) ESTRUTURA RELAÇÃO DE ATIVIDADE
 temPropriedade (TR) INGESTÃO DIÁRIA ACEITÁVEL
 temPropriedade (TR) MODO DE AÇÃO
 temPropriedade (TR) PROPRIEDADES PESTICIDAS
 temPropriedade (TR) RESISTÊNCIA A PESTICIDAS
 temPropriedade (TR) RESISTENCIA QUIMICA
 usaProcesso (TR) APLICAÇÃO FOLIAR
 usaProcesso (TR) TRATAMENTO DE SEMENTES

15 ciclo da cultura

temTraduçãoEN (EQ-TEN) CROP CYCLE
 temSinônimo (UP EQ-SIN) CICLO DA PLANTA
 éParteDe (TGP) AGRONOMIA
 termoEspecíficoGênero (TEG) ESTÁGIO DA CULTURA
 afeta (TR) MANEJO DA CULTURA
 afeta (TR) MANEJO DO SOLO
 afeta (TR) POUSSO
 influencia (TR) CULTURA
 influencia (TR) PERÍODO DE TRATAMENTO

16 cultura anual

temTraduçãoEN (EQ-EN) ANNUAL PLANT
 temSinonímiaAproximada (UP EQ-APR) CULTURA DE CICLO CURTO
 temSinônimo (UP EQ-SIN) PLANTA ANUAL
 termoGenéricoGênero (TGG) CULTURA CÍCLICA
 termoGenéricoGênero (TGG) SISTEMAS DE CULTIVO
 termoEspecíficoGênero (TEG) CULTURA DE CEREAL
 termoEspecíficoGênero (TEG) CULTURA DE GRÃO
 termoEspecíficoGênero (TEG) ERVA DANINHA ANUAL
 termoEspecíficoGênero (TEG) PLANTA ANUAL DE JARDIM
 termoEspecíficoGênero (TEG) PLANTA ANUAL ORNAMENTAL
 temComponente (TEP) PLANTA HERBÁCEA
 afeta (TR) POUSSO
 éTipoRelacionadoA (TR) PORTE DA PLANTA

17 pecuária

temTraduçãoEN (EQ-EN) LIVESTOCK
 temSinonímiaAproximada (UP EQ-APR) ANIMAL DE FAZENDA
 temSinonímiaAproximada (UP EQ-APR) ESTOQUE DE ANIMAL
 temSinonímiaAproximada (UP EQ-APR) PECUÁRIA E AGRICULTURA
 termoGenéricoGênero (TGG) ANIMAL DOMÉSTICO
 termoGenéricoGênero (TGG) ANIMAL
 éParteDe (TGP) AGRONOMIA
 termoEspecíficoGênero (TEG) ANIMAIS MONOGÁSTRICOS
 termoEspecíficoGênero (TEG) AVES
 termoEspecíficoGênero (TEG) BOVINOS E BUBALINOS
 termoEspecíficoGênero (TEG) CAPRINOS E OVINOS
 termoEspecíficoGênero (TEG) EQUÍDEOS
 termoEspecíficoGênero (TEG) PECUÁRIA EXTENSIVA
 termoEspecíficoGênero (TEG) PECUÁRIA INTENSIVA
 termoEspecíficoGênero (TEG) PEIXES
 termoEspecíficoGênero (TEG) SUÍNOS
 temParte (TEP) CIRCUITO PECUÁRIO CENTRO-OESTE
 temParte (TEP) CIRCUITO PECUÁRIO LESTE
 temParte (TEP) CIRCUITO PECUÁRIO NORDESTE
 temParte (TEP) CIRCUITO PECUÁRIO NORTE
 temParte (TEP) CIRCUITO PECUÁRIO SUL
 afeta (TR) INDÚSTRIA PECUÁRIA
 éInfluenciadoPorOuDependeDe (TR) REPRODUÇÃO ANIMAL
 inclui (TR) ANIMAIS DE TRABALHO
 inclui (TR) PECUÁRIA DE CORTE
 inclui (TR) PECUÁRIA DE LÃ
 inclui (TR) PECUÁRIA DE LEITE
 influencia (TR) PRODUÇÃO ANIMAL
 utilizaValor PREÇO DO ANIMAL

18 adubo verde

temNotaDeEscopo (NE) NE-6 Subprodutos de culturas ou restos culturais incorporados ao solo, enquanto ainda não apodrecidos, para melhorar a preparação do solo.
 temTraduçãoEN (EQ-EN) GREEN MANURES
 temSinônimo (UP EQ-SIN) ADUBO ORGÂNICO
 termoGenéricoGênero (TGG) ADUBO
 termoGenéricoGênero (TGG) FERTILIZANTE ORGÂNICO
 éUsadoComo (TR) ADUBAÇÃO VERDE
 éUsadoComo (TR) PLANTA DE COBERTURA
 éUsadoEm (TR) CULTURA
 fazUsoDe (TR) GLIRICIDIA
 fazUsoDe (TR) VIGNA VEXILLATA
 incluídoEm (TR) PRÁTICA DE CONSERVAÇÃO

19 produto agropecuário

temNotaDeEscopo (NE) NE-7 Produtos em seu estado original ou processados, disponíveis para venda(marketing).
 temTraduçãoEN (EQ-EN) AGRICULTURAL PRODUCTS
 temSinonímiaAproximada (UP EQ-APR) COMMODITIE AGRÍCOLA
 temSinonímiaAproximada (UP EQ-APR) MERCADORIAS
 temSinônimo (UP EQ-SIN) PRODUTO AGRÍCOLA
 temSinônimo (UP EQ-SIN) PRODUTO DA AGRICULTURA
 termoGenéricoGênero (TGG) BENS POR ATACADO
 termoGenéricoGênero (TGG) PRODUTO
 termoEspecíficoGênero (TEG) BIOPRODUTO
 termoEspecíficoGênero (TEG) PRODUTO ALIMENTÍCIO
 termoEspecíficoGênero (TEG) PRODUTO DA PESCA
 termoEspecíficoGênero (TEG) PRODUTO DE CELULOSE
 termoEspecíficoGênero (TEG) PRODUTO DE ORIGEM ANIMAL

termoEspecíficoGênero (TEG) PRODUTO DE ORIGEM VEGETAL
 termoEspecíficoGênero (TEG) PRODUTO FEITO DE INSETO
 termoEspecíficoGênero (TEG) PRODUTO FORA DA ÉPOCA
 termoEspecíficoGênero (TEG) PRODUTO PARA ALIMENTAÇÃO ANIMAL
 termoEspecíficoGênero (TEG) PRODUTO PERECÍVEL
 termoEspecíficoGênero (TEG) PRODUTO VEGETAL
 termoEspecíficoGênero (TEG) PRODUTO FLORESTAL
 temComponente (TEP) PRODUTO ARMAZENADO
 temComponente (TEP) PRODUTO FRESCO
 temComponente (TEP) PRODUTO PROCESSADO
 temComponente (TEP) SUBPRODUTO
 afeta (TR) ECONOMIA AGRÍCOLA
 éInsumoPara (TR) AGRONEGÓCIO
 influencia (TR) BOLSA DE MERCADORIA
 usaProcesso (TR) ROTULAGEM DE PRODUTO

20 biomassa

temNotaDeEscopo (NE) NE-8 Quantidade total ou o peso de organismos disponíveis numa dada área.
 temTraduçãoEN (EQ-EN) BIOMASS
 temSinônimo (UP EQ-SIN) BIOMASSA DA CULTURA
 temSinônimo (UP EQ-SIN) BIOMASSA DA PLANTA
 termoGenéricoGênero (TGG) RECURSO BIOLÓGICO
 termoGenéricoGênero (TGG) SUBSTÂNCIAS
 éComponenteDe (TGP) CULTURA ENERGÉTICA
 éComponenteDe (TGP) RECURSO RENOVÁVEL
 éParteDe (TGP) AGRONOMIA
 termoEspecíficoGênero (TEG) BIOMASSA AÉREA
 termoEspecíficoGênero (TEG) BIOMASSA MICROBIANA
 termoEspecíficoGênero (TEG) BIOMASSA SUBTERRÂNEA
 termoEspecíficoGênero (TEG) FITOMASSA
 afeta (TR) ACUMULAÇÃO DE MATÉRIA SECA
 inclui (TR) BIOCARVÃO
 inclui (TR) MATÉRIA ORGÂNICA
 influencia (TR) BIOENERGIA
 influencia (TR) ÍNDICE DE VEGETAÇÃO
 influencia (TR) PRODUTIVIDADE PRIMÁRIA
 produtoDe (TR) PRODUÇÃO DE BIOMASSA
 usaProcesso (TR) BIOREFINAMENTO
 utilizaValor (TR) AREA
 utilizaValor (TR) MASSA
 utilizaValor (TR) PESO
 utilizaValor (TR) VOLUME

TERRITÓRIO E PAISAGEM

21 posse da terra

temTraduçãoEN (EQ-EN) LAND TENURE
 temSinonímiaAproximada (UP EQ-APR) SISTEMA FUNDIÁRIO
 temSinônimo (UP EQ-SIN) SISTEMA DE POSSE
 temVariaçãoOrtográfica (UP EQS-ORT) TERRA, POSSE DA
 termoGenéricoGênero (TGG) ECONOMIA DA TERRA
 termoGenéricoGênero (TGG) POSSE
 termoGenéricoGênero (TGG) PROPRIEDADE RURAL
 termoGenéricoGênero (TGG) TERRITÓRIO E PAISAGEM
 termoEspecíficoGênero (TEG) LATIFUNDIO
 termoEspecíficoGênero (TEG) MINIFUNDIO
 éInfluenciadoPor (TR) ARRENDAMENTO
 éInfluenciadoPor (TR) DESAPROPRIAÇÃO
 éInfluenciadoPor (TR) POSSEIRO
 éInfluenciadoPor (TR) REFORMA AGRÁRIA

éInfluenciadoPor (TR) RENDA DA TERRA
 inclui (TR) LEASING
 influencia (TR) DIREITO À TERRA
 influencia (TR) DIREITO DE POSSE
 influencia (TR) MUDANÇA DE COBERTURA DAS TERRAS
 influencia (TR) USUCAPIÃO

22 escassez de terra

temTraduçãoEN (EQ-EN) LAND SCARCITY
 termoGenéricoGênero (TGG) TERRITÓRIO E PAISAGEM
 éInfluenciadoPor (TR) BIODEGRADAÇÃO
 éInfluenciadoPor (TR) CRESCIMENTO POPULACIONAL
 éInfluenciadoPor (TR) FATORES CLIMÁTICOS
 éInfluenciadoPor (TR) ZONA URBANA
 influencia (TR) CULTURA DE ALIMENTO
 influencia (TR) LATIFÚNDIO
 influencia (TR) POLÍTICAS PÚBLICAS
 influencia (TR) REFORMA AGRÁRIA
 influencia (TR) SEGURANÇA ALIMENTAR

23 cobertura da terra

temTraduçãoEN (EQ-EN) LAND COVER
 termoGenéricoGênero (TGG) CARACTERÍSTICAS FISIOGRÁFICAS
 termoGenéricoGênero (TGG) TERRITÓRIO E PAISAGEM
 termoEspecíficoGênero (TEG) ÁGUA DE SUPERFÍCIE
 termoEspecíficoGênero (TEG) MUDANÇA DE COBERTURA DAS TERRAS
 termoEspecíficoGênero (TEG) PLANTA DE COBERTURA
 termoEspecíficoGênero (TEG) SOLO
 termoEspecíficoGênero (TEG) TIPOS DE COBERTURA
 afeta (TR) ECOLOGIA
 inclui (TR) COBERTURA FÍSICA DA TERRA
 inclui (TR) ESTRUTURA GEOLÓGICA
 inclui (TR) HIDROGRAFIA
 influencia (TR) PAISAGEM
 influencia (TR) USO DA TERRA
 temObjetoDeAtividade (TR) TERRA

24 mudança de cobertura da terra

temTraduçãoEN (EQ-EN) LAND COVER CHANGE
 termoGenéricoGênero (TGG) COBERTURA DA TERRA
 termoGenéricoGênero (TGG) FENÔMENO NATURAL
 termoGenéricoGênero (TGG) TERRITÓRIO E PAISAGEM
 éInfluenciadoPor (TR) AGRICULTURA COMERCIAL
 éInfluenciadoPor (TR) DENSIDADE DEMOGRÁFICA
 éInfluenciadoPor (TR) POSSE DA TERRA

SOCIOECONOMIA

25 Densidade demográfica

temTraduçãoEN (EQ-EN) POPULATION DENSITY
 temSinônimo (UP EQ-SIN) DENSIDADE POPULACIONAL
 termoGenéricoGênero (TGG) CARACTERÍSTICAS POPULACIONAL
 termoGenéricoGênero (TGG) MEDIÇÃO
 termoGenéricoGênero (TGG) POPULAÇÃO
 éParteDe (TGP) SOCIOECONOMIA
 termoEspecíficoGênero (TEG) DENSIDADE VEGETAL
 termoEspecíficoGênero (TEG) POPULAÇÃO HUMANA
 termoEspecíficoGênero (TEG) SUPERPOPULAÇÃO
 inclui (TR) NÚMERO DE HABITANTE

influencia (TR) CAPACIDADE DE TRANSPORTE
 influencia (TR) PRESSÃO DEMOGRÁFICA
 influencia (TR) USO DA TERRA
 utilizaValor (TR) TAXA DE DISTRIBUIÇÃO DE SEMENTES
 utilizaValor (TR) TAXA DE FECUNDIDADE TOTAL
 utilizaValor (TR) TAXA DE MORTALIDADE INFANTIL
 utilizaValor (TR) TAXA DE URBANIZAÇÃO

26 crescimento populacional

temTraduçãoEN (EQ-EN) POPULATION GROWTH
 temSinônimo (UP EQ-SIN) AUMENTO DA POPULAÇÃO
 temSinônimo (UP EQ-SIN) EVOLUÇÃO DA POPULAÇÃO
 termoGenéricoGênero (TGG) DINÂMICA POPULACIONAL
 termoGenéricoGênero (TGG) POPULAÇÃO HUMANA
 éParteDe (TGP) SOCIOECONOMIA
 termoEspecíficoGênero (TEG) EMIGRACAO RURAL E URBANA
 termoEspecíficoGênero (TEG) IMIGRAÇÃO RURAL E URBANA
 termoEspecíficoGênero (TEG) MIGRACAO RURAL E URBANA
 influencia (TR) PRESSÃO POPULACIONAL
 utilizaValor (TR) TAXA DE MORTALIDADE INFANTIL
 utilizaValor (TR) TAXA DE NATALIDADE

27 pressão populacional

temTraduçãoEN (EQ-EN) POPULATION PRESSURE
 temSinônimo (UP EQ-SIN) PRESSÃO DEMOGRÁFICA
 termoGenéricoGênero (TGG) ESTADO E CONDIÇÃO
 éParteDe (TGP) SOCIECONOMIA
 afeta (TR) DESMATAMENTO
 atuaSobre (TR) DESMATAMENTO
 causa (TR) DESMATAMENTO
 influencia (TR) COBERTURA DA TERRA
 influencia (TR) CRESCIMENTO POPULACIONAL
 influencia (TR) DINÂMICA POPULACIONAL
 influencia (TR) USO DA TERRA

ESPAÇO E TEMPO

28 mudança agrícola

temTraduçãoEN (EQ-EN) AGRICULTURAL CHANGE
 temSinonímiaAproximada (UP EQ-APR) DESENVOLVIMENTO AGRÍCOLA
 termoGenéricoGênero (TGG) DESENVOLVIMENTO ECONÔMICO
 termoGenéricoGênero (TGG) ESPAÇO E TEMPO
 termoEspecíficoGênero (TEG) DESENVOLVIMENTO DA PESCA
 termoEspecíficoGênero (TEG) DESENVOLVIMENTO FLORESTAL
 termoEspecíficoGênero (TEG) MECANIZAÇÃO
 termoEspecíficoGênero (TEG) REVOLUÇÃO VERDE
 influencia (TR) COBERTURA DA TERRA
 influencia (TR) FOMENTO AGRÍCOLA
 influencia (TR) POLÍTICA DE DESENVOLVIMENTO
 influencia (TR) PRODUÇÃO
 influencia (TR) USO DA TERRA

29 análise de séries temporais

temTraduçãoEN (EQ-EN) TIME SERIES ANALYSIS
 temSinonímiaAproximada (UP EQ-APR) ANÁLISE DE SÉRIES CRONOLÓGICAS
 termoGenéricoGênero (TGG) ESPAÇO E TEMPO
 termoGenéricoGênero (TGG) ANÁLISE DE DADOS
 termoGenéricoGênero (TGG) MÉTODOS ESTATÍSTICOS
 éInfluenciadoPorOuDependeDe (TR) ANÁLISE ESTATÍSTICA
 éInfluenciadoPorOuDependeDe (TR) ANÁLISE GEOESTATÍSTICA

utilizaValor (TR) PADRÃO TEMPORAL
utilizaValor (TR) SÉRIE TEMPORAL

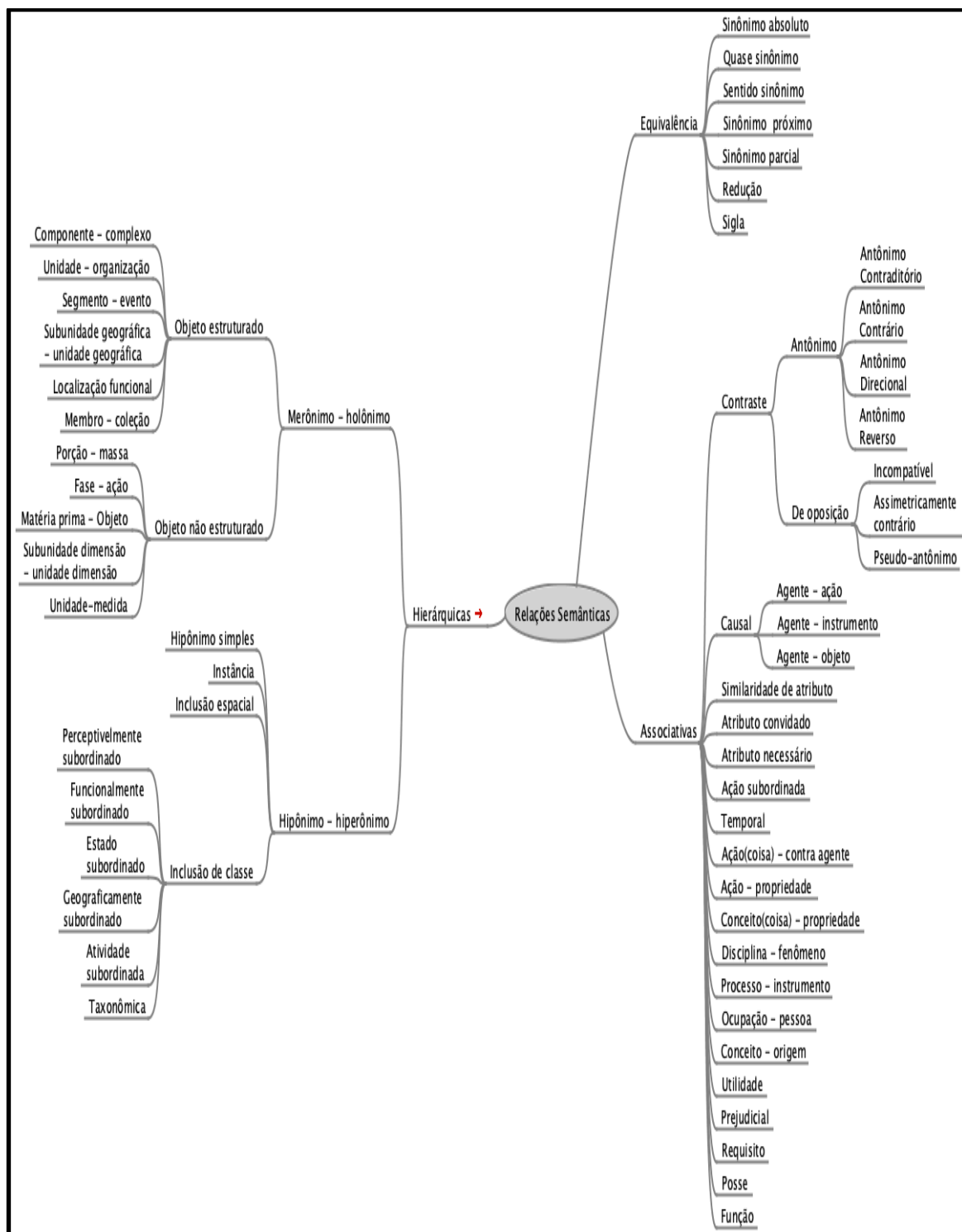
INSTITUIÇÕES

30 Embrapa

temTraduçãoEN (EQ-EN) EMBRAPA
temNomeInstitucional (UP EQ-INS) Empresa Brasileira de Pesquisa Agropecuária
éParteDe (TGP) INSTITUIÇÕES
éParteDe (TGP) INSTITUIÇÕES DE PESQUISA
temParte (TEP) UNIDADES CENTRAIS
temParte (TEP) UNIDADES DESCENTRALIZADAS
temParte (TEP) LABORATÓRIOS VIRTUAIS
temParte (TEP) ESCRITÓRIOS INTERNACIONAIS
influencia (TR) POLITICA DE DESENVOLVIMENTO
influencia (TR) PESQUISA AGROPECUÁRIA
temObjetivoOuProcesso (TR)

FONTE: Maculan (2015, p. 318).

ANEXO D - Modelo de relações semânticas para os instrumentos de organização no contexto da Biblioteconomia e Ciência da Informação



FONTE: Maia (2018, p. 70).

ANEXO E - Esquema de Neville para níveis de correspondências entre tesouros

Tipo de compatibilidade	Descrição da compatibilidade	Exemplos
Correspondência exata entre os descritores.	<p>Os descritores que estão no plural ou singular são considerados idênticos.</p> <p>Os vocabulários dos tesouros A (origem) e B devem representar o mesmo conceito, mesmo em idiomas diferentes.</p> <p>Caso forem utilizados os três vocabulários com o mesmo sentido, é necessário apenas inserir o mesmo código para cada um deles, de modo que fique organizado.</p>	<p>Tesouro A (origem): Campos aéreos Código: 0101 0101 = Campos aéreos</p> <p>Tesouro B: Campo de pouso Código: 0101 0101 = Campos aéreos</p> <p>Tesouro C: Airfield Código: 0101 0101 = Airfield</p>
Descritores ou palavras-chave para indicar que possuem conceitos iguais, ou seja, sinônimos.	Os sinônimos podem ser comparados com o mesmo número de código em diferentes tesouros, ou pelos descritores que possuem o mesmo significado, preservando as palavras-chave de origem e inserir o mesmo código para cada um deles.	<p>Tesouro (origem): Estruturas de subsuperfície, USE Estruturas debaixo da terra</p> <p>0201 = Estruturas debaixo da terra</p> <p>Tesouro B: Estruturas subsuperfície (0201), Estruturas subterrâneas, USE Estruturas subsuperfície, Estruturas enterradas, USE Estrutura de subsuperfície</p> <p>0201= Estruturas debaixo da terra</p>
O tesouro de origem possui uma palavra-chave para um conceito que não existe no tesouro de destino.	No tesouro de destino, é necessário somente incluir um novo código, ou a palavra-chave do tesouro de origem, e/ou é incluída como uma nova palavra-chave em outro tesouro, ou nenhuma inclusão de palavras-chave.	<p>Tesouro (origem): Monte de neve Código: 0301</p> <p>Ou 0301 = Monte de neve</p> <p>Tesouro B: Monte de Neve Código: 0301</p> <p>Ou 0301 = Monte de neve</p> <p>OU Nenhuma inclusão</p>

Tipo de compatibilidade	Descrição da compatibilidade	Exemplos
O tesouro de origem possui uma palavra mais específica do que o tesouro de destino.	A palavra-chave do tesouro de origem é mais específica do que a palavra-chave do outro tesouro, mas vale a pena incluir um termo existente que seja mais amplo.	<p>Tesouro (origem): Monte de Neve (0301)</p> <p>0301 = Monte de neve</p> <p>Tesouro B:Neve Monte de neve (0301), USE neve</p> <p>0301 = Monte de neve</p>
Para combinar dois ou mais elementos de um tesouro, sendo que o de origem utiliza apenas uma palavra-chave para indicar um conceito e o tesouro de destino utiliza um conjunto de duas ou mais palavras-chave para indicar esse mesmo conceito.	Devem-se combinar duas ou mais palavras de dois ou mais tesouros para serem equivalentes.	<p>Tesouro (origem): Gerenciamento de projetos</p> <p>Tesouro B:Gerenciamento de projetos (0401), USE</p> <p>Gerenciamento + Projetos</p> <p>0401 = Gerenciamento de projetos</p>
A combinação de duas ou mais palavras-chave no tesouro de destino forma a palavra-chave do tesouro de origem.	A palavra-chave do tesouro de origem corresponde à combinação de uma ou mais palavras-chave do tesouro de destino.	<p>Tesouro (origem):Termômetro</p> <p>Tesouro B:Termômetro (0503), USE Temperatura + Medição + instrumento</p> <p>0503 = Termômetro</p>
É realizada a distinção de homônimos pelo tesouro de origem, mas não pelo de destino.	Deve-se fazer a inclusão de palavras-chave no tesouro de destino para a reconciliação com o tesouro de destino.	<p>Tesouro (origem):Tanques (veículos de combate) – (0601) Tanques (containers) – (0602)</p> <p>Tesouro B:Tanques, inclusão (combate 0601), (containers - 0602)</p> <p>0601 = Tanques (veículos de combate) 0602 = Tanques (containers)</p>
As palavras-chave do tesouro de origem são separadas para diferenciar termos utilizados em sentidos diferentes, mas isso não acontece no tesouro de destino.	As palavras-chave do tesouro reconciliado (destino) devem ter os mesmos sentidos do tesouro de origem.	<p>Tesouro (origem): Alumínio (material) (0701) Estruturas de alumínio (0702)</p> <p>Tesouro B: Alumínio (material 0701) e (estruturas 0702), OBS: Estes foram incluídos (0701 e 0702).</p> <p>0701= Alumínio (material) 0702 = Alumínio (estruturas)</p>

Tipo de compatibilidade	Descrição da compatibilidade	Exemplos
Ocorre quando termos do tesouro de origem não representam claramente um conceito.	Devem-se excluir os termos que não representam um significado claro do tesouro reconciliado, por um termo preferido entre os sinônimos desses termos.	<p>Tesouro (origem):</p> <p>Aditivos (cimento - 0804) Aditivos de cimento, USE Aditivos</p> <p>0804 = Aditivos (cimento)</p> <p>Tesouro B:</p> <p>Aditivos de cimento</p> <p>0804 = Aditivos de cimento</p>
As palavras-chave do tesouro de origem têm sinônimos no tesouro de destino.	O tesouro de origem deve listar todas as combinações utilizadas para a palavra-chave "Alto", que é combinada por um termo preferido e com as palavras-chave do tesouro de destino que são sinônimas.	<p>Tesouro (origem): Alta</p> <p>(Alta + temperatura - 0801)</p> <p>(Alto + edifícios - 0802)</p> <p>(Alta + força - 0803)</p> <p>Alta temperatura (0801), USE Alta + temperatura</p> <p>Edifícios altos (0802), USE Alto + edifícios</p> <p>Alta resistência (0803), USE Alta + força</p> <p>Edifícios de vários andares (0802), USE Alto + edifícios</p> <p>0801 – Alta + temperatura</p> <p>0802 – Alto + edifícios</p> <p>0803 = Alta + força</p> <p>Tesouro B: Alta temperatura (0801)</p> <p>Alta resistência (0803) Edifícios altos, USE vários andares</p>

Tipo de compatibilidade	Descrição da compatibilidade	Exemplos
Utilização de termo do tesouro apenas para palavras-chave com termos que possuem significado em um local específico.	É sugerido reconciliar para termos mais genéricos, por causa de restrição de localização e conhecimento de pequena parte da população.	<p>Tesouro (origem):Suprimentos de classe 1 (1001) Suprimentos de classe 2 (1001)</p> <p>1001 = Suprimentos de classe 1 Suprimentos de classe 2A</p> <p>Tesouro B:Suprimentos militar do EUA (1001)</p> <p>1001 = Suprimentos militar do EUA</p>
É utilizado um sistema de codificação de modo arbitrário para alguns conceitos no tesouro de origem e que devem ser correspondentes no tesouro de destino.	É identificado no tesouro de origem os termos que combinam com o tesouro de destino.	<p>Tesouro (origem): Vigas de mesa larga (1101)</p> <p>1101 = Vigas de mesa larga</p> <p>Tesouro B:Vigas de mesa larga, USE (1101), USE mesa + 4 cadeiras</p> <p>1101 = Vigas de mesa larga</p>

FONTE: Elaborado pelo autor (2021). Adaptado de Neville (1970, p. 313-331, tradução nossa).