UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Câmpus de Marília

FACULDADE DE FILOSOFIA E CIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

MARCIA CRISTINA DOS REIS

INTERFACE DE BUSCA E RECUPERAÇÃO DE INFORMAÇÃO:

Um Modelo Baseado nos Princípios da Hermenêutica e na Utilização de Ontologias

Marília
2019

MARCIA CRISTINA DOS REIS

INTERFACE DE BUSCA E RECUPERAÇÃO DE INFORMAÇÃO:
Um Modelo Baseado nos Princípios da Hermenêutica e na Utilização de Ontologias

Tese apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista (UNESP), Faculdade de Filosofia e Ciências de Marília, como requisito parcial para obtenção do título de Doutor em Ciência da Informação.

Área de Concentração: Informação, Tecnologia e Conhecimento

Linha de Pesquisa: Informação e Tecnologia

Orientador: Prof. Dr. Edberto Ferneda

Financiamento: Fundação Araucária

Marília
2019

R375i

Reis, Marcia Cristina dos
Interface de busca e recuperação de informação : um modelo baseado nos princípios da hermenêutica e na utilização de ontologias / Marcia Cristina dos Reis. -- Marília, 2019
308 p. : il., tabs., fotos

Tese (doutorado) - Universidade Estadual Paulista (Unesp), Faculdade de Filosofia e Ciências, Marília
Orientador: Edberto Ferneda

1. Recuperação da informação. 2. Ontologia. 3. Hermenêutica. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Filosofia e Ciências, Marília. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

MARCIA CRISTINA DOS REIS

INTERFACE DE BUSCA E RECUPERAÇÃO DE INFORMAÇÃO:
Um Modelo Baseado nos Princípios da Hermenêutica e na Utilização de Ontologias

BANCA EXAMINADORA

Prof. Dr. Edberto Ferneda (Orientador)
Departamento de Ciência da Informação
Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP)

Prof. Dr. Carlos Cândido de Almeida
Departamento de Ciência da Informação
Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP)

Prof. Dr. Walter Moreira
Departamento de Ciência da Informação
Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP)

Prof. Dr. Licínio Gomes Roque
Departamento de Engenharia Informática
Universidade de Coimbra (UC)

Profa. Dra. Luciana Maria Vieira Pöttker
Departamento de Informática
Instituto Federal do Paraná (IFPR)

Marília, 14 de março de 2019.

Tudo tem seu tempo determinado, e há tempo para todo o propósito debaixo do céu: há tempo de nascer e tempo de morrer, tempo de plantar e tempo de colher, tempo de matar e tempo de curar, tempo de derrubar e tempo de construir, tempo de chorar e tempo de sorrir, tempo de espalhar pedras e tempo de juntar pedras, tempo de abraçar e tempo de afastar-se de abraçar, tempo de ganhar e tempo de perder, tempo de estar calado e tempo de falar, tempo de amar e tempo de aborrecer, tempo de guerra e tempo de paz...

(Eclesiastes, 3:1–5)

AGRADECIMENTOS

“Posso todas as coisas naquele que me fortalece.” (Filipenses, 4:13)

O agradecimento é uma forma de reconhecer que as dificuldades da vida não são superadas sozinhas. Muitas são as pessoas que me auxiliaram na construção desta tese e me incentivaram nos momentos difíceis. No entanto, não posso deixar de reconhecer que, sem a orientação e as bênçãos de **Deus** na minha vida, tudo teria sido em vão. Te agradeço Senhor, por mais uma etapa vencida. Obrigada por ter me concedido forças nos momentos difíceis, por ter iluminado os meus caminhos e por ter estado ao meu lado em todos os momentos.

Aos **meus pais**, Gaspar (*in memorian*) e Clarice, com suas orações, apoio e constante motivação, sempre presentes, como estímulos para a caminhada da minha vida. Embora, pela vontade de Deus, não possa ter ambos ao meu lado neste momento, sei que o meu querido pai, onde quer que ele esteja, se lhe for permitido, está muito feliz e orgulhoso por mais esta conquista. Amo vocês mais do que tudo nesta vida. Obrigada por tudo!

Ao **meu orientador**, Prof. Dr. Edberto Ferneda, pelo apoio incondicional, pelos conhecimentos compartilhados, pelo incentivo, pela disponibilidade em sempre ajudar no desenvolvimento da tese, pela confiança em mim depositada e pelo carinho nos momentos mais difíceis. Obrigada por me fazer acreditar que eu seria capaz de vencer mais esta etapa e por me mostrar que uma relação de orientação pode sim ser baseada na amizade, no respeito mútuo e no apoio incondicional. Isso fez toda a diferença durante os quatro anos que trabalhamos juntos! Você é um ser humano incrível, de uma humildade rara e de um coração gigantesco que muitas não cabe no seu peito. Que privilégio ter tido você como meu orientador!

Ao **Prof. Dr. Licínio Gomes Roque**, por me receber na Universidade de Coimbra, pelas suas orientações no período em que estive em Portugal, pelos seus direcionamentos e pelas contribuições valiosas no desenvolvimento desta tese. Serei eternamente grata por tudo o que fez por mim e por tudo o que me ensinou. Deus foi muito generoso comigo quando te colocou no meu caminho!

Aos **professores Dr. Carlos Cândido de Almeida** e **Dra. Luciana Maria Vieira Pöttker**, pela disponibilidade em participar do meu exame de qualificação, pelas discussões construtivas e pelas contribuições que foram de grande importância para a

finalização desta tese.

Ao **Prof. Dr. Walter Moreira**, por ter aceito o convite para fazer parte da banca de defesa, pela leitura minuciosa da minha tese e por todas as correções sugeridas, que possibilitaram o aprimoramento da pesquisa realizada, bem como a visualização de perspectivas de estudos futuros.

Aos **docentes do Programa de Pós-Graduação em Ciência da Informação** que participaram desta jornada, compartilhando suas experiências e contribuindo para a minha formação.

À **Fundação Araucária** pelo financiamento desta pesquisa.

Ao **Instituto Federal do Paraná** (IFPR), que concedeu afastamento integral durante a minha qualificação, permitindo uma maior dedicação às atividades do doutorado.

Aos **professores do Instituto Federal do Paraná** (IFPR) pela colaboração decorrente do afastamento das minhas atividades.

Ao **Fábio Martins Campos**, que esteve ao meu lado nos momentos mais difíceis, que me ofereceu seu amor, seu carinho e seu apoio. Obrigada por fazer parte da minha vida, dos meus planos e dos meus sonhos.

Aos **amigos do Programa de Pós-Graduação em Ciência da Informação** da UNESP de Marília e às pessoas com as quais tive oportunidade de conviver neste período, em especial ao Fabrício Baptista, por todos os momentos compartilhados, pela parceria de todas as horas e por cada palavra de apoio e motivação nos momentos difíceis.

Aos **amigos do Programa de Pós-Graduação em Ciência da Informação da Universidade de Coimbra**, que me acolheram e estiveram comigo durante o desenvolvimento do meu Estágio de Doutoramento, em especial, ao Ariovaldo Veiga de Almeida (Ari) e ao Antonio José Rodrigues Neto (Neto from Brazil). Saudades imensas de vocês!

Aos **amigos mais próximos** que estiveram presentes durante toda esta caminhada, em especial à Andressa Alexandre Mendes, à Josiane Pinheiro, à Maria Aparecida Mariano, à Maria Elvira Souza, à Terezinha Viccioli e ao Leonardo Sturion.

Por fim, agradeço a todos os que contribuíram, direta ou indiretamente, para o desenvolvimento desta tese.

REIS, Marcia Cristina dos. **Interface de busca e recuperação de informação**: um modelo baseado nos princípios da hermenêutica e na utilização de ontologias. 2019. 308 f. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2019.

RESUMO

As interfaces de busca disponíveis no contexto da Web e dos repositórios digitais, apesar da importância que possuem em função da necessidade cada vez maior de recuperação de informação relevante e útil, tiveram mudanças pouco significativas nos últimos anos em relação à forma de interação com o usuário. Além disso, apresentam inúmeros problemas relacionados principalmente à dependência de conhecimento em relação ao assunto investigado, às dificuldades em expressar uma necessidade de informação por meio de palavras-chave, à ambiguidade da linguagem e à composição inadequada de termos de busca. Com base neste panorama, o estudo desenvolvido nesta tese teve como objetivo agregar os recursos oferecidos pelas ontologias e pela hermenêutica às interfaces de busca, na tentativa de sugerir melhorias que pudessem minimizar os problemas observados. Neste sentido, foi possível propor um modelo de interface de busca e recuperação de informação baseado nos princípios da hermenêutica e na utilização de ontologias, com funcionalidades projetadas para proporcionar uma melhor interação do usuário com o sistema e facilitar a análise da similaridade existente entre os documentos e a interpretação de um *corpus* documental. A metodologia adotada para alcançar os objetivos propostos contou com pesquisa exploratória e aplicada, que ofereceu o aporte teórico necessário para a fundamentar o estudo e desenvolver o modelo proposto. Os resultados da avaliação de usabilidade realizada por meio dos protótipos de baixa fidelidade indicaram a viabilidade para a implementação da proposta, bem como das funcionalidades projetadas com fundamento nos princípios da hermenêutica e nas ontologias que, de acordo com a perspectiva dos usuários, podem trazer contribuições significativas para o processo de busca e recuperação de informação, bem como para a análise dos resultados recuperados. O principal diferencial da proposta está na possibilidade de interação do usuário com o sistema por meio de uma interface gráfica, que disponibiliza os termos do domínio pesquisado no formato de mapa de conceitos para que possam ser explorados, analisados e utilizados na elaboração de expressões de busca semânticas. Conclui-se, portanto, que o desenvolvimento de interfaces de busca com recursos elaborados a partir dos fundamentos da hermenêutica e das ontologias possui potencial para contribuir com o processo de recuperação de informação, uma vez que as funcionalidades previstas podem facilitar a formulação de expressões de busca, a análise da similaridade entre os documentos e a interpretação de um *corpus* documental.

Palavras-chave: *Design* de Interação. Hermenêutica. Interface de Busca. Ontologia. Recuperação de Informação.

REIS, Marcia Cristina dos. **Information retrieval and information interface**: a model based on the principles of hermeneutics and the use of ontologies. 2019. 308 f. Thesis (PhD Degree in Information Science) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2019.

ABSTRACT

Searching interfaces available in the context of the Web and digital repositories, despite their importance due to the increasing need to retrieve relevant and useful information, have had few significant changes in the last years regarding mainly the interaction mode with the user. In addition, they present numerous issues related mainly to the dependence of knowledge on the subject investigated, the difficulties in expressing a need of information through keywords, the language ambiguity and the search terms inappropriate composition. Based on this scenario, the study developed in this thesis aimed to aggregate the resources offered by the ontologies and the hermeneutics to the search interfaces, in an attempt to suggest improvements that could minimize the observed faults. In this regard, it was possible to propose a search and retrieval interface model based on the principles of hermeneutics and use of ontologies, with functionalities designed to provide more excellent user interaction with the system and facilitate the analysis of the similarity between documents and interpretation of a documentary corpus. The methodology adopted to achieve the proposed objectives counted on exploratory and applied research, which offered the theoretical contribution necessary to base the study and develop the proposed model. The usability evaluation results performed using the low fidelity prototypes indicated the viability of its implementation, as well as the projected functionalities based on the principles of hermeneutics and ontologies that, according to the users' perspective, can bring significant contributions to the information retrieval and retrieval process, as well as for the recovered results analysis. The main distinction of the proposal is the possibility of user interaction with the system through a graphical interface, which provides the domain terms searched in the format of the concept map, so that they can be explored, analyzed and used in the elaboration of semantic search expressions. It is concluded, therefore, that the development of search interfaces with resources elaborated from the foundations of hermeneutics and ontologies has the potential to contribute to the information retrieval process since the expected functionalities can facilitate the formulation of search expressions, the analysis of the similarity between documents and the interpretation of a documentary corpus.

Keywords: Interaction Design. Hermeneutics. Search Interface. Ontology. Information Retrieval.

LISTA DE FIGURAS

Figura 1 - Áreas envolvidas no estudo 25
Figura 2 - Percurso metodológico para o desenvolvimento da pesquisa 26
Figura 3 - Palavras-chave utilizadas na busca 30
Figura 4 - Representação do processo de recuperação de informação 37
Figura 5 - Combinações booleanas 40
Figura 6 - Representação de uma busca no modelo vetorial 43
Figura 7 - Resultado de uma busca 45
Figura 8 - Resultados de busca do Google 48
Figura 9 - Campos de pesquisa em comportamento informacional 55
Figura 10 - Relacionamentos entre documentos de repositórios variados 66
Figura 11 - Relacionamentos entre objetos e termos 67
Figura 12 - Tipos de ontologias 76
Figura 13 - Método 101 para o desenvolvimento de ontologias 78
Figura 14 - Elementos Gráficos 85
Figura 15 - Princípios de *Gestalt* 86
Figura 16 - Exemplo da aplicação de elementos gráficos e princípios de *Gestalt* 87
Figura 17 - Visualização baseada em lista indentada 88
Figura 18 - Visualização baseada em nó-*link* e árvore 90
Figura 19 - Visualização nó-*link* baseada em *layout* radial 91
Figura 20 - Visualização baseada em aproximação 93
Figura 21 - Visualização baseada em preenchimento de espaço (*layout* retangular) 94
Figura 22 - Visualização baseada em preenchimento de espaço (*layout* radial) 95
Figura 23 - Visualização baseada em contexto+foco e distorção 97
Figura 24 - Visualização baseada em paisagem 99
Figura 25 - Interface de busca do Hermeneus 101
Figura 26 - Interface de busca do MuseumFinland 103
Figura 27 - Interface de busca do SemSearch 104
Figura 28 - Relação entre o *design* de interação e demais abordagens 154
Figura 29 - Ciclo de vida do *Design* de Interação 156
Figura 30 - Participantes e problemas de usabilidade encontrados 161
Figura 31 - Base teórica do modelo de interface de busca 165
Figura 32 - Relacionamentos entre os documentos 173

Figura 33 - Protótipo da tela inicial da interface de busca 178
Figura 34 - *Layout* do modelo de interface de busca 180
Figura 35 - Relacionamentos entre conceitos de uma ontologia 184
Figura 36 - Busca baseada em palavras-chave 188
Figura 37 - Interação para a elaboração de expressões de busca 189
Figura 38 - Seleção de perspectiva de análise de busca 190
Figura 39 - Elaboração de expressões de busca com dois termos 191
Figura 40 - Perspectivas de análise, termos e resultados da busca 192
Figura 41 - Elementos de similaridade entre resultados e expressão de busca 195
Figura 42 - Elementos de similaridade entre resultados da busca 195
Figura 43 - Notas explicativas e áreas de conhecimento 196
Figura 44 - Sinônimos e termos relacionados 198
Figura 45 - Expansão dos resultados da busca 199
Figura 46 - Variações linguísticas 200
Figura 47 - Antônimos e termos conceitualmente contraditórios 201
Figura 48 - Substituição semântica de critérios de busca 202
Figura 49 - Busca baseada em estrutura terminológica 205
Figura 50 - Expansão da estrutura terminológica 206
Figura 51 - Elaboração de expressões de busca 208
Figura 52 - Estrutura terminológica de um domínio especializado 210
Figura 53 - Busca baseada em perspectivas de análise de contexto 212
Figura 54 - Busca baseada em duas perspectivas de análise de contexto 213
Figura 55 - Histórico do autor 214
Figura 56 - Resultados de busca baseada de perspectivas de análise de contexto 215
Figura 57 - Seleção de documento para análise de conteúdo 218
Figura 58 - Documento selecionado para análise de conteúdo 219
Figura 59 - Comentários interpretativos 220
Figura 60 - Erros e falhas 221
Figura 61 - Parágrafos avaliados 222
Figura 62 - Sugestão de pesquisas complementares 223
Figura 63 - Busca baseada em perspectivas de análise de conteúdo 225
Figura 64 - Resultados de busca baseada em perspectivas de análise de conteúdo 225
Figura 65 - Rede de similaridade dos resultados de busca 226
Figura 66 - Grupo de documentos 227

Figura 67 - Inclusão de material na base de dados 229
Figura 68 - Interpretação de um novo documento 230
Figura 69 - Relacionamentos entre material e ontologia de domínio 231
Figura 70 - Inserção de novo termo à ontologia de domínio 232
Figura 71 - Inserção de um novo autor à ontologia de domínio 233
Figura 72 - Definição de conceitos pelos usuários 234

LISTA DE GRÁFICOS

Gráfico 1 - Comparativo de sucesso das tarefas realizadas 247
Gráfico 2 - Comparativo das perspectivas de avaliação 248
Gráfico 3 - Avaliação dos atributos de uso da interface 250
Gráfico 4 - Grau de utilidade dos recursos da interface 251
Gráfico 5 - Avaliação de questões gerais da interface 252
Gráfico 6 - Avaliação da ontologia como componente de interface 253
Gráfico 7 - Avaliação da ontologia como recurso para indexação semântica 255
Gráfico 8 - Avaliação dos recursos de interpretação de um documento ou de um *corpus* 256

LISTA DE TABELAS

Tabela 1 - Tipo de ontologia adotada 112
Tabela 2 - Áreas de desenvolvimento de pesquisa 114
Tabela 3 - Abordagens envolvendo ontologias 116
Tabela 4 - Problemáticas discutidas 117
Tabela 5 - Perfil dos participantes 237
Tabela 6 - Características do roteiro de atividades 238
Tabela 7 - Resultados dos ensaios de usabilidade realizados no primeiro cenário 239
Tabela 8 - Resultados dos ensaios de usabilidade realizados no segundo cenário 242
Tabela 9 - Resultados dos ensaios de usabilidade realizados no terceiro cenário 243
Tabela 10 - Resultados dos ensaios de usabilidade realizados no quarto cenário 244
Tabela 11 - Resultados dos ensaios de usabilidade realizados no quinto cenário 245

LISTA DE QUADROS

Quadro 1 - Percurso metodológico para o desenvolvimento da pesquisa 27
Quadro 2 - Operadores booleanos AND, OR, NOT e sinais modificadores 61
Quadro 3 - Problemas do processo de busca e recuperação de informação na Web 70
Quadro 4 - Sugestões de melhorias para o processo de busca 71
Quadro 5 - Comparativo entre os métodos de visualização de ontologias 107
Quadro 6 - Síntese da utilização de ontologias em interfaces de busca 109
Quadro 7 - Problemas e sugestões de melhorias relacionados à busca 169
Quadro 8 - Problemas e sugestões de melhorias relacionados aos resultados 170
Quadro 9 - Requisitos funcionais do modelo de interface de busca 175
Quadro 10 - Requisitos não funcionais do modelo de interface de busca 176
Quadro 11 - Busca baseada em palavras-chave 187
Quadro 12 - Indexação dos documentos recuperados 193
Quadro 13 - Busca baseada em estrutura terminológica 203
Quadro 14 - Conhecimento implícito das ontologias 207
Quadro 15 - Busca baseada em perspectivas de análise de contexto 211
Quadro 16 - Busca baseada em perspectivas de análise de conteúdo 217
Quadro 17 - Inclusão e análise de material 228
Quadro 18 - Requisitos funcionais analisados nos ensaios de usabilidade 249
Quadro 19 - Recursos do modelo de interface de busca projetados a partir das ontologias 258
Quadro 20 - Princípios hermenêuticos aplicados ao modelo de interface de busca 259

LISTA DE ABREVIATURAS E SIGLAS

2D	Bidimensional (Duas Dimensões)
3D	Tridimensional (Três Dimensões)
ACM	Association for Computing Machinery (Associação para Estudos Computacionais)
BRAPCI	Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação
CC	Ciência da Computação
CI	Ciência da Informação
CML	Context Modelling Language (Linguagem de Modelagem de Contexto)
DAML	Darpa Agent Markup Language (Linguagem de Marcação de Agente)
DBpedia	Dados Estruturados da Wikipedia
DI	*Design* de Interação
Doc	Documento
FA	Facilidade de Aprendizagem
FAF	Facilidade de Acesso à Funcionalidade
F-Logic	Frame Logic (*Frame* Lógico)
Freebase	Base de Conhecimento Livre e Colaborativa
HCI	Human-Computer Interaction (Interação Humano-Computador)
HTML	Hypertext Markup Language (Linguagem de Marcação de Hipertexto)
IA	Inteligência Artificial
IEEE	Institute of Electrical and Electronic Engineers (Instituto de Engenheiros Eletricistas e Eletrônicos)
IFPR	Instituto Federal do Paraná
IHC	Interação Humano-Computador
II	Interpretação da Informação
INFOVIS	Visualização de Informações
KIF	Knowledge Interchange Format (Formato de Intercâmbio de Conhecimento)

OCML	Operational Conceptual Modelling Language (Linguagem de Modelagem Conceitual Operacional)
OIL	Ontology Inference Layer (Camada de Inferência Ontológica)
RC	Representação do Conhecimento
RDF	Resource Description Framework (Estrutura de Descrição de Recursos)
RF	Requisito Funcional
RI	Recuperação de Informação
RV	Realidade Virtual
SCIELO	Scientific Electronic Library Online (Biblioteca Eletrônica Científica *Online*)
SHOE	Simple HTML Ontology Extension (Extensão de Ontologia HTML Simples)
SI	Sistemas de Informação
SIGAI	Special Interest Group on Artificial Intelligence (Grupo de Interesse Especial em Inteligência Artificial)
SIGCHI	Special Interest Group on Computer-Human Interaction (Grupo de Interesse Especial em Interação Humano-Computador)
SIGDOC	Special Interest Group on Design of Communication (Grupo de Interesse Especial em *Design* da Comunicação)
SIGIR	Special Interest Group on Information Retrieval (Grupo de Interesse Especial em Recuperação de Informação)
SIGSOFT	Special Interest Group on Software Engineering (Grupo de Interesse Especial em Engenharia de *Software*)
SIGWEB	Special Interest Group on Hypertext, Hypermedia and Web (Grupo de Interesse Especial em Hipertexto, Hipermídia e Web)
SPARQL	Protocol and RDF Query Language (Protocolo e Linguagem de Consulta RDF)
SRI	Sistemas de Recuperação de Informação
SUMO	Suggested Upper Merged Ontology (Ontologia Superior Padrão)
SUO-KIF	Standard Upper Ontology Knowledge Interchange Format (Formato de Intercâmbio de Conhecimento de Ontologia Superior Padrão
TI	Tecnologia da Informação
UNESP	Universidade Estadual Paulista Júlio de Mesquita Filho

UNIOESTE	Universidade Estadual do Oeste do Paraná
UNISINOS	Universidade do Vale do Rio dos Sinos
UNOPAR	Universidade Norte do Paraná
URLs	Uniform Resource Locator (Localizador Padrão de Recursos)
UTFPR	Universidade Tecnológica Federal do Paraná
VI	Visibilidade da Informação
W3C	World Wide Web Consortium (Consórcio World Wide Web)
XML	Extensible Markup Language (Linguagem Extensível de Marcação)
XOL	XML-based Ontology Language (Linguagem Ontológica Baseada em XML)
YAGO	Yet Another Great Ontology (Ontologia Genérica de Alto Nível)

SUMÁRIO

1 INTRODUÇÃO 20
1.1 CARACTERIZAÇÃO DO PROBLEMA DE PESQUISA 23
1.2 HIPÓTESES 23
1.3 PROPOSIÇÃO 23
1.4 TESE 24
1.5 OBJETIVOS 24
1.5.1 Objetivo Geral 24
1.5.2 Objetivos Específicos 24
1.6 PROCEDIMENTOS METODOLÓGICOS 25
1.7 ESTRUTURA DO TRABALHO 32

2 RECUPERAÇÃO DE INFORMAÇÃO 35
2.1 MODELOS DE RECUPERAÇÃO DE INFORMAÇÃO 39
2.1.1 Modelo Booleano 39
2.1.2 Modelo Vetorial 42
2.1.3 Modelo Probabilístico 44
2.2 RECUPERAÇÃO DE INFORMAÇÃO NA WEB 46
2.2.1 Comportamento de Busca em Sistemas de Recuperação de Informação 54
2.2.2 Características do Processo de Busca 60
2.2.3 Recuperação de Documentos Textuais 65
2.3 RESUMO E DISCUSSÃO DO CAPÍTULO 69

3 ONTOLOGIAS 73
3.1 PRINCÍPIOS PARA A VISUALIZAÇÃO DE ONTOLOGIAS 82
3.2 VISUALIZAÇÃO GRÁFICA DE ONTOLOGIAS 87
3.2.1 Visualização Baseada em Lista Indentada (*Indented List*) 88
3.2.2 Visualização Baseada em Nó-*Link* e Árvore (*Node-Link and Tree*) 89
3.2.3 Visualização Baseada em Aproximação (*Zoomable*) 92
3.2.4 Visualização Baseada em Preenchimento de Espaço (*Space Filling*) 93
3.2.5 Visualização Baseada em Contexto+Foco Distorção (*Context+Focus Distortion*) 96
3.2.6 Visualização Baseada em Paisagem (*Information Landscapes*) 98
3.3 INTERFACES DE BUSCA BASEADAS EM ONTOLOGIAS 100

3.4 RESUMO E DISCUSSÃO DO CAPÍTULO 106

4 UTILIZAÇÃO DE ONTOLOGIAS NA RECUPERAÇÃO DE INFORMAÇÃO 111
4.1 TIPOS DE ONTOLOGIA ADOTADOS 111
4.2 ÁREAS DE PESQUISA E PRINCIPAIS ABORDAGENS 113
4.3 PROBLEMÁTICAS E PROPOSTAS DE MELHORIAS NO ÂMBITO DO USO DE ONTOLOGIAS NA BUSCA E RECUPERAÇÃO DE INFORMAÇÃO 116
4.3.1 Busca Baseada em Ontologias (*Searching*) 117
4.3.2 Navegação Baseada em Ontologias (*Browsing*) 120
4.3.3 Busca Híbrida Baseada em Ontologias (*Searching and Browsing*) 121
4.3.4 Visualização de Informação 123
4.3.5 Outras Perspectivas 124
4.4 VALIDAÇÃO DE PROPOSTAS POR MEIO DE ESTUDOS DE USUÁRIOS 125
4.5 RESUMO E DISCUSSÃO DO CAPÍTULO 127

5 HERMENÊUTICA 129
5.1 PRINCÍPIOS DA HERMENÊUTICA 142
5.1.1 Princípio do Círculo Hermenêutico 143
5.1.2 Princípio da Contextualização 144
5.1.3 Princípio da Interação entre Pesquisadores e Sujeitos 145
5.1.4 Princípio da Abstração e da Generalização 145
5.1.5 Princípio do Raciocínio Dialético 146
5.1.6 Princípio das Múltiplas Interpretações 147
5.1.7 Princípio da Suspeição 148
5.1.8 Princípio do Compartilhamento de Linguagem 148
5.1.9 Interdependência entre Princípios 150
5.2 RESUMO E DISCUSSÃO DO CAPÍTULO 150

6 *DESIGN* DE INTERAÇÃO 153
6.1 *DESIGN* BASEADO EM CENÁRIOS 161
6.2 RESUMO E DISCUSSÃO DO CAPÍTULO 163

7 MODELO DE INTERFACE DE BUSCA PROPOSTO 165
7.1 ETAPAS DE DESENVOLVIMENTO DA PROPOSTA 166
7.2 REQUISITOS DO MODELO DE INTERFACE DE BUSCA 168

7.2.1 Análise do Domínio e Oportunidades de Melhorias 168
7.2.2 Requisitos Funcionais e Não Funcionais 175
7.2.3 Análise do Público Alvo e das Tarefas Previstas 176
7.3 *DESIGN* DE ALTERNATIVAS E PROTÓTIPOS DA INTERFACE 180
7.3.1 *Layout*, Estrutura de Navegação e Recursos Gráficos 180
7.3.2 Personagens e Cenários 187
7.3.2.1 Cenário de busca baseada em palavras-chave 187
7.3.2.2 Cenário de busca baseada em estrutura terminológica 203
7.3.2.3 Cenário de busca baseada em perspectivas de análise de contexto 211
7.3.2.4 Cenário de busca baseada em perspectivas de análise de conteúdo 216
7.3.2.5 Cenário de inclusão e análise de material 228
7.4 AVALIAÇÃO DO MODELO PROPOSTO 235
7.4.1 Pesquisa Piloto 236
7.4.2 Ensaios de Usabilidade 236
7.4.3 Avaliação de Usabilidade 249
7.5 RESUMO E DISCUSSÃO DO CAPÍTULO 256

8 CONSIDERAÇÕES FINAIS 262

REFERÊNCIAS 266

APÊNDICES 287
APÊNDICE A - Roteiro de Tarefas para Avaliação de Usabilidade 288
APÊNDICE B - Questionário Final de Avaliação de Usabilidade 304

1 INTRODUÇÃO

A quantidade excessiva de informações disponibilizadas diariamente na rede tem exigido o desenvolvimento de sistemas de recuperação de informação cada vez mais eficazes, rápidos e precisos, que são projetados com o objetivo de "minimizar as dificuldades do usuário em localizar a informação requisitada, ou seja, diminuir o tempo gasto em um processo de busca até que a informação desejada possa ser acessada" (CRISTOVÃO; DUQUE; SERQUEIRA, 2012, p.5-6). São de grande importância para a sociedade e organizações em geral, que possuem a necessidade de recuperar documentos para a tomada de decisão e resolução de problemas (BEPPLER, 2008).

A Recuperação de Informação (RI) é uma área relevante para a Ciência da Informação (CI) e apresenta inúmeros trabalhos expressivos (SARACEVIC, 1997; GUARINO; MASOLO; VETERE, 1999; GARCIA; SICILIA, 2003; FININ et al., 2005; PAZ-TRILLO; WASSERMANN; BRAGA, 2005; FORD; EAGLESTONE; MADDEN, 2009; FERNEDA, 2013; STYPEREK; CIESIELCZYK; SZWABE, 2014; SAVOLAINEN, 2016), com grande potencial para o desenvolvimento de pesquisas. A área tem estado presente em inúmeros encontros, conferências, fóruns e seminários específicos (ATSRI, 2013; ECIR, 2013; ENANCIB, 2013, 2017; FIRE, 2016; SIGIR, 2017).

> O processo de recuperação de informação consiste em identificar, no conjunto de documentos (*corpus*) de um sistema, quais atendem à necessidade de informação do usuário. O usuário de um sistema de recuperação de informação está, portanto, interessado em recuperar 'informação' sobre um determinado assunto e não em recuperar dados que satisfazem sua expressão de busca, nem tampouco documentos, embora seja nestes que a informação estará registrada. Essa característica é que diferencia os sistemas de recuperação de informação dos Sistemas Gerenciadores de Banco de Dados (ou simplesmente 'banco de dados'), estudados e implementados desde o nascimento da Ciência da Computação (FERNEDA, 2003, p.14-15).

Neste contexto, as interfaces de busca são responsáveis pela comunicação do usuário com os sistemas de recuperação de informação, de forma a auxiliá-los na elaboração de consultas, na escolha entre as fontes de informação disponíveis e na compreensão dos resultados de pesquisa. Por serem destinadas a usuários com conhecimentos diversificados, principalmente no contexto da Web, devem ser projetadas de forma simples, necessitando o mínimo de aprendizado para a sua utilização (HEARST, 2009; BAEZA-YATES; RIBEIRO-NETO, 2013).

As interfaces de busca disponíveis atualmente, a maioria baseadas nas caixas de texto, tiveram mudanças pouco significativas nos últimos anos em relação à forma de interação com o usuário (exemplos: Google, Yahoo Search, Lycos, Bing etc.), apesar das muitas propostas de melhorias que surgiram através das pesquisas científicas realizadas na área (HOBBS; PFITZNER; POWERS, 2002; HOGAN et al., 2011). De forma geral, possuem um modelo padrão de interação com o usuário, tanto para a elaboração das expressões de busca, quanto para a apresentação dos resultados, que tem como principal característica a disponibilização de um campo de busca (no formato de caixa de texto), utilizado pelos usuários para fornecer uma lista de termos ou palavras-chave representativas, de acordo com as suas necessidades de informação. Como resposta à consulta realizada, a interface retorna uma série de resultados, cada um dos quais contendo título da página, *link* de acesso e descrição concisa do conteúdo que ele representa (MAZZA, 2009).

Apesar de ser um modelo já consolidado entre usuários, as interfaces de busca atuais apresentam inúmeros problemas, relacionados principalmente com a interação do usuário na elaboração das expressões de busca e na análise dos resultados retornados, tais como: dificuldade do usuário em traduzir uma necessidade de informação em expressão de busca, falta de conhecimento em relação ao contexto pesquisado, problemas linguísticos que interferem na construção dos termos de busca, sobrecarga de informação, dificuldade de manipular grandes quantidades de resultados, entre outros. Além disso, as ferramentas de busca são, em sua maioria, de natureza sintática, recuperando apenas documentos indexados que correspondam total ou parcialmente com a expressão de busca elaborada pelo usuário. As indexações realizadas de forma automática, referem-se ao documento como um todo e não exploram informações relacionadas à sua estrutura em particular, como seções, subseções, parágrafos e figuras.

Assim como ocorre na Web, os repositórios digitais de informação científica, pertencentes a um contexto mais específico e direcionado a um público diferenciado, também apresentam as mesmas dificuldades observadas nos buscadores *on-line*. Este cenário apresenta desafios sobre o armazenamento, a recuperação e a transformação dessas informações em conhecimento.

O presente trabalho tem como objetivo apresentar um “Modelo de interface de busca e recuperação de informação”, com funcionalidades voltadas para três aspectos importantes no âmbito de uma pesquisa científica: a interação do usuário com o sistema, a similaridade existente entre os documentos retornados em uma busca e a interpretação de um *corpus* documental.

Os recursos para facilitar o processo de interação do usuário com o sistema e a análise da similaridade entre os documentos foram projetados a partir do uso de ontologias, que deverão ser utilizadas como componentes de interface e na estrutura de indexação dos artigos científicos. Os relacionamentos existentes entre os documentos, se explorados corretamente, apresentam grande potencial para a descoberta de informações que podem potencializar a busca e revelar resultados não evidenciados. Para fundamentar o desenvolvimento das funcionalidades referentes à interpretação de um documento ou de um *corpus* documental, recorreu-se a área da Filosofia, mais especificamente, aos princípios da hermenêutica, baseados principalmente no filósofo alemão Hans-Georg Gadamer. O projeto contou ainda com as contribuições da área da Visualização de Informação (VI), que forneceu os recursos gráficos necessários para a visualização das informações apresentadas ao usuário, tanto no âmbito da elaboração da expressão de busca como na interpretação dos resultados.

Além da Ciência da Informação, da Visualização da Informação e da Filosofia, o estudo contou ainda com contribuições advindas da Ciência da Computação (CC), relacionadas, principalmente, com o processo de desenvolvimento de *software.* A importância da aproximação entre a Ciência da Informação com as três áreas que fundamentam esta pesquisa é evidenciada na literatura. Vieira e Correa (2011) afirmam que a Visualização da Informação pode fornecer formas de representação gráficas para a elaboração de sistemas de Recuperação de Informação, com o intuito de facilitar a percepção e o entendimento do usuário, além de permitir a dedução de novos conhecimentos com base na informação apresentada.

Capurro (2003) evidencia a contribuição da Filosofia, mais especificamente no contexto da Hermenêutica, decorrente de autores como Friedrich Schleiermacher (1768-1834), Wilhelm Dilthey (1833-1911), Edmund Husserl (1859-1938), Martin Heidegger (1889-1976) e Hans-Georg Gadamer (1900-2002), para os estudos da informação, enfatizando que eles podem fornecer fundamentos teóricos e métodos de análise para área de CI. A Ciência da Computação, por sua vez, apresenta uma relação significativa de cooperação no desenvolvimento de processos e aplicações que contribuem para o aprimoramento da Ciência da Informação em suas diversas vertentes (SARACEVIC, 1996).

Como principais contribuições, o estudo tem a intenção de minimizar os problemas observados na área, relacionados principalmente com as interfaces de busca e, como isso, proporcionar aos usuários maior capacidade de consulta, maior relevância às respostas obtidas e mais facilidades no processo de organização, análise e exploração de informação científica.

1.1 CARACTERIZAÇÃO DO PROBLEMA DE PESQUISA

Partindo do pressuposto de que as interfaces de busca não identificam as relações existentes entre os documentos e não oferecem subsídios para o desenvolvimento de análises sobre as informações recuperadas, a presente pesquisa norteia-se pelo seguinte questionamento: *quais funcionalidades podem ser implementadas em uma interface de busca, a partir dos princípios da hermenêutica e do uso de ontologias, para apoiar as atividades de um pesquisador na interação com o sistema, na análise da similaridade entre os documentos e na interpretação de um corpus documental?*

1.2 HIPÓTESES

A partir do problema apresentado, as hipóteses que norteiam a presente pesquisa partem dos seguintes pressupostos:

a) A visualização gráfica da ontologia, com a representação dos conceitos e relacionamentos do domínio, amplia a busca do usuário, principalmente quando ele não tem conhecimento sobre o assunto pesquisado, seja por meio da apresentação de termos desconhecidos ou de resultados correlacionados;
b) A anotação das relações existentes entre os documentos, com o objetivo de análise ou síntese, realizada individualmente ou em colaboração, revela ao usuário contextos de similaridade, bem como múltiplas perspectivas de interpretação;
c) A semântica presente nos relacionamentos de uma ontologia favorece a compreensão das terminologias do domínio, bem como das relações existentes entre elas, e torna mais eficaz a descoberta de contextos e associações relevantes;
d) Os princípios da hermenêutica, quando aplicados à uma interface de busca, potencializam a interpretação de um documento ou de um *corpus* documental.

1.3 PROPOSIÇÃO

Estudar os princípios da hermenêutica e os recursos oferecidos pelas

ontologias e aplicá-los como instrumento de auxílio ao processo de busca e recuperação de informações presentes em documentos textuais.

1.4 TESE

A tese (ou premissa) desta pesquisa é comprovar que uma interface de busca baseada nos princípios da hermenêutica e na utilização de ontologias contribui para a melhoria do processo de recuperação de documentos textuais, pois fornece subsídios para facilitar a formulação de expressões de busca, a análise da similaridade entre os documentos e a interpretação de um *corpus* documental a partir de múltiplas perspectivas de análise.

1.5 OBJETIVOS

1.5.1 Objetivo Geral

Com o propósito de contribuir para obtenção de resultados mais significativos no processo de recuperação de informação, o principal objetivo desta pesquisa é o desenvolvimento de um "Modelo de interface de busca e recuperação de informação baseado nos princípios da hermenêutica e na utilização de ontologias", com funcionalidades que favoreçam a interação do usuário com o sistema e auxiliem na análise da similaridade entre os documentos e na interpretação de um *corpus* documental.

1.5.2 Objetivos Específicos

a) Discutir a área de Recuperação de Informação, enfatizando os principais conceitos envolvidos no processo de busca, bem como os fatores que influenciam o comportamento do usuário;
b) Discutir conceitos e características das ontologias, bem como os principais métodos e técnicas de visualização;
c) Analisar a utilização de ontologias na busca e recuperação de informação e identificar suas principais contribuições para a área;
d) Identificar os princípios da hermenêutica para a compreensão e interpretação documental;
e) Identificar recursos e interações, com base nos princípios da

hermenêutica e na utilização de ontologias, que possam potencializar a busca e auxiliar as atividades desenvolvidas pelo usuário;

f) Propor um modelo de interface de busca e recuperação de informação, descrevendo personagens, cenários de uso e protótipos;

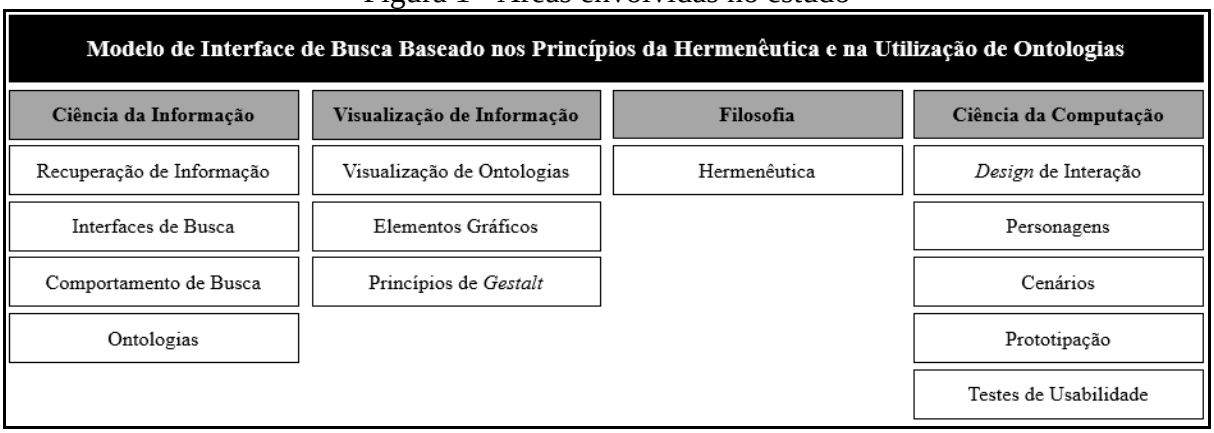
g) Avaliar a usabilidade do modelo proposto.

1.6 PROCEDIMENTOS METODOLÓGICOS

Os procedimentos metodológicos adotados para subsidiar o desenvolvimento da presente pesquisa estão fundamentados nos aspectos teóricos e conceituais da metodologia científica. Segundo Fachin (2003, p.27), o método pode ser caracterizado como "um instrumento do conhecimento que proporciona aos pesquisadores, em qualquer área de sua formação, orientação geral que facilita planejar uma pesquisa, formular hipóteses, coordenar investigações, realizar experiências e interpretar os resultados".

Para a concepção do "Modelo de interface de busca e recuperação de informação" foi necessária a investigação de quatro áreas distintas, que deram subsídios teóricos para o estudo: Ciência da Informação, Visualização de Informação, Filosofia e Ciência da Computação.

Figura 1 - Áreas envolvidas no estudo



Fonte: desenvolvido pela autora

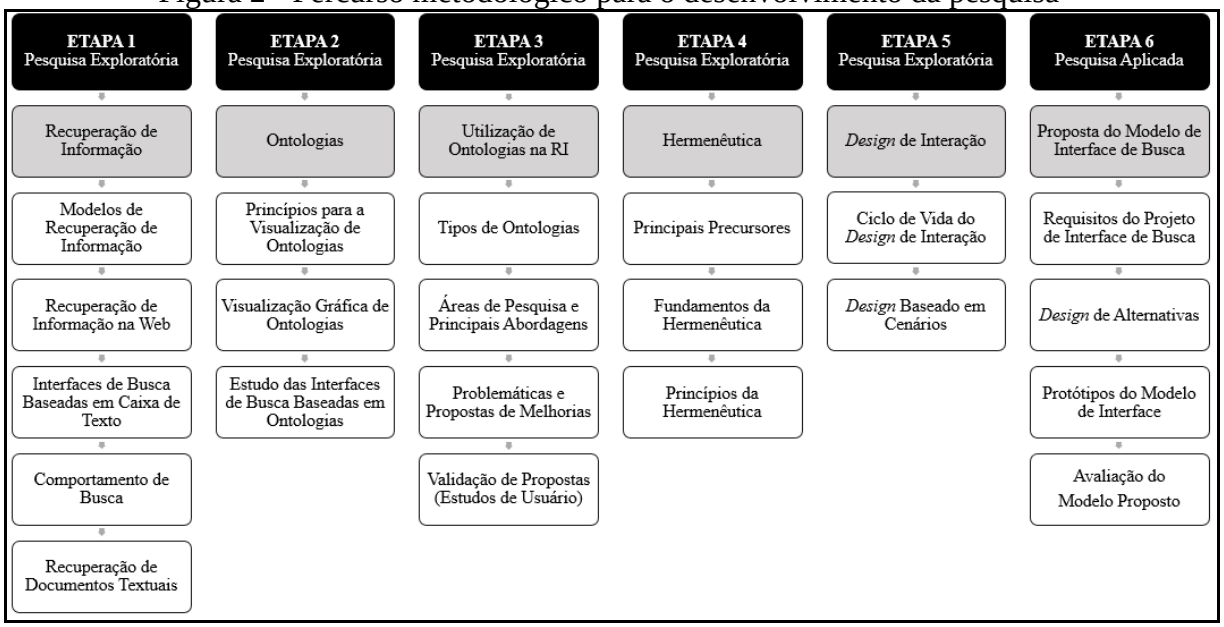
A Ciência da Informação forneceu os principais subsídios para o desenvolvimento desta pesquisa, que teve como principal objetivo apresentar contribuições e melhorias para a área. Desta forma, os estudos realizados em Recuperação de Informação, Interfaces de Busca, Comportamento de Busca e Ontologias, embora de natureza interdisciplinar por envolverem outras áreas como Ciência da Computação e Filosofia, foram

discutidos sob a perspectiva dos problemas observados no campo científico da Ciência da Informação. A área de Visualização da Informação, por sua vez, ofereceu fundamentos teóricos sobre a visualização de ontologias, discutindo princípios, elementos e recursos que foram adotados no projeto de interface de busca para facilitar a análise dos dados apresentados ao usuário.

Já a área da Filosofia, a partir dos estudos relacionados à Hermenêutica de Gadamer, deu suporte para que fosse possível entender o processo de interpretação documental, a partir do qual foi possível projetar recursos na interface para facilitar o desenvolvimento deste processo. E por último, com base na área de Ciência da Computação, foi possível entender as etapas necessárias para o desenvolvimento desta proposta, que foi projetada a partir do *Design* de Interação (DI), com a utilização de Personagens, Cenários, Prototipação e Testes de Usabilidade.

Para que os objetivos pudessem ser alcançados, o estudo foi pautado em pesquisa exploratória e aplicada, de acordo com as seis etapas apresentadas na Figura 2.

Figura 2 - Percurso metodológico para o desenvolvimento da pesquisa



Fonte: desenvolvido pela autora

As primeiras etapas do trabalho foram desenvolvidas a partir de pesquisa exploratória, de caráter bibliográfico, que permitiu a construção do embasamento teórico necessário para a definição do problema e para o levantamento dos requisitos que deram subsídios à proposta apresentada. Foram pautadas nas literaturas clássicas e atuais sobre a temática de estudo, com a realização de buscas em repositórios digitais, artigos *on-line*, dissertações, teses, revistas e periódicos científicos da área.

De acordo com Cervo, Bervian e Da Silva (2007), os estudos exploratórios possuem como finalidade contribuir para o entendimento dos conceitos e ideias que envolvem o tema, proporcionando maior familiaridade com o problema. Desta forma, contribuem para ampliar o entendimento do pesquisador acerca do fenômeno observado e permitir uma formulação clara do problema e das hipóteses que fundamentam a investigação.

As pesquisas de natureza aplicada buscam contribuir para fins práticos, com o propósito de buscar soluções para problemas específicos e concretos. Está voltada principalmente para o desenvolvimento de novos processos, produtos e tecnologias (CERVO; BERVIAN; DA SILVA, 2007; GIL, 2008).

Para facilitar a compreensão do percurso metodológico adotado, a pesquisa foi dividida em seis etapas principais, que são apresentadas detalhadamente no Quadro 1, juntamente com os objetivos e a principal contribuição de cada uma delas.

Quadro 1 - Percurso metodológico para o desenvolvimento da pesquisa

Etapas	**Objetivos**	**Contribuição**
RECUPERAÇÃO DE INFORMAÇÃO		
ETAPA 1 Pesquisa Exploratória	▪ Levantar conceitos e características da área de Recuperação de Informação ▪ Apresentar a área de Recuperação de Informação na Web, com foco no processo de busca ▪ Identificar as variáveis que interferem no comportamento de busca do usuário ▪ Analisar os principais problemas do processo de busca e recuperação de informação na Web • Analisar as principais recomendações presentes na literatura para a melhoria do processo de busca • Identificar as particularidades relacionadas à busca e recuperação de documentos textuais	Levantamento dos fatores que interferem no processo de busca e recuperação de informação
ONTOLOGIAS		
ETAPA 2 Pesquisa Exploratória	▪ Discutir conceitos e características das ontologias ▪ Apresentar os princípios para a visualização de ontologias ▪ Discutir os principais métodos e técnicas de visualização gráfica de ontologias ▪ Identificar as características das interfaces de busca baseadas em ontologias ▪ Analisar contribuições, perspectivas e desafios sobre a utilização de ontologias em interfaces de busca	Levantamento das contribuições do uso de ontologias no processo de busca e recuperação de informação
UTILIZAÇÃO DE ONTOLOGIAS NA RECUPERAÇÃO DE INFORMAÇÃO		
ETAPA 3 Pesquisa Exploratória	▪ Identificar os tipos de ontologias adotadas nas pesquisas ▪ Identificar as áreas de pesquisa e principais abordagens ▪ Analisar problemáticas e propostas de melhorias no âmbito do uso de ontologias ▪ Identificar validação de propostas por meio de estudos de usuário ▪ Analisar problemas que possam ser solucionados por meio do uso de ontologias na Recuperação de Informação	Definição do problema de pesquisa

Quadro 1 - Percurso metodológico para o desenvolvimento da pesquisa (continuação...)

Etapas	Objetivos	Contribuição
HERMENÊUTICA		
ETAPA 4 Pesquisa Exploratória	▪ Identificar os principais precursores da hermenêutica ▪ Identificar os fundamentos para a compreensão da hermenêutica ▪ Levantar os princípios da hermenêutica para a compreensão e interpretação documental	Identificação dos princípios da interpretação documental
***DESIGN* DE INTERAÇÃO**		
ETAPA 5 Pesquisa Exploratória	▪ Apresentar os fundamentos do *design* de interação para o planejamento e desenvolvimento de sistemas ▪ Analisar a relação do *design* de interação e demais abordagens ▪ Identificar as etapas do ciclo de vida básico do *design* de interação ▪ Identificar as ferramentas do *design* baseado em cenários	Definição das etapas e das ferramentas necessárias para o processo de desenvolvimento
MODELO DE INTERFACE DE BUSCA PROPOSTO		
ETAPA 6 Pesquisa Aplicada	▪ Apresentar as etapas para o desenvolvimento da proposta ▪ Identificar os requisitos do projeto da interface de busca ▪ Apresentar o *design* de alternativas e os protótipos da interface ▪ Descrever os cenários de busca que fundamentam o modelo proposto ▪ Avaliar o modelo proposto	Desenvolvimento do modelo de interface de busca proposto

Fonte: desenvolvido pela autora

A motivação inicial para o desenvolvimento deste estudo foi a possibilidade de projetar uma ferramenta que pudesse contribuir para a área de Recuperação de Informação. Para isto, todas as leituras da primeira etapa foram direcionadas para explorar o campo de RI e identificar os principais conceitos envolvidos na busca de informação, em especial, no contexto da Web e dos documentos textuais.

A principal contribuição desta etapa foi o levantamento dos problemas enfrentados pelo usuário durante o processo de busca, envolvendo características dos dados disponibilizados, a interação na formulação da expressão de busca e na interpretação dos resultados e aspectos relacionados ao comportamento do usuário, tais como: problemas de linguagem, questões culturais, variáveis demográficas e outros. Estas informações deram subsídios para a identificação de questões importantes para o projeto da interface, incluindo problemas da área, riscos envolvidos e ideias que pudessem trazer melhorias para o contexto da recuperação de informações digitais. Além disso, permitiu também o levantamento dos primeiros requisitos para o modelo proposto.

Na segunda etapa da pesquisa buscou-se identificar os conceitos e as características das ontologias, os principais métodos de visualização, além de contribuições, perspectivas e desafios sobre a utilização de ontologias em interfaces de busca. As informações levantadas nesta fase permitiram uma melhor compreensão sobre a importância das ontologias, bem como o seu potencial para o desenvolvimento de aplicações e recursos no

âmbito da Recuperação de Informação, indicando um possível caminho de investigação. Além disso, foi possível avaliar alguns projetos que já agregaram, de alguma forma, ontologias às interfaces de busca, com o intuito de levantar contribuições, perspectivas e desafios já apontados por outros autores. Parte das informações levantadas neste estágio da investigação serviu de base para a concepção de requisitos do "Modelo da Interface de Busca e Recuperação de Informação" proposto neste estudo.

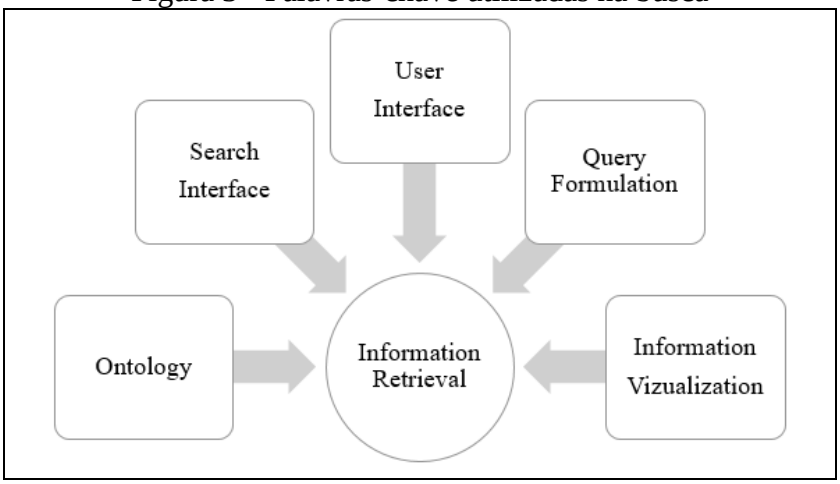
A partir dos resultados observados nos estudos preliminares desenvolvidos na fase anterior, a terceira etapa do estudo teve o objetivo de analisar, de forma aprofundada, como as ontologias estão sendo utilizadas no âmbito da busca e recuperação de informação e quais lacunas ainda podem ser exploradas e requerem aprofundamento de pesquisa. Para que este panorama fosse possível, foi realizada uma pesquisa nas publicações da ACM Digital Library (http://dl.acm.org/) e do Google Scholar (http://scholar.google.com.br/).

A ACM Digital Library foi escolhida por possuir uma coleção significativa de artigos publicados em revistas científicas e em eventos relacionados às áreas de Ciência da Computação e Tecnologia da Informação (TI). Além disso, conta com a colaboração de diversos grupos, envolvendo áreas que trazem contribuições significativas e intercâmbio de ideias para a temática, tais como: *Special Interest Group on Artificial Intelligence* (SIGAI), *Special Interest Group on Computer-Human Interaction* (SIGCHI), *Special Interest Group on Design of Communication* (SIGDOC), *Special Interest Group on Information Retrieval* (SIGIR), *Special Interest Group on Software Engineering* (SIGSOFT) e *Special Interest Group on Hypertext, Hypermedia and Web* (SIGWEB). O Google Scholar, por sua vez, foi escolhido por realizar buscas em materiais científicos selecionados, contendo publicações atualizadas e revisadas. Desta forma, entende-se que as fontes de pesquisa selecionadas possuem materiais relevantes, interdisciplinares e com abrangência suficiente para fundamentar o estudo proposto e oferecer uma visão holística da literatura já publicada.

As palavras-chave utilizadas para realizar a busca dos artigos de estudo são apresentadas na Figura 3. Elas envolvem basicamente três áreas: Ontologias, Interfaces de Busca (*Search Interface*, *User Interface*, *Query Formulation*) e Visualização da Informação. Inicialmente, adotou-se como termo de busca cada palavra isolada e, posteriormente, a combinação entre elas (duas, três e quatro palavras no mesmo termo), que foi realizada por meio da utilização do operador lógico (+). Também foi possível utilizar o recurso de navegação oferecido pelo *ACM Computing Classification System* para realizar buscas e avaliar se os termos adotados haviam recuperado todos os materiais de interesse. Neste ponto, foram analisadas duas categorias: 1) *Human-centered computing*, que apresenta estudos

relacionados com a área de *Human computer interaction* (HCI), *Interaction design* e *Visualization*, e 2) *Information systems*, que traz materiais sobre World Wide Web e *Information retrieval*.

Figura 3 - Palavras-chave utilizadas na busca



Fonte: desenvolvido pela autora

A pesquisa realizada retornou um total de 90 (noventa) artigos, que foram publicados entre os anos de 2000 a 2017. Posteriormente, cada material selecionado foi analisado, sendo que 28 (vinte e oito) deles foram descartados pois, embora relacionados ao tema, não adotavam ontologias em sua estrutura, que é o recurso base deste estudo. Também foram retirados da análise outros 03 (três) artigos por estarem diretamente relacionados ao processo de desenvolvimento de ontologias e, portanto, fora do escopo do trabalho. Os 59 (cinquenta e nove) artigos restantes, todos referentes à utilização de ontologias na área de recuperação de informação, foram analisados de acordo com os seguintes critérios: 1) tipo de ontologia adotada (genérica, domínio, tarefa, aplicação); 2) área de desenvolvimento da pesquisa e principais abordagens; 3) problemáticas e propostas de melhorias; 4) validação por meio de estudos de usuários.

Nesta perspectiva, foi possível traçar um panorama sobre como as ontologias estão sendo utilizadas para a melhoria da busca e recuperação de informação, bem como analisar quais são suas principais contribuições para a área e quais lacunas ainda não foram solucionadas e oferecem oportunidades de novas investigações. Partindo destes resultados, foi possível definir o problema de pesquisa desta tese, bem como as principais hipóteses levantadas para conduzir o desenvolvimento do estudo.

Com base no problema de pesquisa e nas hipóteses levantadas, a quarta etapa do estudo foi baseada no campo da Filosofia e teve como objetivo identificar elementos

que pudessem fornecer entendimento sobre o processo de interpretação e, a partir dele, projetar recursos que pudessem minimizar as dificuldades encontradas pelo usuário na compreensão de um documento ou de um *corpus* documental, relacionadas tanto à indexação dos materiais quanto à interpretação dos resultados apresentados. Entre os diversos caminhos avaliados, recorreu-se à Hermenêutica, com foco nos estudos de Gadamer, que ofereceu fundamentos necessários para que fosse possível entender a forma como o ser humano se comporta no contexto da compreensão e interpretação documental.

A quinta etapa do estudo foi fundamentada na literatura da Ciência da Computação e teve o intuito de apresentar o ciclo de vida de desenvolvimento de *software* proposto pelo *Design* de Interação, bem como as ferramentas que podem ser adotadas para documentar e subsidiar o processo. Além disso, também foram analisados os conceitos e as características do *Design* Baseado em Cenários, que é uma abordagem centrada no usuário, com a principal característica de documentar requisitos e decisões do projeto através de cenários e personagens.

A sexta e última etapa do estudo envolveu pesquisa de natureza aplicada e teve como objetivo apresentar o "Modelo de interface de busca e recuperação de informação". O processo de desenvolvimento do projeto foi fundamentado no *Design* de Interação e contou com um ciclo de vida composto de quatro estágios: 1) definição dos requisitos do projeto de interface de busca; 2) *design* de alternativas; 3) desenvolvimento dos protótipos da interface; 4) avaliação do modelo proposto.

As informações necessárias para definir os requisitos do sistema foram obtidas a partir da realização de entrevistas não estruturadas, realizadas com usuários de sistemas de busca de informação, além da análise da literatura existente sobre o assunto. Entre os dados obtidos, foi possível identificar as características do domínio de estudo, as oportunidades de melhoria, as funcionalidades do sistema, as particularidades do público alvo e as tarefas previstas na interface.

O *design* de alternativas teve como principal objetivo documentar as ideias do projeto elaboradas a partir dos requisitos definidos na fase anterior. Entre as atividades desenvolvidas nesta etapa, destacaram-se: a) definição do *layout*, estrutura de navegação e recursos gráficos adotados na interface; b) desenvolvimento dos personagens e dos principais cenários de busca.

Com base nas informações identificadas durante o levantamento de requisitos e nos elementos definidos no *design* de alternativas, foi possível projetar protótipos de baixa fidelidade e simular as funcionalidades previstas para a interface de busca. Os

protótipos foram elaborados com o auxílio da ferramenta de prototipação Pencil Project, versão 3.0.4 (https://pencil.evolus.vn), que é um *software* gratuito e de código aberto.

A partir da elaboração dos protótipos, foi possível avaliar o modelo proposto através de testes de usabilidade, com o intuito de analisar se os requisitos foram satisfeitos, bem como identificar a viabilidade de aceitação do usuário e possíveis erros de interação. Além disso, através dos testes realizados, pode-se avaliar as hipóteses levantadas no início da pesquisa e identificar quais foram aceitas e quais foram rejeitadas. Para a elaboração desta etapa, adotou-se o método de avaliação proposto por Snyder (2003).

Os testes de usabilidade foram elaborados com a utilização de protótipos de baixa fidelidade em papel e realizados com três participantes na fase de pré-teste e, posteriormente, com dez participantes, na fase dos ensaios, todos pesquisadores e potenciais usuários da interface projetada, com experiência na utilização de ferramentas de busca e no uso de repositórios digitais de informação científica. Para orientar o processo, foram adotados dois instrumentos: um roteiro de atividades sequenciais, previstas para cada cenário elaborado e um questionário de avaliação, composto por questões abertas, fechadas, escala de Likert (5 pontos) e diferencial semântico (7 pontos). Além disso, utilizou-se ainda a técnica de observação direta, que foi realizada durante os testes de usabilidade, com o intuito de observar o comportamento dos usuários durante as atividades propostas, bem como suas principais dificuldades na interação com a interface.

1.7 ESTRUTURA DO TRABALHO

Esta tese está organizada em oito capítulos. O primeiro capítulo referente à **Introdução** apresentou o delineamento da pesquisa, identificado o tema abordado, as contribuições do estudo, a caracterização do problema, as principais hipóteses levantadas, a proposição, a tese, os objetivos do estudo (geral e específicos) e os procedimentos metodológicos adotados para o desenvolvimento do trabalho.

O capítulo 2, denominado **Recuperação de Informação,** apresenta a área de Recuperação de Informação (RI), destacando os modelos booleano, vetorial e probabilístico, que são os três mais difundidos pela literatura. Também discute questões relacionadas à busca e recuperação de informação na Web, abordando fatores que interferem no comportamento do usuário, características de interação presentes no processo e especificidades relacionadas aos documentos textuais.

O capítulo 3, intitulado **Ontologias**, traz uma fundamentação teórica sobre

as ontologias, discutindo conceitos, características e contribuições para a área de recuperação de informação. Também discute a utilização de elementos gráficos, dos princípios de *Gestalt* e dos principais métodos e técnicas para a visualização de ontologias, com a intenção de apresentar elementos que auxiliem o usuário na análise e na extração de informações relevantes, seja no contexto da elaboração da expressão de busca ou da análise de resultados. Por fim, apresenta contribuições, perspectivas e desafios do uso de ontologias em sistemas de recuperação de informação.

O capítulo 4 apresenta um panorama sobre a **Utilização de Ontologias na Recuperação de Informação**, discutindo tipos de ontologias adotadas, áreas de aplicação, principais abordagens, propostas de melhorias, estudos validados por usuários e lacunas de pesquisa ainda existentes que pudessem ser empregadas neste estudo.

O capítulo 5, denominado **Hermenêutica**, apresenta uma breve revisão da literatura sobre a Hermenêutica, que é a área da Filosofia responsável pelo estudo dos fundamentos acerca da compreensão e interpretação de documentos textuais. A partir dos elementos identificados, foi possível definir funcionalidades para a ferramenta projetada neste estudo, visando ampliar a capacidade do usuário de organização, de análise e de interpretação de um *corpus* documental.

O capítulo 6 apresenta o ***Design* de Interação**, que é uma área fundamentada na Ciência da Computação e que discute padrões e ferramentas para o processo de desenvolvimento de produtos e sistemas. O conteúdo apresentado contribuiu para delinear o ciclo de vida de elaboração da proposta, bem como os procedimentos adotados para documentar e subsidiar as etapas previstas.

O capítulo 7, intitulado **Modelo de Interface de Busca Proposto**, apresenta detalhadamente o processo de desenvolvimento da proposta, com base nos fundamentos do *Design* de Interação. Discute os requisitos do projeto (domínio da aplicação, oportunidades de melhoria, funcionalidades da interface, público alvo e tarefas previstas), o *design* de alternativas (*layout*, estrutura de navegação, recursos gráficos, personagens e cenários) e os protótipos de baixa fidelidade projetados para atender os objetivos do usuário. Também discute a avaliação do modelo proposto realizada através de ensaios de usabilidade, bem como as contribuições da hermenêutica e das ontologias no suporte ao processo de recuperação de informação.

Por fim, o capítulo 8 apresenta as **Considerações Finais** do estudo realizado, a partir de uma síntese dos resultados da tese em relação aos objetivos propostos, destacando ainda as contribuições para a área de Recuperação de Informação, as limitações

encontradas e as sugestões de trabalhos futuros que poderão ser desenvolvidos nas próximas pesquisas sobre o tema.

2 RECUPERAÇÃO DE INFORMAÇÃO

Este capítulo apresenta uma revisão de literatura com o objetivo de discutir a Recuperação de Informação, destacando a história da criação do termo e sua interdisciplinaridade com a Ciência da Computação e a Ciência da Informação. Também discute cada um dos elementos pertencentes ao processo de RI e os três modelos mais difundidos pela literatura, chamados de modelos clássicos: booleano, vetorial e probabilístico. Além disso, traz uma discussão sobre a busca e recuperação de informação no contexto da Web, abordando fatores que interferem no comportamento do usuário, características de interação presentes no processo e especificidades relacionadas aos documentos textuais, que é o foco principal deste estudo.

■■■

O termo *Information Retrieval* (Recuperação de Informação) foi apresentado em 1951, pelo cientista da computação Calvin Mooers, que o definiu como um processo no qual um usuário pode converter sua necessidade de informação em uma lista real de citações de documentos armazenados, que contenham informações relevantes (MOOERS, 1951). A evolução da Recuperação de Informação está diretamente relacionada com o advento dos computadores, que passaram a ser utilizados para armazenar e recuperar grandes quantidades de informações (SALTON, 1985). Desta forma, "engloba tanto os aspectos intelectuais da descrição de informações e suas especificidades para a busca, como quaisquer sistemas, técnicas ou máquinas empregadas para o desempenho da operação" (MOOERS, 1951, p.25).

Por ser um termo originário da área da Ciência da Computação e estar diretamente relacionado com os avanços de *hardware* e *software*, principalmente os algoritmos referentes à pesquisa de texto, a Recuperação de Informação pode ser considerada a vertente tecnológica da Ciência da Informação (SARACEVIC, 1996, 2009), apresentando grande interdisciplinaridade entre as duas áreas. Enquanto a Ciência da Informação dedica-se à busca de "conhecimentos relacionados à origem, coleção, organização, armazenamento, recuperação, interpretação, transmissão, transformação, e utilização da informação" (BORKO, 1968, p.3), a Ciência da Computação atua no desenvolvimento de técnicas computacionais e algoritmos que possam viabilizar a implementação e aplicação desses conhecimentos. Saracevic (1996, p.50) demonstra esta relação, quando afirma que:

> A Ciência da Computação trata de algoritmos que transformam informações enquanto a CI trata da natureza da informação e sua comunicação para uso pelos humanos. Ambos os objetos são interrelacionados e não competidores, mas complementares.

Neste contexto, a Recuperação de Informação tem como objetivo principal pesquisar e desenvolver tecnologias que possibilitem o aprimoramento dos sistemas de recuperação de informação, de modo que os usuários possam obter acesso facilitado às informações, com resultados de buscas mais precisos e relevantes. É uma área de pesquisa direcionada para a representação, armazenamento, organização, busca e recuperação de informação, que pode estar contida em documentos, páginas Web, objetos multimídia, catálogos *on-line*, entre outros (SALTON, 1968; BAEZA-YATES; RIBEIRO-NETO, 2013).

Baeza-Yates e Ribeiro-Neto (2013) apresentam inúmeros estudos realizados em Recuperação da Informação a partir da década de 50, destacando contribuições importantes de cientistas que foram pioneiros e dedicaram grandes esforços para a área, tais como: Hans Peter Luhn, Eugene Garfield, Philip Bagley, Allen Kent, Gerard Salton, Karen Spark-Jones, Calvin Mooers, entre outros. Estes conhecimentos formataram a área e possibilitaram o desenvolvimento dos conceitos fundamentais que levaram às modernas tecnologias em RI, além de terem subsidiado a publicação dos primeiros materiais sobre o assunto: o primeiro livro sobre RI escrito por Salton (em 1968), articulações sobre a hipótese do agrupamento (em 1971) e o primeiro congresso sobre RI (em 1978). Desde então, muitos outros livros, artigos e materiais foram publicados sobre o assunto, possibilitando o crescimento e o fortalecimento da comunidade de estudiosos em RI.

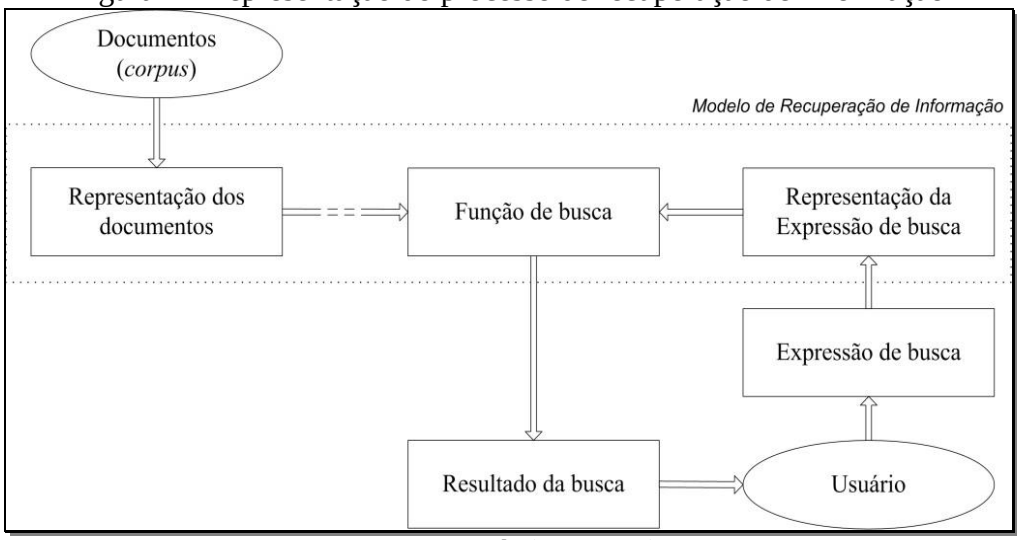
A expansão da Web, no início dos anos 90, também possibilitou um progresso significativo na área e atraiu a atenção de inúmeros pesquisadores, oriundos de áreas diversas. De um lado, houve um crescimento significativo de informações e o surgimento do maior repositório de conhecimento da história (BAEZA-YATES; RIBEIRO-NETO, 2013). Em contrapartida, como consequência imediata, os problemas de RI se multiplicaram e a dificuldade de encontrar uma informação relevante e de interesse do usuário, principalmente no contexto da Web, tornou-se um dos grandes problemas observados e discutidos no âmbito da Ciência da Informação (SARACEVIC, 1995, 1996; LIN; DEMNER-FUSHMAN, 2006).

> A Recuperação da Informação [*Information Retrieval*] é o campo de pesquisa central em Ciência da Informação [...]. O objetivo é estudar e entender os processos de Recuperação da Informação para projetar, construir e testar sistemas de recuperação que possam facilitar a comunicação efetiva

> de informação desejada entre o gerador humano e o usuário humano. (INGWERSEN, 1992, p.49, tradução nossa)

De acordo com Ferneda (2012, 2013), o processo de recuperação de informação pode ser representado conforme a Figura 4 e possui basicamente três elementos: a **representação dos documentos** pertencentes ao *corpus*, a **representação da expressão de busca** e a **função de busca**, que determinará o grau de similaridade entre os dois itens anteriores.

Figura 4 - Representação do processo de recuperação de informação



Fonte: Ferneda (2013, p.47)

A representação dos documentos, que no contexto da Web é realizada por meio de indexação automática, consiste na extração de termos (expressões linguísticas) que têm como objetivo representar, de forma resumida, o seu conteúdo informacional e possibilitar a sua posterior recuperação. Em ambientes de Recuperação de Informação menores que a Web (bibliotecas digitais, por exemplo), a tarefa de representação dos documentos (ou indexação) é realizada por especialistas com a finalidade de atribuir os identificadores relevantes para cada elemento de informação (SALTON, 1980). Tais termos ou identificadores definem os pontos de acesso para a busca e impactam diretamente na eficiência do processo de recuperação de informação (ROWLEY, 2002; FERNEDA, 2013).

Por outro lado, com o objetivo de interagir com o sistema e recuperar documentos relevantes, o usuário tenta expressar linguisticamente suas necessidades de informação, utilizando palavras-chave e termos relacionados com o assunto que deseja pesquisar. Este diálogo entre o usuário e o sistema só é possível por meio de uma interface de busca.

> A fim de ser efetivo em sua tentativa de satisfazer a necessidade de informação do usuário, o sistema de RI deve, de alguma forma, 'interpretar' o conteúdo dos itens de informação, isto é, dos documentos de uma coleção, e classificá-los de acordo com o grau de relevância à consulta do usuário (BAEZA-YATES; RIBEIRO-NETO, 2013, p.4).

Para que essa classificação seja realizada de forma eficiente e retorne um maior número de documentos relevantes, é necessário que exista uma similaridade entre a expressão de busca utilizada pelo usuário e a representação dos documentos pertencentes ao *corpus*. No entanto, os sistemas de RI atuais, principalmente os disponibilizados na Web, não fornecem mecanismos que indiquem quais termos foram utilizados para representar o acervo documental, atribuindo ao usuário a tarefa de testar inúmeras expressões durante a busca até que encontre os resultados esperados. Além disso, são ambientes fortemente influenciados pelos problemas linguísticos, semânticos e sintáticos, o que dificulta ainda mais a tradução das necessidades de informação do usuário em expressões ou palavras-chave.

A função de busca, por sua vez, é responsável pela análise do grau de similaridade existente entre a expressão de busca formulada pelo usuário e cada um dos documentos pertencentes ao *corpus*. É por meio dela que o sistema realiza a recuperação e o ranqueamento, retornando uma lista de documentos que possuem mais probabilidade de serem relevantes para o usuário. Para Baeza-Yates e Ribeiro-Neto (2013, p.21), "esse ranqueamento é possivelmente o processo mais importante em um sistema de RI", uma vez que existe um grau de incerteza e imprecisão devido à subjetividade do conceito de relevância, ou seja, o que é relevante para um usuário pode não ser para o outro, mesmo que as buscas sejam realizadas de forma idêntica.

> Esse algoritmo preditivo é essencialmente a função de ranqueamento utilizada para estabelecer um ordenamento dos documentos recuperados. Documentos que aparecem no topo deste ordenamento são considerados como tendo mais chance de serem relevantes (BAEZA-YATES; RIBEIRO-NETO, 2013, p.22).

A eficiência de um algoritmo preditivo de um sistema de recuperação de informação está diretamente ligada ao modelo que ele utiliza. Ao longo dos anos, uma variedade de modelos foi proposta, oferecendo diferentes estruturas de representação de documentos e de princípios para a construção da função de ordenação. Baeza-Yates e Ribeiro-Neto (2013) classificam os modelos disponíveis na literatura em cinco grupos:

a) clássicos: booleano, vetorial e probabilístico;
b) alternativos da teoria de conjuntos: baseado em conjuntos, booleano

estendido, conjunto *fuzzy*;

c) algébricos alternativos: vetorial generalizado, modelo de indexação semântica latente, redes neurais;
d) probabilísticos alternativos: BM25, modelos de linguagem, divergência da aleatoriedade, redes *bayesianas*;
e) outros modelos (modelo de hipertexto, baseados na Web, recuperação de texto estruturado, recuperação multimídia, busca corporativa e vertical.

Como pode ser observado, existem inúmeros modelos que empregam estratégias de recuperação de informação diferenciadas. Em função da sua simplicidade e relevância para a área de RI, os modelos clássicos (booleano, vetorial e probabilístico) serão o foco deste estudo. Os demais são variações destes três primeiros, nos quais foram incorporados métodos alternativos, tais como: uso de probabilidade, funções polinomiais, regressão logística, ranqueamento probabilístico, lógica *fuzzy*, escalamento multidimensional, redes neurais e outros recursos na tentativa de minimizar as limitações encontradas e melhorar os resultados do processo de recuperação de informação.

2.1 MODELOS DE RECUPERAÇÃO DE INFORMAÇÃO

Esta seção tem como objetivo apresentar as características dos três principais modelos de recuperação de informação presentes na literatura. É de grande importância compreender o seu funcionamento, uma vez que eles subsidiaram o desenvolvimento de diversos outros modelos propostos ao longo dos anos e ainda fundamentam os sistemas de RI atuais e os mecanismos de busca da Web.

2.1.1 Modelo Booleano

O modelo booleano foi o primeiro modelo de recuperação de informação proposto e provavelmente, o mais criticado. É baseado na teoria dos conjuntos e na álgebra booleana e, em função da sua simplicidade, é um dos mais fáceis de ser implementado. Foi adotado por muitos dos primeiros sistemas bibliográficos comerciais (HIEMSTRA, 2009; BAEZA-YATES; RIBEIRO-NETO, 2013).

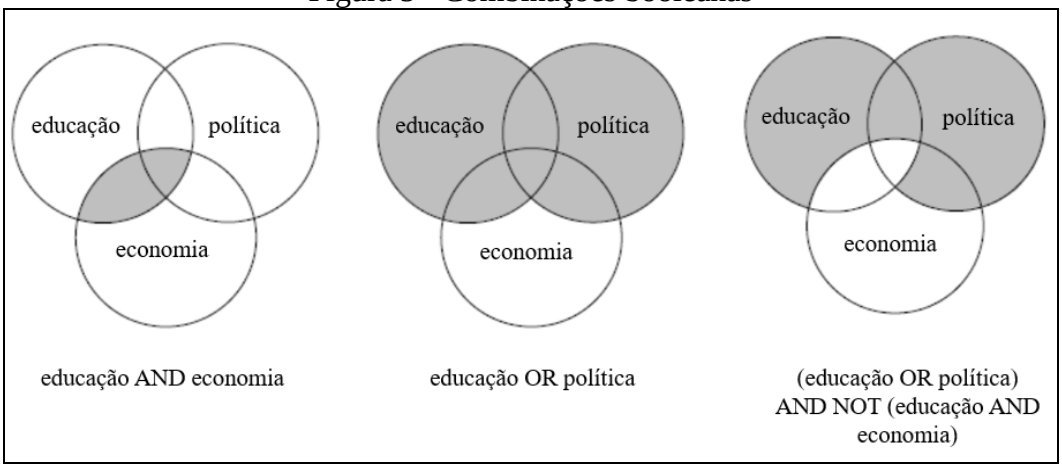
Em um sistema booleano, a relevância é obtida através da aplicação dos três operadores ou funções lógicas: E (AND), OU (OR) e NÃO (NOT). Os documentos são

representados por meio de um conjunto de termos indexados e, desta forma, são recuperados aqueles que contêm os termos que satisfaçam à expressão lógica da consulta, não existindo satisfação parcial, ou seja, os documentos são considerados **relevantes** ou **não relevantes**, dependendo da presença ou ausência dos termos de busca nos documentos (FERNEDA, 2003; BAEZA-YATES; RIBEIRO-NETO, 2013).

> Termos e operadores booleanos podem ser combinados para especificar buscas mais detalhadas ou restritivas. Como a ordem de execução das operações lógicas de uma expressão influência no resultado da busca, muitas vezes é necessário explicitar essa ordem delimitando partes da expressão por meio de parênteses. A definição de expressões complexas exige um conhecimento profundo da lógica booleana. O conhecimento da lógica booleana é importante também para entender e avaliar os resultados obtidos em uma busca (FERNEDA, 2013, p.49).

A Figura 5 apresenta, resumidamente, por meio do Diagrama de Venn, o funcionamento das expressões booleanas AND, OR e NOT. Para cada consulta realizada, são recuperados todos os documentos que possuem os termos nas condições especificadas pelo usuário. Os resultados são representados pela respectiva parte sombreada (áreas destacadas em cinza) de cada diagrama abaixo.

Figura 5 - Combinações booleanas



Fonte: Adaptado de Hiemstra (2009, p.5)

O primeiro diagrama apresenta o funcionamento da expressão booleana AND, exemplificando uma busca realizada com os termos ***educação AND economia***. Neste caso, o resultado da busca será o subconjunto de documentos indexados com a interseção de ambos os termos, ou seja, aqueles que contenham obrigatoriamente ***educação E economia***. É, portanto, uma função restritiva, ou seja, quanto maior o número de termos utilizados na busca, menores serão os resultados encontrados no *corpus* documental.

O segundo modelo apresenta uma consulta realizada com o operador booleano OR, contendo os termos ***educação OR política***. Neste caso, os documentos recuperados seriam aqueles indexados ou pelo termo ***educação OU*** pelo termo ***política***, ou seja, apresenta como resposta a união de ambos os conjuntos. Assim, o conjunto de documentos recuperados é sempre maior do que o conjunto de documentos de qualquer um dos termos individuais.

O terceiro diagrama, por sua vez, apresenta o resultado da busca realizada com a expressão booleana NOT. Neste caso, o operador booleano NOT define os termos dos documentos que NÃO devem ser incluídos nos resultados, ou seja, do subconjunto de resultados retornados na busca ***educação OR política***, não devem ser incluídos aqueles que possuem os termos ***educação AND economia***. Este operador pode ser visto como um operador de diferença de conjuntos, uma vez que subtrai elementos de um conjunto de termos indexados (FERNEDA, 2003, 2013; HIEMSTRA, 2009; BAEZA-YATES, RIBEIRO-NETO, 2013).

Além da utilização das expressões booleanas, o sistema ainda possibilita a combinação de termos e operadores para elaborar consultas mais complexas e melhorar resultados da busca, restringindo ou elevando a quantidade de documentos recuperados (BEPPLER, 2002).

Hiemstra (2009) e Baeza-Yates e Ribeiro-Neto (2013) destacam algumas vantagens referentes ao modelo booleano, tais como o formalismo claro por trás do modelo e sua simplicidade. Em função disso, permite que os usuários tenham um maior controle sobre o sistema, uma vez que possibilita saber exatamente por que um documento foi recuperado para uma determinada consulta.

Em contrapartida, o modelo possui algumas limitações, sendo que a principal delas é o fato de não existir um mecanismo que permita o ranqueamento dos documentos recuperados, dificultando a análise dos resultados da busca. Além disso, a formulação das consultas booleanas exige um conhecimento expressivo sobre a Lógica de Boole, sendo uma barreira para a maioria dos usuários. Outra desvantagem sobre o modelo booleano está no fato de selecionar os documentos apenas de forma binária, ou seja, eles são analisados apenas sob o critério dualista **relevante** ou **não relevante**, não sendo apresentada nenhuma outra espécie de ordenação de resultados que atendam às especificações da consulta.

2.1.2 Modelo Vetorial

O modelo vetorial ou também denominado de modelo espaço vetorial foi proposto em 1968 por Gerard Salton (SALTON; BUCKLEY, 1988). Sua principal característica está na representação dos termos da consulta e do documento como vetores em um espaço n-dimensional, tendo um peso associado a cada termo de indexação. Sua estrutura é composta por um espaço n-dimensional, contendo operações de álgebra linear aplicáveis aos vetores.

Os sistemas baseados no modelo vetorial, diferentemente dos booleanos, não são binários e utilizam termos e pesos na avaliação do grau de similaridade entre os documentos e a consulta do usuário. Os documentos recuperados são ordenados de forma decrescente, de acordo com o grau de similaridade, incluindo os que combinam parcialmente com os termos da consulta, com o objetivo de construir um *ranking*. De acordo com Ferneda (2003, p.30), cada elemento do vetor é normalizado, assumindo valores entre zero (0) e um (1), sendo que os pesos mais próximos de um (1) indicam termos com maior importância para a descrição do documento. Da mesma forma, uma expressão de busca também é representada por um vetor numérico, no qual cada elemento representa o grau de relevância do termo. O cálculo do grau de similaridade é realizado de acordo com a seguinte fórmula:

$$sim(x,y) = \frac{\sum_{i=1}^{t}(w_{i,x} \times w_{i,y})}{\sqrt{\sum_{i=1}^{t}(w_{i,x})^2} \times \sqrt{\sum_{i=1}^{t}(w_{i,y})^2}}$$

Onde:

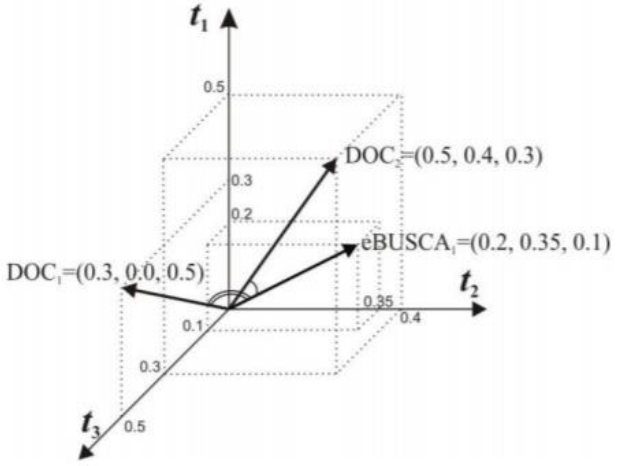
- ***x*** e ***y*** são os vetores;
- ***t*** é o número total de documentos considerados;
- $w_{i,x}$ é o peso do i-esimo elemento do vetor x;
- $w_{i,y}$ e o peso do i-esimo elemento do vetor y;
- ***sim*** é a função de similaridade.

Para atribuir uma pontuação numérica a um documento para uma consulta, o modelo mede a similaridade entre o vetor de consulta e o vetor do documento. A semelhança entre os dois é utilizada como medida de disparidade entre os vetores e o cosseno do ângulo é usada como semelhança numérica. Desta forma, “um *corpus* contendo um número indefinido de documentos e termos de indexação pode ser representado através de uma matriz onde cada linha representa um documento e cada coluna representa a associação de um determinado termo aos diversos documentos” (FERNEDA, 2003, p.29-30).

A Figura 6 apresenta o processo de busca no modelo vetorial, ilustrando a representação dos documentos DOC_1 (0.3, 0.0, 0.5) e DOC_2 (0.5, 0.4, 0.3), bem como a representação da expressão de busca $eBUSCA_1$ (0.2, 0.35, 0.1).

Figura 6 - Representação de uma busca no modelo vetorial

	t_1	t_2	t_3
$eBUSCA_1$	**0.2**	**0.35**	**0.1**
DOC_1	**0.3**	**0.0**	**0.5**
DOC_2	**0.5**	**0.4**	**0.3**



Fonte: Ferneda (2012, p.34)

Desta forma, o modelo vetorial é capaz de ordenar os documentos de acordo com o grau de similaridade de cada um deles em relação à consulta elaborada pelo usuário, ou seja, ao invés de determinar se um documento é relevante ou não, ele o ordena de acordo com o grau de similaridade, possibilitando que a recuperação ocorra mesmo que não satisfaça totalmente a consulta (MANNING; RAGHAVAN; SCHÜTZE, 2008; BAEZA-YATES; RIBEIRO-NETO, 2013).

O modelo vetorial destaca-se por ser a base da maioria dos sistemas de recuperação de informação, principalmente os voltados para a Internet, que agregam ainda outras técnicas para determinar o *ranking* dos documentos. Como principais vantagens, Baeza-Yates e Ribeiro-Neto (2013) enfatizam que:

a) o uso de termos de índices ponderados contribui para a melhoria do desempenho da recuperação de informações;
b) sua estratégia de combinação permite a recuperação de documentos parcialmente relevantes que se aproximam das condições de consulta;
c) após o cálculo da correlação entre o documento e a consulta, é possível ordenar os resultados obtidos conforme o grau de similaridade;
d) possui uma boa estratégia de ranqueamento para coleções genéricas;
e) é um modelo simples e rápido.

Apesar das inúmeras vantagens, Ferneda (2003) destaca que uma importante

limitação do modelo vetorial está no fato de não permitir a formulação de consultas booleanas, o que pode restringir consideravelmente a sua flexibilidade. Além disso, Baeza-Yates e Ribeiro-Neto (2013, p.47) ainda comentam sobre a desvantagem de o modelo vetorial considerar os termos de indexação mutuamente independentes, ou seja, não levar em conta os relacionamentos entre eles. Apesar de destacarem como uma desvantagem, afirmam que, na prática, "considerar as dependências entre termos é desafiador e pode levar a resultados ruins se não for feito de maneira adequada.

2.1.3 Modelo Probabilístico

As primeiras ideias sobre a recuperação de informação probabilística foram apresentadas por Maron e Kuhns, em um artigo publicado em 1960 (MARON; KUHNS, 1960). No entanto, após o estudo de diversos outros pesquisadores, o modelo probabilístico foi proposto em 1976 por Robertson e Sparck Jones e tem como base a teoria das probabilidades (Teorema de *Bayes*) para definir as representações para os documentos e as consultas. Desta forma, é um modelo no qual a recuperação é vista como um problema de estimativa de probabilidade de que a representação de um documento corresponda ou satisfaça a representação de uma consulta, ou seja, ele supõe que exista um conjunto ideal de documentos que atenda a cada uma das possíveis buscas que possam ser feitas pelos usuários (SOUZA, 2006; BAEZA-YATES; RIBEIRO-NETO, 2013).

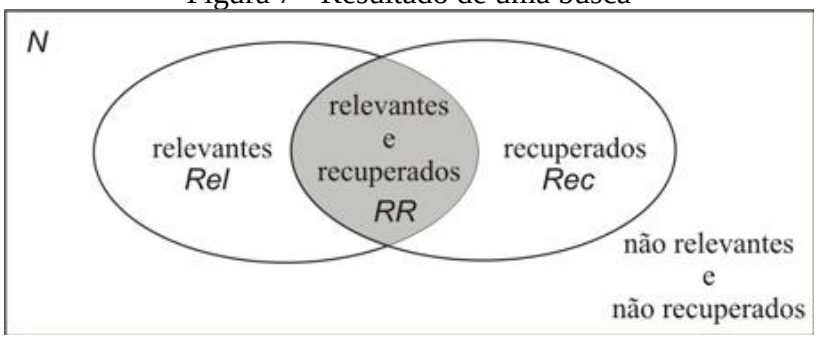
De acordo com Rijsbergen (1979) e Salton (1989), o modelo probabilístico normalmente calcula a probabilidade condicional de um determinado documento ser relevante ou não para uma consulta realizada. Assim, para cada consulta, existe um conjunto de documentos que são considerados relevantes, ou seja, uma resposta ideal para satisfazer as necessidades de informação do usuário. Para determinar este conjunto de documentos relevantes é necessário, no entanto, definir algumas características que possam identificá-los.

Para que isto seja possível, o sistema realiza uma interação com o usuário, por meio do processo de *relevance feedback*, por meio do qual determina a relevância dos documentos e busca a melhoria dos resultados através de cálculos de probabilidade. Deste modo, o usuário visualiza um conjunto de documentos retornados inicialmente pelo sistema (que pode ser selecionado por meio de outras técnicas de outros modelos, como o vetorial) e, a partir deles, decide quais são relevantes e quais não são. Várias interações poderão ser realizadas, com o objetivo de refinar os resultados e encontrar um conjunto ideal de documentos relevantes (BAEZA-YATES; RIBEIRO-NETO, 2013).

> Dada uma consulta de usuário *q* e um documento d_j da coleção, o modelo probabilístico tenta estimar a probabilidade do usuário achar o documento d_j interessante (isto é, relevante). O modelo supõe que essa probabilidade de relevância depende apenas das representações da consulta e do documento, ou seja, das informações disponíveis ao sistema. Além disso, o modelo supõe que exista um conjunto de todos os documentos que o usuário prefira como conjunto resposta para a consulta *q*. Tal conjunto resposta ideal é chamado de *R* e de maximizar a propriedade de relevância para o usuário (BAEZA-YATES; RIBEIRO-NETO, 2013, p.48).

Para Ferneda (2003), pode-se dividir o *corpus* (com *N* documentos) em quatro subconjuntos distintos, que são retornados a partir de uma expressão de busca: ***Rel***: representa o conjunto dos documentos relevantes, ***Rec***: conjunto dos documentos recuperados; ***RR***: conjunto dos documentos relevantes que foram recuperados e conjunto de documentos não recuperados e não relevantes, conforme pode ser observado na Figura 7.

Figura 7 - Resultado de uma busca



Fonte: Ferneda (2003, p.39)

Os documentos considerados pelo sistema como ideais e de interesse do usuário são os que fazem parte da interseção entre os relevantes (Rel) e os recuperados (Rec), pertencentes ao grupo dos relevantes e recuperados (RR).

Baeza-Yates e Ribeiro-Neto (2013) destacam vantagens e desvantagens para o modelo probabilístico. Segundo eles, a maior vantagem está no fato de que os documentos são ranqueados de forma decrescente, de acordo com a sua probabilidade de serem relevantes para a consulta do usuário. Ferneda (2003) acrescenta ainda a importância do processo de *relevance feedback*, uma vez que o modelo reconhece que a atribuição da relevância é uma tarefa do usuário. Além disso, enfatiza ainda a possibilidade de integrar as vantagens dos modelos probabilístico e vetorial e, com isso obter os benefícios de ambos, buscando melhorias no processo de recuperação de informação.

Em contrapartida, Baeza-Yates e Ribeiro-Neto (2013) destacam algumas desvantagens do modelo probabilístico, tais como:

a) necessidade de estimar incialmente o conjunto de documentos relevantes e não relevantes;
b) o método não considera a frequência na qual um termo de indexação em um documento, sendo que todos os pesos são binários;
c) falta de normalização pelo tamanho dos documentos.

Jones, Walker e Robertson (2000) ainda destacam que os resultados realizados com alguns experimentos utilizando o modelo probabilístico, em contextos limitados, com pequena quantidade de documentos, demonstraram resultados pouco superiores ao booleano. No entanto, espera-se que no contexto heterogêneo da Web, esses resultados sejam melhores e mais expressivos. Sua complexidade e dificuldade de implementação desencoraja muitos desenvolvedores de sistemas, restringindo o desenvolvimento de novos protótipos e a avaliação real da sua *performance* de recuperação e informação.

2.2 RECUPERAÇÃO DE INFORMAÇÃO NA WEB

Os sistemas de recuperação de informação na Web têm como principal objetivo "minimizar as dificuldades do usuário em localizar a informação requisitada, ou seja, diminuir o tempo gasto em um processo de busca até que a informação desejada possa ser acessada" (CRISTOVÃO; DUQUE; SERQUEIRA, 2012, p.5-6). Também são responsáveis por "recuperar todos os documentos que são relevantes à necessidade de informação do usuário e, ao mesmo tempo, recuperar o menor número possível de documentos irrelevantes" (BAEZA-YATES; RIBEIRO-NETO, 2013, p.4). São de grande importância para a sociedade contemporânea e organizações em geral, que dependem de documentos e informações para a tomada de decisão, resolução de problemas e construção de novos conhecimentos (BEPPLER, 2008).

Apesar da sua relevância, os buscadores disponíveis atualmente, a maioria baseados nas caixas de texto, apresentam inúmeros problemas e ainda estão distantes de atender às expectativas dos usuários, uma vez que é necessário despender um tempo significativo na tentativa de localizar uma informação desejada e, quando as expressões de buscas são mal formuladas, retornam uma grande quantidade de documentos, atribuindo ao usuário a árdua tarefa de analisar e determinar se eles são ou não relevantes (LAI; SOH, 2004; WOODS, 2004; ALLAN; CARTERETTE; LEWIS, 2005; HOGAN et al., 2011).

Os problemas relacionados à recuperação de informação podem ser

observados a partir de inúmeras perspectivas: relacionados à qualidade e variedade dos dados, relacionados à interação dos usuários com os sistemas e relacionados ao comportamento de busca. Entre os inúmeros aspectos que geram dificuldades e requerem atenção especial dos pesquisadores, Huang (1999), Baeza-Yates e Ribeiro-Neto (2013) e Nayak, Prasad e Senapati (2015) destacam:

a) grande volume de dados: a quantidade de informações disponibilizada é muito ampla, heterogênea e se modifica com muita frequência;
b) dados heterogêneos: os dados são disponibilizados em uma grande variabilidade de formatos (textual, imagens, vídeos, etc.) e idiomas (inglês, português, etc.). Além disso, os conteúdos presentes em um documento podem também serem heterogêneos, podendo conter texto, áudio, vídeo e imagens;
c) dados redundantes: a quantidade de documentos duplicada é volumosa, ou seja, a mesma informação pode ser encontrada em diversas fontes e elas, na maioria das vezes, não estão correlacionadas;
d) dados não estruturados: grande parte dos dados da Web são disponibilizados e duplicados de forma livre, sem a adoção de qualquer tipo de estrutura, linguagem ou padrão de interoperabilidade;
e) alto percentual de dados voláteis: as informações disponibilizadas podem ser adicionadas, removidas ou alteradas com grande facilidade, gerando problemas com *links* suspensos (ou quebrados), mudança de nomes de arquivos e relocação de domínios;
f) qualidade dos dados: muitos dados disponibilizados na Web apresentam problemas relacionados à qualidade, podendo ser imprecisos, obsoletos, inválidos, mal escritos (erros ortográficos, gramaticais), inocentes ou maliciosos.

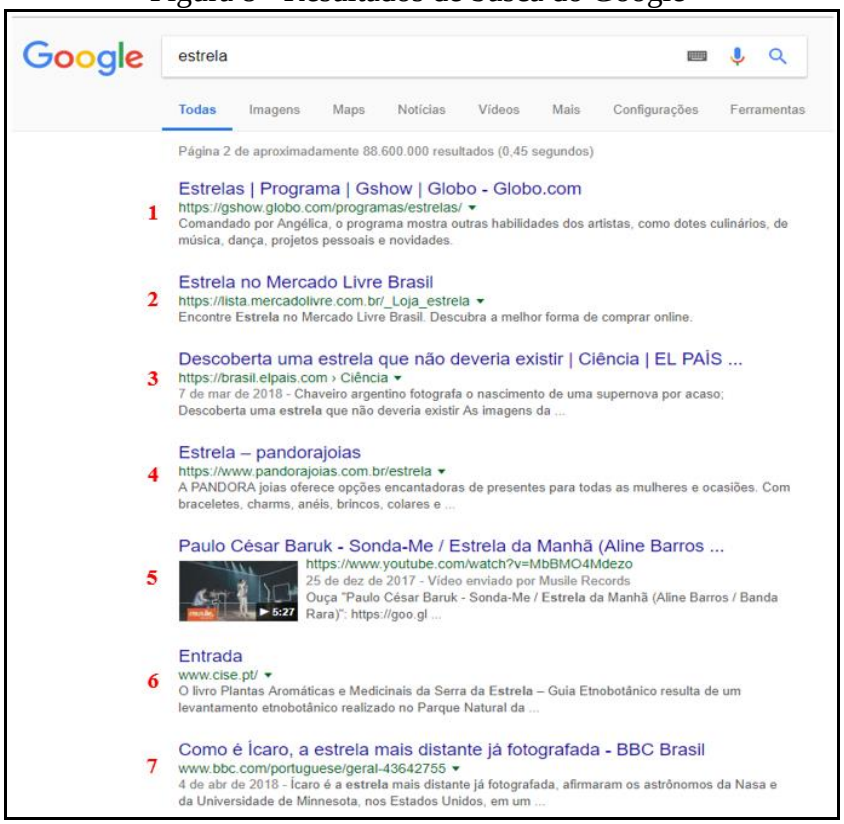
Além dos problemas relacionados à qualidade e variedade de dados disponibilizados, a recuperação de informação na Web também apresenta dificuldades referentes à interação dos usuários com os sistemas, que se manifestam principalmente em dois aspectos: na elaboração da expressão de busca e na interpretação dos resultados recuperados (BAEZA-YATES; RIBEIRO-NETO, 2013). É importante considerar que a recuperação de informação é um processo de comunicação complexo, que ocorre entre o usuário e um acervo documental, realizado por meio de uma interface, com o objetivo principal de satisfazer as necessidades informacionais do usuário (SARACEVIC, 1997;

ARAUJO JUNIOR, 2007).

Para que esta comunicação aconteça de forma satisfatória e que o sistema retorne uma informação parcial ou totalmente relevante ao contexto pesquisado, é necessário que o usuário estabeleça um diálogo com o sistema, na tentativa de expressar suas necessidades, que geralmente são traduzidas no formato de palavras-chave ou expressões de busca. A eficiência do processo de recuperação de informação está relacionada às coincidências terminológicas entre os termos utilizados na expressão de busca e os termos de indexação empregados para representar os documentos.

Assim, tal eficiência depende do conhecimento que o usuário possui da terminologia da área ou assunto que está pesquisando, permitindo-lhe expressar com mais precisão, a sua necessidade de informação. Quanto maior o seu conhecimento em relação ao domínio de interesse, mais opções de vocabulário o usuário terá à sua disposição para enriquecer as buscas, ao mesmo tempo que, quanto menor conhecimento possuir, mais genéricas e inadequadas serão as expressões utilizadas (ANTONIOU; HARMELEN, 2004; BEPPLER, 2008). Como exemplo, a Figura 8 demonstra parte deste processo de comunicação, apresentando os resultados de uma busca realizada no Google a partir do termo **estrela**.

Figura 8 - Resultados de busca do Google



Fonte: Google (2018)

O primeiro ponto que deve ser observado é o número excessivo de resultados retornados na busca (88.600.000), gerando uma grande sobrecarga de informação e tornando praticamente impossível a análise de todo o *corpus* de resultados. Em função disso, os usuários apresentam dificuldade para selecionar os elementos mais importantes e restringem-se, na maioria das vezes, em avaliar apenas as primeiras páginas de resultado, na confiança de que elas apresentam as informações mais relevantes (HUANG, 1999; SPINK et al., 2001).

Além da quantidade excessiva de resultados, a ausência de informações que indiquem a relação existente entre a pesquisa realizada e os resultados recuperados, bem como se existem semelhanças entre eles, também dificulta o processo de busca (HOBBS; PFITZNER; POWERS, 2002). Observando superficialmente os resultados apresentados na Figura 8, sem a realização de uma análise detalhada, algumas perguntas ficam sem resposta:

a) O resultado representado pelo item 3 possui alguma relação com o item 7, apesar de serem de fontes diferentes?
b) Os itens 3 e 7 possuem informações (semelhantes ou divergentes) que possam fundamentar um estudo ou ponto de vista?
c) Os itens 3 e 7 estão relacionados a partir de alguma perspectiva (assunto, autor, fundamentação, problema)?
d) Os itens 2 e 4 referem-se à mesma loja de joias?

Outro problema presente no processo de recuperação de informação discutido acima, refere-se ao contexto da busca, uma vez que a palavra **estrela** apresenta diversos significados. Segundo Ferreira (2004), o termo pode ser empregado a pelo menos três contextos diferenciados: 1) nome que indica astros luminosos; 2) artista de renome; 3) pessoa que se sobressai em determinada área. Além disso, o termo ainda é atribuído a outros elementos, tais como: fabricante de brinquedos, músicas, nomes de empresas, times de futebol, nomes de ruas etc. Neste cenário, é impossível saber qual significado corresponde à intenção de busca do usuário, a menos que o campo semântico (contexto) ao qual se refere seja inserido ou especificado. Como consequência, um grande número de resultados irrelevantes é apresentado, dificultando a análise daqueles que são realmente importantes.

Por ser um ambiente fortemente influenciado pelos problemas linguísticos semânticos e sintáticos, os usuários, muitas vezes, apresentam dificuldades para traduzir suas necessidades de informação em um conjunto de termos de busca. Segundo Bräscher (2002), problemas relacionados ao significado das palavras afetam a qualidade da recuperação de informação e interferem diretamente nos resultados apresentados. O uso de poucas palavras

para formular consultas pode não fornecer contexto suficiente para diferenciar adequadamente os significados opostos das palavras constituintes.

> Efetuar buscas em grandes repositórios de informações não estruturadas e não padronizadas torna-se uma tarefa árdua, levando muitas vezes a ambiguidades, a conteúdos fora de contexto e até mesmo a não recuperação da informação desejada. As ferramentas de busca, ao fazerem uso da linguagem natural, necessitam de conhecimento sobre o significado das expressões utilizadas e das relações entre elas, o que possibilita contextualização e tratamento de fenômenos linguísticos que afetam a qualidade da recuperação (ALTOUNIAN; GOMES, 2016, p.138).

Bräscher (2002) afirma que a ambiguidade ocorre quando palavras ou frases podem gerar mais de uma interpretação de seu significado. O campo da semântica, tradicionalmente explorado na linguística e na filosofia, é responsável pelos estudos referentes ao significado ou ao processo de significação. Neste contexto, quatro propriedades básicas destacam-se na semântica e afetam diretamente as buscas realizadas pelos usuários: sinonímia, homonímia, paronímia e polissemia.

De acordo com Bregagnol e Lara (2012) e Altounian e Gomes (2016), estes elementos podem ser definidos como:

a) Sinonímia é um dos temas mais abordados no campo da semântica e estuda a relação entre duas palavras ou mais que apresentam significados iguais ou semelhantes, ou seja, os sinônimos. Exemplos: ensinamento, lição, experiência / garota, menina / débil, fraco, frágil / distante, afastado, remoto;
b) Homonímia refere-se à relação entre duas ou mais palavras que, apesar de possuírem significados diferentes, possuem a mesma pronúncia e grafia semelhante. Alguns exemplos são: concerto (sessão musical), conserto (corrigir) / cessão (ceder), sessão (reunião), seção (repartição).
c) Paronímia estuda a relação existente entre duas ou mais palavras que possuem significados diferentes, mas são muito parecidas na pronúncia e na escrita. Exemplos: eminente (alto), iminente (próximo) / absolver (perdoar), absorver (assimilar) / geminada (duplicada), germinada (que germinou) / descrição (descrever), discrição (discreto).
d) Polissemia é a propriedade que uma mesma palavra tem de apresentar vários significados. Exemplos: manga (fruta), manga (parte de uma peça do vestuário) / letra (símbolo utilizado na escrita), letra (música ou

canção) / banco (local de realização de operações financeiras), banco (assento).

Além das propriedades apresentadas, a semântica estuda ainda a denotação e a conotação de palavras. Denotação é o uso da palavra com o seu sentido original (exemplos: manga como fruta; boca como o primeiro órgão do aparelho digestivo). A conotação, por sua vez, é a propriedade que uma palavra tem de ampliar-se no seu campo semântico, dentro de um contexto, gerando várias possibilidades interpretativas (exemplos: manga de camisa, boca da noite, coração de pedra).

Outro problema observado no processo de comunicação entre o usuário e uma interface, seja para elaborar uma expressão de busca ou interpretar os resultados apresentados, refere-se às variações linguísticas, que podem estar relacionadas a diversos fatores, incluindo as alterações resultantes da passagem de tempo (ou diacrônicas), variedades geográficas (ou diatópicas) e variedades socioculturais (ou diastráticas).

Segundo Camacho (2004) e Baronas (2011), as alterações resultantes da passagem do tempo (ou diacrônicas) é um processo natural de evolução do homem, assim como ocorre com as ideias, valores e crenças. Como exemplo para este tipo de variação há o pronome **você**, enquanto resultado das mudanças na expressão **Vossa Mercê**. As variações geográficas (ou diatópicas), referem-se às características regionais, que são representadas por diferentes sotaques e variados nomes para o mesmo objeto. Diversos exemplos podem ser apresentados neste contexto:

a) embora ambos países adotem a língua portuguesa, Brasil e Portugal apresentam diferentes nomenclaturas que representam o mesmo conceito, respectivamente:
   - açougue, talho;
   - apostila, sebenta;
   - banheiro, casa de banho;
   - celular, telemóvel;
   - ônibus, autocarro;
   - sanduíche, sandes;
   - suco, sumo;
   - trem, comboio,
b) diferentes nomenclaturas para o mesmo conceito em regiões brasileiras:
   - pão francês, pão carequinha, pão de sal, cacetinho ou pão-cacetinho (Rio Grande do Sul);

- geladinho, sobremesa de sacolé (Rio de Janeiro), din-din (Nordeste), chup-chup (São Paulo), laranjinha (Goiás);
- mexerica, bergamota ou vergamota (Sul), poncan ou mexerica (Goiás), tangerina (Sudeste), mimosa ou laranja-cravo (Nordeste), mandarina, fuxiqueira, manjerica.

Além das variações em função do tempo e das características regionais, existem as variações relacionadas aos diferentes estratos sociais (diastráticas), que podem estar relacionadas à idade, gênero, sexo ou classe social. Neste sentido, Castilho (2010) afirma que as pessoas não falam exatamente da mesma forma e que as que possuem maior nível de escolaridade tendem a utilizar o **português culto**, ou variedade padrão, aprendido na escola ou no ambiente familiar, em detrimento das demais, que adotam o **português popular**, ou variedade não culta. Existem diferenças significativas relacionadas à camada mais jovem da população, que incluem no vocabulário gírias e linguagem informal, e as pessoas mais idosas, que tendem a ser mais conservadoras.

Os problemas linguísticos apresentados são fatores de interferência na recuperação de informação, uma vez que as características do pesquisador (nível de conhecimento, cultura, localização geográfica, estrato social) e a ambiguidade presente na língua escrita (sinonímia, homonímia, paronímia, polissemia e variações linguísticas) tendem a influenciar as terminologias adotadas na escrita dos documentos, no processo de indexação (quando realizado de forma manual), na elaboração da expressão de busca e na interpretação dos resultados apresentados. Isto significa que muitos documentos relevantes podem não ser recuperados em função da não coincidência entre os termos utilizados na busca e na indexação ou por limitação do vocabulário adotado pelo usuário.

Em contrapartida, outras questões ainda podem ser levantadas que também causam complicações para o processo de elaboração das expressões de busca: a) falta de interatividade que as ferramentas atuais oferecem; b) uso de poucas palavras para formular consultas (aproximadamente 2,2 termos); c) resistência e falta de conhecimento para a correta utilização de expressões booleanas; d) dificuldade de traduzir uma necessidade de informação por meio de termos ou expressões de busca; e) obrigatoriedade de conhecimento do usuário sobre o domínio pesquisado, uma vez que, para realizar uma busca utilizando palavras-chave (ou expressões) significativas, é necessário que ele tenha conhecimento sobre o contexto de interesse (ANTONIOU; HARMELEN, 2004; BEPPLER, 2008; NAYAK; PRASAD; SENAPATI, 2015).

Lopes (2002) destaca que, apesar dos problemas observados, os sistemas de

recuperação de informação na Web têm potencial para o desenvolvimento de recursos que podem contribuir para o aperfeiçoamento do processo de busca. Neste sentido, foram observadas sugestões de melhorias que envolvem questões relacionadas à elaboração das expressões de busca, indexação do *corpus* e interpretação dos resultados.

Em relação à elaboração de expressões de busca, Jackson e Moulinier (2002), Woods (2004), Baeza-Yates e Ribeiro-Neto (2013) e Shneiderman et al. (2017) e apresentam as seguintes sugestões para facilitar a interação do usuário com o sistema:

a) desenvolvimento de novas metáforas visuais para a elaboração de expressões de busca e visualização de resultados, com condições de reduzir o trabalho adicional do usuário para a combinação da estrutura e do conteúdo nas consultas;
b) possibilidade de formulação de consultas baseadas em conceitos;
c) possibilidade de formulação de consultas baseadas em conteúdo;
d) ferramentas de busca devem ser capazes de lidar com relacionamentos semânticos;
e) desenvolvimento de recursos que permitam a execução de múltiplas tarefas de recuperação de informação, tais como: selecionar documentos de uma base, agrupá-los em categorias ou classes e deles extrair determinados conjuntos de informações;
f) novas opções devem ser exploradas nas ferramentas de busca, tais como filtragem, filtragem colaborativa, criação de sentido e análise visual.

Baeza-Yates e Ribeiro Neto (2013) trazem duas sugestões importantes para a melhoria dos sistemas de recuperação de informação, relacionadas à indexação (e ranqueamento) dos documentos:

a) desenvolvimento de melhores esquemas de ranqueamento, que explorem tanto o conteúdo quanto a estrutura dos documentos;
b) construção de estruturas de indexação específicas que explorem a estrutura dos documentos.

Do ponto de vista da análise e interpretação de resultados, Branski (2004), Hogan et al. (2011) e Baeza-Yates e Ribeiro-Neto (2013), trazem algumas sugestões de contribuições que podem aperfeiçoar as buscas realizadas pelos usuários:

a) as ferramentas de busca devem procurar ir além da mera identificação, a partir de um termo ou conceito, dos conteúdos que serão apresentados ao usuário;

b) a partir de um termo ou conceito oferecido pelo usuário, a interface deve ser capaz de identificar citações relacionadas e, a partir delas, extrair outros termos ou conceitos que indiquem novas estratégias de busca;
c) é importante incluir recursos que melhorem a navegação, tais como: melhor exploração dos *links*, popularidade de páginas Web, similaridade de conteúdo, colaboração, 3D e realidade virtual, com atenção especial para abordagens de busca/navegação híbridas;
d) as ferramentas de busca devem incorporar tarefas de coleta de informação acerca de um grupo de sites ou documentos. Nas ferramentas atuais, os usuários devem agregar manualmente as informações que precisam de vários locais heterogêneos, cada qual com suas características de formatação e navegação.

Saracevic (1997) afirma que o sucesso ou fracasso de qualquer sistema interativo ou tecnologia depende de como os problemas dos usuários e os fatores humanos são abordados. Portanto, conhecer as características dos diferentes tipos de usuários, as atividades que desenvolvem e de que forma eles buscam, utilizam e transferem a informação em diferentes contextos (GASQUE; COSTA, 2010) pode ser fator determinante para o projeto de sistemas de busca mais eficientes e para a redução dos problemas existentes.

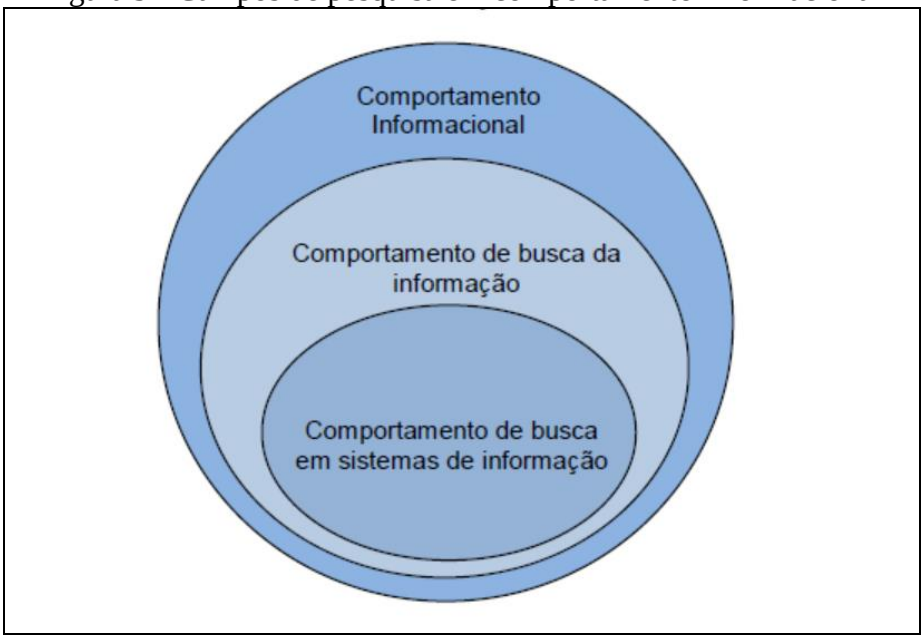
Além dos problemas relacionados à qualidade e variedade de dados disponibilizados e à interação dos usuários com os sistemas, a área de comportamento de informacional, mais especificamente no que se refere ao comportamento de busca em sistemas de recuperação de informação, também discute inúmeras barreiras que interferem na busca realizada pelo usuário, sendo este o principal foco de discussão da próxima seção.

2.2.1 Comportamento de Busca em Sistemas de Recuperação de Informação

Os estudos sobre comportamento informacional são de grande relevância no contexto da Recuperação de Informação, uma vez que incluem investigação sobre "comportamentos humanos relacionados com a busca, coleta, recuperação, organização e uso da informação" (SPINK; COLE, 2006, p.25). Eles estão ligados não apenas com a área de Ciência da Informação, mas também com vários outros domínios, tais como a Psicologia Cognitiva, a Comunicação, a Biblioteconomia, a Arquivística, a Economia, a Antropologia, entre outras (FIGUEIREDO, 1994).

Nesta perspectiva, Wilson (1999) desenvolveu um modelo conceitual, na tentativa de esclarecer os objetivos das pesquisas sobre comportamento informacional e comportamento de busca de informação. Seu estudo é de grande relevância para tentar minimizar os problemas relacionados com o conflito existente entre as perspectivas de análise e a terminologia utilizada pelos autores para caracterizá-los, destacando claramente os campos de atuação de cada área, conforme apresenta a Figura 9.

Figura 9 - Campos de pesquisa em comportamento informacional



Fonte: Wilson (1999, p.263, tradução nossa)

Como pode ser observado, Wilson (1999, 2000) destaca três campos de pesquisa correlacionados, definidos a partir da análise dos estudos desenvolvidos, sendo eles:

a) Comportamento informacional *(information behaviour)* refere-se ao campo mais geral de investigação, que inclui a totalidade de fontes e canais de comunicação, com as buscas de informação passiva (assistir TV) e ativa (comunicação face a face), abrangendo informações do cotidiano;

b) Comportamento de busca de informação *(information seeking behaviour)* é um campo de estudos menor, relacionado à variedade de métodos utilizados pelos indivíduos para acessar fontes de informação, com o objetivo de satisfazer uma necessidade. Estas fontes podem ser sistemas de informação manuais (tais como periódicos, bibliotecas ou outros centros de informação) ou computacionais;

c) Comportamento de busca em sistemas de informação *(information searching behaviour)* refere-se à um subcampo do Comportamento de Busca de Informação que estuda as interações entre usuário (com ou sem intermediários) e sistemas de informação. Este campo abrange o nível da interação humano/computador (por exemplo, o uso do *mouse* e os *clicks* nos *links*) ou o nível intelectual (por exemplo, a adoção de uma estratégia de busca, a utilização de operadores booleanos ou a determinação de critérios para decidir quais materiais serão mais importantes), o que também envolverá atos mentais, tal como o julgamento de relevância dos dados ou informação recuperada.

Partindo desses conceitos, o foco desta pesquisa está no campo mais restrito, direcionado ao comportamento de busca em sistemas de informação. Neste contexto, diversos modelos teóricos foram desenvolvidos e descrevem como os usuários normalmente navegam nas páginas da Web e de que forma eles buscam as informações de que necessitam. Eles variam quanto aos seus pressupostos, estrutura, propósitos, escopo e uso pretendidos (CASE, 2012). Entre os diversos modelos com esta finalidade, pode-se destacar: Wilson (1981, 1999), Krikelas (1983), Ellis (1989), Kuhlthau (1991), Dervin (1992) e Savolainen (1995), dentre outros.

A maior parte destes modelos define o reconhecimento de alguma necessidade (pessoal, social ou profissional), uma lacuna de conhecimento ou uma motivação como ponto de partida para o processo de busca, ou seja, o indivíduo percebe que a informação que possui não é o suficiente para responder seus questionamentos, reduzir suas incertezas, dar sentido à uma situação atual, tomar uma decisão, realizar uma tarefa particular ou resolver um problema (WILSON, 1981; CASE, 2012). Neste sentido, busca de informação pode ser definida por:

> [...] qualquer atividade de um indivíduo que compromete-se a identificar uma mensagem que satisfaça a uma necessidade percebida. Em outras palavras, a busca de informação tem início quando alguém percebe que o seu atual estado de conhecimento é menor do que o necessário para lidar com algum assunto [ou problema] (KRIKELAS, 1983, p.6-7).

Essa necessidade, por sua vez, quando reconhecida pelo usuário, gera a busca pela informação para a mudança do seu estado de conhecimento, consistindo em um conjunto de ações ou estágios, que compõem a busca. Para Kuhlthau (1991), é importante entender cada destes estágios, bem como os aspectos sentimentais e cognitivos que

acompanham o usuário, para que, a partir destas informações, possam ser projetados sistemas de informação capazes de, a cada etapa do processo, auxiliar o usuário na busca por informações de acordo com suas reais necessidades. Cole et al. (2015) afirmam que as atividades desenvolvidas em uma sessão de busca para identificar uma informação, não são aleatórias e que o tipo de tarefa a ser desenvolvida pode afetar o comportamento do usuário durante a pesquisa. As decisões sobre a aquisição e utilização de informações para o progresso em direção ao objetivo de uma tarefa dependem, em parte, de estados mentais complexos, incluindo o nível de conhecimento do usuário sobre a tarefa, o seu conhecimento sobre o domínio, além da natureza e do objetivo da tarefa.

Existem inúmeros fatores que interferem na forma como os indivíduos buscam informação e como interpretam os resultados das buscas realizadas. Wilson (1996) discute três categorias de barreiras que influenciam o comportamento informacional dos indivíduos, denominando-as de **variáveis interferentes**, que envolvem características pessoais, sociais/interpessoais e ambientais.

Em relação às barreiras de ordem pessoal, Wilson (1996) destaca variáveis relacionadas aos seguintes elementos:

a) dissonância cognitiva: são barreiras socioculturais originadas por normas sociais e valores culturais (crenças existentes, orientação política, estereótipos, preferências, preconceitos etc.);
b) exposição seletiva: sugere que os indivíduos, consciente ou inconscientemente, preferem informações que estejam de acordo com os seus interesses, necessidades e atitudes;
c) características fisiológicas, cognitivas e emocionais: inclui limitações de ordem física (problemas auditivos, por exemplo), falta de conhecimento, autopercepção (autoavaliação de conhecimento e habilidades), limitações verbais, nervosismo e variáveis emocionais (aversão ao risco, dúvida, conformismo etc.);
d) nível educacional e base de conhecimento: o nível de educação e de conhecimento estão diretamente relacionados à forma como o indivíduo realiza suas buscas e interpreta os resultados encontrados. Também podem ter relação com a quantidade de informação procurada (pessoas com mais conhecimento tem a sentir menos necessidade de procurar mais informações), com a facilidade de busca (quanto mais conhecimento o indivíduo tiver, mais facilidade terá para adquirir a

informação desejada), com a percepção de conhecimento (os indivíduos tendem a buscar menos informações nos assuntos sobre os quais se sentem bem informados) e com a importância percebida de um tópico (quanto mais importante um assunto for para o indivíduo, mais interesse terá na busca);

e) variáveis demográficas: inclui idade, sexo, estado social e econômico, experiência de trabalho etc.

As barreiras relacionadas às questões sociais e interpessoais referem-se ao papel que o indivíduo exerce na sociedade (normalmente há vários, inclusive o papel profissional: pai, mãe, líder, gerente, empresário, médico, professor, pesquisador) e às variáveis interpessoais (de caráter profissional), envolvendo os seguintes elementos: exigências, regulamentos e limitações; padrões de comportamento estabelecidos (em uma categoria profissional em particular); lugar que uma pessoa ocupa em uma organização ou em um sistema de organizações; hierarquia típica de valores; e nível de responsabilidade. Estas barreiras estão geralmente relacionadas ao contexto do indivíduo e influenciam suas necessidades de informação (as necessidades de informação de um médico são diferentes das necessidades de informação de um empresário, por exemplo), sua percepção em relação às diversas barreiras e a forma como uma necessidade é satisfeita (NIEDŹWIEDZKA, 2003; WILSON, 1996).

Além das barreiras pessoais, sociais e interpessoais, Wilson (1996) também discute as barreiras relacionadas às questões ambientais, que se referem ao macroambiente (sistema sócio-político e econômico de um país ou setor da indústria), ao meso-ambiente (ambiente regional, comunidade local, uma cidade particular) ou ao microambiente de uma única organização. Nesta categoria incluem-se elementos relacionados à legislação, situação econômica, nível de estabilização, estrutura organizacional, cultura informacional, tecnologia da informação, tipo de organização, cultura organizacional.

Neste mesmo contexto, Savolainen (2016) realiza uma análise sobre as principais barreiras socioculturais que interferem no processo de busca, apresentando seis categorias principais relacionadas a: problemas de linguagem, estigma e tabu cultural, características locais, instituição, organização e falta de capital social e econômico. Essas barreiras possuem um impacto negativo na busca da informação, uma vez que restringem as fontes de informação e geram emoções negativas.

As barreiras socioculturais relacionadas aos problemas de linguagem são provenientes da deficiência de conhecimento do usuário em relação à língua dominante (ou

idioma da busca) e, consequentemente, do seu vocabulário limitado. Essas dificuldades são ainda mais acentuadas pelas diferenças culturais, pelos problemas inerentes à linguagem relacionados, principalmente, pela semântica (sinonímia, homonímia, paronímia, polissemia, denotação, conotação, variações linguísticas) e pelas mudanças de sentido que podem ocorrer em função do tempo e do espaço geográfico. Esses fatores interferem diretamente na elaboração dos termos de busca e no acesso às informações disponibilizadas na Web (LORENCE; SPINK, 2004; MEHRA; BILAL, 2007; YI; STVILIA; MON, 2012).

As barreiras geradas a partir do estigma e tabu cultural podem afetar pessoas pertencentes a qualquer nível social, independente do seu grupo linguístico. São problemas gerados em função de características indesejadas que formam um estereótipo ou causam um estigma, tais como: deficiência física, orientação sexual, obesidade excessiva, *status* social, etc. Também podem ser provenientes de restrições culturais (tabus) baseadas em julgamento moral ou crenças religiosas e, por medo ou vergonha, são responsáveis por limitar a busca de informações dos indivíduos, principalmente quando ela ocorre pessoalmente ou em grupos de discussão (CAIDI; ALLARD; QUIRKE, 2010; SAVOLAINEN, 2016).

Diferentemente das restrições relacionadas ao estigma social e tabu cultural, as barreiras associadas às características locais aparecem em nível de comunidade e orientam os indivíduos de acordo com normas específicas e valores culturais daquele grupo particular. Geralmente, são restritas aos integrantes da comunidade (HAYTER, 2006; SAVOLAINEN, 2016).

Savolainen (2016) discute ainda três outras barreiras que interferem na busca de informações: barreiras institucionais, organizacionais e por falta de capital social e econômico. As barreiras institucionais e organizacionais se manifestam de várias formas, consciente ou inconscientemente, e impedem que os indivíduos obtenham a informação que é necessária. Entre as principais, destacam-se: burocracia excessiva para acesso e disponibilização de informação, linguagem de difícil entendimento, censura (sistemas restritivos de informação), dificuldades de acesso a bases de dados, entre outras. As barreiras que ocorrem devido à falta de capital social e econômico, por sua vez, não atingem igualmente toda a população, sendo que as pessoas menos favorecidas têm mais probabilidade de enfrentar este tipo de problema. Em relação às dificuldades de acesso à informação, encontram-se: alfabetização, habilidades analíticas, pobreza de informação, dificuldades de acesso à informação (recursos e equipamentos) entre outras (DERVIN, 1999; CHOWDHURY; GIBB, 2009; YU, 2010).

Embora não fiquem claros os limites existentes entre uma categoria e outra,

tampouco a organização das variáveis dentro do entendimento de cada autor, o quesito de maior relevância para o contexto da presente pesquisa é o levantamento de fatores que pode influenciar no processo de busca e recuperação de informação, pois a partir destes elementos, foi possível entender as principais limitações dos usuários na utilização dos sistemas de busca. Além disso, o conhecimento do pesquisador sobre o domínio de interesse é fundamental para o desenvolvimento de buscas significativas e com resultados relevantes.

2.2.2 Características do Processo de Busca

De acordo com Xie e Joo (2010), a principal interação no processo de busca de informações na Web é a submissão de termos de buscas, ou seja, a maioria dos movimentos ou táticas elaboradas estão relacionadas com a formulação ou reformulação da consulta (ampliando, mantendo, restringindo ou alterando o contexto pesquisado). Além disso, destacam-se também o monitoramento do processo de pesquisa e a análise dos documentos recuperados.

Diferentes definições têm sido empregadas para representar movimentos, táticas e estratégias realizadas durante um processo de busca de informações. Bates (1992) definiu os níveis de movimentos que podem ser realizados pelo usuário em uma atividade de pesquisa: a) um **movimento** indica um pensamento identificável ou uma ação básica no processo de busca de informações; b) uma **tática** consiste em um movimento ou em um conjunto de movimentos, incluindo opções de busca e ações, utilizados pelos usuários para avançar no processo de busca de informações; c) um **estratagema** é composto de várias táticas que exploram uma informação específica de um domínio; d) uma **estratégia** representa um plano que inclui movimentos, táticas, e/ou estratagemas para o processo de pesquisa.

Ainda neste contexto, Marchionini (1995) também apresenta definições sobre movimentos, táticas e estratégias. Os movimentos são ações individuais de baixo nível, as táticas são um grupo de ações aplicados para melhorar a busca da informação e as estratégias são abordagens adaptadas para resolver os problemas, que consistem em uma série ordenada de táticas. As estratégias de pesquisa podem ser divididas em dos tipos essenciais: busca (*searching*) e navegação (*browsing*).

De acordo com Hearst (2009) e Case (2012), a atividade de busca (*searching*) é aquela realizada pelo usuário com um objetivo específico e previamente definido. Neste paradigma, o pesquisador tem o intuito de encontrar um determinado conteúdo e, para isto, utiliza expressões ou palavras-chave referentes ao assunto de interesse.

São geralmente realizadas por meio de caixas de busca, disponibilizados nos buscadores da Web ou em outros ambientes digitais, tais como bibliotecas, repositórios, *sites*, etc.

As atividades de navegação (*browsing*), por sua vez, possuem o objetivo de possibilitar a exploração de informações e, de certa forma, estimular o mecanismo associativo dos usuários e a serendipidade, ou seja, a descoberta de informações importantes por acaso, sem que elas estivessem sendo procuradas. Neste tipo de busca, o usuário não possui um objetivo específico e explora o conteúdo disponibilizado em um *corpus* documental com intenção de visualizar alguma informação relevante. Desta forma, o comportamento de busca do usuário não é planejado e as informações são encontradas sem ação premeditada. Este cenário acontece frequentemente quando o pesquisador se depara com uma grande quantidade de informações, tais como: jornais eletrônicos, revistas, publicidade, quadro de avisos, *sites*, *blogs* e etc.

Pode-se incluir também a navegação por estruturas de *links* ou categorias previamente definidas, alternando de uma visualização para outra, em uma sequência de operações de rastreamento e seleção. Além destes dois paradigmas, é comum o desenvolvimento de ambientes que disponibilizem ao usuário a possibilidade de elaborar pesquisas híbridas, que podem ser desenvolvidas a partir da busca e da navegação.

Para a elaboração de pesquisas a partir do paradigma de busca (*searching*), o usuário pode utilizar operadores booleanos e sinais modificadores para ampliar ou restringir os resultados retornados, conforme apresentam Rowley (2002) e Branski (2004):

Quadro 2 - Operadores booleanos AND, OR, NOT e sinais modificadores

Operador	**Significado**
AND (A and B)	Recupera apenas as páginas ou documentos que contenham A e B
OR (A or B)	Recupera todas as páginas ou documentos que contenham somente A, somente B ou mesmo AB
NOT (A not B)	Recupera apenas as páginas ou documentos que contenham A, ou seja, as que contenham B ou AB não serão identificadas
+A –B	Recupera documentos que contenham o termo A, mas não contenham o termo B
“Recuperação de Informação”	Recupera documentos que contenham a frase buscada

Fonte: Rowley (2002) e Branski (2004)

É importante observar que quanto mais termos ou conceitos forem combinados em uma busca contendo o operador AND, menos registros serão recuperados, ao contrário do operador OR, que quanto mais termos ou conceitos forem combinados em uma busca utilizando este operador, mais registros serão recuperados (ROWLEY, 2002). Além da

utilização dos operadores booleanos e sinais modificadores, é importante destacar que o ambiente Web oferece recursos que proporcionam o surgimento de novas táticas e estratégias de busca, que podem ser utilizadas para a formulação e reformulação de consultas.

Apesar dos recursos de busca disponibilizados pelos buscadores, é importante avaliar ainda as mudanças ocorridas com o surgimento da Web e os principais impactos causados no comportamento do usuário (GRĂDINARU, 2013). Diante da grande quantidade de materiais disponíveis *on-line*, páginas com conteúdos diversos, sites especializados e um número expressivo de periódicos digitalizados, Nicholas (2010) assegura que a simplicidade e rapidez são tão importantes quanto relevância e qualidade, ou seja, a dimensão de um artigo ou material disponibilizado passou a ser fator decisivo para a sua leitura ou não. Se um artigo for muito extenso, existe grande probabilidade de ter apenas o seu resumo lido. Além disso, existe uma grande pressão para que os resumos sejam bem escritos, estruturados e cada vez mais informativos, para serem capazes de indicar sua escolha e uso.

Outra característica do comportamento de busca na era digital é que os usuários estão realizando pesquisas superficiais sobre os conteúdos de interesse (busca horizontal) em detrimento das buscas com análises mais aprofundadas (busca vertical). De forma geral, eles não ficam muito tempo em um mesmo local (*site*, jornal, revista) e buscam informações de forma panorâmica, por meio de uma navegação rápida pelas muitas informações disponíveis, apenas com o objetivo de selecionar algumas partes e não de se aprofundar em um determinado assunto.

Navarro-Prieto, Scaife e Rogers (1999) apresentaram três padrões gerais de busca de informações: estratégia *top-down* (quando a pesquisa se inicia em um grande domínio e vai se restringindo até encontrar a informação específica); estratégia *bottom-up* (quando o usuário, mais experiente ou com domínio do conteúdo a ser pesquisado, procura por uma determinada palavra-chave) e estratégia mista, formada pela junção das duas anteriores.

Outro ponto importante relacionado ao comportamento de busca *on-line* refere-se ao fato de que existe uma grande quantidade de usuários visualizando apenas algumas páginas da Web. Segundo Case (2012), a perspectiva é de que 80% dos usuários utilizem apenas 20% de todo o conteúdo disponibilizado na Web (regra 80/20). A explicação para este comportamento está fundamentada em alguns fatos: existe uma enorme variedade de informações disponíveis na Web, as características pouco interativas das ferramentas de busca, das ferramentas e da interface, a escassez de tempo e o uso de habilidades básicas de recuperação (e consequentemente, uma grande quantidade de resultados irrelevantes).

É possível afirmar ainda que existe uma relação entre a experiência e o comportamento do indivíduo na Web. Isto significa que os níveis de conhecimento de domínio do usuário, sua experiência em relação à utilização da Web e às ferramentas de busca e sua experiência com a elaboração de pesquisas são elementos que podem gerar diferenças na elaboração de pesquisas mais eficazes e bem-sucedidas (COTHEY, 2002; FORD; EAGLESTONE; MADDEN, 2009). Tang (2007) assegura que quanto mais conhecimento o usuário tiver sobre o contexto pesquisado, mais significativas e contextualizadas serão suas expressões de busca, além da possibilidade de explorar os diferentes aspectos de suas necessidades.

Lazonder, Biemans e Woperis (2000) discutem as questões relacionadas com a experiência do usuário em relação à Web e às ferramentas de busca. Segundo eles, os usuários com níveis mais elevados de experiência na utilização destes recursos, tiveram melhores desempenhos na realização de buscas, obtendo resultados mais rápidos, mais precisos e com a necessidade de empregar menos ações (menos *links* acessados até encontrar a informação desejada). Em contrapartida, os baixos níveis de experiência têm sido relacionados com maior tempo gasto na pesquisa e com dificuldades na localização de informações relevantes. Turner, Kaske e Baker (1990) descobriram ainda correlação entre a experiência de busca do usuário e o uso de uma maior variedade de comandos e recursos complexos para elaborar e reformular consultas, ampliando ou restringindo o campo de investigação.

Neste mesmo contexto, Jansen, Spink e Saracevic (2000) analisaram os registros de 51.473 consultas realizadas por 18.112 usuários do buscador Excite e levantaram informações importantes sobre as sessões de busca (tipos de consultas, alterações nas consultas durante uma sessão e número de páginas visitadas), as consultas realizadas (número de termos de busca e utilização de operadores booleanos e modificadores) e os termos mais frequentemente utilizados no processo de busca. O grande diferencial do estudo é que ele foi desenvolvido com usuários em reais situações de busca e não em condições propostas ou direcionadas, como na maioria dos estudos realizados neste campo. Portanto, foi possível traçar um panorama, com estatísticas detalhadas, sobre o comportamento de busca e a interação entre os usuários e os buscadores da Web.

Em relação às sessões de busca, os autores adotaram como **consulta original** a primeira busca realizada pelo usuário, **consulta modificada** como sendo uma busca subsequente à primeira, realizada pelo mesmo usuário, com termos adicionados, removidos ou alterados a partir da consulta original e, por fim, **consultas idênticas** são

aquelas em que o usuário opta pela visualização de páginas subsequentes à primeira (uma página significa um grupo contendo 10 resultados). No Excite, cada vez que um usuário acessa uma segunda página, uma consulta idêntica à primeira é gerada.

As consultas originais ou modificadas representaram 57% de todas as consultas realizadas no contexto pesquisado, ou seja, 29.437 consultas foram desse tipo e 43% foram caracterizadas como consultas idênticas, que indica que o usuário está explorando os resultados da consulta original (páginas subsequentes à primeira). Do total de consultas originais ou modificadas, percebeu-se que 67% dos usuários efetuaram apenas uma consulta em uma sessão, sendo que os outros 33% realizam consultas sucessivas e deste total, 14% realizaram três ou mais consultas. Esses dados sugerem a necessidade de realizar estudos mais detalhados com os 14% de usuários que realizam consultas sucessivas, uma vez que eles podem representar um subgrupo de usuários da Web com mais experiência ou maior motivação para realizar modificações nas expressões de busca, possuindo características diferenciadas dos demais.

Uma sessão média, incluindo os três tipos, foi de 2,84 consultas por sessão. Isto significa que muitos usuários modificam suas consultas ou querem ver resultados subsequentes (ou ambos). A duração de uma sessão média, ignorando consultas idênticas, foi de 1,6 buscas por usuário.

No que diz respeito às consultas modificadas (consultas sucessivas do mesmo usuário que tiveram inclusão, alteração ou redução de termos), que representaram 11.249 do total analisado, os autores observaram que em cerca de 1/3 delas os usuários mantiveram a mesma quantidade de termos em suas consultas sucessivas, ou seja, não aumentam e nem diminuíram os termos utilizados. Nas 7.338 consultas restantes foi observada alteração (acréscimo ou redução) da quantidade de elementos utilizados na busca.

No quesito visualização de resultados, Jansen, Spink e Saracevic (2000) verificaram que o número médio de páginas examinados por usuário foi de 2,35. A maioria dos usuários, 58% deles, não acessam qualquer resultado após a primeira página. Não se pode afirmar, no entanto, se as primeiras buscas foram altamente precisas e, por isso, os usuários ficaram plenamente satisfeitos com os resultados apresentados e descartaram os demais ou se eles simplesmente desistiram ou ficaram cansados de visualizar os demais resultados.

Este panorama foi discutido por Case (2012), que aponta problemas típicos dos usuários de informação, relacionados principalmente com a sobrecarga de informações. Case (2012, p.98) afirma que “sempre que o número de alternativas for superior à 10 elementos, os indivíduos estarão propensos a experimentar uma sobrecarga de informações”.

Neste sentido, é comum os usuários aceitarem como satisfatório os primeiros resultados apresentados, mesmo que eles sejam mais simples, parcialmente relevantes ou menos confiáveis, apenas para reduzirem o tempo de pesquisa. É importante destacar, no entanto, que os propósitos da busca são fatores determinantes para a realização de consultas superficiais (que aceitam quaisquer resultados) ou mais abrangentes e detalhadas (que necessitam de informações mais precisas).

Os dados mostraram ainda que, geralmente, as consultas realizadas são curtas, contendo em média 2,21 termos. Cerca de 62% de todas as consultas analisadas continham um ou dois termos de busca. Menos de 4% das consultas foram elaboradas com mais do que 6 termos. Em relação à estrutura das consultas, foi possível observar que os operadores (AND, OR e NOT) foram pouco utilizados, aparecendo em apenas 8% das consultas realizadas. Já os termos modificadores (+, -, “ ”) foram utilizados mais frequentemente, sendo observados em aproximadamente 15% das consultas elaboradas.

De acordo como Jansen, Spink e Saracevic (2000), estes dados comprovam que apenas uma porcentagem muito pequena de usuários incorpora operadores booleanos ou termos modificadores para elaborar suas buscas, indicando a realização de consultas mais simples e menos sofisticadas. Tais recursos, se utilizados corretamente, teriam condições de trazer melhores resultados de busca aos usuários.

2.2.3 Recuperação de Documentos Textuais

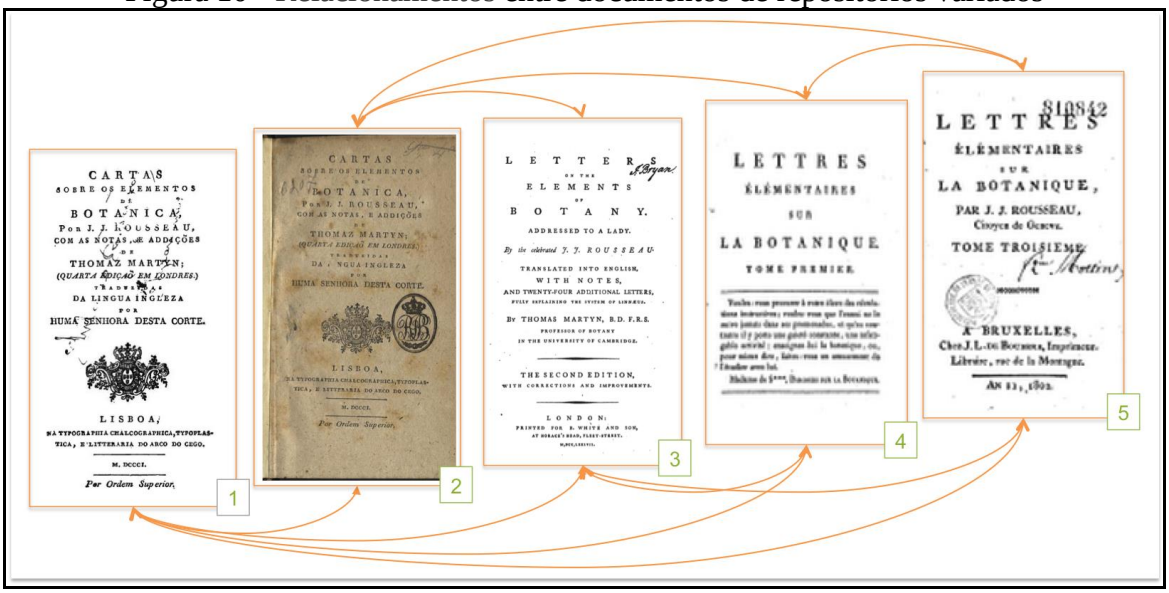
Os Sistemas de Recuperação de Informação (SRI) utilizam recursos de indexação automática ou manual, com o objetivo de representar abreviadamente o conteúdo dos documentos pertencentes ao *corpus* e, posteriormente, facilitar as buscas realizadas pelos usuários e a localização de informações de seu interesse (BAEZA-YATES; RIBEIRO-NETO, 2013; LOHBARE; MESHRAM, 2014). A maioria dos métodos de indexação utilizados são baseados em aspectos sintáticos e estatísticos, considerando a frequência e a distribuição de palavras presentes em documentos. Desta forma, as consultas realizadas pelos usuários devem coincidir, total ou parcialmente, com a indexação realizada, para que os documentos possam ser ranqueados e recuperados pelo sistema, de acordo com a explicitação de uma necessidade de informação.

Apesar dos diversos algoritmos utilizados para apoiar o processo de indexação dos documentos, os modelos de RI consideram os textos como uma *string*, sem nenhuma estrutura em particular, sendo que as informações relacionadas às seções, subseções,

parágrafos e figuras não são utilizadas para o ranqueamento. Desta forma, a busca se concentra apenas em informações apresentadas nos documentos de forma explícita, dispensando todo o conhecimento latente, que é previamente desconhecido, mas de grande relevância para a análise de um *corpus* documental. Além disso, os documentos retornados são apresentados de forma desconectada uns dos outros, mesmo que apresentem elementos ou informações relacionadas (BAEZA-YATES; RIBEIRO-NETO, 2013; PAUL et al., 2014; NAYAK; PRASAD; SENAPATI, 2015).

Souza (2012) apresenta um exemplo sobre as possibilidades de relacionamentos entre cinco documentos, que não seriam evidenciadas nos resultados de um sistema de busca.

Figura 10 - Relacionamentos entre documentos de repositórios variados



Fonte: Souza (2012)

Na Figura 10 podem ser observados cinco documentos digitais, elaborados a partir das *Cartas sobre os Elementos de Botânica*, de J. J. Rousseau, que estão disponibilizados em bibliotecas digitais do Brasil, de Portugal, da Inglaterra e da Suíça. Do ponto de vista da pesquisa científica, o acesso ao conjunto destes materiais é de grande relevância e pode ser fundamental para o desenvolvimento de investigações aprofundadas (por exemplo, um estudo sobre a trajetória da tradução portuguesa).

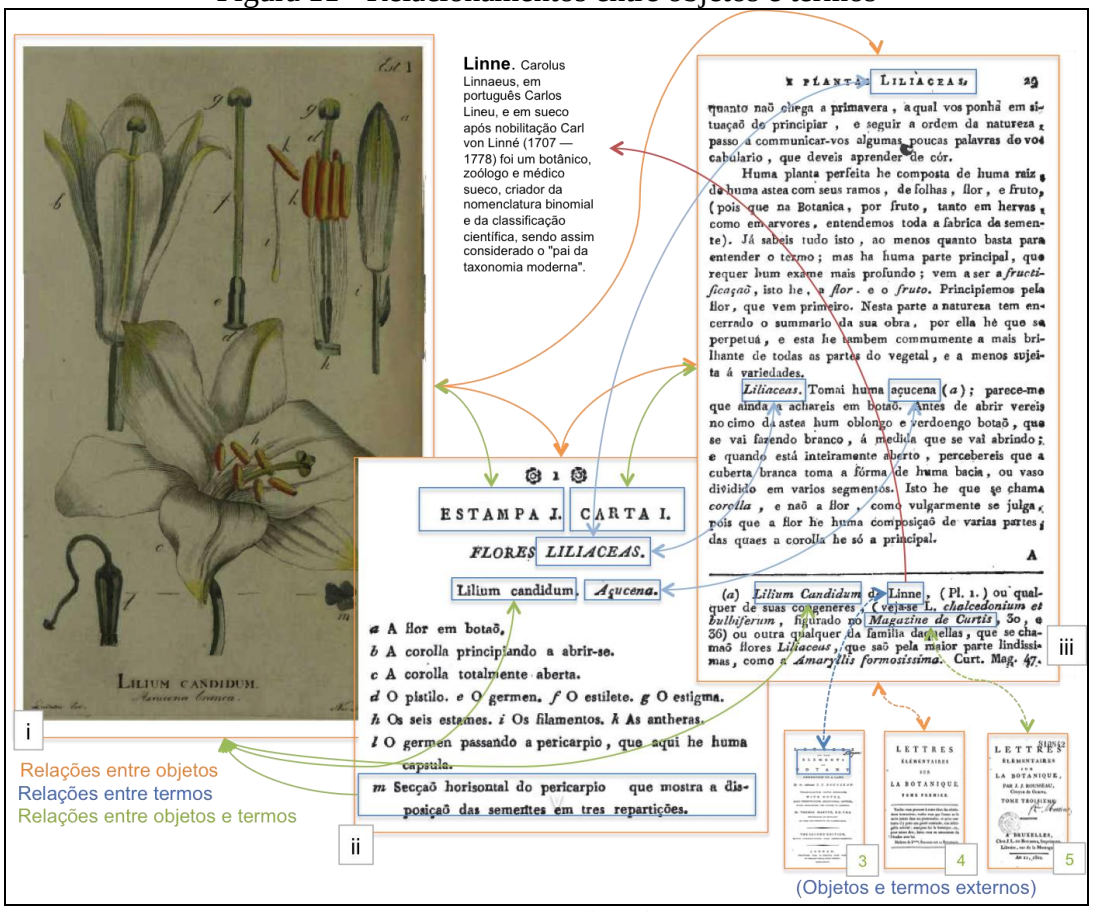
Apesar de todas as edições e respectivas traduções estarem em formato digital, os resultados de uma busca sobre o assunto seriam apresentados de forma desconectada, exigindo que o pesquisador tivesse tempo e conhecimentos suficientes para identificar as relações existentes entre eles. A dificuldade de recuperação dos documentos em

conjunto ainda é agravada pelo fato de terem sido publicados em idiomas diferentes, o que exigiria a realização de diversas consultas para que todos os exemplares fossem recuperados (SOUZA, 2012; PAUL et al., 2014).

O contexto discutido neste exemplo representa a maior parte das pesquisas realizadas na Web e nos repositórios digitais. Existe uma grande quantidade de documentos disponíveis relacionados uns com os outros, que são apresentados aos usuários de forma de isolada e se misturam a outros materiais, muitas vezes irrelevantes, e acabam não sendo encontrados. Desta forma, se houvesse um relacionamento explícito entre os documentos pertencentes a um *corpus*, todos seriam encontrados em apenas uma busca.

Além dos relacionamentos entre objetos externos (documentos digitais diferentes), existem aqueles que se referem a objetos internos (partes de um mesmo documento). De uma forma ainda mais específica, existem também os relacionamentos entre os termos, que podem ser externos (entre termos pertencentes a documentos diferentes) ou internos (entre termos pertencentes a um mesmo documento). Um exemplo destes quatro tipos de relacionamentos pode ser observado na Figura 11.

Figura 11 - Relacionamentos entre objetos e termos



Fonte: Souza (2012)

O documento apresentado refere-se a três páginas do documento *Cartas sobre os Elementos de Botânica*. Neste exemplo é possível observar diversas relações recíprocas entre os elementos: Figura (i), página de legenda (ii), texto da página 29 (iii) e documentos externos. Neste contexto, Souza (2012) destaca algumas relações que podem ser facilmente notadas:

a) entre objetos externos: relação entre o texto da página 29 (iii) e o documento 4;
b) entre objetos internos: relação entre a Figura (i), que representa um lírio e suas partes, e a página de legenda (ii);
c) entre termos externos: relação entre o termo *Linne* da página 29 (iii) e *Elementos da Botânica* (documento 3);
d) entre termos internos: as três páginas do documento estão relacionadas por apresentarem o termo *lilium candidum*, que é o nome científico do lírio. Também existem relações referentes à *flor lilácea* entre a página de legenda (ii) e o texto da página 29 (iii);

Além das relações discutidas, que poderiam auxiliar um usuário durante a análise do documento ou dos resultados de uma busca, é possível ainda agregar informações aos objetos e termos do texto. Um exemplo disto, seria a ligação entre Carlos Lineu e o termo *Linne* da página (iii). Freire, Pinheiro e Feitosa (2016) destacam ainda alguns elementos que também poderiam ser identificados neste processo: a) sentenças completamente diferentes, em assuntos diferentes; b) sentenças não relacionadas, mas que compactuam do mesmo assunto; c) sentenças de certa forma relacionadas, que podem descrever fatos diferentes, mas compartilham alguns detalhes; d) sentenças fortemente relacionadas, que divergem apenas em alguns detalhes; e) sentenças que possuem o mesmo significado.

Embora estas relações sejam óbvias, os sistemas de busca atuais não conseguem apresentar estas informações ao usuário, a menos que elas estejam explicitadas de alguma forma. Isto significa que, caso a pesquisa acima tivesse sido realizada, o usuário só teria acesso a este conjunto de informações, se conseguisse recuperar e analisar cuidadosamente os documentos relacionados ao contexto estudado.

Esta perspectiva de ligação entre os documentos foi amplamente discutida nas décadas de 80 e 90, no âmbito da utilização do hipertexto e da linguagem HTML. Os estudos buscavam o desenvolvimento de recursos que pudessem facilitar a navegação e o acesso à informação, através da utilização de sumários, índices e trajetos predefinidos e recomendados pelo autor do documento (CODINA, 2001; ROVIRA; CODINA, 2003). No

entanto, as abordagens realizadas não incorporavam as conexões semânticas entre os documentos e suas partes, tampouco entre consultas e documentos.

No contexto deste estudo, as ligações entre os termos, documentos e sentenças serão realizadas a partir do uso de ontologias. Os conceitos, características e benefícios deste recurso para os sistemas de recuperação de informação serão abordados no próximo capítulo desta tese.

2.3 RESUMO E DISCUSSÃO DO CAPÍTULO

Este capítulo teve como objetivo apresentar uma revisão de literatura na área de Recuperação de Informação, discutindo conceitos e características da área e dos principais modelos presentes na literatura: booleano, vetorial e probabilístico. Posteriormente, com o intuito de se aproximar do foco do trabalho, abordou a busca e recuperação de informação na Web, os principais fatores que influenciam o comportamento do usuário, as principais características do processo e as dificuldades relacionadas aos documentos textuais.

A literatura apresenta múltiplas abordagens e entendimentos, realizados sob diferentes perspectivas e correntes teóricas, que apresentam características e fatores que interferem no processo de busca do usuário, tais como: influência de modelos mentais, objetivos de pesquisa e personalidade no comportamento de busca de informação, diferenças entre usuários experientes e novatos, estudos dos aspectos afetivos ou cognitivos de usuários, modelos teóricos e descritivos de busca de informação, busca de informação em sistemas computacionais e estudos sobre estratégias de busca de informação (WILSON, 2000).

Neste sentido, a principal contribuição deste capítulo foi o levantamento dos principais fatores que interferem no processo de busca e recuperação de informação na Web, relacionados principalmente com os documentos textuais, conforme pode ser observado no Quadro 3.

O processo de busca e recuperação de informação sofre interferências de naturezas diversas, podendo ser relacionadas à qualidade e variedade dos dados disponibilizados na Web, à interação do usuário com o sistema durante a elaboração de uma expressão de busca ou interpretação dos resultados apresentados ou relacionados ao comportamento de busca do usuário. Os fatores apresentados são importantes para revelar a complexidade da área e a consequente necessidade de projetar ferramentas que possam minimizar os problemas observados, que é o foco da presente pesquisa.

Quadro 3 - Problemas do processo de busca e recuperação de informação na Web

Perspectiva	Problemas Observados	Autores
Dados	a) Grande volume de dados b) Dados heterogêneos c) Dados redundantes d) Dados não estruturados e) Alto percentual de dados voláteis f) Qualidade dos dados	Baeza-Yates e Ribeiro-Neto (2013), Huang (1999), Nayak, Prasad e Senapati (2015)
Interação (Expressão de Busca)	a) Falta de interatividade b) Uso de poucas palavras para formular consultas c) Dificuldade na utilização de operadores lógicos e termos modificadores d) Dificuldade de traduzir uma necessidade de informação em termos ou expressões de busca e) Problemas linguísticos: sinonímia, homonímia, paronímia e polissemia, denotação, conotação e variações linguísticas f) Obrigatoriedade de conhecimento sobre o domínio pesquisado	Antoniou e Harmelen (2004), Baronas (2011), Beppler (2008), Camacho (2004), Nayak, Prasad e Senapati (2015)
Interação (Interpretação dos Resultados)	a) Falta de relevância e precisão dos resultados apresentados b) Número excessivo de documentos retornados (sobrecarga de informação) c) Dificuldade para manipular grandes quantidades de resultados d) Baixa tolerância em investigar a lista de documentos recuperados e) Não existem informações da relação existente entre a pesquisa realizada e os documentos recuperados f) Semelhanças entre documentos não são apresentadas g) Relacionamentos entre termos (internos e externos) não são evidenciados h) Resultados são baseados apenas em informações apresentadas nos documentos de forma explícita, dispensando todo o conhecimento latente	Baeza-Yates e Ribeiro-Neto (2013), Case (2012), Hobbs, Pfitzner e Powers (2002), Huang (1999), Spink et al. (2001)
Comportamento de Busca	a) Barreiras pessoais: dissonância cognitiva, exposição seletiva, características fisiológicas, cognitivas e emocionais, nível educacional e base de conhecimento, variáveis demográficas b) Barreiras sociais e interpessoais: papéis exercidos, cargo ocupado, nível de responsabilidade c) Barreiras ambientais: contexto que regulamenta a vida do indivíduo (legislação, situação econômica, cultura informacional, etc.)	Wilson (2006)
Comportamento de Busca	a) Problemas de linguagem: vocabulário limitado, diferenças culturais, problemas semânticos b) Estigma e tabu cultural: deficiência física, orientação sexual, julgamento moral, crenças religiosas c) Características locais: normas específicas e valores culturais de um grupo particular d) Institucionais e organizacionais: burocracia excessiva, linguagem de difícil entendimento, censura e) Falta de capital social e econômico: alfabetização, habilidades analíticas, pobreza de informação	Savolainen (2016)

Fonte: desenvolvido pela autora

As barreiras apresentadas por Wilson (2006) interferem direta ou indiretamente em todo o processo de busca, sendo responsáveis por parte dos problemas gerados na fase de elaboração das expressões de busca e na interpretação dos resultados

apresentados. As questões pessoais, sociais, interpessoais e ambientais estabelecem as características e o contexto do usuário, afetando as suas necessidades de informação, a tradução dessa necessidade em termos e expressões de busca, bem como a interpretação das diversas informações retornadas. Desta forma, cada usuário interpreta sua realidade (de busca ou de resultados) de acordo com os conhecimentos que possui e com as suas perspectivas de análise.

As barreiras apresentadas por Savolainen (2016) também são de grande relevância para o presente estudo, em especial as relacionadas aos problemas de linguagem que também interferem na elaboração das expressões de busca e na interpretação dos resultados apresentados. No entanto, seus estudos são mais direcionados ao intercâmbio de informações que ocorre entre pessoas, funcionários, organizações e não especificamente ao contexto da busca na Web.

Em relação à recuperação de documentos textuais, um dos principais problemas observados refere-se à ausência de relacionamentos entre os objetos digitais e suas partes. Consequentemente, os resultados de busca são apresentados de forma desconectada, como se um documento não compartilhasse características com os demais, ficando sob a responsabilidade do usuário a árdua tarefa de identificar os relacionamentos existentes no *corpus* de estudo.

Em contrapartida, os estudos realizados neste capítulo também permitiram o levantamento de diretrizes presentes na literatura, que serão parcialmente utilizadas para o desenvolvimento da ferramenta proposta neste estudo. O Quadro 4 sintetiza as principais sugestões de melhorias para o processo de busca e recuperação de informação no contexto da Web, propostas pelos autores que deram subsídios a esta fundamentação teórica.

Quadro 4 - Sugestões de melhorias para o processo de busca

Perspectiva	**Sugestões de Melhorias**	**Autores**
Interação (Expressão de Busca)	a) Desenvolvimento de novas metáforas visuais para a elaboração de expressões de busca e visualização de resultados; b) Possibilidade de consultas baseadas em conceitos; c) Possibilidade de consultas baseadas em conteúdo; d) As ferramentas de busca devem ser capazes de lidar com relacionamentos semânticos; e) Desenvolvimento de recursos que permitam a execução de múltiplas tarefas de RI, tais como: selecionar documentos de uma base, agrupá-los em categorias ou classes e deles extrair determinados conjuntos de informações; f) Novas opções devem ser exploradas nas ferramentas de busca, tais como: filtragem, filtragem colaborativa, criação de sentido e análise visual.	Baeza-Yates e Ribeiro-Neto (2013), Jackson e Moulinier (2002), Shneiderman et al. (2017), Woods (2004)

Quadro 4 - Sugestões de melhorias para o processo de busca (continuação...)

Perspectiva	Sugestões de Melhorias	Autores
Indexação e Ranqueamento	a) Desenvolvimento de melhores esquemas de ranqueamento, que explorem tanto o conteúdo quanto a estrutura dos documentos; b) Construção de estruturas de indexação específicas que explorem a estrutura dos documentos.	Baeza-Yates e Ribeiro-Neto (2013)
Interação (Interpretação dos Resultados)	a) As ferramentas de busca devem procurar ir além da mera identificação, a partir de um termo ou conceito, dos conteúdos que serão apresentados ao usuário b) A partir de um termo ou conceito oferecido pelo usuário, a interface deveria ser capaz de identificar citações relacionadas e, a partir delas, extrair outros termos ou conceitos que indicassem novas estratégias de busca c) É importante incluir recursos que melhorem a navegação, tais como: melhor exploração dos *links*, popularidade de páginas Web, similaridade de conteúdo, colaboração, 3D e realidade virtual, com atenção especial para abordagens de busca/navegação híbridas d) Um dos desafios da busca de informações é encontrar resultados relevantes. Em vez de uma longa lista de títulos, URLs e documentos, os mecanismos de pesquisa visuais podem oferecer apresentações ricas de resultados, muitas vezes visualmente conectadas a termos de pesquisa relacionados e) As ferramentas de busca devem incorporar tarefas de coleta de informação acerca de um grupo de sites ou documentos. Nas ferramentas atuais, os usuários devem agregar manualmente as informações que precisa de vários locais heterogêneos, cada qual com suas características de formatação e navegação	Baeza-Yates e Ribeiro-Neto (2013), Branski (2004), Hogan et al. (2011)

Fonte: desenvolvido pela autora

As sugestões e propostas apresentadas pelos autores estudados envolvem três variáveis relevantes para o desenvolvimento desta tese, sendo relacionadas com a interação na elaboração de expressões de busca, indexação do *corpus* documental e interação na interpretação dos resultados apresentados.

3 ONTOLOGIAS

Este capítulo tem como objetivo discutir conceitos e características das ontologias, destacando sua origem no âmbito Filosófico e suas principais contribuições no campo da Ciência da Informação. Além de apresentar os pressupostos para o seu entendimento, discute ainda métodos e técnicas de visualização de ontologias, bem como exemplos de projetos que as utilizaram para o desenvolvimento de sistemas de recuperação de informação e interfaces de busca, destacando contribuições, perspectivas e desafios encontrados.

■■■

O termo "ontologia" teve sua origem no âmbito da Filosofia, mais especificamente no campo da Metafísica, que estuda os tipos de coisas que existem no mundo (ALMEIDA; BAX, 2003). É proveniente da união de duas palavras gregas: "*onto"* (ser) e "*logia*" (estudo ou conhecimento), que indicam o estudo do ser, com suas propriedades e características. Guarino, Oberle e Staab (2009) explicam que ontologia significa o estudo da natureza e da estrutura da realidade, independente de quaisquer considerações posteriores, inclusive, de sua efetiva existência. Para Chauí (2003), é o estudo dos entes, das coisas como verdadeiramente são, isto é, do "Ser".

As ontologias se desenvolveram através dos estudos de diversos filósofos, tais como Aristóteles, Kant e Husserl, com o principal objetivo de categorizar os objetos do mundo para a organização da realidade, especificando conceitos sobre uma determinada área de conhecimento, bem como suas respectivas relações, restrições e axiomas válidos entre esses conceitos e suas instâncias (GUARINO, 1998; THOMASSON, 2018).

Apesar da procedência filosófica, houve uma apropriação do termo "ontologia" por diversos outros campos de pesquisa, contribuindo assim para um distanciamento da sua acepção tradicional e, consequentemente, para o surgimento de uma multiplicidade de conceitos e aplicações, contendo características diferenciadas e particularidades advindas de cada área de conhecimento. Sua importância é reconhecida em diferentes domínios de investigação, principalmente na Ciência da Computação e Ciência da Informação, apresentando estudos com abordagens diferenciadas: Engenharia do Conhecimento, Representação do Conhecimento, Modelagem Qualitativa, Engenharia da Linguagem, *Design* de Banco de Dados, Modelagem de Informações, Análise Orientada à Objetos, Recuperação e Extração de Informação, Gestão e Organização do Conhecimento,

Sistemas Baseados em Agentes, Processamento de Linguagem Natural, entre outros (GUARINO, 1998).

No entanto, apesar do uso e da ampla difusão do termo, entender o significado de "ontologia" não é uma tarefa simples e ainda existem muitas confusões e divergências relacionadas à sua aplicação (ALMEIDA, 2014). Em alguns casos, o conceito é utilizado apenas como um nome fantasia que indica o resultado de atividades como análise conceitual e modelagem de domínio (GUARINO, 1998).

É importante destacar ainda a existência de uma diferença terminológica presente na literatura em relação à grafia da palavra ontologia: Ontologia com 'O' maiúsculo e ontologia com 'o' minúsculo. A primeira refere-se à uma disciplina acadêmica da filosofia que descreve um sistema particular de categorização (Exemplo: Ontologia de curupiras), sendo também adequado denominá-la de metafísica. Em outra vertente, o segundo tipo de grafia prevalece nas áreas da Ciência da Computação e Ciência da Informação e é utilizado predominantemente como esquema de representação do conhecimento, pertencendo assim ao campo de atuação da epistemologia (GUARINO, 1998; GUARINO; NIRENBURG; RASKIN, 2004; OBERLE; STAAB, 2009).

No âmbito da Ciência da Computação, mais especificamente na Inteligência Artificial (IA), o termo ontologia tem sido utilizado desde a década de 60, principalmente em estudos envolvendo Representação do Conhecimento (RC), com o intuito de se referir a uma estrutura de conceitos representados por um vocabulário lógico (ALMEIDA, 2014). Ferneda (2013) afirma que a primeira referência às ontologias foi realizada no artigo *Another look at data*, de George H. Mealy, em 1967. Desde então, o tema "ontologia" vem despertando o interesse de diversos pesquisadores e está sendo amplamente discutido e utilizado na área, "tendo em vista a possibilidade de melhorar significativamente a representação de um domínio de conhecimento" (FERNEDA, 2003, p.26). Para Branco Neto (2006, p.74), "as ontologias representam o ponto mais elevado já atingido em termos de representação, compartilhamento e reutilização do conhecimento".

Guarino (1998, p.4, tradução nossa) faz uma breve distinção sobre a utilização das ontologias no âmbito filosófico e como ferramenta de representação da informação:

> No sentido do filosófico, nós podemos referir a uma ontologia como um sistema particular de categorias que consideram uma certa visão do mundo. Como tal, este sistema não depende de uma linguagem particular: ontologia de Aristóteles é sempre a mesma, independentemente da linguagem utilizada para descrevê-la. Por outro lado, em seu uso mais predominante na IA, uma

> ontologia é referida como um artefato de engenharia, constituído por um vocabulário específico, utilizado para descrever uma certa realidade e um conjunto de pressupostos explícitos relacionados com o significado pretendido para as palavras do vocabulário.

Na vertente da Ciência da Informação, por sua vez, as ontologias começaram a ser utilizadas no final da década de 90, principalmente por pesquisadores da área de Organização do Conhecimento (FERNEDA, 2013), apresentando-se como uma estrutura de representação, recuperação e intercâmbio de informações. Desta forma, possuem como principal objetivo classificar as coisas em categorias, na perspectiva do sujeito e da linguagem do domínio (VITAL; CAFÉ, 2011).

Além disso, estão cada vez mais presentes em projetos e estudos relacionados aos ambientes digitais, como forma de possibilitar melhorias nos processos de organização e recuperação de informação, "garantindo respostas mais consistentes em relação às solicitações de buscas realizadas pelos usuários" (BOCCATO; RAMALHO; FUJITA, 2008, p.199). Possuem diferenças e similaridades com as taxonomias e os tesauros e, em função de uma falta de clareza conceitual são, muitas vezes, confundidas e abordadas de uma forma equivocada.

> As ontologias se colocam como um novo instrumento a ser incorporado ao arsenal teórico e prático da Ciência da Informação. A aprendizagem de novos conceitos e novos recursos oferecidos pelas ontologias é um desafio para os profissionais da informação, mas que pode ser facilmente enfrentado utilizando toda bagagem teórica acumulada durante a história da Ciência da Informação (FERNEDA, 2013, p.41).

Gruber (1993, p.199) define ontologia como sendo "uma especificação formal e explícita de uma conceitualização compartilhada". Segundo ele, essa conceitualização refere-se ao conjunto de conceitos, relações, objetos e restrições que são definidos para um modelo semântico de um domínio de interesse. As ontologias expressam o formalismo dos conceitos e das relações acerca de um domínio, de modo que possa viabilizar a inferência, pela máquina, da semântica aos significados das informações. Assim, "uma ontologia define os termos usados para descrever e representar uma área de conhecimento" (OWL, 2003), ou seja, ela codifica o conhecimento do domínio e também o conhecimento que se entende do domínio, tornando-o reutilizável.

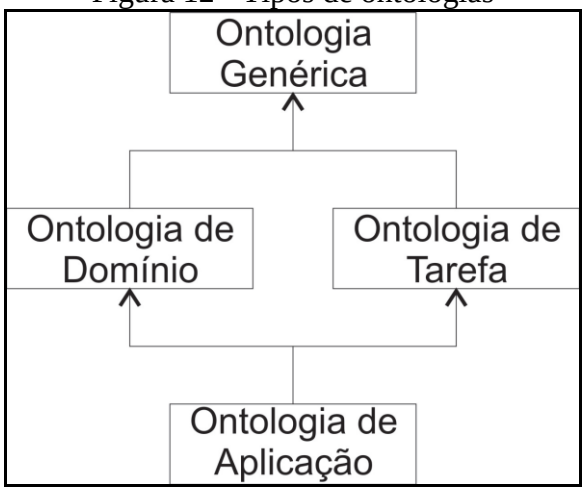
Segundo Gruber (1996), Noy e McGuinness (2001), a ontologia formaliza o conhecimento através da utilização de classes (organizadas em uma taxonomia), relações (representam o tipo de interação entre os conceitos de um domínio), axiomas (usados para

modelar sentenças sempre verdadeiras) e instâncias (utilizadas para representar elementos específicos, ou seja, os próprios dados). Almeida e Bax (2003, p.7) afirmam que uma ontologia "define as regras que regulam a combinação entre termos e relações em um domínio do conhecimento".

Deste modo, pode-se dizer que uma ontologia visa (em alguns aspectos) desenvolver um conjunto de regras que possibilitem a abstração do significado semântico das informações de um determinado domínio disponibilizadas na *Web*. Segundo Fensel (2003), isso oferece vantagens, tais como: possibilitar o compartilhamento e a interoperabilidade do conhecimento entre os domínios, estruturá-los de forma que permita sua compreensão com maior clareza e objetividade e permitir a reutilização dos conceitos abordados.

Guarino (1998) afirma que as ontologias podem ser classificadas de acordo com o seu nível de generalidade ou função para a qual foi desenvolvida, podendo ser genéricas, de domínio, de tarefa e de aplicação, conforme apresenta a Figura 12:

Figura 12 - Tipos de ontologias



Fonte: Guarino (1998, p.9)

Nesta perspectiva, as ontologias de domínio e de tarefa são possíveis especializações de uma ontologia genérica, bem como uma ontologia de aplicação são especializações das duas anteriores. Resumidamente, são apresentadas como:

a) Ontologia de Alto Nível: descrevem conceitos mais amplos, que são independentes de um determinado problema ou domínio, tais como: elementos da natureza, espaço, tempo, etc. De forma geral, são destinadas a grandes comunidades de usuários e possuem o objetivo de construir teorias básicas do mundo, de caráter abstrato, aplicáveis a qualquer domínio (conhecimento de senso comum);

b) Ontologia de Domínio: são os tipos mais comuns e descrevem o vocabulário relacionado a um determinado domínio de conhecimento (medicina, genética, computação) ou a uma atividade genérica (compra, venda), especializando os termos apresentados na Ontologia de Alto Nível;

c) Ontologia de Tarefas: descrevem conceitos que são usados por processos (tarefas e atividades genéricas), sem dependência com um domínio específico, como um processo de compra e venda, por exemplo. Podem contribuir na resolução de problemas, independente do domínio que ocorrem e possuem como principal motivação facilitar a integração dos conhecimentos de tarefa e de domínio em uma abordagem mais uniforme e consistente;

d) Ontologia de Aplicação: descrevem conceitos que dependem tanto de um domínio particular, quanto de uma tarefa específica. São as ontologias mais especialistas e descrevem os papéis que as entidades do domínio assumem quando realizam determinada tarefa.

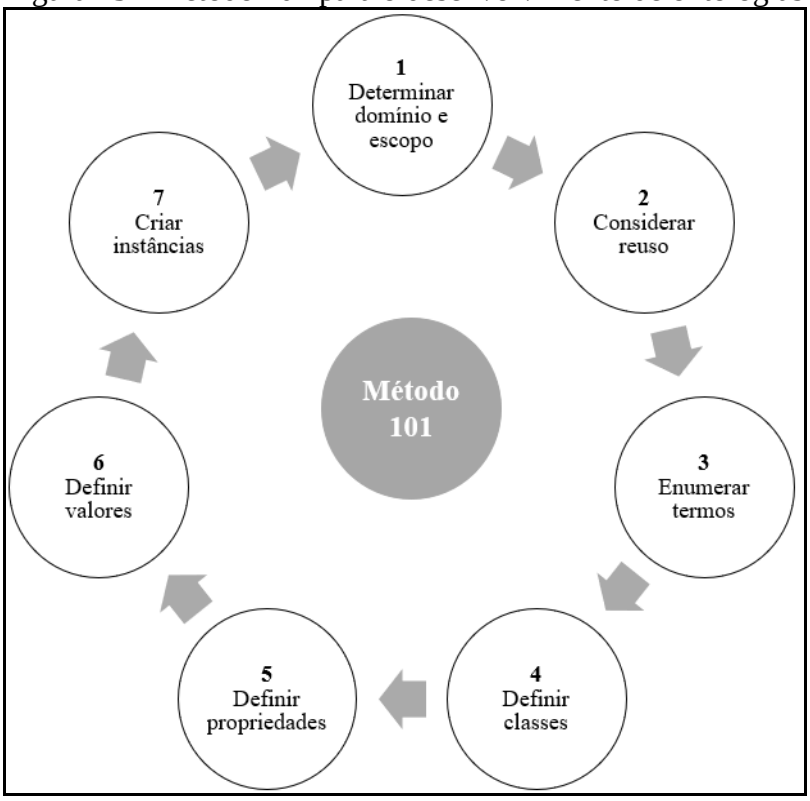
Além da classificação apresentada por Guarino (1998), outras propostas são comumente encontradas na literatura. Alguns autores definem tipos de ontologias relacionando-as ao **grau de formalismo de seu vocabulário**: informal, semiinformal, semiformal e formal (USCHOLD; GRUNINGER, 1996), à sua **aplicação**: de autoria neutra, de especificação e de acesso comum à informação (JASPER; USCHOLD, 1999), à sua **estrutura**: alto nível, domínio e tarefa (HAAV; LUBI, 2001) e ao seu **conteúdo**: terminológicas, de informação, de modelagem do conhecimento, de aplicação, de domínio, genéricas e de representação (VAN HEIJIST; SCHREIBER; WIELINGA, 1997). Apesar de todas essas tipologias, Ruiz e Hilera (2006) alertam que a adoção de apenas um critério de classificação não permite refletir adequadamente sobre a complexidade das ontologias.

Nesta perspectiva, percebe-se que existe uma grande diversidade na literatura acerca das terminologias, conceituações e abordagens sobre ontologias. Os principais autores que estudam o assunto apresentam conceitos e características que são, muitas vezes, semelhantes entre si, mas explicados por meio de nomenclaturas diferenciadas, sem que exista um consenso entre eles. Isto acontece também em relação às metodologias de desenvolvimento das ontologias.

Baskaya, Kekäläinen e Järvelin (2010) afirmam que construir ontologias não é uma tarefa trivial e que a reutilização de conhecimento é sempre muito útil. Existem

inúmeras propostas diferenciadas para apresentar um possível processo de criação de uma ontologia, cada qual com etapas e critérios estabelecidos pelo julgamento do próprio autor, sem a adoção de um padrão oficial de desenvolvimento. Entre elas, pode-se destacar: Método Cyc (REED; LENAT, 2002), Método Sensus (SWARTOUT et al., 1996), Methontology (FERNANDEZ; GOMEZ-PEREZ; JURISTO, 1997) e o Método 101, que é sistematizado na Figura 13 e discutido posteriormente (NOY; McGUINNESS, 2001).

Figura 13 - Método 101 para o desenvolvimento de ontologias



Fonte: Breitman (2005, p.76)

De acordo Noy e Mcguiness (2001), o Método 101 é dividido em quatro etapas principais: i) definição das classes da ontologia; ii) organização das classes em uma hierarquia taxonômica (subclasse e superclasse); iii) definição de relações, atributos e os valores que cada um poderá receber; iv) criação de instâncias pela inserção de valores correspondentes a cada atributo. A metodologia descreve, de forma mais específica, 7 passos imprescindíveis para a construção de ontologias:

a) Passo 1: Determinar o domínio e escopo da ontologia: no início do processo de desenvolvimento, é preciso limitar o escopo e definir claramente os propósitos da ontologia, respondendo os seguintes questionamentos: i) qual é o domínio que a ontologia irá cobrir?; ii)

será desenvolvida para quais propósitos?; iii) para quais informações a ontologia deve fornecer respostas?; iv) quem vai utilizar e manter a ontologia?;

b) Passo 2: Considerar o reuso de ontologias existentes: é importante verificar a existência de outras ontologias no domínio desejado e que possam ser refinadas ou reutilizadas (em sua totalidade ou apenas em parte);
c) Passo 3: Enumerar os termos importantes da ontologia: neste ponto deve-se definir uma relação de termos que representem conceitos do domínio;
d) Passo 4: Definir a hierarquia das classes: refere-se à organização dos termos listados no passo anterior, definindo-se classes e subclasses. Este processo é realizado por meio de três abordagens: topo-para-baixo (*top-down*): inicia-se com a definição das classes mais genéricas do domínio e, posteriormente, com as suas respectivas especializações; baixo-para-cima (*bottom-up*): inicia-se com a definição das classes mais específicas e, posteriormente, com o agrupamento dessas classes em outras mais genéricas; combinação: é a combinação das duas metodologias anteriores (*bottom-up* e *top-down*);
e) Passo 5: Definir propriedades das classes e suas respectivas relações: nesta etapa do processo, define-se as propriedades das classes (atributos), bem como as relações existentes entre elas;
f) Passo 6: Definir os valores das propriedades: neste ponto deve-se definir as características relacionadas à cardinalidade, tipo de dado, domínio e escopo de cada atributo definido da etapa anterior;
g) Passo 7: Criar instâncias: etapa que tem como objetivo a criação de instâncias individuais para cada atributo definido (preenchimento de valores).

Smith et al. (2005) discutem sobre a dificuldade em enumerar as relações de uma ontologia, uma vez que isso implica na aceitação de que algumas são primitivas e independentes de um domínio. Assim, definem três tipos de relações binárias: a relação classe-classe (a relação “é um”), a relação instância-classe (“é instância de”) e a relação instância-instância (por exemplo, a relação “é parte de”). Guarino (1995, p.12) afirma ainda que as relações das ontologias podem ser **estruturantes**, que são as taxonômicas, com o

objetivo de contribuir para a estruturação do domínio e **não estruturantes**, que são aquelas que apresentam relações não-taxonômicas, ou seja, que não oferecem informação estrutural, mas informação adicional aos objetos já identificados.

Como pode ser observado, o processo de desenvolvimento de uma ontologia envolve várias etapas, desde a identificação dos seus objetivos até a sua representação formal (GUARINO; GIARETTA, 1995). Nesta representação é necessária a utilização de uma linguagem que ofereça recursos que suportem a definição de axiomas e possibilitem certas restrições. Diversas linguagens têm surgido com este propósito e, de acordo com Morais e Ambrósio (2007), elas podem ser divididas em três tipos: linguagens de ontologias tradicionais (Cycl, Ontolíngua, F-Logic, CML, OCML, Loom, KIF), linguagens padrão Web (XML, RDF) e linguagens de ontologias Web-based (OIL, DAML+OIL, SHOE, XOL, OWL).

Dentre as inúmeras opções disponíveis, a OWL é recomendada pela W3C desde fevereiro de 2004 como a linguagem padrão para o desenvolvimento de ontologias, oferecendo mecanismos para representar explicitamente o significado dos termos e os relacionamentos existentes entre eles (OWL, 2003). Além disso, expressa mais facilmente significados de semântica que XML, RDF e RDF(s) e representa conteúdo interpretável por máquinas na Web. No entanto, o critério para escolha de uma linguagem ou outra deve ser pautado no tipo de ontologia a ser desenvolvida.

Além de uma linguagem de representação, é necessário ainda a utilização de uma ferramenta de desenvolvimento de ontologias. Almeida e Bax (2003) apresentam inúmeras ferramentas destinadas a este propósito. Entre as mais utilizadas, destacam-se a Protegé, a WebODE e a OntoEdit (MORAIS; AMBRÓSIO, 2007).

> Uma ontologia é criada por especialistas e define as regras que regulam a combinação entre termos e relações em um domínio do conhecimento. Os usuários formulam consultas usando conceitos definidos pela ontologia. O que se busca, em última instância, são melhorias nos processos de recuperação da informação (ALMEIDA; BAX, 2003, p.7).

Uma das características principais das ontologias é que elas possibilitam o compartilhamento da terminologia de um determinado domínio, podendo ser utilizadas por diversas aplicações. Um requisito para que o compartilhamento seja possível é a padronização da linguagem de representação da ontologia. Pickler (2007) afirma que por meio do uso de ontologias é possível elaborar uma grande rede de conhecimento, contemplando o processamento de máquina e melhorando qualitativamente os serviços *web*, sobretudo os de

busca e recuperação de documentos.

> Ontologias são instrumentos de controle terminológico utilizados na recuperação da informação de domínios específicos do conhecimento. No tocante a sistemas informatizados, significam hoje o que há de mais avançado na área de representação do conhecimento e na promoção da interoperabilidade entre sistemas processados por máquinas (DIAS; COSTA, 2011, p.481).

Em relação à aplicabilidade das ontologias, relacionadas principalmente com a área de Recuperação de Informação, McGuinness (2003) e Ferneda (2013) destacam a possibilidade de desenvolvimento de sistemas contendo os seguintes recursos:

a) checagem de consistência: as informações obtidas em uma busca podem ser associadas às definições de uma ontologia para verificar se os valores retornados são válidos ou não;
b) complemento de informações: após obter um pequeno conjunto de informações sobre determinada entidade, as ontologias podem ser utilizadas para gerar novas informações com base nas previamente encontradas;
c) resolução de problemas de linguagem (sinonímia, homonímia, paronímia, polissemia e variações linguísticas): se o mesmo termo aparece em mais de um lugar, ele está relacionado a uma classe e subclasse correspondente que pode ajudar a distinguir entre os vários contextos do termo;
d) indexação automática: a representação de um documento é realizada automaticamente por termos derivados de uma ontologia;
e) expansão de consulta: adição de novos termos à busca do usuário provenientes de uma ontologia;
f) sistemas de recuperação de informação semânticos: os documentos são previamente anotados (marcados) de acordo com uma ontologia de domínio; e
g) interfaces de busca: os conceitos de uma ontologia são apresentados ao usuário que seleciona aqueles que serão utilizados como termos de busca.

Apesar das diversas possibilidades de aplicação de ontologias na recuperação de informação, o tema ainda é considerado incipiente e apresenta grande

potencial para o desenvolvimento de estudos e aplicações na área. As pesquisas apontam que a principal utilização das ontologias está relacionada à representação dos documentos, mais especificamente, ao processo de indexação automática, extração de informação e expansão de consultas. Poucos trabalhos utilizam este recurso para as outras finalidades especificadas (FERNEDA, 2013).

Na próxima seção serão apresentados princípios, métodos e técnicas para a visualização de ontologias, bem como uma breve discussão acerca das vantagens e desvantagens de cada abordagem.

3.1 PRINCÍPIOS PARA A VISUALIZAÇÃO DE ONTOLOGIAS

Os estudos de visualização de ontologias estão presentes em diversas áreas, incluindo a Ciência da Computação e Ciência da Informação (MENDES NETTO; LIMA, 2017). Estão relacionados diretamente com o campo da Visualização de Informação, que fornece subsídios e contribuições para as pesquisas realizadas. Para Freitas et al. (2001), a Visualização da Informação é uma área de aplicação de técnicas da computação gráfica, geralmente interativas, visando auxiliar o processo de análise e compreensão humana de um conjunto de dados, por manipulações das representações gráficas.

Seu principal objetivo é utilizar representações gráficas para facilitar o entendimento e a interpretação de grande quantidade de informação de forma compacta, a partir de diferentes perspectivas e com suporte a vários níveis de detalhamento (CARD; MACKINLAY; SHNEIDERMAN, 1999).

> A área de visualização tem um potencial atraente quando se trata de criar, explorar ou verificar coleções complexas e grandes quantidades de dados, como as ontologias. Em particular, a Visualização de Informações (InfoVis) que lida com informações abstratas e dados não-espaciais, oferece um conjunto de técnicas para representar dados hierárquicos ou semiestruturados. Portanto, não é surpresa que muitas ferramentas de ontologia integrem recursos relacionados à visualização (LANZENBERGER; SAMPSON; RESTER, 2010, p.1036-1037, tradução nossa).

A literatura apresenta inúmeros métodos e técnicas para a visualização de ontologias, que são frequentemente representadas de forma gráfica, uma vez que elas compreendem um conjunto de conceitos interligados que descrevem um domínio específico (KATIFORI et al., 2007; BURCH; LOHMANN, 2015). De acordo com Katifori et al. (2007), é importante que a visualização de ontologias seja eficiente e permita, ao mesmo tempo, a

exibição de informações e uma boa interação com o usuário, o que é uma tarefa bastante complexa de ser desenvolvida. Para que isto seja possível, Burch e Lohmann (2015) destacam a necessidade do desenvolvimento de projetos visuais avançados, que se preocupem com o tamanho e a complexidade crescente das ontologias e que sejam direcionados ao aprimoramento de *layouts*, tornando-os mais rápidos e capazes de lidar com as percepções do usuário e com os problemas visuais de escalabilidade.

Diversas abordagens sobre a visualização de ontologias foram realizadas nos últimos anos, com inúmeras pesquisas desenvolvidas e trabalhos publicados na área, tais como: Mutton e Golbeck (2003), Katifori et al. (2007), Lanzenberger, Sampson e Rester (2009; 2010), Bera, Burton-Jones e Wand (2011), Dudáš, Zamazal e Svátek (2014), Burch e Lohmann (2015), Saghafi (2016), dentre outros. O assunto também tem sido objeto de discussão de *workshops*, tais como o Voila! (2015, 2016, 2017). São estudos de grande relevância, uma vez que possuem como principal objetivo fornecer aos usuários representações visuais e técnicas de interação intuitiva que possam auxiliar na exploração e compreensão dos domínios e conhecimentos representados por ontologias, além de possibilitar mecanismos diferenciados de interação, como é a proposta deste estudo.

A necessidade de uma visualização eficaz, com foco no *design*, gerenciamento e navegação, surgiu com a expansão do uso de ontologias em diversos contextos e áreas de aplicação, que compreendem desde pequenos sistemas, como uma coleção de documentos digitais, a sistemas maiores como os disponibilizados na Web. Também se faz importante no contexto da Web Semântica, mais especificamente em aplicações envolvendo o Linked Open Data, uma vez que os especialistas de domínio precisam ser capazes de compreender as ontologias que fazem parte deste contexto (BOJARS et al., 2016).

Katifori et al. (2007) afirmam que uma boa visualização de ontologias deve ser capaz de possibilitar que o usuário, sem esforço, consiga observar os seguintes elementos:

a) Classes: deve ser capaz de exibir todas as classes da ontologia de uma só vez ou a critério do utilizador, fornecendo pelo menos o nome delas de forma legível;
b) Instâncias: por serem dados reais associados à ontologia, a maior parte dos usuários está interessada nessa informação, embora muitos métodos de visualização não apresentem este recurso. Quando disponibilizadas, são apresentadas no formato de nós conectados a uma classe (nem sempre são eficazes) ou uma lista dentro de uma janela separada;

c) Taxonomia (relações Isa): a apresentação da taxonomia da ontologia é fundamental para o entendimento da relação de herança entre as classes. Embora não sejam fáceis de serem representadas, é importante que se tenha, ao menos, uma visão holística sobre elas;

d) Herança Múltipla: refere-se aos casos em que uma classe tem mais de um pai, e assim como as taxonomias, não são fáceis de serem efetivamente representadas. É importante que a visualização indique os nós com múltiplos pais e forneça a visualização eficaz de todos os seus ascendentes;

e) Relações de função: assim como a taxonomia e a herança múltipla, as relações de função são difíceis de serem representadas, mas são de grande importância. Além do *link*, que deve ser visível, um rótulo com o tipo de função também deve ser exibido, alternativamente, escondido, para evitar exibir desordem;

f) Propriedades: as propriedades associadas a uma entidade também devem ser incluídas na visualização, seja na ontologia principal ou em um espaço separado.

De acordo com Lanzenberger, Sampson e Rester (2010), as ferramentas com este propósito são capazes de exibir desde ontologias simples até bases de conhecimento altamente estruturadas e com relações complexas. Embora a literatura apresente diferentes classificações e nomenclaturas, os métodos de visualização de ontologias costumam ser agrupados de acordo com alguns critérios, tais como: características da apresentação, técnicas de interação, funcionalidade suportada ou dimensões da visualização (KATIFORI et al., 2007).

As técnicas de visualização também são variadas, envolvendo diversos recursos, tais como: hierarquias (listas indentadas), nó-*link* e árvores, aproximação, preenchimento de espaço; técnicas de distorção orientada (efeito *fisheye*, contexto+foco e distorção) e paisagens (KATIFORI et al., 2007; LANZENBERGER; SAMPSON; RESTER, 2010; BURCH; LOHMANN, 2015; SAGHAFI, 2016).

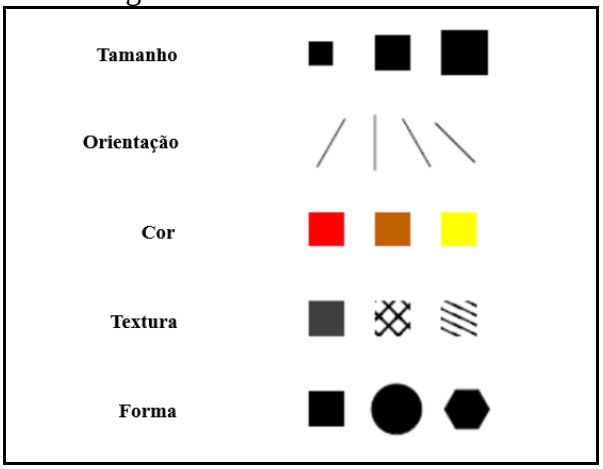
Shneiderman (1996) cita sete tipos de ações possíveis sobre a visualização de informações, que também podem ser adaptadas ao contexto da visualização de ontologias: a) **visão geral**: permitir uma visão geral da ontologia, bem como detalhes de áreas selecionadas para análise; b) **filtro**: possibilidade de informação conforme o interesse do usuário; c) ***zoom***: ampliar os itens de interesse do usuário; d) **detalhes sob demanda**: possibilidade de selecionar um item

ou um grupo de itens e obter informações relevantes sobre ele; e) **relação**: identificar as relações existentes entre os elementos da ontologia, conforme as características selecionadas pelos usuários; f) **histórico**: manter um histórico das ações realizadas e permitir o retorno a um estágio anterior ao atual; g) **extrações**: permitir a elaboração de consultas para recuperar subconjuntos da ontologia, conforme os parâmetros especificados pelo usuário.

Para facilitar a análise e o reconhecimento visual das informações apresentadas nas ontologias, alguns recursos são agregados ao processo de visualização, tais como: tamanho, textura, cor, espessura, valor, orientação, forma, etc. Estes elementos são considerados artefatos cognitivos, na medida em que complementam e fortalecem as habilidades mentais dos usuários e podem contribuir para o processo de RI, considerando que: a) auxiliam os usuários na organização das informações; b) possibilitam a seleção de informações importantes para análise; c) ajudam a identificar as relações existentes entre as informações apresentadas; d) permitem comparações visuais; e) relevam padrões e tendências; f) podem facilitar a descoberta e a inferência de informações; g) podem auxiliar na manipulação dos dados (CARD; MACKINLAY; SHNEIDERMAN, 1999; MEIRELLES, 2013).

Mazza (2009) afirma que os elementos gráficos utilizados na visualização de informações são sensíveis à retina do olho e estão diretamente relacionados às capacidades de percepção humana. Os mais comuns são apresentados na Figura 14.

Figura 14 - Elementos Gráficos

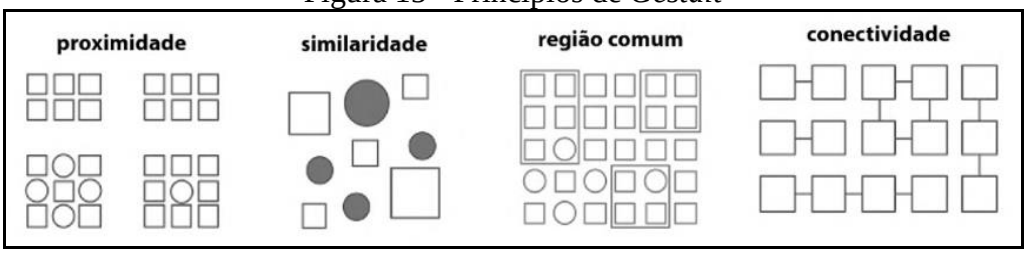


Fonte: Mazza (2009, p.21)

Além dos elementos gráficos, os princípios de *Gestalt* também são importantes para a visualização, uma vez que fornecem uma estrutura para a compreensão da interface apresentada, a partir da perspectiva do usuário. As leis mais funcionais para o processo de análise de ontologias e recuperação de informação são: proximidade,

similaridade, região comum e conectividade (WARE, 2003; KOSHMAN, 2006). A Figura 15 apresenta um exemplo de cada um destes princípios.

Figura 15 - Princípios de *Gestalt*



Fonte: Barbosa e Silva (2010, p.51)

O princípio da proximidade indica que os objetos (ou entidades visuais) que estão localizados próximos uns dos outros, são interpretados como um grupo ou uma unidade. Desta forma, os usuários conseguem distinguir com mais facilidade grupos ou *clusters* de itens quando estão analisando dados de uma ontologia. É um recurso também muito utilizado no *design* de interfaces.

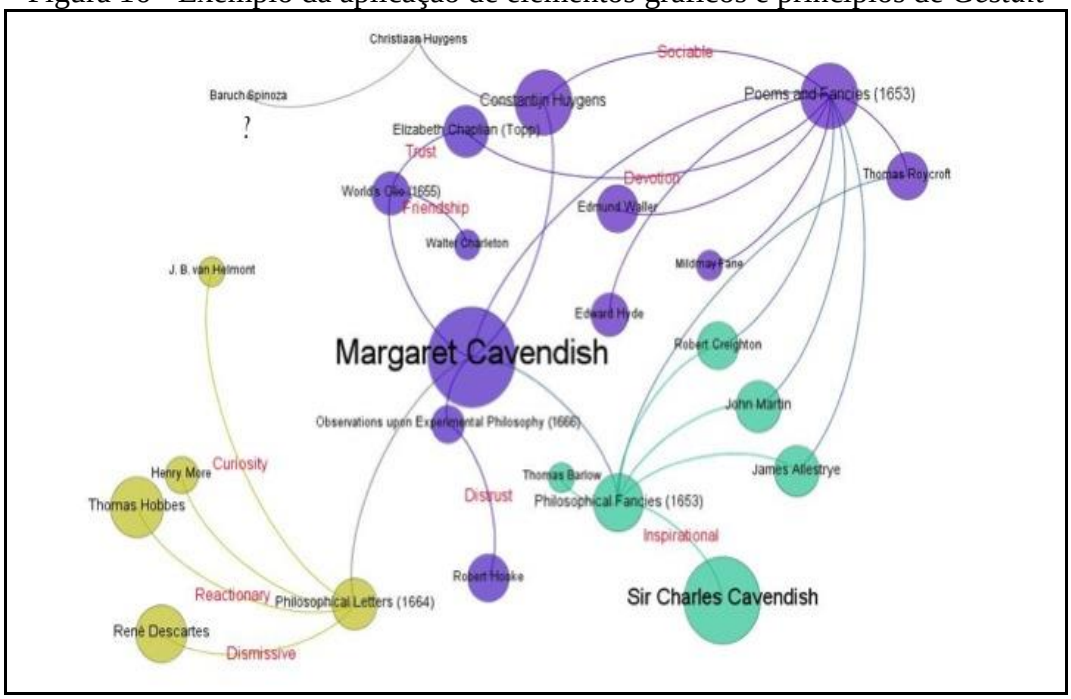
O princípio da semelhança, por sua vez, indica que objetos semelhantes que compartilham os mesmos atributos visuais pertencem a um mesmo grupo. Quando unidos os princípios da proximidade e da semelhança na mesma visualização, a interpretação do usuário sobre o compartilhamento de características do grupo de informações fica ainda mais evidente.

Outro princípio que também é utilizado para indicar elementos pertencentes a um mesmo grupo é o de região comum. De acordo com esta premissa, os objetos apresentados dentro de uma região espacial, são percebidos pelo usuário como um conjunto. Por último, o princípio da conectividade indica que objetos conectados por traços contínuos são entendidos como elementos relacionados (KOSHMAN, 2006; HEARST, 2009; BARBOSA; SILVA, 2010).

A Figura 16 apresenta um exemplo da aplicação de elementos gráficos e de princípios de *Gestalt* na visualização das informações referentes à trajetória de Margaret Cavendish, Duquesa de Newcastle (1623-1673). Os recursos gráficos que podem ser facilmente identificados são os que se referem à cor e tamanho. A cor foi utilizada para destacar a presença de três *clusters* de informações conectados entre si: o roxo, que possui elementos diretamente relacionados ao centro do gráfico e contendo a maior quantidade de nós; o verde e o amarelo. O tamanho foi empregado para destacar a relevância que cada elemento ocupa em relação ao contexto analisado, ou seja, nós maiores são mais

representativos e indicam maior importância (ou maior valor), enquanto que nós menores indicam menor expressividade.

Figura 16 - Exemplo da aplicação de elementos gráficos e princípios de *Gestalt*



Fonte: Moore (2013)

Em relação aos Princípios de *Gestalt*, a proximidade e a conectividade são muito fáceis de serem percebidas. Os nós de cada *cluster* de informações estão agrupados próximos uns dos outros, e conectados por linhas que indicam os relacionamentos existentes entre eles.

A seção a seguir apresenta inúmeras técnicas para a visualização de informações presentes nas ontologias, sendo que os elementos gráficos e os Princípios de *Gestalt* são aplicados, separadamente ou em conjunto, na maior parte delas.

3.2 VISUALIZAÇÃO GRÁFICA DE ONTOLOGIAS

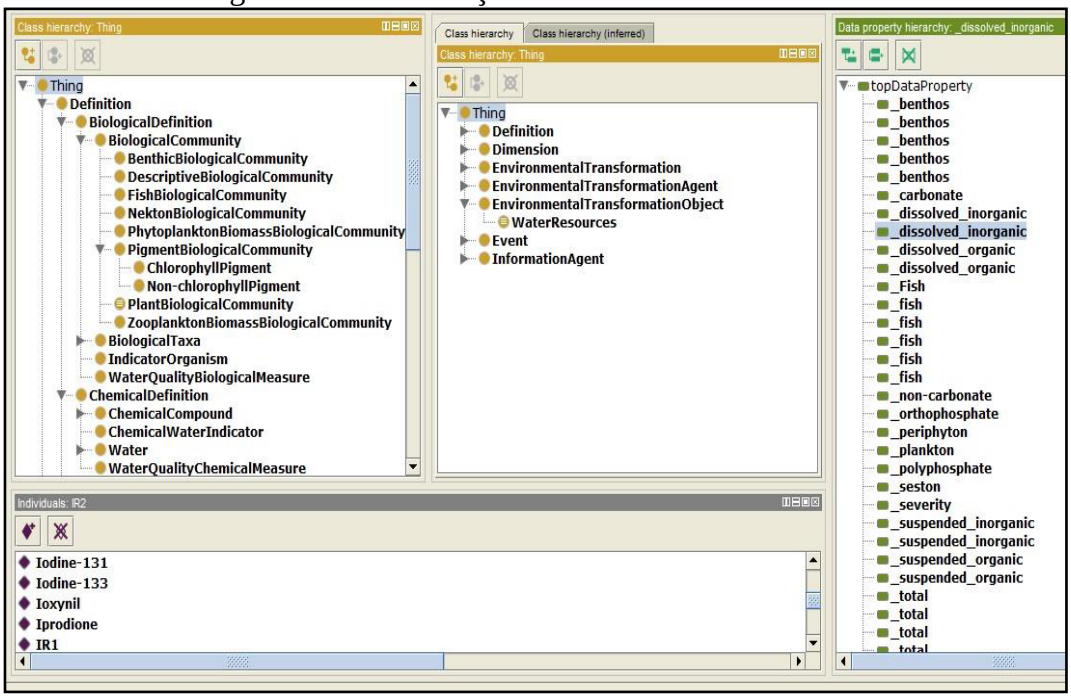
No contexto deste estudo foram analisados os métodos apresentados por Katifori et al. (2007), Lanzenberger, Sampson e Rester (2010), Burch e Lohmann (2015) e Saghafi (2016), que suportam grandes quantidades de dados ontológicos e que podem ser adaptados para o contexto da Web. Foram subdivididos em seis categorias, de acordo com as características predominantes e técnicas suportadas: Visualização Baseada em Lista Indentada, Visualização Baseada em Nó-*Link* e Árvore, Visualização Baseada em Aproximação (*Zoomable*), Visualização Baseada em Preenchimento de Espaço (*Space*

Filling), Visualização Baseada em Contexto+Foco e Distorção (*Context+Focus and Distortion*) e Visualização Baseada em Paisagem (*Information Landscapes*).

3.2.1 Visualização Baseada em Lista Indentada (*Indented List*)

Os métodos de visualização baseados em lista indentada apresentam a taxonomia da ontologia em uma estrutura semelhante à organização dos diretórios de arquivos (pastas) do Windows Explorer, ou seja, as classes são representadas como uma árvore retrátil. Por ser eficaz na recuperação de informação, a maioria das ferramentas de visualização de ontologias possui este recurso como principal técnica de visualização, incluindo a Protegé, OntoEdit e OntoRama (KATIFORI et al., 2007; MENDES NETTO; LIMA, 2017). Sua principal característica é possuir uma interface de navegação simples e rápida, que permite a exploração sistemática de toda a ontologia, bem como a visualização detalhada da estrutura (classes e instâncias) de um elemento específico. Um exemplo deste tipo de visualização pode ser observado na Figura 17.

Figura 17 - Visualização baseada em lista indentada



Fonte: Mendes Netto, Lima e Pierozzi Júnior (2016, p.11)

A visualização baseada em lista indentada apresenta algumas vantagens, entre as quais se destacam: simplicidade de implementação e familiaridade do usuário, uma vez que o conceito de navegação implementado é usado em diversos navegadores de

arquivos, diretórios e páginas da Web. Também se destaca em comparação com quase todas as outras técnicas por não haver sobreposição de rótulos e não ser necessário mover o mouse sobre um item para visualizar o seu rótulo. A retração e expansão de nós é um recurso particularmente útil por possibilitar o foco em partes específicas da hierarquia, principalmente para grandes ontologias, facilitando a localização de uma classe ou instância específica (KATIFORI et al., 2007).

No que diz respeito aos problemas observados, este tipo de visualização não possui representação visual da ontologia. Desta forma, exibe apenas as relações de herança (isa) e não as relações de função. Além disso, os casos de herança múltipla não são muito óbvios. Outro ponto importante refere-se à visualização parcial das grandes ontologias, ou seja, apenas uma parte delas pode ser visível de uma só vez, sendo necessário a utilização de barra de rolagem para a navegação nos demais itens (PLAISANT; GROSJEAN; BEDERSON, 2002; KATIFORI et al., 2007).

Em função de suas características, a visualização baseada em lista indentada não é indicada para tarefas relacionadas à estrutura geral de uma ontologia, como identificar a profundidade da hierarquia ou encontrar nós com muitos filhos ou hierarquias profundas. No entanto, tem sido comprovado em diversas avaliações que este tipo de visualização parece ter um desempenho melhor do que as outras usadas para hierarquias (KOBSA, 2004).

3.2.2 Visualização Baseada em Nó-*Link* e Árvore (*Node-Link and Tree*)

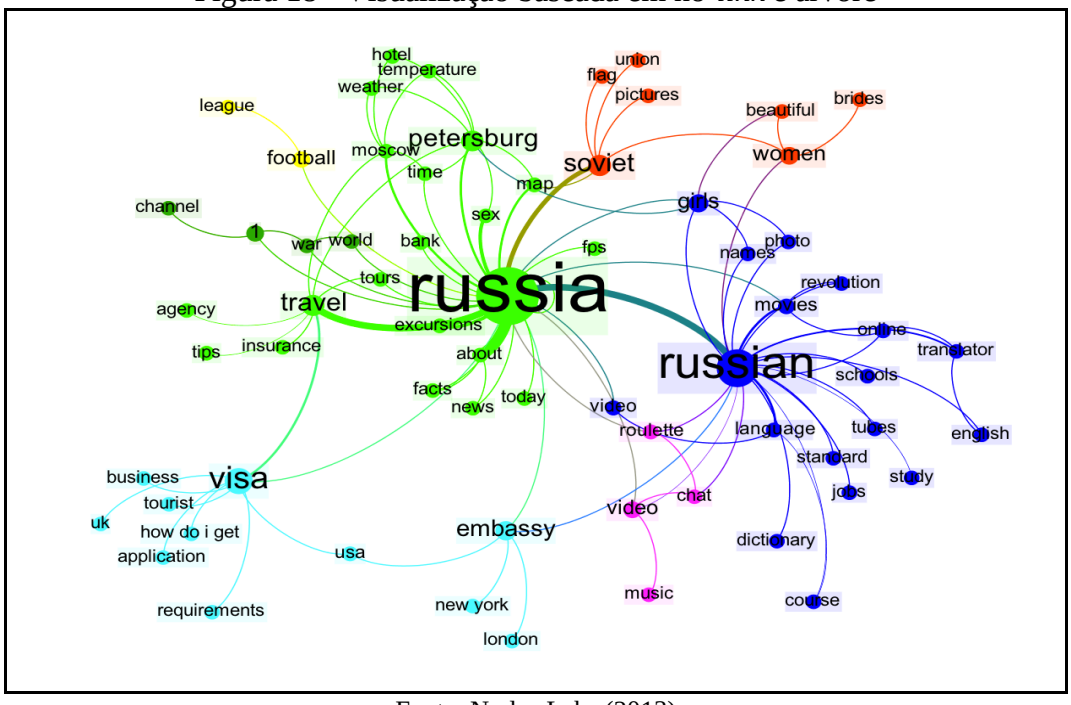
A principal característica da visualização baseada em nó-*link* e árvore é que a ontologia é representada como um conjunto de nós interconectados (grafos) que geralmente podem ser expandidos e retraídos pelo usuário, a fim de organizar e ajustar os detalhes da informação apresentada. Neste tipo de visualização, cada classe da ontologia é considerada um nó e os *links* do grafo representam as relações existentes entre elas. É útil para a tarefa de identificar conceitos semelhantes em uma ontologia de domínio e permite que os usuários tenham flexibilidade para analisar as ontologias e suas relações de forma gráfica e reorganizada suavemente à medida em que as interações vão sendo realizadas. Possui a capacidade de apresentar até 50 nós por vez (SAGHAFI, 2016).

Os métodos de visualização baseados em nó-*link* e árvore utilizam diversos algoritmos para produzir o *layout* da árvore, seja de cima para baixo ou da direita para a esquerda. Entre os mais utilizados, estão os algoritmos do tipo *force-directed* (digirido por força) e *spring-embedder* (impulsão por mola), que atuam com o objetivo de aproximar os nós

com características semelhantes (BRANDES, 2001; MUNZNER, 2014) e, com isso, otimizar o seu posicionamento, principalmente quando existe dificuldade do usuário em entender a estrutura geral do grafo em função do grande número de bordas de cruzamento (MAZZA, 2009).

A visualização baseada em nó-*link* e árvore, como pode ser observado na Figura 18, oferece ao usuário uma boa visão das estruturas hierárquicas, permitindo a fácil distinção de diferentes níveis e características como profundidade e largura. Pode ser apresentada de forma bidimensional (visualização mais simples) ou tridimensional (visualização mais complexa). As ferramentas com recursos bidimensionais geralmente podem apresentar as seguintes informações de uma ontologia: classes, instâncias, propriedades, relações de herança e papéis. Entre elas, destacam-se: OntoViz, IsaViz, SpaceTree, entre outras. As tridimensionais, por sua vez, apresentam melhores recursos de visualização e navegação, permitindo diferentes visões e níveis de detalhamento da ontologia. Algumas ferramentas desta categoria são: Cone Tree, Tree Viewer e OntoSphere (KATIFORI et al., 2007).

Figura 18 - Visualização baseada em nó-*link* e árvore



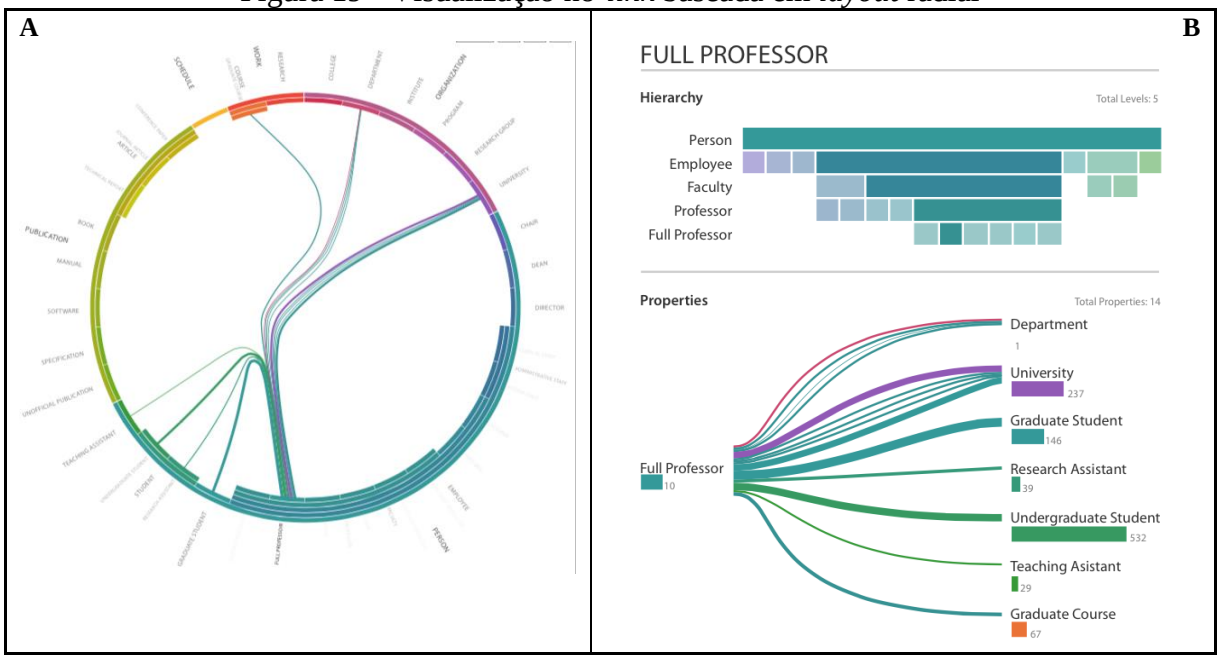
Fonte: Nodus Labs (2013)

Entre os problemas observados, os métodos de visualização nó-*link* e árvore geralmente apresentam sobrecarga de informações na tela, sendo que algumas árvores precisam de múltiplas telas para serem completamente exibidas, pois apenas parte do

diagrama é visível em um mesmo momento. Em função disto, outros obstáculos são identificados pelos usuários, tais como: desordem dos nós, dificuldades na interação e navegação, falta de visibilidade de rótulos e ausência de ordenação (as instâncias não são apresentadas em ordem alfabética ou em qualquer outra ordem). Outro ponto importante que dificulta a interação com as ferramentas é a falta de familiaridade dos usuários com este tipo de visualização (HEARST, 2009).

Apesar das dificuldades apresentadas, Hearst (2009) afirma que existem vários métodos interativos que podem auxiliar na melhoria da visualização das informações, tornando as redes mais compreensíveis. Uma possibilidade importante, principalmente para as redes de tamanho médio (aproximadamente 100 nós), é permitir que o usuário selecione um nó de interesse e faça dele o foco de exibição, reorganizando os demais ao seu redor. Outra abordagem é eliminar a maioria dos nós diferentes daqueles que rodeiam o nó selecionado. Uma terceira alternativa é alterar o formato de visualização do grafo de nó-*link*, organizando os nós em círculo, sendo que os *links* podem cruzar o centro do círculo ou conectar-se a outros nós, conforme apresentado na Figura 19.

Figura 19 - Visualização nó-*link* baseada em *layout* radial



Fonte: Saghafi (2016, p.5)

A figura está dividida em duas partes e apresenta a visualização de uma ontologia em *layout* radial, no formato de círculos hierárquicos conectados. A parte A mostra a ontologia apresentada como um todo e a parte B, a ampliação de uma classe selecionada. Neste tipo de visualização, os elementos externos do círculo representam as classes de nível

superior, sendo que cada uma delas é codificada com uma cor diferente. As subclasses são visualizadas nos círculos internos, codificados com diferentes níveis de saturação da cor dos pais. A relação entre classes também é modelada como *links*, que conectam os diferentes segmentos do círculo. Na parte B, localizada à direita da figura, é possível observar um exemplo de filtro de informações, a partir da seleção de uma pequena seção do círculo, no qual o usuário ampliou a classe "Professor" com o objetivo de visualizar as características relacionadas a ela (SAGHAFI, 2016).

Apesar da variedade de *layouts* disponibilizados na literatura, as visualizações baseadas em nó-*link* e árvore são eficazes apenas para oferecer uma visão geral da estrutura da ontologia, mas não para tarefas relacionadas à localização de nós específicos. Desta forma, inúmeros recursos estão sendo estudados para a melhoria das ferramentas, incluindo: filtros, retração e expansão.

3.2.3 Visualização Baseada em Aproximação (*Zoomable*)

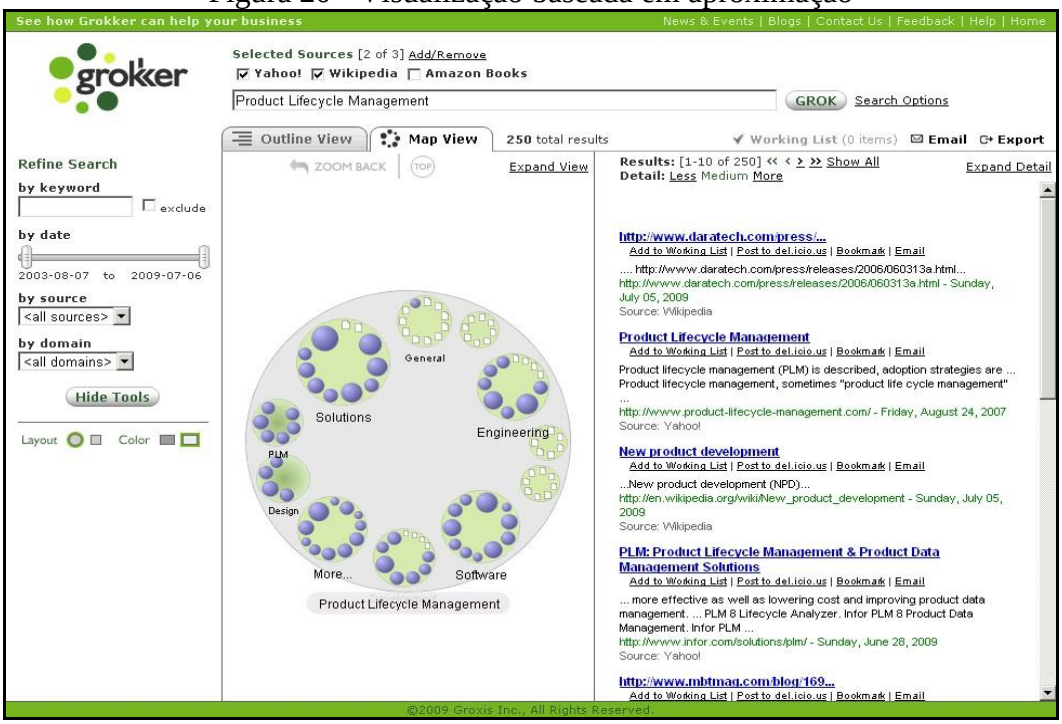
A visualização baseada em aproximação tem como principal característica a apresentação de informações em um plano bidimensional (2D) de diferentes tamanhos, onde os nós secundários (filhos) são agrupados no interior dos nós principais (pai), podendo ser ampliados ou visualizados pelo usuário por meio da utilização do recurso de *zoom*. Katifori et al. (2007) afirmam que avaliações realizadas indicam ser uma técnica eficiente para localizar nós específicos e fornecer uma visão abrangente do nível de hierarquia do usuário. Um exemplo de utilização dessa técnica é a interface do Grokker, apresentada na Figura 20.

O Grokker é um mecanismo de busca visual disponibilizado na Web, que disponibiliza os resultados de forma gráfica, podendo ser classificados por data, fonte, domínio e refinados a partir da seleção ou exclusão de palavras-chave específicas. O mecanismo de agrupamento apresenta os documentos como uma série de círculos aninhados, permitindo que os usuários possam navegar na hierarquia e tornar os conteúdos visíveis por meio de cliques nos círculos (KOSHMAN, 2006). Além do Grokker, existem diversas outras ferramentas que apresentam visualização baseada em aproximação (*zoomable*), destacando-se: Jambalaya, Information Cube, Information Pyramids e GopherVR.

Entre os problemas observados com este tipo de visualização, Katifori et al. (2007) destacam que a navegação pode ser tornar confusa, principalmente quando o usuário não consegue identificar qual nó foi ampliado ou se um determinado nó já foi visitado. Além disso, os rótulos se sobrepõem no caso de muitas instâncias, dificultando a visualização das

informações. Outro ponto importante refere-se à utilização de métodos de navegação não convencionais e pouco intuitivos, que acabam sobrecarregando a carga cognitiva do usuário.

Figura 20 - Visualização baseada em aproximação



Fonte: Dassault Systèmes (2009)

Rivadeneira e Bederson (2003) sugerem que a visualização baseada em aproximação poderia ser melhorada com a implementação de alguns recursos adicionais, tais como a apresentação de pistas de navegação que informariam os usuários quais elementos já foram visitados e pistas hierárquicas que poderiam informar em que nível eles estão e quão profunda é a estrutura.

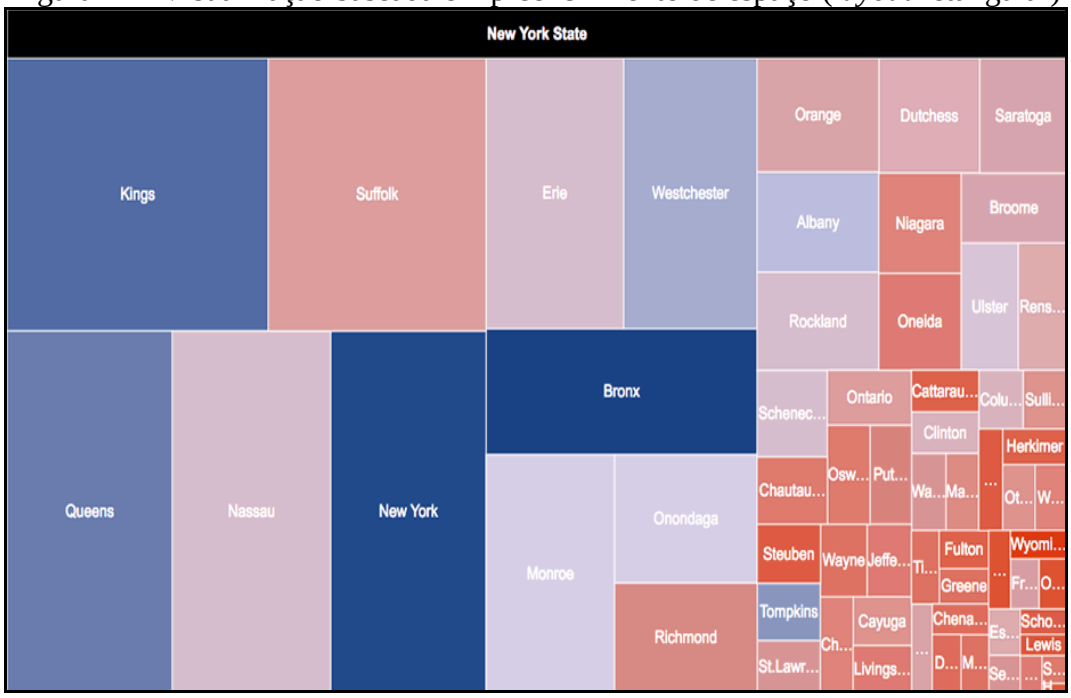
3.2.4 Visualização Baseada em Preenchimento de Espaço (*Space Filling*)

A visualização baseada em preenchimento de espaço tem como principal característica utilizar todo o espaço da tela, subdividindo o espaço disponível entre os nós de uma hierarquia. O espaço atribuído a um nó-pai é subdividido entre os seus nós-filhos. O tamanho e a cor de cada subdivisão correspondem a uma propriedade da informação que se deseja representar.

As visualizações baseadas em preenchimento de espaço podem ser bidimensionais ou tridimensionais. Entre as abordagens que utilizam recursos 2D, encontram-se os métodos: TreeMaps, SequoiaView e Information Slices. Em relação aos recursos 3D, o

principal destaque é para BeamTrees, que utiliza técnicas de preenchimento de espaço, juntamente com a visualização de nó-*link* em 3D (KATIFORI et al., 2007). Podem ainda adotar diferentes técnicas de abordagem de preenchimento da área, incluindo algoritmos de *layout* retangular ou radial (SHNEIDERMAN, 1992; SAGHAFI, 2016). Um exemplo de visualização baseada em preenchimento de espaço de *layout* retangular, do tipo TreeMap, pode ser observado na Figura 21.

Figura 21 - Visualização baseada em preenchimento de espaço (*layout* retangular)



Fonte: Uptown/Downtown Media LLC (2016)

Os dados apresentam o resultado das eleições presidenciais americanas de 2016, no estado de Nova York. O tamanho de cada quadrado indica a porcentagem de votos totais de cada candidato. Os tons de azul representam a quantidade de votos de Hillary Clinton e os tons de vermelho representam os votos a Donald Trump.

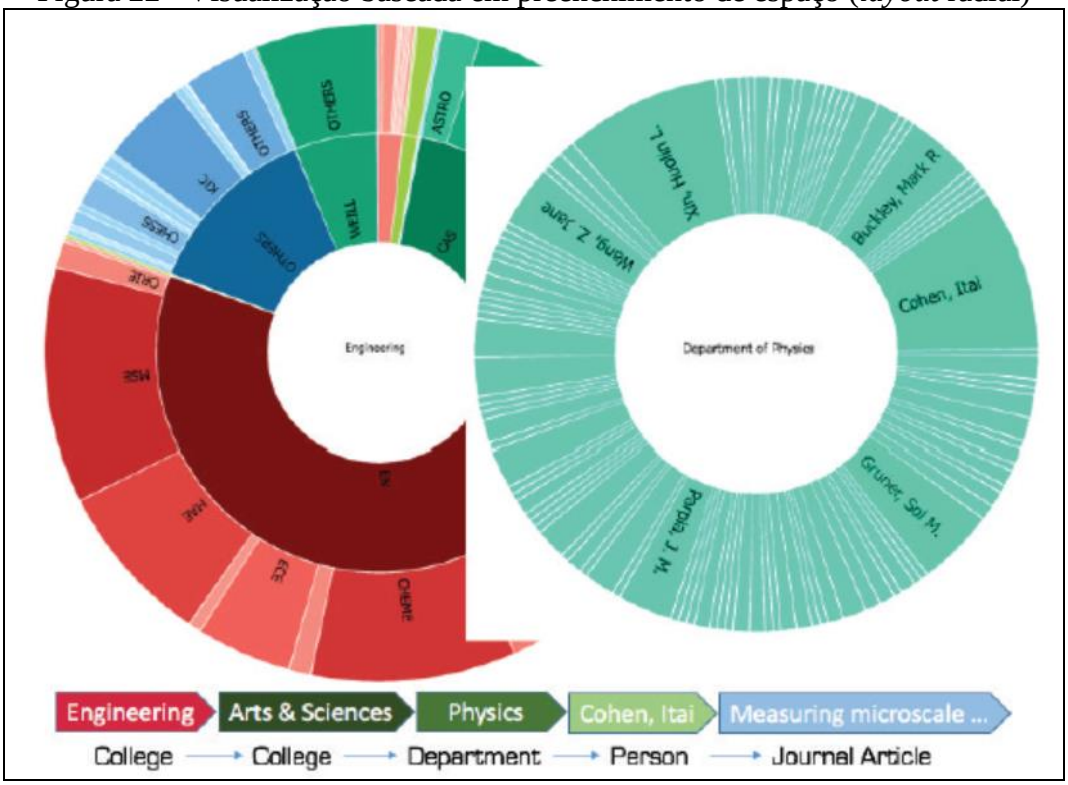
De acordo com Hearst (2009), a técnica de visualização TreeMap ou Mapas de Árvores, que é uma das mais populares e discutidas da categoria, foi desenvolvida inicialmente com a intenção de exibir informações hierárquicas de grande profundidade, usando retângulos aninhados ou camadas. Na prática, ela é utilizada apenas para apresentar dois ou três níveis de hierarquia. Cada nó é representado por um retângulo, que é preenchido por retângulos menores, que são os nós-filhos. O tamanho de um retângulo é proporcional à sua frequência relativa ou à importância do nó subjacente. Quando projetado com cuidado e aplicado a uma tarefa apropriada, o TreeMap pode fornecer uma visão geral útil de uma

coleção de informações, que por si só é de difícil acesso. No entanto, os Mapas de Árvores não foram bem-sucedidos na apresentação de dados textuais.

Entre os problemas observados, Katifori et al. (2007) afirmam que a visualização TreeMap apresenta dificuldades na representação da informação presente nos nós internos da árvore, principalmente quando se trata de hierarquias mais complexas. Além disso, a utilização deste tipo de técnica exige treinamento do usuário, uma vez que eles não estão familiarizados com o *layout* disponibilizado.

Por outro lado, as visualizações baseadas em preenchimento de espaço de *layout* radial apresentam algumas vantagens quando comparadas com as de *layout* retangular, pois facilitam a organização da camada adequada para cada nó e também permitem uma melhor distinção entre as diferentes categorias. Um exemplo deste tipo de visualização é apresentado na Figura 22.

Figura 22 - Visualização baseada em preenchimento de espaço (*layout* radial)



Fonte: Javed et al. (2016, p.86)

A principal característica das visualizações baseadas em preenchimento de espaço de *layout* radial é a utilização de discos completos que permitem a navegação rápida em hierarquias compostas de múltiplos níveis, representados por cores diferenciadas. No exemplo apresentado na Figura 22, é possível observar dados de uma base de conhecimento acadêmica, contendo três camadas de informações: camada de nível universitário (mais

interna), camada de nível departamental (meio) e camada de nível pessoal (mais externa). Para explorar informações relacionadas a um determinado pesquisador, é necessário iniciar a navegação na camada de nível universitário, explorar o departamento de interesse e, posteriormente, selecionar um dos pesquisadores disponíveis, bem como os respectivos trabalhos associados a ele, que podem ter sido desenvolvidos de forma individual ou colaborativa (JAVED et al., 2016).

Katifori et al. (2006) afirmam que as visualizações que utilizam *layout* radial podem apresentar problemas na representação de nós do mesmo nível. Além disso, o ângulo utilizado para representar cada fatia é proporcional às propriedades da ontologia que ela representa, podendo resultar em espaços muito finos que contrapõem a outros mais densos, dificultando assim a leitura dos atributos exibidos.

Desta forma, as técnicas de preenchimento de espaços são adequadas para a representação visual de certas propriedades de uma ontologia ao nível de instância ou quando existe a preocupação apenas com propriedades dos nós de nível hierárquico mais baixo, sem se ater à estrutura da ontologia como um todo, não sendo, portanto, tão eficazes para este propósito.

3.2.5 Visualização Baseada em Contexto+Foco e Distorção (*Context+Focus and Distortion*)

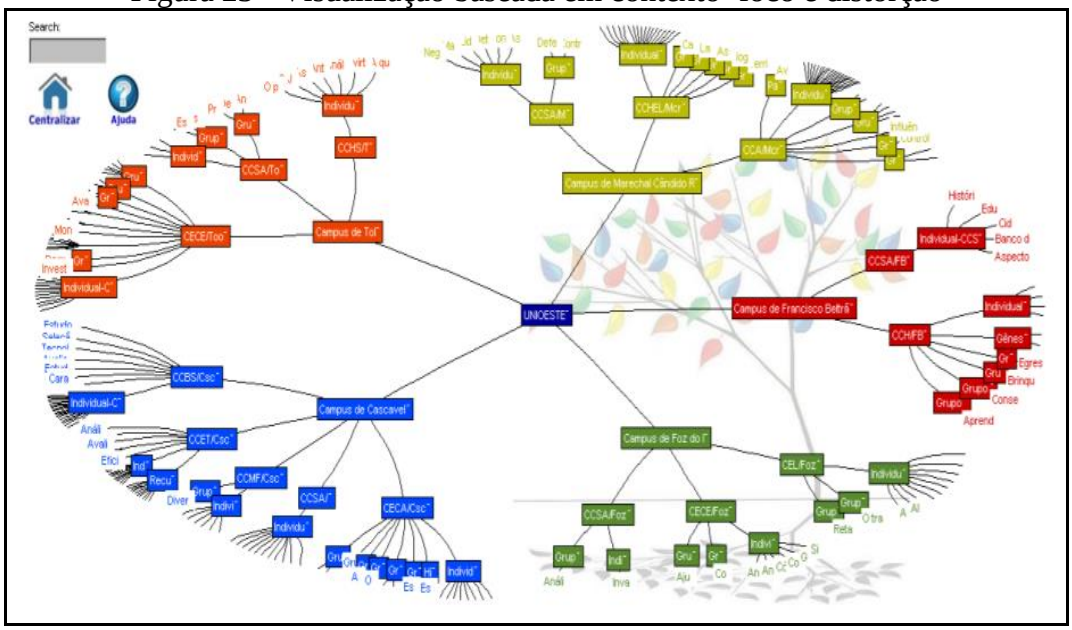
As técnicas de visualização baseadas em contexto+foco e distorção (também conhecida como visão detalhada ou efeito olho de peixe - *fisheye*) são compostas por abordagens que variam escala de ampliação da imagem e possuem como principal característica fornecer simultaneamente ao usuário informações contextuais e detalhadas (foco) na mesma área, sem a utilização de duas visões separadas. Desta forma, permitem uma visualização geral da ontologia, exibindo vários nós ao mesmo tempo, além de detalhes relacionados a um nó específico, sem que se perca o contexto onde ele está inserido na hierarquia. O nó em foco geralmente passa a ser o nó central, sendo que os demais são apresentados ao seu redor, reduzidos em tamanho até não serem mais visíveis. Isso pode ser feito através de técnicas de distorção ou por meio da eliminação dos detalhes nas zonas periféricas. Geralmente, são desenvolvidas com a ajuda de equação hiperbólica (KATIFORI et al., 2007; MAZZA, 2009).

Katifori et al. (2007) afirmam que as visualizações baseadas em contexto+foco e distorção podem ser bidimensionais ou tridimensionais. Entre as ferramentas projetadas a partir da utilização de recursos 2D destacam-se: 2D Hyperbolic Tree, StarTree,

OntoRama, MoireGraphs, TGVizTab e Ozone. Na categoria das ferramentas tridimensionais encontra-se a 3D Hyperbolic Tree, que foi desenvolvida para aplicações na Web, com o intuito de alcançar maior densidade de informação. Neste contexto, a técnica mais utilizada entre as aplicações e recursos desenvolvidos é a árvore hiperbólica, conforme exemplo apresentado na Figura 23.

A interface de visualização das árvores hiperbólicas facilita a exploração de ontologias complexas através de um *layout* radial, permitindo a navegação através dos nós e a ampliação dos elementos de interesse. A utilização dos recursos de contexto+foco são de grande importância, pois reduzem a carga cognitiva dos usuários ao interagir com grandes hierarquias (SAGHAFI, 2016). Inicialmente, o nó principal é apresentado no centro da imagem, enquanto os nós-filhos são posicionados ao seu redor, com um menor nível de detalhamento, apresentando-se distorcidos ou semi-ocultos em detrimento da parte focalizada. É possível, no entanto, selecionar um nó periférico e movê-lo para o primeiro plano, no centro da figura, a partir do qual toda a visualização é reorganizada (MAZZA, 2009).

Figura 23 - Visualização baseada em contexto+foco e distorção



Fonte: Unioeste – Núcleo de Inovações Tecnológicas (2012)

Outro recurso importante neste tipo de visualização refere-se ao *clustering*, que pode ser utilizado quando for necessário reduzir a complexidade visual do gráfico. Esta técnica é importante para descobrir agrupamentos ou classes de dados a partir de uma semântica escolhida. Podem ser utilizados para o agrupamento de informações baseadas em estrutura ou em conteúdo, como também para realizar funções de filtro e pesquisa (HERMAN; MELANÇON; MARSHALL, 2000).

Entre as vantagens relacionadas às árvores hiperbólicas, destacam-se: a) a estrutura da hierarquia de conceitos é de fácil entendimento; b) mesmo que essa técnica seja geralmente utilizada para exibição da taxonomia, ela possibilitaria a sua expansão para uma exibição mais completa das informações da ontologia; c) é uma técnica ideal para a visualização e navegação de grandes hierarquias; d) cada nó de interesse pode ser facilmente movido para o centro e exibido com mais detalhes, sem que o contexto seja perdido; e) permitem a navegação e localização rápida de classes ou instâncias específicas (KATIFORI et al., 2007; PIZZINATO; VIEIRA; RIGO, 2010).

Apesar das vantagens, Katifori et al. (2007) discutem alguns problemas observados no uso das árvores hiperbólicas: a) embora possuam fácil entendimento da hierarquia de conceitos, não oferecem uma representação muito óbvia de toda a estrutura ontologia, principalmente porque nós com a mesma distância do centro não estão necessariamente no mesmo nível da árvore; b) os nós externos, quando menores que um *pixel*, não são exibidos; c) não existe um posicionamento constante dos nós, principalmente quando existe a mudança de foco de visualização, o que pode desorientar o usuário; d) os rótulos podem aparecer desalinhados e sobrepostos, dificultando a visualização; e) a visualização inicial de uma ontologia complexa fornece apenas uma visão parcial da estrutura, uma vez que muitos rótulos e conceitos não podem ser observados.

De forma geral, as técnicas de visualização baseadas em contexto+foco e distorção, em especial as árvores hiperbólicas, parecem ser muito eficazes em fornecer visões globais da ontologia e exibir muitos nós de uma só vez. Pode ser útil para a análise de um determinado nó de interesse, uma vez que é possível movê-lo para o foco da estrutura e visualizar todos os relacionamentos referentes a ele. No entanto, este tipo de visualização não oferece uma representação clara de todas as relações presentes na estrutura ontológica, o que pode comprometer o entendimento do usuário.

3.2.6 Visualização Baseada em Paisagem (*Information Landscapes*)

A principal característica deste tipo de visualização é a utilização de paisagens (*landscapes*) para representar interrelações entre grupos de dados. É muito utilizada para aplicações relacionadas a visualização da similaridade entre documentos texto com base em palavras-chave, bem como em ambientes de Realidade Virtual (RV), onde os documentos são representados por objetos tridimensionais, codificados por cor e tamanho, dispostos em um plano em perspectiva horizontal, simulando uma planície (KATIFORI et al., 2007). Na

visualização de ontologias, a tridimensionalidade da visualização baseada em paisagem pode ser utilizada para fornecer uma dimensão extra da hierarquia, sendo possível representar conexões entre um nó selecionado e os respectivos nós relacionados.

Entre as ferramentas que adotam este tipo de visualização, podem ser destacadas: a) sistema File System Navigator, desenvolvido com o propósito de permitir a exploração de arquivos 3D, em sistemas UNIX; b) Harmony Information Landscape, projetado para documentos hipertexto; c) KartOO, um site de busca na Web (http://www.kartoo.com/) que apresenta os resultados na forma de *landscapes* interativos. Um exemplo de visualização baseada em paisagem pode ser observado na Figura 24 (NASCIMENTO; FERREIRA, 2005; KATIFORI et al., 2007).

Figura 24 - Visualização baseada em paisagem



Fonte: Silva e Freitas (2006, p.180)

A ferramenta KartOO é um metabuscador proposto em 2001, que utiliza princípios cartográficos para apresentar os resultados retornados de uma pesquisa na Web. Os sites retornados são exibidos como ícones de páginas, enquanto “regiões” identificadas através de palavras-chaves indicam temas recorrentes nessas páginas. Os temas aparecem como “vales” dispostos numa representação espacial arbitrária. Linhas surgem conectando o ícone de uma página selecionada aos temas que ela contém, enquanto a seleção de um tema realça as páginas com conteúdo naquela categoria (SILVA; FREITAS, 2006).

Entre os problemas da visualização baseada em paisagem, Katifori et al. (2007) destacam a dificuldade de representar conjuntos de dados em que um nó possui muitos filhos, uma vez que seria necessário um *layout* amplo que não pode ser visto de uma só vez e,

como resultado, não é possível reter o contexto global durante o *zoom*. Além disso, a visualização dos *hiperlinks* não é muito eficaz, devido à desordem e sobreposição de rótulos.

Apesar dos problemas relatados, a visualização baseada em paisagem pode simplificar a navegação ao utilizar representações visuais para mapear propriedades como tamanho e tipo e, assim, apresentar cada nó de forma diferenciada, facilitando a compreensão do usuário. Outra vantagem refere-se à possibilidade de apresentar a visão geral de uma hierarquia, se combinadas com ferramentas de pesquisa e filtragem apropriadas e mecanismos de navegação intuitivos, simples e eficazes. No entanto, ainda não é evidente a utilidade deste tipo de visualização no contexto das ontologias, uma vez que faltam avaliações extensivas de ferramentas com este propósito (KATIFORI et al., 2007).

3.3 INTERFACES DE BUSCA BASEADAS EM ONTOLOGIAS

Para tentar minimizar os problemas observados no processo de recuperação de informação, vários estudos estão sendo direcionados para a utilização de ontologias em interfaces de busca, uma vez que elas possibilitam a construção de componentes interativos, com condições de auxiliar a composição de expressões de busca, permitindo que o usuário de um sistema de recuperação utilize conceitos por ela definidos para traduzir linguisticamente sua necessidade de informação (GARCIA; SICÍLIA, 2003).

Com o acesso direto aos conceitos de uma ontologia, o usuário pode elaborar expressões de busca mais significativas, contextualizadas e obter um melhor entendimento sobre o domínio pesquisado. Além disso, a interação com documentos recuperados pode servir de base para refinar ou iniciar novas buscas. Tang (2007) afirma ainda que a possibilidade de interação direta com conceitos definidos por uma ontologia pode estimular os usuários a explorarem diferentes aspectos de suas necessidades, além de ser um recurso importante para ajudar na elaboração de expressões de busca, principalmente quando o usuário não sabe ao certo o que precisa ou quando não possui conhecimento sobre o assunto pesquisado.

As interfaces de busca baseadas em ontologias podem ser projetadas para interagir com os usuários de formas diferenciadas. Para Paulheim e Probst (2010), as aplicações podem optar por apresentar o conteúdo nos seguintes formatos:

a) nenhuma representação: as ontologias são escondidas completamente, não possibilitando a visualização/interação pelo usuário. Neste caso, são utilizadas apenas na estruturação do sistema ou na camada de dados

e não como componentes de interface;

b) listas de categorias: as ontologias e suas relações são apresentadas ao usuário no formato de listas de seleção, taxonomias, relação de conceitos ou blocos de textos em lugares específicos da interface;

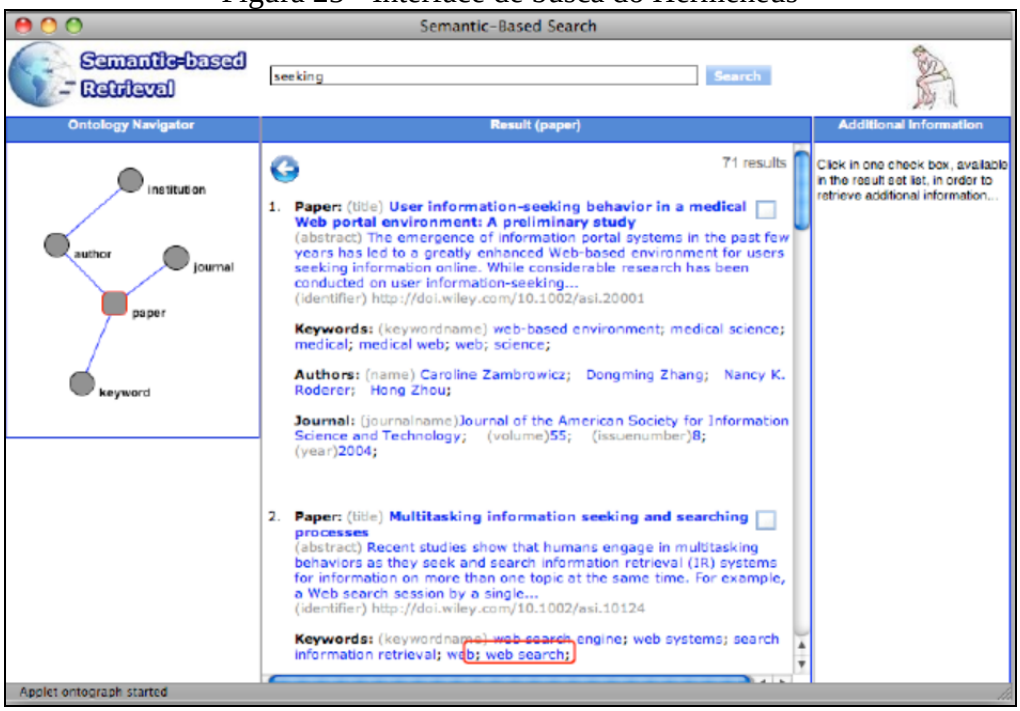
c) gráfico: as relações entre os conceitos são visualizadas como árvores (na maioria das vezes mostrando a taxonomia) ou gráficos (incluindo as relações não taxonômicas). Existem diversas técnicas para visualização de ontologias, abrangendo representações gráficas em 2D e 3D (KATIFORI et al., 2007);

d) verbalizado: a representação textual dos axiomas contidos na ontologia é fornecida ao usuário;

e) código-fonte: o código-fonte da ontologia é mostrado ao usuário (OWL ou F-Logic, por exemplo).

Diversos trabalhos já foram desenvolvidos na tentativa de melhorar o processo de recuperação de informação utilizando ontologias como componente de interface de busca. Beppler (2008) apresenta o Hermeneus, que é um *framework* de recuperação e busca de informação que permite aos usuários navegar nos conceitos e instâncias de uma ontologia de domínio, que é o artefato principal para a criação de módulos e componentes. Sua interface principal é apresentada na Figura 25.

Figura 25 - Interface de busca do Hermeneus



Fonte: Beppler (2008, p.77)

O sistema foi desenvolvimento para ser utilizado em qualquer domínio e é composto por quatro componentes: **a) módulo de indexação**: utiliza uma ontologia de domínio para criar índices semânticos automaticamente; **b) módulo de recuperação**: responsável por recuperar as instâncias da ontologia que estão armazenadas nos índices semânticos em resposta a uma consulta; **c) módulo de apresentação**: utiliza uma ontologia de domínio para disponibilizar um ambiente interativo na interface do usuário. Esse módulo é composto de outros três componentes: navegador de ontologia, instâncias recuperadas e informação adicional; **d) módulo de inferência**: responsável por extrair informação adicional de uma determinada instância recuperada por meio da utilização de técnicas de inferência e regras pré-configuradas sobre a ontologia e sua base de conhecimento.

Segundo Beppler (2008, p.86), o objetivo de uma ferramenta gráfica "é facilitar que os seres humanos consigam adquirir percepções mais aguçadas sobre dados". O Hermeneus foi desenvolvido com objetivo de estimular o usuário a explorar a informação disponibilizada, além de auxiliar na construção de uma expressão de busca mais apropriada, mesmo quando não existe conhecimento sobre o domínio pesquisado. Possibilita ainda a visualização de diferentes perspectivas das informações recuperadas e resultado de busca contextualizado e disponível para utilização dinâmica e interativa.

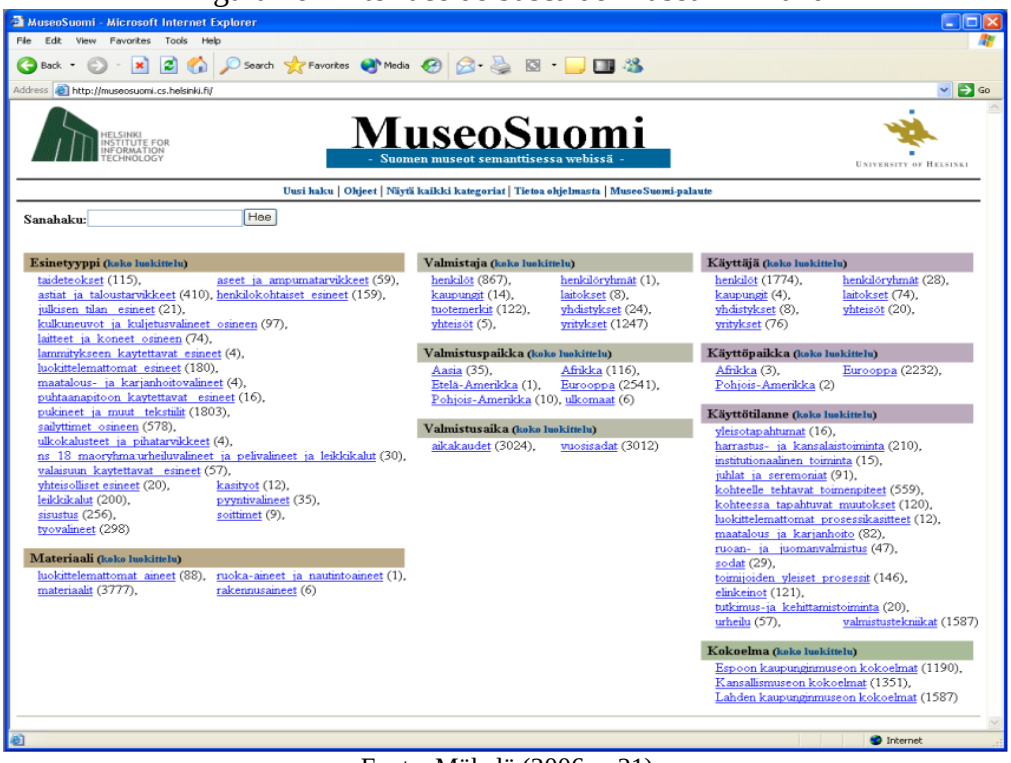
Ainda, de acordo com o autor, os testes realizados com o protótipo desenvolvido foram positivos e apresentaram contribuições importantes para a melhoria do processo de RI, tais como:

a) elaboração de expressões de busca mais precisas, geradas a partir da interação com uma ontologia, por meio de cliques do *mouse*;
b) possibilidade de o usuário adquirir, durante a interação com uma ontologia, novas percepções sobre as suas necessidades de informação;
c) os usuários podem formular consultas menos ambíguas, uma vez que podem expressar claramente o contexto de cada termo utilizado na consulta;
d) com base nos resultados apresentados, os usuários podem ter novas percepções sobre as suas necessidades, além de refinar ou formular novas consultas;
e) as ontologias podem ser usadas para produzir diferentes visões do conhecimento, particularmente para propósitos de exploração e navegação; e
f) possibilidade de o usuário desenvolver uma série de tipos de interações

sobre a interface de busca, que vão desde a reformulação de uma consulta, navegação e refinamento, até a verificação, análise de relevância e aprendizagem.

Outro exemplo importante para o contexto de interfaces baseadas em ontologias é o MuseumFinland, que é um portal semântico utilizado para a publicação de coleções culturais heterogêneas de museus, conforme pode ser observado na Figura 26. Foi apresentado no trabalho de Hyvönen e Mäkelä et al. (2005) e tem como base a utilização de ontologias para garantir a interoperabilidade e possibilitar busca semântica. Além disso, permite ao usuário utilizar os conceitos definidos na ontologia, que são disponibilizados no formato de taxonomia multifacetada, para elaborar consultas, refinar os resultados e navegar nos itens das coleções. Bonino et al. (2004) afirmam que o ponto chave para o processo de refinamento de uma busca semântica está na disponibilidade de uma ontologia de domínio, e na capacidade de compreender as relações semânticas entres os conceitos ontológicos.

Figura 26 - Interface de busca do MuseumFinland



Fonte: Mäkelä (2006, p.21)

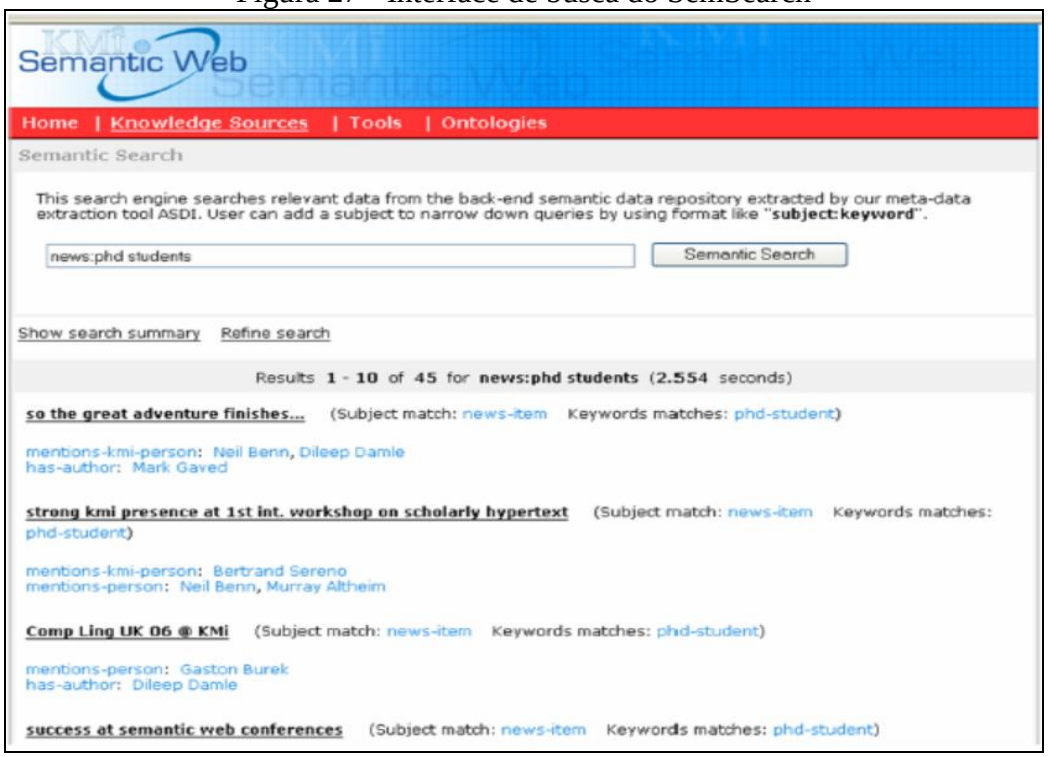
Segundo Hyvönen e Mäkelä et al. (2005), o MuseumFinland se diferencia dos demais portais por apresentar um sistema de busca inteligente, baseado em conceitos ontológicos e não em palavras-chave. Além disso, as relações semânticas implícitas na ontologia possibilitam ao usuário entender o contexto da pesquisa e evitar problemas

relacionados com ambiguidade, como também descobrir associações dentro do conteúdo do portal e utilizá-las para enriquecer a navegação.

Os estudos de Hyvönen e Mäkelä et al. (2005) destacam algumas contribuições para o processo de busca e recuperação de informação: a) o conteúdo apresentado no formato de taxonomia proporciona ao usuário uma visão geral do tipo de informação disponível (ou não) para a consulta; b) as hierarquias podem orientar o usuário na formulação das consultas, uma vez que disponibilizam vocabulário adequado para realizar pesquisas sob diferentes pontos de vista; c) as hierarquias não apresentam problemas de ambiguidade, relacionados principalmente com homônimos; d) as facetas podem ser utilizadas para auxiliar a navegação do usuário em busca de um conteúdo específico; e) o usuário pode refinar as consultas ou formular outras, por meio da seleção de subcategorias ou de novas categorias.

Assim como o MuseumFinland, o sistema de busca SemSearch proposto por Lei, Uren e Motta (2006) foi projetado para disponibilizar ao usuário a possibilidade de realizar buscas semânticas. De acordo com os autores, o sistema possui uma interface no estilo do Google, conforme pode ser observado na Figura 27, que tenta extrair a semântica dos termos digitados em uma consulta para recuperar documentos melhor contextualizados e mais relevantes.

Figura 27 - Interface de busca do SemSearch



Fonte: Lei, Uren e Motta (2006, p.15)

O objetivo do SemSearch foi criar um sistema de fácil utilização, capaz de auxiliar usuários a especificar consultas simples (até dois termos) ou complexas (três termos ou mais), de forma eficiente e com resultados mais precisos, mesmo para aqueles que não tenham familiaridade com o domínio pesquisado, escondendo, porém, a complexidade das buscas semânticas. Desta forma, o sistema possibilita que todos os tipos de usuários (experientes ou não) tenham condições de obter os benefícios das tecnologias da Web Semântica, mesmo sem conhecer linguagens de consulta específicas.

Durante o processo de busca, o SemSearch compara cada termo utilizado na consulta com fontes de dados heterogêneas, na tentativa de encontrar conceitos, relações ou instâncias descritas. “Quando há casamento entre um ou mais termos da consulta com os termos presentes na ontologia, a consulta é alterada, tornando-se uma consulta descrita em linguagem formal, ou seja, uma consulta semântica” (BEPPLER, 2008, p.91).

Uma avaliação inicial do SemSearch apresentada por Lei, Uren e Motta (2006) mostrou resultados positivos para o processo de recuperação de informação: a) o sistema possui uma interface simples, que minimiza os problemas de sobrecarga de informação; b) é capaz de produzir respostas precisas para as consultas realizadas pelos usuários; c) possibilita o desenvolvimento de consultas semânticas, mesmo que o usuário não tenha conhecimento das ferramentas da Web Semântica.

Além dos três experimentos analisados, diversas outras aplicações foram desenvolvidas utilizando ontologias com o propósito de minimizar problemas observados no processo de busca e recuperação de informação: OntoSeek (GUARINO; MASOLO; VETERE, 1999), OntoIR (GARCÍA; SICILIA, 2003), CIRI (AIRIO et al., 2004), Textpresso (MÜLLER; KENNY; STERNBERG, 2004), OnAIR (PAZ-TRILLO; WASSERMANN; BRAGA, 2005), OWLIR (FININ et al., 2005), OntoSmart (FERNEDA, 2013), entre outras.

Apesar dos resultados positivos observados, das contribuições das ontologias no processo de RI e das perspectivas para o desenvolvimento de ferramentas de busca semântica e interfaces diferenciadas na Web, alguns problemas precisam ser analisados cuidadosamente. Segundo Lei, Uren e Motta (2006), as propostas de busca semântica ainda não estão sendo direcionadas para o usuário final, que não estão familiarizados com domínios específicos de dados semânticos, linguagens de consultas e navegação em interfaces guiadas por estruturas terminológicas (ontologias).

Baeza-Yates e Ribeiro-Neto (2013) afirmam ainda que não há um padrão muito bem estabelecido para a visualização dos resultados de busca, principalmente em caso de interfaces gráficas, além da dificuldade de representar respostas de grandes dimensões

(grande quantidade de resultados), como é o caso no ambiente Web. Outra questão comum apresentada pelos autores é a falta de habilidade dos usuários em utilizar estruturas de navegação (ontologias, listas, categorias), uma vez que eles podem perder o contexto quando mergulham profundamente em um caminho específico. Portanto, as ferramentas visuais para a seleção e/ou navegação de conteúdo não foram implantadas em toda a Web porque elas ainda têm que demonstrar valor agregado ao usuário.

Katifori et al. (2007) também discutem alguns desafios encontrados no desenvolvimento de novas ferramentas de busca utilizando ontologias. Segundo eles, não existem avaliações comparativas relacionadas à eficácia dos métodos de visualização de ontologias para tarefas diferenciadas, tais como a busca. Isto significa que não basta projetar e implementar novos mecanismos de busca e recuperação de informação. Portanto, é necessário acrescentar às pesquisas, avaliação rigorosa junto aos usuários para analisar quais tipos de recursos e interfaces são realmente eficazes e podem contribuir significativamente para o desempenho e a melhoria do processo de busca de informações na Web.

Como pode ser observado, a utilização de ontologias em interfaces de busca é promissora e apresenta inúmeras contribuições e perspectivas para a área. No entanto, não foram encontrados autores que comentassem sobre os três sistemas discutidos neste trabalho: o Hermeneus, o MuseumFinland e o SemSearch. Os dados apresentados e discutidos referem-se apenas à opinião dos desenvolvedores. Também não foram apresentados resultados de testes exaustivos realizados com usuários finais, impossibilitando, portanto, uma análise mais aprofundada sobre o desempenho e a eficácia das ferramentas desenvolvidas, bem como sua aceitação pelo usuário final e sua contribuição efetiva no processo de recuperação de informação.

3.4 RESUMO E DISCUSSÃO DO CAPÍTULO

Este capítulo apresentou conceitos e características sobre a utilização de ontologias, destacando as possibilidades de desenvolvimento de novos recursos para a melhoria dos sistemas de recuperação de informação. Além disso, fornece um breve resumo sobre a visualização de ontologias, com foco em problemas de recuperação de informação, discutindo métodos, técnicas e principais ferramentas utilizadas. Também discutiu projetos já desenvolvidos por outros autores que utilizaram ontologias em interfaces de busca, destacando contribuições, perspectivas e desafios encontrados.

O foco principal foi o estudo dos princípios para a visualização de

ontologias, uma vez que os métodos e as técnicas discutidas serão utilizados na proposta apresentada nesta tese. No entanto, a escolha de uma técnica mais adequada não é uma tarefa trivial, visto que existem vantagens e desvantagens em cada uma delas e que a literatura não elege uma que seja mais adequada para representar uma ontologia, bem como para interagir com diferentes tipos de usuários e necessidades informacionais. O Quadro 5 apresenta um breve comparativo entre os métodos de visualização discutidos neste capítulo.

Quadro 5 - Comparativo entre os métodos de visualização de ontologias

Método de Visualização	**Elementos da Ontologia a Serem Representados**				
	Classes e Instâncias	**Taxonomia**	**Múltiplas Heranças**	**Propriedades**	**Axiomas**
Lista Indentada	Mostradas em listas e interfaces separadas	Classes representadas de forma indentada	Replicação de elementos	Apresentadas em interfaces separadas	Apresentados em interfaces separadas
Nó-*Link* e Árvore	Representados por nós em formas geométricas	Representada por proximidade e ligações	Representados por ligações em grafo	Apresentadas em janelas separadas	Rótulos com *links* para acesso
Aproximação	Representados em conjuntos	Tamanhos menores e em conjuntos	Replicação de elementos	Apresentadas em interfaces separadas	Apresentadas em interfaces separadas
Preenchimento de Espaço	Representados por cores e tamanhos distintos	Diferenciadas por tamanhos e cores	Geralmente não representados	Apresentadas em interfaces separadas	Apresentadas em interfaces separadas
Contexto + Foco e Distorção	Representados por rótulos, conforme foco	Distância e ligações entre as classes e instâncias	Replicação de elementos	Apresentadas em interfaces separadas	Apresentadas em interfaces separadas
Paisagem	Distribuídos graficamente em relevos	Linhas que conectam as classes	Replicação de elementos	Apresentadas em interfaces separadas	Apresentadas em interfaces separadas

Fonte: Mendes Netto, Lima, Pierozzi Júnior (2016, p.16)

Além dos métodos e técnicas empregadas para a visualização de ontologias, também foram discutidos elementos gráficos, tais como: tamanho, textura, cor, espessura, valor, orientação, forma. Neste mesmo contexto, os princípios de *Gestalt*, que envolvem elementos relacionados à proximidade, similaridade, região comum e conectividade, podem também ser úteis no processo de recuperação de informação, contribuindo para a organização, análise e seleção de informações, bem como para a elaboração de comparativos e a descoberta de conhecimento latente não foi evidenciado. Estes recursos poderão ser empregados tanto na elaboração de uma expressão de busca, quanto na interpretação dos resultados apresentados.

Outro ponto importante discutido no capítulo refere-se à utilização de ontologias em interfaces de busca. O desenvolvimento de sistemas de recuperação de

informação utilizando ontologias, direcionados principalmente para a Web, vem se consolidando nos últimos anos, devido aos inúmeros benefícios evidenciados na representação, organização, disseminação e recuperação de conteúdos. Os pesquisadores da área têm direcionado esforços para projetar ferramentas que consigam minimizar os problemas observados no processo de busca e possam oferecer resultados mais relevantes aos usuários. Para isto, muitos estudos se concentram nas interfaces de busca baseadas em ontologias, com o propósito de acrescentar novas funcionalidades às interfaces atuais, torná-las mais amigáveis e interativas. O Quadro 6 apresenta uma síntese das principais contribuições, perspectivas e desafios levantados sobre o assunto.

Os resultados apresentados ainda são parciais e referem-se às análises realizadas por meio da revisão da literatura. No entanto, foi possível identificar inúmeras contribuições que podem ser agregadas às interfaces de busca atuais, destacando-se principalmente a possibilidade de disponibilizar ao usuário um vocabulário adequado, baseado nos conceitos, instâncias e relações semânticas de uma ontologia de domínio, para ser utilizado na elaboração de expressões de busca mais significativas, contextualizadas e com a possibilidade de reduzir os problemas relacionados com a subjetividade inerente à língua (sinonímia, homonímia, paronímia, polissemia e variações linguísticas), que tanto dificultam o processo de recuperação de informação.

Outra contribuição significativa está no fato de que, por meio da interação direta com os conceitos presentes na ontologia, o usuário terá mais condições para elaborar expressões de busca eficientes, mesmo que não tenha conhecimento sobre o contexto pesquisado. Também podem oferecer diversos tipos de interação para a formulação e reformulação de uma consulta, refinamento, análise de relevância e aprendizagem sobre o contexto pesquisado.

Embora tenham sido identificadas inúmeras contribuições sobre o assunto, ainda existem questões em aberto relacionadas à utilização de ontologias em interfaces de busca que precisam ser melhor exploradas. Entre os principais pontos levantados, é preciso envolver efetivamente os usuários nos estudos realizados, uma vez que a maioria dos trabalhos encontrados na literatura apresentam resultados baseados apenas na opinião dos autores e desenvolvedores, impossibilitando, portanto, uma análise mais aprofundada sobre o desempenho, a eficácia, a usabilidade e as limitações das ferramentas desenvolvidas, bem como sua aceitação pelo usuário final e sua contribuição efetiva no processo de recuperação de informação.

Quadro 6 - Síntese da utilização de ontologias em interfaces de busca

Contribuições
▪ Permitem a utilização de conceitos da ontologia na elaboração de expressões de busca, minimizando os problemas linguísticos; ▪ Possibilitam a formulação de consultas menos ambíguas, uma vez que os usuários podem expressar claramente o contexto de cada termo utilizado na consulta; ▪ Disponibilizam vocabulário adequado ao usuário para a elaboração de expressões de busca mais significativas e melhor contextualizadas, realizadas a partir dos conceitos definidos na ontologia; ▪ Possibilitam um melhor entendimento sobre os conceitos e instâncias de um domínio pesquisado; ▪ Permitem a interação com documentos recuperados, servindo de base para refinar ou iniciar novas buscas; ▪ Ajudam na elaboração de expressões de busca, principalmente quando o usuário não sabe ao certo o que deseja ou não possui conhecimento sobre o assunto pesquisado; ▪ Estimulam os usuários a explorarem diferentes aspectos de suas necessidades, por meio da interação direta com conceitos definidos na ontologia; ▪ Possibilitam a visualização de diferentes perspectivas das informações recuperadas; ▪ Podem ser implementadas de forma a extraírem o significado semântico do termo digitado pelo usuário e, desta forma, recuperar documentos melhor contextualizados e mais relevantes.
Perspectivas
▪ Desenvolvimento de ferramentas de busca semântica e interfaces diferenciadas na Web; ▪ Desenvolvimento de interfaces de busca com a possibilidade de oferecer ao usuário diversos tipos de interações, que vão desde a reformulação de uma consulta, navegação e refinamento, até a verificação, análise de relevância e aprendizagem; ▪ Desenvolvimento de interfaces de busca que ofereçam aos usuários menores problemas de linguagem (sinonímia, homonímia, paronímia, polissemia e variações linguísticas) ▪ Desenvolvimento de ferramentas que ofereçam busca semântica de forma gráfica, substituindo as palavras-chave, tais como: lista de categorias, hierarquias e taxonomia multifacetada; ▪ Desenvolvimento de sistemas com recursos variados que auxiliem o processo de busca, tais como: indexação automática, expansão de consultas, checagem de consistência e complemento de informações.
Desafios
▪ As propostas de busca semântica ainda não estão sendo direcionadas para o usuário final; ▪ Os usuários não estão familiarizados com domínios específicos de dados semânticos, linguagens de consultas e navegação em interfaces guiadas por ontologias; ▪ Não existe um padrão bem estabelecido para a visualização dos resultados de busca, principalmente no caso de interfaces gráficas; ▪ Ainda existem dificuldades para representar respostas com grande quantidade de resultados; ▪ Falta de habilidade dos usuários em utilizar estruturas de navegação (ontologias, listas e categorias); ▪ As ferramentas visuais para a seleção e/ou navegação de conteúdo ainda precisam demonstrar valor agregado ao usuário; ▪ Ainda não existem avaliações comparativas relacionadas à eficácia dos métodos de visualização de ontologias para tarefas diferenciadas, tais como a busca; ▪ A maioria das ferramentas desenvolvidas não foram testadas com usuários finais. O que se têm são apenas opiniões dos desenvolvedores sobre o seu desempenho e sua eficácia; ▪ É necessário acrescentar avaliação rigorosa junto aos usuários para analisar quais tipos de recursos e interfaces são realmente eficazes.

Fonte: Adaptado de Hyvönen e Mäkelä et al. (2005), Lei, Uren e Motta (2006) e Beppler (2008)

Apesar dos problemas levantados, as ontologias possuem grande potencial para o desenvolvimento de aplicações mais sofisticadas e interfaces de busca contendo recursos gráficos que possam torná-las mais fáceis de serem utilizadas e capazes de auxiliar o

usuário a expressar de forma mais precisa sua necessidade de informação. Além disso, elas podem ser utilizadas para promover a interoperabilidade dos sistemas e implementar outros recursos importantes para a melhoria dos sistemas de RI, tais como: indexação automática, expansão de consultas, checagem de consistência e complemento de informações.

4 UTILIZAÇÃO DE ONTOLOGIAS NA RECUPERAÇÃO DE INFORMAÇÃO

Este capítulo tem como objetivo apresentar uma análise dos principais estudos disponibilizados na literatura sobre a utilização de ontologias na busca e recuperação de informação, com o intuito de ampliar o entendimento acerca das múltiplas possibilidades de aplicação e desenvolvimento de projetos. Neste contexto, foi possível identificar tipos de ontologias adotadas, áreas de aplicação, principais abordagens, propostas de melhorias, estudos validados por usuários e lacunas de pesquisa ainda existentes que pudessem ser empregadas neste estudo.

■■■

A análise apresentada nesta etapa do estudo foi desenvolvida a partir das publicações presentes na ACM Digital Library e no Google Scholar. As palavras-chave adotadas para identificar os materiais de estudo envolvem três áreas específicas: Ontologias, Interfaces de Busca (*Search Interface*, *User Interface*, *Query Formulation*) e Visualização da Informação. A pesquisa realizada retornou um total de 90 (noventa) artigos, publicados entre os anos de 2000 a 2017, dos quais 59 (cinquenta e nove) foram selecionados por estarem de acordo com os propósitos deste trabalho.

A partir da seleção dos artigos, a análise foi realizada de acordo com os seguintes critérios: 1) tipo de ontologia adotada (genérica, domínio, tarefa, aplicação); 2) área de desenvolvimento da pesquisa e principais abordagens; 3) problemáticas e propostas de melhorias; 4) validação por meio de estudos de usuários.

4.1 TIPOS DE ONTOLOGIA ADOTADOS

Em função da grande quantidade de informações disponíveis, principalmente no contexto da Web, inúmeras pesquisas estão sendo direcionadas para o uso de ontologias, com a principal finalidade de representar todo esse conhecimento de forma estruturada e possibilitar uma recuperação mais eficiente. Os trabalhos com este propósito destacam as inúmeras áreas que estão adotando este recurso, bem como a variedade de propostas e aplicações relacionadas.

A maioria dos trabalhos avaliados adota ontologias de domínio, indicando que se trata de abordagens realizadas a uma área de conhecimento específica. As ontologias genéricas, de aplicação e de tarefa aparecem em menor quantidade. Em relação a este quesito,

foi possível perceber que existem conflitos conceituais entre os autores sobre os tipos de ontologias. Os conceitos e as características que definem e distinguem cada um deles, principalmente em relação às ontologias de aplicação e de tarefa, não são bem compreendidos, dificultando a adoção dessas duas terminologias. As ontologias de tarefa e de aplicação, de acordo com Guarino (1998), são especializações das ontologias de domínio, o que faz com que a distinção entre elas seja ainda mais complicada.

Tabela 1 - Tipo de ontologia adotada

Tipo	**Quantidade de Trabalhos**
Ontologia Genérica (*top-level*)	08
Ontologia de Domínio	40
Ontologia de Aplicação	05
Ontologia de Tarefa	03
Tipo de Ontologia não Identificado	03

Fonte: desenvolvido pela autora

Os trabalhos que envolvem **ontologias genéricas** são direcionados principalmente para o desenvolvimento de técnicas de visualização ontológica e exploração de dados na Web Semântica. As principais ontologias genéricas utilizadas são: Yago (BAST et al., 2007; FREITAS; CURRY, 2014; STYPEREK; CIESIELCZYK; SZWABE, 2014), Sumo (BUNTAIN, 2008), Freebase (DELIGIANNIDIS; KOCHUT; SHETH, 2007; DEMIDOVA; ZHOU; NEJDL, 2013) e Dbpedia (FREITAS; CURRY, 2014; STYPEREK; CIESIELCZYK; SZWABE, 2014).

A ontologia Yet Another Great Ontology (YAGO), desenvolvida em parceria pelo Instituto Max Planck de Ciência da Computação da Alemanha e pela Universidade Telecom ParisTech da França, é uma base de conhecimento semântica extraída automaticamente da Wikipédia, WordNet e GeoNames, contendo mais de 10 milhões de entidades (pessoas, organizações, cidades, etc.) em diferentes idiomas e contém mais de 120 milhões de fatos sobre elas (categorias, redirecionamentos e *infoboxes* (MAX PLANCK INSTITUT INFORMATIK, 2017).

A Suggested Upper Merged Ontology (SUMO), desenvolvida pela IEEE, é considerada a maior ontologia pública formal existente. Foi mapeada para todo o léxico WordNet e está escrita na linguagem Standard Upper Ontology Knowledge Interchange Format (SUO-KIF). Possui aproximadamente 25.000 termos e 85.000 axiomas, quando combinadas todas as ontologias de domínio que a constituem (ADAMPEASE, 2017).

A Freebase, por sua vez, refere-se a uma base de conhecimento

colaborativa, que recolhe dados estruturados *on-line* de diversas fontes. Foi desenvolvida pela empresa de *software* americana Metaweb e adquirida pela Google em 2010 (GOOGLE DEVELOPERS, 2017).

Por fim, a DBPedia é uma base de conhecimento também colaborativa, que abrange diversos domínios, mantida por milhares de usuários, que tem como objetivo extrair informações estruturadas da Wikipedia e disponibilizá-las na Web, servindo de base para o desenvolvimento de novos mecanismos de busca, exploração e visualização, bem como para projetos de aperfeiçoamento da própria enciclopédia (DBPEDIA, 2017).

As pesquisas que adotam **ontologias de domínio** são as mais comuns e referem-se a trabalhos relacionados a contextos diversos, tais como: a) Recuperação de Informação na Web (VELART; ŠALOUN, 2007; KONDYLAKIS; PLEXOUSAKIS, 2011; PITTARELLO; GATTO, 2012b; ROZELL et al., 2012; ZHENG et al., 2012); b) Web Semântica (BERNSTEIN et al., 2005; GUPTA; BHOWMIK; GOVINDARAJU, 2008; SOYLU et al., 2016); c) Visualização de Informações (SOUZA; SANTOS; EVANGELISTA, 2003; STOREY, 2007; PAULHEIM; MEYER, 2011); d) Recuperação Multimídia (ALBERTONI; BERTONE; DE MARTINO, 2005; POPESCU; MILLET; MOËLLIC, 2007; WU et al., 2008; BERTINI et al., 2010); e) Desenvolvimento de Sistemas (PAULHEIM, 2009; ZHENG et al., 2012; EL-ATAWY; KHALEFA, 2016); f) Avaliação de Eficácia e Usabilidade de Interface (ZHUHADAR; NASRAOUI, 2010; DUDYCZ, 2015).

As **ontologias de tarefa** e de **aplicação**, por sua vez, são adotadas para o desenvolvimento de pesquisas em contextos específicos e bem definidos, relacionados principalmente às áreas de avaliação de resultados de busca (KAMEL; LEE; POWERS, 2007), dispositivos móveis (SHIN et al., 2013a, 2013b) e recuperação de informação na Web (PAPADAKIS et al., 2009; THAI; HANDSCHUH, 2009; CHEN; CHUA, 2013).

De todos os trabalhos analisados, não foi possível identificar o tipo de ontologia adotada em apenas três deles, por não terem sido citadas pelos respectivos autores ou por tal informação ser indiferente ao projeto apresentado.

4.2 ÁREAS DE PESQUISA E PRINCIPAIS ABORDAGENS

As ontologias estão sendo adotadas em uma multiplicidade de projetos, pertencentes a áreas variadas. Neste sentido, os trabalhos avaliados foram classificados de acordo com a sua abordagem principal, conforme as categorias apresentadas na Tabela 2.

Tabela 2 - Áreas de desenvolvimento de pesquisa

Área de Desenvolvimento de Pesquisas	Quantidade de Trabalhos
Recuperação de Informação na Web	23
Web Semântica	12
Visualização de Informações	07
Recuperação Multimídia	07
Avaliação (eficácia, resultados, usabilidade)	03
Dispositivos Móveis	03
Desenvolvimento de Interfaces e Sistemas	04

Fonte: desenvolvido pela autora

A maior parte das pesquisas está relacionada à área de **Recuperação de Informação na Web**, com propostas de trabalhos que visam a melhoria dos aspectos relacionados à busca nas suas diversas vertentes: 1) suporte à busca e navegação (GROTH; LANNERÖ, 2006; WONG; LIU; BENNAMOUN, 2010; ISSERTIAL; PITTARELLO; GATTO, 2012b; TSUJI, 2013); 2) indexação e representação de documentos (LEE; MIN; CHUNG, 2009; BASKAYA; KEKÄLÄINEN; JÄRVELIN, 2010; SIAHBANI et al., 2013); 3) formulação de *queries*, refinamento e expansão de consultas (STOJANOVIC, 2005; KHATTAK, 2008; ARFAOUI; HIDRI, 2016) e 4) visualização de informações (VELART; ŠALOUN, 2007; LIN et al., 2010).

Os projetos desta área são desenvolvidos em domínios variados, tais como: oceanografia biológica e química, dados públicos abertos, observatório virtual no campo da física solar terrestre (ROZELL et al., 2012), organização de dados pessoais (THAI; HANDSCHUH, 2009), mundos 3D (PITTARELLO; GATTO, 2012a, 2012b), mapas comunitários (MAURO, 2017), educação, mais especificamente cursos *on-line* (TIJERINO; SANATI, 2005; VELART; ŠALOUN, 2007), coleta e navegação de *papers* (ISSERTIAL; TSUJI, 2013), recursos geográficos (ALBERTONI; BERTONE; DE MARTINO, 2005), comércio eletrônico (CARVALHO et al., 2002), documentos XML (ARFAOUI; HIDRI, 2016), jornal eletrônico, com notícias relacionadas à economia, tecnologia e medicina (WONG; LIU; BENNAMOUN, 2010), textos biomédicos (CHUA; KIM, 2012), documentos bilíngues (BASKAYA; KEKÄLÄINEN; JÄRVELIN, 2010) e sistemas e técnicas de busca de informação (GROTH; LANNERÖ, 2006; LEE; MIN; CHUNG, 2009; LIN et al., 2010; SIAHBANI et al., 2013).

Outro destaque no uso de ontologias são os trabalhos direcionados para a **Web Semântica**, que visam desenvolver interfaces de busca, bem como métodos e técnicas de pesquisa, exploração e interpretação de grande quantidade de dados estruturados (BAST et

al., 2007; GUPTA; BHOWMIK; GOVINDARAJU, 2008; SOYLU et al., 2013).

A área de **Visualização de Informações**, assim como a Web Semântica, também se destaca pelos inúmeros problemas relacionados à manipulação e visualização de quantidade volumosa de dados. Os projetos com este propósito buscam analisar aspectos relacionados com a visualização de ontologias de grande porte (TZITZIKAS; HAINAUT, 2006), avaliar métodos, técnicas e ferramentas de visualização (KATIFORI et al., 2007; STOREY, 2007; PAULHEIM; MEYER, 2011; SILVA; FREITAS; SANTUCCI, 2012) e propor novas técnicas e ferramentas de visualização de dados, metadados e ontologias (SOUZA; SANTOS; EVANGELISTA, 2003; BUNTAIN, 2008).

As ontologias também estão sendo adotadas para a melhoria do processo de **Recuperação Multimídia**, que é uma área que desperta o interesse de muitos pesquisadores em função da complexidade de recuperação de informação não textual, presente em arquivos de imagem, áudio e vídeo. As pesquisas desenvolvidas neste contexto buscam agregar os recursos presentes nas ontologias para minimizar os problemas e propor soluções mais eficientes. Entre os recursos utilizados, pode-se destacar: raciocínio baseado em relações semânticas (ALBERTONI; BERTONE; DE MARTINO, 2005), expansão de espaço semântico (WU et al., 2008), busca e anotação baseadas em ontologias (BERTINI et al., 2010), anotação automática (GAO; LUO; FAN, 2006) e recuperação de conteúdo baseado nos conceitos da ontologia (POPESCU; MILLET; MOËLLIC, 2007). Ainda neste contexto, foi encontrada pesquisa comparando a recuperação de imagens baseada em palavras-chave com recuperação baseada em ontologias (WANG; LIU; CHIA, 2006).

Em menor quantidade, destacam-se as pesquisas relacionadas à **Avaliação** de projetos desenvolvidos na área, com o objetivo de identificar a aceitação do usuário e avaliar pontos fortes e fracos. Os trabalhos pertencentes a esta categoria referem-se à avaliação da eficácia e da usabilidade de interface de busca (DUDYCZ, 2015; ZHUHADAR; NASRAOUI, 2010) e à avaliação dos resultados de busca baseada em ontologia (KAMEL; LEE; POWERS, 2007).

Também foram observadas pesquisas relacionadas à utilização de ontologias em aplicações para **Dispositivos Móveis**, com o objetivo de agregar recursos semânticos à busca (SHIN et al., 2013a, 2013b; KO et al., 2014) e como componentes para o **Desenvolvimento de Interfaces e Sistemas**, sendo utilizadas para implementar recursos semânticos variados, bem como para possibilitar reutilização de código, personalização de aplicações, acesso e persistência de dados, mapeamento relacional e integração de ontologias (PAULHEIM, 2009; ROZELL et al., 2012; EL-ATAWY; KHALEFA, 2016).

A Tabela 3 apresenta resumidamente as áreas que possuem aplicações relacionadas ao uso de ontologias. Tais abordagens estão, de alguma forma, relacionadas à melhoria da recuperação de informação.

Tabela 3 - Abordagens envolvendo ontologias

Área	Abordagem	Quantidade
Recuperação de Informação na Web	Interfaces e recursos de suporte à busca e navegação	23
	Indexação e representação de documentos	
	Formulação de *queries*, refinamento e expansão de consultas	
	Visualização de informações	
Web Semântica	Interfaces de busca	12
	Métodos e técnicas de pesquisa, exploração e interpretação de dados	
Visualização de Informações	Avaliação de técnicas e ferramentas de visualização	07
	Proposta de novas técnicas de visualização	
Recuperação Multimídia (imagem, áudio e vídeo)	Raciocínio baseado em relações semânticas	07
	Expansão de espaço semântico	
	Busca e anotação baseadas em ontologias	
	Anotação automática	
	Recuperação de conteúdo baseada em ontologia	
	Comparação entre busca baseada em palavras-chave e em ontologias	
Avaliação	Avaliação da eficácia da interface	03
	Avaliação da usabilidade da interface	
	Avaliação de resultados de busca	
Dispositivos Móveis	Busca semântica em dispositivos móveis	03
Desenvolvimento de Interfaces e Sistemas	Acesso e persistência de dados	04
	Mapeamento relacional	
	Integração de ontologias	

Fonte: desenvolvido pela autora

A partir das áreas e abordagens envolvendo ontologias, a próxima seção apresenta as contribuições e propostas de melhorias em cada uma das principais problemáticas relacionadas à busca de informações na Web.

4.3 PROBLEMÁTICAS E PROPOSTAS DE MELHORIAS NO ÂMBITO DO USO DE ONTOLOGIAS NA BUSCA E RECUPERAÇÃO DE INFORMAÇÃO

Os trabalhos selecionados, todos elaborados a partir de abordagens relacionadas ao uso de ontologias na Recuperação de Informação, foram analisados de acordo com as contribuições apresentadas no âmbito dos seguintes paradigmas de busca: a) busca baseada em ontologia; b) navegação baseada em ontologia; e c) busca híbrida (busca e navegação baseadas em ontologias). Algumas pesquisas não tinham como objetivo discutir

nenhuma destas perspectivas e sim problemáticas relacionadas à visualização de informações e desenvolvimento de *software*, conforme pode ser observado na Tabela 4. Desta forma, foi possível analisar como as ontologias estão sendo utilizadas no suporte à busca de informações e quais são as principais propostas de melhorias apresentadas nesta área.

Tabela 4 - Problemáticas discutidas

Problemática / **Área**	**Paradigmas de Pesquisa**			**Outras Perspectivas**	
	Busca (*Searching*)	**Navegação (*Browsing*)**	**Híbrido (Busca e Navegação)**	**Visualização**	***Software***
Recuperação de Informação na Web	11	03	09	-	-
Web Semântica	08	-	04	-	-
Visualização de Informações	-	-	-	07	-
Recuperação Multimídia	02	01	04	-	-
Avaliação	-	-	03	-	-
Dispositivos Móveis	-	-	03	-	-
Desenvolvimento de Interfaces e Sistemas	-	-	-	-	04
Total	**21**	**04**	**23**	**07**	**04**
Validados (Estudos de Usuários)	**02**	**-**	**04**	**02**	**-**

Fonte: desenvolvido pela autora

Os estudos relacionados à primeira categoria são aqueles desenvolvidos com o propósito de utilizar ontologias para o aprimoramento da atividade de busca (*searching*), que é aquela realizada pelo usuário com um objetivo específico e previamente definido. A segunda categoria refere-se aos trabalhos nos quais as ontologias são empregadas para possibilitar a exploração de informações (*browsing*) e a terceira categoria, por sua vez, agrupou os trabalhos desenvolvidos com o propósito de apresentar melhorias para projetos que englobam a busca híbrida, contendo os dois paradigmas integrados: busca e navegação baseadas em ontologias (*searching and browsing*). Na quarta categoria, foram analisadas pesquisas relacionadas à visualização de informações, com destaque principalmente para a problemática da escalabilidade de grande conjunto de dados e ontologias. Na última categoria, foram agrupados todos os trabalhos que não tinham relação com a busca, navegação e visualização de conteúdos, mas sim com a estrutura de desenvolvimento de *software*.

4.3.1 Busca Baseada em Ontologias (*Searching*)

Dentre os trabalhos analisados, 21 destinavam-se à utilização de ontologias

para o aprimoramento da atividade de busca. Neste grupo, as pesquisas encontradas estão relacionadas às áreas de Recuperação de Informação na Web, Web Semântica e Recuperação Multimídia.

Em relação à **Recuperação de Informação na Web**, as principais contribuições e propostas de melhorias por meio do uso de ontologias são destinadas à formulação de *queries*, refinamento e expansão de consultas, indexação e representação de documentos. Entre as propostas apresentadas, encontram-se:

a) funcionalidades e recursos interativos: diversas funcionalidades e recursos baseados em ontologias estão sendo implementados nas caixas de busca, com o objetivo de auxiliar o usuário a elaborar expressões ou palavras-chave melhores e mais significativas: autosugestão de informação semântica (PAPADAKIS et al., 2009), filtro de domínio, recomendação de termos de busca, correção ortográfica, expansão automática de consultas e recurso de autopreenchimento (CHEN; CHUA, 2013);
b) busca semântica baseada em comportamento do usuário: sugestão de termos de busca (SENDHILKUMAR; GEETHA, 2008) e reformulação de consultas (LIMBU et al., 2006; MAURO, 2017) com base em ontologia personalizada, desenvolvida a partir do contexto, do perfil e do comportamento de busca do usuário;
c) busca semântica baseada em contexto: abordagem de busca que incorpora três técnicas principais: 1) contexto para cada conceito extraído de propriedades e axiomas; 2) interpretação de consulta baseada no contexto extraído; e 3) classificação dos resultados da busca, usando algoritmo de aprendizagem (GAVANKAR; LI; RAMAKRISHNAN, 2015);
d) busca baseada em relacionamentos semânticos: modelo de classificação baseado em ontologias, que considera o peso e número de relações semânticas significativas entre os recursos, para determinar a relevância de um documento (LEE; MIN; CHUNG, 2009);
e) busca baseada em múltiplas versões de ontologias: Kondylakis e Plexousakis (2011) propõem um estudo sobre a evolução de ontologias e apresentam uma plataforma Web que permite que as consultas dos usuários sejam respondidas automaticamente sobre múltiplas versões de

ontologia;

f) expansão de consultas baseada em vizinhança semântica: este recurso permite que a expansão de consultas enviadas pelo usuário seja realizada a partir dos sinônimos e da vizinhança semântica, ou seja, dos conceitos próximos presentes na ontologia (KHATTAK et al., 2008);

g) expansão de consultas baseada em múltiplas ontologias: o processo de expansão de consultas é baseado na integração de múltiplas ontologias de domínio. Essa integração é realizada por meio de técnicas de alinhamento de ontologias, que encontra equivalências entre conceitos de ontologias diferentes (CHUA; KIM, 2012);

h) refinamento de consultas: o processo de refinamento de consultas é realizado por meio de uma ontologia e da análise do *gap* de conhecimento do usuário (lacuna cognitiva, necessidade de informação, informação procurada, preferências e tarefa em questão) (STOJANOVIC, 2005).

No contexto da **Web Semântica**, os trabalhos envolvendo o uso de ontologias são destinados ao desenvolvimento de interfaces de linguagem natural ou refinamento de consultas, com o objetivo de reduzir a complexidade do uso da lógica formal (XML, RDF, SPARQL) e minimizar as dificuldades dos usuários na elaboração de buscas e na utilização dos recursos disponíveis (BERNSTEIN et al., 2005; GUPTA; BHOWMIK; GOVINDARAJU, 2008; ARFAOUI; HIDRI, 2016).

Outras propostas envolvem o desenvolvimento de interfaces gráficas para elaborar buscas na Web Semântica, tais como as apresentadas por Bast et al. (2007), Soylu et al. (2013, 2014) e Styperek, Ciesielczyk e Szwabe (2014). Além do desenvolvimento de interfaces de linguagem natural e visual para a elaboração de consultas, Demidova, Zhou e Nejdl (2013) propõem o desenvolvimento de algoritmos para explorar espaços de interpretação de consultas de palavras-chave, além de novas opções de interação para permitir uma melhor escalabilidade.

As pesquisas relacionadas à **Recuperação Multimídia** também utilizam os recursos das ontologias para propor melhorias para a área. Neste sentido, os trabalhos encontrados com este propósito visam expandir o espaço semântico de uma busca de vídeo, formulada pelo usuário por meio de uma palavra-chave e, a partir deste recurso, apresentar melhores resultados (WU et al., 2008). Wang, Liu e Chia (2006) apresentam um estudo comparativo entre a recuperação de imagem baseada em palavras-chave e a recuperação de

imagem baseada em ontologias. Apesar de não apresentar nenhuma proposta de implementação de novos recursos, o trabalho destaca a eficácia da utilização de ontologias para a recuperação de imagens, bem como apresenta contribuições e desafios para a área.

4.3.2 Navegação Baseada em Ontologias (*Browsing*)

Na categoria referente à navegação baseada em ontologias, foram encontrados 4 trabalhos envolvendo ontologias, com contribuições importantes para as áreas de Recuperação de Informação na Web e Recuperação Multimídia.

Os trabalhos pertencentes ao campo da **Recuperação de Informação na Web**, em sua maioria, destinam-se ao desenvolvimento de interfaces gráficas baseadas em ontologias, que permitem a visualização e exploração do *corpus* documental e seleção de elementos (*tags*, facetas, *links*) para a formulação de *queries* e expressões de busca. Neste contexto, Thai e Handschuh (2009) apresentam uma interface visual, com base em ontologia pessoal, que permite aos usuários a navegação em coleções de texto através de diferentes perspectivas e níveis de detalhamento e, posteriormente, a realização das consultas de interesse.

Issertial e Tsuji (2013), por sua vez, discutem uma interface de usuário, com recursos gráficos para facilitar a navegação e busca de chamadas de artigos (*call for paper*). A proposta utiliza banco de dados estruturado a partir da extração de dados relevantes das chamadas de artigos, via mineração de textos e, posteriormente, enriquecidos através de uma ontologia. Ao lado destes dois estudos, a proposta de Velart e Šaloun (2007) refere-se a uma interface visual para a organização de um curso *on-line* e navegação de materiais de aprendizagem.

Em relação à área de **Recuperação Multimídia**, o trabalho de Albertoni, Bertone e De Martino (2005) apresenta uma abordagem baseada em técnicas de visualização de ontologias e interação gráfica para apoiar a busca de imagens geográficas e facilitar a análise de metadados. A abordagem fornece dois tipos de raciocínios: um baseado em visão e outro baseado em ontologia. O raciocínio visual tem como objetivo possibilitar o reconhecimento de padrões e a interação direta com eles, enquanto que o raciocínio com base em ontologia é proposto para auxiliar a descoberta de relações semânticas entre os dados. Desta forma, o usuário tem uma visão geral e compacta dos dados disponíveis, facilitando a interpretação do conjunto de resultados, propriedades e relacionamentos existentes.

4.3.3 Busca Híbrida Baseada em Ontologias (*Searching and Browsing*)

A maior parte dos trabalhos selecionados foram classificados nesta categoria, pois apresentam contribuições que visam a utilização de ontologias em projetos direcionados à busca híbrida, que contêm ambas funcionalidades: busca e navegação (*browsing*) de informações. Neste cenário, os trabalhos são direcionados a diversas áreas: Recuperação de Informação na Web, Web Semântica, Recuperação Multimídia, Avaliação e Dispositivos Móveis.

No domínio da **Recuperação de Informação na Web**, os estudos agregam as ontologias para melhorar a busca e exploração de informações e facilitar a realização de consultas pelo usuário. Entre as propostas apresentadas, encontram-se:

a) recursos implementados em interfaces de busca semântica: diversos recursos estão sendo implementados para facilitar a elaboração de consultas, tais como: nuvem de termos, onde as informações podem ser filtradas e selecionadas (WONG; LIU; BENNAMOUN, 2010); navegador de árvore de ontologia, conjunto de filtros de documentos, conjunto de painéis de visualização, visualizador de títulos de documentos recuperados, navegador de documentos recuperados (LIN et al., 2010); menu de busca e espaço de navegação (GROTH; LANNERÖ, 2006);

b) ferramenta de busca e recuperação de informação: a proposta apresentada por Siahbani et al. (2013) consiste no desenvolvimento de um sistema de visualização de informações (navegador visual), que pode ser usado para explorar e navegar eventos extraídos de uma grande coleção de texto em linguagem natural, com base na exploração de rótulos de função semântica (quem fez o que a quem, quando e como). No mesmo contexto, Tijerino e Sanati (2005) apresentam um mecanismo de busca que combina métodos de pesquisa estatística e semântica para fornecer uma interface semântica visual para educadores que desejam distribuir, compartilhar e pesquisar materiais de preparação de cursos. Baskaya, Kekäläinen e Järvelin (2010) apresentam uma ferramenta baseada em ontologias multilíngue, com múltiplas funcionalidades, direcionada à busca, exploração, análise e anotação de documentos *on-line*. A ferramenta oferece suporte a: 1)

busca por palavras-chave: na qual o pesquisador pode combinar conceitos da ontologia aos termos que está procurando; 2) **anotação**: documentos recuperados podem ser anotados por meio das ontologias escolhidas, em várias línguas, para identificar conceitos de interesse no documento; 3) ***clustering***: agrupamento e apresentação de documentos recuperados; 4) ***bookmarking***: o pesquisador pode armazenar os *links* de uma ontologia para realizar buscas e análises posteriores; e 5) **integração de funções**: palavras de texto e rótulos de *clusters* identificados, podem ser usados com flexibilidade em uma próxima pesquisa;

c) interface visual para busca de documentos 3D: interface visual destinada à navegação e exploração de mundos 3D, incluindo informações geométricas e textuais. Possui recursos de busca, navegação e anotação de documentos (PITTARELLO; GATTO, 2012a, 2012b);

d) interface de busca dinâmica baseada em modelos de conhecimento: interface de busca dinâmica para o comércio eletrônico, projetada para possibilitar que o usuário navegue rapidamente nas páginas da Web e obtenha o conteúdo semanticamente relevante e relacionado ao seu perfil (CARVALHO et al., 2002).

Para a área da **Web Semântica**, as propostas apresentam estruturas de busca e navegação, elaboradas com o propósito de facilitar o acesso e a análise de grande quantidade de dados. Discutem a problemática da consulta às estruturas heterogêneas da Web Semântica, bem como a escalabilidade de informações (SUH; BEDERSON, 2002; DELIGIANNIDIS; KOCHUT; SHETH, 2007; FREITAS; CURRY, 2014; SOYLU et al., 2016).

No contexto da **Recuperação Multimídia**, as propostas de melhorias são fundamentadas no uso de ontologias e interfaces visuais de navegação para possibilitar melhores resultados na busca e exploração do *corpus* documental. O sistema proposto por Popescu, Millet e Moëllic (2007) foi projetado para realizar a busca de imagens de duas formas: por meio de navegação conceitual e navegação visual. Também permite a navegação do usuário na página de resultados, possibilitando que novas buscas sejam realizadas.

Gao, Luo e Fan (2006) também colaboraram com a área e apresentaram um sistema de busca e navegação de imagem em banco de dados de larga escala, com o recurso

de anotação automática, desenvolvido a partir de três abordagens diferentes: abordagem baseada em segmentação, abordagem baseada em grade e abordagem baseada em relacionamentos.

Outro exemplo importante relacionado à Recuperação Multimídia é o sistema Sírio, discutido por Alisi et al. (2009) e Bertini et al. (2010), que tem como objetivo apresentar um motor de busca de vídeos baseado em ontologias, com aplicações voltadas para a transmissão de notícias, vídeos de vigilância e documentários de herança patrimonial e cultural. Possui recurso de anotação semântica, expansão de consultas e especialização de conceitos, além de uma interface voltada para usuários técnicos e não técnicos, permitindo diferentes modalidades de consulta: texto livre, linguagem natural e composição gráfica de expressões de busca.

Em relação à área de **Avaliação**, foram encontrados trabalhos com o objetivo de analisar a eficácia de uma interface de busca visual (DUDYCZ, 2015), a usabilidade de um sistema de busca (ZHUHADAR; NASRAOUI, 2010) e os resultados de um estudo de usuário comparativo entre a eficácia das consultas na Web formuladas por ontologias com aqueles obtidos usando apenas um motor de busca baseado em palavras-chave (KAMEL; LEE; POWERS, 2007). Os estudos contribuem com a área, uma vez que apresentam resultados positivos sobre a adoção de ontologias no processo busca e recuperação de informação e promovem discussões importantes sobre o assunto.

No âmbito dos **Dispositivos móveis**, os trabalhos apresentados por Ko et al. (2014) e Shin et al. (2013a, 2013b) discutem a implementação de busca semântica para *smartphones*, agregando recursos e funcionalidades para promover resultados mais significativos, entre eles: 1) refinamento e expansão da busca, por meio do uso de ontologia subjacente; 2) pesquisa baseada em intenção e contexto de busca; 3) busca a partir do reconhecimento de voz; e 4) gráfico de respostas, no qual o usuário tem condições de indicar quais informações ele realmente precisa. Neste contexto, a captura do significado pretendido da consulta e a expansão da busca contribuem para que a pesquisa seja mais precisa, uma vez que os conceitos próximos de uma ontologia possuem uma conexão semanticamente forte.

4.3.4 Visualização de Informação

Nesta categoria foram agrupados todos os trabalhos relacionados especificamente à problemática da Visualização de Informação, incluindo manipulação de ontologias e escalabilidade de grande quantidade de dados. Tzitzikas e Hainaut (2006)

analisaram aspectos importantes relacionados à manipulação de grandes ontologias, destacando características de classificação, filtragem, *clustering* e *layouts* de visualização. Katifori et al. (2007), Storey (2007), Paulheim e Meyer (2011) e Silva, Freitas e Santucci (2012) apresentam estudos sobre métodos e técnicas de visualização de ontologias existentes, direcionados principalmente para o *design*, gerenciamento e navegação, destacando características, vantagens e desvantagens de cada um deles.

Além da análise dos aspectos relacionados à manipulação de dados em grande quantidade e dos métodos e técnicas de visualização existentes, novas ferramentas de visualização de dados são propostas: aplicativo de visualização de ontologias, que utiliza a renderização tridimensional (BUNTAIN, 2008) e ferramenta de suporte à visualização de ontologias por meio de árvores hiperbólicas (SOUZA; SANTOS; EVANGELISTA, 2003).

4.3.5 Outras Perspectivas

Somadas às propostas de melhorias relacionadas aos paradigmas de pesquisa (busca, navegação e busca híbrida) e à visualização de informações, a área de desenvolvimento também tem utilizado as ontologias para a implementação de recursos diferenciados em interfaces e sistemas de busca.

El-Atawy e Khalefa (2016) utilizaram as ontologias para reduzir o custo do desenvolvimento e da manutenção de sistemas médicos, bem como para minimizar as dificuldades relacionadas à área: múltiplas especialidades, diversidade de conceitos clínicos, dados temporais, transações financeiras e ampla variação de detalhes em diferentes domínios e configurações. Desta forma, as ontologias foram adotadas para modelar as soluções e possibilitar a reutilização de código, a extensão e a personalização dos sistemas e o tratamento da persistência, acesso e troca de dados. Além disso, as ontologias facilitam a modelagem da lógica de negócios (fluxos e regras) dos sistemas médicos, garantindo a qualidade de dados, o suporte às especialidades da área e o desenvolvimento de sistemas fortemente acoplados a modelos de dados.

Rozell et al. (2012) adotaram as ontologias para a construção de uma estrutura para o desenvolvimento de interfaces de usuários interativas para catálogos de dados. Para isto, utilizaram tecnologias da Web Semântica, incluindo ontologia OWL, para descrever a semântica dos serviços de dados e os componentes da interface de usuário, bem como para possibilitar a extensibilidade e a reutilização. Para testar a proposta, a estrutura foi aplicada no desenvolvimento de três aplicações diferentes: 1) interface facetada de navegação,

integrada com um mapeamento interativo e ferramenta de visualização de dados para a área de oceanografia biológica e química; 2) navegador facetado para mais de 700.000 conjuntos de dados públicos abertos em mais de 100 catálogos mundiais; e 3) interface de usuário para um observatório virtual no campo da física solar terrestre.

Zheng et al. (2012) apresentaram uma ferramenta que fornece uma camada intermediária entre repositórios de ontologias e banco de dados para apoiar operações semânticas, baseadas em linguagens de consulta, diretamente de dentro de um banco de dados convencional. Desta forma, facilita a integração e o uso de ontologias em bancos de dados e melhora o desempenho para a maioria das buscas.

Paulheim (2009) apresentou uma abordagem para a modularização de interfaces do usuário, que visa facilitar o desenvolvimento de aplicações em grande escala e reduzir a dependência entre *plugins*. Utiliza ontologia de aplicação para descrever os requisitos e os componentes da interface, além de ontologia de domínio para descrever os termos do domínio da aplicação. A interoperabilidade entre os *plugins* pode ser alcançada por meio de raciocínio ontológico, mesmo quando eles são atualizados ou recombinados.

4.4 VALIDAÇÃO DE PROPOSTAS POR MEIO DE ESTUDOS DE USUÁRIOS

O desenvolvimento de estudos de usuários para a validação de pesquisas e propostas de melhorias, embora de grande importância, são pouco utilizados na área de Recuperação de Informação. Dentre os 59 trabalhos analisados no contexto da nossa pesquisa, apenas 08 foram validados por usuários, o que representa menos de 15% do total. Isto significa que os dados apresentados e discutidos, na maioria dos trabalhos, referem-se apenas à opinião dos autores e desenvolvedores, impossibilitando, portanto, uma análise mais aprofundada sobre o desempenho, a eficácia, a usabilidade e as limitações para adoção das ferramentas desenvolvidas, bem como sua aceitação pelo usuário final e sua contribuição efetiva no processo de recuperação de informação.

Em relação a este quesito, os trabalhos validados foram encontrados no grupo das abordagens relacionadas a: 1) busca baseada em ontologias, nas áreas de Recuperação de Informação na Web (CHEN; CHUA, 2013) e Web Semântica (BERNSTEIN et al., 2005); 2) busca e navegação do *corpus* baseadas em ontologias, nas áreas Web Semântica (DELIGIANNIDIS; KOCHUT; SHETH, 2007; SOYLU et al., 2016) e Avaliação (KAMEL; LEE; POWERS, 2007; DUDYCZ, 2015); e 3) visualização de informações (SOUZA; SANTOS; EVANGELISTA, 2003; PAULHEIM; MEYER, 2011).

Dentre os trabalhos avaliados, Chen e Chua (2013) realizaram testes com 50 usuários para analisar o *design* de uma interface de busca baseada em ontologias, contendo diversos recursos interativos para auxiliar a busca. Foi possível obter *feedback* sobre os seguintes itens: facilidade de uso, facilidade de aprendizagem, utilidade da correção ortográfica e análise das funções de ajuda aos usuários na formulação de buscas.

O processo de validação do protótipo de uma interface de busca para a Web Semântica apresentada por Bernstein et al. (2005) contou com duas etapas: 1) execução de diversas consultas altamente complexas e comparação do desempenho de recuperação de informação; 2) testes de usabilidade e avaliação de desempenho realizados com 20 usuários, que executaram tarefas com diferentes níveis de complexidade.

Kamel, Lee e Powers (2007) elaboraram testes envolvendo a participação de 10 usuários para comparar a eficácia entre as buscas realizadas na Web (por meio de palavras-chave no Google) e as realizadas em uma ferramenta de busca baseada em ontologias. Entre os dados coletados, foi possível identificar a precisão dos resultados encontrados em cada grupo de consultas e a importância da definição do contexto de busca para a recuperação de resultados relevantes. Foi possível ainda avaliar a visão dos usuários em relação ao *design* da interface.

O estudo de usuário apresentado por Dudycz (2015), com o objetivo de avaliar a eficácia e a usabilidade de uma interface de busca baseada em ontologias, contou com a utilização de questionário de satisfação do usuário e análise de usabilidade, que foi realizada de acordo com as métricas propostas por Shneiderman (1997): 1) tempo de familiarização com o sistema; 2) velocidade de desempenho da realização de tarefas; 3) erros nas tarefas atribuídas; 4) retenção das características do sistema; e 5) satisfação do usuário. O estudo foi realizado com 143 usuários, que foram subdivididos em quatro experimentos com propósitos e características diferenciadas.

Deligiannidis, Kochut e Sheth (2007) desenvolveram estudos envolvendo 10 participantes, com a intenção de avaliar o protótipo de uma ferramenta para a busca e a navegação de dados na Web Semântica. Foi possível identificar o desempenho do usuário em relação ao uso da ferramenta (tempo para completar a tarefa e para encontrar a informação desejada) e sua satisfação em relação à interface (interface amigável, facilidade de uso, facilidade de começar a utilizar e satisfação geral).

Ainda com o objetivo de avaliar uma ferramenta desenvolvida para o contexto da Web Semântica, Soylu et al. (2016) propuseram um estudo com usuários de duas empresas do ramo industrial: Statoil e Siemens. A análise permitiu a avaliação de dois grupos

com características diferenciadas, uma vez que os participantes da Siemens possuíam mais habilidades e conhecimentos sobre a realização de buscas predefinidas que os participantes da Statoil. O estudo teve a finalidade de identificar: 1) necessidades de informação: domínios de interesse, tipos comuns de consultas e a sua complexidade, tipos de tarefas; 2) atributos de qualidade: análise que interferem no tempo de execução, *design* e experiência do usuário, tais como: usabilidade, modularidade, escalabilidade, adaptatividade, adaptabilidade, extensibilidade, interoperabilidade, portabilidade e reusabilidade.

Souza, Santos e Evangelista (2003) realizaram apenas estudos preliminares para avaliar uma ferramenta de visualização de ontologias por meio de árvores hiperbólicas. Os usuários foram treinados para utilizar o protótipo e relatar diretamente sua opinião sobre ele em relação à otimização das atividades e ganho de produtividade. Por fim, Paulheim e Meyer (2011) apresentaram um estudo com 22 usuários voluntários para avaliar um protótipo de visualização de ontologias e redes semânticas. Para o desenvolvimento desta etapa da pesquisa, foram propostas doze tarefas de três níveis diferentes de complexidade. Com a utilização de um questionário, foi possível avaliar o tempo de conclusão das tarefas propostas, a taxa de erro, bem como a experiência do usuário nas seguintes perspectivas: atração, facilidade, perspicácia, eficiência, estimulação e novidade.

4.5 RESUMO E DISCUSSÃO DO CAPÍTULO

As análises realizadas nesta etapa da pesquisa permitiram traçar um panorama sobre a utilização de ontologias no âmbito da busca e recuperação de informação, destacando os tipos de ontologias adotadas, as áreas de desenvolvimento de pesquisas, as principais abordagens, problemáticas, contribuições e propostas de melhorias.

Os dados apresentados permitem uma visão sistematizada sobre como as ontologias estão sendo utilizadas, destacando principalmente a possibilidade de agregar novos recursos e funcionalidades aos sistemas atuais, tais como: recursos interativos de busca e recuperação de informação, busca semântica, expansão e refinamento de consultas, desenvolvimento de interfaces gráficas de navegação, ferramenta visual para a exploração de conteúdos, interface visual para mundos 3D, entre outros. As ontologias também se destacam em projetos direcionados à Web Semântica, principalmente em propostas que apresentam interfaces para minimizar as dificuldades do usuário com a elaboração de buscas e visualização de grande quantidade de dados. Além dessas duas áreas de destaque, também foram encontradas contribuições nas áreas de Visualização de Informações, Recuperação

Multimídia, Dispositivos Móveis e Desenvolvimento de Interfaces e Sistemas.

Apesar das inúmeras propostas e contribuições observadas no contexto da busca e recuperação de informações, foi possível perceber que, em muitos projetos, as ontologias são utilizadas apenas como recurso de filtragem de informações (quando apresentadas como *tags*, facetas e *links*), deixando em segundo plano todo o potencial para o qual elas foram, de fato, desenvolvidas. Neste sentido, uma importante lacuna de pesquisa observada é a necessidade de desenvolvimento de ferramentas que utilizem os recursos das ontologias, principalmente o contexto e o significado existente nas relações entre os conceitos, para amplificar a busca e não apenas para filtrar informações e resultados.

Neste contexto, a partir dos resultados observados, foi possível delimitar o problema de pesquisa que fundamenta este estudo.

5 HERMENÊUTICA

Este capítulo tem como objetivo discutir a Hermenêutica, buscando identificar os fundamentos e os princípios relacionados à compreensão e interpretação documental. Também visa apresentar os principais precursores da área, destacando especialmente as contribuições oferecidas pelos estudos de Hans-Georg Gadamer referente à Hermenêutica Filosófica. Essa escolha se deu em função do entendimento de Gadamer sobre o processo de compreensão e pela sua proximidade em relação aos estudos sobre o comportamento de busca do usuário no contexto da Web. Os conceitos aqui levantados deram subsídios à definição das funcionalidades do "modelo de interface de busca" referentes à análise do *corpus* documental, relacionadas tanto com a indexação dos documentos inseridos na base de dados do repositório digital, como com interpretação dos resultados apresentados ao usuário.

■■■

Do ponto de vista histórico, a hermenêutica surgiu como reflexão teórica-metodológica para a interpretação dos textos sagrados, clássicos (literários) e jurídicos (leis). (DOMINGUES, 2004, p.345). Seu principal objetivo era reconstruir o sentido original dos textos, analisando as condições sob as quais a compreensão ocorria e facilitando o processo de interpretação (SCHLEIERMACHER, 2003).

Com o decorrer dos anos, diversas concepções filosóficas e inúmeras teorias foram associadas para determinar o conjunto de regras que possibilitariam a "arte universal de interpretação" (DILTHEY, 1999, p.15). Compreende um vasto campo com diferentes objetivos e posições filosóficas, assim como inúmeras teorias de interpretação de textos inspiradas nos estudos desenvolvidos por grandes filósofos e pesquisadores da área como Schleiermacher (1768-1834), Dilthey (1833-1911), Weber (1864-1920), Heidegger (1889-1976), Gadamer (1900-2002), Habermas (1929-) e Ricoeur (1913-2005).

Etimologicamente, a hermenêutica deriva do verbo grego *hermeneuein* ou *hemeneutice*, que significa expressar em voz alta, explicar ou interpretar e traduzir. No latim, significa *interpretativo*, que deu origem ao termo *interpretação* em português (SCHMIDT, 2014). Sua constituição originária está associada à mitologia grega, na representação de Hermes, o deus mensageiro, que fazia a mediação entre deuses e humanos, tendo a habilidade de resolver mal-entendidos e apaziguar conflitos. Por esta razão, era considerado o patrono da comunicação e do entendimento humano (PALMER, 1969; INWOOD, 1998).

De acordo com Grondin (2012), a utilização do termo também está relacionada ao teólogo estraburguense Dannhauer, que o associou ao que ele chamava antes de *Auslegungslehre* ou a arte da interpretação. Em 1564 Dannhauer foi o primeiro a utilizar o termo no título de uma obra, em sua *Hermeneutica sacra sive methodus exponendarum sacrarum litterarum*, que resume por si só o sentido clássico da disciplina: a hermenêutica sagrada, ou seja, o método para interpretar textos religiosos.

É considerada um ramo da Filosofia que estuda o problema geral da compreensão, não apenas relacionado ao que diz um texto, mas a que significa compreender um texto (MELO, 2016). Apesar dos diversos entendimentos sobre o conceito de hermenêutica presentes na literatura, o que mais se aproxima do discutido no contexto deste estudo é o de Schmidt (2014, p.152) que afirma que não se trata de propor um método de análise estrutural ou linguística do texto, tampouco, "fazer uma ponte temporal e reconstruir a situação original do texto, mas descobrir o que o texto tem a dizer para nós". Portanto, a hermenêutica pode ser conceituada como sendo o conjunto de teorias voltadas para a interpretação de algo, não somente de um texto escrito, mas de tudo que se possa atribuir significado e sentido (STAGLIANO, 2016).

Schmidt (2014) apresenta os diversos estágios de desenvolvimento da hermenêutica, bem como as diferentes abordagens discutidas pelos estudiosos ao longo do tempo, e afirma que desde a Antiguidade, um grande número de teorias da interpretação foram desenvolvidas para apoiar disciplinas específicas, entre as quais destacaram-se: a *hermenêutica legal*, que teve como objetivo buscar a interpretação e a codificação correta das leis; a *hermenêutica bíblica*, que buscou o desenvolvimento de regras e teorias para a interpretação da Escrituras Sagradas e a *hermenêutica filológica* que, no Renascimento, cresceu e se concentrou na interpretação de textos clássicos.

A partir de propostas filosóficas distintas, é possível identificar especificidades e divergências nas abordagens apresentadas por cada teórico relacionado ao campo da hermenêutica. Na hermenêutica medieval, o pensador de destaque foi Santo Agostinho, que estabeleceu regras para a interpretação das sagradas escrituras, distinguindo o sentido próprio do sentido figurado, com o intuito de facilitar a compreensão das pessoas. Suas propostas repercutiram até o século XIX, quando Schleiermacher conferiu outro enfoque à hermenêutica (GOMES, 2015).

Friedrich Daniel Ernst Schleiermacher (1768-1834), teólogo, protestante, estudioso da Bíblia e de clássicos, foi o proponente de uma nova hermenêutica, denominada de hermenêutica geral ou universal, que teve como principal objetivo unificar as várias teorias

existentes (hermenêutica legal, bíblica e filológica), abrangendo a interpretação de todos os textos, independente do gênero e de doutrinas. Desta forma, Schleiermacher (2005) entende a hermenêutica como a ciência da compreensão linguística, atribuindo a ela o caráter de 'ciência' ou 'arte' da compreensão. Isto faz com que ela seja pensada como sendo uma ciência geral, cujos princípios podem ser aplicados a qualquer discurso e não somente à forma escrita como era a proposta da hermenêutica clássica, tendo como principal objetivo compreender corretamente aquilo que foi expressado por outra pessoa, seja por meio da linguagem oral ou escrita.

Para que isto fosse possível, Schleiermacher atribuiu à hermenêutica um caráter metódico e psicológico, dividindo-a em duas práticas: a interpretação gramatical e a interpretação técnica ou psicológica. A primeira refere-se à compreensão da linguagem usada pelo autor (significado das palavras), incluindo regras gramaticais e semânticas. Isto significa que, de acordo com Schleiermacher (2005, p.44), o "sentido de cada palavra num dado local precisa ser determinado através de sua conexão com as palavras ao seu redor". A segunda é o complemento da primeira e está relacionada à compreensão de como o autor pensava e de como estes pensamentos alcançam a expressão, considerando sua cultura e seu contexto histórico. Isto indica que o sentido apropriado de uma palavra precisa ser determinado a partir do seu contexto. Assim, uma atividade depende da outra para completar a tarefa de interpretação. O objetivo é, portanto, "reconstruir o processo criativo do autor, descobrir o significado intencionado do autor, e talvez compreender o autor melhor do que ele próprio se compreendia", uma vez que, na reconstrução do pensamento do autor, o intérprete terá consciência de influências que o autor não tinha (SCHMIDT, 2014, p.19).

Alguns fatores influenciam a linguagem utilizada pelo autor, tais como seu lugar na história, sua educação, sua ocupação e até mesmo seu dialeto. Desta forma, não é possível compreender um texto por completo a não ser que se conheça a cultura que deu origem ao texto, o que pressupõe o conhecimento dos textos e acontecimentos que constituem a cultura. Isto significa que, quanto melhor o intérprete conhecer o autor, mais fácil será compreender sua linha de pensamento.

> O processo de esquematização na linguagem permite que o intérprete compreenda o que o autor quer dizer, especialmente no uso cotidiano da linguagem. Em casos mais complexos, os métodos divinatórios e comparativo permitem que o intérprete reconstrua o significado do texto. Para aplicar estes métodos, o intérprete precisa conhecer a linguagem do autor e plateia a que ele se dirigia, a vida do autor e sua época, e também o assunto discutido. O método comparativo é usado com os contemporâneos do autor, e especialmente, com nós mesmos. O método divinatório se baseia

> numa semelhança entre os seres humanos e funciona através da analogia e comparação com nós mesmos. A hermenêutica requer talento e experiência da parte do intérprete (SCHMIDT, 2014, p.48).

Para Schleiermacher, existe uma interdependência entre a parte e o todo, conhecida como círculo hermenêutico, no qual "o conhecimento completo está sempre neste círculo aparente, onde cada particular só pode ser compreendido através do geral do qual faz parte, e vice-versa", ou seja, não é possível compreender o todo até as partes tenham sido compreendidas (SCHLEIERMACHER, 2005, p.24). Para que isto seja possível, a partir do conhecimento suficiente da linguagem, é necessário realizar primeiro uma leitura superficial do texto para conseguir uma visão do todo e então recomeçar interpretando cada uma das partes até ser possível reconstruí-lo em sua gênese, estrutura e significado (BATISTA, 2012; SCHMIDT, 2014).

Wilhelm Dilthey (1833-1911), biógrafo de Schleiermacher, na segunda metade do século XIX, ampliou a compreensão hermenêutica a partir da proposta de uma metodologia única para as ciências humanas, que foi apresentada no livro *Introdução às ciências humanas*, de 1883. Segundo ele, o método das ciências naturais não era apropriado para as ciências humanas, uma vez que os eventos da natureza devem ser explicados, mas os fenômenos históricos, os valores e a cultura devem ser compreendidos (DILTHEY, 2010).

Dilthey é importante para a discussão da hermenêutica, uma vez que a sua análise da compreensão incorpora vários elementos da hermenêutica de Schleiermacher e influencia, posteriormente, os estudos de Heidegger e Gadamer. Tal como Schleiermacher, a compreensão vista por Dilthey não é restrita a textos e discursos, mas a qualquer expressão da vida humana (gestos, ações, eventos passados etc.).

Desta forma, dois tipos de compreensão podem ser observados na teoria de Dilthey: elementares e superiores. O primeiro tipo refere-se à compreensão de expressões simples, como um discurso, uma ação, um gesto ou o medo, que só é possível porque existem uma cultura e uma linguagem compartilhadas, não havendo, portanto, nenhuma interferência no processo de compreensão relacionada ao contexto mais amplo da vida para a verificação do sentido apreendido. Assim, depois de aprender os significados que as palavras têm no contexto ordenado da linguagem, é possível compreender a manifestação individual numa sentença enunciada. Para Dilthey (2002, p.230), "uma sentença é inteligível devido aos aspectos comuns que existem numa comunidade linguística sobre o significado das palavras e das formas de inflexão, e sobre o sentido da estrutura sintática".

O segundo tipo de compreensão, por sua vez, é mais elevado e está

relacionado ao processo de transição da interpretação elementar para a superior, na medida em que surgem as incertezas, as contradições, as dificuldades de interpretação. Neste ponto, existe a necessidade de compreender a individualidade do autor para, posteriormente, compreender o que ele diz ou fala. Isto significa que a compreensão mais elevada geralmente envolve a compreensão dos indivíduos e não somente a compreensão geral (INWOOD, 1998; BATISTA, 2012).

Este entendimento relaciona-se diretamente com os elementos de análise e de síntese adotados no processo de compreensão de um documento. Em regra, eles não coincidem e não são equivalentes, uma vez que na "via da análise descobrem-se os elementos da significação que não têm nenhuma relação com as coisas ditas; na via da síntese, revela-se a função da significação que é *dizer*, e finalmente *mostrar*" (RICOUER, 1988, p.64).

De acordo com Schmidt (2014), Dilthey identifica quatro pontos importantes na obra de Schleiermacher para o desenvolvimento da hermenêutica:

a) a análise da compreensão fundamenta a interpretação, ou seja, as regras hermenêuticas para interpretar textos são um caso específico do processo de compreensão em geral;
b) o intérprete e o autor compartilham uma natureza humana comum;
c) o intérprete pode recriar as ideias do autor em função da natureza humana compartilhada;
d) pode-se compreender o significado inteiro de um texto a partir das palavras, ou seja, é possível entender o todo através de sinais individuais determinados apenas relativamente (palavras).

Esses quatro pontos estão relacionados à condição de compreensão que se dá tanto na relação direta do autor com o intérprete, quanto com suas obras, o que Dilthey denomina de "transposição do eu". Essa transposição é possível devido à homogeneidade do espírito e à contribuição do intérprete ao sentido da manifestação interpretada. Nesse ponto, Dilthey concorda com Schleiermacher acerca da possibilidade de se compreender um autor melhor do que ele próprio, pois: "as expressões da vivência contêm mais do que o que existia na consciência do poeta ou do artista e por isso mesmo têm também mais poder de evocação" (DILTHEY, 1984, p.268).

Com base nesta transposição, é possível recriar ou reexperienciar uma determinada realidade. Isto significa que um "historiador não pode realmente experimentar o passado, mas precisa, interpretativamente, compreender os documentos escritos, os fatos históricos, as motivações dos autores principais, e assim por diante", de forma que ao

apresentá-los, seja possível para o leitor reexperimentar o evento e seu significado da forma como o historiador pretendeu (SCHMIDT, 2014).

As contribuições apresentadas por Schleiermacher e Dilthey foram essenciais para entender a virada ontológica da hermenêutica desencadeada nos anos 1920 por Heidegger e Gadamer. Martin Heidegger (1889-1976), filósofo alemão, combinou o método de pesquisa fenomenológica de Husserl com aspectos da teoria da compreensão da vida de Dilthey. Além disso, teve a influência de pensadores como Platão, Aristóteles, Mestre Eckhard e Friedrich Nietzsche. Seu trabalho de maior impacto foi *Ser e tempo*, publicado em 1927 e que se caracterizou por ser uma das obras filosóficas mais importantes do século XX. Seus estudos possibilitaram o avanço da hermenêutica para o centro da reflexão filosófica e permitiram o desenvolvimento da hermenêutica existencial ou da facticidade (BATISTA, 2012; SCHMIDT, 2014).

Para Heidegger, o entendimento não é um método de leitura, tampouco um procedimento de reflexão crítica, mas um "modo de ser", e como tal é uma característica do ser humano. Neste contexto, a hermenêutica seria filosófica, e não científica; ontológica, e não epistemológica; existencial, e não metodológica. Seria responsável, portanto, por procurar a essência da compreensão, e não a normatização do processo compreensivo. O estudo da compreensão poderia ser confundido com o estudo da existência, uma vez que permitiria o conhecimento do Ser. Desta forma, sua teoria deixou de lado a interpretação de textos e as ciências humanas e tomou a própria existência como objeto da hermenêutica (BATISTA, 2012; GOMES, 2015).

A abordagem heideggeriana elevou a hermenêutica de um sentido técnico ou metodológico para adquirir um caráter fenomenológico que se confundia com a própria ontologia. O método de Heidegger é ontológico porque investiga a existência humana, isto é, o *Dasein*. "[...] o Dasein é o ente que compreende o ser, o que significa compreendê-lo em sua existência e entender a existência como possibilidade sua, de ser ou de não ser si mesmo, com a qual está concernido" (NUNES, 2002, p.12). Com outras palavras, Bonfim (2010, p.77) define *Dasein* como sendo a "condição do sujeito de se ver imerso em um contexto histórico-linguístico, condição de possibilidade que molda e fornece um *horizonte de sentidos*".

De acordo com Heidegger (2005), o entendimento, que precede a interpretação, é intuitivo e pré-reflexivo. A ideia de verdade surge como um acordo entre o julgamento do intérprete e o mundo, considerando os seus pressupostos. Desta forma, "o entendimento se dá no seio de uma estrutura de antecipação, de uma antecipação de significatividade, regida pela existência e por sua necessidade de orientação" (GRONDIN,

2012, p.47).

> A compreensão humana se orienta a partir de uma pré-compreensão que emerge da eventual situação existencial que demarca o enquadramento temático e o limite de validade de cada tentativa de interpretação. Pré-estrutura, em sua 'hermenêutica existencial da facticidade', significa que o 'Dasein', o ser-aí humano, se caracteriza por uma interpretação que lhe é peculiar e que se encontra antes de qualquer locução ou enunciado. Em seu conceito de compreensão sustentado na fórmula 'entender-se sobre algo', Heidegger afirma que entender teoricamente de um contexto, fato ou coisa, significa estar em condições de enfrentá-los, levá-los a cabo, poder começar algo com eles. Essa compreensão é designada de compreensão 'existencial', ou seja, como modo de ser por força do qual nós conseguimos e procuramos situar-nos neste mundo (BRITO et al., 2007, p.7).

Com a mudança da ideia de verdade, Heidegger reformulou também a concepção do círculo hermenêutico, que era entendido em termos de uma relação mútua entre o texto e suas partes individuais. O novo conceito trata o círculo hermenêutico de uma forma completamente diferente, referindo-se à interação entre o autoentendimento do intérprete e seu entendimento do mundo. A partir desta virada ontológica, a hermenêutica passa a se preocupar com o sentido (ou a falta de sentido) da vida humana, tornando-se secundário o problema da filologia (BATISTA, 2012).

A lição primordial deixada por Heidegger é a de que palavras não possuem significados fixos e unívocos, desvinculado de qualquer contexto. Desta forma, os significados acumulam-se nas palavras a partir de interrelações que constituem o mundo. Um "martelo" não é simplesmente um "utensílio para bater": o significado da palavra deriva do contexto de bancada, pregos, madeira, oficina e consumidores que constituem o "mundo" do artesão. Assim, o significado de uma palavra depende do mundo em que seu usuário vive, podendo existir ambiguidade relacionada às coisas, aos outros e ao ser de *Dasein* (INWOOD, 1998; SCHMIDT, 2014). Isso significa que para compreender um texto é necessário ir além dos dicionários e das regras gramaticais, de forma a reconstruir o mundo do autor e as possibilidades por ele oferecidas.

Apesar das contribuições apresentadas em sua obra *Ser e tempo*, Heidegger considerou que estava próximo demais da linguagem da metafísica e, a partir disto, repensou sua posição, recuando a uma situação mais original a partir da qual o pensamento tem que começar. Neste ponto, o termo "hermenêutica" foi deixado de lado e não foi mais utilizado em seus últimos escritos.

A hermenêutica de Heidegger foi explorada por inúmeros estudiosos, destacando-se principalmente Hans-Georg Gadamer (1900-2002), que foi seu aluno e um dos

seus seguidos mais próximos. Sua principal obra foi *Verdade e método: traços fundamentais de uma hermenêutica filosófica*, publicada em 1960, que teve como objetivo principal oferecer uma justificação filosófica para "a experiência da verdade que transcende o domínio do método científico" (GADAMER, 1999, p. xxii).

> Hans-Georg Gadamer (1900-2002) é o principal responsável por nosso pensamento sobre a hermenêutica na filosofia contemporânea hoje por três razões. Primeiro, Heidegger abandonou o termo em seus últimos escritos. Segundo, Sartre, que defendeu o existencialismo, que foi parcialmente desenvolvido a partir de S*er e tempo* de Heidegger, não incorporou a hermenêutica. Terceiro, Gadamer, um aluno de Heidegger, desenvolve especificamente a análise da compreensão de *Ser e tempo* em sua obra principal, *Verdade e método*, chamando sua teoria de hermenêutica filosófica (SCHMIDT, 2014, p.21).

A obra *Verdade e método* foi dividida em três partes, que tratam da experiência da verdade na arte, a experiência da verdade na compreensão das ciências humanas e a fundamentação ontológica da hermenêutica na linguagem. Em função das características deste trabalho, o foco de análise dos estudos de Gadamer se concentra na segunda parte do livro, que se refere às estruturas fundamentais da experiência hermenêutica e sua base na linguagem.

A partir de um entendimento diferente de Schleiermacher e Dilthey, que defendiam a hermenêutica orientada para os métodos e para a metodologia, Gadamer não se preocupa diretamente com os problemas práticos da formulação de princípios interpretativos corretos, mas antes pretende esclarecer o próprio fenômeno da compreensão. Isto não significa que ele negue a existência de tais princípios, que são necessários às disciplinas interpretativas. Seu objetivo maior é a indagação sobre uma questão preliminar e fundamental: como é possível a compreensão, não só nas humanidades, mas em toda a experiência humana sobre o mundo? Nesta perspectiva, Gadamer conclui que a verdade não pode ser alcançada metodicamente, mas dialeticamente, uma vez que o método seria incapaz de revelar uma nova verdade, apenas explicitando a verdade já implícita no próprio método (ALCANTARA; PAIVA; BRITO, 2018).

Assim como Heidegger, Gadamer entende a hermenêutica como algo inerente à existência humana, uma vez que a interpretação é uma atividade desenvolvida constantemente, de uma forma ou de outra, realizada pelas pessoas acerca das coisas, de outras pessoas ou delas mesmo. Com isso, a linguagem se torna o horizonte da experiência hermenêutica e, em função do seu caráter especulativo, tem a habilidade de convencer os

participantes do diálogo de sua verdade. Assim, a hermenêutica deixa de ser considerada um método das ciências e se torna uma questão de ontologia, uma vez que é "capaz de articular o sentido e a compreensão da verdade na perspectiva filosófica que ultrapasse o campo do controle da metodologia científica". A compreensão hermenêutica pode, portanto, revelar e garantir verdades que o método científico não consegue (BRITO et al. 2007, p.6).

Do ponto de vista epistemológico, a virada ontológica na hermenêutica filosófica se deu a partir da ruptura da ideia clássica de que o intérprete determina o significado das coisas ou produz uma compreensão meramente reprodutiva, sem se envolver ou se deixar afetar pelo ato de interpretação, como se tudo fosse observado de fora. Para Schwandt (2006, p.198), "a hermenêutica filosófica sustenta que a compreensão não é, em primeiro lugar, uma tarefa controlada por procedimentos ou por regras, mas sim, uma condição do ser humano", o que significa que, sendo uma estrutura básica da experiência do ser humano no mundo, não pode ser uma atividade isolada, uma vez que a compreensão é participativa e dialógica. Gadamer afirma que o significado de uma ação ou de um texto deve ser sempre negociado e consensuado, não podendo ser simplesmente descoberto pelo pesquisador, ou seja, o objetivo da interpretação não é fazer uma ponte temporal e reconstruir a situação original do texto ou aquilo que o autor queria dizer com ele, mas descobrir o que o texto tem a dizer para o intérprete (INWOOD, 1998; BATISTA, 2012).

Nessa perspectiva, a compreensão de um texto, de acordo com Gadamer, envolve os seguintes conceitos: a) a inexistência de uma verdade absoluta; b) uma pré-compreensão (ou preconceitos) historicamente determinada; c) a tradição, que é fonte de significados, crenças e preconceitos; d) a autoridade, que se relaciona com o conhecimento e não com a obediência cega; e) uma fusão de horizontes, relacionados ao passado do texto e ao presente daquele que compreende; f) a aplicação, que está relacionada ao que o texto diz no horizonte do intérprete, agora expandido. Em função desta multiplicidade de elementos interferentes, não é possível ter certeza de que uma interpretação é correta ou melhor que outras interpretações anteriores, pois todas estão sujeitas a revisões futuras (SCHMIDT, 2014; ALCANTARA; PAIVA; BRITO, 2018).

> A interpretação se dá a partir de círculos distintos, sendo que a verdade nada mais é do que a fusão de horizontes (ou de círculos), algo consensual, acordado intersubjetivamente. Entendimento é um processo dialógico entre o passado e o presente no qual eu exponho minhas pressuposições ao outro, sabendo que essas pressuposições são elas mesmas afetadas durante esse encontro, ou seja, estão suscetíveis a revisões futuras (BATISTA, 2012, p.106-107).

A compreensão surge, portanto, como algo produzido no diálogo, e não como algo meramente reproduzido por um intérprete ao se deparar com um texto na intenção de compreendê-lo. Este processo pode ser comparado à elaboração prévia de um projeto, que será constantemente revisado para que se tenha avanço na penetração do sentido. O sentido somente se manifesta "porque quem lê o texto lê a partir de determinadas expectativas e na perspectiva de um sentido determinado" (GADAMER, 1999, p.272).

De acordo com Schmidt (2014), qualquer processo de compreensão parte de uma pré-compreensão (ou preconceitos) do intérprete, que são herdados do seu passado no processo de aculturação. Gadamer emprega o termo "preconceitos" (*vorurteile*) para designar coletivamente as estruturas prévias da compreensão definidas por Heidegger. Ao contrário do que defendiam Schleiermacher e Dilthey, que uma época histórica não deveria ser julgada em termos de outra, Gadamer aponta que não é possível abandonar o presente para entender o passado, de modo que o significado de uma obra não pode ser visto unicamente sob seus próprios termos. Pelo contrário, o significado da obra passada define-se em termos das questões que se lhe colocam a partir do presente. Desta forma, quanto maior for a distância temporal entre intérprete e texto, é mais provável que ela leve a preconceitos diferentes. Isto é importante, uma vez que só é possível questionar preconceitos através do confronto com outros preconceitos diferentes, além de eliminar erros e abrir novas possibilidades de significado.

Gadamer (1999) entende que os preconceitos são estruturas prévias da compreensão herdadas, que incluem tudo o intérprete sabe, de forma consciente ou inconscientemente, compreendendo o significado das palavras, as preferências, os fatos aceitos, seus valores e juízos, os valores sobre a natureza humana e divina, entre outros. Desta forma, o significado do termo não está relacionado, em momento algum, à expressão pré-conceito ou pré-juízo num sentido pejorativo, ou seja, possui uma conotação neutra intencionada, não sendo nem positivo e nem negativo. De acordo com Bonfim (2010, p.77), a proposta do processo de compreensão é "manter um constante interpretar até que os conceitos prévios deixem de sê-los, e ao longo da comunicação, sejam substituídos por outros conceitos novos mais adequados".

Para Gadamer, tem que haver, antes de tudo, a apropriação de opiniões prévias e preconceitos pessoais, pois é fundamental dar-se conta dos próprios pressupostos, a fim de que o próprio texto possa se apresentar em sua alteridade, confrontando assim "a verdade do texto com as opiniões prévias pessoais do intérprete" (GADAMER, 2005, p.358). A compreensão não deve, portanto, ser entendida como um ato subjetivo e sim como um

processo de transmissão em que o passado e o presente são mediados constantemente, ou seja, “compreender é a possibilidade de interpretar a partir da particularidade dos mundos”, que representa todas as experiências vivenciadas pelo intérprete, que contribuíram para a formação do seu conhecimento (BONFIM, 2010, p.78).

> A compreensão implica sempre uma pré-compreensão que, por sua vez, é prefigurada por uma tradição determinada em que vive o intérprete e que modela os seus preconceitos. Assim, todo encontro significa a ‘suspensão’ de meus preconceitos, seja o encontro com uma pessoa com quem aprendo a minha natureza e os meus limites, seja com uma obra de arte ou com um texto; e é impossível contentar-se em ‘compreender o outro’, buscar e reconhecer a coerência imanente aos significados-exigências do outro (GADAMER, 2006, p.13-14).

Schmidt (2014) afirma que os preconceitos podem ser legítimos, quando são baseados nas coisas em si, ou ilegítimos, quando são baseados em ideias ao acaso e concepções populares. Neste sentido, Gadamer (1999) discute que a base para a legitimidade de preconceitos está no reconhecimento da autoridade, que é um ato de razão e se dá a partir do momento em que uma pessoa é considerada como sendo superior em juízo e percepção e, portanto, seu entendimento tem precedência. Um exemplo sobre o reconhecimento da autoridade está na previsão do tempo, uma vez existe um entendimento de que os meteorologistas possuem mais informação e maior conhecimento para a realização de tal tarefa. Da mesma forma, um estudante aceita a autoridade do professor de transmitir informações sobre um determinado assunto, uma vez que reconhece sua capacidade para isto. Por outro lado, é possível perder a autoridade quando existe uma base razoável para suspeitar dela. Assim seria racional, por exemplo, questionar um boletim científico apoiando um determinado medicamento que tenha sido escrito por alguém financiado indiretamente pela companhia farmacêutica.

Além do reconhecimento da autoridade, existe também a tradição, por meio da qual juízos e costumes são repassados de uma geração a outra. Utilizar a tradição para alcançar a compreensão é de fundamental importância, “pois não há uma relação de apropriação pelo intérprete do texto, mas um diálogo, seguido de perguntas e respostas, no qual o intérprete participa, ‘ouvindo’ o que o texto tem a dizer” (BONFIM, 2010, p.80). Gadamer (1999, p.281) afirma que uma tradição “precisa ser afirmada, acolhida, cultivada. Ela é, essencialmente, preservação”. O fato dela ter sobrevivido numa tradição implica que ela provavelmente foi considerada digna por aqueles que a acolheram. No entanto, apesar disto, nem tudo numa tradição é verdadeiro, mas ela é uma possível fonte de preconceitos

legítimos. Desta forma, um texto é considerado clássico pelo fato de demostrar tradição e conter, repetidamente, preconceitos verdadeiros e legítimos, o que não significa, apesar disso, que o futuro fará este mesmo juízo.

> A tarefa da compreensão hermenêutica é diferenciar os preconceitos legítimos de todos os ilegítimos que precisam ser criticados e abandonados. A compreensão precisa começar de nossos preconceitos herdados, e trata apenas de alguns poucos deles, enquanto o resto constitui o pano de fundo não questionado da compreensão (SCHMIDT, 2014, p.149).

Para Gadamer (1999, p.271-272), aquele que quer compreender, fazer uma interpretação correta, adotar um comportamento reflexivo diante da tradição, tem que "proteger-se contra a arbitrariedade da ocorrência de *felizes ideias* e contra a limitação dos hábitos imperceptíveis do pensar, e orientar sua vista às *coisas elas mesmas*", ou seja, aquele que pretende compreender não pode entregar-se antecipadamente ao arbítrio de suas próprias opiniões prévias, mas deve estar disposto a deixar que estas lhe digam algo.

Outro ponto importante da teoria gadameriana é a insustentabilidade de um conhecimento universalmente válido, de uma verdade absoluta que poderia ser generalizada a partir de uma experiência particular e histórica, uma vez que inexiste qualquer método científico que garanta uma verdade hermenêutica, o que se deve, em parte, pelo fato da compreensão estar sempre condicionada por preconceitos. Para Gadamer, a finitude do conhecimento não gera nenhum obstáculo ao próprio conhecimento, pelo contrário (BONFIM, 2010).

No entendimento de Gadamer, toda compreensão é interpretativa, pois acontece dentro do círculo hermenêutico, no qual o intérprete se move de um significado projetado do todo para as partes, e então volta para o todo. "A harmonia de todos os detalhes com o todo é o critério da compreensão correta. Este processo de compreensão é "a interação entre o movimento da tradição e o movimento do intérprete" (GADAMER, 1999, p.293).

> Um exemplo de movimento de tradição são as formas diferentes pelas quais a *República* de Platão teria algo esclarecedor a dizer no decorrer de sua preservação na tradição. O movimento do intérprete inclui não apenas a leitura do texto original, mas também um exame dessa leitura à luz de outras interpretações de Platão com o objetivo de estabelecer uma unidade de significado para o texto (SCHMIDT, 2014, p.150).

Gadamer (1999, p.306) entende o termo "horizonte" como um conjunto herdado de preconceitos, que "constituem, então, o horizonte de um presente particular, pois

eles representam aquilo além do que é impossível ver". Isso significa que o horizonte de um intérprete pode mudar através da adoção de outros preconceitos, pode se expandir a partir da inclusão de mais preconceitos, ou pode diminuir através da exclusão de alguns preconceitos.

Ao tentar compreender um texto, é importante que o intérprete busque um acordo com o autor e não apenas adote a posição dele, ignorando sua posição e seus preconceitos. Isso não é nem desejável, nem possível. Desta forma, o intérprete projeta o horizonte do texto dentro do seu próprio horizonte (SCHMIDT, 2014). Assim, Gadamer (1999, p.306) afirma que "a compreensão sempre é a fusão destes horizontes que supostamente existem por si mesmos".

A compreensão interpretativa envolve, necessariamente, a aplicação na projeção do horizonte do texto. Isto não significa que o interprete primeiro compreende o texto e depois o aplique à sua situação e ao seu contexto. Em vez disto, a aplicação é parte do processo de compreensão do que o texto tem a dizer, ou seja, realiza o significado do texto para a situação concreta do intérprete. Para demonstrar isto, Gadamer usa o exemplo de um historiador legal e afirma que sua compreensão de uma lei não pode ser limitada a seu uso inicial, mas deve incluir como ela foi interpretada desde então porque os casos precedentes são considerados partes do significado total da lei (GADAMER, 1999; SCHMIDT, 2014).

Tal como Heidegger, Gadamer também reconhece a importância da linguagem para que o diálogo entre o texto e o intérprete seja possível. Assim, compreender o que o texto tem a dizer sempre envolve interpretação e aplicação, o que significa que é necessário traduzi-lo para a linguagem do intérprete. Quando texto e intérprete falam linguagens diferentes, é necessário que um tradutor traduza o contexto linguístico para que a comunicação entre ambos seja possível. Desta forma, quanto mais perto estiver a linguagem do texto da linguagem do intérprete, menos perceptíveis e mais automáticas serão as tarefas de interpretação. A linguagem não é, portanto, um instrumento, mas sim um campo de interação (BATISTA, 2012; SCHMIDT, 2014).

Assim como não há uma linguagem perfeita que apresenta o mundo em si, também não há uma interpretação perfeita que apresenta a coisa em si, sendo que ambos os casos indicam a finitude fundamental da compreensão humana. A mesma palavra pode ser traduzida de formas diferentes, representando, portanto, um aspecto ou uma visão de um mundo particular. Isto significa que podem existir interpretações diferentes de um texto, mas todas corretas, advindas de horizontes históricos diferentes, que revelam uma multiplicidade de aspectos relacionados ao assunto expresso no texto (SCHMIDT, 2014).

Além das contribuições de Schleiermacher, Dilthey, Heidegger e Gadamer

discutidas no âmbito deste trabalho, o campo da hermenêutica possui outros grandes estudiosos que possuem trabalhos significativos na área, tais como: Weber (1864-1920), E.D. Hirsch Jr. (1928-), Jüngen Habermas (1929-), Paul Ricoeur (1913-2005) e Jacques Derrida (1930-2004).

A partir do século XX passou a existir uma nítida divergência entre duas tradições hermenêuticas: a primeira, originada em Schleiermacher e Dilthey (século XIX) constitui-se como um corpo geral de princípios metodológicos, pautado pela busca da objetividade e da validação das interpretações construídas. A segunda, por sua vez, enfatiza o caráter histórico da compreensão e, consequentemente, as limitações de toda pretensão a um conhecimento objetivo, como é o caso do entendimento de Heidegger e de Gadamer (SCHMIDT, 2014).

Apesar de terem opiniões divergentes, as ideias dos diversos autores da área ampliaram os horizontes da hermenêutica e, consequentemente, dos princípios relacionados ao processo de compreensão e interpretação documental que serão discutidos a seguir.

5.1 PRINCÍPIOS DA HERMENÊUTICA

O campo da hermenêutica é vasto e em função das características e dos diferentes entendimentos dos estudiosos da área, apresenta, muitas vezes, conceitos divergentes entre si relacionados ao processo de compreensão e interpretação de documentos. Nesse contexto, Klein e Myers (1999) elaboraram uma proposta metodológica para o desenvolvimento e a avaliação de pesquisa interpretativa no campo dos sistemas de informação (SI), com base nos elementos da hermenêutica filosófica, da antropologia e da fenomenologia considerados por eles mais estáveis.

Para Klein e Myers (1999), a aplicabilidade destes princípios pode ser estendida a qualquer pesquisa de natureza interpretativa, incluindo estudos de caso que são tipicamente mais curtos e menos detalhados, desde que haja uma precaução em relação ao seu uso de forma burocrática, sendo necessário, portanto, pensamento criativo para sua aplicação em um determinado contexto de investigação, ressalvando a arbitrariedade e a interdependência entre eles. Assim, é possível que outros autores elaborem conjuntos adicionais de princípios, uma vez que existem inúmeras teorias relacionadas ao interpretativismo.

Klein e Myers (1999) entendem que a pesquisa interpretativa não reconhece a ideia de aplicação mecanicista de um conjunto pré-determinado de critérios, o que não

significa, no entanto, que não existam padrões que possam ser adotados. A adoção de princípios para a elaboração de trabalhos interpretativos minimiza os problemas relacionados à subjetividade inerente ao processo, facilita a aplicação e o entendimento dos conceitos envolvidos e favoreça a utilização em diversas áreas de conhecimento.

Os princípios propostos por Klein e Myers (1999) são os seguintes: Princípio do Círculo Hermenêutico, da Contextualização, da Interação entre Pesquisadores e Sujeitos, da Abstração e da Generalização, do Raciocínio Dialético, das Múltiplas Interpretações e da Suspeição. Além destes, a partir dos estudos realizados no âmbito desta tese, foi proposto um oitavo princípio: do Compartilhamento de Linguagem.

5.1.1 Princípio do Círculo Hermenêutico

O primeiro e mais fundamental de todos os princípios é o círculo hermenêutico, sendo a base para todo trabalho interpretativo de natureza hermenêutica. Em função da sua importância, ele é discutido por diversos estudiosos, incluindo Schleiermacher, Heidegger e Gadamer.

Para Schleiermacher (2005), não é possível compreender o todo até que as partes tenham sido compreendidas, existindo, portanto, uma interdependência entre a parte e o todo. Isto quer dizer que se pode compreender o significado de um texto a partir das palavras, ou seja, é possível entender o todo através de sinais individuais determinados apenas relativamente (palavras). Heidegger, por sua vez, reformulou a concepção do círculo hermenêutico e apresentou um novo conceito completamente diferente dos demais, referindo-se a ele como a interação entre o autoentedimento do intérprete e seu entendimento do mundo (BATISTA, 2012). A partir de outra perspectiva, Gadamer (1999) entende que toda compreensão é interpretativa, pois acontece dentro do círculo hermenêutico, no qual o intérprete se move de um significado projetado do todo para as partes, e então volta para o todo. Desta forma, só é possível entender um todo complexo a partir de preconcepções sobre os significados das suas partes e dos seus interrelacionamentos.

Considerando o entendimento de Gadamer, Klein e Myers (1999) exemplificam o círculo hermenêutico, afirmando que a interpretação de uma frase pode começar a partir das suas partes, que são as palavras. Neste contexto, pode ser necessária interpretações relacionadas a múltiplas dimensões, tais como a perspectiva histórica ou de base temporal, que interpreta a parte como um componente que deverá ser entendido no âmbito do seu contexto histórico (princípio da contextualização). Esses elementos podem

ainda ser as preconcepções ou entendimentos dos pesquisadores e demais participantes de um estudo ou de uma área específica. O todo consistiria nos significados compartilhados que emergem das interações entre eles, devido à apropriação e tradução mútua, ao longo de sucessivas interações através dos círculos hermenêuticos. Esta interpretação deve ser utilizada na aplicação do princípio da interação entre pesquisadores e sujeitos. Desta forma, durante a repetição dos ciclos hermenêuticos, todos os princípios sugeridos podem ser aplicados iterativamente, formando uma rede complexa de interpretações.

5.1.2 Princípio da Contextualização

O princípio da contextualização se fundamenta nos estudos de Schleiermacher e Gadamer e está baseado na diferença de entendimento entre o autor de um texto e o intérprete, resultante da distância temporal ou cultural entre eles. Tem a preocupação de entender o posicionamento do todo no seu contexto histórico.

Schleiermacher se refere ao princípio da contextualização quando aborda a interpretação técnica ou psicológica, afirmando ser necessário compreender como o autor pensava e de como estes pensamentos foram influenciados pela cultura e o pelo contexto histórico vivenciado por ele.

Para Gadamer, a tarefa da hermenêutica não seria a de disfarçar a distância de entendimento entre autor e intérprete, mas de conscientemente a explicitar, de forma a revelar os significados nos diversos contextos, bem como a base constituinte ao qual se referem, indicando como a situação atual emergiu e se desenvolveu ao longo da história (SCHMIDT, 2014). Desta forma, o entendimento é um processo dialógico entre o passado e o presente, no qual texto e intérprete expõem suas pressuposições ao outro, sabendo que elas serão afetadas neste encontro e estarão sujeitas a revisões futuras (BATISTA, 2012).

O princípio da contextualização exige, portanto, que o tema seja posicionado no seu contexto social e histórico, de forma a permitir que leitores e estudiosos interessados em compreender o assunto, possam julgar sobre as razões pelas quais a situação em interpretação terá emergido, bem como da emergência dessa interpretação. Klein e Myers (1999) afirmam que a investigação interpretativa procura entender um algo em movimento, na medida em que cada instância estudada é tratada como uma ocorrência histórica única. Desta forma, alertam para o papel que a própria intervenção de investigação pode ter na formação da história futura, sendo necessário, portanto, que o princípio da contextualização seja explicitamente considerado na descrição destes estudos. Os sujeitos tornam-se produtores de

conhecimento e não apenas produtos da história.

Gadamer (1999) exemplifica o princípio da contextualização quando afirma que a compreensão de um historiador legal referente a uma determinada lei não pode ser limitada apenas ao seu uso inicial, mas deve incluir como ela foi interpretada desde então, uma vez que os casos precedentes são considerados partes do significado total da lei. Klein e Myers (1999), em outra perspectiva, exemplificam este princípio através da Fábrica de Montagem da Fiat, que em 1995 passou por mudanças radicais na organização e nas operações de trabalho. No entanto, os antigos conceitos de produção fordista não foram descartados e continuaram influenciando os novos paradigmas de produção.

5.1.3 Princípio da Interação entre Pesquisadores e Sujeitos

Com base no entendimento de Gadamer (1999), o princípio da interação propõe que o investigador se posicione, juntamente com os demais sujeitos, numa perspectiva histórica e, a partir da interação entre eles, produza dados e conhecimento. Desta forma, o significado de um texto não pode ser visto unicamente a partir das perspectivas e do entendimento de um único intérprete, mas, ao contrário, deve ser sempre negociado e consensuado. Assim, a compreensão surge como algo produzido no diálogo entre os participantes, através de um processo de construção de conhecimento, portanto, uma fusão de horizontes, algo consensual, acordado intersubjetivamente.

Klein e Myers (1999) afirmam que os pesquisadores devem reconhecer que os participantes exercem uma capacidade interativa autônoma, podendo ser reconhecidos como intérpretes ou analistas. Os participantes são considerados intérpretes quando alteram seus horizontes a partir da apropriação de preconceitos usados por pesquisadores com os quais está interagindo e são analistas, na medida em que suas ações são modificadas por seus horizontes alterados. Neste processo, o pesquisador pode ser confrontado com as suas próprias preconcepções, face a outras que são por ele desconhecidas, tornando-se mais autoconsciente e possivelmente mais capaz de as questionar.

5.1.4 Princípio da Abstração e da Generalização

O princípio da abstração e da generalização está baseado em Heidegger e refere-se à possibilidade de produção de resultados de caráter abstrato e genérico, embora os princípios anteriores enfatizem a particularidade ou a singularidade da situação nos estudos

interpretativos. Assim, é comum a tentativa de relacionar as informações que podem ser descritas no princípio da contextualização a categorias muito abstratas, uma vez que instâncias únicas podem estar relacionadas a ideias e conceitos que se aplicam a múltiplas situações. No entanto, é importante que as abstrações teóricas e generalizações sejam cuidadosamente relacionadas aos estudos de campo realizados, bem como na coerência dos raciocínios lógicos que lhe deram origem, de tal forma que o leitor consiga compreender como o pesquisador chegou às suas percepções teóricas (KLEIN; MYERS, 1999).

Em relação ao problema da generalização, Walsham (1993) argumenta que a validade das inferências extraídas de um ou mais casos não depende da representatividade estatística dos casos, mas antes na plausibilidade e coerência do raciocínio lógico usado para a descrição dos resultados dos casos e para retirar as conclusões. Deste modo, existem quatro tipos de generalizações: o desenvolvimento de conceitos, a geração de teoria, a retirada de implicações específicas e a contribuição de perspectivas “ricas”.

Klein e Myers (1999) exemplificam a utilização deste princípio na área dos sistemas de informação através do estudo de Monteiro e Hanseth (1996), que aplica a Teoria Ator-Rede de Latour no desenvolvimento de uma infraestrutura de informação para o setor de saúde. Ao traçar relações teóricas entre o desenvolvimento tecnológico e os problemas organizacionais, eles aplicam os conceitos abordados pela Teoria de Latour em um contexto original e diferenciado.

5.1.5 Princípio do Raciocínio Dialético

O princípio do raciocínio dialético se fundamenta nas teorias de Gadamer e requer que todo pesquisador confronte as preconcepções que o conduziram na sua proposta de pesquisa original com a experiência emergente no decurso do processo de investigação. De acordo com Klein e Myers (1999), o ponto fundamental seria que o pesquisador deixasse claro a base intelectual histórica do seu trabalho, tanto para o leitor quanto para si próprio. Esta base intelectual subjacente à proposta de investigação será a lente através da qual a observação é constituída, documentada e organizada. No caso de resultados não compatíveis com esta base, o pesquisador decide reformulá-los ou abandoná-los por completo.

Para Gadamer, qualquer processo de compreensão parte dos preconceitos do intérprete, que são herdados do seu passado. Os preconceitos incluem todo o conhecimento do intérprete, de forma consciente ou inconsciente, incluindo o significado das palavras, as preferências, os fatos aceitos, seus valores e juízos, os valores sobre a natureza humana e

divina, entre outros. É de fundamental importância, portanto, dar-se conta dos próprios pressupostos, a fim de que o próprio texto possa se apresentar em sua originalidade, confrontando assim a verdade do texto com as opiniões prévias pessoais do intérprete (GADAMER, 1999, 2005; SCHMIDT, 2014).

A hermenêutica reconhece, portanto, que o preconceito é o ponto de partida para qualquer processo de compreensão. É necessário, no entanto, diferenciar os preconceitos legítimos de todos os ilegítimos que precisam ser criticados e abandonados. Esse percurso faz parte da própria historicidade do pesquisador ou intérprete, da qual ele deve ser consciente (KLEIN; MYERS, 1999; SCHMIDT, 2014). Esta regra, quando aplicada repetidamente num processo hermenêutico, permite que os entendimentos alcançados numa fase possam ser os preconceitos da seguinte.

5.1.6 Princípio das Múltiplas Interpretações

O princípio das múltiplas interpretações está fundamentado nos estudos de Gadamer e Ricouer e propõe que o pesquisador examine as influências que o contexto social tem sobre as ações em estudo, de forma a documentar múltiplos pontos de vista juntamente com as razões que os justificam. A análise das razões inclui o entendimento das relações de poder, econômicas ou de valores, bem como dos preconceitos do pesquisador, que conduziram a proposta da pesquisa, conforme anteriormente discutido no princípio do raciocínio dialético (KLEIN; MYERS, 1999).

Para Gadamer, o foco do processo de interpretação deixou de ser o que o autor queria dizer com o texto e passou a ser aquilo que o texto tinha a dizer para o leitor. Isto significa que o objetivo da interpretação não é fazer uma ponte temporal e reconstruir a situação original do texto ou aquilo que o autor queria dizer com ele, mas descobrir o que o texto tem a dizer para o intérprete. Desta forma, haverá diferentes interpretações acerca do mesmo texto, uma vez que cada interprete o compreenderá a partir do seu mundo, ou seja, de todas as experiências vivenciadas por ele e que contribuíram para a formação do seu conhecimento, revelando assim uma multiplicidade de aspectos relacionados ao assunto expresso no texto. Não é possível, portanto, garantir que uma interpretação é correta ou melhor que outras anteriores, pois todas estão sujeitas a revisões futuras (BONFIM, 2010; BATISTA, 2012; SCHMIDT, 2014).

O pesquisador deverá, então, confrontar e entender os conflitos de interpretação nas diversas versões e, a partir delas, rever o seu entendimento referente à

situação estudada. Em comparação com o princípio do raciocínio dialético, não se trata agora de confrontar os preconceitos do pesquisador com a experiência, mas o entendimento dos diversos participantes entre si.

5.1.7 Princípio da Suspeição

O princípio da suspeição (ou da suspeita) foi adaptado das teorias de Gadamer e Ricouer e requer sensibilidade para identificar possíveis “vieses” e “distorções” sistemáticas, bem como o reconhecimento de preconceitos ilegítimos. Embora os princípios já discutidos incentivem várias formas de pensamento crítico, eles estão mais preocupados com a interpretação de significados do que com a descoberta de imprecisão e de erros no processo de compreensão (KLEIN; MYERS, 1999).

De acordo com Klein e Myers (1999), mais do que trabalhar os entendimentos, a ideia é a de revelar os efeitos socialmente criados nas distorções de percepção. Mais comum nos trabalhos de caráter crítico do que interpretativo, o objetivo deste princípio é demonstrar e criticar formas de dominação, assimetria e comunicação distorcida, expondo como as construções sociais da realidade podem favorecer interesses e não reconhecer ou esconder construções alternativas. Este tipo de abordagem propõe-nos uma leitura subliminar do mundo social por detrás dos textos produzidos pelos diversos autores.

5.1.8 Princípio do Compartilhamento de Linguagem

Além dos princípios apresentados e discutidos por Klein e Myers (1999), um oitavo princípio foi proposto para o contexto deste trabalho: o princípio do compartilhamento de linguagem. Sua fundamentação teórica está nos estudos de Schleiermacher, Dilthey, Heidegger e, principalmente, Gadamer. Está baseado na necessidade de compartilhamento de linguagem entre texto e intérprete para que a compreensão seja possível.

De acordo com Gadamer (1999), o processo de compreensão surge a partir de algo produzido no diálogo entre texto e intérprete (ou participantes). Este diálogo, no entanto, só é possível a partir da linguagem. Desta forma, compreender o que o texto tem a dizer sempre envolve interpretação e aplicação, o que significa que é necessário traduzi-lo para a linguagem do intérprete. Quando texto e intérprete falam linguagens diferentes, é necessário um processo de tradução do contexto linguístico para que a comunicação entre

ambos seja possível. Neste sentido, quanto mais perto estiver a linguagem do texto da linguagem do intérprete, menos perceptíveis e mais automáticas serão as tarefas de interpretação e aplicação (BATISTA, 2012; SCHMIDT, 2014).

Schleiermacher se refere a este princípio quando atribui caráter metódico à hermenêutica e afirma que uma das práticas envolvidas no processo de compreensão é a interpretação gramatical, que inclui o entendimento da linguagem usada pelo autor (significado das palavras), das regras gramaticais e semânticas. Segundo ele, alguns fatores influenciam a linguagem utilizada pelo autor, tais como seu lugar na história, sua educação, sua ocupação e até mesmo seu dialeto. Desta forma, o sentido de cada palavra precisa ser determinado através de sua conexão com as palavras ao seu redor, ou seja, a partir do seu contexto (SCHLEIERMACHER, 2005; SCHMIDT, 2014).

Para Dilthey, a compreensão de uma sentença só é possível depois do entendimento do significado que as palavras têm no contexto ordenado da linguagem. Heidegger afirma ainda que as palavras não possuem significados fixos e unívocos, desvinculado de qualquer contexto e que os significados se acumulam nas palavras a partir de interrelações que constituem o mundo. No entendimento de Gadamer, a mesma palavra pode ser traduzida de formas diferentes, representando, portanto, um aspecto ou uma visão de um mundo particular. Isto significa que podem existir interpretações diferentes de um texto, mas todas corretas, advindas de contextos e horizontes históricos diferentes (SCHMIDT, 2014).

Gadamer apresenta uma reflexão acerca dos problemas linguísticos que interferem no processo de compreensão, relacionados principalmente às várias formas de entendimento das diferentes comunidades linguísticas, que são influenciadas por tradições culturais diversificadas. Segundo ele, é necessário entender que uma visão de linguagem é uma visão limitada do mundo, o que explica as diferenças existentes nas linguagens particulares. Como toda linguagem é uma linguagem humana, ela está aberta a inclusão de outras perspectivas, ou seja, pode se expandir para incluir qualquer percepção possível contida em outra linguagem.

O princípio do compartilhamento de linguagem sugere, portanto, a necessidade de entendimento do vocabulário utilizado pelo autor do texto, envolvendo o significado das palavras, as regras gramaticais e semânticas e o contexto no qual as palavras estão inseridas. Também evidencia as diferenças linguísticas existentes, que são resultado de horizontes diferentes e comunidades linguísticas distintas e que podem interferir no processo de compreensão de um texto.

5.1.9 Interdependência Entre Princípios

Os princípios hermenêuticos apresentados por Klein e Myers (1999) não devem ser utilizados como receitas mecanicistas, de forma arbitrária ou independente. Na sua proposta, os autores advertem para a interdependência existente entre eles e para a necessidade de serem reinterpretados em cada contexto particular.

Os princípios do círculo hermenêutico e do compartilhamento de linguagem são os mais gerais sobre os quais os outros seis se expandem. Assim, se um pesquisador estiver decidindo sobre um contexto relevante para ser estudado, deve recorrer ao princípio da contextualização, que depende do princípio da interação entre pesquisadores e sujeitos para entender como elaborar dados em interação com os sujeitos. Na sequência, é necessário recorrer ao princípio da abstração e generalização para adotar uma teoria ou um conjunto de conceitos que darão subsídios ao estudo. O princípio do raciocínio dialético permite que o pesquisador confronte os seus preconceitos com a experiência adquirida durante o processo de investigação e, a partir do princípio das múltiplas interpretações, tem a possibilidade de compreender as diferentes versões abordadas em relação ao assunto pesquisado. E por fim, o princípio da suspeição permite o questionamento crítico da realidade apresentada (KLEIN; MYERS, 1999).

Embora Klein e Myers (1999) recomendem que os princípios discutidos não devem ser utilizados a *la carte*, eles sugerem que a escolha de quais princípios serão utilizados no contexto de uma pesquisa depende do público alvo e do nível de detalhamento adotado pelo pesquisador. Desta forma, é responsabilidade do pesquisador descobrir como e quais princípios devem ser aplicados em uma situação particular.

A proposta de adoção dos princípios hermenêuticos apresentados no contexto de um “Modelo de interface de busca e recuperação de informação” visa contribuir para o desenvolvimento de funcionalidades que auxiliem o pesquisador a compreender um documento ou um *corpus* documental no contexto de uma pesquisa científica.

5.2 RESUMO E DISCUSSÃO DO CAPÍTULO

O objetivo deste capítulo foi discutir a Hermenêutica, que é a área da Filosofia responsável pela compreensão e interpretação documental. A pesquisa realizada permitiu entender o processo de evolução da área, bem como identificar os principais precursores que ainda hoje influenciam estudos e reflexões sobre o assunto. Em função da

importância que possuem para o campo da hermenêutica, o foco da discussão foram os seguintes estudiosos: Schleiermacher (1768-1834) através da proposta da Hermenêutica Universal; Dilthey (1833-1911) que entendeu a hermenêutica como uma Metodologia para as Ciências Humanas; Heidegger (1889-1976) com a Hermenêutica da Facticidade e Gadamer (1900-2002) através da Hermenêutica Filosófica. Apesar de todos apresentarem importantes contribuições para a área, o entendimento de Hans-Georg Gadamer sobre o processo de compreensão é o que mais se aproxima das discussões sobre o comportamento de busca de um usuário no contexto da Web. Por este fato, seus estudos servirão de base para a proposta ora apresentada nesta tese.

Gadamer entende que o significado de um texto deve ser sempre negociado e que o objetivo da interpretação não é fazer uma ponte temporal e reconstruir a situação original do texto ou aquilo que o autor queria dizer com ele, mas descobrir o que o texto tem a dizer para o intérprete. Esta perspectiva vai ao encontro com as teorias de Dervin (1999), Wilson (2006) e Savolainen (2016) que discutem os inúmeros fatores que interferem na forma como os indivíduos buscam informações e constroem conhecimento, envolvendo características pessoais, sociais e ambientais. Isto significa que a elaboração de uma expressão de busca e a interpretação dos resultados obtidos referentes a um documento específico ou a um *corpus* documental estão diretamente relacionados às características individuais de cada usuário.

Assim, é possível afirmar que um texto relacionado à Filosofia, por exemplo, não será interpretado da mesma forma e com o mesmo nível de detalhamento por usuários de áreas diferentes e com níveis de entendimento distintos, ou seja, o seu significado para um filósofo, que possui maiores conhecimentos do domínio, da terminologia, das relações existentes entre os conceitos e de todos os pressupostos envolvidos sempre será diferente do significado para um médico que, teoricamente, possui uma menor compreensão da área. Um texto, portanto, pode ser interpretado segundo diversas perspectivas de análise, geradas com base na realidade, na cultura, nos conhecimentos e na experiência individual de cada usuário. Estes elementos formam o horizonte (ou conjunto de preconceitos) do intérprete, a partir do qual o texto será lido e interpretado.

A história e evolução da hermenêutica está associada a diversas concepções filosóficas e entendimentos, muitas vezes contraditórios uns dos outros. Até o século XVIII foi considerada uma arte de interpretação de textos, no século XIX passou a ser considerada uma metodologia das ciências humanas e no XX se transformou em algo completamente diferente: uma filosofia. Inicialmente tratada como um desafio de ordem técnica, configurou-

se como um problema filosófico: o problema geral da *compreensão*. A partir deste momento, a questão já não se restringiu a *o que dizem* os textos, mas a *o que significa compreender* um texto.

Com base nos principais estudiosos e nos elementos da hermenêutica considerados mais estáveis, Klein e Myers (1999) elaboraram uma proposta metodológica para o desenvolvimento e a avaliação de pesquisa interpretativa no campo dos sistemas de informação (SI), composta de sete princípios: Princípio do Círculo Hermenêutico, da Contextualização, da Interação entre Pesquisadores e Sujeitos, da Abstração e da Generalização, do Raciocínio Dialético, das Múltiplas Interpretações e da Suspeição. Além destes, foi proposto um oitavo princípio: do Compartilhamento de Linguagem. A maioria deles tem embasamento teórico na Hermenêutica Filosófica, a partir dos estudos de Gadamer sobre o processo de compreensão textual.

Este conjunto composto de oito princípios deram subsídios teóricos para o desenvolvimento de recursos e interações do "Modelo de interface de busca e recuperação de informação" proposto neste estudo, com o objetivo de minimizar as dificuldades do usuário na compreensão de um documento ou de um *corpus* documental, que podem estar relacionadas tanto ao processo de organização dos materiais quanto à interpretação dos resultados apresentados.

6 *DESIGN* DE INTERAÇÃO

Este capítulo tem como objetivo apresentar os fundamentos teóricos relacionados ao *Design* de Interação, que forneceu subsídios para o processo de desenvolvimento da interface proposta neste estudo, bem como ofereceu as principais ferramentas adotadas para a identificação dos requisitos, *design* das alternativas, prototipação e avaliação da usabilidade do modelo projetado.

■■■

O *Design* de Interação emergiu em meados de 1990, a partir da popularização da Internet, do crescimento da computação pessoal (doméstica e de lazer) e, consequentemente, do consumo cada vez maior de produtos digitais interativos. Na virada do século, a noção de *design* de interação começou a ganhar popularidade como abordagem para desenvolvimento de produtos digitais, que sejam fáceis de serem utilizados, com interações eficientes e livres de erros para suportar tarefas de trabalho e proporcionar uma experiência de usabilidade agradável ao usuário (LOWGREN, 2013; ROGERS; SHARP; PREECE, 2013).

Moggridge (2007, p.14), *designer* industrial e fundador da IDEO, que é uma empresa internacional de *design* e consultoria em inovação, relembra as primeiras ideias sobre o *design* de interação:

> Senti que havia uma oportunidade de criar uma nova disciplina de *design*, dedicada a criar soluções imaginativas e atraentes em um mundo virtual, onde era possível projetar comportamentos, animações, sons e formas. Isso seria o equivalente do design industrial, mas em software, em vez de objetos tridimensionais. Como o *design* industrial, a disciplina começaria a partir das necessidades e desejos das pessoas que usam um produto ou serviço, e se esforçaria para criar projetos que dariam prazer estético, além de satisfação e prazer duradouros (tradução nossa).

O *design* de interação tem como objetivo "projetar produtos interativos para apoiar o modo como as pessoas se comunicam e interagem em seus cotidianos, seja em casa ou no trabalho" ou, em outras palavras, "criar experiências de usuário que melhorem e ampliem a maneira como as pessoas trabalham, se comunicam ou interagem" (ROGERS; SHARP; PREECE, 2013, p.8). Para Winograd (1997, p.160), o termo pode ser definido como "o projeto de espaços para comunicação e interação humana". Saffer (2010, p.4), por sua vez, enfatiza os aspectos artísticos do *design* de interação, descrevendo-o como "a arte de facilitar

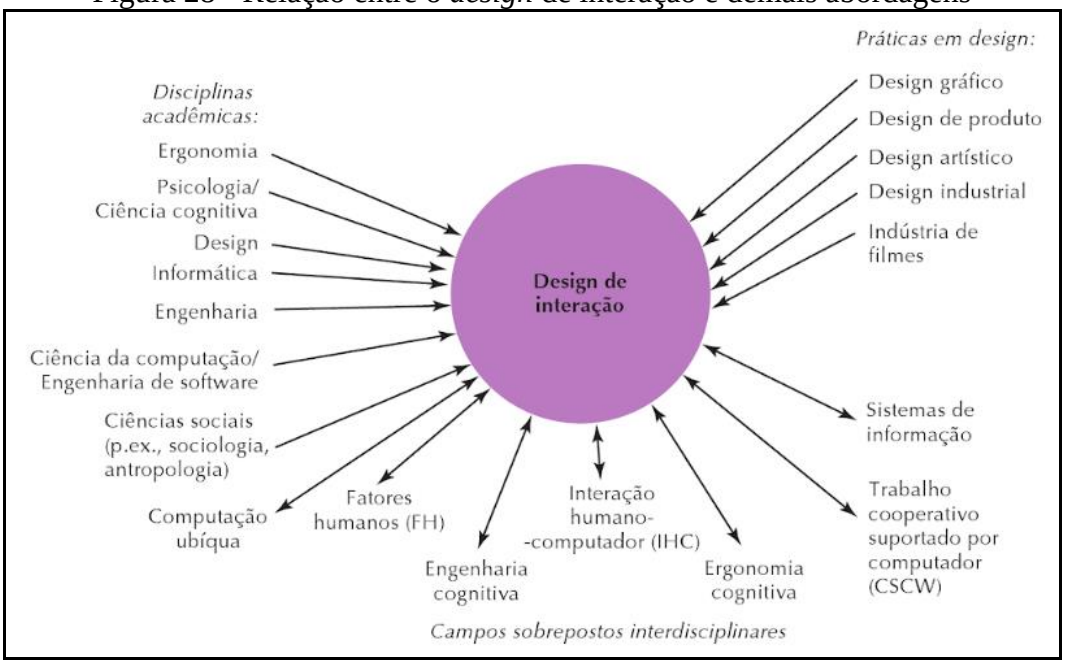
as interações entre os seres humanos por meio de produtos e serviços".

O principal foco de interesse do *design* de interação é o usuário, sendo este o elemento de maior relevância no desenvolvimento de todas as atividades projetadas, buscando-se a compreensão de suas necessidades, restrições, habilidades, aspirações e expectativas. A partir desta perspectiva, o *design* de interação visa reduzir os aspectos negativos da experiência do usuário (frustração, aborrecimento) e ao mesmo tempo, melhorar os positivos (divertimento, motivação, satisfação, interesse), de forma a aumentar o conforto e a eficácia de resultados.

Uma variedade de termos tem sido utilizada na literatura para enfatizar diferentes aspectos do que está sendo projetado: *design* de interface do usuário, projeto de *software*, *design* centrado no usuário, *design* do produto, web *design*, *design* de experiência e *design* de sistemas interativos. Neste contexto, o termo *design* de interação está sendo cada vez mais aceito com um "guarda-chuva", uma vez que abrange todos esses aspectos (ROGERS; SHARP; PREECE, 2013).

O *design* de interação é uma abordagem ampla, que se relaciona com inúmeras disciplinas e diversos campos de atuação, conforme pode ser observado na Figura 28. Para Rogers, Sharp e Preece (2013), a diferença entre *design* de interação e as demais abordagens deve-se, em grande parte, aos métodos, filosofias e pontos de vista adotados para estudar, analisar e projetar soluções para computadores.

Figura 28 - Relação entre o *design* de interação e demais abordagens



Fonte: Rogers, Sharp e Preece (2013, p.10)

Entre as abordagens relacionadas com o *design* de interação, a Engenharia de *Software* e a Interação Humano-Computador (IHC) são especialmente importantes em função da relevância que possuem na área de desenvolvimento de sistemas.

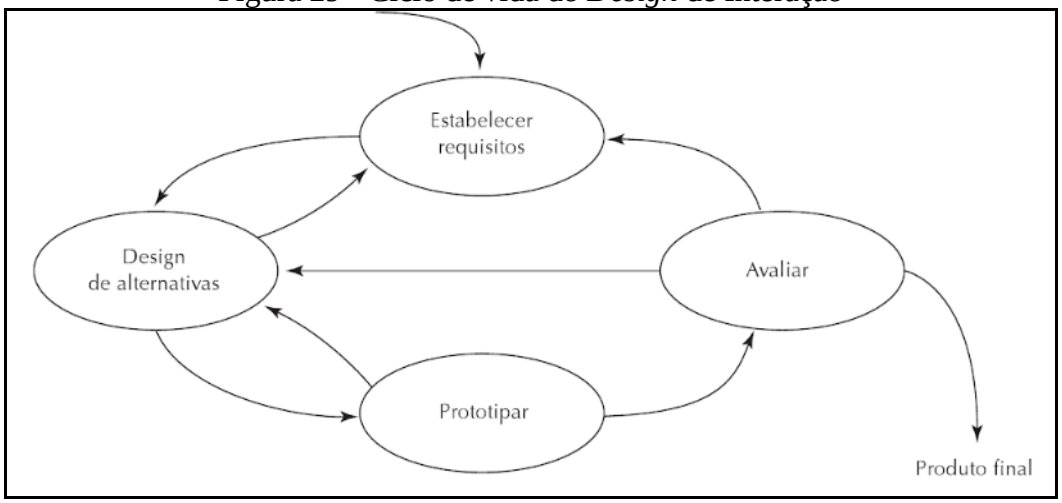
Winograd (1997) explica a diferença entre *design* de interação e engenharia de *software* por meio de uma analogia referente à atuação de arquitetos e engenheiros na construção de uma casa. Neste sentido, segundo o autor, os arquitetos estão preocupados com as pessoas, suas interações e os espaços relacionados com o interior da residência, tais como: proporção certa entre áreas privativas e sociais, proximidade entre as áreas de cozinhar e fazer as refeições e forma como os espaços projetados serão utilizados pelos moradores. Já os engenheiros estão interessados em questões relacionadas com a realização do projeto, ou seja, os aspectos estruturais, ambientes e práticos, mais especificamente, com o custo e a durabilidade, regulamentações contra incêndio e métodos de construção. Assim como existem diferenças entre projetar e construir uma casa, há também uma distinção entre o *design* de um produto interativo e a engenharia de um *software* para isso.

Em relação à IHC, por sua vez, a principal diferença em relação ao d*esign* de interação está no escopo, ou seja, enquanto o DI possui uma visão mais ampla, abordando a teoria, a pesquisa e a prática no *design* de experiências de usuário para todos os tipos de tecnologias, sistemas e produtos, a IHC tem um foco mais restrito, direcionado apenas para o *design*, a avaliação e a implementação de sistemas de computação interativos (ROGERS; SHARP; PREECE, 2013, p.8).

Um projeto de *design* de interação pode ser adotado com o objetivo de substituir ou atualizar um sistema já existente ou desenvolver um projeto totalmente inédito. Seja qual for a situação, é necessário que haja um entendimento amplo acerta do usuário, envolvendo suas atividades e metas atuais, bem como as condições sob as quais o produto será utilizado e as restrições relacionadas ao seu desempenho. De acordo com Rogers, Sharp e Preece (2013), isto representa uma infinidade de análises, escolhas e decisões que deverão ser tomadas durante o ciclo de vida de um produto, procurando compreender claramente alguns pontos importantes: a) quais são as principais habilidades dos usuários (no que eles são bons ou não); b) o que pode auxiliar as pessoas na sua atual maneira de fazer as coisas; c) quais elementos podem proporcionar experiências de usuário com qualidade; d) o que as pessoas querem e como envolvê-las no *design*.

O ciclo de vida básico recomendado pelo *design* de interação é composto de quatro etapas relacionadas, todas centradas no usuário, que representam um conjunto de atividades que devem ser desenvolvidas durante o projeto, conforme mostra a Figura 29.

Figura 29 - Ciclo de vida do *Design* de Interação



Fonte: Rogers, Sharp e Preece (2013, p.332)

A primeira etapa do ciclo de vida consiste em **estabelecer os requisitos** que darão suporte ao projeto e às atividades posteriores. Para isso, é necessário analisar o usuário, levantar dados sobre suas necessidades e oportunidades de melhorias, com o intuito de determinar as características do produto de *design* e identificar informações importantes: quais funcionalidades devem ser incluídas, que tecnologias devem ser utilizadas, que fatores devem ser privilegiados e quais tarefas devem ser apoiadas (BARBOSA; SILVA, 2010).

De acordo com Waslawick (2011), os requisitos podem ser funcionais ou não funcionais. Os requisitos funcionais descrevem as funcionalidades e o comportamento do de um produto ou sistema, enquanto que os não funcionais referem-se às qualidades globais de um *software*, como manutenibilidade, usabilidade, desempenho, custos e várias outras.

Existem diversas técnicas utilizadas para coletar dados, que é parte fundamental do estabelecimento de requisitos e do processo de avaliação. Em relação à atividade de requisitos, o objetivo da coleta de dados é "obter dados suficientes, precisos e relevantes para que um conjunto de requisitos possa ser produzido". No âmbito da avaliação, a coleta de dados é importante para "captar as reações dos usuários e seu desempenho com um sistema ou protótipo" (ROGERS; SHARP; PREECE, 2013, p.222). Dentre as técnicas utilizadas frequentemente, destacam-se: entrevistas, questionários, observação e grupos de foco.

A entrevista é uma das técnicas mais utilizadas de coleta de dados e levantamento de requisitos e requer a participação de um entrevistador perguntando a um ou mais entrevistados um conjunto de questões que podem ser estruturadas ou não estruturadas. As entrevistas estruturadas são elaboradas a partir de um roteiro previamente elaborado, que é seguido pelo entrevistador, não permitindo que novas questões ou tópicos sejam explorados

durante a atividade. Em contrapartida, em uma entrevista não estruturada, o entrevistador tem a possibilidade de realizar perguntas de modo flexível, o que possibilita um maior aprofundamento dos tópicos importantes. Além desses dois tipos, as entrevistas podem ser conduzidas ainda de forma semiestruturada, ou seja, a partir de um roteiro contendo tópicos ou perguntas, mas com liberdade do pesquisador para explorar em maior profundidade as respostas fornecidas pelo entrevistado, desde o foco nos objetivos da entrevista seja mantido (BARBOSA; SILVA, 2010).

Os questionários também são frequentemente utilizados para a coleta de dados. Sua principal característica é a presença de um formulário impresso ou *on-line*, contendo uma série de perguntas elaboradas previamente de acordo com os objetivos do estudo e que devem ser respondidas pelos usuários e demais participantes. Para compor um questionário, pode-se utilizar questões abertas, fechadas, de múltipla escolha e escalas de classificação (*Likert*, diferencial semântico) (BARBOSA; SILVA, 2010; ROGERS; SHARP; PREECE, 2013).

Para Rogers, Sharp e Preece (2013), a técnica da observação é útil para qualquer fase do processo de desenvolvimento e pode ser realizada de forma direta ou indireta. A observação direta consiste no investigador passar algum tempo com os usuários, observando todas as atividades que são realizadas. A indireta, por sua vez, tem como característica elaborar registros das atividades do usuário enquanto elas acontecem para serem estudados posteriormente. No início do processo de *design*, a observação pode ser útil para ajudar os *designers* a compreenderem o contexto dos usuários, suas tarefas e objetivos. É também importante nas etapas finais, principalmente na avaliação, para investigar o quão bem o protótipo desenvolvido apoia as atividades previstas no início do estudo.

Os grupos de foco, por sua vez, se assemelham às entrevistas, mas com o diferencial de serem realizadas por diversas pessoas (geralmente entre 3 e 10) reunidas numa espécie de discussão coletiva. São úteis em diversos contextos: gerar ideias; obter opiniões de pessoas sobre tópicos, conceitos ou demonstrações; obter respostas a uma série de questões; identificar conflitos relacionados a terminologias; identificar expectativas de diferentes grupos de pessoas; e descobrir problemas, desafios, frustrações e preferências. Sua principal vantagem é a possibilidade de obter, em pouco tempo, múltiplas perspectivas e pontos de vista de um grupo de pessoas (BARBOSA; SILVA, 2010).

A segunda etapa do ciclo de vida do *design* da interação é o ***design* de alternativas**. Nesta fase, o principal objetivo é sugerir ideias que atendam aos requisitos levantados no início do processo de desenvolvimento. Esta atividade pode ser dividida em

duas subatividades: *design* conceitual e *design* físico. O modelo conceitual é uma abstração que esboça o que as pessoas podem fazer com um produto e quais conceitos são necessários para entender como interagir com ele. O *design* físico considera os detalhes do produto, incluindo características como: cores, sons, imagens, menus e ícones a serem utilizados (ROGERS; SHARP; PREECE, 2013).

Com base no levantamento dos requisitos e no *design* de alternativas, a terceira etapa do ciclo de vida consiste na elaboração de **protótipos** (versão limitada do produto) com a finalidade de analisar a viabilidade do *design* ou a sua adequação por parte dos usuários. Para a elaboração desta atividade, não é necessário que uma parte do *software* esteja desenvolvida, uma vez que existem diferentes técnicas de prototipação.

Nos estágios iniciais do processo, os protótipos denominados de baixa fidelidade podem ser construídos de papel ou papelão que, de acordo com Rogers, Sharp e Preece (2013, p.392), "são úteis porque tendem a ser simples, baratos e de rápida produção", sendo também fáceis de serem modificados e reformulados, se necessário. Em função da sua flexibilidade, são ferramentas importantes para o processo de *design*, pois dão suporte à exploração de alternativas de *design*, à discussão de ideias e a identificação de problemas. Além dos protótipos em papel, podem ser utilizados ainda para a prototipação outros recursos, tais como: *storyboards* (conjunto de cenários, histórias em quadrinho), esboços desenhados a mãos (*sketch*) e prototipação com cartões.

Além dos protótipos de baixa fidelidade, os de alta fidelidade também são de grande importância para o processo de desenvolvimento. Este tipo de prototipação adota protótipos que apresentam funcionalidades completas, são totalmente interativos e possuem a mesma aparência do produto final. No entanto, possuem algumas desvantagens que podem interferir negativamente na sua utilização: são demorados para serem construídos, podem criar expectativas altas nos clientes, os desenvolvedores podem ser relutantes à mudanças sugeridas, uma vez que podem ser necessárias inúmeras horas de trabalho adicional para isto e, por fim, os revisores tendem a comentar mais sobre aspectos superficiais do que sobre o conteúdo do produto (ROGERS; SHARP; PREECE, 2013).

A próxima etapa do ciclo de vida consiste em **avaliar** o *design* desenvolvido para determinar a sua usabilidade e aceitabilidade. É uma atividade que pode adotar grande variedade de critérios, incluindo o número de erros que os usuários cometeram durante a utilização, o quão atraente ele é, o quanto os requisitos foram satisfeitos, a facilidade de aprendizado e quaisquer outras informações que possam contribuir para a melhoria do *design*. De acordo com Rogers, Sharp e Preece (2013), as avaliações podem ser formativas ou

somativas. As avaliações formativas são elaboradas ao longo de todo o processo de *design* e possuem como principal objetivo compreender e confirmar as necessidades dos usuários, bem como identificar possíveis problemas de interação e de interface. A avaliações somativas são realizadas no final do processo de *design*, quando já existe uma solução parcial ou completa (protótipo de alta fidelidade) e são utilizadas para analisar a qualidade de uso do produto, bem como evidências que indiquem se as metas de usabilidade foram, de fato, alcançadas.

Existem diferentes métodos de avaliação que podem ser adotados e adaptados a circunstâncias específicas. Rogers, Sharp e Preece (2013) apresentam o *framework* DECIDE que fornece uma lista de verificação (*checklist*) para ajudar no planejamento, na execução e na análise de uma avaliação de *design*. É composto por seis itens:

1. ***D****etermine*: determinar os objetivos da avaliação;
2. ***E****xplore*: explorar as questões a serem respondidas pela avaliação;
3. ***C****hoose*: escolher os métodos de avaliação a serem utilizados;
4. ***I****dentify*: identificar as questões práticas a serem consideradas;
5. ***D****ecide*: decidir como lidar com as questões éticas;
6. ***E****valuate*: avaliar, analisar, interpretar e apresentar os dados coletados.

No mesmo contexto, Snyder (2003) apresenta um método específico para avaliação de prototipação em papel, que tem como objetivo avaliar a usabilidade de um *design* representado em papel, através de simulações de uso, com a participação de potenciais usuários. De acordo com Barbosa e Silva (2010, p.358), "esse método é uma opção interessante para uma avaliação formativa junto aos usuários, principalmente para comparar alternativas de *design*". Ele permite avaliar com facilidade soluções parciais de baixa e média fidelidade, que ainda não definem todos os detalhes da interface, com a intenção de identificar problemas de usabilidade antes mesmo do desenvolvimento do sistema. É composto por cinco etapas:

a) Preparação: elaboração do material de apoio e dos protótipos que serão utilizados. As atividades que devem ser desenvolvidas nesta etapa são:
   - Definir tarefas para os participantes executarem;
   - Definir o perfil dos participantes e recrutá-los;
   - Criar protótipos em papel da interface para executar as tarefas;
   - Executar um teste piloto;
b) Coleta de dados: cada usuário deve executar as tarefas propostas, interagindo com os protótipos em papel, mediado por um avaliador, que

deve simular o comportamento do sistema e observar a experiência de uso, com a intenção de identificar. É importante registrar:

- Partes da interface que funcionaram bem e as que não foram bem interpretadas pelo usuário;
- Tarefas concluídas com sucesso;
- Erros cometidos;
- Comentários feitos;
- Quaisquer outros dados que possam auxiliar na avaliação da interface;

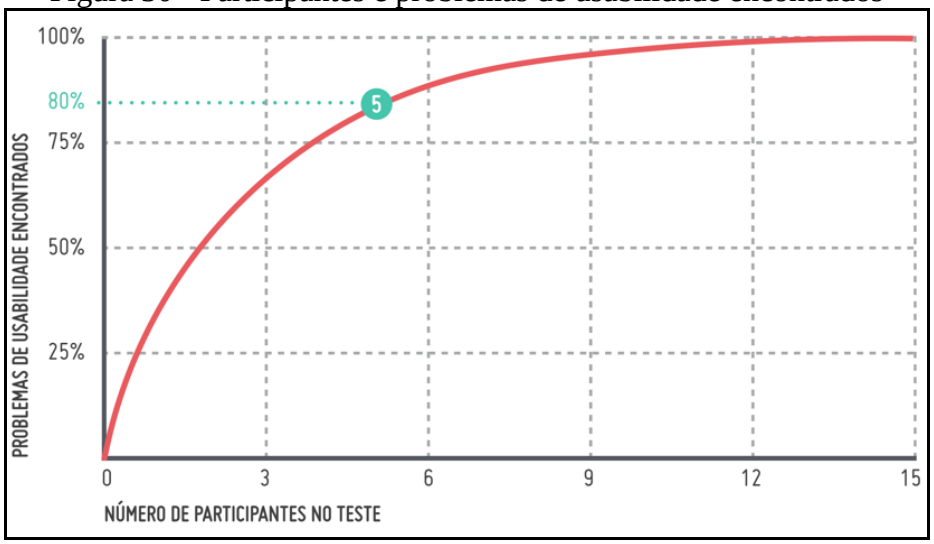
c) Interpretação: realizada após a coleta de dados, tem como objetivo analisar as anotações feitas sobre a experiência de uso, afim de identificar problemas de usabilidade da interface. Suas principais atividades são:
   - Listar os problemas encontrados;
   - Refinar os protótipos em papel para resolver os problemas mais simples;

d) Consolidação dos resultados: etapa na qual o avaliador verifica quais problemas não puderam ser resolvidos e sugerem propostas de correção ou de alternativas que podem ser exploradas para melhorar a interface;

e) Relato dos resultados: etapa final, que tem como objetivo relatar os problemas encontrados e as sugestões de correção. É importante descrever:
   - As tarefas executadas pelos participantes;
   - Tabelas e gráficos que sumarizam as medições realizadas;
   - Lista de problemas corrigidos durante os ciclos de avaliação, *design* de alternativas e prototipação;
   - Lista de problemas ainda não corrigidos;
   - Sugestões de solução.

Barbosa e Silva (2010) afirmam que a avaliação deve ser realizada com participantes que representem o público-alvo do sistema avaliado, ou seja, que possuam características semelhantes aos usuários típicos. Nielsen (2000) sugere que, em média, apenas cinco participantes são necessários para a realização de estudos informais de usabilidade, o que possibilita a identificação de cerca de 80% dos problemas existentes, conforme pode ser observado na Figura 30.

Figura 30 - Participantes e problemas de usabilidade encontrados



Fonte: Nielsen (2000, p.2)

O cenário apresentado por Nielsen (2000) se refere ao fato de que, ao coletar dados de um único participante, é possível levantar quase um terço de tudo o que existe para ser identificado sobre o *design*. Quando o segundo participante é testado, o avaliador descobre novas informações, mas com algum percentual de sobreposição. O terceiro, por sua vez, identificará muitas coisas que já foram observadas pelos dois primeiros participantes, adicionando, portanto, uma quantidade muito pequena de novas informações. Desta forma, à medida que mais participantes são adicionados ao teste, menos o avaliador aprende, pois a maioria dos problemas observados já foi indicada nos testes anteriores.

Por fim, é importante destacar ainda que as atividades previstas no ciclo de vida do *design* de interação não são sequenciais, tampouco realizadas apenas uma vez durante o processo de desenvolvimento da proposta. Isto significa que pode ser necessário analisar os requisitos e propor alternativas de *design* diversas vezes antes de iniciar a prototipação. Além disso, pode ser necessário revisar os requisitos e retomar as etapas após a avaliação do produto. É, portanto, um conjunto de atividades relacionadas e interativas, desenvolvidas de forma evolutiva, até que o produto final atenda às necessidades do usuário e os critérios de usabilidade previstos.

6.1 *DESIGN* BASEADO EM CENÁRIOS

O *Design* Baseado em Cenários é uma metodologia proposta por Rosson e Carroll (2002), com o objetivo de apoiar as atividades previstas nas primeiras etapas do ciclo de concepção do sistema. É uma abordagem centrada no usuário, que consiste em registrar

atividades significativas, com a finalidade de documentar o raciocínio de projetistas sobre cenários de uso e possibilitar a análise de situações sob diferentes perspectivas. Sua utilização motiva os interessados no sistema a participarem do processo de desenvolvimento, de forma a contribuírem direta ou indiretamente, com as decisões de *design* (BARBOSA; SILVA, 2010).

Carroll (2000) sintetiza o conceito, afirmando que cenários são histórias sobre pessoas e suas atividades. São importantes para identificar as necessidades do usuário, prever futuras possibilidades de utilização, identificar novas formas de interação e modelar corretamente as tarefas que serão implementadas. Além disso, ajudam os desenvolvedores a coordenar e refletir sobre as ações do projeto, oferecem múltiplas visões de uma interação, contribuem para a análise das consequências decorrentes de qualquer mudança de *design* e promovem a comunicação entre as partes interessadas, contribuindo para a concepção de sistemas que possam, de fato, atender as necessidades e as preocupações dos usuários (CARROLL, 2000).

> Ao ler, escrever e revisar cenários, a equipe de *design* (incluindo representantes dos usuários) tem a oportunidade de discutir e analisar como as atividades dos usuários são afetadas pela tecnologia existente e como elas poderiam ser afetadas pelo sistema que está sendo desenvolvido. As histórias dos cenários estimulam a imaginação da equipe *design* e encorajam a análise de caminhos alternativos (BARBOSA; SILVA, 2010, p.112).

De acordo com Nielsen (2013), a *persona* é um dos principais elementos dos cenários e pode ser definida como um personagem fictício ou um modelo hipotético de um grupo de usuários reais, que é criado para descrever um usuário típico, representando os agentes ou atores que interagem com o sistema. Entre os benefícios desta técnica, a autora aponta que a utilização de *personas* auxilia os desenvolvedores na compreensão das características de cada grupo de usuários, na análise das necessidades mais relevantes e, por fim, facilita a aproximação dos desenvolvedores com os potenciais usuários. São documentadas com a utilização de foto, nome e objetivos relacionados ao futuro sistema (CARVALHO et al., 2011).

Embora *personas* sejam fictícias, elas devem ser definidas com rigor e detalhamento suficiente para representar usuários típicos, a partir de um processo de investigação que analisa, de forma criteriosa, as características e os perfis dos envolvidos. Quanto mais específicas forem, mais eficientes elas serão como ferramentas de *design* e comunicação. Desta forma, apenas o nome e os detalhes pessoais devem ser inventados.

Nielsen (2013) salienta que uma *persona* não é o mesmo que uma modelo

ou uma pessoa. A descrição de uma *persona* não tem como característica detalhar o usuário como um todo, mas utilizar a área de foco ou domínio como uma lente para destacar as atitudes relevantes e o contexto específico associado à área de trabalho. De acordo com Barbosa e Silva (2010, p.176), as *personas* "devem ser definidas principalmente por seus objetivos, que são determinados num processo de refinamentos sucessivos durante a investigação inicial do domínio de atividade do usuário".

Cooper (1999) recomenda que a prática de ampliar a funcionalidade de um produto para atender a maior parte das pessoas deve ser evitada. Ao invés disso, os desenvolvedores devem tentar projetar especificamente para uma única *persona*. Isto significa que quando uma funcionalidade é estendida para incluir mais grupos de utilizadores, obstáculos podem ser inseridos para todos os usuários, ou seja, os recursos que agradam alguns, podem interferir na satisfação e no desempenho de outros. Assim, tentar satisfazer muitos pontos de vista diferentes pode arruinar um bom projeto.

Além das *personas*, os cenários possuem ainda os seguintes elementos: contexto de uso, metas ou objetivos a serem atingidos, sequência de ações e eventos para atingir os objetivos propostos e avaliação das atividades desenvolvidas. Podem ser representados através de descrições textuais, *storyboards*, vídeos e ferramentas de prototipagem (ROSSON; CARROLL, 2002).

6.2 RESUMO E DISCUSSÃO DO CAPÍTULO

O objetivo deste capítulo foi apresentar as etapas de desenvolvimento do *Design* de Interação, que deu subsídios para a elaboração do "Modelo de interface de busca e recuperação de informação". O ciclo de vida proposto é composto de quatro etapas relacionadas e interativas, que podem ser realizadas inúmeras vezes no contexto de um projeto, sendo elas: identificação dos requisitos, *design* de alternativas, prototipação e avaliação.

Além das etapas previstas pelo *Design* de Interação, o *Design* Baseado em Cenários também foi discutido no âmbito deste capítulo, que é uma abordagem centrada no usuário, com o objetivo de documentar atividades significativas sobre cenários de uso e analisar situações sob diversas perspectivas.

Os cenários são histórias sobre pessoas e suas atividades, que contribuem para que os desenvolvedores possam identificar as necessidades dos usuários, prever futuras possibilidades de utilização, identificar novas formas de interação e modelar corretamente as

tarefas que serão implementadas. Podem ser representados através de descrições textuais, *storyboards*, vídeos e ferramentas de prototipagem e devem ser elaborados a partir dos seguintes elementos: *personas*, contexto de uso, metas ou objetivos a serem atingidos, sequência de ações e eventos para atingir os objetivos propostos e avaliação das atividades desenvolvidas.

O principal elemento de um cenário é uma *persona*, que pode ser caracterizada como um personagem fictício ou um modelo hipotético de um grupo de usuários reais, que é criado para descrever um usuário típico, representando os agentes ou atores que interagem com o sistema.

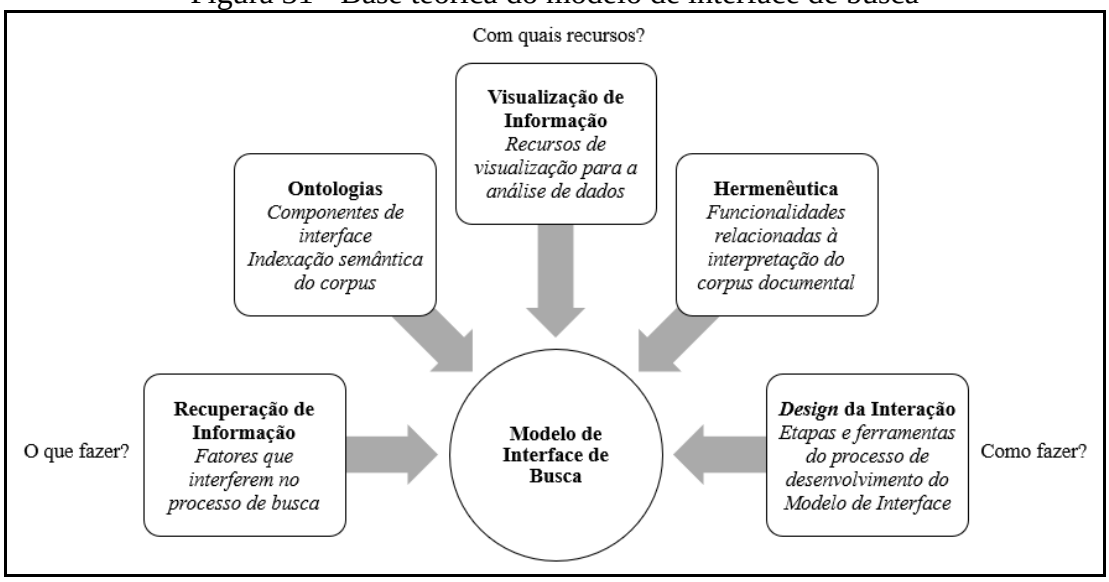
7 MODELO DE INTERFACE DE BUSCA PROPOSTO

Este capítulo tem como objetivo apresentar um "Modelo de interface de busca e recuperação de informação", discutindo detalhadamente as etapas do processo de desenvolvimento, bem como os requisitos levantados, as características do público alvo, as tarefas previstas, os cenários de busca que deram subsídios para o projeto da interface, a análise das interações suportadas em cada contexto, os protótipos que atendem os requisitos identificados e, finalmente, a avaliação de usabilidade.

■■■

O escopo desta pesquisa é a concepção de um "Modelo de interface de busca e recuperação de informação baseado nos princípios da hermenêutica e na utilização de ontologias", com recursos capazes de possibilitar a organização, análise e exploração de documentos científicos, a partir de um mapeamento visual das informações, permitir que o usuário tenha uma melhor compreensão e interpretação do domínio e das relações existentes no *corpus* documental. A base teórica adotada para fundamentar a proposta está resumida na Figura 31.

Figura 31 - Base teórica do modelo de interface de busca



Fonte: desenvolvido pela autora

A área Recuperação de Informação forneceu subsídios para que fosse possível entender os fatores que interferem no processo de busca, bem como os principais problemas envolvidos com a recuperação de informação na Web, relacionados particularmente com o comportamento do usuário e a sua interação na formulação da

expressão de busca e na interpretação dos resultados apresentados. Desta forma, foi possível identificar as principais dificuldades da área e, a partir delas, definir os primeiros requisitos do "Modelo de interface de busca" proposto neste estudo.

As funcionalidades projetadas para minimizar os problemas identificados e facilitar o processo de recuperação e análise de informação foram fundamentadas no uso de ontologias e nos princípios da hermenêutica. As ontologias foram adotadas em dois contextos diferentes: como componentes de interface, com o objetivo de facilitar a interação do usuário com o sistema e como estrutura de indexação dos artigos científicos, com a intenção de possibilitar a análise da similaridade existente entre os documentos pertencentes ao *corpus*.

Os princípios da hermenêutica deram subsídios ao desenvolvimento de funcionalidades que visam auxiliar a interpretação do usuário durante a análise de um documento ou de um *corpus* documental, seja no processo de indexação ou de exploração dos resultados de uma busca. Além dos recursos projetados a partir do uso de ontologias e dos princípios da hermenêutica, a área de visualização de informação também ofereceu algumas contribuições para potencializar a percepção, o entendimento e a análise do usuário em relação à informação apresentada.

Os recursos gráficos utilizados para facilitar a análise das informações disponibilizadas na interface, tanto no contexto da elaboração de buscas quanto na análise dos resultados foram projetados com base na área de Visualização de Informação.

E, por fim, as etapas de desenvolvimento da proposta foram fundamentadas na literatura da Ciência da Computação, mais especificamente no processo de desenvolvimento de *software*, adotando como base o *design* de interação, com foco na utilização de personagens, cenários e prototipação (ROSSON; CARROL, 2002; ROGERS; SHARP; PREECE, 2013).

7.1 ETAPAS DE DESENVOLVIMENTO DA PROPOSTA

O ciclo de vida básico recomendado pelo *design* de interação e adotado no contexto de desenvolvimento do "Modelo de interface de busca e recuperação de informação" é composto de quatro etapas relacionadas, todas centradas no usuário, que representam um conjunto de atividades que devem ser elaboradas durante o projeto, de forma interativa, através de ciclos repetidos de *design*-avaliação-*redesign*. São previstas quatro atividades principais: estabelecer requisitos, *design* de alternativas, prototipação e avaliação, todas envolvendo a participação efetiva do usuário.

A primeira etapa de desenvolvimento da proposta consistiu em **estabelecer os requisitos**, de forma que fosse possível compreender as principais características e funcionalidades da proposta. Para isso, foi necessário analisar detalhadamente o domínio de estudo, as oportunidades de melhoria, os requisitos funcionais e não funcionais, o público alvo e as tarefas previstas.

A coleta dos dados necessários para definir os requisitos iniciais da interface de busca foi realizada por meio de entrevistas não estruturadas com usuários de sistemas de busca de informação, além da análise da literatura existente sobre o assunto. Desta forma, foi possível identificar problemas e sugestões de melhorias relacionados à busca (Quadro 7) e à apresentação dos resultados (Quadro 8). A partir das informações levantadas nesta etapa, foi possível delinear os requisitos funcionais e não funcionais, que são apresentados nos Quadros 9 e 10, respectivamente.

A etapa subsequente, denominada de ***design* de alternativas** é considerada a atividade principal do *design*, e tem como objetivo sugerir ideias para satisfazer os requisitos identificados. Nesta fase, a partir dos requisitos definidos na etapa anterior, foi possível definir cenários para documentar as ideias relacionadas à interface de busca proposta, bem como o *layout*, a estrutura de navegação e os recursos gráficos adotados. Para apoiar as atividades desenvolvidas, utilizou-se a metodologia de *Design* Baseado em Cenários proposta por Rosson e Carroll (2002).

Com base no levantamento dos requisitos e no *design* de alternativas, foi possível elaborar **protótipos** de baixa fidelidade, que tiveram como principal objetivo simular as funcionalidades previstas e documentadas nos cenários, bem como enfatizar características da interface. Os protótipos foram elaborados com o auxílio da ferramenta de prototipação denominada Pencil Project, versão 3.0.4 (https://pencil.evolus.vn), que é um *software* gratuito e de código aberto. A documentação produzida nas etapas de *design* de alternativas e prototipação é apresentada na seção 7.3.

A última etapa do ciclo de vida de desenvolvimento do “Modelo de interface de busca e recuperação de informação” prevê a elaboração de testes de usabilidade para **avaliar** os cenários elaborados, as tarefas previstas e os protótipos desenvolvidos, com o intuito de analisar se os requisitos foram satisfeitos e determinar a usabilidade e aceitabilidade do produto, bem identificar possíveis erros de interação. Para a elaboração desta atividade, adotou-se o método proposto por Snyder (2003).

O processo de avaliação realizado neste estudo teve o intuito de ser formativo, ou seja, foi realizado durante o *design* para verificar se a interface proposta atende

aos requisitos e satisfaz as necessidades do usuário. Como método de avaliação, optou-se pela realização de testes de usabilidade informais, que são menos rígidos, mais rápidos e requerem uma menor quantidade de participantes. No entanto, podem indicar com precisão quais caminhos de *design* devem ser seguidos e revelar problemas de usabilidade assim como outros testes mais rigorosos (HEARST, 2009).

O processo de avaliação foi orientado através de um roteiro de atividades sequenciais, previstas para cada cenário idealizado (Apêndice A). Os testes de usabilidade foram elaborados com o apoio e a utilização dos protótipos de baixa fidelidade em papel e realizados com três participantes na fase de pré-teste e com dez participantes, na fase de ensaios, todos pesquisadores e potenciais usuários da interface projetada, com experiência na utilização de ferramentas de busca e no uso de repositórios digitais de informação científica. Além disso, foi utilizado questionário de avaliação no final de cada teste realizado (Apêndice B), composto por questões abertas, fechadas, escala de Likert (5 pontos) e diferencial semântico (7 pontos). Ainda como técnica de coleta de dados, utilizou-se a observação direta, que foi realizada durante os testes de usabilidade.

7.2 REQUISITOS DO MODELO DE INTERFACE DE BUSCA

A primeira etapa de desenvolvimento do "Modelo de interface de busca e recuperação de informação" consistiu em definir os requisitos necessários para a elaboração do projeto. Para isto, foi necessário o desenvolvimento de três atividades importantes: 1) análise de domínio e oportunidades de melhoria; 2) requisitos funcionais e não funcionais; 3) análise do público alvo e das tarefas previstas. Os dados que deram subsídios para esta etapa foram levantados por meio de entrevistas não estruturadas com usuários de sistemas de busca de informação (potenciais usuários da interface proposta), bem como a partir da análise da literatura existente sobre o assunto.

7.2.1 Análise do Domínio e Oportunidades de Melhorias

A ideia de propor uma interface de busca, com base nos princípios da hermenêutica e na utilização de ontologias, surgiu a partir do levantamento bibliográfico realizado nas duas primeiras etapas deste estudo (etapas 1 e 2), que ofereceu subsídios para a análise das dificuldades enfrentadas pelos usuários durante um processo de busca e recuperação de informações, bem como das inúmeras limitações apresentadas pelas interfaces

atuais e da possibilidade de minimizá-las.

De acordo com Antoniou e Harmelen (2004), Beppler (2008) e Baeza-Yates e Ribeiro Neto (2013), os buscadores disponíveis, a maioria baseados nas caixas de texto, apesar de terem se tornado populares em função da sua extrema simplicidade, apresentam diversos problemas que interferem na processo de recuperação de informação, relacionados principalmente com a elaboração das expressões de busca e com a interpretação dos resultados de uma pesquisa.

Os principais problemas referentes à elaboração das expressões de busca estão fundamentados nos estudos de Antoniou e Harmelen (2004), Camacho (2004), Beppler (2008), Baronas (2011) e Nayak, Prasad e Senapati (2015). As propostas de melhorias encontradas na literatura para a solução dos problemas observados estão embasadas nas pesquisas apresentadas por Jackson e Moulinier (2002), Woods (2004), Hyvönen e Mäkelä et al. (2005), Lei, Uren e Motta (2006), Beppler (2008), Baeza-Yates e Ribeiro-Neto (2013) e Shneiderman et al. (2017). Problemas e respectivas sugestões de melhorias são apresentados resumidamente no Quadro 7.

Quadro 7 - Problemas e sugestões de melhorias relacionados à busca

Problema 1: Dificuldade de expressar uma necessidade de informação por meio de palavras-chave ou termos de busca ▪ O desenvolvimento de novas metáforas visuais para a elaboração de expressões de busca pode trazer melhorias para o processo de recuperação de informação ▪ É necessário o desenvolvimento de ferramentas que ofereçam busca semântica de forma gráfica, substituindo as palavras-chave, tais como: lista de categorias, hierarquias e taxonomia multifacetada; ▪ As ontologias permitem a interação com documentos recuperados, servindo de base para refinar ou iniciar novas buscas
Problema 2: Obrigatoriedade de conhecimento sobre o domínio pesquisado, ou seja, quanto mais o usuário souber sobre o domínio de interesse, mais opções de vocabulário e terminologia terá a sua disposição para enriquecer as buscas, ao mesmo tempo que, quanto menos conhecimento possuir, mais genéricas e inadequadas serão as expressões utilizadas ▪ As ontologias disponibilizam vocabulário adequado ao usuário para a elaboração de expressões de busca mais significativas e melhor contextualizadas, realizadas a partir dos conceitos definidos na ontologia
Problema 3: Dificuldade de compreender as terminologias de um domínio e interpretar as diversas relações existentes entre elas, principalmente quando não existe conhecimento do usuário sobre o assunto pesquisado ▪ As ferramentas de busca devem ser capazes de lidar com relacionamentos semânticos ▪ As ontologias apresentadas em formato gráfico, possibilitam um melhor entendimento sobre os conceitos e instâncias de um domínio pesquisado
Problema 4: São ambientes fortemente influenciados pelos problemas linguísticos (sinonímia, homonímia, paronímia, polissemia, denotação, conotação e variações linguísticas), que dificultam ainda mais a tradução das necessidades de informação do usuário em expressões de busca ▪ Os conceitos de uma ontologia podem ser utilizados na elaboração de expressões de busca, minimizando os problemas linguísticos ▪ As ontologias possibilitam a formulação de consultas menos ambíguas, uma vez que os usuários podem expressar claramente o contexto de cada termo utilizado na consulta

Quadro 7 - Problemas e sugestões de melhorias relacionados à busca (continuação...)

Problema 5: Não oferecem suporte para auxiliar o pesquisador quando ele não possui objetivos de busca bem definidos, ou seja, quando ele está à procura de algo que não sabe exatamente o que é ou quais termos de busca deve utilizar ▪ As ontologias ajudam na elaboração de expressões de busca, principalmente quando o usuário não sabe ao certo o que deseja ou não possui conhecimento sobre o assunto pesquisado
Problema 6: Não apresentam mecanismos que auxiliem o usuário a relembrar termos do domínio ou a utilizar palavras que não sejam conhecidas ▪ As ontologias estimulam os usuários a explorarem diferentes aspectos de suas necessidades, por meio da interação direta com conceitos e relacionamentos definidos na ontologia ▪ A partir de um termo ou conceito oferecido pelo usuário, a interface deve ser capaz de extrair outros termos ou conceitos relacionados, indicando novas estratégias de busca

Fonte: desenvolvido pela autora

Os problemas referentes à interpretação dos resultados de uma pesquisa foram fundamentados em Huang (1999), Spink et al. (2001), Hobbs, Pfitzner e Powers (2002), Case (2012) e Baeza-Yates e Ribeiro-Neto (2013). As respectivas propostas de melhorias presentes na literatura foram encontradas nos estudos de Branski (2004), Hyvönen e Mäkelä et al. (2005), Lei, Uren e Motta (2006), Beppler (2008), Hogan et al. (2011) e Baeza-Yates e Ribeiro-Neto (2013), conforme apresentado no Quadro 8.

Quadro 8 - Problemas e sugestões de melhorias relacionados aos resultados

Problema 7: Os sistemas não apontam as relações existentes entre os documentos que, predominantemente, são considerados objetos desconectados entre si, transferindo ao usuário a necessidade de estabelecer as conexões existentes para compreender o contexto investigado ▪ Desenvolvimento de novas metáforas visuais para a visualização de resultados, com condições de reduzir o trabalho adicional do usuário para a combinação da estrutura e do conteúdo nas consultas ▪ Devem ser implementadas estruturas de indexação específicas que explorem a estrutura dos documentos ▪ É importante incluir recursos que melhorem a navegação, tais como: melhor exploração dos *links*, popularidade de páginas Web, *similaridade de conteúdo, colaboração*, 3D e realidade virtual, com atenção especial para *abordagens de busca/navegação híbridas*
Problema 8: Resultados são baseados apenas em informações apresentadas nos documentos de forma explícita, dispensando todo o conhecimento latente ▪ As ferramentas de busca devem incorporar tarefas de coleta de informação acerca de um grupo de sites ou documentos. Nas ferramentas atuais, os usuários devem agregar manualmente as informações que precisa de vários locais heterogêneos, cada qual com suas características de formatação e navegação ▪ É importante o desenvolvimento de recursos que permitam a execução de múltiplas tarefas de recuperação de informação, tais como: selecionar documentos de uma base, agrupá-los em categorias ou classes e deles extrair determinados conjuntos de informações ▪ Com base em um termo oferecido pelo usuário, a interface deveria ser capaz de identificar citações relacionadas e, a partir delas, extrair outros termos ou conceitos que indicassem novas estratégias de busca ▪ As ontologias possibilitam a visualização de diferentes perspectivas das informações recuperadas
Problema 9: Não fornecem mecanismos que indiquem quais termos foram utilizados para representar o acervo documental ▪ Um dos desafios da busca de informações é encontrar resultados relevantes. Em vez de uma longa lista de títulos, URLs e documentos, os mecanismos de pesquisa visuais podem oferecer apresentações ricas de resultados, muitas vezes visualmente conectadas a termos de pesquisa relacionados

Fonte: desenvolvido pela autora

Além dos problemas observados no contexto da elaboração de expressões de busca e da interpretação dos resultados de uma pesquisa, destacam-se ainda as diferenças

entre os usuários, que estão relacionadas às suas características individuais, sociais e ambientais (WILSON, 2006; SAVOLAINEN, 2016). Isto significa que cada pesquisador, ao utilizar um repositório digital para buscar um determinado material, o fará de acordo com o seu conhecimento, sua compreensão de domínio, seu vocabulário, seus preconceitos e sua percepção em relação ao mundo.

Com base neste cenário, percebeu-se, portanto, a necessidade de desenvolver uma interface de busca, contendo recursos de interação capazes de minimizar os problemas observados e auxiliar os usuários durante o processo de recuperação de informação. Neste sentido, a interface proposta apresenta funcionalidades que visam apoiar as atividades desenvolvidas por um pesquisador em três aspectos principais: na sua interação com o sistema, na análise da similaridade existente entre os documentos retornados em uma busca e na interpretação de um *corpus* documental.

Os problemas observados no contexto da interação do usuário com o sistema (problemas 1, 2, 3, 4, 5 e 6 apresentados no Quadro 7) foram minimizados com a utilização de ontologias como componentes de interface. Entre os benefícios esperados, o uso de ontologias em formato gráfico de mapa de conceitos, dará ao pesquisador condições de aprender a partir das interações realizadas e adquirir conhecimento sobre o conteúdo pesquisado. Outro ponto importante é a possibilidade de compreender as terminologias do domínio da busca e as relações existentes entre elas e, a partir disso, permitir a utilização de termos desconhecidos ou não lembrados pelo usuário, expandindo assim o conjunto de resultados recuperados. Além disso, poderá ainda disponibilizar vocabulário adequado para a elaboração de buscas mais significativas e contextualizadas que serão úteis, inclusive, para auxiliar o pesquisador quando ele não possuir objetivos de pesquisa bem definidos.

Ainda em relação ao uso de ontologias como componentes de interface, outra contribuição significativa refere-se à possibilidade de redução dos problemas linguísticos, uma vez que, ao selecionar um termo de uma ontologia para compor uma expressão de busca, o usuário não estará selecionando apenas uma palavra qualquer, mas sim toda a semântica que a acompanha, ou seja, o seu significado original e o contexto de interpretação ao qual ela está relacionada.

A análise da similaridade existente entre os documentos foi obtida a partir do uso de ontologias na estrutura de classificação dos documentos e das relações, visando a organização e a indexação do *corpus* documental. Este recurso tem como objetivo minimizar os problemas 7, 8 e 9 apresentados no Quadro 8 e possibilitar a construção de uma rede de informações entre os documentos, que influenciará todas as buscas realizadas e auxiliará os

usuários a compreender e interpretar o contexto avaliado.

Além de contribuir para a estruturação dos documentos, as ontologias permitiram ainda uma representação visual dos resultados de uma busca, a análise das semelhanças existentes entre eles, a identificação de *clusters* de documentos e a extração de conjuntos de informações relacionados a eles, a descoberta de conhecimento latente a partir das relações apresentadas e a exploração das conexões semânticas entre os documentos e suas partes. Neste sentido, a partir da análise dos relacionamentos existentes dentro de um contexto, será possível identificar pontos de convergência e divergência existentes entre os elementos de um *corpus*.

As funcionalidades relacionadas à interpretação do *corpus* documental, tanto no contexto da organização dos documentos, quanto na análise dos resultados da busca foram obtidas a partir dos fundamentos oferecidos pela Hermenêutica. O principal objetivo é ampliar a capacidade de compreensão de um *corpus* por parte do usuário, através dos seguintes princípios: círculo hermenêutico, contextualização, interação entre os pesquisadores e sujeitos, abstração e generalização, raciocínio dialético, múltiplas interpretação, suspeição e compartilhamento de linguagem. Também visa reduzir as barreiras referentes ao comportamento do usuário, principalmente as que dizem respeito à dissonância cognitiva, ao nível de conhecimento e aos problemas de linguagem (vocabulário limitado, diferenças culturais e problemas semânticos) que interferem no processo de interpretação dos documentos e dificultam o entendimento do conteúdo analisado por parte do pesquisador.

O estudo foi desenvolvido com o objetivo de ser adotado em um domínio científico, mas possui potencial de aplicação em diversos outros contextos que envolvam a busca, a análise e a recuperação de informações, tais como sistemas organizacionais, empresariais e acadêmicos. Os documentos presentes nesta base de dados referem-se a artigos científicos, artigos de revisão, ensaios e resenhas. Para definir a estrutura de representação destes materiais, adotou-se como referência as seguintes bases científicas: ACM Digital Library, Brapci, Google Scholar, Scielo e Web of Science, que possuem uma organização de conteúdo, em sua maioria, representada pelos seguintes itens: área, autor, afiliação, palavras-chave, objetivo, objeto de estudo, problema, base teórica, procedimento metodológico, conclusão e referência. No contexto deste trabalho, estes elementos serão denominados de **Perspectivas de Análise**.

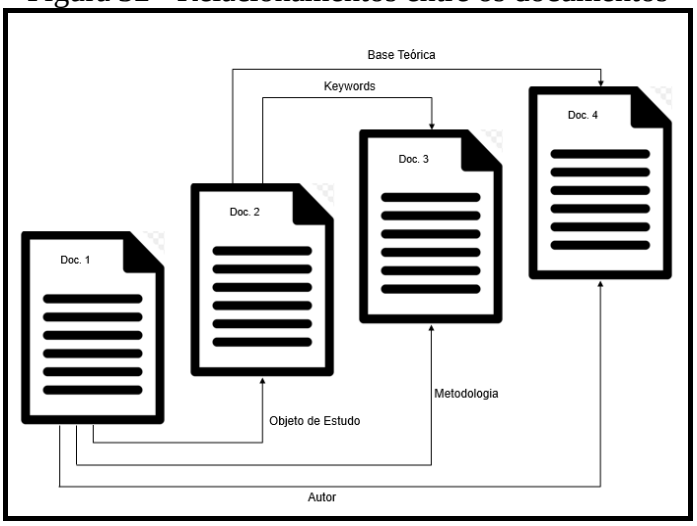
O conceito de perspectivas de análise é fundamentado por Gadamer (1999) quando discute as perspectivas de percepção de uma coisa em si ou de um objeto. Segundo ele, quando alguém anda ao redor de uma mesa, experimenta perspectivas diferentes e

distintas da mesa, ou seja, nenhuma das perspectivas revela a mesa inteira. Desta forma, ao possibilitar a análise de um documento (ou de parte dele), de um *cluster* de documentos ou de um *corpus* a partir de perspectivas de análise diversas, a interface, indiretamente, ampliará as possibilidades de inferência do usuário acerca dos elementos presentes na busca ou nos resultados retornados.

As informações referentes à estrutura de apresentação do conteúdo são intensamente utilizadas pelo usuário, de forma consciente ou inconsciente, durante o processo de análise de conhecimento. Apenas para exemplificar, pode-se imaginar um pesquisador com o objetivo de escrever um artigo do tipo *survey* sobre **Biblioteca Pública**. Para que esta tarefa seja possível, é necessário consultar as bases de dados relevantes sobre o assunto e selecionar artigos que possam colaborar com a análise do pesquisador.

Os materiais encontrados são apresentados pelas ferramentas de busca de forma isolada, apenas atendendo a critérios estabelecidos para a pesquisa, como se não compartilhassem nenhum tipo de informação. Neste caso, cabe ao usuário analisar as publicações, identificar as relações existentes entre elas e organizá-las de acordo com suas semelhanças e diferenças, bem como com as características importantes para atender os objetivos da pesquisa. Um exemplo deste cenário pode ser visualizado na Figura 32.

Figura 32 - Relacionamentos entre os documentos



Fonte: desenvolvido pela autora

Considerando o contexto apresentado, o pesquisador deve analisar as características de cada documento e compará-las com as dos demais, com o intuito de entender o cenário estudado. Ao avaliar os documentos a partir de diversas perspectivas, é possível levantar informações importantes: se pertencem a um mesmo autor, à uma mesma

instituição, se compartilham o mesmo objeto de pesquisa, se adotam uma mesma metodologia, se possuem as mesmas *keywords*, se compartilham uma mesma base teórica, e assim sucessivamente. De forma superficial, as principais relações existentes entre os quatro artigos são:

a) Doc. 1 e Doc. 2: compartilham o mesmo objeto de estudo (um pode ser o complemento do outro);
b) Doc. 1 e Doc. 3: apresentam características semelhantes em relação à metodologia de análise de dados (ambos podem ter adotado um determinado teste estatístico, por exemplo);
c) Doc. 1 e Doc. 4: escritos pelo mesmo autor;
d) Doc. 2 e Doc. 3: compartilham as mesmas palavras-chave (ou palavras-chave que fazem parte de um mesmo contexto semântico);
e) Doc. 2 e Doc. 4: compartilham uma mesma base teórica para fundamentar o estudo que foi desenvolvido.

É importante destacar que os documentos podem compartilhar inúmeras características entre si e devem ser avaliados no âmbito de uma determinada perspectiva. Isto significa que dois artigos podem ter sido escritos pelo mesmo autor, mas não possuírem nenhuma outra ligação em termos de conteúdo. No entanto, é muito provável que dois artigos que possuam inúmeros pontos em comum relacionados a uma mesma Metodologia, tenham adotado métodos de pesquisa semelhantes, mesmo que abordem assuntos diferentes.

Este processo de organização e de análise de um *corpus* documental não é uma tarefa trivial e, para ser elaborado de forma aprofundada, exige tempo, dedicação e conhecimento de domínio por parte do pesquisador. Além disso, também não é possível a automação da representação e organização das informações dos documentos, necessárias para colocar em prática boa parte das perspectivas. Esta complexidade está diretamente relacionada à quantidade de artigos selecionados para análise, ou seja, o trabalho de um investigador ao avaliar quatro artigos é totalmente diferente daquele que avaliará 100 documentos. No contexto deste estudo, a proposta é que a rede de relacionamentos gerada a partir dos conceitos definidos na ontologia, seja construída de forma colaborativa, envolvendo o indexador responsável pela base de dados, os autores e os usuários da interface.

Pelo fato de não existirem ontologias pré-elaboradas que atendam às necessidades de informação previstas pela ferramenta, este recurso também deverá ser construído interativamente, por meio de conceitos e relacionamentos definidos pelos usuários durante a análise ou a busca de documentos. À medida que nós e relacionamentos vão sendo

definidos entre as publicações presentes no *corpus* documental, a ontologia de base é expandida e atualizada. Essas relações serão do tipo estruturante e não estruturante.

7.2.2 Requisitos Funcionais e Não Funcionais

Com base nas informações levantadas sobre o domínio, foi possível identificar os principais problemas da busca e recuperação de informação. A partir destes elementos, pode-se definir os requisitos funcionais do projeto e estabelecer suas principais características e funcionalidades. Os recursos foram fundamentados nos princípios da hermenêutica e no uso de ontologias.

Quadro 9 - Requisitos funcionais do modelo de interface de busca

INTERAÇÃO DO USUÁRIO COM O SISTEMA (ONTOLOGIA COMO COMPONENTES DE INTERFACE)
1. Permitir a elaboração de buscas realizadas a partir de palavras-chave elaboradas pelo usuário
2. Permitir a elaboração de buscas realizadas a partir de uma estrutura terminológica (ontologia)
3. Permitir a elaboração de buscas realizadas a partir de uma ou várias perspectivas (autor, problema)
4. Permitir a elaboração de buscas realizadas a partir de uma ou várias perspectivas de um documento
5. Permitir a seleção de perspectivas de análise (uma, múltiplas ou todas)
6. Apresentar estrutura de navegação hibrida, disponibilizando recursos de *searching* e *browsing*
7. Apresentar recursos de *zoom* para facilitar a navegação e o detalhamento de informações da ontologia
8. Disponibilizar filtros de busca para auxiliar as pesquisas do usuário
9. Evidenciar as perspectivas de análise candidatas à seleção do usuário
10. Evidenciar os conceitos e documentos selecionados pelo usuário
11. Evidenciar o contexto dos termos selecionados pelo usuário
12. Evidenciar os resultados mais relevantes
SIMILARIDADE ENTRE DOCUMENTOS (ONTOLOGIA COMO ESTRUTURA DE INDEXAÇÃO SEMÂNTICA DO *CORPUS*)
13. Permitir a inclusão, alteração e exclusão de documentos científicos na base de dados
14. Permitir a inclusão, alteração e exclusão de conceitos e relacionamentos pertencentes à ontologia
15. Permitir a inclusão, alteração e exclusão de relacionamentos entre os documentos e a ontologia
16. Permitir a exploração dos documentos através de múltiplas perspectivas de análise
17. Permitir a análise da similaridade existente entre expressão e resultados da busca
18. Permitir a análise da similaridade existente entre resultados de busca
19. Permitir a análise de um grupo de documentos selecionados pelo usuário
20. Apresentar a referência completa dos resultados selecionados para facilitar a identificação de documentos
INTERPRETAÇÃO DE UM DOCUMENTO OU *CORPUS* DOCUMENTAL (PRINCIPIOS DA HERMENÊUTICA)
21. Apresentar, através da estrutura terminológica, conceitos e contextos associados
22. Permitir a visualização total e parcial (todo-parte) dos conceitos presentes na estrutura terminológica para potencializar o movimento interpretativo
23. Apresentar notas explicativas relacionadas a termos de busca
24. Apresentar sinônimos relacionados a termos de busca
25. Apresentar variação linguística relacionada a termos de busca
26. Apresentar as diversas áreas de conhecimento relacionadas a termos de busca
27. Apresentar termos antônimos e conceitualmente contraditórios relacionados a termos de busca
28. Permitir a interação entre pesquisadores na interpretação de documentos pertencentes ao *corpus*
29. Apresentar o histórico do autor relacionado com um determinado material
30. Apresentar sugestões de pesquisas complementares contendo ideias e/ou teses contraditórias
31. Apresentar sugestões de pesquisas complementares contendo resultados relacionados ao mesmo contexto
32. Permitir a indicação de erros e falhas nos documentos pertencentes ao *corpus*

Fonte: desenvolvido pela autora

Os recursos funcionais apresentados no Quadro 9 foram relacionados à três perspectivas relevantes: interação do usuário com o sistema, análise da similaridade existente entre os documentos e interpretação de um *corpus* documental.

Além dos requisitos funcionais, também foi possível delinear os requisitos não funcionais que devem ser analisados durante o desenvolvimento do projeto de interface. Estes elementos se referem, principalmente, às características comportamentais e de usabilidade desejadas.

Quadro 10 - Requisitos não funcionais do modelo de interface de busca

INTERAÇÃO DO USUÁRIO COM O SISTEMA (ONTOLOGIA COMO COMPONENTES DE INTERFACE)
1. Facilitar o entendimento sobre o que existe no *corpus* documental para ser pesquisado
2. Permitir que o usuário entenda quais termos foram utilizados para representar um documento
3. Permitir a aquisição de conhecimento do domínio durante as interações realizadas
4. Permitir a compreensão das terminologias de domínio e das relações existentes entre elas
5. Disponibilizar vocabulário adequado para a elaboração de buscas mais significativas e contextualizadas
6. Permitir a utilização de termos desconhecidos pelo usuário e, consequentemente, ampliar seu vocabulário e suas possibilidades de busca
7. Minimizar os problemas linguísticos da busca (ambiguidade, sinonímia, homonímia etc.)
8. Auxiliar o pesquisador quando ele não possui objetivos de busca bem definidos
SIMILARIDADE ENTRE DOCUMENTOS (ONTOLOGIA COMO ESTRUTURA DE INDEXAÇÃO SEMÂNTICA DO *CORPUS*)
9. Permitir que a estrutura dos documentos possa ser explorada
10. Permitir que as conexões semânticas entre documentos e suas partes sejam exploradas
11. Permitir a identificação de *clusters* de documentos e a extração de um conjunto de informações
12. Permitir a identificação de resultados correlacionados, que o usuário não sabia que existiam
13. Permitir a análise da similaridade entre os documentos
14. Permitir a aquisição de conhecimento implícito (latente) por meio da análise das relações indiretas
INTERPRETAÇÃO DE UM DOCUMENTO OU *CORPUS* DOCUMENTAL (PRINCIPIOS DA HERMENÊUTICA)
15. Favorecer a apropriação de linguagem e a aprendizagem do usuário em relação ao domínio pesquisado
16. Auxiliar na interpretação do significado das palavras e dos relacionamentos semânticos
17. Permitir a aquisição de conhecimento adicional relacionado à busca realizada
18. Permitir a aquisição de conhecimento adicional relacionado aos resultados recuperados
19. Facilitar o processo de análise do usuário relacionado à um documento ou a um contexto de informação
20. Ampliar a compreensão de um documento através das múltiplas perspectivas de interpretação
21. Reduzir as barreiras relacionadas com comportamento do usuário (pessoais, sociais, interpessoais)

Fonte: desenvolvido pela autora

Após o delineamento dos requisitos funcionais e não funcionais e, consequentemente, das funções e características necessárias para o projeto de interface, a próxima etapa teve como objetivo analisar as características do público alvo e as tarefas previstas no modelo proposto.

7.2.3 Análise do Público Alvo e das Tarefas Previstas

O público que deverá interagir com a interface de busca proposta é

composto por três tipos de atores: indexador responsável pela base de dados (que também será o supervisor de conteúdo), autores e usuários. Os documentos podem ser disponibilizados no repositório digital tanto pelo indexador, quanto por autores e usuários. Quando disponibilizados por autores e usuários, só serão publicados após a revisão do indexador. Os autores, quando não inserem seus próprios documentos, possuem a responsabilidade de agregar as informações que não foram identificadas pelo indexador e sugerir a criação, ampliação e atualização dos componentes da ontologia para que os materiais desenvolvidos possam ser melhor indexados.

Os usuários, representados geralmente por pesquisadores, professores, estudantes, profissionais, futuros autores e demais interessados, são responsáveis pelo consumo das informações disponibilizadas e pela construção e expansão de conhecimento acerca das relações envolvendo os conceitos e os contextos presentes na ontologia e nos documentos. O indexador, na condição de supervisor de conteúdo, desenvolverá a tarefa de aceitar ou rejeitar as sugestões dos usuários, legitimando as propostas realizadas. Este controle de qualidade será realizado com base em dois critérios: na qualificação do usuário e/ou na concordância entre as informações apontadas por eles.

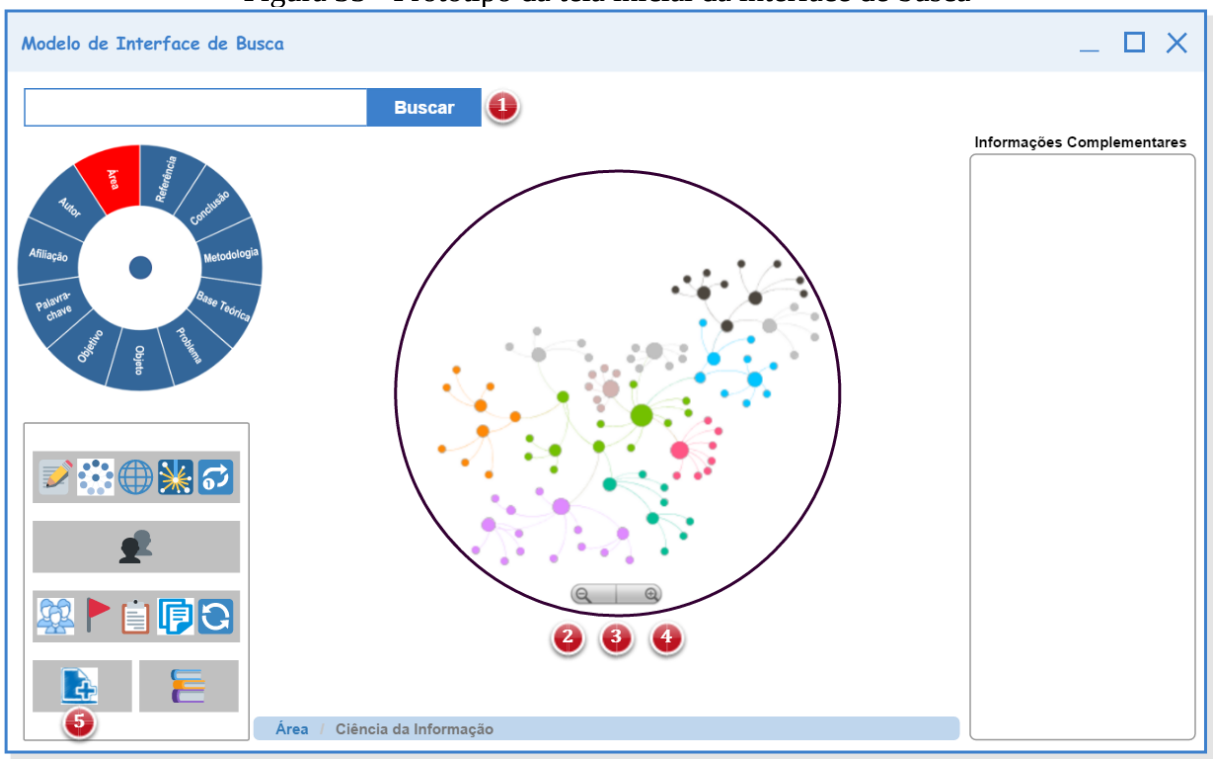
Wilson (2006) afirma que as características **pessoais** (dissonância cognitiva, conhecimento sobre o assunto, sexo, idade, nível educacional, orientação política); **sociais** (categoria profissional, organização, nível de responsabilidade) e **ambientais** (país, legislação, situação econômica, cultura informacional) interferem na busca, nas necessidades de informação e na forma como os usuários analisam e interpretam os resultados. No contexto deste trabalho, os usuários foram classificados em três grupos, de acordo com o nível de conhecimento que possuem em relação ao domínio:

a) Sem conhecimento: não possuem vocabulário suficiente para realizar as buscas e não têm condições de fazer inferência entre os assuntos;
b) Conhecimento parcial: possuem vocabulário parcial para realizar buscas e domínio de conhecimento para fazer algumas inferências entre os assuntos;
c) Conhecimento suficiente: possuem conhecimento de domínio suficiente para realizar buscas e fazer inferências entre os assuntos.

A interface proposta para esta base de dados possibilitará a elaboração de buscas baseadas em ontologias, que poderão ser realizadas por todos os tipos de usuários, especialmente por aqueles que possuem limitações de conhecimento acerca do domínio de pesquisa e, portanto, apresentam dificuldades para explorar os conteúdos disponibilizados. O

protótipo da tela inicial da interface de busca pode ser observado na Figura 33.

Figura 33 - Protótipo da tela inicial da interface de busca



Fonte: desenvolvido pela autora

A interface permitirá a realização de cinco tarefas principais envolvendo busca e navegação, que visam auxiliar o usuário na elaboração de pesquisas relacionadas a um objetivo específico, bem como na compreensão e interpretação das relações existentes entre os documentos disponibilizados na base de dados, sendo elas:

1) Busca baseada em palavras-chave: refere-se à busca realizada a partir de um termo ou palavra-chave digitada pelo usuário, de acordo com o seu conhecimento sobre o domínio de interesse;
2) Busca baseada em estrutura terminológica: refere-se à busca realizada a partir da seleção de termos e conceitos presentes na ontologia de domínio selecionada, que são apresentados graficamente ao usuário para auxiliar na elaboração de expressões de busca mais significativas, principalmente quando o seu conhecimento é incipiente;
3) Busca baseada em perspectivas de análise de contexto: refere-se à busca realizada a partir de critérios definidos na ontologia, relacionados a um contexto específico, com o intuito de identificar documentos contendo uma ou várias características correlacionadas, tais como: artigos

pertencentes a um mesmo autor, problemas abordados em um mesmo cenário, metodologias de estudo adotadas por um determinado grupo, bases teóricas que fundamentam um cenário de pesquisa, entre outras;

4) Busca baseada em perspectivas de análise de conteúdo: refere-se à busca realizada a partir de um documento específico, com o objetivo de identificar os principais relacionamentos existentes entre este material e os demais presentes na base de dados científica. A análise pode ser elaborada a partir de perspectivas diversas, com base em características como área, autor, afiliação, palavras-chave, base teórica, metodologia, conclusão, referências, entre outros. Como consequência, facilita a identificação de materiais que tenham sido escritos pelo mesmo autor, que compartilhem as mesmas palavras-chave, que possua metodologias de pesquisa semelhantes, etc. Esta busca é elaborada a partir de toda a base de dados disponível no repositório digital, mas existe a possibilidade de ser realizada em um contexto específico (um grupo de artigos selecionados pelo autor), por meio da análise de grupos;
5) Inclusão e análise de material: refere-se à inclusão de uma publicação e sua respectiva indexação na base de dados. Pode ser realizada por um autor ou usuário, que durante o desenvolvimento desta atividade, assume a condição de produtor do conhecimento, pois classifica o conteúdo disponibilizado de acordo com o seu conhecimento e sua perspectiva pessoal, social e ambiental, contribuindo também para a organização do *corpus* e a expansão das relações entre os conceitos da ontologia.

Os documentos podem ser classificados de acordo com as diversas perspectivas de análise de um artigo científico (área, autor, afiliação, palavra-chave, objetivo, objeto de estudo, problema, base teórica, metodologia, conclusão e referência), além de outras que poderão ser acrescentadas. Cada uma dessas facetas constituirá uma classe da ontologia de domínio, contendo palavras, termos e relacionamentos definidos pelo indexador, autor ou usuário. A organização da ontologia dará suporte às buscas realizadas, bem como à interpretação de todo o *corpus* documental disponível para a pesquisa. As informações da interface, tanto no contexto do mapa de conceitos, quanto na apresentação dos resultados, poderão ser filtradas de acordo com alguns critérios: ordem alfabética, data de publicação e número de acessos.

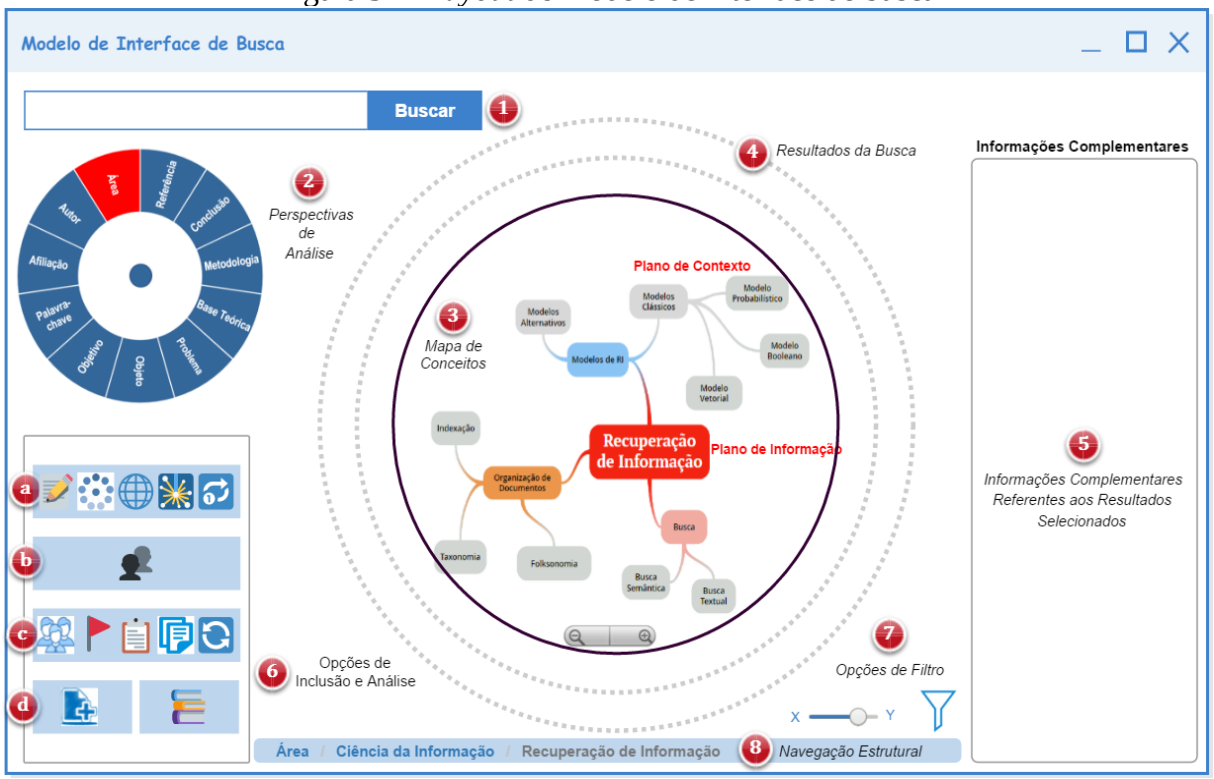
7.3 *DESIGN* DE ALTERNATIVAS E PROTÓTIPOS DA INTERFACE

O *design* de alternativas teve como objetivo documentar as ideias do projeto elaboradas a partir dos requisitos definidos na fase anterior. Entre as atividades desenvolvidas nesta etapa, destacaram-se: a) definição do *layout*, estrutura de navegação e recursos gráficos adotados na interface; b) desenvolvimento dos personagens e dos principais cenários de busca.

7.3.1 *Layout*, Estrutura de Navegação e Recursos Gráficos

O *layout* da interface de busca proposta neste estudo foi projetado para permitir a interação do usuário com o repositório digital por meio de uma estrutura de navegação híbrida, com recursos de busca (*searching*) e navegação (*browsing*), conforme apresentado na Figura 34.

Figura 34 - *Layout* do modelo de interface de busca



Fonte: desenvolvido pela autora

A interface está dividida em oito áreas. A primeira apresenta uma caixa de texto para o desenvolvimento de pesquisas baseadas em palavras-chave, com o objetivo de facilitar a busca de termos de interesse do usuário na ontologia. A segunda área disponibiliza

todas as perspectivas de análise disponíveis, a partir das quais o usuário pode explorar o *corpus* documental.

A área central da interface (área 3) é destinada ao mapa de conceitos, que contém os termos do domínio e as relações existentes entre eles. Este componente de interface foi projetado a partir de uma ontologia de domínio, com o objetivo de permitir a elaboração de pesquisas baseadas em estrutura terminológica e facilitar a entendimento sobre o que existe no *corpus* documental e o que pode ser explorado, bem como para minimizar os problemas relacionados com a dificuldade do usuário em expressar uma necessidade de informação por meio de palavras-chave, uma vez que tem potencial para ampliar o vocabulário utilizado, as possibilidades de pesquisa, os caminhos que podem ser explorados e o conhecimento acerca do assunto pesquisado. Como parte do mapa de conceitos, são apresentados os planos de contexto (contexto da análise) e de informação (foco da análise). Na parte inferior, localizam-se dois botões: *zoom out* (para aumentar o *zoom*) e *zoom in* (para diminuir o *zoom*).

A proposta da quarta área é a apresentação dos resultados da busca, tendo como principal característica a relação entre o grau de relevância dos documentos e o nível de proximidade do mapa de conceitos, ou seja, os mais relevantes, que correspondem totalmente à expressão de busca elaborada, são exibidos mais próximos, enquanto que os menos relevantes, que atendem parcialmente aos critérios estabelecidos, aparecem mais distanciados.

A quinta área apresenta detalhes sobre os resultados que forem selecionados, complementando as informações disponibilizadas na área 4, tais como: referência, resumo, notas explicativas, áreas de conhecimento, elementos de interpretação do documento (comentários, erros e falhas, parágrafos avaliados), resultados contraditórios e resultados relacionados ao mesmo contexto.

Com o objetivo de disponibilizar opções de inclusão e análise de termos e documentos, a sexta área apresenta ao usuário treze funcionalidades relacionadas a quatro contextos diferenciados:

a) termo de busca selecionado: notas explicativas, sinônimos e termos relacionados, variações linguísticas, áreas de conhecimento, antônimos e termos conceitualmente contraditórios;
b) autor selecionado: histórico do autor;
c) resultado de busca selecionado: elementos de interpretação do documento (comentários interpretativos, erros e falhas e avaliação de parágrafos), resultados relacionados ao mesmo contexto, resultados contraditórios;

d) inserção e análise de documentos: inserção de um novo documento e análise de um grupo específico de documentos.

A sétima área, por sua vez, disponibiliza filtros que podem ser utilizados na análise dos resultados recuperados, bem como a quantidade de documentos que o usuário deseja visualizar. Os filtros projetados são: data de publicação, número de comentários e número de acessos.

A última área (área oito) disponibiliza a barra de navegação estrutural (*breadcrumb navigation*), que tem como objetivo auxiliar o usuário na sua localização dentro da hierarquia do mapa de conceitos, bem como no entendimento contextual da informação que está sendo fornecida pela interface durante a realização de uma busca.

A estrutura visual, organizada a partir de uma ontologia de domínio, foi desenvolvida com base na visualização *Nó-Link* e árvore (*Node-link and tree)*, de *layout* radial, que permite que os conceitos sejam apresentados como um conjunto de nós interconectados. Cada classe da ontologia é considerada um nó e os *links* representam as relações existentes entre elas. Esta escolha foi fundamentada em Katifori et al. (2007), que afirma que este tipo de visualização facilita o entendimento das estruturas hierárquicas, bem como permite a elaboração de diferentes visões e níveis de detalhamento das informações, sendo capaz de apresentar classes, instâncias, propriedades, relações de herança e papéis.

Além da estrutura *Nó-link* e árvore, incorpora ainda recursos de visualização baseados em contexto+foco e distorção (*contexto+focus and distortion*), que facilitam a análise do contexto e o detalhamento das informações na mesma área. Desta forma, é possível obter uma visão geral da ontologia, bem como detalhes relacionados a um nó específico. Para minimizar os problemas de sobrecarga de informações, foram disponibilizados filtros de busca (área 7), bem como perspectivas a partir das quais a informação pode ser explorada (área 2).

Para facilitar a análise e o reconhecimento visual das informações apresentadas na interface, foram utilizados ainda dois outros recursos: os elementos gráficos (cor e forma) e os princípios de *Gestalt* (princípio da proximidade, da região comum e da conectividade).

Os elementos gráficos foram fundamentados em Mazza (2009) e podem ser observados em três contextos: no mapa de conceitos, nos resultados recuperados e nas perspectivas de análise. No mapa de conceitos, os recursos de cor indicam os *clusters* de assuntos, sendo cada um deles apresentado em uma cor diferenciada. Nos resultados da busca, a cor é utilizada para auxiliar o usuário a identificar a relevância dos documentos recuperados:

documentos mais relevantes são apresentados em cores mais escuras (do azul escuro para o cinza claro). Neste sentido, os resultados em azul devem ser considerados mais relevantes do que aqueles apresentados em cinza. A tonalidade das cores também deve ser observada, ou seja, dois documentos apresentados na mesma cor, mas com tonalidades diferentes (azul escuro e azul claro, por exemplo), deve-se considerar o mais escuro como sendo o mais relevante.

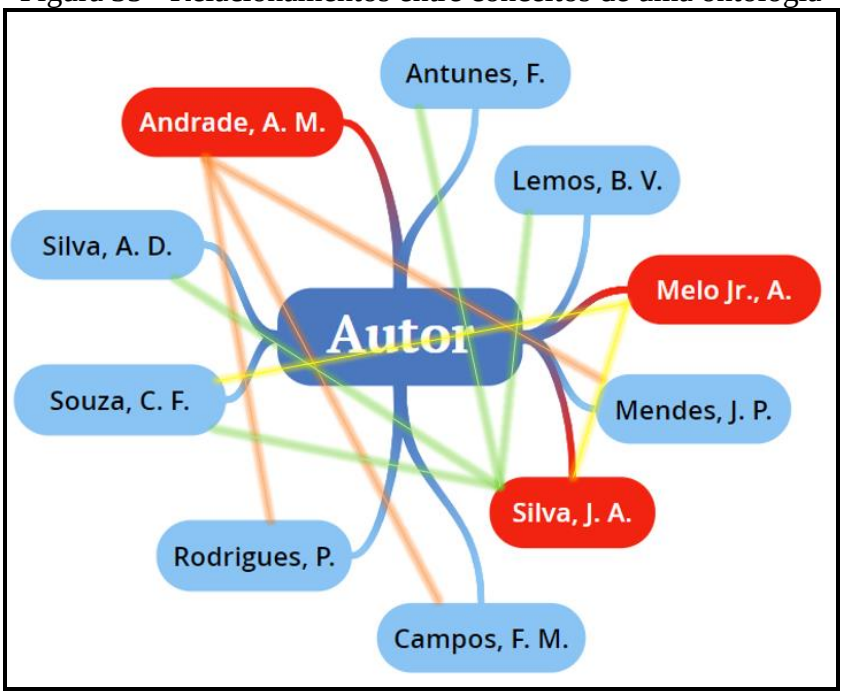
A cor vermelha indica *elementos selecionados pelo usuário* e é utilizada tanto no mapa de conceitos (ontologia) para a seleção de termos, como na análise de resultados para a seleção de documentos. Além disso, é adotada para indicar as perspectivas de análise selecionadas pelo usuário, que se diferenciam das perspectivas candidatas à seleção, que são apresentadas na cor laranja. As cores desbotadas (em cinza) no mapa de conceitos indicam o contexto dos termos selecionados, dando ao usuário a possibilidade de outras opções de busca relacionadas ao assunto pesquisado.

Ainda em relação aos elementos gráficos, o retângulo com bordas arredondadas foi o formato adotado para representar os termos pertencentes ao mapa de conceitos (ontologia de domínio).

Os princípios de *Gestalt*, baseados em Koshman (2006) e Ware (2003), foram adotados também para facilitar a identificação dos *clusters* de conceitos (na ontologia) e de resultados (nos resultados da busca). Desta forma, os elementos que são apresentados próximos uns dos outros, devem ser interpretados como um grupo (princípio da proximidade e da região comum), assim como os elementos conectados devem ser entendidos como elementos relacionados (princípio da conectividade). Em relação aos resultados, os documentos mais próximos do círculo devem ser considerados os mais relevantes, pois possuem um maior número de relações de similaridade com a expressão de busca.

A similaridade existente entre os documentos é um recurso projetado para permitir que o usuário navegue pelos documentos recuperados, com a ideia de que aqueles que compartilham inúmeros pontos de similaridade com um documento de interesse, possuem maior probabilidade de serem também relevantes do que aqueles que são diferentes. A análise da similaridade foi projetada a partir dos conceitos e relacionamentos presentes na ontologia de domínio, bem como da indexação semântica do *corpus* documental. A Figura 35 apresenta a rede de relacionamentos existente entre conceitos e instâncias de uma ontologia de domínio.

Figura 35 - Relacionamentos entre conceitos de uma ontologia



Fonte: desenvolvido pela autora

Os termos da ontologia são conectados uns aos outros por meio dos documentos presentes no *corpus*. Desta forma, ao observar o mapa de conceitos, é possível deduzir que Andrade, A. M. está associado aos autores Mendes, J. P., Campos, F. M. e Rodrigues, P. Entre as principais contribuições, este recurso permite identificar *clusters* de documentos e extrair deles um conjunto de informações; encontrar resultados correlacionados, que o usuário não sabia que existiam; revelar conhecimento implícito por meio da análise das relações indiretas e, por fim, explorar a estrutura dos documentos através de múltiplas perspectivas de análise.

A hermenêutica, fundamentada principalmente em Gadamer (1999), subsidiou o desenvolvimento de funcionalidades que se referem ao processo de interpretação, relacionados à elaboração das expressões de busca e à análise de resultados retornados. Em relação à elaboração das expressões de busca, a interface apresenta recursos projetados para oferecer conhecimento adicional sobre termos selecionados, todos disponibilizados nas opções de inclusão e análise (item a):

a) notas explicativas, sinônimos e termos relacionados, variações linguísticas: funcionalidades desenvolvidas com base no princípio do compartilhamento de linguagem;
b) área de conhecimento de um termo: funcionalidade projetada com base no princípio da abstração e generalização;

c) antônimos e termos conceitualmente contraditórios: funcionalidade desenvolvida com base no princípio do raciocínio dialético;

Outro recurso importante projetado a partir dos princípios da hermenêutica é o mapa de conceitos, que exibe a ontologia de domínio disponível para a busca. Foi inspirado em Schleiermacher (2005), que afirma que o sentido de cada palavra num dado local precisa ser determinado através de sua conexão com as palavras ao seu redor. Desta forma, a intenção foi projetar uma interface que tivesse condições de apresentar ao usuário os conceitos e seus respectivos contextos pertencentes ao domínio de uma busca.

Para facilitar o entendimento e a compreensão do usuário sobre o assunto explorado, foram projetados dois ambientes: o plano de informação e o plano de contexto. O plano de informação destaca o foco do usuário, apresentando conceitos e nós relacionados que estão sendo investigados, seja por meio da navegação ou da seleção. O plano de contexto revela os nós e conceitos que fazem parte do plano de informação.

O mapa de conceitos, através dos planos de informação e contexto, também possui fundamento no princípio do círculo hermenêutico, que é base do processo de interpretação. Para Schleiermacher (2005), pode-se compreender o significado de um texto através das palavras, que são sinais individuais determinados apenas relativamente. O círculo hermenêutico é o contexto no qual o usuário obtém novas ideias ou novos aprendizados sobre a informação que está pesquisando.

Em uma tarefa de busca e recuperação de informação, o pesquisador se depara com classes e instâncias que alteram constantemente o seu interesse, como se estivesse em um ciclo. Este processo favorece a interação entre as classes e instâncias, através de um recurso de *vai* e *vem*, que oferece ao usuário a compreensão do contexto pesquisado. O conhecimento do pesquisador sobre o domínio de uma busca ou sobre um determinado documento vai aumentando à medida em que a rede de conceitos relacionada a ele é explorada.

Além do círculo hermenêutico, o mapa de conceitos também está subsidiado no princípio do compartilhamento de linguagem, uma vez que os recursos das ontologias oferecem mecanismos para minimizar os problemas relacionados com a linguagem que interferem na elaboração das expressões de busca, referentes à ambiguidade, sinonímia, homonímia, paronímia, conotação e variações linguísticas. Também auxiliam os usuários a interpretar o significado das palavras, as regras gramaticais e os relacionamentos semânticos que são a base para o processo de interpretação documental.

O princípio da interação entre sujeitos e das múltiplas interpretações

fundamenta o processo de inclusão e análise de materiais na base de dados (opção disponível na Figura 34, no item d). O processo de interpretação não deve ser considerado como uma atividade meramente reprodutiva, como se tudo fosse observado de fora. A compreensão está relacionada com os preconceitos de cada intérprete, que dizem respeito a tudo aquilo que ele sabe, de forma consciente ou inconsciente. Portanto, cada intérprete compreende um texto a partir do seu horizonte e do seu mundo. Em função disto, não existe uma verdade única e absoluta, podendo existir inúmeras interpretações sobre o mesmo documento. As múltiplas interpretações resultantes da interação entre os diversos pesquisadores que utilizam a base de dados podem contribuir para uma melhor compreensão do usuário, uma vez que revela diferentes aspectos relacionados a um contexto avaliado.

Por outro lado, o processo de análise de resultados também está fundamentado nos princípios da hermenêutica. Neste sentido, alguns recursos foram projetados para oferecer ao usuário informações adicionais ou sugestão de pesquisas complementares relacionadas a resultados de busca selecionados (Figura 34, itens b e c), sendo eles:

a) histórico do autor: funcionalidade desenvolvida com base no princípio da contextualização;
b) comentários interpretativos: funcionalidade desenvolvida com base no princípio da interação entre pesquisadores e sujeitos;
c) indicação de erros e falhas e avaliação de parágrafos: funcionalidades desenvolvidas com base no princípio da suspeição;
d) resultados relacionados ao mesmo contexto: funcionalidade projetada com base no princípio das múltiplas interpretações;
e) resultados contraditórios: funcionalidade desenvolvida com base no princípio do raciocínio dialética.

Os recursos de interface projetados com base nos princípios da hermenêutica, relacionados tanto à elaboração das expressões de busca, quanto à análise de resultados retornados podem ser utilizados isoladamente ou em conjunto, sendo que a escolha por um, pelo outro ou por vários deles, depende dos objetivos e do nível de detalhamento do pesquisador.

Neste contexto, a partir da definição do *layout*, da estrutura de navegação e dos recursos gráficos utilizados, foram projetados cinco cenários que descrevem as tarefas de busca e as interações realizadas na interface. São independentes um do outro e não indicam sequência obrigatória de execução de atividades.

7.3.2 Personagens e Cenários

O modelo de interface apresentado neste estudo propõe quatro cenários de busca que descrevem as necessidades de informação e as principais atividades desenvolvidas por um usuário típico de uma base de dados, neste caso, representado pela Fernanda, que é uma pesquisadora da área de Letras, que tem a intenção de compreender e interpretar um novo domínio de conhecimento.

Também foi elaborado um quinto cenário, representado pelo ator Luís Antônio, que é um pesquisador experiente na área de Ciência da Informação e, portanto, tem condições de propor relacionamentos para a ontologia, indicando a associação entre conceitos (conceitos entre si) e entre documentos (conceitos e documentos), com o objetivo de ampliar a rede de informações entre os documentos e compartilhar conhecimento sobre o *corpus*.

7.3.2.1 Cenário de busca baseada em palavras-chave

O primeiro cenário deste estudo apresenta o contexto de desenvolvimento de buscas baseadas em palavras-chaves digitadas pelo usuário, de acordo com os conhecimentos que possui acerca do domínio pesquisado e a sua necessidade de informação. Podem ser realizadas com o auxílio das informações apresentadas no mapa de conceitos, além da possibilidade de serem expandidas por meio das relações semânticas presentes na ontologia.

Quadro 11 - Busca baseada em palavras-chave

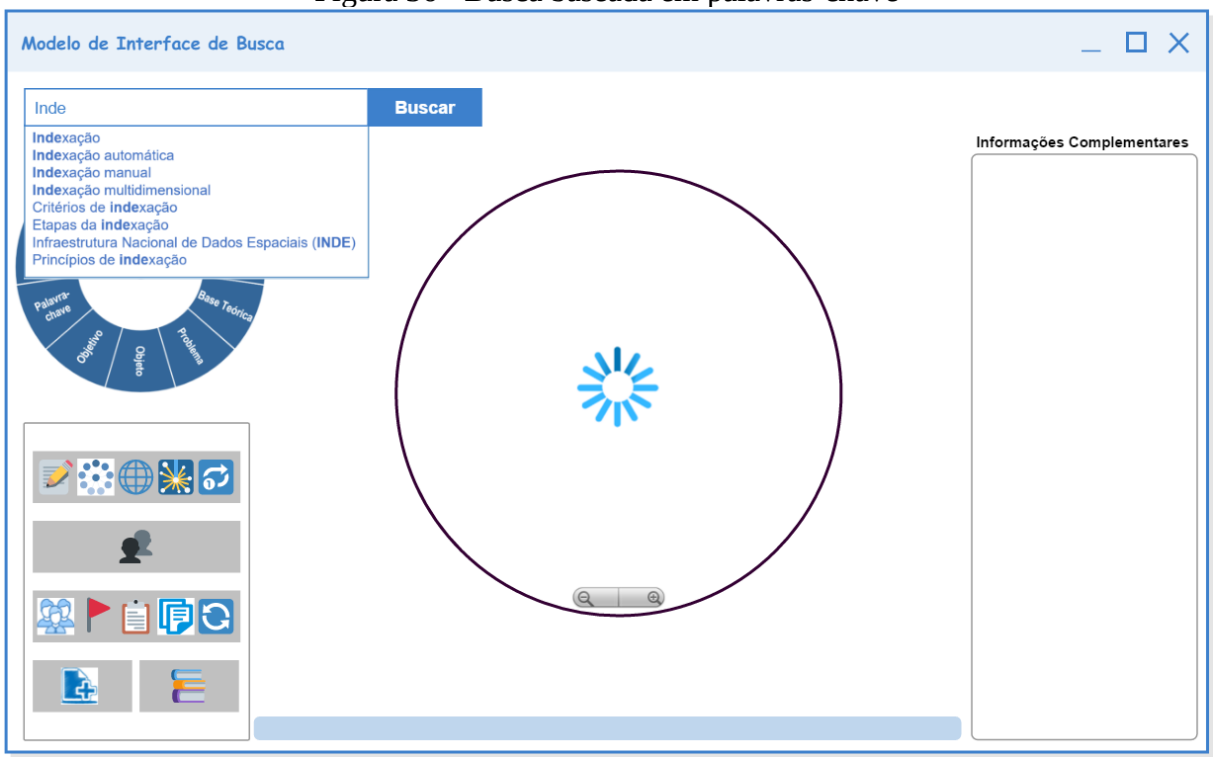
Primeiro Cenário
Fernanda é formada em Letras e atua como professora em uma escola de línguas na cidade de Londrina. Como tem a intenção de aprimorar seus conhecimentos e trabalhar futuramente no ensino superior, está pretendendo elaborar um projeto de pesquisa para concorrer à uma vaga no Mestrado em Ciência da Informação. Por ser uma área desconhecida, ela julga ser importante realizar buscas exploratórias sobre o domínio e, posteriormente, com informações mais precisas, decidir qual assunto deve investigar. Para que isso seja possível, Fernanda quer identificar alguns conceitos importantes da Ciência da Informação e realizar algumas leituras sobre o assunto. O primeiro passo é acessar o repositório digital e utilizar o recurso de pesquisa baseada em palavras-chave. Existe a possibilidade de digitar termos simples ou compostos, que serão localizados na ontologia e apresentados no mapa de conceitos, como forma de indicar ao usuário outras informações correlacionadas. É possível ainda escolher as perspectivas da busca que se deseja realizar, que representam a área, autor, problema, base teórica, metodologia etc. A interface apresenta quatro recursos importantes que auxiliam na elaboração das expressões de busca: recurso de autopreenchimento, interface visual para análise das perspectivas disponíveis, conhecimento adicional para expansão da busca (sinônimos e termos relacionados, variação linguística, antônimos e termos conceitualmente contraditórios) e ambiente gráfico para elaboração de consultas semânticas.

Fonte: desenvolvido pela autora

O cenário refere-se a uma situação típica de um usuário que tem como objetivo explorar um domínio de conhecimento novo e identificar informações importantes sobre ele. Isso acontece em diversas situações do cotidiano acadêmico, organizacional, social e pessoal, tais como: interesse por uma nova área de estudo, mudança de cargo, desenvolvimento de um novo projeto, necessidade de informação para uma tomada de decisão, interesse por um assunto desconhecido e outras tantas circunstâncias. A Figura 36 apresenta o protótipo da interface referente a ele.

Neste caso, Fernanda, usuária fictícia deste cenário, acessa a base de dados científica e realiza diversas pesquisas exploratórias, utilizando termos e palavras-chave de acordo com a sua necessidade de informação. Neste ponto, o processo de elaboração da expressão que será utilizada para a consulta é similar ao do Google, que utiliza um retângulo de busca para permitir a interação do usuário com o sistema.

Figura 36 - Busca baseada em palavras-chave



Fonte: desenvolvido pela autora

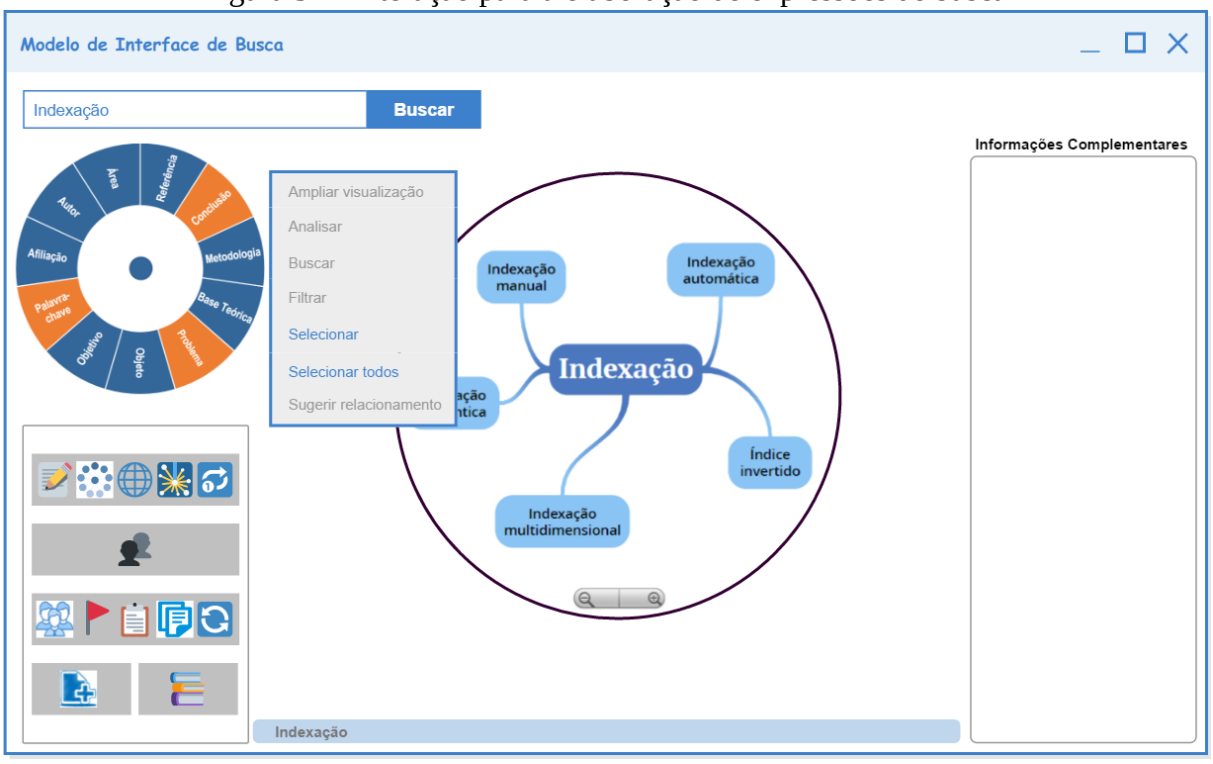
Neste cenário, o recurso de **preenchimento automático** ou **sugestão automática** é de grande importância para auxiliar o pesquisador a elaborar melhores expressões de busca, bem como para apresentar os termos disponíveis na ontologia de domínio que são semelhantes ao que está sendo digitado, contendo ou não o mesmo radical. Consequentemente, ao inserir caracteres no retângulo de busca, sugestões de termos são

oferecidas dinamicamente em uma lista suspensa, permitindo a seleção de uma opção que esteja de acordo com a necessidade de informação. Estes recursos são importantes para o contexto da busca, pois auxiliam os usuários a encontrarem palavras semelhantes aos termos que estão sendo digitados, bem como minimizam erros de digitação.

O recurso de preenchimento automático está presente em diversos buscadores disponíveis na Web, tais como o *Google Suggest*, o *Yahoo*! *Search Assist*, o *Ask.com* e o *Bing* além de outros serviços de busca, como o *Amazon.com* e o *Netflix*.

Para cada termo digitado, a ferramenta apresenta as perspectivas de análise associadas a eles. No caso do exemplo apresentado na Figura 37, o usuário teve a intenção de realizar uma busca usando o termo **Indexação**, que foi encontrado em três contextos diferentes: Palavras-chave, Problema e Conclusão. As perspectivas candidatas à seleção do usuário são destacadas por meio da cor laranja, sendo possível a escolha de uma ou várias perspectivas ao mesmo tempo.

Figura 37 - Interação para a elaboração de expressões de busca



Fonte: desenvolvido pela autora

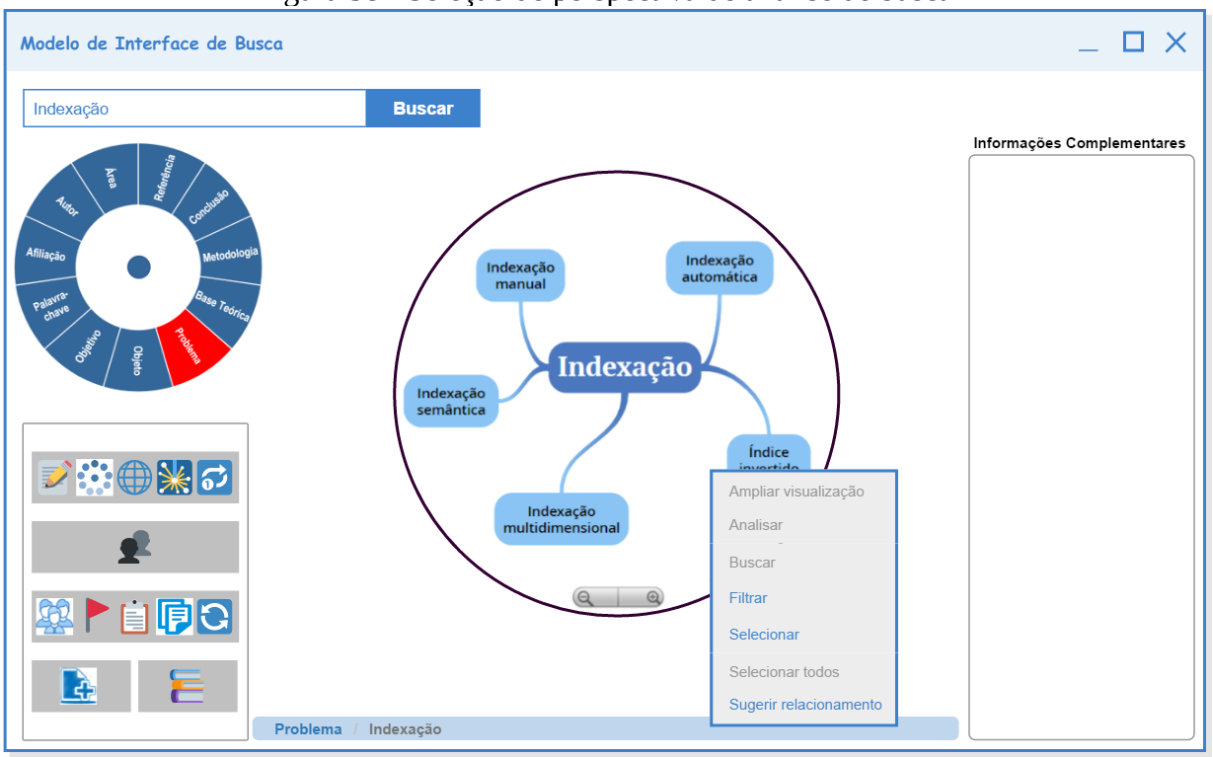
Após a seleção das perspectivas de análise, que pode ser realizada a partir do menu suspenso (opção Selecionar ou Selecionar todos), elas passam a ser apresentadas na cor vermelha, indicando visualmente as opções de busca do usuário. À medida em que as escolhas são feitas, o mapa de conceitos vai sendo atualizado para facilitar a compreensão do

usuário em relação ao domínio que pode ser explorado.

A Figura 38 mostra a seleção da perspectiva Problema. Neste caso, a pesquisa retornará documentos que tenham sido classificados na base de dados a partir do conceito **Indexação**, sendo que os mais relevantes, que são aqueles que correspondem totalmente à expressão de busca digitada pelo usuário (termo Indexação na perspectiva Problema) serão apresentados mais próximos do mapa de conceitos do que os resultados menos relevantes, que são aqueles que possuem o termo Indexação em outras perspectivas.

Neste caso, o menu suspenso oferece três possibilidades de interação: filtrar, selecionar e sugerir relacionamento. A primeira opção se refere ao mapa de conceitos e permite que o usuário selecione as informações que deseja visualizar (por conceito ou por classe). A segunda está relacionada à seleção de um termo para realizar uma busca e a terceira, por sua vez, deverá ser utilizada quando o usuário tiver interesse em sugerir novos termos e relacionamentos para ampliar ou atualizar a ontologia de domínio. Este recurso deve ser acessado por meio de um *click* com o botão direito do *mouse*.

Figura 38 - Seleção de perspectiva de análise de busca

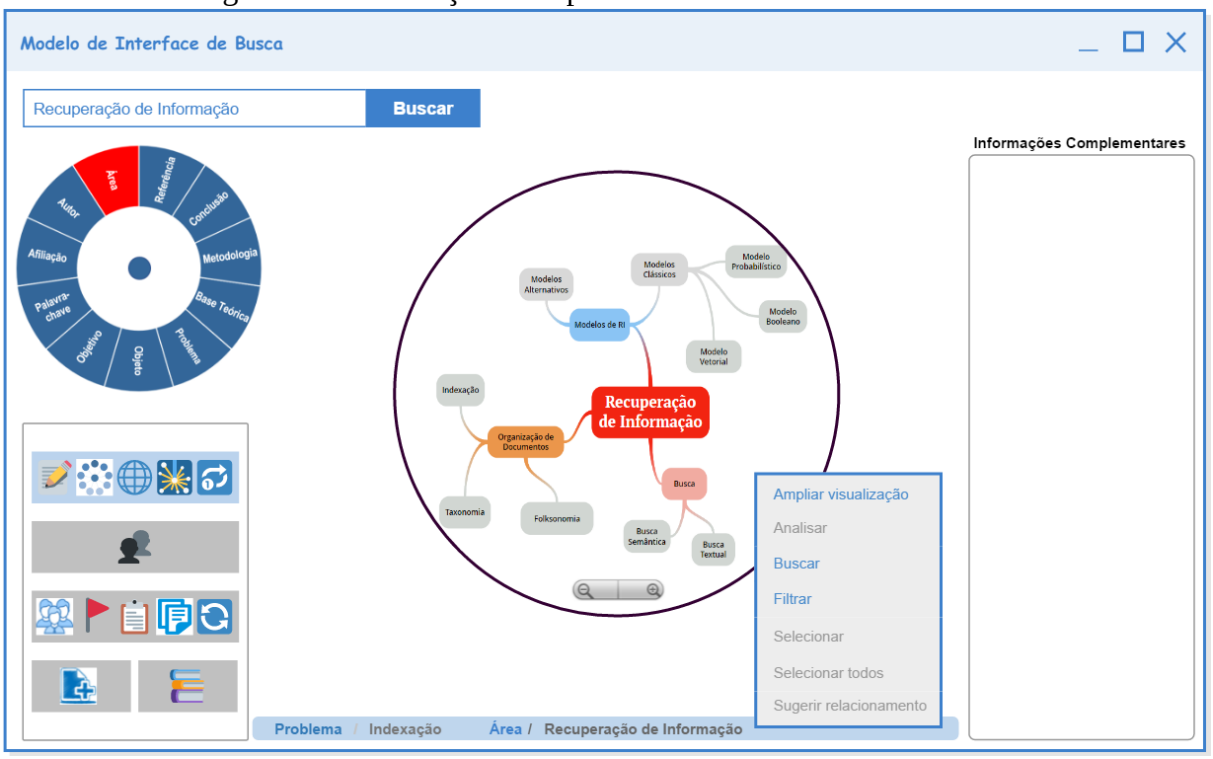


Fonte: desenvolvido pela autora

Existe a possibilidade ainda de adicionar novos termos à pesquisa, como forma de refinar os resultados obtidos. Para que isto seja possível, basta adicionar um novo termo na caixa de busca e repetir todo o processo.

A Figura 39 apresenta a inserção do termo **Recuperação de Informação**, que foi encontrado em quatro outras perspectivas: Área, Base Teórica, Palavra-chave e Conclusão, sendo selecionada apenas a **Área**. Neste caso, a busca será realizada com base nos termos **Indexação** (perspectiva Problema) e **Recuperação de Informação** (perspectiva Área), ou seja, a ferramenta recuperará todos os documentos que tiverem sido classificados por Indexação no contexto do Problema de pesquisa e, ao mesmo tempo, tiverem relação com o termo Recuperação de Informação, na perspectiva Área.

Figura 39 - Elaboração de expressões de busca com dois termos



Fonte: desenvolvido pela autora

Como já existem termos selecionados no mapa de conceitos, as opções habilitadas do menu suspenso foram alteradas: ampliar visualização, buscar e filtrar. A primeira permite a ampliação do espaço para a visualização do mapa de conceitos (opção disponível apenas para mapas contendo uma grande quantidade de conceitos). A segunda deve ser selecionada para realizar uma busca a partir dos termos selecionados e a terceira, caso o usuário queira utilizar um filtro e selecionar apenas alguns conceitos ou classes para visualizar no mapa de conceitos.

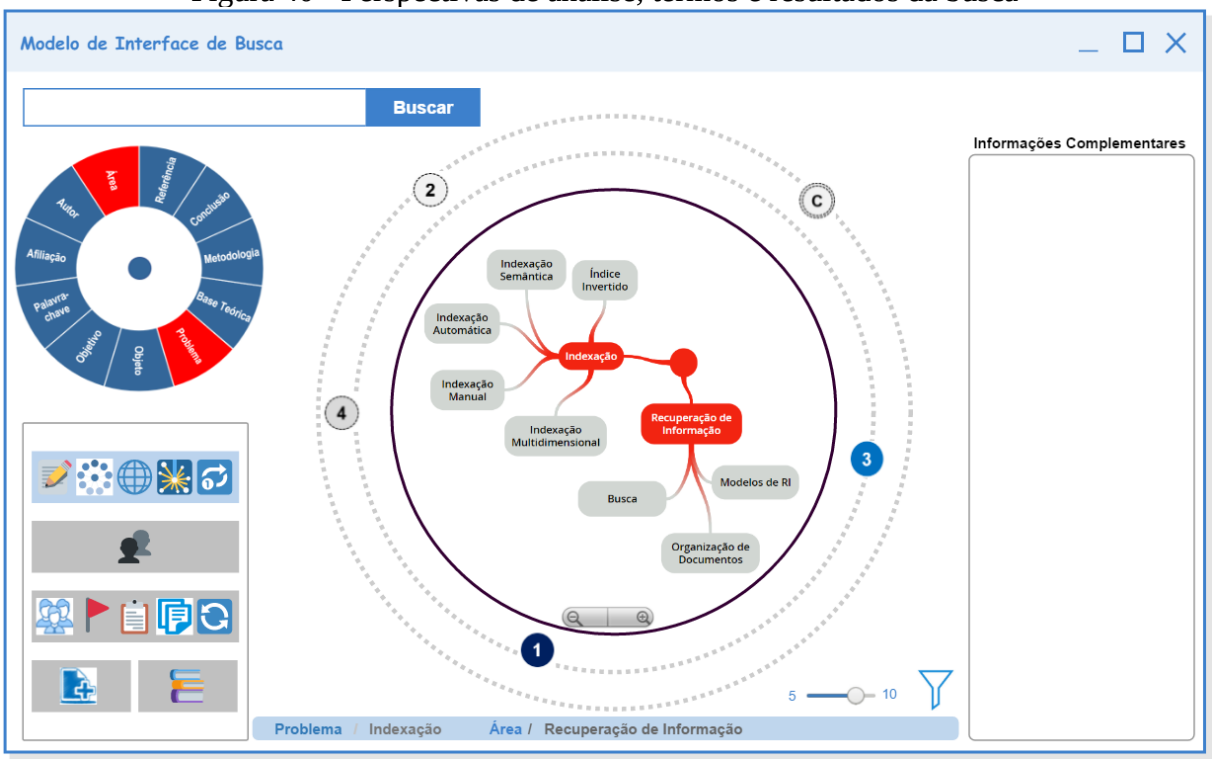
A barra de navegação estrutural (*breadcrumb navigation*), localizada na parte inferior da interface, indica ao usuário as perspectivas e os termos de busca selecionados, bem como o caminho percorrido até o momento (todo e parte). Este recurso está

fundamentado no princípio do Círculo Hermenêutico e visa potencializar o movimento interpretativo, pois fornece um melhor entendimento do contexto que está sendo explorado, contribuindo para o processo de análise e síntese realizado constantemente pelo pesquisador durante a interpretação das informações apresentadas pela interface.

Os resultados da busca são apresentados em torno do mapa de conceitos, no formato de um pequeno círculo, permitindo a identificação visual dos documentos relacionados com cada critério de busca. Os itens em destaque indicam o foco da análise e da atenção do usuário, enquanto os não realçados apresentam o contexto associado a ele. A pesquisa pode ser alterada dinamicamente, sendo que qualquer mudança realizada nos critérios de busca (termos ou perspectivas de análise) interfere diretamente nos planos de informação e de contexto apresentados ao usuário.

Após a elaboração da expressão de busca, o usuário deve utilizar o menu suspenso e escolher a opção **Buscar**, para que os documentos recuperados sejam, então, apresentados. Na Figura 40 podem ser observadas as perspectivas de análise selecionadas, os termos de busca e os resultados obtidos.

Figura 40 - Perspectivas de análise, termos e resultados da busca



Fonte: desenvolvido pela autora

Os resultados da busca realizada são representados por quatro documentos e por um *cluster*, que indica um conjunto de elementos indexados a partir dos mesmos critérios

e igualmente relevantes. Para facilitar a análise do cálculo de relevância, bem como o entendimento da apresentação dos resultados exibidos ao usuário, o Quadro 12 fornece as informações referentes à indexação de todos os documentos recuperados.

Quadro 12 - Indexação dos documentos recuperados

Expressão de Busca: Indexação "Problema" e Recuperação de Informação "Área"										
Perspectivas	**Doc. 1**		**Doc. 2**		**Doc. 3**		**Doc. 4**		***Cluster***	
	Indexaçao	RI	Indexação	RI	Indexação	RI	Indexação	RI	Indexação	RI
Área		**1**				**1**		1		
Autor										
Afiliação										
Palavra-chave	1	1			1					
Objetivo								1		1
Objeto										
Problema	**1**		1		**1**					
Base Teórica		1				1		1		
Metodologia										
Conclusão		1	1					1		1
Referência										
Relacionamentos	6		2		4		4		2	
Data Publicação	20 nov. 2018		05 out. 2017		02 jun. 2014		06 abr. 2018		-	
Num. de Acessos	75		32		44		38		Média = 37	

Fonte: desenvolvido pela autora

Os Docs. 1 e 3 foram considerados relevantes por atenderem mutuamente os dois critérios de busca estabelecidos pelo usuário, ou seja, ambos foram classificados a partir dos termos Indexação, na perspectiva Problema e Recuperação de Informação, na perspectiva Área. Embora os dois sejam relevantes, o Doc. 1 foi apresentado mais próximo do mapa de conceitos e em azul mais escuro por possuir mais contextos relacionados à expressão de busca realizada, indicando uma possibilidade de maior relevância.

Os Docs. 2, 4 e o *cluster* atendem parcialmente à consulta realizada. O Doc. 2 possui dois relacionamentos com o termo Indexação (no contexto do Problema e no contexto da Conclusão). O Doc. 4, por sua vez, possui quatro pontos de conexão com o termo Recuperação de Informação, enquanto que o *cluster* possui dois. Por este motivo, o Doc. 4 está mais próximo do mapa de conceitos e é apresentado em tonalidade de cinza mais escuro do que o *cluster*, indicando, portanto, que o documento pode ser mais relevante.

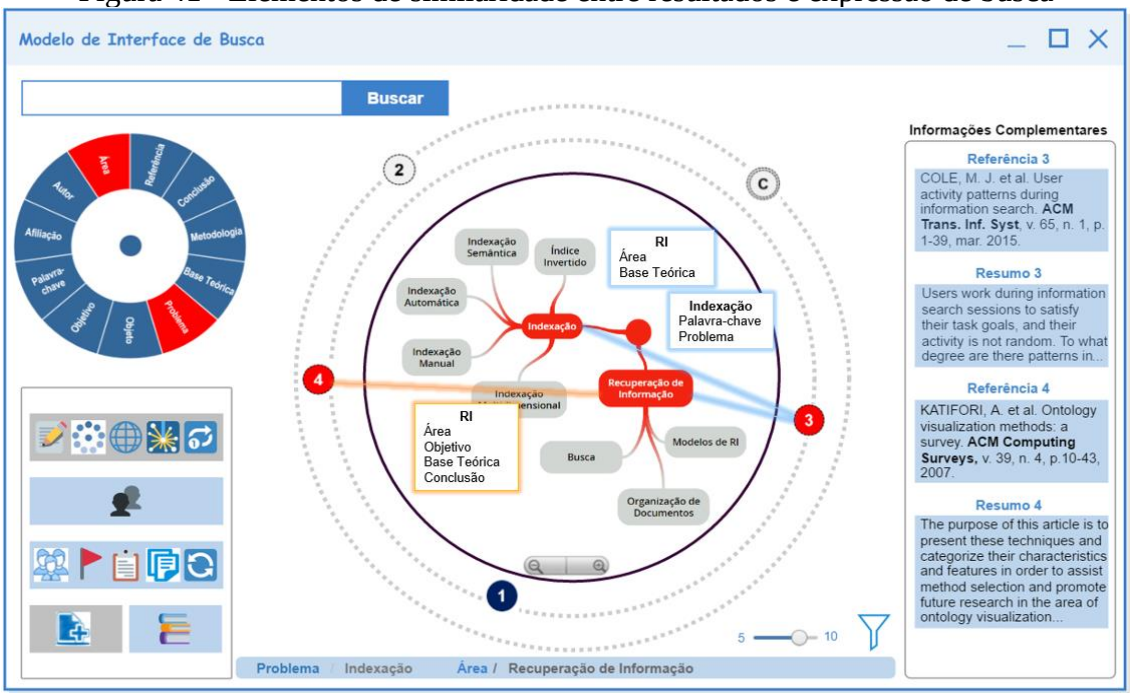
Outros dois critérios considerados para calcular a relevância de um documento e, consequentemente, sua proximidade com o mapa de conceitos e a sua tonalidade de cor, são a data de publicação do documento e o seu número de acessos. Desta forma, caso o número de relacionamentos seja o mesmo entre dois documentos (ou mais), aquele que tiver data de publicação mais recente será considerado mais relevante. Caso possuam a mesma data de publicação, a quantidade de acessos determinará, então, o seu grau de relevância.

Para analisar apenas resultados dentro de um contexto de interesse (os mais recentes, por exemplo), o usuário tem a possibilidade de utilizar filtros para selecionar apenas uma parte dos documentos recuperados, de acordo com os seguintes critérios: data de publicação, número de comentários (interpretativos) e número de acessos. Também pode definir a quantidade de documentos que deseja visualizar, com a intenção de reduzir a carga de informação disponível da interface. Os resultados menos relevantes são os primeiros a serem retirados do campo de visualização.

Ao selecionar um ou vários resultados retornados pela busca, eles ganham foco e são apresentados em vermelho. A referência e o resumo de cada um deles são exibidos no campo de **Informações Complementares**, dando ao usuário mais elementos para julgar a sua relevância. Com o movimento do ponteiro do mouse sobre o resultado selecionado, a interface revela os pontos de conexão entre ele e a expressão de busca, indicando os relacionamentos existentes.

No caso do exemplo da Figura 41, o Doc. 4 foi selecionado pelo usuário, sendo possível analisar os elementos que indicam a sua relação com o termo de busca Recuperação de Informação: possui quatro pontos de conexão nas perspectivas Área, Objetivo, Base Teórica e Conclusão. O Doc. 3, por sua vez, possui relação com os dois termos utilizados na busca: Indexação e Recuperação de Informação. Em relação ao termo Indexação, os pontos de conexão ocorrem nas perspectivas Palavra-chave e Problema. No contexto do termo Recuperação de Informação, os relacionamentos ocorrem em Área e Base Teórica.

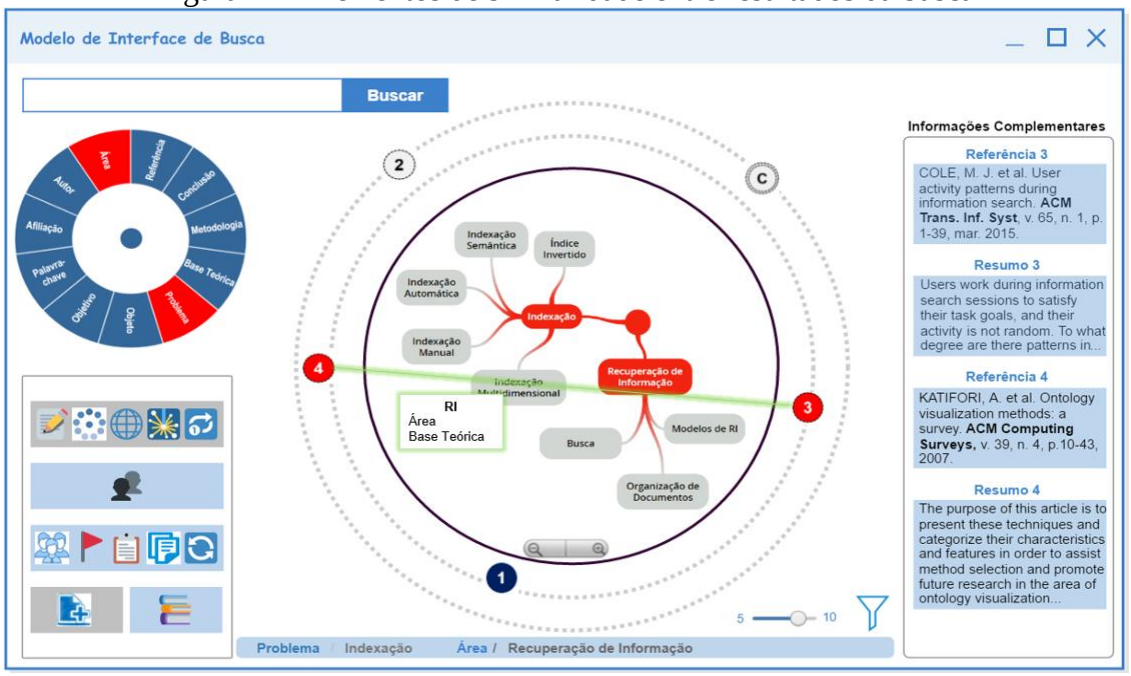
Figura 41 - Elementos de similaridade entre resultados e expressão de busca



Fonte: desenvolvido pela autora

Caso mais de um documento seja selecionado, conforme apresentado na Figura 42, a interface realiza uma análise do grupo e apresenta os pontos de similaridade entre eles em relação aos termos de busca. Desta forma, se os Docs. 3 e 4 fossem selecionados, o usuário seria informado de que ambos documentos são semelhantes entre si em relação ao termo Recuperação de Informação, no contexto das perspectivas Área e Base Teórica.

Figura 42 - Elementos de similaridade entre resultados da busca

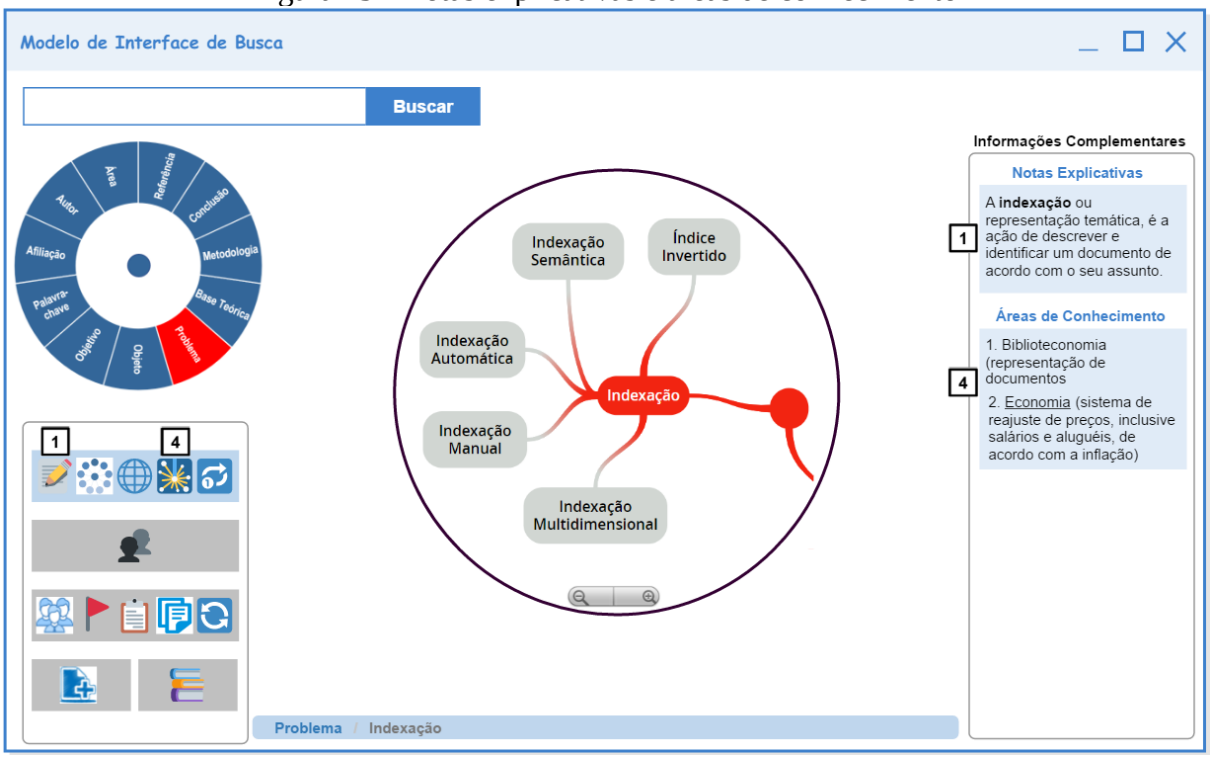


Fonte: desenvolvido pela autora

Caso exista alguma alteração nos termos de busca, os documentos serão automaticamente reorganizados, de forma a atender os novos critérios estabelecidos. Assim, se o usuário cancelasse a seleção do termo Indexação, imediatamente a perspectiva Problema perderia o foco de evidência, sendo apresentada novamente na cor azul. Neste caso, os Docs. 1, 3, 4 e *cluster* seriam apresentados como documentos relevantes, sendo organizados de acordo com a quantidade de relacionamentos com o termo Recuperação de Informação (Docs. 1 e 4 mais próximos do mapa de conceitos). O Doc. 2 não seria recuperado, pois não possuiria nenhum ponto de conexão com o novo termo de busca.

Assim como a função de autopreenchimento, que é um mecanismo que visa melhorar a consulta inicialmente formulada pelo usuário, a interface apresenta ainda cinco outros recursos que fornecem ao usuário conhecimento adicional referente aos termos de busca selecionados, sendo eles: notas explicativas, sinônimos e termos relacionados, variações linguísticas, áreas de conhecimento, antônimos e termos conceitualmente contraditórios. A Figura 43 traz um exemplo do uso das notas explicativas e das áreas de conhecimento relacionadas ao termo **Indexação**, bem como destaca os ícones da interface que representam as respectivas funcionalidades (itens 1 e 4).

Figura 43 - Notas explicativas e áreas de conhecimento



Fonte: desenvolvido pela autora

As **notas explicativas** apresentam ao usuário conceitos e definições

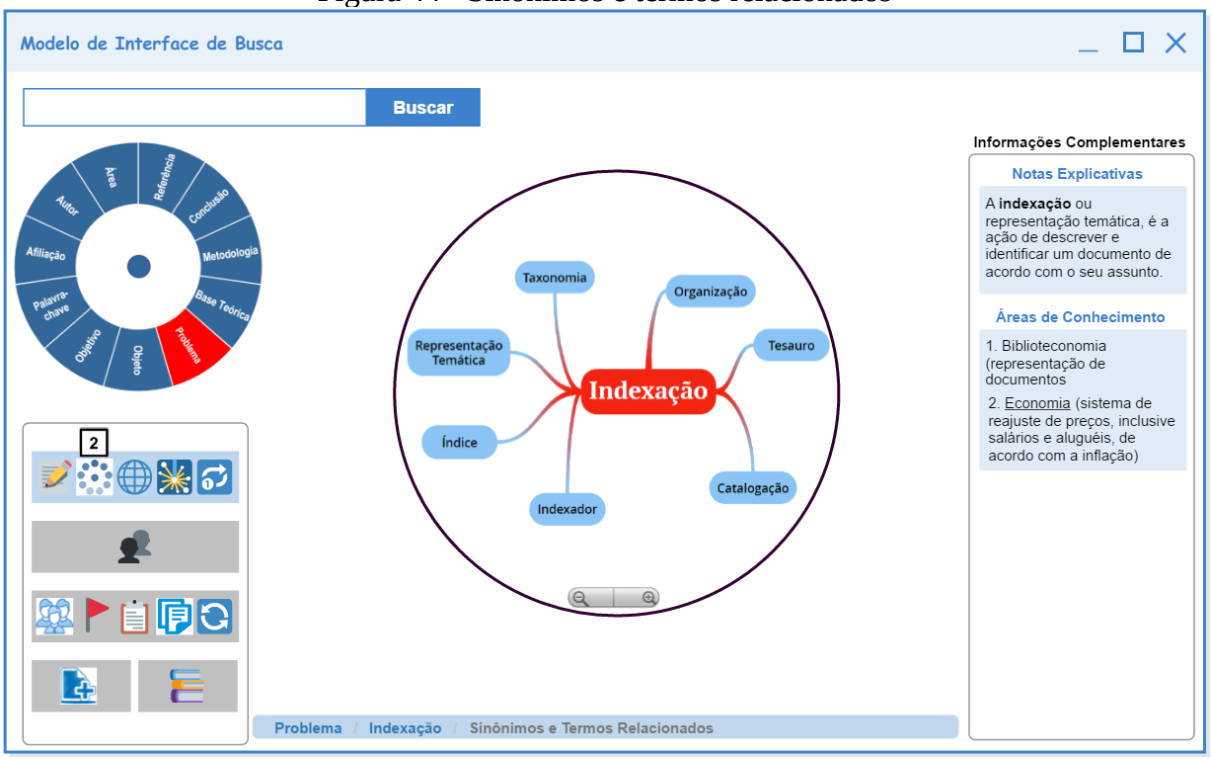
relacionados a um termo de busca. Este recurso é fundamentado no princípio hermenêutico do Compartilhamento de Linguagem, que tem como principal objetivo facilitar o entendimento do intérprete em relação ao vocabulário utilizado pelo autor de um determinado texto. Neste sentido, com um fundamento pedagógico, é uma das funções desenvolvidas para potencializar a apropriação da linguagem e auxiliar na interpretação do significado das palavras, das regras gramaticais e dos relacionamentos semânticos. O acesso à funcionalidade só é disponibilizado após a seleção de um termo no mapa de conceitos. O número 1 indica respectivamente, acesso e resultado apresentado.

Assim como as notas explicativas, a função **áreas de conhecimento** representada pelo número 4, tem como objetivo apresentar ao usuário conhecimento adicional sobre um termo de busca selecionado. Esta funcionalidade foi desenvolvida com base no princípio hermenêutico da Abstração e Generalização, que afirma a possibilidade de aplicação de ideias e conceitos em múltiplos contextos, ampliando a aplicação de um determinado conhecimento em campos e cenários diferenciados. Nas interfaces de busca atuais, esta possibilidade gera problemas relacionados com a polissemia (vários significados para o mesmo termo) e contribui para a recuperação de documentos irrelevantes e fora do contexto de interesse do usuário.

Outro recurso importante para oferecer conhecimento adicional relacionado aos termos de busca refere-se aos **sinônimos e termos relacionados**. Tal como as notas explicativas, é uma funcionalidade projetada a partir do princípio hermenêutico do Compartilhamento de Linguagem e tem como objetivo disponibilizar estratégias pedagógicas para ampliar o conhecimento do usuário e estimular seu aprendizado sobre o domínio pesquisado, uma vez que apresenta todos os relacionamentos envolvidos na consulta. A função será habilitada a partir da seleção de um termo no mapa de conceitos e deve ser acessada por meio do ícone número 2, conforme pode ser observado na Figura 44.

O recurso de sinônimos e termos relacionados, além de contribuir para compreensão do usuário sobre o assunto pesquisado, fornece ainda a possibilidade de expansão semântica dos resultados, pois permite que novos termos sejam agregados à busca realizada, melhorando a precisão dos resultados obtidos. As ontologias são úteis neste processo, uma vez que possuem um conjunto de conceitos interligados por relações semânticas, que possibilitam a aplicação de cálculos de proximidade entre um termo e outro. Os conceitos mais próximos da pesquisa inicialmente desenvolvida podem ser escolhidos para ampliar a busca e aproximar os resultados da intenção do usuário.

Figura 44 - Sinônimos e termos relacionados

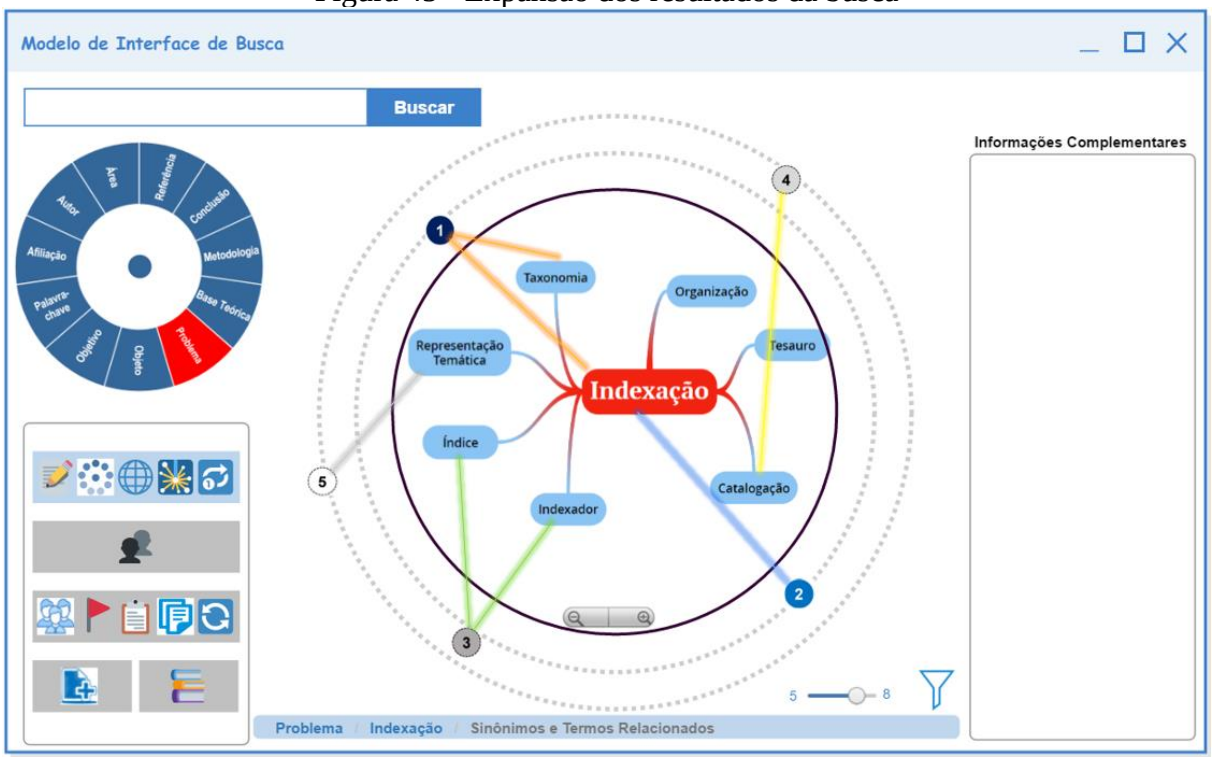


Fonte: desenvolvido pela autora

No exemplo apresentado na Figura 45, pode-se observar, de forma ampliada, os conceitos da ontologia de domínio relacionados ao termo **Indexação**, que é o nó central e corresponde à busca realizada pelo usuário. Ao apresentar um conjunto de conceitos semanticamente relacionados ao termo de busca, aumenta-se a probabilidade de serem encontrados documentos relevantes pertencentes ao mesmo contexto e, consequentemente, do usuário encontrar documentos que atendam suas necessidades de informação.

Isto significa que um usuário interessado em documentos que discutam Indexação, também pode ter interesse em artigos relacionados à Representação Temática, Taxonomia e Catalogação, por exemplo. Em função da sobrecarga de informação no mapa de conceitos, a distância de similaridade padrão para a expansão semântica é de apenas um nó, que é calculado entre o nó principal (termo digitado pelo usuário) e os termos associados a ele. No entanto, é possível configurar outras distâncias, caso seja de interesse do usuário.

Figura 45 - Expansão dos resultados da busca

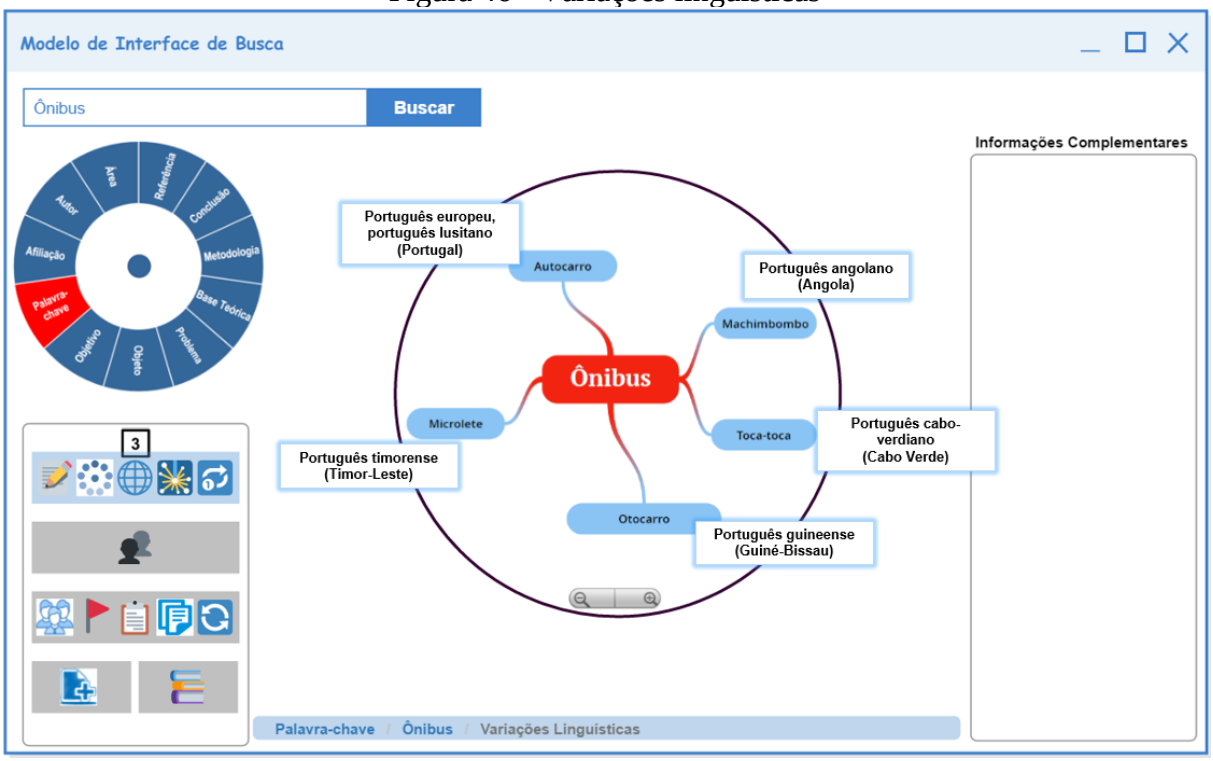


Fonte: desenvolvido pela autora

Este recurso é importante, pois cada autor ou grupo de autores, durante a escrita de um artigo, utilizará palavras que correspondam aos seus interesses, conhecimentos, áreas de estudo e contextos envolvidos, sendo que muitas vezes, embora distintas quanto a sua grafia, possuem o mesmo significado. Isto quer dizer que autores diferentes, possuem vocabulários distintos um do outro e, muitas vezes, documentos relevantes deixam de ser recuperados por terem sido indexados por termos alternativos, embora pertencentes ao mesmo contexto semântico. Esta opção permite, portanto, a recuperação de todos os documentos relacionados direta ou indiretamente, ao termo de busca.

No mesmo contexto, também fundamentado no princípio hermenêutico do Compartilhamento de Linguagem, a funcionalidade relacionada à apresentação de **variações linguísticas** tem como principal objetivo ampliar o vocabulário do usuário e disponibilizar outros termos relacionados ao contexto da busca realizada. A Figura 46 traz um exemplo de como este recurso pode contribuir para o aprimoramento da pesquisa, bem como o ícone que deve ser utilizado para o seu acesso (número 3).

Figura 46 - Variações linguísticas



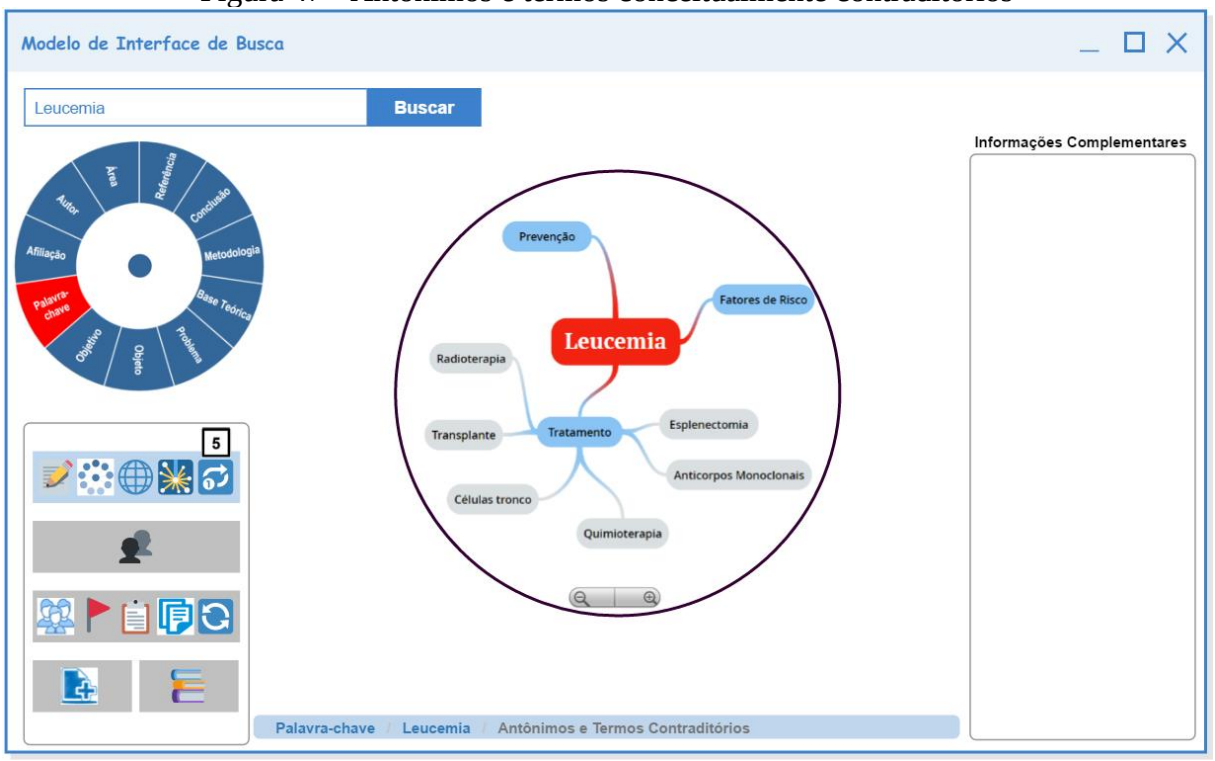
Fonte: desenvolvido pela autora

O exemplo apresentado refere-se à uma busca realizada a partir do termo **Ônibus**, que é um conceito representado por diferentes nomenclaturas em países que adotam a língua portuguesa como idioma padrão. Esta variação linguística é do tipo diatópica (que se refere às características regionais), podendo ocorrer ainda as diacrônicas (em função da passagem de tempo e da evolução da linguagem) e as diastráticas (que estão relacionadas com os diferentes estratos sociais).

Assim como os sinônimos e termos relacionados, os conceitos da ontologia relacionados às variações linguísticas oferecem inúmeras possibilidades para ampliar o contexto da busca. Neste sentido, uma pesquisa realizada pela palavra ônibus pode ser expandida pelos termos autocarro, microlete, otocarro, toca-toca e machimbondo. Este recurso permite, portanto, a expansão de consultas por meio de termos que podem ser totalmente desconhecidos do usuário e que não fazem parte do seu vocabulário, possibilitando a recuperação de resultados correlacionados, que o pesquisador nem sabia que existiam.

A quinta e última funcionalidade que fornece ao usuário conhecimento adicional relacionado a um termo de busca refere-se à apresentação de **antônimos e termos conceitualmente contraditórios**. Assim como as demais, a função só é habilitada a partir da seleção de um termo no mapa de conceitos, e o seu acesso se dá através do ícone número 5, conforme pode ser observado na Figura 47.

Figura 47 - Antônimos e termos conceitualmente contraditórios



Fonte: desenvolvido pela autora

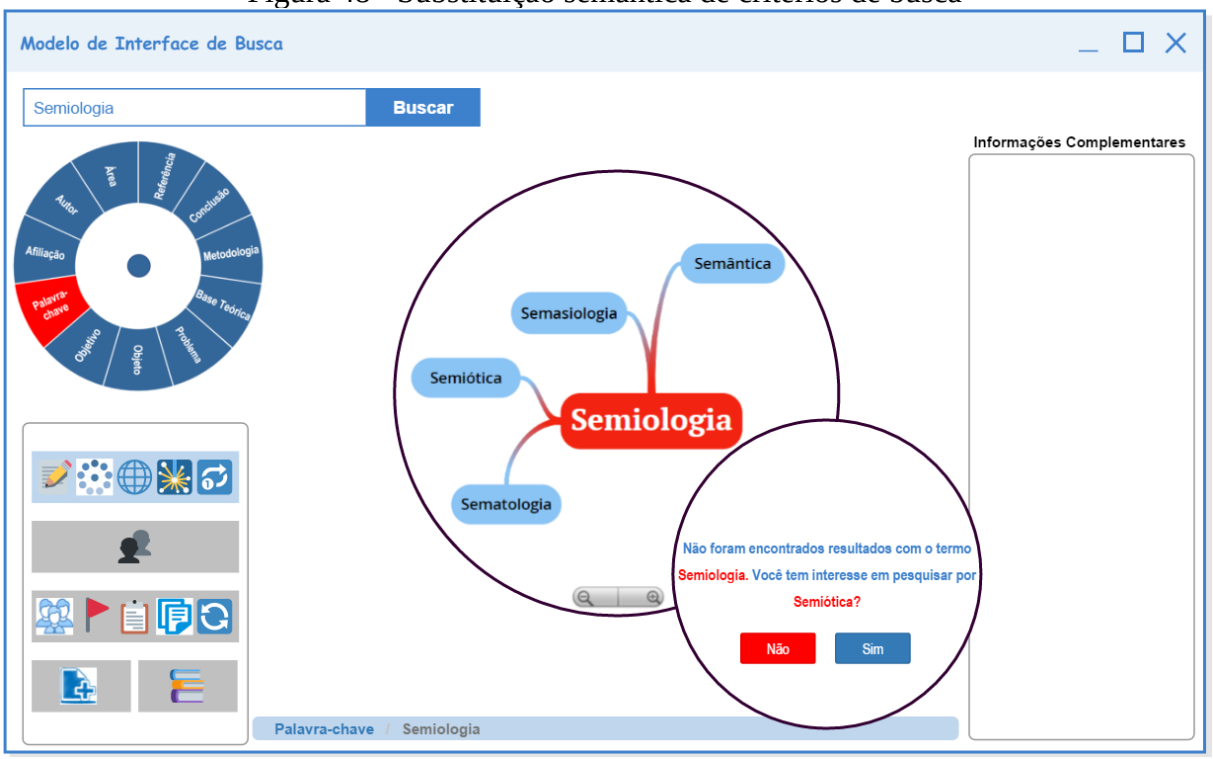
Em um contexto diferenciado dos sinônimos e variações linguísticas, o recurso de antônimos e termos conceitualmente contraditórios tem como base o princípio hermenêutico do Raciocínio Dialético. O principal objetivo desta funcionalidade é oferecer ao usuário elementos contrários à busca realizada, o que seria útil em algumas situações específicas de pesquisa. Neste sentido, ao realizar uma busca pelo termo **Leucemia**, a intenção do usuário pode não ser exatamente obter documentos relacionados à patologia em si, mas ao seu contraditório conceitual, que seria a **Não Leucemia**, ou seja, resultados que se referem ao seu tratamento, fatores de risco e prevenção.

Apesar de ser um recurso aparentemente simples, existe uma certa complexidade para ser desenvolvido. Em um nível básico, ele pode ser implementado a partir de um dicionário de antônimos. Assim, ao selecionar o termo **doença** no mapa de conceitos, a interface teria condições de oferecer várias sugestões de palavras que possuem relação de contrariedade e oposição, tais como: saúde, robustez, vitalidade, bem-estar, vigor, energia entre outros. Em um nível intermediário, é necessário que a ontologia de domínio seja desenvolvida para dar suporte ao contraditório conceitual, não se restringindo apenas aos antônimos perfeitos, mas também aos conceitos que se contradizem em um mesmo nível semântico, tal como o exemplo apresentado na Figura 47.

Além dos recursos já discutidos, caso o usuário elabore uma expressão de

busca utilizando um termo que não possui documentos associados a ele, a interface sugere alternativas de pesquisa a partir de critérios correlacionados.

Figura 48 - Substituição semântica de critérios de busca



Fonte: desenvolvido pela autora

A Figura 48 apresenta um contexto de busca no qual o usuário tem a intenção de encontrar documentos contendo o termo **Semiologia** na perspectiva Palavra-chave. No entanto, não foram encontrados resultados que correspondessem ao critério estabelecido inicialmente. Através dos relacionamentos presentes na ontologia de domínio, a ferramenta tem condições de identificar documentos associados a conceitos semanticamente próximos, que também podem ser de interesse do pesquisador, tal como **Semiótica**. Caso a opção **Sim** seja selecionada, o mapa de conceitos é atualizado, dando destaque ao novo critério de busca, a partir do qual a pesquisa será realizada.

No âmbito do primeiro cenário, a interface de busca possibilita, portanto, o desenvolvimento de pesquisas realizadas a partir de termos, conceitos ou palavras-chaves digitadas pelo usuário. Neste quesito, oferece elementos para auxiliar o pesquisador no desenvolvimento de expressões de busca melhores e mais contextualizadas, bem como para ampliar o seu conhecimento sobre um termo disponível no mapa de conceitos e oferecer opções de expansão para as consultas realizadas. Entre as principais funcionalidades, destacam-se: 1) recurso de autopreenchimento; 2) interação para a elaboração de expressões

de busca, apresentando a possibilidade de inclusão de termos e perspectivas; 3) notas explicativas e definições sobre termos de busca; 4) ambiente gráfico, contendo termos semanticamente relacionados (sinônimos, variação linguística, antônimos e contraditórios), e 5) contextos de análise, com o intuito de revelar as associações e facilitar a interpretação do domínio pesquisado pelo usuário.

É importante destacar que as funcionalidades discutidas neste cenário não pertencem exclusivamente à elaboração de buscas realizadas por meio de palavras-chave, ou seja, são recursos que também poderão ser utilizados nos demais contextos que serão apresentados a seguir.

7.3.2.2 Cenário de busca baseada em estrutura terminológica

Este cenário descreve o desenvolvimento de buscas baseadas em estrutura terminológica. Foi projetado com o intuito de facilitar a pesquisa do usuário, principalmente quando ele não tem conhecimento sobre o domínio explorado ou quando não possui vocabulário suficiente para elaborar expressões de busca significativas, que retornem resultados relevantes. Para que isto seja possível, disponibiliza ao pesquisador, de forma gráfica, toda a terminologia do domínio, contendo os conceitos e os relacionamentos existentes entre eles.

Assim, é possível analisar o domínio da busca e extrair conhecimentos importantes para a sua compreensão, tais como: conceitos desconhecidos, categorias de assuntos, sinônimos, especializações, relações implícitas, relacionamentos entre instâncias e assim por diante. Além disso, a ferramenta permite a elaboração de consultas semânticas a partir da interação direta com os conceitos da ontologia.

Quadro 13 - Busca baseada em estrutura terminológica

Segundo Cenário
Fernanda é formada em Letras e atua como professora em uma escola de línguas na cidade de Londrina. Como tem a intenção de aprimorar seus conhecimentos e trabalhar futuramente no ensino superior, está pretendendo elaborar um projeto de pesquisa para concorrer à uma vaga no Mestrado em Ciência da Informação. Por ser uma área desconhecida, ela julga ser importante realizar buscas exploratórias sobre o domínio e, posteriormente, com informações mais precisas, decidir qual assunto deve investigar. Para que isso seja possível, Fernanda quer identificar alguns conceitos importantes da Ciência da Informação e realizar algumas leituras sobre o assunto. No entanto, está encontrando dificuldades para elaborar palavras-chaves e termos adequados para realizar buscas significativas, uma vez que não possui conhecimento sobre a área. Além disso, não consegue relacionar os conteúdos entre si, tampouco compreender o domínio pesquisado. Para minimizar suas dificuldades, Fernanda acessa a base de dados e utiliza o recurso de busca baseada em

estrutura terminológica, que disponibiliza ao usuário, em formato gráfico, toda a terminologia referente ao campo explorado, contendo os conceitos e os relacionamentos existentes entre eles. Além da visualização, essa funcionalidade permite ainda a elaboração de consultas semânticas, que poderão ser geradas a partir da interação direta com os elementos disponibilizados na ontologia.

Fonte: desenvolvido pela autora

O cenário apresentado no Quadro 13 refere-se à elaboração de buscas em domínios desconhecidos pelo usuário. A maioria das ferramentas disponíveis atualmente não oferece recursos para auxiliar o pesquisador com dificuldades de expressar uma necessidade de informação por meio de palavras-chave. Também não disponibiliza elementos que possam contribuir para que a pesquisa seja bem-sucedida, mesmo quando o usuário não sabe exatamente o que procura ou quando ele não conhece os termos que deve utilizar.

Neste contexto, Fernanda acessa a base de dados científica e indica a perspectiva de interesse a partir da qual as buscas devem ser realizadas. Neste ponto, ao selecionar a área de **Ciência da Informação**, por exemplo, ela está definindo qual será a ontologia que dará subsídios às pesquisas elaboradas. Ao invés de criar termos de busca, o que seria uma tarefa complicada pelo fato de não ter conhecimentos sobre este campo, após a definição da área de busca, a pesquisadora deve selecionar a opção **Buscar** por meio do botão direito do mouse.

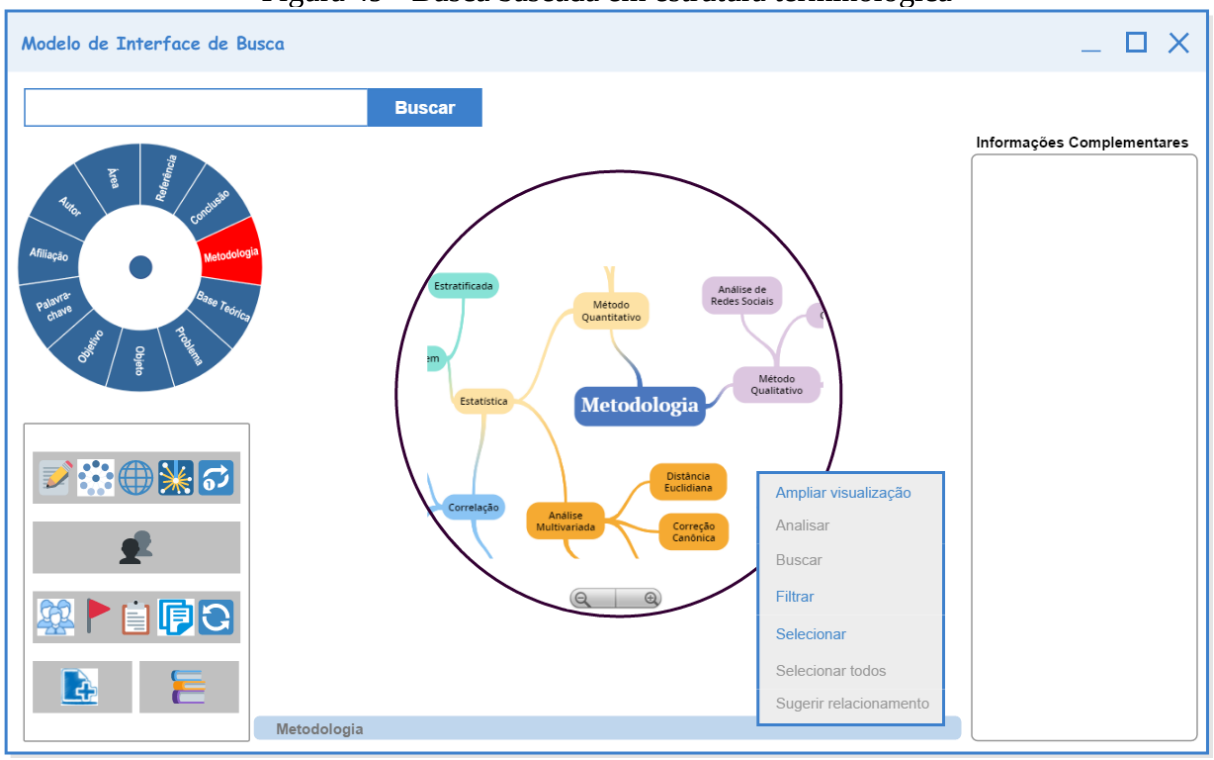
A Figura 49 apresenta a interface de busca baseada em estrutura terminológica. Os conceitos do domínio são apresentados no formato de *Nó-link* e Árvore (grafo), onde cada nó representa um assunto que pode ser explorado. Em função dos recursos de contexto+foco e distorção, é uma interface que permite que os componentes diminuam ou aumentem de tamanho conforme a distância do centro da busca, dando a sensação de que todas as informações disponibilizadas são de fácil acesso. Neste contexto, a pesquisadora selecionou a perspectiva **Metodologia**, a partir da qual as buscas serão realizadas.

A interface apresenta toda a estrutura terminológica disponível para a consulta, relacionada à **Metodologia**, indicando que podem ser explorados elementos relacionados aos **Métodos Quantitativos** e **Qualitativos**. Neste caso, nenhum nó foi selecionado para a busca ainda, ou seja, o usuário está apenas explorando o mapa de conceitos em busca de algo que esteja de acordo com os seus objetivos de pesquisa.

Este recurso é útil, principalmente quando o usuário não tem conhecimento sobre o domínio da busca e, portanto, não consegue compreender quais são os termos relevantes que podem ser utilizados e como eles estão organizados. Também é importante quando os objetivos de busca não estão bem definidos ou para permitir que termos esquecidos

e/ou desconhecidos possam também ser explorados.

Figura 49 - Busca baseada em estrutura terminológica

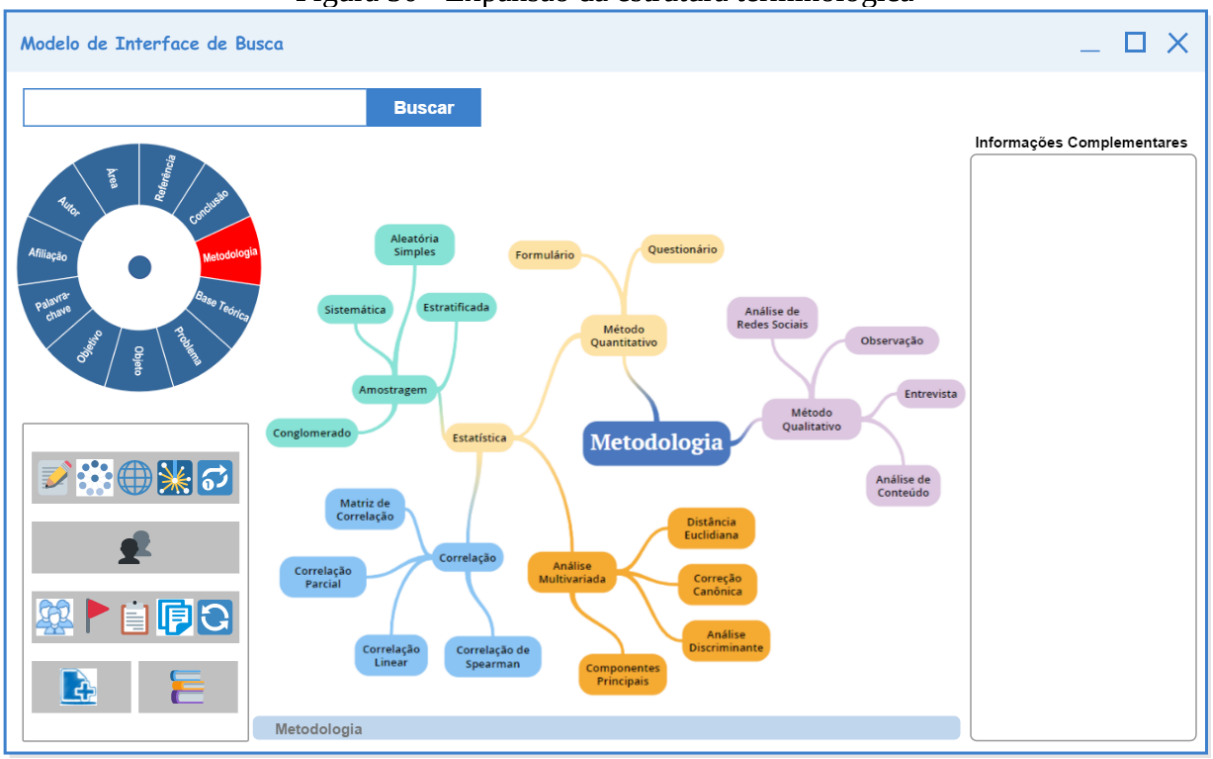


Fonte: desenvolvido pela autora

Como o espaço reservado ao mapa de conceitos é relativamente pequeno e não permite que muitos termos e relacionamentos sejam apresentados de uma única vez, o usuário tem a possibilidade de **Ampliar a visualização** através do botão direito do mouse, possibilitando a sua expansão e, consequentemente, uma maior facilidade de interação. Um exemplo desta funcionalidade pode ser observado na Figura 50.

Os termos são apresentados com cores diferenciadas para cada nó associado ao termo principal. As cores utilizadas pela ferramenta variam de acordo com o contexto da busca, ou seja, se o foco da análise do usuário for áreas de estudo dentro de um determinado domínio, cada área será apresentada por meio de uma cor diferente; se o foco for palavras-chave, cada uma delas terá uma cor diferenciada. Os grupos de palavras que pertencem ao mesmo contexto semântico e que estão correlacionadas, podem ser identificadas também pela cor. Logo, é possível identificar visualmente em que parte do mapa de conceitos encontra-se a informação desejada. Isto pode ser observado na figura: os elementos relacionados à Amostragem são apresentados na cor verde; os que se referem à Correlação são apresentados na cor azul e, assim, sucessivamente.

Figura 50 - Expansão da estrutura terminológica



Fonte: desenvolvido pela autora

O usuário pode interagir diretamente com os componentes da ontologia e utilizá-los tanto para a navegação, quanto para a elaboração de consultas semânticas. A navegação permite a exploração do conteúdo do domínio, bem como a compreensão das relações entre os diversos conceitos envolvidos. Segundo Case (2012), o processo de busca é dinâmico por natureza, ou seja, os conceitos e as necessidades das pessoas mudam à medida que elas interagem com uma nova informação.

A navegação nos conceitos da ontologia, realizada de forma dinâmica e contextualizada, pode proporcionar ao pesquisador conhecimento sobre o domínio explorado, além de desenvolver novas percepções acerca das suas necessidades de informação. Pode contribuir ainda para que o usuário se depare com conceitos que ele não conhecia (ou não se recordava), favorecendo a descoberta casual de novas informações ou estimulando a elaboração de buscas sob outras perspectivas. É também útil para que se possa compreender o que é possível explorar em um determinado domínio.

Outro ponto importante da utilização do mapa de conceitos gerado a partir de uma ontologia de domínio refere-se à possibilidade de o usuário encontrar relacionamentos implícitos e gerar hipóteses potenciais para novas descobertas. Um exemplo é apresentado no Quadro 14.

Quadro 14 - Conhecimento implícito das ontologias

Autor	Objeto	Problema	Metodologia
A1		B	C
A2	A	B	

Fonte: desenvolvido pela autora

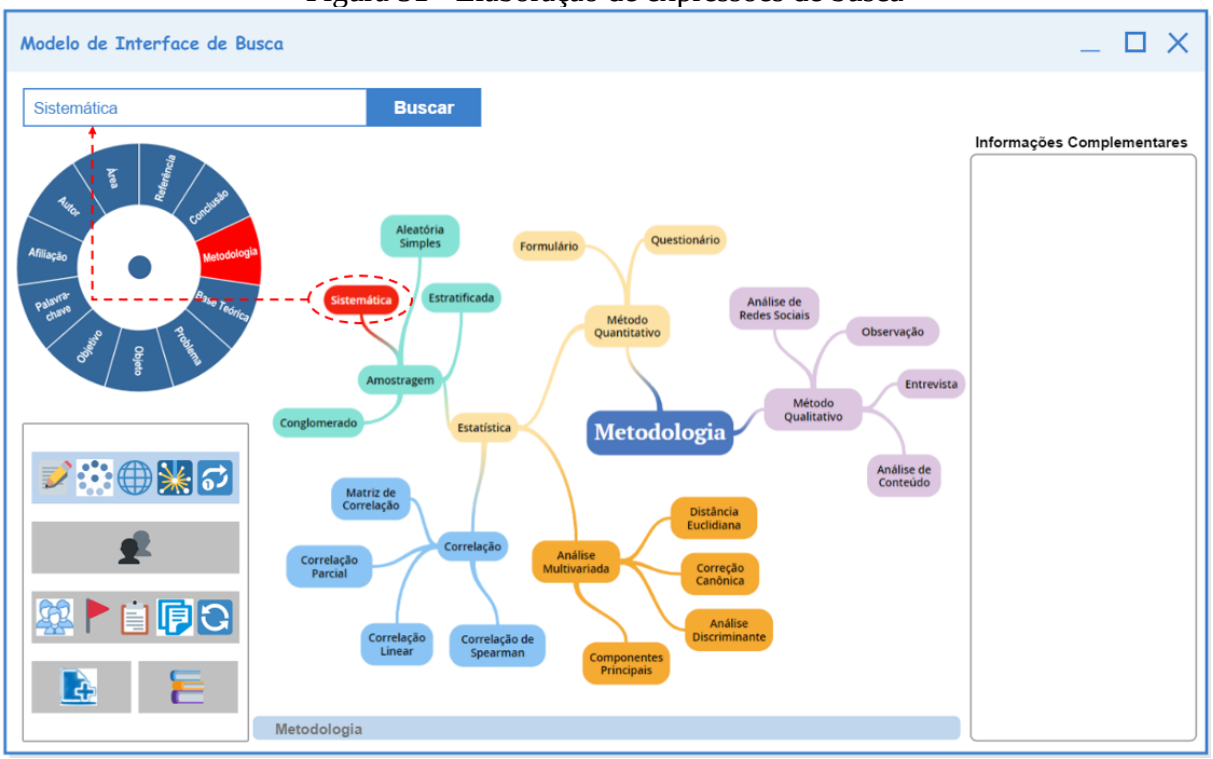
Supondo que o autor A1 tenha abordado um determinado problema B com o auxílio da Metodologia C. Um segundo autor A2, no mesmo contexto, utilizou um objeto de pesquisa A para tentar solucionar o problema B. Um terceiro pesquisador, analisando os elementos que representam este cenário, pode deduzir que existe uma conexão indireta entre A e C, ou seja, que o Objeto A pode ser estudado a partir da Metodologia C. Desta forma, um conceito pode contribuir com outro, mesmo que não estejam conectados diretamente, o que torna a análise do contexto ainda mais importante.

Em relação à elaboração de consultas dinâmicas, a ferramenta permite que o usuário selecione os termos desejados no mapa de conceitos e arraste-os para o retângulo de busca, com o objetivo de elaborar expressões que retornem documentos relevantes para atender as suas necessidades de informação.

Considerando o exemplo apresentado na Figura 51, o pesquisador selecionou um termo no mapa de conceitos e, a partir dele, elaborou uma expressão de busca. Este recurso tem como objetivo auxiliar o usuário sem conhecimentos sobre o domínio e sem opções de vocabulário a elaborar termos de busca significativos. Também reduz os problemas linguísticos, relacionados principalmente com a polissemia, que gera interpretações diferentes para o mesmo vocábulo.

Ao selecionar o termo **Sistemática** para a busca, todo o contexto associado a ele também é inserido na pesquisa, ou seja, o sistema reconhece, por meio das relações semânticas presentes na ontologia, que este conceito está ligado com Amostragem e Estatística. Isto significa que expressões de busca baseadas em ontologias não apresentam problemas de ambiguidade e, mesmo que existam inúmeras interpretações e significados para o termo selecionado, o campo de interesse é claramente definido, impedindo, portanto, que resultados irrelevantes relacionados a outros contextos, sejam recuperados.

Figura 51 - Elaboração de expressões de busca



Fonte: desenvolvido pela autora

Além de minimizar os problemas relacionados com a ambiguidade, a utilização de ontologias como componente de interface para a elaboração de expressões de busca, também reduz as dificuldades discutidas por Altounian e Gomes (2016), Baronas (2011), Bregagnol e Lara (2012) e Camacho (2004), relacionados com:

a) Sinonímia: as palavras sinônimas, que pertencem ao mesmo campo semântico, são apresentadas em conjunto, oferecendo ao usuário múltiplas possibilidades de buscas dentro do mesmo contexto. Assim, se houverem indexações com termos diferentes para o mesmo contexto, isto não impedirá que documentos relevantes sejam recuperados;

b) Homonímia e Paronímia: erros de digitação em função da falta de conhecimento de palavras que possuem a mesma pronúncia ou pronúncias muito parecidas, mas significados diferentes, serão minimizados, pois cada um dos termos estará relacionado à uma classe diferente da ontologia, que será reconhecida pelo usuário em função do seu contexto (concerto – sessão musical; conserto – corrigir);

c) Conotação: a geração de várias possibilidades de interpretação sobre uma determinada palavra também será reduzida, visto que cada variação estará ligada ao seu respectivo contexto;

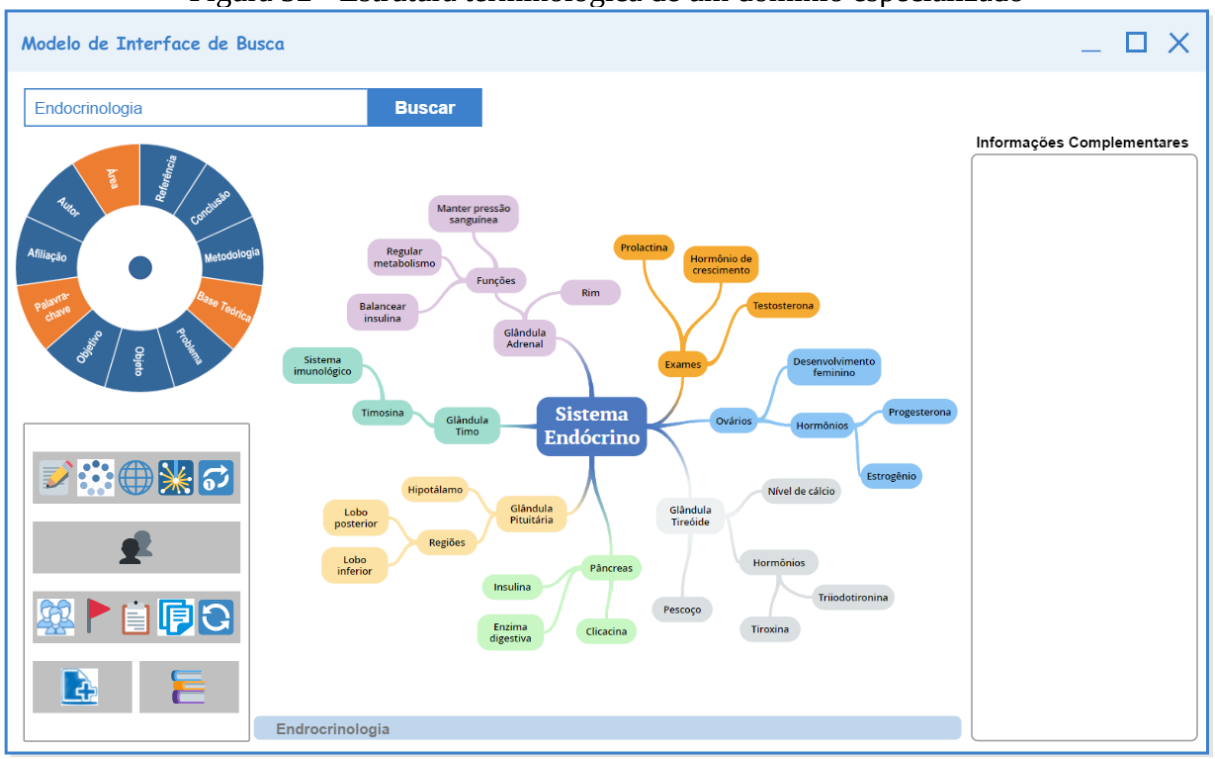
d) Variações linguísticas (diacrônicas, diatópicas e diastráticas): todas as variações linguísticas resultantes da passagem do tempo, das características regionais e dos diferentes estratos sociais podem ser incluídas na ontologia e servirão de base para ampliar o vocabulário do usuário, bem como suas possibilidades de pesquisa. Isto significa que se um documento for escrito por um autor brasileiro e for indexado a partir do termo **apostila** e outro for escrito por um autor português e indexado pelo termo **sebenta**, ambos estarão relacionados em função do seu contexto e serão recuperados, mesmo que o pesquisador não tenha conhecimento acerca destas variações.

Os recursos oferecidos para reduzir os problemas linguísticos na interface de busca proposta estão fundamentados no princípio do compartilhamento de linguagem, baseado nos estudos de Schleiermacher, Dilthey, Heidegger e Gadamer, que discutem sobre a necessidade de compreensão da linguagem para que o processo de interpretação aconteça. Esta compreensão envolve a interpretação gramatical, incluindo o entendimento da linguagem usada pelo autor (significado das palavras), das regras gramaticais e da semântica.

Este princípio justifica também as múltiplas interpretações dos termos, uma vez que a linguagem utilizada por um autor possui influência de vários elementos: seu lugar na história, sua educação, sua tradição, sua cultura e seu dialeto, por exemplo. Assim, os significados se acumulam nas palavras a partir das interrelações que constituem o mundo explicando, portanto, os problemas relacionados com ambiguidade, sinonímia, homonímia, paronímia, conotação e variações linguísticas.

No contexto das buscas baseadas em estrutura terminológica, a interface oferece suporte para a elaboração de termos de busca, mesmo quando o usuário tem poucos conhecimentos sobre o domínio de interesse e, portanto, não possui vocabulário para realizar pesquisas significativas sobre o assunto. Deste modo, toda a terminologia disponível para a busca é apresentada ao pesquisador, no formato de mapa de conceitos, que pode navegar nos componentes gráficos ou arrastá-los para o retângulo de busca. A Figura 52 traz o exemplo de um mapa de conceitos simplificado, relacionado à área da Endocrinologia.

Figura 52 - Estrutura terminológica de um domínio especializado



Fonte: desenvolvido pela autora

Observando o mapa de conceitos apresentado, mesmo sendo de uma área especializada, o usuário tem condições de obter informações detalhadas sobre o domínio, mesmo que não tenha nenhum conhecimento sobre ele. De forma superficial, é possível perceber quais são as glândulas pertencentes ao sistema endócrino: adrenal, timo, pituitária e tireoide. Também pode-se afirmar que ovários, pâncreas, hipotálamo, hormônios e sistema imunológico possuem algum tipo de relação com a endocrinologia, seja ela direta ou indireta.

Neste sentido, além de oferecer opções de vocabulário para usuários que não possuem conhecimento de domínio, também permite que ele tenha acesso aos conceitos da área, aos seus relacionamentos e à sua estrutura de organização. Isto é importante, uma vez que a navegação nos conceitos da ontologia pode revelar termos desconhecidos pelo pesquisador, bem como apresentar outras necessidades de informação até então não identificadas. Também permite a aquisição de conhecimento sobre o domínio, enquanto explora o conteúdo disponibilizado.

No âmbito do segundo cenário, a interface de busca oferece, portanto, recursos projetados para o desenvolvimento de buscas baseadas em estrutura terminológica, tais como: 1) técnica de contexto+foco e distorção para estimular a navegação e a exploração do conteúdo; 2) funcionalidades de *zoom*, para facilitar a visualização e a exploração dos conceitos de interesse; e 3) desenvolvimento interativo de termos e expressões de busca, que

dão subsídios aos usuários para a exploração de conteúdos, independente do conhecimento que possuam sobre o domínio investigado.

7.3.2.3 Cenário de busca baseada em perspectivas de análise de contexto

As ferramentas de busca disponíveis retornam resultados contendo inúmeros documentos desconectados entre si, exigindo do usuário tempo e conhecimento para estabelecer as relações existentes entre eles e, então, compreender as características do domínio explorado. Para minimizar essa dificuldade, o terceiro cenário deste estudo apresenta o desenvolvimento de buscas baseadas em perspectivas de análise de contexto, que possuem o objetivo de auxiliar o pesquisador na identificação de documentos contendo uma ou várias características correlacionadas, bem como na percepção das conexões existentes no *corpus* pesquisado.

Quadro 15 - Busca baseada em perspectivas de análise de contexto

Terceiro Cenário
Fernanda é formada em Letras e atua como professora em uma escola de línguas na cidade de Londrina. Como tem a intenção de aprimorar seus conhecimentos e trabalhar futuramente no ensino superior, está pretendendo elaborar um projeto de pesquisa para concorrer à uma vaga no Mestrado em Ciência da Informação. Por ser uma área desconhecida, ela julga ser importante realizar buscas exploratórias sobre o domínio e, posteriormente, com informações mais precisas, decidir qual assunto deve investigar. Para que isso seja possível, Fernanda quer levantar informações mais aprofundadas da Ciência da Informação, com o intuito de entender como a área está organizada em termos de publicações, quais são os principais autores e instituições que se interessam pela área, qual é o foco das pesquisas realizadas, quais são as metodologias de estudo predominantes, quais problemas são abordados, quais bases teóricas fundamentam os estudos realizados e assim por diante. O primeiro passo é acessar a base de dados, definir a área de pesquisa e, posteriormente, selecionar uma das diversas perspectivas apresentadas. Cada perspectiva revela um contexto de busca, contendo inúmeros critérios que podem adotados para explorar o domínio. Existe a possibilidade de escolher um ou mais critérios, pertencentes a uma ou várias perspectivas. Entre as buscas que são possíveis realizar, Fernanda pode solicitar à interface que retorne todas as publicações da área de Ciência da Informação relacionadas a um determinado autor e que adotem uma mesma metodologia de estudos (Ex. publicações na área de Ciência da Informação, pertencentes ao autor José da Silva e que adotem Análise de Conteúdo como método de pesquisa e levantamento de informações).

Fonte: desenvolvido pela autora

Neste ponto, o usuário tem como objetivo analisar materiais disponibilizados no repositório digital, levantar informações mais aprofundadas e identificar as relações existentes entre eles. Para isto, é necessário ter conhecimento de domínio suficiente para analisar as publicações, entender os conceitos envolvidos em cada contexto e agrupá-los de acordo com critérios de semelhança. No entanto, essa tarefa não é trivial nem

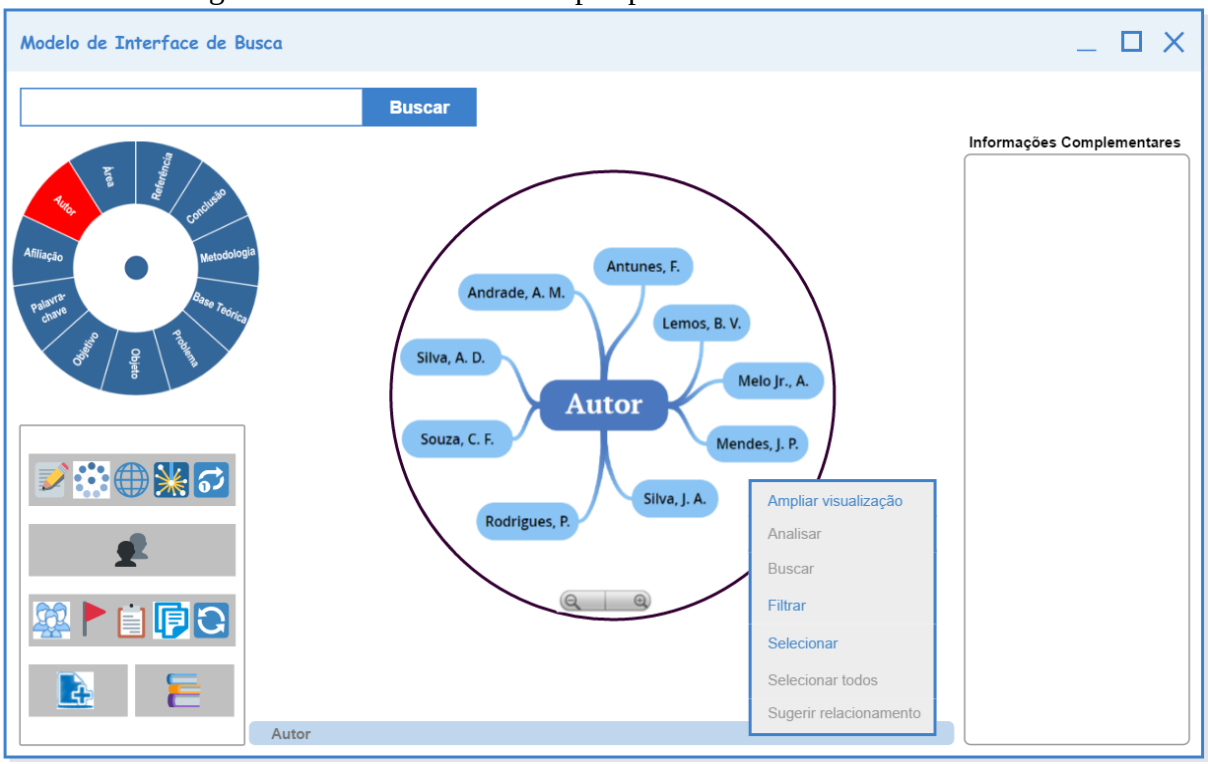
mesmo para pesquisadores que possuem uma visão abrangente e aprofundada do assunto pesquisado, sendo necessário despender um tempo significativo para alcançar este objetivo.

O usuário participante deste cenário tem, portanto, a intenção de elaborar buscas que respondam questões relacionadas a um *corpus* documental, tais como:

a) Quem são os autores que publicam em uma determinada área?
b) Em quais instituições esses autores estão vinculados?
c) Quais publicações possuem objetivos de estudo semelhantes?
d) Quais publicações referem-se a um mesmo objeto de estudo?
e) Quais publicações compartilham uma mesma base teórica?
f) Quais publicações apresentam abordagens metodológicas semelhantes?

Para responder questionamentos como estes, o usuário deve selecionar uma das perspectivas de análise, que revelará todos os conceitos da ontologia associados a ela no mapa de conceitos, dando destaque ao **plano de informação**. A busca pode ser realizada a partir de um ou vários critérios, que podem ser adotados para refinar os resultados. Isto significa que o usuário pode ter interesse em recuperar todos os documentos referentes a um determinado **autor** e que contenham uma **metodologia** específica. As combinações entre os diversos critérios dependem do interesse ou das necessidades de quem está interagindo com a interface. A Figura 53 mostra as opções de busca e análise reveladas pela perspectiva **Autor.**

Figura 53 - Busca baseada em perspectivas de análise de contexto



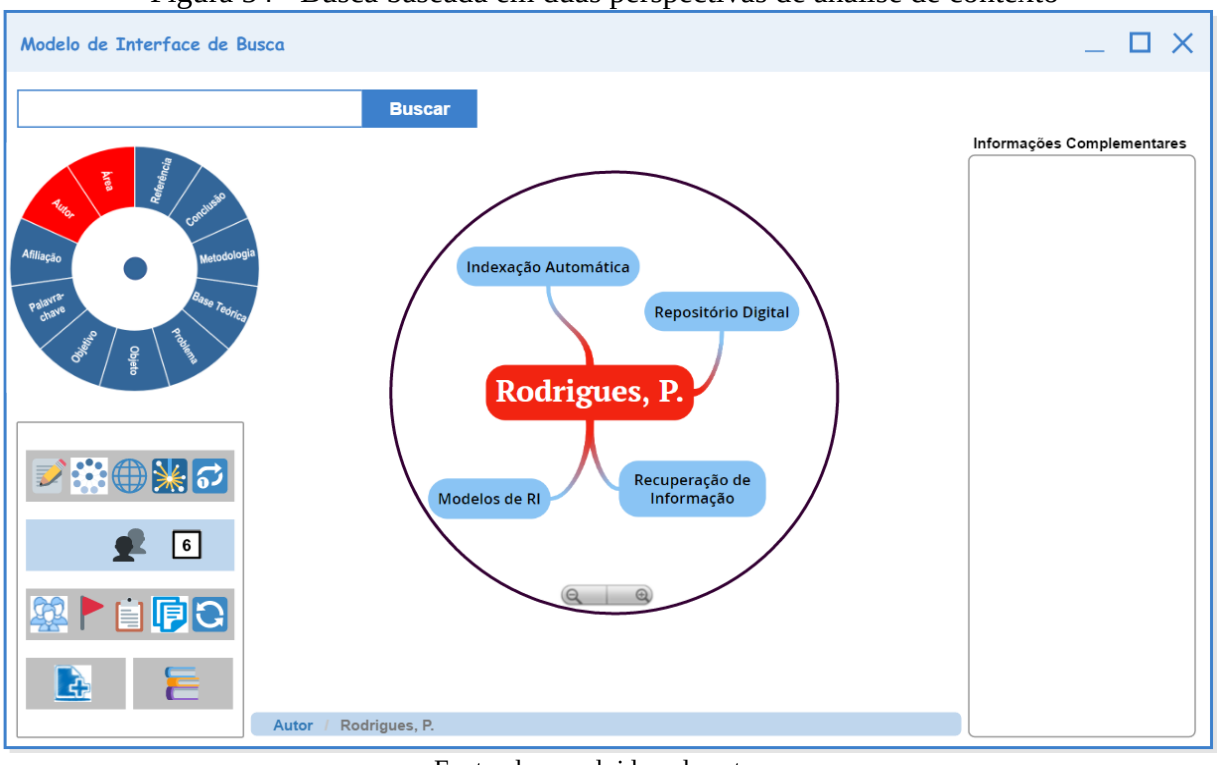
Fonte: desenvolvido pela autora

Além de apresentar as possibilidades de buscas que podem ser realizadas a partir dos autores, é possível filtrar as informações que são apresentadas no mapa de conceitos. A opção **Filtrar**, com acesso através do botão direito do mouse, permite a seleção de autores de acordo com os seguintes critérios: ordem alfabética, data de publicação, número de publicações e número de acessos. Também é possível selecionar apenas um determinado autor e realizar análises mais aprofundadas sobre ele.

Outro ponto importante refere-se à facilidade de realização de buscas baseadas nas diversas perspectivas. Esta funcionalidade permite o cruzamento de informações diversas, viabilizando o desenvolvimento de inúmeras atividades de análise no *corpus* documental, bem como a inferência de informações sobre os contextos avaliados. Por outro lado, apontam ao usuário a rede de relacionamentos existente entre os documentos, o que pode contribuir para a compreensão do contexto analisado.

A Figura 54 apresenta o exemplo de uma busca realizada a partir de duas perspectivas de análise: **Autor** e **Área**. Após a seleção do autor **Rodrigues, P.**, o objetivo do pesquisador foi identificar com quais áreas ele está relacionado. De acordo com as informações apresentadas pelo mapa de conceitos, o autor possui materiais publicados em quatro campos diferenciados: Modelos de RI, Recuperação de Informação, Indexação Automática e Repositório Digital.

Figura 54 - Busca baseada em duas perspectivas de análise de contexto



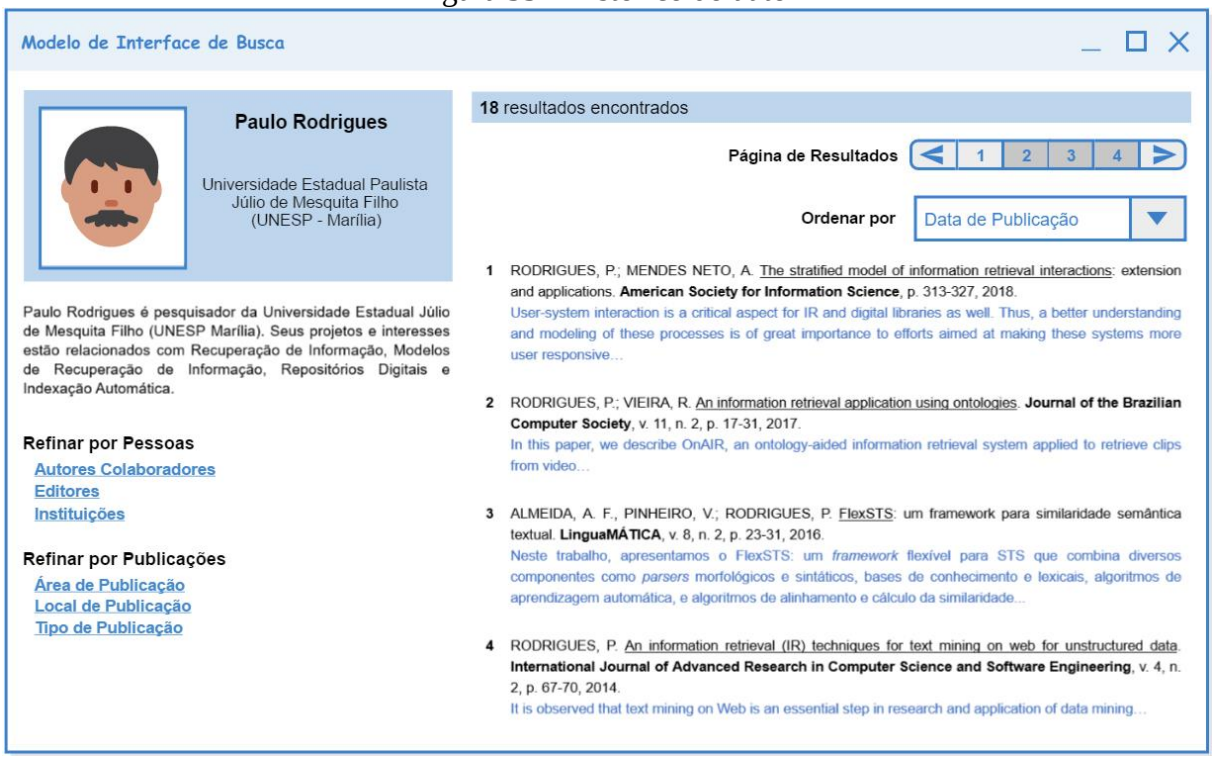
Fonte: desenvolvido pela autora

Neste ponto, o usuário tem a possibilidade de escolher uma ou mais áreas para realizar uma busca ou acessar o **histórico do autor**, por meio do ícone indicado pelo número 6. Este recurso é habilitado todas as vezes que houver a seleção de um autor para a análise, seja por meio de uma perspectiva de busca ou de um documento.

Desenvolvida com base no princípio hermenêutico da Contextualização, esta funcionalidade tem como principal objetivo apresentar ao usuário o histórico de um determinado autor e oferecer elementos para que ele possa ser entendido a partir das suas pesquisas, publicações e áreas de interesse.

A implementação da funcionalidade referente ao histórico do autor pode ser analisada sob duas perspectivas distintas. Em um nível básico, seu propósito é reunir todas as informações referentes ao autor, tal como pode ser observado na Figura 55. Neste exemplo, a interface apresenta um panorama sobre **Paulo Rodrigues**, indicando sua afiliação, uma breve descrição relacionada aos seus projetos e interesses de pesquisa, rede de colaboradores, editores e instituições, áreas de publicação, locais de publicação (periódicos, congressos, simpósios, seminários) e tipos de publicação (artigos, livros, resenhas, resumos, revisões). Além disso, a produção científica completa do autor é disponibilizada para o acesso e a análise do usuário, que pode organizá-la de acordo com os seguintes critérios: data de publicação, número de comentários e número de acessos.

Figura 55 - Histórico do autor

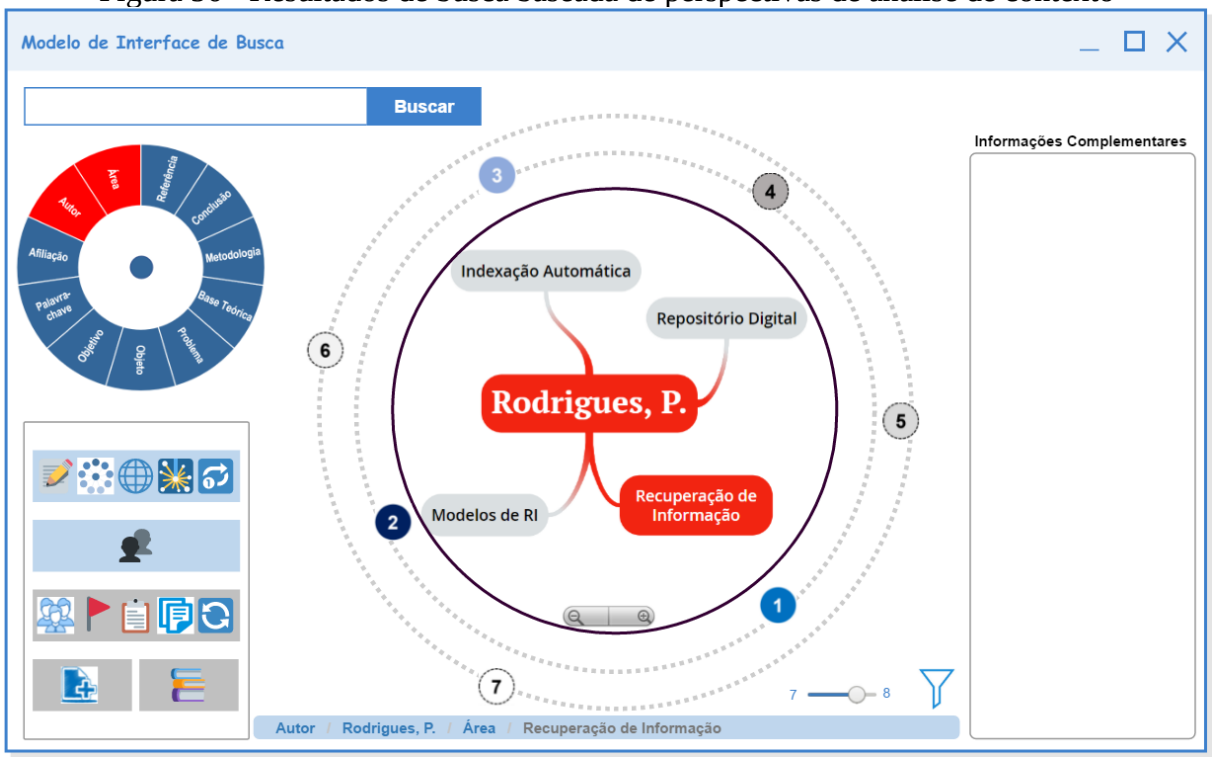


Fonte: desenvolvido pela autora

Em um nível intermediário, o histórico do autor pode ser implementado para dar subsídios à realização de buscas baseadas em aspectos regionais e culturais. Neste sentido, a funcionalidade seria ampliada para oferecer ao usuário sugestão de termos e pesquisas relacionadas ao contexto do autor (região geográfica, grupo de pesquisadores, instituição). Entre as principais contribuições, esta funcionalidade permitiria: 1) analisar termos compartilhados por autores, uma vez que grupos de pesquisadores pertencentes à regiões e/ou instituições diferentes podem utilizar terminologias distintas para a mesma temática; 2) identificar teorias e conceitos convergentes e divergentes entre grupos de pesquisadores; 3) comparar resultados de pesquisas realizadas por pesquisadores de regiões diferentes, que são influenciadas por variáveis locais, tais como: normas, crenças, valores, preconceitos, interesses, nível de desenvolvimento econômico e social, entre outras.

Outra possibilidade de interação do usuário, além de analisar o histórico do autor, seria realizar uma busca com base na seleção de critérios de interesse. A Figura 56 apresenta dois nós destacadas em vermelho e que serão utilizados para efetivar a pesquisa (Autor: **Rodrigues, P**. e Área: **Recuperação de Informação**).

Figura 56 - Resultados de busca baseada de perspectivas de análise de contexto



Fonte: desenvolvido pela autora

A busca retornou documentos que atendem ambos os critérios selecionados, bem como a rede de informações da qual fazem parte. Os Docs. 1, 2 e 3 foram considerados

relevantes, ou seja, foram escritos por Rodrigues, P. e são da área de Recuperação de Informação. Os Docs. 4, 5 e 6 fazem parte da mesma rede de informações, uma vez que estão relacionados, de alguma forma, com a expressão de busca elaborada: ou foram escritos pelo mesmo autor ou possuem vínculo com a área.

É importante destacar que essa rede de relacionamentos pode ser mais complexa e possibilitar a inferência de inúmeras outras informações importantes que podem auxiliar o usuário durante a análise de documentos. Cada documento pode conter um ou mais relacionamentos com uma determinada perspectiva e quanto mais relacionamentos forem compartilhados entre publicações diferentes, mais probabilidade existe de serem semelhantes em relação ao conteúdo avaliado. Essa possibilidade de proximidade entre conteúdos de dois ou mais documentos cresce proporcionalmente à medida em que aumentam o número de relacionamentos compartilhados entre eles.

Embora o exemplo apresentado tenha se limitado em analisar **Autor e Área**, os mesmos passos podem ser aplicados para levantar informações pertencentes a outros contextos: 1) identificar autores que publicam em uma determinada área: Área e Autor; 2) levantar as instituições onde esses autores estão vinculados: Área, Autor e Instituição; 3) identificar publicações que possuem objetivos de estudo semelhantes (Objetivos), entre outras inúmeras possibilidades de análise e interação.

No âmbito das buscas baseadas em perspectivas de análise de contexto, a interface oferece subsídios para que o usuário possa elaborar consultas a partir de cenários diversos (área, autor, afiliação, palavras-chave, objetivo, objeto, metodologia, entre outros), que são apresentados no mapa de conceitos de forma gráfica. Este recurso disponibiliza ao pesquisador uma melhor compreensão dos conceitos e dos relacionamentos existentes entre os termos do domínio, bem como entre os documentos pertencentes ao *corpus* documental. Também revela a rede de relacionamentos existente entre os documentos, possibilitando que informações implícitas possam ser agregadas às análises realizadas pelo usuário.

7.3.2.4 Cenário de busca baseada em perspectivas de análise de conteúdo

A opção de busca baseada em perspectivas de análise de conteúdo, diferentemente dos três primeiros cenários, não pode ser acessada diretamente na interface principal. Os recursos projetados referem-se a análises realizadas a partir de um determinado documento que foi anteriormente recuperado, seja por meio de uma busca baseada em palavras-chave, em estrutura terminológica ou em perspectiva de análise de contexto.

O principal objetivo deste cenário é, portanto, revelar toda a rede de relacionamentos de um determinado documento e permitir buscas realizadas a partir de critérios e características específicas, com o intuito de encontrar outras publicações pertencentes a contextos semelhantes.

Quadro 16 - Busca baseada em perspectivas de análise de conteúdo

Quarto Cenário
Fernanda é formada em Letras e atua como professora em uma escola de línguas na cidade de Londrina. Como tem a intenção de aprimorar seus conhecimentos e trabalhar futuramente no ensino superior, está pretendendo elaborar um projeto de pesquisa para concorrer à uma vaga no Mestrado em Ciência da Informação. Para aprofundar seus conhecimentos, ela julga ser importante ler publicações da área e buscar informações que possam dar subsídios ao assunto que quer investigar. Para iniciar este processo, Fernanda realiza algumas buscas exploratórias (baseada em palavras-chave, em estrutura terminológica ou em perspectiva de análise de contexto) e escolhe uma publicação para ler e analisar. Ao abrir o material, a interface apresenta à pesquisadora a rede de relacionamentos associada a ele, permitindo que outras buscas possam ser realizadas dentro do mesmo contexto. Entre as inúmeras possibilidades, Fernanda pode encontrar outros artigos que possuam as mesmas Palavras-chave (ou palavras-chave que façam parte do mesmo grupo semântico), que tenham sido elaborados pelo mesmo autor (ou grupo de autores e coautores que trabalham em conjunto), que compartilhem uma mesma base teórica, que adotem um mesmo procedimento metodológico, entre outras. Existe ainda a possibilidade de desenvolver análises a partir de grupos específicos de artigos selecionados pelo autor, com o intuito de obter informações específicas sobre ele.

Fonte: desenvolvido pela autora

No âmbito deste cenário, o usuário está interagindo com a base de dados científica com a intenção de obter informações mais aprofundadas sobre um determinado assunto. Para isto, é importante que ele identifique materiais pertencentes à um mesmo contexto, analise as publicações mais relevantes e correlacione-as com outros materiais pertencentes ao *corpus* documental. Neste ponto, três tarefas são bem definidas e podem ser realizadas inúmeras vezes durante o desenvolvimento de uma pesquisa: 1) seleção de um determinado material; 2) análise do material selecionado e 3) correlação deste material com outros materiais do *corpus* que possuam algum tipo de similaridade.

Esta atividade é realizada pelo pesquisador, de forma consciente ou inconsciente, em situações de busca nas quais ele queira responder questões como:

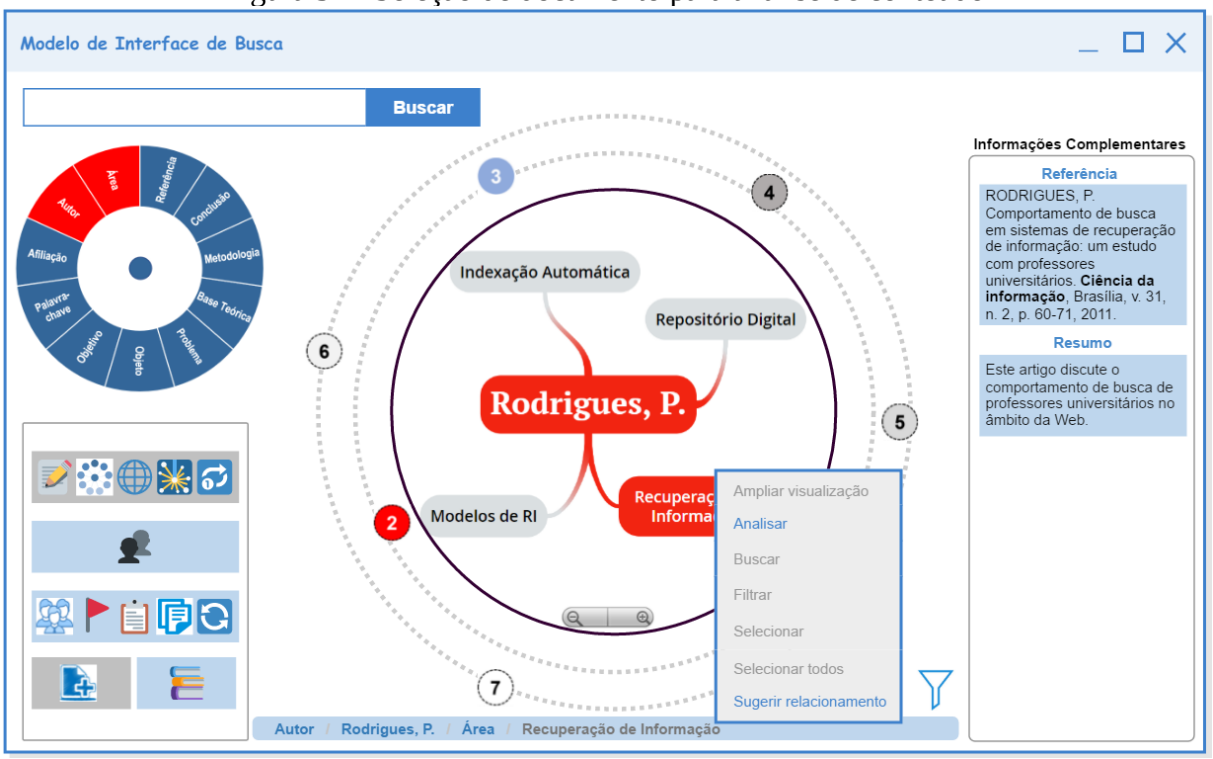
a) Este material possui uma palavra-chave de muito interesse, dentro do contexto da minha busca. Quais outros materiais também foram representados por esta mesma palavra-chave?
b) Quais materiais foram escritos por este mesmo autor (ou grupo de autores)?

c) Quais publicações possuem a mesma base teórica pertencente a este material?

d) Quais publicações também foram desenvolvidas a partir deste mesmo método de pesquisa?

O primeiro passo para o desenvolvimento do cenário é a seleção de um documento presente na base de dados, que tenha sido recuperado em uma busca realizada anteriormente. O acesso ao documento deve ser realizado por meio do botão direito do mouse, através da opção **Analisar**. A Figura 57 apresenta o Doc. 2 selecionado para a análise, bem como sua referência e resumo nas Informações Complementares.

Figura 57 - Seleção de documento para análise de conteúdo



Fonte: desenvolvido pela autora

Com o documento selecionado, todas as informações associadas a ele através da ontologia de domínio ficam disponíveis para a análise do usuário, que pode ser iniciada a partir de três contextos: 1) pelas informações adicionais, que auxiliam o processo de interpretação do documento selecionado: comentários interpretativos, avaliação de parágrafos, indicação de erros e falhas; 2) pela sugestão de pesquisas complementares: resultados relacionados ao mesmo contexto e contraditórios; 3) por uma das perspectivas de análise: Área, Autor, Afiliação, Base Teórica, Metodologia etc.

De acordo com a Figura 58, ao solicitar a análise de um documento, ele

aparece no centro do mapa de conceitos, como forma de indicar ao usuário o contexto pelo qual ele deseja iniciar a busca baseada em perspectivas de análise de conteúdo. As informações adicionais para a interpretação do documento podem ser acessadas pelos ícones 7, 8 e 9 (comentários interpretativos, indicação de erros e falhas, avaliação de parágrafos). As sugestões de pesquisas complementares estão disponíveis nos ícones 10 e 11 (resultados relacionados ao mesmo contexto e contraditórios).

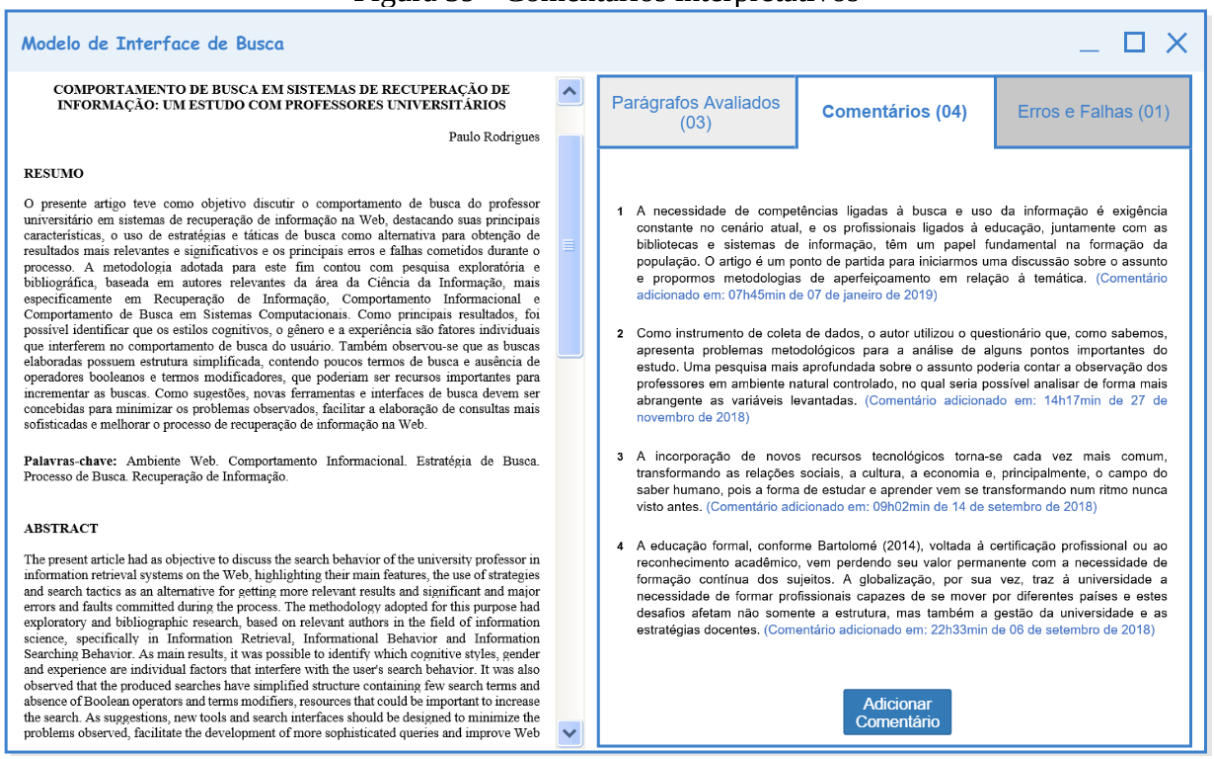
Figura 58 - Documento selecionado para análise de conteúdo



Fonte: desenvolvido pela autora

Ao observar as informações apresentadas pela interface, é possível verificar que o artigo selecionado pelo usuário possui: a) **04 comentários interpretativos**, que se referem às considerações (opiniões, pontos de vista, discussões) de pesquisadores que leram e contribuíram com o texto; b) **01 indicação de erro ou falha**, que diz respeito à algum tipo de problema encontrado, seja de caráter interpretativo, gramatical, metodológico e outros; c) **03 parágrafos avaliados**, que foram observados e analisados por outros usuários, tendo como média 79% de avaliações positivas e 21% de negativas. O documento possui ainda 04 indicações de pesquisas complementares referentes a resultados relacionados ao mesmo contexto.

Figura 59 - Comentários interpretativos



Fonte: desenvolvido pela autora

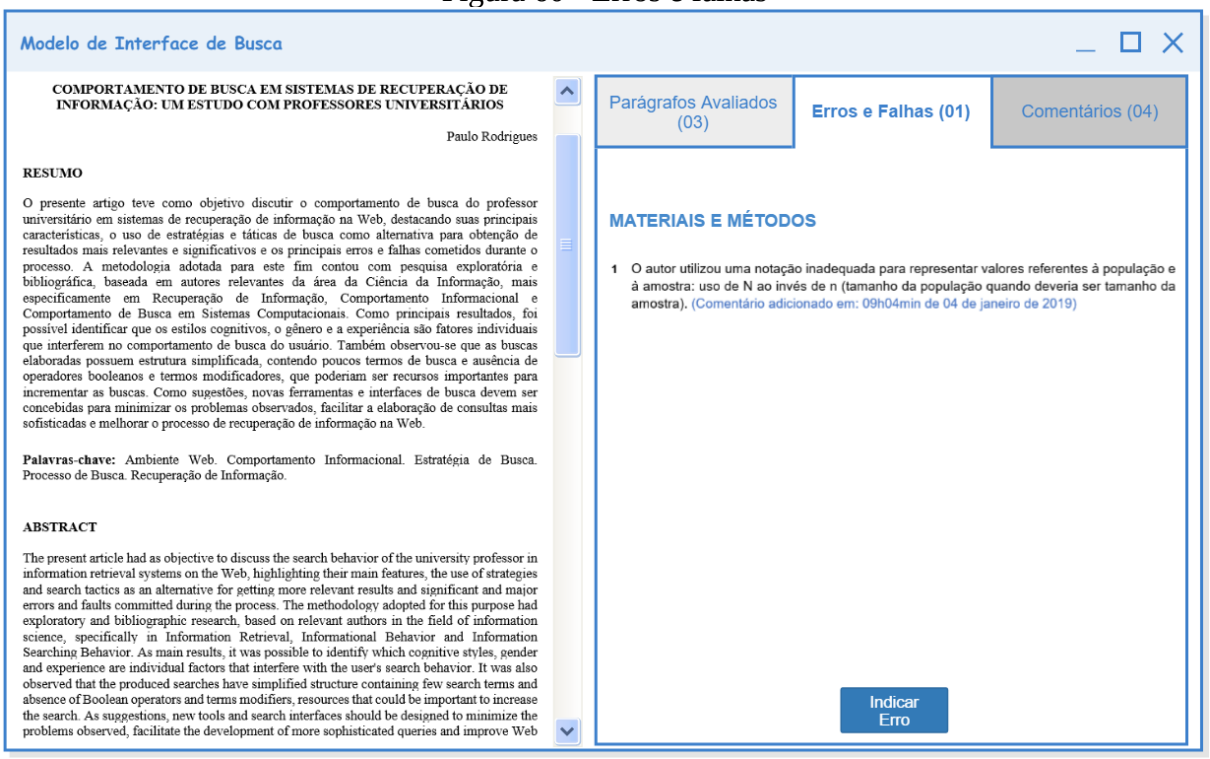
A Figura 59 apresenta os comentários dos usuários relacionados à interpretação do texto selecionado. Este recurso foi desenvolvido com base no princípio hermenêutico da Interação entre Pesquisadores e Sujeitos e tem como objetivo permitir a discussão entre os leitores, como forma de facilitar a interpretação de um determinado material e encontrar um consenso sobre o seu significado.

Cada pesquisador compreende um texto a partir do seu horizonte, que é influenciado pelos seus conhecimentos, preconceitos, experiências, crenças, cultura e toda sua história. Isto significa que um documento pode ser interpretado de diversas formas, com base na perspectiva de cada usuário, o que pode relevar uma multiplicidade de aspectos relacionados ao assunto que está sendo discutido.

Ao possibilitar a interação entre os pesquisadores e usuários, a interpretação de um determinado documento passa a ser um processo de construção de conhecimento, que incorpora informações e pontos de vista a cada nova contribuição, deixando assim de ser uma atividade isolada e vista unicamente a partir das perspectivas e do entendimento do leitor.

A opção **Adicionar Comentário** permite que novas interpretações sejam inseridas no contexto do documento selecionado.

Figura 60 - Erros e falhas



Fonte: desenvolvido pela autora

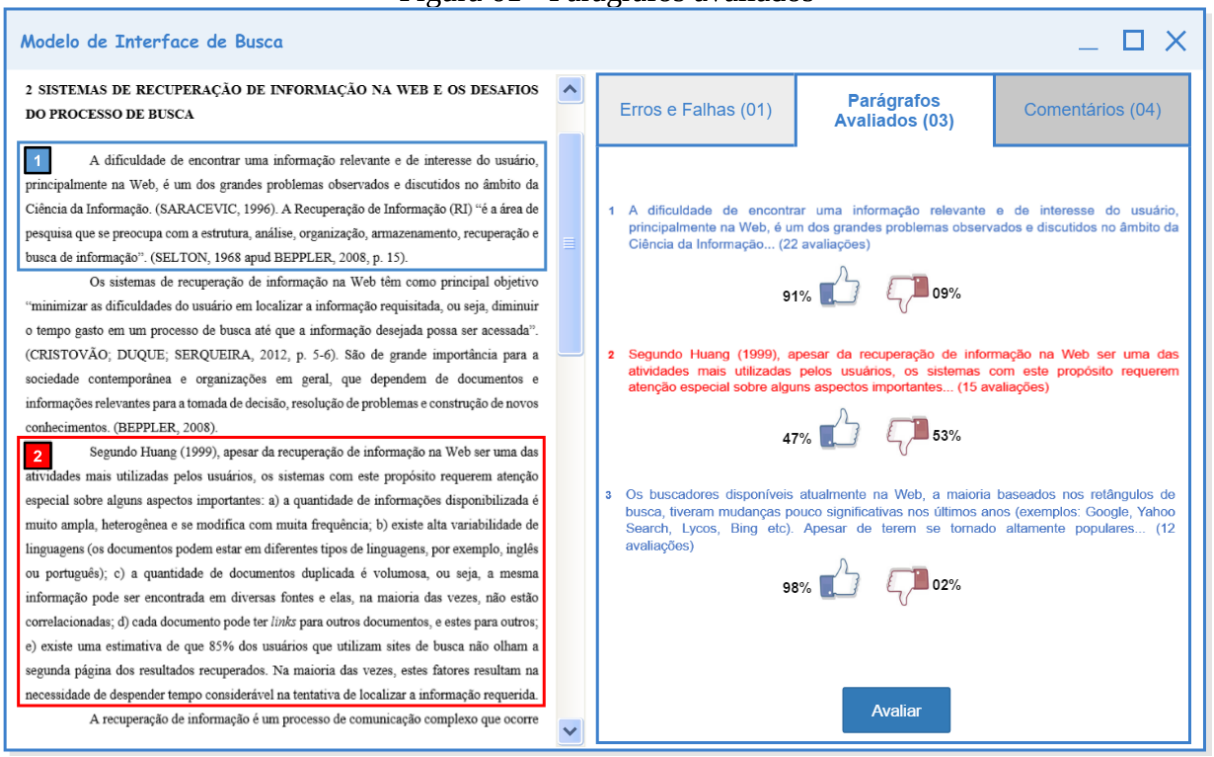
O recurso de erros e falhas, que pode ser observado na Figura 60, foi desenvolvido com base no princípio hermenêutico da Suspeição e tem como objetivo identificar possíveis vieses, distorções e preconceitos ilegítimos no documento selecionado.

A indicação de erros e falhas, assim como os comentários interpretativos, pode ser feita por usuários com a intenção de colaborar com o processo de interpretação do documento. Neste sentido, as suspeitas de equívocos ou imprecisão, independente do caráter que tenham (interpretativo, gramatical, metodológico e outros), pode ser comunicada ao editor ou aos demais leitores. Além disso, este recurso pode ser utilizado ainda para questionar a autoria do texto (ou de partes dele) e indicar a possibilidade de plágio, como também a veracidade dos dados apresentados.

Este recurso pode ser analisado a partir de dois cenários diferentes. Em um nível básico, a funcionalidade pode ser implementada para que os erros e as falhas sejam indicados pelos usuários, em formato textual, tal como observado na Figura 60. Em um nível intermediário, além das avaliações realizadas pelos leitores, o próprio sistema promoveria a análise do documento, por meio da utilização de técnicas de inteligência artificial.

A opção **Indicar erro** permite que problemas possam ser indicados ao editor ou aos usuários.

Figura 61 - Parágrafos avaliados



Fonte: desenvolvido pela autora

A Figura 61 apresenta os parágrafos que foram avaliados no documento selecionado. Esta funcionalidade é um complemento da anterior (Erros e Falhas) e também está fundamentada no princípio hermenêutico da Suspeição. Tem como objetivo promover a avaliação do conteúdo disponibilizado no texto, por meio dos recursos de *like* (gostei, aprovado, interessante) e *dislike* (não gostei, desaprovado, desinteressante).

À medida em que o texto vai sendo avaliado pelos usuários, o sistema gera uma média de cada parágrafo e uma média geral do documento, que foi apresentada na Figura 58 (número 9). Este recurso contribui para que o leitor identifique os elementos mais relevantes do texto, bem como aqueles que foram desaprovados ou que apresentam algum tipo de problema ou inconsistência.

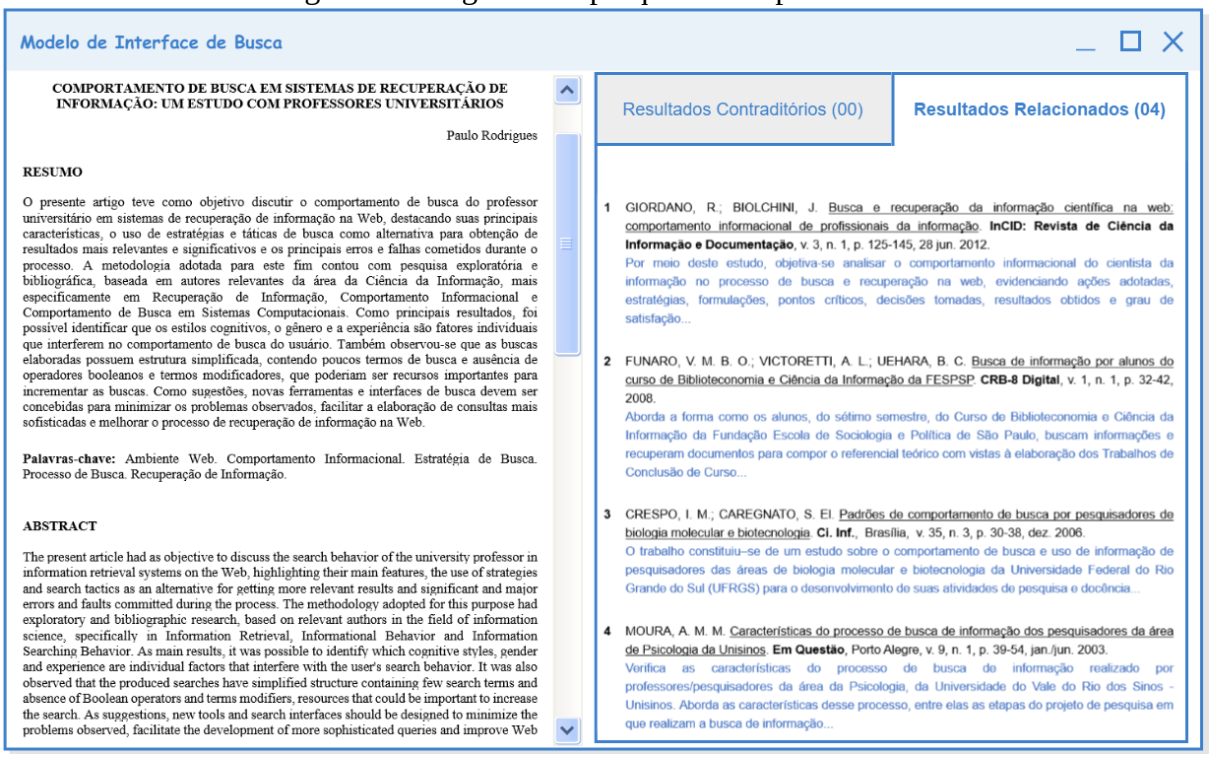
Para facilitar a identificação dos parágrafos e sua respectiva avaliação, optou-se por utilizar as cores azul e vermelho. Os parágrafos com avaliação positiva (igual ou acima de 50%) são destacados em azul, enquanto que aqueles com avaliação negativa (inferior a 50%) são apresentados em vermelho.

Ao avaliar um parágrafo, ele passa a ser apresentado do lado direito da interface, com um número de identificação sequencial, de acordo com a ordem que ele aparece no texto, e um *link* que permite o acesso direto à página onde ele está localizado. A

opção **Avaliar** permite que novas avaliações sejam elaboradas no contexto do material selecionado.

O segundo contexto a partir do qual uma busca baseada em perspectivas de análise de conteúdo pode ser iniciada é pela sugestão de pesquisas complementares, com a intenção de acrescentar informações referentes ao mesmo contexto ou contraditórias ao assunto que está sendo discutido no documento selecionado, conforme mostra a Figura 62.

Figura 62 - Sugestão de pesquisas complementares



Fonte: desenvolvido pela autora

A sugestão de resultados relacionados ao mesmo contexto foi projetada com fundamento no princípio hermenêutico das Múltiplas Interpretações e tem como objetivo apresentar ao usuário sugestões de outras pesquisas desenvolvidas sob a mesma perspectiva ou cenário do artigo selecionado, mas com resultados diferentes.

O material em questão, do autor Paulo Rodrigues, refere-se ao estudo do comportamento de busca com professores universitários. Como sugestões de pesquisas complementares, a interface apresentou: 1) comportamento informacional de profissionais da informação; 2) busca de informação por alunos do curso de Biblioteconomia e Ciência da Informação da FESPSP; 3) padrões de comportamento de busca por pesquisadores de biologia molecular e biotecnologia; 4) características do processo de busca dos pesquisadores da área de Psicologia da Unisinos.

Embora cada artigo relacionado tenha sido desenvolvido com um público diferente, todos discutem o mesmo assunto, existindo grande probabilidade de terem encontrado resultados que se complementam e que auxiliam o leitor a interpretar o contexto abordado de uma forma mais abrangente e aprofundada.

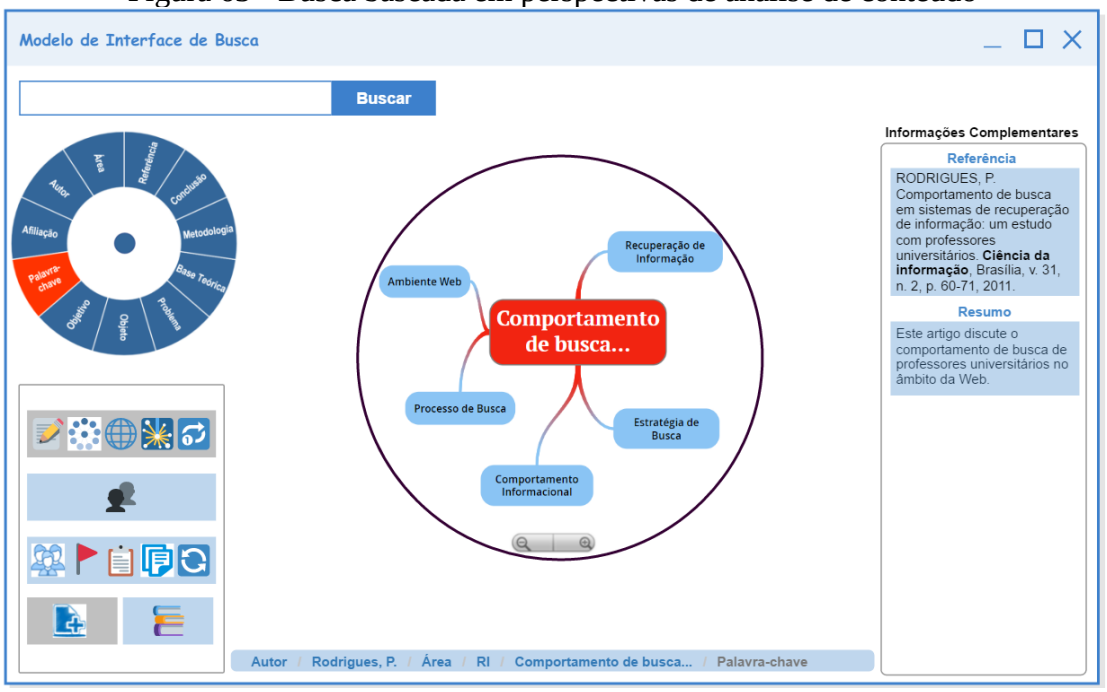
Por outro lado, a sugestão de resultados contraditórios foi desenvolvida com base no princípio hermenêutico da Dialética e tem como proposta apresentar ao usuário sugestões de pesquisas com perspectivas contraditórias, como forma de fortalecer a discussão sobre um determinado assunto e dar subsídios ao leitor para o seu entendimento de forma global e não apenas sob um único aspecto.

Para ilustrar a funcionalidade, o aborto é um assunto polêmico, discutido por inúmeros autores com pontos de vista divergentes: alguns o defendem e apresentam inúmeros argumentos para isto e outros, no entanto, são contrários e se fundamentam em uma série de princípios e razões. Neste contexto, se o documento em análise defendesse a tese da legalidade do aborto, por exemplo, os resultados contraditórios sugeridos seriam relacionados à proibição da interrupção da gravidez.

O desenvolvimento deste recurso pode ser discutido a partir do mesmo princípio dos Antônimos e Termos Contraditórios (funcionalidade apresentada no primeiro cenário). Desta forma, por meio dos relacionamentos presentes na ontologia de domínio, seria possível associar os documentos através dos termos pertencentes ao contraditório conceitual, permitindo assim a sua recuperação em um contexto de oposição. No exemplo mostrado pela Figura 62 não houve nenhuma indicação de materiais contendo teses contrárias ao assunto discutido.

Além das informações adicionais e sugestões de pesquisas complementares relacionadas à um documento recuperado, a busca também pode ser iniciada por uma das perspectivas de análise de conteúdo. No exemplo apresentado na Figura 63, o usuário selecionou a perspectiva de análise **Palavra-chave**, que foi responsável pela atualização do mapa de conceitos, permitindo a escolha de um ou mais termos para compor o critério de busca. As perspectivas de análise de conteúdo estão associadas entre si pela ontologia e oferecem uma rede de informações importantes para subsidiar o processo de recuperação de documentos correlacionados.

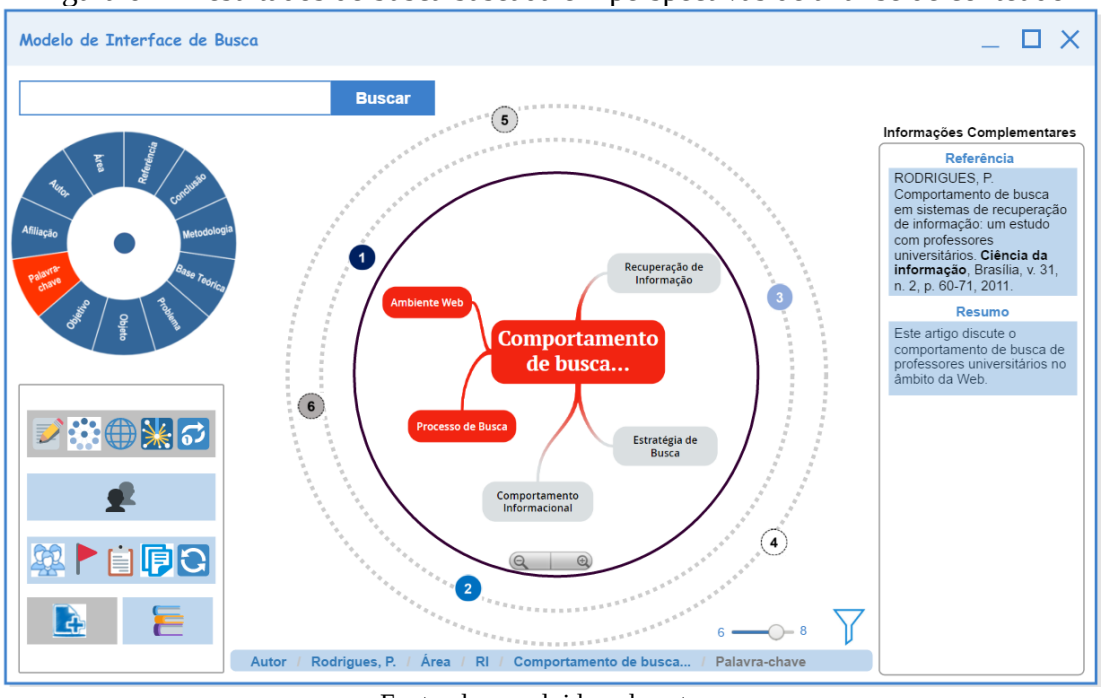
Figura 63 - Busca baseada em perspectivas de análise de conteúdo



Fonte: desenvolvido pela autora

A Figura 64 apresenta os resultados de uma busca realizada a partir de duas palavras-chave: **Ambiente Web** e **Processo de Busca**. O objetivo é recuperar documentos que também tenham sido indexados (na perspectiva Palavras-chave) pelos referidos termos. Se necessário, mais de uma perspectiva pode ser selecionada, acrescentando novos critérios para a pesquisa.

Figura 64 - Resultados de busca baseada em perspectivas de análise de conteúdo



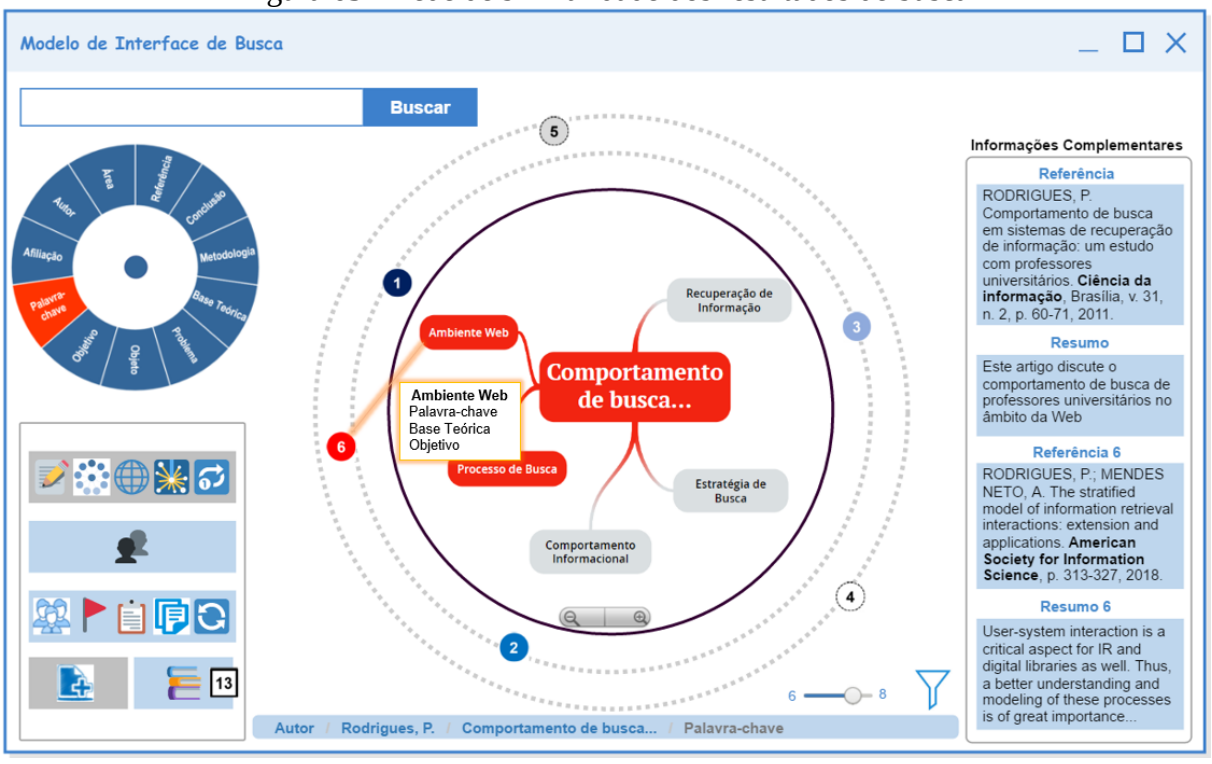
Fonte: desenvolvido pela autora

Entre as informações que a interface pode revelar, mesmo sem o pesquisador realizar qualquer leitura dos materiais pertencentes ao contexto avaliado, é que os Doc. 1, 2 e 3 foram considerados relevantes por possuírem as duas palavras-chave. Os Docs. 4, 5 e 6 possuem relação direta ou indireta com os mesmos termos (ou com apenas um deles), podendo ou não existir relações com outras perspectivas.

Caso o usuário selecione quaisquer resultados recuperados (Doc. 6, por exemplo), a interface apresentará o conjunto de relacionamentos existentes entre os dois documentos escolhidos (Doc. 2, que é o foco na análise e Doc. 6), ou seja, a rede de similaridades existente entre eles. A referência e o resumo dos materiais que estão sendo avaliados também são apresentados na área de Informações Complementares, de acordo com a Figura 65.

Quando um resultado é selecionado para ser analisado por meio das perspectivas de análise de conteúdo, a busca é realizada em todo o *corpus* documental, ou seja, todos os documentos que possuem algum tipo de relação com a expressão de busca serão recuperados, sejam aqueles que coincidem totalmente ou parcialmente. Os resultados menos relevantes têm como objetivo complementar a busca e apresentar informações que, de forma indireta, também estão relacionadas ao contexto pesquisado. Os resultados recuperados podem servir de base para refinar ou iniciar novas buscas.

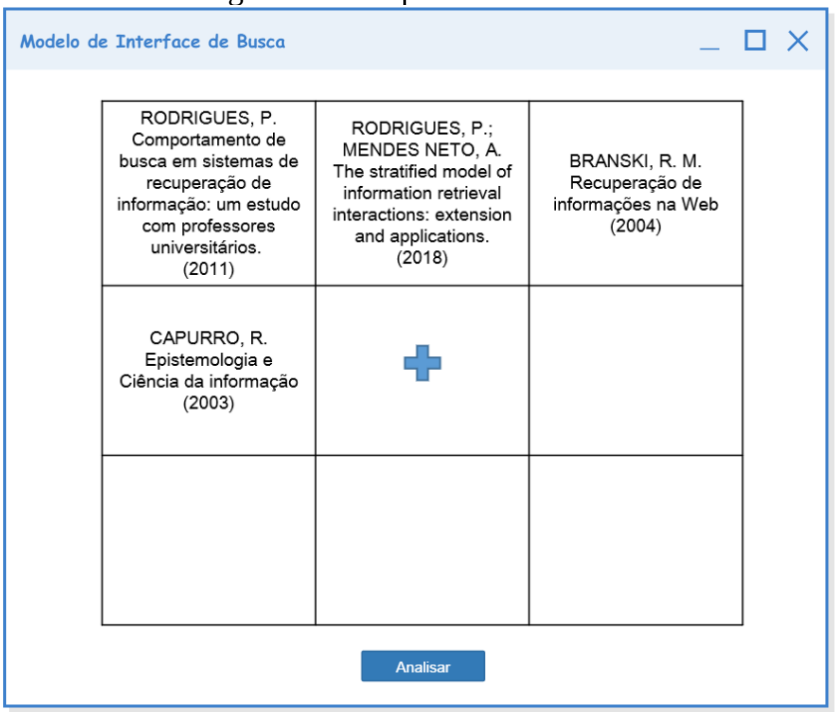
Figura 65 - Rede de similaridade dos resultados de busca



Fonte: desenvolvido pela autora

É possível ainda criar um grupo de análise de documentos específico, que gerará informações relacionadas apenas ao grupo e não ao *corpus* documental. Esta função pode ser utilizada quando o pesquisador tiver a necessidade de desenvolver um estudo sobre um tema específico e quiser realizar uma análise utilizando apenas um conjunto de documentos, com o intuito de levantar informações e fazer inferências sobre ele. Para isto, deve selecionar o item **Análise de Grupo**, disponível no ícone representado pelo número 13 (Figura 65), localizado ao lado esquerdo da interface e, posteriormente, adicionar os documentos que serão analisados. Todos os materiais devem ter sido previamente cadastrados antes do desenvolvimento desta atividade. A Figura 66 apresenta a interface para a seleção dos documentos desejados pelo pesquisador.

Figura 66 - Grupo de documentos



Fonte: desenvolvido pela autora

Este cenário apresenta, portanto, recursos para o desenvolvimento de pesquisas realizadas a partir de um determinado documento, com o intuito de encontrar outros materiais pertencentes à um mesmo contexto e que compartilham características semelhantes entre si. Esta funcionalidade é importante para auxiliar pesquisadores a elaborarem análises sobre o domínio investigado e a identificarem a rede de relacionamentos associadas a uma determinada publicação, permitindo uma análise mais abrangente e aprofundada dos resultados apresentados.

Entre os recursos oferecidos, existe a possibilidade de explorar um

determinado documento a partir de três contextos: pelas informações adicionais (comentários interpretativos, avaliação de parágrafos, indicação de erros e falhas); pela sugestão de pesquisas complementares (resultados relacionados ao mesmo contexto e contraditórios) e por uma ou mais perspectivas de análise, com diversas possibilidades de cruzamento de informações entre elas: 1) resultados do mesmo autor (perspectiva Autor); 2) resultados do mesmo autor, contendo a mesma metodologia (perspectivas Autor e Metodologia); 3) resultados com o mesmo objeto de estudo, pertencente ao mesmo autor (perspectivas Objeto e Autor) e outras.

7.3.2.5 Cenário de inclusão e análise de material

O quinto cenário deste estudo apresenta os recursos da interface de busca destinados à inclusão das publicações que farão parte do *corpus* documental disponibilizado para a pesquisa. As interações previstas têm como objetivo principal construir a rede de relacionamentos existente entre os documentos e, consequentemente, permitir a organização do *corpus* e a ampliação das relações existentes entre os conceitos da ontologia.

Neste contexto, o usuário deixa de ser apenas consumidor de informação e torna-se construtor de conhecimento, compartilhando sua perspectiva sobre um determinado assunto com os demais usuários da base de dados. Este processo deve ser realizado de forma colaborativa, por todos os pesquisadores que tiverem a intenção de inserirem documentos no repositório digital ou que observarem qualquer tipo de relacionamento não documentado durante o desenvolvimento de suas buscas.

Quadro 17 - Inclusão e análise de material

Quinto Cenário
Luiz Antônio é pós-graduando no Doutorado em Ciência da Informação e está desenvolvendo uma pesquisa na área de Recuperação de Informação. Além disso, estuda ainda dois outros assuntos correlacionados com o seu estudo: Ontologias e Interfaces Gráficas. Para escrever e fundamentar um dos capítulos da sua tese, Luiz Antônio selecionou vinte artigos relevantes que, de alguma forma, trazem discussões sobre a área. O primeiro passo é inseri-los na base de dados do repositório digital e, posteriormente, analisá-los com o intuito identificar os problemas abordados por cada um deles, descobrir as suas principais características e, assim, associá-los à rede de relacionamentos já existente. Este processo deverá ser realizado a partir da seleção de perspectivas e termos que melhor representem o conteúdo discutido, de acordo com a percepção e com o entendimento do usuário. A interface permite a atribuição de relacionamentos entre documentos e conceitos e entre conceitos, mas nunca diretamente entre documentos, ou seja, as publicações são sempre associadas umas às outras por meio de um termo presente na ontologia.

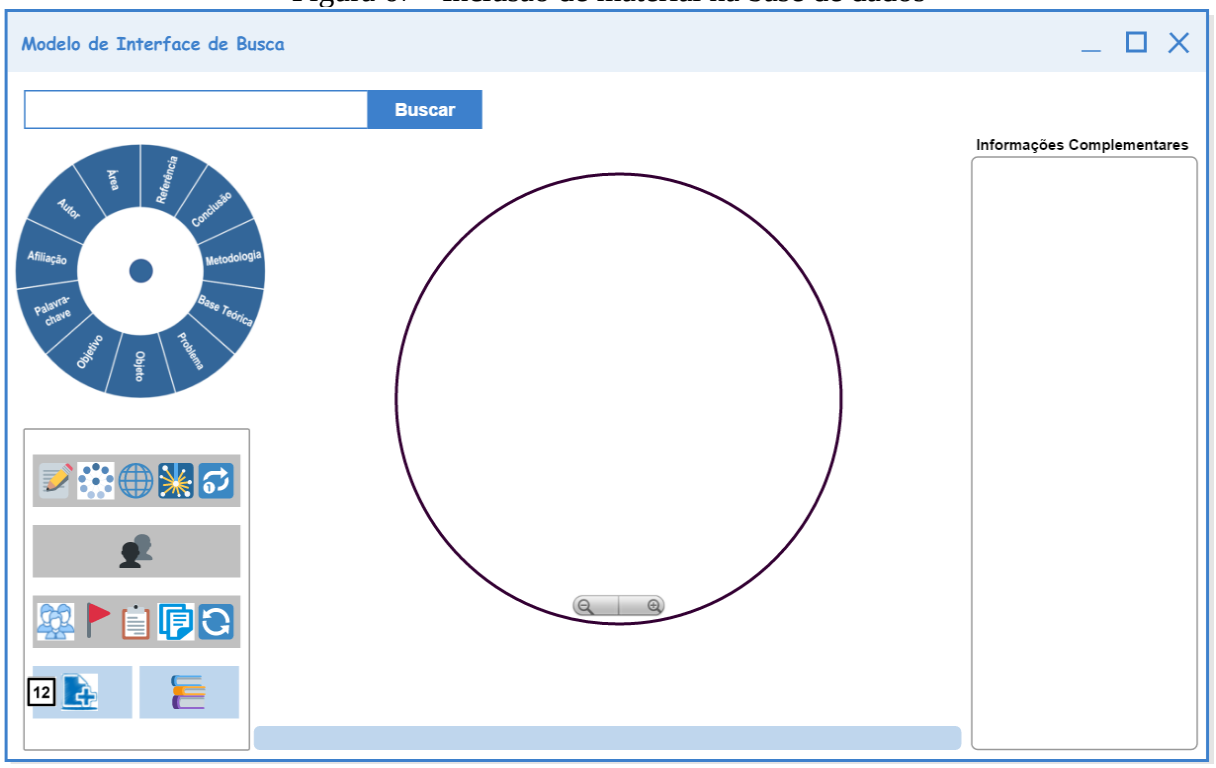
Fonte: desenvolvido pela autora

Neste cenário, o usuário utilizará a ferramenta como instrumento de colaboração e compartilhamento de informações, podendo documentar uma ou várias publicações de seu interesse. Este processo será realizado com o auxílio da ontologia de domínio, que disponibilizará os conceitos necessários para descrever os problemas e as perspectivas associadas a cada publicação.

É possível ainda criar novos conceitos e novos relacionamentos, com a intenção de ampliar ou atualizar as relações semânticas existentes na ontologia. Estes relacionamentos contribuirão para agrupar os documentos com características semelhantes, bem como para facilitar a interpretação e compreensão do usuário em relação ao domínio pesquisado, ampliando assim os resultados das buscas realizadas. As associações poderão ser realizadas entre classes (classe-classe), instâncias e classes (instância-classe) e entre instâncias (instância-instância).

O processo para adicionar um documento inicia-se com a escolha da opção **Inclusão de Material**, representada na Figura 67 pelo ícone de número 12 e, posteriormente, com o *upload* do arquivo que o usuário deseja inserir na base de dados.

Figura 67 - Inclusão de material na base de dados



Fonte: desenvolvido pela autora

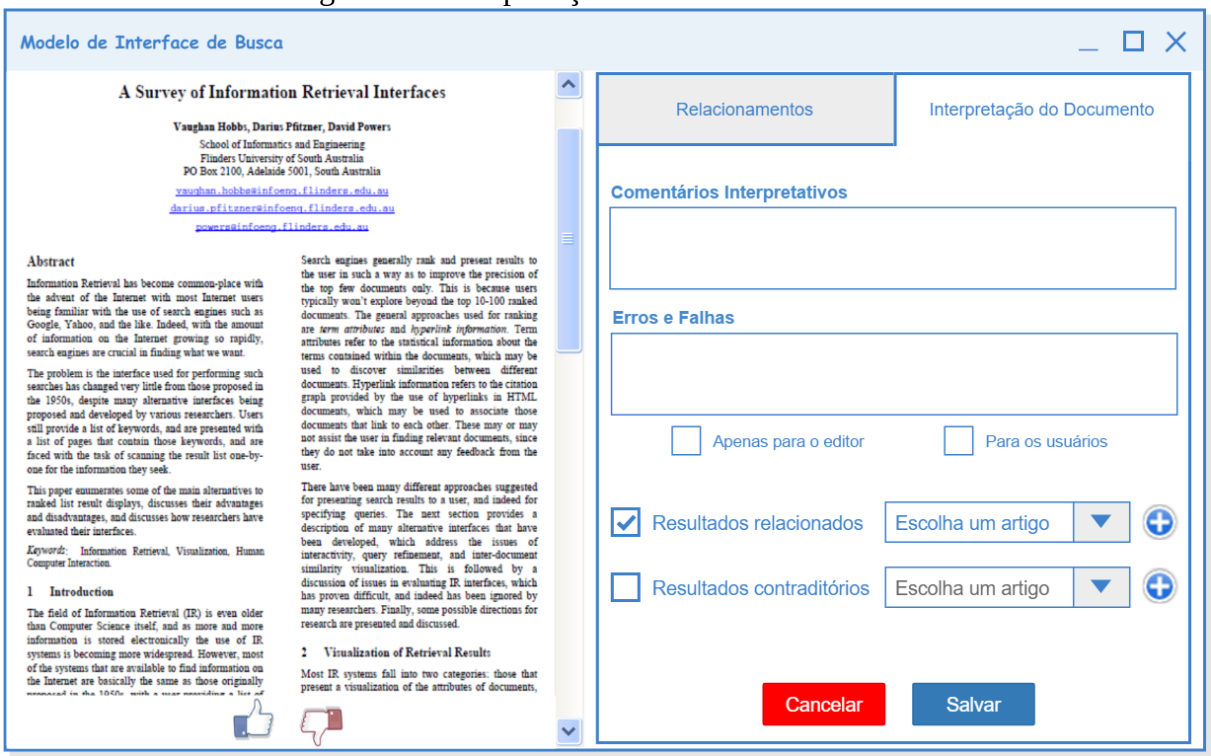
Algumas informações são carregadas automaticamente, tais como autores, palavras-chave, editora, local e tipo de publicação. As demais devem ser acrescentadas a

partir de dois contextos: 1) **interpretação do documento**, que diz respeito às características que auxiliam o processo de compreensão do conteúdo discutido e 2) **relacionamentos**, que são as associações entre o material e a ontologia.

No contexto da interpretação do documento que está sendo inserido, o pesquisador pode documentar a sua percepção sobre o assunto que está sendo discutido, acrescentando comentários interpretativos, erros ou falhas que tenham sido observadas e materiais que contenham resultados relacionados ou contraditórios. Essas informações, embora não sejam obrigatórias, contribuirão para que outros usuários possam utilizá-las durante o processo de leitura e interpretação do texto.

A Figura 68 mostra a interface para a realização desta tarefa, que está dividida em duas partes. A primeira (lado esquerdo), refere-se à área de leitura e permite que o usuário tenha acesso ao material que está sendo inserido na base de dados, e a segunda, por sua vez, possibilita a inserção das informações relacionadas o documento. Outra funcionalidade que também pode ser realizada neste contexto é a **Avaliação dos Parágrafos**. Ao selecionar a aba **Interpretação do Documento**, os ícones de *like* e *dislike* são habilitados e apresentados na parte inferior da área de leitura, permitindo que os parágrafos possam ser avaliados.

Figura 68 - Interpretação de um novo documento

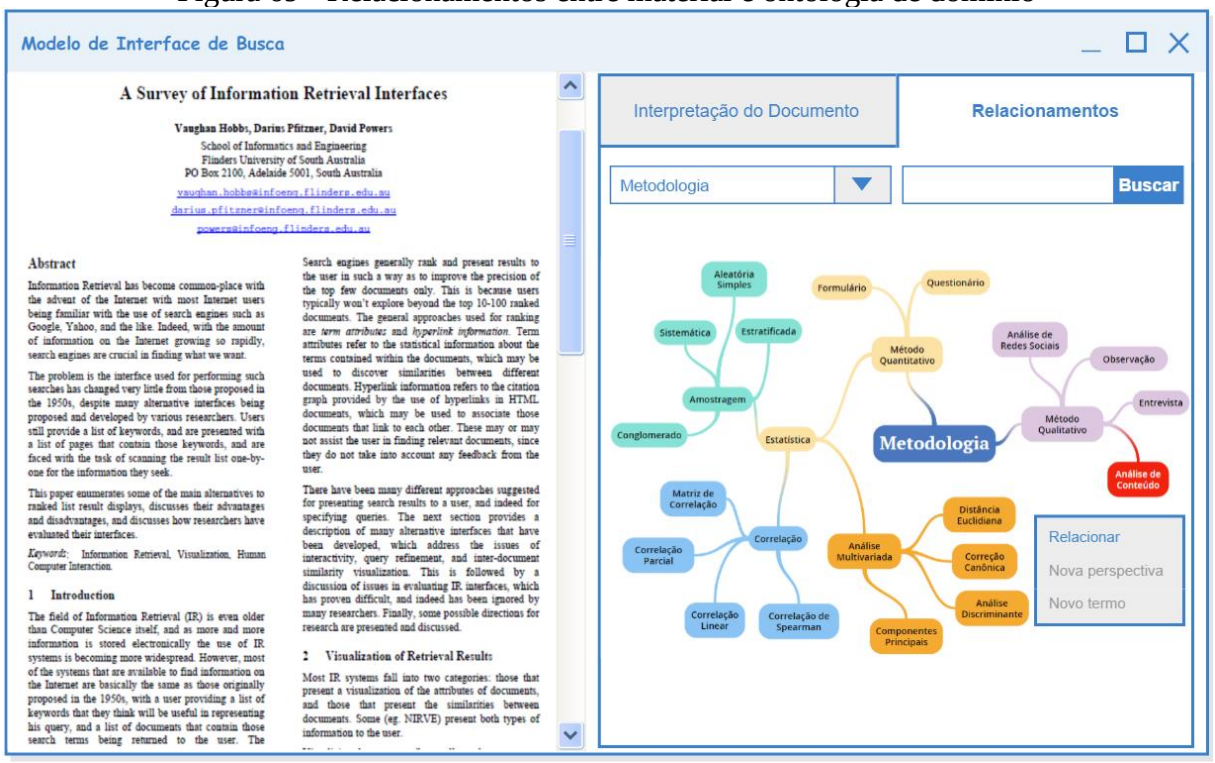


Fonte: desenvolvido pela autora

No caso dos relacionamentos, o pesquisador pode iniciar a documentação por qualquer uma das perspectivas disponíveis ou, se for necessário, incluir outras classes que melhor atendam as características do material que será inserido. Cada classe ou perspectiva de análise refere-se à uma parte da rede semântica construída para identificar pontos de associação e de semelhança entre os documentos.

A Figura 69 apresenta a interface para inserção e análise de relacionamentos de um novo documento, que possibilita a seleção das perspectivas disponíveis para as realizar as associações, como também a inserção de novos elementos na ontologia, tais como: autores, perspectivas e termos. Ao escolher uma determinada perspectiva de análise, a ferramenta atualiza o mapa de conceitos relacionados a ela e mostra ao usuário, de forma gráfica, todos os elementos que podem ser utilizados. É possível ainda buscar um termo através da caixa de busca, independente da perspectiva a qual ele pertença.

Figura 69 - Relacionamentos entre material e ontologia de domínio



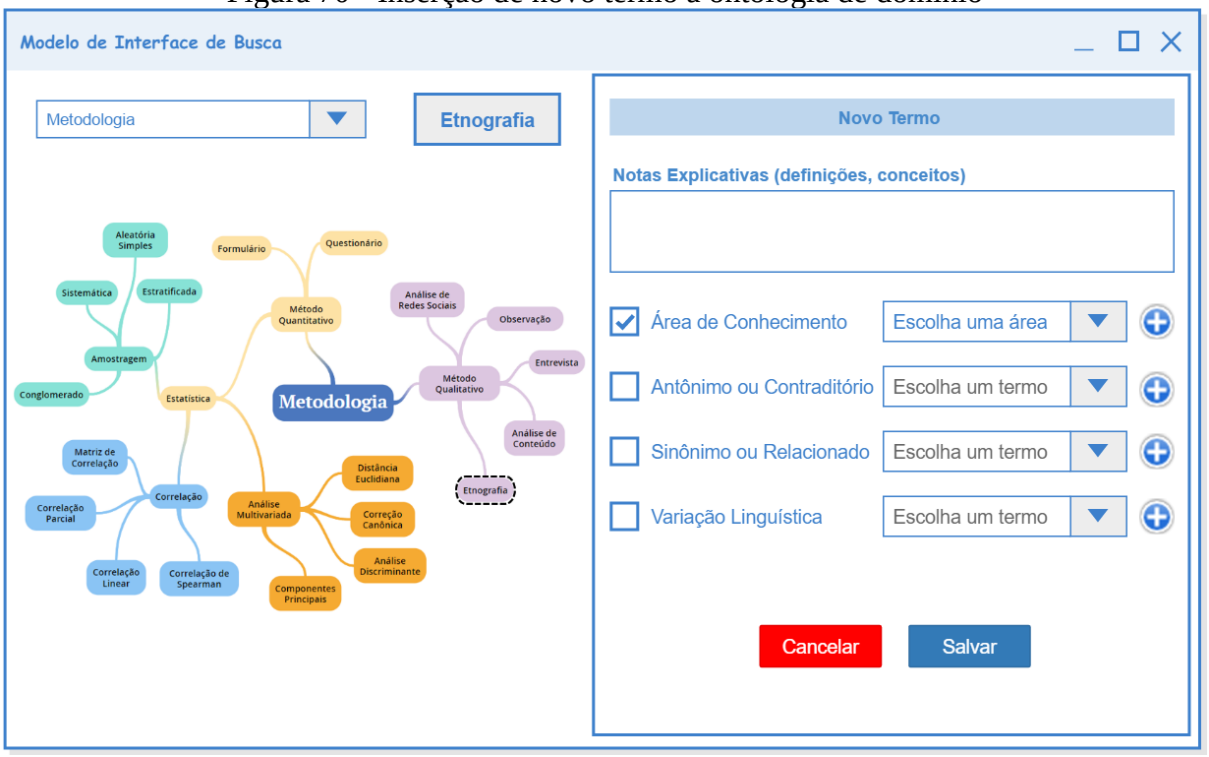
Fonte: desenvolvido pela autora

A partir do mapa de conceitos, o usuário poderá analisar os termos disponibilizados no domínio e selecionar os mais adequados para representar o conteúdo da publicação, de acordo com os seus conhecimentos e a sua percepção sobre as informações discutidas. Para associar o documento à um conceito, basta selecioná-lo e, a partir do menu suspenso, escolher a opção **Relacionar**, conforme pode ser observado na Figura 69. Neste

caso, o usuário optou por associar o documento à perspectiva **Metodologia**, através do termo **Análise de Conteúdo**.

É possível ainda acrescentar novos termos na ontologia e, relacionar um conceito a outro, quando existem uma ou mais características de proximidade entre eles. Para isto, basta selecionar a opção **Novo Termo**, através do menu suspenso.

Figura 70 - Inserção de novo termo à ontologia de domínio

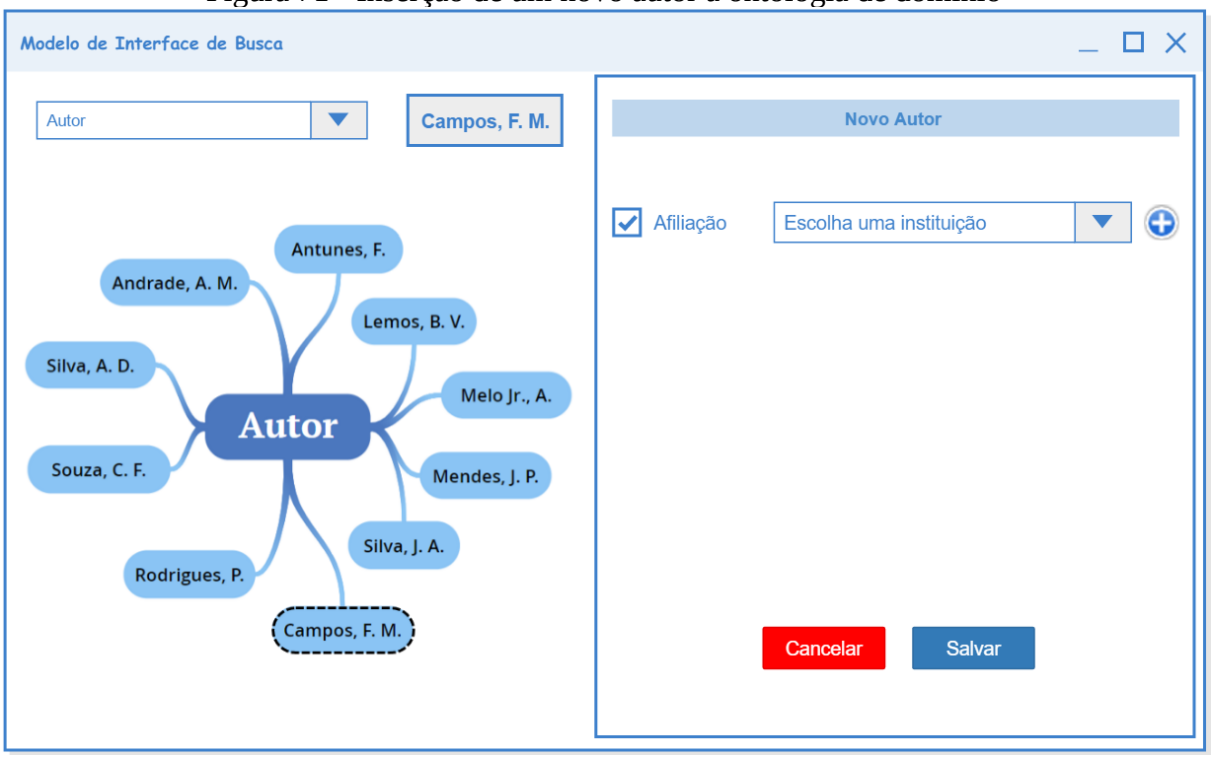


Fonte: desenvolvido pela autora

O primeiro passo para a inserção de um novo termo (**Etnografia**) é a seleção da perspectiva de busca da qual ele fará parte. Por meio do mapa de conceitos, basta selecionar o termo ao qual ele será relacionado e dar dois cliques do *mouse* sobre ele. A ferramenta, então, disponibilizará um novo conceito, que deverá ser nomeado pelo usuário, que também é responsável pelo preenchimento das demais informações sobre ele, tais como: notas explicativas, área de conhecimento, antônimo ou contraditório, sinônimo ou relacionado e variação linguística.

O processo para a inserção de um novo autor é semelhante ao adotado para um novo termo. Para que esta interação seja possível, basta selecionar a perspectiva **Autor**, de acordo com a Figura 71. Apenas autores inseridos no mapa de conceitos poderão ser associados aos documentos disponibilizados na base de dados.

Figura 71 - Inserção de um novo autor à ontologia de domínio



Fonte: desenvolvido pela autora

Para indexar os documentos e oferecer aos usuários múltiplos pontos de acesso, pode-se utilizar termos de síntese ou de análise. Os primeiros estão relacionados com a compreensão elementar e são extraídos do próprio texto, tal como ocorre na indexação automática. Já os termos de síntese referem-se à elementos de significação e interpretação do contexto avaliado e, portanto, não precisam necessariamente fazer parte do documento.

O relacionamento entre elementos de informação é realizado de maneira particular, por cada usuário, e está diretamente relacionado à forma como este sujeito compreende e interpreta os contextos analisados. Essa atribuição de significado a contextos depende das características pessoais (conhecimento sobre o assunto, sexo, idade, nível de escolaridade, orientação política; sociais (categoria profissional, organização) e ambientais (país, legislação, cultura informacional) e, consequentemente, das perspectivas de cada pesquisador em relação ao assunto analisado. Isto quer dizer que, no contexto deste cenário, uma relação entre duas publicações pode ser compreendida e interpretada de maneira completamente diferente, o que possibilita a ampliação das inferências sobre os assuntos e, consequentemente, do conhecimento sobre o domínio.

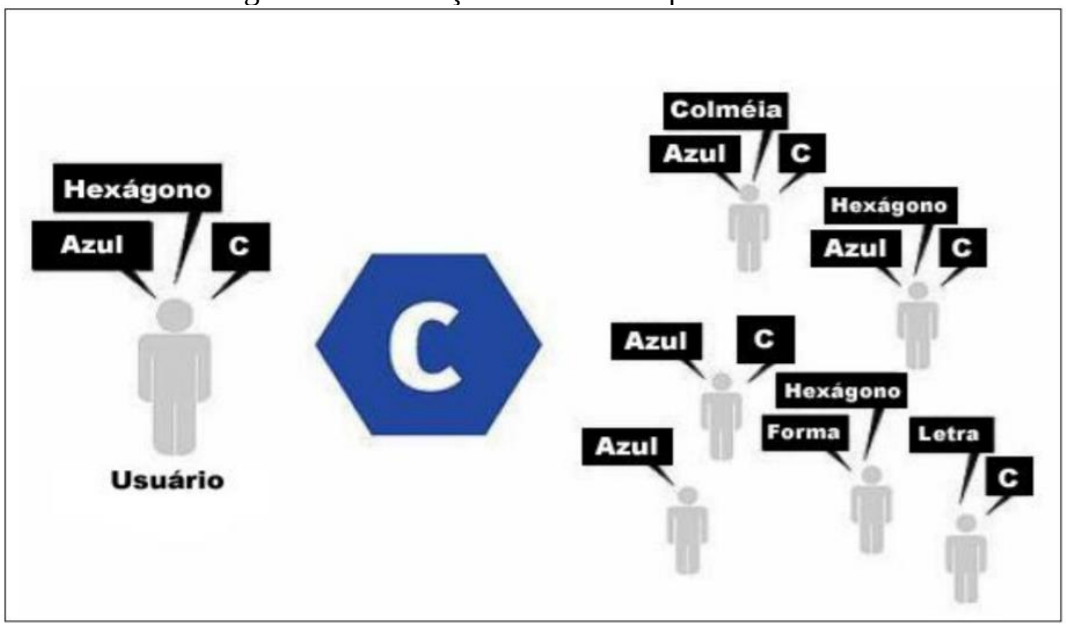
Esta funcionalidade está subsidiada em dois princípios da hermenêutica: da interação entre pesquisadores e sujeitos e das múltiplas interpretações (GADAMER, 1999; KLEIN; MYERS, 1999). O princípio da interação entre pesquisadores e sujeitos discute que

um texto não pode ser visto apenas a partir do entendimento de um único intérprete e que o seu significado deve ser sempre negociado e consensuado com outros pesquisadores, sendo, portanto, um resultado da fusão de vários horizontes. Se assemelha, portanto, a algo produzido em um diálogo e não como algo meramente reproduzido por um intérprete. Esta interação é especialmente útil, pois permite que pesquisadores alterem seus horizontes a partir da apropriação de preconceitos de outros pesquisadores com os quais está interagindo, bem como modifiquem suas ações a partir dos horizontes alterados.

O princípio das múltiplas interpretações afirma que o foco de um processo de interpretação e aquilo que o texto tem a dizer para o pesquisador. E que cada pesquisador possui um horizonte diferente, que é influenciado pelo seu contexto social, preconceitos, valores, nível de conhecimento, *status* social, experiencias vivenciadas, entre outros. Desta forma, cada pesquisador compreenderá um documento a partir do seu mundo, revelando assim, uma multiplicidade de aspectos relacionados a ele.

A Figura 72 apresenta um exemplo de como usuários com perspectivas diferentes enxergam o mesmo contexto, no caso, representado por C.

Figura 72 - Definição de conceitos pelos usuários



Fonte: Nascimento (2008, p.32)

Como o processo é realizado de forma colaborativa, o principal desafio é encontrar um conjunto de termos que melhor represente as diversas perspectivas de um documento. Desta forma, cada interpretação é importante para aumentar a expressividade das informações dentro do contexto avaliado, bem como para melhorar a semântica entre as

publicações ou entre os elementos que a constituem.

A organização dos documentos realizada de forma colaborativa deve interferir positivamente nos resultados das buscas realizadas, uma vez que disponibilizará aos usuários uma multiplicidade de conceitos e interpretações sobre um mesmo assunto, bem como uma rede de relacionamentos entre os documentos, que contribuirão para que sua análise sobre o domínio seja mais ampla e aprofundada. Isto é particularmente útil, principalmente para pesquisadores que não possuem conhecimento de domínio aprofundado e, encontram dificuldades para realizar inferências entre os assuntos e documentos disponibilizados na base de dados.

Todas as sugestões de relacionamentos entre conceitos e documentos é coordenada pelo indexador responsável pelo repositório digital, que decide se serão aceitas ou rejeitadas. As contribuições aceitas serão integradas à ontologia de domínio.

7.4 AVALIAÇÃO DO MODELO PROPOSTO

O “Modelo de interface de busca e recuperação de informação” proposto neste estudo foi avaliado com base no método proposto por Snyder (2003), que possui cinco etapas: preparação, coleta de dados, intepretação, consolidação dos resultados e relato dos resultados. Os ensaios de usabilidade foram realizados com treze participantes, sendo três na fase de pré-testes e dez, na fase de testes. Os três pré-testes tiveram como principal objetivo, revisar o material de apoio desenvolvido, a sequência das atividades propostas e a coerência com os protótipos elaborados. Os testes, por sua vez, tiveram a intenção de analisar se os requisitos foram satisfeitos, identificar a viabilidade de aceitação do usuário e possíveis erros de interação.

Os testes de usabilidade foram organizados com base em um roteiro de atividades sequenciais, previstas para cada cenário elaborado (Apêndice A). Os protótipos utilizados foram elaborados com o auxílio da ferramenta de prototipação Pencil Project. Além disso, adotou-se ainda nesta etapa a observação participante, com o objetivo de levantar maiores informações sobre o comportamento e as dificuldades dos usuários durante a interação com a interface.

Além dos testes de usabilidade, foi aplicado ainda um questionário de avaliação (Apêndice B), com o objetivo de identificar a percepção do usuário sobre a interface de busca proposta neste estudo, bem como referente às funcionalidades que foram projetadas com base nos princípios da hermenêutica e na utilização de ontologias.

7.4.1 Pesquisa Piloto

A pesquisa piloto (ou pré-teste) foi realizada na Universidade de Coimbra, no período compreendido entre 20 e 30 de novembro de 2017, e contou com três participantes do sexo masculino, sendo dois deles pertencentes à faixa etária de 24 a 33 anos e um, de 34 a 43 anos.

A escolha das características dos participantes foi intencional: todos pesquisadores cursando o doutorado, com bons conhecimentos de informática e com ampla experiência no uso de ferramentas de busca, tais como o Google, repositórios digitais e bibliotecas virtuais. Dois deles pertencem ao curso de Engenharia Informática e um ao de Ciência da Informação. Além das atividades acadêmicas, todos estão envolvidos com profissões que exigem a busca constante de materiais científicos para o desenvolvimento de suas atividades, sendo um professor, um programador e um engenheiro de sistemas.

Em função da vasta experiência dos participantes com as ferramentas de busca e desenvolvimento de sistemas, informações significativas foram identificadas nesta etapa (problemas de usabilidade identificados, erros cometidos, comentários realizados durante as atividades, sugestões de melhorias) e deram subsídios para o primeiro ciclo de correções realizadas na interface proposta (avaliação – *design* de alternativas – prototipação). Também foi possível reestruturar o material de apoio desenvolvido para dar suporte aos testes (roteiro de atividades e questionário de avaliação), bem como avaliar se as atividades propostas estavam coerentes com os protótipos desenvolvidos.

7.4.2 Ensaios de Usabilidade

Os testes de usabilidade foram realizados com dez participantes, de acordo com o perfil apresentado na Tabela 5. A maioria deles é do sexo masculino (6), possui até 43 anos (6), está cursando ou já concluiu o mestrado ou doutorado (8), é da área da Ciência da Computação e/ou Informação (5) e exerce atividade profissional como professor (6). Todos afirmam que possuem conhecimentos satisfatórios de informática (avaliados como Ótimo e Bom) e utilizam as ferramentas de busca diariamente.

Por ser um projeto de uma interface para áreas especializadas, assim como os repositórios digitais, optou-se por incluir também participantes com outras formações além de Ciência da Computação e Ciência da Informação, tais como Alimentos, Contábeis, Direito, História e Matemática que, apesar de desenvolverem atividades acadêmicas e profissionais

que exigem a realização de pesquisas com muita frequência, são usuários que costumam não utilizar recursos avançados das ferramentas de busca. Portanto, é importante para o contexto da avaliação poder identificar as dificuldades deste grupo também.

Tabela 5 - Perfil dos participantes

Faixa Etária		De 24 a 33	De 34 a 43	De 44 a 53	54 ou mais	Total
Sexo	Feminino	2		1	1	4
	Masculino	1	3		2	6
Formação	Especialização	2				2
	Mestrado (Cursando)	1	1			2
	Mestrado (Concluído)		1	1	1	3
	Doutorado (Cursando)		1			1
	Doutorado (Concluído)				2	2
Área	Ciência da Computação	1	2			3
	Ciência da Informação				2	2
	Ciência de Alimentos		1			1
	Ciências Contábeis	1				1
	Direito	1				1
	História			1		1
	Matemática				1	1
Profissão	Advocacia	1				1
	Bibliotecário				1	1
	Gestor de Recursos Humanos	1				1
	Professor	1	2	1	2	6
	Técnico		1			1
Instituição	IFPR (Jacarezinho)	1	2			3
	IFPR (Jaguariaíva)		1			1
	UNESP (Marília)				1	1
	UNOPAR (Londrina)				1	1
	UTFPR (Londrina)				1	1
	Empresa Privada (Londrina)	2		1		3

Fonte: desenvolvido pela autora

Os testes foram organizados com base em um roteiro de atividades sequenciais, elaborado na fase de preparação, que é a primeira etapa do processo de avaliação. As tarefas previstas foram subdivididas de acordo com os recursos de cada cenário de utilização da interface. Cada tarefa foi planejada a partir de um objetivo específico, que está associado à um requisito funcional (RF), com a intenção de compreender a interação do usuário em um determinado contexto. Todos os requisitos funcionais definidos no início do projeto da interface foram apresentados no Quadro 9.

Os ensaios de usabilidade foram realizados individualmente com os participantes, no período compreendido entre 20 de setembro e 10 dezembro de 2018. Cada

teste foi iniciado com a apresentação dos principais conceitos e funcionalidades que envolvem a interface de busca, tais como: tipos de busca suportadas, mapa de conceitos, plano de contexto e de informação, perspectivas de análise, área de resultados, entre outros. Posteriormente, as tarefas e os respectivos protótipos foram apresentados em ordem cronológica, de acordo com o roteiro de atividades, permitindo a simulação da interação.

O roteiro de atividades contendo todas as tarefas realizadas durante os ensaios foi disponibilizado no Apêndice A. Suas principais características, incluindo as tarefas envolvidas, recursos avaliados e quantidade de questões dedicadas a cada recurso, são apresentadas na Tabela 6. Algumas tarefas podem avaliar mais de um recurso, no entanto, optou-se por relacionar uma tarefa à apenas uma funcionalidade.

Tabela 6 - Características do roteiro de atividades

Tarefas	Cenário	Recursos Avaliados	Quantidade
De 1 a 19	Cenário 1	Busca baseada em palavras-chave	5
		Resultados da busca	4
		Perspectivas de análise	3
		Mapa de conceitos	2
		Antônimos e termos conceitualmente contraditórios	1
		Áreas de conhecimento	1
		Notas explicativas	1
		Sinônimos e termos relacionados	1
		Variações linguísticas	1
De 20 a 25	Cenário 2	Mapa de conceitos	3
		Busca baseada em estrutura terminológica	2
		Caixa de busca	1
De 26 a 33	Cenário 3	Mapa de conceitos	3
		Resultados da busca	3
		Busca a partir de perspectivas de análise de contexto	1
		Histórico do autor	1
De 34 a 46	Cenário 4	Busca a partir de perspectivas de análise de conteúdo	3
		Resultados da busca	3
		Mapa de conceitos	2
		Comentários interpretativos	1
		Erros e falhas	1
		Parágrafos avaliados	1
		Resultados relacionados	1
		Resultados contraditórios	1
De 47 a 50	Cenário 5	Inclusão e análise de documentos	1
		Inclusão de relacionamentos (termos e documentos)	1
		Inclusão de termos na ontologia	1
		Interpretação dos documentos	1

Fonte: desenvolvido pela autora

Cada tarefa proposta teve a intenção de entender o comportamento do usuário, a partir de uma perspectiva da interface. A primeira perspectiva teve a intenção de analisar a compreensão do usuário em relação à **interpretação da informação** (II), incluindo os conceitos apresentados na interface, o domínio do problema e os resultados recuperados em uma busca. A segunda perspectiva avaliada teve o intuito de verificar a **facilidade de acesso à funcionalidade** (FAF), bem como o entendimento em relação aos elementos do *layout* e à sequência de ações que devem ser realizadas para completar cada tarefa. Na terceira perspectiva buscou-se avaliar a **visibilidade da informação** (VI), ou seja, a acessibilidade das funcionalidades pelos diferentes tipos de usuário. E, por último, a quarta perspectiva teve a intenção de estudar a **facilidade de aprendizagem** (FA) dos recursos da interface, que se refere ao processo pelo qual o usuário aprende a utilizar uma funcionalidade para o desenvolvimento de determinadas tarefas.

Desta forma, durante a elaboração das atividades, foi possível analisar as interações e observar as dificuldades encontradas, os comentários realizados, o entendimento dos termos e do *layout* da interface, a acessibilidade aos recursos projetados e, por fim, se o usuário foi capaz de percorrer a sequência de ações necessária para concluir a atividade e produzir *feedback* sobre a interação, sem a necessidade de intervenção por parte do avaliador.

O primeiro cenário foi o mais explorado nos testes realizados em função da quantidade de recursos discutidos, embora eles não sejam exclusivos deste contexto e possam ser acessados também nos demais. Em função da sua importância, o mapa de conceitos foi analisado em praticamente todos os cenários (exceto no quinto por se tratar de um cenário relacionado à inclusão de materiais).

Tabela 7 - Resultados dos ensaios de usabilidade realizados no primeiro cenário

Tarefa	**RF**	**Recurso Avaliado**	**Perspectiva**	**Tarefa Realizada**		
				Sim, com facilidade	Sim, com dificuldade	Não
01	01	Busca baseada em palavras-chave	FAF	5	3	2
02	01	Busca baseada em palavras-chave	FAF	6	2	2
03	21	Mapa de conceitos	II	7	1	2
04	05	Perspectivas de análise	VI	3	2	5
05	09	Perspectivas de análise	FAF	8	1	1
06	01	Busca baseada em palavras-chave	FA	4	2	4
07	09	Perspectivas de análise	FA	5	2	3
08	01	Busca baseada em palavras-chave	VI	4	2	4
09	12	Resultados da busca	II	5	2	3
10	20	Resultados da busca	FAF	6	2	2
11	17	Resultados da busca	II	4	1	5

Tabela 7 - Resultados dos ensaios de usabilidade realizados no primeiro cenário (continuação...)

Tarefa	RF	Recurso Avaliado	Perspectiva	Tarefa Realizada		
				Sim, com facilidade	Sim, com dificuldade	Não
12	18	Resultados da busca	II	4	2	4
13	11	Mapa de conceitos	FAF	4	3	3
14	23	Notas explicativas	FAF	3	3	4
15	26	Áreas de conhecimento	FAF	5	2	3
16	24	Sinônimos e termos relacionados	II	4	2	4
17	25	Variações linguísticas	II	3	1	6
18	27	Antônimos e termos contraditórios	II	1	2	7
19	01	Busca baseada em palavras-chave	FA	3	2	5
Média de sucesso das tarefas				**4,4**	**2,0**	**3,6**

Fonte: desenvolvido pela autora

Os resultados dos testes de usabilidade realizados no primeiro cenário, referente às buscas baseadas em palavras-chave elaboradas pelo próprio usuário, são apresentados na Tabela 7, de acordo com os resultados de cada tarefa: ordem de aplicação, requisito funcional, principal recurso avaliado, perspectiva de avaliação e análise da interação: se a tarefa foi concluída com facilidade, com dificuldade ou se não foi concluída.

O recurso de **busca baseada em palavra-chave** foi analisado no desenvolvimento das tarefas 01, 02, 06, 08 e 19. Neste contexto, a média de interações realizadas com facilidade foi de 4,4, com algum tipo de dificuldade foi de 2,2 e de não realizadas foi de 3,4. A tarefa que os usuários tiveram maior dificuldade foi a 19, que solicitava a elaboração de uma expressão de busca na área da Endocrinologia. Como nenhum dos participantes tem conhecimentos sobre este campo, foi possível entender as dificuldades da elaboração de expressões de busca em domínios desconhecidos pelo usuário.

Os usuários que conseguiram realizar a tarefa, utilizam termos muito genéricos para compor a expressão de busca, tais como: Endocrinologia, Doenças do Sistema Endócrino e Problemas Tratados pela Endocrinologia. Nenhum dos participantes utilizou conceitos relacionados à área, nem mesmo o termo Tireoide, que é o mais comum. Isto significa que possuir conhecimento de domínio, incluindo termos e relacionamentos, é fundamental para o desenvolvimento de buscas em qualquer área.

Quanto mais informações o usuário tiver sobre o assunto que está sendo pesquisado, mais opções de vocabulário ele possui para realizar as buscas e explorar o conteúdo. Também é necessário compreender como os termos se relacionam dentro de um contexto, para que todos os resultados possam ser entendidos. Assim, se o usuário não souber,

por exemplo, que o sistema endócrino está relacionado com as atividades do pâncreas, através da produção de insulina, dificilmente ele conseguirá interpretar resultados recuperados referentes a estes elementos.

O recurso **resultados de busca** foi observado no desenvolvimento das tarefas 09, 10, 11 e 12. A média de interações realizadas com sucesso (com facilidade ou com algum tipo de dificuldade) foi de 6,5, contrapondo-se à 3,5 de insucesso. A maior dificuldade dos participantes foi compreender o conceito de similaridade entre resultados e termos de busca, bem como entre dois resultados recuperados (Tarefas 11 e 12). Observou-se ainda, em um primeiro momento, confusão no entendimento da relevância dos resultados retornados, que é representada pela gradação das cores (do azul escuro para o cinza) e pela proximidade com o mapa de conceitos (Tarefa 9).

As **perspectivas de análise** foram examinadas durante as tarefas 4, 5 e 7, apresentando uma média de sucesso (interações realizadas com facilidade ou dificuldade) de 7,0 e de insucesso de 3,0. A maior dificuldade observada neste contexto foi relacionada à compreensão do usuário em relação ao acesso às perspectivas de análise, bem como a sua influência nos resultados das pesquisas realizadas (Tarefa 4).

A interação do usuário com **mapa de conceitos** foi analisada nas tarefas 03 e 13, apresentando 7,5 de sucesso e 2,5 de insucesso nas atividades realizadas. De forma geral, os usuários conseguiram interpretar, sem muitas dificuldades, as tarefas apoiadas pelo mapa de conceitos, embora ainda neste momento não tenha sido possível explorar todas as possibilidades de busca que ele oferece.

O mapa de conceitos pode ser avaliado de forma mais criteriosa nas questões que se referem à aquisição de conhecimento adicional sobre um termo de busca selecionado, incluindo os recursos de **notas explicativas**, **áreas de conhecimento**, **sinônimos e termos relacionados**, **variações linguísticas** e **antônimos e termos conceitualmente contraditórios** (Tarefas 14, 15, 16, 17 e 18, respectivamente). Neste contexto, a média de interações realizadas com sucesso foi de apenas 5,2, contrapondo-se à 4,8 de insucesso.

Por serem recursos mais específicos e oferecerem possibilidades de busca inéditas, que não estão presentes nos buscadores mais utilizados, houve uma maior dificuldade por parte dos usuários em compreender os objetivos para os quais foram projetados. A funcionalidade mais complicada de ser interpretada foi a que se refere à sugestão de antônimos e termos contraditórios para complementar a busca realizada (Tarefa 18). Seria necessário, portanto, o desenvolvimento de inúmeras atividades neste âmbito, com

o intuito de possibilitar um melhor entendimento e um maior aprendizado sobre o recurso e sua contribuição nas buscas realizadas.

Tabela 8 - Resultados dos ensaios de usabilidade realizados no segundo cenário

Tarefa	**RF**	**Recurso Avaliado**	**Perspectiva**	**Tarefa Realizada**		
				Sim, com facilidade	Sim, com dificuldade	Não
20	02	Busca (estrutura terminológica)	FAF	2	6	2
21	06	Mapa de conceitos	FAF	7	1	2
22	02	Mapa de conceitos	FAF	5	1	4
23	21	Busca (estrutura terminológica)	II	2	1	7
24	06	Caixa de busca	VI	4	4	2
25	21	Mapa de conceitos	II	7	3	0
Média de sucesso das tarefas				**4,5**	**2,7**	**2,8**

Fonte: desenvolvido pela autora

A Tabela 8 apresenta os resultados dos ensaios de usabilidade realizados no contexto do segundo cenário, que se refere às buscas baseadas em estrutura terminológica. Em função da sua importância, o **mapa de conceitos** foi a funcionalidade mais observada (Tarefas 21, 22 e 25). Tal como foi apontado no primeiro cenário, os usuários não tiveram grandes dificuldades para desenvolver as tarefas a partir deste recurso, apresentando uma média de 7,0 de sucesso contra 3,0 de insucesso, principalmente por elas terem sido apresentadas de forma simplificada, com interações que se resumem à navegação nos conceitos (Tarefa 21), seleção de termos (Tarefa 22) e análise da estrutura terminológica de um domínio (Tarefa 25).

A principal dificuldade foi observada durante o desenvolvimento da tarefa 22, na qual 4 usuários não conseguiram presumir a possibilidade de arrastar um termo do mapa de conceitos para a caixa de busca. De forma geral, eles relacionam a caixa de texto com a interface do Google e consideram apenas a digitação como alternativa de interação com este recurso. No entanto, apesar da dificuldade observada inicialmente, a Tarefa 25 mostrou que a possibilidade de arrastar os termos do mapa de conceitos é uma função facilmente aprendida e memorizada pelos usuários.

Ainda em relação à Tarefa 25, os dados levantados permitiram uma comparação com a Tarefa 19 do cenário anterior, que foi desenvolvida no mesmo contexto. Em ambas atividades, foi solicitado ao usuário a elaboração de expressões de busca sobre uma temática desconhecida, que é a Endocrinologia. No primeiro cenário nenhum recurso de apoio foi oferecido, exigindo a elaboração de palavras-chave pelo próprio participante, sendo que 5 deles não conseguiram completar a tarefa. No segundo cenário, no entanto, uma estrutura

terminológica (mapa de conceitos) deu suporte à atividade, contribuindo para que todos conseguissem realizar a atividade, mesmo sem conhecimentos sobre o assunto pesquisado. Estes dados confirmam o potencial do mapa de conceitos para minimizar as dificuldades do usuário nas buscas realizadas em domínios desconhecidos.

As **buscas baseadas em estrutura terminológica** foram avaliadas no desenvolvimento das tarefas 20 e 23, obtendo uma média de sucesso de 5,5 e de insucesso de 4,5. Entre as dificuldades observadas, alguns usuários ficaram confusos para iniciar a busca por meio do mapa de conceitos, ou seja, a sequência de ações que devem ser realizadas não está evidente (Tarefa 20). Além disso, houve dificuldade para compreender que o contexto de um termo também é adicionado na busca, em função das relações semânticas presentes na ontologia. Neste sentido, 7 usuários não conseguiram deduzir que o termo utilizado na pesquisa (Sistemática), selecionado a partir da estrutura terminológica, não possuía qualquer tipo de ambiguidade. Estes dados refletem o desconhecimento dos usuários em relação ao uso das ontologias e suas respectivas contribuições.

A **caixa de busca** pode ser utilizada tanto para a elaboração de expressões, nas buscas baseadas em palavras-chave, quanto para identificar um conceito no mapa de conceitos, em buscas baseadas em estrutura terminológica. Este recurso foi avaliado na Tarefa 25, sendo que 8 participantes conseguiram completar a atividade.

Tabela 9 - Resultados dos ensaios de usabilidade realizados no terceiro cenário

Tarefa	**RF**	**Recurso Avaliado**	**Perspectiva**	**Tarefa Realizada**		
				Sim, com facilidade	Sim, com dificuldade	Não
26	03	Busca (perspectivas de contexto)	FAF	4	4	2
27	03	Mapa de conceitos	VI	7	2	1
28	16	Mapa de conceitos	II	4	1	5
29	29	Histórico do autor	VI	5	3	2
30	22	Mapa de conceitos	FA	6	1	3
31	17	Resultados da busca	II	5	3	2
32	12	Resultados da busca	II	7	1	2
33	18	Resultados da busca	II	6	2	2
Média de sucesso das tarefas				**5,5**	**2,1**	**2,4**

Fonte: desenvolvido pela autora

Os resultados dos ensaios realizados no terceiro cenário, referente às buscas baseadas em análise de contexto, são apresentados na Tabela 9. O **mapa de conceitos** foi novamente avaliado nas Tarefas 27, 28 e 30, com resultados similares à avaliação realizada no segundo cenário, através de uma média de sucesso de 7,0 e insucesso de 3,0. Neste ponto,

pode-se perceber que os participantes compreendem bem as possibilidades de pesquisa apresentadas pelos planos de informação e de contexto (Tarefas 27 e 30), mas ainda ficam confusos para inserir mais de um termo à busca realizada, tal como também já observado durante o desenvolvimento da Tarefa 6, no âmbito do primeiro cenário.

A interação com os **resultados da busca** também foi mais uma vez avaliada nas Tarefas 31, 32 e 33, apresentando uma média de sucesso (realizadas com facilidade ou com dificuldade de 8,0 e de insucesso (não concluídas) de apenas 2,0, melhores, portanto, do que os resultados observados no primeiro cenário referentes ao mesmo contexto, indicando um possível aprendizado sobre os recursos da interface para a apresentação da informação. As principais dificuldades observadas ainda estão relacionadas com a relevância, que é representada pela proximidade com o mapa de conceitos (Tarefa 32) e com a compreensão do conceito de similaridade entre os termos de busca e documentos (Tarefas 31 e 33).

O **histórico do autor** também foi avaliado neste cenário, por meio da Tarefa 29 e não apresentou problemas para ser interpretado. A única dificuldade observada foi em relação ao ícone de acesso à funcionalidade, uma vez que os usuários não estão familiarizados com todos os recursos da interface.

Tabela 10 - Resultados dos ensaios de usabilidade realizados no quarto cenário

Tarefa	**RF**	**Recurso Avaliado**	**Perspectiva**	**Tarefa Realizada**		
				Sim, com facilidade	Sim, com dificuldade	Não
34	04	Busca (perspectivas de conteúdo)	FAF	5	2	3
35	04	Busca (perspectivas de conteúdo)	II	4	1	5
36	28	Comentários interpretativos	II	5	2	3
37	32	Erros e falhas	II	6	1	3
38	28	Avaliação de parágrafos	II	6	2	2
39	31	Resultados relacionados	II	6	2	2
40	30	Resultados contraditórios	II	7	2	1
41	05	Mapa de conceitos	FAF	5	2	3
42	04	Mapa de conceitos	II	4	2	4
43	10	Busca (perspectivas de conteúdo)	II	9	1	0
44	17	Resultados da busca	II	7	2	1
45	12	Resultados da busca	II	8	1	1
46	17	Resultados da busca	II	6	2	2
Média de sucesso das tarefas				**6,0**	**1,7**	**2,3**

Fonte: desenvolvido pela autora

A Tabela 10 apresenta os resultados dos testes de usabilidade realizados no quarto cenário, que discute o contexto das **buscas baseadas em perspectivas de conteúdo** de um determinado documento. O recurso foi avaliado durante o desenvolvimento das Tarefas

34, 35 e 43, apresentando uma média de sucesso de 7,3 e de insucesso de 2,7. Por ser um recurso novo na interface, as principais dificuldades apresentadas referem-se ao acesso do material que será explorado (Tarefa 34), bem como à compreensão dos participantes sobre as possibilidades de busca a partir das perspectivas de um documento (Tarefa 35).

Em relação aos **resultados da busca**, apesar de ser um cenário diferenciado dos outros três, as atividades dos usuários mantiveram resultados parecidos com as análises realizadas no terceiro cenário, apresentando uma média de sucesso de 8,7 e de apenas 1,3 de insucesso. Nesta etapa dos ensaios, percebeu-se um maior entendimento sobre os critérios que definem a busca, resultados mais e menos relevantes e similaridade existente entre termos e documentos recuperados (Tarefas 44, 45 e 46).

O **mapa de conceitos** foi avaliado novamente sob a perspectiva dos documentos selecionados e, com uma média de sucesso de 7,6 e de insucesso de 2,4, apresentou poucas dificuldades para ser interpretado (Tarefas 41 e 42). Os principais problemas observados referem-se à compreensão do usuário referente à mudança do contexto de busca, ao entendimento das possibilidades de pesquisa geradas a partir de um determinado material (Tarefa 41), bem como à viabilidade de utilização de diversas perspectivas e termos em uma única pesquisa, permitindo inúmeras alternativas de investigação (Tarefa 42).

As funcionalidades desenvolvidas com o objetivo de apresentar informações adicionais para a interpretação de um documento, incluindo **comentários interpretativos**, **erros e falhas**, **avaliação de parágrafos**, **resultados relacionados** e **contraditórios** (Tarefas 36, 37, 38, 39 e 40, respectivamente) também foram avaliadas neste cenário, apresentando uma média de sucesso de 7,8 e de insucesso de 2,2. Neste contexto, não foram observados problemas significativos durante os testes realizados. Os usuários que não conseguiram realizar a tarefa tiveram dificuldades para acessar a funcionalidade e compreender o seu significado no contexto de uma busca.

Tabela 11 - Resultados dos ensaios de usabilidade realizados no quinto cenário

Tarefa	**RF**	**Recurso Avaliado**	**Perspectiva**	**Tarefa Realizada**		
				Sim, com facilidade	Sim, com dificuldade	Não
47	13	Inclusão e análise de documentos	FAF	5	2	3
48	15	Inclusão de relacionamentos	FAF	6	2	2
49	14	Inclusão de termos na ontologia	FAF	3	2	5
50	28	Interpretação de documentos	FAF	5	2	3
Média de sucesso das tarefas				**4,7**	**2,0**	**3,3**

Fonte: desenvolvido pela autora

Os resultados dos ensaios realizados no quinto cenário são apresentados na Tabela 11 e referem-se à inclusão e análise de materiais na base de dados para, posteriormente, serem acessados pelos usuários da interface. A média de sucesso nas atividades desenvolvidas foi de 6,7 contrapondo-se à 3,3 de insucesso. Diferente dos demais cenários avaliados, as atividades analisadas neste contexto não estão relacionadas diretamente com a busca, mas sim com o processo de inserção e análise documental (Tarefas 47, 48, 49 e 50).

Entre as dificuldades observadas, a **inclusão de termos na ontologia** foi o recurso mais difícil de ser interpretado pelos usuários, não ficando evidente a sequência de interações para a realização da tarefa. Além disso, os participantes, em um primeiro momento, não conseguiram relacionar corretamente o termo que estava sendo criado com as informações referentes às suas respectivas áreas de conhecimento, antônimos ou contraditórios, sinônimos ou relacionados e variação linguística, o que limitou o desenvolvimento desta atividade e contribuiu para que 5 usuários não conseguissem concluir a tarefa (Tarefa 49).

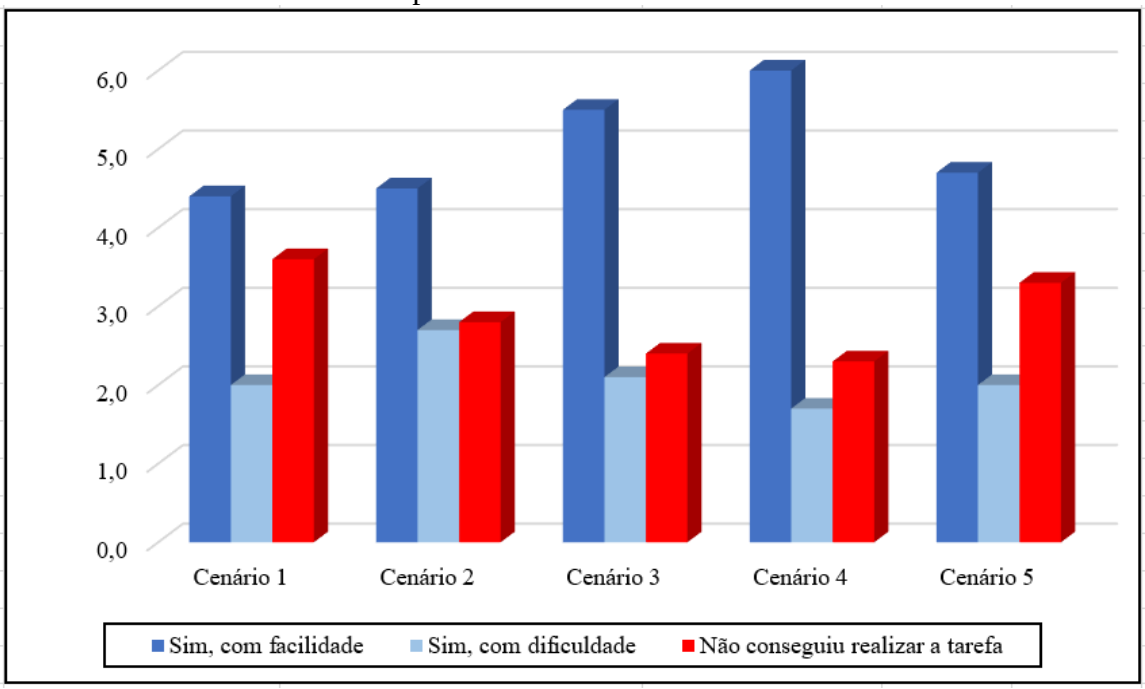
No contexto da **interpretação de documentos**, as principais dificuldades dos usuários referem-se ao relacionamento com os resultados relacionados e contraditórios, bem como com a avaliação dos parágrafos utilizando o recurso *like* e *dislike* (Tarefa 50). Os participantes também ficaram confusos sobre o ícone de **acesso à funcionalidade**, provavelmente por não estarem familiarizados com a interface, mas conseguiram compreender os passos para realizar o *upload* de um documento (Tarefa 47). Em relação à **inclusão de relacionamentos** entre um termo existente e um documento, os usuários não conseguiram deduzir a possibilidade de acesso às funções de relacionamento através de um menu suspenso. No entanto, após a aprendizagem, apenas 2 interações não foram concluídas com sucesso (Tarefa 48).

O Gráfico 1 apresenta um comparativo de sucesso das tarefas realizadas nos ensaios de usabilidade. Apesar das inúmeras dificuldades apresentadas pelos usuários na interação com a interface de busca proposta neste estudo, pode-se observar um crescimento progressivo na quantidade de interações realizadas com sucesso, do primeiro cenário ao quarto, que se referem a contextos de busca diferenciados. As médias de sucesso (tarefas realizadas facilmente ou com dificuldade) verificadas foram: 6,4 (primeiro cenário), 7,2 (segundo cenário), 7,6 (terceiro cenário) e 7,7 (quarto cenário).

Por outro lado, as médias de usuários que não conseguiram realizar tarefas também tiveram um decréscimo gradual: 3,6 (primeiro cenário), 2,8 (segundo cenário), 2,4

(terceiro cenário) e 2,3 (quarto cenário). Estes dados indicam que as dificuldades dos participantes foram minimizadas simultaneamente com o aprendizado adquirido em relação aos recursos, bem como à medida em que o usuário foi se familiarizando com a interface.

Gráfico 1 - Comparativo de sucesso das tarefas realizadas



Fonte: desenvolvido pela autora

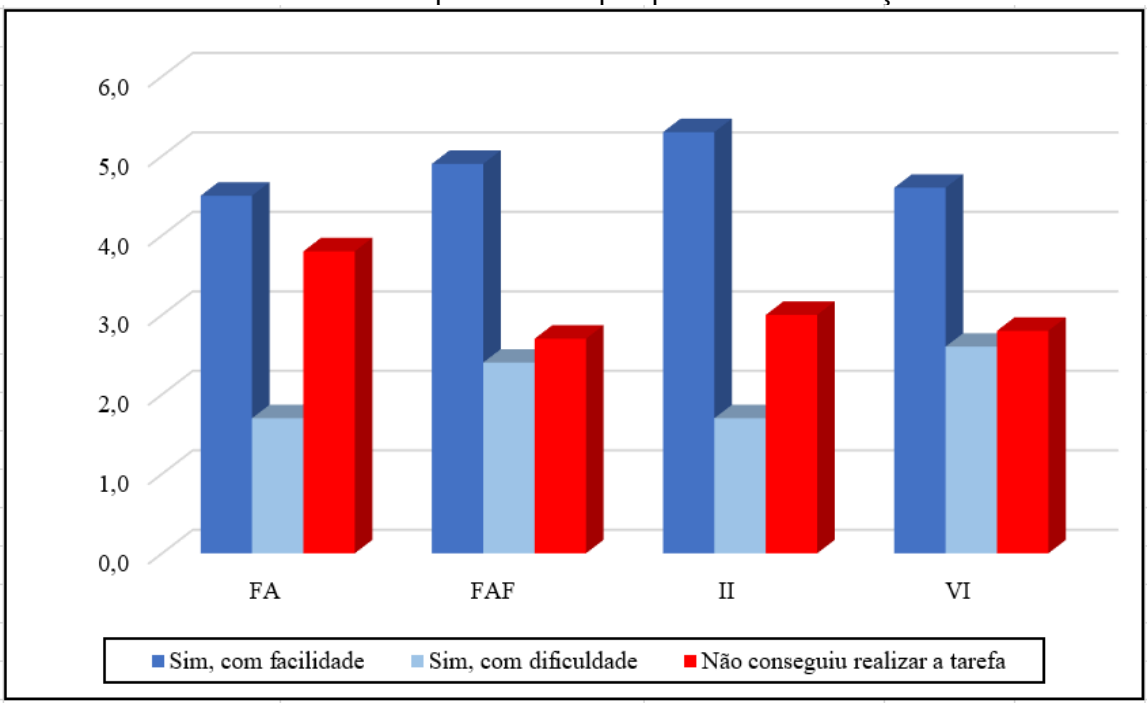
O quinto cenário, apesar de representar um contexto diferenciado de inserção e análise de documentos, também não apresentou dificuldades que pudessem comprometer as atividades desenvolvidas na interface de busca.

O Gráfico 2, por sua vez, mostra um comparativo de sucesso das tarefas realizadas de acordo com as quatro perspectivas de avaliação consideradas no estudo: facilidade de aprendizagem (FA), facilidade de acesso à informação (FAF), interpretação da informação (II) e visibilidade da informação (VI). Embora todos os resultados observados tenham sido acima de 6,0, será necessário aperfeiçoar algumas funcionalidades com o objetivo de minimizar as dificuldades observadas na interação dos usuários.

A melhor avaliação, com uma média de sucesso de 7,3 foi observada no contexto da perspectiva FAF, indicando que os recursos da interface, bem como os elementos de *layout* e a sequência de ações possuem uma quantidade menor de problemas e foram melhor interpretados pelos participantes. Em segundo, a perspectiva visibilidade de informação foi avaliada com uma média de sucesso de 7,2, o que sugere que as funcionalidades não estão suficientemente visíveis e fáceis de serem encontradas pelos

diferentes tipos de usuário, principalmente por aqueles com menos disponibilidade para explorar os recursos. A perspectiva II, por sua vez, obteve uma média de sucesso de 7,0 nas tarefas desenvolvidas, indicando que, apesar dos recursos projetadas com este propósito, ainda existem obstáculos na interpretação da informação, seja no contexto da elaboração das expressões de busca ou na análise dos resultados. E, por último, as atividades sob a perspectiva FA tiveram uma média de sucesso de 6,2, apontando que os recursos não foram assimilados tão facilmente como o esperado e que existe, portanto, necessidade de revisão dos elementos avaliados sob este ponto de vista.

Gráfico 2 - Comparativo das perspectivas de avaliação



Fonte: desenvolvido pela autora

Além das perspectivas de avaliação, cada tarefa do teste de usabilidade foi associada a um requisito funcional, que foi elaborado nas primeiras etapas de desenvolvimento do projeto de interface. Neste contexto, foram testados recursos projetados para auxiliar a interação do usuário com o sistema (requisitos de 01 a 12), para permitir a análise da similaridade existente entre os documentos (requisitos de 13 a 20) e para auxiliar o usuário na interpretação de um documento ou de um *corpus* documental (requisitos 21 a 32). Não foram testados os requisitos 07, 08 e 19, que se referem, respectivamente, aos recursos de *zoom*, aos filtros de busca e análise de grupo de documentos por serem recursos que ficariam limitados nas interações realizadas através de protótipos de baixa fidelidade. As tarefas projetadas para avaliar cada requisito são apresentadas no Quadro 18.

Quadro 18 - Requisitos funcionais analisados nos ensaios de usabilidade

RF	Descrição	Tarefa(s)
01	Permitir a elaboração de buscas realizadas a partir de palavras-chave	01, 02, 06, 08, 19
02	Permitir a elaboração de buscas realizadas a partir de uma estrutura terminológica	20, 22
03	Permitir a elaboração de buscas realizadas a partir de uma ou várias perspectivas	26, 27
04	Permitir a elaboração de buscas realizadas a partir das perspectivas do documento	34, 35, 42
05	Permitir a seleção de perspectivas de análise (uma, múltiplas ou todas)	04, 41
06	Apresentar estrutura de navegação hibrida, com recursos de *searching* e *browsing*	21, 24
09	Evidenciar as perspectivas de análise candidatas à seleção do usuário	05, 07
10	Evidenciar os conceitos e documentos selecionados pelo usuário	43
11	Evidenciar o contexto dos termos selecionados pelo usuário	13, 45
12	Evidenciar os resultados mais relevantes	09, 32
13	Permitir a inclusão, alteração e exclusão de documentos na base de dados	47
14	Permitir a inclusão, alteração e exclusão de relacionamentos da ontologia	49
15	Permitir a inclusão, alteração e exclusão de relacionamentos entre os documentos e a ontologia	48
16	Permitir a exploração dos documentos através de múltiplas perspectivas de análise	28
17	Permitir a análise da similaridade existente entre expressão e resultados da busca	11, 46, 31, 44
18	Permitir a análise da similaridade existente entre resultados de busca	12, 33
20	Apresentar a referência completa dos resultados selecionados	10
21	Apresentar, através da estrutura terminológica, conceitos e contextos associados	03, 23, 25
22	Permitir a visualização total e parcial (todo-parte) dos conceitos da terminologia	30
23	Apresentar notas explicativas relacionadas a termos de busca	14
24	Apresentar sinônimos relacionados a termos de busca	16
25	Apresentar variação linguística relacionada a termos de busca	17
26	Apresentar as diversas áreas de conhecimento relacionados a termos de busca	15
27	Apresentar termos antônimos e contraditórios relacionados a termos de busca	18
28	Permitir a interação entre pesquisadores e usuários no contexto da interpretação	36, 38, 50
29	Apresentar o histórico do autor relacionado com um determinado material	29
30	Apresentar sugestões de resultados contendo ideias e/ou teses contraditórias	40
31	Apresentar sugestões de resultados relacionados	29
32	Permitir a indicação de erros e falhas nos documentos pertencentes ao *corpus*	37

Fonte: desenvolvido pela autora

Ao final de cada teste realizado, os participantes responderam, sem a presença do avaliador, um questionário de avaliação de usabilidade, que teve como objetivo entender, de uma forma mais abrangente e detalhada, a percepção que tiveram em relação à interface proposta e aos recursos desenvolvidos.

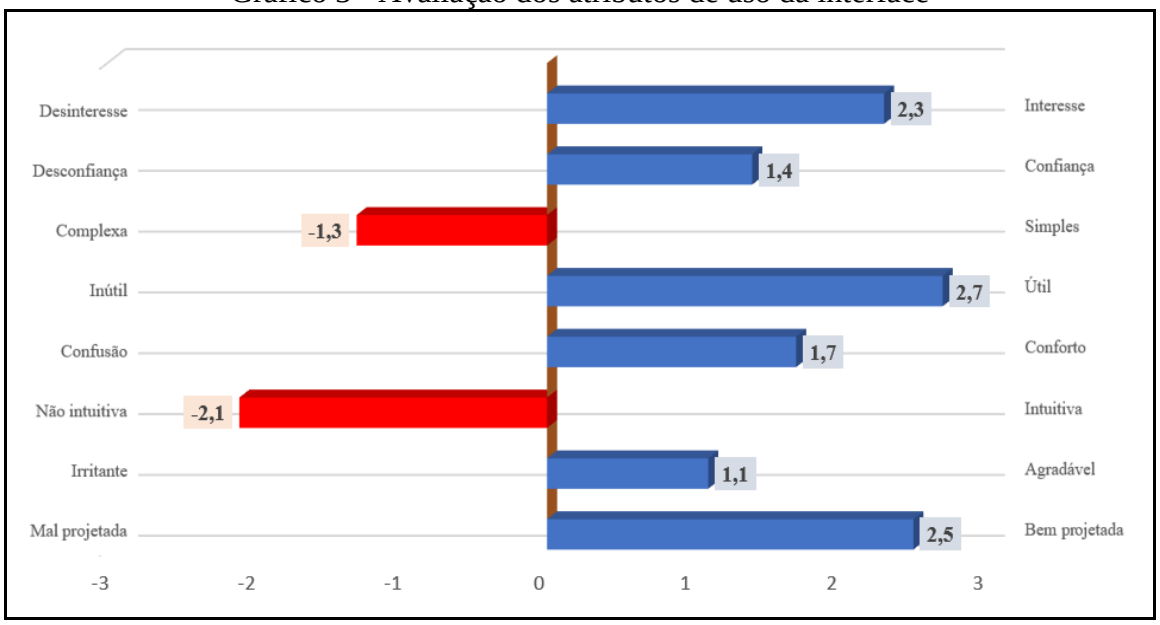
7.4.3 Avaliação de Usabilidade

O questionário final utilizado para a avaliação de usabilidade foi disponibilizado no Apêndice B. Os resultados das questões propostas serão apresentados por meio de gráficos, com a intenção de facilitar a elaboração de comparativos entre as variáveis

analisadas. As questões de 1 a 7 referem-se ao perfil dos participantes e foram discutidas no início desta seção.

A oitava questão teve como objetivo identificar a percepção dos usuários em relação ao uso da interface. As questões com avaliações mais próximas de 3 foram consideradas positivas, enquanto que as mais distantes, foram consideradas negativas.

Gráfico 3 - Avaliação dos atributos de uso da interface



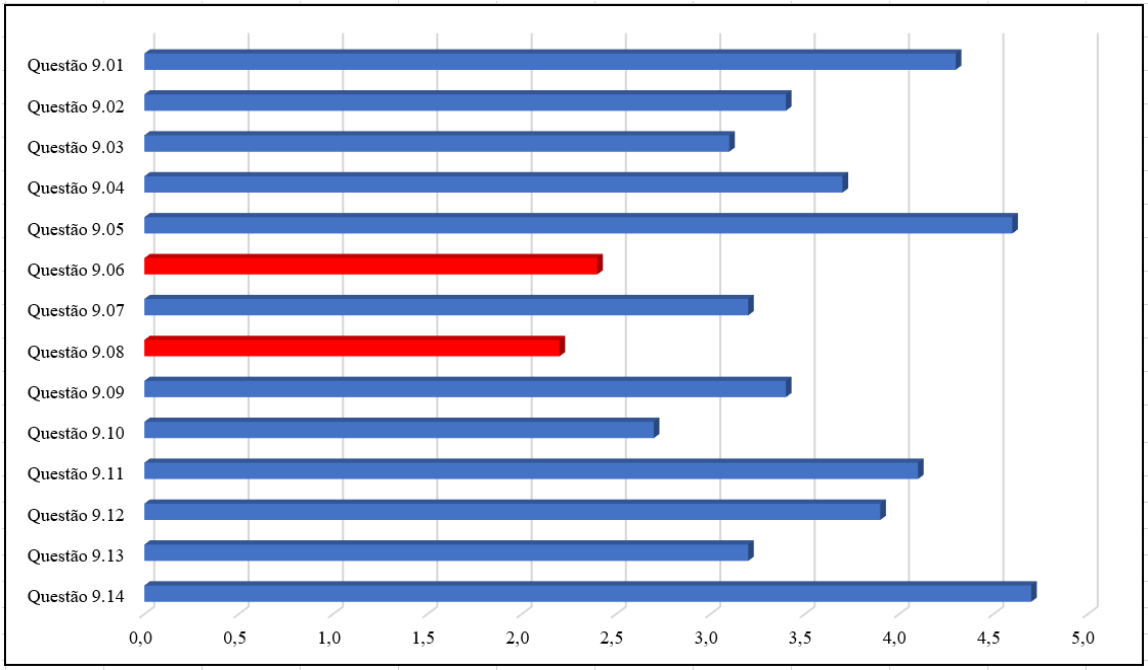
Fonte: desenvolvido pela autora

De acordo com os dados apresentados no Gráfico 3, os participantes consideram a interface de busca como sendo uma ferramenta útil, bem projetada e interessante. Em contrapartida, a qualificam como sendo não intuitiva e complexa. Estes resultados podem estar relacionados com a dificuldade de aprendizagem de alguns recursos, conforme foi apontado nos testes de usabilidade.

O objetivo da questão 9 foi avaliar o grau de utilidade dos principais recursos da interface em atividades de busca e análise de um *corpus* documental. Os recursos avaliados foram: 9.01) busca baseada em palavras-chave; 9.02) busca baseada em estrutura terminológica; 9.03) busca baseada em uma ou várias perspectivas de contexto; 9.04) busca baseada em uma ou várias perspectivas de um documento; 9.05) mapa de conceitos; 9.06) informações complementares; 9.07) plano de contexto; 9.08) barra de navegação estrutural; 9.09) sinônimos e termos relacionados; 9.10) antônimos e termos contraditórios; 9.11) comentários interpretativos dos usuários; 9.12) resultados relacionados ao mesmo contexto; 9.13) indicação de erros, falhas e avaliação de parágrafos e 9.14) representação visual dos resultados da busca.

Para avaliar a questão, os usuários utilizaram uma escala de 1 a 5, onde o 1 foi considerado o menos útil e o 5, o mais útil. Neste contexto, os resultados abaixo de 2,5 foram considerados negativos e os acima, positivos. Os resultados com médias acima de 4,0 foram interpretados como sendo os recursos de maior importância na perspectiva do utilizador.

Gráfico 4 - Grau de utilidade dos recursos da interface



Fonte: desenvolvido pela autora

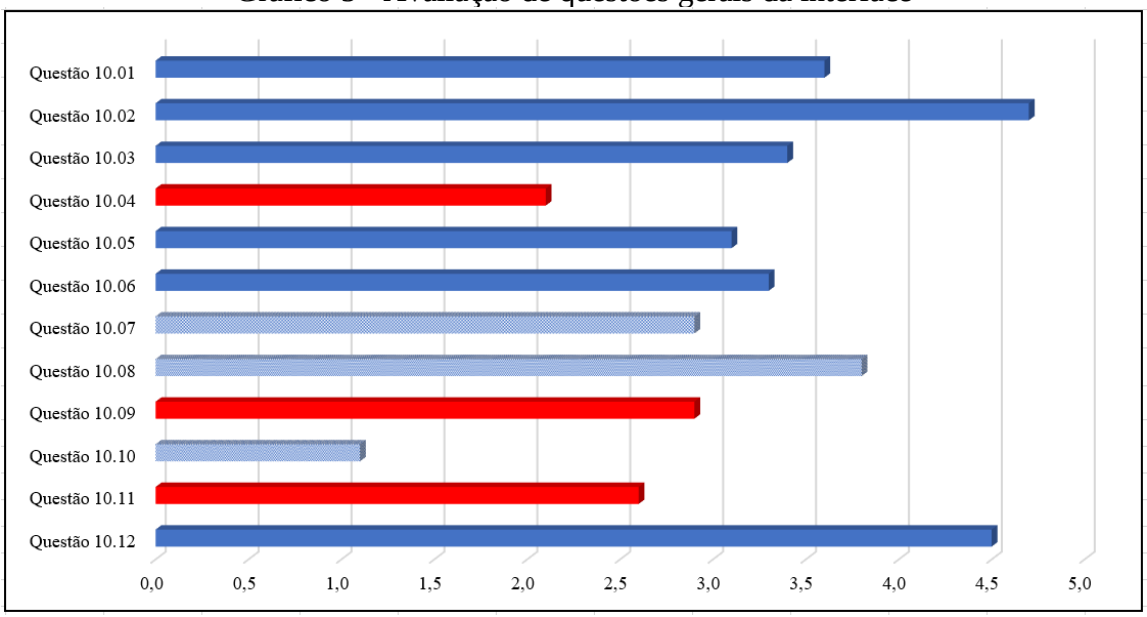
O Gráfico 4 mostra que maioria das funcionalidades da interface tiveram avaliações positivas, com médias acima de 2,5, indicando que possuem relevância na perspectiva do usuário. Os quatro recursos considerados mais importantes para auxiliar na realização de uma busca e na análise de documentos foram: representação visual dos resultados da busca (4,7), mapa de conceitos (4,6), busca baseada em palavras-chave (4,3) e comentários interpretativos dos usuários (4,1). Por outro lado, dois recursos foram considerados pelos usuários como “não tão úteis”, sendo eles: informações complementares (2,4) e barra de navegação estrutural (2,2).

A questão 10 teve como objetivo avaliar questões gerais da interface, tais como: facilidade de interação, recursos de navegação, organização da informação, facilidade da análise de dados, terminologia dos botões, rótulos e títulos, confiança no uso da ferramenta e utilidade para as atividades de pesquisa. Neste cenário, foram apresentadas aos participantes afirmações sobre os elementos em análise, solicitando a concordância ou não com cada enunciado. A escala utilizada foi: 1) discordo totalmente, 2) discordo, 3) sem opinião, 4)

concordo e 5) concordo totalmente. Assim, foram consideradas avaliações positivas as que tiveram média superior a 3,0.

Para avaliar a atenção do usuário com o instrumento de pesquisa e com as respostas que estavam sendo fornecidas, bem como obter mais confiabilidade em relação à uma determinada afirmação, foram adicionadas três questões de controle, que são apresentadas na cor azul clara (Questões 10.07, 10.08 e 10.10). A questão 10.07 contradiz a afirmação da 10.03, a questão 10.08 é contrária à 10.04 e a questão 10.10 se contrapõe à 10.01.

Gráfico 5 - Avaliação de questões gerais da interface



Fonte: desenvolvido pela autora

A análise dos resultados apresentada no Gráfico 5 destaca as duas questões melhor avaliadas pelos participantes: 10.02 com média de 4,7 (eu usaria este sistema com frequência) e 10.12 com média de 4,5 (esta ferramenta seria útil nas minhas atividades de pesquisa).

No caso das questões de controle, o conceito é inverso, ou seja, quanto menores forem as médias, melhor é a avaliação recebida. Neste sentido, a questão 10.07 com média de 2,9 (a análise de dados é confusa) se contrapõe à questão 10.03 com média de 3,4 (foi fácil encontrar a informação que eu precisava). Em função da proximidade entre as duas avaliações, conclui-se que o usuário não tem clareza a respeito deste critério, julgando alguns contextos de informação como sendo fáceis de serem interpretados e outros, no entanto, como confusos. Estes resultados estão de acordo com a avaliação dos atributos de uso da interface na qual o quesito conforto foi avaliado com uma média de apenas 1,7 (Gráfico 3).

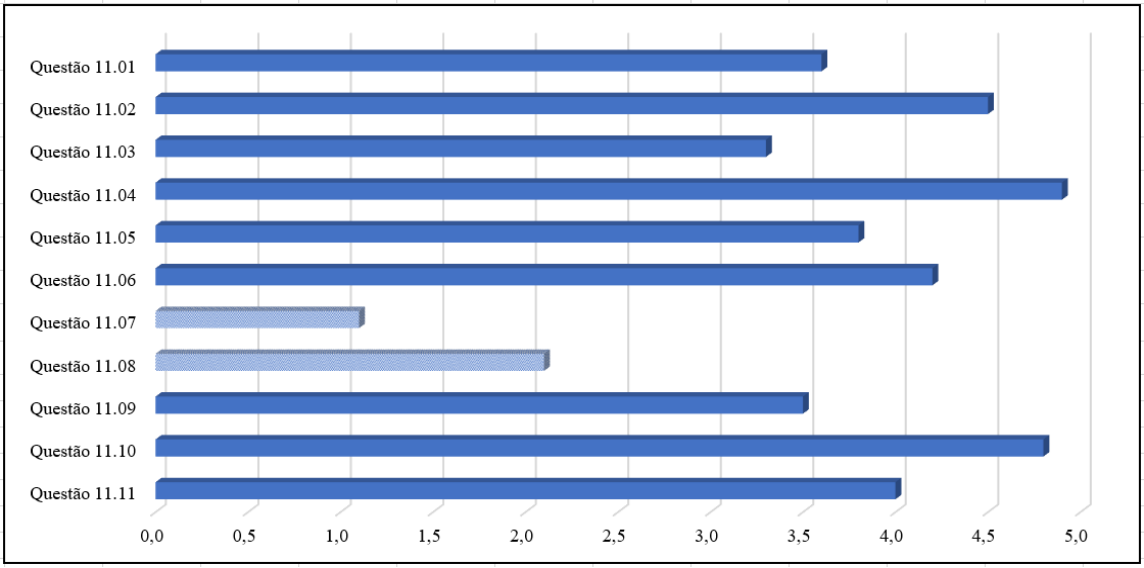
A questão 10.08 com média 3,8 (eu achei a ferramenta complexa, precisei lembrar, pesquisar ou pensar muito para completar as tarefas) contradiz a questão de 10.04 com média 2,1 (as funcionalidades disponibilizadas são de fácil aprendizagem). Neste contexto, como as duas questões apresentaram resultados distintos, mesmo sendo opostas, pode-se concluir que o usuário considera a ferramenta complexa e com recursos de difícil aprendizagem.

Em relação à questão 10.10 com média 1,1 (eu precisaria de apoio de uma pessoa para usar esta interface) se contrapõe à questão 10.1 com média 3,6 (esta ferramenta tem interface de fácil interação). Estes resultados indicam que, apesar das dificuldades observadas durante os testes de usabilidade, os usuários teriam condições de utilizar a interface de busca, sem a necessidade de buscar auxílio.

A análise dos dados mostra ainda três questões com avaliações negativas: 10.09 com média de 2,9 (a terminologia utilizada nos botões, rótulos, títulos etc. é de fácil entendimento), 10.11 com média de 2,6 (eu me senti muito confiante usando esta ferramenta) e 10.04 com média de 2,1 (as funcionalidades disponibilizadas são de fácil aprendizagem).

A questão 11 teve como objetivo avaliar os recursos de interface projetados a partir do uso das ontologias, tais como: mapa de conceitos, plano de contexto, sinônimos, antônimos, variação linguística, círculos de proximidade e relevância de resultados de busca. Os critérios de avaliação foram os mesmos adotados na questão 10.

Gráfico 6 - Avaliação da ontologia como componente de interface



Fonte: desenvolvido pela autora

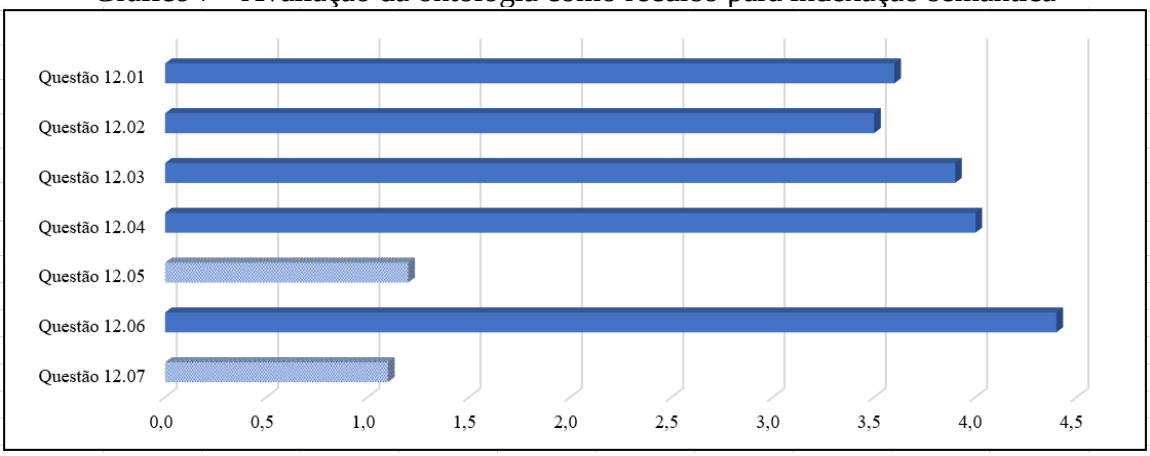
As funcionalidades de interface desenvolvidas a partir do uso de ontologias foram avaliadas com conceitos positivos, conforme os resultados apresentados no Gráfico 6. Neste contexto, quatro questões obtiveram médias de destaque: 11.04 com média de 4,9 (o mapa de conceitos me ajudou a adquirir conhecimento sobre o domínio de pesquisa - termos sinônimos, antônimos, variações linguísticas), 11.10 com média de 4,8 (o mapa de conceitos me ajudou a elaborar uma busca em um domínio totalmente desconhecido), 11.02 com média de 4,5 (o mapa de conceitos foi útil para a elaboração das pesquisas realizadas) e 11.06 com uma média de 4,2 (os círculos de proximidade me ajudaram a perceber quais documentos eram mais relevantes).

Para buscar maiores evidências em relação à percepção dos participantes sobre o mapa de conceitos e os resultados da busca, foram elaboradoras duas questões de controle: 11.07 e 11.08. A questão 11.07 com média de 1,1 (o mapa de conceitos é confuso e dificultou minhas buscas) se opõe à 11.01 com média de 3,6 (o mapa de conceitos me ajudou a entender quais informações estavam disponíveis para serem pesquisadas). Estes resultados indicam que o mapa de conceitos é um recurso útil e com potencial para colaborar com o desenvolvimento das atividades do pesquisador em um ambiente de busca.

A questão 11.08 com média de 2,1 (eu tive dificuldades para compreender quais documentos recuperados eram mais relevantes) é contrária à questão 11.06 com média 4,2 (os círculos de proximidade me ajudaram a perceber quais documentos eram mais relevantes). A análise destes dados, em conjunto com os testes de usabilidade realizados no primeiro cenário, aponta que os participantes tiveram uma dificuldade inicial para compreender os critérios de representação de relevância, incluindo os círculos de proximidade e as cores utilizadas. No entanto, após a assimilação destes conceitos, a interpretação dos resultados da busca ficou menos complicada e mais transparente.

O objetivo da questão 12 foi avaliar o uso de ontologias como recurso para a indexação semântica do *corpus* documental que, entre outras possibilidades, permite a análise da estrutura dos documentos, a recuperação de documentos relacionados semanticamente, a expansão dos resultados da busca e a identificação de contextos de similaridade.

Gráfico 7 - Avaliação da ontologia como recurso para indexação semântica



Fonte: desenvolvido pela autora

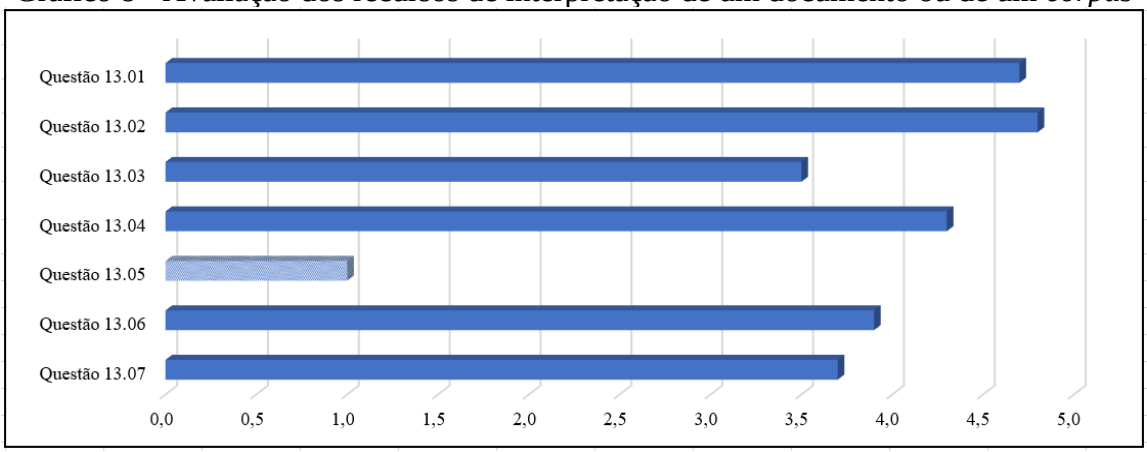
Os resultados apresentados no Gráfico 7 não apontam avaliações negativas no contexto avaliado. A questão de maior destaque foi a 12.06 com média de 4,4 (os resultados correlacionados foram importantes para o contexto da minha pesquisa) e indica que os participantes consideram importante a recuperação de resultados pertencentes ao mesmo contexto, sejam eles associados por sinônimos, antônimos ou variações linguísticas, bem como de pesquisas complementares, com resultados relacionados ou contraditórios.

As questões de controle 12.05 e 12.07 foram elaboradas com a intenção de obter mais informações sobre as perspectivas de análise e a apresentação dos documentos recuperados. A questão 12.05 com média 1,2 (os recursos que permitem a exploração das perspectivas de análise (autor, objetivos, metodologia) não foram úteis para a minha pesquisa) é contrária à 12.01 com média 3,6 (durante as minhas pesquisas foi possível explorar, de forma mais aprofundada, a estrutura dos documentos recuperados). Desta forma, na percepção do usuário, a possibilidade de explorar um documento a partir de diversas perspectivas (autor, palavras-chave, objetivo, metodologia) pode ser um recurso importante em um cenário de busca, recuperação e análise de documentos.

A questão 12.07 com média de 1,1 (os documentos recuperados são apresentados de forma desorganizada e sem critérios) é oposta à 12.4 com média 4,0 (a organização dos documentos recuperados me ajudou a compreender a relação de similaridade entre eles). Isto significa que o usuário avalia de forma positiva a apresentação dos documentos recuperados no contexto de uma busca e que os recursos projetados (cor e proximidade) podem ser úteis para a análise das informações apresentadas.

A questão 13 teve como objetivo avaliar os recursos de interface projetados a partir dos princípios da hermenêutica, tais como: comentários interpretativos, interação entre pesquisadores e sujeitos, sinônimos e sugestões de pesquisas relacionadas.

Gráfico 8 - Avaliação dos recursos de interpretação de um documento ou de um *corpus*



Fonte: desenvolvido pela autora

No contexto da avaliação dos recursos projetados para auxiliar a interpretação de documentos não foram identificadas questões negativas ou que pudessem sugerir problemas relevantes. As três questões de destaque foram: 13.02 com média de 4,8 (quando eu tenho acesso a opiniões diferentes acerca de um mesmo documento ou assunto, é possível conhecer diversos pontos de vista, que estão associados a diferentes crenças, orientações políticas, valores culturais, preconceitos e outros), 13.01 com média de 4,7 (a interpretação de um documento realizada por diversos outros usuários é útil para me ajudar a compreender o assunto pesquisa) e 13.04 com média de 4,3 (foi possível conhecer palavras sinônimas relacionadas com a minha busca).

Com a intenção de avaliar a importância da interação entre os usuários da interface, a questão 10.05 com média de 1,0 (conhecer diferentes pontos de vista sobre um documento ou assunto não contribui para o meu aprendizado) contradiz as questões 10.01 e 10.02 discutidas anteriormente. Estes resultados, avaliados em conjunto, confirmam que os participantes consideram relevantes os recursos projetados para permitir a interação entre os pesquisadores e sujeitos no contexto da interpretação dos documentos e que esta interação pode facilitar a compreensão de um determinado cenário, seja no âmbito conceitual, teórico, interpretativo ou prático.

7.5 RESUMO E DISCUSSÃO DO CAPÍTULO

O principal objetivo deste capítulo foi apresentar as etapas do processo de desenvolvimento do “Modelo de interface de busca e recuperação de informação”, que foi elaborado com base nos princípios da hermenêutica e na utilização de ontologias. Entre as atividades previstas no ciclo de vida do projeto estão: análise do domínio e oportunidades de

melhoria, definição dos requisitos funcionais e não funcionais, estudo do público alvo e das tarefas previstas, *design* de alternativas, desenvolvimento dos protótipos para cada cenário proposto e avaliação de usabilidade.

As atividades desenvolvidas neste capítulo foram fundamentadas no *Design* de Interação, que tem como foco minimizar os aspectos negativos da experiência do usuário (frustração, aborrecimento, tédio) e, ao mesmo tempo, melhorar os positivos (motivação, satisfação, interesse), com a intenção de aumentar a eficácia dos resultados. Além disso, o projeto contou ainda com a contribuição de quatro outras áreas: Recuperação de Informação, Ontologias, Visualização de Informação e Hermenêutica.

A área de Recuperação de Informação forneceu um panorama sobre os fatores que interferem no processo de busca, os principais problemas envolvidos e as principais dificuldades do usuário na formulação de expressões de busca, bem como na interpretação dos resultados apresentados. O estudo das ontologias contribuiu para o projeto de funcionalidades relacionadas à melhoria da interação do usuário com o sistema e à indexação semântica do *corpus* documental. A visualização da informação forneceu elementos e recursos gráficos para auxiliar na análise na extração de informações relevantes, tanto no contexto da elaboração das buscas, quanto na análise dos resultados. E, por fim, a hermenêutica deu subsídios para o desenvolvimento das funcionalidades relacionadas à intepretação de documentos, que pode ser realizada tanto no processo de indexação quanto na análise dos resultados retornados de uma busca.

As funcionalidades desenvolvidas a partir do uso de ontologias foram projetadas em dois contextos diferentes: para facilitar a interação do usuário com o sistema e para permitir a análise da similaridade existente entre os documentos. Os recursos desenvolvidos são apresentados no Quadro 19.

Em relação à interação do usuário com o sistema, as ontologias foram utilizadas como componente de interface de busca e permitiram o delineamento de quatro funcionalidades relacionadas às buscas: 1) palavras-chave; 2) estrutura terminológica; 3) perspectivas de análise de contexto e 4) perspectivas de análise de conteúdo. Além disso, foi possível disponibilizar ao usuário uma estrutura de navegação híbrida, com recursos de *browsing* e *searching* para facilitar a elaboração de buscas nos vários contextos previstos.

Quadro 19 - Recursos do modelo de interface de busca projetados a partir das ontologias

Contexto	Recurso(s)
Interação do Usuário com o Sistema (Ontologia como Componente de Interface)	1. Busca Baseada em Palavra-Chave 2. Busca Baseada em Estrutura Terminológica 3. Busca Baseada em Perspectivas de Análise de Contexto 4. Busca Baseada em Perspectivas de Análise de Conteúdo 5. Estrutura de Navegação Híbrida, com Recursos de *Browsing* e *Searching*
Similaridade entre Documentos (Ontologia como Estrutura de Indexação Semântica do *Corpus*)	1. Inclusão, Alteração e Exclusão de Documentos na Base de Dados 2. Inclusão, Alteração e Exclusão de Conceitos e Relacionamentos na Ontologia 3. Inclusão, Alteração e Exclusão de Conceitos e Relacionamentos entre Documentos e Ontologia 4. Análise da Similaridade entre Termos e Resultados de Busca 5. Análise da Similaridade entre Resultados de Busca

Fonte: desenvolvido pela autora

Além da interação do usuário com o sistema, as ontologias foram adotadas na estrutura de indexação do *corpus* documental, a partir do qual os seguintes recursos foram projetados: inclusão, alteração e exclusão de: 1) documentos na base de dados; 2) conceitos e relacionamentos na ontologia e 3) conceitos e relacionamentos entre documentos e ontologia. Com base nestes elementos, a interface de busca tem condições de oferecer ao usuário a análise de similaridade existente entre termos e resultados de busca.

Para o projeto das funcionalidades subsidiadas pela hermenêutica, com o objetivo de facilitar o processo de interpretação e análise documental, foram adotados os princípios propostos por Klein e Myers (1999), sendo eles: Círculo Hermenêutico, Contextualização, Interação entre Pesquisadores e Sujeitos, Abstração e Generalização, Raciocínio Dialético, Múltiplas Interpretações e Suspeição. Além destes, para o contexto desta pesquisa, foi proposto ainda o princípio do Compartilhamento da Linguagem. O Quadro 20 apresenta resumidamente os recursos da interface de busca elaborados a partir de cada princípio hermenêutico.

Os princípios do círculo hermenêutico e do compartilhamento de linguagem são considerados a base de todo o processo interpretativo e é a partir dos quais os outros seis se relacionam. Com base nestes dois princípios, foram projetados recursos que facilitam a apropriação da linguagem de um determinado domínio, bem como contribuem para o aprendizado do usuário durante as interações realizadas: mapa de conceitos, barra de navegação estrutural, notas explicativas, sinônimos e termos relacionados e variações linguísticas.

Quadro 20 - Princípios hermenêuticos aplicados ao modelo de interface de busca

Princípio	Definição	Recurso(s)
Círculo Hermenêutico	Não é possível compreender o todo até que as partes tenham sido compreendidas.	1. Mapa de Conceitos (Plano de Informação e de Contexto) 2. Barra de Navegação Estrutural
Contextualização	Baseado na ideia da diferença de entendimento entre intérprete e o autor de um texto (resultante da distância histórica (ou cultural) entre eles. Desta forma, tem a preocupação de entender o posicionamento do todo no seu contexto histórico.	Nível Básico 1. Histórico do Autor: reunir informações referentes a um autor Nível Intermediário 2. Histórico do Autor: dar subsídios à realização de buscas baseadas em aspectos regionais e culturais
Interação entre Pesquisadores e Sujeitos	O significado de um texto não pode ser visto unicamente a partir das perspectivas e do entendimento de um único intérprete, mas, ao contrário, deve ser sempre negociado e consensuado	1. Comentários Interpretativos 2. Relacionamentos entre Elementos de Informação
Abstração e Generalização	Refere-se à possibilidade de produção de resultados de caráter abstrato e genérico	1. Área de Conhecimento
Raciocínio Dialético	Requer que todo pesquisador confronte as preconcepções que o conduziram na sua proposta de pesquisa original com a experiência emergente no decurso do processo de investigação	Nível Básico 1. Antônimos e Termos Conceitualmente Contraditórios Baseados em Dicionário de Antônimos 2. Resultados Contraditórios em Nível de Antônimos Nível Intermediário 3. Antônimos e Termos Conceitualmente Contraditórios Baseados em Semântica 4. Resultados Contraditórios em Nível Semântico
Múltiplas Interpretações	Haverá diferentes interpretações acerca do mesmo texto, uma vez que cada interprete o compreenderá a partir do seu mundo, ou seja, de todas as experiências vivenciadas por ele e que contribuíram para a formação do seu conhecimento.	1. Inclusão e Análise de Material 2. Resultados Relacionados ao Mesmo Contexto 3. Relacionamentos entre Elementos de Informação
Suspeição	Propõe o reconhecimento de falsas preconcepções, ou seja, a descoberta de imprecisão e de erros no processo de compreensão	Nível Básico 1. Indicação de Erros e Falhas 2. Avaliação de Parágrafos Nível Intermediário 3. Avaliação de Erros e Falhas realizada automaticamente pelo próprio sistema
Compartilhamento de Linguagem	Está baseado na necessidade de compartilhamento de linguagem entre texto e intérprete para que a compreensão seja possível	1. Mapa de Conceitos 2. Notas Explicativas 3. Sinônimos e Termos Relacionados 4. Variações Linguísticas

Fonte: desenvolvido pela autora

O princípio da contextualização tem como objetivo oferecer informação para que o usuário possa compreender um tema a partir do seu contexto social e histórico.

Para que isto fosse possível, a proposta visa o desenvolvimento de um histórico do autor que, em um nível básico, reunirá todas as informações do autor e, em um nível intermediário, permitirá a realização de buscas baseadas em aspectos regionais e culturais.

A interação entre pesquisadores e sujeitos fundamentou os recursos relacionados aos comentários interpretativos e aos relacionamentos entre elementos de informação, que são aqueles realizados em dois contextos: entre os conceitos da ontologia e na indexação semântica dos documentos (ontologia e documento). Estes recursos possibilitam o diálogo entre os usuários que interagem com a interface, tanto na interpretação quanto indexação de um documento.

O princípio da abstração e da generalização contribuiu a proposta do recurso que apresenta ao usuário as áreas de conhecimento de um determinado termo de busca. Seu objetivo é fornecer subsídios para a produção de resultados de caráter abstrato e genérico, sendo importante no contexto em que são necessárias abstrações teóricas e generalizações, com o intuito de adotar uma teoria ou um conjunto de conceitos para darem subsídios à um determinado estudo.

O princípio do raciocínio dialético sugere o confronto de ideias que conduziram uma proposta de pesquisa, discutindo as preconcepções do pesquisador com a experiência adquirida durante o processo de investigação. Em um nível básico, foram projetados dois recursos a partir deste princípio: 1) apresentação de antônimos e termos conceitualmente contraditórios, sugeridos a partir de um dicionário de antônimos; 2) sugestão de pesquisas complementares com resultados contraditórios a partir de antônimos. Em um nível intermediário, estas duas funcionalidades devem ser elaboradas com base em conceitos que se contradizem semanticamente e não apenas em antônimos perfeitos.

O princípio das múltiplas interpretações oferece ao usuário a possibilidade de compreender as diferentes visões abordadas em relação ao assunto pesquisado. Para que isto fosse possível, foram sugeridos três recursos neste contexto: 1) inclusão e análise de material, 2) resultados relacionados ao mesmo contexto e 3) relacionamentos entre elementos de informação (que foi proposto também com base no princípio da interação entre pesquisadores e sujeitos).

O princípio da suspeição, por sua vez, permite o questionamento crítico da realidade apresentada, visando identificar possíveis vieses, distorções e preconceitos ilegítimos. Neste sentido, em um nível básico foram propostas duas funcionalidades: indicação de erros e falhas e avaliação dos parágrafos de um texto. Em um nível intermediário, a avaliação dos problemas de um documento deve ser realizada

automaticamente pelo próprio sistema, com a implementação de recursos e técnicas de inteligência artificial.

A etapa final de desenvolvimento do “Modelo de interface de busca e recuperação de informação” apresentou os resultados da avaliação de usabilidade, que teve como objetivo identificar a percepção dos usuários sobre a proposta desenvolvida, bem como identificar possíveis erros de interação, *layout* e acessibilidade.

8 CONSIDERAÇÕES FINAIS

O objetivo que norteou o desenvolvimento desta pesquisa foi a elaboração de um "Modelo de interface de busca e recuperação de informação baseado nos princípios da hermenêutica e na utilização de ontologias", com funcionalidades projetadas para proporcionar uma melhor interação do usuário com o sistema e contribuir com a análise da similaridade existente entre os documentos e com a interpretação de um *corpus* documental.

A investigação partiu inicialmente do pressuposto de que uma interface de busca com recursos gráficos elaborados a partir de uma ontologia teria potencial para ampliar a busca do usuário e favorecer a compreensão e aquisição de conhecimento em um determinado domínio, bem como para desenvolver novas percepções acerca das suas necessidades de informação. Em um segundo momento, levantou-se também a hipótese de que os princípios da hermenêutica poderiam trazer contribuições no contexto da interpretação dos resultados recuperados.

Com base neste cenário, as primeiras etapas do estudo foram destinadas à revisão de literatura, que permitiu a identificação dos principais problemas do processo de busca e recuperação de informação, que estão relacionados com os aspectos de comunicação entre o usuário e a interface. Neste sentido, para que o "diálogo" aconteça de forma satisfatória, é necessário que termos e palavras-chave sejam elaboradas na tentativa de informar ao sistema quais são as necessidades de informação. Apesar de parecer um processo simples, ele exige que o usuário possua conhecimentos sobre a terminologia da área investigada, tanto para elaborar expressões de busca significativas, quanto para relacionar os conceitos e resultados dentro do domínio de interesse.

Além da necessidade de conhecimento sobre o domínio da busca, os problemas linguísticos também foram identificados como fatores de interferência na recuperação de informação, principalmente nos aspectos associados à ambiguidade das palavras e à utilização de múltiplas nomenclaturas para representar um mesmo conceito. Estes elementos interferem tanto na busca, quanto na indexação dos documentos, uma vez que o usuário deve observar o contexto associado ao termo utilizado e conhecer as inúmeras variações de um mesmo vocábulo.

Algumas características também estão relacionadas com a forma como os usuários elaboram uma expressão de busca e interpretam os resultados recuperados que, no contexto da pesquisa, foram denominadas de variáveis interferentes. As mais citadas na literatura estudada estão relacionadas com: 1) dissonância cognitiva, que são barreiras

socioculturais produzidas por normas sociais e valores culturais, incluindo crenças, orientação política, preferências, preconceitos etc.; 2) nível educacional e base de conhecimento, que se referem ao conhecimento que o usuário possui sobre um determinado assunto, que é influenciado pela sua formação, experiências, pressupostos teóricos, entre outros; 3) variáveis demográficas, que dizem respeito à faixa etária, sexo, nível social, experiência de trabalho etc.

Outro ponto importante que limita as buscas realizadas nos sistemas de recuperação de informação é que documentos relacionados em diferentes perspectivas são apresentados ao usuário de forma isolada e desconectada, mesmo que existam inúmeros contextos de similaridade entre eles. Desta forma, para que se possa recuperar as várias traduções de uma mesma obra, por exemplo, é necessário que diversas pesquisas sejam realizadas para que este objetivo seja alcançado, uma vez os exemplares podem estar disponibilizados em bases de dados diferentes e indexados por termos distintos.

Para tentar minimizar os problemas identificados, foi realizado um estudo teórico sobre as ontologias, que mostrou as principais abordagens envolvidas, os problemas da área e as propostas de melhorias no âmbito da busca e recuperação de informação. Neste ponto, foi possível identificar uma lacuna de pesquisa: apesar das inúmeras propostas e projetos observados na área, é necessário o desenvolvimento de ferramentas que utilizem, de fato, os recursos oferecidos pelas ontologias, principalmente o contexto e o significado existente nas relações entre os conceitos, com o objetivo de ampliar as possibilidades de busca e não apenas filtrar informações e resultados.

Tão importante quanto as ontologias, os princípios da hermenêutica ofereceram aporte teórico para o entendimento do processo de interpretação documental, que influencia diretamente o comportamento do usuário durante a análise dos resultados recuperados em uma busca, bem como o significado das relações existentes entre eles.

O “Modelo de interface de busca e recuperação de informação” proposto neste estudo foi desenvolvido por meio do ciclo de vida do *Design* de Interação e projetado a partir de funcionalidades fundamentadas nos princípios da hermenêutica e na utilização de ontologias. Apesar dos protótipos terem sido desenvolvidos com foco nos repositórios digitais de informação científica, a interface de busca é genérica e tem potencial para ser aplicada em múltiplos contextos, tais como: sistemas organizacionais, empresariais, acadêmicos, de gestão de projetos entre outros.

O principal diferencial da proposta está na possibilidade de interação do usuário com uma interface gráfica, que permite a elaboração de expressões de busca a partir de um mapa de conceitos (ontologia de domínio) e a análise dos documentos recuperados sob

a perspectiva da similaridade existente entre os documentos, com a intuito de possibilitar uma melhor compreensão e interpretação do domínio e das relações existentes no *corpus* documental. É, portanto, um modelo de interação para o desenvolvimento de pesquisas semânticas, a partir de múltiplas perspectivas e com um maior grau de detalhamento e refinamento.

A avaliação de usabilidade realizada com os protótipos desenvolvidos apresentou resultados positivos e demonstrou que o desenvolvimento do modelo proposto é viável. Na concepção do usuário, a maioria dos recursos foram considerados úteis e de grande importância para a elaboração de atividades de pesquisa e de análise de documentos. No entanto, apesar do valor agregado ser evidente, é uma perspectiva nova de busca e recuperação de informação e necessita, portanto, de um esforço natural para a familiarização dos pesquisadores.

Embora a maioria dos resultados terem sido positivos, foram identificadas inúmeras dificuldades que poderiam limitar o uso da interface. No entanto, como os testes foram realizados apenas com protótipos de baixa fidelidade, a interação do usuário com o sistema foi apenas simulada e não permitiu o levantamento de informações mais detalhadas, tampouco o desenvolvimento de diversas atividades no mesmo contexto para permitir um maior aprendizado sobre as funcionalidades propostas.

Apesar das limitações deste estudo, comprova-se a tese de que interfaces de busca baseadas nos princípios da hermenêutica e na utilização de ontologias possuem potencial para a melhoria do processo de recuperação de documentos textuais. Desta forma, acredita-se que as funcionalidades previstas possam facilitar a formulação de expressões de busca, a análise da similaridade entre os documentos e a interpretação de um *corpus* documental a partir de múltiplas perspectivas de análise.

O estudo sobre interfaces de busca e recuperação de informação desenvolvidas a partir dos princípios da hermenêutica e da utilização de ontologias deve ser aprofundado em pesquisas futuras. Como sugestão de investigações complementares, seria relevante que a interface proposta fosse implementada e que testes mais aprofundados pudessem ser desenvolvidos com usuários em contextos de buscas reais, incluindo também avaliação de desempenho para analisar o tempo de resposta do sistema, principalmente com o aumento da complexidade da ontologia de domínio.

Além disso, as diretrizes identificadas neste trabalho podem servir de base para a elaboração de um instrumento hermenêutico de avaliação de interfaces de busca, permitindo que outros sistemas possam ser avaliados a partir das funcionalidades aqui

previstas. Neste sentido, um número maior de critérios poderia ser definido para possibilitar pelo menos três níveis de classificação: básico, intermediário e avançado.

REFERÊNCIAS

ADAMPEASE. **Suggested Upper Merged Ontology**. 2017. Disponível em: <http://www.adampease.org/OP/>. Acesso em: 10 jun. 2017.

AIRIO, E. et al. CIRI: an ontology-based query interface for text retrieval. In: HYVÖNEN, E.; KAUPPINEN T.; SALMINEN, M.; VILJANEN, L.; ALA-SIURU, P. (Eds). **Proceedings of the 11th Finnish Artificial Intelligence Conference**, 2004.

ALBERTONI, R.; BERTONE, A.; DE MARTINO, M. Visualization and semantic analysis of geographic metadata. In: PROCEEDINGS OF THE 2005 WORKSHOP ON GEOGRAPHIC INFORMATION RETRIEVAL (GIR '05). ACM, New York, NY, USA, 2005. p. 9-16.

ALCÂNTARA, V. C; PAIVA, A. L.; BRITO, M. J. Desvelando 'caixas-pretas' dos textos de estratégia: uma abordagem baseada na hermenêutica crítica. **Organizações & Sociedade**, Salvador, v. 25, n. 84, p. 30-49, jan./mar. 2018.

ALISI, T. et al. Sirio: an ontology-based web search engine for videos. In: MM'09 - PROCEEDINGS OF THE 2009 ACM MULTIMEDIA CONFERENCE, WITH CO-LOCATED WORKSHOPS AND SYMPOSIUMS, 2009. p. 967-968.

ALLAN, J.; CARTERETTE, B.; LEWIS, J. When will Information Retrieval be "Good Enough?". In: PROCEEDINGS OF THE 28TH ANNUAL INTERNATIONAL ACM SIGIR CONFERENCE ON RESEARCH AND DEVELOPMENT IN INFORMATION RETRIEVAL (SIGIR '05). ACM, New York, NY, USA, p. 433-440, 2005.

ALMEIDA, M. B. Uma abordagem integrada sobre ontologias: Ciência da Informação, Ciência da Computação e Filosofia. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 19, n. 3, p. 242-258, jul./set. 2014.

ALMEIDA, M. B.; BAX, M. P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Ciência da Informação**, Brasília, v. 32, n. 3, p. 7-20, set./dez. 2003.

ALTOUNIAN, M. M. de A.; GOMES, B. P. de M. A recuperação semântica da informação no contexto do controle externo. **Revista do TCU**, n. 137, p. 31-41, 2016.

ANTONIOU, G.; HARMELEN, F. Van. **A semantic web primer.** Cambridge: MIT Press, 2004.

ARAUJO JUNIOR, R. H. **Precisão no processo de busca e recuperação da informação**. Brasília: Thesaurus, 2007.

ARFAOUI, O.; HIDRI, M. S. Ontology-based query refinement for XML information retrieval. In: PROCEEDINGS OF THE 6th INTERNATIONAL CONFERENCE ON WEB INTELLIGENCE, MINING AND SEMANTICS (WIMS '16), Rajendra Akerkar, Michel Plantié, Sylvie Ranwez, Sébastien Harispe, Anne Laurent, Patrice Bellot, Jacky Montmain, and François Trousset (Eds.). ACM, New York, NY, USA, Article 35, 4 pages, 2016.

ATSRI. 4th. International Conference on Applied and Theoretical Information Systems Research. Taipei, Taiwan, November 22-24, 2013. Disponível em: <http://atisr2013.internet-

studies.org/>. Acesso em: 15 set. 2013.

BAEZA-YATES, R.; RIBEIRO-NETO, B. **Recuperação de informação**: conceitos e tecnologias das máquinas de busca. 2. ed. Porto Alegre: Bookman, 2013.

BARBOSA, S. D. J.; SILVA, B. S. da. **Interação humano-computador**. Rio de Janeiro: Elsevier, 2010.

BARONAS, J. E. A. Variação linguística na escola: propostas de ação. **SIGNUM: Estud. Ling.**, Londrina, n. 14/2, p. 105-116, dez. 2011.

BASKAYA, F.; KEKÄLÄINEN, J.; JÄRVELIN, K. A tool for ontology-editing and ontology-based information exploration. In: PROCEEDINGS OF THE THIRD WORKSHOP ON EXPLOITING SEMANTIC ANNOTATIONS IN INFORMATION RETRIEVAL (ESAIR '10). ACM, New York, NY, USA, 2010. p. 29-30.

BAST, H. et al. ESTER: efficient search on text, entities, and relations. In: PROCEEDINGS OF THE 30th ANNUAL INTERNATIONAL ACM SIGIR CONFERENCE ON RESEARCH AND DEVELOPMENT IN INFORMATION RETRIEVAL (SIGIR '07). ACM, New York, NY, USA, 2007. p. 671-678.

BATES, M. J. Search and idea tactics. In: WHITE, H. D.; BATES, M. J.; WILSON, P. (Eds.). **For information specialists**: interpretations of reference and bibliographic work. Norwood, NJ: Ablex, 1992. p. 183-200.

BATISTA, M. Hermenêutica filosófica e o debate Gadamer-Habermas. **Critica e Sociedade:** revista de cultura política, v. 2, n. 1, p. 101-118, jan./jun. 2012.

BEPPLER, F. D. **Emprego de RBC para recuperação inteligente de informações**. 2002. 100 f. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal de Santa Catarina, Florianópolis, 2002.

_____. **Um modelo para recuperação e busca de informação baseado em ontologia e no círculo hermenêutico**. 2008. 123 f. Tese (Doutorado em Engenharia e Gestão do Conhecimento) – Universidade Federal de Santa Catarina, Florianópolis, 2008.

BERA, P.; BURTON-JONES, A.; WAND, Y. Guidelines for designing visual ontologies to support knowledge identification. **MIS Quarterly**, v. 35, n. 4, p. 883-908, 2011.

BERNSTEIN, A. et al. Querying ontologies: a controlled english interface for end-users. In: PROC. 4TH INTERNATIONAL SEMANTIC WEB CONFERENCE (ISWC05), Springer-Verlag Berlim Heidelberg, nov. 2005. p. 112-126.

BERTINI, M. et al. Sirio, orione and pan: an integrated web system for ontology-based video search and annotation. In: PROCEEDINGS OF THE 18th ACM INTERNATIONAL CONFERENCE ON MULTIMEDIA (MM '10). ACM, New York, NY, USA, 2010. p. 1625-1628.

BOCCATO, V. R. C.; RAMALHO, R. A. S.; FUJITA, M. S. L. A contribuição dos tesauros na construção de ontologias como instrumento de organização e recuperação da informação em ambientes digitais. **IBERSID: Revista de Sistemas de Información y Documentación**, número monográfico, p. 199-209, 2008.

BOJARS, U. et al. Extending OWL ontology visualizations with interactive contextual verbalization. In: VOILA! 2th. INTERNATIONAL WORKSHOP ON VISUALIZATION AND INTERACTION FOR ONTOLOGIES AND LINKED DATA. Co-located with ISWC 2016, Kobe, Japan, October 17, 2016. p. 5-16.

BONFIM, V. S. Gadamer e a experiência hermenêutica. **Revista CEJ**, Brasília, ano XIV, n. 49, p. 76-82, abr./jun. 2010.

BONINO, D. et al. Ontology driven semantic search. **WSEAS Transaction on Information Science and Application**, v. 1, n. 6, p. 1597-1605, 2004.

BORKO, H. Information science: what is it? **American Documentation**, v.19, n.1, p.3-5, jan. 1968.

BRANCO NETO, W. C. **Web Semântica na construção de sistemas de aprendizagem adaptativos**. 2006. 219 f. Tese (Doutorado em Ciência da Computação) - Universidade Federal de Santa Catarina, Florianópolis, 2006.

BRANDES, U. Drawing on physical analogies. In: KAUFMANN, M.; WAGNER, D. (Eds). **Drawing graphs**: methods and models. Springer LNCS Tutorial 2025, p.71-86, 2001.

BRANSKI, R. M. Recuperação de informações na Web. **Perspectivas da Ciência da Informação**, Belo Horizonte, v. 9, n. 1, p. 70-87, jan./jun. 2004.

BRÄSCHER, M. A ambigüidade na recuperação da informação. **DataGramaZero: Revista de Ciência da Informação**, v. 3, n. 1, fev. 2002.

BREGAGNOL, C. S.; LARA, L. Z. **Uma análise do tratamento da semântica nos livros didáticos**. 2012. Monografia (Especialização em Gramática Ensino de Língua Portuguesa) - Universidade Federal do Rio Grande do Sul, Instituto de Letras, Porto Alegre, 2012.

BREITMAN, K. **Web semântica**: a internet do futuro. Rio de Janeiro: LTC, 2005.

BRITO, R. M. et al. A hermenêutica e o processo de construção do conhecimento. **Dialógica**. Manaus, v. 1, n. 3, p.1-12, 2007.

BUNTAIN, C. 3D ontology visualization in semantic search. In: PROCEEDINGS OF THE 46th ANNUAL SOUTHEAST REGIONAL CONFERENCE ON XX (ACM-SE 46). ACM, New York, NY, USA, 2008. p. 204-208.

BURCH, M.; LOHMANN, S. Visualizing the evolution of ontologies: a dynamic graph perspective. In: VOILA! INTERNATIONAL WORKSHOP ON VISUALIZATION AND INTERACTION FOR ONTOLOGIES AND LINKED DATA. Co-located with ISWC 2015, Bethlehem, Pennsylvania, USA, October 11, 2015. p. 69-76.

CAIDI, N.; ALLARD, D.; QUIRKE, L. Information practices of immigrants. **Annual review of information science and technology**, Medford, NJ: Information Today, v. 44, p. 493-531, 2010.

CAMACHO, R. G. Norma culta e variedades lingüísticas. In: CECCANTINI, J. L. C. T.; PEREIRA, R. F. (Eds.). **Cadernos de formação**: língua portuguesa. São Paulo: Prograd UNESP, 2004. p. 47-60.

CAPURRO, R. Epistemologia e Ciência da informação. In: V ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 5., Belo Horizonte, 2003. **Anais...** Belo Horizonte: Escola de Ciência da Informação da UFMG, 2003.

CARD, S.; MACKINLAY, K. J. D.; SHNEIDERMAN, B. (Eds.). **Readings in information visualization**: using vision to think. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1999.

CARROLL, J. M. Five reasons for scenario-based design. **Interacting with computers**, v. 13, n. 1, p. 43-60, 2000.

CARVALHO, C. R. M. et al. Unindo IHC e negócios através do uso de personas: um estudo de caso no mercado de aplicativos móveis. In: PROCEEDINGS OF THE 10TH BRAZILIAN SYMPOSIUM ON HUMAN FACTORS IN COMPUTING SYSTEMS AND THE 5TH LATIN AMERICAN CONFERENCE ON HUMAN-COMPUTER INTERACTION (IHC+CLIHC '11). Brazilian Computer Society, Porto Alegre, Brazil, 2011. p. 100-104.

CARVALHO, M. et al. Linking dynamic query interfaces to knowledge models. In: PROCEEDINGS OF THE 7th INTERNATIONAL CONFERENCE ON INTELLIGENT USER INTERFACES (IUI '02). ACM, New York, NY, USA, 2002. p. 186-187.

CASE, D. O. **Looking for information**: a survey of research on information seeking, needs, and behavior. 3. ed. Bingley, UK: Emerald Publishing Group, 2012.

CASTILHO, A. T. de. **Nova gramática do português brasileiro**. São Paulo: Contexto, 2010.

CERVO, A. L.; BERVIAN, P. A.; DA SILVA, R. **Metodologia científica**. 6. ed. São Paulo: Pearson Prentice Hall, 2007.

CHAUÍ, M. **Convite à filosofia**. 13. ed. São Paulo: Ática, 2003.

CHEN, L.; CHUA, C. Interactive interface for query formulation. In: PROCEEDINGS OF THE 25th AUSTRALIAN COMPUTER-HUMAN INTERACTION CONFERENCE: AUGMENTATION, APPLICATION, INNOVATION, COLLABORATION (OzCHI '13), Haifeng Shen, Ross Smith, Jeni Paay, Paul Calder, and Theodor Wyeld (Eds.). ACM, New York, NY, USA, 2013. p. 507-510.

CHOWDHURY, S.; GIBB, F. Relationship among activities and problems causing uncertainty in information seeking and retrieval. **Journal of Documentation**, v. 65, n. 3, p. 470-499, 2009.

CHUA, W. W. K.; KIM, J. Semantic querying over knowledge in biomedical text corpora annotated with multiple ontologies. In: PROCEEDINGS OF THE ACM CONFERENCE ON BIOINFORMATICS, COMPUTATIONAL BIOLOGY AND BIOMEDICINE (BCB '12). ACM, New York, NY, USA, 2012. p. 400-407.

CODINA, L. El diseño de la navegación en hipertextos informativos. **Temas de Diseño**, n. 18, p. 38-48, abr. 2001.

COLE, M. J. et al. User activity patterns during information search. **ACM Trans. Inf. Syst**, v. 65, n. 1, art. 1, p. 1-39, mar. 2015.

COOPER, A. **Inmates are running the asylum**: why high-tech products drive us crazy and how to restore the sanity. Sams Publishing, 1999.

COTHEY, V. A longitudinal study of World Wide Web users' information-searching behaviour. **Journal of the American Society for Information Science and Technology**, v. 53, n. 2, p. 67-78, 2002.

CRISTOVÃO, H. M.; DUQUE, C. G.; SERQUEIRA, L. D. Recuperação de informação: uma aplicação na criação e configuração automáticas de cursos virtuais a distância. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB), 13., 2012, Rio de Janeiro. **Anais…** Rio de Janeiro: Fiocruz, 2012.

DASSAULT SYSTÈMES (3D Perspectives). **Visual search engines**. Blog oficial corporativo da Dassault Systèmes. 2009. Disponível em: <http://blogs.3ds.com/perspectives/visual-search-engines/>. Acesso em: 19 abr. 2018.

DBPEDIA. **About Dbpedia.** 2017. Disponível em: <http://wiki.dbpedia.org/about>. Acesso em: 05 jun. 2017.

DELIGIANNIDIS, L.; KOCHUT, K. J.; SHETH, A. P. RDF data exploration and visualization. In: PROCEEDINGS OF THE ACM FIRST WORKSHOP ON CYBERINFRASTRUCTURE: INFORMATION MANAGEMENT IN ESCIENCE (CIMS' 07). ACM, New York, NY, USA, 2007. p. 39-46.

DEMIDOVA, E.; ZHOU, X.; NEJDL, W. Efficient query construction for large scale data. In: PROCEEDINGS OF THE 36th INTERNATIONAL ACM SIGIR CONFERENCE ON RESEARCH AND DEVELOPMENT IN INFORMATION RETRIEVAL (SIGIR '13). ACM, New York, NY, USA, 2013. p. 573-582.

DERVIN, B. From the mind's eye of the 'user': the sense-making qualitative-quantitative methodology. In: GLAZER, J. D.; POWELL, R. R. (Eds.). **Qualitative research in information management**. Englewood: CO: Libraries Unlimited, 1992. p. 61-84.

_____. On studying information seeking methodologically: the implications of connecting metatheory to method. **Information Processing & Management**, v. 35, n. 6, p. 727-750, 1999.

DIAS, E. A. V.; COSTA, H. G. Mapeamento da produção científica no escopo da ontologia. **Sistemas & Gestão**, v. 6, n. 4, p. 481-507, 2011.

DILTHEY, W. A compreensão dos outros e das suas manifestações de vida. In: _____. **Teorias da História**. Organização, introdução e comentários de Patrick Gardiner. Lisboa: Fundação Calouste Gulbenkian, 1984. (Ensaio originalmente publicado no vol. VII dos Gesammelte Schriften).

_____. O surgimento da hermenêutica (1900). **Numen: Revista de estudos e pesquisa da religião**. Juiz de Fora, v. 2, n. 1, p. 11-32, jan./jun., 1999.

_____. **Obras escolhidas**: a fundação do mundo filosófico nas ciências humanas. v. 3, 2002.

_____. Princípios gerais sobre a conexão entre as ciências humanas. In: _____. **A construção do mundo histórico nas ciências humanas**. São Paulo: Editora UNESP, 2010. p.

75-127.

DOMINGUES, I. **Epistemologia das Ciências Humanas**: Positivismo e Hermenêutica. São Paulo: Loyola, 2004.

DUDÁŠ, M.; ZAMAZAL, O.; SVÁTEK, V. Roadmapping and navigating in the ontology visualization landscape. In: JANOWICZ, K.; SCHLOBACH, S.; LAMBRIX, P., HYVÖNEN, E. (Eds.). **EKAW 2014**. **LNCS**, v. 8876, p.137-152, Springer, Heidelberg, 2014.

DUDYCZ, H. Usability of business information semantic network search visualization. In: PROCEEDINGS OF THE MULITIMEDIA, INTERACTION, DESIGN AND INNNOVATION (MIDI '15). ACM, New York, NY, USA, Article 13, 9 pages, 2015.

ECIR. 35th. European Conference on Information Retrieval. Moscow, Russia, March 24-27, 2013. Disponível em: <http://ecir2013.org/>. Acesso em: 15 set. 2013.

EL-ATAWY, S. S.; KHALEFA, M. E. Building an Ontology-Based Electronic Health Record System. In: PROCEEDINGS OF THE 2nd AFRICA AND MIDDLE EAST CONFERENCE ON SOFTWARE ENGINEERING (AMECSE '16). ACM, New York, NY, USA, 2016. p. 40-45.

ELLIS, D. A behavioural approach to information retrieval system design. **Journal of Documentation**, London, v. 45, n. 3, p. 171-212, 1989.

ENANCIB. Encontro Nacional de Pesquisa em Ciência da Informação (XIV ENANCIB). Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina, Out. 29 a Nov. 01, 2013. Disponível em: <http://enancib2013.ufsc.br/index.php/enancib2013/XIVenancib>. Acesso em: 25 set. 2013.

_____. (XVIII ENANCIB). Universidade Estadual Paulista, Marília, São Paulo, Outubro, 23-27, 2017. Disponível em: <http://enancib.marilia.unesp.br/index.php/xviiienancib/ENANCIB>. Acesso em: 16 abr. 2018.

FACHIN, O. **Fundamentos de metodologia**. 4. ed. São Paulo: Saraiva, 2003.

FENSEL, D. **Ontologies:** a silver bullet for knowledge management and eletronic commerce. 2. ed. New York: Springer-Verlag. 2003.

FERNANDEZ, M.; GOMEZ-PEREZ, A.; JURISTO, H. Methontology: from ontological art towards ontological engineering. In: PROCEEDINGS OF AAAI97 SPRING SYMPOSIUM SERIES, WORKSHOP ON ONTOLOGICAL ENGINEERING, Stanford, CA, 1997. p.33-40.

FERNEDA, E. **Introdução aos modelos computacionais de recuperação de informação**. Rio de Janeiro: Ciência Moderna, 2012.

_____. **Ontologia como recurso de padronização terminológica de um sistema de recuperação de informação**. 2013a. 96 f. Relatório de Pesquisa (Pós-Doutorado em Ciência da Informação) – Universidade Federal da Paraíba, João Pessoa, 2013.

_____. **Recuperação de informação**: análise sobre a contribuição da Ciência da Computação para a Ciência da Informação. 2003. 137 f. Tese (Doutorado em Ciências da Comunicação) – Universidade de São Paulo, São Paulo, 2003.

FERREIRA, A. B. H. **Miniaurélio**: o minidicionário da língua portuguesa. 6. ed. Curitiba: Posigraf, 2004.

FIGUEIREDO, N. M. **Estudo de uso e usuários da informação**. Brasília: IBICT, 1994.

FININ, T. et al. Information retrieval and the semantic web. In: PROCEEDINGS OF THE PROCEEDINGS OF THE 38TH ANNUAL HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES (HICSS'05). IEEE Computer Society, 2005.

FIRE. Forum for Information Retrieval Evaluation (8th FIRE). ISI, Kolkata, December, 07-10, 2016. Disponível em: <http://fire.irsi.res.in/fire/2016/home>. Acesso em: 16 maio 2016.

FORD, N.; EAGLESTONE, B.; MADDEN, A. Web searching by the general public: an individual differences perspective. **Journal of Documentation**, v. 65, n. 4, p. 632-667, 2009.

FREIRE, F.; PINHEIRO, V.; FEITOSA, D. FlexSTS: um framework para similaridade semântica textual. **LinguaMÁTICA**, v. 8, n. 2, p. 23-31, 2016.

FREITAS, A.; CURRY, E. Natural language queries over heterogeneous linked data graphs: a distributional-compositional semantics approach. In: PROCEEDINGS OF THE 19th INTERNATIONAL CONFERENCE ON INTELLIGENT USER INTERFACES (IUI '14). ACM, New York, NY, USA, 2014. p. 279-288.

FREITAS, C. M. D. S. et al. Introdução à visualização de informações. **RITA – Revista de Informática Teórica e Aplicada**, Porto Alegre, v. 8, n. 2, p. 143-158, out. 2001.

GADAMER, H.-G. **O problema da consciência histórica**. Organização de Pierre Fruchon. Tradução de Paulo César Duque Estrada. 3. ed. Rio de Janeiro: FGV, 2006.

_____. **Verdade e método**: traços fundamentais de uma hermenêutica filosófica. Tradução de Flávio Paulo Meurer. 3. ed. Petrópolis: Vozes, 1999. (Obra originalmente publicada em 1960).

_____._____. Tradução de Flávio Paulo Meurer. 7. ed. Petrópolis: Vozes, Bragança Paulista: EDUSF, 2005. (Coleção pensamento humano).

GAO, Y.; LUO, H.; FAN, J. Searching and browsing large scale image database using keywords and ontology. In: PROCEEDINGS OF THE 14th ACM INTERNATIONAL CONFERENCE ON MULTIMEDIA (MM '06). ACM, New York, NY, USA, 2006. p. 811-812.

GARCÍA, E.; SICILIA, M. A. Designing ontology-based interactive information retrieval interfaces. In: PROCEEDINGS OF THE WORKSHOP ON HUMAN COMPUTER INTERFACE FOR SEMANTIC WEB AND WEB APPLICATIONS, Springer Lecture Notes in Computer Science 2889, 2003. p. 152-165.

GASQUE, K. C. G. D.; COSTA, S. M. de S. Evolução teórico-metodológica dos estudos de comportamento informacional de usuários. **Ciência da Informação**, Brasília, v. 39, n. 1, p. 21-32, jan./abr. 2010.

GAVANKAR, C.; LI, Y-F.; RAMAKRISHNAN, G. Context-driven concept search across Web Ontologies using keyword queries. In: PROCEEDINGS OF THE 8th

INTERNATIONAL CONFERENCE ON KNOWLEDGE CAPTURE (K-CAP 2015). ACM, New York, NY, USA, Article 20, 4 pages, 2015.

GIL, A. C. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Atlas, 2008.

GOMES, D. Hermenêutica e comunicação: contribuições para compreender a teoria da interpretação e sua aplicação na sociedade midiática. **Temática**, ano XI, n. 4, p. 38-52, 2015.

GOOGLE DEVELOPERS. **Freebase**. 2017. Disponível em: <https://developers.google.com/ freebase/>. Acesso em: 05 jun. 2017.

GOOGLE PESQUISA. **Pesquisa com o termo Estrela**. 2018. Disponível em: <www.google.com.br>. Acesso em: 05 jun. 2018.

GRĂDINARU, C. The information searching behaviour in the digital age. **Argumentum**: Journal the Seminar of Discursive Logic, Argumentation Theory & Rhetoric, v. 11, n. 1, p. 90-101, 2013.

GRONDIN, J. **Hermenêutica**. São Paulo: Parábola Editorial, 2012.

GROTH, K.; LANNERÖ, P. Context browser: ontology based navigation in information spaces. In: PROCEEDINGS OF THE 1st INTERNATIONAL CONFERENCE ON INFORMATION INTERACTION IN CONTEXT (IIiX). ACM, New York, NY, USA, 2006. p. 75-78.

GRUBER, T. R. A translation approach to portable ontology specification. **Knowledge Acquisition**, v. 5, n. 2, p. 199-220, 1993.

_____. What is an ontology? 1996. Disponível em: <http://www.ksl.stanford.edu/ kst/what-is-an-ontology.html>. Acesso em: 12 mar. 2017.

GUARINO, N. Formal ontology in information systems. In: PROCEEDINGS OF THE FIRST INTERNATIONAL CONFERENCE, Trento, Italy, jun. 1998. Amsterdam, IOS Press, p. 3-15. 1998.

GUARINO, N. Formal ontology, conceptual analysis and knowledge representation. **International Journal of Human-Computer Studies**, v. 43, n. 5/6, p. 625-640, 1995.

GUARINO, N.; GIARETTA, P. Ontologies and KBs, towards a terminological clarification. In: MARS, N. (Ed.). **Towards a very large knowledge bases**: knowledge building and knowledge sharing. [S. l.]: IOS Press, 1995. p. 25-32.

GUARINO, N.; MASOLO, C.; VETERE, G. Ontoseek: content-based access to the web. **IEEE Intelligent Systems**, v. 14, n. 3, p. 70-80, 1999.

GUARINO, N.; OBERLE, D.; STAAB, S. What is an ontology? In: STAAB, S.; STUDER, R. (Ed.) **Handbook on ontologies**. 2. ed. Berlin: Springer, 2009.

GUPTA, C.; BHOWMIK, R.; GOVINDARAJU, M. Ontological framework for a free-form query based grid search engine. In: PROCEEDINGS OF THE 17th INTERNATIONAL SYMPOSIUM ON HIGH PERFORMANCE DISTRIBUTED COMPUTING (HPDC '08). ACM, New York, NY, USA, 2008. p. 205-208.

HAAV, H. M.; LUBI, T. L. A survey of concept-based information retrieval tools on the web. In: PROC. OF 5th EAST-EUROPEAN CONFERENCE ADBIS*2001. A. Caplinkas and J. Eder (Eds). Advances in Databases and Information Systems, Vilnius Technika, v. 2, p. 29-41. 2001.

HAYTER, S. Exploring information worlds in a disadvantaged community: a UK perspective. **Canadian Journal of Information & Library Science**, v. 30, n. 1-2, p. 21-35, 2006.

HEARST, M. A. **Search user interfaces**. New York: Cambridge University Press, 2009.

HEIDEGGER, M. **Ser e tempo**. Petrópolis: Vozes, 2005.

HERMAN, I.; MELANÇON, G.; MARSHALL, M. S. Graph visualization and navigation in information visualization: a survey. **IEEE Transactions on Visualization and Computer Graphics**, v. 6, n. 1, p. 1-21, 2000.

HIEMSTRA, D. Information retrieval models. In: GOKER, A.; DAVIES, J. **Information retrieval**: searching in the 21st Century. Wiley, 2009. p. 1–17.

HOBBS, V.; PFITZNER, D.; POWERS, D. A survey of information retrieval interfaces. 2002. Disponível em: <https://www.researchgate.net/publication/241180283_a_survey_of_information_retrieval_interfaces>. Acesso em: 30 maio 2018.

HOGAN, A. et al. Searching and browsing Linked Data with SWSE: the semantic web search engine. **Web Semantics**: Science, Services and Agents on the World Wide Web, v. 9, n. 4, p. 365-401, dez. 2011.

HUANG, L. **A survey on Web information retrieval technologies**. New York, University of New York, 1999.

HYVÖNEN, E., MÄKELÄ, E. et al. MuseumFinland: finnish museums on the semantic web. **Journal of Web Semantics**, v. 3, n. 2-3, p. 224-241, 2005.

INGWERSEN, P. **Information retrieval interaction**. London: Taylor Graham, 1992.

INWOOD, M. **Hermenêutica**. In: CRAIG, E. (Org.). Routledge Encyclopedia of Philosophy. Londres: Routlegde. Tradução de Rogério Bettoni. 1998. Disponível em: <http://criticanarede.com/hermeneutica.html>. Acesso em: 06 jul. 2018.

ISSERTIAL, L.; TSUJI, H. Data management and user interface for a call for paper manager. In: COMPUTER SOFTWARE AND APPLICATIONS CONFERENCE (COMPSAC) 2013 IEEE 37th ANNUAL, July 2013. p. 463-466.

JACKSON, P.; MOULINIER, I. **Natural language processing for online applications**: text retrieval, extraction and categorization. John Benjamins, Amsterdam, Philadelphia, 2002.

JANSEN, B. J.; SPINK, A.; SARACEVIC, T. Real life, real users, and real needs: a study and analysis of user queries on the web. **Information Processing and Management**, v. 36, n. 2, p. 207-227, 2000.

JASPER, R.; USCHOLD, M. A framework for understanding and classifying ontology applications. In: IJCAI-99, ONTOLOGY WORKSHOP. Stockholm, Sweden, jul. 1999.

JAVED, M. et al. Viz-vivo: towards visualizations-driven linked data navigation. In: VOILA! 2016: PROC., ISWC2016 INTERNATIONAL WORKSHOP ON VISUALIZATIONS AND USER INTERFACES FOR ONTOLOGIES AND LINKED DATA, Kobe, Japan. 2016. p. 80-92.

JONES, K.S.; WALKER, S.; ROBERTSON, S. E. A probabilistic model of information retrieval: development and comparative experiments – Part 2. **Information Processing and Management**, v. 36, n. 6, p. 809-840, 2000.

KAMEL, M; LEE, A.; POWERS, E. Ontology-aided vs. keyword-based web searches: a comparative user study. In: LARGE SCALE SEMANTIC ACCESS TO CONTENT (TEXT, IMAGE, VIDEO, AND SOUND) (RIAO '07). LE CENTRE DE HAUTES ETUDES INTERNATIONALES D'INFORMATIQUE DOCUMENTAIRE, Paris, France, 2007. p. 639-652.

KATIFORI, A. et al. Ontology visualization methods: a survey. **ACM Computing Surveys,** v. 39, n. 4, p.10-43, 2007.

KHATTAK, A. M. et al. Intelligent search in digital documents. In: PROCEEDINGS OF THE 2008 IEEE/WIC/ACM INTERNATIONAL CONFERENCE ON WEB INTELLIGENCE AND INTELLIGENT AGENT TECHNOLOGY - Volume 01 (WI-IAT '08), v. 1. IEEE Computer Society, Washington, DC, USA, 2008. p. 558-561.

KLEIN, H., MYERS, M. A set of principles for conducting and evaluating interpretive field studies in Information Systems. **MIS Quaterly**, v. 23, n. 1, p. 67-94, mar. 1999.

KO, J. et al. Semantically enhanced keyword search for smartphones. In: PROCEEDINGS OF THE 23RD INTERNATIONAL CONFERENCE ON WORLD WIDE WEB (WWW '14 Companion). ACM, New York, NY, USA, 2014. p. 327-328.

KOBSA, A. User experiments with tree visualization systems. In: PROCEEDINGS OF THE IEEE SYMPOSIUM ON INFORMATION VISUALIZATION (INFOVIS'04). IEEE Computer Society, Washington, DC, USA, 2004. p. 9-16.

KONDYLAKIS, H.; PLEXOUSAKIS, D. Exelixis: evolving ontology-based data integration system. In: PROCEEDINGS OF THE 2011 ACM SIGMOD INTERNATIONAL CONFERENCE ON MANAGEMENT OF DATA (SIGMOD '11). ACM, New York, NY, USA, 2011. p. 1283-1286.

KOSHMAN, S. Visualization-based information retrieval on the Web. **Library & Information Science Research**, v. 28, n. 2, p. 192-207, 2006.

KRIKELAS, J. Information seeking behavior: patterns and concepts, **Drexel Library Quarterly**, v. 19, n. 2, p. 5-20, 1983.

KUHLTHAU, C. C. Inside the search process: information seeking from the user's perspective. **Journal of the American Society for Information Science**, v. 42, n. 5, p. 361-371, 1991.

LAI, J.; SOH, B. Similarity score for information filtering thresholds. ISCIT 2004. In: IEEE INTERNATIONAL SYMPOSIUM ON COMMUNICATIONS AND INFORMATION TECHNOLOGY, p. 216-221, 2004.

LANZENBERGER, M.; SAMPSON, J.; RESTER, M. Ontology visualization: tools and techniques for visual representation of semi-structured meta-data. **J. UCS**, v. 16, n. 7, p. 1036-1054, 2010.

_____. Visualization in ontology tools. In: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON COMPLEX, INTELLIGENT AND SOFTWARE INTENSIVE SYSTEMS, Fukuoka, Japan, 16-19, March 2009. p. 705-711.

LAZONDER, A. W.; BIEMANS, H. J. A.; WOPERIS, G. J. H. Differences between novice and experienced users in searching information on the World Wide Web. **Journal of the American Society for Information Science**, v. 51, n. 6, p. 576-581, 2000.

LEE, J.; MIN, J-K.; CHUNG, C-W. An effective semantic search technique using ontology. In: PROCEEDINGS OF THE 18th INTERNATIONAL CONFERENCE ON WORLD WIDE WEB (WWW '09). ACM, New York, NY, USA, 2009. p. 1057-1058.

LEI, Y.; UREN, V.; MOTTA, Enrico. SemSearch: a search engine for the semantic web. **3rd European Semantic Web Conference (ESWC)**. Montenegro, 2006.

LIMBU, D. K. et al. Contextual relevance feedback in web information retrieval. In: PROCEEDINGS OF THE 1st INTERNATIONAL CONFERENCE ON INFORMATION INTERACTION IN CONTEXT. ACM, New York, NY, USA, 2006. p. 138-143.

LIN, H. et al. Visualizations for the spyglass ontology-based information analysis and retrieval system. In: PROCEEDINGS OF THE 48th ANNUAL SOUTHEAST REGIONAL CONFERENCE (ACM SE '10). ACM, New York, NY, USA, Article 38, 6 pages, 2010.

LIN, J.; DEMNER-FUSHMAN, D. The role of knowledge in conceptual retrieval: a study in the domain of clinical medicine. In: PROCEEDINGS OF THE 29th ANNUAL INTERNATIONAL ACM SIGIR CONFERENCE ON RESEARCH AND DEVELOPMENT IN INFORMATION RETRIEVAL (SIGIR '06). ACM, New York, NY, USA, 2006. p. 99-106.

LOHBARE, S.; MESHRAM, A. A review on document retrieval from unstructured text. **International Journal of Computer Sciences and Engineering**, v. 2, n. 11, p. 76-80, 2014.

LOPES, I. L. Estratégia de busca na recuperação da informação: revisão da literatura. **Ciência da informação**, Brasília, v. 31, n. 2, p. 60-71, maio/ago. 2002.

LORENCE, D. P.; SPINK, A. Semantics and the medical web: a review of barriers and breakthroughs in effective healthcare query. **Health information and libraries journal**, v. 21, n. 2, p. 109-116, 2004.

LOWGREN, J. Interaction design: brief intro. In: _____. **The encyclopedia of Human-Computer Interaction**. 2. ed. Aarhus, Denmark: The Interaction Design Foundation, 2013. Disponível em: <https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/personas>. Acesso em: 24 nov. 2018.

MÄKELÄ, E. **View-based search interfaces for the semantic web**. Master's thesis, University of Helsinki, 2006.

MANNING, C. D.; RAGHAVAN, P.; SCHÜTZE, H. **Introduction to information**

retrieval. Cambridge University Press, 2008.

MARCHIONINI, G. **Information seeking in electric environments**. New York: Cambridge University Press, 1995.

MARON, M. E.; KUHNS, J. L. On relevance, probabilistic indexing and information retrieval. **Journal of the ACM**, v. 7, n. 3, p. 216-244, 1960.

MAURO, N. Intelligent and personalized community maps. In: PROCEEDINGS OF THE 22nd INTERNATIONAL CONFERENCE ON INTELLIGENT USER INTERFACES COMPANION (IUI '17 Companion). ACM, New York, NY, USA, 2017. p. 181-184.

MAX PLANCK INSTITUT INFORMATIK. **Yago**: a high-quality knowledge base. 2017. Disponível em: <http://www.mpi-inf.mpg.de/departments/databases-and-information-systems /research/yago-naga/yago/>. Acesso em: 10 jun. 2017.

MAZZA, R. **Introduction to information visualization**. New York: Springer Science & Business Media, 2009.

McGUINNESS, D. L. Ontologies come of age. In: FENSEL et al. **Spinning the semantic web**: bringing the World Wide Web to its full potential. Cambridge: MIT Press, 2003. Cap. 6.

MEHRA, B.; BILAL, D. International students' perceptions of their information seeking strategies. In: PROCEEDINGS OF ANNUAL CONFERENCE OF THE CANADIAN ASSOCIATION FOR INFORMATION SCIENCE, 2007.

MEIRELLES, I. **Design for information**: na introduction to the histories, theories, and best practices behind effective information visualizations. Osceola, WI, USA: Rockport Publishers, 2013.

MELO, M. L. A. Contribuições da hermenêutica de Paul Ricoeur à pesquisa fenomenológica em psicologia. **Psicol. USP**, São Paulo, v. 27, n. 2, p. 296-306, 2016.

MENDES NETTO, C.; LIMA, G. A. Visualização de ontologias: estudos e perspectivas. **Inf. & Soc.**, João Pessoa, v. 27, n. 3, p. 59-72, set./dez. 2017.

MENDES NETTO, C.; LIMA, G. N. B. M. O.; PIEROZZI JÚNIOR, I. Modelos de visualização de informação para ontologias de domínio. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB), 17., 2016, Salvador. **Anais...** Salvador, Bahia: Universidade Federal da Bahia, 2016.

MOGGRIDGE, B. **Designing interactions**. Cambridge, MA: MIT Press, 2007.

MONTEIRO, E.; HANSETH, O. social shaping of information infrastructure: on being specific about the technology. In: ORLIKOWSKI, W. J.; WALSHAM, G.; JONES, M. R.; DEGROSS, J. I. (Eds.). **Information technology and changes in organizational work**. London: Chapman and Hall, 1996. p. 325-343.

MOOERS, C. N. Zatocoding applied to mechanical organization of knowledge, **American Documentation**, v. 2, n. 1, p. 20-32, 1951.

MOORE, S. W. **Digital Cavendish Project**. 2013. Disponível em: <http://dayofdh2013.

matrix.msu.edu/smoores/tag/gephi/>. Acesso em: 25 maio 2018.

MORAIS, E. A. M.; AMBRÓSIO, A. P. L. **Ontologias**: conceitos, usos, tipos, metodologias, ferramentas e linguagens. Technical report, Universidade Federal de Goiás, 2007.

MÜLLER, H. M.; KENNY, E. E.; STERNBERG, P. W. Textpresso: an ontology-based information retrieval andextraction system for biological literature. **PLoS Biol**, v. 2, n. 11. 2004.

MUNZNER, T. **Visualization analysis and design**. A K Peters Visualization Series: CRC Press, 2014.

MUTTON, P.; GOLBECK, J. Visualization of semantic metadata and ontologies. In: PROCEEDINGS OF THE 7th INTERNATIONAL CONFERENCE ON INFORMATION VISUALIZATION (IV'03). IEEE Computer Society, London, UK, jul. 2003. p. 300-305.

NASCIMENTO, G. F. C. L. **Folksonomia como estratégia de indexação dos bibliotecários no Del.icio.us**. 2008. 103f. Dissertação (Mestrado em Ciência da Informação). Programa de Pós-Graduação em Ciência da Informação, Universidade Federal da Paraíba, João Pessoa, 2008.

NASCIMENTO, H. A. D.; FERREIRA, C. B. R. Visualização de Informações: uma abordagem prática. In: CONGRESSO DA SOCIEDADE BRASILEIRA DE COMPUTAÇÃO, 25., 2005, São Leopoldo. **Anais...** São Leopoldo: UNISINOS, 2005. p. 1262-1312.

NAVARRO-PRIETO, R., SCAIFE, M.; ROGERS, Y. Cognitive strategies in web searching. In: PROCEEDINGS OF THE 5TH CONFERENCE ON HUMAN FACTORS & THE WEB, 1999. Disponível em: <http://zing.ncsl.nist.gov/hfweb/proceedings/navarro-prieto/index. html>. Acesso em: 5 ago. 2017.

NAYAK, T.; PRASAD, S.; SENAPATI, M. R. A survey on web text information retrieval in text mining. **Research Journal of Applied Sciences, Engineering and Technology**, v. 10, n. 10, p. 1164-1174, 2015.

NICHOLAS, D. Scholarly and professional journals in the digital environment. **Records Management Journal**, v. 20, n. 3, p. 291-300, 2010.

NIEDŹWIEDZKA, B. A proposed general model of information behaviour. **Information Research**, v. 9, n. 1, 2003.

NIELSEN, J. **Why you only need to test with 5 users**. Nielsen Norman Group, March 19, 2000. Disponível em: <http://www.useit.com/alertbox/20000319.html>. Acesso em: 23 mar. 2017.

NIELSEN, L. Personas. In: _____. **The encyclopedia of Human-Computer Interaction**. 2. ed. Aarhus, Denmark: The Interaction Design Foundation, 2013. Disponível em: <https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/personas>. Acesso em: 21 nov. 2018.

NIRENBURG, S.; RASKIN, V. **Ontological semantics**. Cambridge: MIT Press, 2004.

NODUS LABS. **Russia keywords search terms text network graph**. 2013. Disponível em: <https://noduslabs.com/cases/text-network-analysis-seo/attachment/russia-google-keywords-text-network/>. Acesso em: 17 abr. 2018.

NOY, N. F.; McGUINNESS, D. L. **Ontology development 101**: a guide to creating your first ontology. Stanford University, Stanford, 2001.

NUNES, B. **Heidegger & Ser e tempo**. 3. ed. Rio de Janeiro: Zahar, 2002.

OWL. **Web Ontology Language Guide**. 2003. Disponível em: <http://www.w3.org/TR/2003/CR-owl-guide-20030818/>. Acesso em: 10 dez. 2017.

PALMER, R. **Hermenêutica**. Tradução de Maria Luíza Ribeiro Ferreira. Lisboa: Edições 70, 1969.

PAPADAKIS, I. et al. A query construction service for large-scale Web search engines. In: PROCEEDINGS OF THE 2009 IEEE/WIC/ACM INTERNATIONAL JOINT CONFERENCE ON WEB INTELLIGENCE AND INTELLIGENT AGENT TECHNOLOGY - Volume 03 (WI-IAT '09), v. 3. IEEE Computer Society, Washington, DC, USA, 2009. p. 96-99.

PAUL, S. K. et al. An information retrieval (IR) techniques for text mining on web for unstructured data. **International Journal of Advanced Research in Computer Science and Software Engineering**, v. 4, n. 2, p. 67-70, 2014.

PAULHEIM, H. Ontology-based modularization of user interfaces. In: PROCEEDINGS OF THE 1st ACM SIGCHI SYMPOSIUM ON ENGINEERING INTERACTIVE COMPUTING SYSTEMS (EICS '09). ACM, New York, NY, USA, 2009. p. 23-28.

PAULHEIM, H.; MEYER, L. Ontology-based information visualization in integrated UIs. In: PROCEEDINGS OF THE 16th INTERNATIONAL CONFERENCE ON INTELLIGENT USER INTERFACES (IUI '11). ACM, New York, NY, USA, 2011. p. 451-452.

PAULHEIM, H.; PROBST, F. Ontology-enhanced user interfaces: a survey. **International Journal on Semantic Web and Information Systems**, v. 6, n. 2, p. 36-59, 2010.

PAZ-TRILLO, C.; WASSERMANN, R.; BRAGA, P. P. An information retrieval application using ontologies. **Journal of the Brazilian Computer Society**, v. 11, n. 2, p. 17-31, 2005.

PICKLER, M. E. V. Web Semântica: ontologias como ferramentas de representação do conhecimento. **Perspectivas em Ciência da Informação**, v. 12, n. 1, p. 65-83, abr. 2007.

PITTARELLO, F.; GATTO, I. A visual interface for querying ontologically and socially annotated 3D worlds for the web. In: PROCEEDINGS OF THE INTERNATIONAL WORKING CONFERENCE ON ADVANCED VISUAL INTERFACES (AVI '12), Genny Tortora, Stefano Levialdi, and Maurizio Tucci (Eds.). ACM, New York, NY, USA, 2012a. p. 377-381.

_____. An integrated ecosystem of interfaces for annotating, querying and browsing networks of Web3D worlds. In: PROCEEDINGS OF THE INTERNATIONAL WORKING CONFERENCE ON ADVANCED VISUAL INTERFACES (AVI '12), Genny Tortora, Stefano Levialdi, and Maurizio Tucci (Eds.). ACM, New York, NY, USA, 2012b. p. 794-795.

PIZZINATO, P. M.; VIEIRA, R.; RIGO, S. J. Um visualizador online de ontologias utilizando árvores hiperbólicas. In: SEMINÁRIO DE PESQUISA EM ONTOLOGIA NO BRASIL (ONTOBRAS, 2010). Florianópolis, Santa Catarina, 2010. p. 1-10.

PLAISANT, C.; GROSJEAN, J.; BEDERSON, B. SpaceTree: supporting exploration in large node link tree, design evolution and empirical evaluation. In: PROCEEDINGS OF THE IEEE SYMPOSIUM ON INFORMATION VISUALIZATION (INFOVIS'02). IEEE Computer Society, Boston, 2002. p. 57-64.

POPESCU, A.; MILLET, C.; MOËLLIC, P-A. Ontology driven content based image retrieval. In: PROCEEDINGS OF THE 6th ACM INTERNATIONAL CONFERENCE ON IMAGE AND VIDEO RETRIEVAL (CIVR '07). ACM, New York, NY, USA, 2007. p. 387-394.

REED, S. L.; LENAT, B. D. Mapping ontologies into Cyc. In: PROC. AAAI CONFERENCE 2002 WORKSHOP ON ONTOLOGIES FOR THE SEMANTIC WEB, Edmonton, Canadá, 2002.

RICOUER, P. **O conflito das interpretações**: ensaios de hermenêutica. Tradução de M. F. Sá Correia. Porto: Rés, 1988. (Trabalho original publicado em 1969).

RIJSBERGEN, C. J. V. **Information retrieval**. Glasgow, Scotland, UK: University of Glasgow, 1979.

RIVADENEIRA, W.; BEDERSON, B. B. **A study of search result clustering interfaces**: comparing textual and zoomable interfaces, University of Maryland HCIL Tech. Rep. HCIL-2003-36, October, 2003.

ROGERS, Y.; SHARP, H.; PREECE, J. **Design de interação**: além da interação humano-computador. Tradução de Isabela Gasparini. 3. ed. Porto Alegre: Bookman, 2013.

ROSSON, M. B.; CARROLL, J. M. Scenario-based design. In: JACKO, J.; SEARS, A. (Eds.). The **human-computer interaction handbook**: fundamentals, evolving technologies and emerging applications. Lawrence Erlbaum Associates, 2002. cap. 53, p. 1032-1050.

ROVIRA, C.; CODINA, L. La orientación a objetos en el diseño de sedes web: hipertextos y representación de la información. **Revista española de documentación científica**, v. 26, n. 3, p. 267-290, 2003.

ROWLEY, J. **A biblioteca eletrônica**. Brasília: Briquet de Lemos, 2002.

ROZELL, E. et al. S2S architecture and faceted browsing applications. In: PROCEEDINGS OF THE 21st INTERNATIONAL CONFERENCE ON WORLD WIDE WEB (WWW '12 COMPANION). ACM, New York, NY, USA, 2012. p. 413-416.

RUIZ, F.; HILERA, J. R. Using ontologies in software engineering and technology. In: CALERO, C.; RUIZ, F.; PIATTINI, M. (Eds.). **Ontologies for software engineering and software technology**, p. 49-102. Springer, 2006.

SAFFER, D. **Designning for interaction**: creating innovative applications and devices. 2. ed. Berkeley: New Ridgers, 2010.

SAGHAFI, A. Visualizing ontologies: a literature survey. In: HAEMMERLÉ, O.; STAPLETON, G.; FARON ZUCKER, C. (Eds.). Graph based representation and reasoning. ICCS 2016. **Lecture Notes in Computer Science**, v. 9717. Springer, Cham, 2016.

SALTON, G. **Automatic information organization and retrieval**. New York: McGraw-Hill, 1968.

_____. Automatic information retrieval. **IEEE Computer**, v. 13, p. 41-56, 1980.

_____. **Automatic text processing**: the transformation, analysis, and retrieval of information by computer. Boston: Addison-Wesley Longman Publishing Co., 1989.

SALTON, G. Some characteristics of future information systems. In: SIGIR FORUM 18, 2-4 (September 1985). ACM, New York, NY, USA, 1985. p. 28-39.

SALTON, G.; BUCKLEY, C. Term-weighting approaches in automatic text retrieval. **Information Processing & Management**, Oxford, v. 24, n. 5, p. 513-523, 1988.

SARACEVIC, T. Ciência da informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun. 1996.

_____. Information science. In: BATES, M. J.; MAACK, M. N. (Eds.). **Encyclopedia of Library and Information Science**. New York: Taylor & Francis, 2009. p. 2570-2586.

_____. Interdisciplinary nature of information science. **Ciência da Informação**, v. 24, n. 1, p. 36-41, 1995.

_____. The stratified model of information retrieval interactions: extension and applications. **American Society for Information Science**, p. 313-327, 1997.

SAVOLAINEN, R. Approaches to socio-cultural barriers to information seeking. **Library and Information Science Research**, v. 38, n. 1, p. 52-59, 2016.

_____. Everyday life information seeking: approaching information seeking in the context of 'way of life'. **Library & Information Science Research**, v. 17, p. 259-294, 1995.

SCHLEIERMACHER, F. D. E. **Hermenêutica e crítica**. Ijuí: Unijuí, 2005.

_____. **Hermenêutica:** a arte e técnica da interpretação. Bragança Paulista: EDUSF, 2003.

SCHMIDT, L. K. **Hermenêutica**. Tradução de Fábio Ribeiro. 3. ed. Petrópolis: Vozes, 2014.

SCHWANDT, T. A. Três posturas epistemológicas para a investigação qualitativa. In: DENZIN, N. K.; LINCOLN, Y. S. (Orgs.). **O planejamento da pesquisa qualitativa**: teorias e abordagens. Porto Alegre: Artmed, 2006. p. 193-217.

SENDHILKUMAR, S.; GEETHA, T. V. Personalized ontology for web search personalization. In: PROCEEDINGS OF THE 1ST BANGALORE ANNUAL COMPUTE CONFERENCE (COMPUTE '08). ACM, New York, NY, USA, Article 18, 7 pages, 2008.

SHIN, S. et al. Augmenting mobile search engine with semantic web and context awareness. In: PROCEEDINGS OF THE 2013 IEEE/WIC/ACM INTERNATIONAL JOINT

CONFERENCES ON WEB INTELLIGENCE (WI) AND INTELLIGENT AGENT TECHNOLOGIES (IAT) - Volume 03 (WI-IAT '13), v. 3. IEEE Computer Society, Washington, DC, USA, 2013a. p. 278-279.

_____. Semantic search for smart mobile devices. In: PROCEEDINGS OF THE COMPANION PUBLICATION OF THE 2013 INTERNATIONAL CONFERENCE ON INTELLIGENT USER INTERFACES COMPANION (IUI '13 Companion). ACM, New York, NY, USA, 2013b. p. 95-96.

SHNEIDERMAN, B. **Designing the user interface**: strategies for effective human-computer interaction. 3. ed. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1997.

_____. The eyes have it: a task by data type taxonomy for information visualizations. In: PROCEEDINGS OF SYMPOSIUM ON VISUAL LANGUAGENS. Boulder, CO: IEEE, 1996. p. 336-343.

_____. Tree visualization with tree-maps: 2-d space-filling approach. **ACM Trans. Graph**, v. 11, n. 1, p. 92-99, jan. 1992.

SHNEIDERMAN, B. et al. **Designing the user interface**: strategies for effective human-computer interaction. 6. ed. Hoboken: Pearson, 2017.

SIAHBANI, M. et al. Knowledge base population and visualization using an ontology based on semantic roles. In: PROCEEDINGS OF THE 2013 WORKSHOP ON AUTOMATED KNOWLEDGE BASE CONSTRUCTION (AKBC '13). ACM, New York, NY, USA, 2013. p. 85-90.

SIGIR. The 40th International ACM SIGIR conference on research and development in Information Retrieval (SIGIR '17) Shinjuku, Tokyo, Japan, August 07-11, 2017. Disponível em: <http://sigir.org/sigir2017/>. Acesso em: 16 maio 2018.

SILVA, I. C. S.; FREITAS, C. M. D. S. Avaliação de ferramentas de busca na web baseadas em visualização de informações. In: PROCEEDINGS OF VII BRAZILIAN SYMPOSIUM ON HUMAN FACTORS IN COMPUTING SYSTEMS (IHC '06). ACM, New York, NY, USA, p. 178-185, 2006.

SILVA, I. C. S.; FREITAS, C. M. D. S.; SANTUCCI, G. An integrated approach for evaluating the visualization of intensional and extensional levels of ontologies. In: PROCEEDINGS OF THE 2012 BELIV WORKSHOP: BEYOND TIME AND ERRORS - NOVEL EVALUATION METHODS FOR VISUALIZATION (BELIV '12). ACM, New York, NY, USA, Article 2, 7 pages, 2012.

SMITH, B. et al. Relations in biomedical ontologies. **Genome Biology**, v. 6, n. R46, 2005.

SNYDER, C. **Paper prototyping**: the fast and easy way to design and refine user interfaces. San Francisco, CA: Morgan Kaufmann, 2003.

SOUZA, K. X. S.; SANTOS, A. D.; EVANGELISTA, S. R. M. Visualization of ontologies through hypertrees. In: PROCEEDINGS OF THE LATIN AMERICAN CONFERENCE ON HUMAN-COMPUTER INTERACTION (CLIHC '03). ACM, New York, NY, USA, 2003. p. 251-255.

SOUZA, M. C. P. **A rede semântica**. Humanidades Digitais. Grupo de Pesquisa da Universidade de São Paulo (USP). 2012. Disponível em: <https://humanidadesdigitais.org/2012/10/04/a-rede-semantica-i/>. Acesso em: 23 maio 2018.

SOUZA, R. R. Sistemas de recuperação de informações e mecanismos de busca na web: panorama atual e tendências. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 11, n. 2, p. 161-173, maio/ago. 2006.

SOYLU, A. et al. OptiqueVQS: a visual query system over ontologies for industry, **Semantic Web**, IOS Press, p. 1-33, 2016.

_____. OptiqueVQS: towards an ontology-based visual query system for big data. In: PROCEEDINGS OF THE FIFTH INTERNATIONAL CONFERENCE ON MANAGEMENT OF EMERGENT DIGITAL ECOSYSTEMS (MEDES '13). ACM, New York, NY, USA, 2013. p. 119-126.

_____. Why not simply Google? In: PROCEEDINGS OF THE 8th NORDIC CONFERENCE ON HUMAN-COMPUTER INTERACTION: FUN, FAST, FOUNDATIONAL (NordiCHI '14). ACM, New York, NY, USA, 2014. p. 1039-1042.

SPINK, A. et al. Searching the web: the public and their queries. **Journal of the American Society for Information Science**, v. 52, n. 3, p. 226-234, 2001.

SPINK, A.; COLE, C. Human information behaviour: integrating diverse approaches and information use. **Journal of the American Society for Information Science and Technology**, v. 57, n. 1, p. 25-35, 2006.

STAGLIANO, N. **Hermenêutica**: conceitos e características. 2016. Disponível em: <https://simoesstagliano.jusbrasil.com.br/artigos/335787147/hermeneutica-conceitos-e-caracteristicas>. Acesso em: 10 jul. 2018.

STOJANOVIC, N. On the role of a user's knowledge gap in an information retrieval process. In: PROCEEDINGS OF THE THIRD INTERNATIONAL CONFERENCE ON KNOWLEDGE CAPTURE.ACM K-CAP05, 2005. p. 83-90.

STOREY, M-A. D. Navigating documents using ontologies, taxonomies and folksonomies. In: PROCEEDINGS OF THE 2007 ACM SYMPOSIUM ON DOCUMENT ENGINEERING (DocEng '07). ACM, New York, NY, USA, 2007. p. 2-2.

STYPEREK, A.; CIESIELCZYK, M.; SZWABE, A. Semantic search engine with an intuitive user interface. In: PROCEEDINGS OF THE 23rd INTERNATIONAL CONFERENCE ON WORLD WIDE WEB (WWW '14 Companion). ACM, New York, NY, USA, 2014. p. 383-384.

SUH, B.; BEDERSON, B. B. OZONE: a zoomable interface for navigating ontology information. In: PROCEEDINGS OF THE WORKING CONFERENCE ON ADVANCED VISUAL INTERFACES (AVI '02), Maria De Marsico, Stefano Levialdi, and Emanuele Panizzi (Eds.). ACM, New York, NY, USA, 2002. p. 139-143.

SWARTOUT, B. et al. Toward distributed use of large-scale ontologies. In: PROCEEDINGS OF 10TH KNOWLEDGE ACQUISITION FOR KNOWLEDGE-BASED SYSTEMS WORSKHOP. Banff, Canadá, 1996. p. 138-148.

TANG, M. C. Browsing and searching in a faceted information space: a naturalistic study of pubmed users' interaction with a display tool. **Journal of the American Society for Information Science**, v. 58, n. 13, p. 1998-2006, 2007.

THAI, V.; HANDSCHUH, S. IVEA: toward a personalized visual interface for exploring text collections. In: PROCEEDINGS OF THE 14th INTERNATIONAL CONFERENCE ON INTELLIGENT USER INTERFACES (IUI '09). ACM, New York, NY, USA, 2009. p. 479-480.

THOMASSON, A. Categories. **The Stanford Encyclopedia of Philosophy** (Spring 2018 Edition), Edward N. Zalta (ed.). 2018. Disponível em: <https://plato.stanford.edu/archives/spr2018/entries/categories/>. Acesso em: 11 abr. 2018.

TIJERINO, Y. A.; SANATI, R. Onto TEMAS: an ontology-based teaching materials search engine. **J. Comput. Sci**, v. 20, n. 4, p. 177-182, abr. 2005.

TURNER, P. M.; KASKE, N. K.; BAKER, G. S. The effects of baud rate, performance anxiety, and experience on online bibliographic searches. **Information Technology and Libraries**, v. 9, p. 34-42, 1990.

TZITZIKAS, Y.; HAINAUT, J-C. On the visualization of large-sized ontologies. In: PROCEEDINGS OF THE WORKING CONFERENCE ON ADVANCED VISUAL INTERFACES (AVI '06). ACM, New York, NY, USA, 2006. p. 99-102.

UNIOESTE – UNIVERSIDADE ESTADUAL DO OESTE DO PARANÁ. **Anual de procedimentos para inovação no âmbito da Universidade Estadual do Oeste do Paraná – Unioeste**. Núcleo de Inovações Tecnológicas. Cascavel, 2012. Disponível em: <http://projetos.unioeste.br/campi/nit/images/cartilha_nit/cartilha.pdf>. Acesso em: 04 maio 2018.

UPTOWN/DOWNTOWN MEDIA LLC. All Over Albany. **A different view of the Clinton-Trump margin**. Editado por Mary Darcy and Greg Dahlmann. 2016. Disponível em: <http://alloveralbany.com/archive/2016/11/09/new-york-state-presidential-election-2016-county-r>. Acesso em: 02 maio 2018.

USCHOLD, M.; GRUNINGER, M. Ontologies: principles, methods and applications. **Knowledge Engineering Review**, v. 11, n. 2, p. 1-69, 1996.

VAN HEIJIST, G.; SCHREIBER, A. T.; WIELINGA, B. J. Using explicit ontologies in KBS development. **International Journal of Human - Computer Studies**, v. 46, n. 2-3, p. 183-192, fev./mar. 1997.

VELART, Z.; ŠALOUN, P. Ontology based course navigation. In: PROCEEDINGS OF THE EIGHTEENTH CONFERENCE ON HYPERTEXT AND HYPERMEDIA (HT '07). ACM, New York, NY, USA, 2007. p. 151-152.

VIEIRA, J. M. L.; CORREA, R. F. Visualização da informação na construção de interfaces amigáveis para sistemas de recuperação de informação. **Encontros Bibli**: Revista Eletrônica de Biblioteconomia e Ciência da Informação, v. 16, n. 32, p. 73-93, 2011.

VITAL, L. P.; CAFÉ, L. M. A. Ontologias e taxonomias: diferenças. **Perspectivas em Ciência da Informação**, v. 16, n. 2, p. 115-130, 2011.

VOILA! International Workshop on Visualization and Interaction for Ontologies and Linked Data. Co-located with ISWC 2016, Kobe, Japan, October 17, 2016. Disponível em: <http://voila2016.visualdataweb.org/>. Acesso em: 17 fev. 2017.

_____. Co-located with ISWC 2015, Bethlehem, Pennsylvania, USA, October 11, 2015. Disponível em: <http://voila2015.visualdataweb.org/>. Acesso em: 17 fev. 2017.

_____. (VOILA '17), Vienna, Austria. CEUR Workshop Proceedings, v. 1947, CEUR-WS.org, 2017. Disponível em: <http://voila2017.visualdataweb.org/>. Acesso em: 16 abr. 2018.

WALSHAM, G. **Interpreting information systems in organization**. Chichester, UK: Wiley, 1993.

WANG, H; LIU, S.; CHIA, L-T. Does ontology help in image retrieval?: a comparison between keyword, text ontology and multi-modality ontology approaches. In: PROCEEDINGS OF THE 14th ACM INTERNATIONAL CONFERENCE ON MULTIMEDIA (MM '06). ACM, New York, NY, USA, 2006. p. 109-112.

WARE, C. Design as applied perception. In: CORROL, J. M. (Ed.). **HCI models, theories, and frameworks**: toward a multidisciplinary science. San Francisco, CA: Morgan Kaufmann, 2003. p. 11-26.

WASLAWICK, R. S. **Análise e projeto de sistemas de informação orientados a objetos**. 2. ed. Rio de Janeiro: Elsevier, 2011.

WILSON, T. D. Human information behavior. **Informing Science**, Sheffield, v. 3, n. 2, p. 49-53, 2000.

_____. Information behaviour: an interdisciplinary perspective. **Information Processing & Management**, v. 33, n. 4, p. 551-572, 1996.

_____. Models in information behaviour research. **Journal of Documentation**, v. 55, n. 3, p. 249-270, 1999.

_____. On user studies and information needs. **Journal of Documentation**, London, v. 62, n. 6, p. 658-670, dez. 2006.

_____._____. **Journal of Librarianship**, v. 37, n. 1, p. 3-15, 1981.

WINOGRAD, T. From computing machinery to interaction design. In: DENNING, P; METCALFE, R. (Eds.). **Beyond calculation**: the next fifty years of computing. Springer-Verlag, 1997. p. 149-162.

WONG, W.; LIU, W.; BENNAMOUN, M. An ontology-based interface for improving information exploration. In: PROCEEDINGS OF THE FIRST INTERNATIONAL WORKSHOP ON INTELLIGENT VISUAL INTERFACES FOR TEXT ANALYSIS (IVITA '10), Shixia Liu, Michelle X. Zhou, Giuseppe Carenini, and Huamin Qu (Eds.). ACM, New York, NY, USA, 2010. p. 29-32.

WOODS, W. A. **Searching versus finding**: why systems need knowledge to find what you really want. Sun Microsystems Laboratories, 2004. Disponível em:

<http://labs.oracle.com/projects/knowledge/WoodsACMQueueApril2004.pdf>. Acesso em: 08 set. 2017.

WU, P. T. et al. Keyword-based concept search on consumer photos by web-based kernel function. In: PROCEEDINGS OF THE 16th ACM INTERNATIONAL CONFERENCE ON MULTIMEDIA (MM '08). ACM, New York, NY, USA, 2008. p. 651-654.

XIE, I.; JOO, S. Transitions in search tactics during theweb-based. **Journal of the American Society for Information Science and Technology**, v. 61, n. 11, p. 2188-2205, 2010.

YI, Y. J.; STVILIA, B.; MON, L. Cultural influences on seeking quality health information: an exploratory study of the Korean community. **Library & Information Science Research**, v. 34, n. 1, p. 45-51, 2012.

YU, L. How poor informationally are information poor? Evidence from an empirical study of daily and regular information practices of individuals. **Journal of Documentation**, v. 66, n. 6, p. 906-933, 2010.

ZHENG, S. et al. Enabling ontology based semantic queries in biomedical database systems. In: PROCEEDINGS OF THE 21st ACM INTERNATIONAL CONFERENCE ON INFORMATION AND KNOWLEDGE MANAGEMENT (CIKM '12). ACM, New York, NY, USA, 2012. p. 2651-2654.

ZHUHADAR, L.; NASRAOUI, O. Evaluating usability and precision of visual search engine. In: PROCEEDINGS OF THE 2010 SPRING SIMULATION MULTICONFERENCE (SPRINGSIM '10). Society for Computer Simulation International, San Diego, CA, USA, Article 239, 4 pages, 2010.

APÊNDICES

APÊNDICE A - Roteiro de Tarefas para Avaliação de Usabilidade

ROTEIRO DE TAREFAS PARA AVALIAÇÃO DE USABILIDADE

CENÁRIO 1 - BUSCA BASEADA EM PALAVRAS-CHAVE

Os testes realizados neste cenário têm como objetivo analisar a interação do usuário no desenvolvimento de buscas realizadas a partir de termos ou palavras-chave digitadas na caixa de busca. Buscam ainda identificar as dificuldades do usuário na elaboração de buscas em domínios desconhecidos por ele.

TAREFA 1

Protótipo(s) Interface inicial da ferramenta de busca (Figura 67)

Objetivo(s) Entender se o usuário consegue perceber os passos para a realização de uma busca baseada em palavras-chave

Tarefa(s) Você precisa realizar uma pesquisa baseada em palavra-chave, assim como faz no Google. Seu objetivo é explorar o termo Indexação. Como você realizaria esta tarefa? Quais termos você usaria para atingir este objetivo?

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 2

Protótipo(s) Interface de busca baseada em palavra-chave (Figura 36)

Objetivo(s) Entender se o usuário consegue realizar uma busca baseada em palavra-chave e perceber se ele compreende as opções disponíveis para a busca

Tarefa(s) Elabore uma expressão de busca e selecione a opção que condiz com o seu objetivo. Se preferir, você pode digitar apenas as quatro primeiras letras do termo que você quer pesquisar.

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 3

Protótipo(s) Interface para a elaboração de expressões de busca (Figura 37)

Objetivo(s) Entender se o usuário consegue compreender a relação entre a busca realizada e os termos apresentados no mapa de conceitos, bem como se ele percebe as relações existentes entre os conceitos do domínio

Tarefa(s) Neste ponto, quais opções você tem para a busca? Como você compreende o mapa de conceitos?

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 4

Protótipo(s) Interface para a elaboração de expressões de busca (Figura 37)

Objetivo(s) Entender como o usuário interage com as perspectivas de análise e compreender quais são as opções disponíveis para a seleção

Tarefa(s) Analise as perspectivas de busca relacionadas ao termo Indexação. Quais contextos estão disponíveis para o refinamento da sua busca? Como você compreende as perspectivas de busca?

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 5

Protótipo(s) Interface para a elaboração de expressões de busca (Figura 37)
Interface para a seleção de perspectivas de análise de busca (Figura 38)

Objetivo(s) Entender se o usuário consegue selecionar corretamente uma perspectiva de busca

Tarefa(s) Para refinar a busca, você deve escolher a perspectiva de busca Problema

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 6

Protótipo(s) Interface para a elaboração de expressões de busca com dois termos (Figura 39)

Objetivo(s) Entender se o usuário consegue reproduzir os passos da busca anterior (Ontologia) e acrescentar um novo termo à busca que está sendo realizada

Tarefa(s) Você deve acrescentar um novo termo à sua busca (Recuperação de Informação)

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 7

Protótipo(s) Interface para a elaboração de expressões de busca com dois termos (Figura 39)

Objetivo(s) Entender como o usuário percebe as perspectivas de busca disponíveis relacionadas ao novo termo

Tarefa(s) Analise as perspectivas de busca relacionadas ao termo Recuperação de Informação e selecione a perspectiva Área

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 8

Protótipo(s) Interface para a elaboração de expressões de busca com dois termos (Figura 39)

Objetivo(s) Entender a percepção do usuário sobre como solicitar a busca a partir dos termos e das perspectivas selecionadas

Tarefa(s) Realize uma pesquisa contendo dois termos de busca (Ontologia e Recuperação de Informação)

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 9

Protótipo(s) Interface de perspectivas de análise, termos e resultados de busca (Figura 40)

Objetivo(s) Entender como o usuário percebe os resultados da busca realizada

Tarefa(s) Analise os resultados da busca. O que a interface representa para você? Qual arquivo corresponde à busca realizada? Quais arquivos são mais e menos relevantes?

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 10

Protótipo(s) Interface de perspectivas de análise, termos e resultados de busca (Figura 40)

Objetivo(s) Entender a interação do usuário com a interface para obter maiores informações sobre um determinado resultado

Tarefa(s) Você quer obter mais informações sobre o Doc.4 (autores, título, ano, local de publicação). Como você faria isto?

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 11

Protótipo(s) Interface de elementos de similaridade entre resultados e expressão de busca (Figura 41)

Objetivo(s) Entender a percepção do usuário sobre a similaridade existente entre um resultado de busca e a expressão de busca

Tarefa(s) Observe os resultados da busca. Quais são os termos que indexam o Doc.4? Em quais perspectivas eles aparecem?

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 12

Protótipo(s) Interface de similaridade entre resultados de busca (Figura 42)

Objetivo(s) Entender a percepção do usuário sobre a similaridade existente entre dois resultados da busca realizada

Tarefa(s) Observe os resultados da busca. Qual é a relação existente entre os Docs. 3 e 4? Qual é a similaridade existente entre eles

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 13

Protótipo(s) Interface de notas explicativas e áreas de conhecimento (Figura 43)

Objetivo(s) Entender qual é a percepção do usuário em relação ao plano de contexto e as possibilidades de expansão de busca que ele oferece

Tarefa(s) Você realizou uma nova busca utilizando o termo Indexação. Como você interpreta os elementos apresentados no mapa de conceitos? Quais são as suas possibilidades de busca?

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 14

Protótipo(s) Interface de notas explicativas e áreas de conhecimento (Figura 43)

Objetivo(s) Entender qual é a percepção do usuário em relação às funcionalidades que visam apresentar conhecimento adicional referente ao termo de busca utilizado (Notas Explicativas)

Tarefa(s) Você quer maiores informações sobre o termo selecionado (Indexação). Como acessar as notas explicativas? Qual é a definição do termo?

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 15

Protótipo(s) Interface de notas explicativas e áreas de conhecimento (Figura 43)

Objetivo(s) Entender qual é a percepção do usuário em relação às funcionalidades que visam apresentar conhecimento adicional referente ao termo de busca utilizado (Áreas de Conhecimento)

Tarefa(s) O termo Indexação é adotado em alguma outra área de conhecimento e possui algum outro significado?

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 16

Protótipo(s) Interface de sinônimos e termos relacionados (Figura 44)

Objetivo(s) Entender qual é a percepção do usuário em relação às funcionalidades que visam apresentar conhecimento adicional referente ao termo de busca utilizado (Sinônimos e Termos Relacionados)

Tarefa(s) Você quer explorar o termo Indexação para obter Sinônimos e Termos Relacionados para novas possibilidades de busca. Como você interpreta os elementos apresentados no mapa de conceitos? Quais são as suas possibilidades de busca?

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 17

Protótipo(s) Interface de variações linguísticas (Figura 46)

Objetivo(s) Entender qual é a percepção do usuário em relação às funcionalidades que visam apresentar conhecimento adicional referente ao termo de busca utilizado (Variações Linguísticas)

Tarefa(s) O que essa interface significa para você? Quais são as informações apresentadas? Como elas podem contribuir com a sua busca?

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 18

Protótipo(s) Interface de antônimos e termos conceitualmente contraditórios (Figura 47)

Objetivo(s) Entender a percepção do usuário em relação às funcionalidades que visam apresentar conhecimento adicional referente ao termo de busca utilizado (Antônimos e Termos Contraditórios)

Tarefa(s) Você quer explorar o termo Leucemia. Como você interpreta os elementos apresentados no mapa de conceitos? Como as opções apresentadas podem te auxiliar durante a busca?

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 19

Protótipo(s) Interface inicial da ferramenta de busca (Figura 67)

Objetivo(s) Entender as dificuldades do usuário para realizar uma busca baseada em palavras-chave, em um domínio totalmente desconhecido por ele

Tarefa(s) Elabore uma expressão de busca na área de Endocrinologia. Quais são suas principais dificuldades para isto?

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

CENÁRIO 2 – BUSCA BASEADA EM ESTRUTURA TERMINOLÓGICA

Os testes realizados neste cenário têm como objetivo analisar a interação do usuário no desenvolvimento de buscas realizadas a partir de uma estrutura terminológica, neste caso, representado por uma ontologia de domínio (mapa de conceitos). Buscam ainda identificar se este tipo de recurso é útil na aquisição de conhecimentos sobre o domínio explorado e sua estrutura e se ele realmente pode contribuir com a elaboração de buscas em contextos desconhecidos.

TAREFA 20

Protótipo(s) Interface inicial da ferramenta de busca (Figura 67)

Objetivo(s) Entender se o usuário consegue reproduzir os passos realizados na busca baseada em palavras-chave e iniciar uma busca baseada em estrutura terminológica

Tarefa(s) Realize uma busca baseada em estrutura terminológica. Por onde você começaria esta tarefa?

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 21

Protótipo(s) Interface de busca baseada em estrutura terminológica (Figura 49)
Interface de expansão da estrutura terminológica (Figura 50)

Objetivo(s) Entender se o usuário consegue interagir com o mapa de conceitos para realizar tarefas de navegação e seleção de conceitos

Tarefa(s) Navegue no mapa de conceitos e explore os conceitos relacionados à área de Estatística

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 22

Protótipo(s) Interface de expansão da estrutura terminológica (Figura 50)
Interface para a elaboração de expressões de busca (Figura 51)

Objetivo(s) Entender se o usuário consegue elaborar uma busca baseada em estrutura terminológica: selecionar os termos no mapa de conceitos e arrastá-los para a caixa de busca

Tarefa(s) Elabore uma expressão de busca a partir do mapa de conceitos. Seu objetivo é elaborar uma pesquisa relacionada à Amostragem Sistemática. Qual termo você selecionaria? Como você indicaria sua necessidade para a ferramenta de busca?

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 23

Protótipo(s) Interface para a elaboração de expressões de busca (Figura 51)

Objetivo(s) Entender se o usuário consegue perceber que o contexto de um termo também é inserido na busca quando ele realiza uma pesquisa a partir de uma ontologia

Tarefa(s) Você elaborou uma expressão de busca utilizando o termo Sistemática. Como você sabe que os resultados serão relacionados à amostragem sistemática e não à uma pessoa sistemática (metódica)?

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 24

Protótipo(s) Interface para a elaboração de expressões de busca (Figura 51)

Objetivo(s) Entender se o usuário percebe que pode utilizar a caixa de busca para encontrar outros termos, caso os objetivos que ele possui sejam alterados

Tarefa(s) Neste momento, seus objetivos de busca mudam e você tem interesse em pesquisar um novo assunto. Pode onde você começaria?

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 25

Protótipo(s) Interface de estrutura terminológica de um domínio especializado (Figura 52)

Objetivo(s) Entender se o mapa de conceitos é um recurso útil para minimizar as dificuldades do usuário na busca e exploração de domínios desconhecidos

Tarefa(s) Elabore uma expressão de busca na área de Endocrinologia. Como o mapa de conceitos pode ser útil para o desenvolvimento desta tarefa?

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

CENÁRIO 3 – BUSCA BASEADA EM PERSPECTIVAS DE ANÁLISE DE CONTEXTO

Os testes realizados neste cenário têm como objetivo analisar a interação do usuário no desenvolvimento de buscas realizadas a partir de critérios definidos na ontologia de domínio, relacionados à um contexto específico, com o intuito de identificar resultados contendo uma ou várias características correlacionadas. Buscam ainda identificar se o usuário interpreta corretamente as perspectivas que fazem parte de um contexto científico.

TAREFA 26

Protótipo(s) Interface de busca baseada em perspectivas de análise de contexto (Figura 53)

Objetivo(s) Entender se o usuário consegue perceber os passos para explorar o domínio a partir das perspectivas de análise

Tarefa(s) Realize uma busca baseada nas perspectivas Autor e Área

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 27

Protótipo(s) Interface de busca baseada em perspectivas de análise de contexto (Figura 53)

Objetivo(s) Entender a percepção do usuário sobre as informações apresentadas no mapa de conceitos, bem como sobre as possibilidades de busca que são apresentadas

Tarefa(s) Analise as informações apresentadas no mapa de conceitos. Quais são os autores disponíveis para a busca?

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 28

Protótipo(s) Interface de busca baseada em perspectivas de análise de contexto (Figura 53)
Interface de busca baseada em duas perspectivas de análise de contexto (Figura 54)

Objetivo(s) Entender a percepção do usuário sobre a seleção dos termos nas diversas perspectivas de análise, bem como sobre a elaboração de uma busca a partir das informações apresentadas no mapa de conceitos

Tarefa(s) Refine a busca realizada e identifique as publicações do autor Rodrigues, P. Quais são as áreas de publicação do autor?

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 29

Protótipo(s) Interface de busca baseada em duas perspectivas de análise de contexto (Figura 54)
Interface de histórico do autor (Figura 55)

Objetivo(s) Entender a percepção do usuário sobre o histórico do autor

Tarefa(s) Explore o histórico do autor Rodrigues, P. Quais são informações são relevantes?

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 30

Protótipo(s) Interface de busca baseada em duas perspectivas de análise de contexto (Figura 54)

Objetivo(s) Entender a percepção do usuário sobre o refinamento das buscas realizadas

Tarefa(s) Refine a busca novamente e selecione a área de Recuperação de Informação

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 31

Protótipo(s) Interface de resultados de busca baseada em perspectivas de análise de contexto (Figura 56)

Objetivo(s) Entender a percepção do usuário sobre os resultados da busca

Tarefa(s) Analise os resultados da busca realizada. Quais publicações atentem os dois critérios definidos na busca (Autor: Rodrigues, P. e Área: Recuperação de Informação)?

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 32

Protótipo(s) Interface de resultados de busca baseada em perspectivas de análise de contexto (Figura 56)

Objetivo(s) Entender a percepção do usuário sobre a relevância dos resultados (resultados mais relevantes e menos relevantes), bem como sobre os termos que foram utilizados para representar os documentos recuperados

Tarefa(s) Ainda considerando os resultados da busca, quais publicações atendem parcialmente aos critérios de busca estabelecidos?

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 33

Protótipo(s) Interface de resultados de busca baseada em perspectivas de análise de contexto (Figura 56)

Objetivo(s) Entender a percepção do usuário sobre a similaridade existente entre os resultados de uma busca

Tarefa(s) Ainda considerando os resultados da busca, qual é a relação entre os Docs. 1 e 2?

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

CENÁRIO 4 – BUSCA BASEADA EM PERSPECTIVAS DE ANÁLISE DE CONTEÚDO

Os testes realizados neste cenário têm como objetivo analisar a interação do usuário no desenvolvimento de buscas realizadas a partir de um documento específico, com o propósito de identificar resultados relacionados pertencentes ao mesmo contexto. Buscam ainda verificar se o usuário interpreta corretamente as perspectivas que fazem parte de um documento específico.

TAREFA 34

Protótipo(s) Interface de seleção de documento para análise de conteúdo (Figura 57)

Objetivo(s) Entender se o usuário consegue iniciar uma busca a partir de um documento recuperado e explorar as suas diversas perspectivas de análise

Tarefa(s) Realize uma busca baseada no Doc.2, com o propósito de explorar a rede de relacionamentos da qual ele faz parte. Por onde você começaria esta tarefa?

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 35

Protótipo(s) Interface de documento selecionado para análise de conteúdo (Figura 58)

Objetivo(s) Entender a percepção do usuário sobre as possibilidades de busca apresentadas na interface

Tarefa(s) Quais são as opções de busca que você tem à disposição?

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 36

Protótipo(s) Interface de documento selecionado para análise de conteúdo (Figura 58)
Interface de comentários interpretativos (Figura 59)

Objetivo(s) Entender qual é a percepção do usuário em relação às funcionalidades que visam apresentar informações adicionais para a interpretação do documento (Comentários Interpretativos)

Tarefa(s) Acesse os comentários de outros usuários sobre o documento *Comportamento de busca em sistemas de recuperação de informação: um estudo com professores universitários*.

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 37

Protótipo(s) Interface de documento selecionado para análise de conteúdo (Figura 58)
Interface de erros e falhas (Figura 60)

Objetivo(s) Entender qual é a percepção do usuário em relação às funcionalidades que visam apresentar informações adicionais para a interpretação do documento (Erros e Falhas)

Tarefa(s) Acesse o recurso de Erros e Falhas para analisar problemas identificados por outros usuários no documento que você está explorando. Quantos erros foram indicados neste documento?

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 38

Protótipo(s) Interface de documento selecionado para análise de conteúdo (Figura 58)
Interface de parágrafos avaliados (Figura 61)

Objetivo(s) Entender qual é a percepção do usuário em relação às funcionalidades que visam apresentar informações adicionais para a interpretação do documento (Parágrafos Avaliados)

Tarefa(s) Analise o conteúdo do documento por meio da avaliação de parágrafos. Qual é o parágrafo melhor avaliado? Qual possui pior avaliação?

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 39

Protótipo(s) Interface de documento selecionado para análise de conteúdo (Figura 58)
Interface de sugestão de pesquisas complementares (Figura 62)

Objetivo(s) Entender qual é a percepção do usuário em relação às funcionalidades que visam apresentar informações adicionais para a interpretação do documento (Resultados Relacionados)

Tarefa(s) Existem resultados relacionados ao documento que você está analisando? Como eles poderiam contribuir com a sua pesquisa?

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 40

Protótipo(s) Interface de documento selecionado para análise de conteúdo (Figura 58)
Interface de sugestão de pesquisas complementares (Figura 62)

Objetivo(s) Entender qual é a percepção do usuário em relação às funcionalidades que visam apresentar informações adicionais para a interpretação do documento (Resultados Contraditórios)

Tarefa(s) Existem resultados contraditórios ao documento que você está analisando? Como eles poderiam contribuir com a sua pesquisa?

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 41

Protótipo(s) Interface de busca baseada em perspectivas de análise de conteúdo (Figura 63)

Objetivo(s) Entender a percepção do usuário sobre a seleção dos termos nas diversas perspectivas de análise de conteúdo, bem como sobre a elaboração de uma busca a partir das informações apresentadas

Tarefa(s) Realize uma busca baseada em Palavra-chave

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 42

Protótipo(s) Interface de busca baseada em perspectivas de análise de conteúdo (Figura 63)

Objetivo(s) Entender a percepção do usuário sobre as informações apresentadas no mapa de conceitos, bem como sobre as possibilidades de busca que são apresentadas

Tarefa(s) Analise as informações apresentadas no mapa de conceitos. Quais são suas opções de busca?

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 43

Protótipo(s) Interface de busca baseada em perspectivas de análise de conteúdo (Figura 63)

Objetivo(s) Entender a percepção do usuário sobre o refinamento das buscas realizadas

Tarefa(s) Selecione dois termos para refinar a busca: Ambiente Web e Processo de Busca

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 44

Protótipo(s) Interface de resultados de busca baseada em perspectivas de análise de conteúdo (Figura 64)

Objetivo(s) Entender a percepção do usuário sobre os resultados da busca

Tarefa(s) Analise os resultados da busca realizada. Quais publicações atentem os critérios definidos na busca (Autor: Rodrigues, P. e Palavra-Chave: Ambiente Web e Processo de Busca)?

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 45

Protótipo(s) Interface de resultados de busca baseada em perspectivas de análise de conteúdo (Figura 64)

Objetivo(s) Entender a percepção do usuário sobre a relevância dos resultados (resultados mais relevantes e menos relevantes), bem como sobre os termos que foram utilizados para representar os documentos recuperados

Tarefa(s) Ainda considerando os resultados da busca, quais publicações são menos relevantes, mas pertencem ao mesmo contexto da busca? Por qual motivo elas foram recuperadas?

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 46

Protótipo(s) Interface de rede de similaridade dos resultados de busca (Figura 65)

Objetivo(s) Entender a percepção do usuário sobre a similaridade existente entre os documentos retornados e as perspectivas de busca selecionadas

Tarefa(s) Ainda considerando os resultados da busca, qual é a relação entre o Doc.6 e a busca realizada?

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

CENÁRIO 5 – INCLUSÃO E ANÁLISE DE MATERIAL

Os testes realizados neste cenário têm como objetivo analisar a interação do usuário na inclusão de materiais científicos, bem como no desenvolvimento e atualização da rede de relacionamentos existente entre os documentos e a ontologia.

TAREFA 47

Protótipo(s) Interface inicial da ferramenta de busca (Figura 67)

Objetivo(s) Entender a percepção do usuário sobre os passos para a inserir um novo documento na base de dados

Tarefa(s) Faça o *upload* do documento *A survey of information retrieval interfaces*

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 48

Protótipo(s) Interface de relacionamentos entre material e ontologia de domínio (Figura 69)

Objetivo(s) Entender a percepção do usuário sobre a inclusão de relacionamentos entre um termo existente na ontologia e um documento

Tarefa(s) Selecione a perspectiva Metodologia e relacione o documento com o termo Análise de Conteúdo

O usuário conseguiu realizar a tarefa?
( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 49

Protótipo(s) Interface de inserção de novo termo à ontologia (Figura 70)

Objetivo(s) Entender a percepção do usuário sobre a inclusão de termos na ontologia

Tarefa(s) Selecione a perspectiva Metodologia e relacione o documento com o termo Etnografia. Caso o termo não exista, crie um novo termo e associe-o com o termo Entrevista

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

TAREFA 50

Protótipo(s) Interface de interpretação de um novo documento (Figura 68)

Objetivo(s) Entender a percepção do usuário sobre a interpretação dos documentos presentes no *corpus* (comentários interpretativos, avaliação de parágrafos, erros e falhas, resultados relacionados ao mesmo contexto e resultados contraditórios)

Tarefa(s) Interprete o documento inserido, de acordo com os seus conhecimentos e a sua percepção sobre o assunto

O usuário conseguiu realizar a tarefa?

( ) Sim, com facilidade ( ) Sim, com dificuldade ( ) Não conseguiu realizar a tarefa

Dificuldades Observadas
Comentários e Sugestões de Correção
Prioridade para a Correção

APÊNDICE B - Questionário Final de Avaliação de Usabilidade

QUESTIONÁRIO FINAL DE AVALIAÇÃO DE USABILIDADE

PERFIL DOS PARTICIPANTES

1. Sexo
   - ( ) Feminino
   - ( ) Masculino

2. Faixa Etária
   - ( ) De 14 a 23 anos
   - ( ) De 24 a 33 anos
   - ( ) De 34 a 43 anos
   - ( ) De 44 a 53 anos
   - ( ) 54 anos ou mais

3. Qual é sua formação?
   - ( ) Ensino Médio
   - ( ) Ensino Superior
   - ( ) Especialização
   - ( ) Mestrado (Cursando)
   - ( ) Mestrado (Concluído)
   - ( ) Doutorado (Cursando)
   - ( ) Doutorado (Concluído)

4. Qual é a sua área de formação (se for interdisciplinar, detalhe)?_____
_____

5. Qual é a sua profissão? _____

6. Como você avalia os seus conhecimentos em Informática?
   - ( ) Ótimo
   - ( ) Bom
   - ( ) Regular
   - ( ) Ruim

7. Com qual frequência você utiliza os serviços de busca na Web (Google, Bibliotecas Digitais, etc.)?
   - ( ) Diariamente
   - ( ) Semanalmente
   - ( ) Mensalmente
   - ( ) Raramente

AVALIAÇÃO DE USABILIDADE

8. Avalie os atributos apresentados a seguir, de acordo com a sua percepção em relação ao uso da interface.

	-3	-2	-1	0	1	2	3	
Desinteresse								Interesse
Desconfiança								Confiança
Complexa								Simples
Inútil								Útil
Confusão								Conforto
Não intuitiva								Intuitiva
Irritante								Agradável
Mal projetada								Bem projetada

9. Considere os recursos apresentados abaixo e avalie-os de acordo com o grau de utilidade que possuem para facilitar a busca e a análise de um *corpus* documental. Utilize uma escala de 1 a 5, onde 1 é menos útil e 5, mais útil. Escala: 1 – Nada Útil / 2 – Não tão Útil / 3 – Relativamente Útil / 4 – Muito Útil / 5 – Extremamente Útil

Recursos da Interface de Busca	**Importância**
1. Busca baseada em palavras-chave	
2. Busca baseada em estrutura terminológica (realizada a partir do mapa de conceitos)	
3. Busca baseada em uma ou várias perspectivas de contexto (autor, problema, metodologia)	
4. Busca baseada em uma ou várias perspectivas de um documento (autor, problema)	
5. Mapa de conceitos (recurso contendo os conceitos do domínio e as relações entre eles)	
6. Informações complementares (Referência e Resumo dos documentos selecionados)	
7. Plano de contexto (que apresenta o contexto dos termos selecionados para a busca)	
8. Barra de navegação estrutural	
9. Sinônimos e termos relacionados	
10. Antônimos e termos contraditórios	
11. Comentários interpretativos dos usuários (interação entre pesquisadores)	
12. Resultados relacionados ao mesmo contexto	
13. Indicação de erros, falhas e avaliação dos parágrafos	
14. Representação visual dos resultados da busca	

10. Avalie as questões abaixo e indique o grau de concordância/discordância em relação às afirmações sobre a ferramenta testada. Considere o 1 como sendo grau de menor concordância e o 5 o de maior concordância:

1 – Discordo Totalmente
2 – Discordo
3 – Sem Opinião
4 – Concordo
5 – Concordo Totalmente

Questões Gerais	**1**	**2**	**3**	**4**	**5**
1. Esta ferramenta tem interface de fácil interação					
2. Eu usaria este sistema com frequência					
3. Foi fácil encontrar a informação que eu precisava					
4. As funcionalidades disponibilizadas são de fácil aprendizagem					
5. Os recursos de navegação (ícones e botões) estão todos claros e fáceis de achar					
6. A ferramenta fornece todas as informações necessárias para completar as tarefas de forma clara e compreensível					
7. A análise dos dados é confusa					
8. Eu achei a ferramenta complexa. Precisei lembrar, pesquisar ou pensar muito para completar as tarefas					
9. A terminologia utilizada nos botões, rótulos, títulos etc. é de fácil entendimento					
10. Eu precisaria de apoio de uma pessoa para usar esta interface					
11. Eu me senti muito confiante usando esta ferramenta					
12. Esta ferramenta seria útil nas minhas atividades de pesquisa					

11. Avalie as questões abaixo e indique o grau de concordância/discordância em relação às afirmações sobre a ferramenta testada. Considere o 1 como sendo grau de menor concordância e o 5 o de maior concordância:

Ontologia como Componente de Interface	**1**	**2**	**3**	**4**	**5**
1. O mapa de conceitos me ajudou a entender quais informações estavam disponíveis para serem pesquisadas					
2. O mapa de conceitos foi útil para a elaboração das pesquisas realizadas					
3. A cor vermelha utilizada para destacar os conceitos e documentos que eu selecionei foi útil para me ajudar a compreender a informações presentes na interface					
4. O mapa de conceitos me ajudou a adquirir conhecimento sobre o domínio de pesquisa (termos sinônimos, antônimos, variações linguísticas)					
5. O plano de contexto me ajudou a compreender quais eram as minhas outras possibilidades de busca relacionadas ao assunto que eu estava pesquisando					
6. Os círculos de proximidade me ajudaram a perceber quais documentos eram mais relevantes					
7. O mapa de conceitos é confuso e dificultou minhas buscas					
8. Eu tive dificuldade para compreender quais documentos recuperados eram mais relevantes					
9. Eu aprendi termos novos durante a interação com o mapa de conceitos					
10. O mapa de conceitos me ajudou a elaborar uma busca em um domínio totalmente desconhecido					
11. O mapa de conceitos me ajudou a compreender como as informações estão relacionados em um domínio desconhecido					

12. Avalie as questões abaixo e indique o grau de concordância/discordância em relação às afirmações sobre a ferramenta testada. Considere o 1 como sendo grau de menor concordância e o 5 o de maior concordância:

Ontologia como Recurso para Indexação Semântica	**1**	**2**	**3**	**4**	**5**
1. Durante as minhas pesquisas foi possível explorar, de forma mais aprofundada, a estrutura dos documentos recuperados					
2. O mapa de conceitos me ajudou a encontrar documentos relacionados semanticamente com a busca que eu realizei					
3. Eu encontrei resultados relacionados com a minha expressão de busca, que eu nem sabia que existiam					
4. A organização dos documentos recuperados me ajudou a compreender a relação de similaridade entre eles					
5. Os recursos que permitem a exploração das perspectivas de análise (autor, objetivos, metodologia) não foram úteis para a minha pesquisa					
6. Os resultados correlacionados foram importantes para o contexto da minha pesquisa					
7. Os documentos recuperados são apresentados de forma desorganizada e sem critérios					

13. Avalie as questões abaixo e indique o grau de concordância/discordância em relação às afirmações sobre a ferramenta testada. Considere o 1 como sendo grau de menor concordância e o 5 o de maior concordância.

Interpretação de um Documento ou de um *Corpus*	**1**	**2**	**3**	**4**	**5**
1. A interpretação de um documento realizada por diversos outros usuários é útil para me ajudar a compreender o assunto pesquisado					
2. Quando eu tenho acesso a opiniões diferentes acerca de um mesmo documento ou assunto, é possível conhecer diversos pontos de vista, que estão associados a diferentes crenças, orientações políticas, valores culturais, preconceitos e outros					
3. Os termos apresentados no mapa de conceitos me ajudaram a compreender como as palavras estavam relacionados semanticamente em um contexto					
4. Foi possível conhecer palavras sinônimas relacionadas com a minha busca					
5. Conhecer diversos pontos de vista sobre um documento ou um assunto não contribui para o meu aprendizado					
6. O mapa de conceitos me apresentou as diversas nomenclaturas que estão associadas à um mesmo termo					
7. As sugestões de pesquisas relacionadas ao mesmo contexto contribuíram para ampliar as minhas possibilidades de análise sobre o tema					

14. Quais foram as suas principais dificuldades para realizar as tarefas propostas?

15. Você tem alguma sugestão de funcionalidade que poderia ser implementada na ferramenta para auxiliar o desenvolvimento das tarefas realizadas pelo usuário?