

UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE CIÊNCIA DA INFORMAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM GESTÃO & ORGANIZAÇÃO DO
CONHECIMENTO

EVALDO DE OLIVEIRA DA SILVA

**HOMOGENISE: MÉTODO PARA MODELAGEM ONTOLÓGICA DO
CONHECIMENTO EM PESQUISAS QUALI-QUANTI**

Belo Horizonte

2023

EVALDO DE OLIVEIRA DA SILVA

**HOMOGENISE: MÉTODO PARA MODELAGEM ONTOLÓGICA DO
CONHECIMENTO EM PESQUISAS QUALI-QUANTI**

Tese apresentada ao Programa de Pós-Graduação em Gestão & Organização do Conhecimento, Escola de Ciência da Informação da Universidade Federal de Minas Gerais para obtenção do grau de Doutor, área de concentração Ciência da Informação.

Linha de Pesquisa: Gestão e Tecnologia da Informação e Comunicação (GETIC)

Orientador: Dr. Marcello Peixoto Bax
Coorientador: Dr. Paulo Pinheiro da Silva

Belo Horizonte

2023

S586h	<p>Silva, Evaldo de Oliveira da. Homogenise: método para modelagem ontológica do conhecimento em pesquisas quali-quantitativa [recurso eletrônico] : / Evaldo de Oliveira da Silva. - 2023. 1 recurso online (162 f. : il., color.) : pdf.</p> <p>Orientador: Marcello Peixoto Bax. Coorientador: Paulo Pinheiro da Silva</p> <p>Tese (doutorado)– Universidade Federal de Minas Gerais, Escola de Ciência da Informação. Referências: f. 132-141. Apêndice: f. 142-156. Anexo: f. 157-162. Exigência do sistema: Adobe Acrobat Reader.</p> <p>1. Ciência da informação – Teses. 2. Ontologias (recuperação da informação) - Teses. 3. Representação do conhecimento (teoria da informação) - Teses. 4. Saúde mental – Teses. 5. (campo preenchido pela biblioteca) - Teses. I. Bax, Marcello Peixoto. II. Silva, Paulo Pinheiro da. III. Universidade Federal de Minas Gerais. Escola de Ciência da Informação. IV. Título.</p> <p style="text-align: right;">CDU: 025.4.03</p>
-------	---

Ficha catalográfica: Maianna Giselle de Paula – CRB6: 2642

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 - PPGOC

FOLHA DE APROVAÇÃO

HOMOGENISE: MÉTODO PARA MODELAGEM ONTOLÓGICA DO CONHECIMENTO EM PESQUISAS QUALI-QUANTI

IVALDO DE OLIVEIRA DA SILVA

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 Gestão e Tecnologia da Informação e Comunicação.

Aprovada em 11 de dezembro de 2023, por videoconferência, pela banca constituída pelos membros:

Prof(a). Marcello Peixoto Bax (Orientador)
ECI/UFMG

Prof(a). Paulo Pinheiro da Silva (Coorientador)
Instituto Piaget - Portugal

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

Prof(a). Henrique Oliveira Santos
RPI

Prof(a). Maria das Graças da Silva Teixeira
UFES

Prof(a). Patricia Nascimento Silva
ECI/UFMG

Prof(a). Frederico Cesar Mafra Pereira
ECI/UFMG

Belo Horizonte, 11 de dezembro de 2023.



Documento assinado eletronicamente por **Marcello Peixoto Bax, Professor(a)**, em 18/01/2024, às 19:18, 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 **Henrique Oliveira Santos, Usuário Externo**, em 24/01/2024, às 16:56, 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 **Frederico Cesar Mafra Pereira, Professor do Magistério Superior**, em 25/01/2024, às 10:54, 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 das Graças da Silva Teixeira, Usuário Externo**, em 25/01/2024, às 12:26, 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 **Patricia Nascimento Silva, Professora do Magistério Superior**, em 26/01/2024, às 10:39, 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 **Gercina Angela de Lima, Professora do Magistério Superior**, em 02/02/2024, às 16:18, 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 **Paulo Pinheiro da Silva, Usuário Externo**, em 16/02/2024, 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 **2886189** e o código CRC **A8FECCAC**.



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

ATA DA DEFESA DE TESE DO ALUNO

EVALDO DE OLIVEIRA DA SILVA

Realizou-se, no dia 11 de dezembro de 2023, às 14:00 horas, por videoconferência, da Universidade Federal de Minas Gerais, a defesa de tese, intitulada HOMOGENiSE: MÉTODO PARA MODELAGEM ONTOLÓGICA DO CONHECIMENTO EM PESQUISAS QUALI-QUANTI, apresentada por EVALDO DE OLIVEIRA DA SILVA, número de registro 2020659942, graduado no curso de TECNOLOGIA EM INFORMÁTICA EMPRESARIAL, como requisito parcial para a obtenção do grau de Doutor em GESTÃO E ORGANIZAÇÃO DO CONHECIMENTO, à seguinte Comissão Examinadora: Prof(a). Marcello Peixoto Bax - ECI/UFMG (Orientador), Prof(a). Paulo Pinheiro da Silva - Instituto Piaget – Portugal (Coorientador), Prof(a). Gercina Ângela de Lima - ECI/UFMG, Prof(a). Henrique Oliveira Santos - RPI, Prof(a). Maria das Graças da Silva Teixeira - UFES, Prof(a). Patricia Nascimento Silva - ECI/UFMG, Prof(a). Frederico Cesar Mafra Pereira - ECI/UFMG.

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, 11 de dezembro de 2023.

Assinatura dos membros da banca examinadora:



Documento assinado eletronicamente por **Marcello Peixoto Bax, Professor(a)**, em 18/01/2024, às 19:17, 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 **Henrique Oliveira Santos, Usuário Externo**, em 24/01/2024, às 16:55, 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 **Frederico Cesar Mafra Pereira, Professor do Magistério Superior**, em 25/01/2024, às 10:54, 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 das Graças da Silva Teixeira, Usuário Externo**, em 25/01/2024, às 12:27, 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 **Patricia Nascimento Silva, Professora do Magistério Superior**, em 26/01/2024, às 10:39, 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 **Gercina Angela de Lima, Professora do Magistério Superior**, em 02/02/2024, às 16:18, 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 **Paulo Pinheiro da Silva, Usuário Externo**, em 16/02/2024, 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 **2886160** e o código CRC **2AAB7E0B**.

AGRADECIMENTOS

Gratidão a Deus, sempre. Por meio das minhas preces, devo a Ele a determinação que tive para experienciar este momento.

À minha família, especialmente aos meus pais pelo incentivo em vários momentos desta jornada. À minha esposa Víviam por entender as dificuldades inerentes ao desenvolvimento deste trabalho e, por várias vezes, renunciar ao nosso tempo juntos. Ao meu filho Guilherme — minha motivação em buscar sentido na vida.

Ao orientador, Prof. Dr. Marcello Peixoto Bax, por ter acreditado no meu trabalho, pela acolhida, disponibilidade, paciência e ensinamentos esclarecedores que fundamentaram a minha evolução no Doutorado. Ao Prof. Dr. Paulo Pinheiro pelo conhecimento transmitido em várias reuniões e pelas parcerias proporcionadas com diferentes grupos de pesquisa.

Às profissionais, Luciana Januzzi e Thaís Carvalho, que me ajudaram a modelar o conhecimento em saúde mental na primeira etapa deste trabalho.

Ao Centro Universitário UniAcademia de Juiz de Fora (MG), por ter fomentado o projeto aplicado nesta pesquisa com a cooperação de professores e alunos do curso de Psicologia. À coordenadora do curso de Psicologia, Adriana Ventura, por ter me colocado em contato com os professores e ter disponibilizado a infraestrutura da Clínica Escola de Psicologia.

À professora Auxiliatrice Caneschi Badaró pelos conhecimentos transmitidos acerca do domínio em que esta pesquisa foi aplicada. Aos alunos do curso de Psicologia, Eduarda Novais Falcão, Raíssa Barros Dominato, Rubens Amorim Souto Neto, João Vitor Carvalho Teixeira, João Gabriel Garcia Paiva, Natália Maria Silva Costa, Manuela Mayworm Jens e Samuel Henrique Reis do Nascimento, que se dedicaram a apoiar o projeto de pesquisa sob a coordenação da professora Auxiliatrice.

RESUMO

A análise quali-quantitativa é uma técnica de pesquisa que usa dados qualitativos e quantitativos, havendo a necessidade de integrá-los sob os mesmos objetos de estudo. Utilizar diferentes formas de representação desses dados, permite explicitar seus relacionamentos mais complexos e assimilar melhor as informações da pesquisa. O conhecimento sobre os dados contribui para o avanço da pesquisa exigindo que seja explicitado. A Compreensão e a Preparação são atividades iniciais do fluxo de processo de uma pesquisa quali-quantitativa, que podem usar técnicas para detalhar conceitos e relações entre os dados. Neste contexto, ontologias são comumente utilizadas para descrever o conhecimento embutido em coleções de dados e podem facilitar extrair informações, desambiguar termos, formalizar e preservar a semântica das variáveis. Desta forma a seguinte questão de pesquisa é apresentada: “Como compreender e preparar dados usando ontologias para representar o conhecimento a fim de aprimorar pesquisas do tipo quali-quantitativo?”. Para responder à questão, propõe-se um método sistemático baseado na anotação semântica de dados, a fim de facilitar que os dados, informações e conhecimentos gerados pelas pesquisas, sejam explorados por meio de grafos. Uma técnica de dicionarização semântica de dados foi utilizada para gerar grafos a partir de ontologias, metadados e dados, permitindo que especialistas de domínio explorem as informações e os conhecimentos investigados explicitando-os a partir dos dados. Os grafos organizam dados anotados semanticamente para explorar e filtrar informações a fim de responder questões de interesse das pesquisas. Eles facilitam a visualização gráfica dos dados e têm potencial para aprimorar as análises quali-quantitativas. Um estudo de caso na área de Psicologia busca ilustrar e validar o método proposto em uma pesquisa quali-quantitativa que explora dados psicoterápicos. Como contribuição, mostra-se como o uso de grafos aprimoram as análises quali-quantitativas a partir da combinação de dados quantitativos e qualitativos anotados semanticamente e integrados em um mesmo grafo. Além disso, o experimento com o método permitiu criar uma ontologia de domínio na área de pesquisas psicoterápicas. Finalmente, ao executar o método, apontou-se diferentes dificuldades que indicam direções precisas para a sua evolução.

Palavras-chave: Ontologia; Anotação semântica; Grafo de Conhecimento; Análise quali-quantitativa; Saúde mental.

ABSTRACT

Quali-quant analysis is a research technique that uses qualitative and quantitative data, with the need to integrate them under the same objects of study. Using different forms of representation of this data makes it possible to make more complex relationships clear and better assimilate the research information. The knowledge about the data needs to be described. Data Understanding and Preparation are initial activities in the process flow of quali-quant research, which can use techniques to detail knowledge. In this context, Ontologies are commonly used to describe the knowledge embedded in data collections and can facilitate extracting information, disambiguating terms, and formalizing and preserving the semantics of variables. Thus: "How to understand and prepare data using ontologies to represent knowledge to improve quali-quant research?" To answer the question, a systematic method based on semantic data annotation is proposed to explore data, information, and knowledge generated by research through graphs. A semantic data dictionary technique generated graphs from ontologies, metadata, and data. It allowed domain experts to explore the information and knowledge investigated by explaining it from the data. Graphs organize semantically annotated data to explore and filter information to answer research interest questions. They facilitate the graphical visualization of data and have the potential to enhance quali-quant analyses. A case study in the Psychology research area illustrates and validates the proposed method in quali-quant research that explores psychotherapeutic data. As a contribution, it is shown how the manipulation of graphs improves quali-quant analysis by merging quantitative and qualitative data semantically annotated and integrated into the same graph. Furthermore, the experiment with the method created a new domain ontology in psychotherapy research. Finally, when executing the method, different challenges were pointed out that specifically indicate the needs for its evolution.

Keywords: Ontology; Semantic annotation; Knowledge Graph; Quali-quant analysis; Mental health.

LISTA DE FIGURAS

FIGURA 1 - COLETA SIMULTÂNEA DE DADOS QUANTITATIVOS E QUALITATIVOS.....	19
FIGURA 2 - TIPOS DE ONTOLOGIAS.....	27
FIGURA 3 - A TAXONOMIA DA UFO.....	34
FIGURA 4 - REPRESENTAÇÃO DE UMA TRIPLA.....	35
FIGURA 5 - EXEMPLO DE GC CONSTITUÍDO A PARTIR DE FONTES DE DADOS DISTINTAS.....	36
FIGURA 6 - REPRESENTAÇÃO DA ANOTAÇÃO DE DADOS FEITA POR SDD.....	41
FIGURA 7-AÇÕES PARA ESTABELEECER A GT.....	45
FIGURA 8 - ANÁLISE TEMÁTICA DA CONTAMINAÇÃO POR MRSA.....	48
FIGURA 9 - PROJETO DE PESQUISA CONVERGENTE PARA ANÁLISE QUALI-QUANTI.....	53
FIGURA 10 - PROCESSO DE ESTRATIFICAÇÃO DE RISCOS.....	61
FIGURA 11 - CICLO REGULADOR DE WIERINGA (2009).....	65
FIGURA 12 - ESTRUTURA ANINHADA DO PROBLEMA DE ACORDO COM A DSR DE WIERINGA.....	67
FIGURA 13 - ETAPAS DO MÉTODO ODIN.....	72
FIGURA 14 - ATIVIDADES PARA ELABORAÇÃO DA ETAPA DE CONCEPÇÃO E DESENVOLVIMENTO.....	75
FIGURA 15 - ETAPAS DO MÉTODO HOMOGENISE.....	79
FIGURA 16 - DIAGRAMA DA ONTRISCAL COM CONCEITOS DA CONCEITUALIZAÇÃO DO CASO.....	90
FIGURA 17 - DIAGRAMA DA ONTRISCAL COM CONCEITOS DE AÇÃO TERAPÊUTICA E TIPOS DE QUESTIONÁRIOS.....	91
FIGURA 18 – GRAFO DE CONHECIMENTO DA RESPOSTA DA QC 1 (A).....	107
FIGURA 19 – GRAFO DE CONHECIMENTO DA RESPOSTA DA QC 1 (B).....	107
FIGURA 20 – DADOS QUANTITATIVOS DA RESPOSTA DA QC 2.....	108
FIGURA 21 – GRAFO DE CONHECIMENTO DA RESPOSTA DA QC 3.....	109
FIGURA 22 – GRAFO DE CONHECIMENTO DA RESPOSTA DA QC 4.....	110
FIGURA 23 – GRAFO DE CONHECIMENTO DA RESPOSTA DA QC 5.....	111
FIGURA 24 – GRAFO DE CONHECIMENTO DA RESPOSTA DA QC 6.....	111
FIGURA 25 – GRAFO DE CONHECIMENTO DA RESPOSTA DA QC 7.....	112
FIGURA 26 – DADOS QUANTITATIVOS DA RESPOSTA DA QC 8.....	113
FIGURA 27 – DADOS QUANTITATIVOS DA RESPOSTA DA QC 9.....	113
FIGURA 28 – DADOS QUANTITATIVOS DA RESPOSTA DA QC 10.....	114
FIGURA 29 - MATRIZ DE RASTREABILIDADE.....	115
FIGURA 30 - CLASSES DA ONTRISCAL (A).....	117
FIGURA 31 - CLASSES DA ONTRISCAL (B).....	117
FIGURA 32 - CLASSES DA ONTRISCAL (C).....	117
FIGURA 33 - CLASSES DA ONTRISCAL (D).....	117
FIGURA 34 - CLASSES DA ONTRISCAL (E).....	117
FIGURA 35 - CLASSES DA ONTRISCAL (F).....	117
FIGURA 36 - MÉTIRCAS DA ONTRISCAL NO BIOPORTAL.....	118

FIGURA 37 - GC DOS GRUPOS DE QUESTÕES RESPONDIDOS PELO PACIENTE.....	120
FIGURA 38 - GC COM A FAIXA ETÁRIA POR PACIENTE.	121
FIGURA 39 - GC DOS FATORES QUE INFLUENCIAM OS NÍVEIS DE CUIDADOS.....	121
FIGURA 40 - INFERÊNCIA DO ENCAMINHAMENTO PARA NÍVEL TERCIÁRIO.....	123
FIGURA 41 - INFERÊNCIA DE CUIDADOS PARA O NÍVEL TERCIÁRIO.	123

LISTA DE QUADROS

QUADRO 1 - ESPECIFICAÇÃO DO INFOSHEET	38
QUADRO 2 - ESPECIFICAÇÃO DO <i>DICTIONARY MAPPING</i>	38
QUADRO 3 - ESPECIFICAÇÃO DO <i>CODEBOOK</i>	39
QUADRO 4 - ESPECIFICAÇÃO DO TIMELINE	40
QUADRO 5 FASES PARA ANÁLISE TEMÁTICA.....	46
QUADRO 6 - CONJUNTO DE DADOS HIPOTÉTICO DA ESTRATIFICAÇÃO DE RISCO EM ANSIEDADE .	49
QUADRO 7 - CLASSIFICAÇÃO DOS TIPOS DE VARIÁVEIS	50
QUADRO 8 - TRABALHOS CORRELATOS.....	57
QUADRO 9 - RESUMO DOS TRABALHOS CORRELATOS	60
QUADRO 10 - CRONOGRAMA PARA EXECUÇÃO DO ESTUDO DE CASO	84
QUADRO 11- CONCEITOS E PROPRIEDADES EXTRAÍDAS DO DETALHAMENTO DO MINIMUNDO	87
QUADRO 12 - RESTRIÇÕES ENTRE AS CLASSES DE CONCEITOS DETALHADOS DO MINIMUNDO....	88
QUADRO 13 - PREENCHIMENTO DO <i>INFOSHEET</i>	93
QUADRO 14 - ESPECIFICAÇÃO DO DM PARA DADOS EXPLÍCITOS E IMPLÍCITOS	93
QUADRO 15 - ESPECIFICAÇÃO DO DM PARA EXPLICITAÇÃO DOS DADOS	94
QUADRO 16 – <i>CODEBOOK</i>	94
QUADRO 17 - DATASET COM DADOS DE EXEMPLO DA ESTRATIFICAÇÃO DE RISCO.	94
QUADRO 18 - GRAFO RDF NO FORMATO TURTLE	95
QUADRO 19 - ANOTAÇÃO SEMÂNTICA PARA EXPLICITAR DADOS QUALITATIVOS DO SDD QUALI. .	96
QUADRO 20 – EXPLICITAÇÃO DOS DADOS QUALITATIVOS DO SDD QUALI.	96
QUADRO 21 - ANOTAÇÃO SEMÂNTICA PARA EXPLICITAR DADOS QUALITATIVOS DO SDD QUALI DE DOMÍNIO	97
QUADRO 22 – EXPLICITAÇÃO DOS DADOS QUALITATIVOS DO SDD QUALI DO DOMÍNIO.	97
QUADRO 23 - ANOTAÇÃO SEMÂNTICA DOS DADOS QUANTITATIVOS DO QUESTIONÁRIO INICIAL. .	98
QUADRO 24 – EXPLICITAÇÃO DOS DADOS QUANTITATIVOS DO QUESTIONÁRIO INICIAL.....	98
QUADRO 25 - ANOTAÇÃO SEMÂNTICA DOS DADOS QUANTITATIVOS DO QUESTIONÁRIO FINAL....	98
QUADRO 26 - EXPLICITAÇÃO DOS DADOS QUANTITATIVOS DO QUESTIONÁRIO FINAL.....	99

LISTA DE ABREVIATURAS

AT	-	Análise Temática
CAQDAS	-	<i>Computer Aided Qualitative Data Analysis Software</i>
CID	-	Classificação Internacional de Doenças
CSV	-	<i>Comma-Separated Values</i>
DM	-	<i>Dictionary Mapping</i>
DSM	-	<i>Diagnostic and Statistical Manual</i>
DSR	-	<i>Design Science Research</i>
GBD	-	<i>Global Burden Disease</i>
GC	-	Grafo de Conhecimento
GT	-	<i>Grounded Theory</i>
IA	-	Inteligência Artificial
KO	-	<i>Knowledge Organization</i>
KR	-	<i>Knowledge Representation</i>
OMS	-	Organização Mundial da Saúde
OWL	-	<i>Ontology Web Language</i>
QERSM	-	Questionário de Estratificação de Risco em Saúde Mental
RDF	-	<i>Resource Description Framework</i>
REFI-QDA Software	-	<i>Rotterdam Exchange Format Initiative / Qualitative Data Analysis Software</i>
SDD	-	<i>Semantic Data Dictionary</i>
SESA	-	Secretaria de Saúde do Estado do Paraná

- SPARQL - *SPARQL Protocol and RDF Query Language*
- SWRL - *Semantic Web Rule Language*
- TCC - *Terapia Cognitivo-Comportamental*
- UFO - *Unified Foundational Ontology*

SUMÁRIO

1 INTRODUÇÃO	18
1.1 Problema de pesquisa	21
1.2 Objetivos	22
1.3 Justificativa e relevância	23
1.4 Contribuições Previstas	24
1.5 Organização geral da tese	25
2 REFERENCIAL TEÓRICO.....	26
2.1 Ontologias	26
2.1.1 Ontologias em Saúde Mental	27
2.1.2 Abordagens para modelagem conceitual	28
2.1.3 Engenharia de Ontologias	29
2.1.4 Ontologia de alto nível	31
2.2 Grafos de Conhecimento	34
2.3. Anotação de dados com <i>Semantic Data Dictionaries</i>	36
2.3.1 <i>InfoSheet</i>	37
2.3.2 <i>Dictionary Mapping (DM)</i>	38
2.3.3 <i>CodeBook</i>	39
2.3.4 <i>TimeLine</i>	39
2.3.5 <i>Script</i> para geração de grafo de conhecimento	40
2.4 Análise qualitativa	41
2.4.1 Meios de Representação	42
2.4.2. Codificação	42
2.4.3. <i>Grounded Theory</i>	44
2.4.4 Análise Temática	45
2.5 Análise quantitativa	48
2.6 Análise quali-quantitativa	51
2.7 Desafios em pesquisas qualitativas, quantitativas e quali-quantitativa	53
2.8 Trabalhos Correlatos	56
2.9 O Domínio do Estudo de Caso - Aplicação de Estratificação de Riscos a Dados Psicoterápicos	60
3 METODOLOGIA	63
3.1 Uso da <i>Design Science Research</i> como metodologia de pesquisa	63
3.2 Aplicação da DSR ao problema de pesquisa	66
4 O MÉTODO HOMOGENÍSE	70
4.1. Descrever MiniMundo e Identificar Problemas	73
4.2 Definir Objetivos	74
4.3 Concepção e Desenvolvimento	74
4.4 Demonstração	77
4.5 Avaliação	77
4.6 Divulgação	78
4.7 Controle	78
5 APLICAÇÃO DO HOMOGENÍSE EM UM ESTUDO DE CASO COM DADOS PSICOTERÁPICOS	80
5.1 Descrever MiniMundo e Identificar Problemas	81
5.2 Definir Objetivos	83
5.3 Concepção e Desenvolvimento	84
5.4 Demonstração	99
5.4.1. Novas categorias de dados qualitativos baseadas em GCs	102
5.5 Avaliação	103
5.6 Divulgação	104
5.7 Controle	114

6 ONTOLOGIA ONTRISCAL	116
6.1 Ontologia da Estratificação de Riscos em Saúde Mental	116
6.2 Extração de Conhecimento em saúde mental	118
6.3 Inferência de conhecimento	122
7 CONSIDERAÇÕES FINAIS	125
7.1 Contribuições Atingidas	126
7.1.1. Publicações relacionadas ao tema	128
7.2 Dificuldades ao executar o método HOMOGENiSE	129
7.3 Trabalhos Futuros	130
REFERÊNCIAS	132
APÊNDICE A	142
APÊNDICE B	143
ANEXO 1	157
ANEXO 2	160
ANEXO 3	161
ANEXO 4	162

1 INTRODUÇÃO

Diferentes abordagens de pesquisa abrangem métodos, técnicas e planejamento para o avanço de estudos científicos, envolvendo atividades que detalham a coleta, análise e interpretação de dados. Cientistas baseiam-se na motivação, questão de pesquisa, objetos do estudo, estudos de caso e em experiências pessoais, que norteiam a escolha da abordagem de pesquisa mais apropriada (CRESWELL; CRESWELL, 2018). Três abordagens de tratamento de dados são comumente utilizadas em pesquisas científicas: qualitativa, quantitativa ou qualitativa e quantitativa (quali-quantitativa ou método misto) (BAGDONIENĖ; ZEMBLYTĖ, 2005; BAYER; SOUZA, 2007; DAWADI, SHRESTHA; GIRI, 2021; FREITAS; JANISSEK, 2000; FREITAS; MOSCAROLA, 2000).

De acordo com Pathak, Jena e Kalra (2013) em pesquisas científicas na área clínica, o método qualitativo é usado para compreender as crenças, experiências, atitudes, comportamentos e interações das pessoas. Este método é aplicado para analisar dados de grupos focais, entrevistas, observação ou documentos históricos. Os dados coletados na pesquisa qualitativa são termos da linguagem categorizados e organizados para identificar padrões que apoiam as respostas para o problema de pesquisa (PATTON, 1999; POLKINGHORNE, 1995). Esse tipo de dados não possui estrutura que permita o seu tratamento quantitativo, obscurecendo a organização e interpretação dos seus significados, podendo enviesar a observação (FRIESE; SYBING, 2023; GOERTZEN, 2017). Finalmente, a sua visualização sob várias perspectivas é um desafio para expandir o pensamento multidimensional a partir de diferentes variáveis.

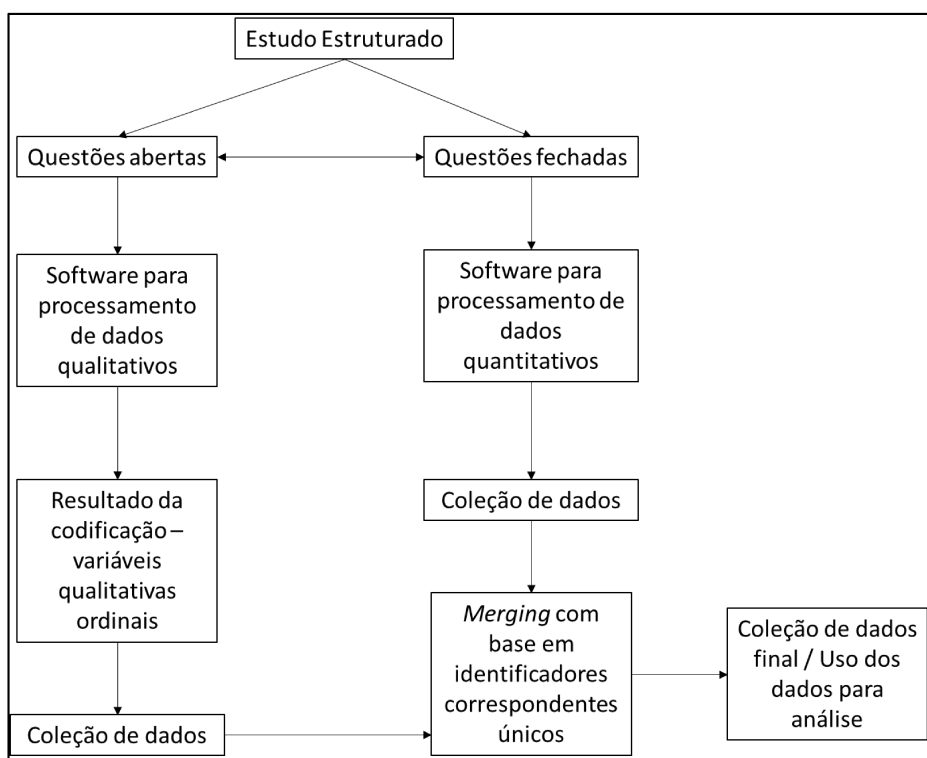
Na pesquisa quantitativa empregam-se técnicas para coletar e analisar dados numéricos para responder à questão de pesquisa, verificar hipóteses, manipular ou mensurar variáveis (CRAMER, 2003). É importante conhecer a proveniência dos dados para sustentar uma tomada de decisão mais eficaz (GOERTZEN, 2017). De forma semelhante à pesquisa qualitativa, na pesquisa quantitativa os especialistas de domínio definem os significados das variáveis.

As pesquisas com o uso de métodos mistos usam a compreensão e preparação dos dados produzidos em “trilhas” distintas de pesquisa (qualitativa e quantitativa). A representação conjunta dos dados produzidos por métodos mistos é um desafio. Diagramas, grafos e outras apresentações auxiliam na captura do conhecimento (SOMEKH; LEWIN, 2005). Ao preparar os dados, as atividades de transformação modificam a forma dos dados, tal como discutido em Driscoll *et al.* (2007).

A Figura 1 apresenta etapas de um projeto de coleta simultânea de dados quantitativos e qualitativos (DRISCOLL *et al.*, 2007). Questões abertas e fechadas são originadas de um

estudo estruturado, e, em seguida, ferramentas de software processam os dados. Os dados qualitativos são oriundos de questões abertas, já os quantitativos são coletados por questões fechadas. O software para processamento de dados qualitativos deriva variáveis qualitativas ordinais a partir da tarefa de codificação (ou *coding*). Ao final, a combinação (*merging*) com base em identificadores únicos constitui uma coleção de dados que integra os dados, quali e quanti, para a sua análise conjunta.

FIGURA 1 - Coleta simultânea de dados quantitativos e qualitativos



Fonte: Adaptado de: (DRISCOLL *et al.* (2007).

Creswell e Creswell (2018) se referem às questões abertas de um questionário de pesquisa como geradoras de dados não-estruturados que podem ser processados por *Computer Aided Qualitative Data Analysis Software* (CAQDAS). Por outro lado, a análise quantitativa é conduzida por meio da coleta de dados de questões fechadas, podendo ser realizada a partir de questionários, pesquisas ou dados estatísticos preexistentes (LABAREE, 2013).

Os autores mencionados anteriormente discutem métodos de pesquisa bem como a compreensão e a preparação dos diferentes tipos de dados e os desafios para organizá-los e transformá-los em informação e conhecimento. A Ciência da Informação oferece técnicas que suportam a organização como a KO (*Knowledge Organization*) e a representação do

conhecimento, a KR (*Knowledge Representation*) (QIN, 2020). Os sistemas de KO são projetados para normalizar vocabulários e sistemas de classificação para organizar e recuperar dados e informações em diferentes tipos de aplicações. A KO enfatiza a representação de conceitos e suas relações em um determinado domínio, construindo vocabulários controlados, *thesauri* e ontologias. A KR é usada, sobretudo, pela Inteligência Artificial (IA) e produz um conjunto de declarações que expressam fatos, relações e condições em linguagens ou esquemas formais sobre os quais o raciocínio pode ser executado para determinar ações ou chegar a conclusões. O componente de raciocínio da KR é talvez a diferença mais importante entre a KR e a KO (QIN, 2020).

Ontologias são comumente utilizadas para organizar o conhecimento sobre coleções de dados (GUARINO, 1998; RASHID *et al.* 2017). Embora possam representar o conhecimento sobre diferentes domínios, somente o modelo ontológico não resolve o problema da organização dos dados. É necessário o uso de abordagens para preparar os dados e que utilizem técnicas de anotação semântica deles, usando, como metadados, os conceitos presentes nas ontologias (RASHID *et al.*, 2017).

Estudos aprimoram a representação e a organização do conhecimento de dados produzidos por pesquisas qualitativas (HOCKER *et al.*, 2021b; MEISTER, 2017; POPPING, 2007). A ontologia QualiCO¹, por exemplo, especifica esquemas de metadados que anotam os dados bem como os documentos produzidos em uma pesquisa qualitativa (HOCKER *et al.*, 2021a). Outro exemplo é o padrão REFI-QDA (*Rotterdam Exchange Format Initiative - Qualitative Data Analysis Software*) que detalha um conjunto de classes de objetos utilizadas para definir dados produzidos em projetos de análise qualitativa. É um padrão aberto e diferentes CAQDAS podem implementá-lo, permitindo que aplicações de software possam “conversar” entre si (REFI-QDA, 2023). O principal objetivo do REFI-QDA é tornar interoperáveis os dados qualitativos entre CAQDAS, abrindo caminho para que os usuários troquem informações sobre os dados processados (EVERS, 2018). Além da QualiCO e do padrão REFI-QDA, grafos de conhecimento têm sido aplicados para organizar o conhecimento dos textos analisados, apoiando na representação dos fatos (dados) gerados pela análise qualitativa (POPPING, 2007).

Outros esforços usam metadados para compartilhar conhecimento com tecnologias semânticas e reutilizar análises qualitativas publicadas em artigos acadêmicos (MEISTER, 2017). A técnica *Semantic Data Dictionnaire* (ou SDD) proposta por Rashid *et al.* (2017) tem propósito semelhante, pois estabelece um conjunto de templates de metadados usando conceitos de ontologias, para que dados possam ser anotados semanticamente. Conforme

¹ <https://github.com/julianhocker/Quali-Codes-Ontology>

discutido por Rashid *et al.* (2020) um SDD é uma coleção de documentos, onde cada um desempenha papel na criação de uma representação de conhecimento, formando um dicionário semântico de dados.

Desse modo, esta pesquisa propõe anotar semanticamente os dados de abordagens mistas, qualitativos e quantitativos, integrando-os por meio de grafo de conhecimento (GC). Com o objetivo de aprimorar as análises quali-quantitativas, apresenta-se o método HOMOGENiSE (*metHod for Ontological Modeling Of knowledGE iN quali-quantitative reSEarch*), com etapas definidas para compreender e preparar dados em métodos mistos. Dentre as técnicas utilizadas pelo HOMOGENiSE, a modelagem ontológica e a anotação semântica, visam integrar a QualiCO, o padrão REFI-QDA e SDDs a fim de gerar GCs que possam aprimorar as análises quali-quantitativas.

O método especifica etapas para integrar dados quali-quantitativos por meio de GCs, e segue a ideia da coleta simultânea de dados quantitativos e qualitativos apresentada em Driscoll *et al.* (2007), onde a mesclagem dos dados será feita por meio de um identificador único (“ID”) dos objetos do domínio. Esta tese aplica as etapas do método em um estudo de caso cujo objetivo é realizar análises quali-quantitativas de dados psicoterápicos. Os tipos de dados provenientes de tratamentos psicoterápicos, usados como domínio de aplicação do estudo de caso, e integrados em um único grafo de conhecimento, aprimoram as análises quali-quantitativas em relação às realizadas sem o apoio do método. Esta é a principal hipótese de pesquisa.

No âmbito deste trabalho, GCs populados com dados quali-quantitativos são gerados a partir de fases que compõem o método. De acordo com Berthold *et al.* (2010), a compreensão e preparação de dados são fases iniciais do processo de análise de dados científicos. Este trabalho propõe abordar técnicas que utilizam a compreensão e a preparação de dados como atividades do método HOMOGENiSE, visando apoiar a representação e explicitação do conhecimento subjacente aos dados quali-quantitativos manipulados por meio de grafo.

1.1 Problema de pesquisa

De acordo com Berthold *et al.* (2010) os procedimentos para análise de dados devem oferecer métodos para desenvolver estudos científicos usando recursos da computação e grande volume de dados. Os autores esclarecem as atividades que contribuem para transformar dados em conhecimento. Abordam a compreensão de dados como atividade que visa entender se existem dados suficientes para responder a diferentes questões de pesquisa. Ela se inicia com a coleta de dados, familiarização e identificação de problemas sobre a qualidade deles, para detectar subconjuntos de interesse da pesquisa científica a fim de

formar hipóteses (DANUBIANU; PENTIUC; TOBOLCEA, 2010). Ontologias auxiliam a compreensão de dados, oferecendo um caminho promissor para resolver problemas de compreensão e extração de informações (BRISSEON; COLLARD, 2008; YAMADA *et al.*, 2020). Elas desambigam termos e preservam a semântica das variáveis do estudo, tornando possível a anotação semântica de dados produzidos por diferentes estudos e pesquisas científicas (BRISSEON; COLLARD, 2008; RASHID *et al.*, 2017). De acordo com Brisson e Collard (2008, p. 6) “ontologias formalizam conceitos de domínio e informações sobre dados”.

De acordo com Paton (2019), a preparação de dados abrange a seleção, limpeza, mapeamento e integração de conjuntos de dados existentes para uma organização de dados adequada para modelagem e análise. Dados devem estar preparados adequadamente para suportar as técnicas de modelagem evitando análises enviesadas (PYLE, 1999).

Este trabalho utiliza a modelagem ontológica como abordagem para compreender o domínio. Ao anotar dados semanticamente, o especialista de domínio usa a ontologia para auxiliar na compreensão de *datasets*, mapeando conceitos e dados. A preparação de dados é executada ao selecionar e integrar *datasets* com o apoio da técnica SDD. Essas atividades, em fases iniciais de um método que antecede a geração de grafos para análise quali-quantitativa, motivam a formulação das duas hipóteses apresentadas a seguir:

- anotar dados semanticamente apresenta-se como abordagem que explicita o conhecimento para análise quali-quantitativa e permite a integração dos dados dos objetos de estudo, auxiliando na compreensão e preparação dos dados;
- dados quali-quantitativos anotados semanticamente por meio de diferentes ontologias e integrados por grafos de conhecimento, aprimoram a análise quali-quantitativa a fim de responder questões relevantes sobre o domínio pesquisado.

Com base na contextualização do tema de pesquisa e no conjunto de hipóteses levantadas, a seguinte questão de pesquisa é apresentada: “Como compreender e preparar dados usando ontologias para representar o conhecimento a fim de aprimorar pesquisas do tipo quali-quantitativo?”. Para responder à questão proposta, desenvolveu-se pesquisa exploratória baseada na literatura e em um estudo de caso.

1.2 Objetivos

Esta seção descreve os objetivos desta pesquisa. Uma vez que a representação conjunta dos dados produzidos por métodos mistos é um desafio, sugere-se o uso de GCs para proporcionar a visualização dos dados sob várias dimensões para aprimorar análises quali-quantitativas. Ontologias estruturam conceitos que podem codificar dados qualitativos e do domínio. A modelagem ontológica em conjunto com a anotação semântica, pode explicitar o

conhecimento dos *datasets* quali-quantitativos, visando prepará-los para gerar GCs.

Deste modo, propõe-se desenvolver e aplicar etapas de um método de modelagem ontológica e anotação semântica para aprimorar a elaboração de análises quali-quantitativas. Além do objetivo geral da pesquisa, os objetivos específicos são:

- modelar o conhecimento sobre Estratificação de Riscos em Ansiedade em psicoterapia, com o apoio de especialistas de domínio, e construir uma ontologia na área;
- utilizar a ontologia para codificar dados qualitativos;
- especificar um método de modelagem ontológica que vise aprimorar análises quali-quantitativas por meio de grafos de conhecimento, gerados com a ontologia criada e dados em psicoterapia;
- aplicar o método especificado, juntamente com especialistas do domínio, para validar as possibilidades de aprimorar análises quali-quantitativas.

O método a ser especificado nesta pesquisa possui etapas detalhadas posteriormente (cf. Capítulo 4), e reúne técnicas utilizadas para alcançar o objetivo geral, tais como: (i) modelagem ontológica para estruturar a codificação de dados qualitativos com o uso de ontologia; (ii) anotação semântica de dados quali-quantitativos; e (iii) geração de GCs com os dados integrados. A anotação semântica dos dados qualitativos será realizada a partir de técnicas de análise qualitativa, com o uso da QualiCO. Classes e propriedades do REFI-QDA são utilizadas com vistas a reutilizar conceitos dos dados qualitativos codificados.

Em síntese, o método utiliza ontologias e técnicas que se complementam para compreender e preparar dados, e representar o conhecimento por meio de GCs. A anotação semântica proposta pelo método, utiliza técnica SDD que sistematiza o processo de anotação dos dados e de geração dos grafos, o qual pode ser executado pelos próprios especialistas de domínio (RASHID *et al.*, 2020). Finalmente, nota-se que o método se situa nas fases iniciais do ciclo de vida do *data science*, que preconiza a organização conceitual na preparação de dados em ciência de dados (MICROSOFT, 2020).

1.3 Justificativa e relevância

A pesquisa relaciona-se com as áreas do Programa de Pós-Graduação em Gestão e Organização do Conhecimento da Universidade Federal de Minas Gerais (PPGGOC/UFMG), mais especificamente com a área de Gestão & Tecnologia da Informação e Comunicação² (GETIC). A GETIC estuda o gerenciamento estratégico da informação e do conhecimento envolvendo sua organização, recuperação e disseminação em contextos digitais e

² Disponível em: <https://ppggoc.eci.ufmg.br/linhas-de-pesquisa/linhas-de-pesquisa/>

organizacionais. Assim, os objetivos da pesquisa estão relacionados a um conjunto de atividades preconizadas pela Ciência da Informação, tais como: organização e representação do conhecimento, modelagem conceitual, ontologias, *web* semântica e metadados.

Espera-se, com este trabalho, enriquecer outras pesquisas que utilizam análise quali-quantitativa de dados. Nesse sentido, entende-se que representar o conhecimento da análise quali-quantitativa permite aprimorar tais análises, independentemente do domínio de aplicação.

Finalmente, cabe ressaltar a relevância do estudo de caso (*cf.* Capítulo 5) no âmbito da saúde mental, pois, de acordo com o GBD (*Global Burden Disease*), dados epidemiológicos e de transtornos mentais devem ser compreendidos, e avaliações comparativas elaboradas para que especialistas possam analisar os fatores de risco associados à saúde mental. O GBD aponta para a necessidade de ter dados mais detalhados nesta área, além do uso de técnicas que garantam a qualidade dos dados para o desenvolvimento dos estudos científicos (WHITEFORD, FERRARI; DEGENHARDT, 2016).

1.4 Contribuições Previstas

Usar ontologias para codificar dados qualitativos, com dicionários semânticos e gerar grafos buscando aprimorar a análise quali-quantitativa é a contribuição principal desta pesquisa. O conhecimento pode ser reutilizado em diferentes estudos a fim de criar bases de dados semânticas para harmonizar variáveis e dados qualitativos. Os GCs populados com dados quali-quantitativos podem ser manipulados por tecnologias específicas para recuperar informações. A recuperação dessas informações possibilita a criação de gráficos, diagramas ou painéis que permitam aprimorar análises quali-quantitativas. Codificar dados qualitativos com ontologias, aprimora a compreensão dos dados e abre caminho para representar o conhecimento, apoiando a aplicação de técnicas de análise qualitativa, tais como, Análise Temática (AT) ou *Grounded Theory* (GT), revisadas nas Seções 2.7.3 e 2.7.4.

Datasets provenientes de domínios distintos podem ser integrados, explicitando-se o conhecimento sobre estes diferentes *datasets*. Por exemplo, o Diabetes Tipo 2 e o Tabagismo são fatores que podem elevar o risco de demência vascular (PASQUIER *et al.*, 2006; RAFFAITIN *et al.*, 2009; OHARA *et al.*, 2011). Percebe-se a possibilidade de analisar dados provenientes de fontes distintas para o tratamento do paciente, podendo existir estruturas heterogêneas e não estarem devidamente documentados. Nesse sentido, o método especificado nesta pesquisa permite representar o conhecimento presente nos dados usando ontologias em um processo incremental.

Finalmente, a modelagem ontológica desenvolvida permitiu a criação da Ontologia ONTRISCAL (Ontologia da Estratificação de Riscos em Saúde Mental). A ONTRISCAL

representa conceitos da codificação de dados qualitativos da estratificação de risco e tratamentos em ansiedade.

1.5 Organização geral da tese

Além do capítulo introdutório, que apresentou o contexto e a questão da pesquisa, os objetivos e as contribuições esperadas, esta tese segue estruturada em 6 capítulos. O Capítulo 2 aprofunda a base teórica que fundamenta os conceitos utilizados para realização da pesquisa, descrevendo aspectos relacionados a ontologias, anotação semântica, grafos de conhecimento, análises qualitativa e quantitativa, desafios encontrados em métodos mistos e trabalhos correlatos.

O Capítulo 3 apresenta a *Design Science Research* (DSR), que ampara metodologicamente este trabalho, decompondo o problema de pesquisa em questões práticas e teóricas mutuamente aninhadas, para atingir o objetivo geral.

O Capítulo 4 traz como contribuição a definição do método HOMOGENiSE, que é especificado por 7 etapas detalhadas como: Descrever MiniMundo e Identificar Problemas; Definir Objetivos; Concepção e Desenvolvimento; Demonstração; Avaliação; Divulgação; e Controle. No Capítulo 5 aplicam-se as etapas especificadas do método por meio de um estudo de caso com dados psicoterápicos e a participação de especialistas de domínio de uma Clínica de Psicologia.

Como resultado da aplicação do método foi criada a ontologia ONTRISCAL . O Capítulo 6 detalha a contribuição da ONTRISCAL, e como novos conhecimentos podem ser extraídos dela.

No Capítulo 7 são abordadas as considerações finais, que evidenciam o cumprimento dos objetivos, dificuldades na execução do método, e as sugestões para trabalhos futuros.

2 REFERENCIAL TEÓRICO

Este capítulo apresenta o levantamento bibliográfico com definições que visam fundamentar teoricamente os temas tratados na pesquisa.

2.1 Ontologias

Uma ontologia é uma especificação da conceituação de um domínio. O termo conceituação diz respeito a uma coleção de conceitos existentes em um determinado domínio e os relacionamentos entre eles, podendo ser uma visão abstrata e simplificada do mundo que se deseja representar (ALMEIDA; BAX, 2003). A conceituação pode ser desenvolvida contendo terminologias e vocabulários, estabelecendo propriedades e permitindo que o conhecimento seja reutilizado, evitando o retrabalho ou a redescoberta de terminologias equivalentes (GUARINO, 1998).

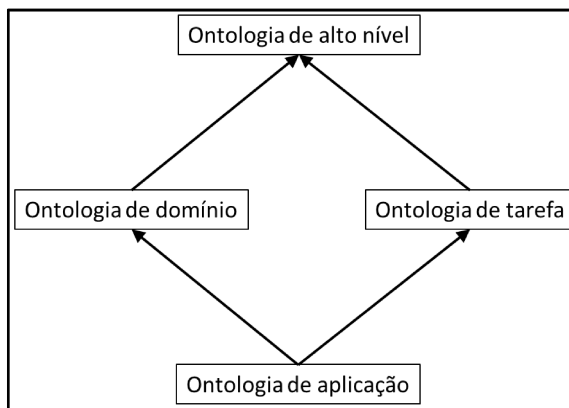
Gruber (1993) define uma ontologia como uma descrição de conceitos e relacionamentos que podem existir por meio de um agente ou uma comunidade de agentes. Esta definição está relacionada com o uso da ontologia como uma coleção de definições, usada, inclusive, na Filosofia. Em diferentes áreas, tais como a Ciência da Informação ou Ciência da Computação, há um consenso no uso de ontologias com o propósito de permitir que o conhecimento seja compartilhado e reutilizado.

Ontologias facilitam o compartilhamento do conhecimento e permitem perguntar e consultar sobre os conceitos e suas relações. Por exemplo, em IA são desenvolvidos agentes de software capazes de acessar a ontologia, sendo que, os conceitos e suas relações podem ser compartilhados permitindo a troca do conhecimento (GRUBER, 1993). De acordo com Guarino (1998, p. 7), as ontologias são divididas em classificações diferentes, que se relacionam segundo os níveis mostrados na Figura 2:

- **Ontologias de alto nível:** também classificadas como ontologias de topo. Incluem classes de conceitos gerais. Não pertencem a um único domínio de aplicação. Esses conceitos são genéricos e podem ser mais especificados por ontologias de domínio ou tarefa.
- **Ontologias de domínio:** representam um determinado domínio, podendo ser reutilizáveis por aplicações do domínio. Os conceitos explicitados no domínio e seus relacionamentos são representados.
- **Ontologias de tarefa:** representam um conjunto de conceitos usados para resolver problemas que podem ser ou não de um mesmo domínio. Inclui conceitos genéricos, podendo ser reaproveitados em diferentes domínios.

- **Ontologias de aplicação:** são criadas a partir de conceitos de um domínio particular, em certos casos, da aplicação de uma tarefa em particular em um domínio específico.

FIGURA 2 - Tipos de Ontologias



Fonte: (GUARINO, 1998).

Dentre os tipos de ontologias definidas, as ontologias de domínio têm sido bastante usadas por pesquisadores da área da saúde mental (BRENAS; SHIN; SHABAN-NEJAD, 2006; CEUSTERS; SMITH, 2010; HASTINGS *et al.*, 2012; BENFARES; IDRISSE; HAMID, 2018; YAMADA *et al.*, 2018). Esses trabalhos serão comentados a seguir.

2.1.1 Ontologias em Saúde Mental

Ceusters e Smith (2010) apresentam uma visão do transtorno mental baseada no realismo ontológico e que segue os princípios incorporados na *Basic Formal Ontology* (BFO) e na aplicação da BFO na *Ontology of General Medical Science* (OGMS). Os autores analisaram afirmações sobre o que é considerado como transtorno mental no contexto da DSM-V. Avaliam se a representação do conhecimento presente nas ontologias BFO e OGMS são adequadas para criar uma ontologia da saúde mental necessária para evitar interpretações erradas sobre um mesmo conceito.

Brenas, Shin e Shaban-Nejad (2019) implementaram uma ontologia formal sobre experiências adversas da infância (ACEs, *Adverse Childhood Experiences*), para integração de dados de estudos científicos na área da saúde mental. Segundo os autores, as ACEs estão relacionadas aos riscos aumentados a partir dos resultados e condições de saúde dos pacientes.

De acordo com Hastings *et al.* (2012) as ontologias podem ser usadas para aplicações sofisticadas de raciocínio automatizado. Os autores descrevem a Ontologia do Funcionamento Mental (MF) e a Ontologia da Doença Mental (MD), duas ontologias baseadas em realismo para a descrição do funcionamento mental humano e das doenças.

Benfares, Idrissi e Hamid (2018) propõem uma arquitetura de sistema de acompanhamento de pacientes, que proporciona o acompanhamento do estado psíquico do paciente. Os autores usam ontologias e sistemas de recomendação para proporcionar aos pacientes o acompanhamento em tempo real. Desse modo, foram desenvolvidas rotinas para recuperar informações de rastreamento dos pacientes com câncer, depressão e transtornos de ansiedade.

Yamada *et al.* (2018) propõem a ontologia OSM (Ontologia de Saúde Mental) desenvolvida para apoiar a integração de dados para saúde mental, a partir de dados mantidos por sistemas em saúde mental.

Rosenberg e Hastings (2020) discutem a importância de estruturar e padronizar a semântica da sintomatologia em saúde mental (ou, a compreensão de sinais e sintomas do paciente). Os autores formalizam uma ontologia usando outras ontologias da área biomédica, tal como a MF. Assim, os autores abordam o esforço de padronização com o método de ontologia aplicada, valendo-se de recursos de código aberto existentes na *Open Biomedical Ontologies* (OBO).

2.1.2 Abordagens para modelagem conceitual

De acordo com Dooley *et al.* (2017) a plataforma *Genomic Epidemiology Entity Mart* (GEEM) é um portal *web* projetado para fornecer aos especialistas de domínio uma interface orientada por ontologia e, assim, examinar padrões de dados em sequências genômicas. O GEEM possui o GenEpiO, um módulo que fornece um conjunto de termos modelados por uma ontologia de aplicação, e que podem ser associados a dados epidemiológicos. A ontologia reutiliza terminologias encontradas na OBO, e tem sido construída usando uma abordagem *bottom-up*, centrada na padronização dos termos necessários para executar os processos relativos aos fluxos de trabalho laboratorial, clínico e epidemiológico.

Abordagens de modelagem *top-down* e *bottom-up* podem ser usadas para a construção de ontologias de domínio. De acordo com Kung, Kung e Gardiner (2012, p. 12) a abordagem *top-down* é definida como:

a obtenção do conhecimento a partir da identificação das entidades, tipos de entidades, e os relacionamentos entre elas. A adoção dessa abordagem, geralmente, inicia-se com um conjunto de requisitos de alto nível, tal como uma narrativa. Esses requisitos descritos iniciam um processo de identificação das entidades necessárias para representar os dados, bem como as propriedades deles e como podem estar organizados em bases de dados.

Em contraste, Kung, Kung e Gardiner (2012, p. 12) ainda definem a abordagem *bottom-up* como: “a aquisição do conhecimento pela compreensão das coleções de dados,

avançando para uma análise inicial de classes de conceitos de nível inferior, como atributos e relacionamentos, e então, procedendo em direção a um modelo conceitual aceitável”.

De acordo com Matteo e Cuel (2005), essas abordagens possuem propósitos específicos para modelar ontologias de domínio:

- **Top-down**: da generalização à especificação do conhecimento. Conceitos complementares podem ser reutilizados de outras ontologias;
- **Bottom-up**: da especificação à generalização do conhecimento. Visa elicitar os conceitos por meio da análise ou processamento de dados estruturados ou não;
- **Middle-out**: inicia-se dos conceitos mais importantes, partindo-se para a sua generalização e especialização.

De acordo com El Ghosh *et al.* (2016) a *middle-out* é uma estratégia que pode combinar as outras duas abordagens em atividades executadas para modelar a ontologia em módulos. Ainda, de acordo com El Gosh *et al.* (2016, p. 2), “a reutilização de outras ontologias que capturam conhecimentos semelhantes ou complementares de ontologias de alto nível, pode contribuir na abordagem *top-down*”. Os autores destacam, ainda, a importância das ontologias de alto nível no reuso do conhecimento. Conforme abordado por Francesconi *et al.* (2010), as abordagens *top-down* e *bottom-up* podem se complementar em projetos que necessitam compreender genericamente o domínio, mas, adicionalmente, explora o conhecimento de diferentes conjuntos de dados para identificar novos conceitos visando, assim, compor uma ontologia.

2.1.3 Engenharia de Ontologias

De acordo com Isotani e Bittencourt (2015) as metodologias para a engenharia de ontologias sistematizam um conjunto de atividades desde a produção até a manutenção de uma ontologia. Consideram atividades de gerenciamento, desenvolvimento (conceitualização ou formalização) e suporte (documentação), presentes no ciclo de vida de uma ontologia. Em geral, o ciclo de vida de uma ontologia possui as seguintes atividades (CALERO, RUIZ e PIATTINI, 2006):

- **Especificação**: início das atividades, apresentação da motivação, seus possíveis usos e os usuários interessados na ontologia.
- **Conceitualização**: organiza e estrutura o conhecimento do domínio sobre o qual se pretende construir a ontologia. As tarefas de aquisição do conhecimento são consideradas aqui;
- **Formalização**: formaliza-se o modelo conceitual, estabelecendo logicamente as relações e axiomas presentes na ontologia;

- **Implementação:** estabelece o uso de alguma linguagem de representação para implementar a ontologia;
- **Manutenção:** com a ontologia implementada em uma linguagem, inicia-se a etapa de manutenção da ontologia.

Além do ciclo de vida mencionado, *softwares* utilizados como editores suportam a implementação de ontologias. Isotani e Bittencourt (2015) apresentam uma comparação entre os seguintes editores: Apollo; OntoStudio; Protégé; TopBraid Composer e; Hozo. Segundo os autores, o Protégé é o mais difundido pela comunidade internacional e possui características como: extensibilidade por meio de *plug-ins*; importação e exportação de modelos por meio dos formatos XML, OWL, RDF e Excel, e; funcionalidades como taxonomia gráfica e trabalho colaborativo.

Dentre as metodologias para engenharia de ontologias, Noy e McGuinness (2001) definem a metodologia 101 (ou *Ontology Development 101*). Esta metodologia é composta por sete etapas que guiam os engenheiros no processo de construção de uma ontologia. Na primeira etapa os engenheiros estabelecem as questões de competência (QC), utilizadas como escopo e base para validar e compreender o propósito da construção da ontologia. Ao aplicar a metodologia 101, as seguintes perguntas norteiam a construção de ontologias:

- Qual o domínio que a ontologia deve cobrir?
- Quais serão os usos da ontologia?
- Quem irá usar e manter a ontologia?
- Quais tipos de questões a ontologia deveria responder?

A segunda etapa está relacionada ao reuso de outras ontologias para apoiar na construção da ontologia de interesse. Após definir quais ontologias poderão ser reusadas, na terceira etapa o ontologista deve enumerar os conceitos que devem estar na ontologia. A etapa seguinte é dedicada à definição das classes e hierarquias de classes. As duas etapas seguintes são para definir as propriedades presentes nas classes, bem como as restrições que cada propriedade e classe devem possuir. Finalmente, é necessário criar as instâncias que materializam os indivíduos representantes dos conceitos presentes na ontologia.

De acordo com Falbo (2024), a SABiO (*Systematic Approach for Building Ontologies*) é uma metodologia usada para desenvolver várias ontologias em diferentes domínios. Ao utilizar a SABiO o ontologista deve considerar cinco processos, a saber: identificar o propósito e requisitos; formalizar a ontologia; *design*; implementar; e testar. Esses processos são suportados por atividades de apoio que permitem controlar a aquisição do conhecimento, documentação, gerência da configuração, avaliação e reuso da ontologia construída.

Costa *et al.* (2023) consideram as QCs como artefatos gerados pela SABiO. Além de servirem de base para identificar os conceitos e as relações que estruturam a ontologia, as QCs estabelecem o controle de qualidade ao validar a ontologia, refletindo como os dados dos usuários podem ser analisados a partir da navegação pelos conceitos presentes na ontologia.

De forma geral, os processos de suporte da SABiO estão presentes durante todas as fases de desenvolvimento de uma ontologia (COSTA *et al.*, 2023; FALBO, 2014). De acordo com Costa *et al.* (2023), no processo de aquisição de conhecimento utilizam-se técnicas convencionais de levantamento de requisitos e de *brainstorming*, podendo reusar conceitualizações já existentes sobre o domínio. O processo de documentação é parte do resultado do desenvolvimento de uma ontologia, devendo ser documentada por algum editor de ontologia. De forma concomitante, o ontologista deve controlar a evolução da ontologia adotando, por exemplo, técnicas de gerenciamento de configuração de *software*. No que se refere à atividade de avaliação, esta compreende as tarefas de verificação e validação. A verificação, por sua vez, visa garantir a qualidade dos artefatos desenvolvidos. E por fim, a validação assegura se a ontologia reflete o domínio, bem como o seu escopo, as questões de competência e os usos pretendidos.

De acordo com Farinelli e Elkin (2017) outras metodologias são propostas na literatura, e no geral, estabelecem as seguintes atividades: especificação de requisitos de ontologia; aquisição de conhecimento; conceitualização e documentação; formalização e implementação. No entanto, os autores discutem que não há um consenso sobre a metodologia mais apropriada para engenharia de ontologias.

2.1.4 Ontologia de alto nível

As ontologias de alto nível representam e estruturam conceitos independentemente do domínio. De acordo com Hoehndorf (2010) em uma ontologia de alto nível, os conceitos são organizados em uma hierarquia “*is-a*”, formando uma taxonomia com classes mais gerais do que outras. A suposição por trás desse tipo de ontologia é que, quando essa generalização é realizada em ontologias de vários domínios diferentes, é possível alcançar um pequeno conjunto de categorias que é o mesmo em todos esses domínios. A maioria dos domínios lidará com objetos, processos, propriedades, relações, espaço, tempo, papéis, funções, indivíduos ou semelhantes, sendo que esses elementos podem estar organizados em categorias mais gerais visando melhorar a interoperabilidade semântica.

A interoperabilidade semântica é a capacidade de diferentes sistemas de *software* serem capazes de trocar informações de forma acessível por meio do significado preciso dos

dados. Vocabulários ou ontologias são estruturas que podem conceituar *datasets* atribuindo-lhes semântica, para facilitar a seleção e integração de dados (HEFLIN; HENDLER, 2000).

Casanova *et al.* (2005) citam o uso de metadados e ontologias como mecanismos capazes de descrever o conteúdo dos dados a fim de proporcionar a interoperabilidade semântica. Segundo os autores, o uso de ontologias facilita a interoperabilidade semântica ao representar o conhecimento do domínio, independentemente do formato dos dados, permitindo a equivalência semântica dos conceitos entre eles.

As ontologias de alto nível, ao serem aplicadas, fornecem interoperabilidade semântica, base ontológica comum para ontologias de domínio. A BFO (*Basic Formal Ontology*) fornece essas características e foi projetada para apoiar na recuperação, análise e integração de informações em domínios científicos. Segundo Spear, Ceusters e Smith (2016), o benefício do uso de ontologias de alto nível é alcançado quando as classes definidas por este tipo de ontologia permitem validar ontologias de domínio com coesão, do mesmo modo que a BFO é usada para validar ontologias na área biomédica.

De acordo Niles e Pease (2001), a SUMO (*Suggested Upper Merged Ontology*) é uma ontologia de alto nível que foi proposta como um documento inicial para o *The Standard Upper Ontology Working Group* por um grupo de trabalho composto por colaboradores das áreas de Engenharia, Filosofia e Ciência da Informação. Tal como a BFO, a SUMO fornece definições para termos de uso geral e atua para validar ontologias de domínio mais específicas.

De acordo com Guizzardi, Falbo e Guizzardi (2008):

Ontologias de Fundamentação (*Foundational Ontologies*) são sistemas de categorias filosoficamente bem fundamentados e independentes de domínio que têm sido utilizados com sucesso para melhorar a qualidade de linguagens de modelagem e modelos conceituais.

As ontologias de fundamentação também são consideradas ontologias de alto nível. Guizzardi (2005) propôs a UFO (*Unified Foundational Ontology*), uma ontologia de fundamentação centrada nas categorias ontológicas de endurantes e perdurantes, que utiliza fundamentos ontológicos para conceitos de modelagem conceitual. O objetivo do UFO, por meio de sua taxonomia, é fornecer bases para a análise de domínio, bem como projetar modelos conceituais com apoio de análise ontológica e explicitar semanticamente o conteúdo embutido em artefatos de representação. Estes, por sua vez, devem também apoiar os usuários humanos em tarefas como a compreensão do domínio (GUIZZARDI *et al.*, 2021).

Ontologistas podem aplicar a UFO utilizando a OntoUML, que é uma linguagem ontologicamente bem fundamentada para modelagem conceitual. A OntoUML é mantida como uma extensão da UML (*Unified Modeling Language*) baseada na UFO. As bases da UFO e OntoUML podem ser encontradas em "*Ontological foundations for structural conceptual*

models”, tese publicada em (GUIZZARDI, 2005). Existe um conjunto de tecnologias que facilitam a fundamentação de ontologias de domínio pela UFO. Neste sentido, ontologistas podem modelar usando a OntoUml por meio de um *plugin* no *Visual Paradigm*. O resultado da modelagem pode ser validado usando a ferramenta Protegé.

Com base nessas características, a UFO apresenta-se como ontologia de alto nível para fundamentar a ontologia ONTRISCAL, e, por ter sido a fundamentação selecionada para o estudo de caso (*cf.* Capítulo 5), ela será mais detalhada a seguir.

Endurantes são indivíduos, do tipo *concrete individual*, que existem no tempo com todas as suas partes. Eles têm propriedades essenciais ou acidentais e, portanto, podem mudar qualitativamente enquanto mantêm sua identidade, ou seja, enquanto permanecem o mesmo indivíduo. Os tipos de mudanças que um endurante pode sofrer enquanto mantém sua identidade, são definidos pelo tipo único que ele instancia. Perdurantes são indivíduos, do tipo *concrete individual*, que se desenvolvem no tempo acumulando partes temporais. Não há identidade através do mundo entre eles e, portanto, eles não podem ser diferentes do que são, por exemplo: Eleições presidenciais no Brasil em 2022; A Guerra do Vietnã; A Final da Libertadores da América 2021 (GUIZZARDI *et al.*, 2021).

A UFO considera entidades existencialmente dependentes ou caracterizadas por momentos, estes últimos denominados aspectos ou particulares abstratos. Os momentos intrínsecos são existencialmente dependentes de um único indivíduo, e podem incluir qualidades, tais como: cor; altura; peso; carga elétrica. Qualidades são entidades abstratas que delimitam o espaço de valores possíveis.

Na UFO a classe *Mode* é um tipo de momento intrínseco que pode suportar seus próprios momentos e variar de maneiras independentes, por exemplo: O amor de João por Maria; O compromisso de José para com Joana de se encontrar para almoçar no próximo domingo (GUIZZARDI, 2005). A classe *Relator* é outro tipo de momento intrínseco dependente de múltiplos indivíduos. Exemplos de instâncias do tipo *Relator* incluem casamentos, matrículas, empregos, contratos e mandatos presidenciais (GUIZZARDI *et al.*, 2021).

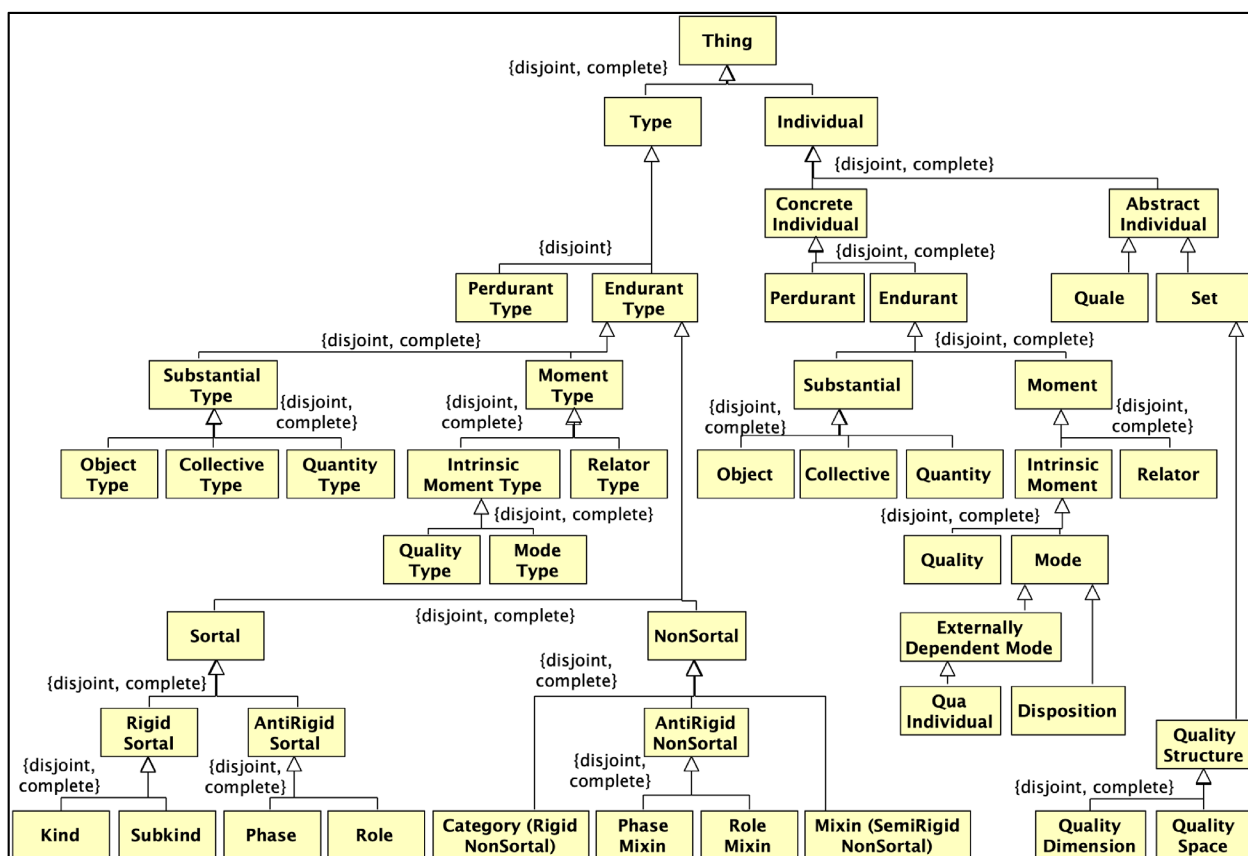
Uma classe do tipo *Kind* é um endurante que fornece princípios uniformes de individuação, identidade e persistência às suas instâncias. Por exemplo, instâncias da classe pessoa aplicadas a indivíduos que existem e, portanto, são tipos rígidos.

Um objeto *Sortal* é um endurante que generaliza os tipos *Rigid Sortal* e *AntiRigid Sortal*. A classe *Subkind* é uma especialização de *Rigid Sortal*, usada para representar especializações rígidas, por exemplo: ‘carro *hatch*’ como um *subkind* de carro. Um tipo *AntiRigid Sortal* possui características intrínsecas denominadas *phase*, por exemplo:

‘adolescente’ como uma fase da pessoa ou ‘dengue hemorrágica’ como uma fase da dengue. A classe denominada *Role* é outro tipo de *AntiRigid Sortal*, por meio da qual todas as instâncias devem seguir o mesmo princípio de identidade (GUIZZARDI, 2005; GUIZZARDI *et al.*, 2021).

A Figura 3 apresenta a taxonomia da UFO proposta em (GUIZZARDI, 2005), e cujos alguns dos tipos modelados foram brevemente comentados em parágrafos anteriores.

FIGURA 3 – A Taxonomia da UFO



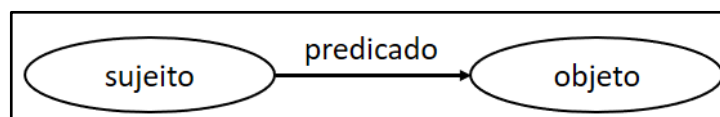
Fonte: (GUIZZARDI *et al.*, 2021).

2.2 Grafos de Conhecimento

Um GC é um grafo multi-relacional composto de entidades como nós (ou vértices) e relações como tipos diferentes de arestas. Uma instância representada por uma aresta é uma tripla denotada como s (sujeito), p (predicado), o (objeto). Ontologias podem representar os conceitos das entidades e relacionamentos instanciados em uma tripla. Além disso, fragmentos de GCs podem ser conectados para estruturar a solução para uma base de conhecimento, facilitando o gerenciamento, recuperação, uso e compreensão das informações (WANG *et al.*, 2014). A relação formada por uma tripla (*cf.* Figura 4) pode ser um

tipo literal para ser manipulado por aplicações de software, e ser representado por um URI³ (*Uniform Resource Identifier*).

FIGURA 4 - Representação de uma tripla



Fonte: Elaborado pelo autor.

Um URI identifica unicamente um recurso definido como um objeto de uma ontologia, sendo que este recurso pode ser um URL (*Uniform Resource Locator*) que representa uma forma de acessar o recurso na web.

De acordo com Pan *et al.* (2017) um GC possui restrições em sua estrutura impostas por meio de ontologias, e, ao serem instanciadas, permitem a inferência de novos fatos enriquecendo o conhecimento e o seu compartilhamento. Ainda de acordo com os autores, os GCs são frequentemente constituídos a partir de várias coleções de dados dependendo das técnicas que permitem extrair e incorporar essas coleções para gerar grafos.

Hogan *et al.* (2020) destacam a importância das ontologias como modelos conceituais utilizados para anotar conjuntos de dados, e, com isso, gerar fragmentos de conhecimento do domínio que podem ser representados por GCs.

Bax e Silva (2020) abordam o uso de ontologias para anotação de dados com vistas a harmonizar e normalizar variáveis de estudos científicos. Apresentam como estudo de caso a anotação semântica de dados fictícios oriundos de Prontuários Eletrônicos de Paciente (PEPs) e como estes dados podem estar organizados em grafos no formato RDF⁴

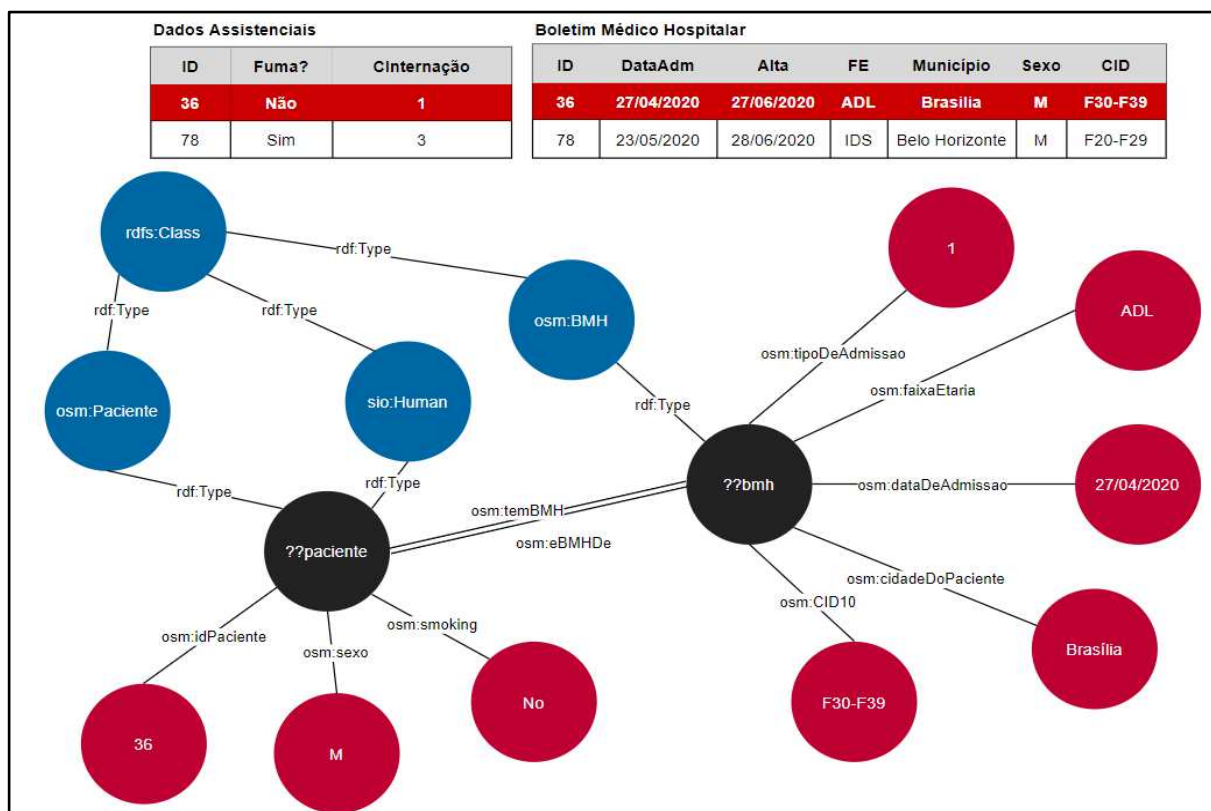
De acordo com Bax e Silva (2020), grafos em RDF podem ser explorados por consultas para originar outros arquivos de dados. O padrão RDF é utilizado para representar GCs, e os padrões RDFS (*Resource Description Framework Schema*) e OWL (*Ontology Web Language*) definem as interpretações semânticas em GC. A Figura 5 apresenta um exemplo de como os GCs podem ser constituídos com base em ontologias, classes de conceitos e fontes de dados distintas.

O GC representado na Figura 5 pode ser explorado para demonstrar a navegação por conceitos e dados. As tabelas listam, respectivamente, dados assistenciais e de internação em hospitais, tais como, o extrato do BMH (Boletim Médico Hospitalar), caracterizados por variáveis como data da admissão, alta faixa etária (FE), sexo e códigos CID.

³ <https://www.w3.org/wiki/URI>

⁴ <https://www.w3.org/RDF/>

FIGURA 5 - Exemplo de GC constituído a partir de fontes de dados distintas



Fonte: (BAX; SILVA, 2020).

Duas ontologias OSM e SIO (*Semanticscience Integrated Ontology*), definem os conceitos sobre paciente (`osm:Paciente`), humano (`sio:Humano`) e BMH (`osm:BMH`). A classe `rdfs:Class` declara um recurso como uma classe para outros recursos, relacionando as classes citadas aqui. Os objetos instanciados, representados em preto, relacionam as classes das ontologias com os dados por meio de propriedades. As triplas podem ser instanciadas para representar o conhecimento sobre os dados presentes no grafo, como, por exemplo: `??paciente osm:temBMH osm:bmh`; e `rdfs:Class rdf:Type osm:Paciente`. Onde:

??paciente e `rdfs:Class`, são classes
`osm:temBMH` e `rdf:Type`, são predicados (relacionamentos)
`osm:bmh` e `osm:Paciente`, são objetos

2.3. Anotação de dados com *Semantic Data Dictionaries*

O termo “anotação” implica, de um modo muito geral, anexar dados a algum outro dado estabelecendo uma relação entre dados anotados. Por exemplo, permite relacionar o texto ou dado “Paris” a uma ontologia, identificando-a como cidade e capital da França. As

ferramentas de anotação manual permitem que os usuários adicionem anotações a recursos na web e as compartilhem com outras pessoas. Ferramentas automáticas podem realizar anotações semelhantes, como reconhecimento de entidade nomeada sem intervenção manual (OREN *et al.*, 2006).

De acordo com Oren *et al.* (2006), três tipos de anotações podem ser implementadas: informais, formais e ontológicas. As anotações informais não são legíveis por máquina porque não usam uma linguagem formal. Por outro lado, as anotações formais são compreensíveis por máquina, mas não usam termos ontológicos. Nas anotações ontológicas, os termos estão associados a um significado, permitindo anotações semânticas, que correspondem a uma conceituação compartilhada por uma ontologia.

Rashid *et al.* (2020) descrevem o uso de tecnologias encontradas na *Web Semântica* onde as colunas das tabelas de dicionários de dados são anotadas com significados que usam ontologias como vocabulários controlados. Esta abordagem visa melhorar a recuperação e o acesso a diversos conjuntos de dados definidos pelo dicionário de dados, tanto por usuários avançados na área da computação, quanto para usuários finais de aplicações computacionais. A partir da anotação semântica os dados são identificados univocamente por meio de URLs. Os autores ressaltam que os dicionários de dados semânticos (ou SDD) podem melhorar o reuso e a interoperabilidade dos conjuntos de dados.

O SDD é um conjunto de padrões de metadados fundamentados em ontologias que descrevem objetos (representados por dados) em classes e relacionamentos (RASHID *et al.*, 2017). A anotação por SDD é feita de forma manual e associa os dados de um *dataset* a conceitos (ou classes) nas ontologias, formalizando a semântica dos dados. A anotação dos dados pode ser realizada por profissionais familiarizados com o domínio e aos *datasets*, com apoio de ontologistas. A formalização do vocabulário abre caminho para a interoperabilidade dos dados que podem ser integrados de fontes diversas. As seções a seguir definem os *templates* de metadados utilizados na anotação por SDD.

2.3.1 InfoSheet

O *InfoSheet* é essencialmente o documento de configuração da estrutura do SDD. É usado para organizar as tabelas de SDD e definir informações sobre o Dicionário de Dados Semânticos, como nome, identificador ou link para a documentação, além da localização das demais tabelas SDD. As tabelas SDD, geralmente, são uma coleção de arquivos CSV (*Comma-Separated Values*) que contêm a maioria das informações sobre o conjunto de dados e seus relacionamentos (RASHID *et al.*, 2017; TWC, 2021). A especificação do *Infosheet* é apresentada no Quadro 1.

Quadro 1 - Especificação do Infosheet

Infosheet Row	Related Property	Description	Example
<i>Code Mapping</i>		Referência à localização da tabela de mapeamento de código	http://...
<i>CodeBook</i>		Referência à localização da tabela Codebook	http://...
<i>Dictionary Mapping</i>		Referência à localização da tabela de mapeamento de dicionário	http://...
<i>Imports</i>	<i>owl:imports</i>	Ontologias que o SDD faz referência	http://semanticscience.org/ontology/sio-subset-labels.owl
<i>TimeLine</i>		Referência à localização da tabela de <i>TimeLine</i>	http://...

Fonte: (TWC, 2021).

2.3.2 Dictionary Mapping (DM)

O DM anota as colunas do *dataset*. Cada linha do DM mapeia uma coluna do *dataset*, formalizando-a conceitualmente, explicitando suas relações com os outros dados. As entradas explícitas contêm mapeamentos para o atributo subjacente, bem como informações de proveniência. As entradas implícitas resultam do esforço de explicitar o conhecimento além dos mapeamentos de atributos. Novas definições e relações entre os dados podem ser determinadas pelas entradas implícitas, avançando acerca do conhecimento do domínio. A entradas implícitas podem, então, ser descritas com tipo, função, relação e outras informações da mesma maneira que as colunas explícitas no conjunto de dados. Ao anotar os dados, as entidades são prontamente reconhecidas e utilizadas para consumo humano. A especificação do DM é apresentada no Quadro 2 (RASHID *et al.*, 2017; TWC, 2021).

Quadro 2 - Especificação do Dictionary Mapping

DM Column	Related Property	Description
<i>Attribute</i>	rdf:type	Classe de entrada de atributo
<i>attributeOf</i>	sio:isAttributeOf	Entidade com o atributo
<i>Column</i>		Cabeçalho da coluna de entrada no conjunto de dados
<i>Comment</i>	rdfs:comment	Comentário para a entrada
<i>Definition</i>	skos:definition	Definição do texto de entrada
<i>Entity</i>	rdf:type	Classe de entrada de entidade
<i>Format</i>		Especifica a estrutura do valor da unidade
<i>inRelationTo</i>	sio:inRelationTo	Entidade à qual a função está vinculada
<i>Label</i>	rdfs:label	Etiqueta para a entrada
<i>Property</i>		Especificação de propriedade de tipo de dados personalizado
<i>Relation</i>		Relação personalizada que substitui inRelationTo
<i>Role</i>	sio:hasRole	Tipo de função da entrada
<i>Time</i>	sio:existsAt	Ponto de tempo de medição
<i>Unit</i>	sio:hasUnit	Unidade de medida para entrada
<i>wasDerivedFrom</i>	prov:wasDerivedFrom	Entidade da qual a entrada foi derivada
<i>wasGeneratedBy</i>	prov:wasGeneratedBy	Atividade a partir da qual a entrada foi produzida

Fonte: (TWC, 2021).

Os nomes das entradas explícitas e implícitas são armazenados em *Column*, que se refere aos nomes das colunas no conjunto de dados. As propriedades de anotação, incluindo comentários, rótulos ou definições, podem ser fornecidas para descrever uma entrada explícita ou virtual em mais detalhes, para o leitor humano. Se uma entrada descreve uma característica, a coluna *Attribute* deve ser preenchida com uma classe apropriada e, se necessário, a coluna *attributeOf* deve ser usada para fazer referência à entidade à qual o atributo pertence. Normalmente, a coluna *attributeOf* contém a entidade implícita para a qual a entrada explícita é uma característica. Supondo um conjunto de dados que possui uma coluna chamada idade. Embora possa ser fácil para o humano ler e entender que a idade se refere à mãe, o conjunto de dados, muitas vezes, não inclui isso explicitamente; se almejamos a caixa como uma representação semântica completa dos dados, essa relação deve ser definida para o computador. No entanto, uma vez que “mãe” não é uma coluna no conjunto de dados, é considerada uma entrada implícita, denotada por ‘??’ na frente da palavra mãe (??mae).

2.3.3 CodeBook

A tabela *Codebook* estrutura os dados categoriais de um *dataset*, mapeando-os para conceitos correspondentes na ontologia. Tem uma função semelhante à tabela DM; enquanto o DM serve para codificar os significados dos cabeçalhos das colunas no conjunto de dados, o *Codebook* contém todos os valores possíveis para cada coluna e seus rótulos associados. Por exemplo, um conjunto de dados que contém informações sobre pessoas, uma das colunas possíveis pode ser gênero, que seria a entrada no DM, enquanto as entradas possíveis para essa coluna, masculino e feminino, seriam entradas no *Codebook*. A especificação do *CodeBook* é apresentada no Quadro 3 (RASHID *et al.*, 2017; TWC, 2021).

Quadro 3 - Especificação do *CodeBook*

Column	Related Property	Description
<i>Class</i>	rdf:type	Classe a que o <i>codebook</i> se refere
<i>Code</i>	sio:hasValue	Valor da entrada do conjunto de dados
<i>Column</i>	-	Cabeçalho da coluna de entrada no conjunto de dados
<i>Comment</i>	rdfs:comment	Comentário para a entrada do <i>codebook</i>
<i>Definition</i>	skos:definition	Definição para a entrada do <i>codebook</i>
<i>Label</i>	rdfs:label	Etiqueta para a entrada do <i>codebook</i>
<i>Resource</i>	rdf:type	URI de recurso ao qual o <i>codebook</i> se refere

Fonte: (TWC, 2021).

2.3.4 TimeLine

Intervalos de tempo personalizados podem ser especificados na planilha chamada

TimeLine. Pode ser usada para anotar a classe e unidade correspondentes relacionadas a uma determinada entrada, bem como os horários de início e término de um evento e uma conexão com conceitos aos quais a entrada pode estar relacionada (TWC, 2021). Ao usar o *template* para preenchimento de *TimeLine*, é importante certificar que a entrada de tempo na tabela seja um intervalo de tempo em vez de uma instância de tempo. Por exemplo, um aniversário não estaria na linha do tempo, mas deveria ser visto como uma característica de um assunto. Por outro lado, no estudo do projeto CHEAR (*The Children's Health Exposure Analysis Resource*), ao aplicar a técnica SDD, McCusker *et al.* (2017) utilizam dados que rastreiam o desenvolvimento infantil em termos de observações feitas em momentos específicos relativos ao nascimento ou concepção da criança. Os autores compararam as medições entre os assuntos para um determinado momento, como “o segundo trimestre da gravidez” (MCCUSKER *et al.*, 2017; TWC, 2021). A especificação do *TimeLine* é apresentada no Quadro 4.

Quadro 4 - Especificação do TimeLine

<i>TimeLine</i>	<i>Related Property</i>	<i>Description</i>
<i>End</i>	sio:hasEndTime	Horário de término associado à entrada da linha do tempo
<i>inRelationTo</i>	sio:inRelationTo	Entrada da linha do tempo está relacionada
<i>Label</i>	rdfs:label	Rótulo para a entrada da linha do tempo
<i>Name</i>	-	Referência à entrada da linha do tempo virtual
<i>Start</i>	sio:hasStartTime	Hora de início associada à entrada da linha do tempo
<i>Type</i>	rdf:type	Classe da entrada da linha do tempo
<i>Unit</i>	sio:hasUnit	Unidade de tempo

Fonte: (TWC, 2021).

2.3.5 Script para geração de grafo de conhecimento

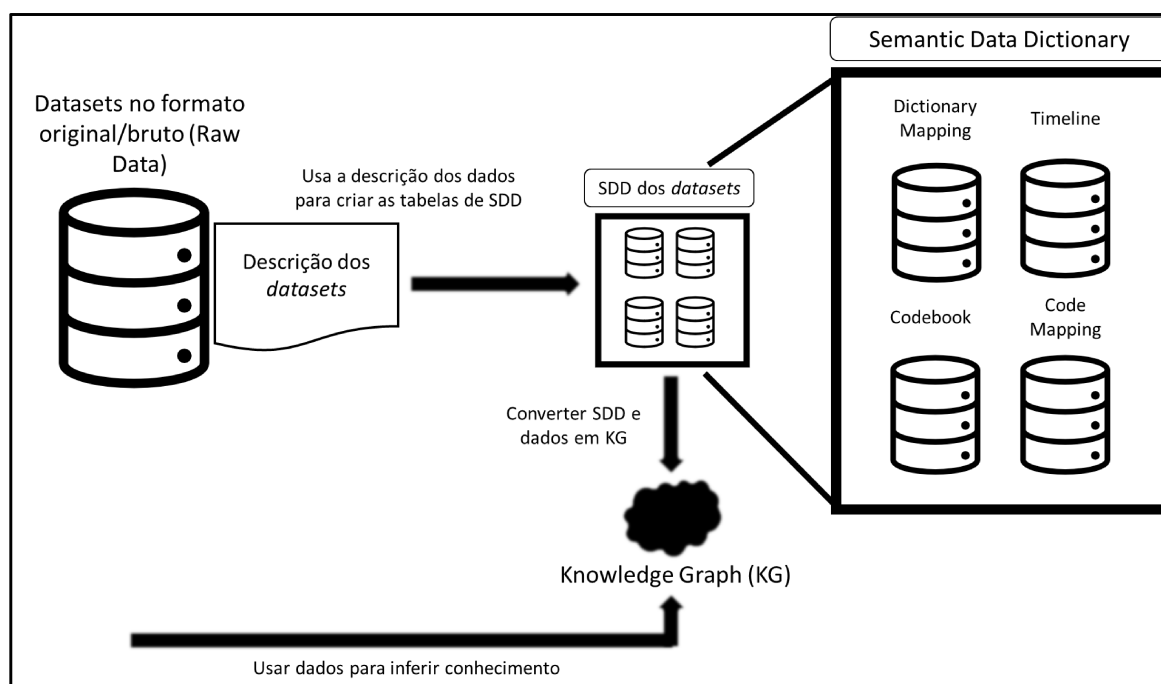
Inicialmente, os dados mapeados para as ontologias pelo SDD são as colunas do próprio *dataset*. Os objetos caracterizados nos *datasets* podem estar implicitamente representados. Os objetos implícitos serão explicitados no SDD e formalizados no grafo final gerado. A explicitação dos objetos implícitos favorece a integração semântica dos dados nos níveis conceituais mais abstratos do domínio do projeto, permitindo normalizar e harmonizar interpretações de conceitos que descrevem dados que se deseja integrar.

Uma vez que a estrutura de anotação esteja pronta, a ferramenta *sdd2rdf* é utilizada para integrar os dados do *dataset* descrito pelo SDD formando um GC persistido em *triplestores*. O *sdd2rdf*⁵ é um *script/software* que interpreta o SDD e converte os dados do *dataset* descrito pelo SDD gerando o GC expresso no padrão RDF (RASHID *et al.*, 2017).

⁵ <https://github.com/tetherless-world/SemanticDataDictionary/tree/master>

Para exemplificar o acesso aos dados anotados no grafo, o *sdd2rdf* cria alguns exemplos de consultas SPARQL. O GC possibilita a inferência de novos fatos enriquecendo o compartilhamento do conhecimento. Consultas SPARQL e regras SWRL permitem inferências. O grafo RDF gerado pelo script *sdd2rdf* se fundamenta em ontologias, e possibilita a integração semântica dos dados bem como a sua interoperabilidade. A Figura 6 representa, de forma simplificada, o processo de anotação feita com o uso de SDDs.

FIGURA 6 - Representação da anotação de dados feita por SDD



Fonte: (RASHID *et al.*, 2017).

2.4 Análise qualitativa

Esta seção define os conceitos relacionados à análise qualitativa e que norteiam a fundamentação teórica para a execução do método HOMOGENiSE. A codificação é abordada nesta seção e serve de base para aplicar as técnicas AT e GT.

Análise é a ação ou efeito de fazer um exame detalhado de algum artefato (ANÁLISE, 2023). O termo qualitativo refere-se à qualidade ou natureza de seres ou objetos, relacionando-se assim com as características de alguém ou de algo (QUALITATIVO, 2023). Deste modo, a análise qualitativa é considerada o ato de examinar detalhadamente a qualidade e as características de algo. De acordo com Sandelowski (1995) a análise qualitativa é um meio de produção de conhecimento suportada pela organização de dados extraídos de diferentes documentos. Estabelece um processo criativo, dependendo dos *insights* e conhecimento do pesquisador com interesse nos dados. Patton (1999) acrescenta

que a análise qualitativa tem como propósito o reconhecimento de padrões apoiando o pesquisador na inferência de novos dados. O autor cita que cientistas realizam a análise qualitativa com o objetivo de identificar padrões de termos para responder questões de pesquisa (PATTON, 1999). Análises qualitativas visam identificar as relações entre categorias de dados, mostrando como as categorias se relacionam entre si (POLKINGHORNE, 1995).

Minayo (2012) propõe um processo que organiza atividades para análise qualitativa, visando melhorar a compreensão dos pesquisadores diante dos dados. O processo visa estabelecer um conjunto de substantivos para classificar os dados, que são analisados baseados na experiência e vivência do pesquisador. A classificação estrutura categorias de dados, promovendo a compreensão deles, e oferecendo um novo processo de teorização ou até mesmo a criação de novas categorias (MINAYO, 2012; SCHENSUL; LECOMPTE, 2012).

2.4.1 Meios de Representação

De acordo com Mayring (2007) é importante selecionar os meios de representação apropriados para análise qualitativa. Os meios de representação são elementos tais como: tabelas, mapas mentais (diagramas), modelos de processos ou gráficos que permitem ordenar os dados para análise qualitativa e explicitar, graficamente, os relacionamentos complexos a fim de facilitar a compreensão da pesquisa.

De acordo com Verdinelli e Scagnoli (2013) elementos visuais ajudam em inferências e considerações sobre como estruturar ou sugerir novos dados. Verdinelli e Scagnoli (2013, p. 3) destacam a “pouca atenção dada às definições, objetivos e usos comuns de diagramas em publicações de pesquisa qualitativa”.

De acordo com Meister (2017) o suporte ao desenvolvimento do trabalho científico por meio de tecnologia semântica, tem despertado interesse dos pesquisadores. Esforços são feitos no fornecimento de metadados estruturados, contribuindo para o desenvolvimento de pesquisas qualitativas mais complexas para melhorar a recuperação e representação de dados por meio de GCs. Meister (2017) ressalta que as tecnologias semânticas têm o potencial de apoiar pesquisas acadêmicas a fim de compartilhar documentos, dados e resultados científicos de forma colaborativa e a reutilização do conhecimento produzido.

2.4.2. Codificação

A codificação é uma estratégia de análise de dados qualitativos que permite ao pesquisador utilizar códigos (*codes*) classificados em categorias, para associá-los aos dados analisados. A decisão sobre como codificar os dados deve ser conduzida por metodologia

própria do pesquisador. Raramente encontra-se guias ou manuais sobre como o pesquisador fará a codificação de acordo com um determinado método. Neste caso, o pesquisador deve tomar a decisão sobre como codificar seu próprio projeto com base no domínio de aplicação (LIBRARY ILLINOIS, 2023).

Segundo Baralt (2012) a codificação na pesquisa qualitativa é o processo analítico para organizar os dados brutos em temas que auxiliem na interpretação dos dados. De acordo com Baralt (2012) a codificação é uma atividade em que o pesquisador utiliza os códigos como “nomes ou símbolos usados para representar um grupo de itens, ideias ou fenômenos semelhantes ao que o pesquisador notou em seu conjunto de dados”.

A codificação é utilizada em um processo para determinar um conjunto de informações pela interação contínua e releitura dos dados. Com a codificação torna-se possível comparar e contrastar os dados, suportando análises para responder à questão de pesquisa. Além disso, ao realizar a codificação, o pesquisador descobre padrões nos dados. Em muitos casos, a codificação de dados qualitativos é congruente com a construção de explicações e até mesmo com a geração de teoria (BARALT, 2012; GIBBS, 2010; SCHENSUL; LECOMPTE, 2012).

Schensul e Lecompte (2012, p. 277) consideram “conceitos organizados em uma taxonomia ou diagrama de árvore, como abordagem mais eficaz para codificar dados qualitativos”. Os dados são codificados por domínios de aplicação em níveis de códigos estruturados pela taxonomia. No entanto, os autores não aconselham codificar usando todos os níveis. Níveis mais inferiores são mais específicos (não genéricos), e dificultam a análise dos dados, consumindo tempo dos codificadores na tentativa de memorizar a estrutura da taxonomia.

A Library Illinois possui a *University of Illinois LibGuides*, uma plataforma que possui guias de recursos elaborados por bibliotecários e auxilia pesquisadores em diferentes assuntos e materiais de pesquisa. Library Illinois (2023) descreve algumas questões a serem consideradas ao planejar a codificação de dados:

- O que codificar? É necessário selecionar e compreender os dados, e definir quais elementos precisam ser codificados, tais como: entrevistas gravadas, textos, ou outros tipos de dados multimídia.
- Qual a proveniência dos códigos? Os códigos usados pela codificação podem ser elaborados a partir de alguma metodologia, ou originarem-se de pesquisas anteriores (codificação dedutiva). Podem ser definidos a partir dos dados, ignorando, tanto quanto possível, o conhecimento prévio do tema em estudo, para desenvolver um esquema fundamentado em seus dados (codificação indutiva).

- Como aplicar os códigos aos dados? O pesquisador pode decidir usar software para codificar seus dados qualitativos. É possível utilizar ferramentas de software, por exemplo, editores de texto ou planilhas eletrônicas ou mesmo CAQDAS. Um CAQDAS deve oferecer funcionalidades que permitam renomear os códigos, mesclá-los ou exportá-los. Alguns softwares incluem funcionalidades para codificar diretamente arquivos de imagem, vídeo e áudio. Exemplos de CAQDAS são: Atlas-Ti⁶, NVivo⁷, MAXQDA⁸ e Taguette⁹,
- O que será codificado? O pesquisador deverá adotar uma abordagem de codificação linha por linha, em textos condensados (mais curtos) usando categorias ou conceitos mais genéricos, ou (em alguns casos específicos) pode aplicar os códigos a segmentos maiores do texto.
- Como explicar o processo de codificação? Independentemente da forma como a codificação é abordada, o processo deve ser devidamente esclarecido ao relatar a pesquisa.
- Como relatar os resultados do processo de codificação? O pesquisador pode utilizar a codificação para gerar relatórios ou planilhas com frequências que contabilizem os códigos. Ou, é possível usar os códigos para ajudar a desenvolver uma descrição dos dados codificados, sem precisar indicar com que frequência o código foi aplicado.
- Como codificar de forma colaborativa? Em um trabalho colaborativo a codificação requer planejamento e execução mais cuidadosos. Um plano de comunicação entre os codificadores deve ser estabelecido, com conversas regulares sobre a execução da codificação.

2.4.3. *Grounded Theory*

Uma “teoria” é uma ideia ou conjunto de ideias que se destina a explicar fatos ou eventos (TEORIA, 2023). A *GT* é um método de pesquisa qualitativa que permite ao pesquisador derivar novas teorias baseadas na coleta e análise iterativas (em ciclos) dos dados. A *GT* apoia a investigação em pesquisas onde a coleção de dados é identificada a partir da leitura reflexiva de textos em conjunto com a técnica de codificação (GIBBS, 2010). Deste modo, ao executar *GT*, pesquisadores aplicam a codificação para gerar dados qualitativos. Os códigos aplicados aos dados são analisados e agrupados. Os pesquisadores fazem o agrupamento dos códigos para criar conexões entre eles, e inferir novos

⁶ <https://atlasti.com/>

⁷ <https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home>

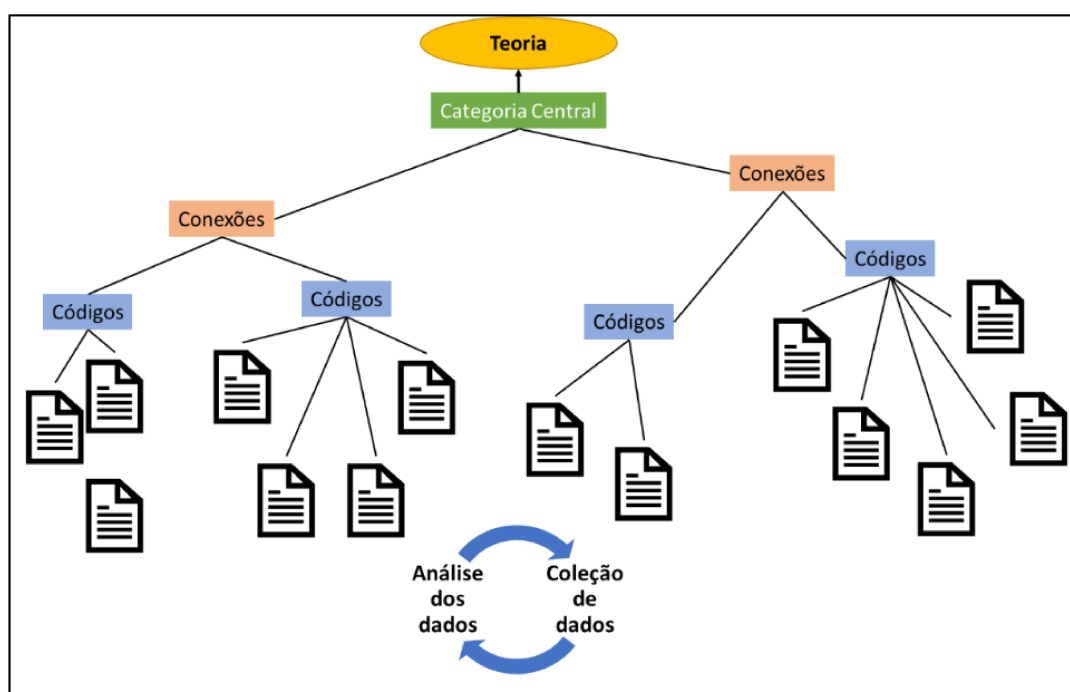
⁸ <https://www.maxqda.com/>

⁹ <https://www.taguette.org/>

conhecimentos. De forma cíclica, novos agrupamentos ocorrem utilizando as conexões em direção a uma teoria (DELVE, 2022; GIBBS, 2010).

Em *GT*, o agrupamento de códigos é uma tarefa chamada codificação axial (*axial coding*) (DELVE, 2022). De acordo com Gibbs (2010) a codificação axial é o meio que auxilia o pesquisador a proceder com a análise dos agrupamentos para definir categorias de dados. Com isso, novos ciclos de coleta e análise possibilitam ao pesquisador perceber novos dados (WALKER; MYRICK, 2007). O pesquisador conecta as categorias de dados (elicitadas pela codificação axial) para identificar uma categoria central potencial (*core category*). Ao determinar a categoria central, o pesquisador define a teoria. A Figura 7 apresenta um esquema que resume as ações para estabelecer a *GT*.

FIGURA 7- Ações para estabelecer a *GT*



Fonte: (DELVE, 2022).

2.4.4 Análise Temática

De acordo com Kiger e Varpio (2020) a AT é um método para identificar e descrever dados qualitativos, para compreender, selecionar e construir temas. Para Braun e Clarke (2006) a AT possui características semelhantes a procedimentos tradicionalmente adotados na análise qualitativa, adotando aspectos como busca por padrões por meio da homogeneidade das categorias de dados. Os padrões ajudam a identificar os temas a serem analisados.

A AT pode ser utilizada amplamente como técnica de análise qualitativa em diferentes questões de estudo e projetos de pesquisa. Embora pesquisadores tenham descrito a AT como pertencente ao domínio da etnografia (ARONSON, 1995) ou como particularmente adequada à fenomenologia (JOFFE, 2011), Braun e Clarke (2006) argumentam que a AT pode ser considerada um método analítico, e aplicada em conjunto com outros métodos de pesquisa qualitativa. Braun e Clarke (2006) fornecem um guia com seis fases para AT. As diferentes fases para análise temática estão resumidas no Quadro 5:

Quadro 5 Fases para análise temática

Fase	Descrição do Processo
1. Familiarizar com os dados	Transcrever e compreender os dados, anotar ideias iniciais.
2. Gerar códigos iniciais	Codificar os dados de forma sistemática, reunindo dados relevantes para cada código.
3. Localizar temas	Agrupar os códigos em temas potenciais, reunindo todos os dados relevantes para cada tema.
4. Revisar temas	Verificar se os temas se relacionam aos dados codificados, gerando um “mapa” temático da análise.
5. Definir e nomear temas	Refinar as especificidades de nomeação de cada tema.
6. Produzir o relatório	Retomar a questão de pesquisa e verificar se a AT permite respondê-la, produzindo um relatório acadêmico da análise.

Fonte: (BRAUN; CLARKE, 2006).

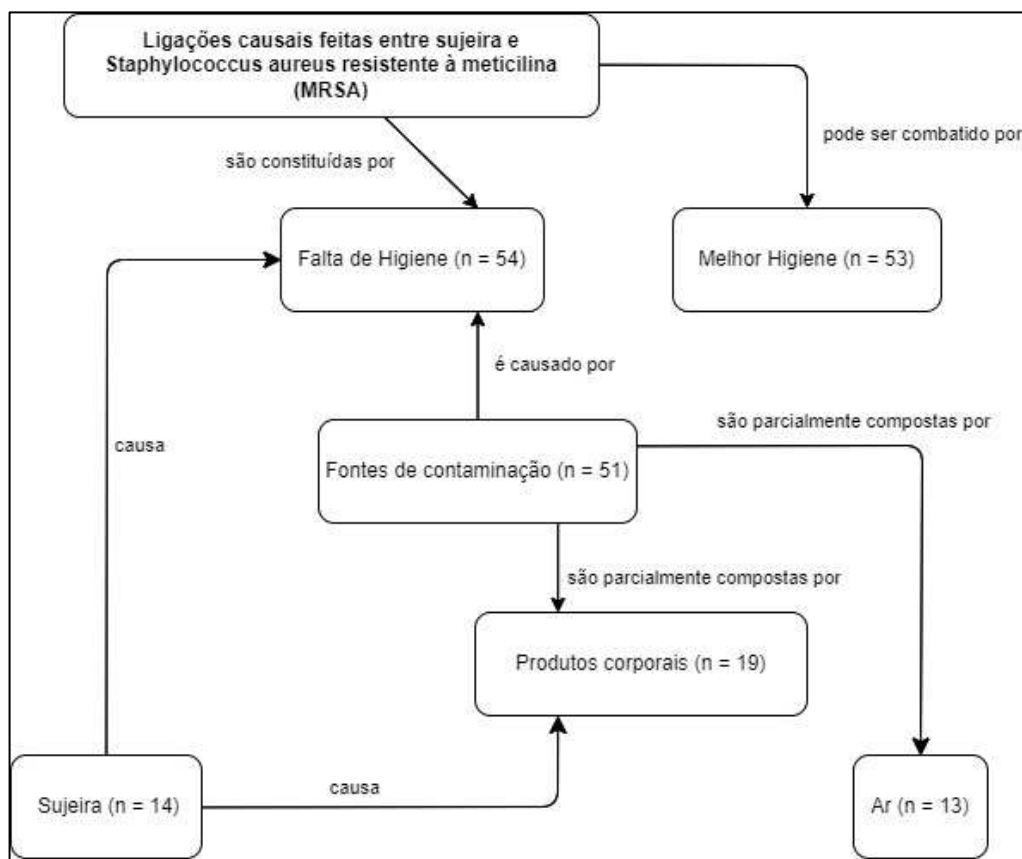
De acordo com Joffe (2011) a AT tem sido usada na saúde mental a fim de explorar as representações sociais. É usada para responder questões associadas a doenças e seus impactos na vida social e pessoal dos pacientes.

Joffe (2011) descreve um guia com métodos de pesquisa qualitativa em saúde mental e psicoterapia. A autora cria um diagrama produto da AT e que representa características de um hospital sem higiene. O objetivo do diagrama foi estabelecer ligações causais entre a falta de higiene e *Staphylococcus aureus* resistente à metilina (MRSA, *Methicillin-resistant Staphylococcus aureus*). A contaminação por MRSA pode contribuir para que os pacientes tenham doenças agravadas ou apareçam outras. No diagrama representado pela Figura 8 a primeira associação é “sujeira”, sendo uma motivação para aplicação da AT. A autora associou a “sujeira” a “infecções e germes”. Os termos “sujeira” e “infecções” foram considerados como temas a serem analisados.

Joffe (2011) entrevistou pacientes que, durante a internação, foram diagnosticados com a MRSA. Utilizou-se a AT para criar o diagrama da Figura 8 e que vincula a sujeira à causa da doença. A autora elaborou os seguintes passos para AT:

- Examinar o conjunto de dados para estruturar a codificação: A estrutura de codificação possui códigos formando um *codebook* para suportar a codificação. O *codebook* estabelece uma taxonomia composta por níveis de códigos. A codificação foi aplicada nos textos das respostas dos pacientes.
- Verificar a confiabilidade do *codebook*. Aplicar o *codebook* por dois pesquisadores independentes. Joffe (2011) aplicou a codificação em uma proporção substancial dos dados, e os pesquisadores aplicaram os códigos em pelo menos 20% do conjunto de dados.
- Realizar a codificação por meio de CAQDAS. Uma vez que *codebook* tenha sido concebido e a confiabilidade verificada, todo o conjunto de dados foi codificado novamente por meio de CAQDAS.
- Analisar os dados. A análise facilita o exame dos temas e suas interconexões e a prevalência dos temas nas amostras. Diferentes CAQDAS permitem que os pesquisadores examinem a padronização dos temas em toda a gama de entrevistas e os caminhos comuns ou cadeias de associação dentro das entrevistas. De acordo com Joffe (2011) as funções de filtragem dos softwares permitiram aos pesquisadores recuperar os padrões de códigos predominantes em grupos específicos (por exemplo, diferentes dados demográficos), e esses padrões puderam ser recuperados como gráficos de frequência, listas de trechos textuais ou visualmente, como redes visuais. Quando o diagrama representado na Figura 8 indica “Melhor higiene (n=53)” como forma de combater o MRSA, de acordo com a autora, significa que 53 das 60 pessoas da amostra disseram, especificamente, que uma melhor higiene ajudaria a combater o MRSA. A Figura 8 apresenta os temas analisados e suas relações com o número de pacientes que responderam questões relacionadas aos respectivos temas. Em particular, mencionou-se que o MRSA é causado pela falta de higiene nos hospitais do Serviço Nacional de Saúde, explicada por déficits nas práticas de lavagem das mãos e escassez de recursos. Em particular, a supervisão do pessoal de limpeza foi considerada deficiente. Com isso, Joffe (2011) considerou que os procedimentos de higiene não foram aplicados.

FIGURA 8 - Análise temática da contaminação por MRSA



Fonte: (JOFFE, 2011).

2.5 Análise quantitativa

De acordo com Labaree (2013), a análise quantitativa utiliza métodos que enfatizam medições, análises estatísticas, matemáticas ou numéricas de dados coletados por meio de enquetes, questionários e pesquisas, ou por manipular dados estatísticos pré-existentes usando técnicas computacionais. A pesquisa quantitativa concentra-se em coletar dados numéricos e generalizá-los em grupos relacionados aos objetos em estudo (tal como pessoas, indivíduos), ou para explicar um fenômeno específico (LABAREE, 2013).

De acordo com Cramer (2003), em pesquisa quantitativa, a coleta de dados visa responder às questões de pesquisa e verificar hipóteses. Dados inconsistentes podem distorcer a leitura das variáveis para a análise quantitativa, e dificultam a operacionalização da pesquisa. De acordo com Sukamolson (2007) uma variável é uma propriedade ou recurso que caracteriza os objetos de pesquisa, variando em qualidade e quantidade. O autor define a pesquisa quantitativa como uma área que quantifica e analisa variáveis para obter resultados, utilizando a análise de dados numéricos e técnicas estatísticas específicas.

De acordo com Cramer (2003), um pesquisador pode estar interessado em explicar as características particulares das variáveis ao aplicar a análise quantitativa, utilizando valores numéricos. Por exemplo, “sexo” é uma variável que pode ter dados atribuídos como feminino ou masculino. Essa variável pode ser representada ou codificada com valores numéricos, onde “feminino” pode ser atribuído com o valor “1” e masculino com “2”.

Entre os diferentes tipos de variáveis utilizadas na análise quantitativa, Cramer (2003) define dois tipos de variáveis: qualitativa e quantitativa. Através da variável qualitativa é possível categorizar dados e aplicar fórmulas para calcular a frequência ou prevalência de termos. Outro exemplo de variável qualitativa é o “estado civil” de uma pessoa, que pode consistir em cinco categorias: (1) nunca foi casado; (2) casado; (3) separados; (4) divorciado; e (5) viúvo. Essas categorias podem ser representadas ou codificadas com qualquer conjunto de cinco números. Os números são usados para se referir às diferentes categorias. O número, frequência ou prevalência de casos que se enquadram em cada uma das categorias só pode ser contado onde o resultado pode ser apresentado em uma variável de frequência. Por exemplo, de 100 pessoas, 30 podem nunca ter sido casadas, 40 podem ser casadas, 8 podem ser separadas, 12 podem ser divorciadas e 10 podem ser viúvas. Além disso, a frequência de casos em uma categoria pode ser expressa como uma proporção ou porcentagem da frequência total de casos (CRAMER, 2003; LAY; REIS, 2005; MORAIS, 2002).

De acordo com Moraes (2002) as variáveis quantitativas podem ser discretas ou contínuas. Uma variável é discreta quando características mensuráveis podem assumir apenas um número contável com valores inteiros. Geralmente, são resultados de contagens, por exemplo: número de filhos ou número de cigarros fumados por dia. Uma variável é contínua quando os valores pertencem a um intervalo de números reais e representam uma mensuração, como por exemplo, altura, peso ou pressão arterial de um paciente. Outro tipo de variável quantitativa é chamado variável ordinal, utilizada quando as categorias de dados podem ser ordenadas (MAYER, 2023).

Abaixo, seguem exemplos de dados com o uso dos tipos de variáveis citados anteriormente. Os exemplos foram adaptados de Bussab e Morettin (2010) com base em um conjunto de dados hipotético da estratificação de risco em ansiedade. Os dados estão reproduzidos no Quadro 6.

Quadro 6 - Conjunto de dados hipotético da estratificação de risco em ansiedade

Paciente	Estado Civil	Instrução	Idade	Renda Familiar	Pontuação Final	Nível de Risco
1	solteiro	1o Grau	26	1.140,00	30	Baixo
2	casado	1o Grau	32	1.312,50	50	Médio
3	casado	Superior	26	1.432,50	20	Baixo
4	solteiro	2o Grau	26	1.565,00	70	Alto
5	solteiro	1o Grau	40	1.665,00	80	Alto

Paciente	Estado Civil	Instrução	Idade	Renda Familiar	Pontuação Final	Nível de Risco
6	casado	Superior	40	1.715,00	25	Baixo
7	solteiro	1o Grau	41	1.847,50	30	Baixo
8	solteiro	2o Grau	41	1.897,50	10	Baixo
9	casado	Superior	26	1.860,00	25	Baixo
10	solteiro	1o Grau	26	2.030,00	40	Médio

Fonte: Elaborado pelo autor. Adaptado de: (BUSSAB; MORETTIN, 2010).

Mayer (2023) utiliza os dados tabulares para fazer análise descritiva, com objetivo de identificar padrões e relações. O primeiro passo é classificar todas as variáveis desse conjunto de dados. O Quadro 7 apresenta um exemplo da classificação das variáveis identificadas no Quadro 6.

Quadro 7 - Classificação dos tipos de variáveis

Variável	Classificação
Paciente	Quantitativa discreta
Estado Civil	Qualitativa nominal
Instrução	Qualitativa ordinal
Idade	Quantitativa discreta
Renda Familiar	Quantitativa contínua
Pontuação Final	Quantitativa discreta
Nível de Risco	Qualitativa nominal

Fonte: Elaborado pelo autor.

As variáveis podem ser aplicadas para análise quantitativa com base nos dados hipotéticos do Quadro 6 (BUSSAB; MORETTIN, 2010; FREITAS *et al.*, 2009; LAY; REIS, 2005; MAYER, 2023). Conforme apresentado no Quadro 7 a variável “Nível de Risco” é qualitativa nominal, valores que expressam atributos, sem nenhum tipo de ordem. É possível obter análises por meio de tabela de frequências e gráfico de barras. Além das representações dos dados, o pesquisador pode aplicar fórmulas estatísticas, tal como, a “moda”, por exemplo, o valor que ocorre com maior frequência. Para distribuir a frequência, basta organizar quantas vezes ocorre cada categoria por meio de uma tabela com uma contagem simples constituindo a frequência absoluta (MAYER, 2023).

Diante dos dados apresentados no Quadro 7, a variável “Instrução” pode exemplificar a análise para variável qualitativa ordinal do grau de instrução dos pacientes, que possui valores que expressam atributos com algum tipo de ordem, ou grau (MAYER, 2023).

Uma variável quantitativa discreta pode assumir valor numéricos como categorias, por exemplo, a “idade” de um paciente. Para variáveis cujos valores possuem ordenação natural (qualitativas ordinais e quantitativas em geral), faz sentido calcular também as frequências acumuladas (BUSSAB; MORETTIN, 2010; FREITAS *et al.*, 2009; MAYER, 2023). A frequência acumulada é obtida pela soma das frequências de todos os valores da variável, menores ou iguais ao valor considerado.

Finalmente, de acordo com o Quadro 7, a “Renda Familiar” foi classificada como variável quantitativa contínua, onde os dados atribuídos a ela pertencem a um intervalo de números reais (naturais ou racionais).

2.6 Análise quali-quantitativa

A análise quali-quantitativa é uma abordagem que integra duas formas para analisar dados e usa *designs* distintos para prepará-los. A suposição dessa abordagem é que a integração de dados quali-quantitativa produz uma visão adicional além da informação fornecida separadamente pelos dados quantitativos ou qualitativos (CRESWELL; CRESWELL, 2018).

Bagdonienė e Zembyltė (2005) relacionam métodos de coleta de dados e suas aplicações pelos pesquisadores:

- Métodos da pesquisa qualitativa: entrevista, observação, análise de documento e análise de literatura científica.
- Métodos da pesquisa quantitativa: questionário ou análise estatística de dados.
- Métodos da pesquisa quali-quantitativa: questionário e entrevista, e questionário e observação.

Para Creswell e Creswell (2018) a análise quali-quantitativa manipula formas de dados que fornecem diferentes tipos de informação, tais como: dados não-estruturados no caso de dados qualitativos, e dados estruturados no caso de dados quantitativos. Os autores ainda consideram alguns pontos positivos da análise quali-quantitativa. Os dados podem ser combinados para desenvolver uma compreensão mais robusta do problema ou questões de pesquisa, explorando amplamente os diferentes tipos de dados, integrando-os.

Para Freitas e Janissek (2000) é importante enriquecer os dados quantitativos com informações qualitativas, de forma a aprimorar as considerações e relatórios de pesquisa. Nesse caso, o desafio é a busca da associação entre o quantitativo e o qualitativo, onde o procedimento exploratório ganha robustez, visto que se poderá multiplicar os dados tratados, a fim de confirmar hipóteses de pesquisa.

Creswell e Creswell (2018) propõem uma lista de verificação de procedimentos para o uso de métodos mistos:

- São fornecidas razões (ou justificativas) para usar ambos dados quantitativos e qualitativos no estudo proposto?
- O leitor tem noção do uso potencial da pesquisa mista?
- O *design* de métodos mistos está identificado?
- Os procedimentos de coleta e análise de dados são mencionados no que se refere ao *design* escolhido?

- As estratégias de amostragem para coleta de dados quali-quantitativos são mencionadas para o projeto?
- Os procedimentos de análise de dados específicos são indicados para o projeto?

Para Dawadi, Shrestha e Giri (2021) a combinação quali-quantitativa pode ser superior a um único método, pois é provável que forneça *insights* enriquecidos sobre os fenômenos de pesquisa que não podem ser totalmente compreendidos usando apenas um dos métodos (qualitativo ou quantitativo). O método quali-quantitativo pode integrar várias fontes de dados que podem ajudar a estudar problemas complexos (POTH; MUNCE, 2020). A aplicação do método quali-quantitativo consolida os dados do propósito do estudo que permite aos pesquisadores buscarem uma visão ampla de suas pesquisas, permitindo-lhes analisar os dados científicos em diferentes perspectivas (SHORTEN; SMITH, 2017).

Dawadi, Shrestha e Giri (2021) formulam um exemplo que reforça a justificativa para realizar a análise quali-quantitativa. Por exemplo, ao compreender os fatores que causam a obesidade em crianças, é necessário avaliar quantitativamente variáveis de exames associadas a dados qualitativos que suportam a análise quantitativa. Creswell e Plano Clark (2017) exemplificam questões de pesquisa que utilizam métodos mistos e destacam variáveis qualitativas e quantitativas, conforme apresentado a seguir:

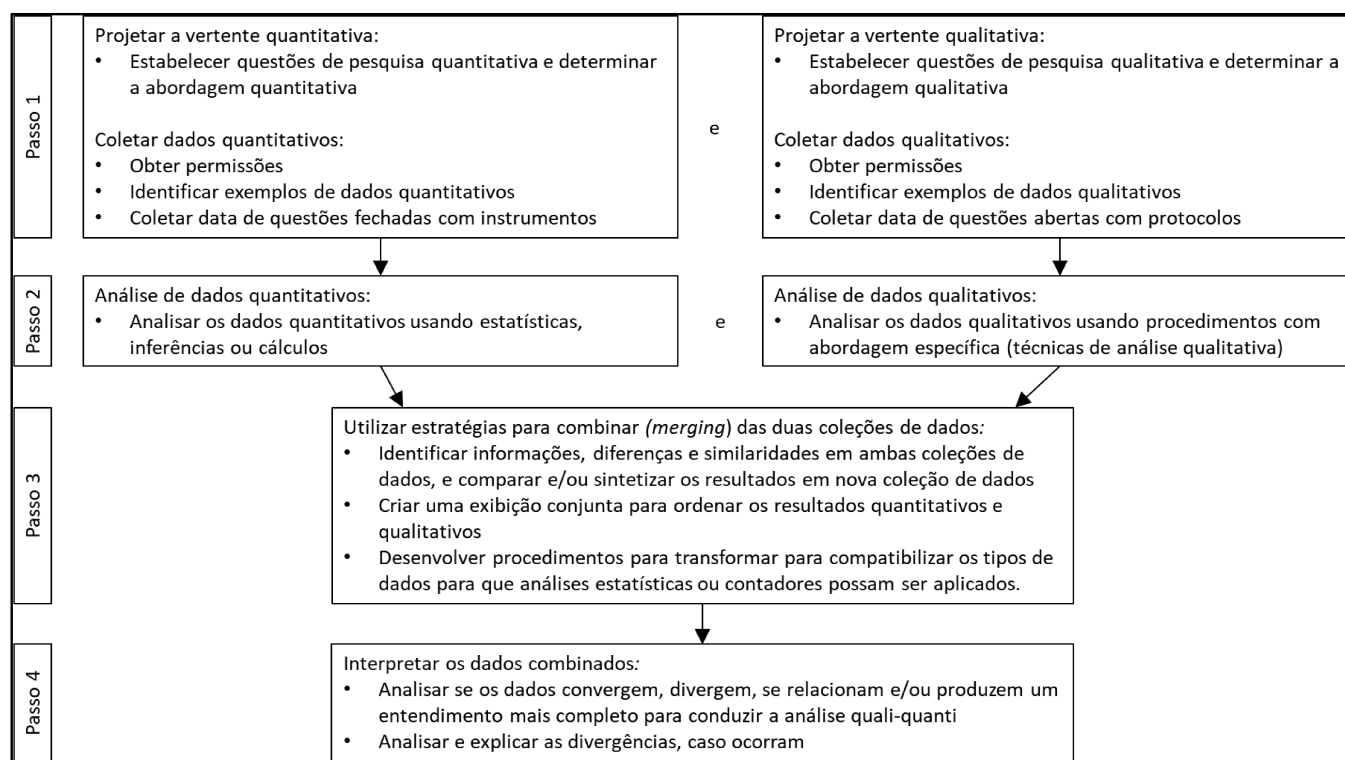
- A frequência de acidentes de trânsito (variável quantitativa) reflete as percepções dos ciclistas sobre a segurança viária (variável qualitativa) em São Paulo?
- Os alunos possuem percepções sobre o ambiente escolar (variável qualitativa) em relação aos resultados dos testes (variável quantitativa)?
- Como as entrevistas sobre satisfação no trabalho na Empresa ABCD (variável qualitativa) justificam o KPI de vendas anual (variável quantitativa) ?

Com base nas questões acima, a análise quali-quantitativa pode ser a escolha apropriada se o processo de pesquisa sugerir que os dados quantitativos ou qualitativos por si só não responderão suficientemente à questão de pesquisa. Ao formular a questão de pesquisa com o uso da análise quali-quantitativa, os métodos qualitativos e quantitativos serão combinados, se a questão não puder ser suficientemente respondida por meio de análises quantitativas ou qualitativas independentes (CRESWELL; PLANO CLARK, 2017).

De acordo com Creswell e Plano Clark (2017) diferentes projetos de pesquisa utilizam a análise quali-quantitativa como método misto. As diferenças entre os projetos de pesquisa dizem respeito ao objetivo da pesquisa, no momento da coleta de dados e à importância dada a cada tipo de dado. Os tipos de projetos de pesquisa que permitem a análise quali-quantitativa, seguem relacionados abaixo (CRESWELL; PLANO CLARK, 2017, p .74):

- **Convergente:** o pesquisador coleta e analisa dois bancos de dados separados – quantitativo e qualitativo – e então mescla os dois bancos de dados com o objetivo de comparar ou combinar os resultados. A Figura 9 representa um esquema que detalha o projeto de pesquisa convergente para análise quali-quanti.
- **Sequencial explicativo:** a coleta e análise de dados quantitativos ocorrem primeiro, seguidas pela coleta e análise de dados qualitativos. Utiliza-se este tipo de projeto especificamente para que os dados qualitativos expliquem as descobertas quantitativas. Pode ser usado para explicar os mecanismos por meio de dados qualitativos que esclarecem porquê os resultados quantitativos ocorreram e como eles podem ser explicados.
- **Sequencial exploratório:** a coleta e análise de dados qualitativos ocorrem primeiro, seguidas pela coleta e análise de dados quantitativos. Utiliza-se os dados quantitativos para testar ou confirmar descobertas qualitativas.

FIGURA 9 - Projeto de pesquisa convergente para análise quali-quanti



Fonte: Adaptado de (CRESWELL; PLANO CLARK, 2017).

2.7 Desafios em pesquisas qualitativas, quantitativas e quali-quanti

Khankeh *et al.* (2015) discutem os desafios em conduzir pesquisas qualitativas e quantitativas em saúde, tendo em vista que o conhecimento clínico consiste em ação interpretativa de fatores que envolvem a comunicação, opiniões compartilhadas e

experiências. Para os autores, os métodos quantitativos de pesquisa dificultam o acesso ao conhecimento em um domínio de aplicação. Os métodos quantitativos são aplicados em questões e fenômenos controláveis, medidos e contáveis. Por outro lado, os autores exemplificam que a pesquisa qualitativa se preocupa com o aumento da compreensão do significado de certas condições para profissionais de saúde e pacientes, e como suas relações são construídas em um determinado contexto social.

De acordo com Khankeh *et al.* (2015) a ausência de compreensão em relação ao domínio de aplicação, experiência e habilidades para fazer pesquisas qualitativas podem dificultar a análise dos dados e a produção de conhecimento científico. Ao contrário dos dados quantitativos, os dados qualitativos não possuem uma estrutura facilmente definida que possa ajudar na análise de dados, e geralmente carecem que os dados estejam organizados para facilitar a compreensão deles.

De acordo com Friese e Sybing (2023) interpretar dados não numéricos é um desafio significativo para pesquisadores qualitativos. Por exemplo, se algum pesquisador medir no mesmo dia a temperatura de duas cidades e obtiver valores distintos, poderá afirmar que uma cidade é mais quente que a outra. Outros estudos podem abordar cidadãos para saber se concordam sobre o que torna uma cidade interessante para morar. No segundo exemplo, o pesquisador precisa esclarecer as possíveis diferenças de entendimento entre as pessoas. Para responder a essas questões, a análise de dados qualitativos deve explicitar os conceitos subjetivos e socialmente construídos para que os dados produzidos sejam compreendidos (FRIESE; SYBING, 2023).

Casterlé *et al.* (2012) discutem desafios para análise qualitativa. Ressaltam para a dependência excessiva de pacotes de software qualitativos. De acordo com Casterlé *et al.* (2012, p. 3) “os pesquisadores muitas vezes não dedicam o tempo necessário para ler e reler o material, refletir sobre o que leram, tentando compreender os temas gerais e as histórias”. Os softwares podem ajudar e promover a análise qualitativa, no entanto, não devem automatizar (sugerir) quais códigos anexar aos segmentos de dados. O pesquisador deve conhecer os materiais utilizados e saber, tecnicamente, como preparar a codificação, para ampliar o seu entendimento sobre como analisar os dados em diferentes perspectivas.

Hunter *et al.* (2002) ressalta a importância de visualizar os dados qualitativos sob várias perspectivas, facilitando o pensamento multidimensional e oferecendo diferentes formas de dar significado aos dados. A visualização elaborada dos dados serve para orientar o pesquisador na compreensão da essência dos resultados da pesquisa. De acordo com Casterlé *et al.* (2012) pesquisadores qualitativos tem dificuldades em analisar dados de entrevistas qualitativas dentro de uma abordagem que utiliza a técnica de *grounded theory*.

Goertzen (2017) trata dos desafios em pesquisa quantitativa sob vários aspectos nas atividades de bibliotecários. Tais profissionais irão tratar um grande volume de informações e diferentes fontes de dados. Neste caso, um dos desafios é estabelecer métodos para a gestão de informações, conhecer a proveniência dos dados e saber usá-los de forma a suportar a tomada de decisão.

Outros desafios são relatados por Goertzen (2017). A falta de padronização entre os conjuntos de dados produzidos por diferentes ferramentas de software dificulta a interpretação do significado dos dados. Essa questão pode criar obstáculos significativos ao executar projetos quantitativos ao necessitar realizar comparações entre variáveis e valores.

Rahman (2020) traz a importância para definir os significados das variáveis utilizadas para a análise quantitativa. Trata-se da necessidade do pesquisador compreender os significados comuns do domínio da pesquisa. Cita como exemplo a aplicação da análise quantitativa envolvendo dados e variáveis, como pessoa, gênero, pessoa por gênero ou atividades por gênero. Rahman (2020) destaca a dificuldade dos pesquisadores em analisar as variáveis com ausência de conceitos e definições de aplicação em um contexto particular. A pesquisa quantitativa possibilita investigar as respostas para as perguntas que iniciam com pronomes interrogativos: com quantos, quanto, em que medida. Rahman (2020, p. 5) reforça que pesquisa quantitativa se concentra nos aspectos que podem ser quantificados, em vez de apenas descobri-los e interpretar seus significados:

[...] Os métodos quantitativos buscam regularidades nas vidas humanas, separando o mundo social em componentes empíricos chamados variáveis, que podem ser representadas numericamente como frequências ou taxas, cujas associações uns com os outros podem ser explorados por técnicas estatísticas e acessados por meio de estímulos introduzidos pelo pesquisador e medições sistemáticas. A pesquisa quantitativa concentra-se nos aspectos do comportamento social que podem ser quantificados e padronizados, em vez de apenas descobri-los e interpretar os significados que as pessoas trazem para suas próprias ações.

De acordo com Somekh e Lewin (2005) o desafio ao aplicar métodos mistos está relacionado às análises iniciadas em “trilhas” qualitativas e quantitativas separadas. Os dados de uma trilha devem ser transformados e cruzados para a outra trilha, para que possam ser analisados posteriormente. Estratégias para análise dados de métodos mistos incluem o uso dos resultados intermediários para obter um conjunto de categorias de dados. As categorias são usadas como uma estrutura para analisar os dados quali-quantitativos. Os autores destacam que um processo de codificação pode ser usado para melhorar a análise de dados quali-quantitativos.

Somekh e Lewin (2005) destacam outros desafios na prática de métodos mistos. É necessário que pesquisadores usem formas representacionais mistas para relatar os resultados da investigação científica usando um método misto. Devem ser desenvolvidas

novas apresentações textuais e tabulares, e novos mecanismos que possam capturar e rerepresentar os resultados da pesquisa com métodos mistos.

2.8 Trabalhos Correlatos

Esta seção utiliza parâmetros para avaliar os trabalhos correlatos. Os seguintes parâmetros baseiam-se nos temas abordados na introdução (considerando delimitação do problema e objetivos) e no referencial teórico: uso de ontologias para análise quali-quantitativa; anotação semântica de dados; análise quali-quantitativa com uso de artefatos para representar os dados anotados; manipulação de dados semânticos.

Para selecionar os trabalhos correlatos, optou-se pela técnica de revisão de literatura orientada pelos parâmetros definidos nesta seção (BIOLCHINI *et al.*, 2005; SNYDER, 2019). De acordo com Biolchini *et al.* (2005) a revisão de literatura estabelece atividades para recuperar, filtrar e organizar trabalhos científicos, visando o mapeamento das produções acadêmicas já existentes sobre um determinado tema, reunindo as principais conclusões sobre ele através desses estudos. Neste sentido, iniciou-se o levantamento de diferentes trabalhos científicos a partir das seguintes bases de dados: Springer, Scopus, Periódicos da CAPES e Google Scholar. A seguinte *string* de busca foi elaborada para recuperar os trabalhos: (*ontology and "semantic annotation" and "knowledge graphs" and "qualitative analysis" and "quantitative analysis"*) or (*ontology and "qualitative coding" and "qualitative analysis" and "semantic web"*).

Ao realizar o levantamento bibliográfico foi possível filtrar a produção acadêmica estabelecendo pontos convergentes e divergentes entre os trabalhos. Critérios de inclusão e exclusão dos trabalhos usaram os parâmetros de avaliação que nortearam a seleção dos trabalhos com o objetivo de eliminar os estudos irrelevantes (BIOLCHINI *et al.*, 2005). A filtragem dos trabalhos obedeceu aos seguintes critérios:

- Critérios de inclusão: uso de ontologias para anotação semântica; uso de ontologias para codificação em análise qualitativa; estudos que abordam o uso de recursos para representação do conhecimento em análise quali-quantitativa; e publicações dos últimos 10 anos.
- Critérios de exclusão: anotação de dados ou codificação por modelos diferentes de ontologias; uso de artefatos para representação do conhecimento diferentes de ontologias e grafos; anotação de dados para análise quali-quantitativa sem o suporte semântico de ontologias.

A seleção dos trabalhos correlatos foi aplicada em três etapas (BIOLCHINI *et al.*, 2005; SNYDER, 2019). Na primeira etapa, a *string* de busca foi executada e possibilitou recuperar os seguintes números de estudos: Springer (388); Scopus (336); Periódicos da CAPES (2); Google Scholar (70). Para cada base de dados foi criada uma lista dos trabalhos selecionados. Os trabalhos repetidos foram eliminados. A segunda etapa estabeleceu a filtragem por meio da leitura dos títulos e resumos obedecendo aos critérios de inclusão e exclusão, obtendo os números de trabalhos a seguir: Springer (15); Scopus (5); Periódicos da CAPES (2); Google Scholar (5). A terceira etapa resultou em uma nova filtragem dos artigos a partir da leitura de cada trabalho obedecendo aos mesmos critérios.

O Quadro 8 apresenta a lista final dos trabalhos correlatos selecionados. Em seguida, são apresentadas considerações a respeito de cada trabalho a fim de correlacioná-los aos parâmetros de avaliação.

Quadro 8 - Trabalhos correlatos

Fonte	Título	Autor(es)	Ano
CAPES	Qualitative analysis of manual annotations of clinical text with SNOMED CT	Jose Antonio Miñarro-Giménez, Catalina Martínez-Costa, Daniel Karlsson, Stefan Schulz, Kirstine Rosenbeck Gøeg	2018
SCOPUS	Linked data models for sentiment and emotion analysis in social networks	C.A. Iglesias, J.F. Sánchez-Rada, G. Vulcu, P. Buitelaar	2017
SCOPUS	Evaluation of automatic annotation by a multi-terminological concepts extractor within a corpus of data from family medicine consultations	Charlotte Siefriid, Julien Grosjeanb, Tatiana Lefebvre, Laetitia Rollin, Stefan Darmoni, Matthieu Schuers	2020
SPRINGER	Ontology for Semantic Data Integration in the Domain of IT Benchmarking	Matthias Pfaff, Stefan Neubig, Helmut Krcmar	2018
SPRINGER	Imaging Biomarker Ontology (IBO): A Biomedical Ontology to Annotate and Share Imaging Biomarker Data	Emna Amdouni, Bernard Gibaud	2018
SPRINGER	Linked Data Supported Content Analysis for Sociology	Tabea Tietz, Harald Sack	2019
SPRINGER	A qualitative and quantitative analysis of open citations to retracted articles: the Wakefield 1998 et al.'s case	Ivan Heibi, Silvio Peroni	2021
SPRINGER	Evaluating QualiCO: an ontology to facilitate qualitative methods sharing to support open science	Julian Hocker, Taryn Bipat, David W. McDonald, Mark Zachry	2021

Fonte: Elaborado pelo autor.

Hocker *et al.* (2021a) utilizaram a QualiCO para anotar dados qualitativos usando templates de metadados. Os autores avaliaram quantitativamente o reuso da QualiCO visando melhorar as tarefas de codificação. No entanto, apesar da anotação semântica pela QualiCO, os autores não abordaram como os dados anotados poderiam estar representados em GCs.

Miñarro-Giménez *et al.* (2018) utilizaram a ontologia SNOMED CT (*SNOMED Clinical Terms*) para codificar dados qualitativos de textos clínicos, a fim de identificar as fontes de divergências com o uso da SNOMED CT. A prevalência dos conceitos anotados contribuiu para revelar as divergências e o nível de compreensão da SNOMED CT para codificar os termos. Apesar dos autores terem usado a codificação e análise quantitativa, não abordaram GCs para apoiar na análise das divergências sobre os dados anotados.

Charlotte Siefridt *et al.* (2020) desenvolveram um extrator multiterminológico para anotar automaticamente textos de consultas médicas usando a SNOMED CT. A acurácia dos termos anotados foi avaliada por meio de cálculos de precisão (com o uso das funções *recall* e *F-measure*). Os autores consideraram a anotação eficaz no domínio de medicina familiar. Porém, não representaram os dados anotados em GCs, ou como a SNOMED CT poderia ser utilizada como estrutura para popular dados do estudo apresentado.

Pfaff, Neubig e Krcmar (2018) modelaram uma ontologia de serviços de Tecnologia da Informação (TI) usando o conhecimento sobre dados de questionários submetidos aos representantes de organizações de TI. Uma rotina em NLP (*Natural Language Processing*) foi implementada para acessar os conceitos da ontologia e anotar textos contidos nas respostas dos questionários. Além dos textos, indicadores (KPIs) usados para avaliar os serviços de TI foram anotados como dados quantitativos. Com os dados armazenados em uma base dados semânticos, consultas SPARQL foram elaboradas para gerar fragmentos de GCs a fim de serem analisados para responder QCs. No entanto, para anotar as fontes de dados (oriundas dos questionários) foi necessário um mapeamento de metadado específico para cada formato de dado.

Iglesias *et al.* (2017) criaram um modelo de *linked data* para análise de sentimentos e emoções em redes sociais. Os termos da análise de sentimentos foram extraídos automaticamente de textos por meio do processamento de vocabulários baseado na linguagem de marcação EmotionML¹⁰. Ao extrair os termos, a rotina de processamento fez a anotação semântica deles usando a ontologia DBpedia. Complementarmente, a prevalência dos conceitos analisados foi tabulada e anotada. Ao final, os autores demonstraram fragmentos de GCs para permitir a análise de sentimentos com dados quali-quantitativos. Porém, os autores não detalharam como a ferramenta foi construída, dificultando o entendimento mais específico sobre como a anotação foi realizada.

Heibi e Peroni (2021) aplicaram a técnica de citação aberta como metodologia para coletar e anotar os dados de artigo científico. Na esteira da web semântica, os autores utilizaram a ontologia *Citation Typing Ontology* (CiTO) e o recurso COCI (*Crossref open DOI-*

¹⁰ <https://www.w3.org/TR/emotionml/>

to-DOI) para definir metadados que descrevem citações em publicações científicas, capturando o objetivo de um autor ao se referir a outro trabalho acadêmico. Os autores realizaram a anotação manual das citações de acordo com conceitos existentes na CiTO. Do ponto de vista quantitativo os autores agruparam áreas temáticas conforme conceitos presente na ontologia, para analisar a prevalência das diferentes áreas que citaram o artigo científico. Apesar de terem utilizado dados anotados semanticamente em análises quali-quantitativas, não representaram os dados anotados em GCs ou como as análises poderiam ser recuperadas por eles.

Tietz e Sack (2019) apresentaram como os princípios de *linked data* e as técnicas de NLP podem ser usados para estruturar, publicar e analisar texto em linguagem natural para pesquisas sociológicas. Os autores descreveram a categorização de textos para análise como problema de pesquisa qualitativa para desenvolver novas teorias ou testar as existentes. Eles propuseram a implementação da anotação semântica com o uso de ontologias, que explicitamente estruturam o conhecimento e definem relacionamentos entre conceitos e indivíduos. Os autores usaram a anotação semântica para o processo de análise de conteúdo em sociologia. No entanto, os autores não integram dados quali-quantitativos em um mesmo GC.

Amdouni e Gibaud (2018) usaram tecnologias semânticas para aprimorar a recuperação de dados de biomarcadores e aumentar a interoperabilidade no campo biomédico. Representaram semanticamente os conceitos no domínio de biomarcadores, que se fundamentam na articulação de três eixos: qualidade medida, ferramenta de medição e a ferramenta de decisão. A ontologia *Imaging Biomarker Ontology* (IBO) foi desenvolvida para suportar a anotação dos dados coletados dos biomarcadores. Os autores ilustraram a utilidade do IBO para anotar dados quantitativos e qualitativos da coleção TCGA (*The Cancer Genome Atlas*) e como os dados podem ser recuperados por GCs. Os autores não especificaram um processo ou método para modelagem ontológica ou como a anotação semântica pode permitir a geração dos GCs com dados quali-quantitativos populados.

O Quadro 9 um resumo dos trabalhos correlatos de acordo com os parâmetros usados para avaliação.

Quadro 9 - Resumo dos trabalhos correlatos

Título	Uso de Ontologia	Anotação Semântica	Análise quali-quantitativa	Manipulação de dados semânticos
Qualitative analysis of manual annotations of clinical text with SNOMED CT	SNOMED CT	Manual	Categorização dos textos clínicos usando codificação. Cálculo da frequência e quantidade de termos por categoria.	Não faz uso de banco de dados semânticos.
Linked data models for sentiment and emotion analysis in social networks	DBpedia, EmotionML	Automática	Categorização de termos/sentenças com base em vocabulários. Cálculo de precisão dos termos categorizados.	Não faz uso de banco de dados semânticos. Porém, exemplifica instruções em SPARQL para recuperar grafos.
Evaluation of automatic annotation by a multi-terminological concepts extractor within a corpus of data from family medicine consultations	SNOMED CT	Automática	Categorização de terminologias. Cálculo de precisão da relação dos termos categorizados.	Não faz uso de banco de dados semânticos.
Ontology for Semantic Data Integration in the Domain of IT Benchmarking	ITBM	Automática/NLP	Categorização de terminologias com base em metadados. Uso de SPARQL para responder questões relacionadas aos KPI anotados.	Utiliza a ferramenta Fortiss Ontology Search Engine para recuperação de grafos via SPARQL.
Imaging Biomarker Ontology (IBO): A Biomedical Ontology to Annotate and Share Imaging Biomarker Data	IBO	Manual	Categorização de terminologias com base em classes da ontologia IBO. Uso de SPARQL para responder questões referentes aos valores dos biomarcadores.	Utiliza a ferramenta CORESE ¹¹ para recuperação de grafos via SPARQL.
Linked Data Supported Content Analysis for Sociology	DBpedia	Automática/NLP	Identificação de documentos da área da sociologia usando uma lista de entidades compatíveis com a DBpedia. Número de entidades anotadas.	Utiliza Blazegraph para recuperação de grafos via SPARQL.
A qualitative and quantitative analysis of open citations to retracted articles: the Wakefield 1998 et al.'s case	CiTO	Manual	Identificação de documentos citados usando entidades compatíveis com a CiTO. Análise do número de citações com base nas entidades anotadas.	Não faz uso de banco de dados semânticos.
Evaluating QualiCO: an ontology to facilitate qualitative methods sharing to support open science	QualiCO	Manual	Uso da tarefa de codificação e preenchimento dos metadados com o uso da QualiCO. Cálculo da pontuação de precisão para cada tarefa de codificação.	Não faz uso de banco de dados semânticos.

Fonte: Elaborado pelo autor.

2.9 O Domínio do Estudo de Caso - Aplicação de Estratificação de Riscos a Dados Psicoterápicos

De acordo com Miller, Reardon e Safi (2001) a estratificação de risco é um processo estatístico que apoia o tratamento em saúde mental ajustado ao risco atribuído ao paciente.

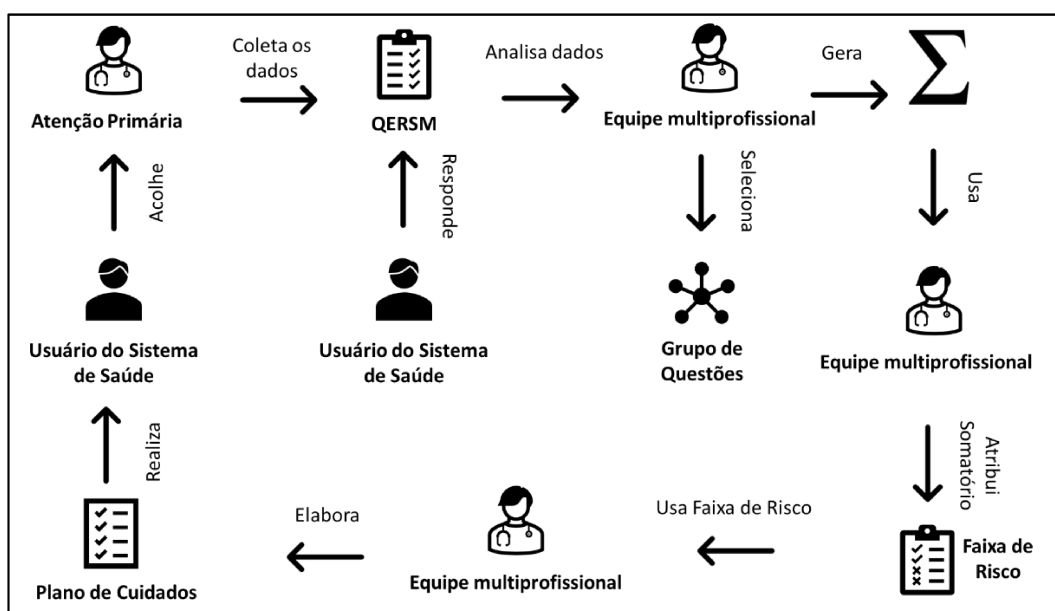
A SESA (Secretaria da Saúde do Estado do Paraná) é um dos órgãos no Brasil que aplica a estratificação por meio de Questionário de Estratificação de Risco em Saúde Mental (QERSM), com questões agrupadas por transtornos mentais classificados pela OMS. Cada

¹¹ <https://project.inria.fr/corese/>

questão é formulada seguindo critérios diagnósticos da CID 10 (Classificação Internacional de Doenças) e do DSM IV-R (*Diagnostic and Statistical Manual*) determinados pelas relações de sinais e sintomas, fatores agravantes e atenuantes identificados por síndromes psicopatológicas (PARANÁ, 2021). Profissionais em saúde mental submetem as questões aos pacientes para avaliar os sinais e sintomas, que são pontuados de acordo com o grau de gravidade, obedecendo a uma escala de pontuação.

A SESA considera situações de risco, tais como: risco de suicídio, risco de heteroagressão e risco de exposição moral. Essas situações apoiam a estratificação dos riscos em diferentes níveis, por exemplo: baixo, moderado ou alto. Riscos baixos a moderados orientam intervenções breves e encaminhamentos para guiar os profissionais no planejamento de cuidados. A forma como a estratificação de risco encontra-se definida aqui, pode ser consultada pelo “Instrutivo para Aplicação do Instrumento de Estratificação de Risco em Saúde Mental” (PARANÁ, 2021). De acordo com Paula (2019), a estratificação obedece a um processo iniciado no acolhimento do paciente pela atenção primária, que utiliza instrumentos QERSM para coletar os dados. Em seguida, uma equipe multiprofissional analisa os dados para atribuir o risco ao paciente. Finalmente, o nível de risco atribuído orienta a elaboração do plano de cuidados. A Figura 10 visa exemplificar um processo de estratificação de risco.

FIGURA 10 - Processo de estratificação de riscos



Fonte: Elaborado pelo autor.

Transtornos mentais causados pela ansiedade podem ser avaliados pela estratificação de risco, visando mitigar as causas de outros transtornos. De acordo com a *Mayo Clinic* algumas das causas da ansiedade são: trauma, acúmulo de estresse, transtorno de

personalidade, ou ainda, parentes próximos com transtorno de ansiedade, ou que fazem uso de drogas ou álcool (MAYO, 2018).

Em tratamentos de ansiedade, a análise da condição mental dos pacientes envolve tanto dados estruturados, extraídos dos instrumentos de avaliação, quanto dados não-estruturados (e.g. evolução dos tratamentos ou textos de sessões terapêuticas) (MELO; LOURENÇO, 2020). Neste contexto, a manipulação de estruturas de dados distintas pode ser um desafio. Dependendo do ambiente de atuação, redes de atenção psicossocial ou hospitais, especialistas de domínio podem necessitar analisar dados quali-quantitativos com o interesse de direcionar a assistência à saúde, a fim de revisar metas e resultados esperados em diferentes intervenções, conforme orientações do Ministério da Saúde (BRASIL, 2014).

A Terapia Cognitivo-Comportamental (TCC) tem sido acolhida como técnica que norteia tratamentos psicoterápicos em ansiedade. De acordo com Melo e Lourenço (2020), no domínio da TCC a ansiedade se expressa em quatro dimensões: através das emoções, do comportamento, dos pensamentos e do corpo. O tratamento deve abranger todas essas dimensões por intermédio de intervenções emocionais, comportamentais, sociais e cognitivas, que devem constar no desenho qualitativo do caso clínico, que estrutura e norteia as sessões de tratamentos por meio da TCC (MELO e LOURENÇO, 2020). O profissional responsável pelo tratamento, registra os dados das evoluções terapêuticas usando o desenho qualitativo do caso clínico. Essa atuação prática da TCC abre caminho para o desenvolvimento de pesquisas baseadas em evidências na área da Psicologia.

Especificamente, a TCC pode ser aplicada como prática baseada em evidência no âmbito da *Clinical Expertise*, em que dados de um caso específico ou de conjunto de casos podem ser utilizados para o desenvolvimento de pesquisas científicas. A experiência clínica inclui as habilidades básicas da prática clínica, bem como a experiência do profissional individual, a fim de gerar evidências relevantes de pesquisa (WIETEN, 2018).

Miller, Reardon e Safi (2001) ressaltam a necessidade de compreender os dados e documentar as variáveis identificadas na estratificação para evitar interpretações ambíguas. Ontologias podem enriquecer a semântica e compreensão de conceitos em saúde mental, a fim de representar o conhecimento. A representação do conhecimento em saúde em mental pode enriquecer o entendimento sobre o domínio, sistematizando os diferentes dados utilizados (BENFARES, IDRISI; HAMID, 2018; BRENAS, SHIN e SHABAN-NEJAD, 2019; YAMADA et al., 2018).

O próximo capítulo aborda a metodologia *Design Science Research* para estruturar o método HOMOGENiSE, fundamentando metodologicamente as etapas do método e o conhecimento para executar esta pesquisa.

3 METODOLOGIA

3.1 Uso da *Design Science Research* como metodologia de pesquisa

Conforme discutido na Seção 1.2, o objetivo geral desta pesquisa consiste em desenvolver e aplicar etapas de um método fundamentado em modelagem ontológica e anotação semântica para aprimorar a elaboração de análises quali-quantitativas. Considera-se que é uma pesquisa aplicada, e, portanto, fundamenta-se metodologicamente na DSR. A execução do método permitirá gerar artefatos que constituem os resultados da pesquisa, tais como, ontologia, templates de metadados, arquivos em formato RDF e consultas SPARQL.

Pesquisas científicas que necessitam produzir artefatos, podem ser vistas como pesquisas muito específicas ou que não contribuem para o desenvolvimento científico. Porém, observa-se que a pesquisa em Ciência da Informação no PPGGOC, atualmente representado pela linha de pesquisa GETIC, à qual esta pesquisa faz parte, pode ser enriquecida com o uso da metodologia abordada neste Capítulo. Linhas de pesquisas semelhantes à GETIC envolvem a produção de artefatos, que visam evidenciar experimentos realizados e que se fundamentam em teorias ou modelos conceituais da área de Ciência da Informação, com o objetivo de solucionar problemas dentro das organizações.

Pesquisas que utilizam artefatos possuem dificuldade de representar ou justificar a geração de conhecimento novo, devido à aplicação de conhecimentos e tecnologias já existentes para criar os artefatos, no entanto, existe um esforço para concebê-los juntamente com a necessidade de resolver problemas teóricos. A dificuldade aqui relatada é recorrente em projetos que envolvem a elaboração de diferentes artefatos, principalmente, artefatos de caráter tecnológico, não caracterizando pesquisa científica (BAX, 2015).

A DSR tem sido utilizada na Ciência da Informação para fundamentar metodologicamente diferentes trabalhos científicos, oferecendo meios ao pesquisador para obter conhecimento ao executar a pesquisa, mesmo que realizem atividades do projeto científico que resultem em produtos ou artefatos. A DSR utilizada como abordagem metodológica envolve uma reflexão tanto sobre o processo de construção dos diferentes artefatos, quanto sobre como esse processo ao produzir o artefato resulta ou gera conhecimento científico.

Com base nas discussões acima, torna-se possível definir a DSR da seguinte forma (HEVNER; CHATTERJEE, 2010, p. 5):

A DSR é um paradigma de pesquisa em que um projetista responde a perguntas relevantes aos problemas humanos por meio da criação de artefatos inovadores, contribuindo assim com novos conhecimentos ao corpo de evidências científicas. Os artefatos projetados são úteis e fundamentais na compreensão desse problema.

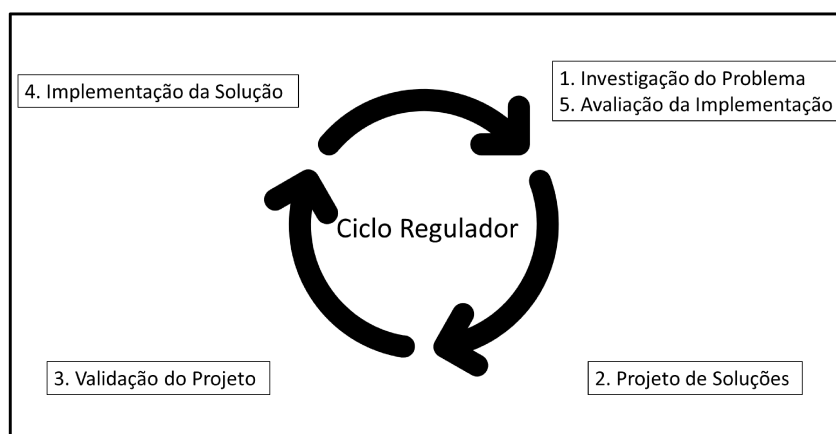
Wieringa (2009) descreve sobre a conexão entre conhecimento e prática, mostrando que pesquisadores podem produzir conhecimento científico projetando coisas úteis. No entanto, sem orientações adicionais, pesquisadores iniciantes que utilizam a DSR tendem a identificar problemas práticos com questões de conhecimento, o que pode levar a projetos de pesquisa metodologicamente incorretos. Para resolver um problema prático, o mundo real é alterado, mas para resolver um problema de conhecimento, adquirimos conhecimento sobre o mundo sem necessariamente mudá-lo. Na DSR, esses dois tipos de problemas são aninhados mutuamente, mas esse aninhamento não deve obscurecer o fato de que sua solução de problemas e justificativa de solução são diferentes.

De acordo com Barbosa (2016) a DSR teve como fundamento inúmeros trabalhos, que definiram estruturas para realizar pesquisas científicas, ou quadros teórico-metodológicos (GREGG *et al.*, 2001; SIMON, 1966; NUNAMAKER, 1994; VASHINAV, 2001; HEVNER *et al.*, 2004; WIERINGA, 2009 apud BARBOSA, 2016).

Bax (2015) ressalta a importância de a metodologia orientar o pesquisador. Quando a metodologia falha em orientar o pesquisador, tem como consequência situações indesejáveis para uma ciência social aplicada: “ou se produz o artefato e não se produz conhecimento, ou se produz conhecimento sem se produzir o artefato”. Um resultado indesejado perde-se no rigor científico, mesmo amparado pelo compromisso apodítico de relevância e severidade na execução da pesquisa. Na Ciência da Informação a concepção de artefatos pode ser questionada quanto à sua natureza científica, ou pelo menos se oferece rigor científico. Isso devido à falta de uma metodologia adequada, que explicita os problemas práticos e as questões de conhecimento.

Para resolver esta questão, Wieringa (2009) propõe o ciclo regulador, conforme pode ser visto na Figura 11. Cada uma das cinco etapas do ciclo é útil para conduzir a parte prática das pesquisas, através de um problema prático, ou gerar novos conhecimentos através de respostas a questões de conhecimento. Ao aplicar a DSR é necessário classificar os problemas de pesquisa classificando-os em problemas práticos e problemas teóricos (ou questões de conhecimento). É possível organizar os problemas em uma estrutura lógica e que consiste em um ciclo regulador, envolvendo atividades de (a) Investigação do problema; (b) Projeto de soluções; (c) Validação do Projeto; (d) Implementação da solução e (e) Avaliação da implementação.

FIGURA 11 - Ciclo Regulador de Wieringa (2009)



Fonte: Elaborada pelo autor, baseado em Wieringa (2009).

A investigação do problema é uma etapa que visa pesquisar um problema prático a partir de uma questão de conhecimento, pois envolve descrever, explicar e até mesmo a prever as possíveis soluções. Essa etapa de investigação requer do pesquisador um melhor entendimento do problema, e os requisitos da solução estabelecidos devem ser levados em consideração.

A segunda etapa é sobre o projeto de soluções, motivada por uma questão prática, sabendo que o projeto define uma solução prática para o problema a ser investigado e definido na primeira etapa. O projeto obtido, ao final desta etapa, não representa necessariamente uma solução definitiva. Podem ser necessárias melhorias no projeto ao serem realizadas as fases de validação e implementação. A solução definida nesta etapa pode ser um artefato preliminar, o qual será aprimorado em outros ciclos (WIERINGA, 2009).

Na etapa de validação do projeto, verifica-se se ele atende aos objetivos estabelecidos na etapa de investigação do problema. Esta etapa busca responder questões de conhecimento relativas à satisfação dos critérios identificados na investigação do problema, e pode determinar a existência de outras soluções que possam satisfazer os critérios ou até mesmo a abrangência da solução em outros contextos de pesquisa.

A quarta etapa diz respeito à implementação da solução projetada a partir das etapas anteriores. Trata-se de uma etapa que se preocupa com a prática no projeto de pesquisa e que artefatos são gerados. Os artefatos são validados na última etapa do ciclo, que pode ser reiniciada com novos problemas a serem investigados. Tais problemas podem ser relacionados a melhorias nos artefatos criados ou na necessidade de novos artefatos para complementar a solução.

Os problemas práticos a serem investigados no início de cada ciclo podem ter níveis diferentes de complexidade. O projeto de um novo artefato, por exemplo, demanda um esforço bem maior do que melhorias em um artefato existente. A solução destes problemas práticos,

independentemente do nível de complexidade, pode ser facilitada com a decomposição deles. Para se fazer essa decomposição do problema, baseando-se ainda no método DSR de Wieringa (2009), é definida uma estrutura aninhada do problema, que consiste em decompor a questão de projeto (ou problema de pesquisa) em problemas práticos (P) e problemas teóricos ou questões de conhecimento (K), mutuamente aninhados. Com base nessa decomposição, projeta-se um fluxo da solução do problema de pesquisa com um conjunto de subproblemas, os quais são classificados em práticos e teóricos. Assim, o pesquisador compreende de forma aprofundada o problema que está tratando e define os objetivos específicos que precisa alcançar com vistas a atingir o objetivo geral. O método de Wieringa prevê ainda, a definição de subtipos para os problemas assim definidos:

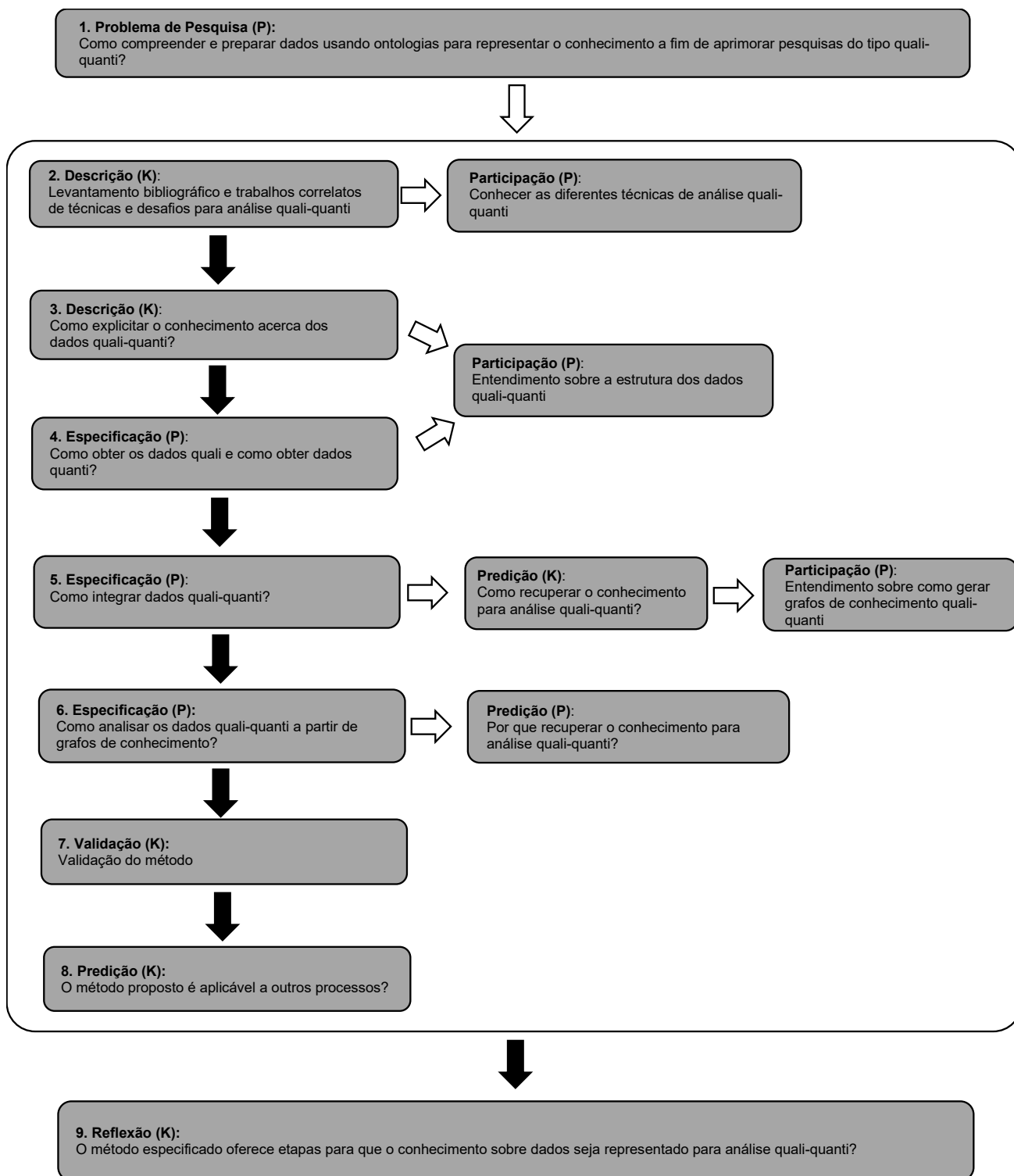
- Problemas práticos:
 - Especificação(P): especificação e desenvolvimento de uma solução;
 - Participação (P): busca de soluções com a participação da equipe;
 - Discussão (P): apresentação dos artefatos, reuniões com os membros envolvidos;
- Problemas teóricos ou questões de conhecimento (K):
 - Descrição (K): descoberta de informações necessárias à investigação.
 - Avaliação (K): observação e diagnóstico dos fatos.
 - Predição (K): estimativa dos efeitos de uma solução.
 - Validação (K): validação de soluções e comparação com critérios.
 - Reflexões (K): questionamentos sobre as lições aprendidas e a geração de conhecimento.

3.2 Aplicação da DSR ao problema de pesquisa

A partir do problema de pesquisa apresentado na Seção 1.1, aplica-se o ciclo regulador de Wieringa (2009), e projeta-se uma solução que visa responder o mesmo. A solução foi projetada e implementada com base nos critérios da DSR. Espera-se estabelecer critérios para avaliação e validação da metodologia. Para o projeto de solução foi utilizada a estrutura aninhada e o ciclo regulador de Wieringa (2009).

A estrutura aninhada utilizada por esta pesquisa é apresentada na Figura 12 e contém a decomposição do problema de pesquisa em subproblemas práticos e teóricos. O problema de pesquisa se encontra denotado como ponto de partida para a decomposição dele. Os demais quadros representam os subproblemas quebrados de acordo com a estrutura aninhada, onde os quadros marcados em cinza representam os subproblemas principais, com os quadros à direita representando a decomposição destes subproblemas em problemas menores. Em cada quadro estão estabelecidos subproblemas K ou subproblemas P.

FIGURA 12 - Estrutura aninhada do problema de acordo com a DSR de Wieringa



Fonte: Elaborada pelo autor

A estrutura aninhada é executada de acordo com o ciclo regulador, o que significa que os problemas podem ser refinados a cada execução do ciclo. A partir do problema de pesquisa estabelecido, “Como compreender e preparar dados usando ontologias para representar o

conhecimento a fim de aprimorar pesquisas do tipo quali-quantitativo”, foram definidos os subproblemas ilustrados na Figura 12, e que estão detalhados abaixo. Os subproblemas listados serão solucionados na pesquisa visando também atingir os objetivos específicos (vide Seção 1.2).

- Levantamento bibliográfico das técnicas de análise quali-quantitativa.
 - (Problema de Conhecimento). Conhecer técnicas de análise quali-quantitativa que possam ser utilizadas para analisar os dados oriundos do domínio de aplicação.
- Como obter os dados qualitativos?
 - (Problema Prático). Realizar a tarefa de codificação como análise qualitativa com uso de software CAQDAS. Exportar e transpor os dados para um formato tabular.
- Como obter os dados quantitativos?
 - (Problema Prático). Os dados quantitativos são obtidos por meio de estruturas de dados tabulares, contendo variáveis com valores numéricos ou variáveis que seja possível aplicar técnicas para análise quantitativa.
- Como explicitar o conhecimento acerca dos dados quali-quantitativos?
 - (Problema de Conhecimento). A explicitação do conhecimento se baseia no levantamento bibliográfico, entendimento do domínio de aplicação, reuniões com especialistas, modelagem ontológica a partir dos conceitos compreendidos. A partir da ontologia, fazer a anotação semântica dos dados quali-quantitativos usando a técnica SDD.
- Como integrar dados quali-quantitativos?
 - (Problema Prático). Os conjuntos de dados quantitativos são anotados semanticamente utilizando o mesmo ID do objeto de estudo, que deve ser encontrado no conjunto de dados qualitativos.
- Como definir o conhecimento dos dados qualitativos?
 - (Problema de Conhecimento). Utiliza-se classes de objetos modeladas no padrão REFI-QDA e na ontologia QualiCO para anotar semanticamente os dados por meio do SDD.
- Como analisar os dados quali-quantitativos a partir de GCs?
 - (Problema Prático). Para solucionar este subproblema é implementada a anotação semântica dos dados quali-quantitativos, geração dos dados em formato RDF, a carga dos dados semânticos, e o uso de soluções de software para representar os GCs.

- Por que recuperar o conhecimento sobre os dados do contexto desta pesquisa?
 - (Problema de Conhecimento). A proposta é utilizar dados quali-quantitativos anotados semanticamente. Os dados poderão ser analisados por meio de GCs para evidenciar estudos científicos específicos.
- Especificação de um método para aplicação da pesquisa.
 - (Problema Prático). A aplicação da DSR abre caminho especificar o método utilizado nesta pesquisa, acrescentando aspectos de validação da modelagem conceitual e controle da evolução dos artefatos. A modelagem conceitual gera um modelo ontológico resultado da compreensão do domínio de aplicação e explicitação dos dados pela anotação semântica. Os *templates* de metadados relacionam-se entre si a partir de anotações preenchidas por profissionais e especialistas de domínio tendo como base o modelo ontológico. A cada execução do *script* para gerar de GCs os *templates* podem ser modificados. É possível que novos dados sejam preenchidos nos *templates* modificando a versão anterior dos documentos para satisfazer novas análises sobre os dados anotados. Com isso, também se propõe construir uma matriz de rastreabilidade. A matriz permite organizar os diferentes documentos (ou artefatos) minimizando o impacto das alterações nos documentos. A cada iteração a matriz de rastreabilidade pode ser versionada a fim de controlar as mudanças realizadas. A definição do método é apresentada no Capítulo 4.
- Validação do método.
 - (Problema de Conhecimento). O método HOMOGENiSE foi aplicado em um estudo de caso real utilizando dados sintéticos extraídos de artigos científicos que possuem dados quali-quantitativos referentes ao domínio de aplicação utilizado nesta pesquisa.
- O método é aplicável a outros domínios?
 - (Problema de Conhecimento). O domínio em que esta pesquisa é aplicada possibilita que as etapas do método sejam reutilizadas em outras áreas de pesquisa (tais como, as áreas educacional ou financeira), ou outras áreas da saúde, a fim de permitir que especialistas do domínio possam ser orientados na análise da evolução do estado de saúde dos pacientes em relação aos riscos estratificados e respectivos planos de cuidados estabelecidos.

No capítulo seguinte o método HOMOGENiSE é especificado com etapas para a modelagem ontológica do conhecimento em pesquisas quali-quantitativas e detalha as técnicas utilizadas para atingir os objetivos apresentados no capítulo introdutório.

4 O MÉTODO HOMOGENiSE

Este capítulo define o método HOMOGENiSE como uma contribuição desta pesquisa, especificando as etapas para fundamentar a modelagem ontológica e anotação semântica para aprimorar a elaboração de análises quali-quantitativas. Apesar deste capítulo não detalhar as técnicas ou tecnologias utilizadas para atingir um determinado resultado, o Capítulo 5 aplica as etapas especificadas por meio de um estudo de caso, detalhando os resultados da aplicação do método.

As etapas do HOMOGENiSE adotam características de um processo iterativo e incremental. Ao executar o método, torna-se possível a construção de uma ontologia para anotar *datasets* de interesse por especialistas do domínio, e para explorar (ou inferir) conhecimento sobre um determinado domínio de aplicação. O método foi elaborado com o objetivo de responder à questão definida nesta pesquisa e outras questões que envolvam explicitar o conhecimento presente nos *datasets*.

Larman e Basili (2003) elaboram um breve histórico do desenvolvimento de *software* iterativo e incremental. Nesse trabalho os autores citam Harlan Mills, um líder do pensamento em Engenharia de *Software* da década de 1970 que trabalhou na divisão de sistemas computacionais da IBM. Mills promoveu o desenvolvimento de projetos iterativos, onde aconselha começar a desenvolver a partir de estruturas de controle de nível superior para baixo, gerando uma sequência de sistemas intermediários de código e subespecificações funcionais para que, a cada etapa, cada sistema (intermediário) possa ser verificado e testado. Define o processo iterativo como um refinamento, mas ele não esclareceu como tratar o tamanho da especificação inicial.

A ideia que fundamenta um processo iterativo é gerar artefatos (ou aplicações) de forma incremental, permitindo que especialistas aproveitem o que foi aprendido durante o desenvolvimento de versões anteriores e incrementais de entrega do artefato. As etapas do processo começam com uma implementação simples de um subconjunto dos requisitos para aprimorar iterativamente a sequência de versões em evolução até que a aplicação completa seja implementada. A cada iteração, modificações no projeto podem ocorrer a partir da adição de novas funcionalidades (BASILI; TURNER, 1975 *apud* LARMAN; BASILI, 2003).

Posteriormente, novas discussões surgiram em torno das definições sobre o processo iterativo. O processo de desenvolvimento usado para criar grandes aplicações deve ser tratado como um processo de aprendizagem e comunicação. Ou seja, um processo de aprendizagem cíclico com alta atenção às habilidades das pessoas, visão comum e problemas de comunicação, em vez de visualizar o esforço como um "processo de fabricação" sequencial. Posteriormente, na década de 90, discutiu-se sobre a melhoria do uso do

processo iterativo tendo como ideias a entrega de soluções com prazos menores e a chance de aumentar a taxa de sucesso. Já no início dos anos 2000, Larman e Basili (2003) citam estudos que reforçam o controle das versões de cada iteração e o aprendizado sobre o desenvolvimento das etapas anteriores.

A partir das discussões citadas anteriormente, observa-se a necessidade de controlar a evolução das aplicações desenvolvidas com base no processo iterativo. No entanto, apesar dessas citações estarem fundamentadas em estudos nas áreas da Engenharia de Software e Métodos Ágeis, o processo iterativo também é aplicado em outras áreas de estudos.

Rittle-Johnson e Siegler (2001) aplicaram experimentos que permitiram observar que os conhecimentos conceitual e procedimental podem se desenvolver de forma iterativa. Os experimentos foram conduzidos com alunos que estavam aprendendo sobre frações decimais. No Experimento 1 observou-se que o conhecimento conceitual inicial das crianças previu ganhos no conhecimento procedimental, e ganhos no conhecimento procedimental previram melhorias no conhecimento conceitual. No Experimento 2, a quantidade de suporte para a representação correta do problema foi manipulada experimentalmente, e as manipulações levaram a ganhos no conhecimento procedimental. Os autores constataram que os conhecimentos conceitual e procedimental se desenvolveram iterativamente.

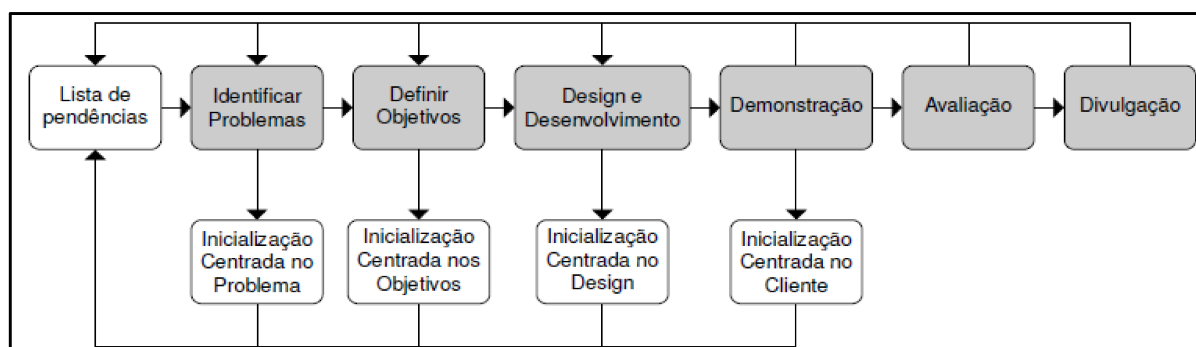
De acordo com Cervone (2011) o processo iterativo pode ser aplicado na gestão ágil de projetos. A partir da revisão contínua e prazos curtos de *design*, a equipe do projeto é mais capaz de adaptar rapidamente os projetos a ambientes em rápida evolução.

Aizenbud-Reshef *et al.* (2006) estabelecem um modelo de rastreabilidade que ajuda as partes interessadas a compreenderem as muitas associações e dependências que existem entre os artefatos criados durante um projeto de desenvolvimento de software. Os autores ainda reforçam que a extensão da prática de rastreabilidade é vista como uma medida da qualidade do sistema e maturidade do processo e é exigida por muitos padrões. Discutem também a importância da revisão da prática de rastreabilidade à luz dos processos ágeis e iterativos. Outros autores discutem a prática da rastreabilidade de artefatos juntamente com processos ágeis (DURAISAMY; ATAN, 2013; JEONG; CHO, 2018; KODALI, 2015).

Gonçalves (2020) propõe o método Odin baseado em um processo iterativo composto por conjunto de atividades voltadas para a execução, em um intervalo de tempo, de um conjunto específico de tarefas. A cada iteração, cria-se uma versão aprimorada do grafo RDF, validada pelos atores envolvidos. Estabelece também quatro tipos de atores que interagem com o processo. A Figura 13 representa as etapas do método Odin. Abaixo seguem as descrições para cada ator:

- Especialistas de Domínio (equipes de profissionais ou pesquisadores): são os interessados em realizar a preparação e organização das coleções de dados de estudo de um dado domínio.
- Ontologistas: profissionais que possuem o conhecimento na criação, evolução e pesquisa de ontologias.
- Desenvolvedores: colaboram no processo com o seu conhecimento em tecnologia da informação, através da criação de pequenos programas ou até mesmo de soluções mais complexas.
- Cientistas de Dados: especialistas em ciências de dados, que poderão auxiliar na obtenção de informações a partir dos dados ingeridos, utilizando técnicas estatísticas, de aprendizado de máquina e outros recursos associados ao tratamento de grandes volumes de dados.

FIGURA 13 - Etapas do Método Odin



Fonte: GONÇALVES (2020).

A utilização do HOMOGEniSE é motivada por um problema que pode ser resolvido a partir da explicitação do conhecimento e anotação semântica de *datasets*. A proposta se baseia em uma abordagem “*bottom-up*”, usando dados e SDDs a partir de uma ontologia existente (ou servindo de base para criar uma ontologia), sendo possível integrar de forma complementar a uma abordagem “*top-down*”, onde a sua aplicação é definida na próxima seção.

O método Odin aborda a implementação de SDDs para modelagem ontológica, e sugere a geração de GCs para demonstrar como especialistas de domínio podem navegar pelos conceitos e dados anotados, a fim de preparar artefatos para divulgação de pesquisas científicas. O método Odin é fundamentado metodologicamente pela DSR, e possui as etapas definidas seguindo a proposta de um método ágil. O Odin é resultado de estudos desenvolvidos em projeto de pesquisa desenvolvido no PPGGOC/UFMG, cujos objetos de estudo (ontologias, anotação semântica, manipulação de GCs) também fazem parte da definição do HOMOGEniSE.

Diante das propostas semelhantes entre os métodos, o método HOMOGENiSE reutiliza as etapas do método Odin. As técnicas listadas a seguir enriquecem as etapas propostas pelo Odin. São elas:

- Descrição do “minimundo” que esclarece o domínio a ser modelado contemplando artefatos que possuem dados quali-quantitativos.
- Explicitação do conhecimento com base em datasets.
- Fundamentação da ontologia de domínio a partir de uma ontologia de alto nível.
- Controle da evolução dos artefatos produzidos a cada etapa do método.
- Validação da ontologia com base em QCs que visam apoiar na elucidação de problemas de pesquisa ou sobre o conhecimento dos dados utilizados como datasets em um estudo.

As seções seguintes descrevem as etapas do método HOMOGENiSE que contemplam as técnicas citadas acima.

4.1. Descrever MiniMundo e Identificar Problemas

Esta etapa é centrada na identificação do problema a ser tratado na iteração, e que abrirá caminho para definir os objetivos para implementar a solução. Para identificar o problema é necessária a descrição de um minimundo que define de forma sucinta o domínio de aplicação. Por meio do minimundo é possível dimensionar a complexidade em termos de tipos de dados ou até mesmo o processo de aquisição deles. Além disso, essa definição ajuda na compreensão dos dados a serem anotados e deve contar com a participação de especialistas de domínio para contribuir e validar o minimundo. Para ampliar o entendimento desta etapa, artefatos devem ser contemplados a fim de caracterizar os dados qualitativos e quantitativos. Por exemplo, questionários, textos, áudios ou até mesmo vídeos.

O resultado das definições servirá para construir uma primeira versão do modelo da ontologia, possibilitando a aplicação de uma abordagem *top-down*. Ao compreender melhor o domínio com a definição do minimundo, ontologistas podem pesquisar outras ontologias para que possam ser reutilizadas a fim de enriquecer o conhecimento sendo modelado.

A identificação de problemas reutiliza a ideia proposta pelas metodologias de engenharia de ontologias ao determinar a elaboração de QCs. O minimundo fundamenta as QCs, que são propostas pelos especialistas de domínio. Recomenda-se a execução desta etapa com o apoio de alguma metodologia de engenharia de ontologia, visando formalizar e orientar a elaboração das QCs.

Desse modo, as QCs são importantes para as próximas etapas do HOMOGENiSE. Elas norteiam a análise dos dados e motivam a anotação semântica. Além disso, o minimundo

pode ser redefinido ao executar as etapas posteriores do método. É possível que novos conceitos, dados ou outras QCs sejam formuladas. Neste caso, as demandas por redefinição do minimundo devem ser priorizadas na Lista de Pendências, para que possa ser discutido se devem ou não ser utilizadas.

4.2 Definir Objetivos

Esta etapa define os objetivos da solução visando orientar a etapa de Concepção e Desenvolvimento. A lista de objetivos tem como base o minimundo, as QCs, artefatos especificados na etapa anterior, e que instigam o desenvolvimento das etapas posteriores. Dentre os objetivos, sugere-se elaborar um planejamento para executar as próximas etapas, e um cronograma pode ser construído.

Os objetivos podem determinar as propriedades e informações necessárias para gerar os GCs, e que visam orientar as respostas das QCs. Alguns dos objetivos definidos podem ser adiados e acrescentados à Lista de pendências. A definição dos objetivos conduz a anotação semântica, pois, ao ser realizada, é necessário preencher os templates de metadados com informações inerentes ao domínio definido.

Ao realizar a definição dos objetivos é possível que o entendimento sobre o domínio possa ser verificado, e, com isso, os atores envolvidos nesta etapa podem suscitar a alteração do minimundo descrito na etapa anterior, havendo a necessidade de executar novamente as etapas iniciais especificadas pelo método.

4.3 Concepção e Desenvolvimento

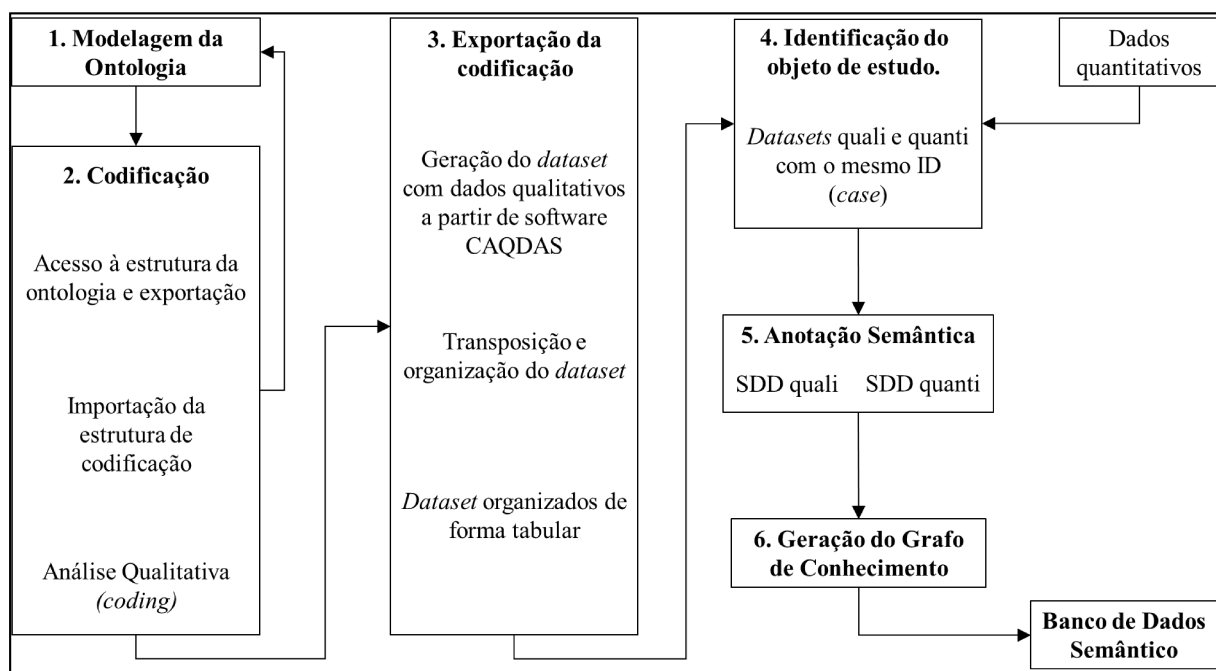
A etapa de Concepção e Desenvolvimento utiliza o minimundo para iniciar a modelagem da ontologia com a estruturação de conceitos, classes e propriedades para responder às QCs. Essas atividades conduzem o *design* da solução em conjunto com a anotação semântica.

Editores de ontologias devem suportar a modelagem ontológica. Sugere-se, ainda, o uso de tecnologias para anotação semântica, tais como: planilhas eletrônicas ou editores de textos. Os metadados preenchidos consolidam a anotação semântica com o uso da técnica SDD, e são considerados artefatos produzidos nesta etapa.

Datasets devem ser preparados para serem anotados semanticamente. A anotação semântica proposta neste método suporta a geração de GCs no formato RDF. Sugere-se a ingestão dos GCs em repositórios de dados semânticos, para que possam ser acessados por instruções a fim de recuperar informações e orientar os especialistas do domínio na análise dos dados.

O método HOMOGENiSE determina, nesta etapa, um processo fundamentado em modelagem ontológica para compreender e preparar dados para análises quali-quantitativas por meio de GCs. O processo é representado na Figura 14, que organiza seis atividades definidas da forma que segue.

FIGURA 14 - Atividades para elaboração da etapa de Concepção e Desenvolvimento



Fonte: Elaborado pelo autor.

- **Atividade 1 - Modelagem da ontologia**: Visa compreender o domínio por meio da modelagem ontológica. É uma atividade que se fundamenta nas discussões de Berthold *et al.* (2010) e Schröer, Kruse, Gómez (2021) ao definir a compreensão em etapas iniciais de um processo para análise de dados científicos. Sugere-se nesta atividade o uso de alguma metodologia para engenharia de ontologia, bem como o uso de alguma ontologia de fundamentação, com objetivo de formalizar o reuso de outras ontologias, a estruturação de classes, propriedades, restrições e hierarquias de classes. Além disso, nesta atividade, é necessário o registro de especificação da ontologia por meio da representação gráfica, a geração do glossário de termos, axiomas e afins. O *minimundo* é usado para fundamentar termos (ou conceitos e significados), representados por sujeitos ou respectivos complementos verbais dentro de uma oração. A eliciação dos termos deve suportar o entendimento e a criação das classes de conceitos e predicados (*objects properties*) da ontologia.
- **Atividade 2 – Codificação**: Um esquema de codificação é estruturado, nesta atividade, para análise qualitativa dos dados. A estrutura hierárquica da ontologia fornece conceitos para que sejam utilizados como códigos (*codes* para o *coding*), formando um *codebook* dentro da

ferramenta de CAQDAS. As definições no entorno da seção 2.4 (cf. Referencial Teórico) fundamentam esta atividade. A codificação é usada como estratégia para permitir a análise de dados qualitativos com o uso de códigos classificados em categorias, para associá-los aos dados analisados. É uma atividade que analisa *datasets* a partir da leitura reflexiva de textos com a aplicação desta técnica. Esta atividade (no processo aqui proposto), visa identificar e descrever dados qualitativos, com o objetivo de compreender, selecionar e construir temas a partir de dados com características semelhantes; investigar padrões por meio da homogeneidade das categorias de dados. A ontologia criada na Atividade 1, modela conceitos organizados em uma taxonomia, por onde o especialista de domínio (codificador) tem acesso ao conhecimento para codificar dados qualitativos (SCHENSUL; LECOMPTE, 2012). Ao utilizar os editores de ontologias, é possível exportar os conceitos para serem importados para a ferramenta de CAQDAS, a fim de formar o *codebook* da codificação. Os conceitos exportados ofereceram suporte aos especialistas fazerem a codificação dedutiva (coding dedutivo), a partir da estrutura consensuada da ontologia como o conhecimento utilizado *a priori* para a codificação. No entanto, durante a tarefa de codificação podem surgir novos códigos (*coding* indutivo) que posteriormente, serão acrescentados na estrutura da ontologia a fim de enriquecê-la e possa ser reutilizada em novas codificações.

- Atividade 3 – Exportação da Codificação: A codificação é exportada pela ferramenta de CQDAS. Os dados qualitativos codificados são selecionados e transpostos de linhas para colunas. Ao transpor os dados, os códigos são caracterizados por variáveis (colunas) no *dataset*, permitindo que estejam organizados para executar as atividades posteriores. Esta atividade expõe a necessidade de unificar os dados qualitativos em um formato tabular e compatível com tabelas de dados originados de *datasets* quantitativos. Conforme sugerido por Driscoll *et al.* (2007) o método misto inclui projetos que modificam uma forma de dados para outra, sendo necessária a definição do procedimento de transformação de dados qualitativos. Esta atividade é proposta em consonância com a preparação de dados como tarefa que antecede a análise de dados a fim de evitar equívocos (ou análises enviesadas) (PATON, 2019; PYLE, 1999).
- Atividade 4 – Identificação do objeto de estudo: O objeto do estudo deve ser identificado univocamente por uma variável em ambos os *datasets* qualitativos e quantitativos. Por exemplo, se o objeto do estudo for um paciente a variável é o identificador do paciente e que deve ser anotada semanticamente por SDDs quali-quantitativos. Esta atividade identifica o objeto de estudo com o objetivo de combinar ou integrar as coleções de dados quali-quantitativos (DRISCOLL *et al.*, 2007).

- Atividade 5 – Anotação Semântica: A técnica SDD é proposta nesta atividade. Os templates de metadados são preenchidos com o apoio dos especialistas de domínio. A ontologia definida na Atividade 1 é utilizada para anotar semanticamente os dados preparados nas atividades iniciais (2, 3 e 4). As QCs servem de base para o *design* da solução e execução da técnica SDD.
- Atividade 6 – Geração do Grafo de Conhecimento: Os GCs são gerados em formato RDF por meio da execução do script `sdd2rdf`, tendo como entradas os metadados e os *datasets*.

Finalmente, esta etapa de Concepção e Desenvolvimento prevê a validação da ontologia. A cada iteração deve ocorrer a validação por meio de uma ontologia de topo, com objetivo de legitimar o modelo ontológico gerado a partir da explicitação do conhecimento sobre os *datasets*.

4.4 Demonstração

Com os GCs gerados na etapa anterior é necessário demonstrar o atendimento às QCs, apresentando aos envolvidos o resultado da iteração. É possível demonstrar o atendimento às QCs por meio de consultas ao grafo. Outras análises realizadas pelos cientistas de dados ou por ferramentas criadas pelos desenvolvedores podem complementar a demonstração. A demonstração é considerada uma etapa que realiza a validação dos artefatos produzidos na etapa de Concepção e Desenvolvimento, se a modelagem ontológica e anotação semântica produzidos atendem a navegação pelos GCs.

A partir de um repositório de dados semânticos, os GCs podem ser extraídos em pequenas e grandes quantidades de dados. Essa extração de dados pode permitir que fragmentos do conhecimento sejam analisados com o objetivo de orientar os especialistas de domínio na inferência de novos conhecimentos. É possível que os grafos analisados sejam criticados, motivando os especialistas na reformulação do conhecimento modelado e explicitado nas etapas de identificação de problemas, definição de objetivos e, concepção e desenvolvimento. Daí a importância de manter o controle da evolução e da matriz de rastreabilidade dos artefatos produzidos para análises dos dados anotados semanticamente.

4.5 Avaliação

Nesta etapa, verifica-se se as etapas anteriormente especificadas estruturam procedimentos e artefatos que visam responder à investigação da pesquisa e QC. Utiliza-se formas de avaliação que visam descrever a capacidade do método em explicitar o conhecimento sobre *datasets* e representá-los conceitualmente por meio de ontologias e grafos é um indicador da validade da abordagem proposta, sobretudo se tal capacidade for

validada no contexto de diferentes estudos ou pesquisas. Avalia-se aqui, não somente a capacidade de anotar semanticamente *datasets*, mas também, a possibilidade de permitir que o método, ao ser executado, possa orientar especialistas de domínio na melhoria da análise de *datasets*. Neste caso, ao realizar pesquisa quali-quantitativa a partir do HOMOGENiSE, hipóteses podem ser confirmadas nesta etapa visando avaliar a estrutura e conteúdo dos artefatos produzidos. A avaliação por meio de hipóteses de pesquisas pode motivar a reformulação das técnicas e atividades executadas anteriormente. Assim, torna-se possível surgir pendências que influenciam na reformulação das etapas anteriores. Além da avaliação apresentada, indicadores são propostos nesta etapa a fim de afirmar a capacidade do método:

- Número de códigos criados. Avalia o reuso dos conceitos da ontologia. Espera-se o número menor de códigos criados a cada execução do método, sugerindo que a ontologia possui completude em conceitos usados para a codificação de dados quali;
- Tempo de codificação. Avalia o tempo de acesso aos conceitos da ontologia. Espera-se o tempo menor, sugerindo que os especialistas de domínio tiveram maior facilidade em compreender e reutilizar a estrutura da ontologia na codificação;
- Tempo de execução. Avalia o tempo de execução do método entre as etapas 1 e 4. Espera-se o tempo menor, sugerindo que o método facilita produzir os artefatos para divulgação.

4.6 Divulgação

De acordo com Gonçalves (2020) a divulgação é um subprocesso que representa o estágio final de uma iteração, no qual as descobertas são compartilhadas com os interessados nas análises dos dados anotados semanticamente. A divulgação pode ser realizada por meio de *dashboards* publicados na internet, por exemplo, onde os dados são combinados em gráficos usando filtros de campos que possibilitem as análises de interesse sobre os dados anotados (SILVA; BAX, 2019).

Outros tipos de divulgação podem ser acadêmicos por meio de revistas científicas. Esta etapa não ocorre em todas as iterações, pois depende do estágio de evolução da integração semântica. Nesta etapa, os problemas, os objetivos da solução e o *design* podem ser revistos de acordo com o *feedback* do público.

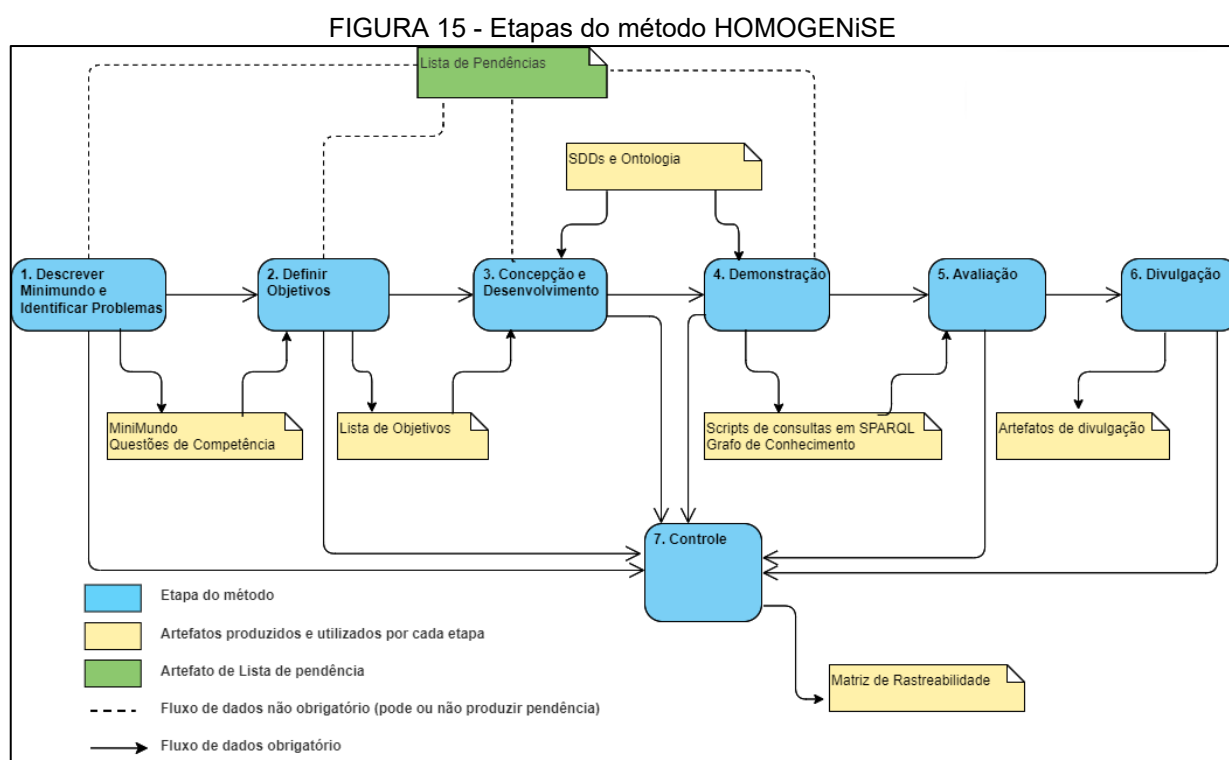
4.7 Controle

Nesta etapa é sugerido o controle da evolução da execução do HOMOGENiSE. A cada iteração os atores envolvidos produzem diferentes artefatos em cada etapa do método, tornando necessário o controle da produção dos artefatos. O controle pode ser por meio de

técnicas de versionamento e configuração de software, ou até mesmo o uso de matriz de rastreabilidade. Para registrar as mudanças a cada iteração são sugeridas as seguintes informações:

- Atores participantes da mudança. Esta informação é preenchida com os nomes completos dos atores.
- QC respondida. Este campo permite relacionar a solução dada com a QC que motivou o desenvolvimento.
- Data/Hora da mudança. Data e Hora da mudança realizada no projeto de anotação semântica.
- Descrição da mudança. Descrição sucinta sobre a mudança descrevendo as tarefas realizadas na modelagem.

A matriz de rastreabilidade contribui para auxiliar no controle da evolução e permite documentar o impacto da mudança entre os diferentes templates de metadados e a construção da ontologia. A Figura 15 representa as etapas do HOMOGENiSE que foram adaptadas com técnicas que permitem controlar a evolução dos artefatos gerados durante a execução do método.



Fonte: Elaborado pelo autor.

No Capítulo 5, o método é aplicado em um estudo para análise de dados quali-quantitativos de tratamentos psicoterápicos. A aplicação tem o propósito de responder à questão de pesquisa e atingir o objetivo geral definidos neste trabalho.

5 APLICAÇÃO DO HOMOGENiSE EM UM ESTUDO DE CASO COM DADOS PSICOTERÁPICOS

Este capítulo apresenta a execução das etapas do método HOMOGENiSE aplicado em um estudo de caso para análise quali-quantitativa de tratamentos psicoterápicos.

A especificação do método HOMOGENiSE teve início em agosto de 2020 com término em setembro de 2021, no Hospital Vila Verde Saúde Mental (HVVSM) em parceria com o PPGGOC/UFMG (cf. ANEXO 1). Essa parceria caracterizou a primeira etapa desta tese até o exame de qualificação, e que permitiu fundamentar os conceitos sobre a estratificação de risco e planos de cuidados em saúde mental, visando apenas a análise quantitativa. Como continuidade desta pesquisa, no domínio da saúde mental, dois projetos foram conduzidos no Centro Universitário UniAcademia em Juiz de Fora, Minas Gerais (cf. ANEXO 2 e ANEXO 4, respectivamente), sendo o primeiro iniciado em maio de 2022, com término em novembro de 2022.

O segundo projeto, iniciado em março de 2023 foi multidisciplinar entre os cursos de graduação de Engenharia de Software e de Psicologia, e possibilitou consolidar e executar as etapas do HOMOGENiSE. O curso de Psicologia possui uma Clínica-escola que atende à sociedade de Juiz de Fora e região com inúmeros projetos de extensão na área da saúde mental. Os alunos de Engenharia de Software apoiaram no desenvolvimento das rotinas computacionais usadas para aplicação tecnológica do processo proposto. Os alunos de Psicologia exerceram atividades como especialistas do domínio sob a supervisão de professores do próprio curso.

Com a execução do projeto em 2023, o principal interesse dos especialistas de domínio foi a geração de evidências de tratamentos conduzidos pela TCC, a fim de avaliar as habilidades básicas da prática clínica e experiência do profissional responsável pelo tratamento, conforme preconizado pela *Clinical Expertise*. Dois artigos científicos (MELO; LOURENÇO, 2020; REMOR, 2000) foram usados para extrair dados sintéticos da análise quali-quantitativa. Foram realizadas reuniões de acompanhamento e repasse de informações sobre como o método HOMOGENiSE deveria ser aplicado. As reuniões forneceram fundamentos para compreender e interpretar o domínio de aplicação e quais artefatos puderam ser utilizados para redefinir o minimundo descrito no início da pesquisa. Os artigos científicos utilizados para extração dos dados sintéticos foram lidos por todos os membros do projeto de pesquisa a fim de compreender o teor dos tratamentos ali descritos e identificar os dados qualitativos e quantitativos.

5.1 Descrever MiniMundo e Identificar Problemas

O minimundo a seguir contém definições formuladas em colaboração com o HVVSM. Apesar dos estudos desenvolvidos no HVVSM terem permitido gerar grafos com dados oriundos de pacientes internados em um hospital, os grafos produzidos para análise quali-quantitativa sob a ótica da TCC também possuem conceitos correlacionados, tais como: questionário para estratificação de risco, nível de risco, questões, tipos de questões, profissional responsável, dados sociodemográficos, sinais e sintomas, pontuação (ou escala) por questão e pontuação final. A nova definição do minimundo é descrita a seguir:

“Os especialistas possuem interesse em analisar dados quantitativos a partir da aplicação da estratificação de risco em ansiedade, que fornece dados sobre a avaliação de sinais e sintomas do paciente. Dados qualitativos são analisados e extraídos a partir da aplicação da técnica TCC. A TCC possui uma estrutura que orienta a obtenção de dados para a evidência do tratamento. Essa estrutura é organizada por meio de um desenho qualitativo clínico, que possui os seguintes dados: dados relevantes do paciente; crenças centrais; crenças intermediárias; estratégias de enfrentamento; padrões identificados (pensamentos, emoções e aspectos comportamentais); aspectos situacionais (fatores que aconteceram durante a semana e impactaram nos padrões observados do paciente); aspectos contextuais (como o paciente lidou com fatores de seu contexto a partir de seu padrão cognitivo); e aspectos emocionais (emoções ativadas e regulação do estado emocional). A estratificação de risco em ansiedade é estruturada a partir da definição de um questionário contendo questões organizadas em grupos. O profissional responsável (psicólogo ou terapeuta) informa os seguintes dados do paciente (dados sociodemográficos): Identificação do Paciente, Data de Início do Tratamento, Nome Completo do Paciente, Sexo e Idade. Ato contínuo, o profissional responsável inicia a estratificação de riscos identificando a gravidade dos sinais e sintomas apresentados e as condições atuais de vida do paciente (aspectos psicossociais) com base nos seguintes dados (dados clínicos): se fuma (SIM ou NÃO), se utiliza substância psicoativa (SIM ou NÃO), tipo de substância psicoativa utilizada e frequência do uso da substância psicoativa (BAIXA, MÉDIA ou ALTA). O paciente é entrevistado a partir das questões organizadas da forma descrita, sendo a resposta limitada em duas opções: SIM ou NÃO. Para cada resposta SIM uma pontuação é atribuída, caso contrário a pontuação é 0. Após todas as questões respondidas ocorre um somatório das pontuações gerando a pontuação final. A pontuação final é atribuída a uma faixa de risco que determina o nível de risco do paciente em BAIXO, MÉDIO ou ALTO. A mensuração da presença e gravidade dos parâmetros dos sinais e sintomas tem a finalidade de auxiliar na condução do tratamento em ansiedade”.

Os *datasets*, com dados sintéticos e oriundos da estratificação, encontram-se anotados neste estudo de caso. A organização dos dados e os códigos-fonte utilizados aqui podem ser vistos em um repositório controlado criado durante a execução desta pesquisa. Os *datasets* podem ser consultados no repositório de controle de versão utilizado para esta pesquisa¹².

A SABI^o foi sugerida como metodologia para execução deste estudo de caso. O minimundo abriu caminho para suportar a aquisição do conhecimento do domínio, e uma lista de QCs foram elaboradas para serem respondidas por meio da ontologia (COSTA *et al.*, 2023; FALBO, 2014). De acordo com Falbo (2014) a SABI^o e UFO se complementam, pois, a metodologia pressupõe a adoção de uma fundamentação ontológica. Desta forma, uma vez que o HOMOGENISE também preconiza a fundamentação ontológica, justifica-se adotar a SABI^o neste estudo de caso bem como a UFO.

Ao executar o estudo de caso, reuniões de acompanhamento do projeto serviram para refinar o entendimento do domínio e verificar se as QCs elaboradas faziam sentido na execução do projeto. Ao compreender o domínio, surgiu a necessidade de estruturar conceitos para codificar os dados qualitativos. Assim, os especialistas estabeleceram um desenho qualitativo com base na TCC para estruturar os conceitos existentes nos tratamentos psicoterápicos (*cf.* definição em Seção 2.9). Além disso, para prosseguir na elaboração das QCs, foi necessário estruturar e consensuar o desenho qualitativo do caso clínico¹³. Após essas atividades, as QCs foram elaboradas.

Conforme discutido por Dawadi, Shrestha e Giri (2021), justifica-se a análise quali-quantitativa, neste estudo de caso, a partir das QCs enumeradas a seguir, sendo necessário avaliar quantitativamente variáveis associadas a dados qualitativos. As QCs seguem listadas abaixo:

1. Houve reduções nos sintomas apresentados no desenvolvimento do processo terapêutico? Se sim, quais foram?
2. Quais foram as estratégias de intervenção de maior impacto nos pensamentos, emoções e comportamentos da(o) paciente no desenvolver do processo? Quais procedimentos se destacaram?
3. Quais as relações entre pensamentos, emoções e crenças da(o) paciente?
4. Quais as demandas e objetivos apresentados pela(o) paciente? De que maneira o processo alcança essa necessidade?
5. Quais os fatores interferiram nos resultados do processo? Foram de cunho particular ou contextual?

¹² <https://github.com/evaldo/homoginise.git>

¹³ <https://tinyurl.com/y7f74c8f>

6. Quais valores influenciam no engajamento do paciente para o tratamento? Quais as funções atribuídas a eles?
7. Quais os sinais, positivos e negativos, de engajamento da(o) paciente diante do processo?
8. Qual a prevalência dos sintomas que influenciam na hipótese clínica?
9. Qual a prevalência dos códigos que influenciam na validação da crença central?
10. Quais intervenções realizadas no tratamento permitiram a redução do nível e valor do risco em ansiedade?

A Lista de Pendências a seguir foi criada para orientar as atividades das próximas etapas do método, e foram registradas após reuniões com os especialistas de domínio:

- Revisar a QC 1 se os dados quantitativos estavam tabulados em questionários distintos (questionário inicial e final), com os respectivos níveis de riscos registrados;
- Modelar os conceitos de engajamento do paciente na ontologia. Engajamento do paciente não foi descrito no minimundo;
- Na próxima reunião discutir quais dados podem ser considerados parte do processo de tratamento, a fim de modelar conceitos na ontologia que permitam responder a QC 5.

A Matriz de Rastreabilidade foi revisada e o impacto da mudança foi em três artefatos produzidos nesta etapa: Minimundo; QCs; e Lista de Pendências. O versionamento ocorreu no repositório do projeto (<https://github.com/evaldo/homoginise>).

5.2 Definir Objetivos

Os artefatos gerados na etapa anterior (minimundo, QCs e desenho qualitativo do caso clínico) serviram para nortear os objetivos do estudo de caso. Apesar da definição dos objetivos necessitar da identificação dos problemas, ao iniciar a escrita dos objetivos, é possível que novas pendências surjam, havendo a necessidade de rever os artefatos anteriores. Os objetivos estabelecem propósitos para estruturar o restante do estudo de caso, aplicar técnicas e tecnologias visando responder às QCs. A seguir são listados os objetivos:

- Criar o cronograma de execução do estudo de caso. O Quadro 10 apresenta o cronograma criado para execução do projeto de pesquisa realizado no UniAcademia (*cf.* ANEXO 4). Conforme mencionado anteriormente, o projeto de pesquisa fundamentou a execução do estudo de caso com base em dados sintéticos.

Quadro 10 - Cronograma para execução do estudo de caso

Atividade/Mês	3	4	5	6	7
1. Modelagem da ontologia ONTRISCAL	X	X			
2. Codificação dos dados qualitativos		X			
3. Seleção dos dados (sintéticos) quantitativos			X		
4. Preparação dos dados quali-quantitativos			X		
5. Anotação semântica			X	X	
6. Geração dos GCs				X	
7. Análise quali-quantitativos				X	X
8. Relatório final (respostas às questões de competência)					X

Fonte: Elaborado pelo autor.

- Permitir que os especialistas do domínio entendam o preenchimento dos templates para anotação semântica dos dados a fim de fundamentar a modelagem ontológica.
- Detalhar o minimundo em conceitos e propriedades e como se relacionam.
- Pesquisar tecnologias para apoiar na modelagem ontológica e visualização dos GCs.
- Preparar o repositório para controle das atualizações dos artefatos.

A Lista de Pendências criada na etapa anterior, foi revisada e todas as pendências sanadas pelos especialistas de domínio. Nesta etapa, duas novas pendências foram identificadas. São elas:

- Separar os especialistas de domínio em grupos. Um grupo para apoiar na modelagem da ontologia, e outro para codificar os dados qualitativos a partir do mês 4. Prazo de uma semana para resolução desta pendência;
- Confirmar a disponibilidade do especialista de domínio “líder” que irá coordenar a revisão do domínio durante a modelagem da ontologia e codificação dos dados qualitativos. Prazo de uma semana para resolução de desta pendência.

A Matriz de Rastreabilidade foi revisada e o impacto da mudança foi em dois artefatos produzidos nesta etapa: Objetivos e Lista de Pendências. O versionamento ocorreu no repositório do projeto (<https://github.com/evaldo/homoginise>).

5.3 Concepção e Desenvolvimento

Os artefatos produzidos nas etapas anteriores servem como *inputs* para executar esta etapa. As atividades especificadas no processo representado pela

4 são aplicadas a seguir. As atividades executadas são:

Atividade 1 - Modelagem da ontologia: O minimundo possui sentenças com termos substantivos e verbos que fundamentam a modelagem da ontologia. Outros artefatos também foram utilizados aqui. Termos originados do desenho qualitativo do caso clínico possuem conceitos que foram modelados como códigos para a tarefa de codificação. Neste caso, seriam códigos dedutivos, uma vez que o esquema de codificação seria originado da elicitação dos termos no minimundo e do desenho qualitativo. Abaixo, segue a lista de termos (em negrito) e concepções verbais (sublinhados) presentes no minimundo e que norteiam a modelagem da ontologia:

1. **A estratificação de riscos**
2. é estruturada a partir de um **questionário** contendo
3. **questões** organizadas em grupos de questões.
4. Os **grupos de questões** são classificados por **sintomas** de **transtornos mentais**.
5. O **profissional responsável** (**médico, enfermeira, psicólogo**) pela **estratificação de riscos** informa:
 - a. **Identificação do Paciente;**
 - b. **Data de Admissão, Data de Internação;**
 - c. **Nome Completo do Paciente, Sexo e Idade.**
6. O **profissional responsável** identifica a **gravidade** dos **sinais e sintomas** apresentados;
7. O **profissional responsável** identifica as **condições atuais de vida do paciente** (**aspectos psicossociais**) com os dados:
8. O **profissional responsável** identifica se o paciente usa tabaco;
9. O **profissional responsável** identifica associa o transtorno mental identificado ao **grupo de CID**;
10. O **profissional responsável** identifica se o paciente utiliza substância psicoativa;
11. O **profissional responsável** identifica o **tipo de substância psicoativa** utilizada;
12. O **profissional responsável** identifica a **Frequência do uso da substância psicoativa** (**BAIXA, MÉDIA ou ALTA**);
13. O **profissional responsável** identifica se o **paciente** foi internado no último ano e **quantas vezes** foi internado no último ano;
14. O **paciente** é entrevistado a partir das questões organizadas da forma descrita, sendo a **resposta** *limitada em duas opções*: **SIM ou NÃO**;
15. Para cada **resposta SIM** uma **pontuação** é atribuída, caso contrário (**NÃO**) a **pontuação** é 0;

16. Após todas as **questões respondidas** ocorre um **somatório das pontuações gerando a pontuação final**;
17. A **pontuação final** é **atribuída a uma faixa de risco**
18. **A Faixa de Risco** determina o **nível de risco** do paciente em **BAIXO, MÉDIO** ou **ALTO**;
19. A **estratificação de riscos** tem a finalidade de auxiliar na escolha dos **parâmetros** que vão definir o nível em que ocorrerá a **assistência em saúde**;
20. O nível de risco determinado orienta ações terapêuticas que são desenvolvidas em sessões;
21. Durante as sessões ocorrem avaliações clínicas para obtenção de dados qualitativos a fim de formular modelo cognitivo de acordo com a seguinte estrutura:
 - a. crenças centrais
 - b. crenças intermediárias
 - c. estratégias de enfrentamento
 - d. padrões identificados (pensamentos, emoções e aspectos comportamentais)
 - e. aspectos situacionais (fatores que aconteceram durante a semana e impactaram nos padrões observados do paciente)
 - f. aspectos contextuais (como o paciente lidou com fatores de seu contexto a partir de seu padrão cognitivo)
 - g. aspectos emocionais (emoções ativadas e regulação do estado emocional).

A lista acima contribuiu para definir as classes de conceitos (termos substantivos) e propriedades (verbos). Após isso, foi possível reestruturar a lista anterior formando outra lista com o propósito de representar as relações entre os conceitos e refinar o entendimento para criação de uma primeira versão da ontologia ONTRISCAL. O entendimento do minimundo e desenho qualitativo foi explorado para fundamentar termos que fazem referência a nomes e/ou aos seus respectivos significados. Os termos podem ser representados por um sujeito ou respectivos complementos verbais dentro de uma oração. Abaixo segue a lista de termos (em negrito) e complementais verbais (sublinhados) presentes nos artefatos citados:

1. O **paciente** possui **dados pessoais**.
2. O **profissional responsável** informa os **dados sociodemográficos**.
3. O **paciente** apresenta **gravidade de sinais e sintomas**.
4. O **paciente** possui **dados de internação**
5. O **paciente** utiliza algum tipo de **substância psicoativa**.
6. O **paciente** utiliza com **frequência a substância psicoativa**.

7. O **paciente** faz uso de tabaco.
8. O **paciente** possui **histórico de internação**.
9. O **profissional de saúde** informa os **dados clínicos**.
10. **Questões** formulam um **grupo de questões**.
11. **Transtornos mentais** se relacionam com **sinais e sintomas**.
12. **Sinais e sintomas** determinam os **grupos de questões**.
13. O **paciente** responde **questões**.
14. Uma **questão** possui duas **opções de respostas**: “SIM” ou “NÃO”.
15. A **opção de resposta** “SIM” recebe uma **pontuação**.
16. A **pontuação final** é determinada por **somatório** das **pontuações** respondidas com “SIM”.
17. O **profissional de saúde** calcula a **pontuação final**.
18. O **paciente** recebe a **pontuação final**.
19. A **profissional de saúde** atribui uma **faixa de risco** à **pontuação final**.
20. A **faixa de risco** determina o **nível de risco**
21. O **nível de risco** é definido em **BAIXO, MÉDIO** ou **ALTO**.
22. O planejamento da **assistência em saúde** considera dados **sociodemográficos** e **dados clínicos**, a presença e a **gravidade dos sinais e sintomas** apresentados.
23. **Nível de Risco** orienta **Ação Terapêutica**
24. **Ação Terapêutica** estabelece **Modelo Cognitivo**
25. **Tratamento psicoterápico** é um tipo de **Ação Terapêutica**

Os Quadro 11 e Quadro 12 especificam as classes de conceitos, as propriedades que os relacionam e restrições. A partir dos quadros mencionadas, torna-se possível implementar a primeira versão da ontologia de forma mais objetiva. Os quadros podem ser úteis para ontologistas ou até mesmo especialistas de domínio com conhecimento mais avançado para modelar a primeira versão da ontologia.

Quadro 11- Conceitos e propriedades extraídas do detalhamento do minimundo

Domínio	Propriedade	Range
Paciente	Possui	Dado Pessoal
Profissional Responsável	Informa	Dado Sociodemográfico
Profissional Responsável	Avalia	Dado Clínico
Gravidade de Sinal e Sintoma	Define	Nível da assistência em saúde
Paciente	apresenta	Gravidade de Sinal e Sintoma
Paciente	possui	Dado de Internação
Paciente	faz uso de	Tipo de Substância Psicoativa
Paciente	utiliza frequência tipo substância psicoativa	Frequência do Tipo de Substância Psicoativa
Paciente	faz uso de	Tabaco
Questão	formula	Grupo de questões

Domínio	Propriedade	Range
Transtorno mental	relaciona com	Sinal e Sintoma
Quantidade de Internações no ano	relaciona com	Internação no último ano
Sinal e Sintoma	determina	Grupo de questões
Paciente	responde	Questão
Questão	possui	Opção de Resposta
Opção de Resposta	recebe	Pontuação
Paciente	possui	Pontuação Final
Pontuação Final	é atribuída a	Faixa de Risco
Faixa de Risco	determina	Nível de Risco
Nível de Risco	Orienta	Ação Terapêutica
Ação Terapêutica	estabelece	Modelo Cognitivo
Tratamento Psicoterápico	é um tipo de	Ação Terapêutica

Fonte: Elaborado pelo autor.

Quadro 12 - Restrições entre as classes de conceitos detalhados do minimundo.

Conceito-Origem	Restrição	Conceito-Destino
1. Paciente	Paciente é uma Pessoa	Pessoa
2. Profissional Responsável	Profissional responsável é uma Pessoa	Pessoa
3. Paciente	Paciente não é Profissional Responsável	Profissional Responsável
4. Paciente	Profissional Responsável atende Paciente	Profissional Responsável
5. Paciente	É observada somente em Paciente	Transtorno
6. Depressão	Depressão é um tipo de Transtorno	Transtorno
7. Transtorno De Ansiedade	Transtorno de Ansiedade é um tipo de Transtorno	Transtorno
8. Paciente	Paciente responde questão	Questão
9. Paciente	Paciente pode ter alguma doença ou comorbidade, e apresentar um sinal e sintoma.	Comorbidade, Doença. Sinal e Sintoma.
10. Questão	Pontuação é atribuída à Questão	Pontuação
11. Questão	Tem Pontuação Atribuída por algum Profissional Responsável	Profissional Responsável
12. Faixa De Risco Baixo	Pontuação para Faixa de Risco Alto está entre 1 e 39	-
13. Faixa De Risco Médio	Pontuação para Faixa de Risco Alto está entre 40 e 69	-
14. Faixa De Risco Alto	Pontuação para Faixa de Risco Alto está entre 70 e 100	-

Fonte: Elaborado pelo autor.

O aplicativo Protégé¹⁴ foi utilizado como aplicativo para modelar a ontologia ONTRISCAL conforme os conceitos e restrições listadas no Quadro 11 e Quadro 12. Além das especificações apresentadas nos quadros, abaixo segue a lista das restrições definidas por lógica de descrição da modelagem ontológica, a fim de formalizar o conhecimento do domínio.

1. Paciente \subseteq Pessoa
2. ProfissionalResponsavel \subseteq Pessoa
3. Pessoa \cap ProfissionalResponsavel = ϕ
4. \forall atendidoPor.ProfissionalResponsavel
5. \exists temDiagnostico.Doenca
6. Depressao \subseteq Doenca

¹⁴ <https://protege.stanford.edu/>

7. $\text{TranstornoDeAnsiedade} \subseteq \text{Doenca}$
8. $\forall \text{responde. Questao}$
9. $(\exists \text{itemDiagnostico. Comorbidade} \vee \exists \text{itemDiagnostico. Doenca}) \wedge \forall \text{apresenta. SinalESintoma}$
10. $\forall \text{recebe. Pontuacao}$
11. $\forall \text{itemAtribuicaoPontuacao. ProfissionalResponsavel}$
12. $\text{FaixaDeRiscoBaixo} \subseteq (\geq 1 \text{item. Valor} \wedge \leq 39 \text{item. Valor})$
13. $\text{FaixaDeRiscoMedio} \subseteq (\geq 40 \text{item. Valor} \wedge \leq 69 \text{item. Valor})$
14. $\text{FaixaDeRiscoAlto} \subseteq (\geq 70 \text{item. Valor} \wedge \leq 100 \text{item. Valor})$
15. $\forall \text{atribuidaSomaDaPontuacao. FaixaDeRiscoAlto} \vee \forall \text{atribuidaSomaDaPontuacao. FaixaDeRiscoMedio} \vee \forall \text{atribuidaSomaDaPontuacao. FaixaDeRiscoBaixo}$
16. $\forall \text{itemAtribuido. NivelDeRisco}$
17. $\forall \text{itemViDeRisco. NivelDeRiscoAlto} \vee \forall \text{itemViDeRisco. NivelDeRiscoMedio} \vee \forall \text{itemViDeRisco. NivelDeRiscoBaixo}$
18. $\text{Maria} \subseteq \text{Pessoa}$
19. $\text{Maria}: (\exists \text{Comorbidade} \subseteq (\exists \text{itemDiagnostico. Depressao} \wedge \exists \text{itemDiagnostico. TranstornoDeAnsiedade}))$

Conforme sugerido na seção 5.1 este estudo de caso utiliza a SABiO, que propõe um processo de documentação e fundamentação da ontologia sendo modelada (COSTA *et al.*, 2023; FALBO, 2014). Deste modo, o restante desta atividade exemplifica como a ONTRISCAL foi fundamentada na UFO, visando melhorar a análise de domínio e contribuindo para explicitar o conhecimento da primeira versão da ontologia. A fundamentação foi realizada utilizando o *plugin* da OntoUML por meio da ferramenta Visual Paradigm¹⁵.

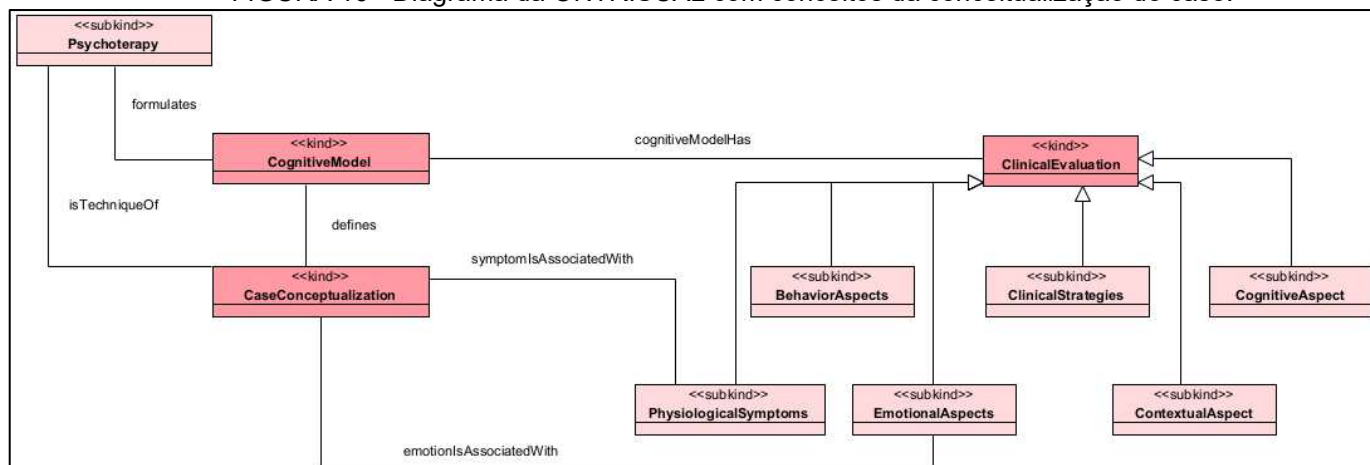
O modelo da ONTRISCAL padroniza conceitos e relações do domínio, resultando em um artefato que pode ser usado para comunicação e compreensão dos especialistas de domínio envolvidos na análise quali-quantitativa de tratamentos psicoterápicos, possibilitando a integração e a expansão de outras representações de conhecimento da área de saúde mental.

A Figura 16 apresenta o diagrama que exemplifica graficamente a ONTRISCAL com os principais conceitos fundamentados na UFO, e usados para definir a conceitualização do caso do paciente. A conceitualização define o modelo cognitivo, o qual formula a ação terapêutica. A avaliação clínica ocorre por aspectos emocionais e sintomas psicológicos

¹⁵ <https://www.visual-paradigm.com/>

associados à conceitualização do caso. As classes *ClinicalEvaluation*, *CognitiveModel*, *CaseConceptualization* estão fundamentadas como *kinds*, e fornecem princípios uniformes de individuação como tipos rígidos. As subclasses de *ClinicalEvaluation* são *subkinds* que representam especializações rígidas, do tipo *Rigid Sortal*.

FIGURA 16 - Diagrama da ONTRISCAL com conceitos da conceitualização do caso.

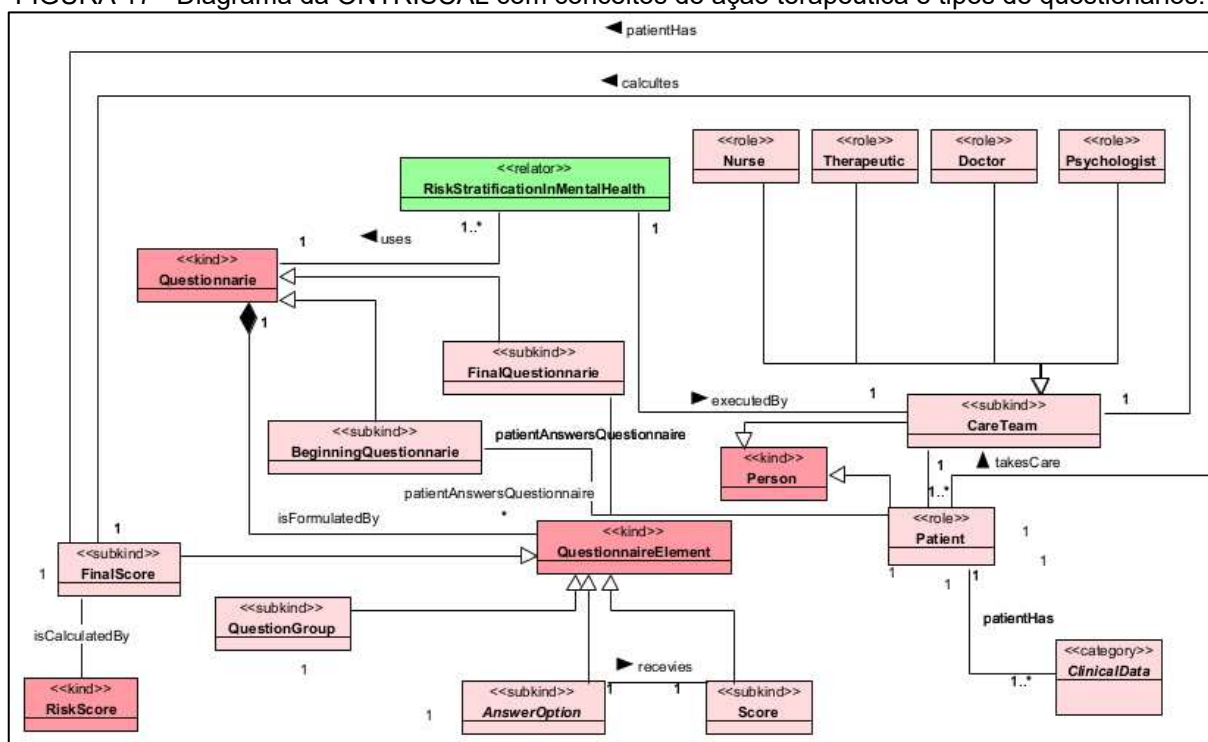


Fonte: Elaborado pelo autor.

A Figura 17 apresenta outros conceitos fundamentados na UFO. A classe *Risk Stratification In Mental Health* é um *relator*, pois é dependente das classes *Questionnaire* e *CareTeam*, tal como definido por Paula (2019), onde a estratificação de risco segue um processo que é iniciado no acolhimento do paciente pela atenção primária (equipe que realiza os cuidados), e que utiliza questionários para coletar os dados de risco.

Outras classes seguem fundamentadas na UFO (cf. Figura 17). As classes *BeginningQuestionnaire* e *FinalQuestionnaire* são *subkinds* de *Questionnaire*, sendo esta última fundamentada como *kind*. As classes *Nurse*, *Therapeutic*, *Doctor*, *Psychologist* estão fundamentadas como *Role*, e, portanto, são do tipo *AntiRigid Sortal*. Um *Role* é uma construção usada para representar especializações anti-rígidas e que proveem princípio de identidade (*i.e.*, *Kind*). A classe *ClinicalData* é um *Category*, que agrega as propriedades que constituem dados clínicos, caracterizando tipos de dados (indivíduos) que seguem diferentes princípios de identidade, tais como: propriedades de saúde do paciente; histórico de doenças pregressas, ou; procedimentos e tratamentos realizados (MILLER, REARDON E SAFI, 2001).

FIGURA 17 - Diagrama da ONTRISCAL com conceitos de ação terapêutica e tipos de questionários.



Fonte: Elaborado pelo autor.

Além das classes anteriormente discutidas, a UFO fundamenta as classes *SignAndSymptom* e *Comorbidity*, ambas do tipo *Mode* (cf. Apêndice A). Essas classes representam momentos intrínsecos podendo variar de formas independentes caracterizando o estado de saúde do paciente durante o tratamento. Os sinais e sintomas (classe *SignAndSymptom*) são identificados no início do tratamento. No entanto, diversos fatores podem elevar a gravidade do transtorno mental, e os sinais e sintomas podem variar fazendo com que o tratamento seja reconduzido com novas estratégias de intervenção. O mesmo pode ocorrer com a comorbidade, ou seja, o agravamento do transtorno de ansiedade pode aumentar, e outras comorbidades podem surgir durante o tratamento.

Atividade 2 – Codificação: Os conceitos da ontologia foram exportados da ferramenta Protégé para que pudessem ser importados para a ferramenta Taguette¹⁶ utilizada como CAQDAS para a codificação. Os especialistas de domínio analisaram e codificaram dados qualitativos (sintéticos) dos artigos intitulados: “Terapia cognitivo-comportamental no tratamento de um caso de transtorno de ansiedade generalizada: um relato de caso” (MELO; LOURENÇO, 2020); e “Tratamento psicológico do medo de viajar de avião, a partir do modelo cognitivo: caso clínico” (REMOR, 2000).

¹⁶ <https://app.taguette.org/login>

Atividade 3 – Exportação da Codificação: A codificação foi exportada pela ferramenta de Taguette em formato .CSV. Os dados foram selecionados e transpostos de linhas para colunas usando o Microsoft Excel. Ao transpor os dados, os códigos (*codes*) foram caracterizados por variáveis (colunas no *dataset*), preparando-os para execução da Atividade 4.

Atividade 4 – Identificação do objeto de estudo: Ao analisar os dados sintéticos dos artigos citados na Atividade 2, dados pessoais foram identificados. No entanto, apesar de serem dados sintéticos, um identificador único foi criado (atribuído aleatoriamente) para cada paciente do respectivo artigo, com o intuito de evitar a manipulação dos dados pessoais nas coleções de dados. Desta forma, os identificadores “Caso 4” e “Caso 6” foram atribuídos, respectivamente, para os pacientes identificados nos artigos Melo e Lourenço (2020) e Remor (2000). Os dados quantitativos foram identificados e selecionados a partir de questionários aplicados para avaliar e atribuir valores numéricos ao nível de risco dos pacientes. Estes últimos também foram registrados com identificadores únicos com o mesmo código do *dataset* qualitativo usado para identificar o paciente.

Atividade 5 – Anotação Semântica: A técnica SDD foi utilizada nesta atividade. Os templates de metadados foram preenchidos com o apoio dos especialistas de domínio. A modelagem da ontologia para apoiar a anotação semântica foi realizada a partir da aplicação da técnica SDD. As QCs serviram de base para o *design* da solução e execução da técnica SDD. Os dados sintéticos permitiram extrair dados quantitativos que possuem pontuações sobre os sinais e sintomas de pacientes com ansiedade. Os dados produzidos na Atividade 2 constituem os *datasets* qualitativos. Além da ONTRISCAL, classes e propriedades do projeto REFI-QDA e da ontologia QualiCO foram utilizadas para anotar semanticamente os dados quali-quantitativos.

Os especialistas de domínio apoiaram a anotação semântica de dois SDDs (SDD quali e SDD quanti). Como estratégia para enriquecer o conhecimento da análise quali-quantitativa, foram criados dois SDDs dos dados qualitativos, definidos abaixo:

- O “SDD Quali” possui anotações semânticas que utilizam classes do projeto REFI-QDA e QualiCO, e define o conhecimento sobre o esquema de codificação com base em classes específicas. As classes utilizadas para este SDD são as seguintes: *refiqda:NameOfDocument*, *refiqda:Textsource*, *refiqda:User*, *refiqda:NameOfUser*, *refiqda:IdProject*, *refiqda:NameOfProject*, *qualico:title*, *qualico:Keywords* e *qualico:researcharea*;
- O “SDD de Domínio” anota os dados com base nas classes da ONTRISCAL.

Os Quadros (Quadro 13, Quadro 14, Quadro 15 e Quadro 16) apresentam, respectivamente, exemplos de *templates* de metadados de acordo com a aplicação da técnica SDD. O objetivo desses exemplos é introduzir, de forma simplificada, o uso da técnica SDD com dados quantitativos extraídos da pesquisa realizada no HVVSM. No entanto, mais adiante nesta seção, o estudo de caso aprofunda na anotação de dados quali-quantitativos dos tratamentos psicoterápicos. Não foi necessário usar o *template TimeLine*, que serviria para mapear conceitos sobre os momentos (eventos) entre a internação e alta do paciente em algum plano de cuidado.

Quadro 13 - Preenchimento do Infosheet

Attribute	Value
<i>Title</i>	Análise quali-quantitativa por meio de grafos
<i>Comment</i>	HOMOGENiSE método para geração de grafos de conhecimento para análise quali-quantitativa
<i>Date Created</i>	03/05/2023
<i>Creators</i>	Evaldo de Oliveira da Silva
<i>Link</i>	https://github.com/evaldo/homogenise.git
<i>Identifier</i>	HOMOGENiSE
<i>Keywords</i>	HOMOGENiSE, Grafos de Conhecimento, Análise Quali-Quantitativa, Ontologia risco saúde mental
<i>File Format</i>	Csv
<i>Dictionary Mapping</i>	ExampleProject/input/DM/exampleDM.csv
<i>Codebook</i>	ExampleProject/input/CB/exampleCB.csv
<i>Code Mapping</i>	ExampleProject/config/code_mappings.csv
<i>Timeline</i>	ExampleProject/input/timeline.csv
<i>Imports</i>	owl:imports

Fonte: Elaborado pelo autor.

Quadro 14 - Especificação do DM para dados explícitos e implícitos

Column	Attribute	attributeOf	Unit
ID do Paciente	ontriscal:IdentificadorDeCadastro	??paciente	schema:Integer
Nome do Paciente	ontriscal:NomeCompleto	??paciente	
Grupo1_1_Sensacao de morte iminente, panico e/ou medo intenso	ontriscal:OpcaoDeResposta	??questao	
pontos_grupo1_1	ontriscal:Pontuacao	??questao	schema:Integer
Total	ontriscal:PontuacaoFinal	??elementoquestionario	schema:Integer
Risco	ontriscal:NivelDeRisco	??paciente	

Fonte: Elaborado pelo autor.

Quadro 15 - Especificação do DM para explicitação dos dados

Column	Entity	Relation	inRelationTo
??questao	ontriscal:Questao	ontriscal:questaoPossui	ontriscal:OpcaoDeResposta
??questao	ontriscal:Questao	ontriscal:fazParte	ontriscal:GrupoDeQuestao
??paciente	ontriscal:Paciente	ontriscal:responde	ontriscal:Questao
??paciente	ontriscal:Paciente	ontriscal:pacientePossui	ontriscal:PontuacaoFinal
??elementoquestionario	ontriscal:ElementoQuestionario	ontriscal:temElementoQuestionario	ontriscal:Paciente
??paciente	ontriscal:Paciente	ontriscal:pacientePossui	ontriscal:NivelDeRisco
??nivelderisco	ontriscal:NivelDeRisco	ontriscal:nivelDeRiscoOrienta	ontriscal:ProjetoTerapeutico
??projeto terapeutico	ontriscal:ProjetoTerapeutico	ontriscal:projetoTerapeuticoEFormadoPor	ontriscal:AcaoTerapeutica
??nivelderiscoalto	ontriscal:NivelDeRiscoAlto	ontriscal:nivelDeRiscoOrienta	ontriscal:AtendimentoMedico
??nivelderisco medio	ontriscal:NivelDeRiscoMedio	ontriscal:nivelDeRiscoOrienta	ontriscal:OficinaDeConvivencia
??nivelderiscobaixo	ontriscal:NivelDeRiscoBaixo	ontriscal:nivelDeRiscoOrienta	ontriscal:InstrumentoAssistencial

Fonte: Elaborado pelo autor.

Quadro 16 – Codebook

Column	Code	Label	Class
RISCO	1	Baixo	ontriscal:NivelDeRiscoBaixo
RISCO	2	Medio	ontriscal:NivelDeRiscoMedio
RISCO	3	Alto	ontriscal:NivelDeRiscoAlto

Fonte: Elaborado pelo autor.

Atividade 6 – Geração do Grafo de Conhecimento: O script para geração do GC utiliza como *inputs* os metadados (cf. Quadro 13, Quadro 14, Quadro 15 e Quadro 16) que anotam dados na primeira linha dos dados apresentados na Quadro 17. O Quadro 18 apresenta um fragmento como exemplo de grafo em RDF, gerado em sintaxe turtle pelo script *sdd2rdf*. Os *datasets* quantitativos, qualitativos, os SDDs e demais artefatos gerados como resultados das atividades realizadas, podem ser consultados no repositório de versionamento desta tese¹⁷.

Quadro 17 - Dataset com dados de exemplo da estratificação de risco.

ID do Paciente	Nome do Paciente	Grupo1_1_Sensacao de morte iminente, panico e/ou medo intenso	pontos_grupo1_1	Total	Risco
4	A_25	Não	15	30	BAIXO
6	A_161	Sim	58	85	ALTO
12	A_130	Não	32	63	MEDIO

Fonte: Elaborado pelo autor.

¹⁷ <https://github.com/evaldo/homoginise.git>

Quadro 18 - Grafo RDF no formato turtle

```

@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix prov: <http://www.w3.org/ns/prov#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix uo: <http://purl.obolibrary.org/obo/UO> .
@prefix sio: <http://semanticscience.org/resource/> .
@prefix ontriscal: <http://www.estrdefineatificacaoderisco/onto/ontriscal#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix dc: <http://purl.org/dc/terms/> .
ontriscal:Grupo1_1_Sensacao_de_morte_iminente__panico_e_
ou_medo_intenso a owl:Class ;
    rdfs:subClassOf [ a owl:Restriction ;
        owl:allValuesFrom ontriscal:questao ;
        owl:onProperty sio:isAttributeOf ],
    ontriscal:Questao .
ontriscal:ID_do_Paciente a ontriscal:ID_do_Paciente,
    ontriscal:IdentificadorDeCadastro ;
    sio:hasUnit sio:Interger ;
    sio:hasValue 4 ;
    sio:isAttributeOf ontriscal:paciente .
ontriscal:Nome_do_Paciente a ontriscal:NomeCompleto,
    ontriscal:Nome_do_Paciente ;
    sio:hasValue "A_25"^^xsd:string ;
    sio:isAttributeOf ontriscal:paciente .
ontriscal:RISCO a ontriscal:NivelDeRisco,
    ontriscal:RISCO ;
    sio:isAttributeOf ontriscal:paciente .
ontriscal:TOTAL a ontriscal:PontuacaoFinal,
    ontriscal:TOTAL ;
    sio:hasUnit sio:Interger ;
    sio:hasValue "BAIXO"^^xsd:string ;
    sio:isAttributeOf ontriscal:elementoquestionario .
ontriscal:nivelderisco a owl:Class ;
    rdfs:label "nivelderisco"^^xsd:string ;
    ontriscal:nivelDeRiscoOrienta
ontriscal:ProjetoTerapeuticoSingular ;
rdfs:subClassOf ontriscal:NivelDeRisco .
    ontriscal:nivelderiscoalto a owl:Class ;
        rdfs:label "nivelderiscoalto"^^xsd:string ;
ontriscal:nivelDeRiscoOrienta ontriscal:AtendimentoMedico ;
    rdfs:subClassOf ontriscal:NivelDeRiscoAlto .
ontriscal:nivelderiscomedio a owl:Class ;
    rdfs:label "nivelderiscomedio"^^xsd:string ;
    ontriscal:nivelDeRiscoOrienta
ontriscal:AtendimentoMedico ;
    rdfs:subClassOf ontriscal:NivelDeRiscoMedio .
ontriscal:paciente a ontriscal:Paciente,
    ontriscal:paciente ;
    rdfs:label "paciente"^^xsd:string ;
        ontriscal:responde ontriscal:Questao .
        ontriscal:fazParte
        ontriscal:GrupoDeQuestao .
        ontriscal:questao a ontriscal:Questao,
            ontriscal:questao ;
            rdfs:label "questao"^^xsd:string ;
            ontriscal:questaoPossui
            ontriscal:OpcaoDeResposta .
            ontriscal:paciente a ontriscal:Paciente,
                ontriscal:paciente ;
                rdfs:label "paciente"^^xsd:string ;
                ontriscal:pacientePossui
                ontriscal:PontuacaoFinal .
                ontriscal:pontos_grupo1_1 a ontriscal:Pontuacao,
                    ontriscal:pontos_grupo1_1 ;
                    sio:hasUnit sio:Interger ;
                    sio:hasValue 30 ;
                    sio:isAttributeOf ontriscal:questao .
                ontriscal:projetoTerapeuticoSingular a owl:Class ;
                    rdfs:label "projetoTerapeuticoSingular"^^xsd:string ;
                    ontriscal:projetoTerapeuticoSingularEFormadoPor
                    ontriscal:AcaoTerapeutica ;
                    rdfs:subClassOf ontriscal:ProjetoTerapeuticoSingular .
                ontriscal:questao a ontriscal:Questao,
                    ontriscal:questao ;
                    rdfs:label "questao"^^xsd:string ;
                    ontriscal:questao a ontriscal:Questao,
                        ontriscal:questao ;
                        rdfs:label "questao"^^xsd:string ;
                        ontriscal:questaoPossui ontriscal:Pontuacao .
                    ontriscal:elementoquestionario a owl:Class ;
                        rdfs:label "elementoquestionario"^^xsd:string ;
                        ontriscal:temElementoQuestionario ontriscal:Paciente ;
                        rdfs:subClassOf ontriscal:ElementoQuestionario .
                    ontriscal:paciente a ontriscal:Paciente,
                        ontriscal:paciente ;
                        rdfs:label "paciente"^^xsd:string ;
                        ontriscal:pacientePossui
                        ontriscal:NivelDeRisco .
                    ontriscal:questao a owl:Class ;
                        rdfs:label "questao"^^xsd:string ;
                        ontriscal:fazParte ontriscal:GrupoDeQuestao ;
                        ontriscal:questaoPossui ontriscal:OpcaoDeResposta,
                            ontriscal:Pontuacao ;
                            rdfs:subClassOf ontriscal:Questao .
                    ontriscal:paciente a owl:Class ;
                        rdfs:label "paciente"^^xsd:string ;
                        ontriscal:pacientePossui
                        ontriscal:NivelDeRisco,
                            ontriscal:PontuacaoFinal ;
                            ontriscal:responde ontriscal:Questao ;
                            rdfs:subClassOf ontriscal:Paciente .

```

Fonte: Elaborado pelo autor.

Para demonstrar os conceitos e como podem ser correlacionados no GC, foram anotados os SDDs quali e quanti. Independentemente do domínio de aplicação, para tornar o GC acessível aos conceitos de codificação da análise qualitativa, foi criado o SDD quali por onde é possível navegar pelo grafo usando os conceitos da REFI-QDA e QualiCO. Este SDD

anota as colunas do *dataset* contendo conceitos da tarefa de codificação. O SDD quali pode ser visto no Quadro 19 e Quadro 20, onde esta última explicita os dados qualitativos sob a luz dos conceitos existentes na REFI-QDA e QualiCO. As classes de conceitos e propriedades da REFIQDA foram modeladas em formato OWL¹⁸. O GC¹⁹ foi gerado a partir das anotações apresentadas a seguir.

Quadro 19 - Anotação semântica para explicitar dados qualitativos do SDD quali.

<i>Column</i>	<i>Attribute</i>	<i>attributeOf</i>	<i>Unit</i>
id_case	hasco:originalID	??case	schema:Integer
document	refiqda:NameOfDocument	??project	schema:String
Source	refiqda:Textsource	??project	schema:String
id_user	refiqda:User	??project	schema:Integer
User	refiqda:NameOfUser	??project	schema:String
id_project	refiqda:IdProject	??project	schema:String
Project	refiqda:NameOfProject	??project	schema:String
TitleOfStudy	qualico:title	??project	schema:String
Keywords	qualico:keywords	??project	schema:String
ResearchArea	qualico: researcharea	??project	schema:String
Fear	refiqda:EmotionalAspect	??note	schema:String
FearOfLosingControl	refiqda:EmotionalAspect	??note	schema:String
FearOfNegativeEvaluationByOthers	refiqda:CognitiveAspect	??note	schema:String
Frustration	refiqda:EmotionalAspect	??note	schema:String
Generalization	refiqda:CognitiveAspect	??note	schema:String
MaximizationOfTheNegative	refiqda:Thinking	??note	schema:String
MinimizationOfThePositive	refiqda:Thinking	??note	schema:String
ObjectPhobic	refiqda:CognitiveAspect	??note	schema:String
Perfectionism	refiqda:Thinking	??note	schema:String
PredictionOfTheFuture	refiqda:CognitiveAspect	??note	schema:String
RuminantThoughts	refiqda:CognitiveAspect	??note	schema:String
SelfAwareness	refiqda:CognitiveAspect	??note	schema:String
NumberOfSessions	refiqda:StructureOfTreatmentProcess	??note	schema:String
PatientHistory	refiqda:StructureOfTreatmentProcess	??note	schema:String
ProcessDuration	refiqda:StructureOfTreatmentProcess	??note	schema:String
SessionDuration	refiqda:StructureOfTreatmentProcess	??note	schema:String
TherapeuticGoals	refiqda:StructureOfTreatmentProcess	??note	schema:String
ClinicalStrategies	refiqda:Treatment	??note	schema:String
CopingStrategies	refiqda:Treatment	??note	schema:String
CopingWithStress	refiqda:Treatment	??note	schema:String
Psychopathology	refiqda:ClinicalHypothesis	??note	schema:String
Suppositions	refiqda:CognitiveAspect	??note	schema:String

Fonte: Elaborado pelo autor.

Quadro 20 – Explicitação dos dados qualitativos do SDD quali.

<i>Column</i>	<i>Entity</i>	<i>Relation</i>	<i>inRelationTo</i>
??case	sio:Human	refiqda:appliesTo	??project
??project	refiqda:Project	refiqda:contains	??case
??note	refiqda>Note		
??project	refiqda:Project	refiqda:contains	??note

Fonte: Elaborado pelo autor.

¹⁸ <https://github.com/evaldo/homoginise/tree/master/Ontologia>

¹⁹ <https://github.com/evaldo/homoginise/tree/master/Dados Sintéticos em TTL>

As especificações de SDDs apresentadas a seguir anotam dados do conjunto de dados sintéticos. Os SDDs quali anotam o mesmo *dataset* qualitativo podendo ampliar a possibilidade de acessar os dados e inferir novos conhecimentos. Em cada SDD utiliza-se a mesma coluna “id_case” com as mesmas anotações semânticas (em cada SDD), a fim permitir ao especialista de domínio navegar pelos diferentes conceitos de um mesmo objeto de estudo. O SDD quali de domínio pode ser visto nos Quadro 21 e Quadro 22. O Quadro 22 explicita os dados qualitativos de domínio tendo como base as classes de conceitos da ontologia ONTRISCAL.

QUADRO 21 - Anotação semântica para explicitar dados qualitativos do SDD quali de domínio.

<i>Column</i>	<i>Attribute</i>	<i>attributeOf</i>	<i>Unit</i>
id_case	hasco:originalID	??case	schema:Interger
document	ontriscal:TherapeuticCase	??psychotherapy	schema:String
Source	ontriscal:MedicalRecord	??psychotherapy	schema:String
id_user	ontriscal:HealthProfessional	??psychotherapy	schema:Interger
User	ontriscal:FullName	??psychotherapy	schema:String
id_project	ontriscal:Treatment	??psychotherapy	schema:String
Project	ontriscal:FullName	??psychotherapy	schema:String
Fear	ontriscal:EmotionalAspect	??caseconceptualization	schema:String
FearOfLosingControl	ontriscal:EmotionalAspect	??caseconceptualization	schema:String
FearOfNegativeEvaluationByOthers	ontriscal:CognitiveAspect	??caseconceptualization	schema:String
Frustration	ontriscal:EmotionalAspect	??caseconceptualization	schema:String
Generalization	ontriscal:CognitiveAspect	??caseconceptualization	schema:String
MaximizationOfTheNegative	ontriscal:Thinking	??caseconceptualization	schema:String
MinimizationOfThePositive	ontriscal:Thinking	??caseconceptualization	schema:String
ObjectPhobic	ontriscal:CognitiveAspect	??caseconceptualization	schema:String
Perfectionism	ontriscal:Thinking	??caseconceptualization	schema:String
PredictionOfTheFuture	ontriscal:CognitiveAspect	??caseconceptualization	schema:String
RuminantThoughts	ontriscal:CognitiveAspect	??caseconceptualization	schema:String
SelfAwareness	ontriscal:CognitiveAspect	??caseconceptualization	schema:String
NumberOfSessions	ontriscal:StructureOfTreatmentProcess	??caseconceptualization	schema:String
PatientHistory	ontriscal:StructureOfTreatmentProcess	??caseconceptualization	schema:String
ProcessDuration	ontriscal:StructureOfTreatmentProcess	??caseconceptualization	schema:String
SessionDuration	ontriscal:StructureOfTreatmentProcess	??caseconceptualization	schema:String
TherapeuticGoals	ontriscal:StructureOfTreatmentProcess	??caseconceptualization	schema:String
ClinicalStrategies	ontriscal:Treatment	??caseconceptualization	schema:String
CopingStrategies	ontriscal:Treatment	??caseconceptualization	schema:String
CopingWithStress	ontriscal:Treatment	??caseconceptualization	schema:String
Psychopathology	ontriscal:ClinicalHypothesis	??caseconceptualization	schema:String
Suppositions	ontriscal:CognitiveAspect	??caseconceptualization	schema:String

Fonte: Elaborado pelo autor.

Quadro 22 – Explicitação dos dados qualitativos do SDD quali do domínio.

<i>Column</i>	<i>Entity</i>	<i>Relation</i>	<i>inRelationTo</i>
??case	sio:Human	ontriscal:isComprehendedBy	??caseconceptualization
??caseconceptualization	ontriscal:CaseConceptualization	ontriscal:definesCognitiveModelOf	??case
??symptom anxiety	ontriscal:PhysiologicalSymptoms	ontriscal:symptomIsAssociatedWith	??caseconceptualization
??emotion	ontriscal:EmotionAspects	ontriscal:emotionIsAssociatedWith	??caseconceptualization
??psychotherapy	ontriscal:Psychotherapy	ontriscal:treats	??case
??caseconceptualization	ontriscal:CaseConceptualization	ontriscal:isTechniqueOf	??psychotherapy

Fonte: Elaborado pelo autor.

Os conceitos dos dados quantitativos são apresentados a partir de dois SDDs (especificações de *Dictionary Mapping*). O primeiro para dados de um questionário aplicado

no início do tratamento psicoterápico, e o segundo aplicado ao final do tratamento. Em cada SDD quanti preserva-se a utilização da coluna “id_case” com as mesmas anotações semânticas utilizadas nos SDDs quali. Os SDDs quanti (questionários inicial e final) podem ser vistos nas próximos quadros (Quadro 23, Quadro 24, Quadro 25 e Quadro 26).

Quadro 23 - Anotação semântica dos dados quantitativos do questionário inicial.

<i>Column</i>	<i>Attribute</i>	<i>attributeOf</i>	<i>Unit</i>
id_case	hasco:originalID	??case	schema:Interger
initial_questionnaire	ontriscal:FullName	??beginningquestionnaire	
warm_sensation	ontriscal:WarmSensation	??beginningquestionnaire	
leg_shaking	ontriscal:LegShaking	??beginningquestionnaire	
inability_to_relax	ontriscal:InabilityToRelax	??beginningquestionnaire	
fear_of_the_worst_happening	ontriscal:FearOfTheWorstHappening	??beginningquestionnaire	
dizziness	ontriscal:Dizziness	??beginningquestionnaire	
accelerated_heart	ontriscal:AcceleratedHeart	??beginningquestionnaire	
balance_disorder	ontriscal:BalanceDisorder	??beginningquestionnaire	
tremor	ontriscal:Tremor	??beginningquestionnaire	
nervous	ontriscal:Nervous	??beginningquestionnaire	
feeling_of_suffocation	ontriscal:FeelingOfSuffocation	??beginningquestionnaire	
hand_tremor	ontriscal:HandTremor	??beginningquestionnaire	
difficulty_breathing	ontriscal:DifficultyBreathing	??beginningquestionnaire	
fear_to_die	ontriscal:FearToDie	??beginningquestionnaire	
get_scared	ontriscal:GetScared	??beginningquestionnaire	
feeling_faint	ontriscal:FeelingFaint	??beginningquestionnaire	
flushed_face	ontriscal:FlushedFace	??beginningquestionnaire	
sweat	ontriscal:Sweat	??beginningquestionnaire	
initial_score_bai_scale	ontriscal:Score	??beginningquestionnaire	
initial_risk_level	ontriscal:RiskLevel	??beginningquestionnaire	

Fonte: Elaborado pelo autor.

Quadro 24 – Explicação dos dados quantitativos do questionário inicial.

<i>Column</i>	<i>Entity</i>	<i>Relation</i>	<i>inRelationTo</i>
??case	sio:Human	ontriscal:patientAnswersQuestionnaire	??beginningquestionnaire
??beginningquestionnaire	ontriscal:Questionnaire	ontriscal:questionnaireAnsweredBy	??case

Fonte: Elaborado pelo autor.

Quadro 25 - Anotação semântica dos dados quantitativos do questionário final.

<i>Column</i>	<i>Attribute</i>	<i>attributeOf</i>	<i>Unit</i>
id_case	hasco:originalID	??case	schema:Interger
final_questionnaire	ontriscal:FullName	??finalquestionnaire	
warm_sensation	ontriscal:WarmSensation	??finalquestionnaire	
leg_shaking	ontriscal:LegShaking	??finalquestionnaire	
inability_to_relax	ontriscal:InabilityToRelax	??finalquestionnaire	
fear_of_the_worst_happening	ontriscal:FearOfTheWorstHappening	??finalquestionnaire	
dizziness	ontriscal:Dizziness	??finalquestionnaire	
accelerated_heart	ontriscal:AcceleratedHeart	??finalquestionnaire	
balance_disorder	ontriscal:BalanceDisorder	??finalquestionnaire	
Tremor	ontriscal:Tremor	??finalquestionnaire	
Nervous	ontriscal:Nervous	??finalquestionnaire	
feeling_of_suffocation	ontriscal:FeelingOfSuffocation	??finalquestionnaire	
hand_tremor	ontriscal:HandTremor	??finalquestionnaire	
difficulty_breathing	ontriscal:DifficultyBreathing	??finalquestionnaire	
fear_to_die	ontriscal:FearToDie	??finalquestionnaire	

<i>Column</i>	<i>Attribute</i>	<i>attributeOf</i>	<i>Unit</i>
get_scared	ontriscal:GetScared	??finalquestionnaire	
feeling_faint	ontriscal:FeelingFaint	??finalquestionnaire	
flushed_face	ontriscal:FlushedFace	??finalquestionnaire	
Sweat	ontriscal:Sweat	??finalquestionnaire	
final_score_bai_scale	ontriscal:Score	??finalquestionnaire	
final_risk_level	ontriscal:RiskLevel	??finalquestionnaire	

Fonte: Elaborado pelo autor.

Quadro 26 - Explicitação dos dados quantitativos do questionário final.

<i>Column</i>	<i>Entity</i>	<i>Relation</i>	<i>inRelationTo</i>
??case	sio:Human	ontriscal:patientAnswersQuestionnaire	??finalquestionnaire
??finalquestionnaire	ontriscal:Questionnaire	ontriscal:questionnaireAnsweredBy	??case

Fonte: Elaborado pelo autor.

A Lista de Pendências criada na etapa anterior, foi revisada e todas foram sanadas pelos especialistas de domínio para que esta etapa pudesse ser concluída. Nesta etapa, nenhuma pendência foi registrada.

A Matriz de Rastreabilidade foi revisada e o impacto da mudança foi em 7 artefatos produzidos nesta etapa. São eles: *Infosheet*; *Dictionary Mapping*; *CodeBook*; ONTRISCAL; coleções de dados; *script* de software; e grafo de conhecimento. O script *sdd2rdf* teve que ser alterado para permitir a geração do grafo para mais de um SDD quali com casos distintos de tratamentos. O versionamento ocorreu no repositório do projeto (<https://github.com/evaldo/homoginise>).

5.4 Demonstração

A etapa anterior implementou o processo que organizou um conjunto de atividades para aprimorar a análise quali-quantitativa. Dados sintéticos do domínio de tratamento psicoterápicos foram compreendidos e preparados. Os SDDs anotaram semanticamente dados para geração de GCs que popularam dados representados por conceitos e as relações entre eles.

Os elementos terminais dos grafos são valores que instanciam propriedades de objetos que, por sua vez, populam o tratamento psicoterápico e são usados para responder às QCs. Os arquivos de dados organizados nos grafos contêm os valores em células tabulares e os templates de metadados são compostos de anotações formais de propriedades de objetos. A ONTRISCAL formaliza a semântica exigindo que as definições sejam declaradas usando termos consensuados no domínio. Considerando os SDDs quali e quanti é possível demonstrar a navegação pelo GC.

Para exemplificar a navegação pelo GC a seguinte QC é utilizada: “Quais foram as estratégias de intervenção de maior impacto nos pensamentos, emoções e comportamentos

da(o) paciente no desenvolvimento do processo? Quais procedimentos se destacaram?”. A questão pode ser respondida a partir do objeto do tipo ??case, uma vez que este objeto é o mesmo para cada SDD (quali e quanti) preenchido. O objeto ??case é um tipo de caso (ou caso clínico) utilizado em um projeto de pesquisa definido pelo objeto ??project anotado pelo SDD quali. O objeto ??project possui atributos definidos na anotação semântica pela propriedade sio:isAttributeOf. O trecho da instrução SPARQL descrito abaixo, demonstra a navegação do objeto ??case para ter acesso ao nome do projeto em que está inserido.

```
?objectcase rdf:type ?case .
?id_case sio:isAttributeOf ?objectcase .
?id_case sio:hasValue ?valuecase .

?objectcase rdf:type ?caseproject .
?caseproject refiqda:appliesTo ?project .

?objectproject rdf:type ?project .
?nameproject sio:isAttributeOf ?objectproject .
?nameproject rdf:type refiqda:NameOfProject .
?nameproject sio:hasValue ?valuenameproject .
```

Ao demonstrar como navegar pelos conceitos do GC é necessário correlacionar o objeto ??case aos dados qualitativos do domínio. Neste caso, utiliza-se os conceitos explicitados no Quadro 22, por onde o objeto ??case está relacionado ao objeto ??caseconceptualization, e à propriedade ontriscal:isComprehendedBy, formando a seguinte tripla: ?case ontriscal:isComprehendedBy ?caseconceptualization. Para navegar pelo GC e recuperar as estratégias de intervenção, utiliza-se o objeto ??caseconceptualization para relacionar aos tratamentos. A instrução SPARQL descrita abaixo exemplifica a navegação entre os objetos ??case, ??caseconceptualization e ??treatment:

```
?case ontriscal:isComprehendedBy ?caseconceptualization .
?treatment sio:hasValue ?valuetreatment .
?treatment sio:isAttributeOf ?caseconceptualization .
?treatment rdf:type ontriscal:Treatment .
```

A QC utilizada para demonstrar a navegação pelo GC tem como interesse conhecer as estratégias de intervenção de maior impacto nos pensamentos, emoções e comportamentos da(o) paciente ao longo do processo. Neste caso, é necessário navegar pelos dados quantitativos a fim de apresentar os níveis de risco atribuídos por meio dos questionários. O objeto ??case pode ser correlacionado aos conceitos que anotam os dados quantitativos por meio do SDD quanti, tanto para o questionário inicial, quanto final. O objeto

??case é explicitado no DM para responder o questionário, onde a navegação no GC pode ser determinada pela seguinte instrução SPARQL:

```
?case ?patientanswersquestionnaire ?beginningquestionnaire .
?beginningquestionnaire rdfs:label "beginningquestionnaire" .

?risklevel sio:hasValue ?valuerisklevel .
?risklevel sio:isAttributeOf ?beginningquestionnaire .
?risklevel rdf:type ontriscal:RiskLevel .

?case ?patientanswersquestionnaire ?finalquestionnaire .
?finalquestionnaire rdfs:label "finalquestionnaire" .

?risklevel sio:hasValue ?valuerisklevel .
?risklevel sio:isAttributeOf ?finalquestionnaire .
?risklevel rdf:type ontriscal:RiskLevel .
```

O GC em formato *turtle* foi armazenado na ferramenta *AllegroGraph*²⁰. É uma ferramenta gratuita e de acesso facilitado, e se enquadra nesta etapa de execução do HOMOGENiSE. Permite a ingestão dos grafos gerados de forma manual. A *AllegroGraph* fornece uma interface chamada “*Gruff*” que possibilita acessar as triplas dos grafos por meio de consultas SPARQL.

Todas as consultas SPARQL para responder às QCs foram construídas e executadas por meio da ferramenta *AllegroGraph*. A partir do repositório de dados semânticos criado no *AllegroGraph*, os GCs podem ser extraídos em pequenas e grandes quantidades de dados. A partir da Seção 5.6 apresentam-se os GCs, resultados das consultas SPARQL e que respondem às QCs elaboradas na Seção 5.1. O Apêndice B apresenta todas as consultas SPARQL executadas para cada QC.

Não houve lista de pendências na etapa anterior. Nesta etapa, uma pendência foi registrada:

- Avaliar a performance do *AllegroGraph* para recuperar as triplas da demonstração realizada nesta seção. O tempo de recuperação satisfatório é de 5 segundos. Caso contrário, testar a performance em outro computador disponível para o projeto.

A Matriz de Rastreabilidade foi revisada e o impacto da mudança foi em 2 artefatos produzidos nesta etapa. São eles: grafo de conhecimento e consulta SPARQL utilizada na demonstração. O versionamento ocorreu no repositório do projeto (<https://github.com/evaldo/homoginise>).

²⁰ <https://allegrograph.com/>

5.4.1. Novas categorias de dados qualitativos baseadas em GCs

Esta seção retoma a discussão que abrange a TCC e a *GT*. Resumidamente, como já mencionado no referencial teórico, a proposta da *GT* baseia-se na codificação e nas conexões entre códigos de forma iterativa, a fim de gerar uma nova teoria (DELVE, 2022; GIBBS, 2010; WALKER; MYRICK, 2007). A *GT* pode conduzir a elucidação de uma categoria de dado inexistente no conjunto de dados qualitativos, podendo sustentar a geração de uma nova teoria. Grafos podem servir de meios de representação a fim de facilitar a análise por meio da *GT*. A importância dos meios de representação na análise quali-quantitativa é discutida por Mayring (2007).

A codificação apresentada nesta tese utilizou o desenho qualitativo originado da estrutura da TCC para criar os códigos. O desenho qualitativo representa a estrutura que orienta os tratamentos psicoterápicos, guiando os especialistas de domínio na elaboração da conceitualização cognitiva do paciente (MELO; LOURENÇO, 2020; REYES; FERMANN, 2017). Trata-se de um modelo no qual o especialista organiza os dados do paciente a fim de entender completamente o caso, direcionando as possíveis perspectivas e ações para o tratamento. A conceitualização cognitiva, sob a luz da TCC, associada à aplicação da *GT*, pode contribuir na organização (agrupamento) dos dados qualitativos coletados a partir das sessões de tratamentos, e, conseqüentemente, direcionar intervenções para melhorar a saúde do paciente.

No Caso 4, inicialmente, o paciente é atendido sem um diagnóstico prévio e que pode implicar na ausência de direcionamento no tratamento. No decorrer do tratamento o especialista coletou e estruturou os dados na tentativa de formar a conceitualização cognitiva do paciente. Esses dados coletados foram usados para a codificação, gerando um *dataset* qualitativo. A tarefa de codificação foi realizada de forma iterativa sendo que os diferentes códigos categorizaram os dados. Os GCs gerados foram populados com dados qualitativos resultantes da tarefa de codificação.

Diante da descrição acima, GCs podem se tornar artefatos importantes para suportar a *GT*. Os GCs possuem os códigos que agrupam dados qualitativos dos tratamentos psicoterápicos. Assim, pode surgir a necessidade do especialista de domínio em determinar novas categorias de dados inexistentes, e que podem ser elicitadas a partir das conexões dos códigos presentes no GCs. Cabe ressaltar que a possibilidade de uso de GCs para *GT*, neste trabalho, não tem o objetivo de propor uma “teoria” para evoluir a TCC como técnica já consolidada na área da psicologia. Mas sim, tornar possível a análise de dados para obtenção de novas categorias que apoiem a conceitualização cognitiva do paciente, propondo “um conjunto de ideias” que melhorem o tratamento do paciente usando a TCC como técnica

norteadora.

Um exemplo do uso de GC com GT seria o agrupamento de códigos relacionados a pensamentos aos códigos que agrupam os sintomas de ansiedade e aspectos comportamentais, onde:

- Códigos que agrupam a categoria de “pensamentos” são: *Catastrophizing*, *Generalization*, *PredictionOfTheFuture*, *FearOfLosingControl*, *MaximizationOfTheNegative* e *MinimizationOfThePositive*;
- Códigos que agrupam a categoria de “sintomas de ansiedade” são: *MuscularTension*, *Nervous*, *ShortnessOfBreath* e *Sweating*;
- Códigos que agrupam a categoria de “aspectos comportamentais” são: *AvoidanceOfThreateningSignsOrSituations*, *BehaviorAspects* e *BehaviorSymptoms*.

Nesse sentido, a conexão das novas categorias (pensamentos e sintomas de ansiedade) conduz ao entendimento para definição de uma hipótese diagnóstica (em consonância com a TCC), que pode ser confirmada (ou não) ao longo do tratamento. Caso não seja confirmada, novas análises de GT deveriam ser formuladas com o objetivo de encontrar novas categorias de dados qualitativos.

Entretanto, uma vez que haja a hipótese diagnóstica, servirá para nortear o especialista no tratamento específico dos sintomas, seguindo protocolos e manuais adequados e existentes na Psicologia. Deste modo, por meio da análise dos códigos e de seus respectivos significados, é possível que seja estabelecida uma hipótese diagnóstica de Transtorno de Ansiedade Generalizada (TAG).

Para implementar a GT juntamente com GCs, seria necessário modelar os novos conceitos das categorias de dados na ontologia ONTRICAL, e o SDD quali domínio deveria ser revisado para anotar as categorias. Assim, a cada ciclo de elaboração da GT, a ontologia e o SDD deveriam ser revisitados.

5.5 Avaliação

As etapas anteriores visaram compreender e preparar dados para análise quali-quantitativa. As etapas iniciais explicitaram o conhecimento do domínio, preparando os dados para análise quali-quantitativa. Com base na execução do HOMOGENiSE, até o momento, observa-se a possibilidade de prosseguir com este estudo de caso a fim de responder à questão de pesquisa e atingir os objetivos definidos. As hipóteses descritas (*cf.* Capítulo 1) são proposições que podem ser confirmadas (ou não), e que estabelecem princípios usados para avaliar o método aplicado até o momento. Desta forma, esta etapa recupera as hipóteses que

objetivam avaliar o método executado até o momento, e os resultados possam ser divulgados, sem a necessidade de reavaliar as etapas anteriores. Abaixo seguem as discussões para cada hipótese:

1. Anotar dados semanticamente apresenta-se como abordagem que explicita o conhecimento para análise quali-quantitativa e permite a integração dos dados dos objetos de estudo, auxiliando na compreensão e preparação dos dados. Discussão: Nesta pesquisa a explicitação do conhecimento inicia-se com a descrição do mundo e as QCs que nortearam a modelagem ontológica da ONTRISCAL. A compreensão se constitui como conhecimento do domínio e necessário para anotar os dados. O estudo de caso apresentou a anotação semântica por SDDs quali e quanti, por onde o conhecimento pode ser explicitado tendo como base as variáveis (colunas de dados) presentes nos datasets. No entanto, a técnica SDD exige um esforço intelectual além, pois novas relações entre as variáveis do estudo podem surgir com o objetivo de responder às QCs, havendo a necessidade de retornar para modelagem ontológica e acrescentar novos conceitos e propriedades elicitados. O estudo de caso desenvolvido nesta pesquisa demonstrou este esforço, e como o conhecimento é explicitado para análise quali-quantitativa, auxiliando na compreensão e preparação dos dados.

2. Dados quali-quantitativos anotados semanticamente por meio de diferentes ontologias e integrados por grafos de conhecimento, aprimoram a análise quali-quantitativa a fim de responder questões relevantes sobre o domínio pesquisado. Discussão: A estratégia de anotação semântica dos SDDs quali e quanti mostrou-se apropriada para permitir a geração de grafos para aprimorar análises quali-quantitativas. Na etapa anterior, foi demonstrada a navegação pelos dados qualitativos e quantitativos identificados univocamente pelo objeto de estudo (objeto ?case). Fragmentos do GC puderam ser recuperados por meio de instruções SPARQL, utilizando como “padrão” uma mesma definição para o objeto ?case. Avalia-se se a estratégia aqui mencionada é importante para recuperação dos GCs.

A etapa seguinte explora o resultado das etapas anteriores com o uso de instruções SPARQL que geram os GCs populados com dados quali-quantitativos, confirmando o avanço da pesquisa a partir da execução do método HOMOGENiSE.

A Matriz de Rastreabilidade foi revisada e esta etapa não produziu modificação em artefatos que justificasse o versionamento.

5.6 Divulgação

A divulgação permite que fragmentos do conhecimento sejam analisados a fim de

orientar os especialistas de domínio na inferência de novos conhecimentos. É a etapa final de uma iteração para apresentar os resultados das análises dos dados, sendo importante que as descobertas sejam compartilhadas. Utiliza-se aqui a ferramenta “*Gruff*” para executar as consultas SPARQL e apresentar graficamente os GCs que permitirão as análises por especialistas de domínio. Esta etapa preconiza o uso de soluções de software que exploram os dados por meio da navegação no grafo, combinando classes por meio do conhecimento representado na ontologia. Neste caso, não há necessidade de implementar a estrutura do dado tal como é feito no modelo relacional. Além disso, no modelo relacional, ao agregar mais dados é necessário alterar a sintaxe e o esquema de dados.

A proposta com o uso dos grafos torna a recuperação de informação independente dos modelos relacionais e permite integrar mais dados na ontologia oriundos de diferentes modelos relacionais, ou arquivos de dados. Com isso, torna-se possível construir futuros repositórios com a flexibilidade adequada para o especialista gerar o conhecimento que quiser para diferentes necessidades do especialista, podendo integrar dados das evoluções dos pacientes.

Nesta etapa, análises são relatadas para apresentar os resultados dos dados anotados semanticamente. As análises fundamentaram-se nos códigos que populam os GCs com dados qualitativos, criados pela tarefa de codificação, e que permite reunir os dados relevantes para análise quali-quantitativa. Para divulgar os resultados, realizou-se análises para responder as QCs listadas na Seção 5.1.

Conforme proposto por Creswell e Creswell (2018) em sua lista de verificação é necessário identificar o projeto de pesquisa para análise quali-quantitativa e os procedimentos para análise dos dados. Deste modo, ao executar as atividades (1 a 5) da etapa de Concepção e Desenvolvimento, este estudo de caso utilizou o projeto de pesquisa convergente para coleta de dados (CRESWELL; PLANO CLARK, 2017; DAWADI; SHRESTHA; GIRI, 2021). A seleção e organização dos dados (Atividades 3 e 4, *cf.* Capítulo 4) ocorreram entre datasets separados (quantitativo e qualitativo). Os dados foram combinados com o objetivo de analisar os resultados em GCs.

Para cada QC necessitou da análise de temas de conceitos relacionados a tratamentos psicoterápicos com base na técnica TCC, tais como: processo terapêutico, técnicas de intervenção, engajamento do paciente, as relações entre pensamentos e emoções que influenciam nas crenças do paciente. Os especialistas de domínio utilizaram os GCs para analisar os temas para responder às QCs. Ao responder as questões fundamentaram-se na AT, especificamente em atividades descritas no Quadro 5: familiarização com os dados; definição de temas; e produção de um relato (BRAUN; CLARKE, 2006; JOFFE, 2011). Os

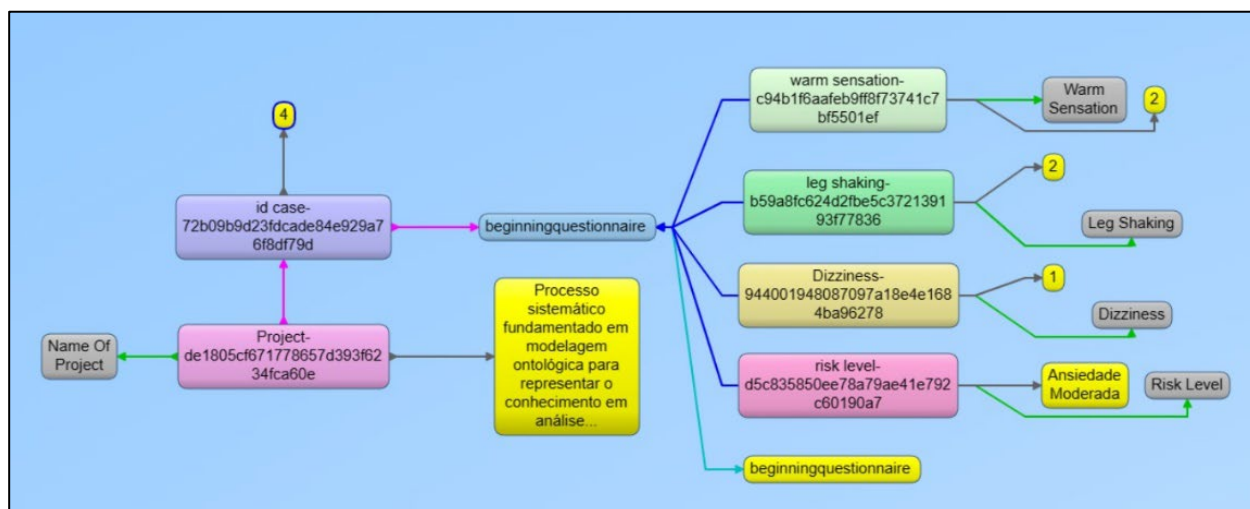
especialistas de domínio compreenderam os dados ao participarem na definição de domínio de aplicação e anotação semântica com SDD. Os temas foram analisados ao responder as QCs. Os relatos da análise são apresentados na resposta de cada QC. Os especialistas de domínio usaram diagramas (gerados pela ferramenta *Gruff*) e tabelas como meios de representação extraídos dos GCs. Os meios de representação são mecanismos utilizados para aprimorar a análise quali-quantitativa e estão definidos conforme discutido em Mayring (2007), Verdinelli e Scagnoli (2013) e Meister (2017).

Ao responder as QCs, os códigos foram organizados como variáveis nominais possibilitando a análise por frequências e percentuais que os códigos foram utilizados. Ao executar a Atividade 3 (etapa do processo definido na seção 5.3) possibilitou o uso dos códigos como variáveis, ao transpor (em colunas) o *dataset* exportado da codificação. As pontuações (valores numéricos) dos sinais de sintomas extraídos dos questionários (inicial e final do tratamento), são analisadas como variáveis quantitativas (BUSSAB; MORETTIN, 2010; FREITAS *et al.*, 2009; MAYER, 2023).

A seguir apresentam-se as QCs. Fragmentos de GCs suportaram a análise quali-quantitativa para os especialistas de domínio responderem as questões. O Apêndice B lista as instruções SPARQL elaboradas para gerar cada fragmento de GC.

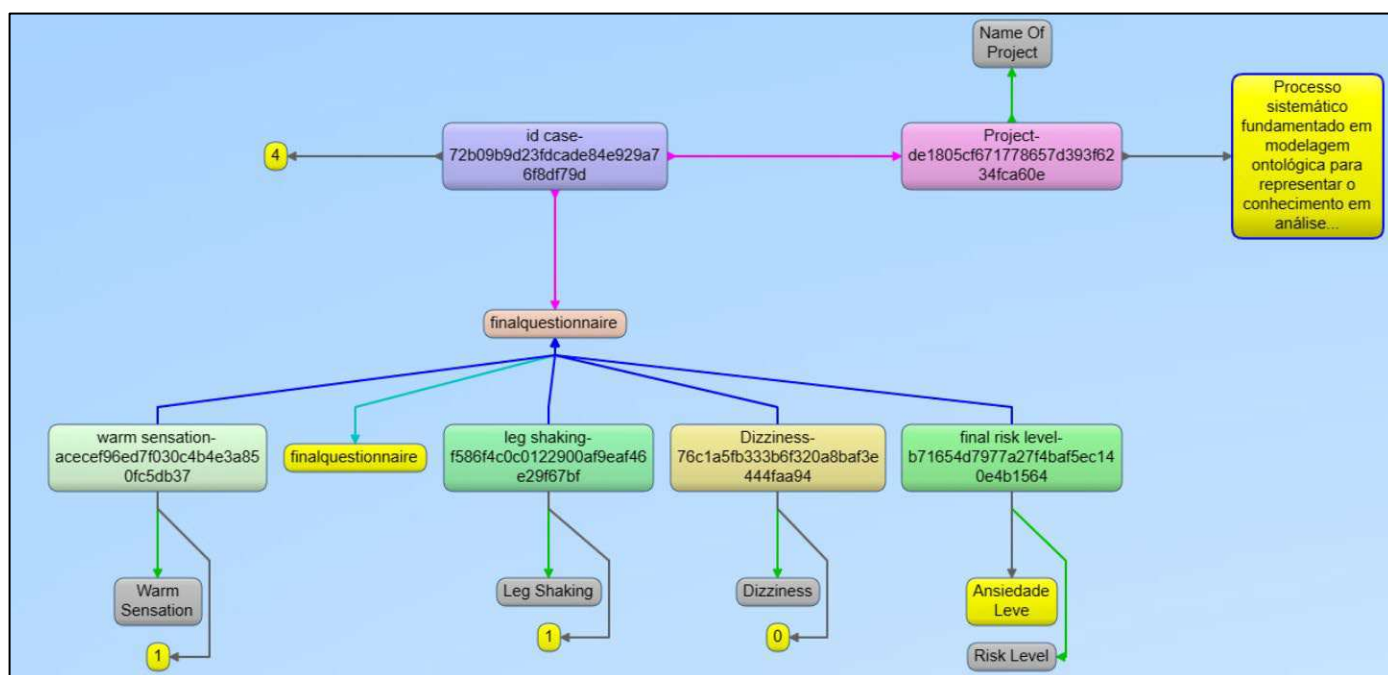
1. Houve reduções nos sintomas apresentados no desenvolver do processo terapêutico? Se sim, quais foram? R. A partir do grafo é possível afirmar que houve uma redução nos sintomas apresentados, uma vez que se compara os dados do questionário inicial com os dados do questionário final. Percebe-se uma redução dos scores de questões que possuem os conceitos: *legshaking*, *warmsensation* e *dizziness*; além da redução da ansiedade de moderada para leve, sendo esta última informação o status do nível de risco analisado pelo grupo de pesquisa e considerado como ponto mais relevante. Os GCs que suportaram a análise da resposta desta questão podem ser vistos na Figura 18 e Figura 19.

FIGURA 18 – Grafo de Conhecimento da resposta da QC 1 (a).



Fonte: Elaborado pelo autor.

FIGURA 19 – Grafo de Conhecimento da resposta da QC 1 (b).



Fonte: Elaborado pelo autor.

2. Quais foram as estratégias de intervenção de maior impacto nos pensamentos, emoções e comportamentos da(o) paciente no desenvolver do processo? Quais procedimentos se destacaram? R. A partir da comparação entre o questionário inicial e final, é possível observar que não só os sintomas de ansiedade diminuíram, como também o próprio “nível”, indo de moderada para leve. Considerando as intervenções e protocolos voltados para o tratamento de ansiedade, as estratégias de enfrentamento em relação ao objeto evocador

da ansiedade costumam ser assertivas e conseqüentemente geram resultados significativos, pois a partir do momento que o paciente enfrenta a situação, ele se vê mais capacitado para lidar com as adversidades que antes lhe causavam medo, assim é possível que haja uma reestruturação em nível cognitivo e comportamental. Nesse caso, o procedimento que mais se destacou foi de fato os *CopingStrategies* (10 vezes) seguido da *Psychoeducation* (4 vezes), procedimentos altamente eficazes para o tratamento da ansiedade. A Figura 20 apresenta os dados quantitativos que suportaram a análise da resposta desta questão.

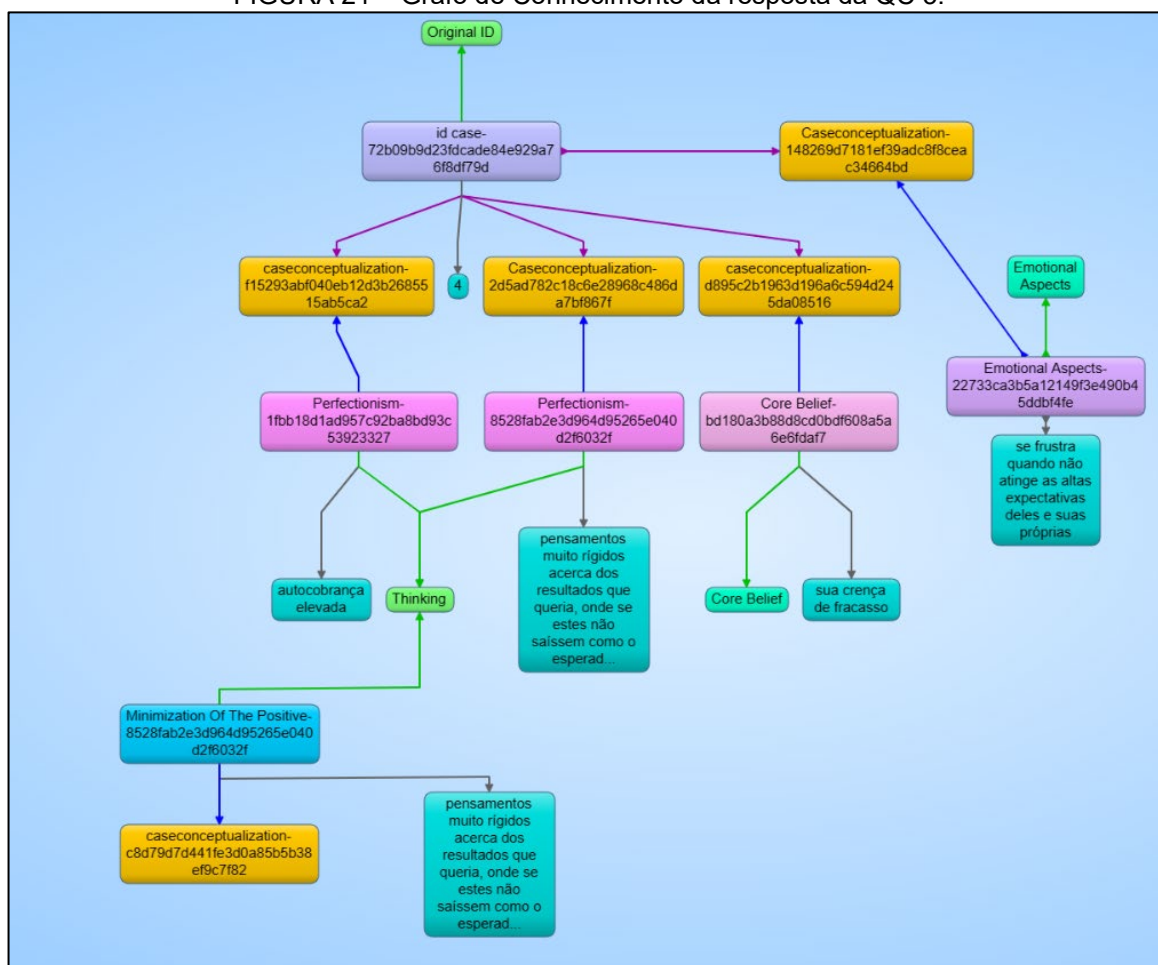
FIGURA 20 – Dados quantitativos da resposta da QC 2.

case	symptomextracted	frequency	questionnariename	valuefearoftheworst happening	valuenervous	valuefeelingofsuffocation	valuerisklevel
4	CopingStrategies	10					
4	ExposureTechniques	2					
4	EvidenceAnalysis	2					
4	Thought	3					
4	Psychoeducation	4					
4	ThoughtStop	3					
4	CopingWithStress	2					
4	Self	1					
4	SocraticQuestioning	3					
4	ClinicalStrategies	3					
4	SystematicDesensitization	2					
4	RecordingDysfunctionalThoughts	2					
4			beginningquestionnaire	2	2	2	Moderate Anxiety
4			finalquestionnaire	1	1	0	Mild Anxiety

Fonte: Elaborado pelo autor.

3. Quais as relações entre pensamentos, emoções e crenças da(o) paciente? R. A partir da análise dos grafos, é possível perceber claramente as relações entre pensamentos, emoções e crenças do paciente. Começando pela CoreBelief “sou um fracasso” é possível estabelecer que se trata de uma crença central de Desvalor. Uma estratégia compensatória muito comum a essa crença é o perfeccionismo e, conseqüentemente, uma autocobrança elevada. Nesse contexto, o paciente se cobra para atingir a perfeição, ao não atingir as altas expectativas, sua emoção tende a ser a frustração, e a crença “sou um fracasso” acaba sendo reforçada (eg. “se eu não atingir a perfeição, logo sou um fracasso”). O GC apresentado na Figura 21 suporta a análise para a resposta desta questão.

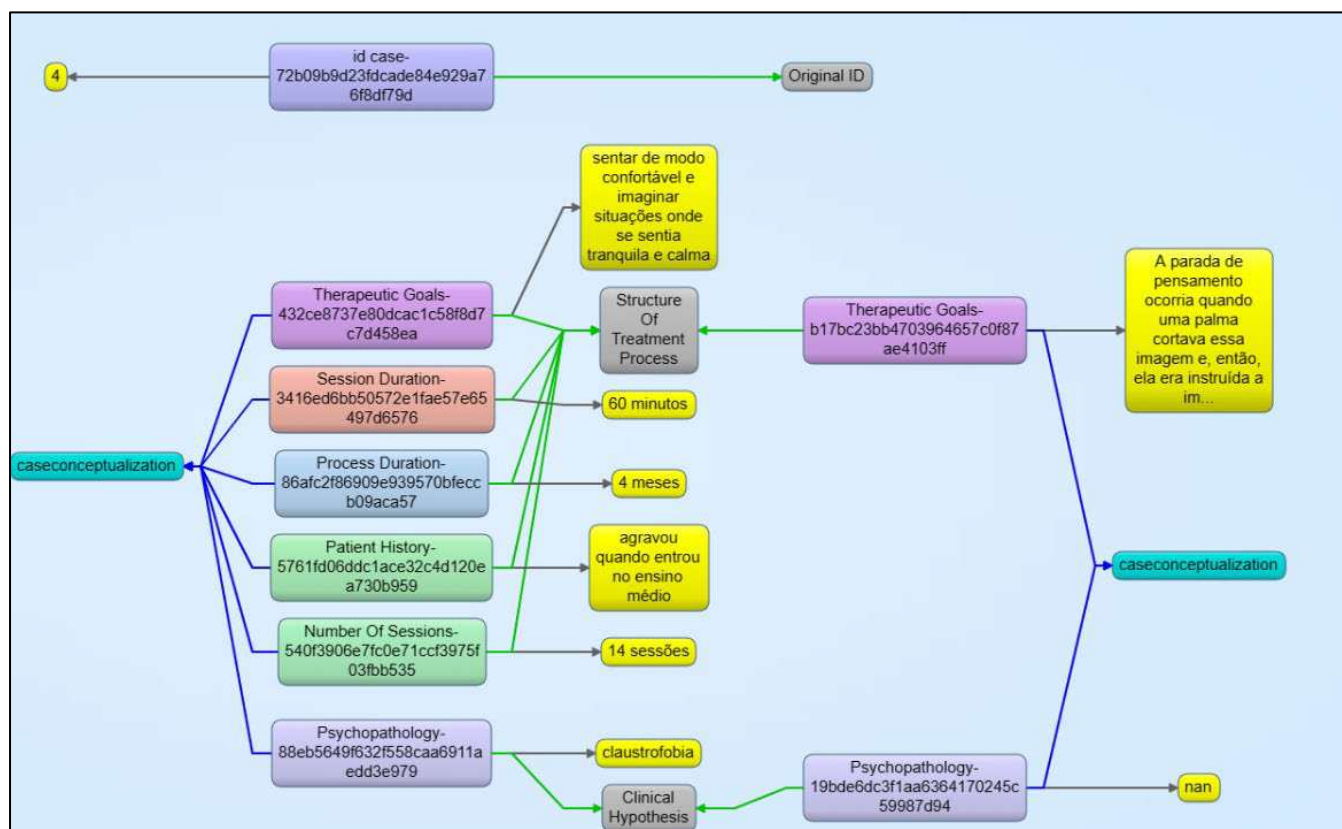
FIGURA 21 – Grafo de Conhecimento da resposta da QC 3.



Fonte: Elaborado pelo autor.

4. Quais as demandas e objetivos apresentados pela(o) paciente? De que maneira o processo os alcança? R. Os objetivos estão bem claros no grafo: “sentar de modo confortável e imaginar situações em que se sentia tranquila e calma”. Já no caso das demandas, ficou pouco claro um mapeamento de todas elas. Temos apenas os objetivos, a hipótese diagnóstica de claustrofobia e o histórico reduzido de vida a respeito do “agravamento no ensino médio”. É possível que o processo alcance as demandas do paciente, entretanto, os dados apresentados no grafo poderiam ser comparados a um protocolo ou manual usado para tratamento da claustrofobia. O GC apresentado na Figura 22 suporta a análise para a resposta desta questão.

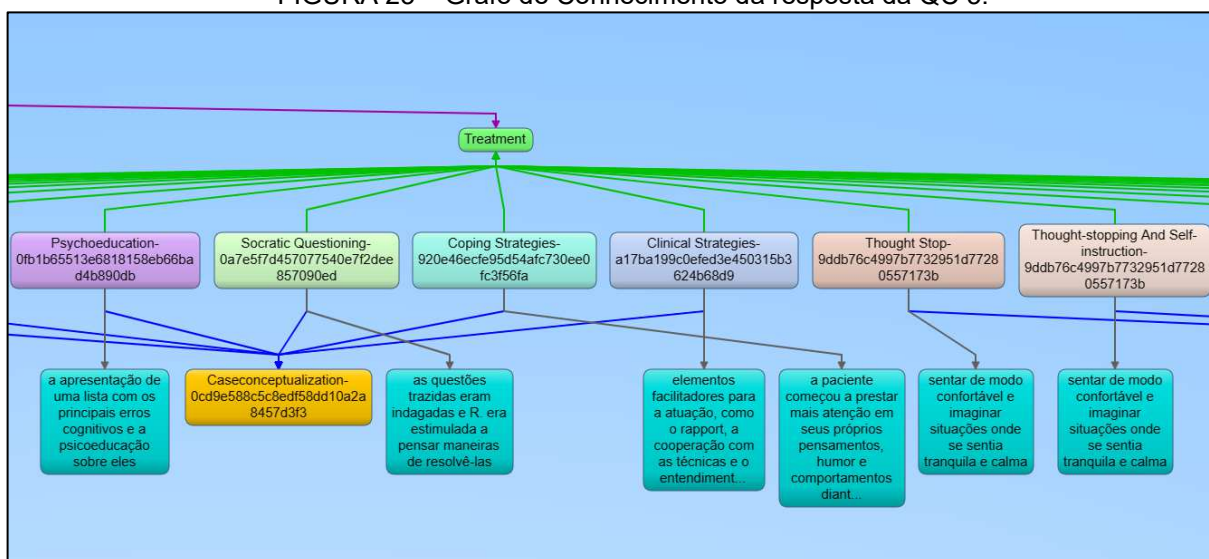
FIGURA 22 – Grafo de Conhecimento da resposta da QC 4.



Fonte: Elaborado pelo autor.

5. Quais os fatores que interferiram nos resultados do processo? Foram de cunho particular ou contextual? R. O tratamento foi realizado a partir de diversas intervenções, com o principal objetivo de auxiliar na promoção da autoeficácia do paciente (termo referente à capacidade de uma pessoa enxergar seu potencial para lidar com situações adversas). Tais fatores são de cunho particular e contextual. Uma vez feita a análise do grafo é possível observar a prevalência de *Coping Strategies*, tal estratégia parte dos dois pontos, o cunho particular quando o paciente sente confiança para realizar a exposição/enfrentamento, mas também o cunho contextual que envolve a situação de exposição e as respostas geradas a partir desse evento. O GC apresentado na Figura 23 suporta a análise para a resposta desta questão.

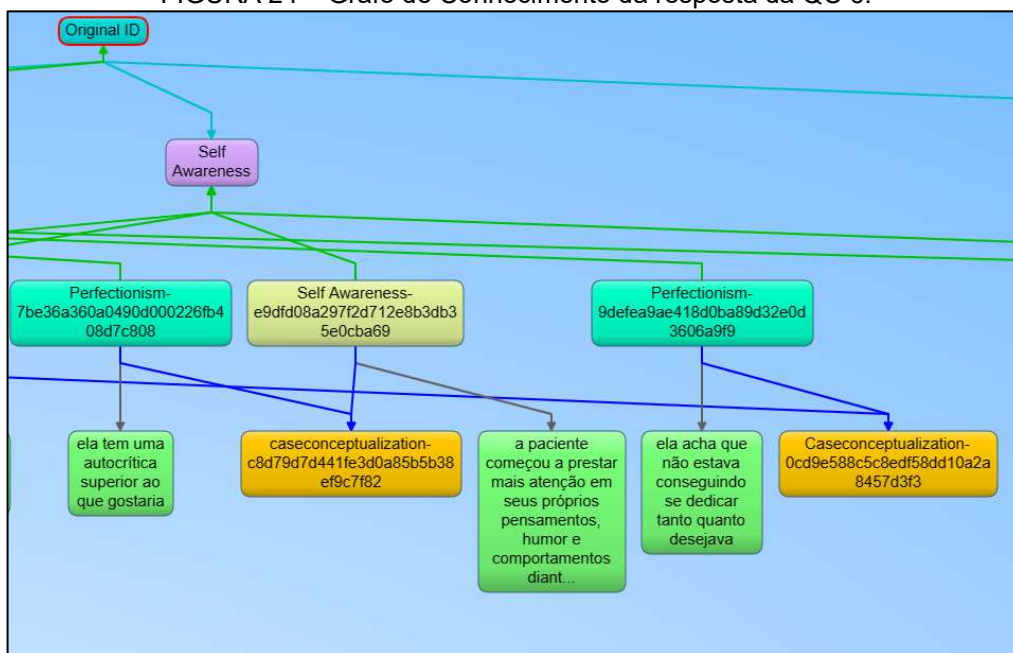
FIGURA 23 – Grafo de Conhecimento da resposta da QC 5.



Fonte: Elaborado pelo autor.

6. Quais valores influenciam no engajamento do paciente para o tratamento? Quais as funções atribuídas a eles? R. O perfeccionismo não é necessariamente algo ruim, entretanto quando há excesso, causa sofrimento significativo no paciente. No caso do paciente citado, por mais que o perfeccionismo gerasse tal sofrimento, pode ser que esse fator tenha influenciado, de certa forma, positivamente para o engajamento da paciente, outro fator importante no caso foi a capacidade de SelfAwareness facilitando o desenvolvimento das intervenções. O GC apresentado na Figura 24 suporta a análise para a resposta desta questão.

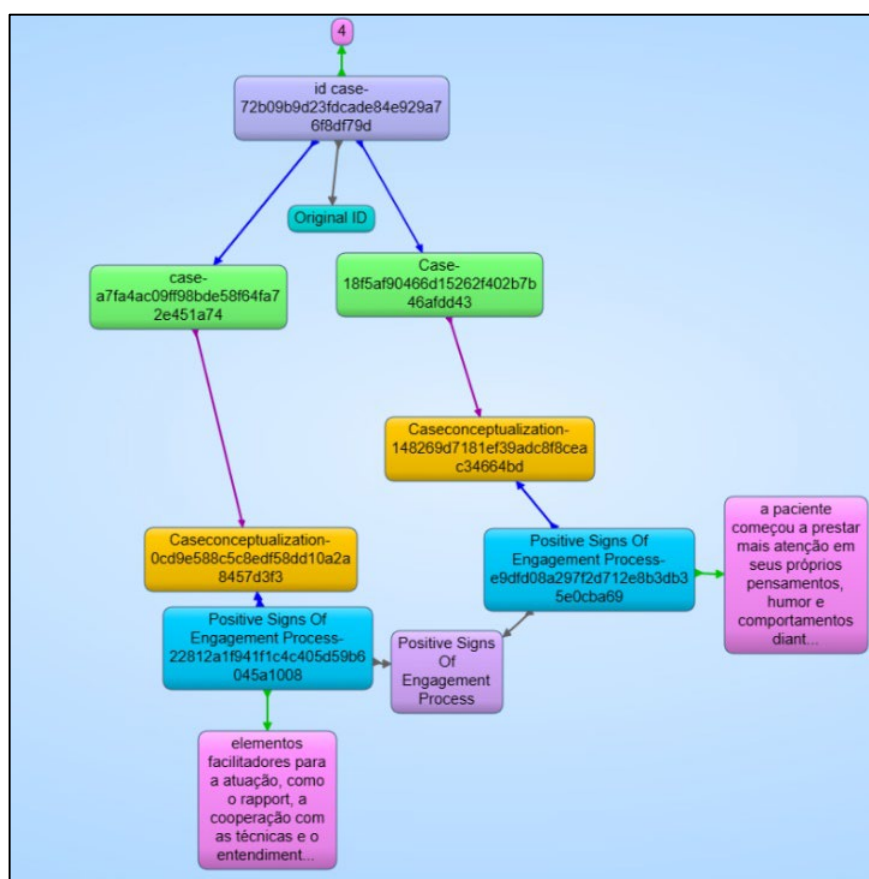
FIGURA 24 – Grafo de Conhecimento da resposta da QC 6.



Fonte: Elaborado pelo autor.

7. Quais os sinais, positivos e negativos, de engajamento da(o) paciente diante do processo? R. A partir do grafo e do *code PositiveSignsOfEngagementProcess*, é possível observar elementos facilitadores para o engajamento, como estabelecimento de *rappor*t, cooperação e entendimento em relação às técnicas, além da paciente ter aprendido a identificar seus próprios pensamentos e emoções e comportamentos. O GC apresentado na Figura 25 suporta a análise para a resposta desta questão.

FIGURA 25 – Grafo de Conhecimento da resposta da QC 7.



Fonte: Elaborado pelo autor.

8. Qual a prevalência dos sintomas que influenciam na hipótese clínica? R. A prevalência dos sintomas psicológicos influencia na hipótese clínica, entretanto, seria necessário identificar mais especificamente quais são esses sintomas para que seja possível relacioná-los diretamente à hipótese. Para um diagnóstico, é provável que seja necessário um número maior de dados. A Figura 26 apresenta os dados quantitativos que suportaram a análise da resposta desta questão.

FIGURA 26 – Dados quantitativos da resposta da QC 8.

case	classsymptoms	symptomextracted	frequency	percentage
4	http://ontriscal#PhysiologicalSymptoms	Nervous	4	32
4	http://ontriscal#PhysiologicalSymptoms	Sweating	2	15
4	http://ontriscal#PhysiologicalSymptoms	MuscularTension	2	15
4	http://ontriscal#PhysiologicalSymptoms	LossOfSpatialPerception	2	15
4	http://ontriscal#PhysiologicalSymptoms	ShortnessOfBreath	3	23

Fonte: Elaborado pelo autor.

9. Qual a prevalência dos códigos que influenciam na validação da crença central? R.

No cenário do caso 4, a prevalência de *AvoidanceOfThreateningSignsOrSituations*, *Fear*, *EmotionalReasoning* e todos os códigos da aba *Thinking*, influenciam e são influenciados diretamente pela crença central. Entretanto, seria mais interessante saber qual a crença central para avaliar de forma mais abrangente os dados. A Figura 27 apresenta os dados quantitativos que suportaram a análise da resposta desta questão.

FIGURA 27 – Dados quantitativos da resposta da QC 9.

Behavior Aspects				
case	classsymptoms	symptomextracted	frequency	percentage
4	http://ontriscal#BehaviorAspects	AvoidanceOfThreateningSignsOrSituations	400000	36
4	http://ontriscal#BehaviorAspects	BehaviorSymptoms	292000	26
4	http://ontriscal#BehaviorAspects	BehaviorAspects	411200	38
EmotionalAspect				
case	classsymptoms	symptomextracted	frequency	percentage
4	http://ontriscal#EmotionalAspect	EmotionalAspects	292000	30
4	http://ontriscal#EmotionalAspect	FearOfLosingControl	205600	21
4	http://ontriscal#EmotionalAspect	EmotionalReasoning	52000	5
4	http://ontriscal#EmotionalAspect	Frustration	138400	14
4	http://ontriscal#EmotionalAspect	Sadness	292000	30
4	http://ontriscal#EmotionalAspect	Fear	4000	0
Thinking				
case	classsymptoms	symptomextracted	frequency	percentage
4	http://ontriscal#Thinking	MinimizationOfThePositive	52000	27
4	http://ontriscal#Thinking	MaximizationOfTheNegative	138400	71
4	http://ontriscal#Thinking	Perfectionism	4000	2

Fonte: Elaborado pelo autor.

10. Quais intervenções realizadas no tratamento permitiram a redução do nível e valor do risco em ansiedade? R. Considerando a análise da prevalência de *AvoidanceOfThreateningSignsOrSituations*, o paciente apresenta um padrão claro de esquiva dos problemas e situações ameaçadoras em geral. No entanto, considerando os fatores listados, ou seja, as intervenções realizadas, todas têm um papel fundamental no tratamento do paciente 4. Entretanto, a prevalência de *CopingStrategies* indica um bom prognóstico, por meio da qual o paciente se esquiva das dificuldades e busca por estratégias de enfrentamento. Esta situação produz uma tendência para a melhora do paciente, à medida que ele se expõe aos eventos, sendo possível aumentar não só seu repertório comportamental, como também

estruturar cognitivamente o paciente acerca das situações vividas diariamente. A Figura 28 apresenta os dados quantitativos que suportaram a análise da resposta desta questão.

FIGURA 28 – Dados quantitativos da resposta da QC 10.

case	classsymptoms	symptomextracted	frequency	percentage
4	http://ontriscal#Treatment	EvidenceAnalysis	292000	10
4	http://ontriscal#Treatment	CopingWithStress	292000	10
4	http://ontriscal#Treatment	Thought	205600	8
4	http://ontriscal#Treatment	Psychoeducation	138400	5
4	http://ontriscal#Treatment	ThoughtStop	205600	7
4	http://ontriscal#Treatment	Self	400000	15
4	http://ontriscal#Treatment	ClinicalStrategies	205600	7
4	http://ontriscal#Treatment	SocraticQuestioning	205600	7
4	http://ontriscal#Treatment	SystematicDesensitization	292000	10
4	http://ontriscal#Treatment	RecordingDysfunctionalThought	292000	10
4	http://ontriscal#Treatment	CopingStrategies	4000	0
4	http://ontriscal#Treatment	ExposureTechniques	292000	11

Fonte: Elaborado pelo autor.

A Matriz de Rastreabilidade foi revisada e o impacto da mudança foi em 2 artefatos produzidos nesta etapa. São eles: grafo de conhecimento e consulta SPARQL criadas para responder às QCs. O versionamento ocorreu no repositório do projeto (<https://github.com/evaldo/homoginise>).

5.7 Controle

Para controlar a mudança dos artefatos e o impacto das alterações é necessário utilizar uma matriz de rastreabilidade. A matriz proposta neste estudo de caso organiza os requisitos para executar a anotação semântica divididos em três grupos: a-) etapas do método HOMOGENiSE; b-) preenchimento dos metadados do SDD; e c-) manipulação dos dados anotados. Os grupos de requisitos estão relacionados com documentos e categorias de tecnologias utilizadas para atingir o objetivo geral de pesquisa, responder à questão de pesquisa e QCs. A Figura 29 apresenta a matriz de rastreabilidade proposta para esta pesquisa. A técnica de versionamento existente na Engenharia de Software foi utilizada nesta etapa. Um repositório com objetos versionados foi criado utilizando a plataforma GitHub.

FIGURA 29 - Matriz de Rastreabilidade.

Documentos e Tecnologias utilizados	Controle dos Artefatos											
	Método HOMOGENISE				Preenchimento dos Metadados					Manipulação dos dados anotados		
	Identificação de problemas	Objetivos de Pesquisa	Design e Desenvolvimento	Demonstração	DM	CodeBook	TimeLine	Anotação	Dados Categóricos	Geração do GC	Aspectos temporais	Ingestão do GC
1 Metadados SDD												
1.1 Infosheet			X		X	X	X					
1.2 Dictionary Mapping	X	X	X		X			X				
1.3 CodeBook	X	X	X			X		X	X			
1.4 TimeLine	X	X	X				X	X			X	
2 Minimundo	X											
3 Perguntas de Competência	X	X						X				
4 Objetivos	X	X										
5 Ontologia	X	X	X					X	X	X	X	
6 Dataset	X		X					X	X		X	
7 Script de software			X							X		X
8 Grafo de Conhecimento	X	X	X							X		X
9 Repositório de Dados Semânticos			X	X								X
10 Consultas SPARQL				X								

Fonte: Elaborado pelo autor.

6 ONTOLOGIA ONTRISCAL

A anotação semântica dos dados quali-quantitativos (*cf.* Capítulo 5), encontra-se sustentada na estrutura de conhecimento modelada pela ontologia ONTRISCAL. O estudo de caso se concentrou em executar as etapas do método HOMOGENiSE, que permitiu usar técnicas de modelagem ontológica originando o conhecimento organizado na ONTRISCAL, criada e atualizada durante a execução do método. Dados anotados semanticamente por SDDs possibilitaram a geração de GCs, que foram usados por especialistas de domínio na análise quali-quantitativa de QCs, especificamente, relacionadas a tratamentos psicoterápicos em ansiedade. Os conceitos da ONTRISCAL foram usados nos SDDs para anotar (enriquecer) semanticamente os dados, facilitando a análise quali-quantitativa.

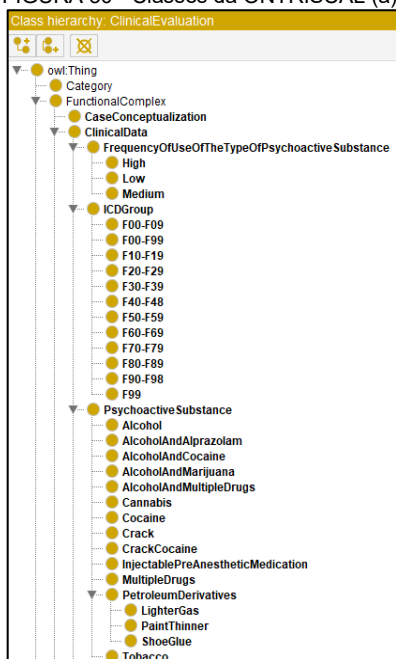
Este capítulo define a ONTRISCAL como contribuição desta tese, apresentando possíveis extrações de conhecimento acerca do domínio da Estratificação de Risco em Saúde Mental, avançando além da aplicação descrita no estudo de caso. E por fim, demonstra como a ontologia pode ser acessada por tecnologias que podem ser manuseadas para representar e inferir novos conhecimentos.

6.1 Ontologia da Estratificação de Riscos em Saúde Mental

Esta seção apresenta a estrutura de conceitos e propriedades da Estratificação de Risco em Saúde Mental modelados na ontologia ONTRISCAL, com base na descrição do domínio de aplicação (*cf.* Seção 2.9), e na modelagem ontológica apresentada no Capítulo 5. Resumidamente, a estratificação visa determinar o nível de risco do paciente por meio da pontuação de sinais e sintomas. Os tratamentos psicoterapêuticos visam mitigar o risco do paciente, produzindo dados qualitativos relatados nas sessões terapêuticas.

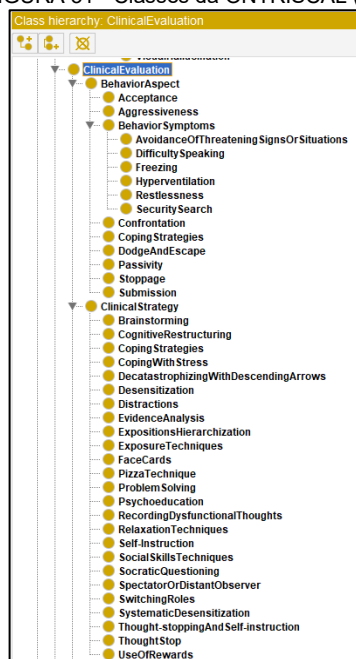
A ONTRISCAL foi criada para apoiar projetos que compartilham conhecimento sobre dados quali-quantitativos produzidos pela Estratificação de Risco em Saúde Mental. Sua estrutura abrange conceitos para anotar dados clínicos, sociodemográficos e sobre níveis de cuidados separados em atendimento clínico, ambulatorial e hospitalar; para compor questionários; descrever profissionais e equipes multiprofissionais (enfermeiros, psicólogos ou médicos). Esses conceitos podem ser reutilizados em aplicações que aplicam conhecimentos em estratificação de risco em saúde mental em diferentes níveis de assistência. A Figura 30, dividida em (a), (b), (c), (d), (e) e (f), exibe a estrutura dentro do Protegé.

FIGURA 30 - Classes da ONTRISCAL (a).



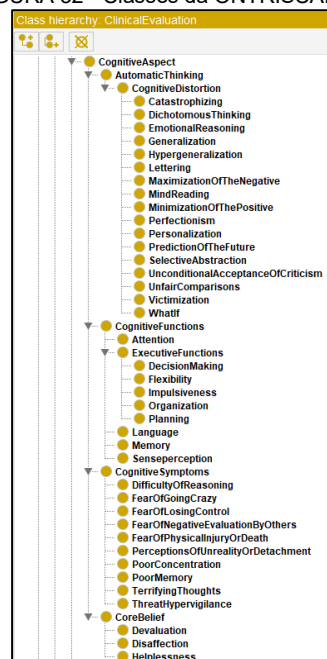
Fonte: Elaborado pelo autor.

FIGURA 31 - Classes da ONTRISCAL (b).



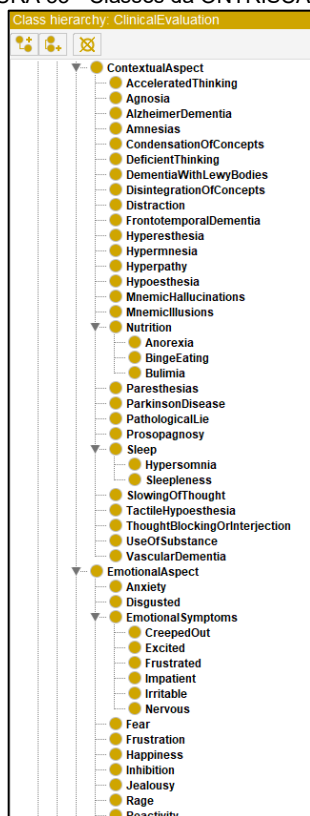
Fonte: Elaborado pelo autor.

FIGURA 32 - Classes da ONTRISCAL (c).



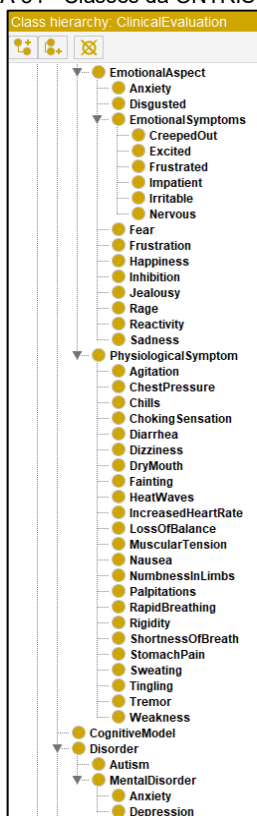
Fonte: Elaborado pelo autor.

FIGURA 33 - Classes da ONTRISCAL (d).



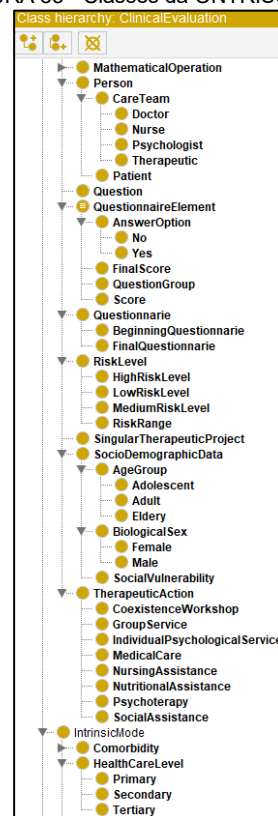
Fonte: Elaborado pelo autor.

FIGURA 34 - Classes da ONTRISCAL (e).



Fonte: Elaborado pelo autor.

FIGURA 35 - Classes da ONTRISCAL (f).



Fonte: Elaborado pelo autor.

A modelagem ontológica foi realizada para todos os conceitos e propriedades do domínio que foram relevantes durante a execução do estudo de caso. Conceitos e

propriedades podem ser combinados com dados ou fatos, abrindo caminho para que novos conhecimentos sejam inferidos.

A semântica das classes e propriedades da ONTRISCAL é descrita formalmente pela linguagem OWL. Ela encontra-se publicada na *web*, onde a documentação e o código-fonte podem ser conferidos na íntegra através dos seguintes ambientes digitais: BioPortal (BIOPORTAL, 2023) e GitHub²¹.

Dessa forma, a ONTRISCAL pode ser consumida por especialistas de domínio, máquinas, banco de dados e aplicações web em geral. É um artefato computacional que compreende 332 classes, 52 relações entre essas classes, 82 instâncias de indivíduos e sua profundidade máxima é 5. A Figura 36 apresenta informações indexadas no portal BioPortal onde toda documentação da ONTRISCAL pode ser encontrada.

FIGURA 36 - Métricas da ONTRISCAL no BioPortal.

Metrics ?	
Classes	332
Individuals	82
Properties	37
Maximum depth	5
Maximum number of children	58
Average number of children	7
Classes with a single child	4
Classes with more than 25 children	3

Fonte: Elaborado pelo autor

Os conceitos da ONTRISCAL são descritos em língua inglesa com anotações semânticas em língua portuguesa, visando explicar termos traduzidos que possuem um significado específico em contexto brasileiro. Por exemplo, o termo “*SingularTherapeuticProject*” é associado a uma definição em língua portuguesa que o relaciona ao conceito de Projeto Terapêutico Singular, proposto como instrumento de organização do cuidado em saúde mental, considerando particularidades de cada paciente (BRASIL, 2014).

6.2 Extração de Conhecimento em saúde mental

Para exemplificar a aplicação da ONTRISCAL, além das QCs respondidas no estudo de caso (*cf.* Capítulo 5), esta seção sugere questões que avançam na extração de

²¹ <https://github.com/evaldo/homoginise/tree/master/Ontologia>

conhecimento em saúde mental sob diferentes aspectos, dentre eles: grupos de questões submetidas ao paciente que avaliam sinais e sintomas; nível de risco associado a dados sociodemográficos; fatores que influenciam níveis de cuidados, e; parâmetros que definem a assistência em saúde mental.

As QCs listadas a seguir, diferentemente daquelas apresentadas no estudo de caso, não proporcionam a análise de risco sobre um transtorno mental específico (*i.e.*, ansiedade), mas detalham problemas relacionados aos aspectos citados anteriormente nesta seção. São elas:

1. Quais sintomatologias estão associadas aos grupos de questões a serem respondidas pelo paciente?
2. Qual o nível de risco por faixa etária de paciente?
3. Quais fatores influenciam os níveis de cuidados em saúde mental?

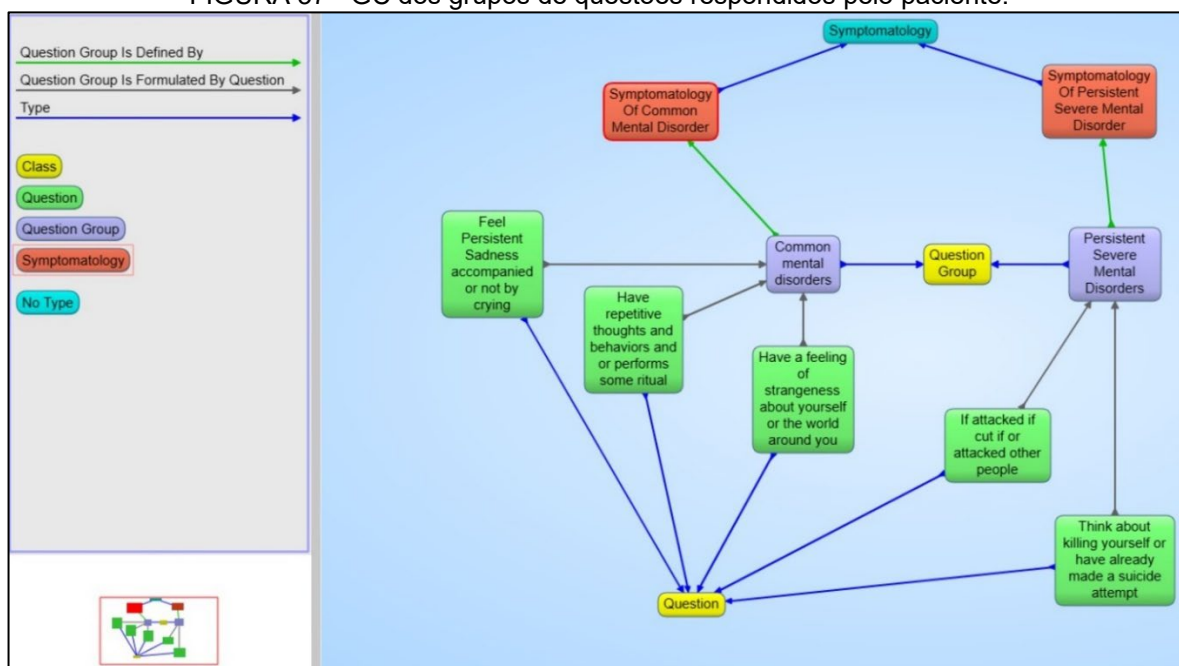
Classes de conceitos presentes na estrutura da ONTRISCAL podem orientar a análise das QCs: *Question*; *QuestionGroup*; *Symptomatology*; *SocioDemographicData*; *HealthCareLevel*, e; *TherapeuticAction*.

A ONTRISCAL oferece, ainda, propriedades que relacionam as classes citadas, facilitando a extração do conhecimento. Instâncias de indivíduos foram adicionadas na ontologia visando exemplificar a recuperação do conhecimento por meio de triplas.

As figuras adiante (Figura 37, Figura 38 e Figura 39), exemplificam GCs gerados a partir dos conceitos e propriedades da ONTRISCAL e que embasam as respostas para cada QC. Para concretizar a extração de conhecimento discutida, a estrutura da ontologia foi exportada da ferramenta Protegé em formato RDF, e, em seguida, foi importada para o *AllegroGraph*. Cada figura representa o GC gerado pela instrução SPARQL elaborada para responder à respectiva questão de competência. A instrução SPARQL a seguir visa responder à QC 1.

```
# Instrução SPARQL para responder a QC.1
PREFIX ontr: <http://ontriscal#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
select distinct ?QuestionGroup ?isFormulatedBy ?type ?Question where {
  ?Question rdf:type ontr:Question .
  ?QuestionGroup rdf:type ontr:QuestionGroup .
  ?Symptomatology rdf:type ontr:Symptomatology .
  ?Question ?isFormulatedBy ?QuestionGroup .
  ?QuestionGroup ?isDefinedBy ?Symptomatology . }
```

FIGURA 37 - GC dos grupos de questões respondidos pelo paciente.



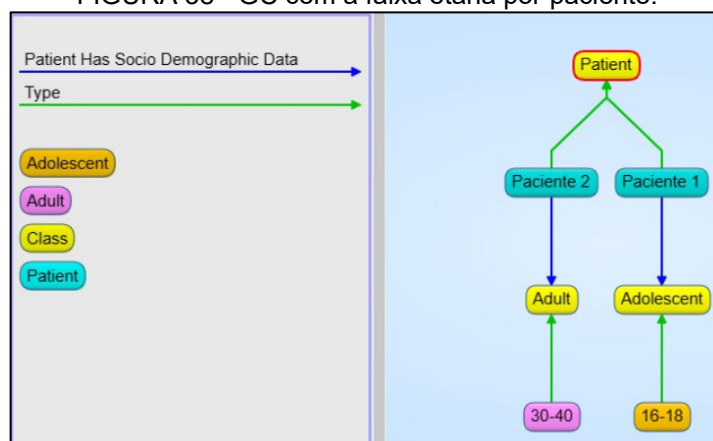
Fonte: Elaborado pelo autor

De acordo com Foa, Stein e Mcfarlane (2006) a sintomatologia é uma coleção de sintomas do paciente, identificada por sinais e sintomas manifestados pelo transtorno mental. A SESA utiliza o agrupamento de questões na estratificação de risco para mapear a sintomatologia do paciente. Esta mesma organização em grupos de questões foi utilizada pelo HVVSM ao executar o projeto de pesquisa descrito no ANEXO 1. A fundamentação teórica e a execução do projeto, abriu caminho para compreender e organizar este conhecimento na ontologia.

Desta forma, a ONTRISCAL modela os conceitos representados na Figura 37, respondendo à QC 1. Ou seja, as sintomatologias relacionadas, respectivamente, à “*Common mental disorders*” e “*Persistent Severe Mental Disorders*”, agrupam questões que norteiam a análise do especialista de domínio visando identificar se o paciente já tentou suicídio (transtorno severo), ou se chora com frequência acompanhado com o sentimento de tristeza. A instrução SPARQL a seguir visa responder à QC 2.

```
# Instrução SPARQL para responder a QC.2
PREFIX ontr: <http://ontriscal#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
select distinct ?Patient ?AgeGroup ?valueAgeGroup where {
  ?Patient rdf:type ontr:Patient .
  ?Patient ontr:patientHasSocioDemographicData ?AgeGroup .
  ?valueAgeGroup rdf:type ?AgeGroup .
}
```


FIGURA 38 - GC com a faixa etária por paciente.

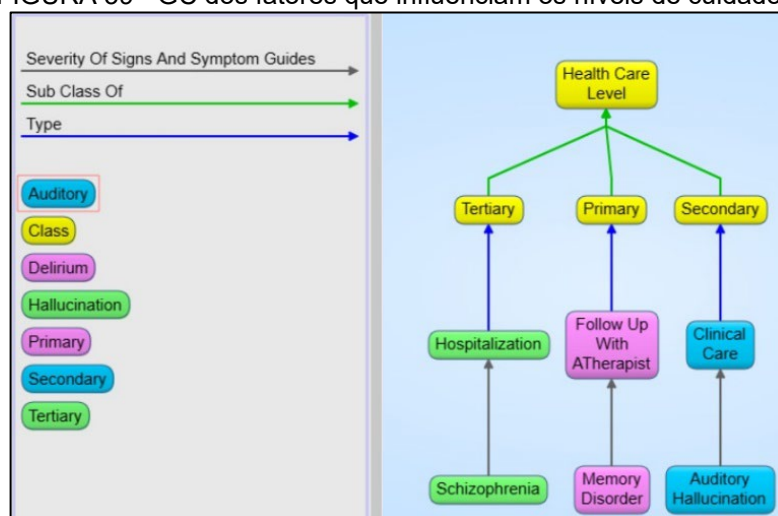


Fonte: Elaborado pelo autor

A modelagem da faixa etária por paciente é importante para o dimensionamento do atendimento em saúde mental. De acordo Rodrigues, Rodrigues e Cardoso (2020) a política de saúde mental estabelecida pelo SUS deve identificar a faixa etária dos pacientes visando estruturar ações e investimentos para consolidar a política no país. Diante da importância das análises de dados sócios demográficos, a ONTRISCAL modela os conceitos representados na Figura 38, respondendo à QC 2. A instrução SPARQL a seguir visa responder à QC 3.

```
# Instrução SPARQL para responder a QC.3
PREFIX ontr: <http://ontriscal#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
select distinct ?levelcare ?signandsymptomseverity ?typecarelevel ?healthcarelevel where {
  ?healthcarelevel rdfs:type ?typecarelevel .
  ?typecarelevel rdfs:subClassOf ?levelcare .
  ?signandsymptomseverity ontr:severityOfSignsAndSymptomGuides ?healthcarelevel .
}
```

FIGURA 39 - GC dos fatores que influenciam os níveis de cuidados.



Fonte: Elaborado pelo autor.

A experiência vivida ao executar o projeto no HVVSM (cf. ANEXO 1), permitiu modelar os conceitos relacionados aos níveis de cuidados de acordo com severidade dos sinais e sintomas do paciente. O HVVSM disponibilizava três níveis de atendimento: primário, com o acompanhamento diário por terapeutas; secundário, atendimento ambulatorial ou clínico, e; terciário, internação em hospital. De acordo com o Ministério da Saúde, a estratificação de risco pode ser uma ferramenta para orientar a assistência à saúde em níveis de atendimento, em relação aos transtornos e faixas etárias identificados (SANTA CATARINA, 2019).

A ONTRISCAL modela conceitos que visam estruturar o nível de atendimento, nível de risco e a severidade dos sinais e sintomas. Os conceitos organizados no GC da Figura 39 respondem à QC 3, e exemplificam como transtornos mentais severos podem estar associados aos níveis de assistência à saúde, indicando o atendimento especializado.

6.3 Inferência de conhecimento

Esta seção sugere axiomas e regras lógicas com o objetivo de estruturar a ONTRISCAL com parâmetros, a fim de tornar possível inferir novos conhecimentos. Para isso, a linguagem SWRL foi utilizada. As regras lógicas propostas encontram-se na forma de afirmações, e que podem ser verdadeiras caso determinadas condições sejam satisfeitas. Utiliza-se o raciocinador Pellet, oferecido pelo Protegé como *plug-in* para aplicar as regras.

Para demonstrar essa aplicação, selecionou-se conceitos e propriedades da ONTRISCAL que tratam do nível de risco em saúde mental, a severidade dos transtornos e a faixa etária para estabelecer os parâmetros. Uma vez que o encaminhamento de assistência à saúde depende de compreender e correlacionar esses conceitos, torna-se importante explorar esse tema nesta seção.

O resultado inferido pelo raciocinador não deve ser considerado um encaminhamento definitivo para assistência à saúde mental, sem a avaliação de especialistas do domínio. Este tipo de funcionalidade estruturada na ONTRISCAL visa fornecer meios ou evidências úteis para auxiliar o especialista a tomar decisões.

A Figura 40, representa o resultado para o encaminhamento à assistência no nível “*Tertiary*”, ou seja, para hospitalização caso o paciente seja adulto, com sinais e sintomas severos e estratificação de risco de nível alto. O seguinte código-fonte em linguagem SWRL foi utilizado para esta inferência:

```
Patient(?p) ^ hasHighRiskLevel(?p, true) ^ hasSeveritySeverityOfSignAndSymptom(?p, true)
^ isAdult(?p, true) -> takesCare(?p, Tertiary)
```

FIGURA 40 - Inferência do encaminhamento para nível terciário.

Fonte: Elaborado pelo autor

O nível de cuidado em saúde mental estabelece cuidados, tratamentos terapêuticos ou assistência médica. No nível primário ocorre o acolhimento, socialização, atividades coletivas ou até mesmo práticas de atividades físicas. No nível secundário o paciente recebe atendimento multiprofissional, atendimento aos familiares e reinserção social. Finalmente, no nível terciário podem ser estabelecido projetos terapêuticos singulares para melhoria da saúde mental de cada paciente e atendimento em hospital especializado em Psiquiatria (GARLA, FUREGATO e SANTOS, 2011; SANTA CATARINA, 2019).

Diante dessas definições, uma vez que o nível de atendimento terciário é confirmado, o paciente deve ser encaminhado para tratamentos específicos para este nível. Tal situação também pode ser inferida pelo raciocinador, e é representada pela Figura 41. O seguinte código-fonte em linguagem SWRL foi utilizado para esta inferência:

```
ontriscal:Patient(?p) ^ ontriscal:hasHealthLevelCare_Tertiary(?p,true) -> recebeis(?p, SingularTherapeuticProject) ^ recevieis(?p, MedicalCare)
```

FIGURA 41 - Inferência de cuidados para o nível terciário.

Fonte: Elaborado pelo autor

Este capítulo definiu a ontologia ONTRISCAL como contribuição desta tese acerca do domínio da Estratificação de Risco em Saúde Mental. Além disso, demonstrou como a ontologia pode ser acessada por tecnologias que podem ser manuseadas para representar e inferir novos conhecimentos.

O próximo capítulo é descrito com as considerações finais deste trabalho e que evidenciam o cumprimento dos objetivos, dificuldades na execução do método HOMOGENiSE, e as sugestões para trabalhos futuros.

7 CONSIDERAÇÕES FINAIS

A pesquisa especificou etapas para definir o método HOMOGENiSE, que utiliza a compreensão e preparação de dados, visando explicitar o conhecimento implícito em dados quali-quantitativos. O HOMOGENiSE propõe sete etapas, entre elas a etapa de Concepção e Desenvolvimento, que descreve um processo fundamentado em modelagem ontológica compreendido por atividades que determinam como compreender e preparar dados para análises quali-quantitativas por meio de GCs. O método foi aplicado com dados sintéticos extraídos de artigos científicos que avaliam tratamentos psicoterápicos. Um estudo de caso demonstrou a anotação semântica por meio de SDDs e a geração de GCs populados com dados, denotando que é possível aprimorar análises quali-quantitativas em relação ao estado da arte. Para tal, no estudo de caso, buscou-se responder às QCs de interesse dos especialistas de domínio em um caso concreto de pesquisa na área da Psicologia.

Os resultados sugerem que o método HOMOGENiSE aprimora a análise quali-quantitativa, visando superar os desafios apresentados ao aplicar essa abordagem de pesquisa. Resumidamente, os obstáculos para o avanço de pesquisas quali-quantitativas são: formalizar técnicas para produzir dados quali-quantitativos; compartilhar o conhecimento sobre as diferentes variáveis; reutilizar o conhecimento para a codificação, e visualizar os dados sob várias perspectivas (CASTERLÉ *et al.*, 2012; DRISCOLL *et al.*, 2007; HUNTER *et al.*, 2002; SOMEKH; LEWIN, 2005; VERDINELLI; SCAGNOLI, 2013).

Ao considerar os resultados sugeridos em relação aos desafios, o método HOMOGENiSE possui etapas que formalizam a produção de dados quali-quantitativos. Ele aborda GCs como artefatos que podem ser compartilhados por pesquisadores, tornando-se possível representar os dados quali-quantitativos sob várias perspectivas e possibilitando a extração de fragmentos do conhecimento a fim de facilitar a análise dos dados. Ao propor o uso de ontologias como estrutura que fornece o conhecimento para codificação em análise qualitativa, o método HOMOGENiSE contribui para ampliar o reuso da codificação, independentemente de CAQDAS.

O estudo de caso teve a participação de especialistas de domínio da área de psicologia que apoiaram na definição do escopo, determinando os conceitos e propriedades para modelagem da ontologia ONTRISCAL. A anotação semântica dos dados implementou três tipos de SDDs (SDD quali, SDD quali domínio e SDD quantitativo), com vistas a gerar um GC que integrasse dados quantitativos e qualitativos, e permitisse análises quali-quantitativas. Os especialistas de domínio apoiaram o preenchimento dos SDDs e anotaram os dados com classes da ONTRISCAL, REFI-QDA e QualiCO, enriquecendo os GCs gerados com conceitos do domínio e conceitos de análise qualitativa, respectivamente.

O objetivo de fortalecer a tese, bem como o interesse dos especialistas na análise quali-quantitativa, motivaram a elaboração de QCs que pudessem ser respondidas a partir de fragmentos de GCs integrando dados quali-quantitativos. A navegação pelos GCs foi feita utilizando consultas SPARQL que possibilitaram gerar tabelas e diagramas com informações extraídas dos GCs, aprimorando as análises quali-quantitativas ao responder as QCs. Após os apontamentos acima, os objetivos específicos puderam ser retomados com as considerações que visam demonstrar como é/foi possível atingi-los:

- modelar o conhecimento sobre Estratificação de Riscos em Ansiedade em psicoterapia, com o apoio de especialistas de domínio, e construir uma ontologia na área. A ONTRISCAL foi criada como ontologia de domínio, com apoio de especialistas de domínio, para anotar semanticamente os dados quali-quantitativos;
- utilizar a ontologia para codificar dados qualitativos. A Atividade 2, definida pela etapa de Concepção e Desenvolvimento, implementou uma proposta para codificar os dados qualitativos;
- especificar um método de modelagem ontológica que vise aprimorar análises quali-quantitativas por meio de grafos de conhecimento, gerados com a ontologia criada e dados em psicoterapia. O método HOMOGENiSE foi especificado e aplicado conforme apresentado, respectivamente, nos Capítulos 4 e 5;
- aplicar o método especificado, juntamente com especialistas do domínio, para validar as possibilidades de aprimorar análises quali-quantitativas. A Seção 5.6 apresentou fragmentos de GC que possibilitaram diferentes análises quali-quantitativas.

7.1 Contribuições Atingidas

A abordagem *bottom-up* utiliza os dados para explicitar o conhecimento neles embutido. Em conjunto com a técnica SDD, esta abordagem pode reduzir o custo decorrente dos desafios enfrentados para compreender e preparar os dados. Ao explicitar o conhecimento presente nos dados por meio dos SDDs, os GCs são gerados sem a necessidade de implementar a estrutura dos dados, tal como ocorre em modelos relacionais, mas sim como os dados estão conceituados e relacionados no domínio.

Modelos relacionais são criados a partir do modelo entidade-relacionamento até sua implementação por meio de um modelo físico, suportado por algum software específico de banco de dados. Grafos podem combinar diversas classes e propriedades, possibilitando a integração de dados com diferentes estruturas (independentemente de um software específico para suportar os dados), não havendo a preocupação com a navegação pela estrutura do dado, diminuindo o esforço de agregar novos dados. A estrutura oferecida por

grafos permite ser enriquecida com diferentes classes, propriedades e dados, caso os especialistas de domínio necessitem aprofundamento no conhecimento representado. A integração dos diferentes dados é feita por meio de anotação semântica e ontologias. Esses dados podem ser oriundos de diferentes tabelas relacionais. Os grafos permitem a implementação de repositórios semânticos de dados com a flexibilidade adequada para o especialista inferir novos conhecimentos no caminho de repositórios mais flexíveis para diferentes requisitos.

Em conjunto com a *bottom-up*, esta pesquisa utiliza a abordagem *top-down* ao definir uma primeira versão da ontologia a partir da descrição e detalhamento do minimundo. Assim, considera-se que este trabalho utiliza as abordagens citadas de forma complementar para a construção do conhecimento, e que permitiu exercitar o entendimento sobre o domínio da aplicação, produzindo artefatos e metadados que, preliminarmente, detalham um modelo conceitual inicial do minimundo definido juntamente com especialistas do domínio. A partir deste primeiro entendimento foi possível criar a primeira versão da ontologia chamada ONTRISCAL. Essas abordagens podem ser aplicadas em conjunto tal como definido por Francesconi *et al.* (2010).

Demonstra-se nesta pesquisa a execução da técnica SDD em conjunto com a ontologia ONTRISCAL para interpretar os significados dos conceitos, a fim de desambiguar e preparar os dados do domínio de aplicação. Os dados são organizados em grafos RDF que podem ser explorados por consultas, com o intuito de darem origem a outros *datasets* ou arquivos de dados que são divulgados utilizando um banco de dados semânticos. Outra contribuição a ser considerada é o uso da UFO para fundamentar a ontologia ONTRISCAL.

As demais contribuições seguem listadas no âmbito da análise quali-quantitativa com o uso de GCs. As etapas de demonstração e divulgação do método HOMOGENiSE apresentam como as QCs elaboradas foram respondidas. A visualização dos GCs tornou possível a análise dos dados conectados às respectivas classes de conceitos, orientando as respostas das QCs. Os dados puderam ser recuperados de forma gráfica e tabular, flexibilizando as análises quali-quantitativas. Desta forma, vale retomar a importância de obter os meios de representação apropriados para análise quali-quantitativa conforme discutido por Mayring (2007). Os meios de representação devem ser elementos tais como tabelas, diagramas, modelos de processos, modelos de contextos, modelos de estruturas ou gráficos que permitem ordenar os dados para análise qualitativa e explicitar graficamente os relacionamentos complexos a fim de facilitar a compreensão da pesquisa.

Ao utilizar a ONTRISCAL na codificação, o conhecimento foi estruturado seguindo a ideia de *coding* dedutivo, com o propósito de viabilizar a reutilização deste conhecimento em diferentes projetos de pesquisa e os conceitos da ontologia como códigos.

Ao retomar Driscoll *et al.* (2007), os autores discutem que os métodos mistos sejam executados para integrar dados quali-quantitativos incluindo técnicas que modifiquem uma forma de dados para outra. A etapa de Concepção e Desenvolvimento define atividades para selecionar, integrar e preparar dados, permitindo implementar (unificar) a estrutura de dados qualitativos compatível com formato tabular.

Ainda no âmbito da anotação semântica, considera-se o uso das classes de conceitos do projeto REFI-QDA e da QualiCO. O resultado da anotação semântica é a geração de GC por onde análises quali-quantitativas podem acontecer com o uso de diferentes conceitos, incluindo aqueles que definem a codificação, e propriedades que enriquecem especificamente as análises.

Finalmente, a navegação pelos códigos populados no GC pode exemplificar o uso da GT (seção 5.4.1) é considerada como contribuição. Uma vez que a ontologia anota semanticamente os dados qualitativos por meio dos códigos, as conexões entre eles podem existir por meio da navegação no GC, sendo possível descobrir novas conexões, e até alcançar uma nova categoria de dado ou teoria.

7.1.1. Publicações relacionadas ao tema

Esta seção aborda as publicações relacionadas ao tema desta pesquisa como contribuições. As publicações seguem enumeradas com título, resumidas em ordem cronológica e classificadas por eventos e periódicos. Além disso, encontram-se inseridas nas referências deste trabalho para maiores detalhes.

1. Título: Compreensão e preparação de dados aplicados na estratificação de risco em ansiedade. Data de publicação: 28 de Nov. de 2022. Periódico: Ciência Da Informação Express. Resumo: A publicação é resultado do projeto de pesquisa realizado em 2022 em conjunto com a Clínica de Psicologia do UniAcademia (*cf.* ANEXO 2). Silva *et al.*(2022) aborda a compreensão e preparação de dados para o desenvolvimento de pesquisas na área da psicologia, a fim de aprimorar análises de casos clínicos, especificamente, no tratamento da ansiedade. Discute ainda a elaboração de procedimentos para compreender dados estruturados identificados a partir estratificação de risco em ansiedade, juntamente com dados textuais (não estruturados) de tratamentos psicoterápicos a fim de prepará-los para análise qualitativa e quantitativa.

2. Título: Processo sistemático fundamentado em modelagem ontológica aplicado à estratificação de risco em saúde mental para análise quali-quant: 06 de Dez de 2022. Evento: Workshop de dados, informação e tecnologia (WIDAT). Resumo: De acordo com Silva, Marques e Bax (2022) abordam o processo sistemático fundamentado na modelagem ontológica para análise quantitativa e qualitativa de dados da estratificação de risco em ansiedade, em conjunto com dados extraídos dos tratamentos psicoterápicos. O processo faz o uso de ontologias e a técnica SDD (*Semantic Data Dictionary*) para compreensão e preparação dos dados. São gerados grafos de conhecimento que contém os dados anotados semanticamente, podendo ser úteis para encontrar relações entre doenças, sintomas e tratamentos além de servir como ferramenta para formação de diagnóstico e auxílio na definição dos níveis de cuidados com o paciente.

3. Título: Uso de dicionários semânticos na estratificação de riscos em saúde mental. Data de publicação: 09 de Dez de 2022. Evento: Encontro nacional de pesquisa e pós-graduação em ciência da informação (ENACIB). Resumo: Silva e Bax (2022) propõe a modelagem ontológica da estratificação de riscos em ansiedade em conjunto com a anotação semântica dos dados utilizando a técnica *Semantic Data Dictionary* (SDD). O artigo sugeriu a geração de GCs a partir da anotação semântica, para facilitar a visualização das informações e a descoberta de relações entre, p.ex., doenças, sinais e sintomas, servindo como ferramenta para aprimorar o processo de estratificação, além de auxiliar na definição dos níveis de cuidados com o paciente.

7.2 Dificuldades ao executar o método HOMOGENiSE

As principais dificuldades encontram-se em obter os dados qualitativos. Embora esta dificuldade seja de ordem geral para todas as pesquisas, já que a tarefa de codificação demanda expertise e tempo para ser realizada, o método HOMOGENiSE estabelece que a codificação deve utilizar a estrutura da ontologia de domínio, uma dificuldade adicional para aqueles que não dominam ontologias. A tarefa de exportar os conceitos da ontologia para criar os códigos foi realizada manualmente, criando uma dificuldade técnica adicional para a codificação dos dados qualitativos. Além disso, os especialistas de domínio necessitaram de apoio técnico para controlar os novos códigos a serem acrescentados na ontologia, visando enriquecê-la possibilitando sua reutilização em outras codificações.

No domínio de aplicação deste trabalho, o entendimento foi realizar a codificação para cada sessão de tratamento psicoterápico. No entanto, equívocos podem ocorrer caso o texto não esteja adequadamente escrito. Observou-se que a tarefa de codificação demanda tempo

para gerar coleções de dados com qualidade, e se o texto estiver inadequado (ou desorganizado em relação à sequência de sessões) a codificação pode gerar ambiguidades e enviesar a análise quali-quantitativa.

Um desafio ao executar o método HOMOGENiSE foi a tentativa de trabalho colaborativo entre os especialistas de domínio. A codificação executada de forma colaborativa possibilita que vários códigos sejam criados por diferentes especialistas. Neste caso, a ontologia deve ser evoluída considerando os novos códigos definidos como classes de conceitos na ontologia. A dificuldade encontra-se em como controlar, ou evoluir a ontologia com os novos conceitos, exigindo um procedimento para sistematizar ou organizar o versionamento da ontologia. Os especialistas de domínio precisariam conhecer boas práticas de versionamento de artefatos. Sistemas de Gerenciamento Eletrônico de Documentos poderiam contribuir para minimizar este problema, oferecendo suporte como repositórios para controlar a evolução da ontologia.

Outra dificuldade encontra-se na compreensão das técnicas para análise quali-quantitativa, ou separadamente, análise qualitativa e quantitativa, e em como aplicá-las com o suporte de GCs. Observou-se a dificuldade dos especialistas de domínio em compreender e aplicar as técnicas, sendo necessário que tenham experiência no desenvolvimento de pesquisas com a análise quali-quantitativa.

7.3 Trabalhos Futuros

Como trabalhos futuros espera-se a execução do método HOMOGENiSE usando dados reais de tratamentos psicoterápicos. A execução do método diversas vezes, por diferentes especialistas, permitirá o refinamento das etapas do método e o enriquecimento da ontologia ONTRISCAL. Durante o desenvolvimento desta pesquisa, o projeto foi submetido para ser aplicado com pacientes voluntários. A submissão foi aceita para que a pesquisa pudesse ter continuidade com dados reais. A aprovação do projeto na Plataforma Brasil²² pode ser vista no Anexo 3.

Ao executar o método HOMOGENiSE, os especialistas de domínio não abordaram uma técnica específica de análise quali-quantitativa para responder as QCs. Desta forma, espera-se aplicar diferentes técnicas de análise quali-quantitativa visando explorar mais amplamente os GCs. É importante definir qual técnica de análise quali-quantitativa será aplicada ao responder as QCs, no intuito de avaliar a utilidade dos GCs para suportar as respostas. Caso os GCs não

22 <https://www.gov.br/pt-br/servicos/submeter-na-plataforma-brasil-de-projetos-de-pesquisa-envolvendo-seres-humanos-para-avaliacao-etica>

gerem as respostas, deve-se reavaliar a anotação semântica por meio dos SDDs e refinar os GCs.

Espera-se aplicar a *GT* com apoio de GCs na análise quali-quantitativa de dados psicoterápicos reais. O HOMOGENiSE será aplicado como continuidade desta pesquisa em um projeto conduzido com pacientes atendidos na Clínica de Psicologia do UniAcademia. O projeto possibilitará manipular um conjunto de dados maior para análise quali-quantitativa. O uso de GCs com dados quali-quantitativos poderá ser uma alternativa para aplicar a *GT* mais amplamente, podendo trazer contribuições tanto para área de GC quanto para *GT*.

Como continuidade desta pesquisa pretende-se sistematizar a execução do método HOMOGENiSE, principalmente das atividades definidas na etapa de Concepção e Desenvolvimento. Essa sistematização envolve integrar atividades de forma que possam ser executadas por meio de uma aplicação de software.

REFERÊNCIAS

- AIZENBUD-RESHEF, Neta et al. **Model traceability**. IBM Systems Journal, v. 45, n. 3, p. 515-526, 2006.
- ALMEIDA, M. B.; BAX, M. P. **Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção**. Ciência da Informação, Brasília DF, v. 32, n. 3, p. 7-20, 2003.
- AMERICAN PSYCHIATRIC ASSOCIATION, A. P. et al. **Diagnostic and statistical manual of mental disorders (DSM-IV)**. Washington: Washington, DC: American psychiatric association, 1994.
- AMDOUNI, Emna; GIBAUD, Bernard. **Imaging Biomarker Ontology (IBO): A Biomedical Ontology to Annotate and Share Imaging Biomarker Data**. Journal on Data Semantics 7, no. 4 (2018): 223-236.
- ANÁLISE. In: MICHAELIS, **Dicionário Brasileiro da Língua Portuguesa**. Brasil: Melhoramentos, 2023. Disponível em: < <https://tinyurl.com/ycyu5ysw>>. Acesso em: 23/01/2023.
- ARONSON, Jodi. **"A pragmatic view of thematic analysis."** The qualitative report 2, no. 1 (1995): 1-3.
- BAGDONIENĖ, Liudmila; ZEMBLYTĖ, Jurgita. **"Service research: Advantages and limitations of quantitative and qualitative approaches."** Socialiniai mokslai 4 (2005): 26-37.
- BARALT, Melissa. **"Coding Qualitative Data."** Research methods in second language acquisition (2012): 222.
- BARBOSA, Daniel Mendes. **Um modelo de gestão da informação e do conhecimento para o contexto da avaliação de cursos de graduação**. Saberes científicos da biblioteconomia em diálogo com as ciências sociais e humanas. 163p. Tese (Doutorado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2016.
- BAX, M. P. (2015). **Design science: filosofia da pesquisa em ciência da informação e tecnologia**. Ciência Da Informação, 42(2). Recuperado de <http://revista.ibict.br/ciinf/article/view/1388>.
- BAX, M. P. ; SILVA, Evaldo de Oliveira . **Dicionários Semânticos de Dados para Integrar Dados de Prontuários Eletrônicos de Pacientes** (Semantic Data Dictionary to Integrate Data of Electronic Medical Records). In: Ontobras, 2020, Vitória. Proceedings of the XIII Seminar on Ontology Research in Brazil and IV Doctoral and Masters Consortium on Ontologies (ONTOBRAS 2020). Aachen: CEUR, 2020. v. 2728. p. 298-303.
- BAYER, Fábio Mariano; SOUZA, Adriano Mendonça. **"Análise de Correspondência entre clínicas médicas e indicadores hospitalares: o caso do Hospital Universitário de Santa Maria."** IV Simpósio de Excelência em Gestão e Tecnologia. http://www.aedb.br/seget/historico_seget2007.html (acessado em 10/Jun/2011) (2007).

BENFARES, Chaymae; EL IDRISSEI, Younès El Bouzekri; HAMID, Karim. **Personalized healthcare system based on ontologies**. In: International Conference on Advanced Intelligent Systems for Sustainable Development. Springer, Cham, 2018. p. 185-196.

BEREIJJO, Antonio. **THE CATEGORY OF “APPLIED SCIENCE” AN ANALYSIS OF ITS JUSTIFICATION FROM “INFORMATION SCIENCE” AS DESIGN SCIENCE**. In: Scientific Realism and Democratic Society. Brill Rodopi, 2011. p. 327-350.

BERTHOLD, M.R., BORGELT, C., HÖPPNER, F., KLAWONN, F.: **Guide to intelligent data analysis: how to intelligently make sense of real data**. Springer (2010).

BIOLCHINI, Jorge; MILAN, Paula Gomes; NATALI, Ana Candida Cruz; TRAVASSOS, Guilherme Horta Travassos. **Systematic review in software engineering**. System engineering and computer science department COPPE/UFRJ, Technical Report ES 679, no. 05 (2005): 45.

BIOPORTAL. **Ontology of Risk Stratification in Mental Health**. Disponível em <https://bioportal.bioontology.org/ontologies/ONTRISCAL>. Acesso em 28 de dez de 2023.

BRASIL. Ministério da Saúde. **Departamento de Atenção Básica. Cadernos de Atenção Básica: Núcleo de Apoio à Saúde da Família - Volume 1: ferramentas para a gestão e para o trabalho cotidiano**. Brasília: Ministério da Saúde; 2014. Disponível em: http://bvsmms.saude.gov.br/bvs/publicacoes/nucleo_apoio_saude_familia_cab39.pdf

BRAUN, Virginia; CLARKE, Victoria. **"Using thematic analysis in psychology"**. Qualitative research in psychology 3, no. 2 (2006): 77-101.

BRENAS, Jon Hael; SHIN, Eun Kyong; SHABAN-NEJAD, Arash. **Adverse childhood experiences ontology for mental health surveillance, research, and evaluation: advanced knowledge representation and semantic web techniques**. JMIR mental health, v. 6, n. 5, p. e13498, 2019.

BRISSON, L.; COLLARD, M. (2008). **An ontology driven data mining process**. In International Conference on Enterprise Information Systems (pp. 54-61).

BUSSAB, Wilton de O.; MORETTIN, Pedro A. **"Estatística básica"**. In Estatística básica, pp. xvi-540. 2010.

CALERO, C.; RUIZ, F.; PIATTINI, M. **Ontologies for Software Engineering and Software Technology**. Berlin: Springer, 2006.

CASANOVA, M. A.; BRAUNER, D. F.; CÂMARA, G.; JÚNIOR, P. D. O. L. 2005. **Integração e interoperabilidade entre fontes de dados geográficos**. CASANOVA, MA et al. Banco de dados geográficos. Curitiba: Mundogeo, 315-352.

CERVONE, H. Frank. **Understanding agile project management methods using Scrum**. OCLC Systems & Services: International digital library perspectives, 2011.

CEUSTERS, Werner; SMITH, Barry. **Foundations for a realist ontology of mental disease**. Journal of biomedical semantics, v. 1, n. 1, p. 1-23, 2010.

CHARLOTTE, Siefridt; GROSJEAN, Julien; LEFEBVRE, Tatiana; ROLLIN, Laetitia; DARMONI, Stefan; SCHUERS, Matthieu. **Evaluation of automatic annotation by a multi-**

terminological concepts extractor within a corpus of data from family medicine consultations. International Journal of Medical Informatics 133 (2020): 104009.

CRAMER, Duncan. **Advanced quantitative data analysis.** McGraw-Hill Education (UK), 2003.

CRESWELL, J.W.; CRESWELL, J.D. 2018. **Research Design. Qualitative, Quantitative, and Mixed Methods Approaches.** Fifth Edition. Los Angeles: Sage.

CRESWELL, John W; PLANO CLARK, Vicki L. **Designing and conducting mixed methods research.** Sage publications, 2017.

DANUBIANU, Mirela; PENTIUC, Ștefan Gheorghe; TOBOLCEA, Iolanda. **CRISP-DM Model Applied for Knowledge Discovery in Speech Disorders Therapy Area.** 2010. Disponível em: <https://fiesc.usv.ro/wp-content/uploads/sites/17/2021/01/Danubianu1.pdf>. Acesso em 02 de junho de 2022.

DAWADI, Saraswati; SHRESTHA, Sagun; GIRI, Ram A. Giri. **"Mixed-methods research: A discussion on its types, challenges, and criticisms."** Journal of Practical Studies in Education 2, no. 2 (2021): 25-36.

CASTERLÉ, DIERCKX Bernadette de; GASTMANS, Chris; BRYON, Els; DENIER, Yvonne. **"QUAGOL: A guide for qualitative data analysis."** International journal of nursing studies 49, no. 3 (2012): 360-371.

DELVE, Intro to Grounded Theory Analysis with Delve. **Qualitative Data Analysis Tips.** YouTube, 2022. Disponível em: <https://www.youtube.com/watch?v=0wcv4amjfk0&t=393>

DOOLEY, Damion M; GRIFFITHS, Emma J.; GOSAL, Gurinder; BRINKMAN, Fiona SL; HSIAO, William WL. **The Genomic Epidemiology Ontology and GEEM Ontology Reusability Platform.** In JOWO. 2017.

COSTA, Danielli dos Reis; GUARNIER, Thayza Sacconi; SANTOS, Andréia Soprani; BUBACH, Susana; RISSINO, Sílvia das Dores; TEIXEIRA, Maria das Graças da Silva. **Construção de uma ontologia de referência para o domínio da HIV/aids.** Revista Eletrônica de Comunicação, Informação & Inovação em Saúde 17, no. 3 (2023).

DRISCOLL, David L; APPIAH-YEBOAH, Afua; Philip Salib; RUPERT, Douglas J. **"Merging qualitative and quantitative data in mixed methods research: How to and why not."** (2007).

DURASAMY, Gunavathi; ATAN, Rodziah. **Requirement traceability matrix through documentation for scrum methodology.** Journal of Theoretical & Applied Information Technology, v. 52, n. 2, p. 154-159, 2013.

EL GHOSH, Mirna; NAJA, H.; ABDULRAB, H.; KHALIL, M. **Towards a middle-out approach for building legal domain reference ontology.** International Journal of Knowledge Engineering 2, no. 3 (2016): 109-114.

EVERS, Jeanine C. **Current issues in qualitative data analysis software (QDAS): A user and developer perspective.** The Qualitative Report 23, no. 13 (2018): 61-73.

FALBO, R.A. **SABiO: Systematic Approach for Building Ontologies.** In: Joint Workshop ONTO.COM / ODISE on Ontologies in Conceptual Modeling and Information Systems

Engineering (Co-located with 8th International Conference on Formal Ontology in Information Systems, FOIS 2014). Rio de Janeiro – RJ, Brazil. 2014

FARINELLI F; ELKIN PL. **Construção de ontologia na prática: um estudo de caso aplicado ao domínio obstétrico**. Ciência da Informação. 2017 Dec 28;46(1).

FOA, E. B., STEIN, D. J., & MCFARLANE, A. C. (2006). **Symptomatology and psychopathology of mental health problems after disaster**. J Clin Psychiatry, 67(Suppl 2), 15-25.

FRANCESCONI, Enrico; MONTEMAGNI, Simonetta; PETERS, Wim; TISCORNIA, Daniela. **Integrating a bottom-up and top-down methodology for building semantic resources for the multilingual legal domain**. Springer Berlin Heidelberg, 2010.

FREITAS, Henrique; JANISSEK, Raquel. **"Análise léxica e análise de conteúdo: técnicas complementares, seqüenciais e recorrentes para exploração de dados qualitativos."** Porto Alegre: Sagra Luzzatto (2000).

FREITAS, Henrique; JANISSEK-MUNIZ R.; COSTA, R.; ANDRIOTTI, F.; FREITAS, P. **Guia Prático Sphinx**. Canoas RS 1 (2009).

FREITAS, H.; MOSCAROLA, J. (2000). **Análise de dados quantitativos e qualitativos: casos aplicados usando o Sphinx**. Porto Alegre: Sphinx.

FRIESE, Susanne; SYBING, Roehl. **Qualitative vs Quantitative Research**. Disponível em: <https://atlasti.com/research-hub/qualitative-vs-quantitative-research>. Acesso em 31 de jan de 2023.

GARLA, C.; FUREGATO, R. F.; SANTOS, L. F. Profissionais de Ambulatórios de Saúde Mental: Perfil, práticas e opiniões sobre as políticas. **Cadernos Brasileiros de Saúde Mental/Brazilian Journal of Mental Health**, [S. l.], v. 2, n. 4-5, p. 73–93, 2011. DOI: 10.5007/cbsm.v2i4-5.68459. Disponível em: <https://periodicos.ufsc.br/index.php/cbsm/article/view/68459>. Acesso em: 30 dez. 2023.

GIBBS, Graham R. **"Grounded theory."** (2010).

GOERTZEN, Melissa J. **"Introduction to quantitative research and data."** Library Technology Reports 53, no. 4 (2017): 12-18.

GONÇALVES, José Eugênio de Assis et al. **Método ágil de integração semântica de dados científicos baseado em ontologias**. Tese de Doutorado. 2020.

GRUBER, T. R. **A Translation Approach to Portable Ontology Specifications**. 1993. Disponível em < http://www-ksl.stanford.edu/KSL_Abstracts/KSL-92-71.html>. Acesso em 04 de agosto de 2021.

GUARINO, Nicola. **Formal Ontology and Information Systems**. In: Proceedings of the First Int. Conference on Formal Ontology in Information Systems, Trento, Italy, Junho 1998.

GUIZZARDI, Giancarlo. **Ontological foundations for structural conceptual models**. 2005.

GUIZZARDI, G.; FALBO, R. A.; GUIZZARDI, R. S. S. **A importância de Ontologias de Fundamentação para a Engenharia de Ontologias de Domínio: o caso do domínio de Processos de Software**. IEEE Latin America Transactions, v. 6, n. 3, p. 244-251, 2008.

GUIZZARDI, Giancarlo et al. **UFO: Unified Foundational Ontology**. *Applied Ontology*, in. Preprint, p. 1-44, 2021.

HASTINGS, Janna et al. **Representing mental functioning: Ontologies for mental health and disease**. 2012.

HEIBI, Ivan; PERONI, Silvio. **A qualitative and quantitative analysis of open citations to retracted articles: the Wakefield 1998 et al.'s case**. *Scientometrics* 126, no. 10 (2021): 8433-8470.

HEVNER, Alan; CHATTERJEE, Samir. **Design science research in information systems**. In: *Design research in information systems*. Springer, Boston, MA, 2010. p. 9-22.

HEFLIN, Jeff; HENDLER, James. **Semantic interoperability on the web**. In *Proceedings of extreme markup languages*, vol. 2000, pp. 111-120. 2000.

HOCKER, J.; BIPAT, T.; MCDONALD, D. W.; ZACHRY, M. (2021a). **Evaluating QualiCO: an ontology to facilitate qualitative methods sharing to support open science**. *Journal of Internet Services and Applications*, 12(1), 1-29.

HOCKER, J., BIPAT, T., MCDONALD, D. W., & Zachry, M. (2021b). **Developing an Ontology for Qualitative Coding Schemas - QualiCO**. Disponível em: <https://eera-ecer.de/ecer-programmes/conference/26/contribution/50889/>. Acesso em 28 de jun de 2022

HOEHNDORF, Robert. **What is an upper level ontology?. Ontogenesis**, 2010. Disponível em: <http://ontogenesis.knowledgeblog.org/740/?kblog-transclude=2>. Acesso em 16 dez 2021.

HOGAN, Aidan, BLOMQVIST, Eva, COCHEZ, Michael, D'AMATO, Claudia, MELO Gerard de, GUTIERREZ, Claudio, GAYO, José Emilio Labra, KIRRANE, Sabrina, NEUMAIER, Sebastian, POLLERES, Axel, NAVIGLI, Roberto, NGOMO, Axel-Cyrille Ngonga, RASHID, Sabbir M., RULA, Anisa, SCHMELZEISEN, Lukas, SEQUEDA, Juan, STAAB, Steffen, ZIMMERMANN, Antoine. **Knowledge Graphs**. arXiv preprint arXiv:2003.02320, 2020.

HUNTER, A., LUSARDI, P., ZUCKER, D., JACELON, C., CHANDLER, G., 2002. **Making meaning: the creative component in qualitative research**. *Qualitative Health Research* 12 (3), 388–398.

IGLESIAS, Carlos A.; SANCHEZ-RADA, J Fernando; VULCU, Gabriela; BUITELAAR, Paul. **Linked data models for sentiment and emotion analysis in social networks**. In *Sentiment Analysis in Social Networks*, pp. 49-69. Morgan Kaufmann, 2017.

ISOTANI, Seiji; BITTENCOURT, Ig Ibert. **Dados abertos conectados: em busca da web do conhecimento**. Novatec Editora, 2015.

JEONG, Serin; CHO, Heetae; LEE, Seonah. **Agile requirement traceability matrix**. In: *Proceedings of the 40th International Conference on Software Engineering: Companion Proceedings*. 2018. p. 187-188.

JOFFE, Helene. **"Thematic analysis"**. *Qualitative research methods in mental health and psychotherapy* 1 (2012): 210-223.

KHANKEH, H., RANJBAR, M., KHORASANI-ZAVAREH, D., ZARGHAM-BOROUJENI, A., & JOHANSSON, E. (2015). **Challenges in conducting qualitative research in health: A conceptual paper**. *Iranian Journal of Nursing and Midwifery Research*, 20(6), 635.

KIGER, Michelle E.; VARPIO, Lara. **"Thematic analysis of qualitative data: AMEE Guide No. 131."** Medical teacher 42, no. 8 (2020): 846-854.

KODALI, Manvisha. **Traceability of requirements in scrum software development process.** 2015.

KUNG, Hsiang-Jui; KUNG, L.; GARDINER, Adrian. **Comparing top-down with bottom-up approaches: teaching data modeling.** In Proceedings of the Information Systems Educators Conference ISSN, vol. 2167, p. 1435. 2012.

LABAREE, Robert V. **"Organizing your social sciences research paper."** (2013). Disponível em: <https://libguides.usc.edu/writingguide/quantitative>

LARMAN, Craig; BASILI, Victor R. **Iterative and incremental developments. a brief history.** Computer, v. 36, n. 6, p. 47-56, 2003.

LAY, Maria Cristina Dias, REIS, Antonio Tarcisio da Luz. **"Análise quantitativa na área de estudos ambiente-comportamento."** Ambiente construído: revista da Associação Nacional de Tecnologia do Ambiente Construído. Vol. 5, n. 2 (abr./jun. 2005), p. 21-36 (2005).

LIBRARY ILLINOIS. **Qualitative Data Analysis: Coding.** Disponível em: <https://guides.library.illinois.edu/qualitative>. Acesso em 25 de jan de 2023.

MATTEO, Cristani; CUEL, Roberta. **A survey on ontology creation methodologies.** International Journal on Semantic Web and Information Systems (IJSWIS) 1, no. 2 (2005): 49-69.

MAYER, Fernando de Pol. **Análise exploratória de dados.** Disponível em: <https://www.inf.ufsc.br/~andre.zibetti/probabilidade/aed.html#>. Acesso em 26 de jan de 2023.

MAYO. **Anxiety Disorders.** 2018. Disponível em: <https://www.mayoclinic.org/diseases-conditions/anxiety/symptoms-causes/syc-20350961>. Acesso em 31 de ago de 2023

MAYRING, Philipp. **"On generalization in qualitatively oriented research."** In Forum Qualitative Sozialforschung/Forum: Qualitative Social Research, vol. 8, no. 3. 2007.

MCCUSKER, J., RASHID, S.M., LIANG, J., LIU, Y., CHASTAIN, K., PINHEIRO, P., STINGONE, J.A., and MCGUINNESS, D.L.: Broad, **Interdisciplinary Science** In Tela: An Exposure and Child Health Ontology. In Proceedings of 9th International ACM Web Science Conference 2017. Troy, NY, USA (2017)

MEISTER, Vera G. **"Towards a Knowledge Graph for a Research Group with Focus on Qualitative Analysis of Scholarly Papers."** In SemSci@ ISWC, pp. 71-76. 2017.

MELO, Ana Luísa Caetano; LOURENÇO, Lélío Moura. **Terapia cognitivo-comportamental no tratamento de um caso de transtorno de ansiedade generalizada: um relato de caso.** Psicologia. (2020). Disponível em <https://www.psicologia.pt/artigos/textos/A1408.pdf>. Acesso em 03 de jul de 2023.

MICROSOFT. **O estágio de entendimento empresarial do ciclo de vida do Processo de Ciência de Dados da Equipe.** 2020. Disponível em: <https://docs.microsoft.com/pt-br/azure/machine-learning/team-data-science-process/lifecycle-business-understanding>. Acesso em: 17 de maio de 2021.

MILLER, Charles C.; REARDON, Michael J.; SAFI, Hazim J. **Risk stratification: a practical guide for clinicians**. Cambridge University Press, 2001.

MINAYO, Maria Cecília de Souza. "**Análise qualitativa: teoria, passos e fidedignidade.**" *Ciência & saúde coletiva* 17 (2012): 621-626.

MIÑARRO-GIMÉNEZ, José Antonio; MARTÍNEZ-COSTA, Catalina; KARLSSON Daniel; SCHULZ Stefan; GØEG, Kirstine Rosenbeck. **Qualitative analysis of manual annotations of clinical text with SNOMED CT**. *Plos one* 13, no. 12 (2018): e0209547.

MORAIS, Carlos. "**Descrição, análise e interpretação de informação quantitativa.**" Disponível em: <http://www.ipb.pt/~cmmm/discip/ConceitosEstatistica.pdf>. Acesso em 09 de set de 2023.

NILES, Ian; PEASE, Adam. **Towards a standard upper ontology**. In: *Proceedings of the international conference on Formal Ontology in Information Systems-Volume 2001*. 2001. p. 2-9.

NOY, Natalya F; MCGUINNESS, Deborah L. **Ontology development 101: A guide to creating your first ontology**. 2001.

OHARA, Tomoyuki et al. **Glucose tolerance status and risk of dementia in the community: the Hisayama study**. *Neurology*, v. 77, n. 12, p. 1126-1134, 2011.

OREN, Eyal et al. **What are semantic annotations**. *DERI Galway*, v. 9, p. 62, 2006.

PAN, Jeff Z. et al. (Ed.). **Exploiting linked data and knowledge graphs in large organisations**. Heidelberg: Springer, 2017.

PARANÁ (Estado). **Rede de Saúde Mental**. Disponível em: <https://www.saude.pr.gov.br/Pagina/Saude-Mental>. Acesso em 09 de abr de 2021.

PASQUIER, Florence et al. Diabetes mellitus and dementia. **Diabetes & metabolism**, v. 32, n. 5, p. 403-414, 2006.

PATON, Norman. **Automating data preparation: Can we? should we? must we?**. In: *21st International Workshop on Design, Optimization, Languages and Analytical Processing of Big Data*. 2019.

PATTON, M. Q. (1999). **Enhancing the quality and credibility of qualitative analysis**. *Health services research*, 34(5 Pt 2), 1189

PAULA, George Luiz Costa de. **Classificação de risco em saúde mental: implicações clínicas, éticas e sócio-políticas**. 2019.

PATHAK, V.; JENA, B.; KALRA, S. (2013). **Qualitative research. Perspectives in Clinical Research**, 4(3), 192-194

PFUFF, Matthias; NEUBIG, Stefan; KRUMHOLTZ, Helmut. **Ontology for semantic data integration in the domain of IT benchmarking**. *Journal on data semantics* 7 (2018): 29-46.

POLKINGHORNE, D. E. (1995). **Narrative configuration in qualitative analysis**. *International Journal of Qualitative Studies in Education*, 8(1), 5-23.

POPPING, Roel. "Text analysis for knowledge graphs." *Quality & quantity* 41 (2007): 691-709

POTH, Cheryl; MUNCE, Sarah EP. **Commentary--Preparing today's researchers for a yet unknown tomorrow: Promising practices for a synergistic and sustainable mentoring approach to mixed methods research learning.** *International Journal of Multiple Research Approaches* 12, no. 1 (2020).

PYLE, Dorian. **Data preparation for data mining.** Morgan Kaufmann, 1999.

QUALITATIVO. In: DICIO, **Dicionário Online de Português.** Portugal: 7Graus, 2020. Disponível em: <<https://www.dicio.com.br/qualitativo/>>. Acesso em: 14/01/2023.

QIN, Jian. **Knowledge organization and representation under the AI lens.** *Journal of Data and Information Science*, v. 5, n. 1, p. 3-17, 2020.

RAFFAITIN, Christelle et al. **Metabolic syndrome and risk for incident Alzheimer's disease or vascular dementia: the Three-City Study.** *Diabetes care*, v. 32, n. 1, p. 169-174, 2009.

RAHMAN, Md Shidur. "The advantages and disadvantages of using qualitative and quantitative approaches and methods in language "testing and assessment" research: A literature review." (2020).

RASHID, Sabbir M. et al. **The Semantic Data Dictionary Approach to Data Annotation & Integration.** In: *SemSci@ ISWC.* 2017. p. 47-54.

RASHID, S. M., MCCUSKER, J. P., PINHEIRO, P., BAX, M. P., SANTOS, H., STINGONE, J. A., ... & MCGUINNESS, D. L. (2020). **The Semantic Data Dictionary--An Approach for Describing and Annotating Data.** *Data Intelligence*, 443-486.

REFI-QDA. **Rotterdam Exchange Format Initiative.** Disponível em : <https://www.qdasoftware.org/>. Acesso em 11 de abril de 2023.

REMOR, Eduardo Augusto. **Tratamento psicológico do medo de viajar de avião, a partir do modelo cognitivo: caso clínico.** *Psicologia: reflexão e crítica* 13 (2000): 205-216.

REYES, A. N.; FERMANN, I.L.; **Eficácia da terapia cognitivo-comportamental no transtorno de ansiedade generalizada.** *Revista Brasileira de Terapias Cognitivas.* 2017.

RITTLE-JOHNSON, Bethany; SIEGLER, Robert S.; ALIBALI, Martha Wagner. **Developing conceptual understanding and procedural skill in mathematics: An iterative process.** *Journal of educational psychology*, v. 93, n. 2, p. 346, 2001.

RODRIGUES, T. A. D. S., RODRIGUES, L. P. D. S., & CARDOSO, Â. M. R. (2020). **Adolescentes usuários de serviço de saúde mental: avaliação da percepção de melhora com o tratamento.** *Jornal Brasileiro de Psiquiatria*, 69, 103-110.

ROSENBERG, Rasmus Larsen; HASTINGS, Janna. **Mapping the patient's experience: An applied ontological framework for phenomenological psychopathology.** *Phenomenology and Mind* 18 (2020): 200-219.

SANDELOWSKI, M. (1995). **Qualitative analysis: What it is and how to begin.** *Research in nursing & health*, 18(4), 371-375.

SANTA CATARINA. **LINHA DE CUIDADO PARA ATENÇÃO À SAÚDE MENTAL**. 2019. Disponível em: <http://tinyurl.com/38c2r3a4>. Acesso em 29 dez 2023.

SCHENSUL, Jean J. e LECOMPTE, Margaret Diane. **Essential ethnographic methods : A Mixed Methods Approach**. Vol. 2. Rowman Altamira, 2012.

SCHRÖER, C.; KRUSE, F.; GÓMEZ, J. M. **A systematic literature review on applying CRISP-DM process model**. Procedia Computer Science, 181, 526-534. 2021. Disponível em: <https://doi.org/10.1016/j.procs.2021.01.199>. Acesso em 02 de junho de 2022.

SHORTEN, Allison; SMITH; Joanna. **"Mixed methods research: expanding the evidence base."** Evidence-based nursing 20, no. 3 (2017): 74-75.

SILVA, Evaldo de Oliveira; BAX, Marcello Peixoto. **Dicionário Semântico de Dados: abordagem de anotação de dados aplicada à geração de indicadores de desempenho**. Ciência da Informação, v. 49, n. 3, 2019.

SILVA, Evaldo de Oliveira; BAX, Marcello Peixoto; BADARÓ, A. C.; DOMINATO, R. B.; ASSIS, C. M. G. de; JENS, M. M.; FALCÃO, E. N.; SOUTO NETO, R. A.; CARVALHO, J. V. .; PAIVA, J. G. G.; NASCIMENTO, S. H. R. do. (2022). **Compreensão e preparação de dados aplicados na estratificação de risco em ansiedade**. Ciência Da Informação Express, 3, 1–10. <https://doi.org/10.60144/v3i.2022.54>.

SILVA, Evaldo de Oliveira; BAX, Marcello Peixoto. **Uso de dicionários semânticos na estratificação de riscos em saúde mental**. Encontro nacional de pesquisa e pós-graduação em ciência da informação. 2022. Acesso em 18 de dez de 2023. Disponível em: <http://hdl.handle.net/1843/59669>.

SILVA, Evaldo de Oliveira; MARQUES, Yuri Bento; BAX, Marcello Peixoto. **Processo sistemático fundamentado em modelagem ontológica aplicado à estratificação de risco em saúde mental para análise quali-quantitativa**. Workshop de dados, informação e tecnologia (WIDAT). 2022.

SNYDER, Hannah. **Literature review as a research methodology: An overview and guidelines**. Journal of business research 104 (2019): 333-339.

SPEAR, Andrew D.; CEUSTERS, Werner; SMITH, Barry. **Functions in basic formal ontology**. Applied Ontology, v. 11, n. 2, p. 103-128, 2016.

SOMEKH, Bridget; LEWIN, Cathy, eds. **Research methods in the social sciences**. Sage, 2005.

SUKAMOLSON, S. (2007). **Fundamentals of quantitative research**. Language Institute Chulalongkorn University, 1(3), 1-20.

TEORIA. In: Cambridge University, **Cambridge University Dictionary**. Cambridge: Cambridge University Press, 2023. Disponível em: <https://dictionary.cambridge.org/dictionary/english/theory>. Acesso em: 24 de jan de 2023.

TIETZ, Tabea; SACK, Harald. **Linked Data Supported Content Analysis for Sociology**. In Semantic Systems. The Power of AI and Knowledge Graphs: 15th International Conference, SEMANTiCS 2019, Karlsruhe, Germany, September 9–12, 2019, Proceedings 15, pp. 34-49. Springer International Publishing, 2019.

TWC. **Tetherless World Constellation - SDD Documentation**. Disponível em <https://tetherless-world.github.io/sdd/documentation>. Acesso em 18 de dez de 2021.

VERDINELLI, Susana, e SCAGNOLI, Norma I. "**Data display in qualitative research**." International Journal of Qualitative Methods, no. 1 (2013): 359-381.

WALKER, Diane, and MYRICK, Florence. "**Grounded theory: An exploration of process and procedure**." Qualitative health research 16, no. 4 (2006): 547-559.

WANG, Zhen et al. **Knowledge graph embedding by translating on hyperplanes**. In: Proceedings of the AAAI Conference on Artificial Intelligence. 2014.

WHITEFORD, Harvey; FERRARI, Alize; DEGENHARDT, Louisa. **Global burden of disease studies: implications for mental and substance use disorders**. Health Affairs, v. 35, n. 6, p. 1114-1120, 2016.

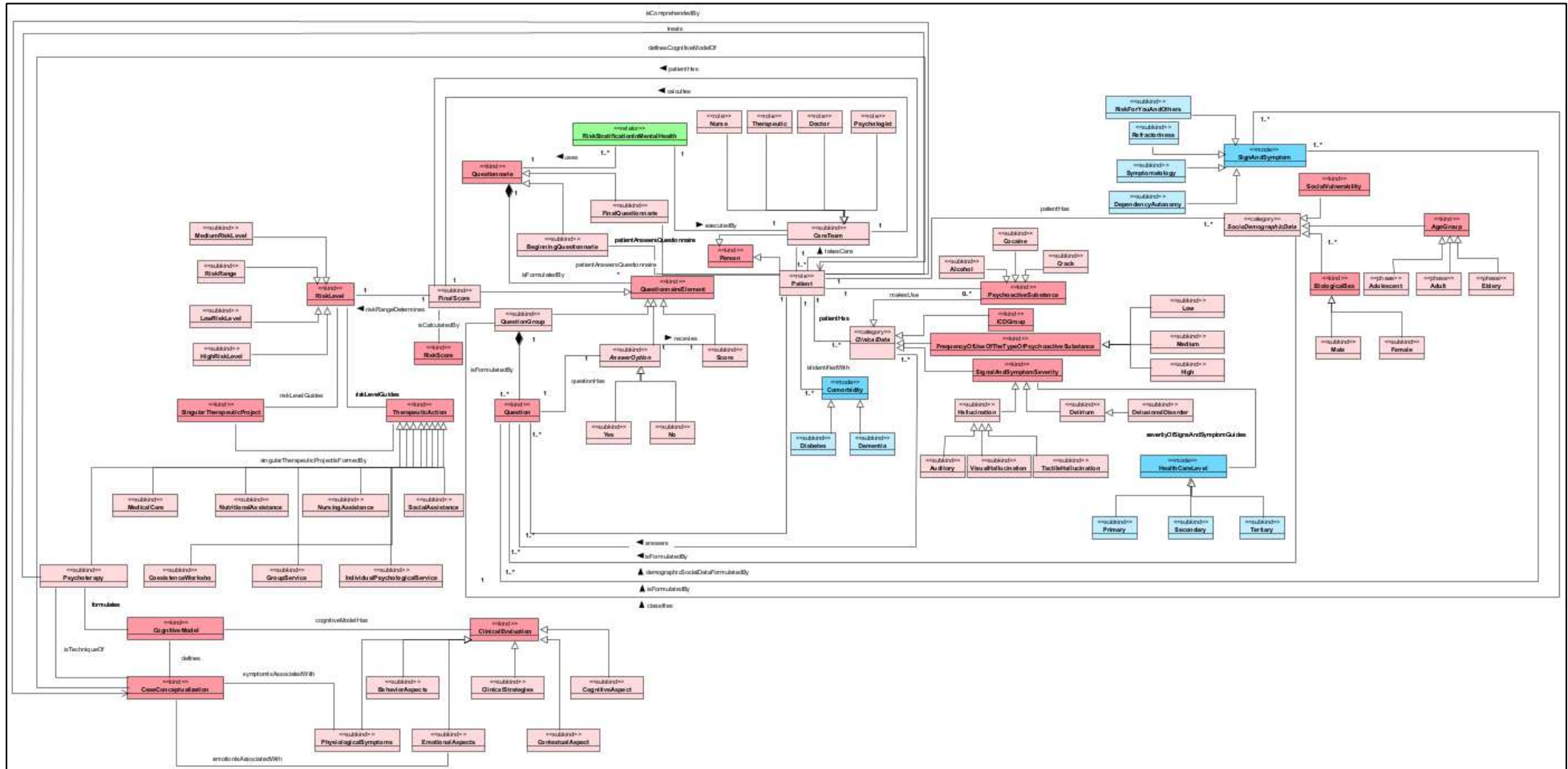
WIETEN, Sarah. **Expertise in evidence-based medicine: a tale of three models**. Philosophy, ethics, and humanities in medicine 13 (2018): 1-7.

WIERINGA, Roel. **Design science as nested problem solving**. In: Proceedings of the 4th international conference on design science research in information systems and technology. 2009. p. 1-12.

YAMADA, Diego Bettioli et al. **Proposal of an ontology for Mental Health Management in Brazil**. Procedia computer science, v. 138, p. 137-142, 2018.

YAMADA, Diego Bettioli et al. **Ontology-Based Inference for Supporting Clinical Decisions in Mental Health**. In: International Conference on Computational Science. Springer, Cham, 2020. p. 363-375.

APÊNDICE A



Obs.: É possível acessar o modelo da ontologia em melhor definição e fundamentado pela UFO²³

²³ Link: <https://github.com/evaldo/homoginise/tree/master/Ontologia>

APÊNDICE B

Questão 1. Houve reduções nos sintomas apresentados no desenvolver do processo terapêutico? Se sim, quais foram?

```

prefix prov: <http://www.w3.org/ns/prov#>
prefix xsd: <http://www.w3.org/2001/XMLSchema#>
prefix uo: <http://purl.obolibrary.org/obo/UO_>
prefix sio: <http://semanticscience.org/resource/>
prefix stato: <http://purl.obolibrary.org/obo/STATO_>
prefix example-kb: <http://example.com/kb/example#>
prefix chear: <http://hadatac.org/ont/chear#>
prefix hasco: <http://hadatac.org/ont/hasco#>
prefix obo: <http://purl.obolibrary.org/obo/>
prefix skos: <http://www.w3.org/2008/05/skos#>
prefix osm: <http://www.w3.org/2008/05/osm#>
prefix oersm: <http://hadatac.org/ont/oersm#>
prefix ontriscal: <http://hadatac.org/ont/ontriscal#>
prefix qualico: <http://hadatac.org/ont/qualico#>
prefix refiqda: <http://hadatac.org/ont/refiqda#>
prefix kpionto: <http://hadatac.org/ont/kpionto#>

```

```

SELECT distinct ?id_case ?nameproject ?questionnaire ?warmsensation ?dizziness
?legshaking ?risklevel ?valuecase ?valuenameproject ?valuewarmsensation
?valuedizziness ?valuelegshaking ?valuerisklevel
where
{
  ?objectcase rdf:type ?case .
  ?id_case sio:isAttributeOf ?objectcase .
  ?id_case sio:hasValue ?valuecase .

  ?objectcase rdf:type ?caseproject .
  ?caseproject refiqda:appliesTo ?project .

  ?objectproject rdf:type ?project .
  ?nameproject sio:isAttributeOf ?objectproject .
  ?nameproject rdf:type refiqda:NameOfProject .
  ?nameproject sio:hasValue ?valuenameproject .

  ?case ?patientanswersquestionnaire ?beginningquestionnaire .
  ?beginningquestionnaire rdfs:label "beginningquestionnaire" .

  ?warmsensation sio:hasValue ?valuewarmsensation .
  ?warmsensation sio:isAttributeOf ?questionnaire .
  ?warmsensation rdf:type ontriscal:WarmSensation .

  ?legshaking sio:hasValue ?valuelegshaking .
  ?legshaking sio:isAttributeOf ?questionnaire .
  ?legshaking rdf:type ontriscal:LegShaking .

  ?dizziness sio:hasValue ?valuedizziness .
  ?dizziness sio:isAttributeOf ?questionnaire .
  ?dizziness rdf:type ontriscal:Dizziness .

  ?risklevel sio:hasValue ?valuerisklevel .
  ?risklevel sio:isAttributeOf ?questionnaire .
  ?risklevel rdf:type ontriscal:RiskLevel .

  ?questionnaire rdfs:label "beginningquestionnaire" .
}

```

```

prefix prov: <http://www.w3.org/ns/prov#>
prefix xsd: <http://www.w3.org/2001/XMLSchema#>
prefix uo: <http://purl.obolibrary.org/obo/UO_>
prefix sio: <http://semanticscience.org/resource/>
prefix stato: <http://purl.obolibrary.org/obo/STATO_>
prefix example-kb: <http://example.com/kb/example#>
prefix chear: <http://hadatac.org/ont/chear#>
prefix hasco: <http://hadatac.org/ont/hasco#>
prefix obo: <http://purl.obolibrary.org/obo/>
prefix skos: <http://www.w3.org/2008/05/skos#>
prefix osm: <http://www.w3.org/2008/05/osm#>
prefix oersm: <http://hadatac.org/ont/oersm#>
prefix ontriscal: <http://hadatac.org/ont/ontriscal#>
prefix qualico: <http://hadatac.org/ont/qualico#>
prefix refiqda: <http://hadatac.org/ont/refiqda#>
prefix kpionto: <http://hadatac.org/ont/kpionto#>

```

```

SELECT distinct ?id_case ?nameproject ?questionnaire ?warmsensation ?dizziness
?legshaking ?risklevel ?valuecase ?valuenameproject ?valuewarmsensation
?valuedizziness ?valuelegshaking ?valuerisklevel
where
{
    ?objectcase rdf:type ?case .
    ?id_case sio:isAttributeOf ?objectcase .
    ?id_case sio:hasValue ?valuecase .

    ?objectcase rdf:type ?caseproject .
    ?caseproject refiqda:appliesTo ?project .

    ?objectproject rdf:type ?project .
    ?nameproject sio:isAttributeOf ?objectproject .
    ?nameproject rdf:type refiqda:NameOfProject .
    ?nameproject sio:hasValue ?valuenameproject .

    ?case ?patientanswersquestionnaire ?finalquestionnaire .
    ?finalquestionnaire rdfs:label "finalquestionnaire" .

    ?warmsensation sio:hasValue ?valuewarmsensation .
    ?warmsensation sio:isAttributeOf ?questionnaire .
    ?warmsensation rdf:type ontriscal:WarmSensation .

    ?legshaking sio:hasValue ?valuelegshaking .
    ?legshaking sio:isAttributeOf ?questionnaire .
    ?legshaking rdf:type ontriscal:LegShaking .

    ?dizziness sio:hasValue ?valuedizziness .
    ?dizziness sio:isAttributeOf ?questionnaire .
    ?dizziness rdf:type ontriscal:Dizziness .

    ?risklevel sio:hasValue ?valuerisklevel .
    ?risklevel sio:isAttributeOf ?questionnaire .
    ?risklevel rdf:type ontriscal:RiskLevel .

    ?questionnaire rdfs:label "finalquestionnaire" .
}

```

Questão 2. Quais foram as estratégias de intervenção de maior impacto nos pensamentos, emoções e comportamentos da(o) paciente no desenvolver do processo? Quais procedimentos se destacaram?

```

prefix prov: <http://www.w3.org/ns/prov#>

```



```

prefix xsd: <http://www.w3.org/2001/XMLSchema#>
prefix uo: <http://purl.obolibrary.org/obo/UO_>
prefix sio: <http://semanticscience.org/resource/>
prefix stato: <http://purl.obolibrary.org/obo/STATO_>
prefix example-kb: <http://example.com/kb/example#>
prefix chear: <http://hadatac.org/ont/chear#>
prefix hasco: <http://hadatac.org/ont/hasco#>
prefix obo: <http://purl.obolibrary.org/obo/>
prefix skos: <http://www.w3.org/2008/05/skos#>
prefix osm: <http://www.w3.org/2008/05/osm#>
prefix oersm: <http://hadatac.org/ont/oersm#>
prefix ontriscal: <http://hadatac.org/ont/ontriscal#>
prefix qualico: <http://hadatac.org/ont/qualico#>
prefix refiqda: <http://hadatac.org/ont/refiqda#>
prefix kpionto: <http://hadatac.org/ont/kpionto#>

SELECT distinct ?valuenamproject ?valuecase ?symptomextracted ?numberofappearance
?questionnariename ?valuefearoftheworsthappening ?valuenervous
?valuefeelingofsuffocation ?valuerisklevel
  where
  {{
    SELECT ?id_case ?valuecase ?iscomprehendedby ?caseconceptualizationextracted
?hastypeoftreatment ?symptomextracted (count( distinct ?valueesyptom) as
?numberofappearance)
      where
      {
SELECT ?id_case ?valuecase ?iscomprehendedby ?caseconceptualizationextracted
?hastypeoftreatment ?symptomextracted ?valueesyptom
where {
?id_case sio:isAttributeOf ?case .
?id_case rdf:type hasco:originalID .
?id_case sio:hasValue ?valuecase .

?case ontriscal:isComprehendedBy ?caseconceptualization .

?symptom sio:hasValue ?valueesyptom .
?symptom sio:isAttributeOf ?caseconceptualization .
?symptom ?type ?classsymptoms .

bind(strbefore(strafter(str(?symptom), "#"), "-") as ?symptomextracted) .
bind(uri(strbefore(str(?caseconceptualization), "-" )) as
?caseconceptualizationextracted) .
bind(uri("ontriscal:isComprehendedBy") as ?iscomprehendedby) .
bind(uri("ontriscal:conceptualizationHasTypeOfTreatment") as ?hastypeoftreatment) .

FILTER(?classsymptoms=ontriscal:Treatment) .
FILTER(?type=rdf:type) .

} order by ?symptom

} group by ?id_case ?valuecase ?iscomprehendedby ?caseconceptualizationextracted
?hastypeoftreatment ?symptomextracted

} union {
  ?objectcase rdf:type ?case .
  ?id_case sio:isAttributeOf ?objectcase .
  ?id_case sio:hasValue ?valuecase .

  ?objectcase rdf:type ?caseproject .
  ?caseproject refiqda:appliesTo ?project .

  ?objectproject rdf:type ?project .
  ?nameproject sio:isAttributeOf ?objectproject .
  ?nameproject rdf:type refiqda:NameOfProject .

```

```

?nameproject sio:hasValue ?valuenameproject .

?case ?patientanswersquestionnaire ?beginningquestionnaire .
?beginningquestionnaire rdfs:label "beginningquestionnaire" .

?fearoftheworsthappening sio:hasValue ?valuefearoftheworsthappening .
?fearoftheworsthappening sio:isAttributeOf ?questionnaire .
?fearoftheworsthappening rdf:type ontriscal:FearOfTheWorstHappening .

?nervous sio:hasValue ?valuenervous .
?nervous sio:isAttributeOf ?questionnaire .
?nervous rdf:type ontriscal:Nervous .

?feelingofsuffocation sio:hasValue ?valuefeelingofsuffocation .
?feelingofsuffocation sio:isAttributeOf ?questionnaire .
?feelingofsuffocation rdf:type ontriscal:FeelingOfSuffocation .

?risklevel sio:hasValue ?valuerisklevel .
?risklevel sio:isAttributeOf ?questionnaire .
?risklevel rdf:type ontriscal:RiskLevel .

?questionnaire rdfs:label "beginningquestionnaire" .

    bind(strbefore(strafter(str(?questionnaire), "#"), "-") as
?questionnariename) .

} union {

?objectcase rdf:type ?case .
?id_case sio:isAttributeOf ?objectcase .
?id_case sio:hasValue ?valuecase .

?objectcase rdf:type ?caseproject .
?caseproject refiqda:appliesTo ?project .

?objectproject rdf:type ?project .
?nameproject sio:isAttributeOf ?objectproject .
?nameproject rdf:type refiqda:NameOfProject .
?nameproject sio:hasValue ?valuenameproject .

?case ?patientanswersquestionnaire ?finalquestionnaire .
?finalquestionnaire rdfs:label "finalquestionnaire" .

?fearoftheworsthappening sio:hasValue ?valuefearoftheworsthappening .
?fearoftheworsthappening sio:isAttributeOf ?questionnaire .
?fearoftheworsthappening rdf:type ontriscal:FearOfTheWorstHappening .

?nervous sio:hasValue ?valuenervous .
?nervous sio:isAttributeOf ?questionnaire .
?nervous rdf:type ontriscal:Nervous .

?feelingofsuffocation sio:hasValue ?valuefeelingofsuffocation .
?feelingofsuffocation sio:isAttributeOf ?questionnaire .
?feelingofsuffocation rdf:type ontriscal:FeelingOfSuffocation .

?risklevel sio:hasValue ?valuerisklevel .
?risklevel sio:isAttributeOf ?questionnaire .
?risklevel rdf:type ontriscal:RiskLevel .

?questionnaire rdfs:label "finalquestionnaire" .

    bind(strbefore(strafter(str(?questionnaire), "#"), "-") as
?questionnariename) .

```

```
}}
```

3. Quais as relações entre pensamentos, emoções e crenças da(o) paciente?

```
prefix prov: <http://www.w3.org/ns/prov#>
prefix xsd: <http://www.w3.org/2001/XMLSchema#>
prefix uo: <http://purl.obolibrary.org/obo/UO_>
prefix sio: <http://semanticscience.org/resource/>
prefix stato: <http://purl.obolibrary.org/obo/STATO_>
prefix example-kb: <http://example.com/kb/example#>
prefix chear: <http://hadatac.org/ont/chear#>
prefix hasco: <http://hadatac.org/ont/hasco#>
prefix obo: <http://purl.obolibrary.org/obo/>
prefix skos: <http://www.w3.org/2008/05/skos#>
prefix osm: <http://www.w3.org/2008/05/osm#>
prefix oersm: <http://hadatac.org/ont/oersm#>
prefix ontriscal: <http://hadatac.org/ont/ontriscal#>
prefix qualico: <http://hadatac.org/ont/qualico#>
prefix refiqda: <http://hadatac.org/ont/refiqda#>
prefix kpionto: <http://hadatac.org/ont/kpionto#>
```

```
SELECT distinct ?id_case ?valuecase ?caseconceptualization ?cognitiveaspect
?valuecognitive
where {
  {
    ?id_case sio:isAttributeOf ?case .
    ?id_case rdf:type hasco:originalID .
    ?id_case sio:hasValue ?valuecase .

    ?case ontriscal:isComprehendedBy ?caseconceptualization .

    ?cognitiveaspect sio:hasValue ?valuecognitive .
    ?cognitiveaspect sio:isAttributeOf ?caseconceptualization .
    ?cognitiveaspect rdf:type ontriscal:Thinking .

    FILTER (?valuecognitive="autocobrança elevada" ||
?valuecognitive="pensamentos muito rígidos acerca dos resultados que queria, onde
se estes não saíssem como o esperado, eram considerados sem valor.") .

  } union {
    ?id_case sio:isAttributeOf ?case .
    ?id_case rdf:type hasco:originalID .
    ?id_case sio:hasValue ?valuecase .

    ?case ontriscal:isComprehendedBy ?caseconceptualization .

    ?cognitiveaspect sio:hasValue ?valuecognitive .
    ?cognitiveaspect sio:isAttributeOf ?caseconceptualization .
    ?cognitiveaspect rdf:type ontriscal:EmotionalAspects .

    FILTER (regex(?valuecognitive, "se frustra")) .

  } union {
    ?id_case sio:isAttributeOf ?case .
    ?id_case rdf:type hasco:originalID .
    ?id_case sio:hasValue ?valuecase .

    ?case ontriscal:isComprehendedBy ?caseconceptualization .
```

```

?cognitiveaspect sio:hasValue ?valuecognitive .
?cognitiveaspect sio:isAttributeOf ?caseconceptualization .
?cognitiveaspect rdf:type ontriscal:CoreBelief .

FILTER (?valuecognitive="sua crença de fracasso") .

} union {

?id_case sio:isAttributeOf ?case .
?id_case rdf:type hasco:originalID .
?id_case sio:hasValue ?valuecase .

?case ontriscal:isComprehendedBy ?caseconceptualization .

?cognitiveaspect sio:hasValue ?valuecognitive .
?cognitiveaspect sio:isAttributeOf ?caseconceptualization .
?cognitiveaspect rdf:type ontriscal:IntermediateBelief .

}

}

```

4. Quais as demandas e objetivos apresentados pela(o) paciente? De que maneira o processo os alcança?

```

prefix prov: <http://www.w3.org/ns/prov#>
prefix xsd: <http://www.w3.org/2001/XMLSchema#>
prefix uo: <http://purl.obolibrary.org/obo/UO_>
prefix sio: <http://semanticscience.org/resource/>
prefix stato: <http://purl.obolibrary.org/obo/STATO_>
prefix example-kb: <http://example.com/kb/example#>
prefix chear: <http://hadatac.org/ont/chear#>
prefix hasco: <http://hadatac.org/ont/hasco#>
prefix obo: <http://purl.obolibrary.org/obo/>
prefix skos: <http://www.w3.org/2008/05/skos#>
prefix osm: <http://www.w3.org/2008/05/osm#>
prefix oersm: <http://hadatac.org/ont/oersm#>
prefix ontriscal: <http://hadatac.org/ont/ontriscal#>
prefix qualico: <http://hadatac.org/ont/qualico#>
prefix refiqda: <http://hadatac.org/ont/refiqda#>
prefix kpionto: <http://hadatac.org/ont/kpionto#>

```

```

SELECT distinct ?id_case ?valuecase ?caseconceptualization ?therapeuticgoal
?valuetherapeuticgoal ?clinicalhypothesis ?valueclinicalhypothesis
where {

```

```

?id_case sio:isAttributeOf ?case .
?id_case rdf:type hasco:originalID .
?id_case sio:hasValue ?valuecase .

?case ontriscal:isComprehendedBy ?caseconceptualization .

?therapeuticgoal sio:hasValue ?valuetherapeuticgoal .
?therapeuticgoal sio:isAttributeOf ?caseconceptualization .
?therapeuticgoal rdf:type ontriscal:StructureOfTreatmentProcess .

?clinicalhypothesis sio:hasValue ?valueclinicalhypothesis .
?clinicalhypothesis sio:isAttributeOf ?caseconceptualization .
?clinicalhypothesis rdf:type ontriscal:ClinicalHypothesis .

FILTER (?valuetherapeuticgoal!='nan') .

}

```

5. Quais os fatores que interferiram nos resultados do processo? Foram de cunho particular ou contextual?

```

prefix prov: <http://www.w3.org/ns/prov#>
prefix xsd: <http://www.w3.org/2001/XMLSchema#>
prefix uo: <http://purl.obolibrary.org/obo/UO_>
prefix sio: <http://semanticscience.org/resource/>
prefix stato: <http://purl.obolibrary.org/obo/STATO_>
prefix example-kb: <http://example.com/kb/example#>
prefix chear: <http://hadatac.org/ont/chear#>
prefix hasco: <http://hadatac.org/ont/hasco#>
prefix obo: <http://purl.obolibrary.org/obo/>
prefix skos: <http://www.w3.org/2008/05/skos#>
prefix osm: <http://www.w3.org/2008/05/osm#>
prefix oersm: <http://hadatac.org/ont/oersm#>
prefix ontriscal: <http://hadatac.org/ont/ontriscal#>
prefix qualico: <http://hadatac.org/ont/qualico#>
prefix refiqda: <http://hadatac.org/ont/refiqda#>
prefix kpionto: <http://hadatac.org/ont/kpionto#>

```

```

SELECT distinct ?id_case ?valuecase ?caseconceptualization ?treatment
?valuetreatment
where {

    ?id_case sio:isAttributeOf ?case .
    ?id_case rdf:type hasco:originalID .
    ?id_case sio:hasValue ?valuecase .

    ?case ontriscal:isComprehendedBy ?caseconceptualization .

    ?treatment sio:hasValue ?valuetreatment .
    ?treatment sio:isAttributeOf ?caseconceptualization .
    ?treatment rdf:type ontriscal:Treatment .

    FILTER (?valuetreatment!='nan') .

}

```

6. Quais valores influenciam no engajamento do paciente para o tratamento? Quais as funções atribuídas a eles?

```

prefix prov: <http://www.w3.org/ns/prov#>
prefix xsd: <http://www.w3.org/2001/XMLSchema#>
prefix uo: <http://purl.obolibrary.org/obo/UO_>
prefix sio: <http://semanticscience.org/resource/>
prefix stato: <http://purl.obolibrary.org/obo/STATO_>
prefix example-kb: <http://example.com/kb/example#>
prefix chear: <http://hadatac.org/ont/chear#>
prefix hasco: <http://hadatac.org/ont/hasco#>
prefix obo: <http://purl.obolibrary.org/obo/>
prefix skos: <http://www.w3.org/2008/05/skos#>
prefix osm: <http://www.w3.org/2008/05/osm#>
prefix oersm: <http://hadatac.org/ont/oersm#>
prefix ontriscal: <http://hadatac.org/ont/ontriscal#>
prefix qualico: <http://hadatac.org/ont/qualico#>
prefix refiqda: <http://hadatac.org/ont/refiqda#>
prefix kpionto: <http://hadatac.org/ont/kpionto#>

```

```

SELECT distinct ?id_case ?valuecase ?caseconceptualization ?cognitiveaspect
?valuecognitive
where {

```

```

{
  ?id_case sio:isAttributeOf ?case .
  ?id_case rdf:type hasco:originalID .
  ?id_case sio:hasValue ?valuecase .

  ?case ontriscal:isComprehendedBy ?caseconceptualization .

  ?cognitiveaspect sio:hasValue ?valuecognitive .
  ?cognitiveaspect sio:isAttributeOf ?caseconceptualization .
  ?cognitiveaspect rdf:type ontriscal:Perfectionism .

} union {

  ?id_case sio:isAttributeOf ?case .
  ?id_case rdf:type hasco:originalID .
  ?id_case sio:hasValue ?valuecase .

  ?case ontriscal:isComprehendedBy ?caseconceptualization .

  ?cognitiveaspect sio:hasValue ?valuecognitive .
  ?cognitiveaspect sio:isAttributeOf ?caseconceptualization .
  ?cognitiveaspect rdf:type ontriscal:PatientHistory .

  FILTER(?valuecognitive!='nan')

} union {

  ?id_case sio:isAttributeOf ?case .
  ?id_case rdf:type hasco:originalID .
  ?id_case sio:hasValue ?valuecase .

  ?case ontriscal:isComprehendedBy ?caseconceptualization .

  ?cognitiveaspect sio:hasValue ?valuecognitive .
  ?cognitiveaspect sio:isAttributeOf ?caseconceptualization .
  ?cognitiveaspect rdf:type ontriscal:SelfAwareness .

  FILTER(?valuecognitive!='nan')

}
}

```

7. Quais os sinais, positivos e negativos, de engajamento da(o) paciente diante do processo?

```

prefix prov: <http://www.w3.org/ns/prov#>
prefix xsd: <http://www.w3.org/2001/XMLSchema#>
prefix uo: <http://purl.obolibrary.org/obo/UO_>
prefix sio: <http://semanticscience.org/resource/>
prefix stato: <http://purl.obolibrary.org/obo/STATO_>
prefix example-kb: <http://example.com/kb/example#>
prefix chear: <http://hadatac.org/ont/chear#>
prefix hasco: <http://hadatac.org/ont/hasco#>
prefix obo: <http://purl.obolibrary.org/obo/>
prefix skos: <http://www.w3.org/2008/05/skos#>
prefix osm: <http://www.w3.org/2008/05/osm#>
prefix oersm: <http://hadatac.org/ont/oersm#>
prefix ontriscal: <http://hadatac.org/ont/ontriscal#>
prefix qualico: <http://hadatac.org/ont/qualico#>
prefix refiqda: <http://hadatac.org/ont/refiqda#>

```

```

prefix kpionto: <http://hadatac.org/ont/kpionto#>

SELECT distinct ?case ?id_case ?valuecase ?caseconceptualization ?cognitiveaspect
?valuecognitiveaspect ?thinking ?valuethinking ?process ?valueprocess
where {

    ?id_case sio:isAttributeOf ?case .
    ?id_case rdf:type hasco:originalID .
    ?id_case sio:hasValue ?valuecase .

    ?case ontriscal:isComprehendedBy ?caseconceptualization .

    ?cognitiveaspect sio:hasValue ?valuecognitiveaspect .
    ?cognitiveaspect sio:isAttributeOf ?caseconceptualization .
    ?cognitiveaspect rdf:type ontriscal:PositiveSignsOfEngagementProcess .

    FILTER(?valuecognitiveaspect!='nan') .

}

```

8. Qual a prevalência dos sintomas que influenciam na hipótese clínica?

```

prefix prov: <http://www.w3.org/ns/prov#>
prefix xsd: <http://www.w3.org/2001/XMLSchema#>
prefix uo: <http://purl.obolibrary.org/obo/UO_>
prefix sio: <http://semanticscience.org/resource/>
prefix stato: <http://purl.obolibrary.org/obo/STATO_>
prefix example-kb: <http://example.com/kb/example#>
prefix chear: <http://hadatac.org/ont/chear#>
prefix hasco: <http://hadatac.org/ont/hasco#>
prefix obo: <http://purl.obolibrary.org/obo/>
prefix skos: <http://www.w3.org/2008/05/skos#>
prefix osm: <http://www.w3.org/2008/05/osm#>
prefix oersm: <http://hadatac.org/ont/oersm#>
prefix ontriscal: <http://hadatac.org/ont/ontriscal#>
prefix qualico: <http://hadatac.org/ont/qualico#>
prefix refiqda: <http://hadatac.org/ont/refiqda#>
prefix kpionto: <http://hadatac.org/ont/kpionto#>

select ?id_case ?valuecase ?classsymptoms ?symptomextracted ?numberofappearance
?total (round((?numberofappearance/?total*100)) as ?percentofnumberofappearance)
where {
SELECT ?id_case ?valuecase ?classsymptoms ?symptomextracted (count(distinct
?valueesympptom) as ?numberofappearance) ?total
  where {
    SELECT ?id_case ?valuecase ?classsymptoms ?symptom ?symptomextracted
?valueesympptom ?total
      where {

        ?id_case sio:isAttributeOf ?case .
        ?id_case rdf:type hasco:originalID .
        ?id_case sio:hasValue ?valuecase .

        ?case ontriscal:isComprehendedBy ?caseconceptualization .

        ?symptom sio:hasValue ?valueesympptom .
        ?symptom sio:isAttributeOf ?caseconceptualization .
        ?symptom rdf:type ?classsymptoms .

        bind(strbefore(strafter(str(?symptom), "#"), "-") as
?symptomextracted) .

        {select (count(distinct ?valueesympptom)-1 as ?total)
          where {
            ?id_case sio:isAttributeOf ?case .
            ?id_case rdf:type hasco:originalID .

```

```

?id_case sio:hasValue ?valuecase .

?case ontriscal:isComprehendedBy ?caseconceptualization .

?symptom sio:hasValue ?valueesyptom .
?symptom sio:isAttributeOf ?caseconceptualization .
?symptom rdf:type ontriscal:PhysiologicalSymptoms .

}} .

FILTER(?symptomextracted != 'PhysiologicalSymptoms') .
FILTER(?classsymptoms = ontriscal:PhysiologicalSymptoms)

} order by ?symptom

} group by ?id_case ?valuecase ?classsymptoms ?symptomextracted ?total
}

```

9. Qual a prevalência dos códigos que influenciam na validação da crença central?

```

-----
--BehaviorAspects
-----
-----

```

```

prefix prov: <http://www.w3.org/ns/prov#>
prefix xsd: <http://www.w3.org/2001/XMLSchema#>
prefix uo: <http://purl.obolibrary.org/obo/UO_>
prefix sio: <http://semanticscience.org/resource/>
prefix stato: <http://purl.obolibrary.org/obo/STATO_>
prefix example-kb: <http://example.com/kb/example#>
prefix chear: <http://hadatac.org/ont/chear#>
prefix hasco: <http://hadatac.org/ont/hasco#>
prefix obo: <http://purl.obolibrary.org/obo/>
prefix skos: <http://www.w3.org/2008/05/skos#>
prefix osm: <http://www.w3.org/2008/05/osm#>
prefix oersm: <http://hadatac.org/ont/oersm#>
prefix ontriscal: <http://hadatac.org/ont/ontriscal#>
prefix qualico: <http://hadatac.org/ont/qualico#>
prefix refiqda: <http://hadatac.org/ont/refiqda#>
prefix kpionto: <http://hadatac.org/ont/kpionto#>

select ?id_case ?valuecase ?classsymptoms ?symptomextracted ?numberofappearance
?total (round((?numberofappearance/?total*100)) as ?percentofnumberofappearance)
where {
SELECT ?id_case ?valuecase ?classsymptoms ?symptomextracted (count( ?valueesyptom)
as ?numberofappearance) ?total
  where {

      SELECT ?id_case ?valuecase ?classsymptoms ?symptomextracted ?valueesyptom
?total
  where {

      ?id_case sio:isAttributeOf ?case .
      ?id_case rdf:type hasco:originalID .
      ?id_case sio:hasValue ?valuecase .

      ?case ontriscal:isComprehendedBy ?caseconceptualization .

      ?symptom sio:hasValue ?valueesyptom .
      ?symptom sio:isAttributeOf ?caseconceptualization .
      ?symptom rdf:type ?classsymptoms .

```



```

        bind(strbefore(strafter(str(?symptom), "#"), "-") as
?symptomextracted) .

        (select (count( ?valueesympptom) as ?total)
            where {
                ?id_case sio:isAttributeOf ?case .
                ?id_case rdf:type hasco:originalID .
                ?id_case sio:hasValue ?valuecase .

                ?case ontriscal:isComprehendedBy
?caseconceptualization .

                ?symptom sio:hasValue ?valueesympptom .
                ?symptom sio:isAttributeOf ?caseconceptualization .
                ?symptom rdf:type ontriscal:BehaviorAspects .

            }

        } .

        FILTER(?classssyptoms =   ontriscal:BehaviorAspects) .

    } order by ?symptom

} group by ?id_case ?valuecase ?classssyptoms ?symptomextracted ?total
}

-----
--EmotionalAspect
-----

prefix prov: <http://www.w3.org/ns/prov#>
prefix xsd: <http://www.w3.org/2001/XMLSchema#>
prefix uo: <http://purl.obolibrary.org/obo/UO_>
prefix sio: <http://semanticscience.org/resource/>
prefix stato: <http://purl.obolibrary.org/obo/STATO_>
prefix example-kb: <http://example.com/kb/example#>
prefix chear: <http://hadatac.org/ont/chear#>
prefix hasco: <http://hadatac.org/ont/hasco#>
prefix obo: <http://purl.obolibrary.org/obo/>
prefix skos: <http://www.w3.org/2008/05/skos#>
prefix osm: <http://www.w3.org/2008/05/osm#>
prefix oersm: <http://hadatac.org/ont/oersm#>
prefix ontriscal: <http://hadatac.org/ont/ontriscal#>
prefix qualico: <http://hadatac.org/ont/qualico#>
prefix refiqda: <http://hadatac.org/ont/refiqda#>
prefix kpionto: <http://hadatac.org/ont/kpionto#>

select ?id_case ?valuecase ?classssyptoms ?symptomextracted  ?numberofappearance
?total (round((?numberofappearance/?total*100)) as ?percentofnumberofappearance)
where {
SELECT ?id_case ?valuecase ?classssyptoms ?symptomextracted (count( ?valueesympptom)
as ?numberofappearance) ?total
    where {

        SELECT  ?id_case ?valuecase ?classssyptoms ?symptomextracted ?valueesympptom
?total
        where {

                ?id_case sio:isAttributeOf ?case .
                ?id_case rdf:type hasco:originalID .
                ?id_case sio:hasValue ?valuecase .

```

```

?case ontriscal:isComprehendedBy ?caseconceptualization .

?symptom sio:hasValue ?valueesympptom .
?symptom sio:isAttributeOf ?caseconceptualization .
?symptom rdf:type ?classsymptoms .

bind(strbefore(strafter(str(?symptom), "#"), "-") as
?symptomextracted) .

(select (count( ?valueesympptom) as ?total)
  where {
    ?id_case sio:isAttributeOf ?case .
    ?id_case rdf:type hasco:originalID .
    ?id_case sio:hasValue ?valueecase .

    ?case ontriscal:isComprehendedBy ?caseconceptualization
.

    ?symptom sio:hasValue ?valueesympptom .
    ?symptom sio:isAttributeOf ?caseconceptualization .
    ?symptom rdf:type ontriscal:EmotionalAspect .

  }

} .

FILTER(?classsymptoms = ontriscal:EmotionalAspect) .

} order by ?symptom

} group by ?id_case ?valueecase ?classsymptoms ?symptomextracted ?total
}

```

```
-----
--Thinking
-----
```

```

prefix prov: <http://www.w3.org/ns/prov#>
prefix xsd: <http://www.w3.org/2001/XMLSchema#>
prefix uo: <http://purl.obolibrary.org/obo/UO_>
prefix sio: <http://semanticscience.org/resource/>
prefix stato: <http://purl.obolibrary.org/obo/STATO_>
prefix example-kb: <http://example.com/kb/example#>
prefix chear: <http://hadatac.org/ont/chear#>
prefix hasco: <http://hadatac.org/ont/hasco#>
prefix obo: <http://purl.obolibrary.org/obo/>
prefix skos: <http://www.w3.org/2008/05/skos#>
prefix osm: <http://www.w3.org/2008/05/osm#>
prefix oersm: <http://hadatac.org/ont/oersm#>
prefix ontriscal: <http://hadatac.org/ont/ontriscal#>
prefix qualico: <http://hadatac.org/ont/qualico#>
prefix refiqda: <http://hadatac.org/ont/refiqda#>
prefix kpionto: <http://hadatac.org/ont/kpionto#>

select ?id_case ?valueecase ?classsymptoms ?symptomextracted ?numberofappearance
?total (round((?numberofappearance/?total*100)) as ?percentofnumberofappearance)
where {
SELECT ?id_case ?valueecase ?classsymptoms ?symptomextracted (count( ?valueesympptom)
as ?numberofappearance) ?total
  where {

    SELECT ?id_case ?valueecase ?classsymptoms ?symptomextracted ?valueesympptom
?total
  where {

```

```

?id_case sio:isAttributeOf ?case .
?id_case rdf:type hasco:originalID .
?id_case sio:hasValue ?valuecase .

?case ontriscal:isComprehendedBy ?caseconceptualization .

?symptom sio:hasValue ?valueesympptom .
?symptom sio:isAttributeOf ?caseconceptualization .
?symptom rdf:type ?classsymptoms .

bind(strbefore(strafter(str(?symptom), "#"), "-") as
?symptomextracted) .

(select (count( ?valueesympptom) as ?total)
where {
?id_case sio:isAttributeOf ?case .
?id_case rdf:type hasco:originalID .
?id_case sio:hasValue ?valuecase .

?case ontriscal:isComprehendedBy ?caseconceptualization
.

?symptom sio:hasValue ?valueesympptom .
?symptom sio:isAttributeOf ?caseconceptualization .
?symptom rdf:type ontriscal:Thinking .

}

} .

FILTER(?classsymptoms = ontriscal:Thinking) .

} order by ?symptom

} group by ?id_case ?valuecase ?classsymptoms ?symptomextracted ?total
}

```

10. Quais intervenções realizadas no tratamento permitiram a redução do nível e valor do risco em ansiedade?

```

prefix prov: <http://www.w3.org/ns/prov#>
prefix xsd: <http://www.w3.org/2001/XMLSchema#>
prefix uo: <http://purl.obolibrary.org/obo/UO_>
prefix sio: <http://semanticscience.org/resource/>
prefix stato: <http://purl.obolibrary.org/obo/STATO_>
prefix example-kb: <http://example.com/kb/example#>
prefix chear: <http://hadatac.org/ont/chear#>
prefix hasco: <http://hadatac.org/ont/hasco#>
prefix obo: <http://purl.obolibrary.org/obo/>
prefix skos: <http://www.w3.org/2008/05/skos#>
prefix osm: <http://www.w3.org/2008/05/osm#>
prefix oersm: <http://hadatac.org/ont/oersm#>
prefix ontriscal: <http://hadatac.org/ont/ontriscal#>
prefix qualico: <http://hadatac.org/ont/qualico#>
prefix refiqda: <http://hadatac.org/ont/refiqda#>
prefix kpionto: <http://hadatac.org/ont/kpionto#>

select ?id_case ?valuecase ?classsymptoms ?symptomextracted ?numberofappearance
?total (round((?numberofappearance/?total*100)) as ?percentofnumberofappearance)
where {
SELECT ?id_case ?valuecase ?classsymptoms ?symptomextracted (count( ?valueesympptom)
as ?numberofappearance) ?total
where {

```

```

SELECT ?id_case ?valuecase ?classsymptoms ?symptomextracted ?valueesympom
?total
  where {

    ?id_case sio:isAttributeOf ?case .
    ?id_case rdf:type hasco:originalID .
    ?id_case sio:hasValue ?valuecase .

    ?case ontriscal:isComprehendedBy ?caseconceptualization .

    ?symptom sio:hasValue ?valueesympom .
    ?symptom sio:isAttributeOf ?caseconceptualization .
    ?symptom rdf:type ?classsymptoms .

    bind(strbefore(strafter(str(?symptom), "#"), "-") as
?symptomextracted) .

    {select (count( ?valueesympom) as ?total)
      where {
        ?id_case sio:isAttributeOf ?case .
          ?id_case rdf:type hasco:originalID .
        ?id_case sio:hasValue ?valuecase .

        ?case ontriscal:isComprehendedBy ?caseconceptualization
.

        ?symptom sio:hasValue ?valueesympom .
        ?symptom sio:isAttributeOf ?caseconceptualization .
        ?symptom rdf:type ontriscal:Treatment .

      }
    } .

    FILTER(?classsymptoms=ontriscal:Treatment ) .

  } order by ?symptom
} group by ?id_case ?valuecase ?classsymptoms ?symptomextracted ?total
}

```

ANEXO 1



**INSTRUMENTO NORMATIVO DE COLABORAÇÃO ENTRE O
PPGGOC-UFGM E O HOSPITAL VILA VERDE DE SAÚDE MENTAL
QUE REGULAMENTA ATIVIDADES DE PESQUISAS
ACADÊMICAS CONJUNTAS.**

O Programa de Pós-Graduação em Gestão & Organização do Conhecimento (PPGGOC-UFGM), situado à Avenida Antônio Carlos, 6.627, na Escola de Ciência da Informação – Sala 3004, Belo Horizonte, Minas Gerais, neste ato representado pelo Prof. Marcello Peixoto Bax, CI M3.929.284, e o Hospital Vila Verde de Saúde Mental, situado à Alameda Cruzeiro Sto. Antônio, 555 - São Pedro, Juiz de Fora, Minas Gerais, neste ato representado por seu Coordenador do Programa de Residência Médica, Clovis Melo, assinam este **instrumento normativo de colaboração**, sob as seguintes cláusulas:

CLÁUSULA I – DO OBJETO

O objetivo desta colaboração é promover, entre as instituições nomeadas acima, o intercâmbio de informações com vistas à realização de pesquisas acadêmicas. As pesquisas se darão no âmbito de um Projeto Acadêmico envolvendo a Gestão de Dados do Hospital Vila Verde e serão apoiadas tecnicamente pela Plataforma HADatAc - *Human-Aware Data Acquisition framework*. A colaboração tem por objeto utilizar a Plataforma no apoio à condução de estudos de casos exploratórios que busquem analisar aspectos característicos dos serviços do Hospital Vila Verde. O HADatAc é um repositório (plataforma de software) utilizado para integração e curadoria de dados. Por meio da Plataforma os pesquisadores fazem a aquisição e a preparação de seus dados para análise posterior. A Plataforma é mantida por pesquisadores do *Tetherless World Constellation (TWC)* do *Rensselaer Polytechnic Institute (RPI)*, situado em Troy, NY, EUA. No Brasil, o HADatAc tem apoio do PPGGOC-UFGM, sob liderança do Prof. Marcello Peixoto Bax.

CLÁUSULA II – DO PERÍODO DE EXECUÇÃO E ATIVIDADES

No Projeto estão previstas atividades de pesquisa acadêmica no período entre Agosto de 2020 e Agosto de 2021. As atividades serão realizadas por meio de contato direto entre os alunos do PPGGOC-UFGM e profissionais do Hospital Vila Verde, responsáveis pelo apoio à análise dos dados e validação dos resultados. Regras de anonimização, que garantam a não identificação dos indivíduos, deverão ser aplicadas previamente aos dados. A anonimização dos dados será validada pelos profissionais do Hospital Vila Verde envolvidos na pesquisa. Os dados utilizados nas pesquisas serão mantidos somente durante a vigência deste Instrumento, e o seu descarte deverá ser acompanhada pelos responsáveis por esta colaboração.

CLÁUSULA III – DAS ATIVIDADES EXCEDENTES



Outras atividades de interesse das partes convenientes, além do objeto estabelecido na Cláusula I, poderão ser consideradas mediante conhecimento prévio da liderança do projeto no PPGGOC-UFMG e da Diretoria do Hospital Vila Verde, adicionando cláusulas a este documento.

CLÁUSULA IV – DAS DESPESAS

As atividades de pesquisa realizadas no âmbito desta colaboração, objeto da Cláusula I, não prevêm quaisquer despesas de quaisquer naturezas para as instituições em colaboração. O PPGGOC oferecerá, a título de colaboração, a infraestrutura e instrumentos de apoio necessário para o desenvolvimento das pesquisas. Além disso, estudo preliminar mostrou que o Hospital Vila Verde detém instrumentos tecnológicos relevantes para a colaboração, podendo por exemplo, caso aprovado pelos responsáveis, hospedar a Plataforma HADatAc em sua infraestrutura tecnológica.

CLÁUSULA V – DAS ALTERAÇÕES EVENTUAIS

O PPGGOC e o Hospital Vila Verde podem modificar o presente Instrumento, a qualquer tempo e por entendimento mútuo, adicionando cláusulas ou anexos que venham a propiciar melhorias para esta colaboração.

CLÁUSULA VI – DA RESCISÃO

Este instrumento de colaboração poderá ser anulado por qualquer das partes que o assinam, sem investigação judicial, o que pode ocorrer inesperadamente devido à impossibilidade, acordada entre as partes, de concluir o projeto, porque se tornou impraticável ou foi cancelado por acordo das partes, a qualquer momento.

CLÁUSULA VII – DA UTILIZAÇÃO DA MARCA EM DIVULGAÇÕES ACADÊMICAS

Esta colaboração prevê o uso do nome, logotipo e marca das instituições parceiras exclusivamente para a divulgação das atividades do projeto e seus resultados, conforme descrito na Cláusula I. Os meios de divulgação das atividades poderão ser: páginas web; reportagens de televisão; revistas e jornais impressos; revistas eletrônicas; artigos científicos; relatórios técnicos; apresentações em conferências; redes sociais; conferências e seminários.

CLÁUSULA VIII – DA UTILIZAÇÃO DOS DADOS DO HOSPITAL VILA VERDE

Esta colaboração prevê a permissão para o uso, após anonimização, dos dados constantes do Banco de Dados do Hospital Vila Verde, EXCLUSIVAMENTE pelos membros autorizados das equipes no âmbito desta colaboração, com fins exclusivos previstos na Cláusula I, durante a vigência prevista na Cláusula II, e para publicação nos meios de comunicação mencionados na Cláusula VII.

CLÁUSULA IX – DA VIGÊNCIA



A Colaboração entrará em vigor na data de sua assinatura e terá por vigência o prazo estabelecido na Cláusula II.

As partes em colaboração assinam este documento em 3 (três) cópias de mesmo conteúdo e forma, juntamente com as testemunhas abaixo.

Juiz de Fora, 03 de Agosto de 2020.

Clovis Melo
Coordenador do Programa de
Residência Médica

Marçello Peixoto Bax
Professor do PPGOC

Testemunhas:

ANEXO 2



Certificado

O Centro Universitário Academia, através do Centro de Extensão e Pesquisa, certifica que **Evaldo de Oliveira da Silva** coordenou o Projeto de Iniciação Científica intitulado **“Processo sistemático fundamentado em modelagem ontológica aplicado à estratificação de risco em saúde mental”**, ocorrido de maio a novembro de 2022.

Juiz de Fora, 07 de fevereiro de 2023.

Marco Antônio Pereira Araújo
Coordenador de Pesquisa e Extensão

Giovânio Aguiar
Reitor

ANEXO 3

CENTRO UNIVERSITÁRIO
ACADEMIA - UNIACADEMIA



Continuação do Parecer: 5.771.735

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1952217.pdf	15/11/2022 23:04:34		Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	Termo_de_consentimento_livre_e_esclarecido_adulto_atendimento_parecer_5747624.docx	15/11/2022 23:04:09	IVALDO DE OLIVEIRA DA SILVA	Aceito
Outros	MetodologiaProposta_Atendimento_Parecer.pdf	10/10/2022 11:40:33	IVALDO DE OLIVEIRA DA SILVA	Aceito
Outros	TCUD.pdf	10/10/2022 11:34:38	IVALDO DE OLIVEIRA DA SILVA	Aceito
Projeto Detalhado / Brochura Investigador	Projeto_de_Pesquisa_Detalhado.pdf	19/08/2022 10:23:07	IVALDO DE OLIVEIRA DA SILVA	Aceito
Folha de Rosto	folhaderosto_assinada.pdf	19/08/2022 10:19:02	IVALDO DE OLIVEIRA DA SILVA	Aceito
Outros	Declaracao_Concordancia_ClinicaPsicologia.pdf	19/08/2022 09:51:49	IVALDO DE OLIVEIRA DA SILVA	Aceito
Declaração de concordância	Declaracao_Concordancia_CentroDeEnsino_e_Extensao.pdf	19/08/2022 09:49:25	IVALDO DE OLIVEIRA DA SILVA	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	Termo_de_consentimento_livre_e_esclarecido_adulto.docx	16/08/2022 00:19:52	IVALDO DE OLIVEIRA DA SILVA	Aceito
Outros	Desenho_qualitativo_do_caso_clinico.docx	16/08/2022 00:18:30	IVALDO DE OLIVEIRA DA SILVA	Aceito
Outros	Plano_de_intervencao_16_sesoes.docx	16/08/2022 00:17:41	IVALDO DE OLIVEIRA DA SILVA	Aceito
Outros	Instrumento_de_Estratificacao_de_Riscos_Ansiedade.docx	16/08/2022 00:16:57	IVALDO DE OLIVEIRA DA SILVA	Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

JUIZ DE FORA, 23 de Novembro de 2022

Assinado por:

THIAGO LUIZ BERZOINI MACHADO
(Coordenador(a))

Endereço: HALFELD 1179 . PRÉDIO DO UNIACADEMIA, SALA 215
Bairro: CENTRO CEP: 36.016-000
UF: MG Município: JUIZ DE FORA
Telefone: (32)3250-3800 Fax: (32)3250-3842 E-mail: comiteetica@uniacademia.edu.br

ANEXO 4



DECLARAÇÃO

O Centro Universitário Academia, através do Centro de Extensão e Pesquisa, declara que o Professor Evaldo de Oliveira da Silva está Coordenando atualmente o Projeto de Iniciação Científica “Processo sistemático fundamentado em modelagem ontológica para representar o conhecimento em análise quali-quantitativa”, iniciado em abril de 2023 com previsão de finalização em dezembro de 2023 e carga horária semanal de 12 horas durante o período de ocorrência.

Marco Antônio Pereira Araújo
Coordenador de Extensão e Pesquisa do UniAcademia

**Endereços**

Unidade Academia Rua Halfeld, 1179 - Centro
Unidade Arnaldo Janssen Av. Luz Interior, 345 - Estrela Sul
Seminário Santo Antônio Av. Rio Branco, 4516 - Passos

Contatos

☎ 3250-3800
www.uniacademia.edu.br
f @ @uniacademiajf