

UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE CIÊNCIA DA INFORMAÇÃO

MARCEL FERRANTE SILVA

**PROPOSTA DE MODELO DE COLABORAÇÃO
PARA CATÁLOGO *WEB* FACETADO**

Belo Horizonte

2013

MARCEL FERRANTE SILVA

**PROPOSTA DE MODELO DE COLABORAÇÃO
PARA CATÁLOGO *WEB* FACETADO**

Tese apresentada ao Programa de Pós-Graduação em Ciência da Informação da Escola de Ciência da Informação da Universidade Federal de Minas Gerais para obtenção do grau de Doutor em Ciência da Informação.

Linha de pesquisa: Organização e Uso da Informação

Orientadora: Prof^ª. Dra. Gercina Ângela Borém de Oliveira Lima

BELO HORIZONTE

2013

Silva, Marcel Ferrante.

S586p

Proposta de modelo de colaboração para catálogo *web* facetado [manuscrito] / Marcel Ferrante Silva. – 2013.
269 f.: il., enc.

Orientadora: Gercina Ângela Borém de Oliveira Lima.
Tese (doutorado) – Universidade Federal de Minas Gerais,
Escola de Ciência da Informação.

Referências: f. 193-204

Apêndices: f. 205-269

1. Ciência da informação – Teses. 2. Sistemas de recuperação da informação – Teses. 3. Ferramentas de busca na web – Teses. 4. Classificação facetada – Teses. 5. Folksonomia – Teses. I. Título. II. Lima, Gercina Ângela Borém de Oliveira. III. Universidade Federal de Minas Gerais, Escola de Ciência da Informação.

CDU: 025.49:004.55



UFMG

Universidade Federal de Minas Gerais
Escola de Ciência da Informação
Programa de Pós-Graduação em Ciência da Informação

FOLHA DE APROVAÇÃO

"PROPOSTA DE MODELO DE COLABORAÇÃO PARA CATÁLOGO WEB FACETADO"

Marcel Ferrante Silva

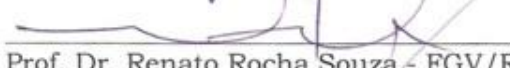
Tese submetida à Banca Examinadora, designada pelo Colegiado do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Minas Gerais, como parte dos requisitos à obtenção do título de "**doutor em Ciência da Informação**", linha de pesquisa "**Organização e Uso da Informação**".

Tese aprovada em: 16 de outubro de 2013.

Por:



Prof. Dra. Renata Maria Abrantes Baracho Porto - ECI/UFMG (presidente)



Prof. Dr. Renato Rocha Souza - FGV/RJ



Prof. Dr. Eduardo José Wense Dias - Prof. Aposentado - ECI/UFMG

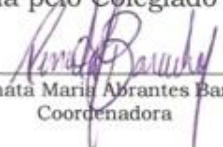


Profa. Dra. Maria Aparecida Moura - ECI/UFMG



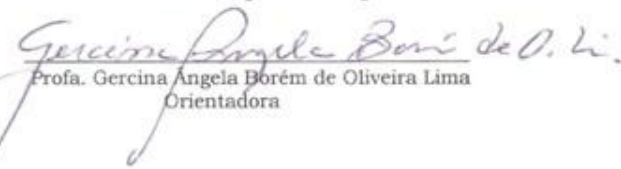
Prof. Dr. Mauricio Barcellos Almeida - ECI/UFMG

Aprovada pelo Colegiado do PPGCI



Prof. Renata Maria Abrantes Baracho Porto
Coordenadora

Versão final Aprovada por



Profa. Gercina Angela Borém de Oliveira Lima
Orientadora



UFMG

Universidade Federal de Minas Gerais
Escola de Ciência da Informação
Programa de Pós-Graduação em Ciência da Informação

ATA DA DEFESA DE TESE DE **MARCEL FERRANTE SILVA**, matrícula: 2008651740

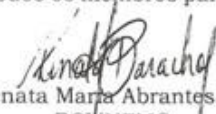
Às 14:00 horas do dia 16 de outubro de 2013, reuniu-se na Escola de Ciência da Informação da UFMG a Comissão Examinadora aprovada pelo Colegiado do Programa de Pós-Graduação em Ciência da Informação em 11/10/2013, para julgar, em exame final, o trabalho intitulado **Proposta de modelo de colaboração para catálogo web facetado**, requisito final para obtenção do Grau de DOUTOR em CIÊNCIA DA INFORMAÇÃO, área de concentração: Produção, Organização e Utilização da Informação, Linha de Pesquisa: Organização e Uso da Informação. Abrindo a sessão, a Presidente da Comissão, Profa. Dra. Renata Maria Abrantes Baracho Porto, após dar conhecimento aos presentes do teor das Normas Regulamentares do Trabalho Final, passou a palavra ao candidato para apresentação de seu trabalho. Seguiu-se a arguição pelos examinadores com a respectiva defesa do candidato. Logo após, a Comissão se reuniu sem a presença do candidato e do público, para julgamento e expedição do resultado final. Foram atribuídas as seguintes indicações:

Profa. Dra. Renata Maria Abrantes Baracho Porto - presidente	APROVADO
Prof. Dr. Renato Rocha Souza	APROVADO
Prof. Dr. Eduardo José Wense Dias	APROVADO
Profa. Dra. Maria Aparecida Moura	APROVADO
Prof. Dr. Mauricio Barcellos Almeida	APROVADO

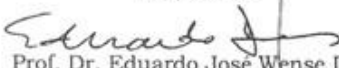
Pelas indicações, o candidato foi considerado APROVADO.

O resultado final foi comunicado publicamente ao candidato pela Presidente da Comissão. Nada mais havendo a tratar, a Presidente encerrou a sessão, da qual foi lavrada a presente ATA que será assinada por todos os membros participantes da Comissão Examinadora.

Belo Horizonte, 16 de outubro de 2013


Profa. Dra. Renata Maria Abrantes Baracho Porto
ECI/UFMG

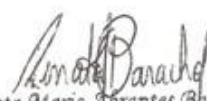

Prof. Dr. Renato Rocha Souza
RGV/RJ


Prof. Dr. Eduardo José Wense Dias
Prof. Aposentado - ECI/UFMG


Profa. Dra. Maria Aparecida Moura
ECI/UFMG


Prof. Dr. Mauricio Barcellos Almeida
ECI/UFMG

Obs: Este documento não terá validade sem a assinatura e carimbo da Coordenadora.


Profa. Renata Maria Abrantes Baracho Porto
- coordenadora do Programa de
Pós-Graduação em Ciência da
Informação - ECI / UFMG

AGRADECIMENTOS

Em uma longa jornada é importante contar com pontos de apoio no caminho, e graças a Deus, eu tive com quem contar em ambas às cidades em que vivi. Partindo das cidades e depois para meu ambiente de convívio, realizo os agradecimentos a seguir.

Gostaria de agradecer a minha tia Rosa, ao meu tio Eduardo e ao meu primo Dudu por terem me acolhido nos primeiros momentos em Belo Horizonte. Eles foram muito importantes para mim e os estimo muito.

Em primeiro lugar gostaria de agradecer a minha orientadora, Profa. Gercina, por ter me acolhido de braços abertos, pela paciência por me guiar por uma área totalmente nova para mim, pelos ótimos conselhos (nem sempre ouvidos da primeira vez), e pela capacidade de transitar entre os campos desse vasto universo da Ciência da Informação.

Gostaria de agradecer, também, a Benildes e ao Flávio pelas conversas, que foram fundamentais para o meu amadurecimento em relação à tese. Também agradeço a Fernanda, a Graciane e a Alessandra, cujos trabalhos me ajudaram a edificar minha própria pesquisa.

Não poderia deixar de agradecer a Gisele, Nely e a Carolina do programa de pós-graduação da UFMG pelo apoio, por serem sempre prestativas e por aturar minhas amolações frequentes e aos ilustres professores do programa de pós-graduação da Escola de Ciência da Informação - ECI/UFMG: Maria Aparecida Moura, Lídia Alvarenga, Alcenir, Marcello Bax (que me apresentou à Profa. Gercina), Ricardo Barbosa, Beatriz Cendón, Maria Guiomar, Maurício Almeida, Mônica Nassif, Marta Araújo, Renata Baracho, Hélio Kuramoto e Renato Rocha. E a Profa. Maria Luiza Campos, cuja tese me influenciou profundamente e que tive a oportunidade de conhecer.

Aos meus grandes amigos Daniel, Nely, Cleber, Sandro, Maria Rita e Lora, amigos com quem pude trocar infinitas ideias e que me estimularam a seguir a diante e em especial ao Pedro Barreto por me indicar o caminho na ECI/UFMG.

E claro, também, ao casal de ouro, Thales e Priscila, que me acolheram na reta final do estudo, sou imensamente grato por terem me suportado naquele período crucial. Enfim, todos de Belo Horizonte, não hesitem em me contatar caso venham a Goiânia, pois faço questão de ser seu anfitrião.

Gostaria de expressar minha gratidão a Érica e a Profa. Lindomar pela imensa ajuda na revisão da tese, e a Profa. Eliany e o Prof. Arnaldo por me abrirem as portas da Universidade Federal de Goiás para a realização da etapa final da pesquisa, o que não teria conseguido sem vocês.

Gostaria de agradecer também a todos que me influenciaram de alguma forma para atingir esse objetivo e que não foram mencionados aqui.

Concluindo, é impossível descrever o quanto meus pais me ajudaram, o quanto eu sou grato e tenho sorte de ser seu filho, dedico esse estudo a eles, essa conquista é nossa.

Só depois que a tecnologia inventou o telefone, o telégrafo, a televisão, a Internet, foi que se descobriu que o problema de comunicação mais sério era o de perto.

Millôr Fernandes

RESUMO

O presente estudo aponta um modelo de colaboração para a indexação e busca de registros em um catálogo *web* facetado colaborativo (*Facetlog*), cujo conteúdo é criado pelo próprio usuário. O estudo objetiva facilitar ou melhorar a qualidade da indexação, oferecendo uma estratégia complementar à etiquetagem, através da seleção de termos de uma taxonomia facetada. Visa, também, aperfeiçoar a recuperação da informação em catálogos *web*, através de uma interface de busca que combina a navegação facetada com a busca palavra-chave. Para a concepção do modelo de colaboração foi realizada uma pesquisa bibliográfica das seguintes temáticas: a) processo e modelos de recuperação da informação; b) indexação através de vocabulários controlados; c) classificação facetada; d) taxonomia facetada; e) indexação automática; f) interface para navegação facetada; g) catálogos *web* facetados; h) modelos de colaboração da *Web 2.0*; i) sistemas colaborativos; j) fluxo de trabalho; k) moderação; l) etiquetagem social; m) *wiki*; e, n) sistemas de votação. Na revisão de literatura foram pesquisados trabalhos sobre a etiquetagem facetada, classificação facetada colaborativa e catálogos facetados de endereços de empresas. Para a construção do modelo de colaboração foi utilizado o processo de desenvolvimento de *software* RUP, os diagramas da linguagem UML, prototipação de baixa fidelidade e *softwares* livres como o *Wordpress*. Para a avaliação do modelo de colaboração foi realizado um teste de usabilidade para a validação de requisitos. Este, por sua vez, permitiu constatar a preferência do usuário pela busca por palavra-chave ao invés do uso da navegação facetada. Também mostrou a inclinação do usuário para a realização da etiquetagem em relação à seleção de termos na taxonomia facetada. De qualquer maneira, o estudo revelou que tanto a navegação facetada quanto a taxonomia facetada são estratégias válidas para complementar, respectivamente, a recuperação da informação e a indexação, sendo utilizadas, em geral, em um terço das ações dos usuários. O estudo recomenda a utilização da função de autocompletar para selecionar as etiquetas e os termos da taxonomia facetada tanto na interface de busca quanto na interface de indexação. Também recomenda que, conforme o grau de impacto de uma ação no catálogo *web* facetado, maior seja a moderação, com a associação de uma moderação pré-aprovada para ações como o comentário, votação e indexação e uma moderação por aprovação para as ações de gestão da taxonomia facetada.

Palavras-chave: navegação facetada, folksonomia, *Web 2.0*, colaboração, catálogo *web*

ABSTRACT

This study proposes a collaborative model for indexing and searching for items in a collaborative faceted web catalog (Facetlog), where its own users create the content. It aims at enhancing or improving the quality of indexing by offering a complementary strategy to the tagging through the selection of terms from a faceted taxonomy. It also aims at perfecting the information retrieval in web catalogs by using a searching interface that combines faceted and keyword search. In order to design the collaborative model, a review of the literature was conducted on the following topics: a) processes and models of information retrieval; b) indexing through controlled vocabulary; c) faceted classification; d) faceted taxonomy; e) automatic indexing; f) faceted searching interface; g) faceted web catalogs; h) Web 2.0 collaboration models; i) collaborative systems; j) workflow; k) moderation; l) social tagging; m) wiki; and n) voting systems. Studies on faceted tagging, collaborative faceted classification and faceted yellow pages were also researched in the review of the literature. The collaborative model was created by using the RUP software development process, the UML language diagram, low-fidelity prototyping and open source softwares like the WordPress. The evaluation of the collaborative model, on the other hand, was performed the usability test to validate the requirements. This test revealed the users's preference to use the keyword search instead of the faceted search. It also showed their inclination to use tagging instead of faceted taxonomy. However, the study also revealed that both the faceted search and the faceted taxonomy are valid complementary strategies to information retrieval and indexing, given that both are used in one-third of the users's actions. This study concludes by recommending the use of the autocomplete function for selecting tags and categories in the searching and indexing interfaces. It also recommends that the level of moderation should vary according to the degree of impact of an action on the faceted web catalog. The study suggests the following pattern: voting, comment and classification should be associated to a pre-approved moderation and the taxonomy management to an approved moderation.

Keywords: faceted search, folksonomy, web 2.0, collaboration, web catalog

LISTA DE FIGURAS

FIGURA 1 - Estrutura dos capítulos da tese	30
FIGURA 2 - Estrutura do capítulo de fundamentação teórica e metodológica	31
FIGURA 3 - Etapas da recuperação da informação	33
FIGURA 4 - Arquitetura de um SRI computadorizado.....	34
FIGURA 5 - Representação de objeto de conteúdo.....	36
FIGURA 6 - Formas de visualização de uma hierarquia	39
FIGURA 7 - Relação gênero-espécie	39
FIGURA 8 - Classificação enumerativa versus classificação facetada.....	41
FIGURA 9 - Exemplos de taxonomias facetadas	45
FIGURA 10 - Indexação automática	48
FIGURA 11 - Coletivo versus colaborativo	49
FIGURA 12 - Etiquetagem versus categorização.....	50
FIGURA 13 - Tarefa de recuperação da informação.....	52
FIGURA 14 - Navegação "um nível por tela"	53
FIGURA 15 - Árvore expansível.....	54
FIGURA 16 - Menu <i>dropdown</i>	55
FIGURA 17 - Caixas de seleção encadeadas	55
FIGURA 18 - Tipos de seleções na navegação facetada	57
FIGURA 19 - Operação <i>OR</i> e <i>NOT</i> na navegação facetada	57
FIGURA 20 - Interface do Flamenco.....	58
FIGURA 21 - FACTIC (<i>Facet Image Browser</i>)	60
FIGURA 22 - <i>National Health Promotion Portal</i>	61
FIGURA 23 - Navegador facetado em base de dados RDF	62
FIGURA 24 - Superfacetas do <i>Getty Images</i>	63
FIGURA 25 - Interface do <i>site</i> Yelp.....	64
FIGURA 26 - Estrutura padrão de uma interface facetada	65
FIGURA 27 - Interface de busca do Koah	73
FIGURA 28 - Aprovação de etiquetas sugeridas pelo usuário.....	73
FIGURA 29 - Principal interface do usuário do <i>Venturi Portal</i>	75
FIGURA 30 - Interface do FaThumb	75
FIGURA 31 - Modelo de colaboração 3C	77
FIGURA 33 - Classificação de sistemas segundo o modelo de colaboração 3C.....	79
FIGURA 34 - Modelo de Referência para um WfMS.....	81
FIGURA 35 - Serviços da <i>Web 2.0</i>	82

FIGURA 36 - <i>Web tradicional versus Web 2.0</i>	85
FIGURA 37 - Obtenção de consenso na Wikipédia	88
FIGURA 38 - Interface do sistema Intense Debate	90
FIGURA 39 - Modelos prescritivos de processo de desenvolvimento de <i>software</i>	91
FIGURA 40 - Fases do processo de desenvolvimento RUP	93
FIGURA 41 - Fases do processo de desenvolvimento RUP	94
FIGURA 42 - Exemplo de um diagrama de caso de uso	98
FIGURA 43 - Exemplo de fluxo de trabalho representado através um diagrama de atividades	99
FIGURA 44 - Protótipo de baixa e alta fidelidade respectivamente	101
FIGURA 45 - Revisão de uma interface através de um protótipo de baixa fidelidade	102
FIGURA 46 - Prototipação em papel	103
FIGURA 47 - Estratégia de teste de <i>software</i>	106
FIGURA 48 - Modelo V para teste de <i>software</i>	107
FIGURA 49 - Curva de problemas detectados versus participantes	111
FIGURA 50 - Estrutura da revisão de literatura	113
FIGURA 52 - Interface para etiquetagem facetada do <i>Faccete</i>	115
FIGURA 53 - Interface do <i>delicious.com</i>	115
FIGURA 54 - Interface para navegação facetada do <i>Faccete</i>	116
FIGURA 55 - Etiquetagem facetada de <i>bookmarks</i> do <i>Facetag</i>	117
FIGURA 56 - Navegação facetada no <i>Facetag</i>	118
FIGURA 57 - Navegação facetada por <i>tagclouds</i> do <i>TagExplorer</i>	119
FIGURA 58 - Mapeamento automático das etiquetas para as categorias do <i>Wordnet</i>	119
FIGURA 59 - Interface para classificação facetada colaborativa	121
FIGURA 60 - Etapas da pesquisa	124
FIGURA 61 - Etapas do desenvolvimento do protótipo	132
FIGURA 62 - Modelo de recuperação da informação do <i>Facetlog</i>	133
FIGURA 63 - Processo de prepação e indexação do <i>Facetlog</i>	134
FIGURA 64 - Triângulo do modelo de colaboração 3C do <i>Facetlog</i>	139
FIGURA 65 - Criar catálogo	140
FIGURA 66 - Adicionar registro	141
FIGURA 67 - Buscar e votar em registro	142
FIGURA 68 - Visualizar e indexar dos registros	143
FIGURA 69 - Tratar as contestações	144
FIGURA 70 - Coordenação com fluxo de trabalho de aprovação	145
FIGURA 71 - Interface original de cadastro do módulo <i>WP Idea Stream</i>	149
FIGURA 72 - Interface original de navegação do módulo <i>WP Idea Stream</i>	149

FIGURA 73 - Diagrama do banco de dados do <i>Wordpress</i>	150
FIGURA 74 - Prototipação de baixa fidelidade na navegação facetada no módulo <i>WP Idea Stream</i>	153
FIGURA 75 - Criar Catálogo	155
FIGURA 76 – Interface de busca	156
FIGURA 77 - Adicionar item ao catálogo.....	157
FIGURA 78 - Usuário entra no <i>site</i> ou se cadastra	158
FIGURA 79 - Adicionar registro no catálogo	159
FIGURA 80 - Visualização de um item do catálogo	160
FIGURA 81 - Contestação de classificação de um item do catálogo	160
FIGURA 82 - Tratamento das contestações.....	161
FIGURA 83 - Analisando contestação	161
FIGURA 84 - Alteração da classificação conforme sugestão do usuário	162
FIGURA 85 - Votação, ordenação e ranking de autores.....	162
FIGURA 86 - Estrutura da análise do teste de usabilidade.....	163
FIGURA 87 - Exemplo do fluxo de trabalho de um EMX	180
FIGURA 88 - Moderação versus cooperação.....	182
FIGURA 89 - Configuração da coordenação.....	184
FIGURA 90 - Moderação associada ao objeto/ação no catálogo	187
FIGURA 91 - Trabalhos futuros	188
FIGURA 92 - Primeiro protótipo de baixa fidelidade da interface facetada	210
FIGURA 93 - Ciclos iniciais de desenvolvimento	215
FIGURA 94 - Criar catálogo	219
FIGURA 95 - Realizar navegação facetada	219
FIGURA 96 - Modelo de colaboração <i>wiki</i>	220
FIGURA 97 - Navegação facetada e modelo de colaboração	221
FIGURA 98 - Modelo de dados	222
FIGURA 99 - inserção de facetas coloridas	224
FIGURA 100 - ordenação de resultados.....	225
FIGURA 101 - A inserção de caixa de busca nas facetas	226
FIGURA 102 - Facetas sem barra de rolagem	227
FIGURA 103 - Facetas com barra de rolagem automática	227
FIGURA 104 - Facetas com agrupamento pelas letras do alfabeto	228
FIGURA 105 - Paginação dos indivíduos	229
FIGURA 106 - Ações colaborativas	230
FIGURA 107 - Realizar navegação facetada	233
FIGURA 108 - Modelo de colaboração <i>wiki</i>	234

FIGURA 109 - Navegação facetada	235
FIGURA 110 - Modelo de colaboração <i>wiki</i>	236
FIGURA 111 - Colaboração na taxonomia da faceta	238
FIGURA 112 - Invalidar categorias	239
FIGURA 113 - Alimentando o catálogo.....	239
FIGURA 114 - Usuários podem validar/invalidar o material inserido.....	240
FIGURA 115 - Adicionar nova classificação	241
FIGURA 116 - Ordenação dos registros recuperados	241
FIGURA 117 - Busca por palavra-chave nos registros	242
FIGURA 118 - Busca por palavra-chave nas facetas.....	242
FIGURA 119 - Consultas complexas envolvendo categorias, classificações ou registros inválidos	246
FIGURA 120 - Realizar navegação facetada	249
FIGURA 121 - Modelo de colaboração <i>wiki</i>	250
FIGURA 122 - Navegação facetada	251
FIGURA 123 - Modelo de colaboração <i>wiki</i>	252
FIGURA 124 - Combinação de categorias (<i>AND</i>)	255
FIGURA 125 - Operação <i>OR</i>	256
FIGURA 126 - Operação <i>NOT</i>	256
FIGURA 127 - Busca por palavra-chave nos registros	257
FIGURA 128 - Busca por palavra-chave nas facetas.....	257
FIGURA 129 - Filtrando recursivamente ou apenas naquele nível hierárquico	258
FIGURA 130 - Paginação e resultados.....	258
FIGURA 131 - Ordenação de resultados por características ou votação, crescente ou decrescente.....	259
FIGURA 132 - Adicionar registro	259
FIGURA 133 - Invalidar registro	260
FIGURA 134 - Mostrar registro inválido e validar registro.....	260
FIGURA 135 - Adicionar classificação	261
FIGURA 136 - Invalidar classificação de um registro.....	261
FIGURA 137 - Adicionar categoria.....	262
FIGURA 138 - Invalidar categoria	262
FIGURA 139 - Mostrar categorias inválidas	263
FIGURA 140 - Validar uma categoria inválida.....	263
FIGURA 141 - Dar nota para um registro em uma determinada votação.....	264

LISTA DE GRÁFICOS

GRÁFICO 1 - Resultados do teste com usuários do <i>FaThumb</i>	76
GRÁFICO 2 – Idade dos participantes do teste de usabilidade	164
GRÁFICO 3 – Formação acadêmica dos participantes	164
GRÁFICO 4 – Tempo de utilização da Internet	165
GRÁFICO 5 – Experiência com catálogo de endereços <i>web</i> na Internet.....	165
GRÁFICO 6 - Cadastros de registros por usuário	166
GRÁFICO 7 – Utilização da classificação facetada <i>versus</i> etiquetas por usuário	167
GRÁFICO 8 – Total da utilização da classificação facetada <i>versus</i> etiquetas	167
GRÁFICO 9 – Utilização das categorias <i>versus</i> etiquetas	168
GRÁFICO 10 – Encontrar empresas para hospedagem em um bairro	169
GRÁFICO 11 – Encontrar empresas que vendem determinado produto perto de um ponto de referência	170
GRÁFICO 12 – Empresas para reforma em casa	170
GRÁFICO 13 – Total da utilização da navegação facetada <i>versus</i> busca	171
GRÁFICO 14 – Estratégias de busca que mais permitiram os usuários encontrarem resultados.....	171
GRÁFICO 15 – Votação nos registros por usuário	172
GRÁFICO 16 – Usuários revisam a classificação dos registros	173
GRÁFICO 17 – Cadastro: categorias facetadas <i>versus</i> etiquetas	174
GRÁFICO 18 – Navegação facetada <i>versus</i> busca por palavra-chave.....	175
GRÁFICO 19 – Avaliação do resultado da votação	175
GRÁFICO 20 – Contestação.....	176

LISTA DE QUADROS

QUADRO 1 – Mapeamento das facetas do PMEST	43
QUADRO 2 - Opções de <i>design</i> para uma interface facetada.....	65
QUADRO 3 - Taxonomia de sistemas colaborativos	77
QUADRO 4 - Análise através do modelo de colaboração 3C.....	138
QUADRO 5 - Mapeamento do catálogo para o banco de dados.....	151
QUADRO 6 - Moderação padrão sugerida	181
QUADRO 7 - Associação de moderação para cada perfil de usuário e ação	183
QUADRO 8 – Vocabulário controlado da tese	211
QUADRO 9 – Legenda.....	214
QUADRO 10 – Modelo de dados para o armazenamento das relações (tripas)	224
QUADRO 11 – Modelo de dados com as categorias filhas	237
QUADRO 12 – Modelo de dados com a linhagem	253
QUADRO 13 – Modelo de dados com a linhagem com valor de verdade.....	254
QUADRO 14 – Avaliação dos ciclos de desenvolvimento	265
QUADRO 15 – Diferenças na plataforma e escopo do projeto nos ciclos de desenvolvimento	267
QUADRO 16 – Ações para gestão da taxonomia	269

LISTA DE ABREVIATURAS

ACM	-	<i>Association for Computing Machinery</i>
AJAX	-	<i>Asynchronous JavaScript and XML</i>
BCI	-	Biblioteconomia e Ciência da Informação
Bepel	-	<i>Business Process Execution Language</i>
BPMN	-	<i>Business Process Management Notation</i>
CDD	-	Classificação Decimal de Dewey
CDU	-	Classificação Decimal Universal
CMS	-	<i>Content Management System</i>
CRG	-	<i>Classification Research Group</i>
CSS	-	<i>Cascading Style Sheets</i>
ECI	-	Escola de Ciência da Informação
EMX	-	Esquema de Moderação Extensível
<i>Facetlog</i>	-	<i>Faceted Collaborative Catalog</i>
FACTIC	-	<i>Faceted Image Browser</i>
FIC	-	Faculdade de Informação e Comunicação
GPL	-	<i>General Public License</i>
HTML	-	<i>HyperText Markup Language</i>
HTTP	-	<i>Hypertext Transfer Protocol</i>
IBICT	-	Instituto Brasileiro de Informação em Ciência e Tecnologia
IBIU	-	Interface de Busca e Indexação Unificada
IBU	-	Interface de Busca Unificada
IR	-	<i>Information Retrieval</i>
ITA	-	Instituto Tecnológico de Aeronáutica
LAMP	-	Linux, Apache, MySQL e PHP
LCA	-	Listas de Cabeçalhos de Assunto
MCF	-	Modelo de Colaboração Facetado
MCU	-	Modelo Colaborativo Unificado
MIT	-	<i>Massachusetts Institute of Technology</i>
MTA	-	Agente de transferência de <i>e-mail</i>
NDLTD	-	<i>Networked Digital Library of Theses and Dissertations</i>
NISO	-	<i>National Information Standards Organization</i>
ODP	-	<i>Open Directory Project</i>
OPAC	-	<i>Online Public Access Catalogue</i>

PMEST - *Personality, Matter, Energy, Space, Time*
PHP - *PHP: Hypertext Preprocessor*
RDF - *Resource Description Framework*
RUP - *Rational Unified Process*
SciELO - *A Scientific Electronic Library Online*
SPO - *Subject-Predicate-Object*
SQL - *Structured Query Language*
SRI - *Sistema de Recuperação da Informação*
UCM - *Unified Collaborative Model*
UFG - *Universidade Federal de Goiás*
UCC - *User Created Content*
UI - *User Interface*
UFMG - *Universidade Federal de Minas Gerais*
UML - *Unified Modeling Language*
URI - *Uniform Resource Identifier*
URL - *Uniform Resource Locator*
URN - *Uniform Resource Name*
USI - *Unified Search Interface*
XML - *eXtensible Markup Language*
XPDL - *XML Process Definition Language*
W3C - *World Wide Web Consortium*

SUMÁRIO

1	INTRODUÇÃO.....	22
1.1	Problema e justificativa	25
1.2	Objetivos.....	28
1.2.1	Objetivo geral.....	28
1.2.2	Objetivos específicos	29
1.3	Estrutura da tese.....	29
2	FUNDAMENTAÇÃO TEÓRICO-METODOLÓGICA	31
2.1	Sistema de recuperação da informação	32
2.1.1	O processo de recuperação	33
2.1.2	Indexação	35
2.2	Vocabulários controlados	36
2.2.1	Taxonomia	38
2.2.2	Taxonomia facetada.....	40
2.3	Linguagem natural	47
2.3.1	Indexação automática	47
2.3.2	Folksonomia e etiquetagem	48
2.4	Interface de busca	51
2.4.1	Navegação hierárquica	52
2.4.2	Navegação facetada	56
2.4.3	Usabilidade	68
2.5	Do catálogo ao catálogo facetado	70
2.5.1	Catálogo - breve histórico	70
2.5.2	Catálogo web facetado	72
2.6	Sistemas colaborativos e modelo de colaboração 3C	76
2.6.1	Fluxo de trabalho (workflow)	80
2.7	Modelos de colaboração da <i>Web 2.0</i>	81
2.7.1	Princípios da <i>Web 2.0</i>	81
2.7.2	Colaboração na <i>Web 2.0</i>	84
2.7.3	Wiki.....	86
2.7.4	Sistemas de votação.....	89
2.8	Processo de desenvolvimento de <i>software</i>	91

2.8.1	RUP	93
2.8.2	UML	95
2.8.3	Prototipação.....	100
2.8.4	Teste de software	103
3	REVISÃO DA LITERATURA.....	113
3.1	Etiquetagem facetada	113
3.2	Classificação facetada colaborativa	120
4	METODOLOGIA.....	122
4.1	Classificação da pesquisa.....	123
4.2	Etapas da pesquisa.....	124
4.2.1	Pesquisa bibliográfica	125
4.2.2	Desenvolvimento do protótipo	127
4.2.3	Teste de usabilidade	128
5	DESENVOLVIMENTO DO CATÁLOGO <i>WEB</i> FACETADO COLABORATIVO	132
5.1	Concepção.....	132
5.1.1	Objetivos e escopo do projeto.....	133
5.1.2	Atores	136
5.2	Elaboração.....	136
5.2.1	Requisitos funcionais	136
5.2.2	Requisitos não funcionais	139
5.2.3	Diagramas UML	140
5.2.4	Escolha da ferramenta	146
5.2.5	Customização do módulo.....	148
5.2.6	Modelo de dados	150
5.2.7	Modelagem de taxonomia facetada	151
5.2.8	Prototipação de baixa fidelidade da interface facetada	152
5.2.9	Construção	153
5.2.10	Criação do ambiente	154
5.2.11	Módulos instalados	154
5.2.12	Implementação da interface de busca.....	155
5.2.13	Implementação do modelo de colaboração.....	157
6	ANÁLISE DOS RESULTADOS E MELHORIAS.....	163

6.1	Perfil dos usuários	164
6.2	Análise da captura do vídeo.....	166
6.2.1	Análise da indexação dos registros.....	166
6.2.2	Análise da recuperação de registros	168
6.2.3	Análise da votação dos usuários.....	172
6.2.4	Análise da revisão da indexação dos registros.....	172
6.3	Análise do questionário de satisfação	173
6.3.1	Análise da indexação dos registros.....	174
6.3.2	Análise da recuperação de registros	174
6.3.3	Análise da votação dos usuários.....	175
6.3.4	Análise da revisão da indexação dos registros.....	176
6.4	Síntese dos erros e deficiências do protótipo	176
6.5	Revisão do modelo de colaboração	178
7	CONSIDERAÇÕES FINAIS.....	185
7.1	Trabalhos futuros	187
	REFERÊNCIAS	193
	APÊNDICE A - TERMO DE CONSENTIMENTO	205
	APÊNDICE B - QUESTIONÁRIO PARA IDENTIFICAÇÃO DO PERFIL DO PARTICIPANTE	206
	APÊNDICE C - LISTA DE TAREFAS.....	207
	APÊNDICE D - QUESTIONÁRIO DE SATISFAÇÃO DO USUÁRIO	209
	APÊNDICE E - PRIMEIRO PROTÓTIPO DA INTERFACE FACETADA.....	210
	APÊNDICE F - VOCABULÁRIO CONTROLADO DA TESE	211
	APÊNDICE G - CICLOS DE DESENVOLVIMENTOS ANTERIORES	215

1 INTRODUÇÃO

No âmbito da Biblioteconomia e Ciência da Informação (BCI), catálogos são instrumentos tradicionalmente utilizados para organização de documentos. Remontam à época das primeiras bibliotecas, quando surgiram como um conjunto de registros, que dispunham de índices com o título da obra ou nome do autor listados sequencialmente ou em ordem alfabética.

Com o passar tempo e aumento do tamanho da coleção de documentos, aperfeiçoamentos foram agregados aos índices como o número de localização do livro na estante, o tamanho do livro ou o número de livros existentes na coleção (KAFURE, 2004).

Novos instrumentos foram elaborados para melhorar a recuperação da informação, como as listas de cabeçalho de assunto (LCA), que passaram a ser usadas na indexação do documento, ou seja, na representação temática ou do conteúdo do documento. Esse tipo de instrumento veio para auxiliar na elaboração dos registros bibliográficos, sugerindo termos para representação dos assuntos do documento. Assim, através da leitura de um registro bibliográfico, é possível selecionar, entre os vários documentos da coleção, um documento que pode satisfazer ou sanar a necessidade de informação do usuário, sem precisar ler todo o documento, para ver se o mesmo o atende. Esse instrumento, atua como um vocabulário controlado, que é um conjunto de termos que devem ser usados preferencialmente em detrimento de outros na representação ou na indexação de um documento, e, portanto, representa apenas uma parcela de todo o vocabulário disponível na linguagem natural. Isso permite, entre outras vantagens, agrupar documentos similares, com os mesmos assuntos expressados por termos diferentes ou diferenciar documentos que possuem termos homógrafos¹.

Historicamente, com a invenção da imprensa por Gutenberg em meados do século XV, houve uma explosão na publicação de documentos em relação aos séculos anteriores (DIAS, 1999). Os documentos não mais precisavam ser copiados a mão, pois podiam ser duplicados mecanicamente e em larga escala. Desse modo, houve uma necessidade de melhorar o processo de representação terminológica desses documentos, e, com isso, a indexação, como um processo de representação temática, foi evoluindo com o tempo. Apesar de o vocabulário controlado ser um conjunto menor de termos em relação aos disponíveis na linguagem natural, ele foi ficando cada vez maior, com o aumento do número de títulos disponíveis. O vocabulário controlado passou a ser impresso também.

¹ Um termo homógrafo é aquele que é grafado da mesma maneira ao qual se atribui mais de um significado, com por exemplo Peru (país) e Peru (ave) (LANCASTER, 1993).

O ato de indexar um documento é complexo, pois, além de toda a subjetividade da análise prévia, em que são identificados os assuntos principais do documento (análise conceitual), o indexador precisa representá-lo através dos termos preferenciais contidos no vocabulário controlado utilizado pelo Sistema de Recuperação da Informação (SRI). No decorrer dos séculos posteriores ao advento da imprensa, o vocabulário controlado utilizado na indexação dos documentos tornou-se ainda mais extenso, causando muitas dificuldades também na outra ponta do processo: a recuperação pelos usuários, que precisavam, então, percorrer uma extensa lista de termos contidos no vocabulário para localizar o documento na estante.

Outros tipos de vocabulário controlado, alternativo à lista de cabeçalhos de assuntos, são os sistemas de classificação e taxonomias, que possuem os termos organizados em uma estrutura hierárquica, geralmente representando a relação gênero-espécie e todo-parte. Segundo Piedade (1983), esses sistemas têm origem nas classificações filosóficas, cujo precursor foi Aristóteles (384-322 a.C.).

Na BCI, os sistemas de classificação enumerativos são os mais difundidos. Entre eles estão o Classificação Decimal de Dewey (CDD) e o Classificação Decimal Universal (CDU). Esses sistemas foram criados no final do século XIX e se difundiram durante o século XX e trouxeram a vantagem de agrupar hierarquicamente os documentos que apresentam temas relacionados, trazendo a possibilidade de mapear os termos de uma mesma área. Sendo assim, o usuário pode conhecer os termos mais gerais que são fornecidos inicialmente na raiz da hierarquia. Ao selecionar um termo, é possível acessar os termos inferiores, mais específicos que o termo superior, e prosseguir sucessivamente, até encontrar o termo com o assunto e especificidade desejados. Isso também permite a descoberta do conhecimento, em que o usuário, durante o processo de busca, toma conhecimento de termos que ele não conhecia anteriormente. Entretanto, Batista e Carvalho (2003, p. 40) apontam: “Uma estrutura hierárquica é fácil de compreender e usar, porém não é flexível”. Corroborando com essa afirmativa, Broughton e Slavic (2007) afirmam que a estrutura hierárquica nem sempre é flexível o suficiente para acomodar novos assuntos ou assuntos compostos.

Como alternativa a essa estrutura pouco flexível, surgiu um pouco mais tarde, a classificação facetada. O princípio da análise facetada foi idealizado por Ranganathan e teve início em 1933, principalmente após a publicação do livro *Colon Classification* (Classificação de Dois Pontos) e, posteriormente, com a publicação do livro *Prolegomena*, em 1937. Nessa proposta, Ranganathan propõe que os termos sejam divididos em facetas, ao invés de ter uma única grande hierarquia que acomoda todos os termos. Assim, uma faceta é uma lista de termos com as mesmas características, entretanto, cada faceta exhibe

uma relação diferente com o assunto geral. Os assuntos compostos são gerados na indexação ou recuperação do documento, tornando as facetas mais sintéticas. Segundo LIMA (2004, p. 78): “[...] em uma área do conhecimento, o número de assuntos específicos que podem ser selecionados numa classificação é infinito”.

Uma desvantagem da indexação manual é alto custo e lentidão no processamento humano, que necessita de treinamento e aprendizado do domínio para escolha dos termos mais adequados para função de representação e recuperação da informação (CARROLL; ROELOFFS, 1969 citado por VIEIRA, 1988). Paralelamente, a evolução exponencial da tecnologia computacional impulsionou a área de catalogação com o surgimento, no final da década de 1950, de uma nova técnica para a representação dos documentos: a indexação automática. Ao invés de ser realizada manualmente por um indexador, a mesma é realizada pelo computador, por meio da linguagem natural. A indexação automática utiliza métodos que realizam a extração das palavras contidas no texto que formam o índice para, posteriormente, ser utilizado na recuperação. Permite a busca por palavra-chave (*text search*), na qual procura realizar casamento entre as palavras digitadas pelo usuário e as palavras do índice, que apontam para os documentos que as contêm. Tal representação tem maior exaustividade, com uma grande cobertura dos assuntos tratados no documento.

Na década de 90, viu-se a Internet se popularizar após a invenção da *World Wide Web* pelo cientista Tim Berners-Lee. De repente, observou-se uma revolução significativa, em que uma pessoa poderia produzir e disponibilizar um documento para ser acessado em qualquer lugar do mundo de forma quase instantânea. Dias (1999) compara isso à invenção da imprensa. O número de *sites* aumentou em ritmo acelerado. No final da década de 90, dois projetos se destacaram como instrumentos para encontrar *sites* na Internet: a gigante *Yahoo!*, que iniciou em 1994 com o *site Jerry and David's Guide to the World Wide Web*, sendo o catálogo de *sites* seu primeiro serviço oferecido; e o *Open Directory Project* (ODP) lançado pela *Netscape*. Ambos consistiam em uma grande hierarquia utilizada para indexar os *sites* e aceitavam sugestões fornecidas pelos usuários. No início dos anos 2000, esses catálogos ultrapassaram a marca de um milhão de *sites* catalogados. Entretanto, outro movimento estava prestes a acontecer.

Em 2001, ocorreu a explosão da bolha da Internet com a queda de ações, fechamentos e fusões das empresas de tecnologia da informação baseadas na Internet, as chamadas empresas "pontocom". Em 2004, a partir da percepção de características comuns às empresas que sobreviveram à bolha, teve origem o conceito *Web 2.0*. Uma dessas características é apostar na colaboração dos usuários ou na inteligência coletiva que eles podem agregar.

Novamente, a evolução tecnológica impactou a área, permitindo que os próprios usuários, de maneira livre, sem regras, políticas ou vocabulários controlados, realizassem a indexação. O nome que se dá a esse tipo de indexação é folksonomia caracterizada pela etiquetagem, que é a associação de etiquetas (termos na linguagem natural) a recursos de informação. É uma característica comum aos *sites* de maior sucesso da era *Web 2.0*, tais como o *Delicious*, *Digg*, *Wikipedia*, *Flickr*, e *Youtube*. A recuperação da informação nesse caso baseia-se no casamento entre os termos usados na marcação do recurso *web* com os termos usados na busca pelo usuário. É considerada uma indexação social (GUEDES, 2010). Age de maneira similar à busca por palavra-chave e indexação automática, que também usa a linguagem natural. É adequada ao ambiente *web*, devido à "diminuição de custos e tempo para o usuário. Isto acontece pelo fato de não existirem hierarquias complexas de termos e nomenclaturas que o usuário tenha que aprender" (GUEDES, 2010, p. 96).

1.1 Problema e justificativa

A motivação para a realização desta extensa pesquisa, que decorreu durante quase uma década, teve origem no ano de 2002, quando o presente pesquisador trabalhava no setor de Pesquisa e Desenvolvimento de uma empresa de desenvolvimento de equipamentos eletrônicos. Nesse período foi possível para o pesquisador ter uma visão holística de todo o processo de desenvolvimento de *sistemas*, desde a fase inicial de elaboração de requisitos, a criação de casos de teste para validar os requisitos, até a realização de testes e a confecção de manuais para o usuário e o administrador. Todo esse processo é documentado, e, a esta documentação, soma-se a especificação da arquitetura do sistema e o código-fonte desenvolvido. Dispunha-se de uma grande quantidade de documentos, elaborados por várias pessoas, ao longo de vários anos, e praticamente todos eles estavam disponíveis em meio digital. O intercâmbio dos documentos entre a equipe acontecia majoritariamente por *e-mail* e através de pastas compartilhadas no servidor de arquivos, pelo tradicional programa *Windows Explorer*.

A organização da documentação de desenvolvimento era um trabalho complexo. A orientação da gerência era que a documentação fosse organizada pela versão do produto. Entretanto, a estrutura de pastas, que agrupava documentos do mesmo tipo, por versão do produto, muitas vezes se mostrava ineficiente. Um exemplo dessa situação é o tratamento dos requisitos: cada versão do produto tinha novos requisitos a serem implementados e

havia um prazo a ser cumprido, pois tratava-se de um projeto. Um requisito originava um ou mais casos de testes. Se fosse aprovado, o requisito impactava nos manuais e, no caso contrário, impactava na arquitetura do sistema e código-fonte. Com o passar do tempo, e à medida que o prazo final para o término do projeto se aproximava, surgiam as perguntas: Quais requisitos foram implementados? Quais requisitos foram implementados e aprovados? Quais requisitos de alta prioridade não foram aprovados? Mas a complexidade da problema era bem maior e não se restringia apenas a essas questões, tendo em vista que novos requisitos podem impactar em requisitos antigos e pode ser necessário reaplicar testes a todos os requisitos do sistema, com o intuito de assegurar sua qualidade. Muitas vezes, nesse processo de atualização, acontecia a liberação de versões do produto, sem que requisitos antigos estivessem funcionando. Isso desperta novas questões referentes à recuperação da informação, como por exemplo: Quais são os requisitos, herdados de todas as versões do produto, que estão vigentes no momento? Quais foram descartados? Quais novos requisitos, com menor prioridade, podem ser descartados ou adiados para a próxima versão, para foco nos requisitos de alta prioridade? Essas necessidades de informação levavam a extenuantes buscas. A organização de toda essa documentação através de uma única grande hierarquia de pastas mostrava-se ineficiente e custosa. Por exemplo, para recuperar todos os requisitos vigentes, era necessário buscar a pasta de cada projeto e, manualmente, somar os requisitos implementados em cada versão do produto. A frustração na execução dessa tarefa fez nascer uma inquietação: deveria haver uma maneira mais adequada de organizar essa informação.

Foi a partir daí que as ideias e reflexões acerca da questão foram se encorpando. Entretanto, a falta de conhecimento da terminologia correta fez com que a pesquisa se estendesse ao longo dos anos de 2002 e 2003, quando o autor encontrou a tese de doutorado de Maria Luiza Campos, intitulada "*A Organização de Unidades do Conhecimento em Hiperdocumentos*" (CAMPOS, 2001). Através desse trabalho, o pesquisador teve contato com a Teoria de Classificação Facetada, que seria de suma importância para o desenvolvimento de todo seu trabalho posterior. Esse novo conhecimento parecia ser endereçado exatamente aos problemas visualizados no cotidiano e colocados em questão inicialmente. O crescente interesse pela área de Gestão e Organização da Informação fez o pesquisador ingressar, em 2004, no curso de Gestão Estratégica de Informação nas Empresas e, posteriormente, no mestrado na linha de Organização e Uso da Informação na Escola de Ciência da Informação (ECI/UFMG). Em 2005, à luz desses novos conhecimentos, foi criado um protótipo da interface do sistema que seria a solução para os problemas que haviam sido encontrados no passado (APÊNDICE E).

Nessa época, estimuladas pelo desenvolvimento da Internet e da *Web 2.0*, também surgiram pesquisas relacionadas com a classificação facetada, e esse assunto ganhou uma nova roupagem no mundo digital. Como aponta Hedden (2010, p. 8), "Hoje com o uso de computadores e bancos de dados, a Classificação de Dois Pontos não é necessária para implementar classificação facetada, e busca por facetas tornou-se fácil de implementar e usar".

Com a aplicação da classificação facetada na organização no meio digital surgiu a possibilidade de combinação dos termos de facetas diferentes para a formação dos assuntos compostos no ato de uma dada busca, que se denominou navegação facetada. Pode-se afirmar que a navegação facetada é uma estratégia eficiente de recuperação da informação e apresenta várias vantagens, conforme evidencia a revisão de literatura feita por Fagan (2010). A navegação facetada tornou-se bastante comum, principalmente em *sítes* de comércio eletrônico, catálogos de endereços e bancos de imagens (DENTON, 2003).

Entretanto, a classificação facetada também pode apresentar obstáculos, conforme aponta Maly, Wu e Zubair:

[...] para coleções que crescem em volume e variedade, o maior desafio envolvendo a classificação facetada é reclassificar os objetos quando ocorre uma modificação no sistema de classificação facetada. Um sistema de classificação com uma gestão centralizada, frequentemente apresenta dificuldades para ser adaptado em grandes coleções que crescem continuamente [...] (MALY; WU; ZUBAIR, 2010, p. 70, tradução nossa).

A etiquetagem social apresenta-se como uma alternativa para tratar o problema acima, mas ela também apresenta limitações. Como exposto por Lai (2009), é difícil navegar em uma nuvem de etiquetas, formada por um conjunto superior a 100 (cem) etiquetas, pois apenas algumas etiquetas ficam em evidência, ofuscando as demais. Além disso, o autor expõe que as ambiguidades no uso das etiquetas podem remeter o usuário a informações pouco relevantes.

Assim, nos anos 2000, surgiram pesquisas que apontam a aplicação da classificação facetada como forma de auxiliar a estruturação da etiquetagem e facilitar a recuperação de informação. Como exemplo desses estudos, podem ser citados os protótipos *Faccete* (LAI, 2009) e o *FaceTag* (QUINTARELLI; RESMINI; ROSATI, 2007) que permitem a etiquetagem facetada, orientando o usuário no momento de atribuir as etiquetas, dando a ele a possibilidade de escolher em qual faceta o termo será inserido. Em outro estudo, Spiteri (2011) aponta o emprego das facetas na etiquetagem, como uma possibilidade de melhorar a qualidade da indexação e recuperação da informação.

Maly, Wu e Zubair (2010), por sua vez propuseram um sistema de classificação facetada colaborativo, cuja vantagem é permitir que o sistema se reorganize

automaticamente conforme os interesses dos usuários, ajudando-os a navegar pela coleção de forma rápida e intuitiva. Entretanto, a etiquetagem é uma abordagem para indexação já consolidada na *web*, e é usada pelos maiores *sites* da Internet. Alterar ou trocar a etiquetagem por outra forma de indexação pode significar um risco para os administradores dos *sites* que já obtêm resultados satisfatórios com a abordagem atual. Assim, é preferível manter a etiquetagem e disponibilizar uma forma opcional de indexação. Um exemplo é o *site Youtube*, que permite a indexação tanto através da etiquetagem, quanto através de categorias já pré-determinadas.

Outra opção é manter a etiquetagem e permitir, de forma alternativa, a indexação através da seleção de múltiplos termos de uma taxonomia facetada. Essa abordagem pode trazer para o usuário uma maior facilidade no ato da indexação, além de permitir, na saída do sistema, que o usuário navegue por meio de tal estrutura facetada, o que traz ganhos na recuperação da informação.

Partindo dos pontos levantados por estudos anteriores e questões apontadas em trabalhos de outros autores, a presente pesquisa tem como expectativa trazer respostas às seguintes questões: Como melhorar a qualidade da indexação realizada pelos usuários em um catálogo *web*? A presença de uma taxonomia facetada facilita ou guia os usuários na indexação de registros em um catálogo *web*? Como envolver colaborativamente os usuários na indexação de registros de através de uma taxonomia facetada? Ao disponibilizar a taxonomia facetada para os usuários, qual é a influência na forma como eles realizam a etiquetagem? Quais os ganhos que essa combinação, utilização da etiquetagem e da taxonomia facetada, traz na recuperação da informação?

1.2 Objetivos

1.2.1 Objetivo geral

Propor um modelo de colaboração para a indexação e busca de registros em um catálogo *web* facetado.

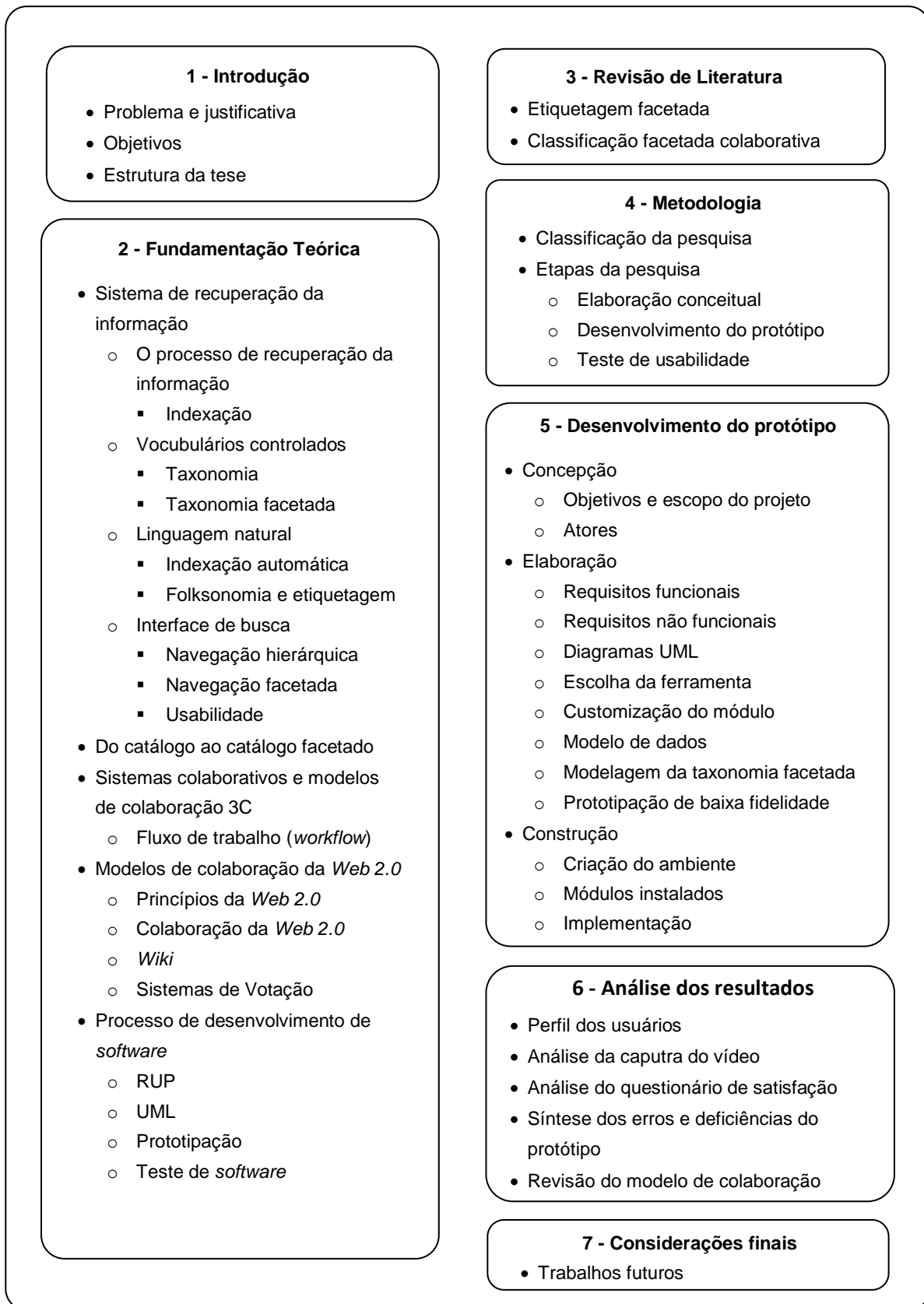
1.2.2 Objetivos específicos

- 1) Verificar a preferência do usuário no processo de indexação por meio da etiquetagem ou através de seleção de termos de uma taxonomia facetada;
- 2) Identificar a estratégia de busca preferida em uma interface com navegação facetada e busca por palavra-chave.

1.3 Estrutura da tese

O capítulo 1 é a introdução, contendo a contextualização, a justificativa, o problema e os objetivos da pesquisa. O capítulo 2 é a fundamentação teórica, apresentando os conceitos e fundamentos que serão usados neste trabalho. O capítulo 3 apresenta uma revisão de literatura com estudos similares ao presente trabalho. O capítulo 4 descreve a metodologia utilizada no desenvolvimento da proposta e o teste para sua validação. No capítulo 5, descreve-se o desenvolvimento do *software* e o plano de tese para avaliação do protótipo realizado. No capítulo 6, são apresentados os resultados da avaliação do protótipo e uma interpretação dos resultados. E, finalmente, no capítulo 7, são feitas as considerações finais relacionando com os objetivos iniciais da pesquisa, bem como são apresentados novos caminhos que esta pesquisa pode seguir, o apontamento de melhorias e aplicações em outras áreas. Este trabalho apresenta, ainda, apêndices que complementam com informações necessárias ao entendimento dessa pesquisa (FIG. 1).

FIGURA 1 - Estrutura dos capítulos da tese

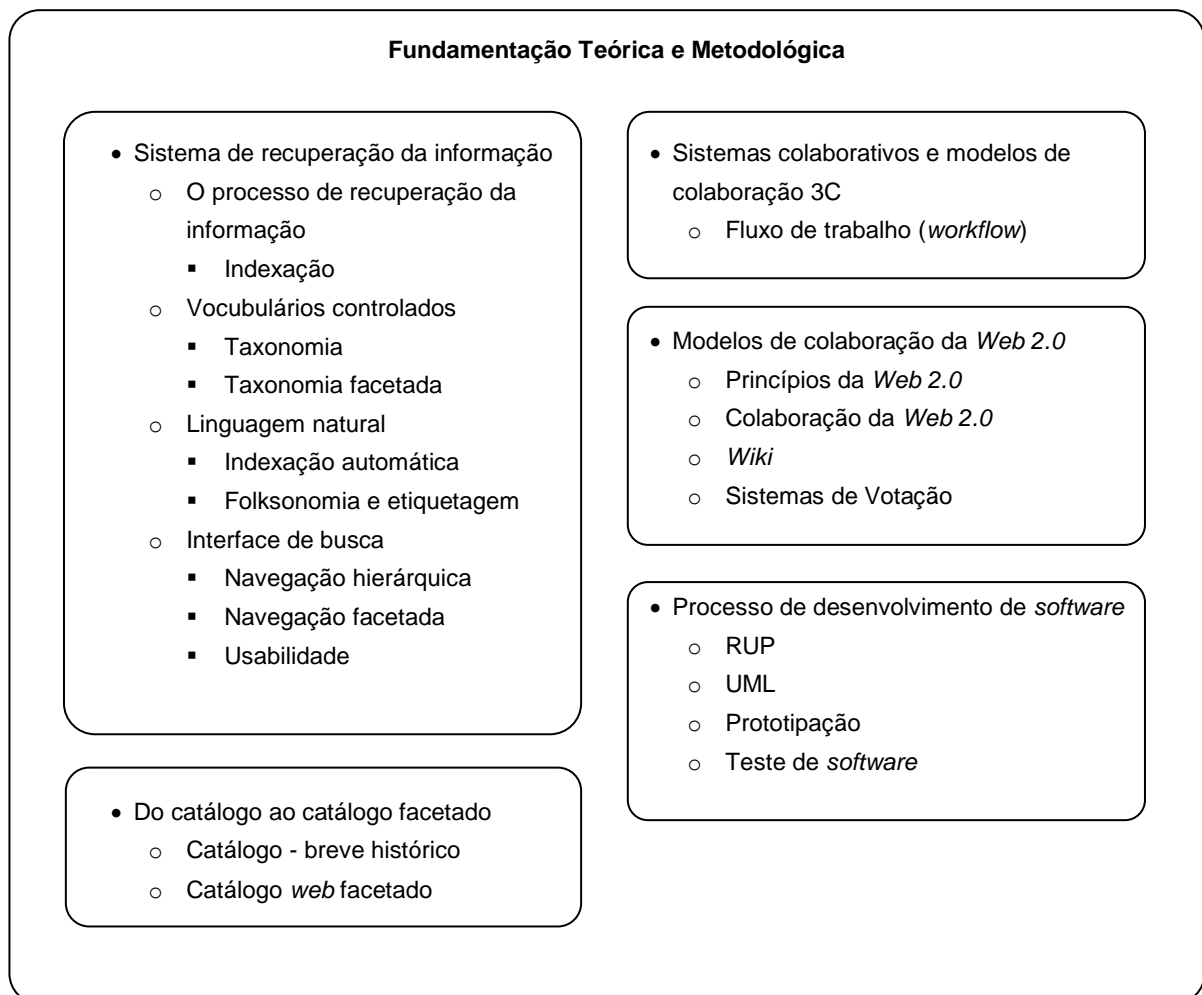


Fonte: Elaborado pelo autor.

2 FUNDAMENTAÇÃO TEÓRICO-METODOLÓGICA

A fundamentação teórico-metodológica (FIG. 2) deste trabalho foi dividida nas seguintes seções: recuperação da informação, com o processo de recuperação e a indexação; vocabulário controlado, com a taxonomia e a taxonomia facetada; linguagem natural, com a indexação e a folksonomia; interface de busca, com a navegação hierárquica e facetada; catálogo, até sua evolução para o catálogo *web* facetado; modelos de colaboração da *Web 2.0*, com o *wiki* e o sistema de votação; e o processo de desenvolvimento de *software*, com o processo unificado da *Rational* (RUP), linguagem de modelagem unificada (UML) e teste de *software*.

FIGURA 2 - Estrutura do capítulo de fundamentação teórica e metodológica



Fonte: Elaborado pelo autor.

2.1 Sistema de recuperação da informação

O termo *recuperação da informação* pode ser compreendido de diversas maneiras. A recuperação da informação é um campo de estudo investigado tanto pela BCI quanto pela ciência da computação. Manning, Raghavan e Schütze (2009, p. 1) afirma que a recuperação da informação pode ser entendida como "encontrar o material (geralmente documentos) de natureza não estruturada (geralmente texto) que satisfaz uma necessidade de informação dentro de grandes coleções (geralmente armazenadas em computadores)". Mas a recuperação da informação é ainda mais abrangente, como o próprio autor anterior aponta, pois, além de cobrir a recuperação de documentos, permite aos usuários navegar ou filtrar documentos.

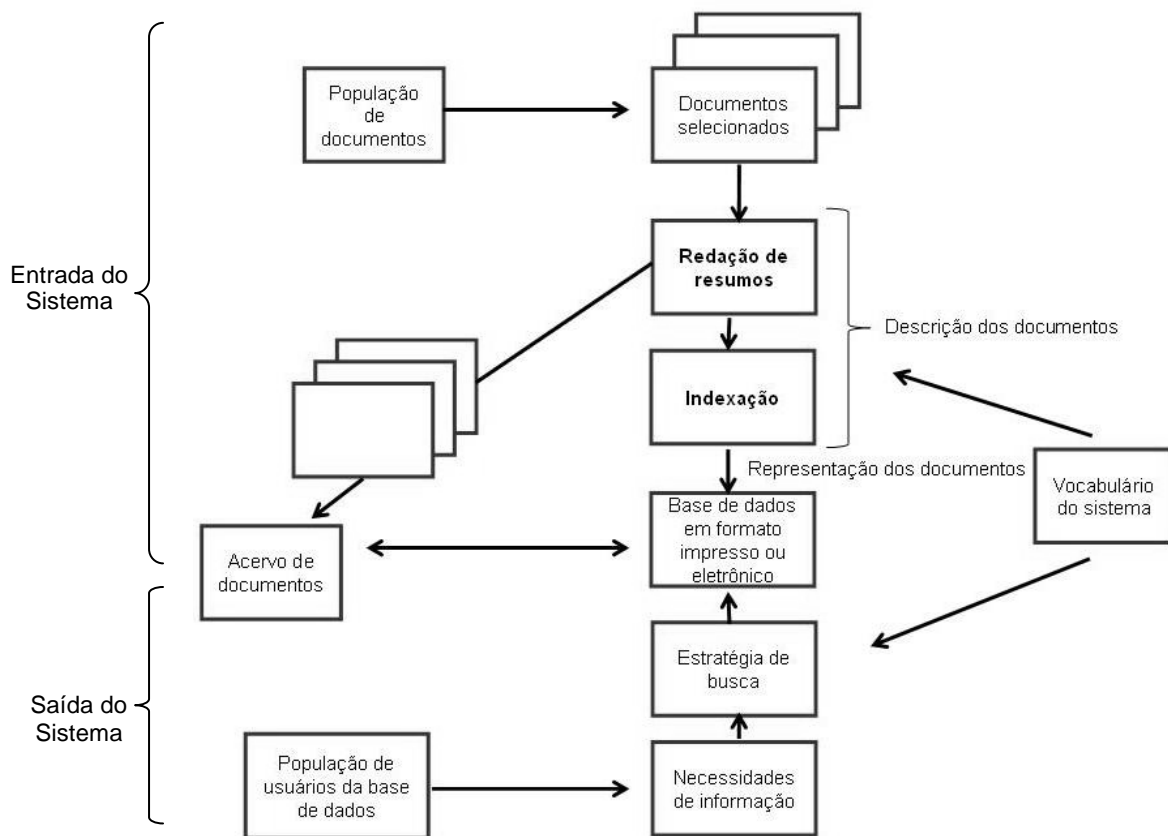
Já um sistema de recuperação da informação (SRI) é definido na norma NISO (2005) como um equipamento associado a um conjunto de operações que consegue selecionar e apresentar registros em resposta a uma requisição do usuário. Segundo Baeza-Yates e Ribeiro-Neto (1999, p. 1) "o objetivo principal de um sistema de IR é recuperar informações que podem ser úteis ou relevantes para o usuário". Para Dias (2001, p. 1), o objetivo desse mecanismo é "de facilitar o acesso à informação necessitada por uma determinada comunidade de usuários".

Manning, Raghavan e Schütze (2009) expõe que os sistemas de recuperação da informação podem ser divididos pelas escalas em que atuam: na *web*, que é a maior e mais volumosa escala, em que existem bilhões de documentos distribuídos em milhões de computadores, cujo principal desafio é fazer com que o processo seja eficiente (ex. fornecendo rapidamente uma resposta) e eficaz (ex. trazendo resultados relevantes); o nível pessoal, que é o outro extremo da recuperação da informação, em que há uma grande variedade de tipos de documentos, programas (como o *e-mail*) e sistemas operacionais (como o *Linux*, *Windows* e *Mac OS*), no qual a tarefa de recuperar a informação deve ser pouco dispendiosa, se possível gratuita, e prejudicando minimamente o desempenho do computador pessoal; e o nível empresarial ou de um domínio específico, voltada para coleções específicas, como documentos de uma instituição, um repositório de artigos de uma área do conhecimento ou uma base de dados do governo.

2.1.1 O processo de recuperação

O processo de recuperação da informação é frequentemente dividido em duas partes: a entrada do sistema, que trata da formação da coleção e organização dos documentos (indexação); e a saída do sistema, na qual o usuário acessa uma parcela dos documentos mediante uma estratégia de busca. Lancaster (1993) fez um diagrama genérico para esquematizar a recuperação da informação, tanto em uma base de dados impressa quanto eletrônica (FIG. 3). Foi feita uma adaptação para destacar as duas fases.

FIGURA 3 - Etapas da recuperação da informação

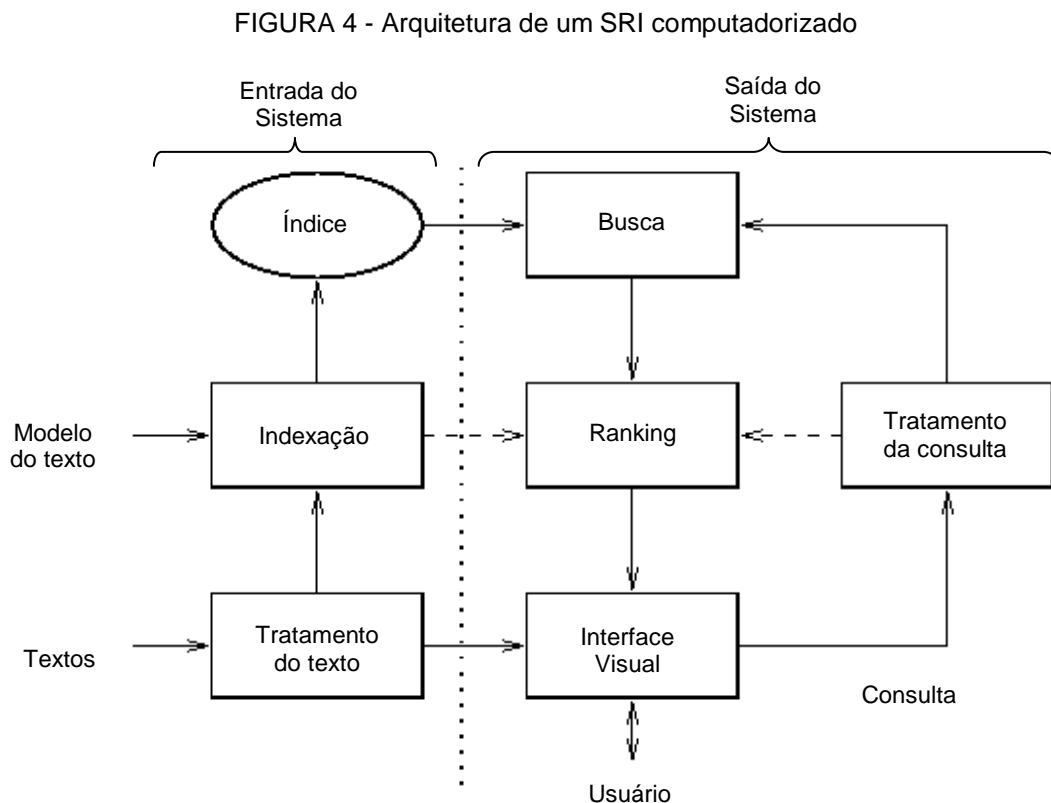


Fonte: Adaptado de Lancaster (1993, p. 2).

Na entrada do sistema, tem-se as seguintes fases: 1) a seleção dos documentos, no qual apenas os que atendem a um determinado critério ou pertencem a uma área específica são selecionados para serem incluídos na base de dados; 2) a representação dos documentos, na qual é realizado um resumo do mesmo (que geralmente apresenta extensão menor que o objeto original); e 3) sua indexação, na qual são extraídos os assuntos e outros metadados como o seu origem, língua e o tipo do documento. Esses metadados são armazenados na base de dados ou índice, que será usado para a efetiva

recuperação da informação na fase seguinte. Na saída do sistema, a etapa mais importante é a estratégia de busca, na qual o usuário deve expressar sua necessidade de informação, que será o ponto de partida para a seleção dos documentos da coleção. A estratégia de busca é uma etapa vital do SRI que permite encontrar, em meio a uma vasta coleção, apenas aqueles documentos úteis para o usuário, que podem sanar uma determinada necessidade de informação. Na FIG. 3 é possível perceber que o vocabulário controlado é um instrumento que pode servir a ambas as fases.

Baeza-Yates e Ribeiro-Neto (1999) apontam uma arquitetura genérica de um SRI computadorizado voltado para uma coleção de documentos textuais. Destacam-se, novamente, as duas fases do sistema, a entrada e a saída. É interessante perceber no esquema proposto pelos autores, e apresentado na FIG. 4 a seguir, que existem blocos comuns ao esquema anterior, tais como o índice e a indexação, e outros que foram adicionados, ou destrinchados, como a interface visual (que formula a consulta e apresenta os resultados), as operações na consulta (que trata a consulta antes de entregá-la para a busca), a busca e o *ranking* (que percorre o índice e ordena os resultados com base em algum modelo de recuperação).



Fonte: Adaptado de Baeza-Yates e Ribeiro-Neto (1999, *on-line*).

Nas próximas seções, são detalhados blocos que fazem parte de um de SRI: a indexação, com os vocabulários controlados (taxonomia e taxonomia facetada) e linguagem natural (indexação automática e a folksonomia); e a interface de busca (navegação hierárquica e facetada).

2.1.2 Indexação

Para Pinto (1985, p. 170) "a indexação consiste na indicação do conteúdo informativo de um documento através da determinação de um ou mais termos que representarão esse conteúdo." A norma NISO (2005, p. 12) apresenta uma definição mais ampla, na qual a indexação "cobre qualquer sistema ou procedimento em que o seleção e organização dos termos exige decisões intelectuais humanas, em algum momento no processo".

Nesse sentido Naves, aponta que:

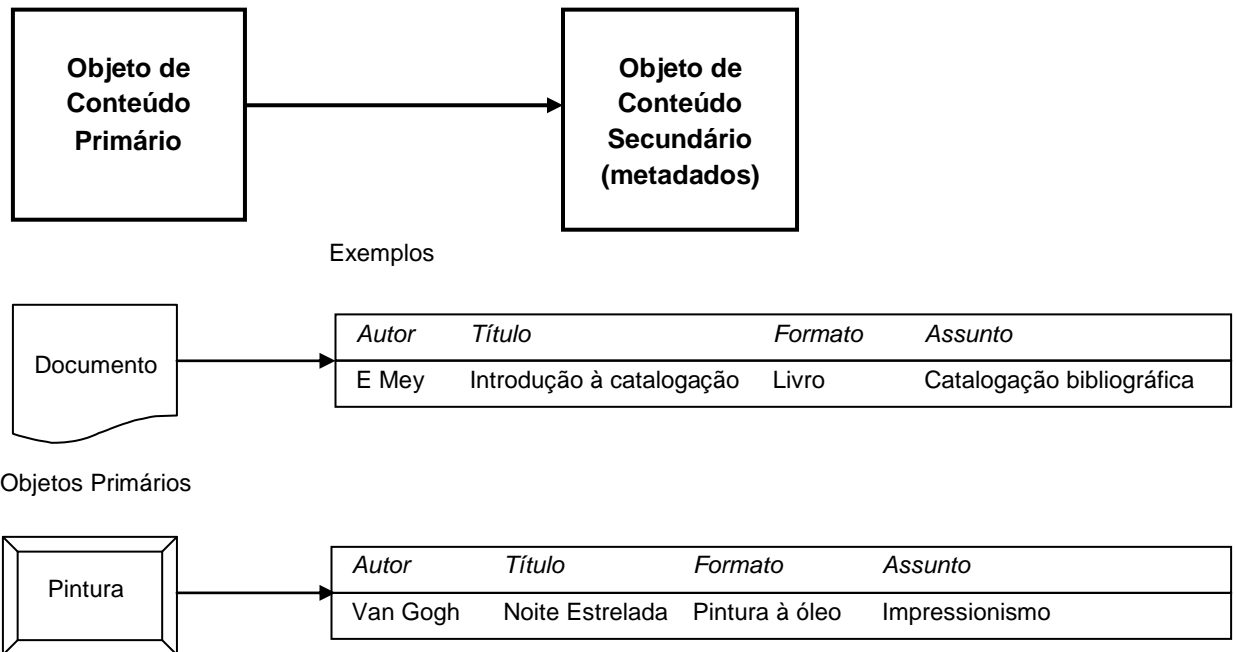
No campo do tratamento da informação, o termo indexação apresenta dois sentidos: um mais amplo, quando se refere à atividade de criar índices, seja de autor, de título, de assunto, tanto de publicações (livros, periódicos), quanto de catálogos ou de banco de dados, em bibliotecas ou centros de informação. O outro sentido, mais restrito, refere-se à indexação, classificação ou catalogação de assuntos das informações contidas em documentos (NAVES, 2004, p. 3).

A indexação também pode ser definida como "um conjunto de procedimentos com objetivo de expressar/representar o conteúdo de documentos pelas linguagens de indexação ou documentárias, visando à recuperação posterior" (GUEDES, 2010, p. 35).

O processo de indexação (FIG. 5) envolve a criação de uma representação (objeto de conteúdo secundário) de algum item (objeto de conteúdo primário), sendo que o item pode ser qualquer tipo de documento, físico ou eletrônico (NISO, 2005). No contexto deste estudo, o item pode ser qualquer objeto do mundo real ou abstrato, e sua representação é realizada por meio da criação de um registro.

Pode-se fazer a indexação dos metadados bibliográficos do item, com a especificação do autor, título e formato (FIG. 5) e dos assuntos do item. Segundo Lancaster (1993), a indexação de assuntos do item, realizada por meio de um processo manual, tem duas etapas: a análise conceitual, que é a extração dos assuntos principais do documento; e a tradução, na qual são escolhidos termos dos vocabulários controlados para a representação dos documentos.

FIGURA 5 - Representação de objeto de conteúdo



Fonte: NISO (2005, p. 12)

A próxima seção trata da construção de vocabulários controlados e da indexação manual realizada por meio deles.

2.2 Vocabulários controlados

A indexação por atribuição, realizada por seres humanos, é a representação dos assuntos do documento por meio de termos de um vocabulário controlado (LANCASTER, 1993). De posse desse conceito, é necessário diferenciar *controle de vocabulário* de *vocabulário controlado*. O controle de vocabulário é considerado um processo que apresenta os seguintes objetivos principais: melhorar a eficácia de um sistema de recuperação da informação e contribuir para uma representação ou descrição mais coesa dos itens, tais como documentos físicos, recursos *web*, entre outros (NISO, 2005). Para Smit e Kobashi (2003), o controle de vocabulário é uma forma de produzir confiança e consistência na representação.

O vocabulário controlado, por sua vez, é um instrumento que pode ser definido como um conjunto ou lista de termos autorizados (LOPES, 2002; HEDDEN, 2010;

LANCASTER, 1993). Esses termos devem "ter uma definição livre de ambiguidade e redundância" (NISO, 2005, p. 5). Diz-se que é controlado, pois, dentre os vários termos existentes na linguagem natural para designar um determinado assunto, apenas aqueles que pertencem ao vocabulário controlado devem ser usados, e a entrada de novos termos na lista obedece a algum critério ou política (HEDDEN, 2010). Os vocabulários controlados são construídos para "admitir somente uma forma de interpretação, ou seja, de significado, além de possibilitar uma maior padronização e rigor de utilização de termos" (BORGES, 2009, p. 37).

Naves explica a função do vocabulário controlado:

Normalmente, fará o controle de sinônimos e quase-sinônimos, para evitar que indexadores diferentes usem termos diferentes para expressar assuntos idênticos. Indicando o sinônimo preferido, o vocabulário evita separação de documentos similares e indica, ao usuário, os caminhos que deve seguir na busca do assunto. Irá distinguir, ainda, homógrafos, como, por exemplo, as palavras planta, manga, que têm mais de uma conotação (NAVES, 2004, p. 9).

Guedes (2010) aponta os seguintes termos como sinônimos para vocabulário controlado: linguagem documental, linguagem documentária, linguagem formal, linguagem de informação, linguagem de indexação e linguagem controlada. Pode-se acrescentar à lista o termo linguagem artificial (BORGES, 2009). No contexto deste estudo, optou-se por usar o termo *vocabulário controlado* devido à maior quantidade de referências que utilizam essa terminologia, a partir de um levantamento na base de dados *Google Acadêmico*.

Pode-se dizer que a presente escolha configura um caso de realização simples de um controle de vocabulário, que se evidencia pela utilização de um termo (no caso deste trabalho, "vocabulário controlado") preferencialmente em relação a outros (no caso do exemplo, "linguagem documental" e as outras citadas acima), a partir do critério utilizado na autorização do termo (neste caso, a maior ocorrência da utilização desse termo na literatura acadêmica). Com intuito de exemplificar o seu uso, foi criado no APÊNDICE F um vocabulário controlado para padronizar a utilização dos termos ao longo da presente tese e dirimir dúvidas de interpretação.

Apresentados esses conceitos, é pertinente destacar que existem dois tipos de vocabulários controlados (NAVES, 2004):

- não estruturados, tais como as listas de cabeçalhos de assunto (LCA), que são formatadas com os cabeçalhos de assunto em ordem alfabética;
- estruturados, em que os termos são estruturados de maneira hierárquica. Subdividem-se em:

- pré-coordenados, em que os assuntos compostos são elaborados durante a indexação. As taxonomias e os sistemas de classificação enumerativos se enquadram nessa categoria;
- pós-coordenados, cuja indexação utiliza assuntos simples, que são combinados durante a busca pelo usuário, formando os assuntos compostos, tais como os sistemas de classificação facetada e tesouros.

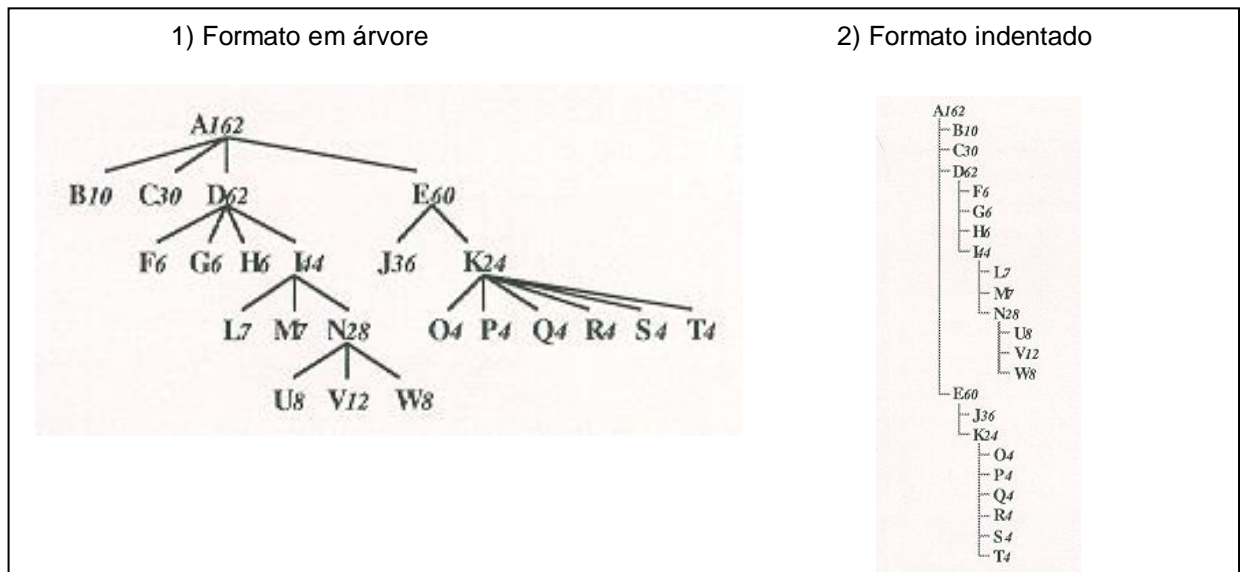
É importante dizer que o vocabulário controlado também tem outras funções relacionadas à saída do SRI. São elas (NISO, 2005): navegação, em que o usuário pode percorrer a hierarquia de termos, ajudando-o na localização dos registros com características semelhantes ou desejadas; indicação de relações entre os termos e a terminologia usada em um determinado domínio; e auxílio ao SRI na busca de registros a partir da consulta do usuário, por exemplo, estendendo a busca para termos sinônimos. A FIG. 3 (p. 33) mostra que o vocabulário controlado pode ser usado tanto na entrada quanto na saída do SRI.

Nas seções seguintes serão apresentados dois tipos de vocabulários controlados: a taxonomia e a taxonomia facetada.

2.2.1 Taxonomia

A taxonomia é considerada um tipo de vocabulário controlado (BORGES, 2009). Na norma para construção de vocabulários controlados, a taxonomia é definida como um "vocabulário controlado constituído de termos preferenciais, que são ligados numa hierarquia ou polihierarquia" (NISO, 2005, p. 18). Em uma taxonomia existe um ou vários termos na sua raiz, que são conectados com um ou mais termos inferiores, organizados em uma única estrutura hierárquica (HEDDEN, 2010). A hierarquia é geralmente visualizada, ou graficamente representada, através de uma árvore invertida (FIG. 6.1) ou uma lista indentada (FIG. 6.2).

FIGURA 6 - Formas de visualização de uma hierarquia



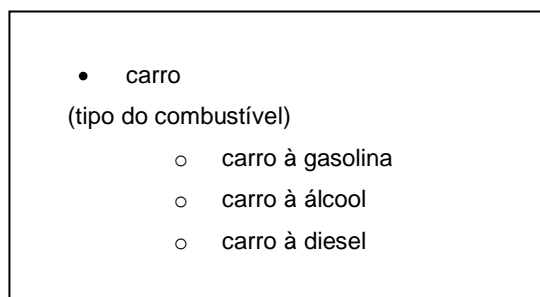
Fonte: Adaptado de JOHNSON e SHNEIDERMAN (1991).

Em uma taxonomia, as relações hierárquicas típicas entre a classe e a subclasse são: de gênero e espécie, ex. País/Federação; relação de classe e instância, ex. País/Austrália; e relação de todo e parte, ex. Continente/País ou Oceania/Austrália (NISO, 2005).

Na relação de gênero e espécie, os indivíduos, seres ou objetos que possuem determinados atributos em comum pertencem à mesma classe. Uma subclasse, que é a espécie, representa o subgrupo de indivíduos que possuem pelo menos um atributo a mais, que não é comum a todos os indivíduos da classe superior (gênero). O tipo do atributo ou característica que forma as subclasses é conhecido como diferença (ARAUJO, 2005).

A FIG. 7 abaixo, apresenta um exemplo de taxonomia, a característica usada na formação das subclasses é o tipo do combustível, e os atributos que formam as subclasses são gasolina, álcool e diesel. As subclasses formadas são carro à gasolina, carro à álcool e carro à diesel.

FIGURA 7 - Relação gênero-espécie



Fonte: Elaborado pelo autor.

Na próxima seção aborda-se a taxonomia facetada, vocabulário controlado que se torna cada vez mais comum em sites na *web* hoje em dia.

2.2.2 Taxonomia facetada

Elaborada a partir da década de 1930 por Shiyali Ramamrita Ranganathan, a Teoria da Análise Facetada - TAF possui "46 cânones, 13 postulados e 22 princípios, e está apresentada, basicamente, em cinco obras: *Five Laws of Library Science*, 1931, *Colon Classification*, 1933, *Prolegomena to Library Classification*, 1937, *Philosophy of Book Classification*, 1951, e *Elements of Library Classification*, 1962" (MACULAN, 2011, p. 48).

A TAF quebrou paradigmas na BCI, trazendo uma nova abordagem para a área de classificação e, ainda hoje, é um propulsor de pesquisas, transbordando para outros campos, como o da Internet, das ontologias, interfaces, entre outros.

Através da análise facetada é possível elaborar um sistema de classificação facetada que foi proposto como uma alternativa aos tradicionais sistemas de classificação enumerativos da época, tais como a Classificação Decimal de Dewey (CDD), desenvolvida por Melvil Dewey (1851–1931) em 1876, e a Classificação Decimal Universal (CDU), criada pelos bibliotecários belgas Paul Otlet e Henri la Fontaine em 1905. Esses sistemas de classificação enumerativos são reconhecidos por serem grandes hierarquias e frequentemente precisam ser adaptados e modificados. Em um sistema de classificação enumerativo, os assuntos gerais são subdivididos em assuntos mais específicos, gerando um grande número de assuntos compostos (TRISTÃO; FACHIN; ALARCON, 2004).

Ranganathan "idealizou uma estrutura mais elástica do que as já existentes, isto é, que permitisse a síntese, o agrupamento de vários componentes para a especificação do assunto" (ARAUJO, 2005, p. 71). Na classificação facetada as entidades são representadas através de várias dimensões ou perspectivas, denominadas facetas (PONTES, 2013). Os termos presentes nas facetas são assuntos simples. Os assuntos compostos são gerados durante a navegação com a combinação de termos de diferentes facetas em um processo chamado síntese (BATISTA; CARVALHO, 2003). Devido a esse fato, a classificação facetada é considerada uma linguagem controlada pós-coordenada, diferentemente de sistema de classificação enumerativa, que é considerado pré-coordenado.

A diferença entre uma classificação enumerativa e uma classificação facetada é apresentada na FIG. 08. É possível notar na classificação enumerativa, a existência de assuntos compostos tais como *respiração de animais aquáticos*, enquanto na classificação

facetada esse assunto é decomposto nas facetas processo e habitat animal. Isso torna o sistema de classificação facetado mais enxuto, pois conforme Lima (2004, p. 79) aponta "o número de assuntos específicos que podem ser selecionados numa classificação é infinito". Como pode ser observado na FIG. 08, uma classificação facetada é mais escalável e pode ser mantida mais facilmente. Geralmente, a faceta é menor e mais compacta, pois não precisa combinar as classes para obter os assuntos compostos. A adição e a remoção de termos, bem como mudanças estruturais, são mais fáceis de fazer e de implementar. A adição de novos termos nas facetas implica a possibilidade de novas combinações de termos para objetos indexados no domínio. Da mesma forma, ao apagar ou renomear um termo, tem impacto apenas naquele nó da faceta, enquanto em uma taxonomia tradicional seria necessário apagar ou alterar vários termos na estrutura em árvore (TZITIKAS *et al.*, 2002).

FIGURA 8 - Classificação enumerativa versus classificação facetada

Classificação Enumerativa	Classificação facetada
Fisiologia	(Faceta de processo)
Respiração	Fisiologia
Reprodução	Respiração
Animais aquáticos	Reprodução
Fisiologia de animais aquáticos	(Faceta do habitat animal)
Respiração de animais aquáticos	Animais aquáticos
Reprodução de animais aquáticos	Animais terrestres
Animais terrestres	(Faceta da taxonomia zoológica)
Fisiologia de animais terrestres	Invertebrados
Respiração de animais terrestres	Insetos
Reprodução de animais terrestres	Vertebrados
Invertebrados	Répteis
Fisiologia de invertebrados	

Fonte: Adaptado de Pietro-Dias (1987 citado por BATISTA; CARVALHO, 2003, p. 41).

Broughton (2006, p. 50, tradução nossa) aponta outras vantagens da classificação facetada em relação aos sistemas de classificação tradicionais:

- A capacidade de expressar de forma sintética a complexidade do conteúdo típica de documentos digitais;
- Uma sintaxe que permite uma gestão regular e consistente;
- Uma estrutura rigorosamente lógica presente em qualquer nível;
- Uma estrutura que é compatível com uma interface gráfica para navegação e formulação de consulta;
- A facilidade através da variação ou rotação da ordem de citação para permitir abordagens a partir de um número de ângulos;
- Uma estrutura e metodologia que permite a conversão para outros formatos de índice, como listas de cabeçalho de assunto e tesouros;
- Facilidade de modificar a palavra-chave da busca através do mapeamento em vocabulários controlados

Em uma classificação facetada, as facetas são os pilares da organização. As facetas são "agrupamentos de termos (...) servem para descrever o conteúdo de múltiplos ângulos, perspectivas, ou atributos" (HEDDEN, 2010, p. 8). Uma faceta pode conter uma lista simples de termos ou uma hierarquia, principalmente quando os termos são numerosos ou possuem uma relação hierárquica. As facetas são mutuamente exclusivas, ou seja, os termos que aparecem em uma faceta não podem ocorrer na outra.

Para compreender melhor a classificação facetada, apresenta-se a seguir seus conceitos principais reunidos:

Classe (...) é um conjunto de coisas ou idéias que possuem vários atributos, predicados ou qualidades comuns; Categorias (...) são as maiores classes de fenômenos, as classes mais gerais que podem ser formadas e que podem ser empregadas para reunir outros conceitos; Característica é a qualidade ou atributo escolhido para servir de base à classificação, sendo que, a partir dela, geralmente, formam-se renques e cadeias; Cadeias são séries de classes, geradas por subdivisões sucessivas, que se movem de forma descendente, de um assunto geral para um assunto específico, formando as relações hierárquicas dos assuntos; Renques são classes formadas a partir de uma única característica de divisão, e que formam uma divisão em fileira de assuntos correlatos; Termos são as representações verbais dos conceitos em uma linguagem natural; Conceito é qualquer unidade de pensamento de qualquer nível de complexidade (LIMA, 2004, p.81).

O sistema de classificação facetada foi desenvolvido inicialmente para ser usado na organização da Biblioteca da Universidade de Madras, na Índia. Ranganathan determinou que cinco facetas ou categorias fundamentais se formam em qualquer área do conhecimento: Personalidade, Matéria, Energia, Espaço e Tempo (PMEST). Assim, cada faceta representa um aspecto do domínio:

"Personalidade", que é o assunto ou objeto; "Matéria", que é a decomposição desse assunto ou objeto em diferentes atributos ou características (materiais, propriedades, qualidade); "Energia", que são as ações, atividades ou processos em relação ao assunto ou objeto investigado; "Espaço", que é a ambientação do assunto ou objeto, tal como seu lugar geográfico; e "Tempo", que geralmente é a data ou período histórico-contextual do objeto estudado, ou mesmo o ano de produção de tal conhecimento (MACULAN, 2011, p. 81).

Para exemplificar o uso do PMEST, no QUADRO 1.A é apresentada uma aplicação do PMEST na classificação de um documento hipotético, uma dissertação de mestrado, que trata da "história dos partidos do Brasil antes e após a ditadura na década de 60". Foram atribuídos os seguintes termos a cada uma das facetas: história (Personalidade), partido (Matéria), ditadura (Energia), Brasil (Espaço), década de 60 (Tempo). Esse exemplo é simples e apenas ilustrativo, no qual a representação resultou em um termo em cada faceta, mas não é necessário que todas as facetas sejam preenchidas e é possível que a mesma faceta receba mais de um termo (PONTES, 2013).

Segundo Aranalde (2009 citado por SILVA, 2010) as facetas do PMEST podem ser associadas à níveis de concretividade que se iniciam do maior, a Personalidade, depois Matéria, até o menor que é o Tempo (QUADRO 1.B). As mesmas facetas também podem ser associadas às questões fundamentais: 'por que', 'o quê', 'como', 'onde' e 'quando' (MOSS, 1964, citado por SILVA, 2010). Essas questões podem ajudar na determinação dos termos mais adequados para cada uma das facetas durante a indexação (QUADRO 1.C).

Maculan (2011) aponta que as cinco facetas do PMEST podem ser correlacionadas com as facetas determinadas pelo Classification Research Group (CRG): Tipos de Produto Final, Partes, Materiais, Propriedades, Formas de apresentação, Processos, Operações, Agentes, Espaço e Tempo. Dessa forma, o esquema facetado proposto pelo CRG pode ser visto como desdobramento das facetas propostas originalmente por Ranganathan (QUADRO 1.D).

QUADRO 1 – Mapeamento das facetas do PMEST

	Facetas	Personalidade	Matéria	Energia	Espaço	Tempo
A	Exemplo de representação temática	história	partidos	ditadura	Brasil	década de 60
B	Concretividade ²	1 (Mais concreto)	2	3	4	5 (Menos concreto)
C	Questões fundamentais ³	Por quê?	O quê?	Como?	Onde?	Quando?
D	CRG ⁴	Tipos de Produto Final	Partes Materiais Propriedades Formas de apresentação	Processos Operações Agentes	Espaço	Tempo
E	Ontologia	classe e subclasse	todo e parte atributos	relações	atributo origem	atributo data
F	Exemplo de representação descritiva	dissertação	formato pdf	publicador UFMG	BH	2011

Fonte: Elaborado pelo autor.

Para o presente autor dessa pesquisa, as facetas do PMEST também podem ser relacionadas aos conceitos elementares de uma ontologia (QUADRO 1.E): classe e subclasse (Personalidade), todo e parte / atributos (Matéria) e relações (Energia). Sendo que a origem ou localização (Espaço) e a data de criação (Tempo) são atributos especiais, presentes em objetos de natureza física (coisas) e, portanto, teriam facetas reservadas para

² Aranalde (2009 citado por SILVA, 2010)

³ Moss (1964, citado por SILVA, 2010).

⁴ Maculan (2011)

os mesmos. Nesse sentido, em uma ontologia cada tipo de propriedade pode ser usada como uma faceta.

Pontes (2013) aponta que na literatura existem basicamente duas interpretações da classificação facetada. A primeira interpretação, mais fiel a sua aplicação original, a classificação facetada é vista como um esquema para a classificação dos conhecimentos que os documentos contém. Refere-se à representação temática do documento (QUADRO 1.A). A segunda interpretação, a classificação facetada é entendida como um esquema de classificação que permite organizar um conjunto de entidades quaisquer de acordo com suas propriedades que são usadas como facetas. No QUADRO 1.F é dado um exemplo de classificação do mesmo documento disposto na linha A, a partir dessa interpretação, utilizando seus atributos. Esse exemplo refere-se à representação descritiva do documento.

O instrumento resultante da classificação facetada é um vocabulário controlado, que pode ser chamado de esquema de classificação facetada (*faceted classification scheme*) ou sistema de classificação facetada (*faceted classification system*) ou taxonomia facetada (*faceted taxonomy*). É comum na literatura o sistema de classificação facetada ter uma notação usada para representar as facetas e seus termos na organização física das entidades, no caso os livros, aplicação original desse instrumento. Na taxonomia facetada não existe a necessidade da notação. Por esse motivo opta-se em utilizar o termo *taxonomia facetada* para designar o instrumento usado ao longo dessa pesquisa. Dessa forma, é importante defini-la:

Rigorosamente, uma taxonomia facetada consiste em um conjunto finito de facetas. Cada faceta consiste de uma terminologia, que possui um conjunto finito de palavras ou termos, estruturados por meio de uma relação hierárquica. Cada faceta é desenhada separadamente e modela um aspecto distinto do domínio (TZITZIKAS *et al.*, 2002, p. 3).

Em outras palavras, em uma taxonomia facetada, os termos de mesma natureza são agrupados em facetas, que agem como critérios ou aspectos para organização dos itens. Durante a busca os termos atuam como filtros que podem ser combinados para restringir os resultados na busca.

Para Hedden (2010), a taxonomia facetada e a busca facetada funcionam melhor se os registros a serem recuperados contiverem atributos em comum. Não é necessário que todos os registros tenham os mesmos atributos, mas é recomendável usar os atributos em comum na determinação da facetas e, quanto mais semelhantes os registros, mais fácil ou maior o número de facetas que podem ser obtidas. A taxonomia facetada oferece melhores resultados em um domínio específico do que em um domínio amplo. Por exemplo, uma taxonomia facetada das obras de artes do mundo poderia ter as facetas: tipo da obra de arte (escultura, gravura, pintura); estilo (barroco, renascentista,

modernismo); localização (EUA, Brasil, França). Tal taxonomia facetada poderia englobar as pinturas produzidas por artistas brasileiros, entretanto, uma taxonomia facetada das pinturas dos artistas brasileiros poderia ter as facetas: técnica da pintura (a óleo, aquarela, guache); tipo da moldura (de madeira, de metal, vidro); origem (São Paulo, Rio de Janeiro, Bahia), que funcionariam melhor para indexar e recuperar estas pinturas.

Hedden, em seu livro, cita várias taxonomias facetadas usadas em *sites*, que são apresentadas na FIG. 9.

FIGURA 9 - Exemplos de taxonomias facetadas

<p>Para sites de comércio eletrônico:</p> <ul style="list-style-type: none"> • Tópico • Produto • Audiência • Localização • Preço 	<p>No site Perfect Car Finder :</p> <ul style="list-style-type: none"> • Preço (período) • Tipo do veículo (novo/usado) • Categoria • Fabricante • Numero de assentos • Milhas por galão (período) • Tamanho do veículo (período)
<p>No site MyRecipes.com:</p> <ul style="list-style-type: none"> • Ingredientes principais • Exclusão • Cursos • Ocasões • Cozinha • Conveniências • Métodos de cozinha • Dietas • Publicações • Patrocinadores 	<p>No site de serviços Monster:</p> <ul style="list-style-type: none"> • Estado (EUA) • Indústria • Categoria • Data da publicação (período) • Nível do cargo • Anos de experiência (período) • Nível da educação • Tipo do serviço

Fonte: Adaptado de Hedden (2010, p. 8).

Devido à complexidade da teoria proposta por Ranganathan, Hedden, afirma que a adesão a mesma foi baixa: "A *Classificação de Dois Pontos* não se popularizou, principalmente porque era muito complicada para o usuário médio" (HEDDEN, 2010, p. 8, tradução nossa). Na década de 90, Spiteri criou um modelo simplificado para classificação facetada, sintetizando os cânones e postulados de Ranganathan em sete princípios no plano das ideias e dois princípios para ordem de citação das facetas e dos focos. Em seu popular artigo publicado em 2003, *How to Make a Faceted Classification and Put It On the Web*, Denton trouxe esses estudos para o contexto atual do meio digital e Internet. Ele revisita os passos sugeridos por Vickery e o modelo simplificado de Spiteri e cria uma "receita de bolo" para realização da classificação facetada, sugerindo as seguintes etapas (DENTON, 2003):

1. Criar uma amostra representativa do domínio: selecione entidades suficientes para cobrir todas as possibilidades previstas.
2. Listar as entidades: separe as descrições em partes e frases isolando os conceitos básicos.
3. Criar as facetas: examine os termos resultantes, veja os que são mais genéricos e as categorias de alto nível que aparecem em todas as entidades. Crie as facetas como conjuntos de termos que são mutuamente excludentes e que podem ser unidos exaustivamente. É possível usar categorias do PMEST ou do CRG como ponto de partida. Faça com que cada faceta represente um aspecto das entidades.
4. Rearranjar as facetas: reexamine e reorganize os termos distribuindo-os nas facetas. Faça testes de classificação das entidades verificando se esses termos são suficientes para representá-las e caso contrário complemente as facetas.
5. Ordenar a citação: é referente a notação, utilizada para identificar as facetas e termos. Aplica-se apenas nas situações em que é preciso organizar fisicamente as entidades (tais como livros em uma biblioteca). Caso esteja elaborando uma taxonomia facetada essa etapa não é necessária.
6. Classificação: Use o instrumento elaborado para classificar as entidades do domínio e veja se está adequado para descrevê-las.
7. Rever, testar e manter: Nessa etapa é possível perceber algumas incoerências como a inexistência de termos ou até mesmo de facetas. Volte nos passos anteriores repetindo o processo até que o resultado seja satisfatório. Acrescente novos termos ou substitua-os quando julgar necessário.

As taxonomias facetadas se tornam cada vez mais frequentes em sites de comércio eletrônico, repositórios de mídia e bibliotecas digitais. Na próxima seção são apresentadas outras formas para representação dos documentos que também se tornaram muito populares na internet utilizando a linguagem natural: a indexação automática e a folksonomia.

2.3 Linguagem natural

Lancaster (1993) define a linguagem natural como a que é utilizada habitualmente na escrita e na fala. A indexação por meio da linguagem natural é caracterizada pela extração das palavras existentes nos documento para representá-lo. Também é considerado indexação por linguagem natural, quando as próprias palavras do usuário são atribuídas ao registro, em sua forma original.

A seguir, são abordados dois tipos de indexação realizados por meio da extração de termos da linguagem natural: a indexação automática, realizada mecanicamente por sistemas de recuperação; e a folksonomia, que é uma indexação manual e colaborativa, para um domínio amplo em escala global.

2.3.1 Indexação automática

O termo “indexação automática” foi introduzido por Gerard Salton, na década de 1960, com o aparecimento do computador de grande porte para grandes corporações e com fins militares. A indexação automática é caracterizada pela utilização dos termos contidos no próprio documento, extraídos de forma computadorizada (indexação por extração), em oposição à indexação manual ou por atribuição (LANCASTER, 1993).

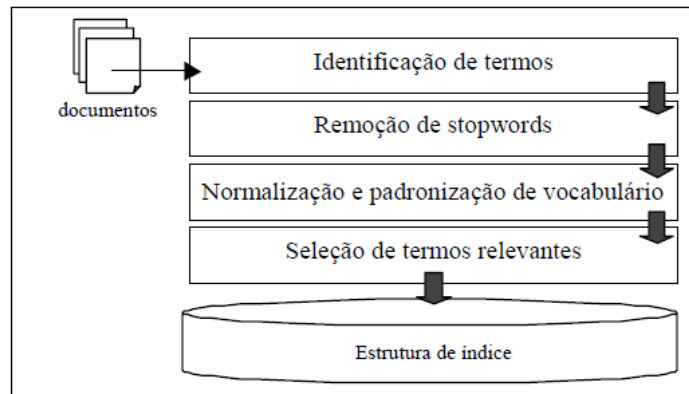
O desenvolvimento da indexação automática foi motivado, principalmente, pelos seguintes fatores: custo e morosidade da indexação manual; contínuo aumento da capacidade e barateamento da tecnologia computacional; e aumento da eficácia no tratamento de alguns problemas da recuperação da informação (BORGES, 2009).

Apesar da ausência de um especialista para fazer a seleção dos melhores termos para representar um documento, a indexação automática permite uma maior cobertura dos assuntos presentes nos documentos. A indexação automática geralmente é realizada em duas etapas (FIG. 10): o tratamento do texto, no qual é gerado um conjunto de termos a partir do texto na íntegra do documento; e a geração do índice.

A etapa do tratamento do texto, na indexação automática, pode ser dividida nas seguintes etapas: identificação de termos, quando se separam apenas as palavras referente ao texto do documento (símbolos e palavras referentes à formatação são descartados); remoção das *stop-words*, com o descarte de palavras utilizadas com o intuito de conectar as frases, tais como artigos, conjunções, preposições, entre outras; normalização e

padronização de vocabulário, com a passagem de todas as palavras para a grafia em minúsculas e no singular, remoção de acentos, entre outros processos; e a seleção de termos relevantes, descartando-se termos que aparecem demasiadamente (WIVES, 2002).

FIGURA 10 - Indexação automática



Fonte: Wives (2002, p. 41).

O segundo processo que compõe a indexação é a geração do índice a partir dos termos que serão usados na indexação, conhecidos como descritores. O índice mais comum é o arquivo invertido (*inverted file*), no qual os termos extraídos formam uma imensa lista, na qual cada termo aponta para um conjunto de documentos. A seguir será apresentado um processo recente para a indexação, principalmente de recursos *web* (qualquer documento acessível por uma URL), que se tornou bastante popular na Internet, e que, ao contrário dos tradicionais vocabulários controlados, usa a linguagem natural e não depende de especialistas, é a chamada Folksonomia.

2.3.2 Folksonomia e etiquetagem

A folksonomia é um termo criado por Thomas Vander Wal em 2004 e resulta da união do termo *folk*, que significa povo, com o termo taxonomia, ou seja, é a taxonomia criada pelo povo. Na verdade, o termo foi criado a partir do termo *folk classification*, sugerido por Eric Scheid, que significa *classificação do povo*.

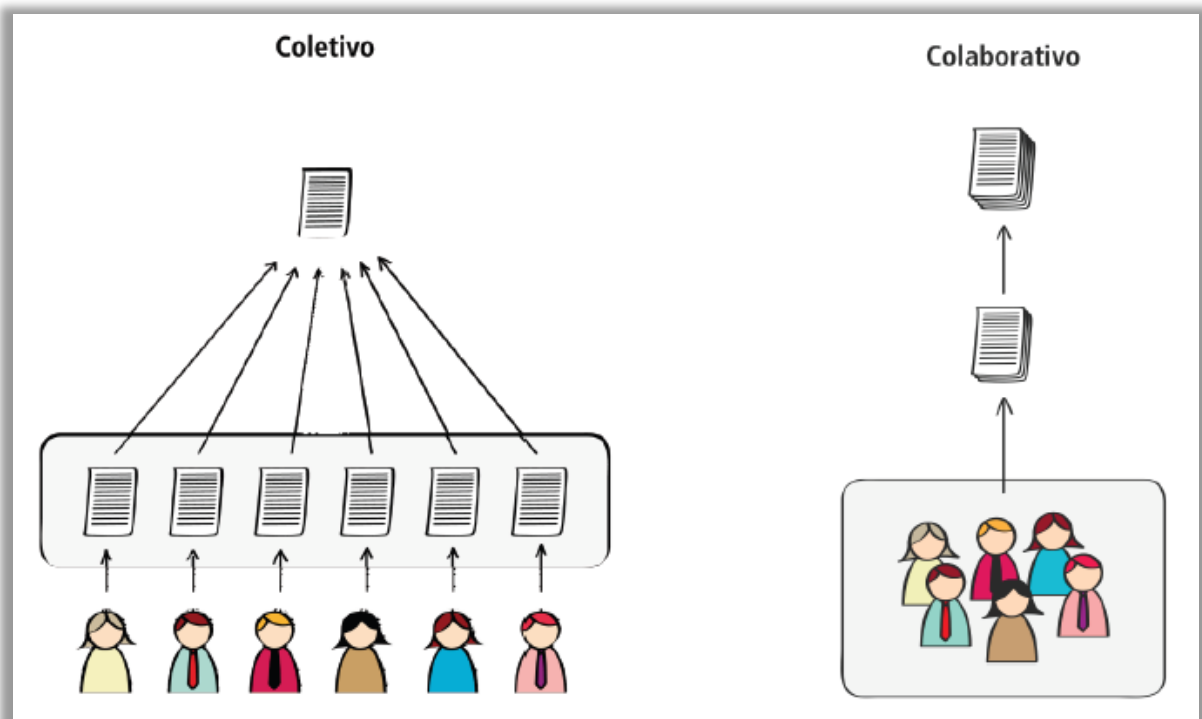
Catarino e Baptista (2007) fizeram uma extensa revisão de literatura sobre os usos do termo folksonomia e termos similares e com base nesse estudo, definiu-se folksonomia como o produto da etiquetagem de um recurso de informação, que no contexto da *web* é um recurso *web*. Um recurso *web* é qualquer item (página, imagem, vídeo, música) que tenha um URI (*Uniform Resource Identifier*). O URI pode ser uma URL (*Uniform*

Resource Locator), que é um endereço de Internet, ou um URN (*Uniform Resource Name*), que pode ser um endereço de *e-mail*.

A etiquetagem (*tagging*) é o processo ou a atribuição pessoal de etiquetas (*tags*) com o objetivo de facilitar sua recuperação. Ela representa um método de indexação livre e inovador, porque não é um profissional da indexação que indexa o recurso da *web*, e, sim, o próprio usuário (CATARINO; BAPTISTA, 2007). A etiquetagem social (*social tagging*) é considerada por alguns autores como sinônimo de folksonomia. Entretanto, no presente estudo, etiquetagem social é definida como um tipo de etiquetagem (processo) em que as etiquetas ficam publicas em um espaço compartilhado, como um *wiki* por exemplo. Outros conceitos têm significado bem próximos ao da folksonomia, tal como a classificação social (*social classification*). Há também a marcação social (*social bookmarking*), que pode ser entendido como uma criação social de favoritos, cujo precursor é também um dos pioneiros do fenômeno da *Web 2.0*, o *site Del.icio.us*, em que a comunidade de usuários integrantes do *site* compartilham entre si seus *links* favoritos (AQUINO, 2007).

Segundo Quintarelli (2005) folksonomia pode ser definida como uma "classificação gerada pelo usuário, através de um consenso emergente de baixo para cima". Entretanto, Wal (2005) expõe que diferente dos wikis que são colaborativos, as folksonomias são coletivas (FIG.11).

FIGURA 11 - Coletivo versus colaborativo

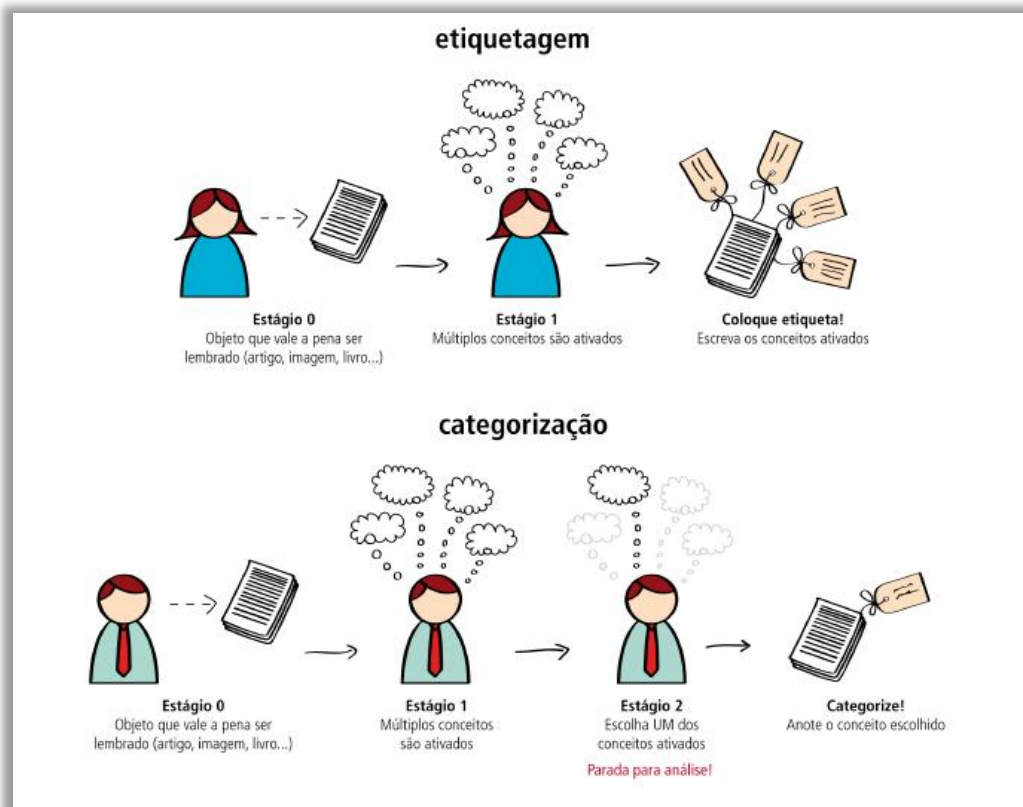


Fonte: Brandt (2009, p. 40).

A principal diferença é que a etiquetagem é um trabalho realizado individualmente que é posteriormente agregado. Na colaboração, várias pessoas se juntam para a construção de algo, com o intuito de se obter um consenso. Não é o caso das etiquetas, pois as mesmas não correspondem a uma visão consensual dos melhores termos para a representação de um determinado recurso web e sim cada etiqueta é apenas um termo que o usuário insere para que o mesmo possa recuperar aquele recurso web no futuro, sendo que essa marcação é aproveitada para outro usuário recuperar o recurso web conforme a FIG. 12 (BRANDT, 2009).

Na área de BCI, a folksonomia é o oposto da ferramenta usual: o vocabulário controlado, na qual as variações de gênero, tempo verbal, grafia do termo são definidos por especialistas da área e são usados preferencialmente em relação aos outros termos similares na indexação de um dado recurso (LOPES, 2002). Nesse sentido, Aquino (2007) aponta que a folksonomia se assemelha a um vocabulário descontrolado: qualquer usuário pode etiquetar da maneira que lhe convém, mesmo sem a garantia de uma correspondência fiel ou adequada do termo ao recurso associado. A FIG. 12 a seguir ilustra a diferença do processo de construção de uma folksonomia (etiquetagem) em relação ao processo de construção de um vocabulário controlado (categorização).

FIGURA 12 - Etiquetagem versus categorização



Fonte: Adaptado de Brandt (2009, p. 66).

Pode-se dizer que o vocabulário controlado prioriza a precisão, recuperando mais documentos relevantes na busca, enquanto a folksonomia prioriza a revocação (aumento da quantidade de documentos recuperados). De qualquer maneira, em um cenário que não é possível a existência de especialistas para a classificação de todos os recursos compartilhados, a folksonomia serve para fins de recuperação da informação, em que o usuário ao formular uma consulta consegue obter uma lista de documentos que pode sanar sua necessidade de informação.

De acordo com Wal (2005 apud WELLER 2007) uma folksonomia pode ser classificada como: larga (*broad*), quando vários usuários podem atribuir etiqueta a um recurso web; e estreita (*narrow*) quando apenas um usuário, geralmente o autor ou quem compartilha, pode atribuir etiquetas ao recurso web ou determina quem pode atribuir etiquetas ao recurso web compartilhado. Nesse sentido, a folksonomia larga faz um uso mais intenso da inteligência coletiva, permitindo que mesmo recurso seja representado por várias visões e perspectivas diferentes, refletindo em uma maior revocação (maior quantidade de resultados) em relação à folksonomia estreita.

A etiquetagem é uma indexação útil, principalmente quando aplicada em um domínio amplo (ex. *Youtube*), já que, caso fosse substituída por uma taxonomia, esta seria muito grande, causando uma dificuldade para o usuário encontrar o termo correto para sua busca. Uma taxonomia muito extensa é também onerosa para realizar a manutenção, pois a adição de novos termos pode impactar em toda a estrutura da hierarquia.

Na próxima seção, aborda-se a parte referente à saída do SRI, que é composta pela interface de busca, cujos tipos são detalhados em seguida.

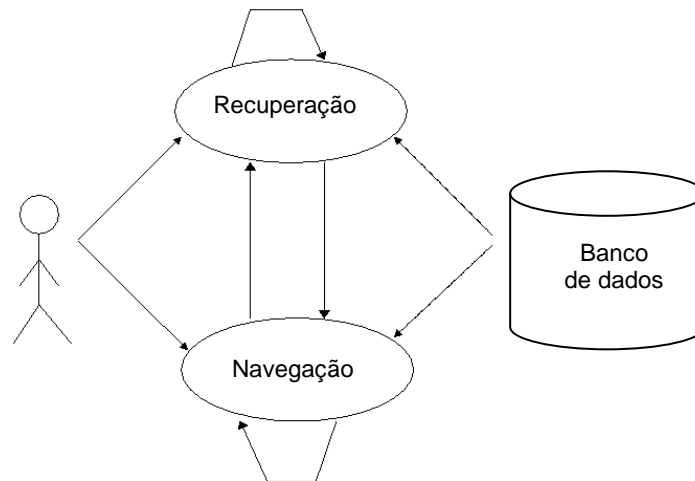
2.4 Interface de busca

Em um sistema de recuperação da informação, "o usuário precisa expressar sua necessidade de informação através de uma consulta na linguagem fornecida pelo sistema" (BAEZA-YATES; RIBEIRO-NETO, 1999). Entretanto, o interesse do usuário ou a necessidade de informação podem não estar claramente definidos no início. Os objetivos do usuário podem ir mudando durante a interação com o sistema. Baeza-Yates e Ribeiro-Neto (1999) dividem a tarefa de busca do usuário em dois tipos: recuperação e navegação (FIG. 13).

A tarefa de recuperação é conhecida como busca textual (*text search*) ou busca por palavra-chave (*keyword search*), em que o usuário, após digitar uma consulta, visualiza

os resultados relevantes, que, geralmente, se apresentam como uma lista de documentos com o título, descrição resumida e endereço ou código de localização do documento. A tarefa de navegação é uma atividade comum na Internet, sendo uma das principais vantagens do hipertexto, através dos *hiperlinks*. A navegação implica clicar e receber resultados sucessivamente até encontrar o que era desejado.

FIGURA 13 - Tarefa de recuperação da informação



Fonte: Baeza-Yates e Ribeiro-Neto (1999, p. 4).

Nota-se na ilustração da FIG. 13 que o usuário pode intercalar atividades de recuperação e navegação. Entretanto, conforme Baeza-Yates e Ribeiro-Neto (1999, p. 4) apontam, "a combinação de recuperação e de navegação não é ainda uma abordagem bem estabelecida e não é o paradigma dominante". Na época em que o livro foi escrito (1999), vislumbrava-se que esse modelo poderia ser o paradigma de busca no futuro, sendo que catálogos de bibliotecas digitais modernas (OPAC) e interfaces *web* de sistemas já começam a exibir tais recursos simultaneamente para consultas avançadas. O presente estudo compartilha dessa visão, de que é possível combinar os dois recursos, propondo uma solução que mescla as duas estratégias. A seguir apresenta-se dois tipos de navegação comuns em interfaces de busca.

2.4.1 Navegação hierárquica

A navegação hierárquica em *sites* da Internet pode ter vários formatos. A presente revisão se faz pertinente, dado o objeto deste estudo, já que um dos objetivos é construir uma interface de busca para o catálogo *web*. Nesse contexto, a FIG.14, a seguir,

apresenta a interface de um catálogo de endereços do *site* Telelistas, acessível via Internet. Pode-se observar a forma de visualização denominada de "um nível por tela", na qual o usuário vê apenas um nível hierárquico de cada vez, navegando nos níveis inferiores a cada clique (HEDDEN, 2010). Esse é um método fácil de implementar e rápido para carregar, pois não precisa conter toda a árvore de assuntos desde o início. Por outro lado, o método exige que o usuário realize mais etapas para atingir o objetivo desejado, e esse pode ser um inconveniente.

FIGURA 14 - Navegação "um nível por tela"



Fonte: *Site da Telelistas*⁵.

Outra forma de navegação hierárquica são as chamadas "migalhas de pão" (Breadcrumbs). No artigo intitulado "*Navigation Bars for Hierarchical Web Sites*" ou "Navegação em barras para sites hierárquicos", de Bowler e Shawartz (2001), elas são definidas como representações textuais da estrutura hierárquica de um *site*, caracterizadas por apresentar os nós estruturais do *site* com *hiperlinks* até a sua raiz. Esses nós são separados por um delimitador textual, geralmente o sinal de maior (>). Um exemplo desse formato é apresentado a seguir: [Home](#) > [Livros](#) > [Infantil](#) > [Até 9 anos](#).

Nota-se que esse é um recurso de fácil implementação, pois apresenta um processamento rápido e um carregamento leve da página em navegadores de Internet. Por isso, o recurso difundiu-se tanto em pequenos *sites* quanto em grandes portais. Além disso, os *breadcrumbs* atendem a requisitos que são consenso entre especialistas de usabilidade, como as heurísticas determinadas por Ben Sheidermam e Nielsen Jacob, que são: a)

⁵ Disponível em: <http://www.telelistas.net/templates/busca_assunto.aspx>.

fornecer atalhos para usuários *experts*; b) permitir a reversão de ações; c) oferecer *feedback* informativo; d) reduzir a carga de memorização; e) evitar a desorientação do usuário. Em geral, as funcionalidades responsáveis por esse conjunto de características nos *Breadcrumbs* são:

Oferecer *links* de atalho para os usuários "saltar" para as páginas visualizadas anteriormente sem usar o botão Voltar [Requisito A e B]
 Oferecer informações aos usuários quanto ao local onde eles estão localizados dentro do *site* e o caminho até a raiz [Requisitos C, D e E] (ROGERS; CHAPARRO, 2003, p. 1).

Entretanto, o objetivo do *breadcrumbs* não é ser a principal forma de navegação em um *site*, mas servir, apenas, como uma técnica de navegação alternativa e complementar à navegação principal.

Uma alternativa para a navegação hierárquica em *sites* é a lista indentada expansível, na qual, geralmente, o usuário ao clicar em um sinal de mais (+), expande o nível hierárquico, exibindo as subclasses de uma classe (FIG. 15). Esse é um método mais intuitivo, no qual o usuário pode aplicar seus conhecimentos prévios. Esse modelo se assemelha ao sistema de exibição de arquivos do computador. É também um método interativo e permite que vários níveis hierárquicos sejam exibidos ao mesmo tempo. Além disso, ele permite que o usuário reverta sua ação, comprimindo o nível expandido, a partir de um clique no sinal de menos (-). Apesar dessas vantagens, seu carregamento é mais lento que os métodos expostos anteriormente e sua aplicação requer uma experiência avançada em desenvolvimento de *sites* (HEDDEN, 2010).

FIGURA 15 - Árvore expansível



Fonte: Site do tribunal de contas de Minas Gerais.

Por sua vez, os denominados menus *dropdown* (FIG. 16) rivalizam com a árvore expansível no quesito de usabilidade, por serem intuitivos e interativos. Eles não permitem a abertura de vários níveis hierárquicos simultaneamente, mas, por outro lado, não é preciso realizar os cliques para que sejam mostrados os níveis inferiores, é suficiente que se posicione o ponteiro do *mouse* sobre o item do menu por alguns instantes para que ele seja exibido. Esse método se assemelha à árvore expansível no que se refere à velocidade do carregamento e ao conhecimento técnico necessário para sua aplicação.

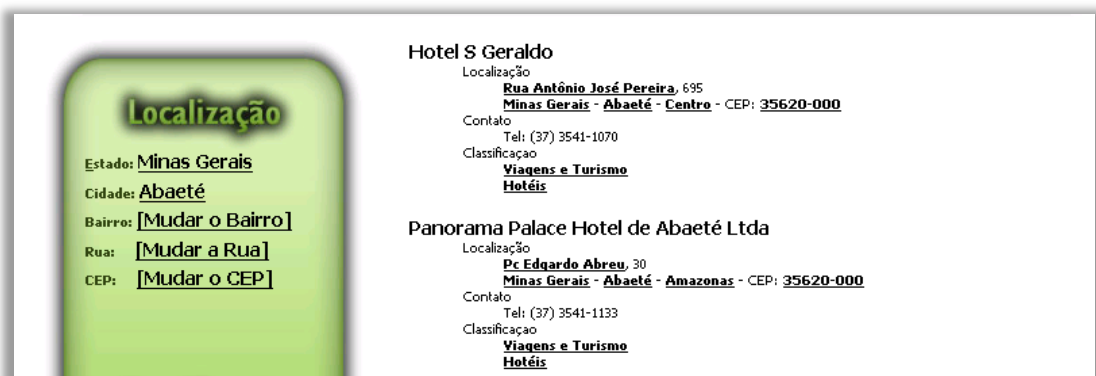
FIGURA 16 - Menu *dropdown*



Fonte: Site Filament Group⁶.

Outro recurso disponível para a visualização de dados hierarquizados são as caixas de seleção encadeadas, como a que é visualizada na interface de outro catálogo de endereços *web* (FIG. 15).

FIGURA 17 - Caixas de seleção encadeadas



Fonte: Site Axaqui⁷.

⁶ Disponível em: <<http://www.filamentgroup.com/examples/menus/flyout.php>>.

⁷ Disponível em: <<http://www.axaqui.net>>.

Nessa forma de navegação hierárquica, o usuário tem acesso, no primeiro momento, às classes do primeiro nível, que, no exemplo abaixo, corresponde aos nomes dos Estados. Após a seleção, aparecem as classes do segundo nível, que no exemplo são as cidades que pertencem ao estado selecionado. Desse modo, o usuário navega através dos níveis hierárquicos gradativamente, a partir da seleção que ele mesmo faz das classes. Nesse modelo, cada caixa de seleção representa um nível hierárquico.

Uma desvantagem desse modelo é que ele pode ser mais cansativo para o usuário, que deve esperar o carregamento das opções do próximo nível. Em contrapartida, essa abordagem pode ser implementada facilmente nos *sites*, é intuitiva e ocupa uma área relativamente pequena na interface, pois apresenta apenas a classe que foi escolhida dentro de determinado nível hierárquico. Na próxima seção, aborda-se aspectos da interface para navegação facetada e são apresentadas técnicas de visualização da taxonomia facetada, para que sejam obtidos os melhores resultados na busca dos registros.

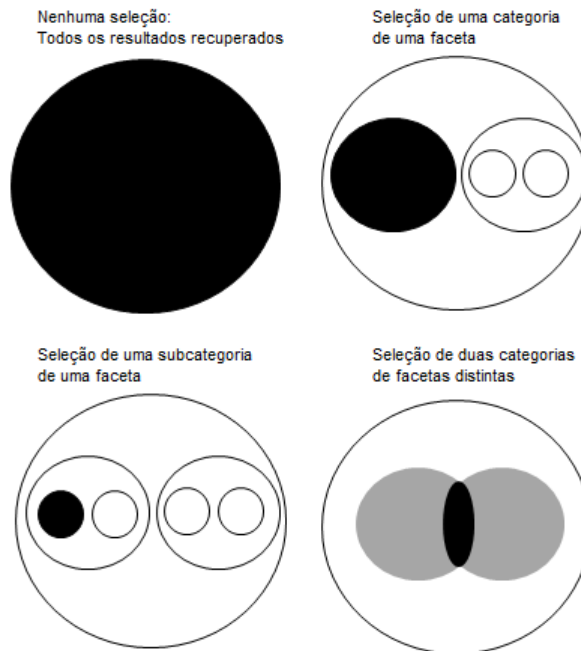
2.4.2 Navegação facetada

A navegação facetada pode ser definida como uma busca que reduz os resultados recuperados a cada iteração, pela seleção de classes pertencentes às facetas (SUOMINEN; VILJANEN; HYVÖNEN, 2007). O termo “navegação facetada” é conhecido em inglês por *faceted browsing* ou *faceted navigation*. São comuns na literatura termos como *faceted search* (busca facetada), *faceted browser* (navegador facetado), *faceted interface* (interface facetada) ou *faceted explorer* (explorador facetado). A diferença entre esses conceitos é que eles se referem: à técnica da navegação facetada (o engenho); ao ato do usuário ao realizar a navegação facetada (ação de recuperar a informação); ou à interface ou aplicativo usado para realizar a navegação facetada (instrumento).

A navegação facetada também pode ser considerada um paradigma visual de consultas e está apoiada na teoria de conjuntos (OREN; DELBRU; DECKER, 2006). Durante a busca, o sistema retorna apenas os documentos relacionados com uma ou várias classes selecionadas simultaneamente, que atuam como filtros dos resultados. Esse tipo de navegação é mediado pela interface de um sistema de recuperação de informação, daí o conceito “interface facetada”. Em uma interface facetada, geralmente são combinadas classes de facetas distintas, ou seja, ao selecionar um termo de uma faceta, a interface está assumindo que o mesmo está sendo adicionado à consulta, utilizando a operação booleana E (*AND*). Desse modo, é realizada uma interseção nos resultados, retornando para o

usuário apenas os registros relacionados a ambas as classes. A FIG. 18 demonstra o processo de busca que ocorre na navegação facetada padrão.

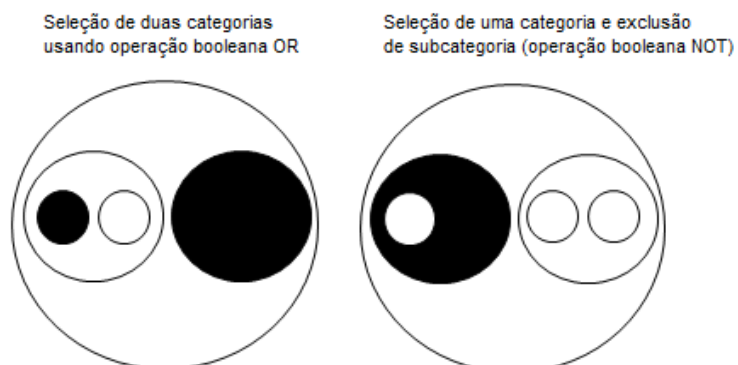
FIGURA 18 - Tipos de seleções na navegação facetada



Fonte: Elaborado pelo autor.

Interfaces com uma navegação facetada mais avançada permitem outras operações booleanas tais como *OR* (OU), que seleciona itens relacionados a um ou outro termo, e *NOT* (NÃO), que seleciona itens não estão relacionados a um termo, conforme a FIG. 19.

FIGURA 19 - Operação *OR* e *NOT* na navegação facetada



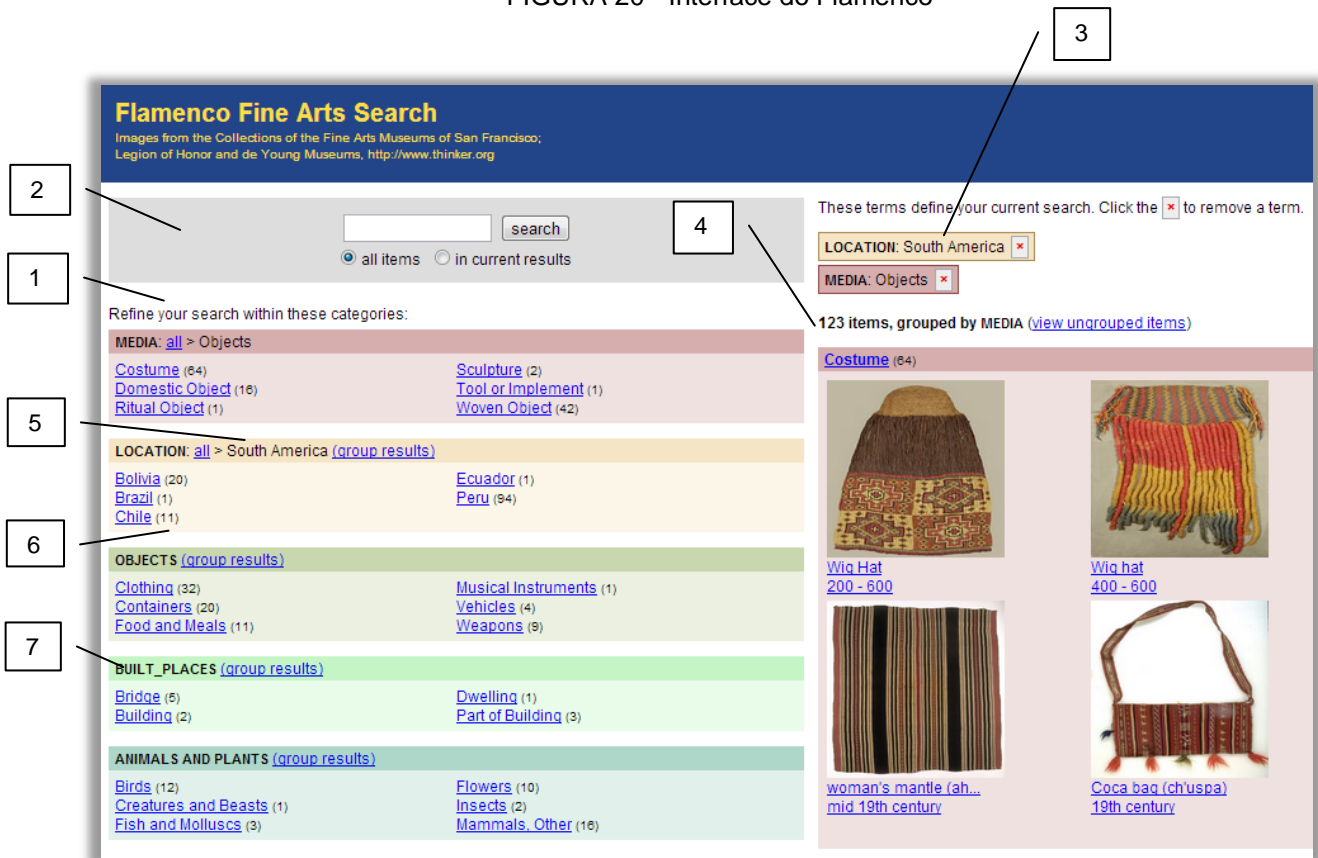
Fonte: Elaborado pelo autor.

Nas próximas seções, serão apresentados estudos e protótipos que acrescentaram melhorias e inovações às interfaces para a navegação facetada.

2.4.2.1 Flamenco

O projeto de interface de busca *Flamenco*⁸, desenvolvido por um grupo de pesquisadores da Universidade de Berkeley, foi um dos primeiros protótipos de navegação facetada na *web*. O projeto teve início em meados dos anos 2000 e inspirou o desenvolvimento de inúmeras interfaces facetadas. Tal influência se deve a sua estrutura, que oferece várias funcionalidades que foram aperfeiçoadas gradativamente, e que resultaram no protótipo apresentado na FIG. 20.

FIGURA 20 - Interface do Flamenco



Fonte: *Site do Flamenco*⁹.

As principais funcionalidades da interface do *Flamenco* estão indicadas através dos números na periferia da imagem que estão descritos a seguir:

1) As subclasses das facetas (blocos do lado esquerdo da interface) podem ser selecionados e combinados para filtrar as imagens recuperadas do lado direito;

⁸ Disponível em: <<http://flamenco.berkeley.edu>>.

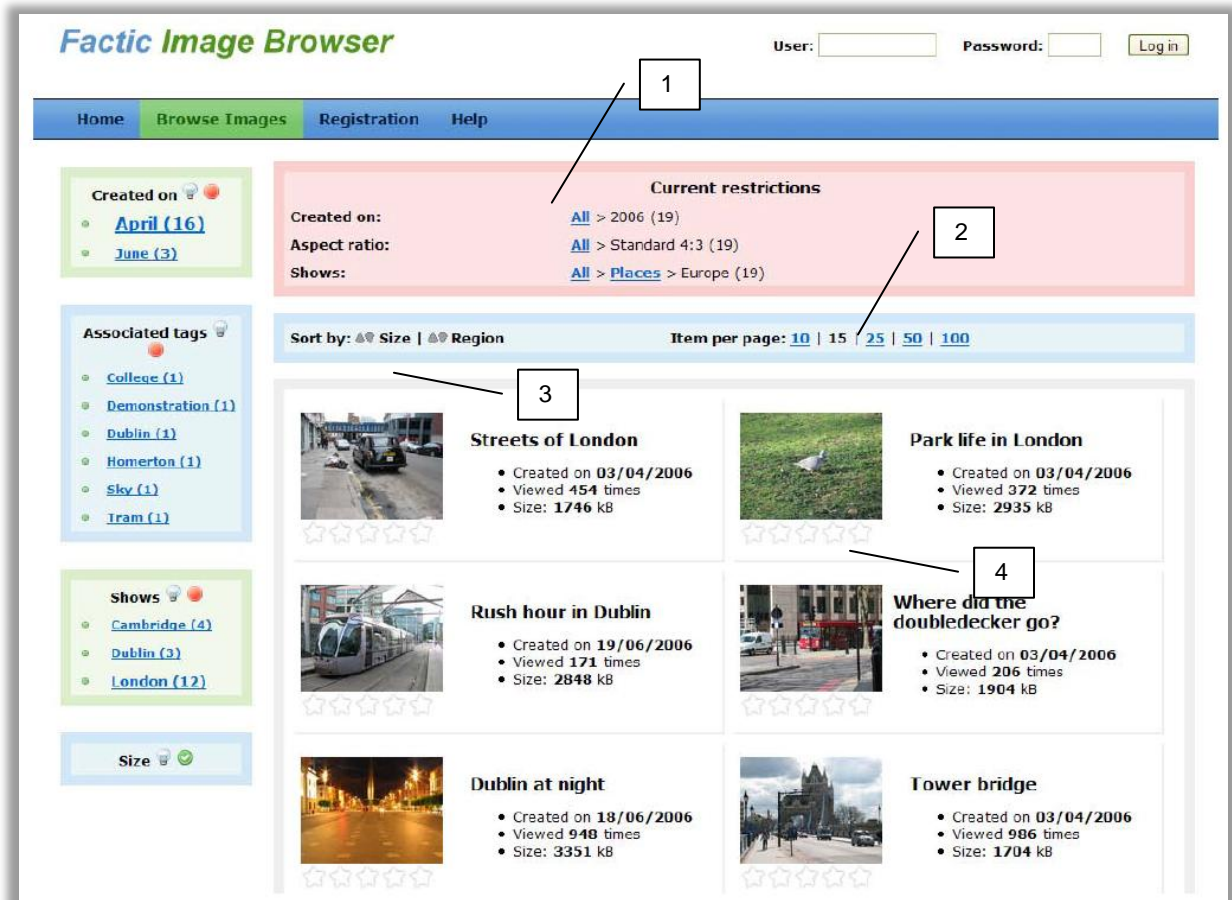
⁹ Disponível em: <<http://flamenco.berkeley.edu/demos.html>>.

- 2) É possível realizar a busca por palavra-chave em todo o âmbito do sistema ou apenas dentro dos resultados atuais;
- 3) As facetas e subclasses selecionadas são evidenciadas para facilitar a visualização;
- 4) Apresenta-se o número de registros recuperados a cada iteração, informando ao usuário se a filtragem está mais ou menos restritiva que a filtragem anterior;
- 5) São usados *breadcrumbs* para representar a hierarquia das facetas;
- 6) Em cada classe é mostrado a quantidade de registros associados, o que permite ao usuário saber as classes com mais registros, e mostra também a gradativa diminuição da quantidade a cada novo filtro;
- 7) as facetas são distinguidas através de cores, o que torna mais intuitiva sua utilização.

2.4.2.2 FACTIC

O FACTIC também é um protótipo de interface facetada. O objetivo dessa interface é realizar a recuperação da informação em grandes coleções, tal como repositórios de imagens, diretórios de *links* e redes sociais. Como princípio básico, esse protótipo usa um sistema de classificação que descreve importantes aspectos do domínio em que será utilizado. Vale explicar que estes aspectos são usados para restringir o número de recursos encontrados, uma vez que o usuário pode aumentar as restrições da pesquisa, filtrando cada vez mais os resultados obtidos (TVAROZEK; BIELIKOVÁ, 2007).

Esse protótipo adiciona as seguintes vantagens sobre as tradicionais interfaces facetadas: facetas automaticamente adquiridas no modelo do domínio utilizado; múltiplas visões para as necessidades correntes dos usuários; e modo de seleção avançado de facetas. A visualização desta interface pode ser vista na figura a seguir (FIG. 21). As funcionalidades em destaque são: 1) reunião das facetas combinadas por meio de *breadcrumbs* no painel *Current Restrictions*; 2) escolha de quantos itens o usuário vê por página; 3) possibilidade de escolha do critério de ordenação das imagens; e 4) possibilidade de o usuário emitir sua opinião quanto à qualidade do registro recuperado através da atribuição de uma nota.

FIGURA 21 - FACTIC (*Facet Image Browser*)

Fonte: Tvarozek e Bielíková (2007, p. 5).

2.4.2.3 HealthFinland

Outro trabalho relacionado à busca facetada na *web*, encontrado durante a revisão bibliográfica para o presente estudo, foi realizado pelo Grupo de Pesquisa em Computação Semântica da Universidade de Helsinki, da Finlândia. Esse grupo de estudos propôs um protótipo destinado a um Portal de Apoio à Saúde, para a busca ou navegação facetada em ontologias facetadas.

Nesse trabalho foram criadas ontologias simples, ortogonais e com estrutura hierárquica. Cada faceta é uma hierarquia e, juntas, elas resultam em uma ontologia facetada (SUOMINEN; VILJANEN; HYVÖNEN, 2007). A interface desse protótipo pode ser visualizada na FIG. 22. No lado esquerdo (1) ficam as facetas principais, hierárquicas, dispostas na vertical.

FIGURA 22 - National Health Promotion Portal

The screenshot shows the National Health Promotion Portal (tervesuomi.fi) with a faceted search interface for 'Smoking'. The interface includes a navigation menu on the left, a search bar at the top, and a main content area with facets for 'Health effects', 'Smoking in Finland', 'How to quit', and 'Law & regulations'. A search results table is displayed below the facets, and a sidebar on the right contains 'Test yourself', 'Health services', 'Research data', and 'Educational resources' sections.

Navigation Menu (Left):

- Nutrition & Food: 230
- Exercise: 178
- Weight control: 142
- Sexuality: 124
- Family & Children: 132
- Addiction & Intoxicants: 252
 - Alcohol: 56
 - Smoking: 65
 - Health effects: 34
 - How to quit: 23
 - Smoking in Finland: 17
 - Law & regulations: 19
 - Drugs: 43
 - Other addictions: 16
- Catastrophes & Epidemics: 86
- Mental health: 88
- Violence & Crises: 112
- Diseases & Symptoms: 79
- Health services: 153
- Environment: 132
- Health at work: 86

Search Results Table:

Type	Publisher	Year	Audience
any	any	any	any

Search Results:

- Tupakoinnin lopettaminen**
Tupakoinnin lopettaminen on merkittävin yksittäinen terveyttä edistävä päätös. Tupakoinnin lopettaminen on hyvin arkinen ja käytännönläheinen juttu ...
[nuuska](#) (3), [tupakan kulutus](#) (10), [Suomi](#) (12)
[Kansanterveyslaitos](#) 8.3.2004
- Suunnitteletko lopettamista?**
Suurin osa tupakoitsijoista haluaa lopettaa. Noin 40 prosenttia yrittää vuosittain päästä eroon tupakasta, mutta läheskään kaikki eivät onnistu heti. Sinnikäs
[savu](#) (14), [keuhkot](#) (10), [ympäristö](#) (12)
[Tupakkaverkko](#) 15.3.2006
- Tupakkatuotteiden kulutus**
Tupakkatuotteiden kulutus on vähentynyt 1990-luvulla, ja nykyisin suomalaiset polttavat keskimäärin noin 1300 savuketta vuodessa eli 3,5 savuketta päivässä.
[nuuska](#) (3), [tupakan kulutus](#) (10), [Suomi](#) (12)
[Kansanterveyslaitos](#) 8.3.2004

Fonte: Suominen, Viljanen e Hyvönen (2007, p. 12).

Nesse modelo, as facetas hierárquicas podem ser combinadas com facetas "secundárias", não hierárquicas, representadas por um elemento de interface conhecido como menu *dropdown*, um *widget web*, dispostas na horizontal (2). É uma abordagem interessante para a uma interface facetada, que expande o horizonte na utilização de *widgets web* para a representação das facetas.

As facetas selecionadas ficam em evidência acima das facetas secundárias, informando ao usuário a consulta que está sendo executada no momento (2). Suominen, Viljanen e Hyvönen (2007) afirmam que a navegação facetada permite separar as facetas do conteúdo indexado pela ontologia. Percebe-se, assim, que as principais vantagens nessa estratégia são: a interface de usuário torna-se mais intuitiva e útil; e os metadados da

ontologia podem ser reutilizados para diferentes casos de uso e interfaces, sem precisarem de alteração.

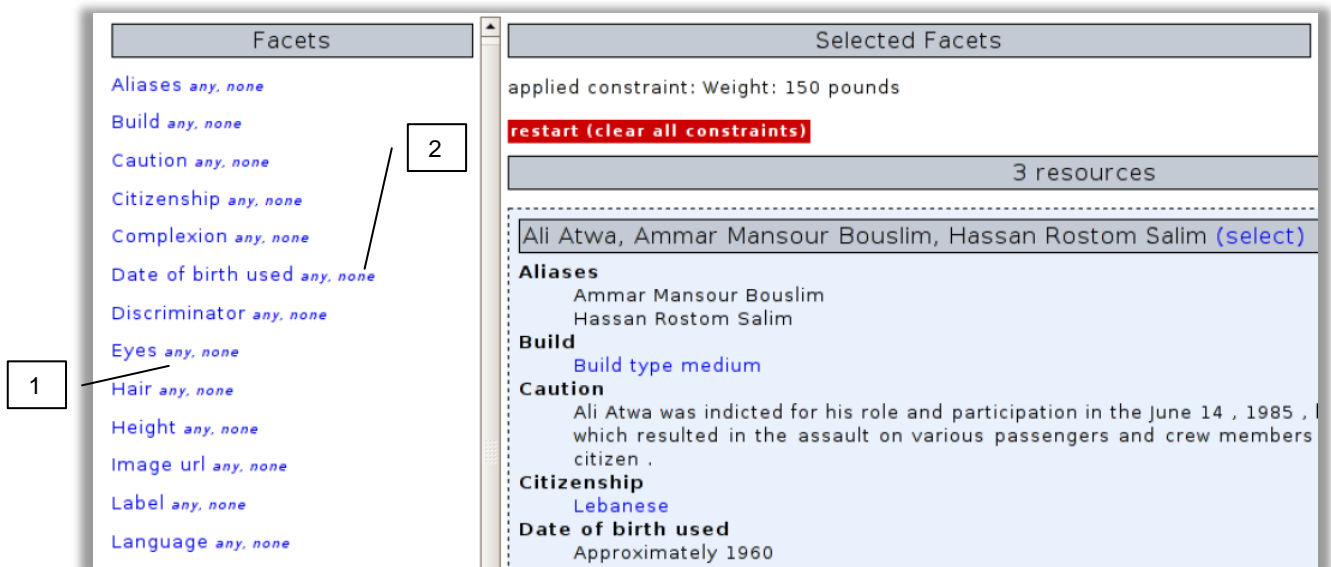
2.4.2.4 BrowseRDF

O estudo intitulado *Extending faceted navigation for RDF data*, elaborado por Oren, Delbru e Decker (2006), desenvolve um protótipo que permite a navegação facetada em uma base de dados no formato RDF, permitindo um poder maior na formulação de consultas. A navegação facetada acontece em um banco de dados do FBI, que traz os criminosos mais procurados dos EUA.

No primeiro painel, à esquerda, os usuários podem escolher as características que os criminosos devem possuir. No segundo painel, à direita, é exibida a lista de criminosos recuperados. É importante esclarecer que, quanto mais características das facetadas forem combinadas, mais restritivo fica o filtro e menos informações (criminosos) serão recuperadas.

As interfaces facetadas tradicionais têm uma capacidade limitada de formulação de consultas. Atento a esse aspecto, o estudo expande restrições ou operadores na navegação facetada. Na interface apresentada na FIG. 23, fica evidenciado como o usuário pode executar as operações booleanas *OR* (1) ou *NOT* (2).

FIGURA 23 - Navegador facetado em base de dados RDF



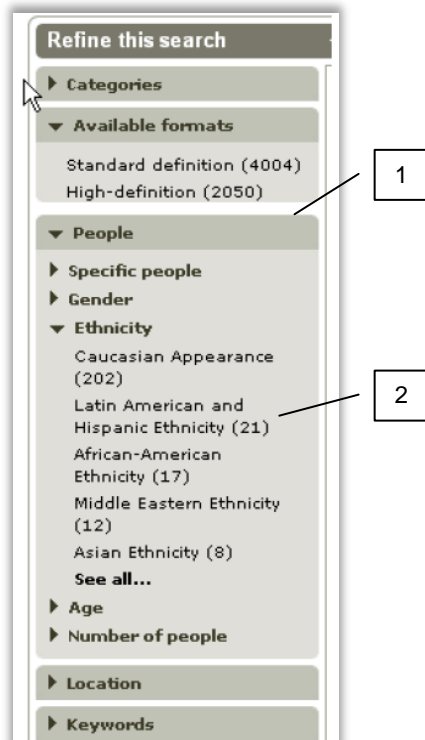
Fonte: Oren, Delbru e Decker (2006, p. 3).

2.4.2.5 Avanços recentes em navegação facetada

Hearst (2008), em seu artigo, apresenta propostas que impulsionam avanços recentes nas interfaces para navegação facetada. Seus estudos focalizam e buscam alternativas para principalmente para tratar dois problemas: opções pobres no *design* da interface podem piorar a usabilidade da interface; e a interface tem dificuldade em lidar com facetas com uma grande quantidade de classes. A primeira interface apresentada é a do *site Getty Images* (FIG. 24).

Na interface acima, as facetas são agrupadas em super-facetadas. Dessa forma, pode-se ver as super-facetadas (FIG. 24.1): categorias, formatos disponíveis, pessoas, localização, palavras-chave, estilo. E dentro da super-facetada Pessoas (*People*), vê-se as facetas: Gênero (*Gender*), Etnicidade (*Ethnicity*), Idade (*Age*), Número de pessoas (*Number of people*). Outra característica observada é que, ao invés de mostrar as categorias em ordem alfabética dentro da facetas, elas são mostradas em ordem decrescente de itens relacionados: as que apresentam mais itens são apresentadas no início da lista, seguidas das que apresentam menos itens (FIG. 24.2). Entretanto, essa interface apresenta dificuldades ao lidar com facetas muito grandes, pois desloca as demais facetas para baixo, ocultando-as da visão do usuário, que precisa rolar a página para vê-las.

FIGURA 24 - Superfacetadas do *Getty Images*

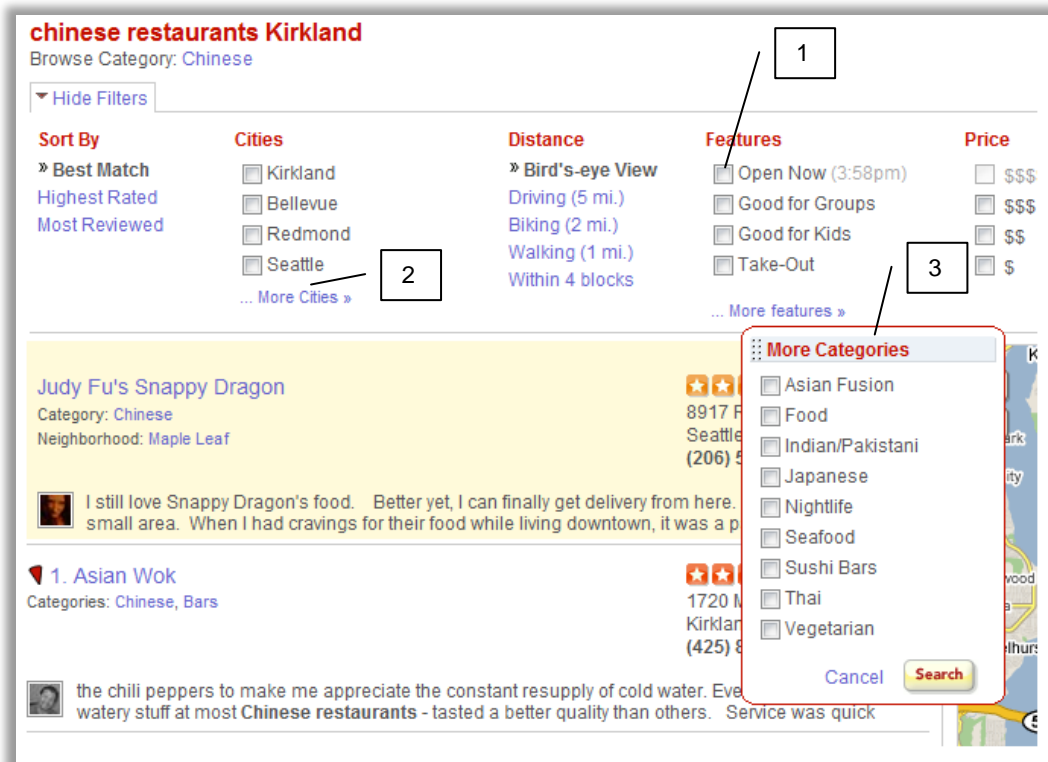


Fonte: Hearst (2008, p. 2).

Uma outra interface com inovações é apresentada na próxima imagem (FIG. 25). O *site Yelp* converte a navegação do *site* de comércio eletrônico *Ebay* em uma navegação facetada. Ela permite a seleção de várias categorias de uma mesma faceta (operação OR) através de *check-box* (1). Outra inovação é lidar com grandes facetas através do *link More* (2), que abre as demais facetas em um painel auxiliar (3), que fica sobre a interface.

A vantagem desse esquema é não aumentar demasiadamente o painel da faceta expandida, evitando prejudicar a visualização das demais. Entretanto, essa interface apresenta uma desvantagem: não é apresentada ao usuário, de forma clara, a consulta que está em vigor, ao serem realizadas múltiplas seleções na mesma faceta (usando OR), combinadas com seleções de categorias em outras facetas (usando AND).

FIGURA 25 - Interface do *site* Yelp



Fonte: Hearst (2008, p. 3).

2.4.2.6 Técnicas e padrões observados

Araújo (2009) identificou em seu trabalho, três blocos usuais em uma interface facetada: 1) facetas, que permite a navegação e seleção de classes das facetas; 2) facetas selecionadas, que na verdade permite visualizar as classes selecionadas de cada faceta e que compõe a consulta realizada; e 3) recursos filtrados, uma área em que são apresentados os registros recuperados com a atual consulta.

FIGURA 26 - Estrutura padrão de uma interface facetada



Fonte: Araújo (2009, p. 50).

A seguir, apresenta-se o QUADRO 2, em que estão sintetizadas técnicas que podem ser utilizadas na construção de cada bloco de uma interface para navegação facetada. É importante ressaltar que existem técnicas concorrentes para representar o mesmo elemento de interface, por exemplo, o formato das facetas pode ser uma lista indentada ou uma lista com um nível hierárquico por vez. Dessa forma, o *designer* da interface precisa fazer escolhas durante o desenvolvimento da navegação facetada. O objetivo do QUADRO 2 é apresentar ao *designer* as opções existentes para a elaboração de uma interface facetada, oferecendo-lhe, também, informações sobre as vantagens e desvantagens para o usuário que cada opção apresenta (não foram abordadas as dificuldades de implementação, que variam conforme o desenvolvedor).

QUADRO 2 - Opções de *design* para uma interface facetada

(continua)

Área	Técnicas concorrentes	Vantagens/Desvantagens
Facetas		
Cor	Facetas Monocromáticas	Deixam o <i>design</i> "mais limpo" e mais uniforme em relação ao <i>design</i> do <i>site</i> em que está sendo usado
	Facetas Coloridas	Ajudam a identificar que as facetas representam esquemas organização diferentes
Tamanho Fixo	Facetas com barra de rolagem automática	Permitem mostrar em um espaço reduzido uma grande taxonomia. Rolar e encontrar um item pode ser cansativo para o usuário
	Facetas com opção Mais (+)	Mostram os itens mais usados em cima e clicando em mais são mostrados todos os itens, o que pode deixar uma faceta muito grande em relação às demais
	Faceta com opção Mais e Bloco Flutuante	Técnica que mostra todos os itens de uma faceta em outro bloco flutuante, que desaparece automaticamente, sem distorcer o tamanho da faceta em relação às demais
	Faceta com agrupamento automático alfabético	Acima de certa quantidade de itens no mesmo nível hierárquico, agrupa os itens de A a Z. A desvantagem está em adicionar mais um passo na navegação

(continua)

Área	Técnicas concorrentes	Vantagens/Desvantagens
Facetas		
Tamanho Variável	Conforme o número de itens	Facetas podem ficar demasiadamente grandes em relação às demais
	Conforme preferência do usuário	O usuário tem que arrastar para aumentar o tamanho
	Com Zoom. Ao utilizar, elas expandem, trazendo elas para frente na interface	Ocupa pouco espaço na interface e permite mostrar grandes taxonomias. Entretanto, o usuário não vê imediatamente as categorias que ele pode navegar
Formato	Lista indentada expandível (estilo sistemas de arquivos)	Ideal para grandes taxonomias. Tem foco naquilo que é desejado
	Lista indentada expandida (é rígida, sempre aberta)	Permite já observar todas as categorias sem o esforço de ficar abrindo uma a uma, mas ocupa em grande espaço na interface
	Lista simples	Mais apropriado para facetas não hierárquicas. Em facetas hierárquicas, mostra apenas um nível hierárquico por vez a cada navegação.
	Menus dropdown	É intuitivo, interativo e não necessitam que a categoria seja clicada para mostrar as categorias inferiores
	Sequência de palavras	Não quebra a linha ao listar as categorias. Vai separando-as por algum caráter como uma vírgula
	Treemaps	Recentemente, esta sendo aplicada na navegação facetada em dispositivos móveis
	Caixa de seleção encadeada	Ocupa pouco espaço na interface. Aconselhável para facetas com poucos itens
Ordenação das categorias	TagCloud	Realça as categorias mais utilizadas graficamente
	Em ordem alfabética	É o padrão em sistemas de navegação
	As mais utilizadas (Ex. Top 10)	Prioriza as categorias mais utilizadas
Quantidade de recursos	Quantidade total de recursos	Mais fácil de implementar e com menos gasto computacional
	Quantidade de recursos naquela combinação de filtros	Evita que usuário clique em "categorias vazias." Tem um alto custo computacional, pois para cada novo filtro, a quantidade de todas as categorias devem ser recalculadas
Categorias Seleccionadas	Negrito ou realce nas categorias dentro da faceta	Em taxonomias muito grandes, as categorias seleccionadas podem ficar "escondidas"
	Cada categoria seleccionada é apresentada como um pequeno bloco fora das facetas	Deixa bem evidente o que está sendo filtrado
	Breadcrumbs	Mostra a categoria seleccionada e da onde ela veio.

(conclusão)

Área	Técnicas concorrentes	Vantagens/Desvantagens
Facetas		
Disposição	Na horizontal, lado a lado, acima dos resultados	Da ênfase na navegação como a principal forma de busca
	Na vertical a esquerda dos resultados	Torna a navegação como uma forma de busca alternativa, em segundo lugar.
Operação Booleana	Sempre <i>AND</i> ao selecionar as categorias	Mais simples de implementar
	<i>AND</i> , <i>OR</i> e <i>NOT</i>	É bem mais complexo de implementar e não é muito intuitivo para o usuário usar (poucos chegam a usar funções complexas)
Busca por palavra-chave	Em todos os resultados combinado a navegação facetada	Modelo padrão na navegação facetada
	Busca para cada faceta	Permite encontrar categorias em facetas com muitos itens. Entretanto, exige o esforço do usuário de trocar de campo de digitação.
	Busca separada em cada faceta com autocomplete	Recurso acima permitindo encontrar a categoria durante a digitação, o que é ótimo para o usuário. Mais difícil de implementar
Registros Recuperados		
Ordenação	Pelos mais recentes	Estimula a inserção de novos registros
	Pelos mais populares	Valoriza os registros mais ativos
	Por ordem alfabética	É o padrão em sistemas de navegação
	Pelos mais votados	Valoriza os recursos com maior qualidade
Paginação	Fixa	Mais fácil de implementar
	Definida na interface	Deixa o usuário livre para lista a quantidade de resultados que ele prefere
Formato	Em lista	Mais apropriado para recursos textuais como notícias e artigos
	Em tabela	Mais apropriado para mídia do tipo imagem e vídeo

Fonte: Elaborado pelo autor.

2.4.3 Usabilidade

É desejável que uma interface seja fácil de usar, qualidade expressada pelo termo "amigável" (*user-friendly*) bastante utilizado nas décadas de 1980 e 1990. Entretanto, não se deseja que a interface seja amiga do usuário e sim que ele a opere da melhor maneira possível. Assim, atualmente o termo usabilidade é preferencialmente utilizado para referenciar essa propriedade da interface (PEREIRA, 2011).

Conforme Nielsen (2003, p. 1), a usabilidade é definida como "um atributo de qualidade que avalia o quão uma interface é fácil para o usuário utilizar. A palavra usabilidade também se refere a métodos para melhorar facilidade de uso durante o processo de design". Assim, a usabilidade poder ser aferida através de cinco componentes de qualidade:

Aprendibilidade: O quanto é fácil para os usuários a realizar tarefas básicas de primeira vez que o usuário tem contato com a interface.

Eficiência: uma vez que os usuários tenham aprendido a concepção, com qual rapidez eles podem executar tarefas?

Memorização: quando os usuários retornam ao sistema após um período de não utilização, quão facilmente se pode restabelecer a competência?

Erros: quais são os principais erros que os usuários fazem, como graves são esses erros, e quão facilmente se pode recuperar dos erros?

Satisfação: quão agradável é a interface para ser usada?

(NIELSEN, 2003, p.1, tradução nossa).

Shneiderman, pioneiro na área de design de interfaces, aponta oito regras de ouro para interfaces *eficientes*, propostas no livro *Designing The User Interface* (SHNEIDERMAN, 1998): 1) consistência: refere-se a manter um padrão, no visual (cores, layout e fontes), na sequência de ações entre procedimentos similares, e na terminologia usada na internet, tais como menus; 2) atalhos para usuários experts: navegação, teclas de atalho, macros e entre outros, são funcionalidades desejadas por usuários experientes, pois facilitam e agilizam tarefas repetitivas; 3) *feedback*: toda e qualquer ação do usuário requer uma resposta do sistema, a qual será mais ou menos explicativa, dependendo do tipo de ação a ser executada; 4) diálogos que indiquem término da ação: as sequências de ações do sistema devem ser organizadas de tal forma que o usuário consiga entender os passos e saiba quando cada um deles for executado com sucesso; 5) prevenção e tratamento de erros: a interface não pode permitir que o usuário cometa erros graves; caso eles ocorram, deve haver mecanismos que os corrijam, na medida do possível; se essa correção estiver fora do alcance da interface, o usuário deve, ao menos, ser instruído com uma possível solução; 6) reversão de ações: sempre que possível, as ações devem ser reversíveis, de forma que tranquilize o usuário e lhe encoraje para explorar o sistema; 7) controle: os usuários mais experientes devem ter a sensação de que eles dominam os processos do

sistema e que este apenas responde a suas ações; e 8) baixa carga de memorização: o sistema deve conter uma interface simples para memorização. Para isso, requer uma boa estrutura e equilíbrio para relacionar elementos e facilitar a memorização subjetiva das telas, sem exigir esforço.

De modo a complementar essas recomendações, recorre-se a Nielsen, que elaborou uma lista com Dez Heurísticas de Usabilidade para orientar *designers* na construção de interfaces, como apresentado a seguir:

Visibilidade do status do sistema: O sistema deve sempre manter os usuários informados sobre o que está acontecendo, através de um *feedback* apropriado em tempo razoável.

Correspondência entre o sistema e o mundo real: O sistema deve falar a linguagem dos usuários, com palavras, frases e conceitos familiares ao usuário, ao invés de termos orientados ao sistema. Seguir convenções do mundo real, fazendo as informações aparecerem de forma natural e lógica.

Controle do usuário e liberdade: os usuários freqüentemente escolhem funções do sistema por engano e vão precisar de uma forma clara saída de emergência marcada "*para sair do estado indesejado sem ter que ir através de um diálogo alargado*". Suporte desfazer e refazer.

Consistência e padrões: os usuários não devem ter surpresas com o uso de diferentes palavras, situações ou ações significam a mesma coisa. Siga as convenções da plataforma.

Prevenção de erro: ainda melhor do que boas mensagens de erro é um design cuidadoso que impede que um problema ocorra em primeiro lugar. Trata-se de eliminar condições de erro ou buscar por eles. Um exemplo é apresentar uma opção de confirmação antes de se realizar a ação.

Reconhecimento ao invés de lembrar: minimizar a carga de memória do usuário, tornando os objetos, ações e opções visíveis. O usuário não deve ter que lembrar informação de uma parte do diálogo para outra. Instruções para utilização do sistema devem ser visíveis ou facilmente recuperáveis sempre que apropriado.

Flexibilidade e eficiência de utilização: aceleradores - despercebidos pelo usuário iniciante - podem agilizar a interação para o usuário avançado de tal forma que o sistema possa atender tanto usuários inexperientes e experientes. Permitir que os usuários criem atalhos para ações frequentes.

Estética e design minimalista: os diálogos não devem conter informação irrelevante ou raramente necessária. Cada unidade extra de informação em um diálogo compete com unidades de informação relevantes e diminui sua visibilidade relativa.

Ajude os usuários a reconhecer, diagnosticar e recuperar de erros: as mensagens de erro devem ser expressas em linguagem simples (sem códigos), indicar com precisão o problema e sugerir uma solução construtiva.

O manual Ajuda e a documentação: mesmo que seja melhor que o sistema possa ser usado sem documentação, pode ser necessário fornecer ajuda e documentação. Qualquer uma dessas informações deve ser fácil de pesquisar, focalizada na tarefa do usuário, com a lista de ações concretas a realizar, e não deve ser muito grande (NIELSEN, 1995, p.1, tradução nossa).

2.5 Do catálogo ao catálogo facetado

2.5.1 *Catálogo - breve histórico*

Segundo Mey (1995, p. 9) um catálogo pode ser definido como "um canal de comunicação estruturado, que veicula mensagens contidas nos itens e sobre os itens do(s) acervo(s), apresentado-as sob forma codificada e organizada, agrupada por semelhanças".

Um catálogo pode ser considerado uma ferramenta para a recuperação da informação e é um dos instrumentos mais antigos usados em bibliotecas (KAFURE, 2004). Uma prova dessa afirmação é uma lista de livros antiga, escrita em um tablete de argila, que data de 2000 a.C., que contém 62 títulos (MEY, 1995).

Quando utilizado em bibliotecas, um catálogo pode ser definido como: "um conjunto de registros que descrevem os documentos (itens) pertencentes a um acervo ou a vários acervos" (KAFURE, 2004, p. 18). Sua função primordial é permitir "encontrar um item desejado através de pontos de acesso, palavras-chave, nome do autor ou apenas o título da obra" (SANTANA, 2012, p. 45). O catálogo tradicional utiliza o papel como suporte de registro da informação e pode se apresentar no formato de um livro ou em fichas que passaram a ser utilizadas em bibliotecas no final do século XIX (CUNHA, 1999).

Os catálogos podem ser considerados fontes de informação secundárias, ou seja, são "interpretações ou avaliações de fontes primárias" (PINHEIRO, 2006, p. 2), sendo que as fontes primárias são os materiais originais, tais como livros, anais de congressos, periódicos, entre outros. Os catálogos também podem ser classificados como: fonte de informação técnica, tais como catálogos de patentes ou normas técnicas; fonte de informação para negócios, tais como os catálogos de empresas ou páginas amarelas; ou fonte de informação científica, tais como catálogos de livros ou teses e dissertações (MONTALLI; CAMPELLO, 1997).

Com o desenvolvimento tecnológico, verificou-se, igualmente, uma evolução dos catálogos. Sua forma de apresentação foi se alterando, originando diferentes tipos de catálogos, cuja história apresenta a seguinte evolução: catálogos manuais, com fichas bibliográficas escritas a mão; impressos, com listas do acervo impressas; semi-automatizados, que dispunham de mecanismos elétrico ou óticos para sua consulta; e automatizados, ou seja, armazenados e acessados pelo computador.

No início, os catálogos computadorizados ou catálogos automatizados tinham apenas a proposta de "facilitar a recuperação da informação, de um modo mais rápido e

eficiente pelos usuários" (SANTANA, 2012, p. 46). Esse objetivo foi traçado diante da realidade desses catálogos, pois, além da velocidade com que trazem os resultados, também permitem a combinação na busca, com palavras presentes no título, descrição ou palavras-chave. Entretanto, no princípio, o acesso a esses catálogos era restrito às dependências da biblioteca. Depois, com o advento das redes de computadores e da Internet, foi criado o OPAC (*Online Public Access Catalogue*), que passou a dar acesso via Internet. Os OPACs tiveram 3 gerações segundo Balby:

Na primeira geração, são considerados catálogos automatizados com pontos de acesso somente para autor e título e como forma de busca apenas a combinação exata de palavras ou frases. Na segunda geração surgem os índices, a visualização e as relações entre termos pelos usuários, registro bibliográfico em formato completo, possibilitando a expansão da busca através de *links*, agrupamentos de registros em bases lógicas e ligações hierárquicas e horizontais entre registros, entre outros. Já a terceira geração, além das características da segunda, possibilita o acesso a banco de dados comerciais, relação do registro bibliográfico com os arquivos digitais e mecanismos de busca. (BALBY, 2002 citado por OLIVEIRA, 2008, p. 47).

O nome dado a um catálogo automatizado acessível por um navegador de Internet é catálogo baseado na Internet (*web-based*) ou simplesmente catálogo *web*, ou, ainda, diretório *web*. Um OPAC acessível pela Internet é um catálogo *web*. Os catálogos *web* do *Yahoo! Directory* e o *Open Directory Project* (ODP) estão entre os mais populares da Internet. O ODP, que é um catálogo de *sites*, tem mais de um milhão de categorias e cinco milhões de *sites*. Exemplos de catálogos *web* são: catálogos de músicas e artistas, como o *musicmoz.com* e o *musicbrainz.com*; catálogos de endereços ou as chamadas páginas amarelas, como *telelistas.net*, *guiamais.com*; catálogos de *softwares*, tais como *cnet.com*, *sourceforge.com*, entre outros. Os catálogos *web* "utilizam vocabulários controlados, como as taxonomias, para indexar objetos de um domínio. Eles se tornaram muito úteis para navegar e consultar" (TZITZIKAS *et al.*, 2002, p. 1).

Em síntese, neste estudo, define-se um catálogo *web* como um conjunto de registros que descrevem itens, que podem ser entidades do mundo real ou abstratas. Os registros são acessíveis pela Internet e organizados sob algum critério, como ordenação alfabética e classificação por autor, ou um vocabulário controlado, como uma taxonomia.

Na seção a seguir, será abordada mais uma evolução na história do catálogo: a utilização da taxonomia facetada, que origina o catálogo *web* facetado, ou simplesmente catálogo facetado.

2.5.2 Catálogo web facetado

Na revisão de literatura feita por Kafure (2004) em sua tese de doutorado, é visível a importância de se aumentar a usabilidade do OPAC, que é tema presente em mais de trinta estudos. Uma preocupação comum é que se tenha diferentes níveis de busca para distintos tipos de usuários, e que haja novas estratégias para se encontrar a informação (KAFURE, 2004).

Indo ao encontro desses objetivos, uma das formas de se aprimorar a recuperação de informação em um catálogo é a implantação da navegação facetada. Serão apresentados vários estudos pioneiros no assunto, que aplicam essa técnica em catálogos automatizados, dando origem aos catálogos facetados.

Fagan (2010) fez uma vasta pesquisa de estudos de usabilidade e aponta que a navegação facetada é uma facilidade comum em novas interfaces de catálogo de bibliotecas. Com as análises realizadas, Fagan concluiu que há uma evidência empírica de que a navegação facetada melhora a performance dos OPACs, mas que isso não necessariamente implica o sucesso do usuário ao usar um OPAC facetado. Como foi visto na seção anterior, existem várias alternativas para a construção de uma interface facetada e isto implica na sua eficácia na recuperação da informação.

No estudo realizado por Niu (2012), emprega-se a navegação facetada nos seguintes catálogos *on-line* de bibliotecas: *McMaster University Library* e *WorldCat*. No estudo, são analisados dois catálogos facetados de biblioteca baseados no aplicativo Endeca, um dos pioneiros na implementação da navegação facetada. Foi feito um registro das ações dos usuários (*log*) por dois meses durante a utilização dos catálogos de biblioteca, com o intuito de analisar seu comportamento em um ambiente real e monitorar suas estratégias de busca. Os meios tradicionais de busca através de palavra-chave são mantidos e disponibilizados em conjunto com a navegação facetada. O estudo destaca que os usuários usam as facetas para dois propósitos: o refinamento da busca e a navegação pelo catálogo, entretanto, grande parte dos usuários são conservadores e utilizam estratégias tradicionais como a busca por palavra-chave no início da pesquisa.

Yang e Hofmann (2010) faz um estudo similar, com três Sistemas Integrados de Biblioteca destaca o Koha, *software* de código aberto, que no trabalho é mais avançado e inovador, por conter, entre outros recursos, a navegação facetada e a contribuição dos usuários. Em sua interface de busca, podem-se observar as facetas dispostas à esquerda (FIG. 27.1), a nota que usuário pode atribuir ao registro (FIG. 27.2), e a indexação do registro através de etiquetas, que o usuário pode realizar (FIG. 27.3).

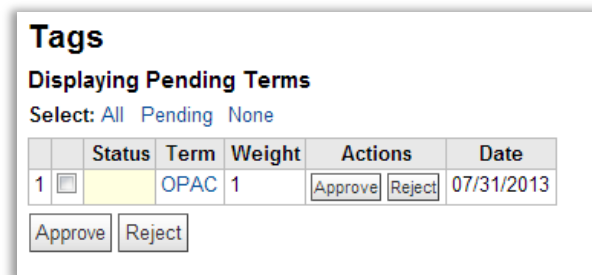
FIGURA 27 - Interface de busca do Koah



Fonte: Site Koah¹⁰.

O Koah trata a indexação realizada pelo usuário, a etiqueta, como uma sugestão, que é aprovada por um membro administrador do sistema. Na FIG. 28 é apresentada a interface para aprovação de etiquetas marcadas pelo usuário. Apenas após a aprovação, a etiqueta aprovada aparece vinculada ao registro para os demais usuários do sistema. A partir desse momento, os usuários podem encontrar o documento através da nuvem de etiquetas, mas elas não são usadas na navegação facetada. Na faceta Tópicos apenas são listados assuntos oriundos do CDD (Classification Decimal de Dewey) que são relacionados à palavra-chave digitada. O modelo de colaboração com aprovação permite as contribuições dos usuários sem comprometer a qualidade da indexação dos registros. Esse modelo, apresenta uma desvantagem, que é a centralização do trabalho na equipe responsável pela administração do sistema, já que ela também se encarrega da gestão dos sistemas de classificação disponíveis.

FIGURA 28 - Aprovação de etiquetas sugeridas pelo usuário



Fonte: Site Koah¹⁰.

¹⁰Disponível em: <http://koha-community.org/demo/>.

O estudo *Extended Faceted Taxonomies for Web Catalogs*, propõe a aplicação da classificação facetada em um catálogo *web*, o ODP. As principais vantagens propostas em relação ao catálogo original são: maior facilidade para a compressão da hierarquia; redução da hierarquia; maior facilidade e eficiência nas operações de manutenção do vocabulário controlado (TZITZIKAS *et al.*, 2002).

O estudo aplica um método que gera a taxonomia facetada a partir da taxonomia original, que, na época do estudo (2002), continha mais de 300 mil termos. O trabalho destaca a economia na taxonomia proporcionada pela metodologia, pois a característica combinatória da taxonomia facetada consegue gerar os assuntos compostos que antes ocupavam um lugar na taxonomia original.

A seção a seguir reúne os principais conceitos apresentados até o presente: taxonomia facetada, navegação facetada e o catálogo *web*, combinando-os e aplicando-os em um catálogo *web* facetado de empresas, ou páginas amarelas, que é o núcleo do objeto desta pesquisa.

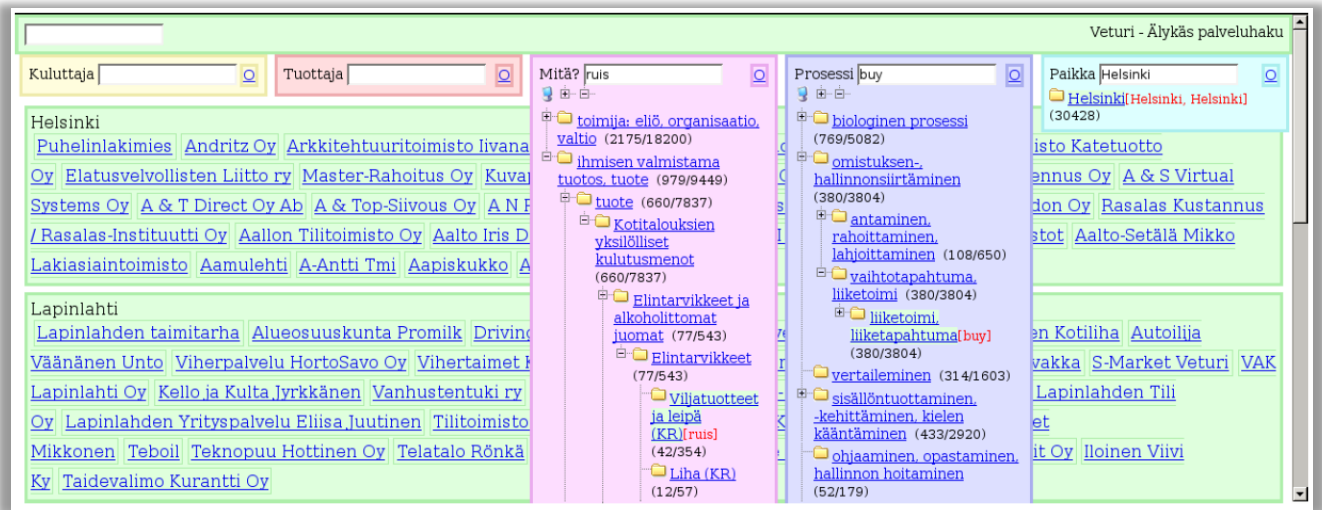
2.5.2.1 Catálogo facetado de empresas

Na pesquisa realizada pelo Grupo de Computação Semântica da Universidade de Helsinki, na Finlândia (SeCo)¹¹, foi desenvolvido um protótipo de um catálogo facetado de empresas. A solução é apresentada na FIG. 29 a seguir (MAKELA *et al.*, 2005).

A taxonomia facetada utilizada pelo catálogo é constituída pelas seguintes facetas: Consumidor, Fabricante, Alvo, Tipo de Serviço e Localização. Quando o usuário clica em cima da caixa de texto da faceta, aparece, embaixo dela, uma árvore que lhe permite navegar hierarquicamente pelos termos da faceta. Ao digitar alguma palavra na caixa de texto que está no alto, imediatamente são filtrados e destacados os termos correspondentes (função de autocompletar) à palavra digitada.

No alto do canto esquerdo, existe uma busca geral que pode ser combinada com a seleção dos termos da faceta. Os registros desse catálogo foram extraídos de uma base de dados com mais de 220 mil empresas e depois representados semanticamente através de um modelo de dados baseado em ontologia.

¹¹Disponível em: <<http://www.seco.tkk.fi/>>.

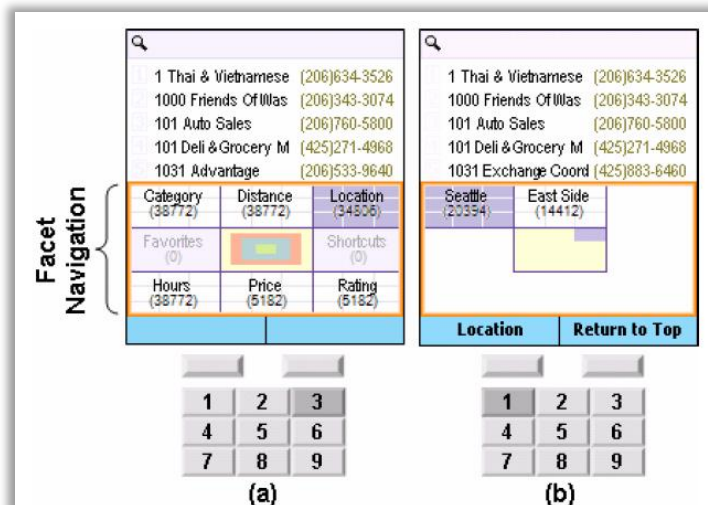
FIGURA 29 - Principal interface do usuário do *Venturi Portal*

Fonte: Makela *et al.* (2005, p. 2).

O artigo *FaThumb: A Facet-based Interface for Mobile Search* aborda a navegação facetada como uma estratégia de busca em dispositivos móveis. O estudo prioriza a navegação e a filtragem de dados interativa em relação a busca por palavra-chave. Para validar a proposta, foi construído um protótipo que usa um catálogo de endereços do estado de Washington.

Na FIG. 30 é apresentada a ilustração do protótipo. O usuário utiliza o teclado para realizar a navegação facetada. Para selecionar a faceta Localização, o usuário pressiona o dígito 3 (FIG. 30.a). Com isso aparecem duas cidades: Seattle e East Side. Depois pressionando o dígito 1, o usuário seleciona a cidade Seattle (FIG. 30.b).

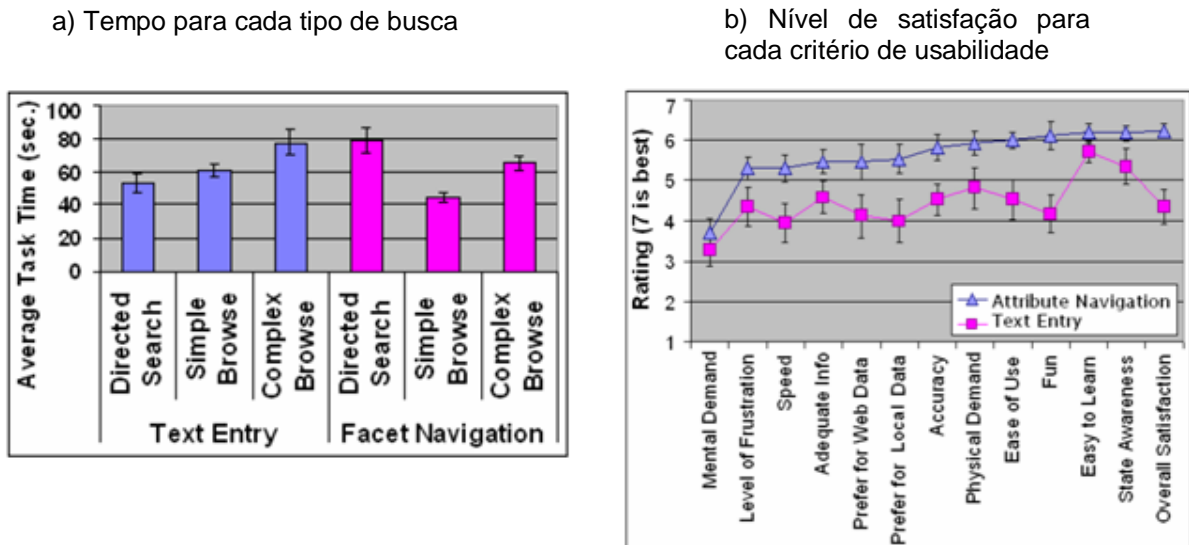
FIGURA 30 - Interface do FaThumb



Fonte: Karlson *et al.* (2006, p. 714).

No teste de usabilidade com os usuários, o estudo conclui que, para pesquisas diretas, a busca por palavra-chave tem um gasto menor que a navegação facetada. Já para pesquisas complexas, a navegação facetada apresenta melhores resultados. No caso da satisfação do usuário, a busca por navegação facetada apresentou índices mais altos em relação aos da busca por palavra-chave (GRÁF. 1).

GRÁFICO 1 - Resultados do teste com usuários do *FaThumb*



Fonte: Karlson *et al.* (2006, p. 718).

2.6 Sistemas colaborativos e modelo de colaboração 3C

Sistemas colaborativos são definidos como: "sistemas baseados em computador que suportam grupos de pessoas envolvidas em uma tarefa (ou meta) em comum e que fornecem uma interface para um ambiente compartilhado" (ELLIS; GIBBS; REIN, 1991, p. 40). Sistemas colaborativos apresentam definição muito semelhante à de *software* social, que é definido como "qualquer *software* que permita a duas ou mais pessoas, em locais diferentes, atuar de forma colaborativa" (DAMES, 2004, citado por BARBOSA; SEPÚLVEDA; DA COSTA, 2009, p. 19), e, no presente estudo, esses dois termos são tratados como sinônimos. Nesse contexto, é apropriado apontar a definição de colaboração:

1 Ato de colaborar; cooperação; ajuda. 2 Trabalho feito pelos colaboradores. 3 Conjunto dos colaboradores. 4 Reunião de duas ou mais pessoas que trabalham juntas para produzir ou utilizar uma aplicação multimídia (MICHAELIS, 2013. p.1).

Sistemas colaborativos podem ser classificados como: síncronos, quando a colaboração entre os usuários ocorre em tempo real; e assíncronos, quando ocorre em tempos diferentes (DAMES, 2004). A colaboração também pode ser realizada no mesmo lugar ou em lugares diferentes (QUADRO 3).

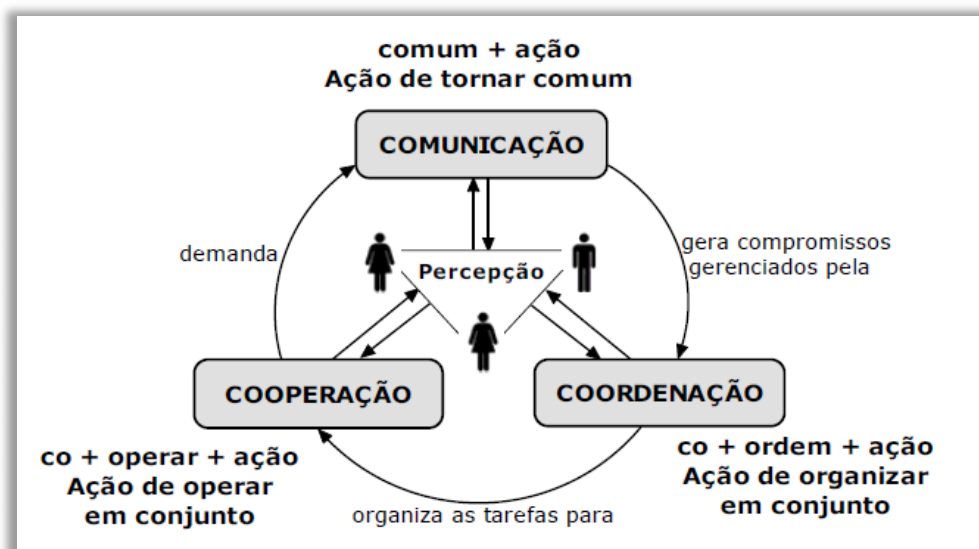
QUADRO 3 - Taxonomia de sistemas colaborativos

Espaço x Tempo	Mesmo tempo	Tempos diferentes
Mesmo lugar	Interação face-a-face	Interação assíncrona
Lugares diferentes	Interação síncrona distribuída	Interação assíncrona distribuída

Fonte: Ellis, Gibbs e Rein (1991, p. 41).

Os sistemas colaborativos podem ser modelados e classificados através do modelo de colaboração 3C, conceito introduzido por Ellis, Gibbs e Rein (1991, p. 40), no qual "o objetivo de um *groupware* é auxiliar grupos na comunicação, na colaboração e na coordenação de suas atividades". Por razões terminológicas, a dimensão colaboração foi alterada para cooperação, e se torna, juntamente com a comunicação e coordenação, uma das atividades da colaboração (GEROSA, 2006). O modelo de colaboração 3C representa a colaboração em sistemas colaborativos através de três componentes: comunicação, cooperação e coordenação, que são apresentados a seguir (FIG. 31).

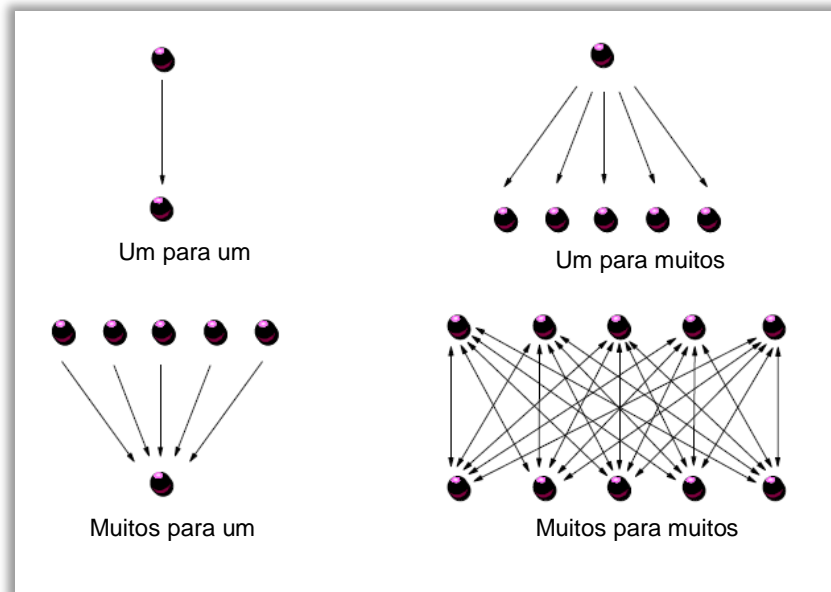
FIGURA 31 - Modelo de colaboração 3C



Fonte: Gerosa (2006, p. 77).

A comunicação, "comum + ação, é a ação de tornar comum, trocar mensagens objetivando o entendimento mútuo, conversar, dialogar" (PIMENTEL, 2006, p. 2), ou seja, a comunicação é a transmissão da informação. As variáveis de tempo, espaço e número de participantes determinam o sistema de comunicação mais apropriado em uma determinada situação (LOTUS, 1995). A comunicação em sistemas colaborativos pode ser classificada como (FIG. 32): um para um; um para muitos, muitos para um e muitos para muitos. Uma das mais importantes contribuições da tecnologia é a eliminação das limitações de tempo e espaço.

FIGURA 32 - Tipos de comunicação em sistemas colaborativos



Fonte: Lotus (1995, p. 9).

A cooperação, "co + operar + ação, é a ação de operar conjuntamente. Os membros do grupo atuam em conjunto, num espaço compartilhado, para a realização de tarefas definidas e organizadas durante a coordenação" (PIMENTEL, 2006, p. 2). O espaço compartilhado é um ambiente essencial para administrar e integrar os conhecimentos individuais dos participantes (LOTUS, 1995).

A coordenação, "co + ordem + ação, é a ação de dispor segundo uma certa ordem e método, organizar, arranjar" (PIMENTEL, 2006, p. 2). É requerida para guiar as pessoas envolvidas em prol de um objetivo em comum. Envolve atividades como estruturar tarefas em uma sequência específica (fluxo de trabalho). Significa coordenar as várias estratégias de comunicação e cooperação balanceando processos estruturados e não estruturados (LOTUS, 1995).

Sistemas colaborativos podem ser classificados de acordo com o modelo de colaboração 3C, conforme funções presentes nos mesmos (FIG. 33). Nessa ilustração os

sistemas são posicionados conforme o grau de realização de cada uma das ações do modelo de colaboração 3C.

FIGURA 33 - Classificação de sistemas segundo o modelo de colaboração 3C



Fonte: Borghoff e Schlichter (2000, p. 125).

É importante ressaltar que o termo “modelo” tem uma extensa lista de significados. Para Sayão (2001, p. 83) modelos são “representações simplificadas e inteligíveis do mundo, que permitem vislumbrar características essenciais de um domínio ou campo de estudo”. Segundo Stachowiak (1972, citado por SAYÃO, 2001), um modelo tem as seguintes características básicas: mapeamento, é uma representação de algo original; redução, não mapeiam apenas os atributos relevantes do original que está sendo representado; e pragmatismo, serve para um propósito em um dado contexto, que é diferente do original. O modelo de colaboração 3C realiza a representação da colaboração através das dimensões comunicação, coordenação e cooperação, e isso quer dizer que a colaboração nos sistemas colaborativos não é representada de forma integral (em todos os seus aspectos de forma definitiva) através do modelo 3C, mas de qualquer maneira, o modelo 3C serve para analisar, representar e classificar tais sistemas.

No dicionário Aulete, no qual existem dezessete definições para modelo, a primeira definição se faz importante no contexto deste trabalho: “modelo é qualquer pessoa ou objeto de atributos especiais, que se reproduz por imitação ou que serve de referência para criação” (AULETE; VALENTE, 2008, p. 1). Sendo assim, o modelo de colaboração que está sendo proposto é algo que pode servir de referência no desenvolvimento de um catálogo *web* facetado e que pode ser adaptado para sistemas colaborativos em geral.

2.6.1 Fluxo de trabalho (*workflow*)

Um fluxo de trabalho ou *workflow* é definido como: "atividades envolvendo a execução coordenada de múltiplas tarefas realizadas por diferentes entidades. Uma tarefa define algum trabalho para ser realizado por uma pessoa, *software* ou ambos" (CASATI *et al.*, 1995, p. 1). Para Georgakopoulos, Hornick e Sheth (1995, p. 123), um fluxo de trabalho é "um conjunto de tarefas organizadas para realizar algum processo de negócio". Um processo de negócio são atividades de uma organização para criar, gerenciar e prover informação. Segundo Hollingsworth (1995, p. 51), processo de negócio é definido como "um tipo de processo no domínio da estrutura organizacional com a finalidade de alcançar objetivos de negócio" e fluxo de trabalho é uma forma de automatizar ou facilitar a execução de um processo de negócio, por exemplo, o tratamento de uma ordem de compra ou um pedido de seguro realizado pelo telefone (GEORGAKOPOULOS; HORNICK; SHETH, 1995).

Existem vários modelos para representação de fluxo de trabalho, tais como: modelo por Gatilhos; modelo de Casati/Ceri; modelo Wide; o modelo da WFMC¹² (*Workflow Management Coalition*); entre outros. Um fluxo de trabalho também pode ser representado através das seguintes linguagens ou notações: *Workflow Process Definition Language* (WPDL); *XML Process Definition Language* (XPDL)¹³; *Business Process Execution Language* (Bpel)¹⁴; *Business Process Management Notation* (BPMN)¹⁵; e diagrama de atividades, que é detalhado na seção 2.8.2.2.

A WFMC também definiu um modelo de referência (FIG. 34) para um sistema de gestão de um fluxo de trabalho (WFMS – *Workflow Management System*). O núcleo do sistema é composto pelo motor ou mecanismo que efetivamente executa o fluxo de trabalho. Para complementar as funções e permitir a interoperabilidade do motor do fluxo de trabalho, deve haver cinco interfaces que são: 1) importação e exportação de definições de processos, que podem ser especificados por uma ferramenta externa e visual de modelagem de fluxo de trabalho; 2) interação com aplicativos clientes do fluxo de trabalho; 3) invocação de ferramentas ou aplicativos de *software*; 4) interoperabilidade entre diferentes sistemas de gerenciamento de *workflow*; e 5) funções de administração e monitoramento (HOLLINGSWORTH, 1995).

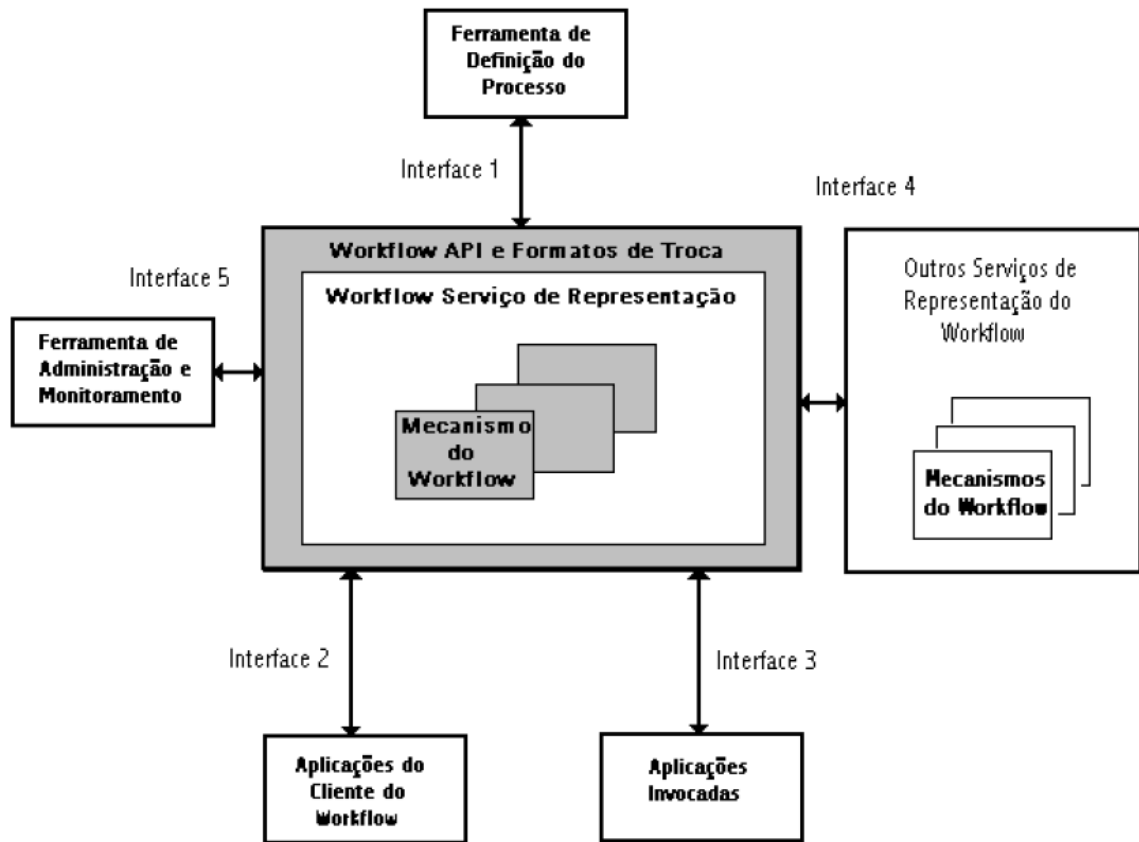
¹² Disponível em: <<http://www.wfmc.org>>.

¹³ Disponível em: <<http://www.xpdl.org>>.

¹⁴ Disponível em: <<http://docs.oasis-open.org/wsbpel/2.0/wsbpel-v2.0.pdf>>.

¹⁵ Disponível em: <<http://www.bpmn.org>>.

FIGURA 34 - Modelo de Referência para um WfMS



Fonte: Hollingsworth (1995, citado por WEINFURTER, 2001, p. 25).

Na seção seguinte são abordados modelos de colaboração específicos que contribuíram para dar origem ao fenômeno da *Web 2.0*.

2.7 Modelos de colaboração da *Web 2.0*

2.7.1 Princípios da *Web 2.0*

A *Web 2.0* teve início a partir de uma segunda geração de serviços *on-line*, caracterizada por ambientes para colaboração, interação e participação comunitária tais como comunidades, como os *sites* de redes sociais, *blogs* e *wikis* (BLATTMANN; SILVA, 2007). O termo *Web 2.0* foi criado para caracterizar a segunda geração de sites da *web*. Na FIG. 35 são apresentadas empresas que exemplificam o conceito da *Web 2.0*:

FIGURA 35 - Serviços da *Web 2.0*

Web 1.0	>	Web 2.0
DoubleClick	>	Google AdSense
Ofoto	>	Flickr
Akamai	>	BitTorrent
mp3.com	>	Napster
Britannica Online	>	Wikipédia
personal websites	>	Blogging
evite	>	upcoming.org and EVDB
domain name speculation	>	Search engine optimization
page views	>	cost per click
screen scraping	>	web services
publishing	>	Participation
content management systems	>	Wikis
directories (taxonomy)	>	tagging ("folksonomy")
stickiness	>	Syndication

Fonte: O'Reilly (2005, p. 1).

Tim O'Reilly, no artigo intitulado "*What is Web 2.0*", lançado em 2005, consolidou o termo *Web 2.0*, e definiu seus setes princípios (O'REILLY, 2005).

O primeiro dele aponta a *web* como plataforma: a *web* começou como apenas um imenso repositório de documentos e informações que podiam ser compartilhados e lidos em qualquer parte do mundo a qualquer momento. A *Web 2.0* impactou no modelo tradicional de desenvolvimento de *software* como aplicativo local ou cliente/servidor. Vários programas passam a ser desenvolvidos apenas para *web*, que passa a ser adotada como plataforma principal. A visão geral é de que ela é um grande meio para disponibilizar serviços *on-line*. É importante observar que as informações não são descartadas, pelo contrário, o conteúdo agrega valor aos serviços. Segundo O'Reilly (2005, p. 1), "o valor do *software* é proporcional à escala e dinamismo dos dados que ajuda a gerir".

O segundo princípio é a questão do uso da inteligência coletiva: na primeira fase da *web*, muitas empresas apenas armazenavam as informações dos usuários. Na *Web 2.0*, muito *sites* estimulam os usuários a produzir mais informações coletivamente e, quanto mais e em maior escala isso for realizado, mais valor e impacto o *site* tem. Mais do que apenas produzir informação, um dos objetivos principais desses sites é utilizar da colaboração de cada usuário para a melhoria contínua da qualidade do serviço, o que foi chamado de uso da inteligência coletiva. Por exemplo, na Wikipédia, a inteligência coletiva é aplicada para a construção de uma enciclopédia; no *Youtube*, para a construção de um portal de canais de vídeo; no *Digg*, um portal de notícias. Com base nesses exemplos, conclui-se que "os

efeitos de rede de contribuições dos usuários são a chave para o domínio do mercado na era da *Web 2.0*" (O'REILLY, 2005, p. 2).

O terceiro princípio aponta para o dado como o elemento mais importante: inúmeras empresas de sucesso na *web* têm como principal diferencial competitivo os dados ou o conteúdo que disponibilizam. Seja um banco de dados especializado, tal como o *Google Maps*, ou um repositório de informações produzidas pelos próprios usuários, tal como a Wikipédia ou *Youtube*, quanto mais abrangente e confiável for, mais valor o *site* na *web* tem. Os direitos autorais sobre esses dados têm sido uma questão cada vez mais importante e, apesar de, em muitos casos, o usuário inserir o conteúdo, tais empresas reclamam para si os direitos autorais para veiculação de determinadas informações, como revisões de produtos e respostas a questões. Além disso, o usuário, ao inserir os dados nessas bases, e concordando com os termos de uso, pode estar sujeito a ter suas informações sendo usadas para *marketing* e outras atividades comerciais. De uma maneira geral, quanto mais informações forem agregadas, produzidas por especialistas ou usuários, mais valor e importância o *site* tem. O dado chega a ser superior que a tecnologia empregada na construção da mesma, já que isto pode ser copiado.

O quarto princípio trata do Fim do Ciclo de Lançamento de *Software*: tradicionalmente, empresas como a *Microsoft* ficam anos desenvolvendo produtos, para que eles sejam lançados e consumidos por seus clientes. Esse é o modelo tradicional de desenvolvimento de *softwares* desde o seu início. Na *Web 2.0*, temos um novo paradigma: ao invés de fornecer produtos, as empresas desenvolvedoras de *softwares* fornecem serviços. Tais serviços são constantemente aprimorados e lançados em um ritmo muito maior que dos *softwares desktops* tradicionais. Por exemplo, o *Flickr*, *site* de compartilhamento de imagens, já revelou que fazem a publicação de atualizações no sistema a cada meia hora. É comum dizer que essas aplicações acessadas pela *web* são um "beta perpétuo", que nunca finalizam seu desenvolvimento ou o ciclo de lançamento do produto.

O quinto princípio aborda os modelos leves de programação: o "beta perpétuo" foi estimulado também pelas linguagens de programação interpretada (*scripts*), pois não é preciso compilar o código-fonte para o sistema funcionar. Isso é também uma grande diferença em relação às empresas de desenvolvimento de *softwares* tradicionais e permite a construção de sistemas dinâmicos que estão em constante mudança. As linguagens de programação tais como *Perl*, *PHP*, *Python* e *Ruby* foram fortemente adotadas. Na *web*, um modelo robusto e complexo cede seu lugar a um modelo mais simples, flexível e fracamente acoplado, permitindo uma maior independência dessas aplicações e uma curva de

aprendizado no desenvolvimento mais rápida, ou seja, com pouco tempo de estudo, já é possível dar seus primeiros passos, utilizando e adaptando.

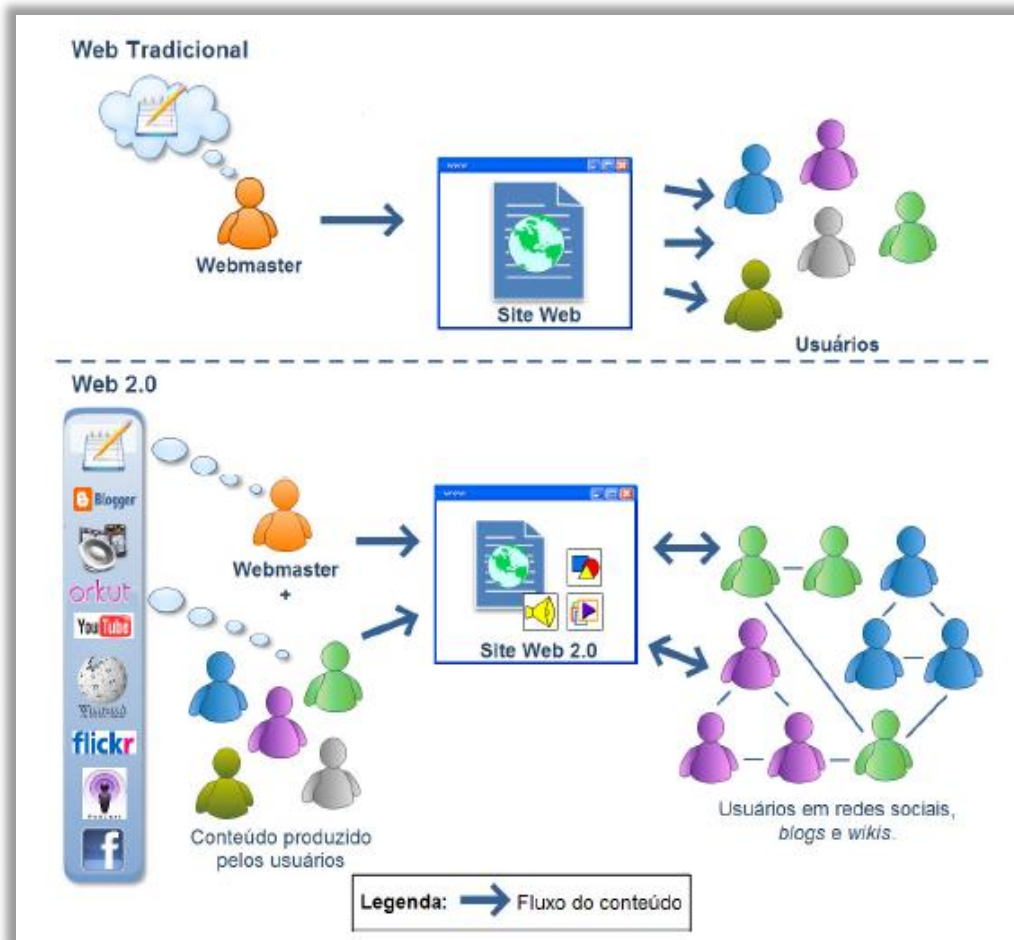
O sexto princípio propõe o desenvolvimento e adaptabilidade do *software* em qualquer plataforma: as aplicações da *Web 2.0* são feitas para rodar em qualquer plataforma de computador, independentemente do sistema operacional. Desde que o navegador de Internet esteja de acordo com os padrões da W3C, mesmo em um dispositivo móvel, isto será transparente para a aplicação *web*.

O sétimo e último princípio estimula a produção de uma interface de usuário rica: no início, a *web* era um suporte para leitura de documentos estáticos. Com o *Javascript* foi possível adicionar capacidade de processamento no cliente, ou seja, o computador do usuário é responsável pelo processamento de parte da lógica da aplicação. Posteriormente com a utilização do XML, houve o surgimento da metodologia AJAX, acrônimo para *Javascript* Assíncrono e XML, para o desenvolvimento de interface de aplicações *Web 2.0*, que se parecem com aplicações *desktop*, que são rápidas, interativas e intuitivas. Com o AJAX, é possível a atualização de dados na interface da aplicação *web* sem a necessidade de o navegador de Internet recarregar toda a página.

2.7.2 Colaboração na Web 2.0

Como apresentado na seção anterior, principalmente relacionado aos princípios da inteligência coletiva e da importância do dado, uma das formas de agregar valor aos *sites* da *Web 2.0* é a captação do conteúdo criado pelos próprios usuários (*User-Created Content* – UCC). Nesse novo paradigma introduzido pela *Web 2.0*, os usuários não são apenas meros leitores do conteúdo disponibilizado, eles passam a exercer o papel de avaliadores ou geradores do conteúdo (FIG. 36).

FIGURA 36 - Web tradicional versus Web 2.0



Fonte: CIRILO (2011, p. 23).

Pode-se definir UCC através das suas características (VICKERY; WUNSCH-VICENT, 2007) que são: publicação requerida, no qual o conteúdo é publicado seja de livre acesso na Internet ou para um grupo de usuário (ou seja, *e-mail* e *chat* não estão inclusos); esforço criativo, que implica que foi realizado um algum novo elemento foi adicionado ao conteúdo original; escopo não profissional, no qual o conteúdo não é obrigado a seguir rotinas ou normas, é muitas das vezes é produzido por não profissionais sem a expectativa de remuneração.

Neste trabalho considera-se um sistema para UCC como uma especialização de um sistema colaborativo, sendo que este último tem um escopo mais amplo, incluindo o *e-mail* e o *chat*, por exemplo.

Para assegurar certo nível de qualidade do conteúdo publicado, existem os seguintes tipos de moderação presente em *sites* UCC (VICKERY; WUNSCH-VICENT, 2007):

- Pré-produção, no qual o conteúdo enviado pelos usuários não é publicado até ser revisado por um especialista ou usuário com tal poder;

- Pós-produção, no qual o conteúdo enviado é publicado imediatamente, mas guardando a possibilidade de ser editado, revisado ou apagado por algum especialista;
- Baseado em pares, no qual o conteúdo enviado é publicado imediatamente, mas pode ser editado, revisado ou apagado por algum especialista, por algum ou todos os usuários, sendo que devem ser mantidas as versões do conteúdo, permitindo a reversão da ação.

A moderação baseada em pares é considerada a melhor forma para estimular a criação de conteúdo pelo usuário, tendo uma alta escalabilidade, e é o formato adotado pela Wikipédia, por exemplo. A moderação nos *sites* UCC atua como a coordenação do modelo de colaboração 3C. Para prover a UCC, são utilizados os seguintes tipos de sistema ou plataforma (VICKERY; WUNSCH-VICENT, 2007): *blogs*; *wikis*; *sites* com comentários; *sites* com publicação de artigos; *sites* com etiquetagem social, tal como o *Digg* ou *Delicious*; *podcasting*; redes sociais; mundos virtuais; e *sites* de compartilhamento.

Nas próximas seções, serão trabalhados sistemas colaborativos amplamente usados na *Web 2.0*: o *wiki* e os sistemas de votação.

2.7.3 Wiki

Wiki é um termo proveniente do havaiano e significa “rápido” ou “veloz”. Foi utilizado pela primeira vez por Ward Cunningham, que foi o autor do primeiro *wiki* disponibilizado em 1995. Sua intenção era desenvolver um *site* no qual o conteúdo fosse gerado pelos próprios usuários (BLATTMANN; SILVA, 2007).

A Wikipédia (2012, p.1) define de forma simples e clara um sistema *wiki*: "*software* colaborativo, que permite a edição coletiva dos documentos, usando um sistema que não necessita que o conteúdo seja revisto antes da sua publicação¹⁶."

Segundo Leuf e Cunningham (2001) um sistema *wiki* pode ser definido como:

uma coleção livremente expansível de páginas *Web* interligadas num sistema de hipertexto para armazenar e modificar informação – um banco de dados em que cada página é facilmente editada por qualquer usuário com um *browser* (LEUF; CUNNINGHAM, 2001 citado por BLATTMANN; SILVA, 2007, p. 201).

¹⁶Disponível em: <<http://pt.Wikipedia.org/wiki/Wiki>>.

Segundo Vieira, Carvalho e Lazzarin (2008), o *wiki* tornou-se um modelo para construção do conhecimento de forma colaborativa devido:

à possibilidade de um conteúdo ser postado, complementado e, até mesmo corrigido, não necessitando da permissão do autor que o postou, anteriormente. Vale ressaltar que esta ferramenta torna possível que alterações ocorridas possam ser consultadas na sua versão anterior ou, recuperadas caso tenham sido excluídas (VIEIRA; CARVALHO; LAZZARIN, 2008, p. 6).

Em 1995, Cunningham (2007 citado por BLATTMANN; SILVA, 2007), estabeleceu os princípios do sistema *wiki*:

- a) Aberto: qualquer leitor pode ter acesso à página, podendo alterar seu conteúdo quando considerar que esteja incompleto ou mal organizado, bem como editar uma nova página;
- b) Incremental: as páginas podem apresentar *links* para outras páginas do próprio *wiki*, inclusive para as páginas que não foram escritas ainda;
- c) Orgânico: a organização estrutural do *site* e dos textos estão abertas à evolução.
- d) Universal: os mecanismos da edição e de organização são os mesmos, de modo que todo escritor seja automaticamente um organizador e um editor;
- e) Preciso: cada página possui um título editável num campo específico;
- f) Tolerante: o comportamento interpretativo é preferido às mensagens de erro;
- g) Observável: as atividades desenvolvidas no *site* podem ser observadas e revisadas por todos;

(Cunningham 2007 citado por BLATTMANN; SILVA, 2007, p. 202)

Uma aplicação para ser considerada como *Web 2.0* não precisa atender, simultaneamente, a todos os princípios apresentados na seção anterior, mas a um subconjunto deles (CIRILO, 2011). Dessa forma, o *wiki* se enquadra em pelo menos três desses princípios:

- a) A Internet como plataforma para processar, produzir ou consumir informação, onde um computador conectado a ela é ferramenta básica e principal de trabalho.
- b) Permite que usuários comuns, que até então não possuíam conhecimentos necessários para publicar conteúdo na Internet - pela ausência de ferramentas de uso simplificado - publicassem e consumissem informação de forma rápida e constante. Notadamente têm-se os blogs e *wikis* como expoentes desta massificação.
- c) Valorização do conteúdo colaborativo e da inteligência coletiva: o conteúdo deve ser produzido e consumido por qualquer um, de forma simples e direta

(BLATTMANN; SILVA, 2007, p. 200).

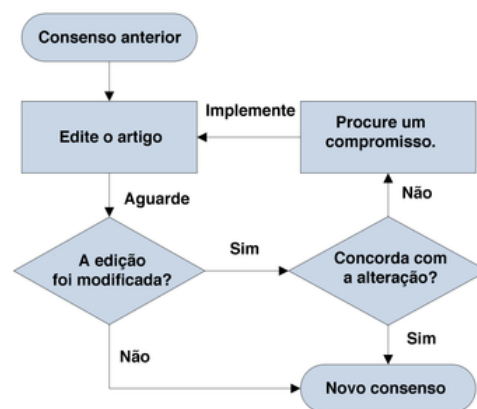
2.7.3.1 Wikipédia

A Wikipédia é o maior caso de sucesso de aplicação de um *wiki*. Ela foi lançada em 2001, por uma organização sem fins lucrativos, a *Wikimedia Foundation*¹⁷, apenas em língua inglesa. Em 2011, possuía aproximadamente 20 milhões de artigos em mais de 250 línguas. O sucesso da Wikipédia se deve em grande parte à:

filosofia *wiki* no desenvolvimento de um ambiente capaz de construir colaborativamente uma grande enciclopédia digital livre, que ao contrário das versões impressas das enciclopédias tradicionais, pudesse receber colaborações da inteligência do coletivo e não somente de especialistas e, além disto, se mantivesse em constante atualização (CAMPOS, 2008, p. 6).

A Wikipédia permite que seus usuários editem livremente seus artigos, mas, em algumas situações, editores têm pontos de vista divergentes referentes ao conteúdo de um artigo, essa situação é chamada de guerra de edições¹⁸ e pode ser caracterizada por usuários que ficam revertendo o artigo para a sua versão indefinidamente. A seguir, é apresentado um fluxograma sugerido para obtenção de consenso na Wikipédia (FIG. 37).

FIGURA 37 - Obtenção de consenso na Wikipédia



Fonte: *Site Wikipédia*¹⁹.

O fluxograma anterior apresenta o processo habitual para obtenção de consenso no conteúdo de um artigo da Wikipédia. Um editor pode realizar uma melhoria editando o conteúdo de um artigo. Entretanto, um editor pode não concordar com sua alteração, revertendo para a versão anterior ou reeditando a sua contribuição. Se o editor concordar com essa alteração é estabelecido um novo consenso. Caso contrário, para evitar reversões

¹⁷Disponível em: <<http://wikimediafoundation.org/wiki/Home>>.

¹⁸Disponível em: <http://pt.wikipedia.org/wiki/Wikipédia:Guerra_de_edições>.

¹⁹ Disponível em: <<http://pt.wikipedia.org/wiki/Wikip%C3%A9dia:Consenso>>.

sucessivas, deve-se criar uma nova seção na página de discussão do artigo. É nessa página de discussão que deve ser obtido um compromisso ou um acordo de qual seria uma solução razoável para o impasse, com a elaboração de um novo conteúdo que agrade os editores envolvidos (WIKIPÉDIA, 2012a). Entretanto, mesmo realizando o fluxo de trabalho acima, pode não ser obtido consenso. Nesses casos, a Wikipédia (2012b) trata a guerra de edição em três instâncias: negociação, mediação, e arbitragem, através dos processos abaixo²⁰:

Opinião de terceiros: é possível solicitar a opinião de terceiros para decidir se uma versão ou outra deve ficar no ar;

Mediação informal: neste processo, é aberta uma mediação na página e um grupo de mediadores voluntários pode participar na resolução da disputa. Para tal, é necessário postar o problema na página “café de mediadores”. É uma mediação informal, ou seja, eles podem ou não ajudar a resolução do problema;

Mediação formal: neste processo, é solicitada uma mediação a um comitê de mediação. Assim, uma pessoa neutra é designada para tratar o problema;

Pedidos de comentários: pode ser criado um pedido para solicitar comentários de usuários sobre o assunto;

Administradores: um administrador pode ser convocado para resolver uma disputa urgente que ataca ou denigre uma pessoa. O administrador é um usuário com, no mínimo, seis meses de participação no site e 2000 páginas editadas;

Processo de arbitragem: é a última etapa na resolução de disputas. É formado por um comitê de administradores, que é diferente para a Wikipédia existente em cada uma das línguas em que o site está disponível. É voltado tanto para conteúdo como para banir usuários de má fé.

Wikipédia (2012b, *on-line*, tradução nossa).

2.7.4 Sistemas de votação

Além da colaboração em *wikis* e a folksonomia, os sistemas de votação são de fundamental importância para proporcionar a qualidade do conteúdo nos ambientes colaborativos.

Segundo Lerman (2007), *sites* sociais de mídia como *Youtube*, *Flirck*, *Delicious*, permitem que seus usuários: (1) criem ou contribuam com conteúdo; (2) classifiquem o conteúdo com etiquetas; (3) avaliem o conteúdo; e (4) criem redes sociais com outros usuários com interesses semelhantes (LERMAN, 2007).

O *Google* revolucionou a busca na *web* aplicando um conceito já usado na bibliometria, no qual um documento produzido por um autor mais citado geralmente é mais relevante para o usuário do que um documento de um autor com menor número de citações.

²⁰ Disponível em: <http://en.wikipedia.org/wiki/Wikipedia:Dispute_resolution>.

Esse conceito assemelha-se a um sistema de votação implícito, baseado na estrutura da *web*, no qual um *link* (URL) para aquele documento é considerado como um voto para ele.

Um sistema de votação é uma estratégia similar a essa, no qual o usuário dá um voto explícito a um recurso. Essa votação pode ser uma nota escalar para um recurso de zero a cinco, ou a votação pode ser um voto positivo ou negativo (*Like* e *Dislike*), procedimento mais simples para o usuário. Um exemplo de um sistema de debate baseado nesse tipo de votação é apresentado na FIG. 38. Um modelo ainda mais simples foi o proposto pelo *Facebook*, no qual existe apenas o voto positivo (*Like*).

FIGURA 38 - Interface do sistema Intense Debate



Fonte: *Site Intense Debate*²¹.

Alguns consideram a votação como um sistema de filtragem social ou um sistema de filtragem colaborativo (LERMAN, 2007). Ela atua como uma espécie de jogo, no qual os usuários competem entre si para terem seus artigos nos primeiros lugares do *ranking*. Essa é uma forma de obter uma lista de resultados mais relevantes, aumentando a precisão. Assim, a ordenação dos documentos, que, no caso do *Digg*, são notícias, se dá conforme a apreciação dos usuários. Outro exemplo é o *Yahoo! Answers*, que exhibe as respostas mais bem votadas, geralmente consideradas mais corretas, no topo da lista. A votação é uma maneira de lidar com a liberdade proporcionada pelo *wiki* ou folksonomia, que pode gerar muito conteúdo para um *site*, mas sem credibilidade ou relevância. Do ponto de vista da BCI, a votação do usuário, nota atribuída ao registro, pode ser considerada como

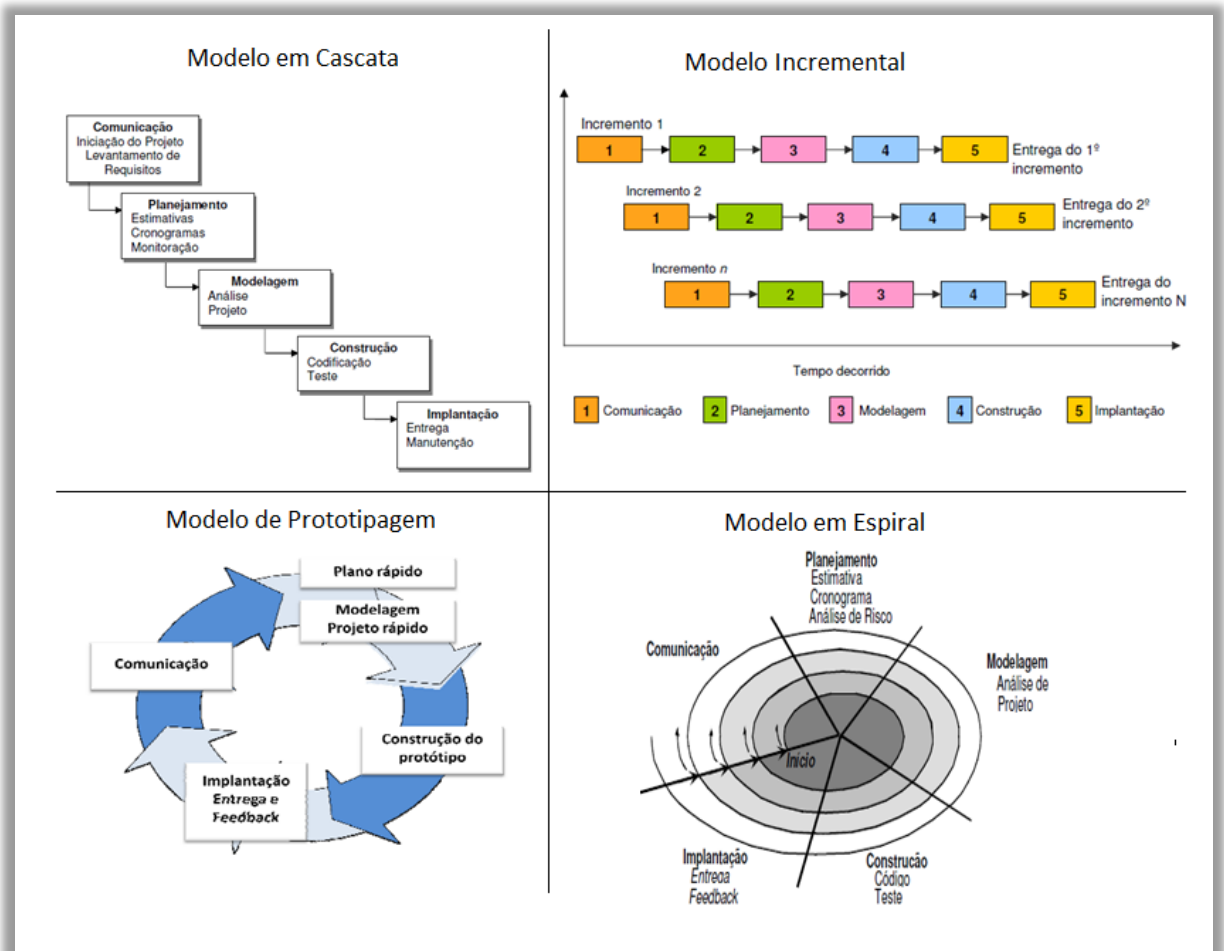
²¹ Disponível em <<http://intensedebate.com>>.

uma indexação. Sua utilização está ligada diretamente em prover informação com maior relevância para o usuário.

2.8 Processo de desenvolvimento de *software*

Nesta seção, serão apresentados modelos de processo de desenvolvimento de *software*. Na FIG. 39, podem ser visualizados os principais modelos prescritivos, que apontam etapas, elementos a serem gerados e como eles devem se relacionar (PRESSMAN, 2006).

FIGURA 39 - Modelos prescritivos de processo de desenvolvimento de *software*



Fonte: Adaptado de Ferreira (2008).

O modelo em cascata, o mais tradicional da engenharia de *software*, é reconhecido por definir etapas que precisam ser seguidas sequencialmente de forma

sistemática. Segundo Ferreira (2008), o modelo em cascata tem as fases de especificação de requisitos pelo cliente, planejamento, modelagem, construção, implantação e manutenção progressiva do *software*. Esse modelo criticado por seu aspecto rígido, já que é necessário aguardar o fim de uma etapa para que outra seja iniciada. Na prática, muitos projetos não seguem uma sequência linear no desenvolvimento das etapas: os requisitos não estão muito claros no início do projeto, e o cliente apenas pode experimentar um resultado nos estágios finais do desenvolvimento, o que pode causar frustração, caso haja um entendimento inadequado dos requisitos ou os requisitos não correspondam às necessidades reais do cliente.

Por sua vez, o modelo incremental assume que "*software* desenvolvido pode sempre crescer e agregar novas funcionalidades" (FERREIRA, 2008, p. 26). Geralmente, apenas requisitos essenciais são desenvolvidos no primeiro incremento, resultando em um produto simplificado que pode ser apresentado ao cliente. Depois, novos incrementos são adicionados gradativamente, modificando e completando o resultado final. Esse modelo tem a vantagem de permitir a alteração de requisitos durante o desenvolvimento do *software*.

Quando se está indeciso sobre os requisitos, principalmente em *softwares* inovadores, no qual se pretende criar algo que ainda não foi feito, sobre o que não existe uma referência, a prototipação permite que a equipe de desenvolvimento e o cliente desenvolvam enxerguem melhor o que precisa ser construído, verificando, ainda em estágios iniciais, se o resultado produzido atende às necessidades do usuário. A prototipação é um modelo iterativo que "permite a detecção precoce de problemas, reduzindo custos e melhorando a qualidade do produto" (PÁDUA, 2007, p. 2).

Outro modelo proposto é o modelo espiral, que é caracterizado por ser:

um modelo iterativo, como o modelo de prototipagem, e sistemático, como o modelo em cascata. Essas características facilitam com que sejam lançadas versões utilizáveis do projeto ao final de cada iteração do modelo, similar ao modelo incremental (FERREIRA, 2008, p. 30).

No modelo espiral, em um ciclo, o processo de desenvolvimento passa por todas as etapas, como no modelo em cascata, resultando em um produto entregue ao cliente na fase final, reiniciando o processo no próximo ciclo. A cada ciclo ou iteração, o *software* vai ficando mais próximo do desejado. Essa abordagem diminui o risco de o produto final não atender aos objetivos esperados, pois, a cada ciclo, o produto é validado com o cliente. Ele também prevê o uso da prototipação para minimizar o risco (PRESSMAN, 2006).

2.8.1 RUP

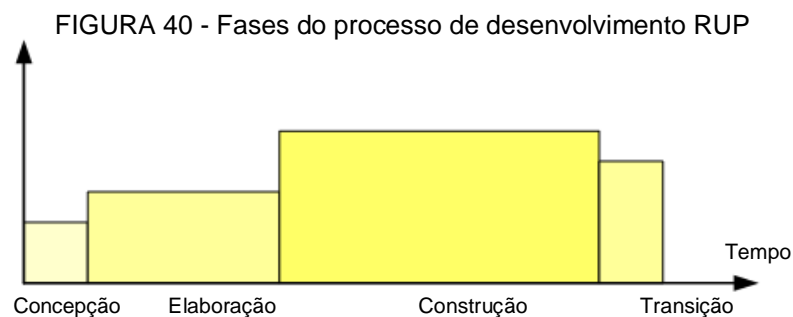
O processo de desenvolvimento de *softwares* denominado *Rational Unified Process-RUP* foi criado por Ivar Jacobson, James Rumbaugh e Grady Booch, em 1999. Conforme Pressman (2006, p. 51):

é uma tentativa de apoiar-se nos melhores recursos e características dos modelos convencionais de processo de *software*, mas caracterizá-los de um modo que implemente muitos dos melhores princípios de desenvolvimento ágil de *softwares*.

Esse processo valoriza a comunicação com o cliente, sugere um processo iterativo e incremental e utiliza amplamente a linguagem *Unified Modeling Language (UML)* durante o processo de desenvolvimento. As fases do processo de desenvolvimento RUP são divididas em:

- **Concepção:** definição do escopo do projeto, levantamento preliminar de requisitos, estimativa de prazo e riscos;
- **Elaboração:** detalhamento dos requisitos, elaboração da arquitetura do projeto, prototipação, resolução dos riscos e atualização do planejamento do projeto;
- **Construção:** implementação dos requisitos com o desenvolvimento dos componentes de *software*, iteração com a modelagem do negócio e requisitos;
- **Transição:** realização de testes, correções e ajustes, implantação no ambiente real de utilização, treinamento.

A FIG. 40 apresenta, através da largura das barras, a proporção entre as fases do Processo Unificado de um projeto típico quanto à sua duração e recursos consumidos em casa fase.

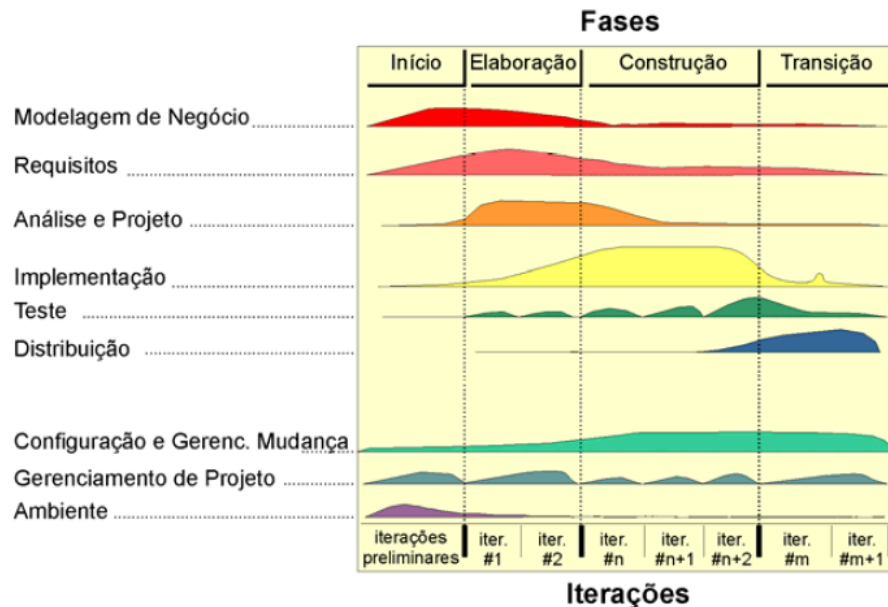


Fonte: Site Wikipédia²².

²² Disponível em: <http://en.wikipedia.org/wiki/Unified_Process>.

Cada fase é composta de iterações. Durante uma iteração são realizadas as disciplinas (FIG. 41), que determinam os artefatos que precisam ser gerados durante o desenvolvimento.

FIGURA 41 - Fases do processo de desenvolvimento RUP



Fonte: Pimentel (2006, p. 85).

Um artefato pode ser um documento, um modelo, um código-fonte ou um plano de teste. As disciplinas que compõem o RUP são (KRUCHTEN, 2003):

- Modelagem de negócios: visa a análise do domínio no qual o sistema será utilizado, entender a dinâmica e a estrutura da organização em que o mesmo será implantado e assegurar que os clientes, usuários e desenvolvedores tem a mesma visão do escopo do projeto;
- Requisitos: visa a elaboração dos requisitos do sistema, ou seja, estabelecer um acordo entre o cliente e os desenvolvedores do que o sistema deve fazer. Um requisito é uma funcionalidade ou condição que o sistema deve atender (KRUCHTEN, 2003). Os requisitos são divididos em: funcionais, especifica o que o sistema faz ou como deve se comportar; e não funcionais, que são parâmetros que o sistema deve atender, tal como performance ou robustez;
- Análise e design: visa descrever como o sistema deve ser implementado a partir dos requisitos. Nesta etapa são usados os diagramas da UML para modelagem do sistema;

- Implementação: visa implementar ou codificar o sistema baseado nas especificações produzidas nas etapas anteriores. Nessa etapa também são realizados testes de unidade para testar se os componentes desenvolvidos estão funcionando;
- Teste: visa verificar se o sistema está funcionando como um todo e se está de acordo com os requisitos;
- Implantação: visa instalar e treinar os usuários para a utilização, caracterizando a entrega do sistema.

Como foi apresentado, a UML é utilizada na fase de Análise e Projeto para a modelagem do sistema, é abordada na próxima seção.

2.8.2 UML

A UML, atualmente regulamentada pela *Object Management Group* (OMG), surgiu em 1996, elaborada pelos mesmos criadores do RUP. Trata-se de uma linguagem não proprietária, utilizada para análise, desenho e implementação de sistemas de computador, modelagem de processos de negócios ou processos em geral. A UML é uma notação que especifica diagramas cujo objetivo é representar elementos definidos na modelagem orientada a objetos. A UML divide os diagramas nos seguintes tipos (OMG, 2011; WIKIPÉDIA, 2011):

- Diagramas estruturais: são diagramas que descrevem, principalmente, a parte estrutural ou estática do sistema, ou que não se altera durante seu funcionamento;
- Diagramas comportamentais: são diagramas que descrevem o que ocorre no sistema modelado ou seu comportamento dinâmico, seja internamente ou com entidades externas;
- Diagramas de interação: são uma especialização dos diagramas comportamentais; descrevem o fluxo de dados no sistema modelado.

Os diagramas estruturais são divididos em:

- Diagrama de classe: descreve uma visão estática do sistema, através de uma modelagem orientada a objeto, que consiste em classes, objetos, atributos e relacionamento entre as classes;

- Diagrama de componente: descreve os componentes do sistema e seus relacionamentos. Os componentes são estruturas mais complexas, sendo constituídos de classes, interfaces ou colaborações;
- Diagrama de composição: descreve um conjunto interconectado de elementos, tais como partes, portas, conectores e colaborações. Representa instâncias em tempo de execução, colaborando para atingir objetivos em comum;
- Diagrama de implantação: representa a associação de artefatos de *software* em nós, que representam dispositivos de *hardware* ou ambiente de execução do *software*. É a disposição física dos artefatos de *software* em um cenário do mundo real. São utilizados na implantação de um sistema;
- Diagrama de objetos: representa como as instâncias de classes, ou objetos, estão relacionados e usados em tempo de execução;
- Diagrama de pacote: representa o sistema dividido em grupos lógicos e os relacionamentos entre eles em um alto nível.

Os diagramas comportamentais são:

- Diagrama de caso de uso: descreve as funcionalidades visíveis através de seus atores, casos de uso e relacionamentos;
- Diagrama de atividade: descreve o comportamento ou fluxo de trabalho do sistema, através da sequência de ações estruturadas por relacionamentos;
- Diagrama de estados: modelam comportamentos discretos do sistema através de máquinas de estado finitas. Descrevem estados em que um objeto pode se encontrar no decorrer da execução de processos de um sistema.

Os diagramas de interação são:

- Diagrama de comunicação: descreve interações. Especifica uma rede e a sequência de mensagens entre objetos em tempo de execução;
- Diagrama de sequência: é uma outra forma de descrever as interações, mostrando a sequência de mensagens entre os objetos por meio de uma linha vertical;
- Diagrama de interação: combina o diagrama de atividades com o diagrama de sequência, apresentando sequências de mensagens durante o fluxo de trabalho;

- Diagrama de tempo: é um tipo específico de diagrama de interação, usado para representar a mudança no estado ou condição de um elemento estrutural ao longo do tempo.

Entre os diagramas disponíveis na UML, este trabalho utilizou diagramas de casos de uso, para modelagem da interação do usuário com o sistema, e diagramas de atividades, para a modelagem da coordenação e da colaboração, que são apresentados a seguir.

2.8.2.1 Diagrama de casos de uso

O diagrama de casos de uso é uma forma de especificar as interações do usuário com o sistema ou usos do sistema (OMG, 2011). Os diagramas de casos de uso "capturam o comportamento do sistema, subsistema, classe ou componente" (RUMBAUGH; JACOBSON; BOOCH, 2004, p. 77). É uma notação que representa graficamente os casos de uso textuais. Na FIG. 42 é fornecido um exemplo de um diagrama de caso de uso, que é composto pelos seguintes elementos principais:

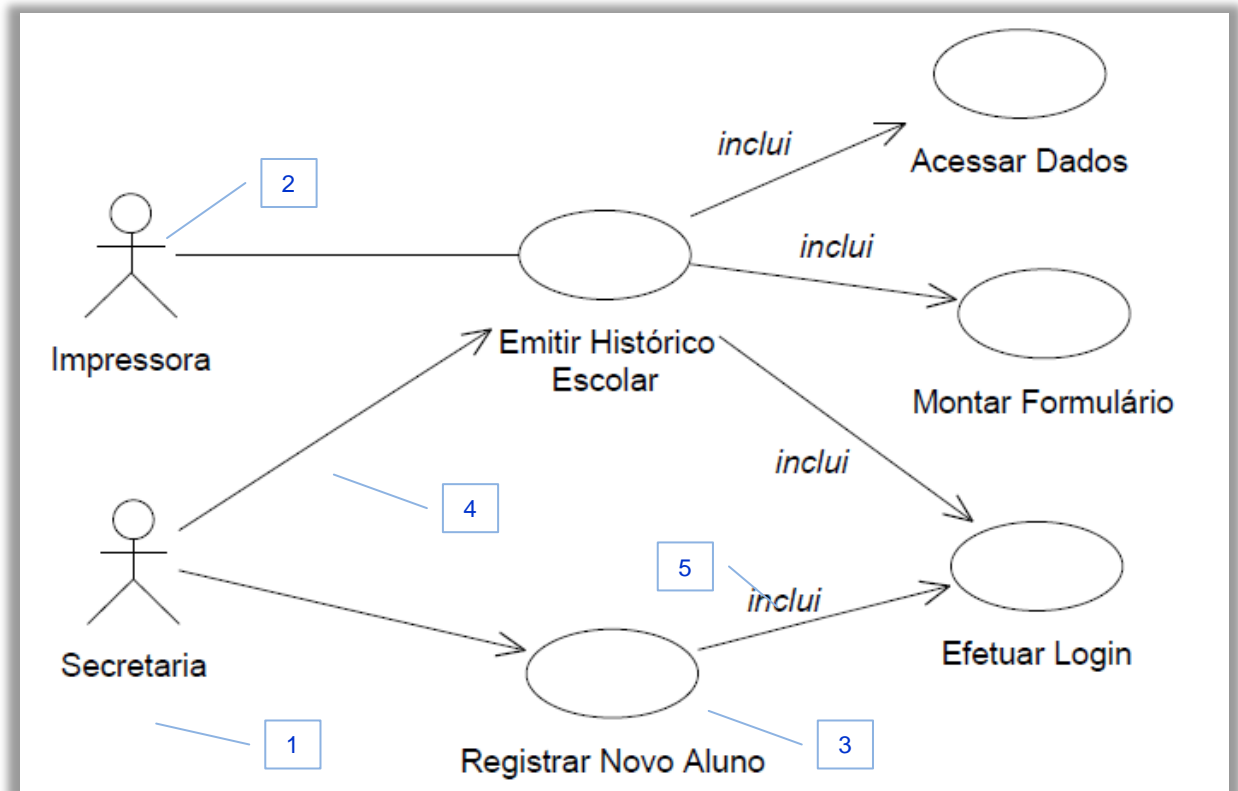
- Ator, a representação de um papel desempenhado por uma entidade externa que interage com o sistema. Os atores podem ser humanos (FIG. 42.1), outros sistemas (FIG. 42.2), ou alguma outra entidade externa. Um ator pode participar de um ou mais casos de uso;
- Caso de uso, a representação de uma funcionalidade realizada pelo sistema, visível externamente e expressa por uma sequência de mensagens entre o sistema e os atores. Cada caso de uso descreve uma parte do comportamento do sistema sem mencionar sua estrutura interna (FIG. 42.3);
- Relacionamento, representa a relação entre os elementos acima.

Os relacionamentos podem ser uma relação entre:

- Ator e um caso de uso, que sempre expressa uma comunicação ou interação entre eles (FIG. 42.4).
- Ator com outro ator, que pode ser uma interação entre atores ou uma relação de generalização. Um exemplo de relação de generalização entre atores é quando há o ator Usuário e o ator Aluno, que adiciona novos atributos além dos atributos do usuário;
- Caso de uso com outro caso de uso, que pode ser do tipo: incluir, usada quando um caso de uso é formado por outros casos de uso (FIG. 42.5); estender, quando um caso de uso acrescenta funcionalidades adicionais

se alguma condição for atendida; e generalização, que segue o mesmo princípio mencionado anteriormente, sendo que o caso de uso mais específico estende os atributos do caso de uso mais genérico.

FIGURA 42 - Exemplo de um diagrama de caso de uso

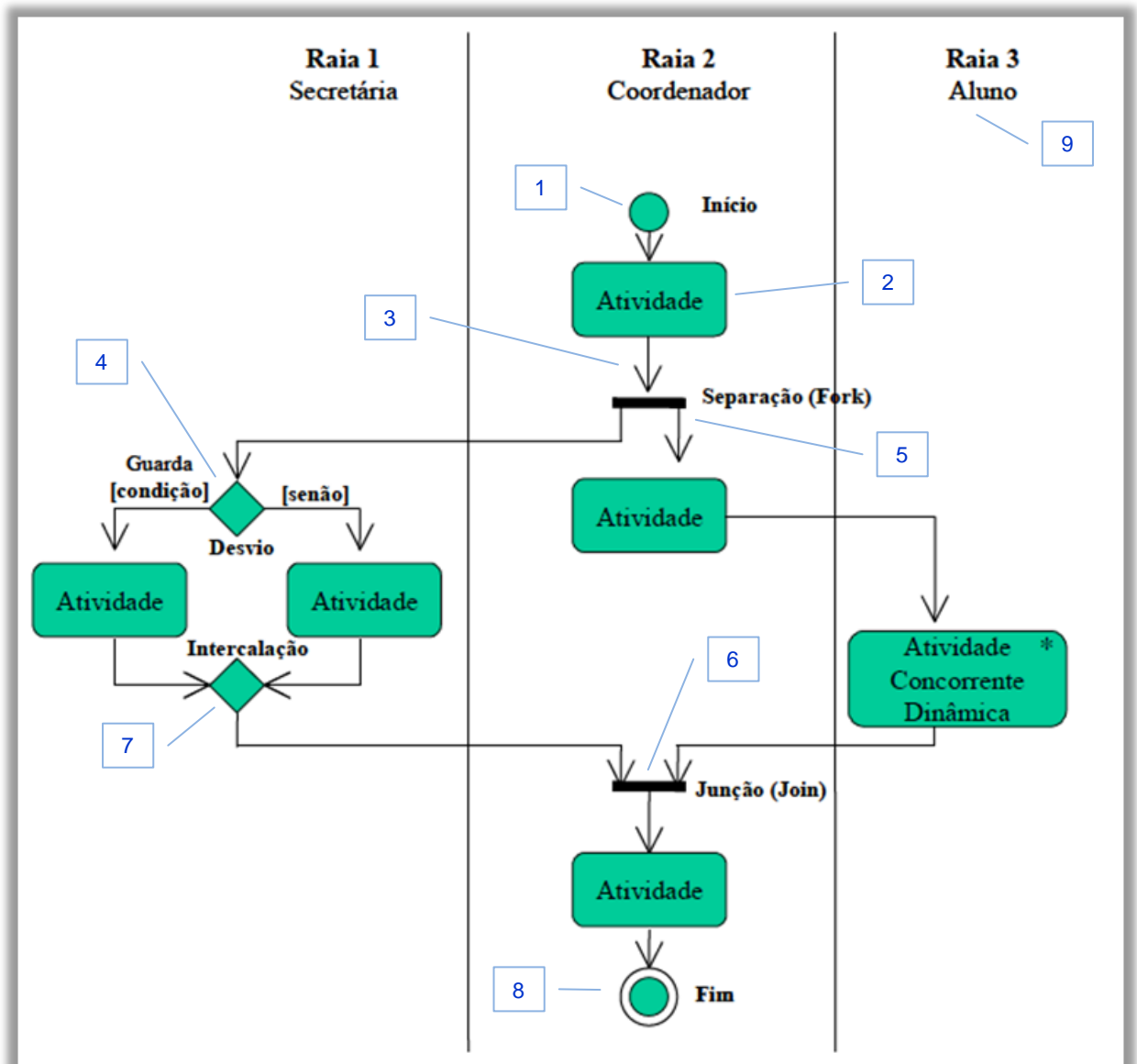


Fonte: Stadzisz (2002, p. 7).

2.8.2.2 Diagrama de atividades

Um diagrama de atividades descreve uma ou mais atividades. Uma atividade representa um comportamento no qual algo é realizado, e é composta por uma sequência de ações estruturadas através de nós de decisão e outros elementos. Um diagrama de atividades pode ser visualizado na FIG. 43.

FIGURA 43 - Exemplo de fluxo de trabalho representado através um diagrama de atividades



Fonte: Nicolao (2004, p. 47).

A seguir são apresentados alguns elementos principais (GUDWIN, 2010):

- Estado inicial: onde inicia a atividade (FIG. 43.1);
- Ação: um passo elementar da atividade, que não pode ser decomposto, tem uma entrada e uma saída. Não se passa para a ação seguinte até que seja realizada a ação atual (FIG. 43.2);
- Arcos: interligam as ações. A seta indica o fluxo (FIG. 43.3);
- Nó de decisão: possui uma entrada única e várias saídas, das quais apenas uma pode ser tomada com base em uma condição (FIG. 43.4);
- Nó de separação: indica a partir daquele ponto a execução de fluxos de controle em paralelo (FIG. 43.5);

- Nó de junção: é um ponto de sincronização, em que somente se as ações dos fluxos de controle em paralelo estiverem concluídas, passa-se para a ação seguinte (FIG. 43.6);
- Nó de intercalação: agregam vários fluxos de controle em um só. Diferentemente do nó de junção, ele não espera que as ações dos fluxos em paralelo sejam concluídas para passar adiante (FIG. 43.7);
- Estado final: onde termina a atividade (FIG. 43.8);
- Partições ou raias: indicam as ações que são realizadas por cada ator em uma determinada atividade. É possível que o ator seja um usuário ou o sistema, fazendo com que o diagrama de atividade possa representar o processo que ocorre em um caso de uso (FIG. 43.9).

O diagrama de atividades também pode ser usado na especificação de fluxos de trabalho (DUMAS; HOFSTED, 2001; PEIXOTO *et al.*, 2008; NICOLAO, 2004). Na norma UML, existem outros elementos que permitem a elaboração de um diagrama de atividades mais sofisticado. No contexto deste estudo, o diagrama de atividades é utilizado para representar a coordenação do modelo de colaboração do sistema colaborativo.

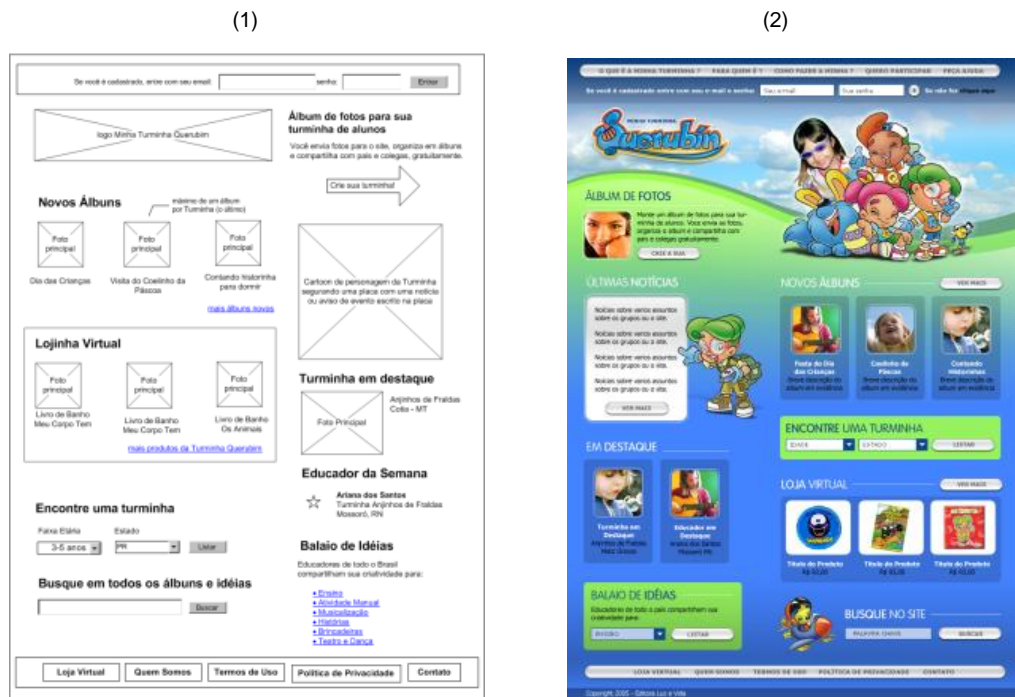
2.8.3 Prototipação

Um protótipo é "uma representação limitada de um produto que permite aos usuários interagir com ele e explorar a sua adequabilidade" (PREECE, 2005, p. 241). Para Pádua (2012, p. 11) é "uma versão parcial de um produto desenvolvida com um mínimo de custo com o objetivo de prover uma visão mais concreta de algum aspecto do produto antes de sua conclusão". A segunda definição em relação a primeira acrescenta duas importantes variáveis: o custo e o tempo. Ou seja, um protótipo tem um custo consideravelmente menor e é concebido bem antes do produto acabado.

No desenvolvimento de *software*, a utilização de protótipos serve para permitir a validação antecipada do produto (antes da conclusão). Preferencialmente, deve ser realizado o mais cedo possível, nos estágios iniciais dentro do ciclo de desenvolvimento, para evitar gastos desnecessários nas demais fases do projeto, por meio da detecção precoce de problemas (PÁDUA, 2012). Permite o refinamento do produto baseado nas impressões (*feedback*) dos usuários (FIG. 45).

A prototipação é a técnica empregada para a elaboração de protótipos. A prototipação pode ser classificada como: de baixa fidelidade, na qual são elaborados esboços que não se parecem muito e que usam materiais bem diferentes dos empregados no produto final, tais como papel e desenhos esquemáticos no computador (FIG. 44.1); e de alta fidelidade, que parecem muito e que usam materiais próximos aos utilizados no produto final (FIG. 44.2).

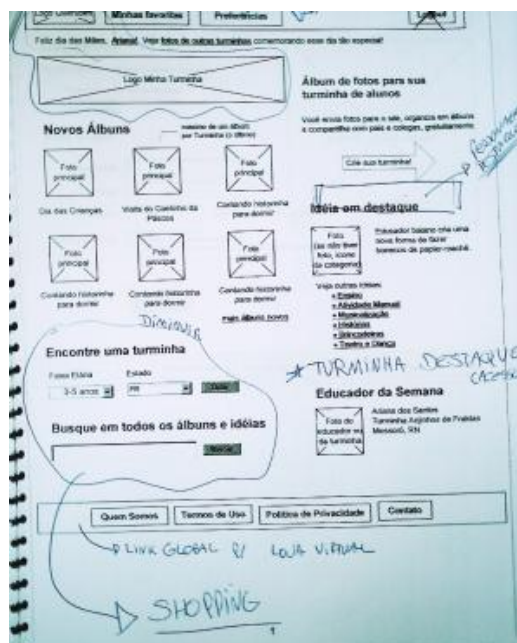
FIGURA 44 - Protótipo de baixa e alta fidelidade respectivamente



Fonte: Blog Usabilidoido²³.

Preece (2005) aponta as vantagens e desvantagens do protótipo de baixa fidelidade e alta fidelidade e recomenda a utilização do protótipo de baixa fidelidade, que, apesar de suas limitações nos testes e detecção de erros, tem custo e tempo de implementação reduzidos e demanda muito menos esforço para ser adaptado (FIG. 45).

²³ Disponível em: <http://usabilidoido.com.br/quanto_mais_simples_o_wireframe_melhor.html>.

FIGURA 45 - Revisão de uma interface através de um protótipo de baixa fidelidade²⁴

Fonte: Blog Usabilidoido²⁵.

Quando a técnica de prototipação, pode ser usada: em papel, na qual os elementos gráficos são desenhados e recortados em papel para montar a interface. Tem a vantagem de poderem ser alterados em tempo de execução, durante o teste (FIG. 46); e navegacional, através de *softwares* para construção de *wireframe* ou *mockup* (FIG. 44.1), que permitem a navegação funcional no protótipo. Com relação aos *softwares*, tem-se as categorias: *desktop*, que são instalados no computador localmente, tais como o *Microsoft Visio*²⁶, *Oversite*²⁷, *Justinmind*²⁸ (gratuito), *SmartyDraw*²⁹, *Axure*³⁰, entre outros; e *on-line*, que não precisam de instalação, podem ser acessados pela Internet de qualquer lugar, a qualquer momento, tais como *Gliffy*³¹, *Mockflow*³², *Wireframe*³³ e *Mockingbird*³⁴, *Moqups*³⁵,

²⁴ Sugestões para alterações na interface estão caracterizados pelos rabiscos no protótipo de baixo nível que foi impresso.

²⁵ Disponível em: <http://usabilidoido.com.br/quanto_mais_simples_o_wireframe_melhor.html>.

²⁶ Disponível em: <office.microsoft.com/pt-br/visio>.

²⁷ Disponível em: <<http://taubler.com/oversite>>.

²⁸ Disponível em: <<http://www.justinmind.com>>.

²⁹ Disponível em: <<http://www.smartdraw.com/specials/wireframe-software.htm>>.

³⁰ Disponível em: <<http://www.axure.com/is/wireframe-software>>.

³¹ Disponível em: <<http://www.gliffy.com>>.

³² Disponível em: <<http://www.mockflow.com>>.

³³ Disponível em: <<https://wireframe.cc>>.

*Creately*³⁶, entre outros. Tanto protótipos em papel quanto por *wireframe* permitem a realização de testes de usabilidade, pois, através deles, os usuários conseguem realizar uma abstração (simplificada) do sistema computadorizado, testando parcialmente a interação nos mesmos (BASSO; CHEIRAN; SANTAROSA, 2009).

FIGURA 46 - Prototipação em papel



Fonte: Basso, Cheiran e Santarosa (2009, p. 4).

A seguir, aborda-se uma etapa que compõe uma fase essencial dentro do processo de desenvolvimento de *software*: o teste de *software*.

2.8.4 Teste de software

Pode-se definir teste de *software* como "o processo de execução de um programa com a intenção de achar defeitos" (MYERS, 2004, p. 11). Outra definição para o procedimento: verificar se o *software* está fazendo o que deveria ser feito, de acordo com os seus requisitos, e também, se não está realizando o que não deveria fazer (RIOS; MOREIRA, 2003). Os testes também são uma forma de incrementar a qualidade do *software* e diminuir o risco na execução do projeto, antes de ele ser disponibilizado para utilização pelos usuários.

Uma abordagem para o teste de *software* é a realização da verificação e validação (V&V). Segundo Pressman (2006, p. 291): "verificação se refere a um conjunto

³⁴ Disponível em: <<https://gomockingbird.com>>.

³⁵ Disponível em: <<https://moqups.com/>>.

³⁶ Disponível em: <<http://creatly.com/Online-UI-Mockups-and-Wireframes>>.

de atividades que garantem que o *software* implementa corretamente uma função específica e a validação se refere a um conjunto de atividades diferentes que garante que o *software* construído corresponde aos requisitos do cliente”.

Em outras palavras, na verificação, indaga-se: "Estamos construindo direito este produto?". Na validação, indaga-se: "Estamos construindo o produto correto?" (BOEHM, 1981). Em linhas gerais, a verificação está mais relacionada ao projeto: se ele foi implementado corretamente, se está funcionando sem erros, e a validação, por sua vez, está mais relacionada aos requisitos funcionais do sistema: se ele está atendendo às necessidades reais dos usuários. O teste de *software* pode ser classificado pelas várias facetas são apresentadas a seguir.

Pode-se classificar teste de *software* quanto ao seu tipo:

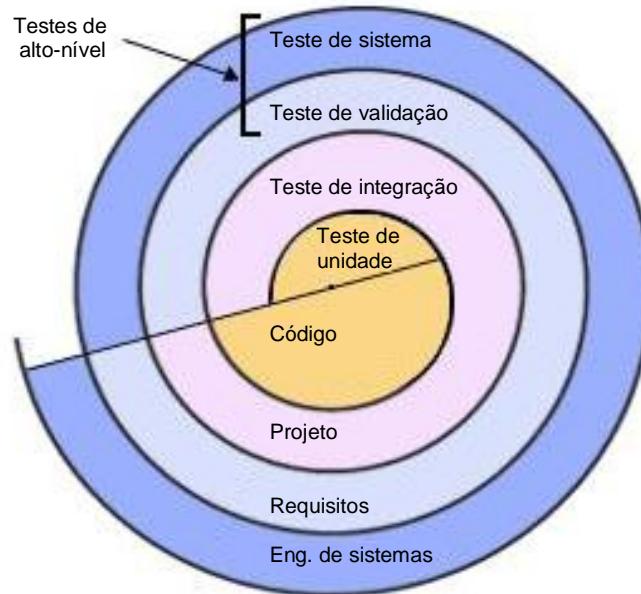
- Caixa-branca ou caixa-preta: teste caixa-preta é aquele no qual o testador trata o que está sendo testado, seja um módulo ou o *software*, como uma caixa-preta: eles verificam se uma entrada especificada tem a saída esperada, não se importando com a lógica ou o que está ocorrendo dentro da caixa-preta. O teste caixa-branca é baseado na análise do código em si e na lógica interna do *software*. Ele é, geralmente, conduzido por programadores. Um teste caixa-branca também é conhecido pelos termos “caixa de vidro” ou “caixa transparente ou estrutural” (JENKINS, 2008);
- Funcional (*functional*) ou não-funcional (*non-functional*): teste funcional é o teste baseado em requisitos, no qual o testador avalia se o *software* realiza a função que foi especificada. Muitos autores consideram o teste funcional como sinônimo de teste caixa-preta, enquanto alguns tratam como apenas um tipo de teste caixa-preta. O teste não-funcional é o teste que não é baseado em um requisito. Testes de *design*, de performance ou de usabilidade são considerados como não-funcionais (JENKINS, 2008);
- Baixo-nível ou alto-nível: teste baixo-nível são aqueles ligados ao módulo implementado como o teste de unidade ou integração. Teste de alto-nível são aqueles ligados aos requisitos, como o teste de validação e teste de sistema (FIG. 47);
- Estático ou dinâmico: o teste estático examina o código fonte à procura de erros lógicos ou verificação de sintaxe. Não é realizado com o *software* funcionando. A técnica mais conhecida é a inspeção de código.

O teste dinâmico é realizado com o *software* em funcionamento e em sua forma final. Testes funcionais ou testes de performance são testes dinâmicos (JENKINS, 2008);

- Manual ou automático: teste manual é aquele executado por um testador e requer a experiência de quem executa para ser bem sucedido. O teste automático é executado por um *software* de forma automatizada. Uma vantagem é poder realizar o teste repetidamente, mas há também uma desvantagem, que é a possibilidade de erros passarem despercebidos por não contarem com a interpretação do testador (JENKINS, 2008).

Pode-se classificar o teste de *software* quanto à fase (FIG. 47):

- Teste de unidade, no centro da espiral, realiza o teste focando um módulo ou componente de *software*, verificando se ele tem erros ou se se comporta como o esperado;
- Teste de integração, usado para verificar se os módulos estão trabalhando de forma correta em conjunto, considerando o *software* como um todo, testando se as entradas e saídas estão de acordo e não apresentam erros. Existem duas abordagens para o teste de integração: descendente (*top-down*), que começa do módulo de controle principal e vai descendo, incorporando gradativamente os módulos subordinados; e ascendente (*bottom-up*), que começa a partir dos módulos atômicos, de nível baixo no programa, assim os componentes são integrados de baixo para cima;
- Teste de validação, que verifica se aquilo que foi construído satisfaz razoavelmente o cliente, ou seja, se cumpre os requisitos estipulados;
- Teste de sistema, que inclui, não apenas o *software*, mas todos os elementos de um sistema baseado em computador, envolvendo o *hardware*, a rede, o pessoal, o conteúdo, analisando se estão operando da maneira coesa (PRESSMAN, 2006).

FIGURA 47 - Estratégia de teste de *software*

Fonte: Adaptado de Pressman (2006, p. 291).

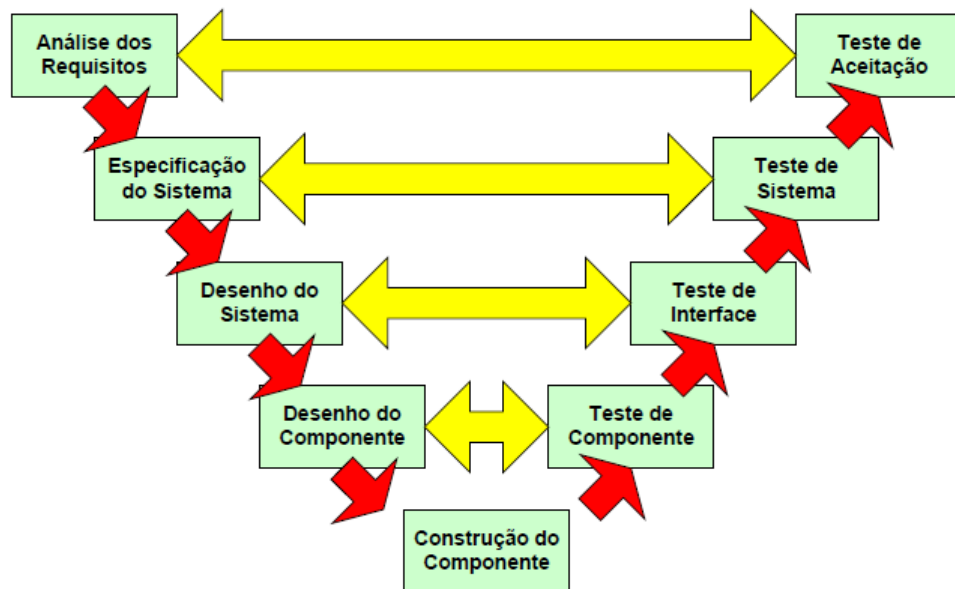
Da mesma maneira que o processo de desenvolvimento de *software*, o teste de *software* pode adotar um modelo. O modelo V para teste de *software* está de acordo com o processo de desenvolvimento de *software* RUP e prescreve um teste para cada estágio de desenvolvimento de um *software* (FIG. 48). Dessa forma, existem vários níveis de teste relacionados com cada fase de desenvolvimento que são (RODRIGUES, 2006):

- Teste de componente: que é realizado após a implementação de um módulo ou componente de *software* e é realizado pelos próprios desenvolvedores;
- Teste de interface, que afere se vários componentes estão se comunicando e funcionando juntos corretamente;
- Teste de sistema, que ocorre após ter sido verificado que todos os componentes do *software* estão trabalhando integrados sem erros. O teste de sistema exercita o *software* como se ele estivesse em produção e afere se todos os requisitos foram implementados e se estão de acordo com o que foi especificado nas etapas iniciais do projeto.
- Teste de aceitação, que também é baseado nos requisitos, mas que é realizado pelo próprio usuário no ambiente de produção. Neste nível, é verificado se o *software* faz realmente o que foi pedido e se atende às necessidades reais do usuário.

É importante lembrar que os testes podem ser aplicados após o término de cada fase de desenvolvimento (setas amarelas na FIG. 48). Por exemplo, após a especificação dos requisitos, pode ser feita uma inspeção para verificar se eles realmente atendem às necessidades do usuário, se não estão vagos ou incompletos, realimentando a fase de desenvolvimento e evitando que problemas só sejam detectados em fases avançadas do projeto, o que gera custos muito altos, já que é necessário realizar a correção dos erros.

Segundo Boehm (1981), um erro encontrado na fase de especificação custa dez vezes menos para ser corrigido do que se ele for encontrado na fase de codificação, e centenas de vezes menos do que um erro encontrado depois do lançamento. Boehm, por meio da análise de vários projetos, criou a proporção de custo 1:6:10:1000 para expressar a relação entre a correção de erros nas fases de requisitos, análise, implementação e utilização, respectivamente.

FIGURA 48 - Modelo V para teste de *software*



Fonte: Rodrigues (2006, p. 29).

É possível classificar teste de *software* quanto à técnica. Diversos autores comentaram essa questão e propõem soluções diferentes.

Segundo Balci (1995, citado por FERREIRA, 2008) tem-se: Auditoria, Validação de Face, Inspeções, Revisões, Teste de Turing, *Walkthroughs*, Checagem de Inconsistência, Análise de Fluxo de Dados, Análise Baseada em Grafos, Análise Semântica, Análise Estrutural, Análise Sintática, Monitoração de Execução, Execução de Perfil, Execução de Traço, Testes de Campo, Comparação, Gráfica, Validação Preditiva, Teste de

Regressão, Análise Sensitiva, Técnicas Estatísticas, Teste de Estresse, Testes de Sub-Modelo, *Debug* Simbólico, Visualização, Simbólica, Gráfico de Causa Efeito, Análise de Partição, Análise de Caminhos, Execução Simbólica, Checagem de Suposições, Análise de Limites, Suposição Indutiva, Indução, Inferência, Dedução Lógica, Cálculos Preditivos, Transformação Preditiva, Prova de Exatidão.

Para avaliação de interfaces, Preece (2005) divide as técnicas em: 1) "quick and dirty" (rápido e sujo); 2) testes de usabilidade; 3) levantamentos; e 4) avaliação preditiva. Para avaliação de usabilidade, segundo PEREIRA (2011), pode-se dividir as técnicas de avaliação da usabilidade da interface em: métodos empíricos, quando envolve a observação dos usuários; e analíticos, quando é baseado no julgamento de especialistas.

Após essa exposição de diversos tipos de teste de *software*, são apresentadas técnicas voltadas para a verificação em fases iniciais do desenvolvimento, tais como o teste *Quick and Dirty*, para um retorno prévio após a prototipação da interface, e o teste *Ad-hoc*, para o teste de unidade, no qual se avalia se o requisito foi implementado. Em fases avançadas do desenvolvimento são apresentadas técnicas voltadas para a validação, como o teste de usabilidade, para aferir se necessidades reais do usuário são alcançadas pelo *software* desenvolvido.

2.8.4.1 Teste *Ad-hoc*

O teste *Ad-hoc* é realizado desde os primórdios do desenvolvimento de *software* (LUO, 2001). É considerado um método informal ou não estruturado, pois não são elaborados casos de teste. Apesar disso, é um teste que pode ser realizado rapidamente. Trata-se de um método baseado na improvisação, bastante dependente da experiência do testador. Entretanto, é útil em fases iniciais do desenvolvimento, e a intenção é realizá-lo apenas uma vez, a menos que sejam encontradas falhas (WIKIPEDIA, 2013).

2.8.4.2 Teste *Quick and Dirty*

Para avaliação de interface, Preece (2005, p. 341) cita uma técnica usual realizada durante o processo de desenvolvimento, chamada de *Quick and Dirty*, ou rápido e sujo, que é similar ao teste *Ad-hoc*, mas deve ser realizado em conjunto com o usuário ou cliente. Essa técnica é chamada de "rápida e suja", porque é feita de maneira informal e em um curto espaço de tempo.

Nessa técnica, os *designers* podem, informalmente, obter um *feedback* dos usuários ou de consultores para confirmar que suas ideias estão de acordo com as necessidades requeridas pelos usuários. Uma grande vantagem é que essas avaliações

podem ser feitas em qualquer fase do projeto, permitindo uma rápida alimentação do processo de desenvolvimento, em oposição a outras técnicas que têm alto custo e demandam muito tempo para sua realização e documentação.

Para realizá-la, geralmente, são feitas reuniões com os usuários. É como se fosse "colhida" uma impressão do usuário, sem a necessidade de seguir rigorosamente um roteiro de testes ou um *checklist*. No início do projeto, os desenvolvedores podem receber um retorno sobre a concepção do projeto. Com o avanço do desenvolvimento, pode-se obter um *feedback* sobre a interface, pontos a serem melhorados. Essas impressões dos usuários são anotadas durante a reunião mediante uma conversa informal. É uma técnica popular em *design* para *web* e em projetos com uma equipe pequena.

2.8.4.3 Avaliação de usabilidade

A avaliação da usabilidade tem como objetivo principal a "avaliação da qualidade da interface como instrumento da interação usuário-computador" (PÁDUA, 2012, p. 124).

Para Pereira (2011), a avaliação de usabilidade pode ser dividida em dois grandes grupos: métodos empíricos e métodos analíticos. Alguns métodos empíricos são:

- Observação direta, quando um avaliador monitora as ações do usuário. Ela tem a desvantagem de ser considerada muito invasiva, pois a observação pode constranger o usuário;
- Observação indireta, quando o monitoramento é feito por câmeras ou captura de tela das ações do usuário. Tem a vantagem de ser menos invasiva, pois há um distanciamento da observação;
- Uso de entrevistas e questionários, através dos quais os usuários podem expressar sua opinião sobre o sistema;
- Grupo focal, que é a realização de uma reunião com mais de cinco usuários, que discutem sobre a interface, sob a regência de um moderador;
- *Thinking-aloud* (protocolo verbal), no qual o usuário verbaliza suas opiniões e sentimentos durante a realização do teste;
- Método de medida de desempenho, em que são registrados parâmetros da interação do usuário, tais como o tempo para realização de uma tarefa ou a quantidade de acertos ou erros.

Alguns métodos analíticos de avaliação de usabilidade são:

- Avaliação heurística, na qual especialistas, com base em heurísticas de usabilidade, realizam uma inspeção da interface. É um procedimento fácil e rápido de conduzir;
- Revisão de *guidelines*, que é similar à avaliação heurística, mas difere-se dela pela utilização de uma lista de recomendações na inspeção da interface;
- Percurso cognitivo, em que especialistas simulam um caminho que o usuário realizaria para execução de determinadas tarefas na interface;
- Inspeção de consistência, em que é observado se a terminologia, cores, ícones, símbolos estão padronizados em uma família de interfaces.

Os métodos analíticos de avaliação de usabilidade tendem a ser mais fáceis de realizar e a demandarem menos tempo do que os métodos empíricos, que envolvem a participação de usuários. Entretanto, os analíticos são mais subjetivos, e dependem do nível de conhecimento do especialista, de forma que nem sempre são menos dispendiosos.

2.8.4.4 Teste de usabilidade

O teste de usabilidade é um tipo de avaliação de usabilidade, baseada na participação dos usuários. É um método empírico ou experimental e é considerado confiável e indicado (HIX; HARTSON, 1993 citado por PÁDUA, 2012).

Rubin (1994) define o teste de usabilidade como o "processo que emprega usuários na realização dos testes que são uma parcela representativa da audiência para avaliar um produto com relação a critérios de usabilidade específicos" (RUBIN, 1994).

Segundo Pádua (2010), o teste de usabilidade tem os seguintes objetivos:

Avaliação da qualidade da interface em termos de sua usabilidade; Identificação de problemas visando melhorias durante o desenvolvimento do produto; Troca de experiência; Promoção da participação/comprometimento dos usuários; Facilita decidir sobre alternativas importantes para o sucesso da interface; Pode ser associada à prototipação, tornando a avaliação do protótipo mais efetiva (PÁDUA, 2010, p. 4).

Para a realização de um teste de usabilidade é necessário elaborar um plano de teste, que tem etapas que serão detalhadas a seguir (RUBIN, 1994).

2.8.4.4.1 Objetivos

Descrever, em alto nível, as razões para a realização do teste sob o ponto de vista da organização que usará a interface. Um exemplo de bom objetivo é: conhecer como

os usuários efetivamente estão interagindo com a interface ou identificar os principais problemas que existem na interface.

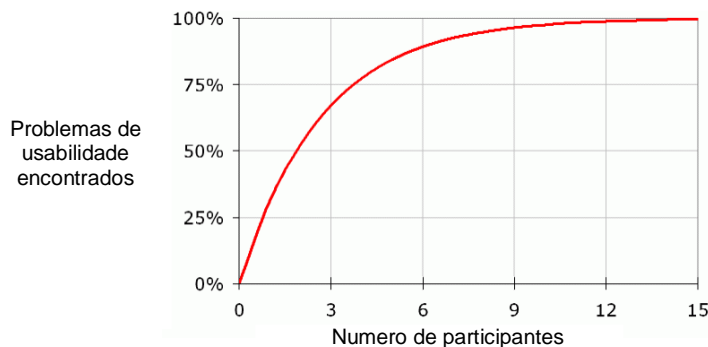
2.8.4.4.2 *Design* do teste

Conhecido também como *design* do teste, "configura as técnicas selecionadas para a avaliação" (PÁDUA, 2012, p. 134). Define, por exemplo, se será realizado a observação direta ou indireta dos participantes, os parâmetros que serão observados, a infraestrutura necessária, os participantes, o roteiro, o tempo de realização, a preparação do ambiente, entre outros parâmetros do teste.

2.8.4.4.3 Participantes

É importante selecionar participantes com um perfil que representa os usuários da interface, ou o público-alvo. Dessa forma, características como idade, gênero, habilidades, entre outras, devem ser consideradas. No caso de avaliação de *sites*, Nielsen (2007) recomenda que os usuários tenham pelo menos um ano de experiência com o uso da *web*, pois, em caso de inexperiência do usuário, as dificuldades em lidar com o próprio navegador de Internet, sistema operacional, entre outros, prejudicam a utilização do *site* que se quer avaliar. Com relação ao número de participantes, segundo Nielsen (2000), com cinco participantes é possível detectar em torno de 85% dos problemas de usabilidade (FIG. 49). Como testes de usabilidade são dispendiosos, é melhor realizar mais testes com cinco participantes do que menos testes com o dobro de participantes, devido ao fato de que cada teste realimenta o desenvolvimento, indicando prioridades para melhoria da usabilidade.

FIGURA 49 - Curva de problemas detectados versus participantes



Fonte: Nielsen (2000, p. 1).

2.8.4.4.4 Ambiente do teste e preparação

Nesta etapa é determinado o ambiente em que o teste será realizado. A determinação disso é uma combinação do *design* de teste almejado com os recursos e o

tempo disponíveis. O ambiente de teste pode ser o ambiente de laboratório ou o ambiente real de utilização do usuário; local ou em várias regiões do globo; com apenas um usuário de cada vez ou com vários usuários ao mesmo tempo no laboratório; pode-se realizar a observação direta pelo moderador, ou a eletrônica, através de um *software* que faça a captura do vídeo do usuário. Na preparação do ambiente, deve-se certificar o pleno funcionamento do sistema a ser testado, dos *softwares* auxiliares que devem existir para a realização das tarefas e coletas de dados.

2.8.4.4.5 Lista de tarefas

A lista de tarefas deve conter aquelas que seriam executadas normalmente na utilização do sistema. É desejável que a descrição da tarefa seja feita da maneira mais clara possível, para evitar a intervenção do moderador do teste durante sua realização.

2.8.4.4.6 Dados a serem coletados

Esse item refere-se à determinação dos dados que se desejam coletar durante o teste, que podem ser: quantitativos, tais como desempenho, taxas de erro, tempo de execução de uma tarefa, número de cliques para se realizar uma tarefa; ou qualitativas, como perguntas de opinião ou satisfação do usuário na realização de cada tarefa.

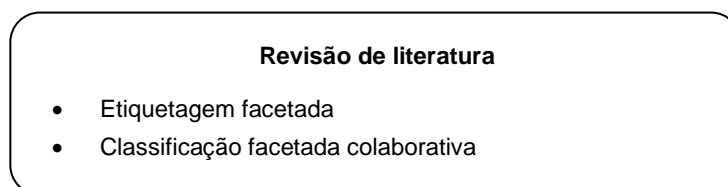
2.8.4.4.7 Análise dos dados

Essa etapa refere-se ao relatório fruto da análise dos dados coletados durante o teste. É voltado principalmente para a equipe de desenvolvimento. É recomendável a utilização de tabelas e gráficos para permitir a percepção de tendências ou comportamentos.

3 REVISÃO DA LITERATURA

A seguir, são apresentados trabalhos que propõem modelos colaborativos para a classificação facetada com objetivos similares aos propostos nesta pesquisa (FIG. 50).

FIGURA 50 - Estrutura da revisão de literatura



Fonte: Elaborado pelo autor.

3.1 Etiquetagem facetada

O primeiro protótipo relevado pela pesquisa bibliográfica que propõe uma etiquetagem facetada é o *Faceted Index Internet Gateways* de 2002 (DEVADASON *et al.*, 2002), um sistema que implementa um catálogo *web* similar ao *Yahoo! Directory*. Além dos metadados do *Dublin Core*, esse estudo estabelece quatro campos para realização da indexação dos cabeçalhos de assunto do recurso *web* (FIG. 51) nas seguintes facetas ou categorias elementares: disciplina/base, campos de estudo convencionais e termos relacionados; entidade/núcleo, objeto principal do estudo dentro da disciplina; propriedade/ação, denota o conceito de atributo, qualitativo ou quantitativo; e qualificador comum, forma, ambiente, tempo e espaço.

Esse estudo é interessante, pois aplica uma estrutura facetada na indexação especificamente do assunto do recurso *web*, inspirado diretamente pela análise facetada de Ranganathan. Esse processo é denominado indexação facetada. Por exemplo, ao indexar o recurso intitulado "Na tecnologia de couro, a cura da pele de porco com sal seco na Tailândia", em cada faceta, poderiam ser usados os termos: tecnologia de couro (disciplina), pele (entidade), curar (ação), Tailândia (qualificador de lugar).

Entretanto, é necessário o preenchimento de vinte e um campos de dados, sendo que a maioria é do *Dublin Core* e, nos campos para indexação facetada, é possível associar vários termos, ou seja, é uma proposta que exige um pouco de empenho do

usuário. No estudo, cita-se que, em alguns casos, o usuário encontra dificuldades para reconhecer as categorias e o autor estipula um treinamento para a formação do indexador.

FIGURA 51 - Indexação facetada do *Faceted Index Internet Gateways*

18.Resource's Faceted Subject Heading (Based on DSIS, each element category can add more than one if there are Standard Terms or synonyms, one per box)

1.Discipline/Base (Conventional fields of study, any aggregate of such fields or artificially created analogous fields Example:Physics, Space Science, HIV)

Arts

Subdivisions of Discipline/Base

Species/Type

2.Entity/Core (Main Object of Study under the Discipline/Base)

Subdivisions of Entity/Core

Species/Type

ADD DELETED

3.Property & Action (Property:Denoting the concept of attribute, qualitative or quantitative; Action:Denoting the concept of doing)

Property

Subdivisions of Property/Action

Property ADD ADDSUB DELETE

4.Common Modifier (Form, Environment, Time, Place)

Form Modifier

Subdivisions of Common Modifier

Fonte: Adaptado de Devadason *et al.* (2002, p. 13).

O protótipo denominado *Faccete* foi desenvolvido por Lai (2009) em sua dissertação de mestrado no MIT (*Massachusetts Institute of Technology*). Esse protótipo propõe uma etiquetagem facetada, modelo similar à tradicional etiquetagem social, em que o usuário digita livremente etiquetas para marcar um recurso.

A grande diferença entre a etiquetagem tradicional e a etiquetagem facetada é que, nesta última, o usuário realiza uma marcação segmentada em facetas. Tal marcação não é hierárquica, as etiquetas são posicionadas na raiz da faceta. Esse modelo é aplicado a um *site* de compartilhamento de *bookmarks* (*links* de Internet favoritos), similar ao *delicious.com*, que é um ícone da *Web 2.0*. Na FIG. 52, apresenta-se uma ilustração da interface de marcação de um *link* de Internet.

As etiquetas são inseridas nas facetas Tipo de objeto (*Type of object*), Nome (*Name*), Sobre (*About*) e Conteúdo (*Contains*). Em relação ao estudo anterior, *Faceted*

Index Internet Gateways, percebe-se que no *Faccete*, o processo de indexação foi simplificado, tornando a tarefa mais fácil para o usuário.

FIGURA 52 - Interface para etiquetagem facetada do *Faccete*

faceted tags

facets	tags		
Type of Object	aggregator	blog >	X
Name	digg	---	X
About	culture	---	X
Contains	news	---	X
more			

unfaceted tags

digg > links > news > web2.0 > 796 project technology >
 blog > tech > social daily > community ---

notes

tags

Fonte: Lai (2009, p. 28).

Na FIG. 53 é exibida a interface tradicional do *delicious.com*. Pode-se compará-la a interface de busca do protótipo *Faccete* (FIG.54). Percebe-se que o *Faccete* permite, além da busca por palavra-chave, a navegação pelas facetas na página inicial do protótipo (painel escuro do lado esquerdo abaixo).

FIGURA 53 - Interface do *delicious.com*

Type a tag

Bookmarks 285
Sorted by Most Recent

Tags
▼ Top 10 Tags

design	71
webdesign	58
tools	28
blog	27
tutorial	26
resources	26
development	22
free	21
inspiration	20
css	19
▼ All Tags	767
review whentlineperms	1
10	1
18"	1
360	2
4blog	1
psd	1

12 AUG 09 30+ New Useful WordPress Tricks & Hacks | Wordpress SAVE 236

Laptop Backpacks and Laptop Bags by booo SAVE 997

Johnston & Murphy SAVE

YouTube - How to Copy Your Online WordPress Blog to XAMPP on Your Local Computer or USB Drive SAVE

10 AUG 09 Raindance Italian Greyhounds, breeder in NY, puppies occasionally available SAVE

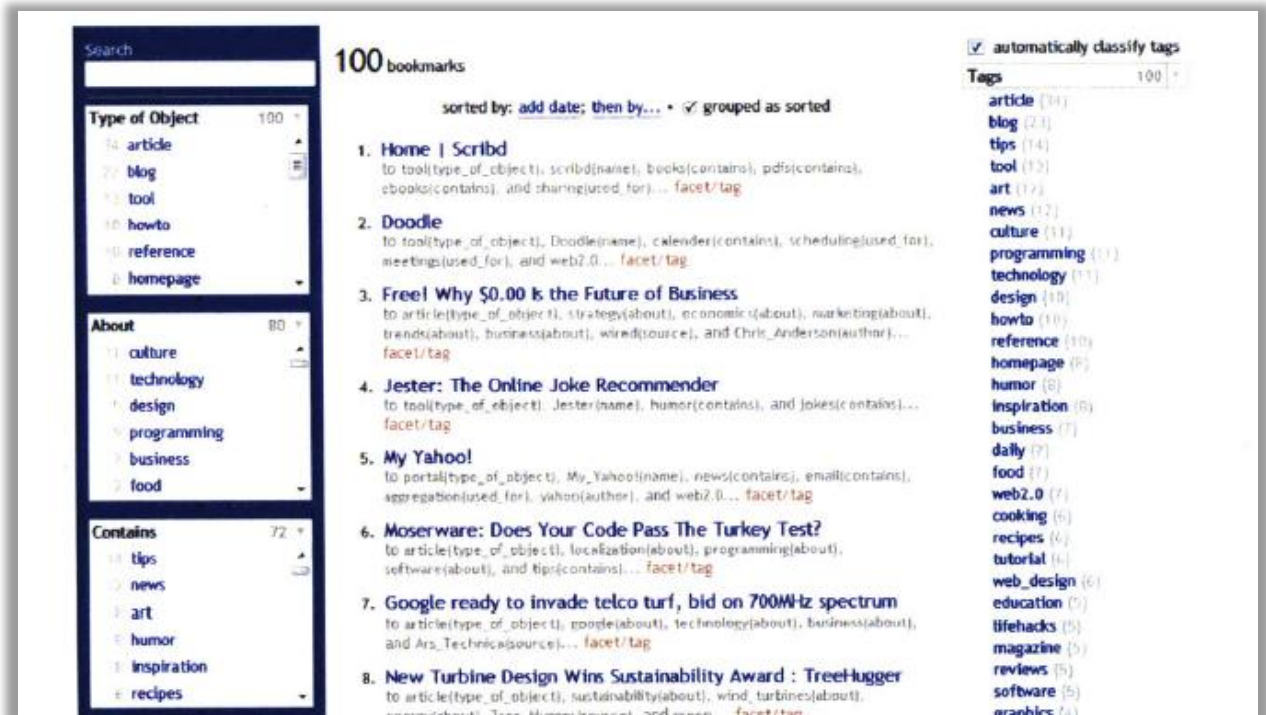
06 AUG 09 Art in the Age SAVE 304

05 AUG 09 Home Page SAVE 2

04 AUG 09 36 Beautiful and "NEW" Free Wordpress Themes | Dzine Blog SAVE 328

02 AUG 09 ThinkGeek :: Mimo Mini USB Monitor SAVE 4

Fonte: Lai (2009, p. 25)

FIGURA 54 - Interface para navegação facetada do *Faccete*

Fonte: Lai (2009, p. 26).

Um projeto similar ao apresentado acima é o *FaceTag*, proposto por Quintarelli, Resmini e Rosati (2007), que também permite a etiquetagem facetada para o compartilhamento de *bookmarks*. O protótipo visa "mostrar como as etiquetas geradas pelo usuário podem efetivamente serem combinadas com um rico esquema de classificação facetada para melhorar a arquitetura de informação de um sistema de etiquetagem social" (QUINTARELLI; RESMINI; ROSATI, 2007, p. 12, tradução nossa).

Neste estudo as facetas foram inspiradas nas trezes categorias fundamentais do CRG, e são: Resource Type (Tipo do Recurso), *Language* (Língua) *Themes* (Temas), *People* (Pessoas), *Purposes* (Propósito) e *Date* (Data). A visualização da etiquetagem proposta pelo *Facetag* é apresentada na FIG. 55. As facetas aparecem abaixo do campo Descrição (*Description*) na imagem. Ao inserir etiquetas em uma faceta, é possível criar hierarquias através do símbolo (>). Além disso, foi usada a técnica com autocompletar para auxiliar a inserção de termos nas facetas.

FIGURA 55 - Etiquetação facetada de *bookmarks* do *Facetag*

New bookmark

Title
The Information Architecture Institute

URL
http://www.lainstitute.org/

Description

Summary Extended text Lithuanian translation

The Information Architecture Institute is a non-profit volunteer organization dedicated to advancing and promoting information architecture. Founded in 2002, the Institute has over 1000 members in 60 countries.

Design HTML

Resource type
web site

Themes
information architecture, information design

People
Reiss, Surla, Starmer, Merholz

Purposes

IA
IA
Information architecture
conferences > Information architecture > IA Summit

Save bookmark

Fonte: Quintarelli, Resmini e Rosati (2007, p. 14).

Na FIG. 56 é apresentada a interface para navegação facetada no *Facetag*. Observam-se vantagens dessa interface em relação à proposta anterior. Com relação a interface para navegação facetada as facetas são associadas a cores (da mesma forma como é feito no *Flamenco*). As etiquetas selecionadas são destacadas em cinza abaixo das facetas. Também é informado o número de *bookmarks* que cada etiqueta contém.

FIGURA 56 - Navegação facetada no *Facetag*

The screenshot shows the Facetag interface with a search bar containing "information architecture". Below the search bar, there are five facets: Language (all), Resource type (article (3), blog (1)), Themes (16 items like deliverables, design, diagrams, folksonomies, intranets, navigation, scent of information, social classification, tag, tagging, usability, navigation), People (5 items like quintarelli, jared spool, stephen turbek), and Purposes (7 items like conference, deliverable, interface design, intranet design, myproject, research). Below the facets, there are 5 bookmarks listed with their titles, authors, dates, and tags.

Language	Resource type (4)	Themes (16)	People (5)	Purposes (7)
all	article (3), blog (1)	deliverables + (2), design (1), diagrams (1), folksonomies (1), intranets (1), navigation design (1), scent of information (1), social classification (2), tag (1), tagging (2), usability (2), navigation (1)	quintarelli (2), jared spool (1), stephen turbek (2)	conference (1), deliverable (1), interface design (1), intranet design (1), myproject (1), research (2)

5 bookmarks

order by insertion (newest / oldest) | order alphabetically (ascending / descending)

Folksonomies: power to the peoples
Folksonomies are web-based collaborative systems for building shared databases of items, enriched by a flat metadata vocabulary that can be used to perform metadata-driven queries,
posted by [andrea](#) at 23 may 2007 20:30 in: [article](#) [folksonomies](#), [information architecture](#), [social classification](#), [tag](#), [tagging](#) [quintarelli](#)
<http://www.infospaces.it/> - cached - mail it - blog this

Real Wireframes Get Real Real Results - Boxes and Arrows: The design behind the design
posted by [Luca](#) at 23 may 2007 20:30 in: [article](#) [information architecture](#), [deliverables](#), [wireframes](#) [stephen turbek](#) [myproject](#)
<http://www.boxesandarrows.com/> - cached - mail it - blog this

GUUUI - Navigation blindness
posted by [emanuele](#) at 23 may 2007 20:30 in: [blog](#) [guuui](#) [information architecture](#), [navigation design](#), [usability](#), [design](#)
<http://www.guuui.com/> - cached - mail it - blog this

Intranet Portals and Scent are Made for Each Other
posted by [emanuele](#) at 23 may 2007 20:30 in: [article](#) [information architecture](#), [usability](#), [intranet](#), [scent of information](#) [jared spool](#)
<http://www.uie.com/> - cached - mail it - blog this

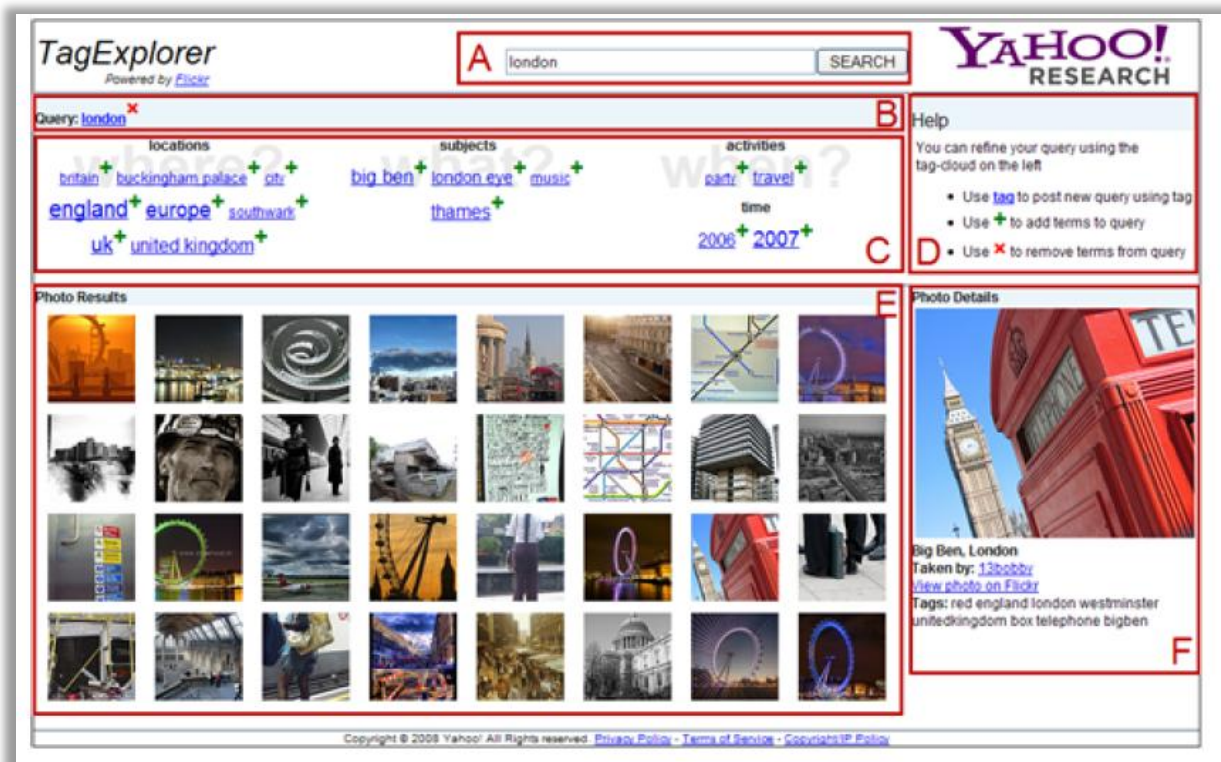
The Lazy IA's Guide to Making Sitemaps - Boxes and Arrows: The design behind the design
Tips for creating site maps in tabular and visual format
posted by [andrea](#) at 23 may 2007 20:30 in: [blog](#) [information architecture](#), [deliverables](#), [sitemaps](#), [diagrams](#) [stephen turbek](#)
<http://www.infospaces.it/> - cached - mail it - blog this

Fonte: Quintarelli, Resmini e Rosati (2007, p. 14).

Esses protótipos, *Faccete* e *FaceTag*, utilizam uma abordagem pragmática para permitir a navegação facetada alimentada por uma comunidade de usuários. É importante observar que, em ambos, as facetas são fixas e diferentes em cada um dos estudos.

O estudo apresentado a seguir é denominado *TagExplorer*, e foi criado em 2010 pelo Laboratório Técnico do *Yahoo!*. Esse estudo propõe a utilização de *tagclouds* facetadas. A proposta foi aplicada a um *site* de compartilhamento de imagens, o *Flickr*. Foram determinadas as facetas: localização (*locations*), assuntos (*subjects*), atividades (*activities*) e data (*time*). A interface do protótipo é apresentada na FIG. 57.

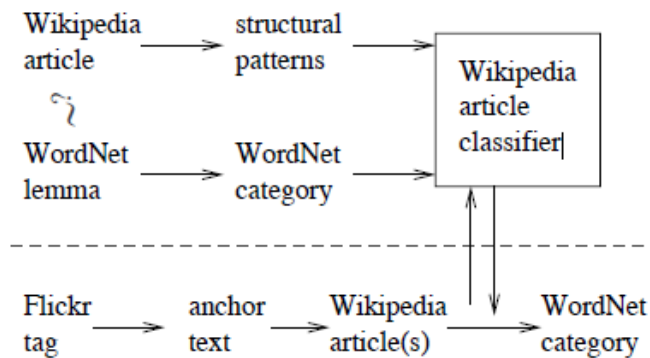
FIGURA 57 - Navegação facetada por *tagclouds* do *TagExplorer*



Fonte: Sigurbjörnsson e Zwol (2010, p. 11).

Pode-se perceber que é possível combinar a busca por palavra-chave (no caso acima, *London*, em B) com a navegação pelas *tagclouds* facetadas (C). Os resultados filtrados aparecem em E. As etiquetas usadas para indexar as imagens são mapeadas para facetas automaticamente, aproveitando a etiquetagem que já foi realizada. O modelo para o mapeamento automático das etiquetas nas facetas combina a utilização do *Wordnet*, um dicionário de língua inglesa, com artigos da Wikipédia. O esquema é mostrado na FIG. 58.

FIGURA 58 - Mapeamento automático das etiquetas para as categorias do *Wordnet*



Fonte: Sigurbjörnsson e Zwol (2010, p. 7).

Uma das grandes vantagens dessa abordagem é que, se existe um repositório de recursos etiquetados pelos usuários, as marcações já existentes podem ser aproveitadas para a construção de uma navegação facetada.

3.2 Classificação facetada colaborativa

O protótipo apresentado a seguir tem o intuito de ser um sistema colaborativo para categorização facetada (*Collaborative Faceted Categorization System*). A ferramenta usada para sua construção foi o sistema de gestão de conteúdo, *Joomla!*, que usa a linguagem de programação PHP e banco de dados MySQL, e é fruto do trabalho do grupo de estudos sobre navegação facetada da Universidade Old Dominion³⁷, de Norfolk, Estados Unidos.

Seu objetivo é permitir que "através de esforços coletivos dos usuários, o esquema de classificação facetado evolua junto com os interesses do usuário e, assim, ajude-os a navegar através da coleção de forma rápida e intuitiva" (MALY; WU; ZUBAIR, 2010, p. 70).

Sua aplicação se deu em uma coleção de imagens da história da África. As imagens são organizadas, primeiramente, em um esquema global de classificação facetada com as facetas: Período de Tempo, Evento, Localização, Gênero, Fotografia e Coleções (FIG. 59).

Esse estudo pode ser considerado aquele que mais se aproxima da presente pesquisa, devido a uma série de características: seu objetivo; as tecnologias utilizadas; a função de permitir que os usuários construam e mantenham a classificação facetada colaborativamente; e o recurso de avaliação dos recursos por parte dos usuários, o que auxilia na manutenção da qualidade dos resultados recuperados na navegação facetada.

³⁷ Site: <<http://facet.cs.odu.edu>>

FIGURA 59 - Interface para classificação facetada colaborativa

The screenshot displays a web interface for collaborative faceted classification. At the top, there are tabs for 'Global' and 'Personal'. The main title is 'MARY MCLEOD BETHUNE, half-length, wearing white blouse and jacket.' with a 71% rating. Below the title is a photograph of Mary McLeod Bethune. To the left of the image are two facet panels:

- Facets (Left):**
 - Time Period: Before 1850(0), 1850 - 1879(2), 1880 - 1899(7), 1900 - 1999(5), 2000 - Present(1), Unknown Time Period(1)
 - Event
 - Location
 - Genre
 - Photographer
 - Collections
- Facets (Right):**
 - Date: Before 1990(2), 1990 - 1999(1), 2000 - Present(1)
 - Location
 - Image Size
 - Image Format
 - Source

Below the image, the following information is displayed:

- Title: MARY MCLEOD BETHUNE, half-length, wearing white blouse and jacket.
- Courtesy of http://www.loc.gov/rr/print/list/083_afr.html
- Tags: [add tags](#)
- Time Period >> 1900 - 1999 [100%] 2
- Event >> Civil Rights [100%] 0
- Location >> South Carolina [100%] 0
- Genre >> People [100%] 0
- Date >> Before 1990
- Location >> South Carolina
- Image Size >> Smaller Than 4x6
- Image Format >> JPEG
- Source >> 20th Century African American Activists

Fonte: Maly, Wu e Zubair (2010, p. 100).

Nesse estudo, o que se propõe, de maneira geral, é deixar o usuário criar as facetas e classes livremente no nível pessoal. No nível global, a manutenção das taxonomias facetadas e a indexação dos recursos *web* através de processos automatizados. Por meio de um processo de cocorrência de etiquetas e categorias, um algoritmo transforma as etiquetas em subcategorias. Nesse processo, ele usa o *Wordnet* para tentar identificar se uma categoria é irmã (mesmo nível hierárquico) ou uma filha (subcategoria).

Ele também usa as etiquetas dos recursos *web* para classificá-los automaticamente nas classes das taxonomias facetadas. O usuário pode votar se a classificação está correta ou incorreta, e tal fato pode acarretar, inclusive, na remoção da indexação.

4 METODOLOGIA

Uma pesquisa pode ser definida como "procedimento formal, com método de pensamento reflexivo, que requer um tratamento científico e se constitui no caminho para conhecer a realidade ou para descobrir verdades parciais" (MARCONI; LAKATOS, 2003, p. 155). Outra definição, que complementa a anterior, é "pesquisa é um conjunto de ações, propostas para encontrar a solução para um problema, que têm por base procedimentos racionais e sistemáticos" (SILVA; MENEZES, 2001, p. 20). As etapas para o desenvolvimento de uma pesquisa são:

1. Seleção do tópico ou problema para a investigação.
 2. Definição e diferenciação do problema.
 3. Levantamento de hipóteses de trabalho.
 4. Coleta, sistematização e classificação dos dados.
 5. Análise e interpretação dos dados.
 6. Relatório do resultado da pesquisa.
- (MARCONI; LAKATOS, 2003, p. 155)

Os estudos sobre a aplicação de modelos de colaboração na classificação e navegação facetada são raros até o momento. Pode-se dizer que esta pesquisa objetiva responder o problema: "como fazer um catálogo *web* facetado que possa ser mantido por uma comunidade de usuários?". O tratamento singular para este problema, proposto no presente estudo, caracteriza sua originalidade, o que é necessário neste nível de trabalho final de pós-graduação.

Para ter cunho científico, uma pesquisa precisa empregar um método científico, que pode ser definido como o "conjunto das atividades sistemáticas e racionais que, com maior segurança e economia, permite alcançar o objetivo – conhecimentos válidos e verdadeiros – traçando o caminho a ser seguido, detectando erros e auxiliando as decisões do cientista" (LAKATOS; MARCONI, 1992, p. 83). Um método científico tem as seguintes etapas:

- a) descobrimento do problema ou lacuna num conjunto de conhecimentos. Se o problema não estiver enunciado com clareza, passa-se à etapa seguinte; se o estiver, passa-se à subsequente;
- b) colocação precisa do problema, ou ainda a recolocação de um velho problema, à luz de novos conhecimentos (empíricos ou teóricos, substantivos ou metodológicos);
- c) procura de conhecimentos ou instrumentos relevantes ao problema (por exemplo, dados empíricos, teorias, aparelhos de medição, técnicas de cálculo ou de medição). Ou seja, exame do conhecido para tentar resolver o problema;
- d) tentativa de solução do problema com auxílio dos meios identificados. Se a tentativa resultar inútil, passa-se para a etapa seguinte; em caso contrário, à subsequente;
- e) invenção de novas ideias (hipóteses, teorias ou técnicas) ou produção de novos dados empíricos que prometam resolver o problema;

- f) obtenção de uma solução (exata ou aproximada) do problema com auxílio instrumental conceitual ou empírico disponível;
- g) investigação das consequências da solução obtida. Em se tratando de uma teoria, é a busca de prognósticos que possam ser feitos com seu auxílio. Em se tratando de novos dados, é o exame das consequências que possam ter para as teorias relevantes;
- h) prova (comprovação) da solução: confronto da solução com a totalidade das teorias e da informação empírica pertinente. Se o resultado é satisfatório, a pesquisa é dada como concluída, até novo aviso. Do contrário, passa-se para a etapa seguinte;
- i) correção das hipóteses, teorias, procedimentos ou dados empregados na obtenção da solução incorreta. Esse é, naturalmente, o começo de um novo ciclo de investigação (LAKATOS; MARCONI, 1992, p. 46).

4.1 Classificação da pesquisa

A presente pesquisa, quanto à natureza, pode ser classificada como pesquisa aplicada, pois "objetiva gerar conhecimento para aplicação prática dirigida a um problema específico" (SILVA; MENEZES, 2001, p. 20). Quanto aos objetivos, como uma pesquisa exploratória, pois este tipo de pesquisa:

proporciona maior familiaridade com o problema, com vistas a torná-lo mais explícito ou a constituir hipóteses em como objetivo principal o aprimoramento de ideias ou a descoberta de intuições. Seu planejamento é, portanto, bastante flexível, de modo que possibilite a consideração dos mais variados aspectos relativos ao fato estudado (GIL, 2002, p. 41).

Nesta pesquisa, foi usado o método comparativo, no qual os dados coletados podem servir para análises qualitativas ou quantitativas. Além disso, por meio desses dados concretos, é possível ampliar e deduzir elementos constantes e gerais (LAKATOS; MARCONI, 1983, p. 82).

Quanto à forma de abordagem do problema, pode-se classificá-la como uma pesquisa qualitativa, pois

considera que há uma relação dinâmica entre o mundo real e o sujeito, isto é, um vínculo indissociável entre o mundo objetivo e a subjetividade do sujeito que não pode ser traduzido em números. A interpretação dos fenômenos e a atribuição de significados são básicas no processo de pesquisa qualitativa. Os pesquisadores tendem a analisar seus dados indutivamente. O processo e seu significado são os focos principais de abordagem (SILVA; MENEZES, 2001, p. 20).

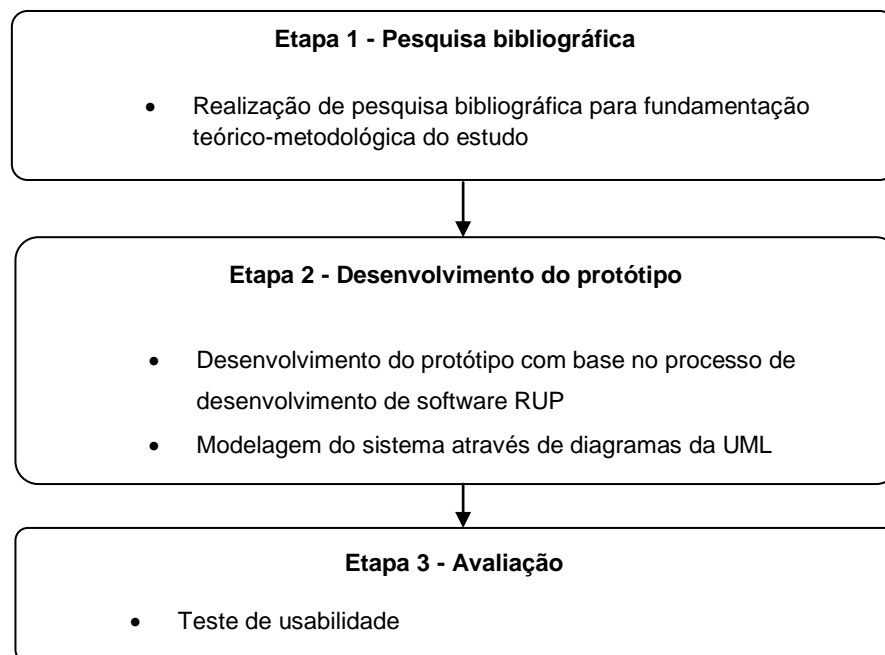
Quanto aos procedimentos técnicos, é classificada como uma pesquisa experimental, visto que "consiste em determinar um objeto de estudo, selecionar as variáveis que seriam capazes de influenciá-lo, definir as formas de controle e de observação dos efeitos que a variável produz no objeto" (GIL, 2002, p. 47).

Contém, também, elementos: 1) de uma pesquisa bibliográfica, pois foi necessário realizar uma vasta busca pela bibliografia da área, como livros, artigos e outros materiais disponíveis, para se conhecer os principais requisitos de uma interface facetada e de sistemas colaborativos; e 2) de uma pesquisa documental, com a realização de uma imersão em diversos artefatos produzidos e trocados durante o desenvolvimento, tais como documentos do projeto, planilhas, *e-mails*, anotações em papel, conversas no bate-papo, entre outras fontes de informação.

4.2 Etapas da pesquisa

Com base na fundamentação teórico-metodológica descrita no capítulo anterior, o percurso metodológico para atingir os objetivos da presente pesquisa (Seção 1.2) é ilustrado na FIG. 60.

FIGURA 60 - Etapas da pesquisa



Fonte: Elaborado pelo autor.

As etapas da pesquisa foram: 1) pesquisa bibliográfica, em que foram usados os conceitos da indexação, da navegação facetada e dos modelos de colaboração da *Web 2.0*; 2) desenvolvimento do protótipo, em que se utilizou o processo de desenvolvimento de *software* RUP, prototipação e a linguagem UML; e 3) avaliação, com o teste de usabilidade

dos usuários no protótipo construído. A seguir, nas próximas seções, as etapas de pesquisa bibliográfica, desenvolvimento do protótipo e avaliação são descritas detalhadamente.

4.2.1 Pesquisa bibliográfica

A pesquisa bibliográfica, primeira etapa dessa pesquisa, foi fundamentada no estudo já apresentado no capítulo 2, tendo sido utilizadas as seguintes bases de dados e repositórios a seguir. À nível nacional: a biblioteca digital de periódicos SciELO³⁸; a biblioteca digital de teses e dissertações (BDTD) do IBICT³⁹; e a BDTD da UFMG⁴⁰. À nível internacional, o arquivo de acesso livre E-LIS⁴¹ com documentos no formato eletrônico na área de Biblioteconomia e Ciência da Informação, a *Networked Digital Library of Theses and Dissertations* (NDLTD)⁴²; a ferramenta de busca especializada para documentos científicos *Scholar* do Google⁴³; a biblioteca digital e motor de busca *CiteseerX*⁴⁴; e a biblioteca digital ACM⁴⁵ para publicações na área de computação e tecnologia da informação. A seguir apresentamos as temáticas abordadas na pesquisa bibliográfica.

A primeira temática abordada foi a recuperação da informação, que se subdividiu nos seguintes temas: processo e modelos de recuperação da informação (LANCASTER, 1993; BAEZA-YATES e RIBEIRO-NETO, 1999); indexação através de vocabulários controlados (NISO, 2005; NAVES, 2004); classificação facetada (CAMPOS, 2001; BATISTA e CARVALHO, 2003; LIMA, 2004; ARAUJO, 2005); taxonomia facetada (TZITZIKAS *et al.*, 2002; HEDDEN, 2010; MACULAN, 2011); indexação por meio da indexação automática (WIVES, 2000; BORGES, 2009); interface de busca, em especial, a interface para navegação facetada, com os tradicionais estudos da área: *Flamenco* (YEE; SWEARINGEN; HEARST, 2003); *Endeca* (TUNKELANG, 2006); e *Facetmap* (SMITH *et al.*, 2006). Foram observados estudos recentes sobre navegação facetada em bases de dados RDF: *FATIC* (TVAROZEK; BIELIKOVÁ, 2007); *HealthFinland* (SUOMINEN; VILJANEN; HYVÖNEN,

³⁸ Disponível em: <<http://www.scielo.org>>.

³⁹ Disponível em: <<http://bdttd.ibict.br>>.

⁴⁰ Disponível em: <<http://www.bibliotecadigital.ufmg.br>>.

⁴¹ Disponível em: <<http://eprints.rclis.org>>.

⁴² Disponível em: <<http://www.ndltd.org>>.

⁴³ Disponível em: <<http://scholar.google.com.br>>.

⁴⁴ Disponível em: <<http://citeseer.ist.psu.edu/>>.

⁴⁵ Disponível em: <<http://dl.acm.org>>.

2007); *BrowseRDF* (OREN; DELBRU; DECKER, 2006); e *Explorator* (ARAÚJO, 2009, p. 50). Foi visualizado o artigo de Hearst (2008), que traz uma revisão de avanços recentes em navegação facetada; os catálogos (MEY, 1995; KAFURE, 2004; SANTANA, 2012) e sua evolução nos catálogos *web* facetados: Niu (2012), Yang (2010); Tzitzikas *et al.* (2002); Makela *et al.* (2005); e *FaThumb* (KARLSON *et al.*, 2006). A qualidade desses estudos nos permitiu reconhecer o formato característico de uma interface facetada, bem como técnicas e requisitos para a navegação facetada eficiente (Seção 2.4.2.6). É através da navegação facetada que se espera ter melhores resultados na recuperação da informação no catálogo *web*.

Na segunda temática abordada foram os modelos de colaboração da *Web 2.0*. Esses modelos são a chave para o sucesso de muitos *sítes* em escala global, que tem o conteúdo criado pelos usuários (VICKERY; WUNSCH-VICENT, 2007). Com relação à etiquetagem social, foram consultados os estudos de Catarino e Baptista (2007); Lopes (2002); Aquino (2007); e Guedes (2010). Com relação ao modelo *wiki*, a própria Wikipédia foi fonte de consulta, com a resolução de disputas (seção 2.7.3.1), e também os artigos: Blattmann e Silva (2007); Vieira, Carvalho e Lazzarin (2008); e Cirilo (2011). Para assegurar uma maior relevância dos resultados recuperados, pesquisou-se sobre sistemas de votação (LERMAN, 2007) e funcionalidade, observada também em várias interfaces facetadas apresentadas. Para representar, classificar e analisar a colaboração da *Web 2.0* foram abordados estudos referentes a sistemas colaborativos (ELLIS; GIBBS; REIN, 1991; DAMES, 2004), o modelo de colaboração 3C (LOTUS, 1995; GEROSA, 2006; PIMENTEL, 2006) e fluxo de trabalho (CASATI *et al.*, 1995; HOLLINGSWORTH, 1995; GEORGAKOPOULOS; HORNICK; SHETH, 1995). A partir desses estudos, buscou-se identificar as características de um modelo de colaboração para um catálogo *web* que pudesse expandir e ser mantido por uma comunidade de usuários.

Foram pesquisados estudos que fazem a união das duas temáticas anteriores, principalmente a navegação facetada e a colaboração na *Web 2.0*, para compor a revisão de literatura. As seguintes pesquisas encontradas apontam uma abordagem para a construção colaborativa de repositórios de mídia ou catálogos, através da etiquetagem facetada: *Faccete* (LAI, 2009), *FaceTag* (QUINTARELLI; RESMINI; ROSATI, 2007); e o *Faceted Index Internet Gateways* (DEVADASON *et al.*, 2002). Os estudos *TagExplorer* (SIGURBJÖRNSSON; ZWOL, 2010) e *Collaborative Faceted Categorization System* (MALY; WU; ZUBAIR, 2010, p. 70). Tais pesquisas trazem uma abordagem diferente, e acrescentam à etiquetagem social, a utilização da base de dados léxica *Wordnet* para o mapeamento de termos nas hierarquias das facetadas.

O presente estudo faz uso de vários elementos desses estudos e aponta outra abordagem, combinando a etiquetagem social com a taxonomia facetada para indexação dos registros por meio de um modelo de colaboração, que é detalhado nas próximas seções.

4.2.2 Desenvolvimento do protótipo

Foi adotado o processo de desenvolvimento de *software* RUP (seção 2.8.1) e a linguagem de modelagem UML (seção 2.8.2) para a implementação do protótipo. O processo de desenvolvimento também foi inspirado por um estudo similar, a tese de mestrado intitulada *Um Modelo Para Desenvolver Editores Colaborativos Hierárquicos*, realizada no ITA, que tinha o objetivo de criar um ambiente que permitisse a elaboração colaborativa de grandes documentos através de um sistema de gestão de conteúdo (SANTOS, 2009). Descrevem-se as fases para o desenvolvimento do protótipo nas subseções a seguir.

4.2.2.1 Concepção

Nessa fase foi planejada a elaboração dos objetivos e escopo do projeto com base nas definições da seção anterior (4.2.1). Isso é detalhado no capítulo seguinte (cap. 5), do desenvolvimento do protótipo.

4.2.2.2 Elaboração

A fase de elaboração foi composta pelas atividades a seguir:

- Especificação dos requisitos funcionais e não-funcionais do sistema, com base nos objetivos e no escopo do projeto da fase de concepção;
- Escolha da ferramenta, *software* livre que será usado como *framework*, que é um conjunto de classes e funções comuns para uma família de aplicações (SAUVÉ, 2000). Dessa forma, é possível fazer uso dessa plataforma, para agilizar o desenvolvimento, implementando apenas o que é específico da sua aplicação;
- Elaboração de diagramas de casos de uso, com a especificação da interação dos usuários com o sistema;
- Elaboração de diagrama de atividade com o fluxo de trabalho dos componentes do sistema;

- Prototipação de baixa fidelidade, que permite visualizar como pode ser a interface do sistema, permitindo ajustes antes da implementação.

A fase de construção no qual o protótipo é implementado é descrita a seguir.

4.2.2.3 Construção:

Essa fase foi constituída pelas seguintes atividades:

- Criação do ambiente, com a instalação do *software* livre adotado, configuração e instalação dos módulos necessários;
- Customização do módulo, que será adaptado para alcançar os objetivos do sistema;
- Implementação da interface para navegação facetada, codificação das facetas e outros elementos da interface facetada;
- Implementação do modelo colaborativo, codificação das funcionalidades relacionadas ao modelo colaborativo para indexação por meio da taxonomia facetada e da folksonomia;
- Testes de unidade, com a realização de testes *Ad-Hoc* (seção 2.8.4.1) ou *Quick and Dirty* (seção 2.8.4.2), para verificar se o requisito foi implementado, realizando ajustes e correções, se necessário.

4.2.2.4 Transição

Essa etapa envolve a implantação do sistema em um ambiente real para utilização pelos usuários. No contexto desta pesquisa, isso é realizado através do teste de usabilidade, que é tratado na próxima seção.

4.2.3 Teste de usabilidade

Dentre as técnicas para teste de *software* apresentadas na seção 2.5.3, foi escolhido realizar o teste de usabilidade, em que são observadas as ações dos usuários na realização de tarefas. O plano de teste é apresentado a seguir.

4.2.3.1 Objetivos

Em conformidade com os objetivos da seção 1.2, optou-se por realizar o teste de usabilidade por duas razões principais: busca-se conhecer o comportamento do usuário na realização da tarefa de indexação em uma interface que oferece duas estratégias opcionais, a etiquetagem ou a taxonomia facetada; deseja-se obter conhecimento sobre a preferência do usuário durante o ato de recuperação da informação em uma interface que oferece a busca por palavra-chave e a navegação facetada. Além desses dois objetivos, deseja-se verificar se registros com as maiores notas são considerados mais relevantes pelos usuários.

4.2.3.2 Ambiente de teste e preparação

O ambiente de teste escolhido foi o laboratório de informática da Faculdade de Informação e Comunicação-FIC/UFG, devido à facilidade de acesso aos participantes. O teste foi projetado para ser realizado com todos os participantes ao mesmo tempo.

Trata-se de um teste de usabilidade que utiliza o método de observação indireta, por ser menos invasivo, no qual as ações dos usuários são gravadas através da captura do vídeo do computador. Para a realização do teste cogitou-se utilizar o *software* Morae, da Camtasia, que é dedicado a testes de usabilidade. Entretanto, era necessário gravar as ações dos usuários nos computadores do laboratório de informática da FIC, cujos computadores têm o sistema operação Linux, com a distribuição Ubuntu. Assim, após reunião com a equipe de informática da FIC, foi descoberto o programa *RecordMyDesktop* que funciona nesse sistema operacional e que faz a captura das telas do computador com suporte a vários formatos de vídeo. Os vídeos foram gravados no formato aberto ogv. A instalação do *RecordMyDesktop* foi muito simples, com a busca do programa pela própria ferramenta de pesquisa de aplicativos do Ubuntu. Após a instalação, aparece um ícone no canto esquerdo inferior da tela, que dá acesso ao aplicativo.

Para realização das tarefas, o protótipo do catálogo *web* facetado colaborativo foi acessado através do navegador de Internet *Mozilla Firefox*, disponível na distribuição do Ubuntu. Foi elaborado um termo de consentimento (APÊNDICE A) a ser entregue para os participantes antes do início da sessão de testes para a autorização do uso das informações coletadas durante os mesmos.

4.2.3.3 Participantes da pesquisa

Como mencionado no tópico anterior, para minimizar o efeito que o conhecimento de área, por parte do usuário, poderia exercer sobre o teste, optou-se por utilizar como participantes do teste alunos do curso superior, com o intuito de trabalhar com

um perfil mais uniforme e com pessoas com um conhecimento mínimo do assunto. Foram selecionados dez alunos voluntários do quarto período do curso de graduação em Gestão da Informação da Universidade Federal de Goiás (UFG). Tal fato permitiu que os usuários escolhidos tivessem um conhecimento básico, com experiência de, no mínimo, dois anos em navegação pela Internet. Esse número de participantes foi suficiente e representou uma margem satisfatória em caso de indisponibilidade ou falha de algum computador (mesmo se a metade dos participantes não pudesse comparecer ao teste, o número ainda seria razoável, conforme seção 2.8.4.4.3).

4.2.3.4 *Design* do teste

Para a avaliação do catálogo *web* facetado colaborativo, o protótipo foi configurado para servir como um catálogo de empresas, ou páginas amarelas na Internet, a mesma aplicação do estudo *Semantic Yellow Page Service Discovery: The Veturi Portal*, referenciado na seção 2.5.2.1. Nesse catálogo, os registros catalogados representam as empresas.

4.2.3.5 Lista de tarefas

A lista de tarefas foi prevista para ser executada em aproximadamente 50 minutos. Sendo assim, o voluntário foi instruído a:

1) Cadastrar empresas a partir de uma lista de empresas fornecidas. Não foi fornecida a categoria da empresa, apenas o bairro em que ela se localizava. A categoria ou área de atividade da empresa deveria ser deduzida a partir do seu nome. Os outros dados fornecidos sobre a empresa foram: telefone, endereço, bairro, CEP e cidade. Foi fornecida uma lista com cerca de vinte empresas de regiões e áreas de atividade aleatórias. Foram dados 10 minutos para o usuário cadastrar a quantidade de empresas que conseguisse e foi explicitamente mencionado que o usuário poderia utilizar as categorias fornecidas ou usar suas próprias palavras (palavras-chave) para descrevê-las.

2) A segunda tarefa foi de recuperação da informação. Nesse teste, as tarefas de busca se dividem em dois tipos: busca simples, quando aquilo que se quer encontrar tem o mesmo nome daquilo que é digitado na busca; e busca complexa, quando, a partir da necessidade, é preciso abstrair a palavra-chave a ser digitada.

Os candidatos dispuseram de 15 minutos para a execução das duas tarefas, e foram esclarecidos de que não era o usuário que estava sendo avaliado, e, sim, a interface. Esta tarefa foi dividida nos seguintes tipos de buscas reais : a) Encontrar cinco empresas para hospedagem no bairro Jardim Vitória, em Goiânia (busca complexa na

área de atuação e busca direta no bairro); b) Encontrar três empresas que vendem sanduíche nos bairros próximos à UFG (busca complexa na área de atuação e no bairro); e c) Encontrar duas empresas voltadas para reforma de casas nas especialidades de pintura, janela de alumínio e instalação elétrica (busca simples e complexa na área de atuação).

3) A tarefa seguinte foi a votação, momento em que cada usuário elegeu as melhores empresas do tipo faculdade ou universidade.

4) A última tarefa solicitava que o usuário revisasse as classificações que um outro usuário realizara. O usuário revisor poderia apontar qualquer erro na classificação que fosse detectado, sendo que, na última tarefa, um terço das empresas havia sido, propositalmente, cadastrada de maneira indevida.

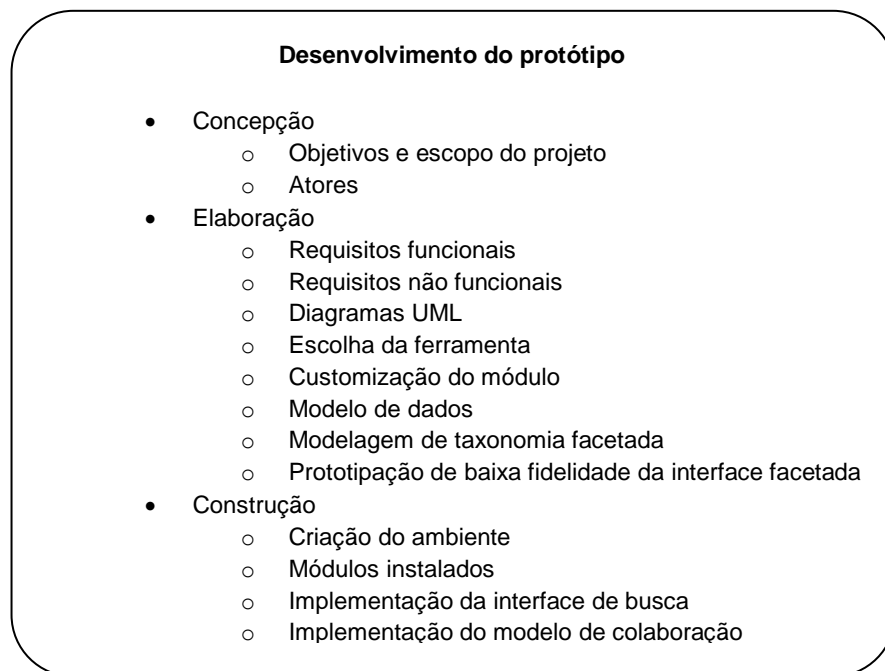
4.2.3.6 Análise de dados

Após a realização dos testes, o procedimento para análise de dados teve início com a transcrição dos questionários, fichas e ações dos usuários em planilhas eletrônicas no computador. Com base nesses registros, foram geradas tabelas sintetizadoras com estatísticas e gráficos para evidenciar padrões que são detalhados no capítulo 6.

5 DESENVOLVIMENTO DO CATÁLOGO WEB FACETADO COLABORATIVO

O desenvolvimento do catálogo *web* facetado colaborativo (*Facetlog*) foi baseado no processo de desenvolvimento de *software* RUP e seguiu o caminho metodológico definido na seção 4.2.2. Foi realizado um desenvolvimento iterativo e incremental, através do qual se obtém, a cada ciclo de desenvolvimento, uma nova versão do *software*, que é avaliado, apontando-se suas deficiências e melhorias possíveis. Tal fato confere subsídios para gerar o escopo de desenvolvimento do próximo ciclo, repetindo o processo sucessivamente, até atender às necessidades do projeto. A seguir na FIG. 61 são descritas as etapas e resultados do último ciclo de desenvolvimento. Os ciclos de desenvolvimento anteriores estão descritos no APÊNDICE G.

FIGURA 61 - Etapas do desenvolvimento do protótipo



Fonte: Elaborado pelo autor.

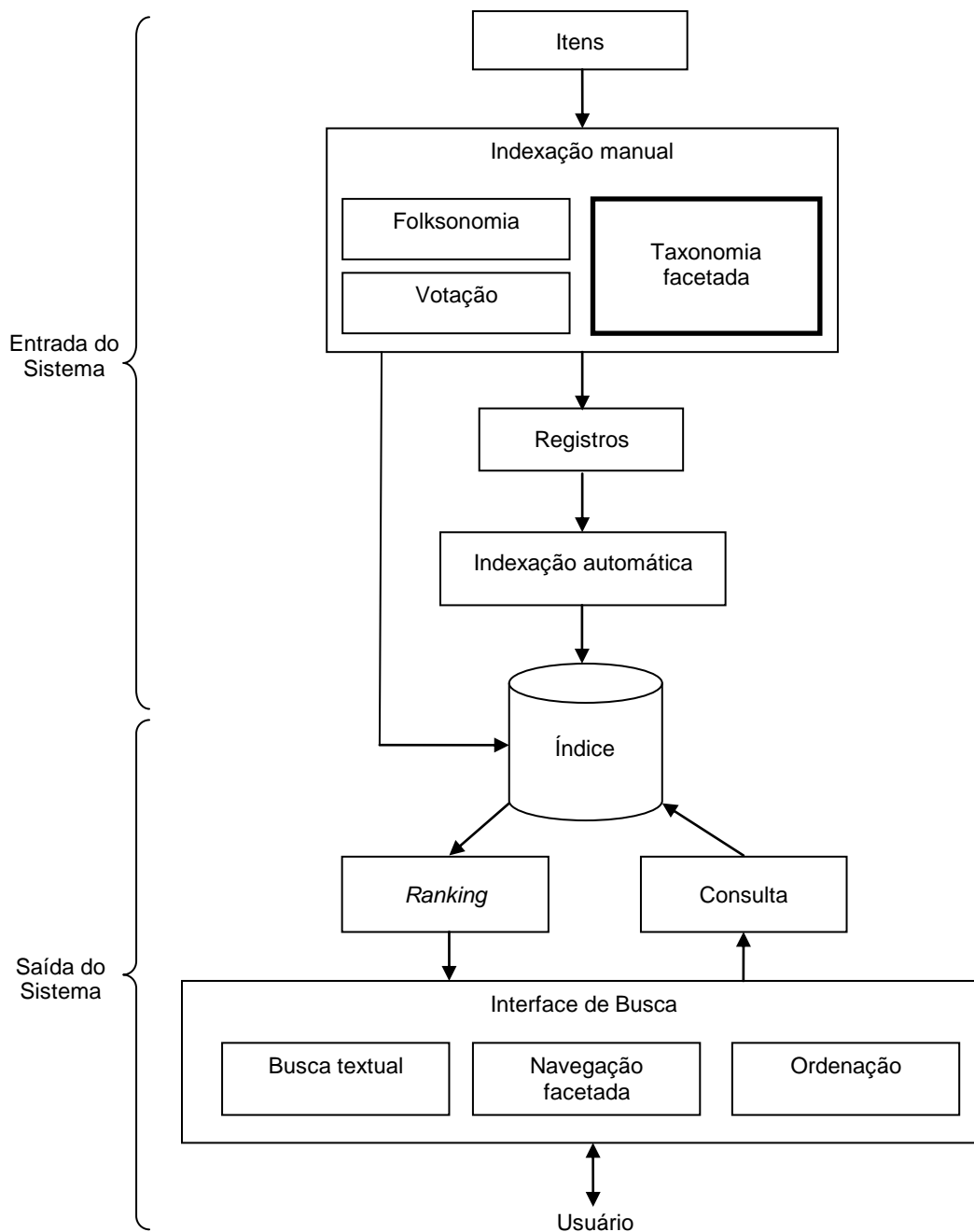
5.1 Concepção

Essa fase é voltada para a elaboração de uma visão geral do sistema, com a delimitação do seu escopo, objetivos e atores do sistema.

5.1.1 Objetivos e escopo do projeto

Com base nos processos de recuperação da informação da seção 2.1.1, foi elaborado um diagrama que representa o modelo de recuperação da informação proposto no *Facetlog*, objeto do presente estudo, que é ilustrado na FIG. 62.

FIGURA 62 - Modelo de recuperação da informação do *Facetlog*

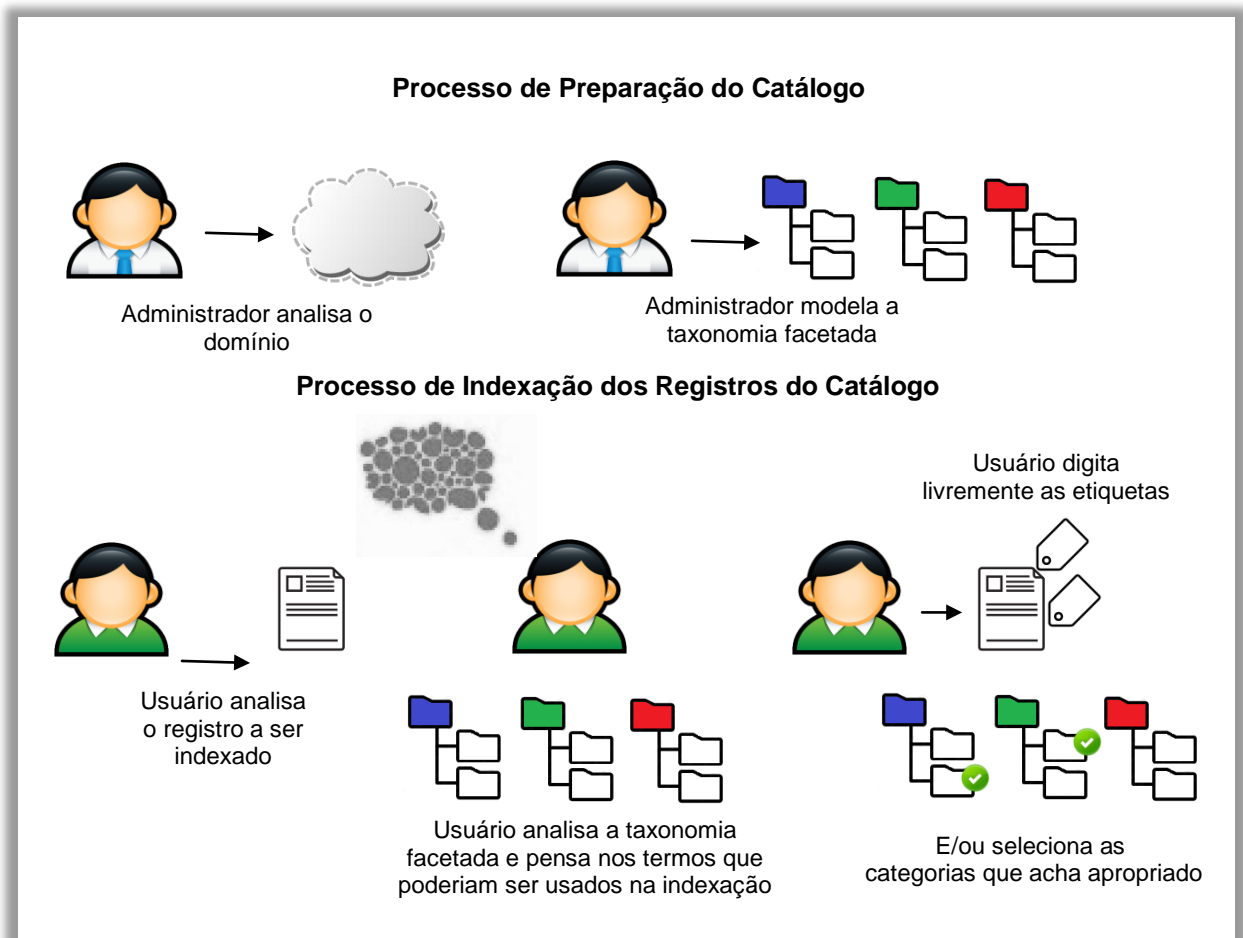


Fonte: Elaborado pelo autor.

O diagrama anterior foi elaborado com base na pesquisa bibliográfica das temáticas apresentadas na seção 4.2.1 e nos objetivos principais da pesquisa (seção 1.2) que são: permitir aos membros da comunidade de usuários realizar a indexação dos registros do catálogo *web* por meio de um modelo de colaboração; e permitir a busca por palavra-chave e a navegação facetada como estratégia de busca no catálogo *web*.

A principal inovação desse modelo de SRI em relação aos modelos apresentados na seção 2.1.1 é a presença da taxonomia facetada, como sendo parte opcional do processo de indexação na entrada do SRI combinando os processos apresentados na seção 2.1.2 (FIG. 12). Previamente a esse processo, é necessário que o administrador determine as facetas e suas categorias conforme o domínio de aplicação do catálogo. É pressuposto que o administrador do catálogo tenha conhecimentos sobre os princípios da classificação facetada e indexação para a modelagem da taxonomia facetada, aprovação ou repreensão (moderação) das contribuições dos usuários. Os usuários não atuam colaborativamente na modelagem da taxonomia facetada e sim na indexação dos registros cadastrados no catálogo (FIG. 63).

FIGURA 63 - Processo de preparação e indexação do *Facetlog*



Fonte: Elaborado pelo autor.

Na saída do sistema, a existência da busca por palavra-chave com a navegação facetada representa uma tendência das novas interfaces de busca. Dessa forma, a busca por palavra-chave está associada à folksonomia e à indexação automática, e a navegação facetada está associada à indexação por meio da taxonomia facetada. Entretanto, o usuário pode realizar consultas que combinam os dois filtros ao mesmo tempo, por exemplo, digitando uma palavra-chave e selecionando um termo na taxonomia facetada.

Com relação à colaboração, na pesquisa bibliográfica identificam-se os seguintes modelos de colaboração:

- Pré-aprovados, como no sistema *wiki*, no qual, por padrão, o usuário tem o poder de realizar uma ação, que pode ser revertida por outros usuários;
- Por permissão, no qual apenas usuários com determinado perfil podem executar a ação;
- Por aprovação, que realiza um fluxo de trabalho antes de executar ou publicar o produto da ação realizada.

Os modelos de colaboração por aprovação podem, ainda, ser divididos em:

- Hierárquicos, no qual um usuário com um determinado perfil tem poder para aprovar a ação de usuários com perfis com menores poderes, por exemplo, o administrador ou moderador aprovar a ação de um recente usuário que se tornou membro;
- Através de sistemas de votação, que é um processo mais horizontal e democrático, no qual os usuários, em conjunto, aprovam, ou não, uma determinada ação realizada por outro usuário.

O presente estudo opta por usar, no catálogo *web* facetado, o seguinte modelo de colaboração:

- Por permissão, no qual apenas os administradores realizam a gestão do catálogo e da taxonomia facetada, para garantir uma qualidade mínima na taxonomia facetada disponibilizada;
- Pré-aprovado, no qual qualquer usuário membro do *site* que estiver logado pode realizar a inserção de um novo registro no catálogo. Qualquer visitante pode se tornar um membro do *site*, o que permite a maior quantidade de contribuições possíveis;
- Por aprovação e hierárquico, no qual qualquer membro pode contribuir com a indexação de registros de outros membros. Entretanto, essa ação deve ser aprovada pelos administradores, de maneira similar ao que ocorre no *software Koha* (p. 73). A aprovação foi escolhida para garantir

um maior controle sobre aquilo que é vinculado a registros de outros membros, coibindo o vandalismo e informações não relevantes ou inadequadas;

- Por votação, no qual qualquer membro pode avaliar a qualidade de um registro, atribuindo-lhe uma nota. Em seguida, calcula-se uma média das notas atribuídas. Isso tem o intuito de promover ganhos na recuperação da informação com a ordenação por meio do critério “mais bem votados”.

5.1.2 Atores

Os atores do sistema (perfis de usuário) são descritos a seguir:

- Super administrador, usuário responsável pela gestão das configurações do sistema;
- Administrador, usuário responsável pela criação, configuração do catálogo, taxonomia facetada e aprovação de contribuições dos membros;
- Membro, usuário cadastrado e logado no sistema;
- Visitante, usuário que não está cadastrado ou não está logado no sistema. Também conhecido com usuário anônimo.

5.2 Elaboração

Na segunda fase do desenvolvimento foi detalhada a arquitetura do sistema, com a elaboração dos requisitos funcionais e não funcionais, diagramas de caso de uso e atividade, escolha da ferramenta, modelo de dados e prototipação de baixa fidelidade da interface.

5.2.1 Requisitos funcionais

Conforme os objetivos propostos na seção 1.2, foram formulados os seguintes requisitos funcionais do catálogo *web* que o tornam colaborativo e facetado:

Referente à gestão do catálogo (apenas para administradores):

- Criar um catálogo
 - Determinar o nome do catálogo, o tipo de item que será cadastrado (ex. obras literárias, filmes, músicas, endereços);
 - Indicar o plural do item cadastrado e se é masculino ou feminino;
 - Escolher as facetas do catálogo dentre as categorias já existentes ou criar novas facetas;
 - Criar as categorias iniciais de cada faceta;
 - Explicitar o tipo de votação que está sendo realizada (ex. melhor, pior, mais barato, mais econômico).

Referente à interface de busca (para todos os usuários):

- Buscar um registro, filtrando os registros através da seleção das categorias das facetas, ou das palavras-chave digitadas na caixa de busca, ou pela combinação de ambas as estratégias. Atende ao objetivo de combinar a navegação facetada com busca por palavra-chave;
- Ordenar os registros da busca pelo critério “data de publicação”, “mais avaliados”, “mais populares”. Permite ao usuário ver os recursos mais acessados ou considerados mais importantes pelos próprios membros do *site*;
- Navegar por uma nuvem de etiquetas (*tagcloud*) na tela principal de navegação. Permite ao usuário navegar pelas etiquetas.

Referente à contribuição no catálogo (para todos os membros):

- Adicionar um registro, utilizando os termos das facetas, através da etiquetagem livre ou ambas as estratégias para classificação;
- Contestar a indexação de um registro. Permite ao usuário revisar a indexação do mesmo adicionando novas categorias das facetas, ou informar eventual classificação erroneamente aplicada (deve enviar um *e-mail* para o administrador do catálogo). Assim, o usuário pode contribuir com o catálogo, melhorando a qualidade da indexação. O administrador deve aprovar a contestação realizada para que seja efetivada;
- Denunciar abuso de um registro no catálogo. Voltado principalmente para questões de vandalismo ou conteúdo impróprio, permite ao usuário informar se um registro cadastrado livremente deve ser retirado do ar. O administrador ou moderador pode aprovar ou rejeitar a contestação;
- Adicionar comentários aos registros;

- Criar um *ranking* dos membros com maior quantidade de registros cadastrados.

Referente ao sistema de votação (para todos os membros):

- Votar, atribuindo uma nota a um registro para um determinado critério avaliado (ex. melhor, mais barato). Tal ação permite que a própria comunidade eleja os registros segundo o critério, permitindo uma ordenação na busca e, conseqüentemente, maior precisão.

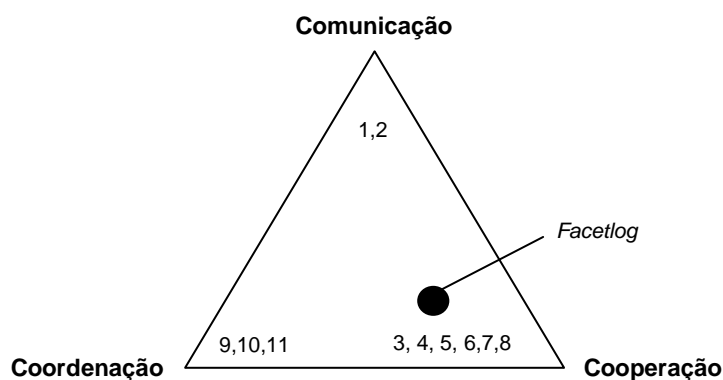
Analisando a colaboração proposta através do modelo de colaboração 3C, de forma similar ao realizado no estudo de Oliveira, Gerosa e Marco (2011), pode-se classificar os requisitos estipulados, com exceção da gestão do catálogo restrita aos administradores, conforme o QUADRO 4 a seguir.

QUADRO 4 - Análise através do modelo de colaboração 3C

Atividade (modelo 3C)	Requisito
Comunicação	1. Alertar nova contestação através de envio de <i>e-mail</i>
	2. Adicionar comentário ao registro
Cooperação	3. Pesquisar registro através da busca por palavra-chave e/ou navegação facetada
	4. Ordenar registros
	5. Navegar pela nuvem de etiquetas
	6. Adicionar e indexar registro
	7. Revisar a indexação do registro de terceiros (contestação)
	8. Listagem do ranking dos usuários
Coordenação	9. Aprovação ou não de nova classificação pelo administrador
	10. Aprovação ou não de exclusão classificação pelo administrador
	11. Denunciar abuso (solicitar remoção do registro impróprio)

Fonte: Elaborado pelo autor.

Na FIG. 64 foram mapeados os requisitos do *Facetlog* para o triângulo do modelo de colaboração 3C. Posteriormente, o *Facetlog* foi situado mais próximo à aresta que tinha, proporcionalmente, a maior quantidade de requisitos. Percebe-se que o *Facetlog* dá ênfase, principalmente, ao espaço de compartilhamento da informação caracterizado pela interface de busca (cooperação). Em segundo lugar, tem-se a coordenação, utilizada para melhorar a qualidade da informação através da moderação e, assim, maior relevância dos resultados. E, em terceiro lugar, a comunicação, a dimensão menos presente no *Facetlog*.

FIGURA 64 - Triângulo do modelo de colaboração 3C do *Facetlog*

Fonte: Elaborado pelo autor.

5.2.2 Requisitos não funcionais

Os requisitos não funcionais do catálogo *web* facetado colaborativo são apresentados a seguir: o catálogo deve oferecer a possibilidade de acesso *on-line* por meio de um navegador de Internet utilizado como programa cliente, que solicita, recebe e interpreta o código em HTML, *Javascript* e CSS. Permite que o protótipo seja acessado de qualquer lugar através da Internet e por vários usuários ao mesmo tempo; devem ser utilizados *softwares* livres, que permitem a realização, o desenvolvimento do aplicativo de forma mais rápida, dando continuidade às pesquisas já realizadas e contribuindo para a comunidade de *software* livre. Além disso, evita custos com licenças de *software* tanto no ambiente do usuário, quanto no ambiente do desenvolvedor:

- O *framework*, a ser escolhido, deve possuir as seguintes características: deve ser simples, rápido e flexível para customizar com uma boa curva de aprendizado; permitir que suas funcionalidades sejam estendidas através do desenvolvimento e instalação de novos módulos; deve ser bem documentado do ponto de vista de usuário, administrador ou desenvolvedor; possuir funcionalidades nativas tais como: gestão de recursos *web*, páginas e *links*; indexação dos recursos *web* através da etiquetagem social e taxonomias; ferramenta de busca com indexação automática; cadastro de usuários, controle de acesso e perfis por um super administrador do sistema; e um sistema de temas para formatar a informação impressa na tela;

- O sistema deve ter um tempo de resposta inferior a dois segundos para não tornar lentas e cansativas as tarefas de busca e navegação.

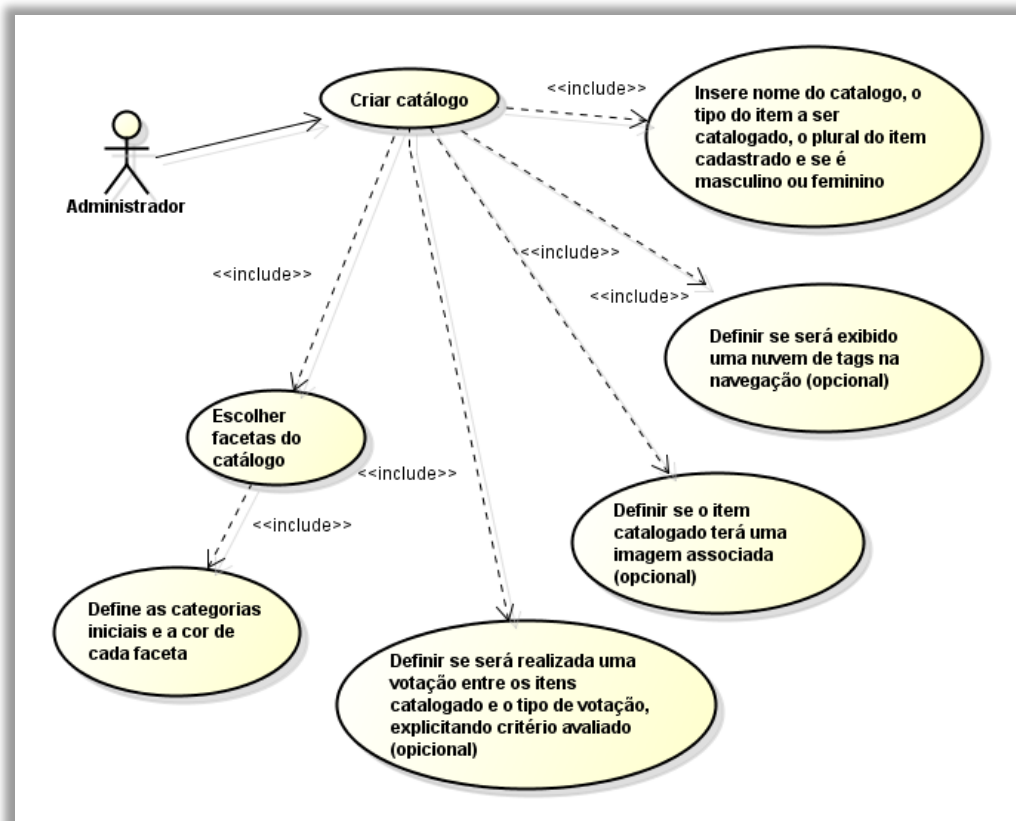
5.2.3 Diagramas UML

Nessa seção são apresentados os diagramas de caso de uso e de atividade realizados para orientar a implementação neste ciclo de desenvolvimento.

5.2.3.1 Diagramas de caso de uso

Os diagramas de caso de uso foram usados para detalhar as funcionalidades e ações do usuário no catálogo *web* facetado colaborativo. O diagrama de caso de uso a seguir refere-se à função Criar Catálogo (FIG. 65).

FIGURA 65 - Criar catálogo

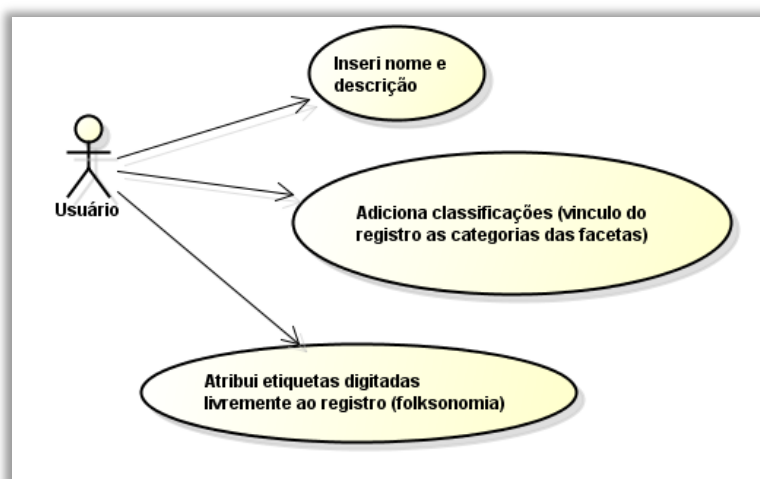


Fonte: Elaborado pelo autor.

Na função Criar Catálogo, o administrador define as configurações do catálogo, que devem ser divididas em configurações: a) ontológicas, definição do tipo do item a ser registrado, se masculino ou feminino e o seu plural; b) de interface, cores das facetas, exibição ou não da nuvem de etiquetas, da imagem associada ao registro e da votação entre os registros; e c) da taxonomia facetada, com a criação das facetas do catálogo e suas categorias.

Após a criação do catálogo, o usuário pode adicionar um registro (caso de uso da FIG. 66). Essa função é composta pelas subfunções: representação descritiva, com a inserção do nome e descrição do registro; representação temática do mesmo, com a sua indexação através da etiquetagem e seleção de categorias das facetas (classificações).

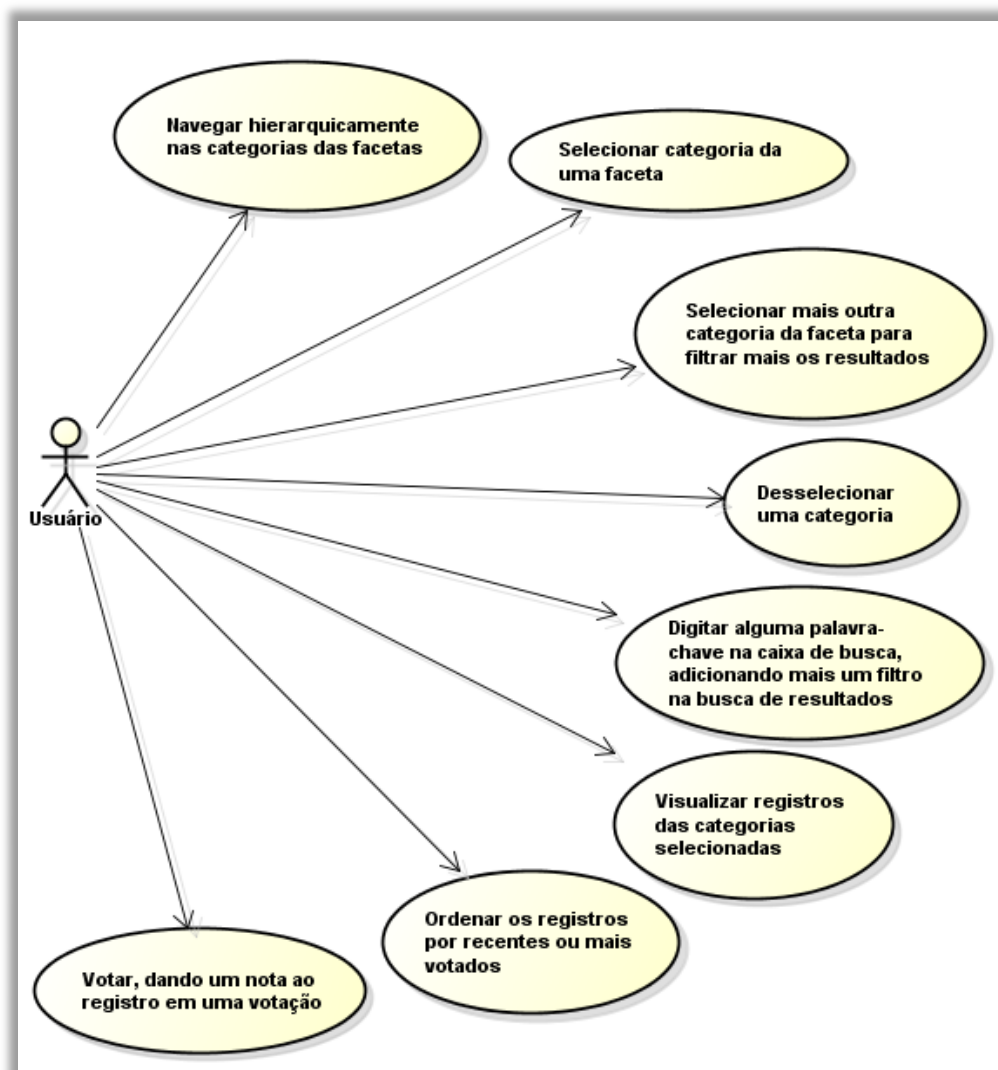
FIGURA 66 - Adicionar registro



Fonte: Elaborado pelo autor.

Após a inserção de registros no catálogo pode ser realizada a função Buscar Registro, ilustrada no diagrama de caso de uso na FIG. 67, sendo composto das seguintes subfunções: navegar hierarquicamente pelas categorias da faceta; selecioná-las e filtrar os registros recuperados; digitar palavras-chave para recuperar registros que contenham em sua representação descritiva (indexação automática) ou em suas etiquetas (etiquetagem social); combinar filtros para restringir mais a consulta; realizar a ordenação dos registros sob o critério de mais recentes, mais populares ou mais avaliados para recuperar os mais relevantes; e, votar para eleger registros mais relevantes para uma determinada consulta. É importante observar que essa função não é parte da estratégia de busca, mas, ocasionalmente, permitir que o usuário possa atribuir uma nota ao registro, realizando uma indexação referente à qualidade daquele do mesmo. A última subfunção é referente à remoção dos filtros, sejam eles categorias selecionadas ou palavras-chave digitadas.

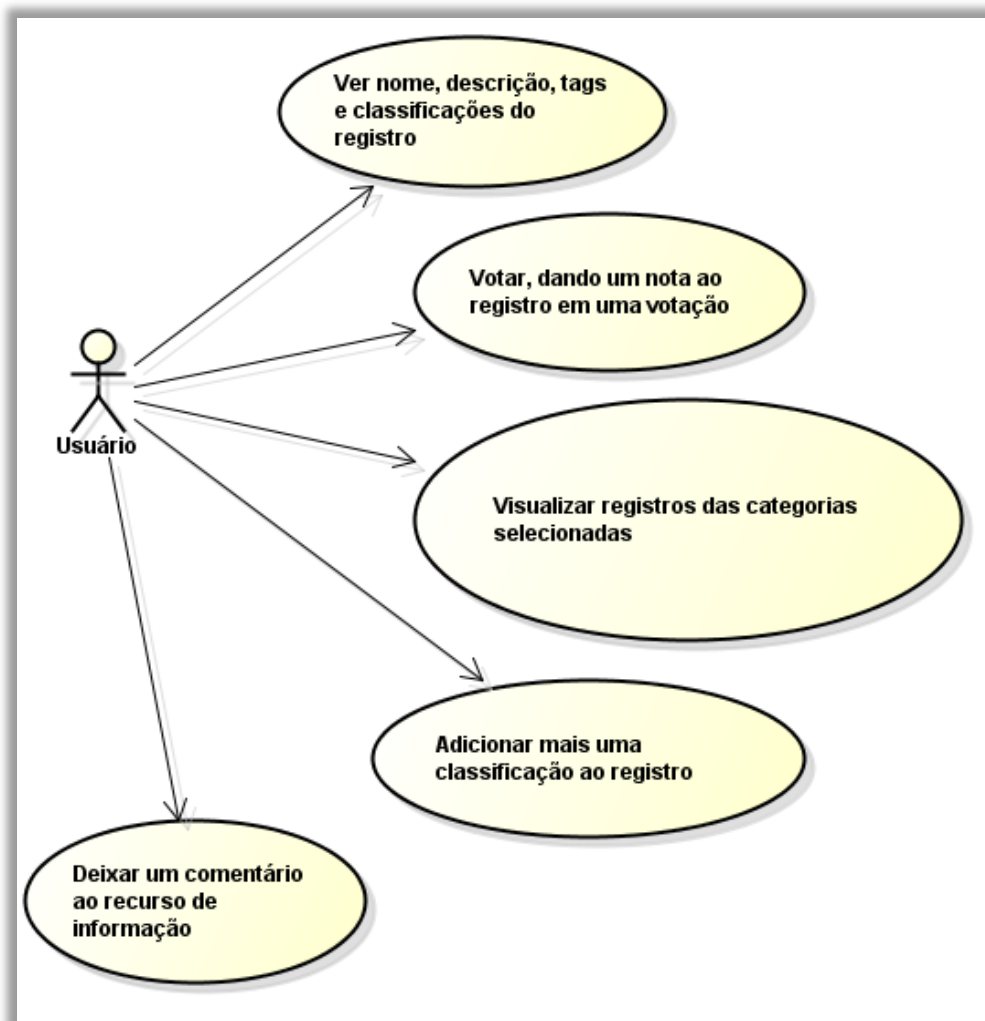
FIGURA 67 - Buscar e votar em registro



Fonte: Elaborado pelo autor.

O diagrama de caso de uso a seguir (FIG. 68) trata da visualização do registro, evento que ocorre após a recuperação e a verificação de que é relevante para o usuário. Qualquer usuário pode visualizar registros da mesma categoria a que este pertença. Entretanto, existem funções reservadas apenas aos membros: votar, quando é atribuída uma nota ao registro; contestar suas classificações; e, adicionar comentários a ele. Ao contestar as classificações do registro, o membro pode adicionar uma nova classificação ou solicitar a remoção de uma ou mais classificações atuais que o mesmo possui. O membro pode apresentar o motivo da adição ou remoção de uma classificação, através do campo de descrição, com o intuito de embasar a sua solicitação. Entretanto, o preenchimento desse campo é opcional.

FIGURA 68 - Visualizar e indexar dos registros

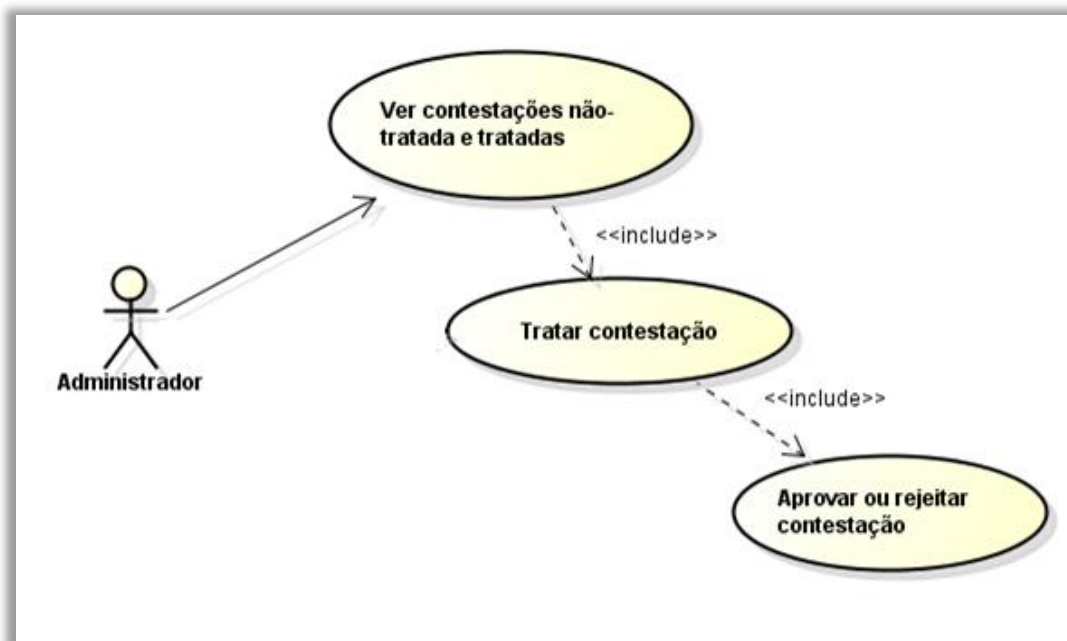


Fonte: Elaborado pelo autor.

O diagrama de caso de uso da FIG. 69 completa o ciclo do modelo de colaboração iniciado quando um membro contesta a classificação de um registro. O tratamento das contestações é realizado apenas pelo administrador do catálogo, que realiza as seguintes interações: verifica as contestações não-tratadas e tratadas; trata a contestação.

Ao tratar uma contestação, realiza duas subfunções: aprovar ou rejeitar a contestação, que pode ser do tipo adicionar uma nova classificação ou remover uma classificação existente; e, descrever o motivo que levou à tomada de decisão (opcional).

FIGURA 69 - Tratar as contestações



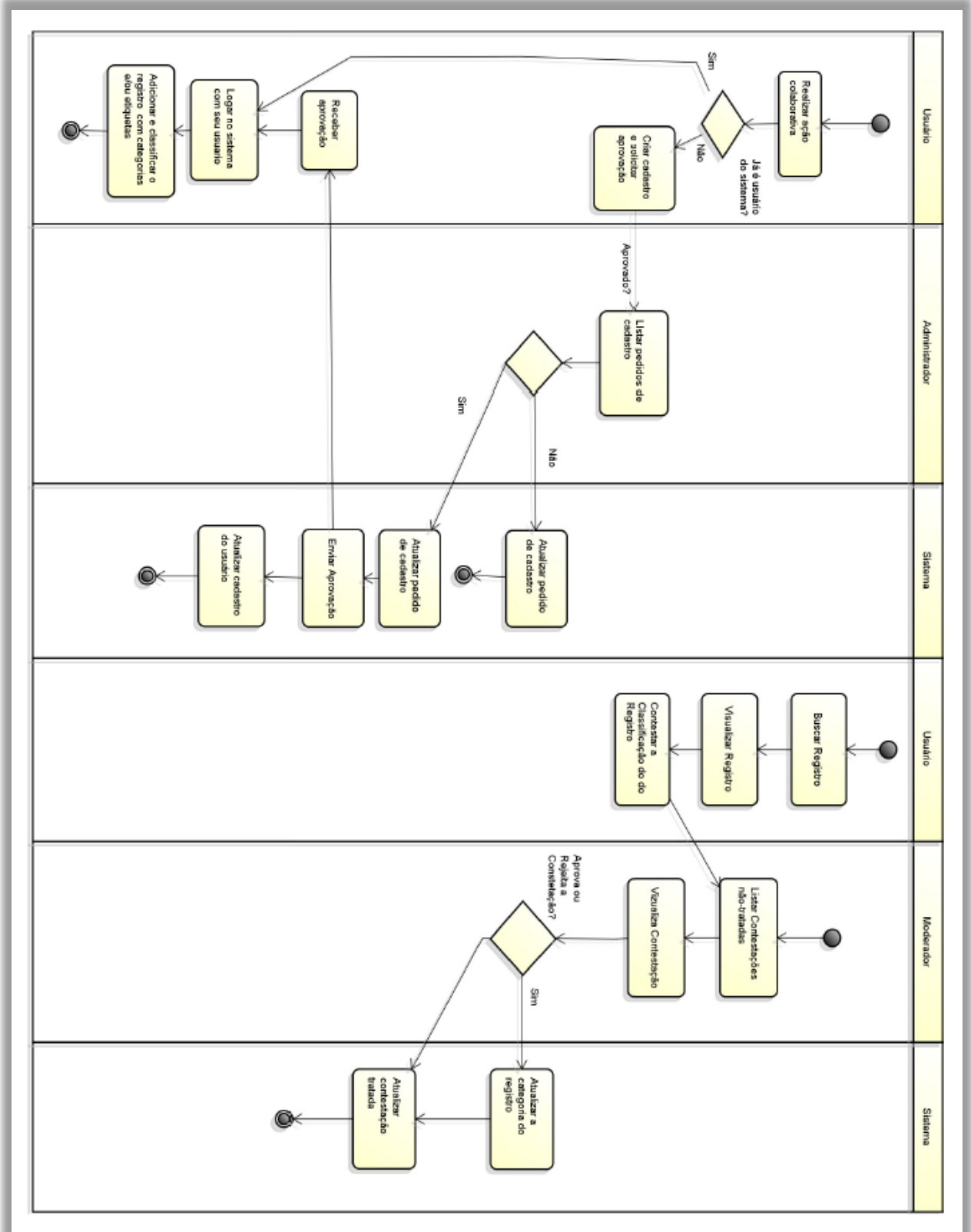
Fonte: Elaborado pelo autor.

5.2.3.2 Diagramas de atividade

A seguir, na FIG. 70, é apresentado o diagrama de atividade para o *Facetlog*. Esse diagrama ilustra a coordenação do modelo de colaboração desenvolvido.

A colaboração no catálogo inicia-se, primeiramente, com o usuário que deseja alimentar o catálogo. Para realizar essa tarefa, é preciso entrar no sistema. Se não tiver um nome de usuário, o mesmo deve se registrar no sistema. Logo após entrar, o usuário pode adicionar um registro ao catálogo, classificando-o da maneira que julgar mais adequada.

FIGURA 70 - Coordenação com fluxo de trabalho de aprovação



Fonte: Elaborado pelo autor.

O fluxo que se inicia na quarta guia, na figura anterior, mostra o usuário navegando e encontrando um registro cuja indexação deseja revisar (contestação). Outro usuário pode contestar a validade de uma classificação atual, podendo sugerir outra para substituí-la. Quando o usuário salva as alterações, as contestações pendentes ficam reunidas em uma área da administração para serem tratadas o mais rapidamente possível. Dessa forma, um ou mais usuários administradores podem acessar essa área administrativa e analisar a contestação. O administrador, então, decide se a contestação é válida ou não. Pressupõe-se que o administrador que analisará as contestações tenha conhecimento do domínio do catálogo. A remoção ou adição de uma categoria do registro poderia ser efetuada imediatamente, sem a necessidade de aprovação. Entretanto, como isso impacta nos recursos recuperados, ao selecionar uma categoria, opta-se por deixar o veredito final para o administrador do catálogo.

5.2.4 Escolha da ferramenta

Para a escolha da ferramenta para a elaboração do *Facetlog* foi realizada uma pesquisa, levando-se em consideração os requisitos funcionais e não funcionais das seções anteriores. Foi adotada a plataforma LAMP, que utiliza o sistema operacional *Linux*, servidor *web Apache*, servidor de banco de dados MySQL e linguagem de programação PHP. Essas tecnologias são muito populares na *web* e foram usadas por diversas *startups* no início da *Web 2.0*.

Também há um grande número de desenvolvedores e fontes de consulta. Outro ponto positivo, que favoreceu a escolha de tais recursos, é que todas as tecnologias envolvidas são gratuitas, inclusive o ambiente de desenvolvimento, que pode ser o *software* NetBeans ou Eclipse. Essa plataforma é apoiada pelo governo do Brasil e amplamente adotada no meio acadêmico e empresarial.

Na plataforma PHP, existem mais de 18.000 (dezoito mil) soluções no repositório *Sourceforge*⁴⁶ para diversas finalidades. Dois tipos de sistemas, devido às suas características, se mostraram adequados para atuar como catálogo *web*: sistemas de gestão de conteúdo (Content Management System-CMS), que podem ser usados em catálogos diversos, principalmente nos que envolvem mídia; e sistemas de gestão de biblioteca digital. Optou-se por um sistema de gestão de conteúdo devido à sua versatilidade e capacidade de lidar com vários tipos de recursos *web*.

⁴⁶ Disponível em: <<http://sourceforge.net>>.

Entre os sistemas de gestão de conteúdo escritos com a linguagem PHP que utilizam o banco de dados MySQL, os mais populares são o *Joomla*, *Drupal* e *Wordpress* (SHREVES, 2011). Os três participam do concurso *Open Source CMS Awards*⁴⁷ e já foram eleitos como o melhor CMS (em anos diferentes). No trabalho "*Colaborative Faceted Classification*" apresentado na seção 3.2 foi utilizado o *Joomla*. Na pesquisa "Modelo Para Desenvolver Editores Colaborativos Hierárquicos" (SANTOS, 2009), optou-se pelo *Drupal*.

No presente estudo foi escolhido o *Wordpress*, visando atender aos requisitos funcionais e não funcionais com o mínimo esforço de desenvolvimento possível. As características que motivaram a sua escolha foram:

- É um *software* livre sob a licença GPL (*General Public License*);
- É um projeto consolidado, que teve início em 2003, sendo um *fork* ou ramificação de outro *software* livre, o *b2/cafeblog*;
- Realiza a gestão de conteúdo, ou seja, gerencia de forma nativa (não é necessário implementar): páginas, *links*, artigos, termos, etiquetas, categorias, usuários e perfis de usuários;
- É bem flexível, tem a curva de aprendizagem mais rápida dos três *softwares* livres cogitados;
- Também é o sistema de gestão de conteúdo mais popular do mundo atualmente. Estima-se que mais de 18% dos *sites* da Internet funcionam sob o *Wordpress*⁴⁸. No *site* de estatísticas da comunidade no registro do dia 20 de março de 2013, havia em torno de 62.755.062 *sites* em *Wordpress* no mundo;
- Tem uma boa documentação oficial⁴⁹, além de ter milhares de comunidades de desenvolvedores para tirar dúvidas e inúmeras referências de consulta;
- Permite ser estendido com a instalação de extensões ou *plugins*, adicionando recursos específicos para uma determinada aplicação. No repositório de extensões da comunidade oficial do *Wordpress*⁵⁰, existem mais de 21.000 (vinte e um mil) *plugins* que podem ser usados gratuitamente. Além desses, ainda existem os *plugins* pagos. Com um

⁴⁷ Disponível em: <<http://www.packtpub.com/article/open-source-awards-previous-winners>>.

⁴⁸ Disponível em: <<http://techland.time.com/2013/07/29/the-word-from-wordpress-catching-up-with-matt-mullenweg/>>.

⁴⁹ Disponível em: <<http://codex.wordpress.org/>>.

⁵⁰ Disponível em: <<http://wordpress.org/plugins/>>.

repertório tão vasto de extensões, na maioria dos casos, ao aplicar o *Wordpress*, não é necessário desenvolver um módulo desde o início, basta escolher um módulo já existente e customizá-lo, apenas alterando ou incluindo novas funções para atender suas necessidades específicas.

5.2.5 Customização do módulo

Após uma extensa pesquisa no repositório de extensões do *Wordpress*, foi descoberto um *plugin* que já possuía algumas funcionalidades descritas nos casos de uso especificados nas seções anteriores. Esse *plugin* é denominado *WP Idea Stream*.

Tal *plugin* pode ser baixado na sua página oficial⁵¹, no *site* da comunidade do *Wordpress* e instalado gratuitamente. Em sua página oficial, é mencionado que o propósito desse *plugin* é adicionar um catálogo de ideias a um *site* em *Wordpress*.

Assim, os membros de um *site* com esse *plugin* instalado podem cadastrar ideias em categorias pré-determinadas (pelos administradores do *site*) ou atribuir livremente uma etiqueta a uma ideia (etiquetagem social). As ideias podem ser recuperadas por meio de uma busca por palavra-chave ou através da navegação pelas categorias, que mostram todas as ideias relacionadas. O próprio *site* da comunidade do *Wordpress* utiliza esse módulo para captar sugestões dos usuários para melhorias no sistema⁵².

A primeira customização no *plugin* foi transformar a ideia em algo genérico, que pode ser qualquer registro. A seguir é apresentada a interface original do módulo. Na FIG. 71, pode-se visualizar a interface para cadastro da ideia e na FIG. 72 é apresentada a interface original para navegação pelas ideias através das etiquetas (do lado direito).


⁵¹ Disponível em: <<http://Wordpress.org/extend/plugins/wp-idea-stream>>.

⁵² Disponível em: <<http://wordpress.org/ideas/>>.

FIGURA 71 - Interface original de cadastro do módulo *WP Idea Stream*

Submit your idea!

Title of your idea

Content of your idea


Caminho: p

Choose at least one category
 Galeria Área de Atuação Bairros teste teste1 Automóvel Marca Lanches Setor Aeroporto Floresta Campus Universitário Vila Romana Setor Oeste Instrumentos Vila Cristina Jardim Bela Vista Parque Anhangüera Vila Boa Vista Chácara Califórnia Jardim Santa Cecília Vila Matilde Parque Amendoeiras C Jardim Setor Garavelo Jogos Vila Santa Cruz Jardim Real Setor Goiânia 2 Parque Das Nações Jardim Liberdade Conjunto Residencial Jardim Leblon Jardim Luz Parque Maracanã Vila São Paulo Vila Jaraguá Jardim Abapuru Jardim Paraguassu Vila Diamantina Jardim Ipe Res Eldorado Conjunto Guadalajara Vila Martins Setor Faicalville Vila Vera Cruz Condomínio Parque Dos Cisnes Portal Do Sul Feliz S. Coimbra Jardim

Add your tag(s)
 Type your tag, then hit the return or space key to add it

Fonte: Elaborado pelo autor.

FIGURA 72 - Interface original de navegação do módulo *WP Idea Stream*

NUMBER OF IDEAS SO FAR: 20676

CATEGORIAS

Área de Atuação
 Agricultura - Criação De Animais
 Agroindústria
 Alimentos - Bebidas - Fumo
 Lanches
 Pizzeria
 Animais De Estimação
 Arte - Cultura
 Artesanato
 Artigos - Equipamentos Para Indústria
 Artigos - Equipamentos Para O
 Comércio
 Artigos - Insumos
 Atividades Esportivas
 Carros - Outros Veiculos
 Casamento
 Ciências
 Cinema - Rádio - Tv
 Consultorias
 Couro - Arço
 Economias - Finanças - Serviços Legais
 Educação Básica
 Educação Superior e Pós-Graduação
 Eletricidade - Eletrônica
 Eletro-eletrônica
 Engenharia
 Escritório
 Espiritualidade - Religião
 Esporte - Lazer
 Estética
 Festas - Eventos
 Fotografia
 Governos - Serviços Públicos - Sociais
 Gráficas - Papel - Embalagens
 Hobbies - Coleções
 Imóveis
 Indústria Alimentícia
 Indústria Química - Farmacêutica
 Informática - Internet -
 Telecomunicações
 Instalações Rurais

Faculdade Cambury
 Posted on março 4, 2013 by admin
 ★★★★★
 Leave a reply
 Avenida T 2, 3531 Setor Sol Nascente - Goiânia - GO - CEP: 74215-010
 Categories: Educação Superior e Pós-Graduação, Empresas de Goiânia - Goiás, Vila Sol Nascente | Edit

Escola Infantil São José
 Posted on março 4, 2013 by admin
 ★★★★★
 Leave a reply
 Do Baby (2 anos) ao 5º Ano. Período Matutino e Vespertino. R 54, s/n Qd B Lt 21/22 Jardim Goiás - Goiânia - GO Tel: (62) 3245-1811 | 3541-3023
 Categories: Educação Superior e Pós-Graduação, Empresas de Goiânia - Goiás, Jardim Goiás | Edit

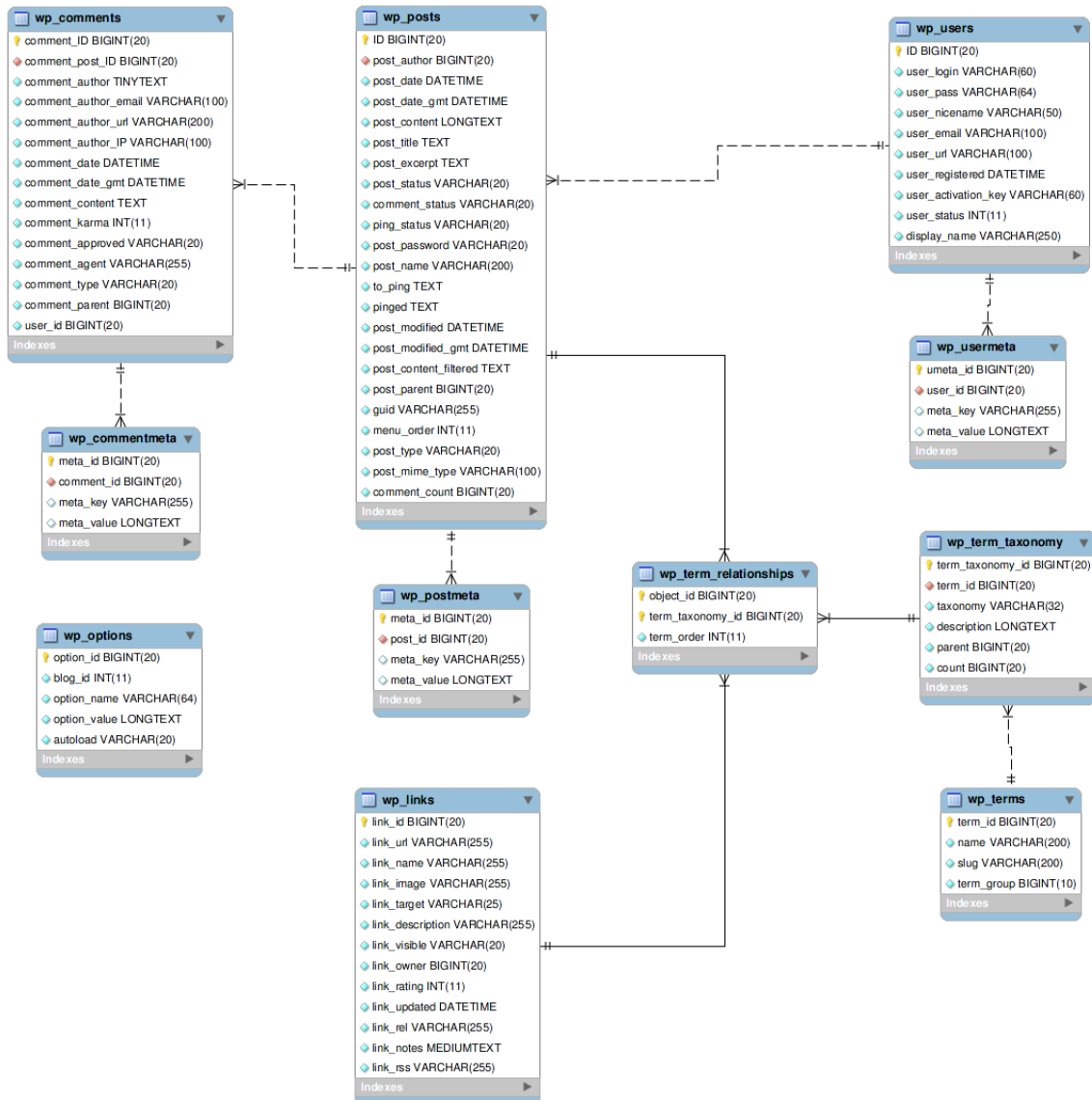
IFG - Instituto Federal de Educação de Goiás
 Posted on março 3, 2013 by admin
 ★★★★★
 Leave a reply
 Rua C 137, s/n qd 567 It Goiânia - GO
 Categories: Educação Superior e Pós-Graduação, Empresas de Goiânia - Goiás, Setor Central | Edit

Fonte: Elaborado pelo autor.

5.2.6 Modelo de dados

O modelo de dados do módulo utilizava as tabelas nativas existentes no Wordpress. O diagrama da base de dados está representado na FIG. 73.

FIGURA 73 - Diagrama do banco de dados do Wordpress



Fonte: Site Wordpress Codex⁵³

⁵³ Disponível em: <http://codex.wordpress.org/Database_Description>.

No QUADRO 5 a seguir são mapeadas as entidades do banco de dados para os objetos do catálogo: os itens registrados são armazenados na tabela *wp_posts*; as facetas, categorias e etiquetas na tabela *wp_terms*; as taxonomias das facetas são armazenadas na tabela *wp_term_taxonomy*; e as classificações dos itens registrados na tabela *wp_term_relationships*. Um catálogo é um tipo de categoria e uma contestação é um tipo de *post* (tabela *wp_posts*). Os atributos dos itens registrados na tabela *wp_postmeta*. Assim, o módulo *WP Idea Stream* utilizado, já usava as tabelas nativas do *Wordpress*.

QUADRO 5 - Mapeamento do catálogo para o banco de dados

Objetos do catálogo	Entidades do banco de dados do <i>Wordpress</i>
itens registrados	<i>posts</i>
facetas, categorias e etiquetas	<i>terms</i>
taxonomia	<i>term_taxonomy</i>
classificação do registro	<i>term_relationships</i>
catálogo	<i>term_taxonomy</i>
contestação	<i>posts</i>
atributos dos itens registrados e contestações	<i>postmeta</i>
usuário	<i>user</i>
perfil do usuário	<i>usermeta</i>

Fonte: Elaborado pelo autor.

5.2.7 Modelagem de taxonomia facetada

Após definir o modelo de dados para armazenamento dos objetos do catálogo, a próxima etapa foi utiliza-lo em um catálogo *web* de empresas de uma cidade, aplicação escolhida para a avaliação do protótipo.

Foi realizada uma classificação de entidades, especificamente empresas, para a construção de um catálogo *web*. Não houve necessidade de uma notação para localização física das entidades. Para a realização da classificação facetada foram usados os atributos

das empresas. Dessa forma, baseado na fundamentação teórica, entende-se que a taxonomia facetada construída refletiu a segunda interpretação da classificação facetada mencionada na seção 2.2.2, em que são classificados objetos ao invés de conhecimentos.

Foi gerada uma lista de empresas por área de atividade de Goiânia e conforme os exemplos de catálogos *web* de empresas da seção 2.5.2.1 foram determinadas as seguintes facetas para a taxonomia facetada: Área de Atividade e Bairro. Pode-se entender que essas facetas correspondem respectivamente às categorias fundamentais Personalidade e Espaço do PMEST de Ranganathan. A modelagem ficou limitada nessas duas facetas devido à falta de outros atributos das empresas na listagem que foi obtida.

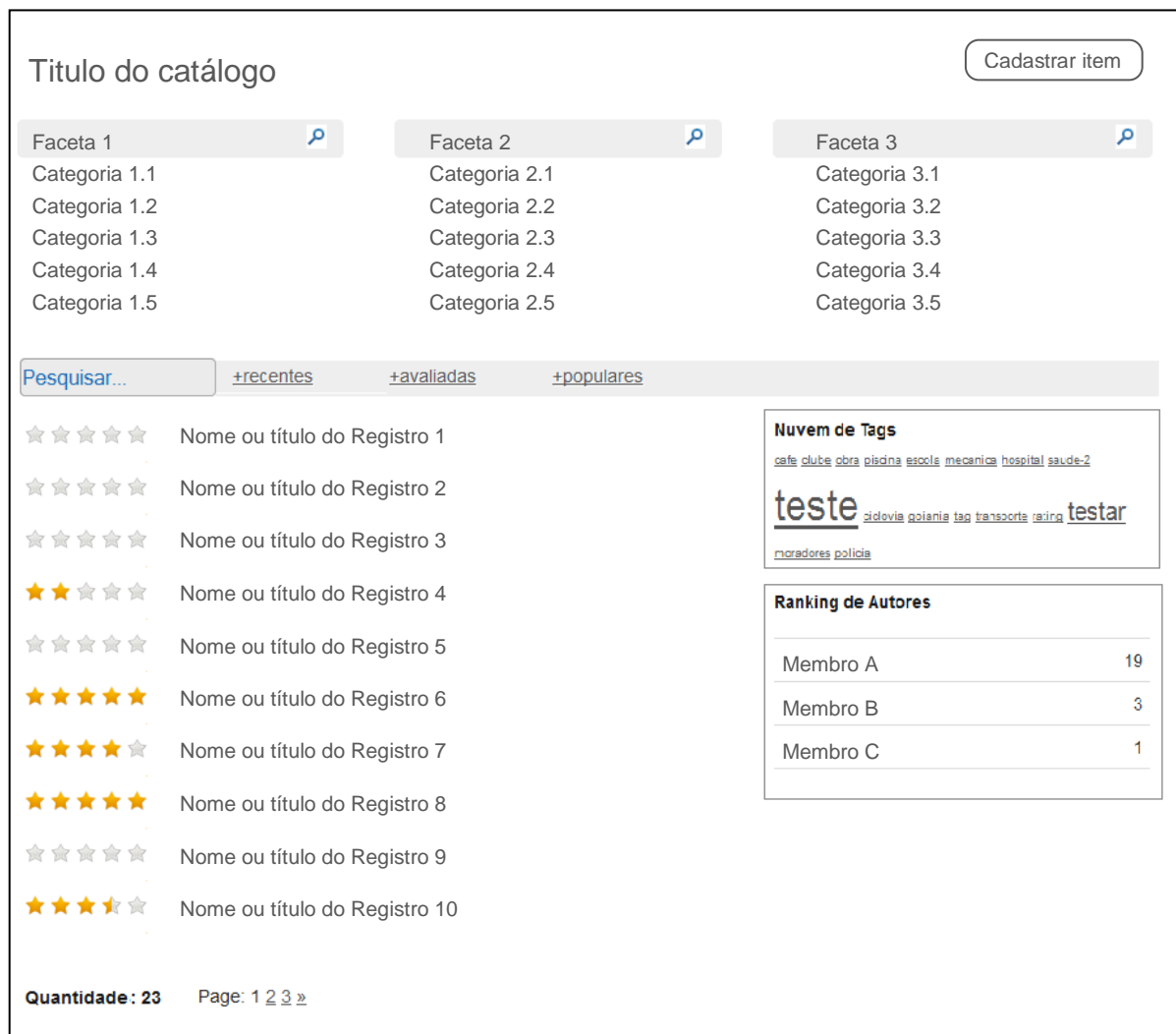
No teste de usabilidade, o usuário poderia indexar ou recuperar uma empresa usando as categorias de ambas facetas. Posteriormente, o registro de cada empresa da lista que estava em um arquivo do tipo texto (txt) foi importado para o banco de dados relacional (MySQL) do sistema. Foram importados em torno de vinte mil registros para que a tarefa de busca fosse relativamente difícil e cerca de 2 mil registros não foram importados propositalmente, para serem utilizados na tarefa de indexação por parte dos usuários participantes no teste do processo de indexação.

No contexto desse experimento é importante salientar que os usuários não realizaram uma representação temática, dado que os objetos catalogados não eram documentos. Entretanto, a seleção de categorias da área de atividade foi uma tarefa subjetiva, na qual o usuário deveria avaliar em quais categorias cada empresa cadastrada melhor se enquadrava. De forma alternativa, o mesmo poderia atribuir etiquetas para a representação da empresa.

5.2.8 Prototipação de baixa fidelidade da interface facetada

Primeiramente, o desenvolvimento da interface iniciou-se a partir da customização da interface principal do módulo para permitir a navegação facetada. O módulo teve sua interface original radicalmente alterada.

Como praticado nos ciclos de desenvolvimento anteriores, foi realizada uma prototipação de baixa fidelidade (*wireframe*) com o intuito de ter uma visão preliminar da interface antes de se realizar a programação (FIG. 74).

FIGURA 74 - Prototipação de baixa fidelidade na navegação facetada no módulo *WP Idea Stream*

Fonte: Elaborado pelo autor.

5.2.9 Construção

Na terceira fase do desenvolvimento foi realizada a implementação ou programação do sistema no ambiente de desenvolvimento, amparada pela arquitetura do sistema composta pelos artefatos (documentos) produzidos nas fases anteriores.

5.2.10 Criação do ambiente

O ambiente no qual o sistema de gestão de conteúdo (*Wordpress*) realiza o papel de catálogo *web* facetado possui as seguintes características:

- Sistema operacional: *Linux* (preferencialmente)
- Servidor *web*: *Apache* na versão 2.0 ou mais atual
- Módulo para a linguagem: PHP na versão 5.2.4 ou mais atual
- Servidor de Banco de Dados: MySQL na versão 5.0 ou mais atual
- Agente de transferência de *e-mail* (MTA): *Sendmail*

O sistema *Wordpress* rodando em um servidor *web* (*Apache*) possui as seguintes bibliotecas instaladas:

- *Mod_rewrite*: para suportar URLs amigáveis
- *cURL*: biblioteca que permite conectar e comunicar com vários tipos diferentes de servidores e com muitos tipos diferentes de protocolos
- *GDLIB*: para tratamento de imagens
- *Eacellerator*: para *cache* e maior desempenho do servidor *web*

O servidor *web* (*Apache*) deve estar configurado da seguinte maneira para um melhor funcionamento do sistema:

- Memória Limite para execução de *script* em PHP (*memory_limit*): mínimo de 64 MB
- Envio máximo de arquivos (*max_upload*): 100 MB

5.2.11 Módulos instalados

Para o cumprimento dos requisitos foram instalados módulos que adicionam funcionalidades ao *Wordpress*, complementando-o, a fim de atingir os objetivos propostos. Os módulos (*plugins*) instalados são:

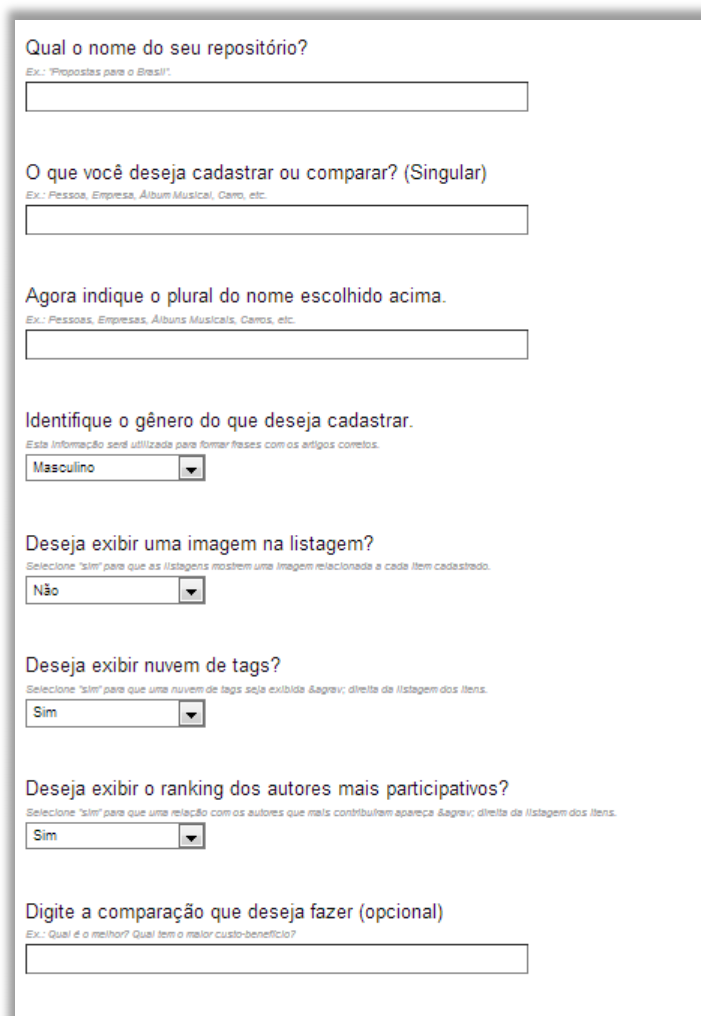
- *Simple Social Buttons*: visa a inserir nas páginas do *Wordpress* os botões *Facebook* (Curtir), *Google Plus* (+1) e de compartilhamento do *Twitter*;
- *Simplr User Registration Form*: adiciona na página desejada um formulário para o usuário se registrar ou *login* no sistema;
- *WP Idea Stream* permite aos usuários compartilharem ideias.

5.2.12 Implementação da interface de busca

5.2.12.1 Criação de um catálogo *web* facetado

Esse é o primeiro passo para disponibilizar um catálogo *web* facetado. A criação de um catálogo pode ser realizada somente pelos administradores do *site*, conforme requisitos da seção 5.2.1. A interface de criação do catálogo pode ser visualizada na FIG. 75 abaixo.

FIGURA 75 - Criar Catálogo



Qual o nome do seu repositório?
Ex.: "Propostas para o Brasil"

O que você deseja cadastrar ou comparar? (Singular)
Ex.: Pessoa, Empresa, Álbum Musical, Carro, etc.

Agora indique o plural do nome escolhido acima.
Ex.: Pessoas, Empresas, Álbuns Musicais, Carros, etc.

Identifique o gênero do que deseja cadastrar.
Esta informação será utilizada para formar frases com os artigos corretos.

Masculino ▾

Deseja exibir uma imagem na listagem?
Selecione "sim" para que as listagens mostrem uma imagem relacionada a cada item cadastrado.

Não ▾

Deseja exibir nuvem de tags?
Selecione "sim" para que uma nuvem de tags seja exibida à grav; direita da listagem dos itens.

Sim ▾

Deseja exibir o ranking dos autores mais participativos?
Selecione "sim" para que uma relação com os autores que mais contribuírem apareça à grav; direita da listagem dos itens.

Sim ▾

Digite a comparação que deseja fazer (opcional)
Ex.: Qual é o melhor? Qual tem o maior custo-benefício?

Fonte: Elaborado pelo autor.

5.2.12.2 Visão geral da interface de busca

Conforme tabela apresentada na seção 2.1, na qual foi feita uma revisão de interfaces para navegação facetada, foram selecionados as funcionalidades e características para permitir os melhores resultados na recuperação da informação. A interface de busca pode ser visualizada na FIG. 76.

FIGURA 76 – Interface de busca

The screenshot displays a search interface for 'Empresas de Goiânia - Goiás'. The interface includes a search bar (1), a search button (3), and a filter section (4) with two active filters: 'Área de Atuação: Alimentos - Bebidas - Fumo' and 'Bairros de Goiânia - Goiás: Conjunto Itatiaia'. Below the filters, there are two columns of category and neighborhood lists. The main content area (2) shows a list of companies, with the first one being 'Pizzaria D Emílio – Goiânia – Goiás'. This entry includes a star rating, author information, and a list of tags (6). To the right, there are two sidebars: 'NUVEM DE TAGS' (8) featuring the tag 'Madeira' and 'RANKING DE AUTORES' listing users and their counts. At the bottom, there is a pagination bar (7) showing 'Numero de empresas para esta consulta: 11' and 'Páginas: 1 2 »'.

Fonte: Elaborado pelo autor.

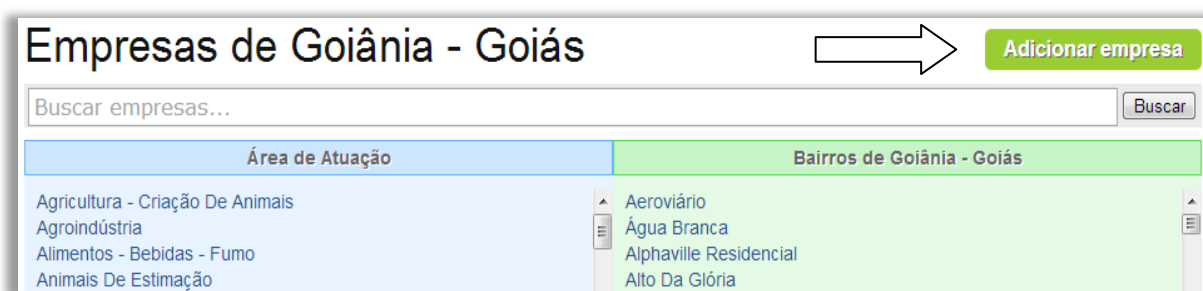
As principais funcionalidades da interface de busca desenvolvida (FIG. 53) são: (1) as facetas selecionadas são evidenciadas; (2) as categorias dos blocos podem ser selecionadas e combinadas para filtrar as empresas recuperadas abaixo; (3) é possível realizar a busca por palavra-chave combinada com navegação facetada; (4) as facetas são distinguidas por meio de cores; (5) há uma barra de rolagem automática para facetas com muitas categorias; (6) ordenação dos resultados; (7) o número de resultados para a consulta é apresentado com paginação; (8) e uma *tagcloud* é combinada com a navegação facetada.

5.2.13 Implementação do modelo de colaboração

5.2.13.1 O visitante se torna membro, entra no *site* e alimenta o catálogo

O primeiro momento da utilização do modelo é caracterizado pela entrada do visitante, que acessa o catálogo e decide contribuir, inserindo um registro. Para tal ação, deve clicar em “Adicionar”. As figuras que acompanham essa seção ilustram as situações descritas, na ordem em que os eventos se apresentam. O primeiro momento é ilustrado pela FIG. 77 abaixo.

FIGURA 77 - Adicionar item ao catálogo



Fonte: Elaborado pelo autor.

Para efetuar qualquer ação no catálogo é necessário que o visitante, antes, entre no sistema (FIG. 78). Ao acessar o endereço é redirecionado para a página de *login*, para que preencha com seu nome de usuário e senha e tenha acesso às funções do sistema. Caso não seja membro pode registrar-se, clicando no *link* "registre-se aqui".

FIGURA 78 - Usuário entra no *site* ou se cadastra

A imagem mostra uma interface web para login. No topo, há uma barra azul com o texto "ENVIAR SUA EMPRESA!". Abaixo, uma caixa amarela contém a mensagem: "Você precisa estar logado para enviar uma empresa. Caso não seja membro, [registre-se aqui](#)". O formulário de login centralizado possui os seguintes elementos: um campo de texto rotulado "Nome de usuário", um campo de texto rotulado "Senha", uma opção "Lembrar" com uma caixa de seleção desmarcada, e um botão "Login" no canto inferior direito. Uma seta horizontal aponta para o formulário, e uma seta vertical aponta para o link "registre-se aqui".

Fonte: Elaborado pelo autor.

Após se registrar e fornecer seu usuário e senha, o mesmo é redirecionado para adicionar um registro no catálogo, evento que conforme seu diagrama de caso de uso (p. 141) consiste em: inserir o nome, a descrição, as etiquetas do registro e selecionar as categorias da taxonomia facetada que o registro pertence (FIG. 79).

FIGURA 80 - Visualização de um item do catálogo

< Empresas de Goiânia - Goiás

Escola do Maternal A 4ª Série Ensino Fundamental

Postado em 2 de março de 2013 por admin
★ ★ ☆ ☆ ☆

Escola Gênese Do Maternal A 4ª Série Ensino Fundamental

Endereço: Av Goiânia Q 68, 10 It 13, 74680-830

Fax: (62) 3207-6909

Jardim Guanabara III, Goiânia, Goiás

Área de Atuação: **Educação Superior e Pós-Graduação** (Avisé-nos se esta informação estiver incorreta.)
Bairros de Goiânia - Goiás: Jardim Guanabara III (Avisé-nos se esta informação estiver incorreta.)
[Editar](#)

Fonte: Elaborado pelo autor.

Clicando nessa opção, o usuário é direcionado para a página de contestação, na qual ele informa a categoria correta do item em questão (FIG. 81).

FIGURA 81 - Contestação de classificação de um item do catálogo

< Empresas de Goiânia - Goiás < Escola Gênese Do Maternal A 4ª Série Ensino Fundamental - Goiânia - Goiás

Contestações para "Área de Atuação" em: Escola do Maternal A 4ª Série Ensino Fundamental (0)

NENHUMA CONTESTAÇÃO

Envie sua contestação!

Informe a classificação correta:

Área de Atuação

- Agricultura - Criação De Animais
- Agroindústria
- Alimentos - Bebidas - Fumo
- Cinema - Rádio - Tv
- Educação Básica
- Educação Superior e Pós-Graduação
- Eletricidade - Eletrônica
- Eletro-eletrônica
- Engenharia
- Escritório
- Espiritualidade - Religião
- Esporte - Lazer
- Estética
- Festas - Eventos

Descrição da contestação

B I U ABC [List of icons]

advanced.path: p

Envie sua contestação →

Fonte: Elaborado pelo autor.

Na área de administração, que somente o administrador do catálogo tem acesso, pode-se observar as contestações informadas pelos usuários do catálogo. A interface exibe, por padrão, as solicitações que ainda não foram tratadas pelo administrador (FIG. 82).

FIGURA 82 - Tratamento das contestações

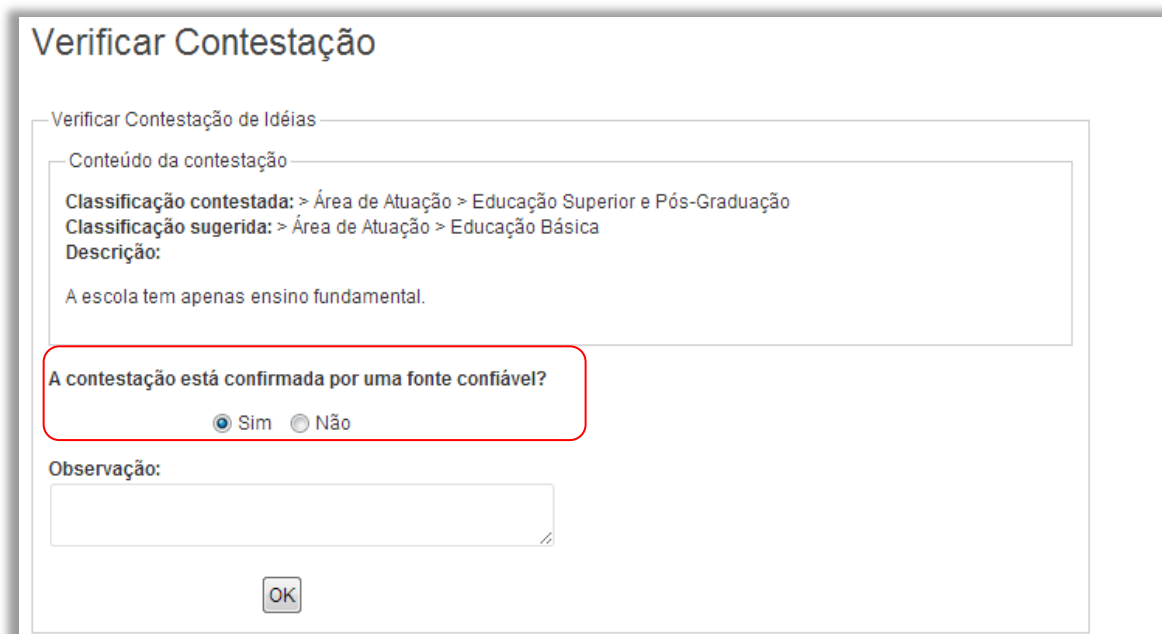


Contestações			
Não tratadas (4) Tratadas (7)			
Idéia	Contestação	Galeria	Situação
Escola Maternal A 4ª Série Ensino Fundamental	A escola tem apenas ensino fundamental. [...]	Empresas de Goiânia - Goiás	Não Verificada
Básico Materiais Construção	Área da Construção civil [...]	Empresas de Goiânia - Goiás	Não Verificada
Condomínio Do Edifício West Plaza	imóveis [...]	Empresas de Goiânia - Goiás	Não Verificada
Condomínio Do Edifício West Plaza	setor bueno [...]	Empresas de Goiânia - Goiás	Não Verificada

Fonte: Elaborado pelo autor.

Clicando na contestação, o administrador do catálogo aprova ou rejeita a contestação sugerida pelo usuário. A contestação pode servir para adicionar ou remover uma classificação a um registro (FIG. 83).

FIGURA 83 - Analisando contestação



Verificar Contestação

Verificar Contestação de Idéias

Conteúdo da contestação

Classificação contestada: > Área de Atuação > Educação Superior e Pós-Graduação
Classificação sugerida: > Área de Atuação > Educação Básica
Descrição:
 A escola tem apenas ensino fundamental.

A contestação está confirmada por uma fonte confiável?
 Sim Não

Observação:

OK

Fonte: Elaborado pelo autor.

Após a aprovação, a classificação do registro é alterada para todos os usuários do catálogo (FIG. 84).

FIGURA 84 - Alteração da classificação conforme sugestão do usuário

< Empresas de Goiânia - Goiás

Escola do Maternal A 4ª Série Ensino Fundamental

Postado em 2 de março de 2013 por [admin](#)
 ★★☆☆☆☆

Escola Gênesis Do Maternal A 4ª Série Ensino Fundamental

Endereço: Av Goiânia Q 68, 10 It 13, 74680-830

Fax: (62) 3207-6909

Jardim Guanabara III, Goiânia, Goiás

Área de Atuação: **Educação Básica (1 Contestação)** [\(Avisar-nos se esta informação estiver incorreta.\)](#)

Bairros de Goiânia - Goiás: [Jardim Guanabara III](#) [\(Avisar-nos se esta informação estiver incorreta.\)](#)

[Editar](#)

Fonte: Elaborado pelo autor.

5.2.13.3 Votação e engajamento

Um usuário do catálogo pode eleger os melhores registros para uma determinada categoria, ou combinação de categorias, por meio de uma nota dada através das estrelinhas (FIG. 85).

FIGURA 85 - Votação, ordenação e ranking de autores

+ recentes + avaliadas + populares

Escolha as melhores empresas de cada atividade na sua localidade!

IFG – Instituto Federal de Educação de Goiás

Autor: [admin](#)

★★★★★

[Educação Superior e Pós-Graduação](#) [\(Avisar-nos se esta informação estiver incorreta.\)](#)
[Setor Central](#) [\(Avisar-nos se esta informação estiver incorreta.\)](#)
[Editar](#)

Universidade Federal De Goiás – Goiânia – Goiás

Autor: [admin](#)

★★★★★

[Educação Superior e Pós-Graduação](#) [\(Avisar-nos se esta informação estiver incorreta.\)](#)
[Governos - Serviços Públicos - Sociais](#) [\(Avisar-nos se esta informação estiver incorreta.\)](#)
[Jardim Lageado](#) [\(Avisar-nos se esta informação estiver incorreta.\)](#)
[Editar](#)

NUVEM DE TAGS

[Decoração](#) [Madeira](#) [proteção redes](#)
[clínica médica](#) [Saúde](#) [madeira](#) [Elétrica](#) [alumínio](#)
[esquadrias](#) [Banho](#) [Cama](#) [Enxovais](#) [Mesa](#) [distac](#)

RANKING DE AUTORES

usuario1	16
usuario10	15
usuario9	9
usuario8	7
usuario3	6
usuario4	5
usuario2	4
usuario6	4
usuario7	3

Fonte: Elaborado pelo autor.

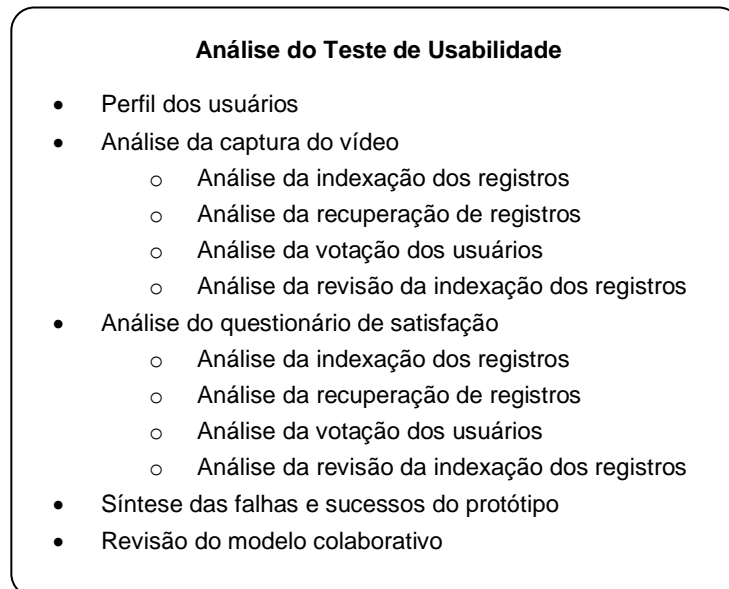
Essa ultima ilustração da implementação da interface encerra a fase de construção do protótipo. A próxima etapa, de avaliação do protótipo através do teste de usabilidade, é detalhada no próximo capítulo.

6 ANÁLISE DOS RESULTADOS E MELHORIAS

Descreve-se, a seguir, como foi a execução dos testes e a observação de seus resultados. Foi explicado aos participantes, antes do teste, que se tratava de uma sessão de teste do catálogo *web* de endereços de Goiânia e também lhes foi informado que seria utilizado o navegador de Internet *Mozilla Firefox*, disponível nativo no sistema operacional *Ubuntu*. Enfatizou-se que o que estava sendo avaliado era a interface do catálogo de endereços, e não o usuário, e que este último estava contribuindo para que os desenvolvedores da interface pudessem detectar erros e perceber quais melhorias eram necessárias para que o modelo satisfaça às necessidades dos usuários.

Na FIG. 86 é apresentada a estrutura deste capítulo que se subdivide nas seções: perfil dos usuários; análise da captura do vídeo dos computadores dos usuários; análise do questionário de satisfação; a síntese das falhas e sucessos do protótipo; e revisão do modelo de colaboração. Tanto na observação da captura do vídeo quanto na análise das respostas dos questionários de satisfação foram avaliadas as tarefas de: indexação dos registros; recuperação dos registros; votação dos usuários; e revisão da indexação dos registros.

FIGURA 86 - Estrutura da análise do teste de usabilidade

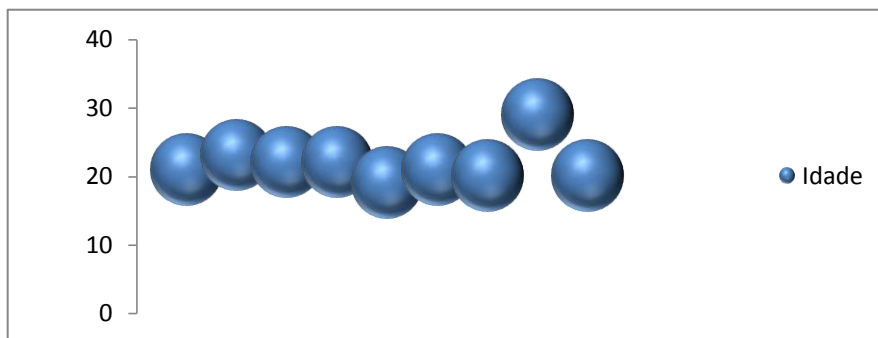


Fonte: Elaborado pelo autor.

6.1 Perfil dos usuários

Para o teste de usabilidade foi prevista a participação de dez voluntários, com a participação efetiva de oito deles, sendo que conforme seção 2.8.4.4.3, esse número é suficiente para descobrir os principais problemas de uma interface (NIELSEN, 2000). Antes de analisar o comportamento dos usuários em cada tarefa realizada, foi levantado, através do formulário de identificação do participante (APÊNDICE B), o perfil do usuário, quanto à sua idade, formação acadêmica, experiência de uso da Internet e catálogos *web*. Como resultado, detectou-se que o grupo era jovem, com idade média de 22 anos (GRÁF. 2).

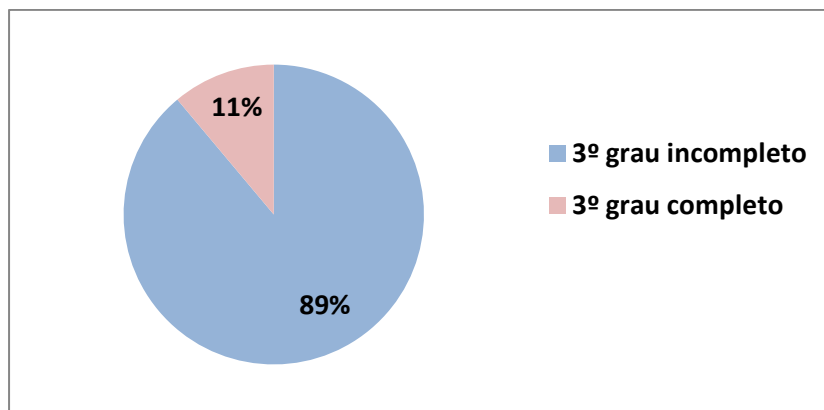
GRÁFICO 2 – Idade dos participantes do teste de usabilidade



Fonte: Elaborado pelo autor.

Todos os participantes do teste, quanto ao nível de escolaridade, tinham pelo menos o terceiro grau incompleto (GRÁF. 3). Os resultados observados no teste realizado por esse grupo podem divergir do resultado obtido através de testes realizados com grupos com características diferentes.

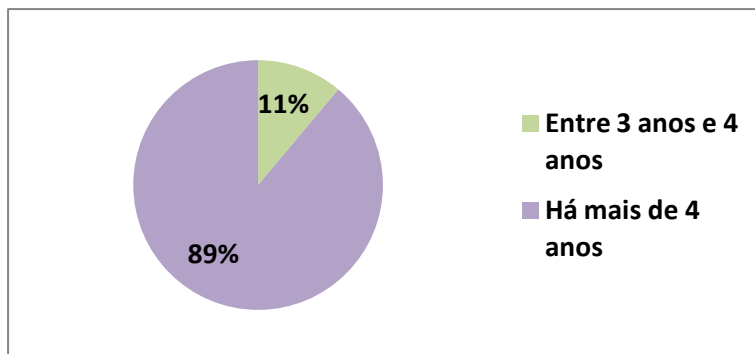
GRÁFICO 3 – Formação acadêmica dos participantes



Fonte: Elaborado pelo autor.

Notou-se, também, que a maioria dos participantes do grupo tinha uma boa experiência com a Internet (tempo de experiência superior a 4 anos, GRÁF. 4)

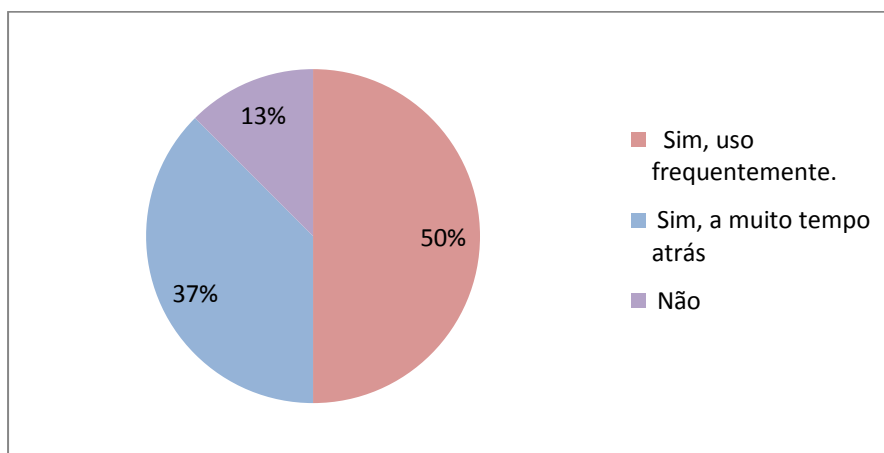
GRÁFICO 4 – Tempo de utilização da Internet



Fonte: Elaborado pelo autor.

Grande parte dos participantes da pesquisa afirmou estar habituada a usar catálogos de endereços *web* (GRÁF. 5).

GRÁFICO 5 – Experiência com catálogo de endereços *web* na Internet



Fonte: Elaborado pelo autor.

Percebe-se que os participantes do teste tinham uma capacidade para lidar com novas interfaces acima da média dos usuários na Internet, pois eram, em sua maioria, jovens adultos, com curso superior incompleto, com experiência na Internet e catálogos *web*. Dessa forma, provavelmente, as dificuldades encontradas pelos mesmos também seriam encontradas, de maneira mais acentuada, em usuários com menor experiência ou tempo de uso da Internet, ou com baixa instrução.

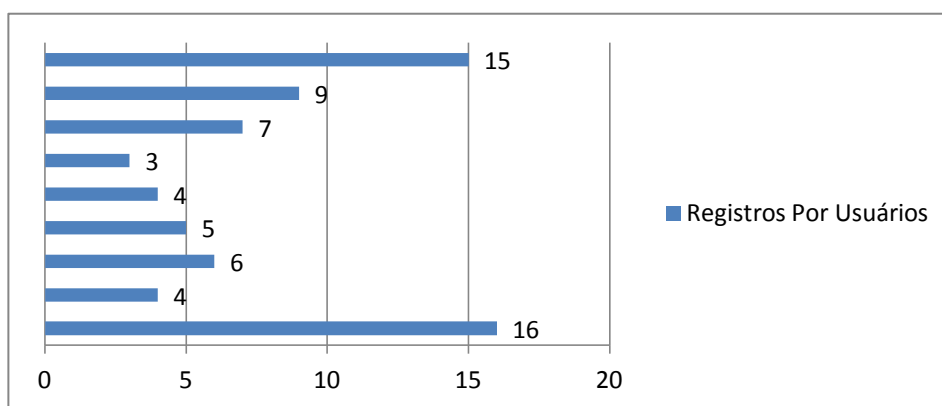
6.2 Análise da captura do vídeo

A análise do teste de usabilidade teve início com a análise das ações do usuário durante a execução da lista de tarefas (APÊNDICE C). Isso foi realizado através da observação da gravação da captura do vídeo do computador de cada um dos participantes.

6.2.1 Análise da indexação dos registros

A primeira atividade do teste de usabilidade que cada usuário foi orientado a realizar foi adicionar empresas ao catálogo de endereços. Cada usuário recebeu uma lista que continha apenas o nome da empresa, o endereço e dados de contato, e deveria cadastrá-la no sistema. Através do GRÁF. 6 a seguir, percebe-se que houve uma grande diferença na produtividade dos usuários, uma vez que um deles foi capaz de cadastrar cinco vezes mais itens quando comparado com o usuário que realizou menor número de cadastros.

GRÁFICO 6 - Cadastros de registros por usuário

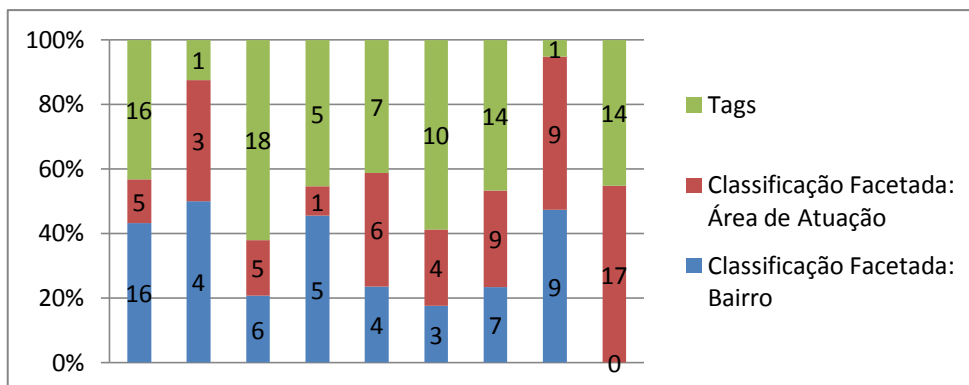


Fonte: Elaborado pelo autor.

Cada usuário que participou do teste foi avisado de que o cadastro da empresa poderia ser feito da maneira como preferisse, quer fosse utilizando etiquetas ou selecionando categorias das facetas. Como apresentado anteriormente, existiam duas facetas: uma que se referia à área de atuação, e outra referente ao bairro em que se localiza a empresa.

A primeira observação com relação à utilização de termos da taxonomia facetada (classificação facetada) *versus* a utilização de etiquetas livremente digitadas pelos usuários é que, para o cadastro do bairro, os usuários utilizaram as categorias em todos os casos. Uma possível explicação para isso é o fato de o bairro ser uma informação fornecida previamente e que apresentava correspondência absoluta com uma das categorias da faceta “bairros”. Com relação à área de atuação, o usuário deveria procurar na taxonomia da faceta qual categoria era mais apropriada para tal registro, ou poderia digitar etiquetas, usando a linguagem natural. O GRÁF. 7 evidencia que a maioria dos usuários preferiu utilizar as etiquetas a usar a classificação facetada para registrar esse tipo de informação.

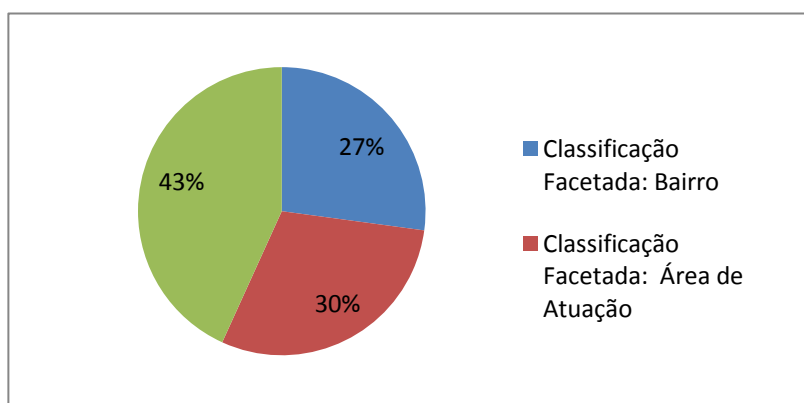
GRÁFICO 7 – Utilização da classificação facetada *versus* etiquetas por usuário



Fonte: Elaborado pelo autor.

No que se refere ao cadastro das áreas de atuação, em 43% dos casos o cadastro foi realizado por meio de etiquetas, contra 30% em que houve uma seleção de categorias da faceta. Não houve nenhum usuário que usou etiquetas para informar o bairro da empresa (GRÁF. 8).

GRÁFICO 8 – Total da utilização da classificação facetada *versus* etiquetas



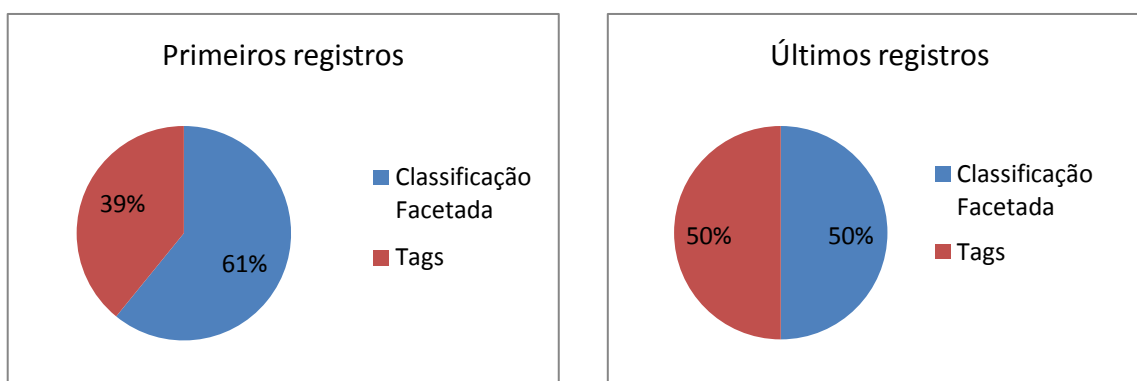
Fonte: Elaborado pelo autor.

6.2.1.1 Alteração no comportamento do usuário

Uma observação importante feita através da análise do resultado dos testes foi a mudança no comportamento do usuário durante a realização do cadastro das empresas. Observou-se que, nos últimos registros, houve grande preferência pela utilização das etiquetas. Nesses, a maioria dos usuários atribuiu um termo do nome próprio da empresa na sua etiquetagem (por exemplo: depósito, imobiliária ou madeireira). Isso pareceu ser mais simples que escolher e selecionar uma categoria na faceta (GRÁF. 9), na indexação do registro.

Outra consideração que merece destaque é que 44% dos usuários procuraram categorias na taxonomia das áreas de atuação, mas não encontraram nenhuma adequada para a indexação. Isso pode ter se dado devido à ausência da categoria na taxonomia ou devido à maneira como o usuário fez a busca, seja falta de atenção ou pressa na execução da tarefa. Nessa situação, o usuário optou por criar as etiquetas.

GRÁFICO 9 – Utilização das categorias *versus* etiquetas



Fonte: Elaborado pelo autor.

6.2.2 Análise da recuperação de registros

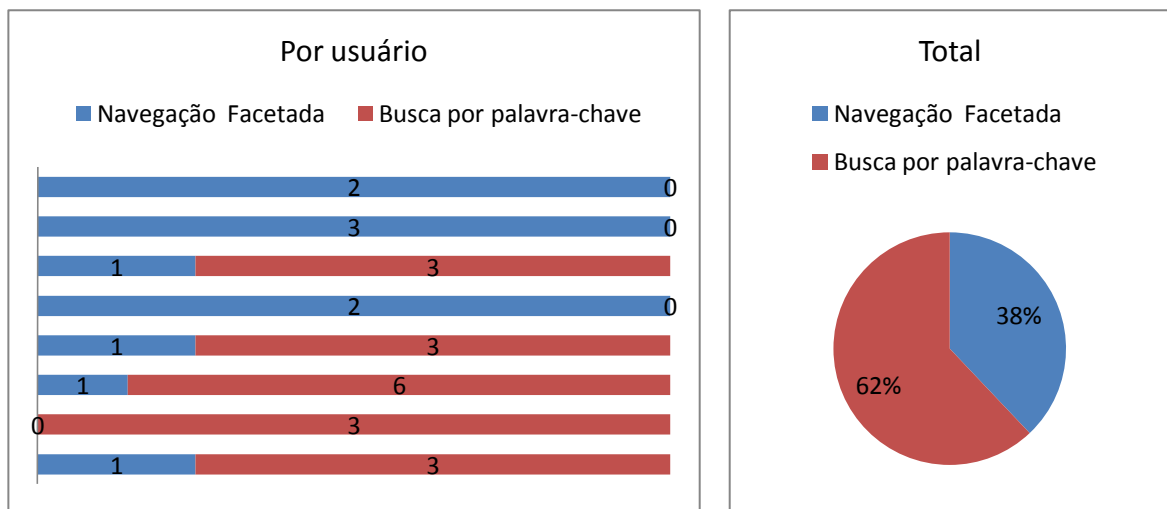
A recuperação de registros foi dividida em três tarefas, que tinham características e níveis de dificuldade diferentes. A intenção ao se criar essa variação foi observar o comportamento dos usuários em relação às estratégias de busca oferecidas e avaliar os passos para obtenção de resultados relevantes.

Na primeira atividade foi requisitado ao usuário que encontrasse cinco empresas de hospedagem no bairro Jardim Vitória em Goiânia. Uma empresa de hospedagem pode ser um hotel, uma pousada ou um albergue. Essas empresas estavam vinculadas à

categoria Viagem e Turismo. Dessa forma, foi avaliado como eles se comportavam ao lidar com uma tarefa de busca que não mencionava diretamente o termo da área de atuação ou o nome da empresa (busca complexa).

A partir do GRÁF. 10, é possível perceber que parte dos usuários preferiram usar apenas a navegação facetada e outros preferiram usar a busca. A maior parte dos usuários usou mais a busca por palavra-chave (62%) que a navegação facetada para encontrar as empresas, no caso da tarefa em questão.

GRÁFICO 10 – Encontrar empresas para hospedagem em um bairro

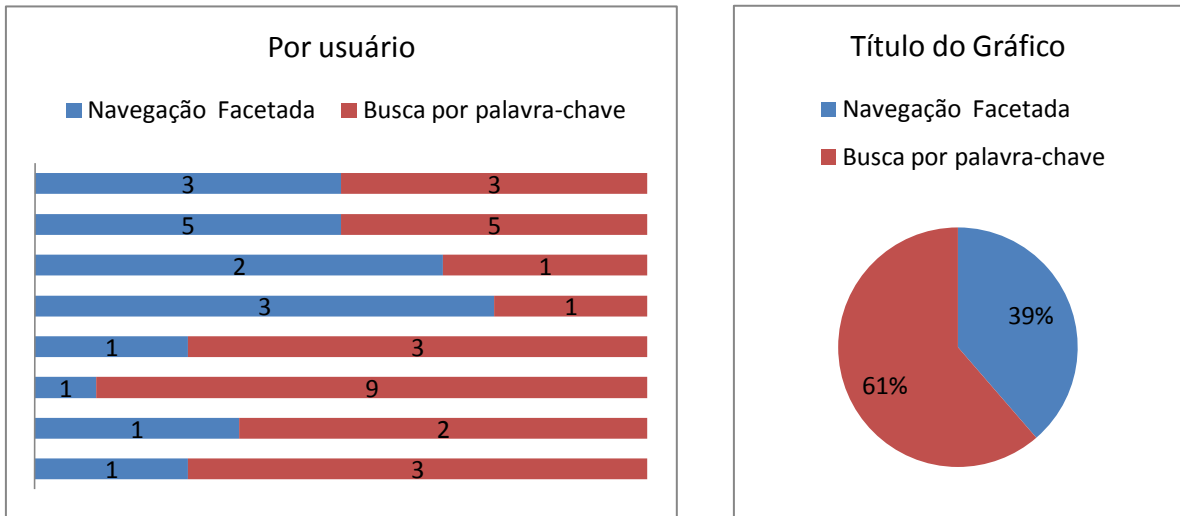


Fonte: Elaborado pelo autor.

A atividade seguinte foi encontrar empresas que tem um tipo de produto próximo a um lugar. Para tal, foi formulado o enunciado: "Encontre 3 empresas que vendem sanduíche nos bairros próximos à UFG". Essa tarefa era mais complexa, pois, além de referenciar o produto, e não a categoria ou o nome da empresa, ainda exigia do usuário a utilização da consulta aos bairros próximos da universidade na qual eles estudavam. Dessa forma, seriam necessárias várias consultas para se obter os resultados requeridos (o que configura uma busca complexa na área de atuação e no bairro).

Nessa atividade, os usuários tentaram uma estratégia de busca diferente da usada anteriormente. Usuários que tinham usado mais a navegação facetada usaram a busca, e vice-versa. Mesmo com essas variações, na média, a busca por palavra-chave, que ocorreu em 61% das ações totais, continuou a ser mais usada que a navegação facetada (GRÁF. 11).

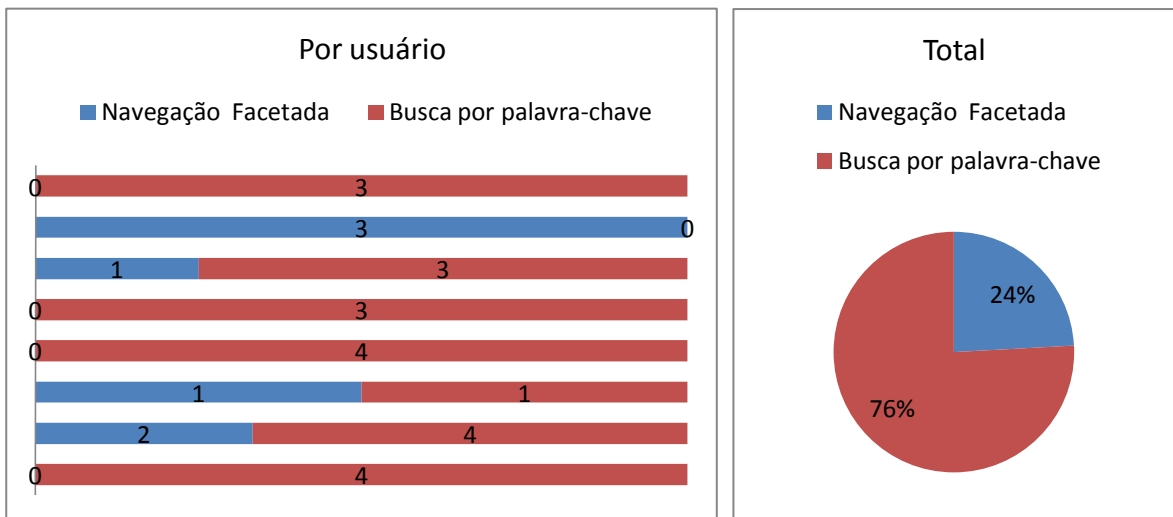
GRÁFICO 11 – Encontrar empresas que vendem determinado produto perto de um ponto de referência



Fonte: Elaborado pelo autor.

Na tarefa seguinte, os usuários deviam encontrar duas empresas em cada especialidade (pintura, janelas de alumínio e instalação elétrica) para reforma em casa. Nessa atividade, o bairro não foi requisitado, os usuários tinham a liberdade de selecionar qualquer empresa da cidade. Nessa consulta, ficou clara a preferência dos usuários pela busca por palavra-chave em uma busca direta (GRÁF. 12).

GRÁFICO 12 – Empresas para reforma em casa

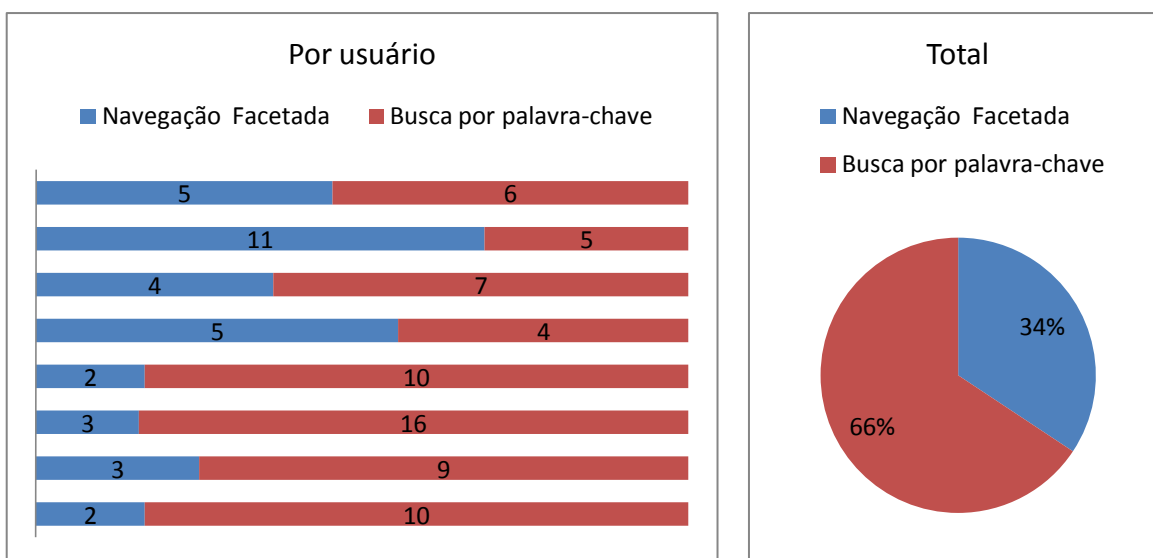


Fonte: Elaborado pelo autor.

6.2.2.1 Total Geral

Ao realizar um somatório de todas as ações de busca realizadas percebe-se que a busca por palavra-chave é preferida em relação à navegação facetada. Em linhas gerais, a navegação facetada foi usada em apenas um terço das buscas, como está mostrado no GRÁF. 13, abaixo.

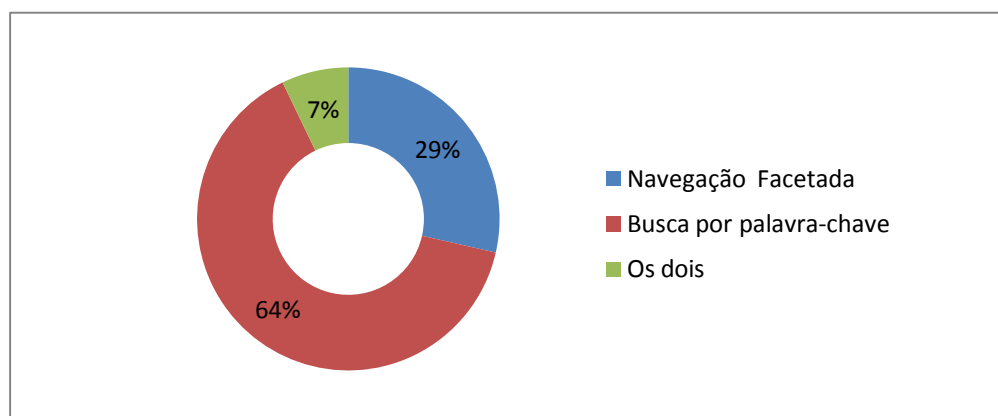
GRÁFICO 13 – Total da utilização da navegação facetada *versus* busca



Fonte: Elaborado pelo autor.

Com relação às estratégias de busca que mais revelaram resultados para os usuários, a busca por palavra-chave mantém a mesma proporção em relação à navegação facetada e à busca que combinou as duas ferramentas (GRÁF. 14).

GRÁFICO 14 – Estratégias de busca que mais permitiram os usuários encontrarem resultados

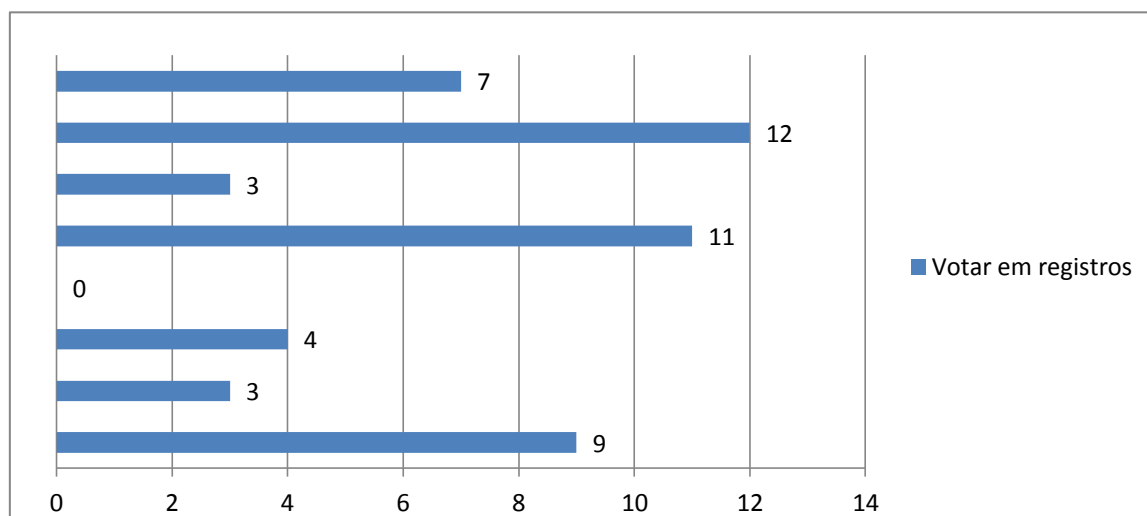


Fonte: Elaborado pelo autor.

6.2.3 Análise da votação dos usuários

No GRÁF. 15 a seguir, verifica-se a quantidade de registros avaliados e que receberam notas dadas pelos usuários. Mais uma vez, percebe-se uma discrepância grande na colaboração dos usuários (mais de 400%), com alguns usuários votando em mais de uma dezena de registros, e outros emitindo sua opinião em poucos registros.

GRÁFICO 15 – Votação nos registros por usuário

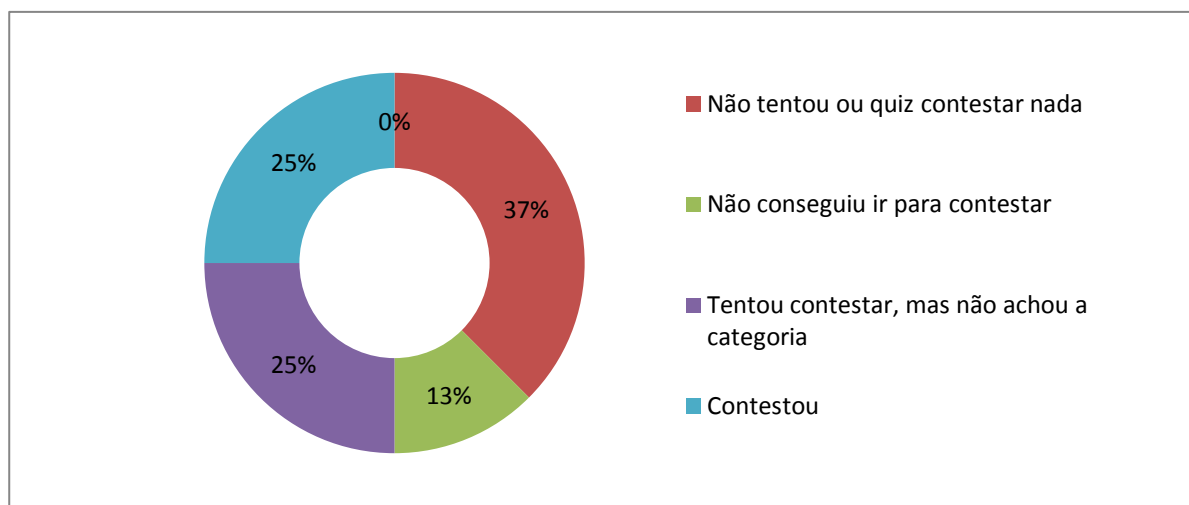


Fonte: Elaborado pelo autor.

6.2.4 Análise da revisão da indexação dos registros

A tarefa de revisar e apontar um erro na classificação de um registro ou adicionar uma nova classificação ao mesmo era realizada por meio do *link* "Avise-nos se esta informação estiver incorreta ou incompleta". Internamente, na administração, tal ação recebeu o nome de "contestação", que deveria ser tratada pelo administrador do catálogo. No GRÁF. 16, vê-se uma estatística referente ao posicionamento/comportamento dos usuários diante da solicitação de realizar essa revisão. Menos de um terço dos usuários (25%) conseguiu realizar essa ação. Atribui-se esse baixo número à dificuldade que tiveram em saber como realizar a ação, como sugerir uma classificação melhor ou como identificar uma classificação errada.

GRÁFICO 16 – Usuários revisam a classificação dos registros



Fonte: Elaborado pelo autor.

Entretanto, nota-se que a maioria dos usuários (63%) tiveram a intenção de contestar. Isso significa que uma grande parcela conseguiu identificar erros ou possíveis melhorias na indexação dos registros. Assim, pode-se esperar melhores resultados futuros, em modelos que estimulem e facilitem a realização dessa ação, com a inserção de estratégias de engajamento de usuários e melhoria contínua da usabilidade da interface para que essa tarefa seja desempenhada com o mínimo de esforço possível e de maneira mais intuitiva. E mesmo com uma baixa adesão, isso significa em soluções de larga escala, é possível ter um batalhão de usuários realizando a indexação gratuitamente, um processo subjetivo e dispendioso quando usado o modelo tradicional através de especialistas. Isso demonstra um potencial futuro nas aplicações do modelo proposto no estudo.

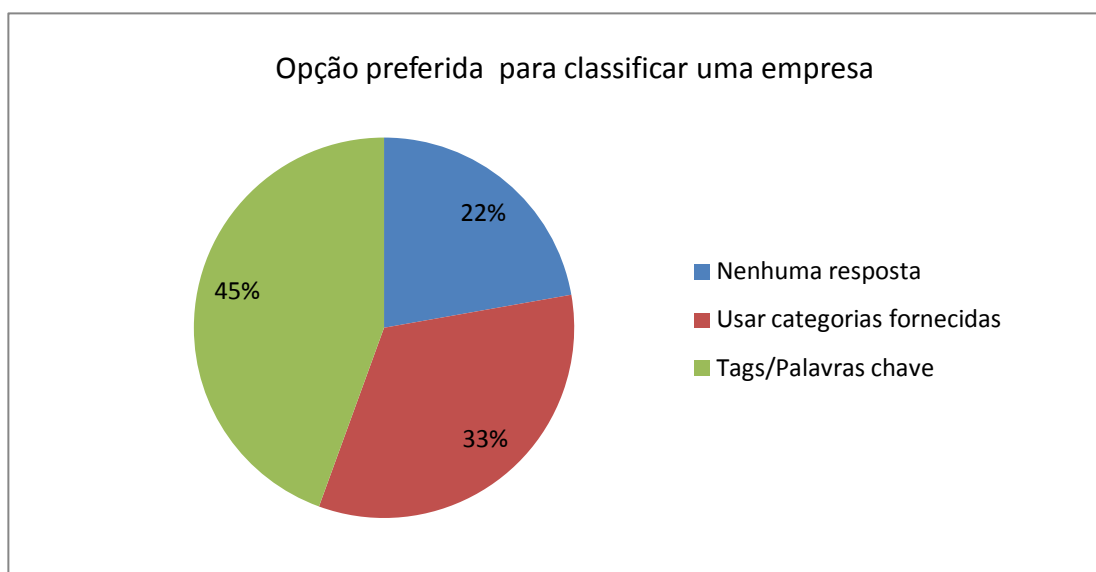
6.3 Análise do questionário de satisfação

A segunda etapa da análise dos dados foi extrair e organizar em planilhas, as respostas dos questionários de satisfação preenchidos por cada um dos usuários (APÊNDICE D). Dessa forma, foi possível estabelecer uma comparação entre as ações observadas e a opinião dos participantes.

6.3.1 Análise da indexação dos registros

A partir do questionário de satisfação do usuário, foi elaborado o GRÁF. 17 que confirma que a preferência da maioria dos usuários era digitar as etiquetas, ao invés de usar as categorias já existentes para o cadastro da empresa.

GRÁFICO 17 – Cadastro: categorias facetadas versus etiquetas

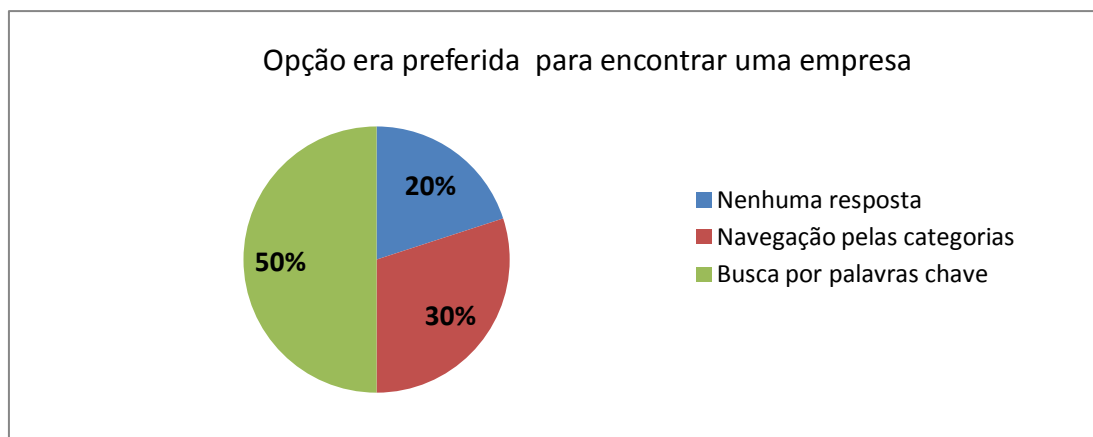


Fonte: Elaborado pelo autor.

6.3.2 Análise da recuperação de registros

Na recuperação dos registros, observou-se a preferência da busca por palavra-chave em relação à navegação facetada. Esse resultado reafirmou o que foi observado no vídeo com as ações dos usuários (seção 6.2.2). O GRÁF. 18 ilustra essa proporção.

GRÁFICO 18 – Navegação facetada versus busca por palavra-chave

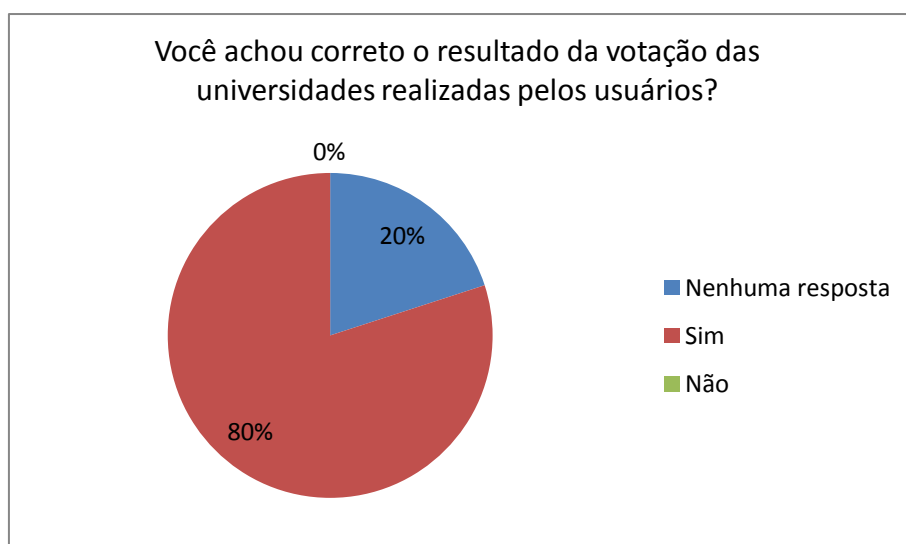


Fonte: Elaborado pelo autor.

6.3.3 Análise da votação dos usuários

Cada usuário descreveu sua impressão após realizar a votação dos registros pela interface do sistema. Foi importante observar que a votação reflete um consenso por parte dos usuários, sobre os itens com maior relevância em uma dada filtragem (GRÁF. 19).

GRÁFICO 19 – Avaliação do resultado da votação

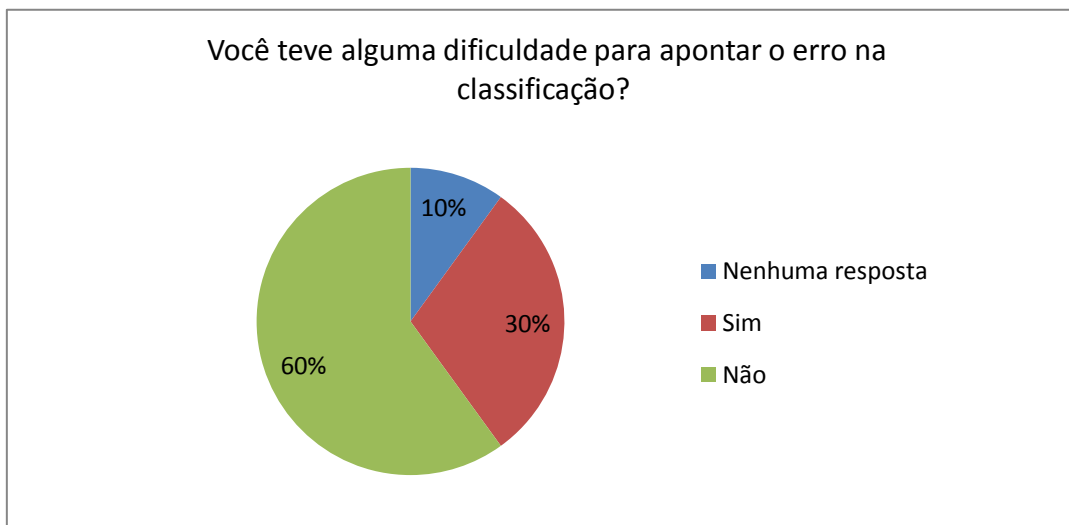


Fonte: Elaborado pelo autor.

6.3.4 Análise da revisão da indexação dos registros

No GRÁF. 20, percebe-se que cerca de um terço dos usuários apresentaram ter tido dificuldade ao realizar a contestação (pedido de revisão da classificação de um registro). Tal fato confirma o comportamento dos usuários durante a captura do vídeo de suas ações e indica a necessidade de melhorias na interface para a realização dessa ação com maior eficiência.

GRÁFICO 20 — Contestação



Fonte: Elaborado pelo autor.

6.4 Síntese dos erros e deficiências do protótipo

Com base nos testes de usabilidade realizados foram classificadas as ocorrências encontradas em erros e deficiências de usabilidade. Os erros na interface que foram percebidos durante a execução das tarefas realizadas durante o teste de usabilidade são:

- Alguns usuários tentaram atribuir uma nota com meio ponto (p. ex.: 4,5). Entretanto, após a tentativa do usuário, a interface não registrou o meio ponto, o que levou os usuários a tentar votar novamente;
- O usuário, ao tentar salvar o registro sem descrição, retornava um caráter estranho dentro do conteúdo do texto;

- Os conceitos da votação, relacionados às notas (péssimo, ruim, bom, ótimo, excelente), não estavam traduzidos;
- A palavra *thumbnail*, que significa “imagem em miniatura”, não estava traduzida, o que obrigou o usuário a fazer uma pesquisa para descobrir o que era. Também faltou a informação de que isso era opcional;
- Ao fazer uma busca por palavra-chave e depois remover a palavra buscada por meio do sinal X, o usuário perdia ou tirava a seleção das categorias que haviam sido selecionadas anteriormente. Tal ação deveria remover da consulta apenas a palavra buscada, não as categorias que haviam sido selecionadas em conjunto.

As deficiências na interface que foram percebidas durante a execução das tarefas realizadas durante o teste de usabilidade são:

- Durante a tarefa de adicionar um registro para cadastrar uma empresa, alguns usuários usaram a busca do navegador (Localizar) para achar uma categoria, principalmente o bairro;
- Após cadastrar uma empresa, o usuário visualizava o registro adicionado. Muitas vezes, percebe-se que os usuários voltavam para a página inicial do catálogo e depois clicavam no botão para cadastrar uma nova empresa. Um *link* para cadastrar uma nova empresa na interface de visualização do registro facilitaria essa ação;
- Ao cadastrar uma empresa, pelo menos uma categoria deveria ser escolhida, seja o bairro ou a área de atuação. Essa obrigatoriedade fez um usuário selecionar aleatoriamente uma categoria na área de atuação que não tinha nenhuma relação com a atividade da empresa cadastrada, apenas para que a tarefa fosse concluída;
- Um usuário tentou cadastrar várias empresas no mesmo registro. Seria interessante se houvesse uma opção para cadastro em massa de vários registros ao mesmo tempo;
- Durante a tarefa de encontrar um registro, vários usuários realizaram busca por palavra-chave, sendo que uma categoria de mesma natureza já estava selecionada. Por exemplo, a categoria bairro Jardim Vitória estava selecionada e, ainda assim, eles buscavam pela palavra-chave Jardim Vitória, ou Itatiaia, que era outro bairro, o que fazia retornar um conjunto vazio de resultados;

- Alguns usuários, ao tentar apontar um erro de classificação (contestação), encontraram uma barreira, que era um campo de descrição, no qual o usuário deveria descrever os motivos ou as referências que ele possuía para justificar a alteração ou adição de uma nova classificação. Apesar de ser uma informação válida para o administrador tomar uma decisão, o fato de esse campo ser obrigatório inibiu dois usuários de realizarem a contestação. Ou seja, é melhor retirar esse campo de descrição, ou torná-lo opcional;
- O usuário, ao fornecer uma nota para um registro, muitas vezes não via uma efetiva diferença. Isso se deve ao fato de a nota ser uma média entre os votos realizados. Assim, com muitos votos, o peso da sua nota poderia ser pouco significativo. Isso parecia, aos olhos do usuário, que seu voto não havia sido computado. Isso também o estimulava a votar em registros com nenhum voto. Outro problema encontrado foi o fato de que, após votar, o usuário não podia reverter sua ação, alterando sua nota;
- Muitos usuários, na busca por palavra-chave, digitaram consultas com os termos no plural. Isso ocorreu principalmente com o termo "universidades". Um tratamento para normalizar as palavras do plural para o singular aumentaria a eficiência da busca.

6.5 Revisão do modelo de colaboração

A partir da análise realizada na seção anterior, percebe-se que a moderação por aprovação pelo administrador é excessivamente conservadora e centralizadora em alguns aspectos (no teste de usabilidade, por exemplo, muitos usuários que não encontravam uma categoria para classificar o registro resolveram a questão criando uma etiqueta nova ou escolhendo categorias inadequadas para a situação). Mas, permitir aos usuários realizarem todas as ações com moderação pré-aprovada, mesmo com a possibilidade de reversão da ação, pode ser desastroso (principalmente a exclusão), como foi percebido nos ciclos de desenvolvimento anteriores (APÊNDICE G). Dessa forma, a proposta final é de um modelo de colaboração unificado (MCU) que tenha uma coordenação flexível, configurável e extensível, que realize a combinação do que há de melhor nos dois mundos, que é a moderação pré-aprovada ou por aprovação, com a possibilidade de inclusão de novos fluxos

de trabalho, configurável para cada uma das ações e perfis de usuário do catálogo. Definem-se os conceitos formulados a seguir:

- Modelo de Colaboração Facetado (MCF)⁵⁵ é a aplicação de facetas em um espaço compartilhado (cooperação) com a implantação da navegação facetada e indexação através de taxonomia facetada regida por um esquema de moderação extensível (coordenação) que visa equilibrar a contribuição e qualidade de informação. A comunicação é utilizada principalmente para servir a moderação (coordenação) e o engajamento de usuários (cooperação);
- Modelo de Colaboração Unificado (MCU)⁵⁶ é o MCF que inclui a utilização da busca por palavra-chave na interface de busca e a etiquetagem na indexação (ambos fazem parte da espaço compartilhado ou cooperação);
- Esquema de Moderação Extensível (EMX)⁵⁷ é aquele no qual cada tipo de contribuição é associado a uma moderação que pode ser do tipo pré-aprovada ou por aprovação. A moderação por sua vez está associada a um fluxo de trabalho. A moderação pré-aprovada significa que ao realizar uma contribuição, ela imediatamente está válida ou visível para um determinado grupo de usuários. Isto não impede que ela seja associada a um fluxo de trabalho que revise seu status posteriormente. A moderação por aprovação significa que a contribuição é criada com o *status* inválido, ou seja, está oculta até que seja realizado o fluxo de trabalho, que pode torná-la válida ou não, para um grupo de usuários. É extensível, pois deve permitir a adição de novos tipos de fluxos de trabalho para a determinação do *status* da contribuição;
- Interface de busca unificada (IBU)⁵⁸ é aquela interface que disponibiliza a busca por palavra-chave ou navegação facetada integradas de forma orgânica utilizando principalmente o recurso de autocompletar;
- Interface de busca e indexação unificada (IBIU) adiciona a IBU a capacidade de indexar os registros através da funcionalidade de arrastar e soltar e janelas modais.

⁵⁵ Em inglês: Faceted Collaboration Model (FCM).

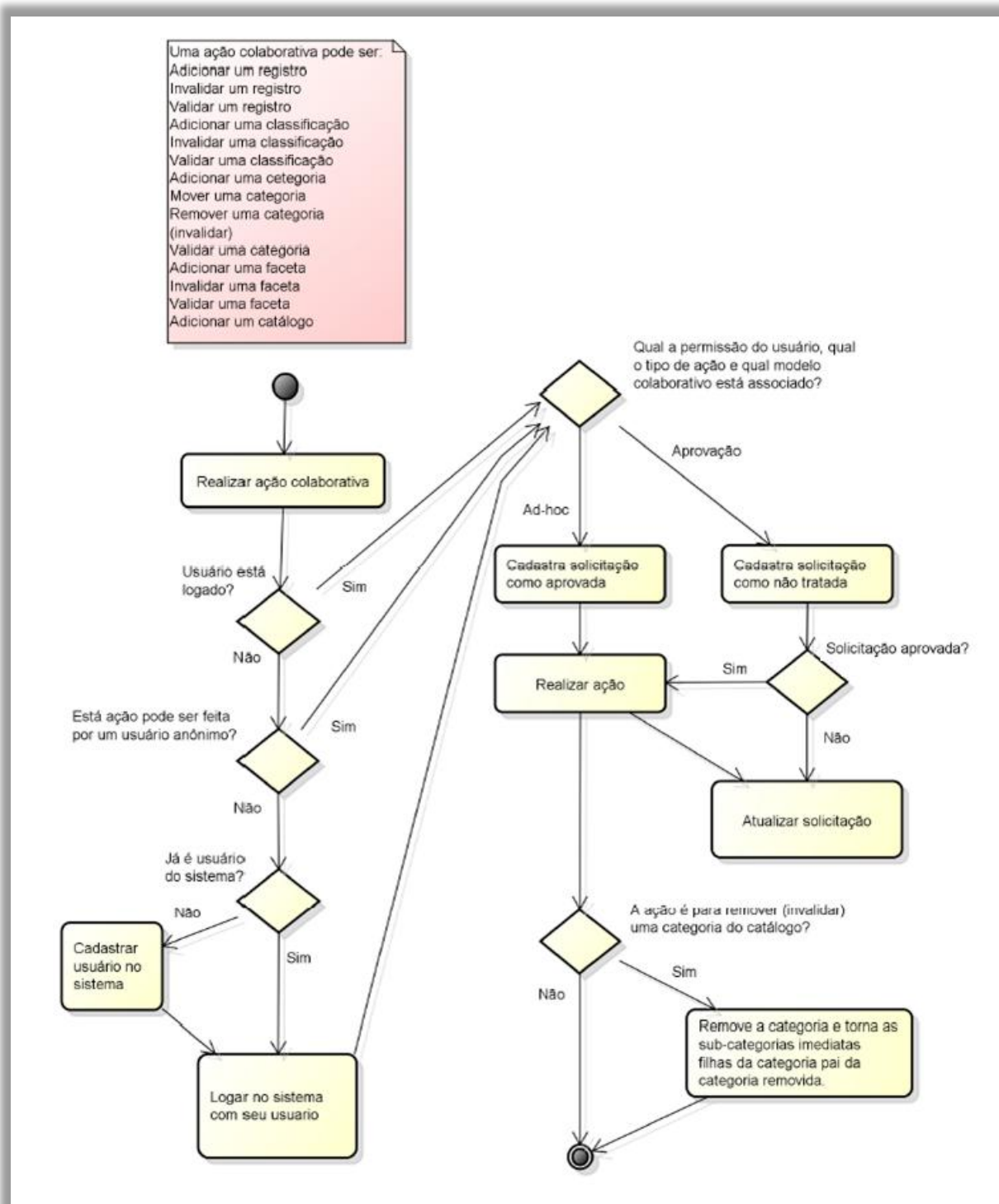
⁵⁶ Em inglês: Unified Collaboration Model (FCM).

⁵⁷ Em inglês: eXtensible Moderation Schema (XMS).

⁵⁸ Em inglês: Unified Search Interface (USI).

Um exemplo de esquema de moderação é exibido na FIG. 87: ao executar a ação, o sistema verifica qual é o tipo do usuário, qual é a ação que ele deseja realizar e qual é o tipo de moderação associada.

FIGURA 87 - Exemplo do fluxo de trabalho de um EMX



Fonte: Elaborado pelo autor.

Esse fluxo de trabalho permite configurar um comportamento colaborativo para cada perfil de usuário, permitindo a participação, também, de usuários anônimos. Em muitos casos, é preferível receber algum tipo de contribuição, a nenhuma contribuição, e a necessidade do *login* afasta muitos usuários (um exemplo é a Wikipédia que aceita colaboração de usuário anônimo). Essa baixa moderação é indicada para ações de baixo impacto no catálogo, como adicionar uma classificação. A proposta é que o administrador do catálogo configure a moderação de seu catálogo conforme seu propósito e comunidade de usuários. Assim, ele pode regular seu catálogo à medida que o mesmo é construído, podendo deixá-lo mais aberto em algumas situações (caso perceba que está havendo pouca colaboração) ou tornando-o mais rigoroso (caso ocorram ações frequentes de vandalismo). Uma sugestão de configuração padrão da moderação é apresentada no QUADRO 6. Nesse esquema, ações com menor impacto são permitidas para usuários anônimos e ações de grande impacto são exclusivas para o administrador do catálogo. A grande diferença em relação aos modelos anteriores é que cada conjunto de configurações das ações resulta em um novo modelo de colaboração para o catálogo.

QUADRO 6 - Moderação padrão sugerida

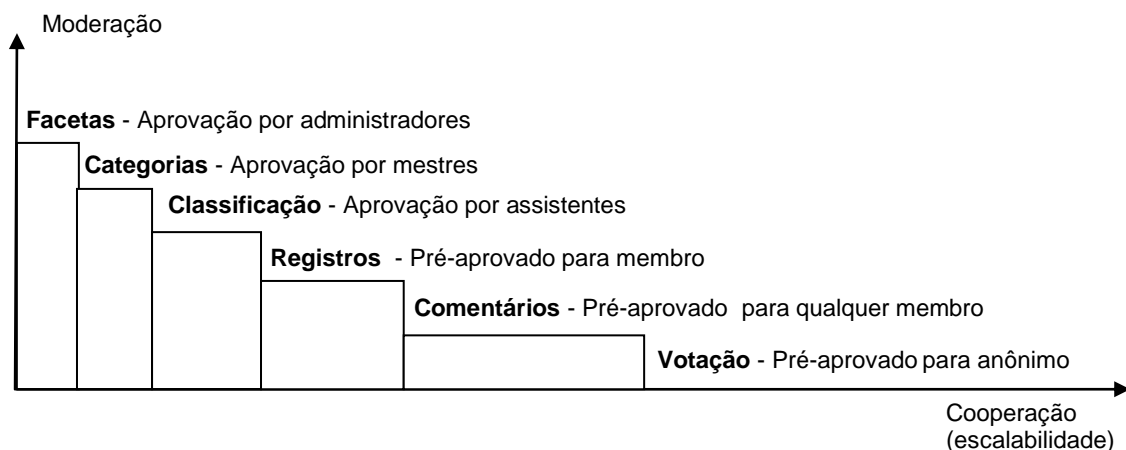
Ações	Anônimo	Membro do site	Moderador	Administrador do catálogo
Adicionar registro	Não permitido	Pré-aprovado	Pré-aprovado	Pré-aprovado
Remover registro	Por aprovação	Por aprovação	Pré-aprovado	Pré-aprovado
Adicionar classificação	Pré-aprovado	Pré-aprovado	Pré-aprovado	Pré-aprovado
Remover classificação	Por aprovação	Por aprovação	Pré-aprovado	Pré-aprovado
Adicionar categoria	Por aprovação	Por aprovação	Pré-aprovado	Pré-aprovado
Remover categoria	Por aprovação	Por aprovação	Pré-aprovado	Pré-aprovado
Votação do registro	Pré-aprovado	Pré-aprovado	Pré-aprovado	Pré-aprovado
Editar categoria	Não permitido	Não permitido	Por aprovação	Pré-aprovado
Adicionar faceta	Não permitido	Não permitido	Por aprovação	Pré-aprovado
Remover faceta	Não permitido	Não permitido	Por aprovação	Pré-aprovado

Fonte: Elaborado pelo autor.

Em um modelo de colaboração, uma moderação alta, mais centralizadora ou mais rigorosa exibe um dinamismo menor, ou seja, é mais lenta para mudar e apenas determinados usuários podem alterá-la. Nesse sentido, o esquema de moderação flexível exibido no QUADRO 6 representa um padrão progressivo, no qual, quanto mais importante for o elemento do catálogo, maior a coordenação ou moderação associada, e menor a escalabilidade da cooperação, ou seja, um número menor de usuários pode estar envolvido ou participar dela. Por exemplo, a moderação por aprovação do administrador do catálogo fica restrita à capacidade de apenas esse usuário tratar as contestações, situação que seria inviável em um *site* com milhões de usuários, como o *Youtube*, por exemplo. Além disso, ao designar todo o poder a um único usuário, o modelo torna-se menos atraente aos outros participantes. Isso diminui a cooperação, ou incentiva menos o compartilhamento de conteúdo pelos usuários comuns. Ao adicionar um novo perfil de usuário com um nível de poder intermediário, por exemplo, como moderador, a responsabilidade na aprovação de determinada ação é destinada a um grupo de usuários, o que amplia a capacidade da cooperação.

A escalabilidade máxima na cooperação é obtida pela moderação pré-aprovada para usuários anônimos, que não está associada a nenhum fluxo de trabalho para revisão, e como ilustrado na FIG. 88, foi associada a votação. Na sequência, foram posicionados: os comentários, registros, classificação (etiquetas ou categorias associadas ao registro de terceiros), categorias e facetas, sendo que, no que se refere às facetas, apenas o administrador do catálogo pode aprovar alterações. Esse modelo padrão também representa uma opção do autor. Em outros contextos, a votação poderia ter uma moderação maior do que os comentários ou registros, mas espera-se que, em geral, as facetas e categorias tenham uma moderação mais alta que os demais objetos do catálogo.

FIGURA 88 - Moderação versus cooperação



Fonte: Elaborado pelo autor

O esquema de moderação extensível representado no QUADRO 7 permite alterar o comportamento do catálogo para cada tipo de ação e perfil de usuário. É expansível, pois prevê a possibilidade de associar novos fluxos de trabalho à moderação: seria possível, por exemplo, incluir uma enquete com uma votação para aprovar uma determinada ação baseada em uma eleição democrática entre os moderadores do catálogo, ou, quando houver uma inserção pré-aprovada, emitir um alerta (*e-mail*) para um moderador, para que ele avalie se a inserção deve continuar a ser visualizada, de maneira similar à do modelo pós produção dos *sites* UCC (seção 2.7.2).

QUADRO 7 - Associação de moderação para cada perfil de usuário e ação

Ações	Perfil do usuário (anônimo, membro do <i>site</i> , administrador ou outro perfil de usuário)
Votação do registro	Não permitido, pré-aprovado ou por aprovação (fluxo de trabalho associado)
Adicionar registro	Não permitido, pré-aprovado ou por aprovação (fluxo de trabalho associado)
Remover registro	Não permitido, pré-aprovado ou por aprovação (fluxo de trabalho associado)
Adicionar classificação	Não permitido, pré-aprovado ou por aprovação (fluxo de trabalho associado)
Remover classificação	Não permitido, pré-aprovado ou por aprovação (fluxo de trabalho associado)
Adicionar categoria	Não permitido, pré-aprovado ou por aprovação (fluxo de trabalho associado)
Remover categoria	Não permitido, pré-aprovado ou por aprovação (fluxo de trabalho associado)
Editar categoria	Não permitido, pré-aprovado ou por aprovação (fluxo de trabalho associado)
Adicionar faceta	Não permitido, pré-aprovado ou por aprovação (fluxo de trabalho associado)
Remover faceta	Não permitido, pré-aprovado ou por aprovação (fluxo de trabalho associado)
Outra ação no catálogo	Não permitido, pré-aprovado ou por aprovação (fluxo de trabalho associado)

Fonte: Elaborado pelo autor.

Na FIG. 89 foi apresentada a prototipação da interface de configuração do esquema de moderação flexível. Nesse protótipo, pode-se observar as ações e o tipo de

moderação associado. Caso esteja configurado como Anônimo, qualquer visitante do catálogo (usuário sem estar logado) realiza a ação sem necessitar de aprovação. Se estiver configurado como Membro, apenas usuários cadastrados e logados podem realizar a ação pré-aprovada. E, se estiver configurado Aprovação, aquela ação realizada deve ser aprovada pelo administrador para ser efetivada. Como foi mencionado, novos fluxos de trabalho podem ser adicionados à listagem da caixa de seleção, expandindo o modelo. Essa interface tem um papel semelhante à ferramenta de definição do fluxo de trabalho de um WFMS apresentado na seção 2.6.1.

FIGURA 89 - Configuração da coordenação

Permissões	
<i>Escolha a permissão para cada uma das ações abaixo.</i>	
Criar categoria	Membro
Editar categoria	Aprovação
Excluir categoria	Aprovação
Adicionar classificação	Anônimo
Excluir classificação	Aprovação

Fonte: Elaborado pelo autor.

O modelo de colaboração proposto, inicialmente restrito apenas ao cadastro e à indexação dos registros, foi expandido no MCU para incluir a gestão da taxonomia facetada e outras opções para a coordenação da colaboração. No capítulo seguinte, é feita uma análise das dificuldades encontradas durante o desenvolvimento, uma síntese dos resultados alcançados em relação aos objetivos propostos e uma projeção dos trabalhos futuros.

7 CONSIDERAÇÕES FINAIS

Essa pesquisa passou por uma grande fase de experimentação, conforme pode ser observado no APÊNDICE G. O desenvolvimento foi dividido em dois principais objetivos: a obtenção de uma interface de busca com navegação facetada para a recuperação de registros; e a criação de um modelo de colaboração que permitisse aos usuários do catálogo contribuir com a manutenção do mesmo.

Durante o desenvolvimento da interface percebeu-se que apenas a tarefa de construir uma interface facetada eficiente, a partir das boas práticas da literatura, exige avançados conhecimentos em programação *web*. O fato de não se ter um escopo claro no início do projeto fez com que muitos avanços ocorressem a partir de tentativas e erros. Foi importante adotar um processo de desenvolvimento iterativo e incremental como o RUP, pois ele permite que o modelo de negócios e requisitos evolua a cada ciclo a partir das falhas e deficiências percebidos na versão anterior, que depende da realização de um teste de *software* adequado. O uso do modelo V para teste de *software* é recomendável para se obter um retorno do projeto em fases iniciais do desenvolvimento, poupando esforços na implementação de requisitos que vão ser reformulados ou descartados, por exemplo.

Adequar um modelo de colaboração a uma interface facetada é uma tarefa que pode ser realizada de diversas formas, e um desafio ao se realizar essa tarefa é disponibilizar tantas informações e ações na interface de forma intuitiva e que não seja cansativa para o usuário. A sugestão é que o modelo de colaboração com a indexação facetada seja incorporado aos poucos, e que, inicialmente, os usuários realizem a indexação facetada em seus próprios registros, depois passem a ter a opção de contribuir na indexação de registros de outros usuários e, finalmente, que possam participar da gestão colaborativa da taxonomia facetada.

Com base no teste de usabilidade com os usuários, ficou claro que, principalmente com usuários jovens habituados a realizar buscas na Internet, a busca por palavra-chave é a estratégia preferida em uma tentativa de sanar alguma necessidade de informação. Contudo, se apenas a navegação facetada não é suficiente em uma interface de busca, a mesma é uma estratégia complementar viável, que pode ser útil em uma parcela significativa de tentativas de recuperação da informação (34% das vezes conforme apontado no GRÁF. 13, seção 6.2.2.1). É usada principalmente quando o usuário não souber o termo correto a ser digitado, em buscas por assuntos compostos, ou quando a busca por palavra-chave não der resultados satisfatórios.

No caso da indexação através da taxonomia facetada foi usada em 57% das vezes em relação à etiquetagem (conforme apontado no GRÁF. 8, seção 6.2.1) evidenciando que a seleção de categorias em uma taxonomia facetada pode ser mais fácil que a digitação de etiquetas. Entretanto, essa preferência pode variar em outro contexto, pois é afetada por fatores, tais como: a qualidade da taxonomia facetada; a usabilidade da interface para indexação facetada; e a bagagem de conhecimento dos usuários.

O modelo de colaboração em um catálogo *web* facetado abrange várias ações, e o seu sucesso depende de um equilíbrio entre restringir o suficiente para garantir certa qualidade sem prejudicar o recebimento de contribuições. O modelo de colaboração unificado permite configurar o comportamento do catálogo conforme a resposta de sua comunidade de usuários. Com relação a catálogos de empresas e instituições de alta credibilidade, espera-se que sejam predominantemente compostos pelo modelo de colaboração por aprovação, enquanto catálogos de grupos e pessoas físicas tendem a ser mais livres, com mais ações pré-aprovadas.

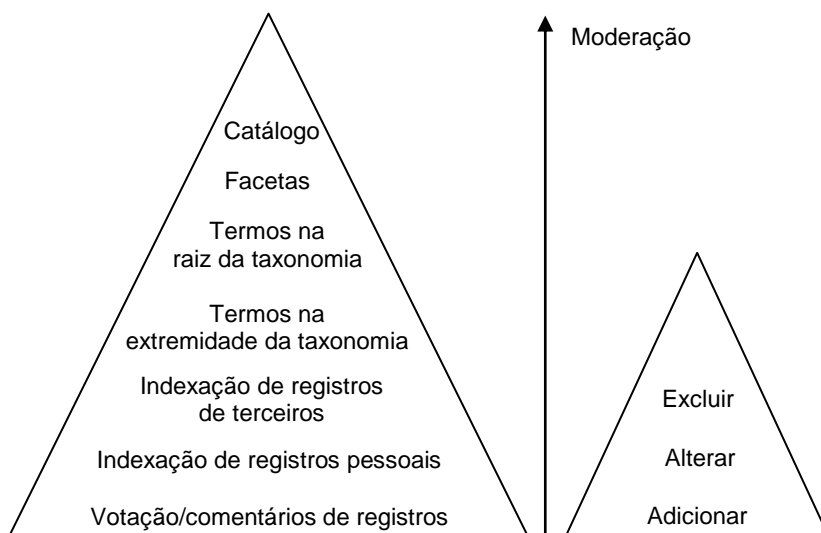
De maneira análoga, espera-se que, conforme o grau de impacto de uma ação em um catálogo *web* facetado, maior seja a coordenação (alta moderação). Assim, em um catálogo *web* facetado colaborativo supõe-se que ações de indexação facetada de registros sejam mais livres que na gestão da taxonomia facetada, e que as categorias nas extremidades da taxonomia devem apresentar maior facilidade para serem incluídas ou alteradas do que as categorias na raiz. E, por último, espera-se que a faceta, que atua como se fosse um pilar da estrutura, seja o item cuja gestão apresenta maior rigor, conforme ilustrado na FIG. 90. Da mesma forma, a ação de adicionar (seja um registro, classificação ou categoria) deve ser mais livre que a ação de alterar, que, por sua vez, deve ser mais livre que a ação de excluir. Nessa linha de raciocínio, a ação de exclusão de uma faceta em um catálogo é a ação que causaria o maior impacto em um catálogo e que deve, portanto, ocorrer na menor frequência possível.

No desenvolvimento do trabalho, demonstrou-se que é possível inserir a indexação através da taxonomia facetada em um catálogo tradicional que apresentava apenas a etiquetagem social. Isso permitiu a inclusão da navegação facetada em conjunto com a busca por palavra-chave na saída do SRI. Essa abordagem é particularmente importante para *sites* que já possuem a etiquetagem social e querem aperfeiçoar a recuperação da informação sem causar uma ruptura no modelo atual.

O presente estudo aponta as seguintes recomendações principais para a implantação do modelo proposto em um catálogo *web* facetado colaborativo: 1) elaborar uma interface de busca que permita a busca por palavra-chave e que convide o usuário a realizar a navegação facetada, se possível, com o uso do autocompletar; 2) desenvolver

uma interface de indexação que permita a indexação através da taxonomia facetada e etiquetagem social; 3) e, ter uma coordenação, na qual as contribuições dos usuários possam ser moderadas, dependendo do contexto e do uso.

FIGURA 90 - Moderação associada ao objeto/ação no catálogo



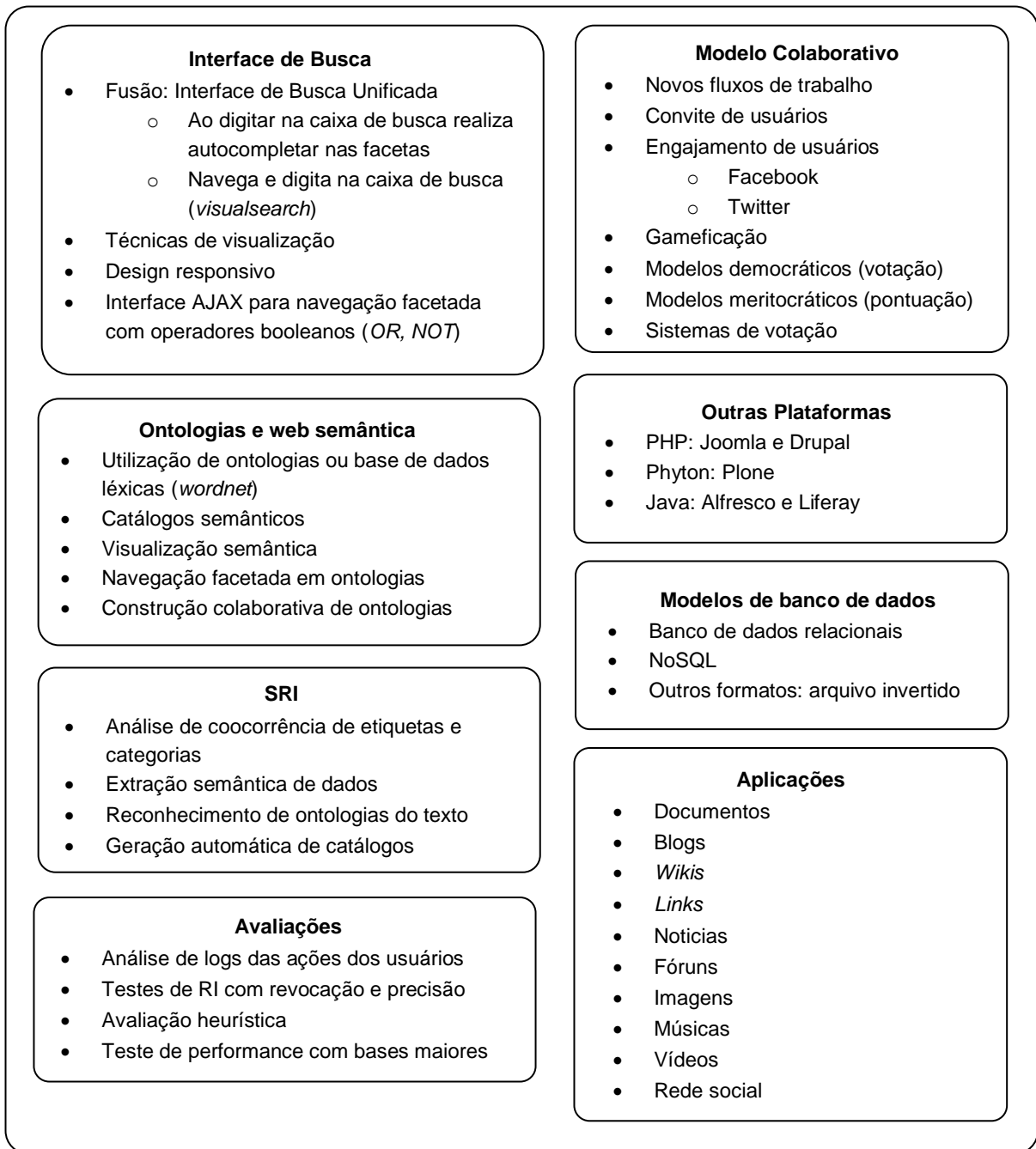
Fonte: Elaborado pelo autor.

Percebe-se que, essa pesquisa não esgota as possibilidades de estudo, mas, pelo contrário, abre-se a novos e instigantes caminhos de investigação.

7.1 Trabalhos futuros

Esse estudo, que se materializa como um objeto do mundo real, um documento científico, é apenas mais um nó que se interliga a vários outros nós, fazendo parte de um enorme emaranhado de *links* que interligam todos nós. Cada vez mais, mais nós e *links* serão explicitados, representados, interconectados, recuperados e comparados. À nossa disposição sempre estiveram as listas e as árvores como suporte para a organização da informação. Agora, para complementá-las, chegam as facetas. Ao mesmo tempo, complementado este estudo, são representados seus desdobramentos através das facetas a seguir (FIG. 91).

FIGURA 91 - Trabalhos futuros



Fonte: Elaborado pelo autor.

Os trabalhos futuros que venham melhorar, estender ou utilizar elementos do *Facetlog* foram divididos nas seguintes temáticas:

Interface de busca: conforme o teste de usabilidade constatou-se que grande parte dos usuários está habituada a digitar um texto na caixa de busca, como estratégia de busca principal. Entretanto, nas ações subsequentes, a navegação facetada pode ser uma opção alternativa para uma parcela significativa dos usuários, principalmente para refinar ou

depurar os resultados. Dessa forma, uma Interface de Busca Unificada, que consiga realizar de forma integrada e fluída a busca por palavra-chave e a navegação facetada pode ser uma alternativa para interfaces de busca em geral. Foram descobertos os seguintes caminhos para o desenvolvimento da IBU: 1) caixa de busca com autocompletar nas facetadas, que destaque ou filtre simultaneamente as categorias das facetadas quando uma palavra for digitada na caixa de busca. A categoria selecionada aparece em destaque fora da caixa de busca, e seu conteúdo é usado apenas na busca por palavra-chave; e 2) navegação facetada com autocompletar na caixa de busca, em que o usuário navega nas categorias das facetadas dentro da caixa de busca, além de digitar termos para a busca por palavra-chave. Esta proposta foi lançada pelo projeto *VisualSearch*⁵⁹, em que a caixa de busca também serve para mostrar a consulta atual com a permanência das categorias e termos até sua efetiva exclusão pelos usuários.

Outra pesquisa que pode ser realizada são as facetadas no contexto. Essa ideia é similar à das facetadas dinâmicas (SACCO; TZITZICAS, 2009), na qual a seleção de uma categoria influencia na visualização das demais categorias da taxonomia facetada. No caso das facetadas no contexto, selecionar uma categoria pode fazer surgir novas facetadas específicas para aquele domínio. Isso já pode ser observado em alguns *sites* de comércio eletrônico, como, por exemplo, o *site* MercadoLivre, que tem facetadas comuns, tal como a localização, e exibe algumas facetadas específicas, tais como as facetadas modelo em celulares⁶⁰ e processador em *notebooks*⁶¹.

A aplicação da técnica de programação AJAX em interfaces facetadas ainda demanda muita pesquisa. Um exemplo do uso é a apresentação dos resultados sem que seja necessário recarregar a página. Mas é possível pensar em aplicações mais amplas. No desenvolvimento *web* atual, existe um movimento denominado *Single Page Interface*⁶² no qual os desenvolvedores buscam criar *sites* e aplicações com uma experiência mais fluída para o usuário, de maneira similar aos aplicativos *desktops*⁶³. Nessas aplicações, um único carregamento de página seria suficiente para responder às ações dos usuários. Assim, a utilização de *frameworks* em *Javascripts* como *Jquery*⁶⁴, *Backbone.js*⁶⁵ e *Node.js*⁶⁶ são, hoje,

⁵⁹ Disponível em: <<http://documentcloud.github.io/visualesearch>>.

⁶⁰ Disponível em: <<http://celulares.mercadolivre.com.br>>.

⁶¹ Disponível em: <<http://notebooks.mercadolivre.com.br/notebooks-laptops>>.

⁶² Disponível em: <http://itsnat.sourceforge.net/php/spim/spi_manifesto_en.php>.

⁶³ Disponível em: <http://en.wikipedia.org/wiki/Single-page_application>.

⁶⁴ Disponível em: <<http://jqueryui.com>>.

⁶⁵ Disponível em: <<http://backbonejs.org>>.

⁶⁶ Disponível em: <<http://nodejs.org>>.

uma necessidade para interfaces mais ricas. Por exemplo, uma grande evolução no *Facetlog* seria que a interface de busca permitisse, sem sair da página, a adição ou indexação de registros através de janelas modais⁶⁷, preservando o *status* das categorias e termos usados na busca e reaproveitando-os para fazer a indexação. Outra necessidade é disponibilizar os operadores booleanos *OR* e *NOT* na navegação facetada.

A aplicação de outras técnicas de visualização da informação, como mapas hiperbólicos, *treemaps* e *web widgets*, pode trazer ganhos na recuperação e aprendizagem da taxonomia facetada. Além disso, técnicas do *design* responsivo podem ser usadas para a elaboração de uma interface de busca que se adapte ao tamanho da tela de uma ampla gama de dispositivos (de TVs à dispositivos móveis).

Modelo de colaboração: utilização de fluxos de trabalho mais descentralizados e escaláveis tais como: criação de perfis de usuário para dividir o esforço no tratamento de contestações; ferramenta de desenho do fluxo de trabalho personalizado para cada catálogo; envio de convites para usuários poderem gerenciar partes da taxonomia facetada ou todo o catálogo, são alguns modelos de colaboração que podem ser desenvolvidos. Para estimular o engajamento de usuários, pode ser realizado o compartilhamento de contribuições dos usuários no catálogo em redes sociais como o *Twitter* e *Facebook*, além de técnicas de gamificação⁶⁸, que utilizam mecanismos típicos de jogos em contextos que não são de jogos com o objetivo de tornar o sistema mais atraente e estimular o engajamento dos usuários. Essas técnicas podem ser úteis, se aplicadas aos catálogos. Por exemplo, as contribuições dos usuários podem ser associadas a um esquema de pontos, em que a pontuação recebida é proporcional à importância da contribuição e, no caso de contribuições reprovadas ou corrigidas, é atribuída uma pontuação negativa. Atingir determinadas quantidades de pontos pode significar a mudança do perfil do usuário, aumentando os poderes que ele passa a ter no catálogo. A utilização de modelos democráticos ou meritocráticos, no qual algumas contribuições são aprovadas através dos votos da maioria dos usuários ou na eleição de registros vencedores em determinadas categorias também pode gerar resultados positivos. Com relação aos sistemas de votação, é necessário fazer muitos testes e verificar maneiras diferentes de efetivar tais sistemas, seja com a aplicação de múltiplas votações com critérios diferentes em um mesmo catálogo ou votação com prazo, por exemplo.

Web semântica: permitir o relacionamento semântico entre os objetos do catálogo visando à importação e exportação de ontologias em catálogos semânticos. Na

⁶⁷ Disponível em: <<http://jqueryui.com/dialog>>.

⁶⁸ Disponível em: <<http://www.gamification.org>>.

área de visualização semântica, a adaptação de técnicas de visualização, como os diagramas hierárquicos para representação dos elementos de uma ontologia. A elaboração de um modelo padrão de interface facetada para a representação e navegação em ontologias⁶⁹, principalmente quando se visa à recuperação de indivíduos. E a adaptação do modelo de colaboração aplicado no catálogo na construção de ontologias por uma comunidade de usuários.

Sistemas de recuperação de informação: a conversão de etiquetas em categorias através da análise da coocorrência e a utilização de ontologias ou base de dados lexicais como o *Wordnet*⁷⁰. A extração de dados de *sites* e banco de dados para popularizar os catálogos através de *crawlers* personalizados. O reconhecimento de ontologias a partir de textos para a captura de conceitos, hierarquias, instâncias e propriedades, e geração automática de catálogos distribuídos.

Avaliações: podem-se realizar outros tipos de teste, tais como: empírico, com a análise de *logs* das ações dos usuários ao aplicar a indexação facetada com a navegação facetada em um *site* que usa apenas a etiquetagem social; revocação e precisão, com a verificação dos ganhos na recuperação de informação em coleções especializadas; avaliação heurística, com recomendações que os gurus das interfaces podem acrescentar à interface facetada; e os testes de performance, para analisar o desempenho do *Facetlog* em bases de dados de milhões ou bilhões de registros.

Modelos de banco de dados: o *Facetlog*, usado para atender a demandas como *Big Data*⁷¹ precisa de aperfeiçoamentos no modelo do banco de dados. Os caminhos de pesquisa vislumbrados para atender a essa demanda são: aumento da performance em bancos de dados relacionais, com a análise de modelo de dados mais eficientes; utilização de bancos de dados NoSQL como o MongoDB e CouchDB, que podem ter uma arquitetura mais horizontal e escalável (DIANA; GEROSA, 2010); e outros formatos, como o arquivo invertido, usados tradicionalmente em máquinas de busca, podem ser adaptados para permitir consultas booleanas recursivas nas hierarquias das facetas.

Outras Plataformas: sugere-se a adaptação do *plugin* do *Facetlog* para outros CMS em PHP, tais como o *Joomla*⁷² e *Drupal*⁷³, para aumentar a visibilidade da proposta. Com isso, é possível também avaliar as vantagens e desvantagens na utilização desses

⁶⁹ Disponível em: <http://www.w3.org/wiki/SemanticWebTools#RDF_and_2For_OWL_Browsers>.

⁷⁰ Disponível em: <<http://wordnet.princeton.edu>>.

⁷¹ Disponível em: <<http://oglobo.globo.com/infograficos/bigdata>>.

⁷² Disponível em: <<http://joomla.org>>.

⁷³ Disponível em: <<http://drupal.org>>.

framework em relação ao *Wordpress*. Portar o *Facetlog* para outras linguagens, tais como *Phyton* e *Java* e outros frameworks, como o *Plone*⁷⁴, *Alfresco*⁷⁵ e *Liferay*⁷⁶ para gestão de conteúdo e o *Dspace* para bibliotecas digitais também é um caminho para experimentação do modelo de colaboração proposto.

Aplicações: é importante ressaltar que o presente estudo pode ser utilizado com diversas finalidades, além do catálogo *web*, tais como: *sites* de notícias, *sites* de compartilhamento de *links*, *blogs*, imagens, vídeos, bibliotecas digitais, fóruns, perguntas e respostas, gestão de documentos, listas de *e-mail*, *wikis*, redes sociais, ontologias, entre outros.

Os temas da FIG. 91 são facetadas na organização dos trabalhos futuros e podem ser combinados na realização de um estudo, por exemplo: "Interface de busca unificada associada a modelo democrático para construção semi-automática de ontologias, a partir de documentos jurídicos, armazenadas em bancos de dados orientados a documentos". O catálogo *web* facetado colaborativo, *Facetlog*, está disponível no endereço <<http://facetlog.com>>, para uso em qualquer área do conhecimento, e existe a intenção de se aprimorar continuamente a solução proposta.

⁷⁴ Disponível em: <<http://plone.org>>.

⁷⁵ Disponível em: <<http://alfresco.com>>.

⁷⁶ Disponível em: <<http://www.liferay.com>>.

REFERÊNCIAS

- AQUINO, M. C. Hipertexto 2.0, Folksonomia e Memória Coletiva: Um estudo das tags na organização da web. **Revista da Associação Nacional dos Programas de Pós-Graduação em Comunicação**, p. 1-18, ago. 2007. Disponível em: <<http://www.compos.org.br/seer/index.php/e-compos/article/viewFile/165/166>>. Acesso em: 13 mar. 2010.
- ARANALDE, M. M. Reflexões sobre os sistemas categoriais de Aristóteles, Kant e Ranganathan. **Ciência da Informação**, Brasília, v. 38, n. 1, p. 86-108, jan./abr. 2009.
- ARAÚJO, Carlos Alberto Ávila. **Análise temática da produção científica em comunicação no Brasil baseada em um sistema classificatório facetado**. 2005. 427f. Tese (Doutorado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2005.
- ARAÚJO, Samur F. C. de. **Explorator: ferramenta para exploração de dados RDF baseado em uma interface de manipulação direta**. 2009. 127f. Dissertação (Mestrado em Informática) – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2009. Disponível em: <http://www.maxwell.lambda.ele.pucRio.br/Busca_etds.php?strSecao=resultado&nrSeq=13792@1>. Acesso em: 19 abr. 2011.
- AULETE, F. J. C.; VALENTE, A. L. dos S. **iDicionário Aulete**. Rio de Janeiro: Lexikon Editora Digital, out. 2008. Disponível em: <<http://aulete.uol.com.br/modelo>>. Acesso em: 20 mar. 2013.
- BAEZA-YATES, R; RIBEIRO-NETO, B. **Modern Information Retrieval**. New York: Addison-Wesley, 1999.
- BALBY, C.N. **Estudos de uso de catálogos online (OPACs): revisão metodológica e aplicação da técnica de análise de log de transações a um OPAC de biblioteca universitária brasileira**. 2002. Tese (Doutorado em Ciências da Comunicação) - Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2002.
- BALCI, Osman. Principles and Techniques of Simulation Validation, Verification, and Testing. In: 27th Winter Simulation Conference, **Proceedings...**Arlington, Virginia, 1995. p. 147-154.
- BARBOSA, R. R.; SEPÚLVEDA, M. I. M; COSTA, M. U. P. Gestão da informação e do conhecimento na era do compartilhamento e da colaboração. **Inf. & Soc: Est.**, João Pessoa, v. 19, n. 2, p. 13-24, mai./ago. 2009.
- BASSO, L. O.; CHEIRAN, J. F. P.; SANTAROSA, L. M. C. Testes de usabilidade com prototipação em papel: validação de ferramenta de AVA acessível a PNEs. In: XII Taller Internacional de Software Educativo - TISE: Nuevas Ideas en Informática Educativa, 2007, Santiago de Chile, Chile. **Anais...**Universidad de Chile, 2006, v. 5, p. 149-156.

BATISTA, E. A. **Uma Taxonomia Facetada para Técnicas de Elicitação de Requisitos**. 2003. 150f. Dissertação (Mestrado em Ciência da Informação) - Instituto de Comunicação, Unicamp, Campinas, 2003.

BERNERS-LEE, Tim *et al.* The semantic web. **Scientific American**, New York, p. 1-18, maio 2001.

BLATTMANN, U.; SILVA, F. C. C. Colaboração e interação na web 2.0 e biblioteca 2.0. **Revista ACB: Biblioteconomia em Santa Catarina**, Florianópolis, v. 12, n. 2, p. 191-215, jul./dez., 2007

BOEHM B., **Software Engineering Economics**. Prentice-Hall, 1981.

BOOTH, P. A. **An introduction to human-computer interaction**. Hove: Psychology Press, 1989.

BORGES, Graciane S. B. **Indexação automática de documentos textuais: proposta de critérios essenciais**. 2009. 111f. Dissertação (Mestrado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2009. Disponível em: <http://www.documentar.com.br/wpcontent/uploads/2011/10/Dissertacao_Graciane_2009.pdf>. Acesso em: 10 set. 2012.

BORGHOFF, U. M; SCHLICHTER, J. H. **Computer supportend cooperative Work: Introduction to Distributed Applications**. Tokyo: Springer, 2000. 529p.

BOWLER D., W. N.; SCHWARTZ, P. **Navigation Bars for Hierarchical Websites**. Maryland: University of Maryland, 2001. Disponível em: <<http://www.otal.umd.edu/SHORE2001/navBar/index.html>>. Acesso em: 02 abr. 2010.

BRANDT, M. B. **Etiquetagem e folksonomia: uma análise sob a óptica dos processos de organização e recuperação da informação na web**. Dissertação (Mestrado em Ciência da Informação) - Universidade de Brasília, Brasília, 2009. Disponível em: <<http://bd.camara.gov.br/bd/handle/bdcamara/4165>>. Acesso em 12 ago. 2011.

BROUGHTON, Vanda. The need for a faceted classification as the basis of all methods of information retrieval. **New Information Perspectives**, v. 58, n. 1/2, p. 49-72, 2006.

BROUGHTON, Vanda; SLAVIC, Aida. Building a faceted classification for the humanities: principles and procedures. **Journal of documentation**, 2007.

CAMPOS, M. L. de A. **A organização de unidades do conhecimento em hiperdocumentos: o modelo conceitual como um espaço comunicacional para realização da autoria**. 2001. 186f. Tese (Doutorado em Ciência da Informação) - Instituto Brasileiro de Informação em Ciência e Tecnologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2001.

CAMPOS, Aline de. A enciclopédia (semi) livre: conflitos, relações hierárquicas e a reconfiguração dos processos na Wikipédia. In: IX Intercom – Congresso de Ciências da

Comunicação na Região Sul, 2008. **Anais...** Porto Alegre: Sociedade Brasileira de Estudos Interdisciplinares da Comunicação, 2008.

CARROLL, J. M.; ROELOFFS, R. Computer selection of keywords using word-frequency analysis. **American Documentation**, p. 227-33, jun. 1969.

CASATI, F. *et al.* Conceptual Modeling of WorkFlows. In: OOER'95 - 14th International Object-Oriented and Entity-Relationship Modelling Conference, 1995. **Proceedings...** Springer-Verlag, 1995, p. 341-354. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=C6637627DCCEF756EA34494EE6F704E4?doi=10.1.1.45.657&rep=rep1&type=pdf>>. Acesso em: 1 jan. 2013.

CATARINO, M. E.; BAPTISTA, A. A. Folksonomia: um novo conceito para a organização dos recursos digitais na Web. **DataGramZero** - Revista de Ciência da Informação, Rio de Janeiro, v. 8, n. 3, jun. 2007. Disponível em: <http://www.dgz.org.br/jun07/Art_04.htm>. Acesso em: 21 ago. 2007.

CIRILO, C. E. **Model Driven Richubi – Processo Dirigido A Modelos Para A Construção De Interfaces Ricas De Aplicações Ubíquas Sensíveis Ao Contexto**. 2011. 203f. Dissertação (Mestrado em Ciência da Computação) - Centro de Ciências Exatas e de Tecnologia, Universidade Federal de São Carlos, 2011.

CUNHA, Murilo Bastos da. Desafios na construção da biblioteca digital. **Ciência da Informação**, v. 28, n. 3, p. 255-266, set./dez.1999.

CUNNINGHAM, W. **Wiki design principles**. Cunningham & Cunningham, Inc. Disponível em: <<http://c2.com/cgi/wiki?WikiDesignPrinciples>>. Acesso em: 22 mar. 2007.

DAMES, K. M. **Social software in the library**. LLRX, jul. 2004. Disponível em: <<http://www.llrx.com/features/socialsoftware.htm>>. Acesso em: 15 dez. 2007.

DENTON, William. **How to make a faceted classification and put it on the web**. Miskatonic University Press, nov. 2003. Disponível em: <<http://www.miskatonic.org/library/facet-web-howto.html>>. Acesso em: 30 ago. 2010.

DEVADASON, Francis J. *et al.* Faceted Indexing Based System for Organizing and Accessing Internet. **Knowledge Organization**, v. 29, p. 65-77, 2002. Disponível em: <<http://www.reocities.com/devadason.geo/FactDSIS.pdf>>. Acesso em: 20 mar. 2009.

DIANA, Mauricio De; GEROSA, Marco Aurélio. NOSQL na Web 2.0: Um Estudo Comparativo de Bancos Não-Relacionais para Armazenamento de Dados na Web 2.0. In: IX Workshop de Teses e Dissertações em Banco de Dados, 2010. **Anais...** Belo Horizonte, 2010. Disponível em: <http://www.lbd.dcc.ufmg.br/colecoes/wtdbd/2010/sbbd_wtd_12.pdf>. Acesso em: 23 ago. 2012.

DIAS, Cláudia Augusto. Hipertexto: evolução histórica e efeitos sociais. **Ciências da informação**, v. 28, n. 3, p. 267, set./dez. 1999.

DIAS, Eduardo Wense. Contexto digital e tratamento da informação. **DataGramZero - Revista de Ciência da Informação**, v. 2, n. 5, out. 2001.

DUMAS, Marlon; HOFSTEDE, Arthur H.M. ter. UML Activity Diagrams as a Workow Specification Language Cooperative Information Systems. In: 4th Int. Conference on the Unified Modeling Language, 2001. **Proceedings...**Toronto: Springer-Verlag, out. 2001. p. 76-90.

ELLIS, C.A.; GIBBS, S.J.; REIN, G. Groupware - Some Issues and Experiences. **COMMUNICATIONS OF THE ACM**, v. 34, n. 1, p. 38-58, 1991.

FAGAN, Jody C. Usability studies of faceted browsing: A literature review. **Information Technology and Libraries**, v. 29, p. 58-66, 2010.

FERREIRA, B. **Uma técnica para validação de processos de desenvolvimento de software**. 2008. 146f. Dissertação (Mestrado em Modelagem Matemática e Computacional) - Centro Federal de Educação Tecnológica de Minas Gerais, Belo Horizonte, 2008.

GEORGAKOPOULOS, Diimitrios; HORNICK, Mark; SHETH, Amit. An Overview of Workflow Management: From Process Modeling to Workflow Automation Infrastructure. **Distributed and Parallel Databases**, Boston, v.3, o. 119-153, 1995.

GEROSA, M. A. **Desenvolvimento de groupware componentizado com base no modelo 3c de colaboração**. 2006. 275f. Tese (Doutorado em Informática) - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2006.

GIL, A. C. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002. 175p.

GUDWIN, R. R. **Diagramas de Atividade e Diagramas de Estado**. Campinas: UNICAMP, 2010. 13p. Aula. Disponível em: <<http://www.dca.fee.unicamp.br/~gudwin/ftp/ea976/AtEst.pdf>>. Acesso em: 13 ago. 2013.

GUEDES, R. M. **A abordagem dialógica na indexação social**. 2010. 186f. Dissertação (Mestrado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2010.

HEARST, M. A. UIs for Faceted Navigation: Re-cent Advances and Remaining Open Problems. In: Workshop on Human-Computer Interaction and Information Retrieval, 2008. **Proceedings...** Redmond: Microsoft Research, 2008. p.13-17. Disponível em: <<http://research.microsoft.com/en-us/um/peo-ple/ryenw/hcir2008/doc/H CIR08-Proceedings.pdf>>. Acesso em: 5 jan. 2011.

HEDDEN, H. **The Accidental Taxonomist**. Medford, NJ: Information Today, 2010. 472p.

HIX, D.; HARTSON, H. R. **Developing User Interfaces: ensuring usability through product & process**. New York, NY: John Wiley and Sons, 1993.

HOLLINGSWORTH, David; **Workflow Management Coalition: The Workflow Reference Model**. Hampshire: Workflow Management Coalition, 1995.

JENKINS, Nick. A Software Testing Primer. **An Introduction to Software Testing**, 2008. Disponível em: <<http://www.nickjenkins.net/prose/testingPrimer.pdf>>. Acesso em: 03 mar. 2013.

JOHNSON, Brian; SHNEIDERMAN, Bem. Tree-maps: a space-filling approach to the visualization of hierarchical information structures. In: Visualization'91, 1991. **Proceedings...** IEEE Conference, 1991. p. 284-291.

KAFURE, Ivette. **Usabilidade da imagem na recuperação da informação no catálogo público de acesso em linha**. 2004. 311f. Tese (Doutorado em Ciência da Informação) - Universidade de Brasília, Brasília, 2004.

KARLSON, A. K., *et al.* FaThumb: A Facet-based Interface for Mobile Search. **Navigation: Mobiles & Audio**, v. 22, n. 27, p. 711-720, apr. 2006.

KRUCHTEN, P. **Introdução ao RUP: Rational Unified Process**. Rio de Janeiro: Ciência Moderna, 2003. 255p.

LAI, Peter. **Facette: Using Facets to Improve Tag-based Bookmarking**. 2009. 52f. Dissertação (Master of Engineering in Electrical Engineering and Computer Science) - Massachusetts Institute of Technology, 2009.

LAKATOS, E. V.; MARCONI, M. A. **Metodologia científica**. 2. ed. São Paulo: Atlas, 1983. 250p.

LAKATOS, E. V.; MARCONI, M. A. **Metodologia do trabalho científico: procedimentos básicos, pesquisa bibliográfica, projeto e relatório, publicações e trabalhos científicos**. 2. ed. São Paulo: Atlas, 1992.

LANCASTER, F. W. **Indexação e resumos: teoria e prática**. Brasília: Briquet deLemos, 1993. 347p.

LERMAN, Kristina. **Dynamics of Collaborative Document Rating Systems**. In: 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis. **Proceedings...** New York: ACM, 2007. p. 46-55.

LEUF, Bo; CUNNINGHAM, Ward. **The Wiki Way: Quick Collaboration on the Web**. Boston: Addison-Wesley, 2001. 435p.

LIMA, Gercina. A. B. O. **Mapa hipertextual (MHTX): um modelo para organização hipertextual de documentos**. 2004.199f. Tese (Doutorado em Ciência da Informação) -

Escola de Ciência de Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2004.

LOPES, I. L. Uso das linguagens controlada e natural em bases de dados: Revisão da literatura. **Ciência da Informação**, Brasília, v. 31, n. 1, p. 41-52, jan./abr. 2002.

LOTUS DEVELOPMENT CORPORATION. **Groupware: Communication, Collaboration and Coordination**. Cambridge: Lotus Development Corporation, 1995. 41p.

LUO, L. **Software Testing Techniques, Technology Maturation and Research Strategies**. Pittsburgh: Carnegie Mellon University, 2001. Disponível em: <<http://mcahelpline.com/tutorials/testing/testing.pdf>>. Acesso em: 13 mar. 2012.

LUZZARDI, Paulo Roberto Gomes. **Critérios de Avaliação de Técnicas de Visualização de Informações Hierárquicas**. 2003. 248f. Tese (Doutorado em Ciência da Computação) - Instituto de Informática, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2003.

MACULAN, Benildes Coura Moreira dos Santos. **Manual de normalização: padronização de documentos acadêmicos do NITEG/UFMG e do PPGCI/UFMG**. 2. ed. atual. e rev. Belo Horizonte: UFMG, 2011. E-Book. ISBN 978-85-914076-0-6. Disponível em: <<http://www.eci.ufmg.br/normalizacao/>>. Acesso em: 13 abr. 2012.

MACULAN, Benildes Coura Moreira dos Santos. **Taxonomia facetada navegacional: construção a partir de uma matriz categorial para trabalhos acadêmicos**. 2011. 191f. Dissertação (Mestrado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2011.

MAKELA, E. *et al.*, Semantic Yellow Page Service Discovery: The Veturi Portal. In: 4th International Semantic Web Conference, 2005. **Proceedings...** Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.214.5211&rep=rep1&type=pdf>>. Acesso em: 02 set. 2011.

MALY, Kurt; WU, Harris; ZUBAIR, Mohammad. A collaborative faceted categorization system - user interactions. In: 14th International Conference on Electronic , 2010, Helsinki, Finland **Proceedings...** Transforming the Nature of Communication: Turid Hedlund and Yasar Tonta, 2010. Disponível em: <http://elpub.scix.net/data/works/att/106_elpub2010.content.pdf>. Acesso em: 19 abr. 2012.

MANNING, C. D.; RAGHAVAN, P.; SCHÜTZE, H. **An Introduction to Information Retrieval**. England: Cambridge University Press Cambridge, 2009. 544p. Disponível em: <<http://nlp.stanford.edu/IR-book/pdf/irbookprint.pdf>>. Acesso em: 12 out. 2012.

MARCONI, M. A.; LAKATOS, E. M. **Fundamentos de metodologia científica**. 5. ed. São Paulo: Atlas, 2003. 312p.

MEY, E. S. A; SILVEIRA, N. **Introdução à catalogação**. Brasília: Briquet de Lemos, 1995.

MICHAELIS. Dicionário. Disponível em: <<http://michaelis.uol.com.br/>>. Acesso em: 20 mar. 2013.

MONTALLI, Katia Maria Lemos; CAMPELLO, Bernardete dos Santos. Fontes de informação sobre companhias e produtos industriais: uma revisão de literatura. **Ciência da Informação**, Brasília, v. 26, n. 3, set. 1997. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19651997000300014&lng=en&nrm=iso>. Acesso em: 05 jan. 2013.

MYERS, G. J. **The Art of Software Testing**. 2. ed. New Jersey-EUA: John Wiley & Sons, 2004. 151p.

NAVES, M. M. L. **Curso de indexação**: princípios e técnicas de indexação, com vistas à recuperação da informação. Belo Horizonte: UFMG. 2004. Material. 23p. Apostila.

NICOLAO, M. **Adaptação em um curso na web usando workflow**. 2004. 249f. Tese (Doutorado em Ciência da Computação) - Universidade Federal do Rio Grande do Sul, Instituto de Informática, Porto Alegre, 2004.

NIELSEN, Jakob. **Ten Usability Heuristics**. Useit.com. Jan. 1995. Disponível em: <http://www.useit.com/papers/heuristic/heuristic_list.htm>. Acesso em: 22 mar. 2010.

NIELSEN, Jakob. **Why You Only Need to Test with 5 Users**. Useit.com. Mar. 2000. Disponível em: <<http://www.useit.com/alertbox/20000319.html>>. Acesso em: 07 fev. 2007.

NIELSEN, Jakob. **Usability 101: Introduction to Usability**. Useit.com. 2003. Disponível em: <<http://www.useit.com/alertbox/20030825.html>>. Acesso em: 19 mar. 2009.

NIELSEN, Jakob; LORANGER, Hoa.; FURMANKIEWICZ, Edson. **Usabilidade na Web**: projetando websites com qualidade. Rio de Janeiro: Elsevier, Campus, 2007. 406p.

NISO-NATIONAL INFORMATION STANDARDS ORGANIZATION. **Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies**. Z39.19. Maryland: ANSI/NISO, 2005. 172p.

NISO-NATIONAL INFORMATION STANDARDS ORGANIZATION. **Guidelines for the Construction, Format, and Management of Monolingual Thesauri**. Z39.19. Maryland: ANSI/NISO, 1993. 69p.

NIU, X. **Beyond text queries and ranked lists: faceted search in library Catalogs**. 2012. 211p. Tese (Doctor of Philosophy) - School of Information and Library Science, University of North Carolina at Chapel Hill, 2012.

O'REILLY, Tim. **What Is Web 2.0** Design Patterns and Business Models for the Next Generation of Software. O'REILLY. 2005. Disponível em: <<http://oreilly.com/web2/archive/what-is-web-20.html>>. Acesso em: 14 out. 2009.

OLIVEIRA, Carla Cristina Vieira de. **A interação de usuários com o catálogo online do sistema *pergamum***. 2008. 200f. Dissertação (Mestrado em Ciência da Informação) - Universidade Federal de Minas Gerais, Belo Horizonte, 2008.

OLIVEIRA, L. S.; GEROSA, Marco A. Collaborative Features in Content Sharing Web 2.0 Social Networks: A Domain Engineering based on the 3C Collaboration Model. In: 17th Conference on Collaboration and Technology, 2011. **Proceedings...** CRIWG, 2011.

OMG-OBJECT MANAGEMENT GROUP. **OMG UML Superstructure**. Version 2.4.1. OMG, 2011. 230p.

OREN, Eyal; DELBRU, Renaud; DECKER, Stefan. **Extending faceted navigation for RDF data**. In: 5th International Semantic *Web* Conference, 2006. **Proceedings...** ISWC, 2006. Disponível em: <<http://iswc2006.semanticweb.org/items/Oren2006nx.pdf>>. Acesso em: 10 de jul. 2007.

PÁDUA, C. I. P. S. **Engenharia de Usabilidade**: Material de Referência. Belo Horizonte: Instituto de Ciências Exatas Departamento de Ciência da Computação, UFMG, 2012. 193p. Apostila.

PÁDUA, C. I. P. S. **Teste de usabilidade com usuários**. Belo Horizonte: Instituto de Ciências Exatas Departamento de Ciência da Computação, UFMG, 2010. Slide. Disponível em: <<http://homepages.dcc.ufmg.br/~clarindo/arquivos/disciplinas/eu/material/transparencias/topicos/12-2-teste-com-usuario.pdf>>. Acesso em: 3 nov. 2010.

PÁDUA, Clarindo Isaías Pereira da Silva e. **Prototipação e Engenharia de Usabilidade**. Belo Horizonte: Laboratório Gestus/Synergia, Departamento de Ciência da Computação, ICEx/ UFMG, 2007. Slide.

PEIXOTO, Daniela C.C. *et al.* A Comparison of BPMN and UML 2.0 Activity Diagrams. In: VII Simpósio Brasileiro de Qualidade de Software, 2008, Florianópolis. **Anais...** Belo Horizonte: Departamento de Ciência da Computação, UFMG. 14p. Disponível em: <<http://www.synergia.dcc.ufmg.br/publicacao01.html>>. Acesso em: 14 de jun. 2009.

PEREIRA, F. **Avaliação de usabilidade em bibliotecas digitais**: um estudo de caso. 2011. 123f. Dissertação (Mestrado em Ciência da Informação) - Universidade Federal de Minas Gerais, Escola de Ciência da Informação, Belo Horizonte, Minas Gerais, 2011.

PIEIDADE, M. A. R. **Introdução à teoria da classificação**. 2. ed., Rio de Janeiro: Interciência, 1983. 221p.

PIMENTEL, M. **RUP-3C-Groupware**: um processo de desenvolvimento de groupware baseado no Modelo 3C de Colaboração. 2006, p. 178. Tese (Doutorado em Informática) - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2006.

PINHEIRO, L. V. R. Fontes ou recursos de informação: categorias e evolução conceitual. **Pesquisa Brasileira em Ciência da Informação e Biblioteconomia**, v. 1, n. 1, 2006.

SACCO, G. M.; TZITZICAS, Y. **Dynamic Taxonomies and Faceted Search: theory, practice, and experience**. Berlin Heidelberg: Springer-Verlag, 2009. 340p.

SANTANA, V. A. **MEMÓRIA ESQUECIDA: uma análise da organização e representação da informação étnico-racial no OPAC da Biblioteca Central da Universidade Federal da Paraíba**. 2012. 139f. Dissertação (Mestrado em Ciência da Informação) - Centro de Ciências Aplicadas, Universidade Federal da Paraíba, João Pessoa, 2012.

SANTOS, Edna Maria dos. **Um Modelo para Desenvolver Editores Colaborativos Hierárquicos**. 2009. 96f. Dissertação (Mestrado em Informática) - Instituto Tecnológico de Aeronáutica, São Jose dos Campos, 2009.

SAUVÉ, Jacques Philippe. **Frameworks: O que é um framework?**. Campina Grande: Universidade Federal de Campina Grande, 2000. Disponível em: <<http://www.dsc.ufcg.edu.br/~jacques/cursos/map/html/frame/oque.htm>>. Acesso em: 2 jul. 2012

SAYÃO, L.F. Modelos teóricos em ciência da informação; abstração e método científico. **Ciência da Informação**, Brasília, v. 30, n. 1, p. 82-91, jan./abr. 2001.

SHNEIDERMAN, Ben . **Designing the user interface: strategies for effective human-computer-interaction**. Reading: Addison-Wesley, 1998.

SHREVES, R. **Open Source CMS Market Share Report**. Singapore: Water & Stone, 2011.

SIGURBJÖRNSSON, B., ZWOL, R. **Tagexplorer: Faceted browsing of flickr photos**. Yahoo! Research, 2010. Technical Report.

SILVA, Alessandra Rodrigues da. **Estudo dos princípios de categorização na Biblioteconomia e Ciência da Informação: Ranganathan - entre a teoria clássica e a abordagem cognitiva contemporânea**. 2010. 193f. Dissertação (Mestrado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2010.

SILVA, E. L.; MENEZES, E. M. **Metodologia da pesquisa e elaboração de dissertação**. 3. ed. Florianópolis: Laboratório de Ensino a Distância da UFSC, 2001. Disponível em [http://projetos.info.ufsc/arquivos/metodologia da pesquisa.3 edicao.pdf](http://projetos.info.ufsc/arquivos/metodologia%20da%20pesquisa.3%20edicao.pdf)>. Acesso em: 08 ago. 2012.

SMIT, J. W.; KOBASHI, N. Y. **Como elaborar vocabulário controlado para a aplicação em arquivos**. São Paulo: Imprensa Oficial do Estado de São Paulo, 2003. 56p.

SMITH, G., *et al.* FacetMap: A Scalable Search and Browse Visualization. **IEEE Transactions on Visualization and Computer Graphics**, v. 12, n. 5, p. 797-804, 2006.

SPITERI, Louise. **Faceted navigation of social tagging applications**. Dalhousie Repository. Halifax: Dalhousie University, 2011. Disponível em: <<http://dalspace.library.dal.ca/handle/10222/14297>>. Acesso em: 13 ago. 2012.

STACHOWIAC, Herbert. **Models. Scientific thought: concepts, methods and procedures**, Paris, p. 145-166, 1972.

STADZISZ, Paulo César. **Projeto de Software usando a UML**. Versão 2002. Curitiba: CEFET-PR, 2002. 69p. Apostila.

SUOMINEN, Osmo; VILJANEN, Kim; HYVÖNEN, Eero. User-centric Faceted Search for Semantic Portals. **The Semantic Web: Research and Applications**. Springer Berlin Heidelberg, 2007. p. 356-370. Disponível em: <<http://gate.ac.uk/sale/dd/related-work/2007+suominen-et-al-user-centric-faceted-search.pdf>>. Acesso em: 09 ago. 2011.

TRISTÃO, A.M.D.; FACHIN G.R.B.; ALARCON, O.E. Sistema de classificação facetada e tesouros: instrumentos para organização do conhecimento. **Ciência da Informação**, Brasília, v.33, n.2, p.161-171, maio/ago. 2004.

TUNKELANG, D. Dynamic category sets: An approach for faceted search. In: ACM SIGIR Workshop On Faceted Search , 2006. **Proceedings...** Seattle, WA, 2006. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.83.4202&rep=rep1&type=pdf>>. Acesso em: 20 nov. 2012.

TVAROZEK, Michal; BIELIKOVAR, Mária. Personalized faceted navigation for collections. In: Semantic Media Adaptation and Personalization, Second International Workshop, 2007. **Proceedings...** IEEE, 2007. p. 104-109.

TZITZIKAS, Yannis. *et al.* Extended Faceted Taxonomies for Web Catalogs. In: Web Information Systems Engineering - WISE, Third International Conference, 2002. **Proceedings...** IEEE, 2002. p. 192-201.

VICKERY, G.; WUNSCH-VINCENT, S. **Participative Web And User-Created Content: Web 2.0 Wikis and Social Networking**. Paris: Organization for Economic Cooperation and Development - OECD, 2007. 129p. Disponível em: <www.oecd.org/publishing/corrigenda>. Acesso em: 10 mai. 2011.

VIEIRA, S. B. INDEXAÇÃO AUTOMÁTICA E MANUAL: REVISÃO DE LITERATURA*. **Ciência da Informação**, Brasília, v. 17, n. 1, p. 43-57, 1988. Disponível em: <http://repositorio.unb.br/bitstream/10482/12901/1/ARTIGO_IndexacaoAutomaticaManual.pdf>. Acesso em: 23 jan. 2012.

VIEIRA, D. V.; CARVALHO, E. B.; LAZZARIN, F. A. Uma Proposta de modelo baseado na Web 2.0 para as bibliotecas das Universidades Federais. In: 9 ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 2008. **Anais...** São Paulo: ANCIB, 2008. Disponível em: <<http://www.bax.com.br/teaching/courses/bibliotecasdigitais/referencias/PROPOSTA.pdf>>. Acesso em: 21 jan. 2011.

WAL, T. V. **Folksonomy Definition and Wikipedia**. Vanderwal. 2005. Disponível em: <<http://www.vanderwal.net/random/category.php?cat=153>>. Acesso em: 11 nov. 2008.

WEINFURTER, M. R. **Modelagem de Sistemas de Informação na Internet como Sistemas de Workflow sobre Documentos Estruturados**. 2001. 105p. Dissertação

(Mestrado em Ciência da Computação) - Instituto de Informática, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2001.

WEINSTEIN, P. C. Ontology-based metadata: transforming the MARC Legacy. In: 3th ACM DIGITAL LIBRARY CONFERENCE, 1998. **Proceedings...** Pittsburgh, PA: ACM, 1998.

WELLER, K. *et al.* Folksonomies and Ontologies: Two New Players in Indexing and Knowledge Representation. **Applying Web 2.0**, v. 2, p. 108-115, 2007.

WIKIPÉDIA. **Unified Modeling Language**. 2011. Disponível em: <http://en.wikipedia.org/wiki/Unified_Modeling_Language>. Acesso em: 30 out. 2011.

WIKIPÉDIA. **Consenso**. 2012a. Disponível em: <<http://pt.wikipedia.org/wiki/Wikip%C3%A9dia:Consenso>>. Acesso em: 10 jan. 2012.

WIKIPÉDIA. **Dispute resolution**. 2012b. Disponível em: <http://en.wikipedia.org/wiki/Wikipedia:Dispute_resolution>. Acesso em: 02 mar. 2012.

WIKIPEDIA. **Ad hoc testing**. 2013. Disponível em: <http://en.Wikipédia.org/wiki/Ad_hoc_testing>. Acesso em: 01 fev. 2013.

WILKINSON, K. *et al.* Efficient Storage and Retrieval in Jena2. In: WORKSHOP ON SEMANTIC WEB AND DATABASES, 2003. **Proceedings...** Berlin: 2003. Disponível em: <http://www.cs.uic.edu/~ifc/SWDB/papers/Wilkinson_etal.pdf>. Acesso em: 14 ago. 2013.

WIVES, Leandro. **Tecnologias de Descoberta de Conhecimento em Textos aplicadas à Inteligência Competitiva**. 2002. 100f. Exame de Qualificação - Pós-Graduação em Computação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2002.

YANG, Sharon Q.; HOFMANN, Melissa A. The next generation library catalog: A comparative study of the OPACs of Koha, Evergreen, and Voyager. **Information Technology and Libraries**, v. 29, n. 3, p. 141-150, 2013.

YEE, K-P.; SWEARINGEN, K.; HEARST, M. Faceted metadata for image search and browsing. In: SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, 2003. **Proceedings...** ACM, 2003. p. 401-408. Disponível em: <<http://kevinli.net/flamenco-chi03.pdf>>. Acesso em: 11 mai. 2012.

APÊNDICE A - TERMO DE CONSENTIMENTO

Eu _____, RG, _____, autorizo o pesquisador Marcel Ferrante Silva, doutorando da ECI/UFMG, a utilizar-se das informações obtidas no teste de usabilidade do qual participo, por meio do protocolo escrito e do mapeamento das telas do computador, obedecendo aos critérios da ética de pesquisa, em que está assegurado o total anonimato.

Declaro-me ciente e concordo com o acima exposto.

Assinatura do Participante

_____/_____/_____

Data

APÊNDICE B - QUESTIONÁRIO PARA IDENTIFICAÇÃO DO PERFIL DO PARTICIPANTE

O objetivo deste questionário é identificar o perfil adequado dos participantes para a realização do teste de usabilidade da interface do catálogo de endereços de Goiânia. As informações fornecidas são essenciais para o aprimoramento do catálogo. Por favor, leia com atenção as questões a seguir e em caso de dúvida, solicite esclarecimento com o avaliador.

1) Uso de catálogo de endereços

Você já utilizou um catálogo de endereços *web* na Internet ?

- a. () Sim, uso frequentemente
- b. () Sim, a muito tempo atrás
- c. () Não

2) Informações Pessoais

Qual é a sua idade? _____ anos.

3) Informações Educacionais

Qual é o seu grau de instrução?

- a. () 3º grau incompleto
- b. () 3º grau completo
- c. () Mestrado incompleto
- d. () Mestrado completo

4) Experiência Profissional

Qual é a sua profissão?

5) Experiência Computacional

Há quanto tempo você utiliza computador?

- a. () Há menos de 1 ano.
- b. () Entre 1 anos e 2 anos
- c. () Entre 3 anos e 4 anos
- d. () Há mais de 4 anos.

APÊNDICE C - LISTA DE TAREFAS

Tarefa 1 - Cadastrar

Abra o navegador de Internet e acesse o portal www.oouniverse.com.br e clique no numero correspondente ao usuário que lhe foi designado.

Esta lista de empresas deve ser cadastradas no catálogo de empresas de Goiânia. Abra um outra janela do navegador de Internet e digite a seguinte URL: www.oouniverse.com.br/catalogo.

Para você cadastrar uma empresa, Clique no botão Adicionar Empresa, e faça o login no sistema, sendo que a senha é igual ao seu usuário que lhe foi fornecido.

Para cadastrar as empresas, você pode usar as categorias fornecidas ou usar suas próprias palavras para descrevê-las. Você tem 10 minutos para cadastrar a quantidade de empresas que conseguir.

Tarefa 2 - Encontrar

Realize as pesquisas abaixo da maneira que preferir, pelo campo de busca por palavra-chave ou pelo painel de navegação nas categorias, as seguintes empresas abaixo.

Se possível, ajude-nos a qualidade melhorar o catálogo através dos recursos da interface. Pode fazer essa tarefa com tranquilidade, você tem 15 minutos para realizá-la. Se você não conseguir realizar a tarefa completamente, não tem problema.

- 1) Encontre 5 empresas para hospedagem no bairro Jardim Vitória em Goiânia
- 2) Encontre 3 empresas que vedem sanduíche nos bairros próximos a UFG
- 3) Encontre por 2 empresas para reforma em casa com as especialidades: Pintura, Janela de alumínio e Instalação Elétrica

Tarefa 3 - Votar

- 1) Eleja as melhores faculdades/universidades de Goiânia

Tarefa 4 - Revisar

Aponte erros na classificação das empresas feitas pelo seu colega com o número de usuário consecutivo no catálogo: www.oouniverse.com.br/catalogo

Para acessar a lista de empresas que seu colega cadastrou digite no navegador de Internet: www.oouniverse.com.br/seu_usuario+1

O último usuário deve apontar os erros feitos pelo usuário 1.

APÊNDICE D - QUESTIONÁRIO DE SATISFAÇÃO DO USUÁRIO

Qual opção você achou mais fácil ou preterida para classificar as empresas: usar as categorias fornecidas ou descrevendo com suas próprias palavras?

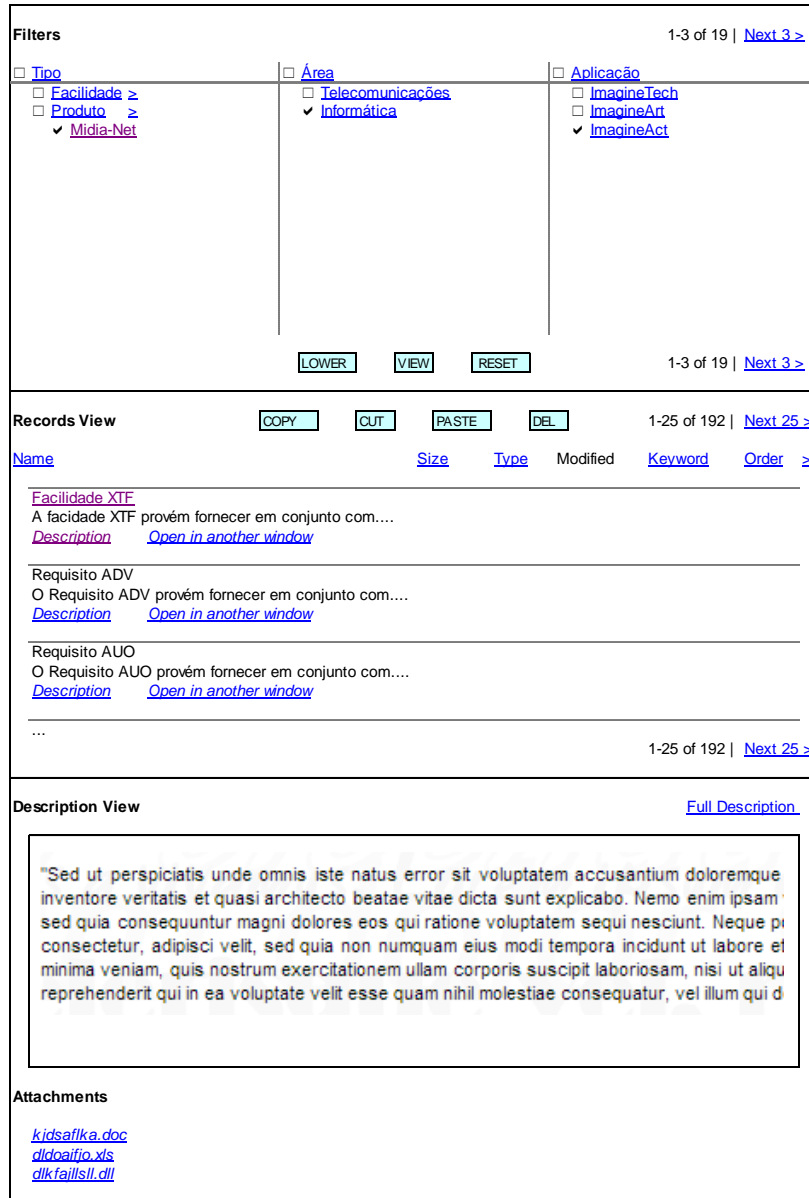
Qual opção você achou mais fácil ou preterida para encontrar as empresas: navegar e selecionar as categorias fornecidas ou buscar com suas próprias palavras?

Você achou correto o resultado da votação das universidades realizadas pelos usuários?

Você teve alguma dificuldade para apontar o erro na classificação?

APÊNDICE E - PRIMEIRO PROTÓTIPO DA INTERFACE FACETADA

FIGURA 92 - Primeiro protótipo de baixa fidelidade da interface facetada



Fonte: Elaborado pelo autor.

APÊNDICE F - VOCABULÁRIO CONTROLADO DA TESE

A seguir, no QUADRO 8, é apresentado o vocabulário controlado da tese, com os principais termos empregados ao longo do texto. Os termos preferenciais e seus relacionamentos representam uma opção do autor. A legenda está logo após o vocabulário controlado.

QUADRO 8 – Vocabulário controlado da tese

(continua)

Termos em português	Termos em inglês
Navegação facetada UP Busca facetada TR Interface facetada	Faceted search UF Faceted navigation UF Faceted browse UF Fielded search
Interface facetada UP Navegador facetado UP Explorador facetado TR Navegação facetada	Faceted interface UF Faceted navigator UF Faceted explorer UF Faceted browser TR Faceted search
Indexação de assuntos UP Classificação temática UP Categorização TG Indexação	Subject Indexing UF Thematic classification UF Categorization BT Indexing
Indexação NE usado para designar o processo da indexação UP Classificação UP Catalogação TE Indexação manual TE Etiquetagem TE Etiquetagem social TE Indexação automática TE Indexação de assuntos	Indexing UF Classification UF Cataloging NT Manual Indexing NT Tagging NT Tagging social NT Automatic Indexing NT Subject Indexing
Classificação NE é o produto da indexação TE Folksonomia UP Indexação	Classification NT Folksonomy UF Indexing
Recurso <i>web</i> UP Recurso na <i>web</i> UP Recurso da <i>web</i> TG Recurso de informação	Web resource BT Information Resource

(continua)

<p>Item</p> <p>UP Objeto primário</p> <p>TE Recurso de informação</p> <p>TE Recurso <i>web</i></p> <p>TR Catálogo</p>	<p>Item</p> <p>UF Primary object</p> <p>NT Information Resource</p> <p>NT Web resource</p> <p>RT Catálogo</p>
<p>Registro</p> <p>NE é a representação de um item</p> <p>UP Objeto Secundário</p> <p>TR Catálogo</p>	<p>Record</p> <p>UF Secondary object</p> <p>RT Catalog</p>
<p>Vocabulário controlado</p> <p>UP Linguagem artificial</p> <p>UP Linguagem controlada</p> <p>UP Linguagem de indexação</p> <p>UP Linguagem documentária</p>	<p>Controlled vocabulary</p> <p>UF Artificial language</p> <p>UF Controlled language</p> <p>UF Indexing language</p> <p>UF Documentary language</p>
<p>Termo</p> <p>TE Etiqueta</p> <p>TE Categoria</p>	<p>Termo</p> <p>TE Tag</p> <p>TE Category</p>
<p>Etiqueta</p> <p>TG Termo</p>	<p>Tag</p> <p>BT Term</p>
<p>Etiquetagem</p> <p>NE usado para designar o processo de indexação</p> <p>TG Indexação manual</p> <p>TR Marcação</p>	<p>Tagging</p> <p>BT Manual Indexing</p> <p>RT Bookmarking</p>
<p>Etiquetagem social</p> <p>NE usado é a etiquetagem publica ou compartilhada</p> <p>UP Etiquetagem colaborativa</p> <p>UP Indexação social</p> <p>UP Indexação colaborativa</p> <p>UP Etnoclassificação</p> <p>TG Etiquetagem</p> <p>TR Marcação social</p>	<p>Social tagging</p> <p>UF Collaborative tagging</p> <p>UF Social indexing</p> <p>UF Collaborative indexing</p> <p>UF Social classification</p> <p>BT Tagging</p> <p>RT Social Bookmarking</p>
<p>Folksonomia</p> <p>NE usado para designar o produto da indexação</p> <p>UP Etnoclassificação</p> <p>TG Classificação</p> <p>TR Etiquetagem</p>	<p>Folksonomy</p> <p>UF Social classification</p> <p>BT Classification</p> <p>RT Tagging</p>
<p>Busca por palavra-chave</p> <p>UP Busca textual</p> <p>UP Busca direta</p>	<p>Keyword search</p> <p>UF Text search</p> <p>UF Direct search</p>

(continua)

Faceta UP Categoria fundamental UP Categoria experimental TR Etiquetação	Facet UF Fundamental category UF Elementary category RT Etiquetação
Categoria UP Classe TG Termo	Category UF Class BT Term
Taxonomia facetada TG Vocabulário controlado TR Sistema de classificação facetada	Faceted taxonomy BT Controlled vocabulary RT Faceted classification system
Sistema de classificação facetada NE é similar a taxonomia facetada acrescentando-se códigos para a organização dos documentos UP Esquema de classificação facetada TG Vocabulário controlado TR Taxonomia facetada	Faceted classification system UF Faceted classification schema BT Controlled vocabulary RT Faceted taxonomy
Web 2.0 TR Web participativa	Web 2.0 TR Participative web
Sistema colaborativo UP Sistema cooperativo UP Sistema colaborativo	Collaborative system UF Cooperative system UF Collaborative software UF Collaborative platform UF Groupware UF Social software UF Computer-Supported Cooperative Work UF CSCW
Conteúdo criado pelo usuário UP Conteúdo gerado pelo usuário UP CGU UP CCU	User created content UF User generated content UF UCC UF UGC
Protótipo de baixa fidelidade UP Protótipo de alto nível	Low fidelity prototype UF High level prototype
Protótipo de alta fidelidade UP Protótipo de baixo nível	High fidelity prototype UF Low level prototype
Ação TE Navegação TE Busca TE Contribuição TE Adicionar registro TE Contestação	Action NT Navigation NT Search NT Contribution NT Add a record NT Contestation

(conclusão)

Adicionar registro NE ação de representação de um item através de um registro UP Inserir registro UP Cadastrar registro UP Criar registro TG Contribuição TR Registro	Add a record UF Insert a record UF Register a record UF Create a record BT Contribution RT Record
Contestação NE ação de revisão de um registro ou da taxonomia faceta UP Solicitação UP Revisão TG Contribuição	Contestation UF Solicitation UF Revision BT Contribution

Fonte: Elaborado pelo autor.

QUADRO 9 – Legenda

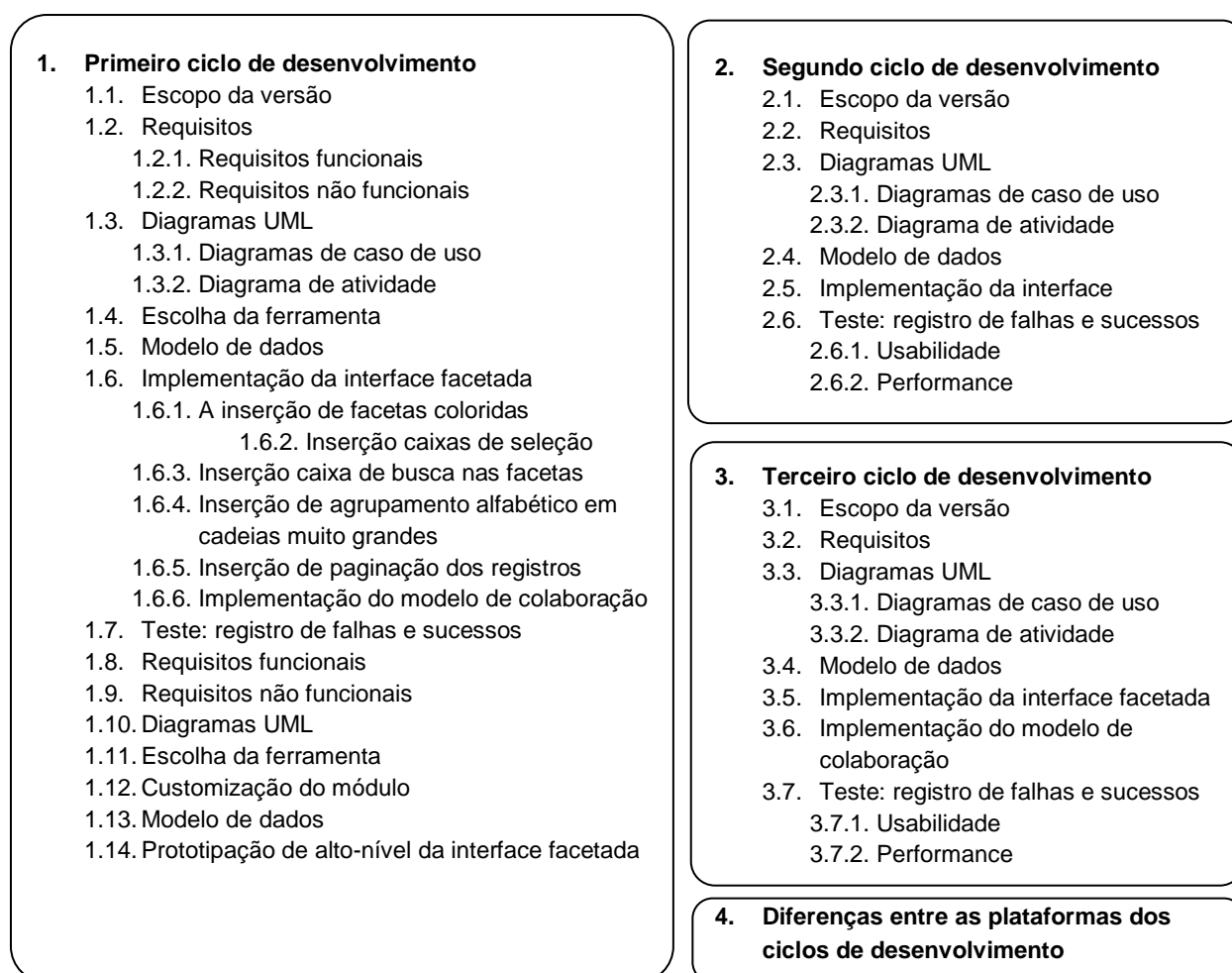
Português	Inglês
NE - Nota de Escopo	SN - Scope Note
USE - Termo Preferencial	U - Use
UP - Use Para - Termo Não Preferencial	UF - Use For
TG - Termo Genérico	BT - Boader Term
TGP - Termo que representa o todo (ex. carro)	TGP - Boader Term Partitive
TE - Termo Específico	NT - Narrower Term
TEP - Termo que representa a parte (ex. roda)	NTP - Narrower Term Partitive
TR - Termo Relacionado	RT - Related Term

Fonte: NISO (1993)

APÊNDICE G - CICLOS DE DESENVOLVIMENTOS ANTERIORES

O desenvolvimento do catálogo *web* facetado passou por vários ciclos de desenvolvimento. A cada ciclo, o modelo de colaboração ficava menos vago e o escopo do projeto ficava mais detalhado, conforme o processo iterativo de desenvolvimento de *software*. Na FIG. 93 é apresentado o que ocorreu nos primeiros ciclos, em uma plataforma que foi descontinuada, mas que permitiu um grande aprendizado para a realização do protótipo descrito no capítulo 5.

FIGURA 93 - Ciclos iniciais de desenvolvimento



Fonte: Elaborado pelo autor.

G.1 Primeiro ciclo de desenvolvimento

G.1.1 Escopo da versão

Os objetivos principais dessa versão do catálogo *web* facetado foram: desenvolvimento de um catálogo com navegação e classificação facetada (seção 2.1); permitir aos membros da comunidade de usuários realizar a manutenção do catálogo *web* através de um modelo de colaboração *wiki* (seção 2.7.3).

Em linhas gerais, o modelo de colaboração *wiki* permite a publicação imediata de um item criado ou alterado, sem a necessidade de revisão. Entretanto, para evitar vandalismos e garantir alguma qualidade, uma contribuição pode ser anulada por outro membro da comunidade. E esta anulação pode ser revertida (seção 2.7.3.1). Assim, para os fins desse trabalho, além ação Adicionar que é livre para qualquer usuário logado, em qualquer momento, foram criadas mais duas ações *wiki* básicas: invalidar e validar. O invalidar tem o mesmo efeito de remover, pois envia o item para a lixeira, ocultando o mesmo. O validar significa uma reversão ao ato de invalidar.

G.1.2 Requisitos

G.1.2.1 Requisitos funcionais

Os requisitos funcionais do catálogo *web* são:

Referente à administração (apenas para o usuário administrador):

- Criar um catálogo
 - Determinar o nome do catálogo, o tipo de item que será cadastrado (ex. obras literárias, filmes, músicas, endereços);
 - Escolher as facetas do catálogo dentre as categorias já existentes ou criar novas facetas;
 - Deverá abrir um painel com a estrutura de categorias existentes no sistema de gestão de conteúdo;
 - Criar as categorias iniciais de cada faceta;
 - Explicitar o tipo de votação que está sendo realizada (ex. melhor, pior, mais barato, mais econômico, entre outras).

Referente à navegação facetada (para qualquer usuário do catálogo, sem a necessidade de *login*):

- Navegar hierarquicamente pelas categorias de uma faceta, através de um lista indentada
- Selecionar uma categoria de cada faceta
- Ao escolher a categoria, deixá-la em evidência (como por exemplo em negrito)
- Ao clicar em visualizar, devem ser mostrados os registros que são simultaneamente classificados com as categorias selecionadas em cada faceta
 - Deverá ser feita uma paginação dos registros mostrados, exibindo uma quantidade de registros por página e uma listagem de páginas (exemplo: < 1 2 3 4 ... 8 >)
- Ao entrar no catálogo pela primeira vez deve mostrar todos os registros, como se todas as categorias raízes estivessem selecionadas por padrão
- Possibilidade de ordenar os registros recuperados pela ordem de criação ou pela votação

Referente ao modelo de colaboração *wiki* (apenas para usuário logado e membro da comunidade de usuários):

- Adicionar um registro no catálogo, inserindo seu nome, descrição e classificando o registro com uma categoria de cada faceta;
- Adicionar o vínculo de um registro a uma categoria, ação denominada de adicionar classificação;
 - Deve ser exibida uma lista indentada das categorias;
- Invalidar classificação. Atua como se fosse um voto negativo do usuário, afirmando que o vínculo do registro com a categoria não deveria ocorrer, ocultando-o;
- Validar classificação. Atua como se fosse um voto positivo do usuário, afirmando que o vínculo do registro com a categoria deveria ocorrer, exibindo-o.

Referente ao sistema de votação:

- Listar a nota que cada item tem em relação a votação;
 - Esta nota deve ser simbolizada através de estrelinhas, indo de zero a cinco pontos;
- Exibir o tipo de votação que está sendo realizada;

- Permitir ao usuário votar, atribuindo uma nota ao registro.

G.1.2.2 Requisitos não funcionais

Os requisitos não funcionais do ambiente do catálogo *web* facetado colaborativo são apresentados a seguir:

- Acesso à *web* por meio de um navegador de Internet utilizado como programa cliente, que solicita, recebe e interpreta o código em HTML, *Javascript* e CSS. Permite que o protótipo seja acessado de qualquer lugar do mundo, e por vários usuários ao mesmo tempo.
- Utilização de *softwares* livres, que permite a realização, o desenvolvimento do aplicativo de forma mais rápida. Isso evita que "reinvente-se" a roda, dando continuidade às pesquisas já realizadas e contribuindo para a comunidade de *software* livre. Além disso, evita custos com licenças de *software*, tanto no ambiente do usuário quanto no ambiente do desenvolvedor.
- O *framework* deveria possuir ferramentas nativas como cadastro de usuários, controle de acesso e perfis por um super-administrador do sistema que também configura o catálogo. Um sistema de temas para formatar a informação que é impressa na tela também é requerido.
- Deve ser simples, rápido e flexível de implementar com uma boa curva de aprendizado, principalmente na customização ou implementação de novos módulos para o sistema.
- Deve ser bem documentado, do ponto de vista de usuário, administrador e desenvolvedor.
- O sistema deveria ter um tempo de resposta inferior a 2 segundos, para não tornar lentas e cansativas as tarefas de busca e navegação.

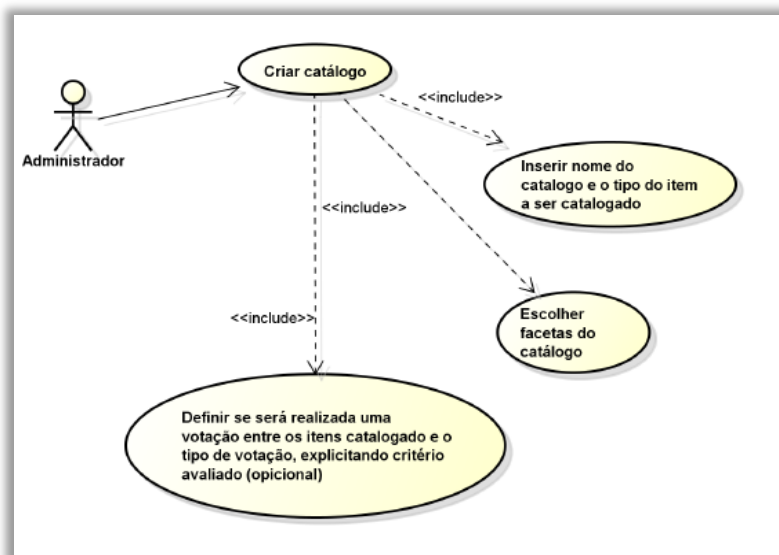
G.1.3 Diagramas UML

Nesta seção são apresentados os diagramas de caso de uso e de atividade realizados para orientar a implementação nesse ciclo de desenvolvimento.

G.1.3.1 Diagramas de caso de uso

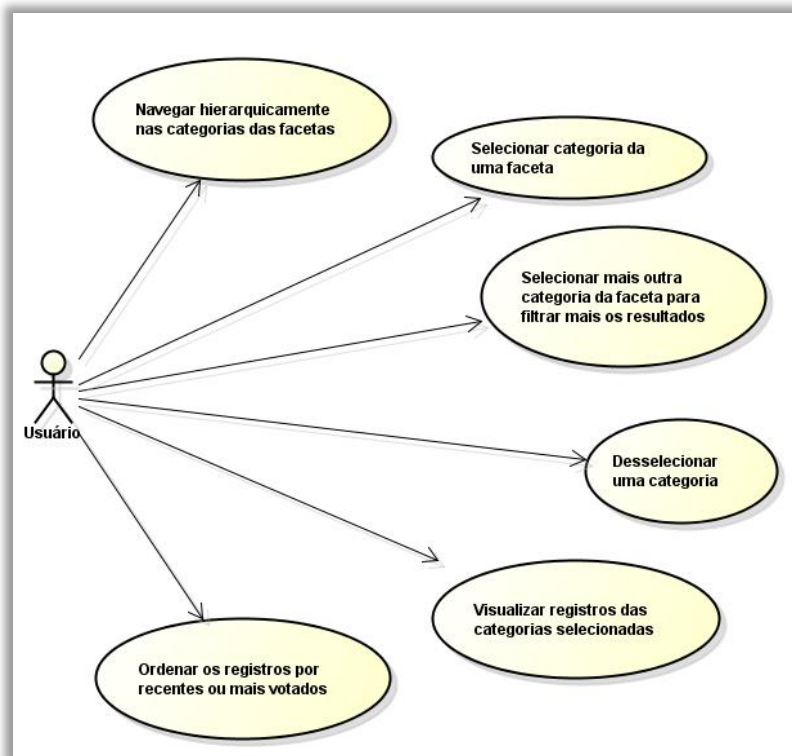
Os diagramas de caso de uso que detalham as funcionalidades e ações do usuário no catálogo *web* facetado colaborativo são: criar catálogo (FIG. 94), realizar navegação facetada (FIG. 95) e modelo de colaboração *wiki* (FIG. 96).

FIGURA 94 - Criar catálogo

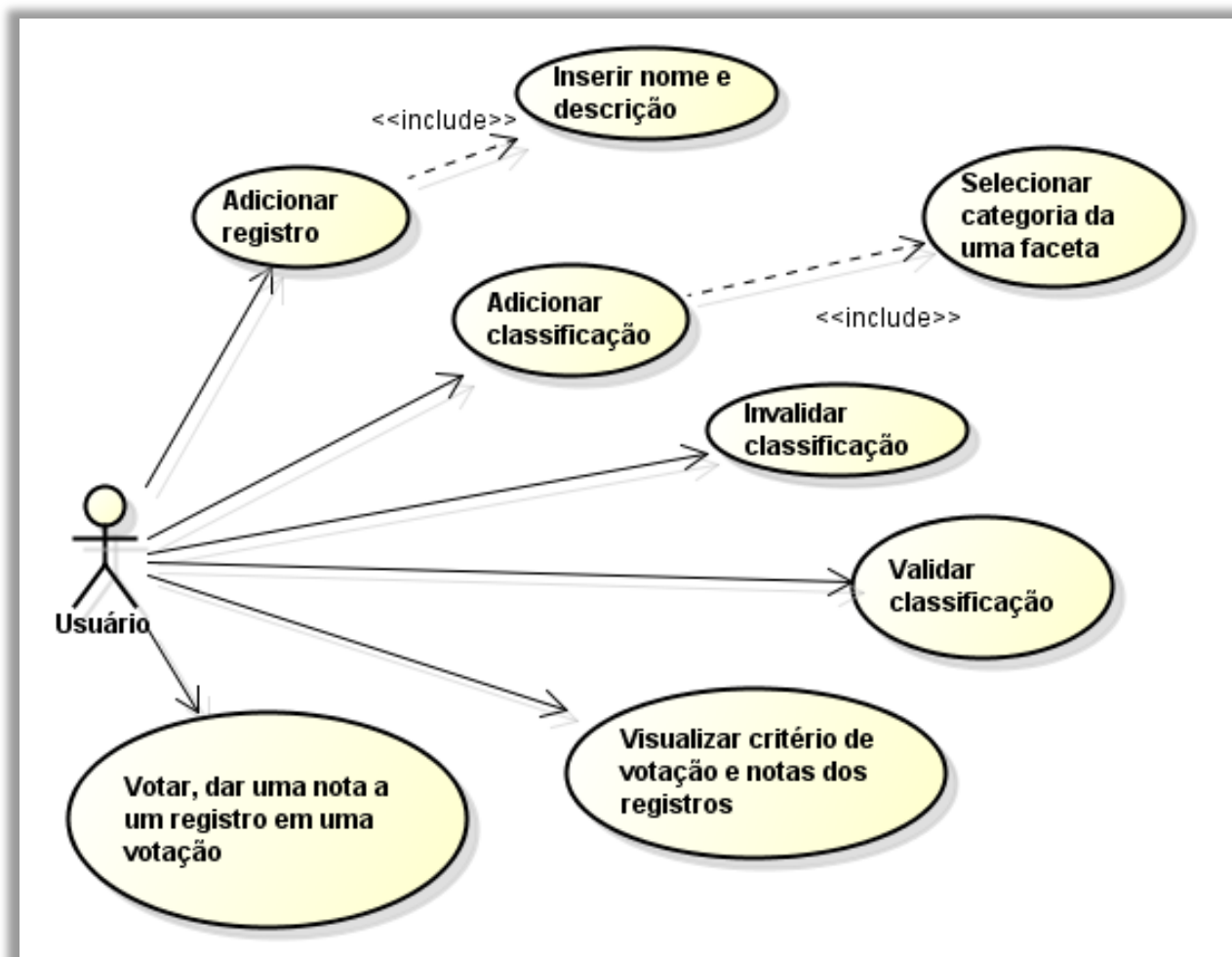


Fonte: Elaborado pelo autor.

FIGURA 95 - Realizar navegação facetada



Fonte: Elaborado pelo autor.

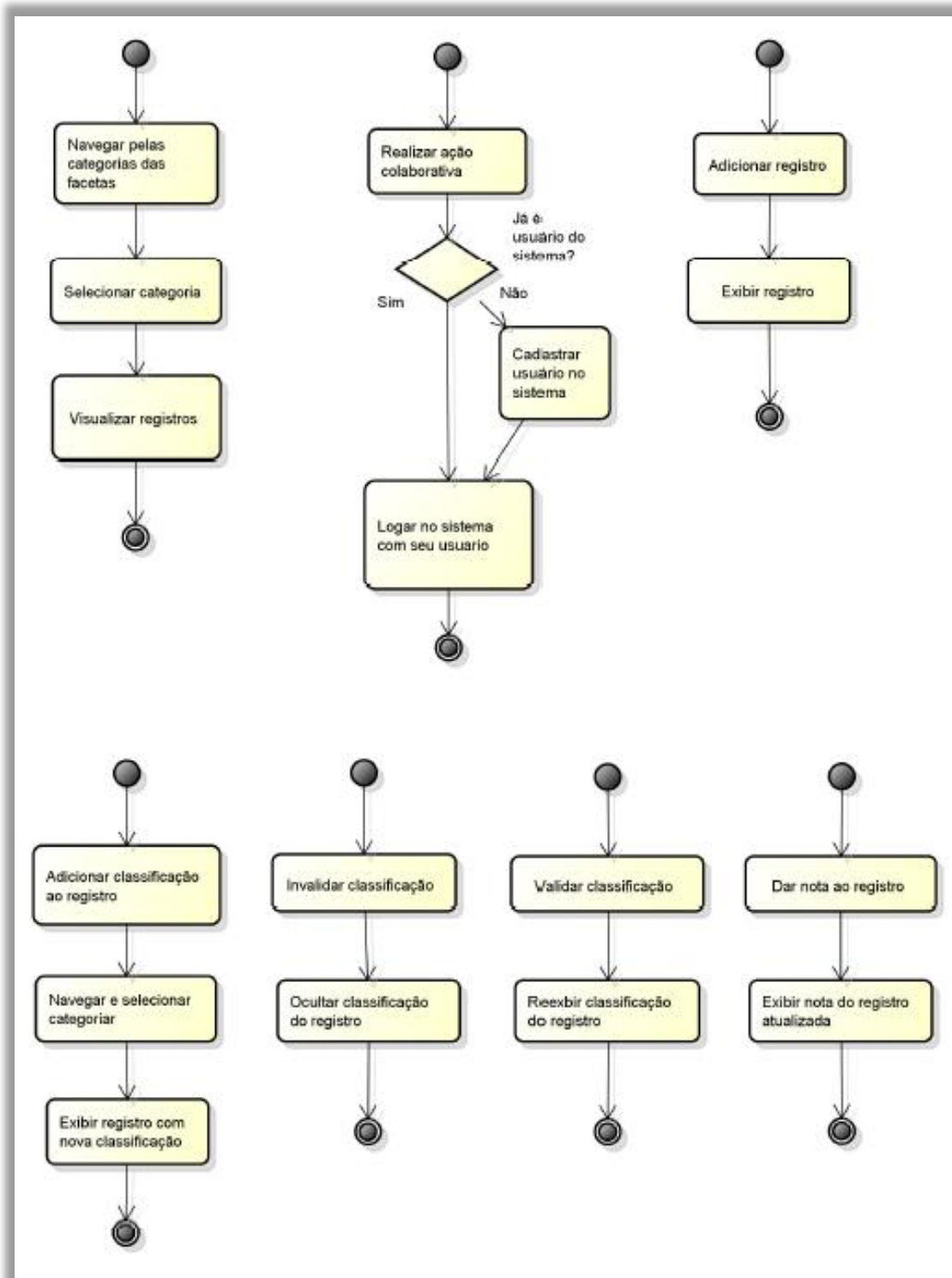
FIGURA 96 - Modelo de colaboração *wiki*

Fonte: Elaborado pelo autor.

G.1.3.2 Diagrama de atividade

O diagrama de atividade na FIG. 97 mostra os fluxos de controle no catálogo *web* facetado.

FIGURA 97 - Navegação facetada e modelo de colaboração



Fonte: Elaborado pelo autor.

G.1.4 Escolha da ferramenta

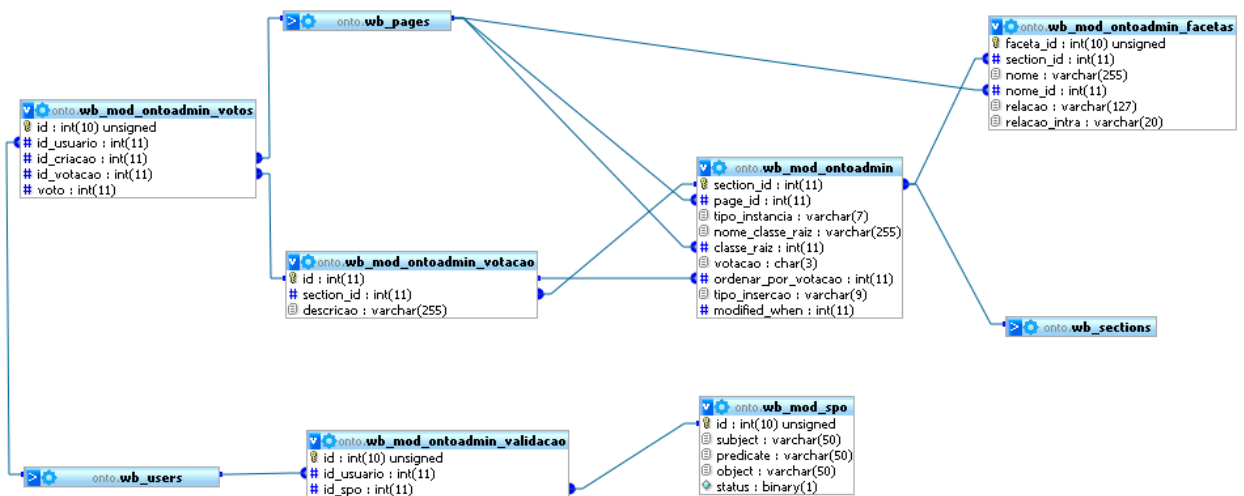
Dentre os sistemas de gestão de conteúdo escritos com a linguagem PHP, utilizando o banco de dados MySQL, foi escolhido o *WebsiteBaker*. O *WebsiteBaker* é um *software* livre sob a licença GPL, com mais de meio milhão de páginas construídas com este sistema⁷⁷. É um CMS que prima pela facilidade de uso, com uma interface simples e amigável, versatilidade no design através de temas e com funções extensíveis através de módulos.

Além disso, este CMS realiza uma organização hierárquica do conteúdo do *site*, ou seja, as páginas do *site* são criadas na raiz ou abaixo de uma página pai. Uma página pode ter várias seções. A seção faz o vínculo da página com o conteúdo a ser exibido, tais como um texto, um vídeo, uma notícia. Desse modo, existem vários tipos de seções, sendo que cada tipo de seção é vinculada a um módulo, que faz o controle e armazena um conteúdo específico. Para esta pesquisa foi criado um módulo que realiza a função de catálogo facetado, exibindo a interface de navegação na seção da página em que o mesmo é vinculado.

G.1.5 Modelo de dados

O modelo do banco de dados relacional do módulo que realiza a função do catálogo facetado é apresentado na FIG. 98.

FIGURA 98 - Modelo de dados



Fonte: Elaborado pelo autor.

⁷⁷ Segundo o *site* da comunidade que é <<http://www.websitebaker.org>>.

A tabela *ontoadmin* é a responsável por guardar as configurações do catálogo, fazendo relação direta ou indireta com as demais tabelas. Foi utilizado um sistema de gestão de conteúdo, cuja entidade principal é a página; que é usada para representar tanto as categorias quanto os registros.

Para relacionar as categorias e os registros, utiliza-se como base o modelo RDF⁷⁸, que é uma recomendação da W3C⁷⁹ para descrição dos metadados de recursos da Internet. É considerado um dos pilares da *web* semântica. Sua unidade fundamental é "tripla", constituída de Sujeito-Predicado-Objeto (SPO).

Segundo Berners-Lee *et al.* (2001), o próprio criador do conceito da *web* semântica, pode-se definir a tripla como: "parecidas com uma sentença linguística elementar constituída de sujeito, verbo e objeto (...) isto é versátil o suficiente para descrever a maioria dos dados processados pelos sistemas (...). Com triplas em RDF é possível formar redes de informação sobre coisas relacionadas ". Em linhas gerais, a tripla é uma forma de conectar um sujeito, que é um recurso da Internet, a um objeto, que também pode ser outro recurso da Internet ou um dado, através de uma relação que é o predicado.

Como os registros e as categorias do catálogo são páginas no sistema de gestão de conteúdo e, conseqüentemente, são recursos da Internet com uma URL (*Uniform Resource Locator*), o modelo RDF se mostra adequado para registrar as relações entre eles. Isto também traz benefícios aos catálogos, tornando-os mais semânticos, facilitando por exemplo, a exportação dos dados do mesmo para outros sistemas, aumentando a interoperabilidade.

Dessa forma, inspirado no modelo RDF (*Resource Description Framework*) foi elaborada a tabela SPO (QUADRO 10). Foi criada a relação "é_subcategoria_de" (predicado), para relacionar uma subcategoria a uma categoria, e a relação "tem_classificacao" para relacionar um registro a uma categoria. Não foi criada uma tabela para os predicados, inserindo seu valor literal no atributo predicado da tabela SPO (modelo denormalizado), ao invés de utilizar os IDs do predicados (modelo normalizado). Segundo Wilkinson *et al.* (2003), este esquema, apesar de ocupar mais espaço, garante uma melhor performance no armazenamento de RDF em banco de dados relacionais. A maior contribuição do presente estudo nesse modelo foi a criação do atributo **Status**, designado para dizer se a tripla é válida ou não. Ou seja, o *status* é o valor de verdade⁸⁰ da tripla,

⁷⁸ Disponível em: <<http://www.w3.org/RDF>>.

⁷⁹ Disponível em: <<http://www.w3.org>>.

⁸⁰ Para saber mais: <http://pt.Wikipédia.org/wiki/Valor_de_verdade>.

determinado por meio do consenso entre os membros da comunidade de usuários do catálogo, conforme o modelo de colaboração *wiki*.

QUADRO 10 – Modelo de dados para o armazenamento das relações (triplas)

ID	Sujeito	Predicado	Objeto	Status
1	Categoria 2	e_subcategoria_de	Categoria 1	1
2	Registro 1	tem_categoria	Categoria 2	1

Fonte: Elaborado pelo autor.

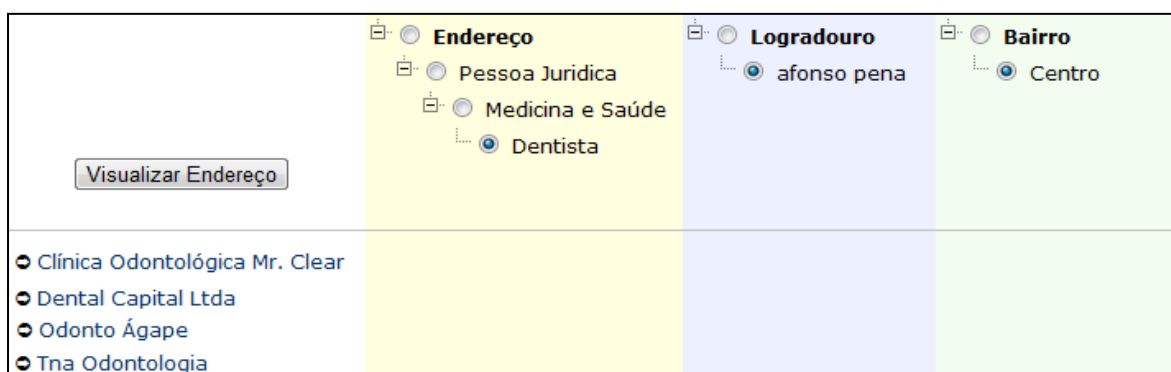
G.1.6 Implementação da interface facetada

Na seção 2.4.2.6, são apresentadas várias recomendações de design para a construção de uma de interface facetada, Dessa forma, o desenvolvimento foi focado na inserção de elementos de interface, a fim de prover melhores resultados na recuperação da informação. A seguir, é apresentado o primeiro protótipo do catálogo. Detalha-se nas próximas seções cada um dos requisitos inseridos.

G.1.6.1 A inserção de facetas coloridas

A interface facetada desenvolvida foi dividida em dois macro-painéis: um com as facetas (classes) e o outro com as instâncias. Cada faceta ocupa uma coluna no painel com as facetas. Assim, cada uma dessas colunas foi preenchida com uma cor para diferenciar o tipo da faceta (FIG. 99).

FIGURA 99 - inserção de facetas coloridas

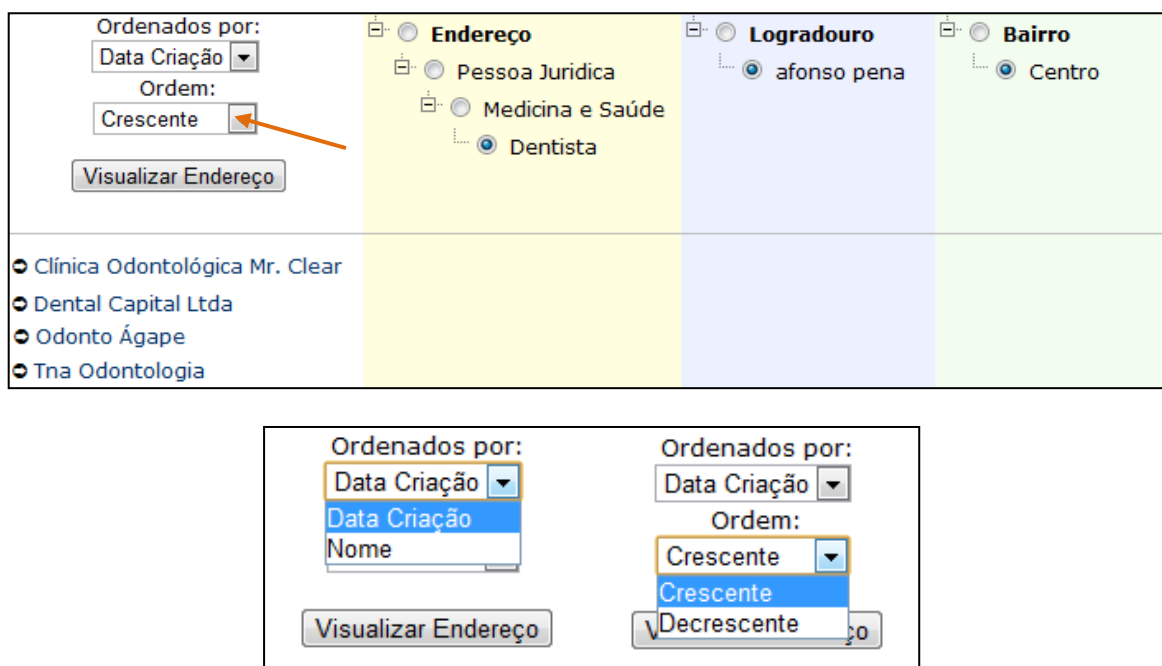


Fonte: Elaborado pelo autor.

G.1.6.2 Inserção de caixas de seleção para ordenação de resultados

Durante a utilização da interface, o usuário, ao procurar os registros, pode ter a necessidade de visualizá-los por data de criação e por ordenação alfabética. Percebeu-se que esta ordenação também pode ser crescente ou decrescente. Em um primeiro momento, foi feita uma caixa de seleção com todas as possibilidades: data de criação - crescente, data de criação – decrescente, nome – crescente, nome – decrescente. Mas, além da caixa de seleção ficar muito grande, se houver mais algum critério de ordenação, como por exemplo data de modificação, a quantidade de itens no controle iria ser multiplicada por dois, repetindo o crescente e o decrescente para todos os critérios. Assim, estas opções foram separadas em duas caixas de seleção distintas (FIG. 100).

FIGURA 100 - ordenação de resultados



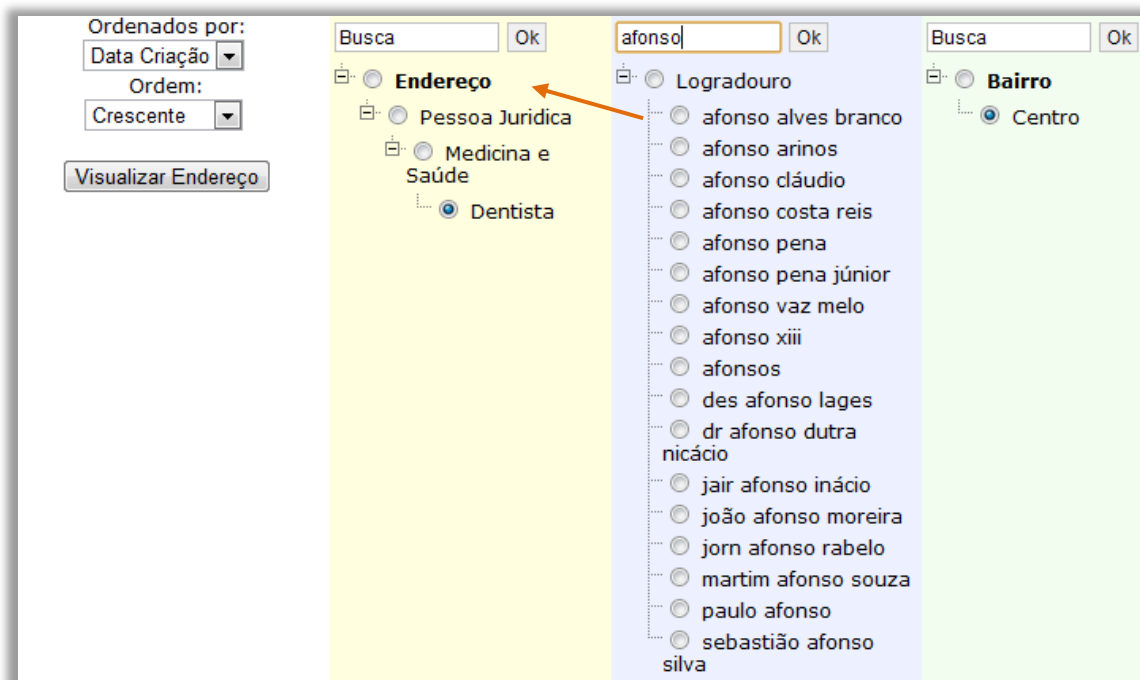
Fonte: Elaborado pelo autor.

G.1.6.3 Inserção da caixa de busca nas facetas

Após várias consultas no catálogo de endereços, percebe-se que em algumas facetas, o usuário já sabia a categoria que queria escolher; mas, algumas vezes, era lento procurar na hierarquia de cada faceta. Neste caso, quando o usuário sabe o nome da categoria desejada, a opção natural é a busca por palavra-chave. Assim, foi preciso

desenvolver uma maneira de combinar a busca por palavra-chave com a navegação dentro da faceta. Esta solução pode ser visualizada na FIG. 101.

FIGURA 101 - A inserção de caixa de busca nas facetas



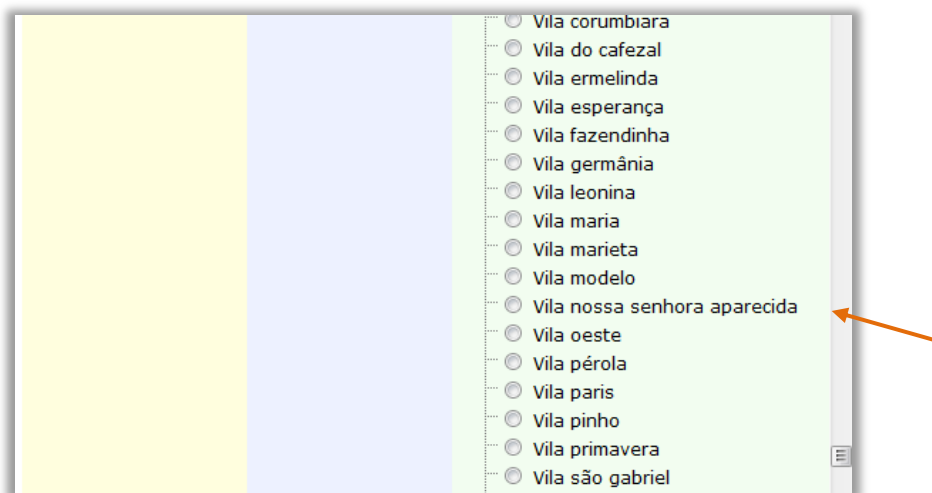
Fonte: Elaborado pelo autor.

G.1.6.4 Inserção de agrupamento alfabético em cadeias muito grandes

O catálogo de endereços foi uma base de dados interessante pela quantidade de registros que possui. Como são mais de vinte mil endereços organizados em centenas de categorias, representa de fato uma aplicação no mundo real.

O problema foi lidar com essa quantidade de categorias em uma faceta. Ao abrir uma lista com mais de cem categorias, a hierarquia de uma faceta ficava com um comprimento muito maior que as demais facetas. Tal fato fazia o usuário ter que “rolar” bastante a página procurando a categoria em uma faceta e escondendo as outras facetas com menos categorias (FIG. 102).

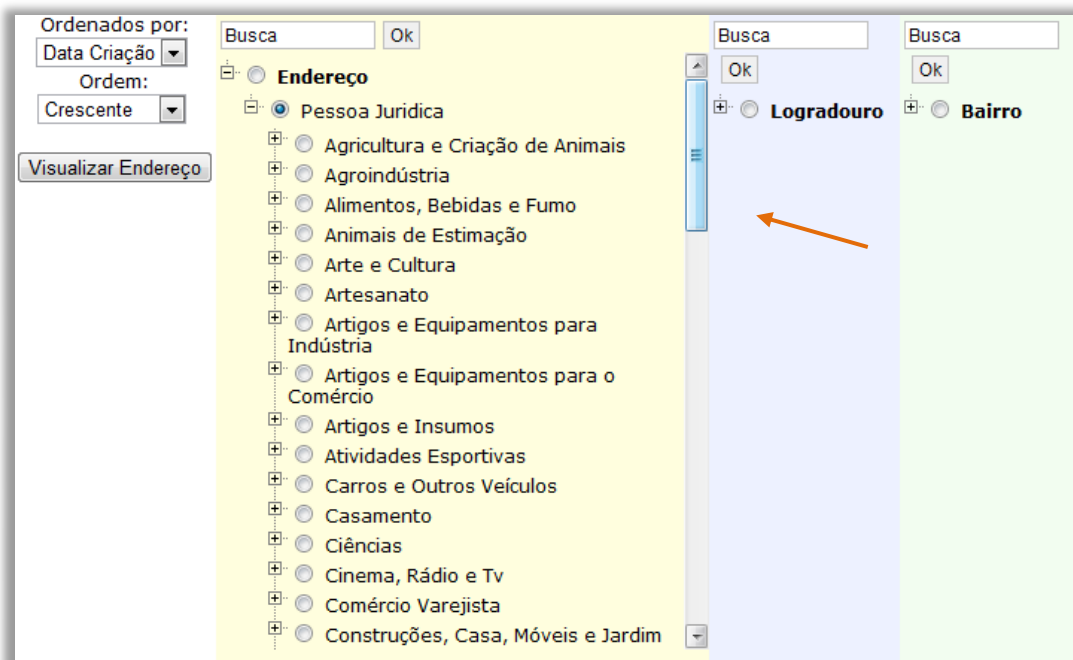
FIGURA 102 - Facetas sem barra de rolagem



Fonte: Elaborado pelo autor.

A primeira medida para contornar essa situação foi colocar uma barra de rolagem automática em cada uma das facetas. Esta barra de rolagem ocorre sempre que a lista atinge uma altura acima de certo tamanho. Tal fato permite continuar visualizando todas as facetas, independentemente se uma delas se expandiu demais (FIG. 103).

FIGURA 103 - Facetas com barra de rolagem automática



Fonte: Elaborado pelo autor.

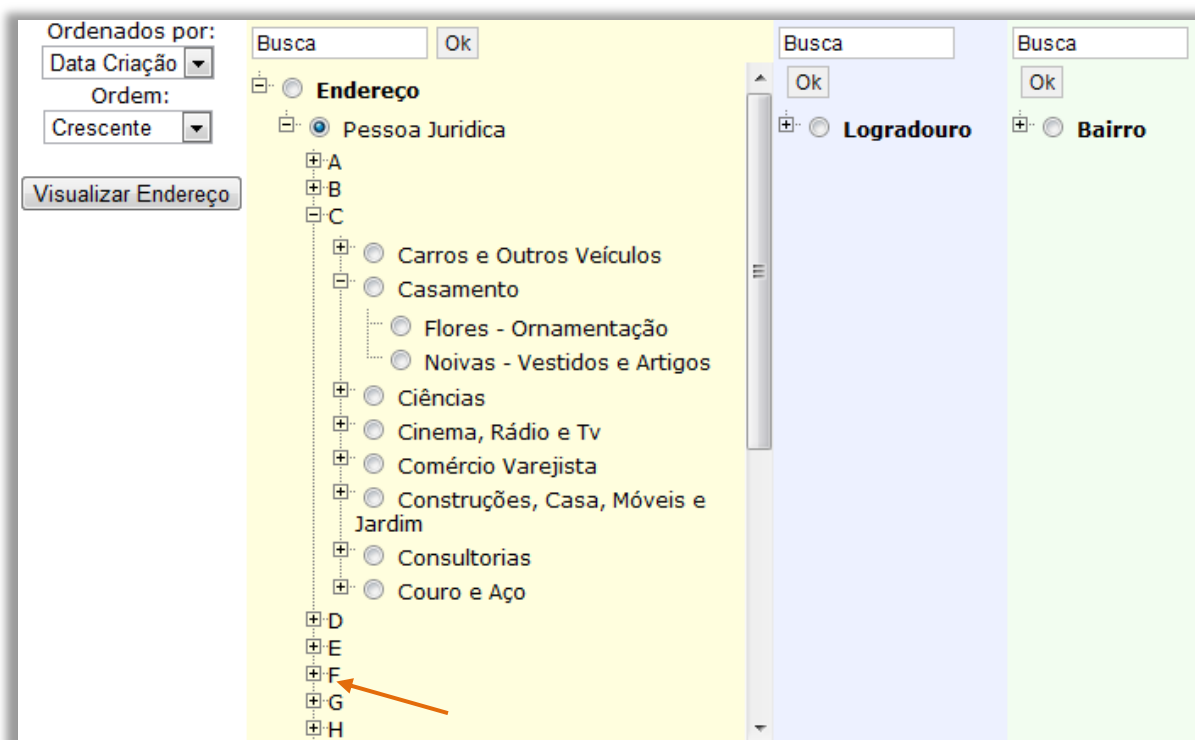
Embora esse recurso tenha melhorado muito o aproveitamento da área da interface, em algumas situações a procura em uma cadeia muito extensa é frustrante para o

usuário. Percebeu-se que mais de cem registros já tornaria essa tarefa incômoda. Para driblar esse problema, recorreu-se a um recurso tradicionalmente utilizado em suportes físicos, tais como dicionários e manuais, ou seja, o índice em ordem alfabética.

Assim, toda vez que a cadeia excede uma determinada quantidade de itens, agrupou-se os registros pelas letras do alfabeto. A quantidade máxima de itens para ativar o agrupamento alfabético pode ser configurável e o valor definido para essa ativação foi cem registros, que são mais de quatro vezes a quantidade que seria exibida quando é ativado o agrupamento (com o agrupamento são exibidos 24 itens correspondentes as letras do alfabeto).

Na FIG. 104 pode ser visto como ficou otimizado o espaço de exibição da faceta endereço, mantendo a visualização das outras facetas. A desvantagem é que acrescenta mais um passo na navegação.

FIGURA 104 - Facetas com agrupamento pelas letras do alfabeto

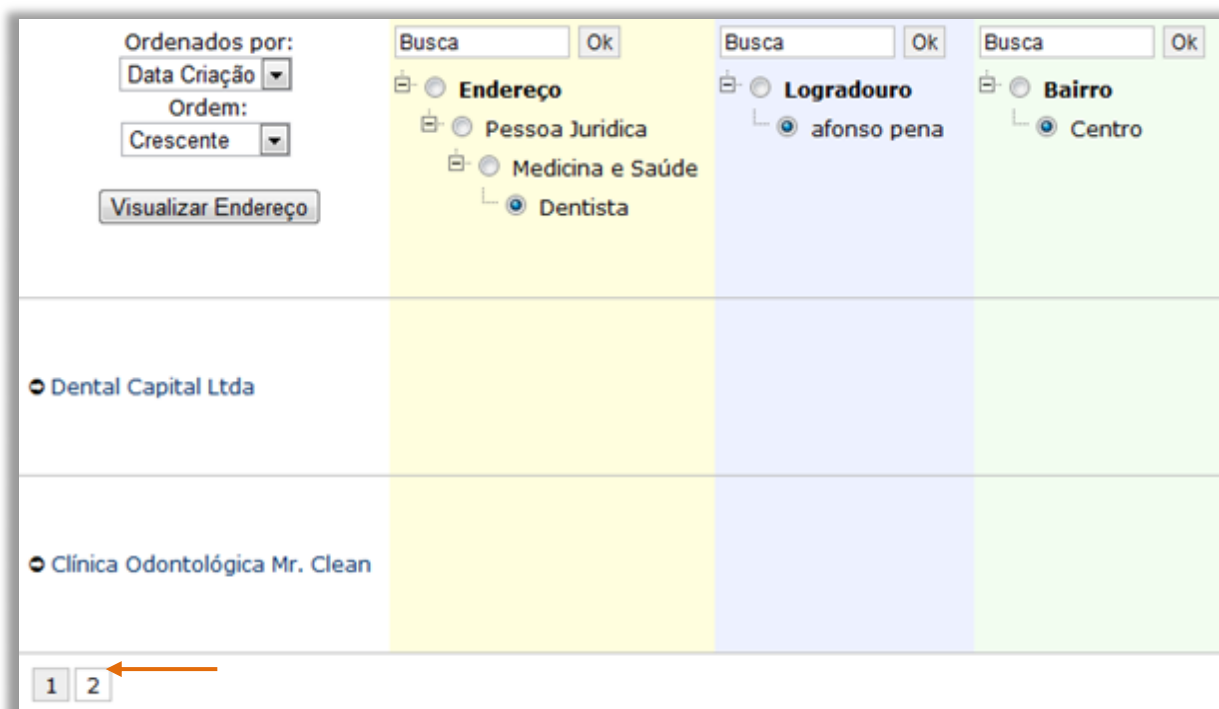


Fonte: Elaborado pelo autor.

G.1.6.5 Inserção de paginação dos registros

Apesar do número de categorias ser grande (centenas de registros), o número de registros é bem maior (dezenas de milhares). Tal fato faz com que algumas buscas retornem uma quantidade muito grande de registros. A forma tradicional para tratar esse tipo de caso é por meio da paginação de registros. Sendo assim, foi incluída uma paginação que também é configurável na área de administração do sistema, conforme pode ser visualizado na FIG. 105 (para fins de ilustração foram configurados dois registros por página na administração).

FIGURA 105 - Paginação dos indivíduos

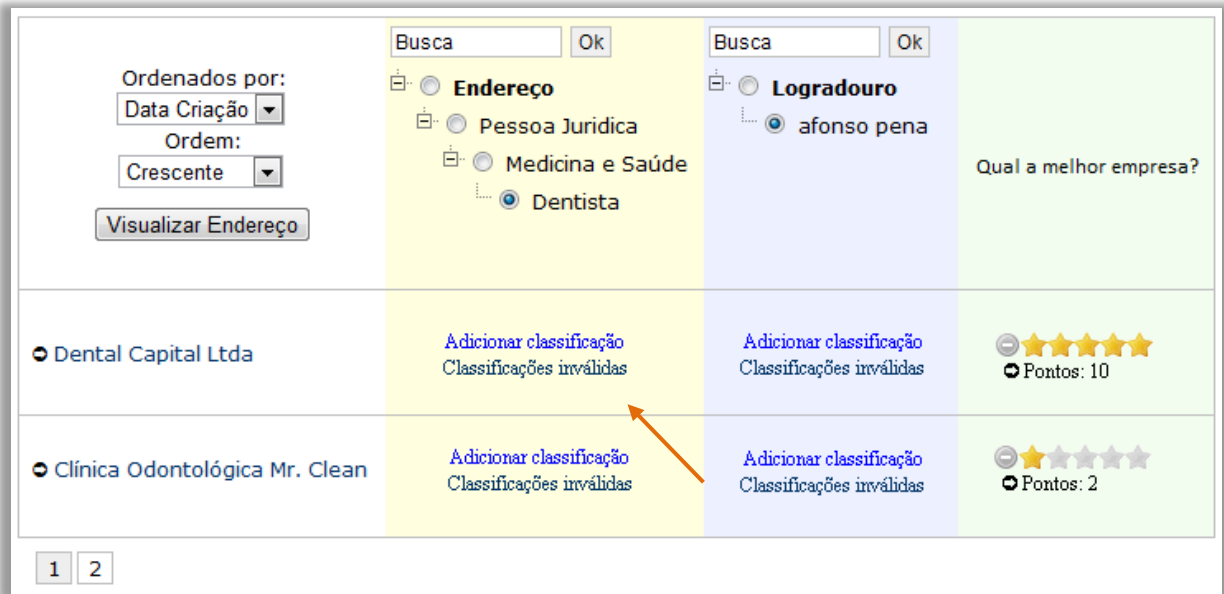


Fonte: Elaborado pelo autor.

G.1.6.6 Implementação do modelo de colaboração

O modelo de colaboração foi implementado com a inserção dos *links* para adicionar uma classificação a um registro, bem como ver classificações inválidas, o que permitia ao usuário verificar as classificações que tinham sido invalidadas e revalidá-las. Além disso, é possível atribuir uma nota ao registro referente ao critério da votação (FIG. 106).

FIGURA 106 - Ações colaborativas



Fonte: Elaborado pelo autor.

G.1.7 Teste: registro de falhas e sucessos

Após o desenvolvimento do primeiro protótipo, foi realizado uma bateria de testes funcionais baseados nos requisitos. O primeiro problema detectado foi que a navegação facetada não estava trabalhando com a eficiência que era esperada. Ao combinar as facetas, a interface recuperava apenas os indivíduos que pertenciam exatamente àquelas categorias selecionadas.

Tal fato não trazia todos os resultados esperados. Para melhorar sua revocação era preciso que a mesma trabalhasse de forma recursiva (navegação recursiva), ou seja, ao combinar as categorias de cada faceta, deviam ser recuperados os registros que pertencessem aquela categoria ou a qualquer subcategoria daquela faceta. No exemplo do catálogo de endereços, se o usuário estivesse procurando por uma empresa, e se selecionasse a categoria Medicina e Saúde deveriam retornar todas as empresas que também estivessem classificadas com a categoria Dentista.

Um incômodo é que o usuário precisava escolher as categorias depois clicar em Visualizar para exibir os registros abaixo. Percebe-se que seria mais natural clicar em uma categoria, selecionando-a, e imediatamente já filtrar os registros abaixo. Assim, a medida que o usuário fosse selecionando as categorias, a quantidade de registros recuperados iria diminuindo progressivamente.

Outro problema é que ao exibir 10 ou 20 registros na listagem, as ações colaborativas começavam a poluir muito a interface de navegação, com a exibição de dezenas de *links*. Não existia espaço na interface para o usuário colaborar com as categorias das facetes. Diante de tais problemas surgiu o escopo do segundo ciclo de desenvolvimento.

G.2 Segundo ciclo de desenvolvimento

G.2.1 Escopo da versão

A partir dos testes realizados no ciclo de desenvolvimento anterior ficou evidente a necessidade de melhorar a usabilidade da interface facetada, e também realizar a ampliação do modelo de colaboração *wiki* para as categorias das facetes.

G.2.2 Requisitos

Os requisitos da versão anterior foram mantidos com a inclusão dos seguintes requisitos:

Referente à navegação facetada:

- Os registros vinculados às subcategorias da categoria selecionada também devem ser exibidos de forma recursiva;
- Remover o botão visualizar. Ao clicar em uma categoria, deve filtrar imediatamente, trazendo os registros abaixo;
- Determinar a quantidade de registros exibidos por página sendo o *default* igual a 10.

Referente ao modelo de colaboração *wiki* (apenas para usuário logado e membro da comunidade de usuários):

- As ações colaborativas somente apareceriam ao passar o *mouse* sobre o elemento. Isto deixaria a interface mais limpa, sem a repetição das mesmas opções para cada um dos registros, e permitiria, também, colocar essas opções disponíveis para as categorias das facetes;
- Adicionar nova categoria à uma faceta;
- Invalidar a categoria da faceta;

- Validar uma categoria inválida da faceta;
 - Para executar essa ação é necessário ter uma opção para mostrar as categorias inválidas, como se fosse a visualização de itens na lixeira, pois senão não é possível saber o que foi invalidado para reverter a ação;
- Mostrar classificações inválidas;
- São considerados válidas as categorias e classificações, cujo somatório de votos válidos seja maior ou igual ao votos inválidos. Nota: optou-se por considerar válidos os registros de votos válidos iguais aos inválidos (somatório igual a zero), dando preferência aos usuários que alimentam o catálogo.

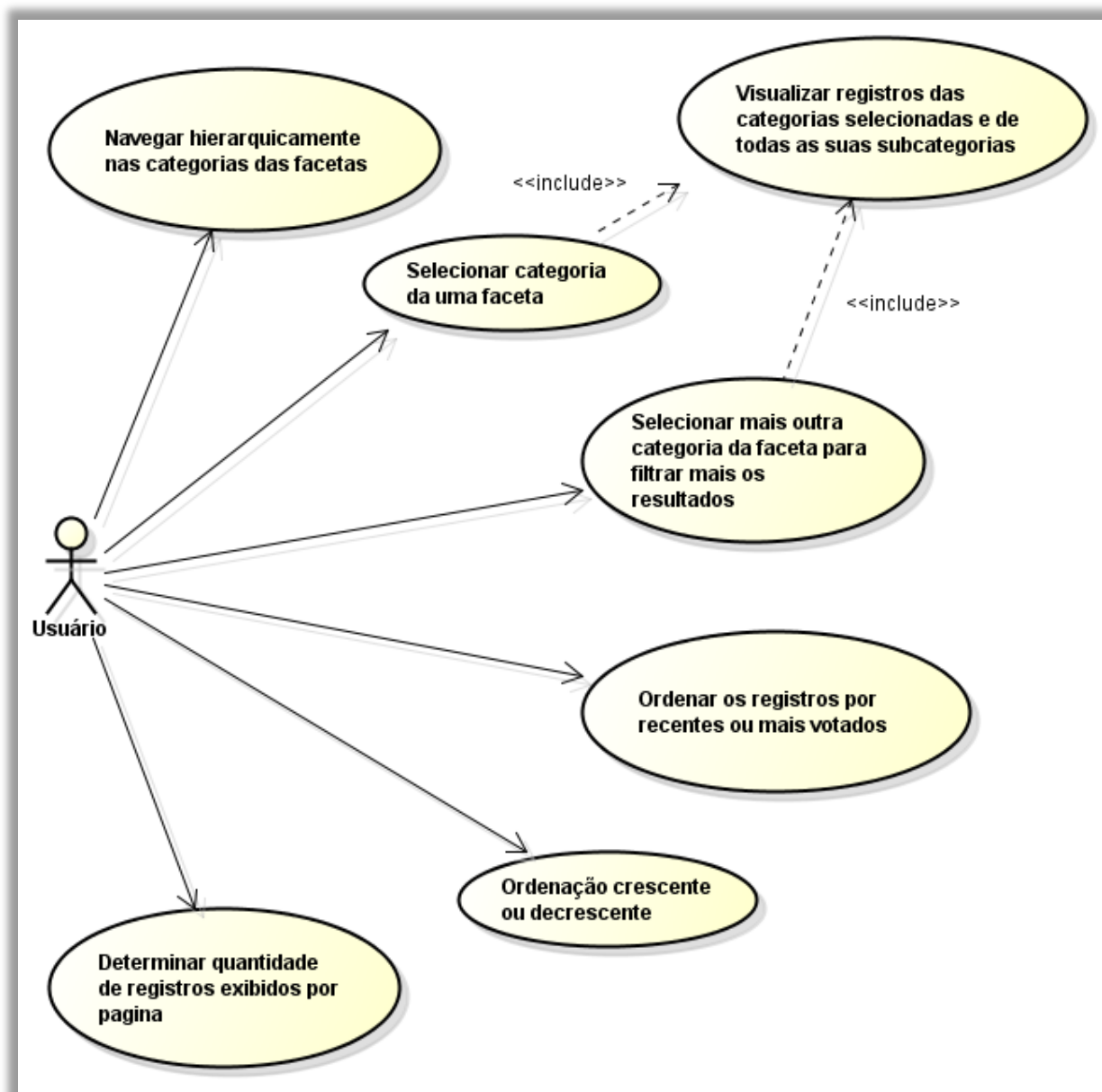
G.2.3 Diagramas UML

Nesta seção são apresentados os diagramas de caso de uso e de atividade realizados para orientar a implementação nesse ciclo de desenvolvimento.

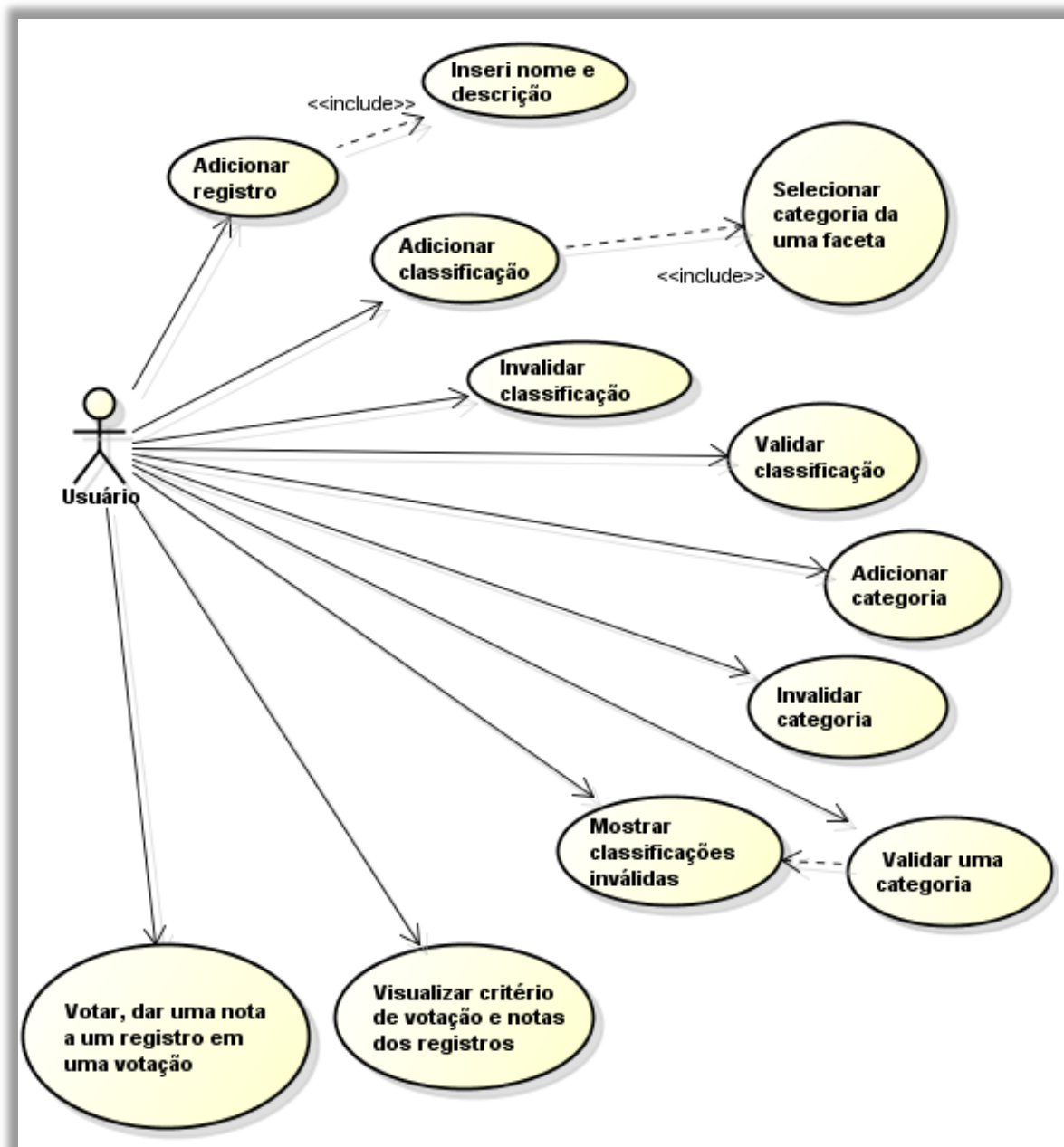
G.2.3.1 Diagramas de caso de uso

Os diagramas de caso de uso atualizados para o escopo deste ciclo de desenvolvimento são apresentados a seguir. Percebe-se que eles foram complementados a partir dos testes realizados na versão anterior (FIG. 107 e FIG. 108).

FIGURA 107 - Realizar navegação facetada



Fonte: Elaborado pelo autor.

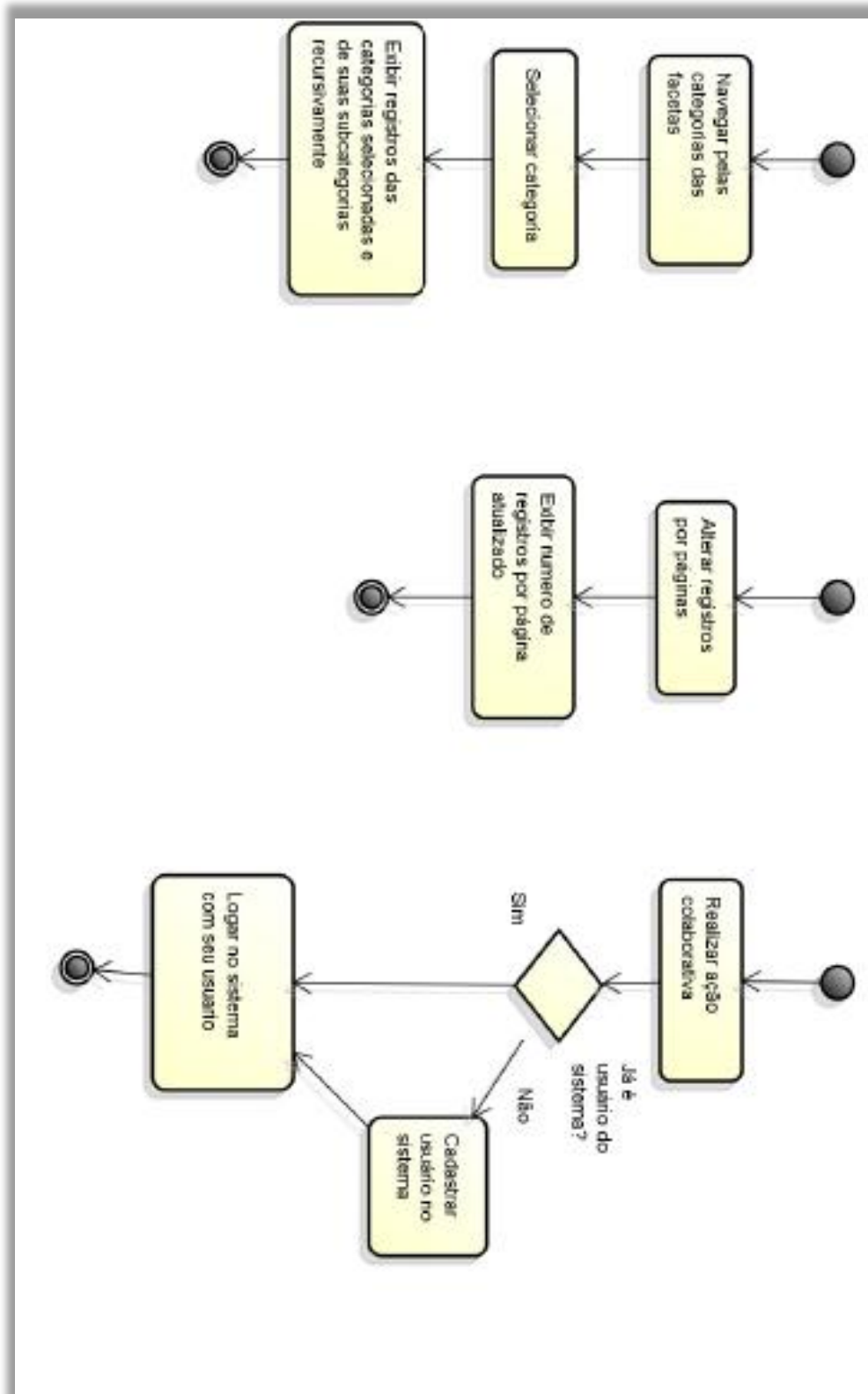
FIGURA 108 - Modelo de colaboração *wiki*

Fonte: Elaborado pelo autor.

G.2.3.2 Diagramas de atividade

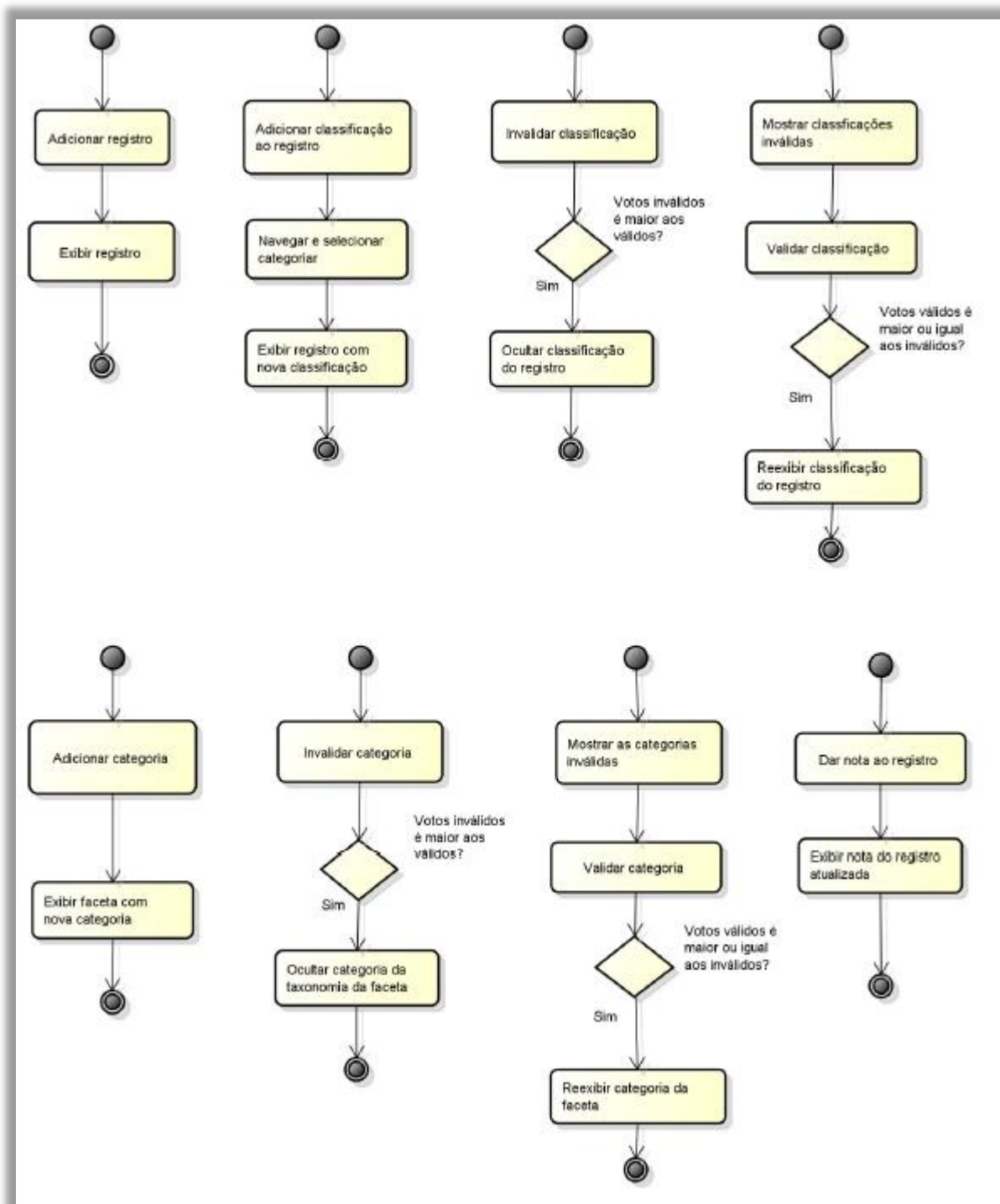
São apresentados a seguir os diagramas de atividade do catálogo, divididos em: navegação facetada (FIG. 109) e modelo de colaboração *wiki* (FIG. 110).

FIGURA 109 - Navegação facetada



Fonte: Elaborado pelo autor.

FIGURA 110 - Modelo de colaboração *wiki*



Fonte: Elaborado pelo autor.

G.2.4 Modelo de dados

No momento da implementação do modo recursivo da navegação facetada, percebeu-se, imediatamente, um problema referente a performance ou eficiência do sistema na recuperação dos registros. Se uma faceta, por exemplo, Endereço, tem centenas de subcategorias, para cada categoria selecionada o sistema precisa percorrer, de forma recursiva, todos os nós da árvore; o que demanda um alto custo computacional, tornando lenta a resposta do sistema.

Para que o modo recursivo opere com eficiência, verificou-se que era necessário uma alteração no modelo de dados. A alteração é ilustrada no QUADRO 11. Foi criado o campo "Filhos", que inclui todas as categorias descendentes (filhas, netas e bisnetas). Tal fato facilitava a consulta, pois ao selecionar uma categoria, por exemplo, Cat2, era possível saber os registros relacionados às suas subcategorias; bastava, então, fazer uma segunda consulta SQL usando as categorias Cat3 e Cat4, recuperando, assim, os registros B,C e D.

Assim, com duas consultas SQL era possível recuperar todos os registros relacionados com aquela categoria ou suas subcategorias, sem a necessidade de realizar um algoritmo recursivo. Após obter a lista dos registros de cada categoria selecionada, bastava fazer uma operação booleana *AND*, selecionando aqueles registros que eram comum aos vetores (*arrays*).

QUADRO 11 – Modelo de dados com as categorias filhas

ID	Sujeito	Predicado	Objeto	Status	Filhos
1	Cat2	e_subcategoria_de	Cat1	1	Cat3, Cat4
2	Cat3	e_subcategoria_de	Cat2	1	Cat4
3	Cat4	e_subcategoria_de	Cat3	1	
4	Registro A	tem_categoria	Cat1	1	
5	Registro B	tem_categoria	Cat2	1	
6	Registro C	tem_categoria	Cat3	1	
6	Registro D	tem_categoria	Cat4	1	

Fonte: Elaborado pelo autor.

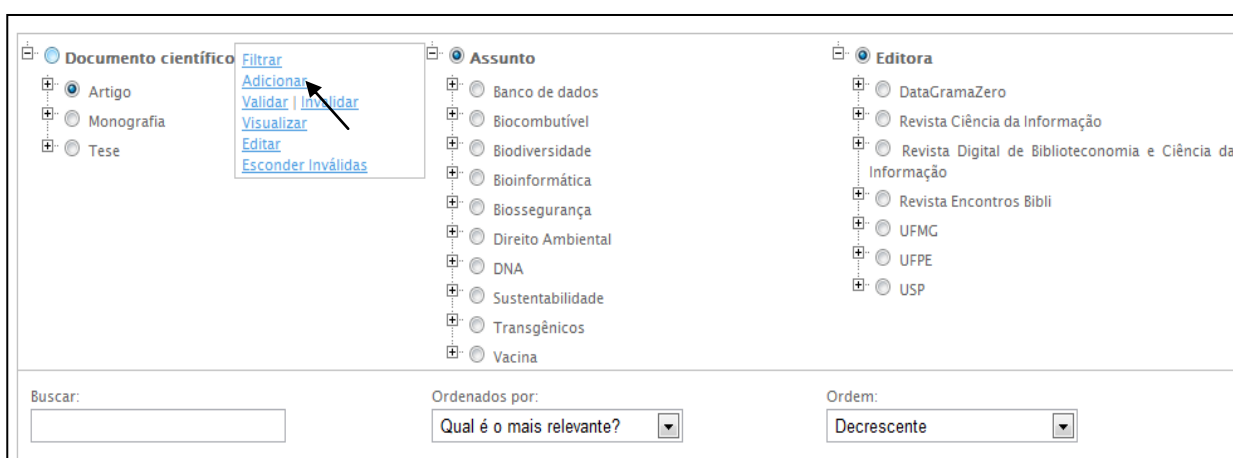
G.2.5 Implementação da interface

Com a reformulação da interface foi removido o botão Visualizar e os registros eram filtrados clicando diretamente na categoria de uma faceta. Neste ciclo, o principal

esforço passou da realização de uma interface facetada para a implementação do modelo de colaboração *wiki*, com a introdução de um Menu de Ações que aparece ao deixar o *mouse* sobre uma categoria por alguns instantes.

Este menu, que visava atender aos requisitos dessa versão, proporcionou um enorme ganho de espaço na interface. Os usuários poderiam contribuir na construção da taxonomia de assuntos inserindo novas categorias, validando ou invalidando a categoria, exibindo ou ocultando as categorias inválidas. A FIG. 111 mostra um usuário adicionando uma nova categoria.

FIGURA 111 - Colaboração na taxonomia da faceta

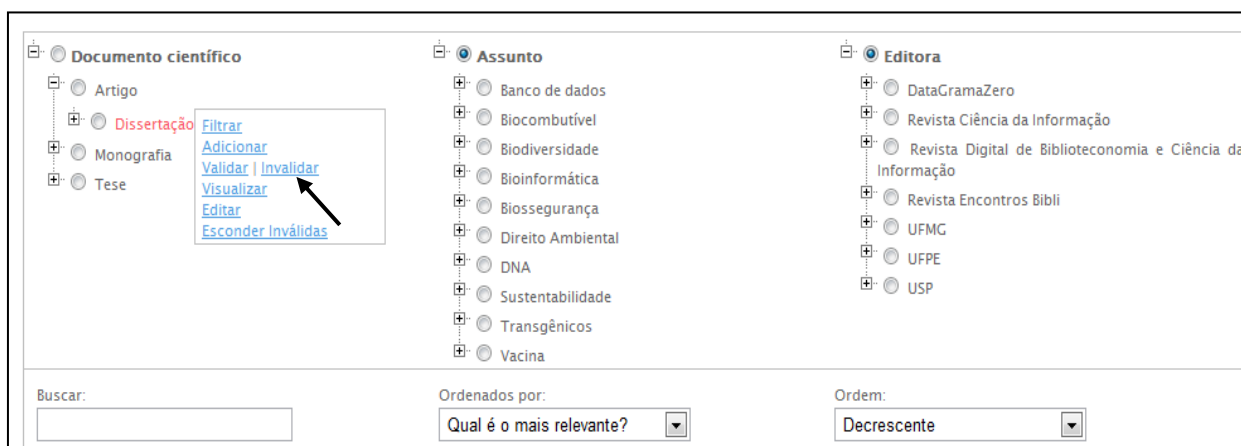


Fonte: Elaborado pelo autor.

Com a inserção de novas categorias por usuários, ao seguir o modelo de colaboração *wiki*, deve ser possível remover contribuições inseridas erroneamente. Ao invalidar, as categorias inválidas eram caracterizadas em vermelho.

As categorias inválidas não são apagadas, pois é essencial manter um histórico de contribuições dos usuários. Elas são apenas ocultadas. Dessa forma, para serem visualizadas é necessário clicar em “Mostrar Inválidas” no Menu de Ações (FIG. 112).

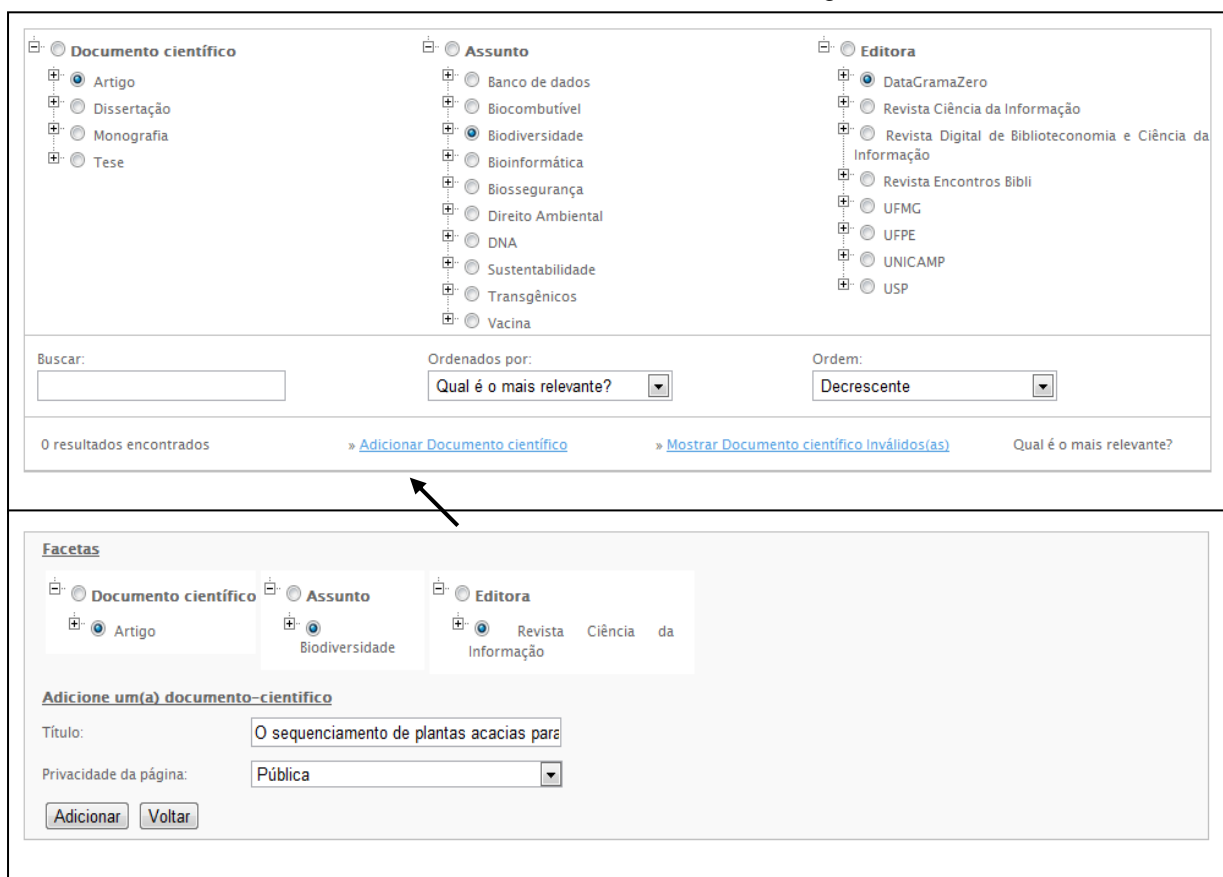
FIGURA 112 - Invalidar categorias



Fonte: Elaborado pelo autor.

Como esperado, os usuários poderiam alimentar livremente o catálogo, adicionando novos registros. Ao fazerem tal ação, realizavam a classificação facetada do registro, conforme pode ser visualizado na FIG. 113.

FIGURA 113 - Alimentando o catálogo



Fonte: Elaborado pelo autor.

Nesta nova versão, os registros também exibiam um Menu de Ações. Dessa forma, era possível invalidar a classificação de um registro. Ao realizar tal ação, a categoria é diferenciada em vermelho (FIG. 114).

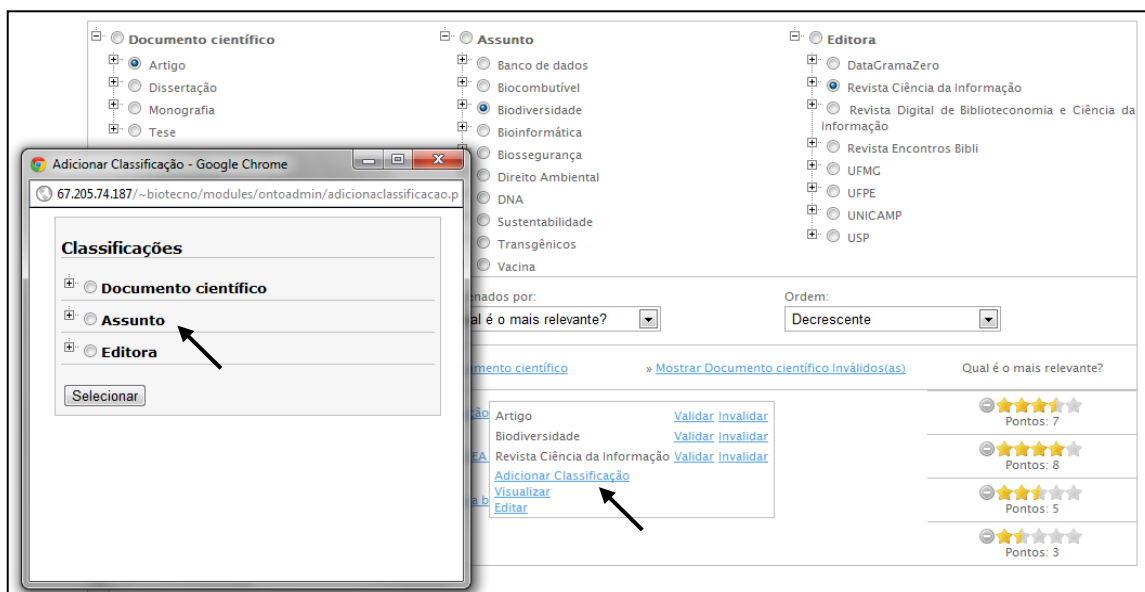
FIGURA 114 - Usuários podem validar/invalidar o material inserido



Fonte: Elaborado pelo autor.

O usuário também pode adicionar uma nova classificação a um registro. Assim, pode contribuir na manutenção do catálogo ao melhorar a relevância dos registros recuperados durante a navegação (FIG. 115).

FIGURA 115 - Adicionar nova classificação



Fonte: Elaborado pelo autor.

O usuário pode ordenar pela data de criação do registro ou pela votação e, também, de forma crescente ou decrescente, os registros conforme ilustrado na FIG. 116.

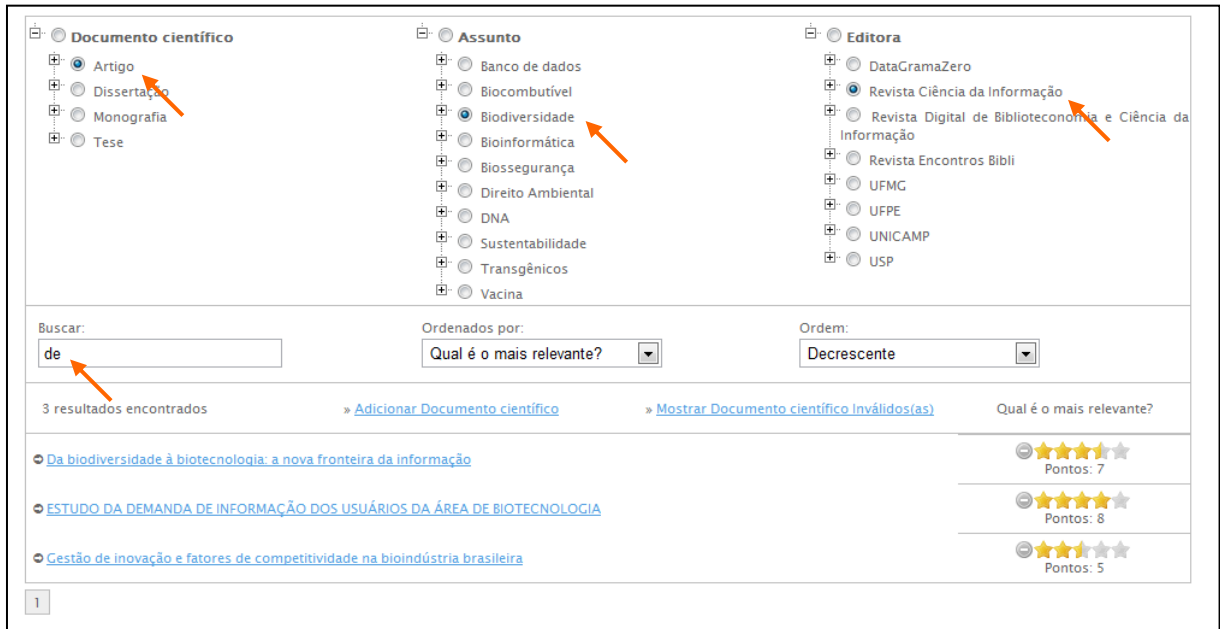
FIGURA 116 - Ordenação dos registros recuperados



Fonte: Elaborado pelo autor.

Para complementar a navegação facetada, o usuário pode combiná-la com a busca por palavra-chave (FIG. 117).

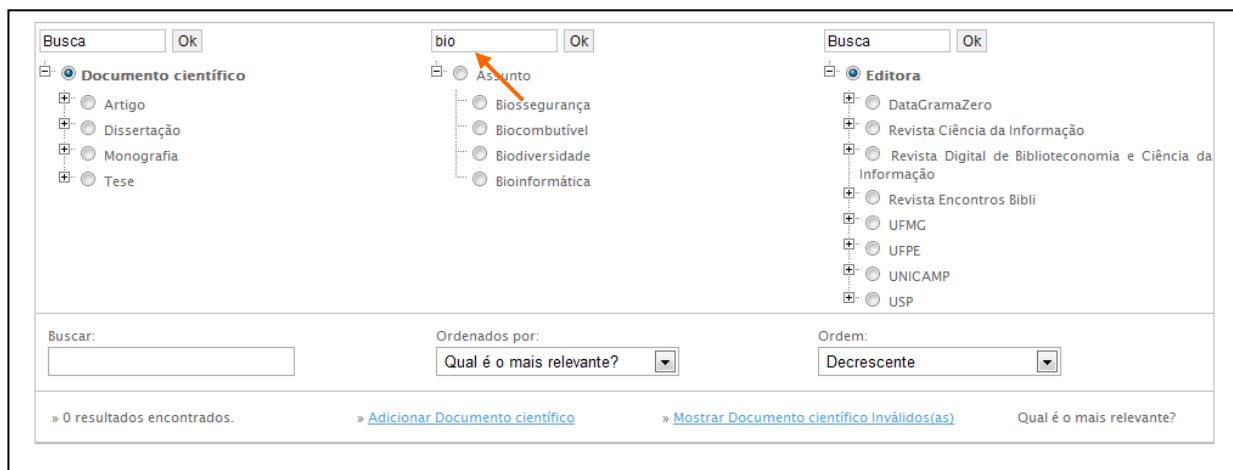
FIGURA 117 - Busca por palavra-chave nos registros



Fonte: Elaborado pelo autor.

O usuário também pode filtrar as categorias da faceta através do campo de busca (FIG. 118).

FIGURA 118 - Busca por palavra-chave nas facetas



Fonte: Elaborado pelo autor.

G.2.6 Teste: registro de falhas e sucessos

Nesta versão do catálogo facetado foram realizados testes de unidade do modelo de colaboração funcionando em conjunto com a navegação facetada. A partir dessa versão obteve-se a verdadeira dimensão do tamanho do problema enfrentado, cujos erros estão listados abaixo:

- A taxonomia das facetas mostrada no Adicionar Classificação não está mostrando a partir do terceiro nível. Deveria carregar nível após nível dinamicamente;
- Ao buscar por nada na caixa de pesquisa da faceta, as categorias somem e não tem mais como exibir todas;
- Ao tentar inserir uma categoria ou registro sem estar logado, exibe um erro no qual a imagem para fechar da caixa exibida, não está sendo encontrada;
- A navegação recursiva deveria ser uma opção do usuário ao navegar e não uma opção na administração, ou seja, poderia ter uma caixa de marcação (*check-button*) com a possibilidade de habilitar ou não a navegação recursiva;
- Ao clicar para adicionar uma categoria, a árvore fica sempre contraída, não expande;
- Quando vários catálogos compartilham a mesma classe raiz, o total de resultados exibido está somando tudo (fica diferente do que é mostrado);
- Ao clicar em esconder inválidas volta-se para a raiz do módulo, perdendo toda a sequência da navegação;
- Ao clicar em esconder/mostrar instâncias invalidadas e/ou validadas, não funciona;
- Validar/invalidar no catálogo recursivo não está fazendo efeito nenhum, continua exibindo sempre;
- Não deixa o próprio usuário revalidar, tanto nas categorias quanto nos registros, ou seja, o usuário deveria poder voltar atrás no seu voto. Poderia validar ou invalidar a qualquer momento. O empate (somatório igual a zero) deixaria válido. Depois do usuário validar, apenas a opção invalidar deveria ficar disponível (e vice-versa);
- A ordenação não está funcionando;

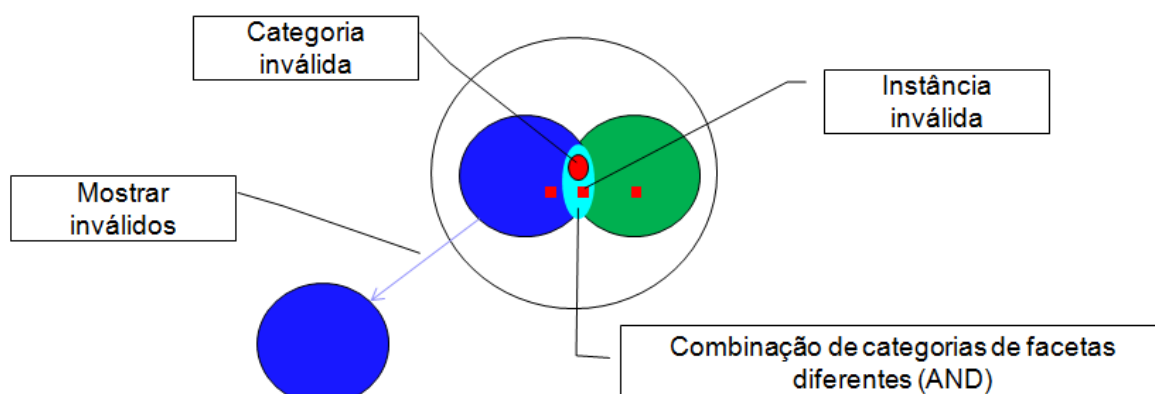
- Ao adicionar uma subcategoria em uma faceta, a verificação está sendo feita de forma errada, por exemplo, se tenho a faceta carro e a subcategoria carro esportivo e insiro, sem querer, carro utilitário dentro de esportivo, terei o *link* /carro/carro-esporte/carro-utilitario. Assim, se eu quiser inserir /carro/carro-utilitario deveria deixar. Dessa forma, o teste de validação de existência da categoria deve ser feito através do *link* completo e não do título da página, pois apenas com o título da página não é possível tratar a polissemia, ex: /fruta/manga /camisa/manga. Do jeito atual, apenas um tipo de manga poderia ser inserido;
- Ao remover um registro que já se encontra na lixeira, está se removendo o mesmo da tabela *pages*, mas não se remove as facetas na tabela *ontoadmin_facetas* e nem as triplas na tabela SPO;
- No cadastro de um registro, as categorias deveriam ser mostradas expandidas, pois quando o usuário escolhe uma categoria na página anterior parece que não tem nada que foi escolhido;
- Na administração, ao tentar alterar o tipo de registro criado era exibida uma tela branca e não se alterava;
- Ao clicar em cima da categoria, não está filtrando correto. É preciso clicar duas vezes para o filtro ocorrer;
- Na navegação recursiva, ao cadastrar um item como Carro Esporte e Marca Honda, e ao filtrar por Carro Utilitário, o item continua a ser exibido
- No catálogo em "Relação entre as categorias da faceta" está faltando a opção "Nova relação" e só aparece duas opções. Deveria ter as mesmas opções da caixa de seleção "Relação dos indivíduos com as facetas";
- Deixar, por padrão, o primeiro nível das facetas do catálogo expandido (criar opção na administração para configurar isso), simulando o usuário clicando para abrir a classe raiz.

Grande parte dos problemas sempre envolvia o validar e invalidar juntamente com a navegação facetada que era recursiva (FIG. 119). Seguem alguns erros encontrados:

- O usuário invalidava um vínculo de um registro com uma categoria (classificação). Ao clicar nesta categoria, o registro não aparecia mais, era ocultado. Entretanto, ao clicar na categoria pai dessa categoria inválida, o registro tornava a aparecer, pois a navegação era recursiva e não era verificado se as relações das subcategorias estavam válidas;

- O usuário invalidava uma categoria. Os registros vinculados às subcategorias (que continuavam válidas) continuavam a aparecer;
- Ao invalidar ou validar uma categoria, o sistema atualizava a interface para mostrar as categorias válidas. Isto fazia perder o *status* da navegação facetada. Por exemplo, se o usuário estivesse com várias categorias selecionadas, ao realizar essa ação começava tudo do início, com todas as facetas, sem nenhuma categoria selecionada, o que era muito frustrante;
- Mostrar as categorias inválidas e depois voltar a mostrar apenas a categorias válidas retornava resultados inconsistentes;
- Outro problema ocorria ao pedir para mostrar os registros inválidos. Isto fazia a lista de registros relacionados a uma categoria aumentar, exibindo registros inválidos em vermelho. Entretanto, quando realizada a paginação ou ordenação voltava a exibir apenas os registros válidos;
- Não estava claro o que devia ser feito ao mostrar categorias inválidas. Esta opção deveria mostrar as subcategorias do nível abaixo inválidas, mostrar todas as subcategorias inválidas ou mostrar todas as subcategorias e registros inválidos ?

FIGURA 119 - Consultas complexas envolvendo categorias, classificações ou registros inválidos



Fonte: Elaborado pelo autor.

G.2.6.1 Usabilidade

A ordenação dos resultados por meio de caixas de seleção não era muito fácil de usar. Não eram todos que entendiam que poderia ser combinado o critério de ordenação (mais recentes e mais votadas) com o tipo de ordenação. É mais comum ver esta função associada a uma coluna na qual seu título tem uma seta que indica a ordenação.

O botão para cadastrar novos registros não estava em destaque na interface. é O mesmo é bastante importante, pois é a porta de entrada da colaboração do usuário, assim era necessário dar maior ênfase ao mesmo.

O menu de ações associado à categoria e ao registro mostrava sempre as opções validar e invalidar. Não era muito consistente mostrar ambas as ações. A ação de invalidar deveria ser exibida quando o registro ou categoria estivesse válido e vice-versa.

G.2.6.2 Performance

Para uma coleção acima de 10.000 registros, o tempo de resposta da interface era muito lento, e o pior, o custo computacional estava principalmente no ato ao entrar no catálogo, pois era executado o algoritmo recursivo em todas as facetas; isso porque as categorias fundamentais de cada faceta estavam selecionadas por padrão, ou seja, como não tinha nenhum filtro, assumia-se mostrar todos os resultados.

3 Terceiro ciclo de desenvolvimento

G.3.1 Escopo da versão

O escopo deste ciclo de desenvolvimento tinha o seguinte objetivo: aumentar a usabilidade e o poder da interface facetada; tratar as inconsistências do modelo de colaboração, atuando juntamente com a navegação facetada; e, aumentar a performance do catálogo.

G.3.2 Requisitos

A implementação anterior revelou a necessidade dos seguintes requisitos:

- Ao invalidar uma categoria, ela deve ser ocultada da taxonomia;
- Ao visualizar as subcategorias inválidas de uma categoria, elas devem aparecer em vermelho;
- Ao invalidar uma classificação, ela deve ser ocultada da lista de categorias que aquele registro possui;
- Ao visualizar as classificações inválidas de um registro, elas devem aparecer em vermelho.

Além desses requisitos já implementados, baseados nos testes de unidade, este ciclo de desenvolvimento exige:

Referente à navegação facetada:

- Disponibilizar as facetadas na posição vertical, como as pastas de um sistema de arquivos para ganhar espaço na interface;
- Disponibilizar a data de criação, votação e qualquer critério de ordenação em colunas, com a ordenação crescente ou decrescente no título da coluna;
- Incluir as operações booleanas NOT, AND e seleção não recursiva (extensão rasa);
- Evidenciar as categorias selecionadas.

Referente ao modelo de colaboração *wiki*:

- Deixar mais visível a ação de adicionar um registro;
- Invalidar um registro do catálogo;
- Validar um registro inválido;

- Mostrar registros inválidos;
- Ao invalidar uma categoria, invalidar também todas subcategorias e classificações de registros. Assim, se uma categoria superior fosse selecionada, os registros vinculados à todas as subcategorias inválidas deveriam ser ocultados na listagem de resultados durante a navegação facetada;
- Se a categoria ou registro estivesse válido, o menu de ações deveria mostrar a opção de invalidar. Se estivesse inválido, deveria aparecer a opção de validar;
- Ao clicar em mostrar todos, deve ser possível visualizar todas as subcategorias e registros inválidos daquela categoria (o *status* era ignorado). Dessa forma, deveria ser possível validar uma categoria inválida;
- Ao realizar a paginação, por exemplo, clicando na segunda página de resultados, manter o *status*, mostrando apenas os registros válidos ou todos os registros. Idem para ordenação;
- Ao adicionar uma nova categoria que tem o mesmo nome de uma categoria já existente, exibir um aviso alertando e mostrando onde a categoria está criada. Se for uma categoria inválida o usuário pode validá-la;
- Ao validar ou invalidar uma categoria manter o *status* de navegação.

Referente ao sistema de votação:

- Permitir múltiplas votações em um catálogo. Um conjunto de registros deve poder ser avaliado sobre diversos critérios.

G.3.3 Diagramas UML

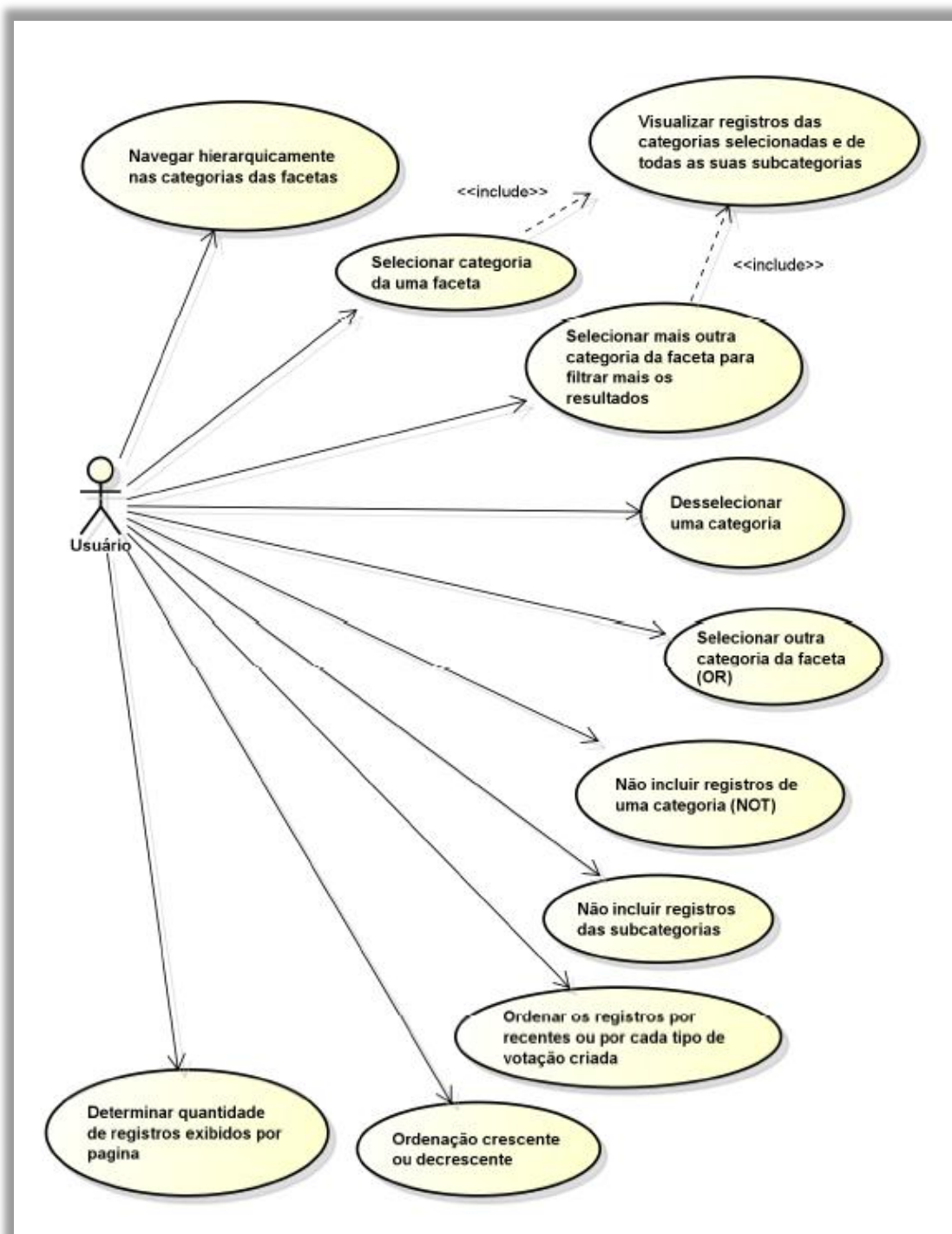
Nesta seção são apresentados os diagramas de caso de uso e de atividade realizados para orientar a implementação nesse ciclo de desenvolvimento.

G.3.3.1 Diagramas de caso de uso

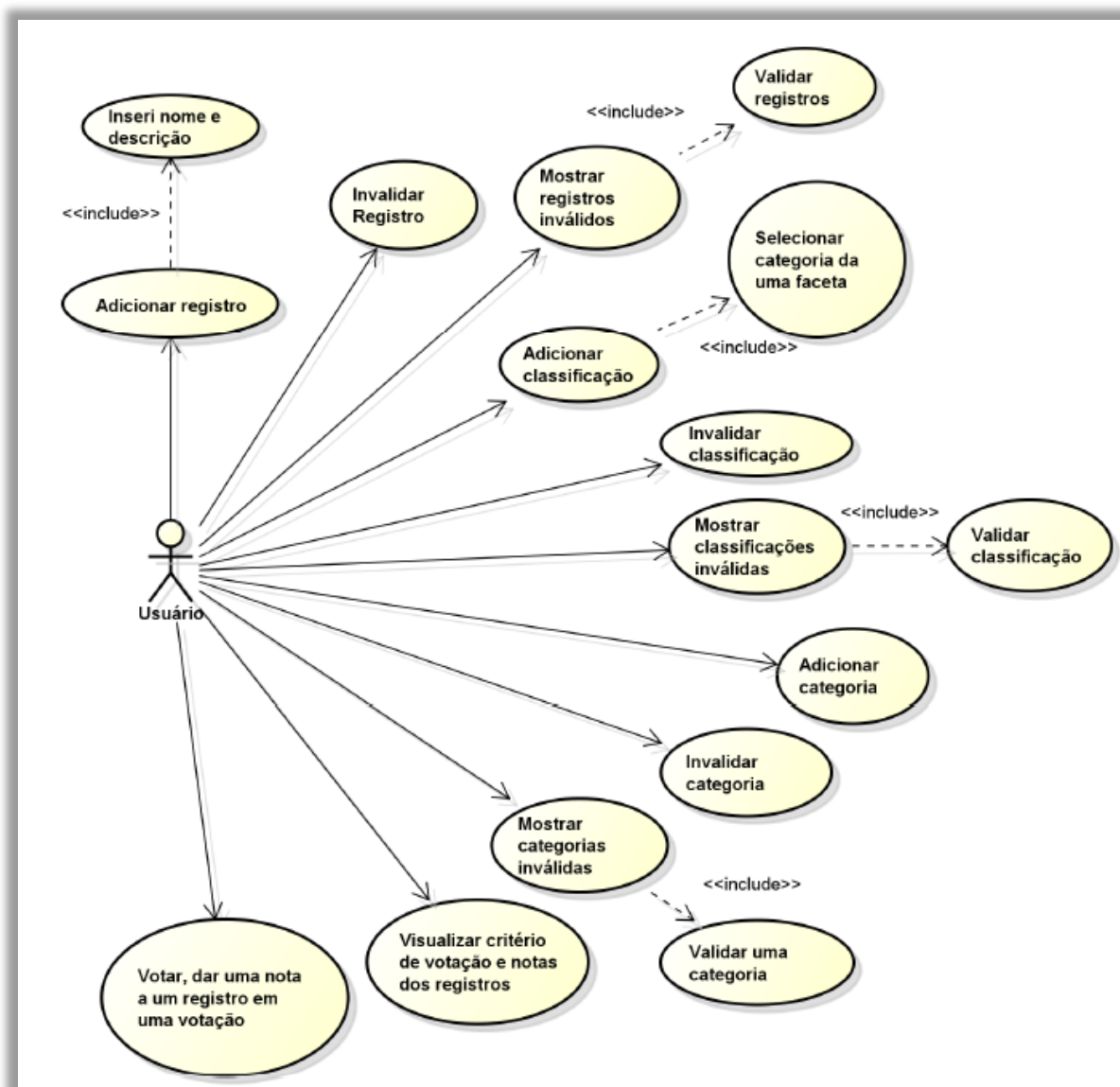
A seguir, nas FIGs. 120 e 121, são apresentados os diagramas de caso de uso atualizados para o escopo desse ciclo de desenvolvimento. É visível que a cada ciclo de

desenvolvimento a complexidade ou detalhamento dos diagramas aumentam, o que era esperado.

FIGURA 120 - Realizar navegação facetada



Fonte: Elaborado pelo autor.

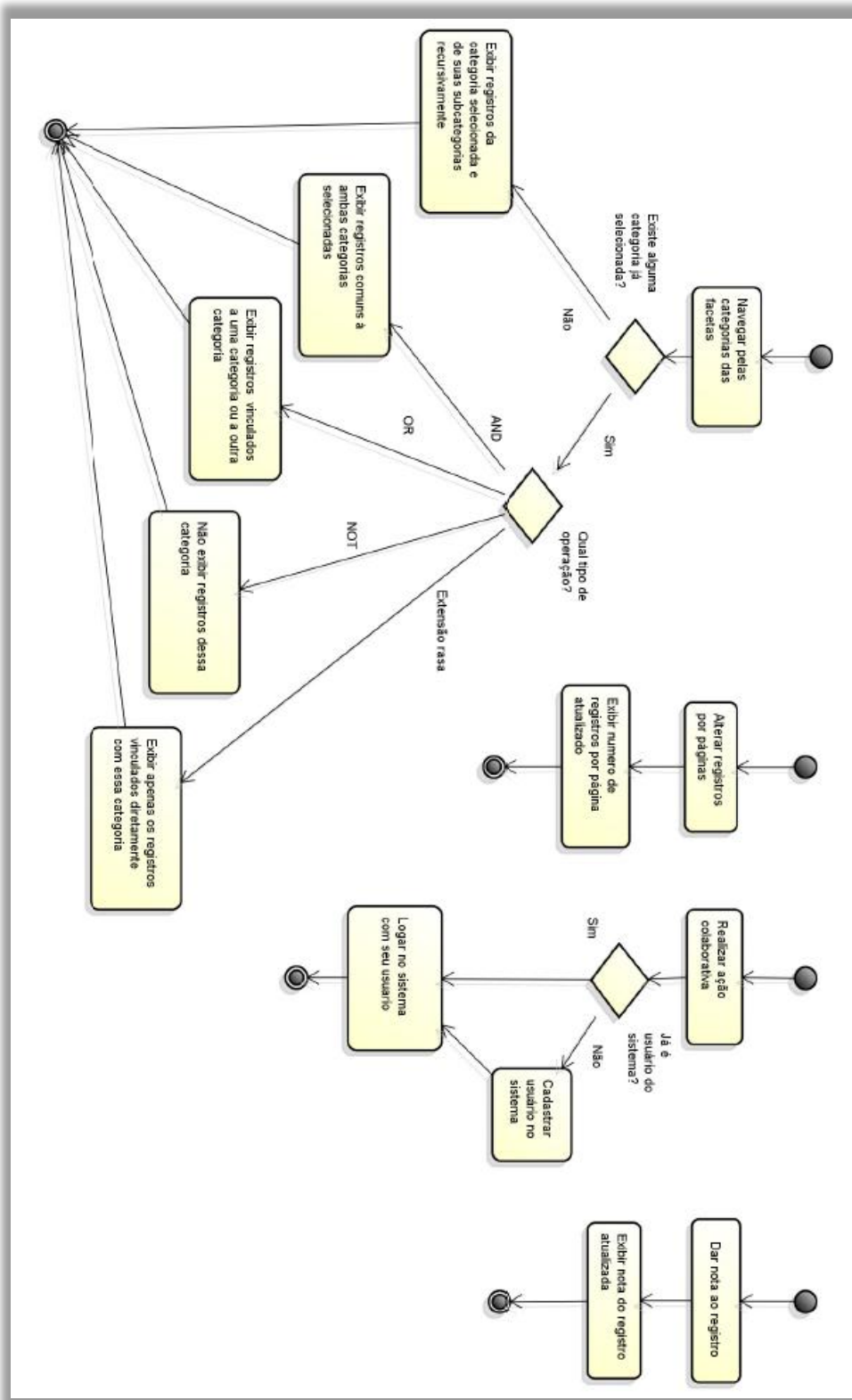
FIGURA 121 - Modelo de colaboração *wiki*

Fonte: Elaborado pelo autor.

G.3.3.2 Diagramas de atividade

Nota-se na FIG. 122 que o diagrama da navegação facetada torna-se mais complexo, principalmente com a inclusão das operações *OR* e *NOT* na interface.

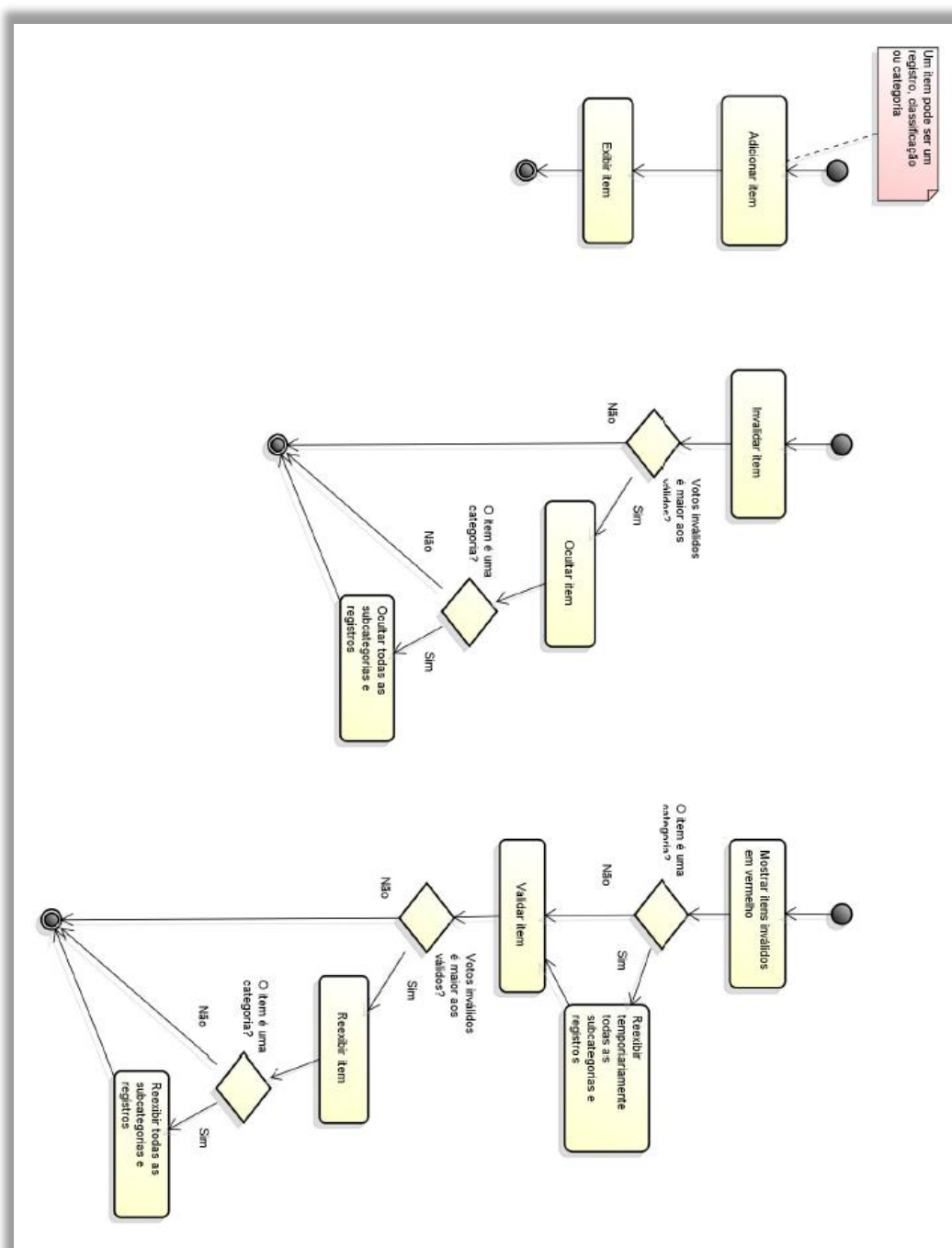
FIGURA 122 - Navegação facetada



Fonte: Elaborado pelo autor.

Neste diagrama de atividade do modelo de colaboração *wiki* (FIG. 123), generalizou-se, aplicando o mesmo modelo para tratar as ações em um item, que pode ser um registro, classificação ou categoria.

FIGURA 123 - Modelo de colaboração *wiki*



Fonte: Elaborado pelo autor.

G.3.4 Modelo de dados

A reformulação do modelo de dados visava resolver o problema de performance e, ao mesmo tempo, tratar as inconsistências dos *status* dos registros, classificações e categorias durante a navegação facetada que é recursiva.

Primeiramente, para tratar o problema de performance, a primeira abordagem cogitada foi utilizar o conceito do arquivo invertido (*inverted file*), estrutura de dados amplamente usada em máquinas de buscas, na qual ao invés de ter uma lista de documentos e seus termos, é criada uma lista de termos, sendo que para cada termo são listados todos os documentos que o contem (BAEZA-YATES; RIBEIRO-NETO, 1999).

Entretanto, como estava sendo usado o banco de dados relacional MySQL, e como havia a intenção de recuperar, além dos documentos, as subcategorias para montar a taxonomia das facetadas, foi utilizada a abordagem proposta no artigo *Hierarchical data in MySQL: easy and fast*⁸¹, no qual foi introduzido o atributo linhagem, que guarda os "antepassados" ou as categorias pais da categoria ou registro.

Dessa forma, através de uma única consulta SQL, utilizando LIKE '%categoria%', é possível recuperar todas as subcategorias ou registros. O exemplo a seguir, uma consulta pela categoria C2, retorna o conjunto com os registros RC, RB, RD e RE, o conjunto com as categorias C3, C4 e C5, e os níveis de cada uma delas para montar a árvore (QUADRO 12).

QUADRO 12 – Modelo de dados com a linhagem

ID	Sujeito	Predicado	Objeto	Status	Nível	Linhagem
1	C2	subcategoria_de	C1	1	1	C1
2	C3	subcategoria_de	C2	1	2	C1,C2
3	C4	subcategoria_de	C2	0	3	C1,C2
4	C5	subcategoria_de	C4	1	3	C1,C2,C4
5	RA	tem_categoria	C1	1		C1
6	RB	tem_categoria	C2	1		C1,C2
7	RC	tem_categoria	C3	1		C1,C2,C3
8	RD	tem_categoria	C4	1		C1,C2,C4
9	RE	tem_categoria	C5	1		C1,C2,C4,C5

Fonte: Elaborado pelo autor

Apesar desse modelo ser mais eficiente que o anterior para a navegação facetada, não era suficiente para englobar o modelo de colaboração representado pelo

⁸¹ Disponível em: <<http://www.ferdychristant.com/blog//archive/DOMM-7QJPM7>>.

status de válido e inválido. No exemplo anterior, note-se que a categoria C4 (linha 3) está com o *status* igual a zero, o que significa que foi invalidada pelo usuário. Dessa forma, ao selecionar a categoria C2 deveriam ser recuperados apenas os registros RB e RC, pois o registro RD estava relacionado à categoria C4, e o registro RE estava relacionado a categoria C5. Para solucionar essa questão foi usado um sinal de válido ou inválido, simbolizado pela letra "v" ou "i", na categoria inválida da linhagem (QUADRO 13). Dessa forma, bastava uma única consulta para obter registros e categorias válidas. Como exemplo, ao selecionar C2 a consulta seria:

```
SELECT *
      FROM SPO
      WHERE SPO.Linhagem LIKE 'C1v,C2v,%'
      AND NOT LIKE 'C1v,C2v,%i%'
      AND Predicado='tem_categoria'
```

Outra vantagem é que para mostrar todos os registros, independente do status das subcategorias da categoria selecionada, bastava retirar o NOT LIKE 'linhagem,%i%' da consulta. O atributo nível permitia também determinar a profundidade da consulta, se necessário. E para revalidar a categoria C4, bastava alterar as ocorrências de C4i para C4v e todas suas subcategorias e respectivos registros voltavam a aparecer, sem perder o *status* anterior dos mesmos, antes da invalidação. Se fosse propagada a ação de invalidar através do campo Status para os descendentes tornando-os com Status igual a 0, depois não seria possível garantir a reversão da ação para o Status original.

QUADRO 13 – Modelo de dados com a linhagem com valor de verdade⁸²

ID	Sujeito	Predicado	Objeto	Status	Nível	Linhagem
1	C2	subcategoria_de	C1	1	1	C1v
2	C3	subcategoria_de	C2	1	2	C1v,C2v
3	C4	subcategoria_de	C2	0	3	C1v,C2v
4	C5	subcategoria_de	C4	1	3	C1v,C2v,C4i
5	RA	tem_categoria	C1	1		C1
6	RB	tem_categoria	C2	1		C1v,C2v
7	RC	tem_categoria	C3	1		C1v,C2v,C3v
8	RD	tem_categoria	C5	1		C1v,C2v,C4i
9	RE	tem_categoria	C5	1		C1v,C2v,C4i,C5v

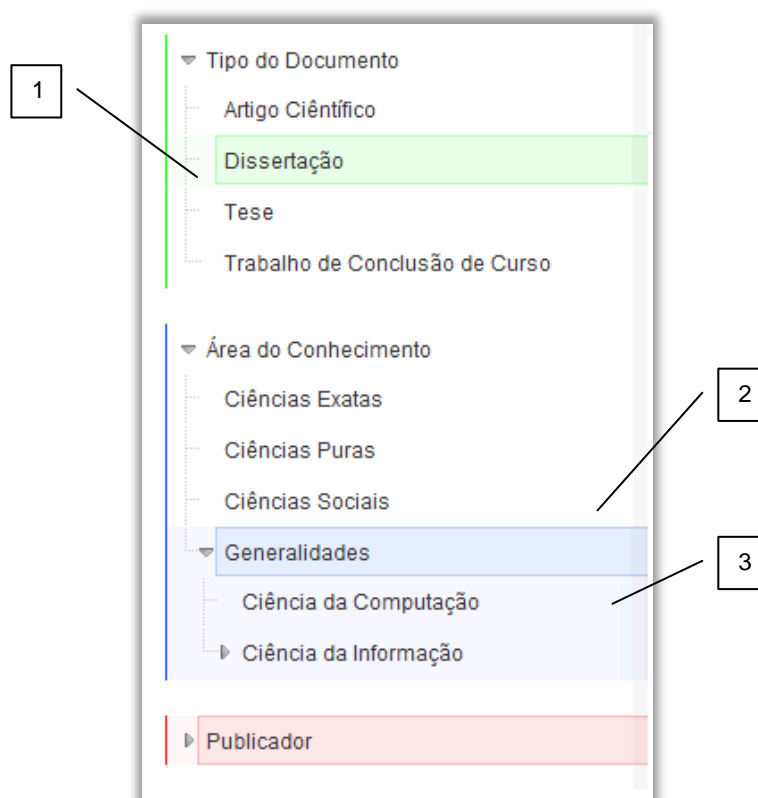
Fonte: Elaborado pelo autor.

⁸² Em negrito, todos os registros a serem recuperados sem considerar a validade. Em vermelho, os registros removidos. Em azul, os registros válidos recuperados.

G.3.5 Implementação da Interface facetada

O terceiro protótipo (FIG. 124) foi marcado por um grande avanço na interface facetada. Foram realizadas várias pequenas melhorias que solucionavam problemas antigos e acrescentavam mais semântica: 1) diferenciação entre as facetas, por meio da primeira linha colorida a esquerda; 2) as categorias selecionadas em cada faceta ficaram dentro da caixa colorida; e, 3) o conjunto que a categoria selecionada engloba com cor mais clara.

FIGURA 124 - Combinação de categorias (AND)



Fonte: Elaborado pelo autor.

Uma nova funcionalidade na navegação facetada foi incluída nesta nova versão da interface (FIG. 125). Ao selecionar uma outra categoria na mesma faceta (1), era assumida a operação *OR* ao incluir os registros relacionados a essa categoria e subcategorias nos resultados (2).

FIGURA 125 - Operação OR

The screenshot shows a search interface with a left sidebar for categories and a main content area for search results. The sidebar includes sections for 'Tipo do Documento', 'Área do Conhecimento', and 'Publicador'. Under 'Área do Conhecimento', 'Ciências Exatas' is selected. A gear icon next to it has a dropdown menu with options: 'Selecione também 'Ciências Exatas'', 'Include to the filter all instances of 'Ciências Exatas'', and 'Adicione uma subcategoria'. The main area shows two search results, both with 0 votes and a date of 18 de May de 2012. The first result is 'TCC Linagem' and the second is 'ESTUDO COMPARATIVO ENTRE INTERFACES HIPERTEXTUAIS PARA A REPRESENTAÇÃO DO CONHECIMENTO'. A search bar at the top right contains the text 'pesquise' and a 'URL' icon. A pagination bar at the bottom of the results shows 'PÁGINA 1 DE 1 PÁGINA' and '10 RESULTADOS POR PÁGINA'.

Fonte: Elaborado pelo autor.

Na FIG. 126, percebe-se que a operação *NOT* ou de exclusão também foi incluída (1). Ao realizar essa ação, a categoria aparecia tachada na interfaces e os registros abaixo dela eram excluídos dos resultados (2).

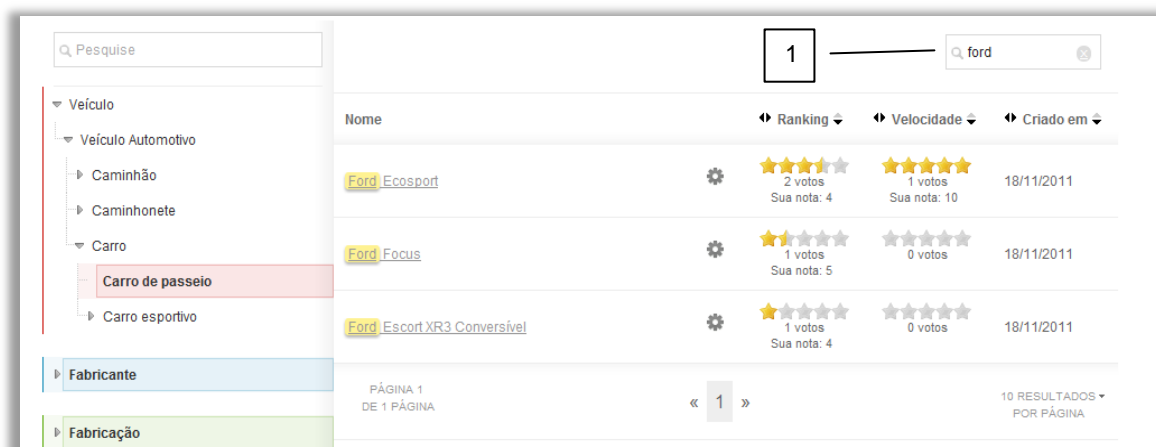
FIGURA 126 - Operação NOT

This screenshot shows the same search interface as FIG. 125, but with 'Organização da Informação' selected under 'Área do Conhecimento'. A gear icon next to it has a dropdown menu with options: 'Desselecione 'Organização da Informação'', 'Não incluir no filtro as instâncias de 'Organização da Informação'', and 'Adicione uma subcategoria'. The main area shows the same search results as FIG. 125. The search bar at the top right contains the text 'pesquise' and a 'URL' icon. A pagination bar at the bottom of the results shows 'PÁGINA 1 DE 1 PÁGINA' and '10 RESULTADOS POR PÁGINA'.

Fonte: Elaborado pelo autor.

Foi inserida na interface uma busca por palavra-chave nos registros (FIG. 127.1). Essa busca era combinada com o filtro representado pelas categorias selecionadas.

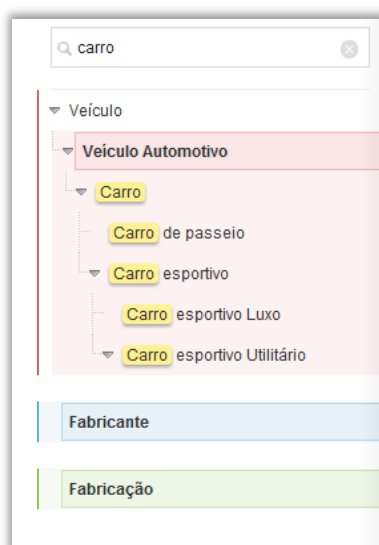
FIGURA 127 - Busca por palavra-chave nos registros



Fonte: Elaborado pelo autor.

A interface também tinha uma busca por categorias em todas as facetas. Isto era mais prático para o usuário que digitar uma busca em cada uma das facetas FIG. 128.

FIGURA 128 - Busca por palavra-chave nas facetas



Fonte: Elaborado pelo autor.

A opção Não Incluir as Subcategorias (FIG. 129) era uma opção que permitia uma maior precisão (pesquisa por extensão rasa), recuperando apenas os registros diretamente classificados com a categoria selecionada.

FIGURA 129 - Filtrando recursivamente ou apenas naquele nível hierárquico



Fonte: Elaborado pelo autor.

Nesta versão da interface facetada, o próprio usuário podia escolher a quantidade de registros por página. Também eram informados quantas páginas foram recuperadas com a paginação dos registros (FIG. 130).

FIGURA 130 - Paginação e resultados

Veículo	Nome	Ranking	Velocidade	Criado em
Fabricante	Ford Ecosport	2 votos Sua nota: 4	1 votos Sua nota: 10	18/11/2011
Fabricação	Honda Civic	2 votos Sua nota: 10	1 votos Sua nota: 10	18/11/2011
	Peugeot 206	1 votos Sua nota: 1	1 votos Sua nota: 9	18/11/2011
	Fiat Punto T-jet	1 votos Sua nota: 1	1 votos Sua nota: 6	18/11/2011
	Fusca	1 votos Sua nota: 1	1 votos Sua nota: 1	18/11/2011

PÁGINA 1 DE 1 PÁGINA « 1 » 10 RESULTADOS POR PÁGINA

Fonte: Elaborado pelo autor.

A FIG. 131 demonstra como era feita a ordenação. Os registros eram ordenados, em primeiro lugar, pela primeira coluna da esquerda para direita. Nas setas do título da coluna, era possível escolher entre ordenação crescente e decrescente.

FIGURA 131 - Ordenação de resultados por características ou votação, crescente ou decrescente

Nome	Ranking	Velocidade	Criado em
Ford Ecosport	2 votos Sua nota: 4	1 votos Sua nota: 10	18/11/2011
Honda Civic	2 votos Sua nota: 10	1 votos Sua nota: 10	18/11/2011
Troller T4	2 votos Sua nota: 10	1 votos Sua nota: 6	18/11/2011
Toyota Corolla	1 votos Sua nota: 5	1 votos Sua nota: 10	18/11/2011

Fonte: Elaborado pelo autor

G.3.6 Implementação do modelo de colaboração

No modelo de colaboração *wiki* proposto foram criadas as seguintes ações: adicionar registro; invalidar registro; mostrar registros inválidos; adicionar classificação; validar classificação; invalidar classificação; mostrar classificações inválidas; validar classificação; adicionar categoria; invalidar categoria; validar categoria; e mostrar categorias inválidas. Na FIG. 132 é exibida a interface para adicionar um registro. A proposta é que o registro seja inserido nas categorias das facetas que estivessem selecionadas naquele momento.

FIGURA 132 - Adicionar registro

Veículo

- Veículo Automotivo
 - Caminhão
 - Caminhonete
 - Carro**
 - Carro de passeio
 - Carro esportivo
 - Carro esportivo Luxo
 - Carro esportivo Utilitário
 - Carro Utilitário

Nome

[Ford Ecosport](#) [config]

[Honda Civic](#) [config]

PÁGINA 1 DE 1 PÁGINA

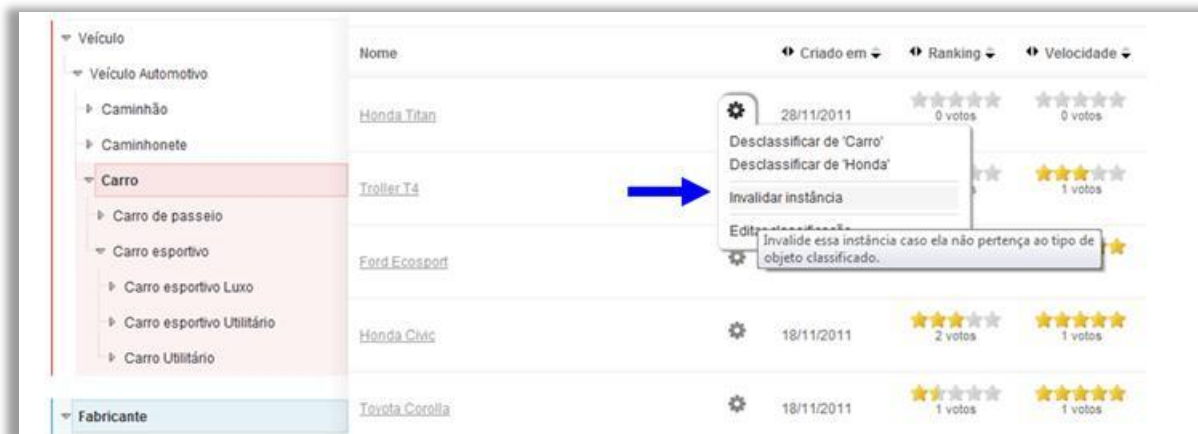
« 1 »

Honda|Titan Adicionar

Fonte: Elaborado pelo autor.

Na nova interface, existia um menu de ações acessado através do ícone de ferramentas que ficava ativo ao passar o *mouse*. Desse modo, o usuário poderia invalidar um registro que considerasse inválido para aquele catálogo (FIG. 133).

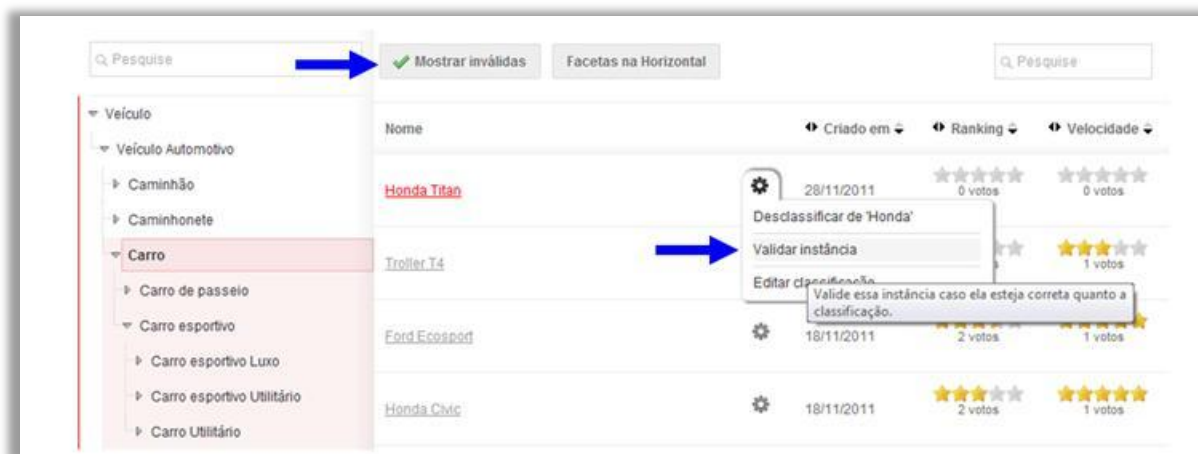
FIGURA 133 - Invalidar registro



Fonte: Elaborado pelo autor.

Um usuário, ao perceber que um legítimo registro fora invalidado indevidamente, poderia clicar em *Mostrar Inválidos* e depois revalidar o registro (FIG. 134).

FIGURA 134 - Mostrar registro inválido e validar registro



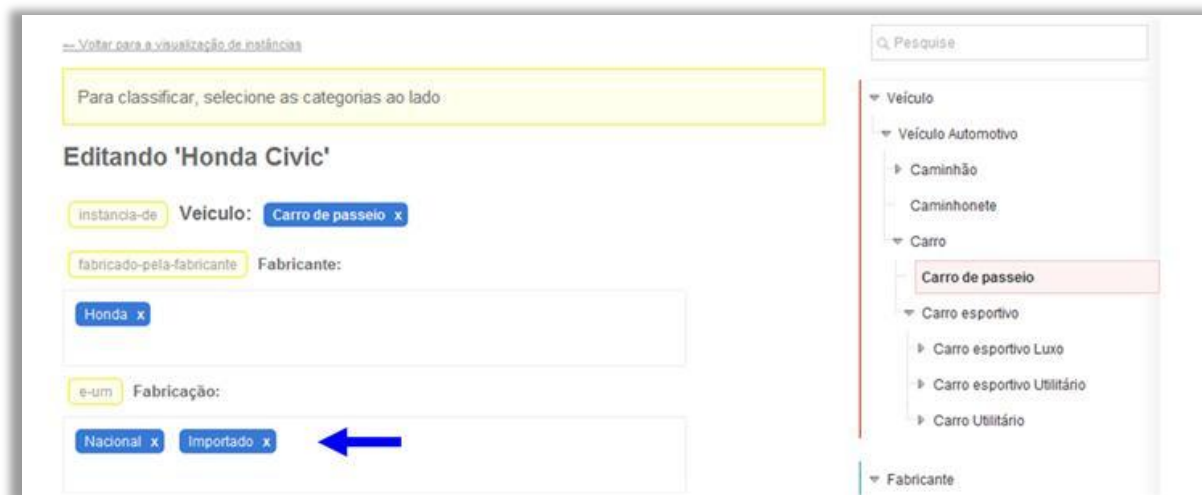
Fonte: Elaborado pelo autor.

Era permitido ao usuário, adicionar novas classificações a um registro. Ao realizar essa ação o usuário é direcionado a uma nova página que exibe as facetas do lado oposto (FIG. 135).

Tal comportamento foi implementado para diferenciar o ato da navegação facetada do ato de classificação facetada, sendo que eram usadas as mesmas taxonomias.

Como foi usado um modelo baseado em triplas SPO no armazenamento, era possível especificar na interface as relações semânticas entre o registro e as categorias.

FIGURA 135 - Adicionar classificação



Fonte: Elaborado pelo autor.

Um usuário que detectasse uma classificação errônea em um registro poderia, invalidá-lo para que o mesmo não aparecesse mais em uma determinada filtragem (FIG. 136).

FIGURA 136 - Invalidar classificação de um registro

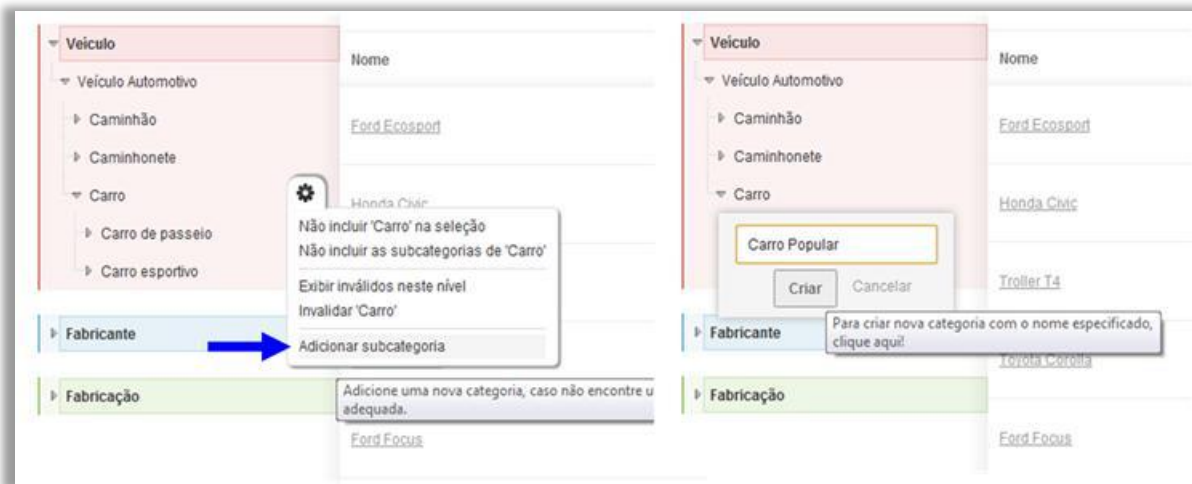


Fonte: Elaborado pelo autor

As ações que exigiam maior processamento no servidor eram as que envolviam as categorias do catálogo. A ação de adicionar uma categoria sanava a necessidade do usuário no momento em que desejava fazer uma classificação e não encontrava a categoria. Poderia ser apenas uma contribuição direta, para deixar a taxonomia facetada mais

completa. Com esta ação o usuário também criava um novo subconjunto de votação de registros (FIG. 137).

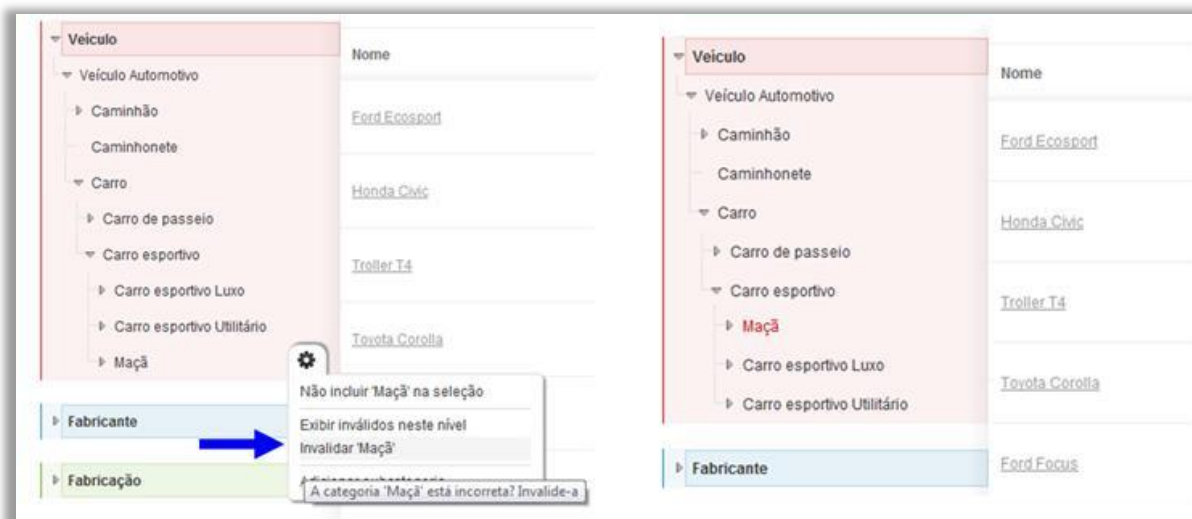
FIGURA 137 - Adicionar categoria



Fonte: Elaborado pelo autor.

A ação de invalidar a categoria é o principal recurso para algum controle de qualidade da taxonomia criada pela comunidade de usuários (FIG. 138). Categorias inseridas em atos de vandalismos ou, sem pensar, poderiam ser ocultadas por membros mais experientes (no exemplo abaixo Maçã em catálogo de veículos).

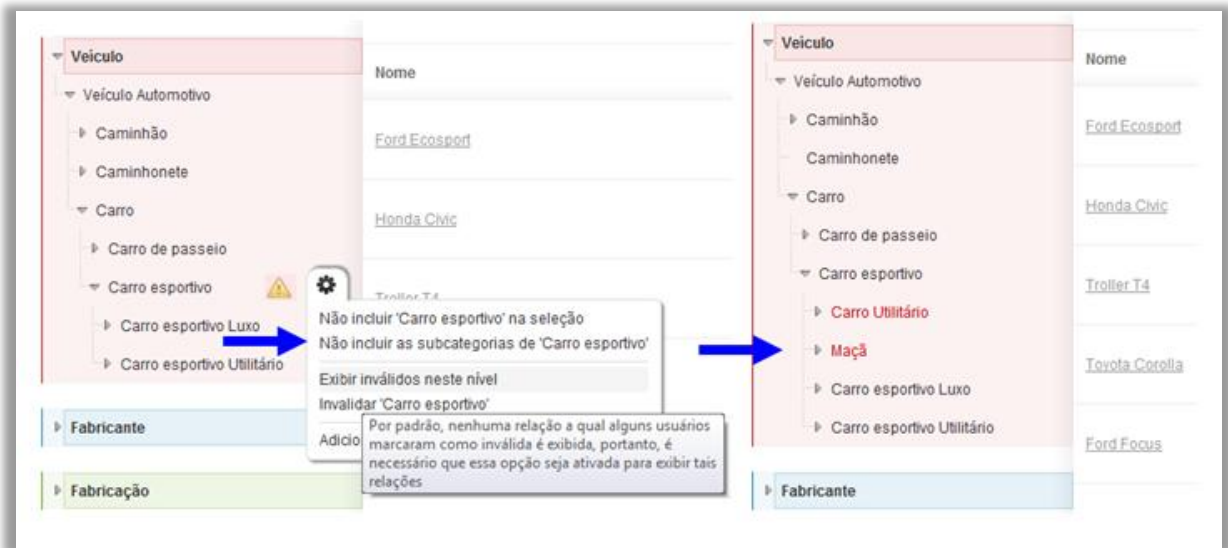
FIGURA 138 - Invalidar categoria



Fonte: Elaborado pelo autor.

Nesta nova versão, existia um ícone para simbolizar que haviam subcategorias inválidas em uma determinada categoria (triângulo com exclamação). As categorias inválidas eram diferenciadas pela cor vermelha (FIG. 139).

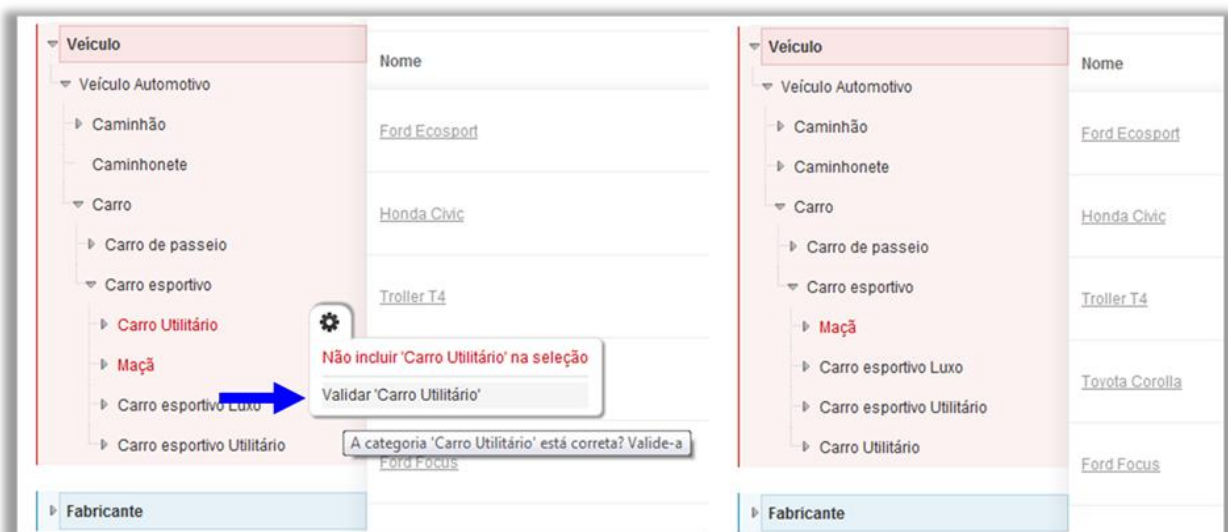
FIGURA 139 - Mostrar categorias inválidas



Fonte: Elaborado pelo autor.

Após solicitar mostrar categorias inválidas era possível revalidar a categoria ao atribuir um voto positivo e anular o voto negativo, anteriormente dado, ou seja, ações de invalidação errôneas poderiam ser revertidas (FIG. 140).

FIGURA 140 - Validar uma categoria inválida



Fonte: Elaborado pelo autor.

Um usuário poderia atribuir uma nota a um registro em uma dada votação. Utiliza-se a cor vermelha para diferenciar a nota dada pelo usuário da média das notas dos usuários que o registro recebeu (FIG. 141).

FIGURA 141 - Dar nota para um registro em uma determinada votação



Fonte: Elaborado pelo autor.

G.3.7 Teste: registro de falhas e sucessos

A interface facetada deste ciclo de desenvolvimento teve vários aperfeiçoamentos, com uma interface AJAX⁸³ e as melhores práticas vistas em outras interfaces facetadas. Isso foi fruto de contínuo do desenvolvimento iterativo realizado nos ciclos passados.

Entretanto, após construir uma interface facetada com boa usabilidade, que respondia de forma fluída aos comandos do usuário, foram notadas questões pertinentes ao modelo de colaboração adotado que era puramente *wiki*.

No QUADRO 14 apresenta-se uma síntese da evolução obtida em cada ciclo de desenvolvimento.

⁸³ Asynchronous *JavaScript* and XML. Disponível em <<http://www.w3schools.com/ajax>>.

QUADRO 14 – Avaliação dos ciclos de desenvolvimento

Ciclo de desenvolvimento	Foco da implementação	Avaliação
1	Início do desenvolvimento do catálogo e navegação facetada, e <u>modelo de colaboração wiki apenas para classificação dos registros</u>	Interface de navegação facetada com baixa usabilidade Modelo de colaboração <i>wiki</i> incompleto, restrito apenas a classificação dos registros
2	Melhoria na usabilidade da interface de navegação facetada Ampliação do <u>modelo de colaboração wiki para classificação dos registros e gestão das categorias</u>	Inconsistências na navegação facetada com o modelo de colaboração <i>wiki</i> Interface com baixa usabilidade Problema de performance
3	Alteração no modelo de dados para melhoria da performance <u>Tratamento das inconsistências</u> do modelo de colaboração <i>wiki</i> com a navegação facetada Melhoria na usabilidade da interface facetada e inclusão da busca por palavra-chave	Modelo de colaboração <i>wiki</i> não é adequado para a gerir toda a taxonomia facetada Evolução na interface e performance

Fonte: Elaborado pelo autor.

Nesse ciclo de desenvolvimento foi realizado o teste *Quick and Dirty*, com alguns usuários, estagiários da área de design, e com um especialista da área, o Prof. Dr. Dagobert Soergel da Universidade de Búfalo de Nova Iorque. Nos testes foram realizadas ações no catálogo que resultaram em uma série de recomendações. A seguir, lista-se vários problemas constatados nesta abordagem.

O modelo de colaboração dava uma liberdade muito grande (em excesso) ao usuário. Qualquer usuário poderia adicionar uma categoria que não fazia sentido em qualquer nível hierárquico. Mesmo com a possibilidade de invalidar aquela categoria, corria-se o risco daquela categoria errada ficar exposta por muito tempo, confundindo usuários que estavam apenas querendo fazer a navegação facetada no sistema. Usuários leigos tem muita dificuldade em entender conceitos como classe e subclasse (gênero e espécie), ou diferenciar essa relação todo e parte. Tal fato prejudicava a qualidade da taxonomia criada. Outros criavam livremente as categorias para marcar os recursos, no estilo da folksonomia, sem se importar onde estavam criando as categorias. Tal fato não refletia em boas categorias.

A questão de invalidar a categoria poderia provocar impactos enormes e não desejados. Exigia uma responsabilidade maior que o ato de adicionar uma categoria. A ação de invalidar uma categoria se propagava para todas as suas subcategorias e recursos relacionados ao estado de inválido da mesma. Assumia-se que aquela categoria e toda a

sua linhagem não deveria existir. Não apenas as categorias filhas e subcategorias, mas, também, todos os recursos relacionados às mesmas (o que foi chamado de linhagem).

Uma situação que poderia ocorrer é: o usuário criar uma categoria em um nível hierárquico e outra pessoa criar a mesma categoria em outro nível hierárquico (havia uma proteção para que não fosse criado no mesmo nível). O problema é que se outros usuários criassem subcategorias em ambas às categorias, e se depois fosse decidido que uma delas estava na posição errada, ao ser invalidada, suas subcategorias também seriam invalidadas.

Nesse caso, o certo era fazer uma fusão entre as categorias e não invalidar uma delas; o que seria uma tarefa bem complexa. Outra tarefa complexa é alterar uma categoria de lugar: o ato de mover significava mudar toda a linhagem de lugar.

Um detalhe dessa implementação foi que todas as taxonomias eram compartilhadas com todos os catálogos, ou seja, seguiu-se a filosofia *wiki*, no qual as contribuições são públicas. A vantagem principal era permitir o reuso das taxonomias. Assim, quem estivesse fazendo um catálogo de carros, poderia compartilhar a faceta "Localização" com outro usuário que tem um catálogo de móveis. Dessa forma, os usuários de ambos os catálogos poderiam contribuir na construção da mesma faceta. Mas na prática foi percebido que essa não era uma boa estratégia. Por exemplo, o administrador do catálogo de carros poderia querer organizar os carros usando a seguinte estrutura em Localização: Brasil > Estado > Cidades. Já o administrador do catálogo de imóveis poderia querer organizar assim: Brasil > Regiões > Estado > Cidades. Nenhuma das duas formas estavam erradas. Outra situação: se um usuário alterasse uma categoria da taxonomia da faceta de um catálogo, isso refletiria em todos os outros catálogos que compartilhavam aquela taxonomia, afetando comunidades de usuários que nem se poderia supor.

A folksonomia não era um recurso disponível na interface, o que causava uma insatisfação nos usuários que não tinham outra opção senão classificar um recurso através das categorias existentes, ou criar alguma categoria aleatoriamente, o que causava distorções na taxonomia, como exposto acima.

G.4 Diferenças entre as plataformas dos ciclos de desenvolvimento

Em virtude dos problemas encontrados após os testes do terceiro ciclo de desenvolvimento, o escopo do projeto foi completamente reformulado. No QUADRO 15 apresenta-se as diferenças dos ciclos de desenvolvimento apresentados neste apêndice (primeira plataforma), em relação ao ciclo apresentado no capítulo 5 (segunda plataforma).

QUADRO 15 – Diferenças na plataforma e escopo do projeto nos ciclos de desenvolvimento

	Primeira Plataforma	Segunda Plataforma
Sistema de Gestão de Conteúdo (CMS)	<i>WebSiteBaker</i>	<i>Wordpress</i>
Moderação	<i>Wiki</i> puro: pré-aprovada para todas as ações do catálogo com possibilidade de reversão	Mista: pré-aprovado para ações com menor impacto e por aprovação em ações com maior impacto)
Tipo de navegação da interface	Não recursiva ou recursiva	Apenas recursiva
Modelo de dados	Inspirado no modelo RDF para representação em bando de dados relacional (tabela SPO)	Representação de taxonomia através do atributo <i>parent</i> na tabela do banco de dados
Compartilhamento da informação	Todas as taxonomias eram compartilhadas para todos os catálogos (estilo <i>wiki</i>)	Cada catálogo tem suas próprias taxonomias
Interface	A mesma interface realizava a navegação facetada, gestão da taxonomia e registros	Interfaces separadas para a navegação facetada, indexação e gestão da taxonomia
Indexação	Apenas através da taxonomia facetada	Etiquetagem social e taxonomia facetada
Tratamento das ações na gestão da taxonomia, validação e navegação	Criação de funções próprias para lidar com o modelo RDF	Utilização das funções nativas do <i>Wordpress</i>

Fonte: Elaborado pelo autor.

Na primeira plataforma, na qual foi utilizado o CMS *WebSiteBaker*, havia uma tabela única para representar todos os objetos do catálogo (categorias e registros) através do modelo SPO, com a indicação dos antecessores (linhagem). Assim, em cada registro continha a referência de todos os pais até a raiz. Essa representação da primeira plataforma era mais semântica, permitindo a construção colaborativa de catálogos semânticos, pois indicava o relacionamento entre o sujeito e o objeto. Mas isso adicionou uma maior complexidade nas consultas SQL.

Na segunda plataforma, na qual foi utilizado o CMS *Wordpress*, foi usada uma tabela para representar a taxonomia (com o ID da categoria pai). Uma tabela para representar as classificações, com os vínculos dos registros com as categorias. Na segunda plataforma, o custo computacional das ações de gestão da taxonomia, por exemplo, remover categoria, é menor do que na primeira abordagem, que precisava alterar todos os registros que continham uma determinada categoria modificada. Na segunda plataforma se

alterava apenas um registro. Ou seja, o modelo SPO com a linhagem era boa para a navegação facetada e poderia representar uma poli-hierarquia ou um grafo, mas tinha um custo computacional maior na colaboração e gestão da taxonomia.

Com o tempo percebeu-se, também, que a escolha da ferramenta não tinha sido ideal. Apesar do *WebSiteBaker* ter uma interface de administração amigável, este CMS não era um *framework* robusto, com classes para manipulação de seus objetos. O *WebSiteBaker* também não tinha suporte nativo a taxonomias e etiquetas, um recurso essencial para a proposta e disponível em outros CMS como o *Wordpress*. Além desse fato, o *WebSiteBaker* tem uma comunidade muito menor e com pouca documentação, o que dificultava o tratamento de dúvidas durante a programação.

A transição de um CMS para o outro ajudou a perceber que o sistema deve assumir um comportamento quando fossem realizadas ações de gestão das categorias da taxonomia. No QUADRO 16 foram apresentadas as diferenças no tratamento dessas ações da primeira plataforma (*WebSiteBaker*) para a segunda plataforma (*Wordpress*).

Na primeira plataforma todas as ações para gerenciar as categorias do catálogo foram desenvolvidas do início. As principais ações foram: criar categoria, remover categoria, mover categoria e alterar o nome da categoria.

Na segunda plataforma foi mais fácil gerenciar as categorias da taxonomia, pois este é um recurso nativo do *Wordpress*, com um modelo exaustivamente testado e aperfeiçoado há quase 10 anos por milhares de desenvolvedores e usuários. Já existiam funções nativas para todas as ações de gestão de categorias, indexação dos registros e busca.

A ação mais crítica era quando um usuário queria remover uma categoria. A principal diferença na segunda plataforma era que apagava o registro ao invés de apenas invalidar. Assim, ao invés de considerar as subcategorias inválidas, na segunda plataforma, as categorias netas tornaram-se filhas da categoria avó. As classificações com a categoria apagada também seriam removidas (QUADRO 16). Ou seja, foi assumido um comportamento na primeira plataforma, sem perceber que havia outras opções para tratar uma determinada ação de gestão da taxonomia. De qualquer maneira, os primeiros ciclos permitiram um grande aprendizado referente aos requisitos de uma interface facetada amigável e eficiente e de um modelo de colaboração adequado. Apesar dos problemas relatados a terceira versão apresentou soluções para que a interface pudesse realizar as operações *OR* e *NOT* de forma amigável para o usuário, sem causar um excesso de informação para o mesmo (as opções somente apareceriam ao passar o mouse sobre a categoria).

QUADRO 16 – Ações para gestão da taxonomia⁸⁴

Criar categoria	Primeira Plataforma	Segunda Plataforma
<p>O que fazer quando já existe uma como mesmo nome? Ex: Cria uma nova categoria C</p> <pre> A - 1 B - 2 C - 3 D - 4 <u>C</u> E - 5 </pre>	<p>Bloquear se for na mesma pasta ou em qualquer outro nível</p> <p>Ex:</p> <pre> A - 1 B - 2 C - 3 D - 4 € E - 5 </pre>	<p>Avisar o usuário, que já tem uma categoria como essa, mas deixa criar, se quiser (com identificador diferente) Ex:</p> <pre> A - 1 B - 2 C - 3 D - 4 <u>C2</u> E - 5 </pre>
Remover categoria	Primeira Plataforma	Segunda Plataforma
<p>Ex: Se C for removido, o que ocorre com D?</p> <pre> A - 1 B - 2 C - 3 D - 4 E - 5 </pre>	<p>Propagar a exclusão ou tornar inválido todas as subcategorias, classificações e registros</p> <p>Ex:</p> <pre> A - 1 B - 2 C - 3 D - 4 E - 5 </pre>	<p>Remover as classificações dos registros com a categoria excluída e tornar suas subcategorias filhas da categoria pai</p> <p>Ex:</p> <pre> A - 1 B - 2 D - 4 E - 5 - 3 </pre>
Mover categoria	Primeira Plataforma	Segunda Plataforma
<p>Ex: Mover C para abaixo de E</p> <pre> A - 1 B - 2 C - 3 D - 4 E - 5 </pre>	<p>Tornar a categoria inválida naquela posição e criar uma outra categoria levando seus descendentes para a nova hierarquia</p> <p>Ex:</p> <pre> A - 1 B - 2 C - 3 E - 5 C - 3 D - 4 </pre>	<p>Mover a categoria pai da categoria alterada. Todos os seus descendentes mudam automaticamente.</p> <p>Ex:</p> <pre> A - 1 B - 2 E - 5 C - 3 D - 4 </pre>
Alterar nome da categoria	Primeira Plataforma	Segunda Plataforma
<p>Deve haver o mesmo tratamento existente na ação de criar uma categoria com o mesmo nome.</p>	<p>Bloquear se for na mesma pasta ou em qualquer outro nível</p> <p>Ex:</p> <pre> A - 1 B - 2 C - 3 D - 4 € </pre>	<p>Avisar o usuário, que já tem uma categoria com esse nome. Cria se ele quiser.</p> <p>Ex:</p> <pre> A - 1 B - 2 C - 3 D - 4 <u>C2</u> </pre>

Fonte: Elaborado pelo autor.

⁸⁴ As letras representam as categorias e os números os registros.