

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM TECNOLOGIA E SOCIEDADE

JOSEFINA APARECIDA SOARES GUEDES

**INDICADORES QUALITATIVOS PARA AVALIAÇÃO DO POTENCIAL
SEMÂNTICO DE PORTAIS DO CONHECIMENTO DE UNIVERSIDADES**

TESE

CURITIBA
2022

JOSEFINA APARECIDA SOARES GUEDES

INDICADORES QUALITATIVOS PARA AVALIAÇÃO DO POTENCIAL SEMÂNTICO DE PORTAIS DO CONHECIMENTO DE UNIVERSIDADES

QUALITY INDICATORS FOR ASSESSING THE SEMANTIC POTENTIAL OF UNIVERSITY KNOWLEDGE PORTALS

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Tecnologia e Sociedade, da Universidade Tecnológica Federal do Paraná, como requisito parcial para a obtenção do grau de Doutor em Tecnologia e Sociedade. Linha de Pesquisa Tecnologia e Desenvolvimento. Área de Concentração: Tecnologia e Sociedade.

Orientadora: Dra. Faimara do Rocio Strauhs

CURITIBA
2022



[4.0 Internacional](https://creativecommons.org/licenses/by/4.0/)

Esta licença permite compartilhamento, remixe, adaptação e criação a partir do trabalho, mesmo para fins comerciais, desde que sejam atribuídos créditos ao(s) autor(es). Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.



JOSEFINA APARECIDA SOARES GUEDES

**INDICADORES QUALITATIVOS PARA AVALIAÇÃO DO POTENCIAL SEMÂNTICO DE PORTAIS DO
CONHECIMENTO DE UNIVERSIDADES**

Trabalho de pesquisa de doutorado apresentado como requisito para obtenção do título de Doutora Em Tecnologia E Sociedade da Universidade Tecnológica Federal do Paraná (UTFPR). Área de concentração: Tecnologia E Sociedade.

Data de aprovação: 04 de Fevereiro de 2022

Prof.a Faimara Do Rocio Strauhs, Doutorado - Universidade Tecnológica Federal do Paraná

Prof Alain Hernandez Santoyo, Doutorado - Universidade Federal de Alfenas (Unifal-Mg)

Prof Egon Walter Wildauer, Doutorado - Universidade Federal do Paraná (Ufpr)

Prof Rodrigo Muller, Doutorado - Universidade Federal da Bahia (Ufba)

Prof Silvestre Labiak Junior, Doutorado - Universidade Tecnológica Federal do Paraná

AGRADECIMENTOS

À luz divina que ilumina firme e incansavelmente meus passos

A minha orientadora, Profa. Dra. Faimara do Rocio Strauhs, pela objetividade e ótima condução dos trabalhos.

Ao meu esposo, Eduardo H. Guedes pelo amor e companheirismo

Ao meu amado filho Murilo H. Guedes, pelo incentivo.

Aos meus pais João Soares e Onorfa B. Soares *in memoriam*, pela vida, amor e base sólida de educação que me ofertaram.

Aos colegas da UFPR e UTFPR por compartilharem os mesmos ideais.

À Banca de avaliação deste estudo, pelas contribuições e compartilhamento do seu conhecimento.

À Universidade Tecnológica Federal do Paraná (UTFPR), ao Programa de Pós-graduação em Tecnologia e Sociedade (PPGTE) e aos demais docentes por disseminarem o conhecimento.

“Não vá aonde o caminho pode levar, vá
aonde não há caminho e deixe um rastro”
(EMERSON, 2020, p. 1).

“Uma rede não é feita de fios de *nylon*,
palavras ou substâncias duráveis; ela é o
traço deixado por um agente em
movimento”
(LATOURE, 2012, p. 192, grifo nosso).

RESUMO

GUEDES, Josefina A. S. **Indicadores qualitativos para avaliação do potencial semântico de portais do conhecimento de universidades**. Orientadora: Faimara do Rocio Strauhs. 2021. 229 f. Tese (Doutorado em Tecnologia e Sociedade) – Programa de Pós-Graduação em Tecnologia e Sociedade, Universidade Tecnológica Federal do Paraná. Curitiba, 2022.

Os portais semânticos, como artefatos sociotécnicos, são regidos por um conjunto de pessoas e de tecnologias que se interconectam para disponibilizar informação de alto valor agregado. Os portais universitários podem se tornar semânticos quando dispõem de tecnologias e de condições específicas que os levem a se tornar portais do conhecimento e podem ser avaliados com o uso de indicadores. O objetivo desse estudo é o de propor indicadores qualitativos para a avaliação do potencial semântico de portais do conhecimento de universidades de acordo com os preceitos da Teoria Ator-Rede. A pesquisa bibliográfica, conduzida neste intuito, incluiu levantamento bibliométrico, análise sistemática e de conteúdo e uma pesquisa de campo com o envio de questionário para coordenadores de TI das 69 universidades federais brasileiras, a fim de identificar as tecnologias e os recursos humanos utilizados no portal de cada instituição. Os questionários foram aplicados no período de junho a setembro de 2021, sendo obtidas 42 respostas. O tratamento dos dados coletados foi efetuado mediante análise estatística descritiva resultando em 40 variáveis elencadas que permitiram a construção de um rol de indicadores; as variáveis foram validadas por um painel de especialistas que ratificou os indicadores propostos. O mapeamento da rede de atores dos portais foi efetuado para elencar os atores humanos e não humanos, seus interesses e inter-relações. A análise dos traços da rede de portais evidenciou laços fracos e fortes relacionados com a credibilidade dos parâmetros refletidos na prática diária e também aqueles relacionados com a inovação tecnológica e que são fundamentais para a evolução dos portais como artefatos sociotécnicos. Como resultado, foi proposto um conjunto de 18 indicadores qualitativos que podem ser utilizados como instrumento de gestão para criação ou melhoramento de portais do conhecimento de universidades. Infere-se que o uso desses indicadores pode melhorar a recuperação da informação que embasa a construção de conhecimento e fortalecer também os processos de comunicação e compartilhamento a partir da interface dos portais.

Palavras-chave: portais do conhecimento de universidades; portais semânticos; Teoria Ator-Rede; indicadores de potencial semântico; recuperação da informação

ABSTRACT

GUEDES, Josefina A. S. **Quality indicators for assessment of the semantic potential of university knowledge portals.** 2021. 229 f. Thesis (PhD in Technology and Society) - Graduate Program in Technology and Society, Federal Technological University of Paraná. Curitiba, 2022.

Semantic portals, as sociotechnical artifacts, are oriented by a set of people and technologies interconnected to provide information with high added value. For university portals to become semantic, it is necessary to have semantic technologies and specific conditions that can be analyzed using indicators. The aim of this study was to propose qualitative indicators to assess the semantic potential of university knowledge portals according to the precepts of the Actor-Network Theory. The research included a bibliometric survey, systematic and content analysis and a field research sending a questionnaire to technology and information coordinators from the 69 Brazilian federal universities, in order to identify the technologies and human resources available on the portal of each institution. The questionnaires were applied from June to September 2021, with 42 responses being obtained. Data processing was performed through descriptive statistical analysis and the 40 variables listed were validated by a panel of experts who ratified the content of the proposed indicators. The mapping of the portal actors' network was carried out to list the human and non-human actors, their interests and interrelationships. The analysis of the features of the portal network evidenced the weak and strong ties related to the credibility of the parameters reflected in daily practice and also those related to technological innovation and which are fundamental for the evolution of portals as a socio-technical system. As a result, a set of 18 qualitative indicators was proposed that can be used as a management tool for creating or improving university knowledge portals. It's inferred that the use of these indicators can improve the retrieval of information that supports the construction of knowledge and also strengthen the communication and sharing processes from the portals interface.

Keywords: university knowledge portal; semantic portal; Actor-Network Theory; semantic potential indicators; retrieval information

LISTA DE FIGURAS

Figura 1 - Classificação metodológica da pesquisa.	35
Figura 2 - Relacionamento de pesquisa na lógica booleana.	42
Figura 3 - Tipos de ontologia.	46
Figura 4 - Arquitetura da Web Semântica.	50
Figura 5 - Roteiro para avaliação de portal semântico.	65
Figura 6 - Condições mínimas para estabelecimento de potencial semântico de portais.	82
Figura 7 - Potenciais indicadores qualitativos semânticos.	83
Figura 8 - Fases da Tradução.	91
Figura 9 - Fases de desenvolvimento da pesquisa.	99
Figura 10 - Fases do levantamento bibliométrico.	100
Figura 11 - Diagrama representativo da combinação de palavras-chaves usadas na pesquisa nas bases de dados e a interface de resultado (central em preto).	101
Figura 12 - Fluxo da informação com as fases da revisão sistemática.	104
Figura 13 - Nuvem de palavras gerada para a categoria de contexto “Recuperação da Informação”.	112
Figura 14 - Nuvem de palavras gerada para a categoria de contexto “Portal Semântico”.	112
Figura 15 - Nuvem de palavras geradas para a categoria de contexto “Sistemas Sociotécnicos”.	113
Figura 16 - Etapas de tratamento dos dados.	128
Figura 17 - Estrutura do Capítulo 6.	129
Figura 18 - Distribuição das universidades federais nas regiões do país com indicação do número de respostas recebidas ao Questionário.	131
Figura 19 - Rede simples de tradução do portal universitário.	134
Figura 20 - Rede de atores do portal universitário semântico.	135
Figura 21 - Administradores na rede do portal semântico.	137
Figura 22 - Desenvolvedores na rede do portal semântico.	139
Figura 23 - Tecnologias na rede do portal semântico.	142
Figura 24 - Usuários na rede do portal semântico.	145
Figura 25 - Conjunto de indicadores qualitativos de potencial semântico.	170

LISTA DE GRÁFICOS

Gráfico 1 - Distribuição das 33 fontes de estudo por data de publicação.....	105
Gráfico 2 - Fontes de informação utilizadas no estudo.	105
Gráfico 3 - Distribuição de fontes de informação por tema abordado	106
Gráfico 4 - Número de citações por autor.....	107
Gráfico 5 - Utilização de tecnologias de padrão aberto.....	148
Gráfico 6 - Sistema de banco de dados utilizado	148
Gráfico 7 - Servidor Web utilizado.....	149
Gráfico 8 - Sistema de segurança para nuvem e rede	149
Gráfico 9 - Sistema utilizado para pesquisa no portal	150
Gráfico 10 - Sistema utilizado para pesquisa no portal	151
Gráfico 11 - Forma de gestão de conteúdo adotada.....	153
Gráfico 12 - Forma de edição de itens de informação.....	153
Gráfico 13 - Forma de atualização das informações.....	154
Gráfico 14 - Forma de processamento da informação	154
Gráfico 15 - Forma de publicação de informações.....	155
Gráfico 16 - Gerenciador de conteúdo usado	156
Gráfico 17 - Mecanismo de inferência usado	156
Gráfico 18 - Linguagens de programação e modelo de dados usados	157
Gráfico 19 - Uso de editor de ontologia.....	158
Gráfico 20 - Uso de tecnologias para importação e exportação.....	158
Gráfico 21 - Membros com permissões para escrever conceitos, instâncias, atributos e colaborar na construção de ontologias.....	159
Gráfico 22 - Processamento semântico ou estrutura da ontologia com materiais para consulta.....	159
Gráfico 23 - Informação semanticamente completa e consistente.....	160
Gráfico 24 - Profundidade de cobertura das informações	160
Gráfico 25 - Gráfico - Uso de vocabulários.....	161
Gráfico 26 - Maturidade de implantação/manutenção do portal.....	161
Gráfico 27 - Administradores do portal.....	162
Gráfico 28 - Desenvolvedores do portal.....	163
Gráfico 29 - Participação dos usuários na construção/ melhoramento do portal...	164
Gráfico 30 - Informações relevantes do domínio específico da comunidade acadêmica.....	165
Gráfico 31 - Mapa de tópicos pessoais para usuários.....	165
Gráfico 32 - Recurso de colaboração, comunicação e compartilhamento	166
Gráfico 33 - Mecanismos de ajuda.....	166
Gráfico 34 - Laços fortes e fracos da rede de portais de universidades.....	177

LISTAS DE QUADROS

Quadro 1 - Características de portais tradicionais e de portais semânticos – comparativo.....	79
Quadro 2 - Arcabouço conceitual.....	93
Quadro 3 - Matriz de consistência.....	98
Quadro 4 - Estratégia única de pesquisa	102
Quadro 5 - Número de citações por autor e base de referência.....	107
Quadro 6 - Recorte da análise sistemática.....	109
Quadro 7 - Categorias de Contexto, de análise e Unidades de Contexto e de Registro para a Análise de Conteúdo.....	114
Quadro 8 - Estrutura do questionário.....	118
Quadro 9 - Questões sobre armazenamento de dados e tecnologias de padrão aberto.....	118
Quadro 10 - Questões sobre formas e gestão da informação.....	119
Quadro 11 - Questões sobre gestão de ontologia.....	119
Quadro 12 - Questões sobre gestão e avaliação do portal.....	120
Quadro 13 - Protocolo de coleta de dados – Objetivo específico 1.....	122
Quadro 14 - Protocolo de coleta de dados – Objetivo específico 2.....	123
Quadro 15 - Protocolo de coleta de dados – Objetivo específico 3.....	124
Quadro 16 - Protocolo de coleta de dados – Objetivo específico 4.....	125
Quadro 17 - Protocolo de coleta de dados – Objetivo específico 5.....	127
Quadro 18 - Atores da rede de portais semânticos	141
Quadro 19 - Tecnologias para estruturação semântica de portais	146
Quadro 20 - Conjunto de indicadores qualitativos de potencial semântico.....	170
Quadro 21 - Parâmetros para laços fortes e fracos.....	176
Quadro 22 - Indicadores qualitativos com parâmetros semânticos e sociotécnicos.....	179

LISTA DE TABELAS

Tabela 1 - Busca por palavras-chave isolada por bases de dados.....	29
Tabela 2 - Resultados das estratégias de busca em bases de dados.....	30
Tabela 3 - Resultados do Levantamento bibliométrico.....	103
Tabela 4 - Respostas do questionário por estratos da população.....	132
Tabela 5 - Variáveis de interesse.....	168
Tabela 6 - Validação dos dados.....	171
Tabela 7 - Variáveis de interesse e validação de indicadores.....	172

LISTA DE SIGLAS E ABREVIATURAS

API	- Application Programming Interface
BIREME	- Centro Latino-Americano e Caribe de Informação em Ciências da Saúde
DAML	- DARPA Agent Markup Language
FOAF	- Friend of a Friend
HTML	- Hyper Text Markup Language
NS	- Name Space
OIL	- Ontology Inference Layer
OWL	- Web Ontology Language
OWL DL	- Web Ontology Language Description Logic
RDF	- Resource Description Framework
RIF	- Rule Interchange Format
RuleML	- Rule Markup Language
SciELO	- Scientific Electronic Library Online
SWAD	- Semantic Web Advanced Environmental Directory
SWRL	- Semantic Web Rule Language
TAR	- Teoria Ator-Rede
URI	- Uniform Resource Identifier
URL	- Uniform Resource Locator
Web	- World Wide Web
W3C	- World Wide Web Consortium
XML	- Extensible Markup Language

SUMÁRIO

1 INTRODUÇÃO	14
1.1 TEMA	14
1.2 DELIMITAÇÃO DA PESQUISA.....	18
1.3 PROBLEMÁTICA E PREMISSAS.....	20
1.4 OBJETIVOS	24
1.4.1 Objetivo geral	25
1.4.2 Objetivos específicos.....	25
1.5 JUSTIFICATIVAS TEÓRICO-PRÁTICAS.....	25
1.5.1 Justificativa teórica	25
1.5.2 Justificativa prática	31
1.6 MARCO TEÓRICO.....	33
1.7 ABORDAGEM METODOLÓGICA DA PESQUISA.....	34
1.8 REVISÃO DE LITERATURA	36
1.9 ESTRUTURA DO TRABALHO	37
2 INFORMAÇÃO: das formas de recuperação aos modelos de representação.	39
2.1 MODELOS DE RECUPERAÇÃO DA INFORMAÇÃO	39
2.1.1 Modelos tradicionais de recuperação da informação	41
2.1.2 Modelos contemporâneos de recuperação da informação.....	43
2.2 ONTOLOGIA COMO MODELO DE RECUPERAÇÃO SEMÂNTICA	45
2.2.1 A ontologia formal	45
2.2.2 Representação semântica da informação.	49
2.2.3 Dados vinculados e dados abertos vinculados.....	54
2.2.4 Sistemas complexos e <i>big data</i>	57
3 PORTAIS SEMÂNTICOS	62
3.1 CONCEITO DE PORTAIS.....	62
3.2 ARQUITETURA DE PORTAIS SEMÂNTICOS	64
3.3 TECNOLOGIAS SEMÂNTICAS	66
3.3.1 Linguagem de representação semântica e ferramentas de Ontologia	67
3.3.2 Interfaces de programação de gerenciamento de ontologias.....	68
3.4 PORTAIS SEMÂNTICOS DE BUSCA – OBJETIVOS, CARACTERÍSTICAS, INDICADORES DE AVALIAÇÃO E EVOLUÇÕES.....	70
3.4.1 Objetivos, características e evolução dos portais semânticos.....	71

3.4.2 Indicadores de avaliação do potencial semântico	80
4 ALINHAMENTO CONCEITUAL	84
4.1 SISTEMAS SOCIOTÉCNICOS	84
4.2 TEORIA ATOR-REDE	87
4.3 PORTAIS SEMÂNTICOS COMO ARTEFATOS SOCIOTÉCNICOS	94
5 PROCEDIMENTOS METODOLÓGICOS DA PESQUISA	97
5.1 CLASSIFICAÇÃO, DESENVOLVIMENTO E PROTOCOLO DE PESQUISA.....	97
5.1.1 Levantamento bibliométrico.....	99
5.2 ANÁLISE SISTEMÁTICA E ANÁLISE DE CONTEÚDO.....	108
5.2.1 Análise sistemática.....	108
5.2.2 Análise de conteúdo	110
5.2.2.1 Pré-Análise.....	110
5.2.2.2 Exploração do Material.....	111
5.3 COLETA, TRATAMENTO, INTERPRETAÇÃO DOS DADOS E INFERÊNCIA.	115
5.3.1 População e Amostra	115
5.3.2 Instrumento de coleta de dados	116
5.3.3 Tratamento e validação dos dados	127
6 APRESENTAÇÃO DOS DADOS E DISCUSSÃO DOS RESULTADOS DA PESQUISA.....	130
6.1 APRESENTAÇÃO DOS DADOS DA PESQUISA	130
6.2 REDE DE ATORES DOS PORTAIS DO CONHECIMENTO DE UNIVERSIDADES	132
6.3 ANÁLISE E VALIDAÇÃO DE DADOS.....	146
6.3.1 Dados coletados.....	147
6.3.1.1 Armazenamento de dados e sua gestão.....	147
6.3.1.2 Gestão da Informação.....	150
6.3.1.3 Gestão de Ontologia.....	158
6.3.1.4 Rede de Atores - Caracterização Sociotécnica	162
6.3.2 Proposta de Indicadores.....	167
6.3.3 Validação dos indicadores propostos	171
7 CONSIDERAÇÕES FINAIS	181
7.1 ATENDIMENTO AOS OBJETIVOS E À PERGUNTA DE PESQUISA.....	181
7.2. LIMITAÇÕES DA PESQUISA E TRABALHOS FUTUROS	184
REFERÊNCIAS.....	186

APÊNDICE 1 – ANÁLISE SISTEMÁTICA DA LITERATURA	205
APÊNDICE 2 – QUESTIONÁRIO DA PESQUISA.....	210
APÊNDICE 3 – QUESTIONARIO PARA VALIDAÇÃO DE DADOS	219
APÊNDICE 4 – ÍNDICE ONOSMÁTICO	223

1 INTRODUÇÃO

Este Capítulo aborda o tema pesquisado, sua delimitação e o principal objeto de estudo, a problemática da pesquisa, os objetivos e a justificativa do tema, considerando a sua relevância e aplicação. Serão delineados, ainda, os procedimentos metodológicos da pesquisa e apresentada a estruturação do trabalho desenvolvido.

1.1 TEMA

O conceito de tecnologia social implica desenvolvimento de tecnologias voltadas para resolver problemas que a sociedade enfrenta nos mais diversos contextos e locais (LINSINGEN; BAZZO; PEREIRA, 2003; VACCAREZZA, 2011).

O aparato técnico isoladamente é considerado um fator parcial, mas quando aplicado e utilizado pelos grupos sociais, torna-se parte integral da tecnologia, assim, o indivíduo constrói e manipula artefatos tecnológicos que lhe permite operar nos mais diversos ambientes (MARCUSE, 1998).

Nesse enquadramento, artefatos tecnológicos, como portais universitários, devem ser planejados de modo a considerar os interesses e as necessidades de todos os atores envolvidos no desenho de uma tecnologia, não se limitando a procedimentos e requisitos essencialmente técnicos (FORLANO, 2017; IYAMU; MGUDLWA, 2018; LAW, 1992; NOVAES; DAGNINO, 2004; ROSE; SCHEEPERS 2001).

Os portais, especialmente os educacionais, são caracterizados por disponibilizar informações a partir de um ponto único de acesso, o que favorece o aprendizado, a socialização, a disseminação e a construção do conhecimento (NONAKA; TAKEUCHI, 1997; STRAUHS *et al.*, 2012). Para tanto, devem ser dotados de um sistema de recuperação da informação que possibilite rapidez e qualidade de resultado das buscas (FERNÁNDEZ *et al.*, 2011). A recuperação semântica da informação, fundamentada em modelagem ontológica, que pressupõe o uso de ontologia formal, aumenta a precisão e a rapidez nesse acesso à informação (WIMMER; YOON; RADA, 2013).

A modelagem ontológica se refere ao uso de uma ontologia formal na estrutura de portais, juntamente com tecnologias que expressam semântica (GRUBER, 1996; TELNOV, 2015). A ontologia inclui um vocabulário controlado formado por termos

estruturados a partir de regras, sendo possível a recuperação semântica da informação que vai além das relações sintáticas de palavras-chaves comuns usadas na busca em portais. A pesquisa semântica permite que o resultado da busca seja de maior qualidade, com recuperação de informação relevante para o usuário (GRUBER, 1996; MALHOTRA; NAIR, 2015; NGO; CAO, 2018; TELNOV, 2015).

Destarte, a semântica, associada às técnicas de recuperação da informação, traz benefícios à sociedade em geral, consumidora de informação, ao mesmo tempo em que possibilita um avanço tecnológico para os portais.

Neste sentido, o uso de ontologias na estruturação de portais favorece a construção de um ambiente colaborativo e interdisciplinar, integrando processos e pessoas envolvidas com a informação e o conhecimento. A modelagem ontológica, então, possibilita maior precisão e rapidez no acesso à informação, além de favorecer troca de experiências e inovação (FERNÁNDEZ *et al.*, 2011; HUANG; MAIDMENT; TIAN, 2011; WIMMER; YOON; RADA, 2013).

Aprofundando, destaca-se a necessidade de avaliação do potencial semântico de portais educacionais a partir da análise das características e das ferramentas tecnológicas adotadas, com ênfase na criação de artefatos tecnológicos socialmente construídos. Neste estudo, o potencial semântico pode ser entendido como sendo a qualidade de resposta na busca de informações em um portal. Está associado ao uso de ferramentas, de tecnologias e de serviços aos usuários previstos na estruturação da Web Semântica conforme difundido por Tim Bernes-Lee, desde 2001. Por conseguinte, um portal semanticamente estruturado, facilita o uso e o reuso da informação, favorecendo a construção de novos conhecimentos, o que caracteriza uma evolução destes para portais do conhecimento (GARGIULO *et al.*, 2014; WORLD WIDE WEB - W3C, 2019).

Os portais educacionais simplesmente disponibilizam informações, já portais que tenham potencial semântico possibilitam a criação de conhecimento e podem ser chamados de portais do conhecimento (HAUKE; OWOC; PONDEL, 2015). Assim nesta pesquisa, a partir desse ponto, será utilizada a designação de portais do conhecimento de universidades, para aqueles portais com potencial semântico.

Os portais, de modo geral, como artefatos sociotécnicos, estão sujeitos à aplicação de um conjunto de normas técnicas e de um sistema de regras, de leis e de princípios normativos, tais como o uso legítimo ou aceitável da tecnologia (BURNS, 2006). Sob esta ótica, a construção de portais envolve uma rede de atores humanos

e não humanos que articulados possibilitam a construção social de um portal da Web, conforme preconizado na Teoria Ator Rede – TAR (CALLON, 1986c), usufruindo de vários de seus conceitos, mas especialmente da simetria, da tradução (CORDELLA; 2010; ELBANNA, 2012; PRADO; BARANAUSKAS, 2012) e do interessamento (HU, 2011; MACOME, 2008), tais conceitos serão discutidos no texto posteriormente.

Para Latour (2001), os artefatos da tecnociência não devem ser vistos como produtos acabados já que não dependem apenas de seus atributos, mas de como a sociedade os avalia, portanto, esses artefatos precisam ser legitimados pela sociedade. Para tanto se faz necessária articulação entre os atores internos e os externos, materiais e imateriais, envolvidos no desenvolvimento de uma tecnologia a fim de que os objetivos sejam alinhados em prol do interesse de todos os atores (EROFEEVA; 2017; PEDRO, 2010).

Os portais educacionais, exemplificados pelos portais do conhecimento de universidades, constituem pontos de entrada para a apresentação e o intercâmbio de informações na Web. Para ser dinâmico um portal educacional deve propiciar o compartilhamento de informações (BERNERS-LEE; HENDLER; LASSILA, 2001; HAWKE *et al.*, 2013; GARGIULO *et al.*, 2014), facilitando a criação de conhecimento novo que, por sua vez, constitui-se um fator estratégico para as instituições (MAROUF; KHALIL, 2015; NONAKA; TAKEUCHI, 1997**Erro! Indicador não definido.**; POLANYI, 1966**Erro! Indicador não definido.**; STRAUHS *et al.*, 2012**Erro! Indicador não definido.**; TATTO; BORDIN, 2016).**Erro! Indicador não definido.**

Para que sejam efetivos, os portais do conhecimento de universidades precisam dispor de uma estrutura que favoreça o acesso às informações de forma rápida e precisa. Tal agilidade e rapidez estão condicionadas às tecnologias e às ferramentas utilizadas, bem como ao tipo de modelo de recuperação de informação usado no portal (HAWKE *et al.*, 2013; LAUSEN *et al.*, 2005; SHIRISHA; SUBBARAO; KAVITHA, 2015).

A construção dos portais do conhecimento de universidades, do ponto de vista de um artefato técnico, portanto, constitui uma rede que envolve ação multitarefa, englobando instrumentos, recursos, profissionais, usuários da informação e outros atores com diferentes interesses que precisam estar alinhados em prol de um objetivo comum (FERNÁNDEZ *et al.*, 2011; SILVELLO *et al.*, 2017). O propósito final consiste em se ter uma plataforma para disponibilizar e recuperar informações de acordo com os interesses dos responsáveis por alimentar o portal, mas, sobretudo, atendendo aos

interesse dos usuários finais (LI *et al.*, 2017; MALGAONKAR; DEVALE, 2016; SILVELLO *et al.*, 2017). É preciso salientar, no entanto, que todos esses elementos estão interconectados e não podem ser vistos isoladamente sob pena de não atender as necessidades de todos os atores da rede dos portais, incorporando os conceitos de Latour (2001).

Portais do conhecimento de universidades, se construídos a partir dos fundamentos da teoria Ator-Rede (TAR), proposta por Michel Callon (1986), John Law (2010) e Bruno Latour (2012), podem ser considerados uma rede sociotécnica. Essa rede é formada por um conjunto de atores heterogêneos, compostos por elementos humanos e não-humanos que interagem no processo de concepção, de produção e de difusão de conhecimentos, dando origem a definições tecnológicas obtidas no processo de solução de problemas (LATOURE, 2001).

Segundo Callon (1986c), em uma rede de atores ocorrem alianças capazes de definir e de modificar componentes mediante múltiplas conexões e entradas na rede, estabelecendo um campo de tensões heterogêneas que associam competências a equipamentos, textos a saberes táticos humanos a não humanos (CALLON, 1989). Assim, na rede analisa-se simetricamente não só o erro e o acerto, mas todo e qualquer efeito das negociações que ocorrem nesta, incluindo a natureza e a sociedade, tal como descrito no princípio de simetria generalizada proposto por Latour (1996).

Os portais do conhecimento de universidades, embora sejam fundamentados em critérios de usabilidade na sua construção e que incluem além da facilidade de uso, a aprendizagem, a colaboração e a satisfação de seus usuários, (FERNÁNDEZ *et al.*, 2011; VITAL; CAFÉ, 2011), não preveem a participação do usuário durante a sua estruturação e os problemas de recuperação da informação ficam mais evidentes quando o portal já está funcionando. Este fato distancia os portais tradicionais das teorias que regem a construção social da tecnologia, pois deixa à margem um importante ator, o usuário da comunidade dos portais.

Um desafio da pesquisa direta na interface dos portais é que estes devem dispor de técnicas de recuperação de informação capazes de acessar de forma eficaz e eficiente os dados desejados, explorando ao mesmo tempo características de descrição de metadados e a semântica relacionada para recuperação da informação de modo a corresponder às necessidades e às preferências temáticas ou filtros de pesquisa de interesse do usuário (AMATO *et al.*, 2017). Sendo assim, torna-se

importante avaliar a estrutura dos portais a fim de verificar se as tecnologias utilizadas, os recursos e os serviços disponibilizados favorecem a busca de informações.

Nesse sentido, a avaliação de determinado problema ou questão pode ser efetuada mediante a utilização de indicadores qualitativos, já que estes podem ajudar a traduzir as dimensões da realidade a ser investigada empiricamente em valores para fins de acompanhamento e comparação (ANDRADE; VAITSMAN; FARIAS, 2010).

O tema deste estudo, no contexto esboçado é, então, a facilitação do acesso às informações, a partir da análise de portais semânticos das universidades federais, apoiado pela construção de indicadores qualitativos contendo parâmetros sociotécnicos com base nos preceitos da TAR e da *Web 3.0*. Os indicadores podem contribuir tanto para a avaliação semântica como para a construção dos portais que ofereçam maior qualidade no retorno das buscas.

A *Web 3.0* é uma descrição alternativa para a *Web Semântica*, cujas normativas preveem a atribuição de significado definido à informação, ampliando-se a interação entre computadores e indivíduos (BERNERS-LEE; HENDLER; LASSILA, 2001; SILVA; SANTOS; FERNEDA, 2013). Dessa perspectiva, os portais semânticos podem ser considerados uma tecnologia social, pois, incluem tecnologias, processos e pessoas com interesse em recuperar informação relevante para uso em diferentes contextos (IYAMU; MGUDLWA, 2018), tais contextos tem seus delineamentos iniciados a partir da Seção seguinte

1.2 DELIMITAÇÃO DA PESQUISA

Da revisão da literatura depreende-se que os portais semânticos podem ser concebidos a partir de uma rede que liga ontologia e tecnologias da *Web semântica* em uma gama de atores humanos e não humanos (FORLANO, 2017; IYAMU; MGUDLWA, 2018; LATOUR, 2001; MAATOUK, 2021).

A ontologia formal, conforme fundamentada por Gruber (1996) e Guarino (1997), estabelece relações semânticas entre conceitos, facilitando a recuperação de informações em portais, com o uso de tecnologias como motores de busca e outras tecnologias que compõe o domínio da Inteligência Artificial. Já a *Web Semântica*, proposta por Berners-Lee, Hendler e Lassila (2001) associada à ontologia de domínio possibilita atribuir um significado definido à informação, propiciando melhor interação entre computadores e pessoas (SILVA; SANTOS; FERNEDA, 2013). Desta forma, a

inteligência artificial facilita o compartilhamento de informações, automatizando tarefas de rotina dos usuários. Neste sentido, os resultados de busca em portais estruturados com tecnologias semânticas são mais precisos e o usuário pode selecionar as informações com maior rapidez (LI *et al.*, 2017; MALGAONKAR; DEVALE, 2016; SILVELLO *et al.*, 2017).

Esta pesquisa abordou a elaboração de indicadores qualitativos para avaliação de portais semânticos e a facilitação ao acesso de informações. Acredita-se que indicadores qualitativos podem consistir em uma ferramenta para uso na prática diária no contexto dos portais do conhecimento de universidades, enfatizando-se a construção social dos artefatos tecnológicos com foco na produção científica.

Destacados os aspectos de objeto e de campo de pesquisa avança-se para o detalhamento do espaço pesquisado. Sobre a territorialidade, Albagli (2004) conceitua o território geográfico como um espaço apropriado por atores, definido e delimitado pelas relações de poder que ocorrem no campo social, econômico, político e cultural. A dimensão sociopolítica envolve as interações sociais que ocorrem em um espaço em que atores podem dominar ou influenciar a rede de relações sociais.

Neste contexto, os portais das universidades federais brasileiras constituem um território e o espaço geográfico e social desta pesquisa, posto que em universidades o portal é usado para disseminação de informações acadêmicas, mas também de interesse para a comunidade (ALBAGLI, 2004; HAUKE; OWOC; PONDEL, 2015).

O quadro das universidades federais brasileiras é atualmente composto por 69 unidades, localizadas em todas as regiões do país. As primeiras universidades no Brasil surgiram em 1920, a exemplo da Universidade do Rio de Janeiro, da Universidade de São Paulo e da Universidade do Paraná (SOUZA; MIRANDA; SOUZA; 2019). Em 2019 novas instituições foram criadas, dentre estas a Universidade Federal do Delta do Parnaíba e a Universidade Federal de Catalão (MASSALLI, 2019; SOUZA; MIRANDA; SOUZA; 2019; UNIVERSIDADE FEDERAL DO PARANÁ, 2021).

As universidades federais constituem, portanto, um ambiente propício para levantar as características de indicadores semânticos que norteiem a estruturação e a avaliação de portais, com vistas à construção de artefatos sociotécnicos de qualidade em termos de recuperação semântica da informação e foram, também, o local de validação dos dados coletados.

A pesquisa incluiu o levantamento das condições tecnológicas em que os portais foram estruturados tendo por base os preceitos da Web semântica. A ênfase foi nos requisitos tecnológicos e nas condições necessárias para que os portais ofereçam modalidade de pesquisa semântica, de forma rápida e eficiente. Não fará parte do escopo desta pesquisa o desenvolvimento de portais, mas sim apresentar indicadores qualitativos que orientem a configuração e a criação destes artefatos de modo que os objetivos dos usuários finais sejam contemplados.

Salienta-se que a interação que ocorre entre atores humanos e não humanos em uma rede é marcada por laços fracos e fortes entre os atores (GRANOVETTER, 1983). Esses traços deixados em um ambiente o delineiam como um artefato sociotécnico, conforme destacaram Burns (2006), Callon (1986 a, b, c), Hughes (1989), Latour (2000) e Thomas (2008). Assim, os portais semânticos construídos partir desta abordagem, levantados problemas e senões, podem ter por objetivo fornecer informações pertinentes e situadas acerca do universo acadêmico ou das universidades.

1.3 PROBLEMÁTICA E PREMISSAS

A despeito da facilidade de acesso ao conhecimento via Internet, a seleção de informações ainda é dificultada pelo próprio acúmulo destas e pelo modo como estas informações são recuperadas na Web (ABDULLAR; ZAMIL, 2018; LEE, *et al.*, 2018). Para que um portal seja efetivo, conforme já mencionado, faz-se necessária a avaliação de sua estrutura a fim de saber se esta contempla as necessidades de seus usuários. Esta avaliação pode ser facilitada pelo uso de indicadores qualitativos, por exemplo.

As ferramentas e as tecnologias usadas nos sistemas tradicionais de recuperação da informação dificultam a recuperação precisa e rápida em uma busca na Web (AL BALUSHI; SAMPAIO; LOUCOPOULOS, 2013; BENABDERRAHMANE *et al.* 2017; LEE, *et al.*, 2018; LI *et al.*, 2017; SILVELLO *et al.*, 2017). Com a explosão das informações da Internet, o algoritmo tradicional de mapeamento de classificação das páginas da Web não consegue lidar com a classificação complexa das páginas ou *sites* da Web (LI *et al.*, 2017), dificultando a seleção de informações relevantes.

O uso de tecnologias semânticas melhora a recuperação da informação e equaliza o interesse dos usuários e o conteúdo da Web (LI *et al.*, 2017;

MALGAONKAR; DEVALE, 2016). Nesse contexto, a crescente demanda pelo desenvolvimento de sistemas de informação complexos que reduzam custos e garantam a confiabilidade leva à adoção de diferentes paradigmas com abordagens que incluem compartilhamento de conhecimento, interoperabilidade, integridade, qualidade e reutilização de dados que são os pressupostos da Web Semântica ou Web. 3.0 (GARGIULO *et al.*, 2014; WORLD WIDE WEB - W3C, 2019).

O cenário tecnológico, portanto, demanda que um portal do conhecimento de universidade reúna e apresente informações relevantes para a comunidade acadêmica e também para a sociedade em geral, já que disponibiliza dados em sistemas abertos (AZANI, 2009). Entretanto, portais concebidos a partir de estruturação tradicional são geralmente ineficientes em atender aos requisitos de qualidade, o que pode gerar insatisfação do usuário (AMATO *et al.*, 2017; BELLANDI *et al.*, 2012; LI *et al.*, 2017; MARTINS; SORAES; FRANCELIN, 2012; MOURA *et al.* 2015; SILVELLO *et al.*, 2017), evoluindo-se para uma demanda de portais de conhecimento, especialmente os portais universitários.

A maioria dos portais tradicionais disponibiliza informações que são recuperadas por meio de palavras-chave indexadas em categorias pré-definidas, com uma hierarquia única de classificação que pode não corresponder ao assunto que o usuário procura, tornando a pesquisa imprecisa (CORREA, 2012). Além disso, inexistente um padrão na apresentação das informações com ausência de separação entre o conteúdo dos documentos e a apresentação das informações o que está associado à grande quantidade de retorno de documentos irrelevantes nas pesquisas (AMATO *et al.*, 2017; BELLANDI *et al.*, 2012; CORREA, 2012; FERNÁNDEZ *et al.*, 2011; SILVELLO *et al.*, 2017).

As técnicas de recuperação da informação usadas nos portais tradicionais geralmente incluem modelo hierárquico, como a taxonomia ou o modelo de árvore de prefixos, como o Practical Algorithm To Retrieve Information Coded In Alphanumeric¹ – PATRICIA (MORRISON, 1968). Contudo, tais sistemas são desprovidos de recursos para tratar a polissemia e a sinonímia², fenômenos linguísticos que reduzem fortemente a precisão da consulta por palavra-chave nos portais, causando a recuperação de diversos documentos irrelevantes (LI *et al.*, 2017; MALGAONKAR;

¹ Algoritmo prático para recuperar informações codificadas em alfanumérico.

² Polissemia é a propriedade que uma palavra tem de assumir vários sentidos. Sinonímia é a relação de sentido entre palavras da mesma categoria gramatical que podem ter o mesmo ou quase o mesmo significado (MARTINS; SOARES; FRANCELIN, 2012).

DEVALE, 2016; MARTINS; SOARES; FRANCELIN, 2012). Assim, acredita-se que os interesses do usuário final não estariam sendo contemplados, o que constitui uma falha na construção de portais do conhecimento de universidades como artefatos sociotécnicos.

Com a proposta da Web Semântica é possível, no entanto, estruturar portais com ferramentas que possibilitam inserir significado nos documentos para que as máquinas também possam decodificar o conteúdo apresentado em um texto na Web (AMATO *et al.* 2017; BERNERS-LEE; HENDLER; LASSILA, 2001; PAIVA; COSTA; SILVA, 2014; TELNOV, 2015). Essa estrutura inclui o uso de ontologias de domínio que definem as regras que regulam a combinação entre os termos e suas relações (GUARINO, 1997), sendo desenvolvidas por engenheiros de computação para recuperação no meio digital (BELLANDI *et al.*, 2012; FERNÁNDEZ *et al.*, 2011).

Aproximando o conceito de rede sociotécnica e da Web Semântica ou Web 3.0 como é conhecida (BIJKER; HUGH; PINCH, 1989; BURNS, 2006; IYAMU; MGUDLWA, 2018; SHIRISHA; SUBBARAO; KAVITHA, 2015), visualizam-se interações, uma vez que estas propõem uma articulação de tecnologias e de condições específicas que alinhadas pela ação humana aumentam a qualidade da recuperação de informações para o usuário do portal universitário (BERNERS-LEE; HENDLER; LASSILA, 2001; CALLON, 1986a, GUO *et al.*, 2018; LAW, 2010; SILVELLO *et al.*, 2017).

As tecnologias semânticas ampliam as possibilidades de recuperação da informação contida nos portais acadêmicos já que incluem agentes inteligentes que, com o auxílio de ontologias de domínio, "entendem" o significado da informação e, assim, recuperam e manipulam informações pertinentes, disponibilizando-as ao usuário (BELLANDI *et al.*, 2012; BERNERS-LEE; HENDLER; LASSILA, 2001; GAVRILOVA, 2011; LEE *et al.*, 2018; LI *et al.*, 2017; MALGAONKAR; DEVALE, 2016; SILVELLO *et al.*, 2017; WIMMER; YOON; RADA, 2013).

A semântica é, portanto, o elemento que traduz os interesses dos usuários dos portais do conhecimento de universidades. As tecnologias semânticas associadas ao uso de ontologia formal como modelo de recuperação da informação possibilitam maior precisão e rapidez no acesso às informações contidas nos portais, quando comparado aos portais tradicionalmente estruturados (BERNERS-LEE; HENDLER; LASSILA, 2001; GUO *et al.*, 2018; HAWKE *et al.*, 2013).

Em um estudo sobre o potencial semântico dos portais do conhecimento de universidades brasileiras Guedes (2015) verificou, no entanto, que nenhum dos portais analisados incluía a ontologia em sua estruturação, embora alguns deles utilizassem isoladamente algumas tecnologias semânticas. O modelo semântico de recuperação implica no uso de ontologias de domínio para que o usuário obtenha resultados esperados e adequados às buscas efetuadas nos portais educacionais (BELLANDI *et al.*, 2012; WIMMER; YOON; RADA, 2013).

Diante do exposto, denota-se a importância da avaliação da estrutura dos portais e a verificação se esta contempla as expectativas dos usuários na recuperação da informação. O uso de indicadores qualitativos pode facilitar a avaliação aludida, perpassando as tecnologias e as condições mínimas para que o portal seja um artefato sociotécnico eficiente, melhorando o acesso à informação e a qualidade dos resultados retornados nas buscas.

Indicadores podem ser utilizados para avaliação de determinado problema ou questão, pois, são formados por métricas qualitativas e quantitativas que procuram agregar diferentes dimensões de uma unidade de análise e “expressam em uma única medida diferentes ordens de fenômenos” (ANDRADE; VAITSMAN; FARIAS, 2010, p. 523). Para Van Bellen (2005), os indicadores constituem valores, medidas que sintetizam as características de um sistema. Por isso, sua elaboração parte de certos pressupostos teóricos a partir dos quais são definidas e ponderadas as dimensões da realidade a serem investigadas empiricamente e traduzidas em valores para fins de acompanhamento e comparação (ANDRADE; VAITSMAN; FARIAS, 2010). Assim, indicadores baseados nos preceitos semânticos de recuperação da informação e na estruturação técnica semântica auxiliariam a construção de portais semânticos com busca e recuperação eficiente da informação.

Indicadores qualitativos de avaliação semântica podem auxiliar a tomada de decisão dos gestores na construção de portais, como artefatos sociotécnicos, atendendo tanto às normas e às regras técnicas, quanto aos princípios semânticos de recuperação da informação (BURNS, 2006; THOMAS, 2008). Desta forma, um rol de indicadores qualitativos semânticos poderia ser utilizado como um *check list* para a formatação técnica de portais baseados nos pressupostos da Web Semântica e nos sistemas sociotécnicos alinhados com as diretrizes da recuperação da informação que, sua vez, está alicerçada na otimização da busca e na qualidade dos resultados nos portais (AMATO *et al.*, 2017; BENABDERRAHMANE *et al.* 2017; LI *et al.*, 2017;

MALHOTRA; NAIR, 2015; SILVELLO *et al.*, 2017; SHIRISHA; SUBBARAO; KAVITHA, 2015).

Diante do exposto, questiona-se:

Quais indicadores qualitativos podem ser usados na avaliação do potencial semântico de portais do conhecimento de universidades?

Com base no alinhamento das teorias da Web Semântica, da TAR e da Recuperação da Informação infere-se que existem parâmetros para formulação de indicadores para avaliação do potencial semântico de portais. Sendo assim, levanta-se a tese de que a formulação de indicadores para avaliação semântica dos portais do conhecimento de universidades pode melhorar o acesso à informação e a qualidade dos resultados retornados nas buscas. Uma ferramenta neste escopo, em decorrência, pode ser de grande valia na gestão de portais voltados para ações de criação, compartilhamento e difusão do conhecimento dentro ou fora do âmbito acadêmico (GUEDES, 2015).

Diante do exposto, acredita-se que uma análise semântica e sociotécnica enfatizando a participação dos usuários finais no desenho dos portais de conhecimento de universidades, constitui uma forma de melhorar a qualidade de recuperação da informação. A construção de indicadores com parâmetros semânticos e sociotécnicos pode facilitar estas análises e a construção de portais do conhecimento de universidades de melhor qualidade.

Sendo assim, esta Tese tem como um dos seus pressupostos que um conjunto de indicadores para avaliação do potencial semântico de portais constituirá em uma ferramenta de gestão para construção e implementação de portais voltados para ações de criação, compartilhamento e difusão do conhecimento dentro ou fora do âmbito acadêmico.

1.4 OBJETIVOS

Nesta Seção são apresentados o objetivo geral e específicos.

1.4.1 Objetivo geral

Propor indicadores qualitativos contendo parâmetros para a avaliação do potencial semântico de portais do conhecimento de universidades, de acordo com os conceitos da TAR.

1.4.2 Objetivos específicos

- a. Levantar as condições básicas para o estabelecimento do potencial semântico de portais, como um artefato sociotécnico.
- b. Mapear os atores humanos e não humanos que compõem a rede estrutural de portais semânticos e suas funções.
- c. Determinar indicadores qualitativos, considerando os atores envolvidos, seus interesses, funções e características técnicas.
- d. Validar os indicadores a partir de um levantamento (*survey*) em universidades e das percepções de um painel de especialistas.
- e. Apresentar indicadores qualitativos para avaliação de portais.

1.5 JUSTIFICATIVAS TEÓRICO-PRÁTICAS

Nesta Seção apresentam-se as justificativas teórica e prática deste estudo, com descrição das justificativas conceituais, metodológicas e das lacunas de pesquisa.

1.5.1. Justificativa teórica

Diante do acúmulo de informações na Web torna-se difícil encontrar informações relevantes em *sites* que usam tecnologias tradicionais de recuperação, conforme já tratado anteriormente. A recuperação da informação, portanto, precisa ser rápida e eficiente para que os objetivos do usuário final sejam contemplados (ABDULLAR; ZAMIL, 2018; LEE *et al.*, 2018).

Nesse sentido, as tecnologias semânticas podem ampliar as possibilidades de recuperação da informação contida em portais (BERNERS-LEE; HENDLER; LASSILA, 2001). Isto porque estruturas técnicas semânticas utilizam agentes

inteligentes associados à ontologia de domínio que possibilitam a recuperação de informações pertinentes para o usuário (BELLANDI *et al.*, 2012; BERNERS-LEE; HENDLER; LASSILA, 2001; GAVRILOVA, 2011; GUO *et al.*, 2018; HAWKE *et al.*, 2013; LEE *et al.*, 2018; LI *et al.*, 2017; MALGAONKAR; DEVALE, 2016; SILVELLO *et al.*, 2017; WIMMER; YOON; RADA, 2013).

A manipulação automatizada de conteúdo mediante a interação entre ferramentas semânticas e desenvolvedores tornam os portais semânticos, aumentando a eficiência das buscas e facilitando a recuperação de informações relevantes (BELLANDI *et al.*, 2012; BENABDERRAHMANE *et al.* 2017; LI *et al.*, 2017; SILVELLO *et al.*, 2017; WIMMER; YOON; RADA, 2013). Quando se considera o usuário um ator no processo, o portal torna-se um artefato sociotécnico, construído com base em normas técnicas, mas também em princípios que resguardam interesses de todos os atores envolvidos, independente do seu porte (BURNS, 2006; CALLON, 1986 a, c; HUGHES, 1989; LATOUR, 2001; PEDRO, 2010; THOMAS, 2008).

Destarte, os portais semânticos são, portanto, artefatos de uma rede sociotécnica de interação entre humanos e não humanos que convergem para um objetivo comum que é o de proporcionar o acesso rápido às informações com qualidade no retorno das buscas (BELLANDI *et al.*, 2012; FERNÁNDEZ *et al.*, 2011), abrindo possíveis caixas-pretas³ e criando conhecimentos.

Os portais do conhecimento de universidades semânticos constituem um veículo importante para a construção do conhecimento na medida em que constituem um canal de compartilhamento de informações e oferecem facilidades para que os usuários localizem informações pertinentes e de acordo com suas necessidades (FERNÁNDEZ *et al.*, 2011; LAUSEN *et al.*, 2005; LI *et al.*, 2017; MALGAONKAR; DEVALE, 2016).

Entretanto, segundo a literatura pesquisada, não foi possível identificar um instrumento qualitativo consolidado para indicar a potencialidade semântica de um portal. Resgatando-se, no estudo de Lausen *et al.* (2005) foi proposto um roteiro para avaliação de portais semânticos, porém, os parâmetros arrolados não foram tratados como indicadores qualitativos de potencial semântico. Abriu-se, então, uma lacuna de pesquisa, pois ao combinar os dados coletados mediante a aplicação de

³ Caixa-preta representa algo acabado e estabilizado (CALLON, 1986) e, no contexto desta Tese, a caixa-preta é representada pelas tecnologias convencionais usadas nos de portais.

questionários com a literatura revisada foi possível encontrar respostas para esta questão.

Para que ocorra a recuperação semântica do conteúdo de um portal do conhecimento de universidades, infere-se que este deva incluir recursos e materiais estruturados mediante diferentes interações homem-máquina, de acordo com os preceitos da TAR (CALLON, 1986 a, b, c; GAO 2005; SHIM; SHIN, 2016). O mapeamento da rede de atores envolvidos na construção de um portal do conhecimento de universidade semântico possibilita compreender alguns dos procedimentos necessários para disponibilizar uma interface de pesquisa que tem por finalidade promover o acesso a informações acadêmicas com rapidez e precisão (BELLANDI *et al.*, 2012; BERNERS-LEE; HENDLER; LASSILA, 2001; LAUSEN *et al.*, 2005). Nesse cenário, os preceitos da TAR, postulados por Latour (2001), propiciam um caminho para a descrição de problemas e a identificação de possíveis lacunas, fornecendo subsídios que apontam soluções fundadas nas relações entre os atores humanos e não-humanos da rede de portais, a fim de que estes ofereçam qualidade na recuperação de informações.

A semântica é uma alternativa para melhorar a precisão da busca nos portais, já que possibilita compreender o significado contextual de termos digitados nos campos de pesquisa. O uso de ontologia faz parte de abordagens para metodologias de pesquisa semântica. Neste tipo de pesquisa um motor de busca possibilita a interpretação das relações hierárquicas dos termos e também as relações complexas que ocorrem entre os termos definidos dentro de ontologias (PAIVA; COSTA; SILVA, 2014).

Considerando que as técnicas tradicionais de recuperação da informação tais como a busca por palavra-chave, são ineficazes para tratar o exponencial volume de dados disponíveis na Web, o desafio consiste em projetar e desenvolver sistemas complexos e facetados por meio de metodologias que possibilitem mensurar o desempenho alcançado, avaliar o impacto de alternativas e de novas ideias e garantir os níveis de eficácia necessários para atender às expectativas dos usuários por informações relevantes (SILVELLO *et al.*, 2017). Os portais do conhecimento de universidades inserem-se neste cenário por arrolarem um grande volume de dados abertos e mesmo envolvendo um domínio específico, as informações disponibilizadas são de interesse da população tornando-os aparatos sociotécnicos (BURNS, 2006).

Os sistemas sociotécnicos além de serem compostos por normativas técnicas também devem seguir regras e preceitos como os princípios organizadores sociais destinados a regular as atividades dos atores que operam e gerenciam a tecnologia (BURNS, 2006; CALLON, 1986b; LATOUR, 2001). Para que um portal seja considerado um artefato sociotécnico faz-se necessário que o usuário final esteja envolvido em sua construção, tornando a busca mais eficiente.

Assim, a análise da rede de atores que compõe um portal universitário, partindo do pressuposto que este transcende sua materialidade, pode contribuir na descrição de problemas e na identificação de lacunas ou controvérsias para alinhar os atores no sentido de seguirem um mesmo objetivo (CALLON, 1986c; LATOUR, 2012; LAW, 2010), que no caso dos portais é oferecer qualidade na recuperação de informações.

Para testar a validade desta afirmativa e aprofundar estudos na área, um levantamento bibliométrico, que é primeira fase da Revisão Sistemática da Literatura, foi realizado com o intuito de verificar a ocorrência na literatura de estudos que abordassem TAR, portais do conhecimento de universidades, recuperação semântica da informação e construção de indicadores.

O uso da TAR como metodologia adicional, favorece a compreensão do envolvimento entre ciência e a tecnologia, a partir das relações sociais que permeiam os ambientes de criação e de inovação científica (CALLON, 1986a; LATOUR, 2012; LAW, 2010). Tal concepção alinha-se com os estudos do uso e do impacto das tecnologias na sociedade, já que os artefatos técnicos são socialmente construídos (FEENBERG, 1999; MARCUSE, 1998). Assim, é possível verificar se a criação de um portal foi efetuada como planejado, identificando também possíveis falhas e eventuais lacunas deixadas que podem afetar o desempenho dos portais semânticos. Desta forma, harmoniza-se a convivência do homem com os objetos técnicos que afetam o modo de vida na sociedade contemporânea (LATOUR, 1994a).

As bases de dados eleitas para a pesquisa foram: Web of Science - Coleção Principal, Library, Proquest (Information Science & Technology Abstracts, Computer and Information Systems Abstracts, Applied Social Sciences Index & Abstracts; Technology Collection, Scopus e SciELO Citation Index (Web of Science). Essas bases incorporam fontes de pesquisa que fornecem cobertura para o tema e os subtemas da pesquisa.

As palavras-chave utilizadas na busca foram: ontologia, teoria ator-rede, recuperação semântica da informação, portal universitário e potencial semântico.

Esses termos foram traduzidos para a língua inglesa com o intuito de se pesquisar em bases internacionais e recuperar uma quantidade de títulos adequada para estudo. Foi utilizada a seguinte estratégia de busca em todas as bases de dados selecionadas para estudo: (((*"information retrieval"* AND (*"Web portals"* OR *"university portal"* OR *"academic portal"* OR *"educational portal"* OR *"semantic portal"* AND (*"semantic potential"* OR *"semantic evaluation"*)) AND *Ontology* AND *"actor-network theory"*))). A pesquisa teve como espaço temporal os últimos 5 anos.

O levantamento bibliométrico inicial com os termos de pesquisa isoladamente recuperou um número significativo de documentos que justifica pertinência e atualidade do tema (Tabela 1). Exemplificando-se, a partir da expressão de busca “recuperação da informação” foram arrolados 3.064 itens; o termo “ontologia” foi associado a 193.278 itens; para “portais universitários” e suas variantes foram recuperados 478 itens; a expressão “*semantic potential*” recuperou 216 itens e a expressão de busca “Teoria Ator-Rede” recuperou 9.607 mil itens. No total recuperou-se mais de 200 mil documentos associados às palavras chaves ou expressões de busca utilizados nas bases de dados arroladas.

Tabela 1 - Busca por palavras-chave isolada por bases de dados.

DESCRITORES	QUANTIDADE DE ARTIGOS POR BASE			
	Web of Science	Scielo	Proquest	Scopus
<i>"retrieval information"</i>	71	4	2.446	543
<i>Ontolog*</i>	33.343	761	11.116	148.058
<i>"university portals"</i> OR <i>"academic portals"</i> OR <i>"educational portals"</i>	34	2	208	234
<i>"semantic potential"</i>	61	3	51	101
<i>"Actor-network Theory"</i>	1.201	75	4412	3.919
Total	34.710	845	18.229	152.855

Fonte: Autoria própria (2021).

Por outro lado, não houve resultado positivo para a estratégia de busca combinada com todas as palavras-chave, indicando potenciais espaços de pesquisa, quando do enfoque pretendido. De outra parte, os resultados obtidos em combinações com alguns grupos isolados de palavras-chave, a exemplo da estratégia *"retrieval information AND ontology"* que recuperou 218 itens de informação e que serviram de subsídios para o desenvolvimento da pesquisa, são significativos. Foram recuperados na segunda fase da busca, pós Qualificação, 1.002 itens, incluindo artigos,

documentos de eventos e conferências para análise de relevância frente ao tema em estudo (Tabela 2). Adicionalmente, uma pesquisa com os termos de busca efetuada na Biblioteca Digital de Teses e Dissertações (BDTD) e no Catálogo de Teses e Dissertações (CAPES) não retornou nenhum estudo, quando da busca combinada.

Tabela 2 - Resultados das estratégias de busca em bases de dados.

ESTRATÉGIAS DE BUSCA	QUANTIDADE DE ARTIGOS POR BASE			
	Web of Science	Scielo	Proquest	Scopus
<i>("retrieval information") AND "actor-network theory"</i>	0	0	2	0
<i>("retrieval information") AND ontology</i>	14	0	139	65
<i>("retrieval information" AND ("university portals" OR "academic portals" OR "educational portals"))</i>	0	0	1	0
<i>("retrieval information" AND ("semantic potential" OR "semantic evaluation"))</i>	4	0	9	0
<i>("semantic potential" OR "semantic evaluation") AND Ontology</i>	8	0	100	10
<i>("university portals" OR "academic portals" OR "educational portals") AND ontology</i>	1	0	4	22
<i>ontology AND "actor-network theory"</i>	181	7	314	113
TOTAL	218	7	567	210

Fonte: Autoria própria (2021).

O levantamento bibliométrico, efetuado com o propósito de verificar a existência de indicadores para avaliação de potencial de portais, utilizando-se a expressão de busca *"semantic index" AND "evaluation" AND "portals"* não retornou resultados nas bases Scielo, Web of Science e Scopus.

Na base Proquest foram encontrados 12 resultados, entretanto, verificou-se que apenas alguns termos da busca apareciam nos resumos das obras. Isso foi comprovado com a substituição do operador booleano *AND* pelo *NEAR/5* que aproxima os termos da expressão de busca e evidencia a relação entre os termos em um mesmo documento. O operador *NEAR* é usado para consultas de conteúdo em que a proximidade de termos identifica a relação entre eles, isto é, um documento corresponderá à consulta se as palavras estiverem a uma distância de até 50 palavras uma da outra no documento (BIBLIOTECA REGIONAL DE MEDICINA - BIREME, 2022). Quanto mais próximas as palavras estiverem, mais elevada será a classificação atribuída ao documento no conjunto de resultado (WEB OF SCIENCE, 2020). Assim, apesar de se utilizar *NEAR/5*, para aproximar os termos de busca a uma

distância de cinco palavras, não foi recuperado nenhum documento. Cabe ressaltar que na atualização da pesquisa bibliométrica para o período de 2021/2022, não foi possível acessar os dados da Proquest porque esta base não está mais disponível no catálogo do Portal de Periódicos da Capes.

Diante desse resultado, comprovou-se a inexistência de indicadores formatados para avaliação semântica de portais, nas bases pesquisadas, o que torna esta pesquisa aderente às necessidades dos gestores dos portais do conhecimento das universidades e também para cobrir a lacuna de pesquisa existente nesse campo de conhecimento.

1.5.2. Justificativa prática

As justificativas práticas desta pesquisa perpassaram pela importância do estudo para a sociedade, e para Programa de Pós-Graduação em Tecnologia e Sociedade – PPGTE, quer seja em termos da linha de estudo, do grupo de estudo relacionado, mas também aos interesses da orientadora e desta pesquisadora.

A abordagem envolvendo indicadores semânticos e portais do conhecimento de universidades constitui uma forma de fortalecer os pressupostos do campo de Ciência, Tecnologia e Sociedade – CTS, evidenciando o envolvimento de atores humanos e não humanos na construção de artefatos técnicos. Na prática, os indicadores qualitativos, objeto deste estudo, podem ser utilizados na avaliação semântica de portais, consistindo assim em um instrumento para a melhoria destes artefatos tecnológicos.

Em relação ao PPGTE, este estudo enfatizou a construção social dos artefatos tecnológicos, especialmente os artefatos sociotécnicos, visando destacar a importância da participação dos usuários ou consumidores finais no planejamento e na construção de tecnologias já que estas são projetadas para o uso da sociedade (MARTÍNEZ-GARCÍA; HERNÁNDEZ-LEMUS, 2013; WALSHAM,1997). O estudo articulou-se com a linha de pesquisa de Tecnologia e Desenvolvimento, com ênfase no tema Redes Sociotécnicas que integram a esfera de ação de produção do conhecimento que visa destacar a tecnologia como uma construção social (PROGRAMA DE PÓS-GRADUAÇÃO EM TECNOLOGIA E SOCIEDADE, 2021).

A pesquisa também tem o propósito de contribuir para os estudos do Grupo de Pesquisa da Conhecimento e Sustentabilidade Organizacional do PPGTE, com

especial atenção à territorialidade e à sustentabilidade que no caso desta pesquisa envolve os portais do conhecimento de universidades como um território sustentável de estudos e aplicações tecnológicas em prol do acesso facilitado à informação e à criação de conhecimentos. Adicionalmente, estes portais associados às tecnologias semânticas, inserem-se na modalidade de Redes Sociotécnicas cuja representação gráfica conecta nós ou segmentos que viabilizam o fluxo de informações (ALBAGLI, 2004).

Para Marcuse (1998), a tecnologia pode ser utilizada beneficentemente como meio de transformação na medida em que propicie a democratização de funções e facilite o desenvolvimento humano em todos os ramos do trabalho, diminuindo o tempo e o esforço dos indivíduos em suas variadas atividades, tais como a recuperação de informações na Web e a estruturação de portais do conhecimento de universidades, por exemplo.

Bijker, Hughes e Pinch (1989) também destacaram que quando a tecnologia é situada no campo social, poderá contribuir nos processos de inovação e de transformação tecnológica, com vantagens para todas as nações e as diferentes camadas da sociedade.

Os sistemas sociotécnicos são também caracterizados por incluírem objetos técnicos complexos que interagem e são regulados no campo social. Dependem de estruturas administrativas, investimentos em ciência e tecnologia e são temporariamente estáveis (BURNS, 2006). A análise de um sistema sociotécnico envolve o estudo tanto do conjunto de normas técnicas, como dos princípios organizadores sociais, já citados anteriormente, destinados a regular as atividades dos atores que operam e gerenciam a tecnologia (BURNS, 2006). Deste modo, os portais do conhecimento de universidades enquanto artefatos tecnológicos fazem parte de uma rede sociotécnica, já que são construídos por dispositivos complexos de interação social (BURNS, 2006).

Segundo Callon (1986c) e Latour (2001), a TAR descreve os artefatos tecnológicos nos contextos sociais e naturais em que atuam e nesta concepção, tais objetos são socialmente construídos e inseridos em uma rede sociotécnica. Esta rede é composta por conjuntos de atores heterogêneos e elementos humanos e não humanos que interagem no processo de concepção e de produção e que dão origem a definições tecnológicas obtidas no processo de solução de problemas tais como a

efetiva recuperação de informações em portais, tema que se conecta com os interesses da orientação deste estudo.

As redes propiciam interações e conexões que intensificam as relações sociais conectando localidades distantes, e assim, os portais do conhecimento de universidades aliados às tecnologias de rede mundial, possibilitam o compartilhamento de informações de domínio acadêmico para comunidades locais, nacionais e internacionais (ALBAGLI, 2004). Deste modo, os portais formam uma rede, um enraizamento de diferentes laços conforme também conceituou Granovetter (1983).

Ainda no contexto da territorialidade, nesta pesquisa, a Web Semântica é o ambiente que define as relações ou condições para que os portais se tornem semânticos. Desta forma, a semântica configura um elemento que envolve os portais em um entrelaçamento essencial de pontos de aprendizagem (MIOR, 2006) e de pessoas que praticam sua territorialidade (SACK, 1986).

Desta maneira, acredita-se que a elaboração de indicadores qualitativos para avaliar o potencial semântico de um portal é uma estratégia válida para melhorar a recuperação de informação e que contribui no sentido de oferecer uma ferramenta prática para a gestão de portais semânticos, favorecendo o aprendizado, o compartilhamento e a aplicação de conhecimento específico e interdisciplinar.

1.6 MARCO TEÓRICO

Neste estudo foi proposta a construção de indicadores qualitativos para avaliação do potencial semântico de portais de universidades, na modalidade portais do conhecimento. Para explicitar os fundamentos da Web Semântica e de portais semânticos apoiou-se nos estudos de Berners-Lee, Hendler e Lassila (2001), Koivunen e Miller (2001) e Maedche e Staab (2001). Para tratar da ontologia formal esta pesquisa se fundamentou nos estudos de Gruber (1996) e Guarino (1997).

As características dos portais semânticos foram extraídas dos estudos de Reynolds, Shabajee e Cayzer (2004), Lausen *et al.* (2005), Correa (2012) e de Guedes (2015) que elencou as condições básicas necessárias para o estabelecimento do potencial semântico de portais mediante a análise da Gestão da Informação, Gestão de Ontologia e das tecnologias utilizadas nos portais.

Na área de estudos sobre TAR e sistemas sociotécnicos, os estudos de Latour (2001) Law (1992), Callon (1986 a, b, c), Burns, 2006 e Thomas (2008) formaram as bases teóricas desse tema nesta pesquisa. Ainda, o embasamento metodológico foi alicerçado nas diretivas de Bardin (2011), Creswell e Clark (2013), Sampieri, Collado e Lucio (2010), Gil (2010) e Marconi e Lakatos (2010 a, b).

1.7 ABORDAGEM METODOLÓGICA DA PESQUISA

Esta pesquisa caracterizou-se como uma descrição e interpretação de fenômenos que ocorrem na vida social, incluindo o desenho de artefatos, cujos significados podem ser compartilhados – uma pesquisa de paradigma interpretativo. A pesquisa interpretativa conecta o mundo natural e social em que as pessoas vivem e o pesquisador interpretativo é o próprio instrumento de pesquisa, auxiliado por ferramentas e técnicas que permitem a construção de um caso de estudo (MOREIRA; CALEFE, 2006). O paradigma interpretativo, ressalte-se, baseia-se na reflexão, na objetividade e na imparcialidade do investigador ao analisar o contexto que explica os processos sociais abordados. Assim, dados quantitativos podem ser utilizados, mas sempre contextualizados pelos detalhes da análise qualitativa (MOREIRA; CALEFE, 2006).

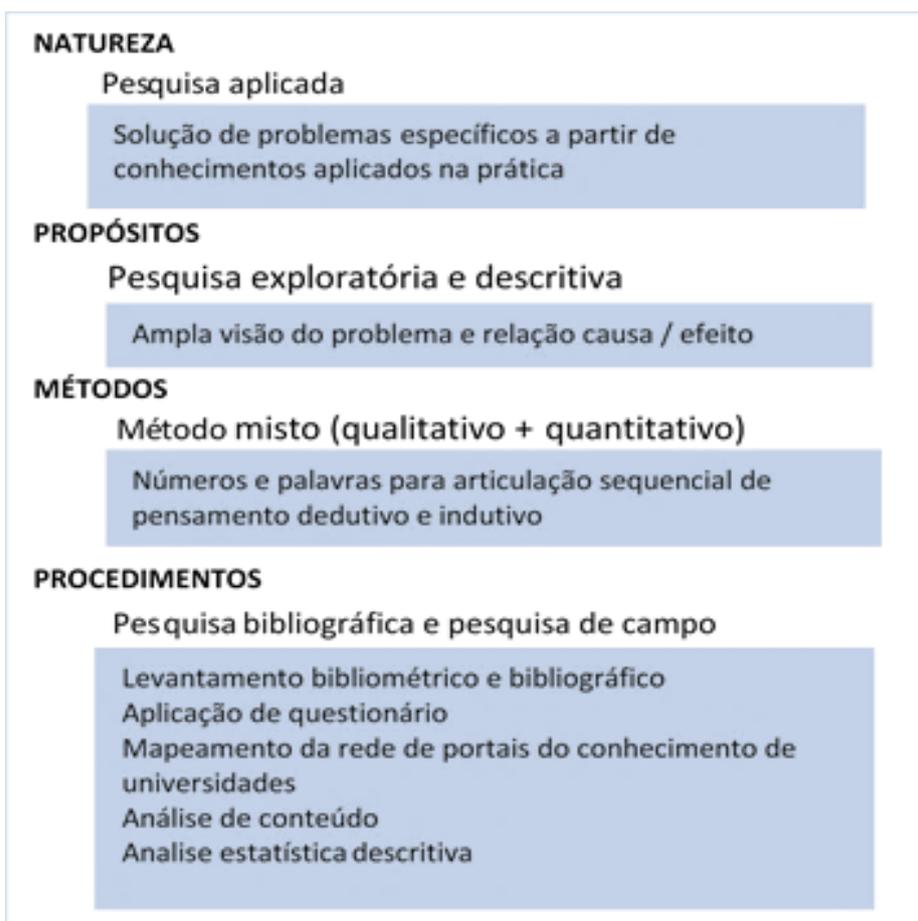
Quanto à natureza a pesquisa foi considerada do tipo aplicada, pois trata de problemas específicos que ocorrem no cotidiano, sendo possível utilizar os conhecimentos gerados para a solução de problemas práticos da vida social (GIL, 2010; MARCONI; LAKATOS, 2010b). Com relação aos propósitos da pesquisadora, a pesquisa pode ser vista como exploratória e explicativa correlacional, pois elucidam e quantificam as relações entre variáveis e, de forma estruturada, geram um sentido de entendimento sobre o problema (SAMPIERI; COLLADO; LUCIO, 2010).

Adicionalmente, baseando-se em Triviños (1987), esta pesquisa teve caráter crítico-descritivo, pois busca conhecer a realidade social, a partir de percepções, reflexões e intuição. A síntese da consecução da pesquisa que ocorreu de acordo com os procedimentos e classificações metodológicas estão indicadas na Figura 1.

Do ponto de vista do método de tratamento dos dados a pesquisa foi considerada mista, de acordo com os postulados de Creswell e Clark (2013), Sampieri, Collado e Lucio (2010), pois, envolve dados qualitativos e quantitativos.

Nesse tipo de pesquisa é possível utilizar números e palavras para articulação sequencial de pensamento dedutivo e indutivo com o uso simultâneo de ambos os métodos, com igual ênfase no estudo. No entanto, nesta pesquisa, salienta-se, privilegiou-se a análise qualitativa.

Figura 1 - Classificação metodológica da pesquisa.



Fonte: Autoria própria, fundamentada em Gil (2010), Marconi e Lakatos (2010) e Creswell e Clark (2013).

A coleta de dados ocorreu em momentos distintos: levantamento bibliométrico, revisão da literatura; *survey* (questionário eletrônico) respondido por técnicos envolvidos na estruturação dos portais de universidades, com descrição aprofundada no Capítulo 5.

A pesquisa da literatura foi efetuada mediante consulta em bases de dados científicas, abordando os temas em estudo disponíveis no Portal de Periódicos da Capes, abrangendo um período de cinco anos retrospectivamente, conforme já especificado. A análise de conteúdo foi efetuada conforme os preceitos de Bardin (2011) e incluindo artigos e documentos científicos obtidos no levantamento

bibliométrico. A análise estatística descritiva, foi efetuada mediante o uso do Microsoft Excel e para a análise bibliográfica utilizou-se o gerenciador bibliográfico EndNote Basic.

A TAR, como metodologia, foi usada neste estudo para desenhar a rede de portais a partir da análise da cadeia de portais do conhecimento de universidades selecionados e da literatura. O desenho da rede possibilitará seguir os atores e apontar a dinâmica social da ciência, a partir de uma perspectiva que subtraia a neutralidade e a dicotomia entre tecnologia e sociedade. A partir desta abordagem foi possível verificar se os portais cumprem seus objetivos de forma adequada, tanto do ponto de vista de sua estruturação, quanto do interesse dos usuários (CALLON; LAW; RIP, 1986).

1.8 REVISÃO DE LITERATURA

A partir da Revisão de Literatura foi possível identificar tendências, reafirmar argumentos, identificar controvérsias e lacunas de pesquisa (MOREIRA; CALEFFE, 2006) e, assim, embasar a investigação em curso.

Os portais corporativos são considerados instrumentos de Gestão do Conhecimento que facilitam a disseminação e a recuperação de informações, propiciando a formação de redes de conhecimento que integram elementos tácitos e explícitos que se traduzem em vantagens competitivas para as instituições (RAUTENBERG; TODESCO; STEIL, 2010; HAUKE; OWOC; PONDEL, 2015).

Um portal semântico é baseado nos preceitos da Web Semântica que, por sua vez, inclui tecnologias específicas e ontologia de domínio em sua estruturação (BERNERS-LEE; HENDLER; LASSILA, 2001; GRUBER, 1996; GUARINO, 1997; STANESCU, 2018; LAWAN; RAKIB, 2019; HYVÖNEN *et al.*, 2021). Com o uso dessa estrutura é possível desenvolver um sistema de recuperação da informação baseado em regras que modulam a combinação entre os termos, evitando a polissemia e a sinonímia que estão associadas a deficiências no resultado das buscas na Web tornando-as mais rápidas e efetivas (GRUBER, 1996; LEE *et al.*, 2018; GAVRILOVA, 2011; GUARINO, 1997; FERNÁNDEZ *et al.*, 2011; MAEDCHE; STAAB, 2001; MAATOUK, 2021).

Para avaliar os portais foi preciso compreender o desenho estrutural do artefato técnico e como ocorre o processo de construção do mesmo, sua dinâmica e as inter-

relações dos atores envolvidos na sua criação (BURNS, 2006; THOMAS, 2008). Nesse âmbito, a TAR propõe a análise de artefatos dentro do contexto social e natural em que atuam, considerando a função de cada elemento envolvido. Assim, em uma concepção sociotécnica, além do conjunto de normas técnicas, a construção dos artefatos deve considerar também as regras que envolvem arranjos culturais e institucionais inerentes aos sistemas sociotécnicos em que a tecnologia está inserida (BURNS, 2006; FORLANO, 2017; IYAMU; MGUDLWA, 2018; THOMAS, 2008).

Diferentemente dos métodos científicos tradicionais que consideram apenas os elementos de uma sociedade circulante em que o poder científico se desenvolve, a TAR enfatiza a relação entre ciência, tecnologia e sociedade de forma não linear, mas sim com destaque para as conexões e as articulações entre os atores envolvidos (CALLON, 1986c; LATOUR, 1994a).

Na abordagem metodológica foram utilizados métodos mistos, com análise de dados qualitativos e quantitativos (GIL, 2010; MARCONI; LAKATOS, 2010a; SAMPIERI; COLADO; LÚCIO, 2010). De acordo com Creswell e Clark (2013), e Sampieri, Colado e Lúcio (2010), este tipo de pesquisa possibilita analisar tanto números quanto palavras para articular o pensamento dedutivo e indutivo de forma sequencial, ampliando-se a exploração dos dados obtidos das duas abordagens.

1.9 ESTRUTURA DO TRABALHO

O trabalho foi composto por sete capítulos principais descritos sequencialmente. Neste Capítulo 1 apresentou-se a Introdução, incluindo o tema, a delimitação e o problema da pesquisa, os objetivos, a justificativa, o marco teórico e a abordagem metodológica.

No Capítulo 2 serão apresentados os conceitos e os modelos de recuperação semântica da informação, com destaque para os modelos ontológicos. No Capítulo 3 caracteriza-se os portais semânticos incluindo o conceito, a arquitetura e as tecnologias semânticas, com exemplos de portais que apresentam estruturação e pesquisa semântica. Discorre-se ainda sobre a classificação e os objetivos dos indicadores qualitativos como elementos norteadores para avaliação do potencial semânticos dos portais.

O alinhamento conceitual será apresentado no Capítulo 4, conciliando as teorias identificadas e associadas aos sistemas sociotécnicos e à Teoria Ator-Rede com os objetivos da pesquisa.

Os procedimentos metodológicos adotados, partindo-se da problemática e dos objetivos, serão descritos no Capítulo 5, incluindo o desenvolvimento da pesquisa, o método de revisão da literatura, a forma de coleta e de tratamento dos dados. A análise e a discussão de resultados, a partir da literatura, da coleta de dados e da formulação de indicadores qualitativos semânticos serão apresentados no Capítulo 6.

Por fim, no Capítulo 7 serão apresentadas as considerações finais acerca do estudo realizado para propor indicadores qualitativos semânticos para avaliação de portais do conhecimento, as limitações e as sugestões para futuras pesquisas.

2 INFORMAÇÃO: das formas de recuperação aos modelos de representação

Neste Capítulo realiza-se a revisão da literatura abordando a necessidade de se incorporar técnicas de recuperação de informação que representem o conhecimento a partir de regras, fatos ou uma classificação hierárquica dos objetos. Inicialmente, serão abordados os métodos tradicionais de recuperação da informação e a ontologia como modelo de recuperação da informação.

2.1 MODELOS DE RECUPERAÇÃO DA INFORMAÇÃO

A produção científica aumentou exponencialmente nos últimos anos, tornando-se difícil lidar com grande quantidade de informações projetadas não estruturadas e obter informações específicas pode levar muito tempo (FERNÁNDEZ *et al.*, 2011; GUO *et al.*, 2018; LEE *et al.*, 2018; RODRIGUES, 2020; SILVELLO *et al.*, 2017). De outra parte, segundo Maniega-Legarda, Pérez-Salmerón e Guerrero-Torres (2008), o usuário deveria poder consultar qualquer tipo de informação disponibilizada sem precisar ter conhecimento específico de técnicas de busca ou alto conhecimento tecnológico.

Os sistemas de recuperação de informação constituem uma alternativa para resolver o problema, no entanto não fornecem a solução perfeita (ABDULLAR; ZAMIL, 2018). Os sistemas tradicionais de recuperação de informação não são eficazes na obtenção de documentos com alta probabilidade de conter respostas, e a recuperação de maior quantidade de documentos só aumenta o número de documentos irrelevantes (LEE *et al.*, 2018). Assim, a pesquisa em sistemas baseados na correspondência com palavras-chave simples é incapaz de fornecer resultados precisos de acordo com as necessidades de informações do usuário (RAZA *et al.*, 2018).

Para Ngo e Cao (2018) os sistemas tradicionais de recuperação de informações representam documentos e consultas por conjuntos de palavra-chave. Entretanto, o conteúdo de um documento ou de uma consulta é definido principalmente por palavras-chave e entidades nomeadas que ocorrem nele. Entidades nomeadas possuem características ontológicas, a saber, os *aliases* – os apelidos (diferentes nomes atribuídos a uma entidade), classes e identificadores, que estão ocultos de sua aparência textual. Além disso, o significado de uma consulta pode implicar entidades

nomeadas latentes relacionadas às palavras aparentes na consulta. Os sistemas tradicionais não incorporam essas características ontológicas por não utilizarem ontologias formais em sua estruturação. Tais características são possíveis com o uso da ontologia formal incorporada na estruturação técnica dos sistemas de busca, o que possibilita a recuperação semântica da informação (TELNOV, 2015; GUO *et al.*, 2018; BARBA-GONZÁLEZ *et al.*, 2019).

A ontologia formal é a “especificação de uma conceituação” (GRUBER, 1993, p.199) que pode ser usada em sistema de recuperação da informação para especificar os termos de busca, relacionando as propriedades, as funções, as restrições e os axiomas destes, de modo que se recupere o conteúdo dos documentos com mais precisão (GRUBER, 1993; GUARINO, 1996).

Um sistema tradicional de recuperação de informação, segundo Raza *et al.* (2018), não considera os relacionamentos semânticos entre as palavras-chave de consulta do usuário, portanto, não podem interpretar o contexto da consulta com precisão. A técnica de expansão de consulta aprimora o desempenho do processo de recuperação de informações, expandindo a consulta do usuário com conceitos significativos relevantes (RAZA *et al.*, 2018).

Aksoy (2016) destaca que a pesquisa por palavra-chave é uma técnica popular para pesquisar dados na Web porque libera o usuário de aprender a estrutura das fontes de dados e de uma linguagem de consulta complexa. No entanto, a conveniência da pesquisa por palavras-chave vem acompanhada de desvantagens que incluem a imprecisão das consultas que geralmente resulta em um grande número de resultados, dos quais apenas poucos são relevantes. Outra desvantagem da pesquisa de palavras-chave é a incapacidade de expressar precisamente a intenção do usuário, já que a resposta da consulta pode conter diferentes tipos de resultados significativos mesmo que o usuário esteja interessado em apenas alguns deles. Diante do grande volume de dados disponíveis e a importância da pesquisa faz-se necessário encontrar estratégias metodológicas para compreender visões mais abrangentes, ou macro visões, das principais tendências de pesquisa desenvolvidas em específico.

A ontologia representa a semântica de um domínio portanto, e surge como um recurso importante para a expansão da consulta, já que viabiliza a incorporação de técnicas de recuperação de informação baseadas em regras e fatos em classificação hierárquica de assuntos. As técnicas controlam a validade e a precisão e, com o

auxílio da inteligência artificial, essas técnicas possibilitam maior qualidade da informação recuperada (MALHOTRA; NAIR, 2015). Na Seção seguinte serão tratados o modelo booleano, o vetorial e o probabilístico, considerados mais tradicionais e, ainda, alguns modelos contemporâneos de recuperação da informação.

2.1.1 Modelos tradicionais de recuperação da informação

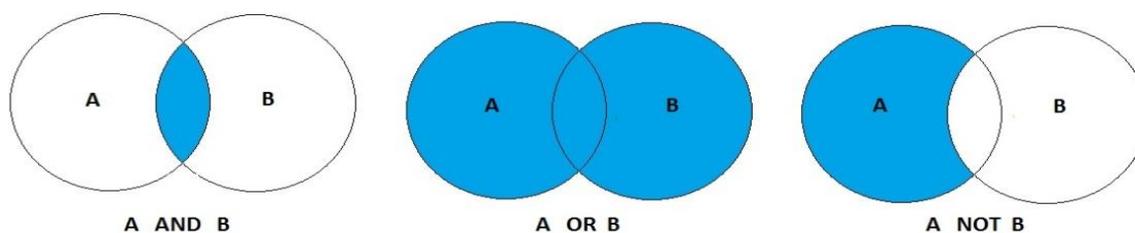
O modelo booleano é baseado no uso de caracteres especiais em estratégias de busca que fazem a função de conectores de palavras-chave na tentativa de melhorar a relevância dos resultados de pesquisas em portais da Web (FERNEDA, 2003; GAVRILOVA, 2011).

O modelo booleano de recuperação da informação foi proposto por Salton em 1968, baseado na teoria dos conjuntos e assume que os documentos são representados por um conjunto de termos de índice (ou seja, $t \in d$, se o documento é indexado pelo termo t) (SALTON, 1968). As consultas são combinações lógicas de termos usando os operadores booleanos AND (\wedge), OR (\vee) e NOT (\neg). Formalmente, as consultas são construídas no conjunto de termos de índice $\{t_i, i = 1, 2, 3, \dots\}$ (ZERARGA; DJOUADI, 2018, p.15).

O uso da lógica formal na recuperação da informação é principalmente devido ao seu poder de modelar o processo de recuperação, que é interpretado como inferência. Para uma determinada consulta, o processo de recuperação consiste em atribuir um valor booleano (0 ou 1) a cada documento. Assim, um documento é recuperado se os termos de índice no documento satisfizerem a expressão lógica na consulta especificada (ZERARGA; DJOUADI, 2018).

O uso dos operadores possibilita três condições básicas de relacionamentos de pesquisa (BIBLIOTECA REGIONAL DE MEDICINA - BIREME, 2022), conforme Figura 2: 1- AND - Intersecção de A e B: as palavras selecionadas para busca devem estar presentes em todos os artigos; $A \cap B$. 2- OR - União de A e B: o interesse está em A ou B: são palavras relacionadas ou que estão em outros idiomas. Entretanto, não é necessário que todas as palavras estejam presentes em um mesmo artigo). Na teoria dos conjuntos tem-se que $A \cup B$; 3- NOT - o interesse está no A. Documentos que possuam o B são excluídos da pesquisa.

Figura 2 - Relacionamento de pesquisa na lógica booleana.



Fonte: Adaptada de BIREME (2022).

Adicionalmente, operadores de proximidade, usado entre termos de busca tais como *WITH*, *SAME* e *NEAR/x* e caracteres especiais de truncagem, usados para omitir ou cortar parte de textos, a exemplo do asterisco (*) e aspas (“ ”) podem ser utilizados em pesquisa na Web para se obter resultados mais precisos sobre o tema pesquisado (FERNEDA, 2003; LOPES, 2002;).

As vantagens do modelo booleano incluem expressividade completa – se o usuário souber exatamente o que quer –, facilidade de programação e exatidão (SILVA; SANTOS; FERNEDA, 2013). Entretanto, o modelo booleano apresenta duas desvantagens principais, ou seja, nenhuma capacidade de classificação de resultados e nenhum índice de ponderação de termos. Conseqüentemente, o modelo booleano clássico tende a produzir resultados que não se percebem facilmente por intuição, devido a essa característica de falta de meio termo – tudo ou nada. Essas dificuldades se traduzem na recuperação parcial do conhecimento, retorno nulo de resultados e resultados com pouco ou muitos documentos. Além disso, não é possível atribuir pesos diferenciados aos termos, assumindo-se que todos têm a mesma importância (BAEZA-YATES; RIBEIRO-NETO, 1999; FERNÁNDEZ *et al.*, 2011).

Já, o modelo vetorial propõe a atribuição de pesos aos termos de indexação e termos de busca. Esses pesos são utilizados para calcular o grau de similaridade entre a expressão de busca formulada pelo usuário e cada um dos documentos do repositório. Este modelo é a base da maioria de sistemas de recuperação de informações na Web para determinar uma lista de documentos relevantes (FERNÁNDEZ *et al.*, 2011; SILVA; SANTOS; FERNEDA, 2013).

O modelo vetorial apresenta limitações que incluem a ausência de relacionamento entre os termos de busca e a dificuldade de utilização em conjunto com o modelo booleano, o que torna seu uso restrito (BAEZA-YATES; RIBEIRO-

NETO, 1999), embora tenha sido aplicado em motores de busca de primeira geração no ambiente da Web (PINHEIRO, 2004).

Ngo e Cao (2018) propuseram um modelo vetorial generalizado baseado em ontologias para a busca de texto semântico. O objetivo foi explorar recursos ontológicos de entidades nomeadas e latentes para revelar a semântica⁴ de documentos e consultas. O modelo incluiu também uma estrutura para combinar ontologias diferentes para obter vantagens complementares para anotação semântica e pesquisa. Esses autores concluíram, a partir de experiências em um conjunto de dados de referência, que o modelo apresentou boa qualidade no retorno de pesquisas.

Ainda, o modelo probabilístico é usado em sistemas de recuperação de informação com o objetivo de saber a probabilidade de um documento ser ou não relevante para uma consulta. Este modelo considera pesos binários que representam a presença ou a ausência de termos para descrever documentos. Com base no cálculo da probabilidade, um documento pode ser relevante para uma consulta (PINHEIRO, 2004).

Apesar do princípio probabilístico de ordenação se constituir uma vantagem, o método é considerado altamente complexo, o que se traduz em desvantagens quando comparados com os demais métodos tradicionais (BAEZA-YATES; RIBEIRO-NETO, 1999).

2.1.2 Modelos contemporâneos de recuperação da informação

Os modelos tradicionais de recuperação de informações recuperam documentos com base na correspondência baseada em palavras-chave (NGO, CAO, 2018) e são incapazes de fornecer resultados de acordo com as necessidades de informações do usuário, diante do crescimento descoordenado da informação na Web (RAZA *et al.* 2018). Abdullar e Zamil (2018) destacaram a dificuldade de se lidar com grande quantidade de informações não estruturadas, pois é uma tarefa que demanda muito tempo, sendo também uma missão difícil obter informações específicas.

⁴ Semântica – relativo ao significado das palavras e interpretação de frases ou enunciados, possibilitando a recuperação mais precisa da informação portais (PAIVA; COSTA; SILVA, 2014)

De outra parte, a produção científica aumentou exponencialmente nos últimos anos tornando-se necessário encontrar estratégias metodológicas para compreender visões abrangentes – ou macrovisão – das principais tendências da pesquisa desenvolvidas em campos específicos (GUO *et al.*, 2018; RODRIGUES, 2020). Diante disso, novos métodos ou sistemas combinando métodos, têm sido propostos para melhorar os resultados das buscas dentre estes, a lógica *fuzzy*, a mineração de dados, as redes neurais e a pesquisa baseada em modelo conceitual e descrição semântica (SILVA; SANTOS; FERNEDA, 2013; ZERARGA-; DJOUADI, 2018).

A lógica *fuzzy* (difusa ou nebulosa) advém da teoria dos conjuntos, assim como a booleana (SHAW; SIMÕES, 1999). O objetivo desta lógica é capturar e operar com a diversidade, a incerteza e as verdades parciais dos fenômenos da natureza de uma forma sistemática e rigorosa (SIMÕES; SHAW, 2007). O uso na recuperação da informação é interessante já que em alguns casos o pesquisador não tem ideia clara, exata, da informação que busca ou do termo adequado para pesquisar. Além disso, o próprio conceito de relevância é vago, impreciso e pessoal, o que torna um documento relevante para alguém e irrelevante para outra pessoa (SILVA; SANTOS; FERNEDA, 2013).

A técnica de mineração de dados ou Data Mining, consiste em extrair conhecimento de grandes volumes de base de dados estruturados, a fim de encontrar padrões consistentes, como regras de associação ou sequências temporais, para detectar relacionamentos sistemáticos entre variáveis, detectando assim novos subconjuntos de dados. É uma técnica que integra a Descoberta de Conhecimento em Banco de Dados ou Knowledge Discovery in Databases (KDD) (BRAGA, 2005).

De acordo com Ferneda (2006) as redes neurais constituem um campo da ciência da computação ligado à inteligência artificial, que se baseia em modelos matemáticos que se assemelhem às estruturas neurais. A utilização de redes neurais em sistemas computacionais de recuperação de informação permite atribuir um caráter dinâmico a tais sistemas, dado que as representações dos documentos podem ser reavaliadas e alteradas de acordo com a especificação de relevância atribuída pelos usuários aos documentos recuperados. A habilidade de um ser humano em realizar funções complexas e principalmente a sua capacidade de aprender advém do processamento paralelo e distribuído da rede de neurônios do cérebro.

Zerarga e Djouadi (2018) propuseram uma teoria axiomática incluindo a lógica booleana, a *fuzzy* e a vetorial para recuperar informações. Para cada modelo, uma

interpretação foi proposta e a satisfação da teoria foi comprovada. Tal abordagem configura uma possível inovação dos sistemas de recuperação da informação.

A despeito de diferentes técnicas existentes para recuperação da informação, a pesquisa semântica, baseada em tecnologias específicas e tendo a ontologia formal em sua estruturação, possibilita resultados mais precisos e de maneira mais rápida para o usuário da informação (LI *et al.*, 2017; MALGAONKAR; DEVALE, 2016). A seguir apresenta-se a ontologia como componente dos sistemas de recuperação da informação na Web.

2.2 ONTOLOGIA COMO MODELO DE RECUPERAÇÃO SEMÂNTICA

Nesta Seção aborda-se o conceito de ontologia e sua respectiva aplicação na recuperação da informação, conforme preceitos da Web Semântica. Aborda-se também a representação semântica da informação para recuperação dados abertos vinculados e disponíveis em sistemas complexos e *big data*.

2.2.1 A ontologia formal

A ontologia formal captura conhecimento fornecendo termos para que possam ser usados de uma maneira processável por máquina (GRUBER, 1996; NOY; GUINNESS, 2005). Dependendo das diferentes aplicações, a estrutura da ontologia é construída e projetada em diferentes modelos (TA; THI, 2015a). Boas ontologias levam diretamente a um maior grau de reutilização e uma melhor cooperação sobre os limites de aplicativos (TA; THI, 2015a).

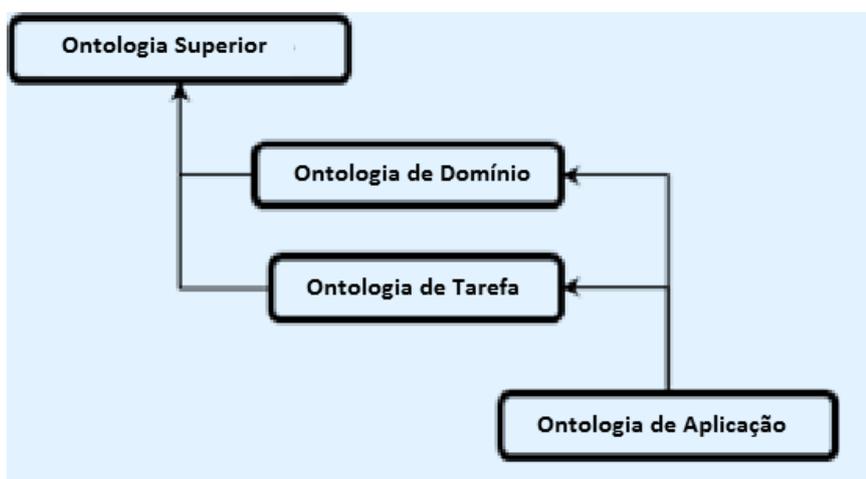
O fundamento da ontologia advém do campo da filosofia, mas quando aplicada à ciência da computação fornece uma especificação formal a um termo e suas relações, associadas a outros termos em um determinado domínio do conhecimento (GRUBER, 1996). A partir dessa especificação se constrói um vocabulário para definir os termos no domínio especificado e os usuários desse domínio, juntamente com *software* que compartilham essa estrutura para analisar, partilhar e reutilizar o conhecimento (BARBA-GONZÁLEZ *et al.*, 2019; GUO *et al.*, 2018; GUARINO, 1997; GRUBER, 1996, TELNOV, 2015).

No contexto da inteligência artificial, uma ontologia é uma especificação formal, uma especificação explícita de uma concepção de um domínio de interesse

(GRUBER, 1996). A ontologia fornece a representação do conhecimento compartilhável e reutilizável e inclui descrições de conceitos e propriedades de um domínio do conhecimento, as relações entre conceitos, as restrições e sobre como as relações podem ser usadas (TELNOV, 2015).

Uma ontologia formal pode ser classificada em três tipos: 1 - ontologia superior: relaciona conceitos amplos e gerais; 2 - ontologia de domínio ou de tarefa: incluem conceitos que estão relacionados a um domínio específico, como por exemplo, neurociência e medicina, ou para uma tarefa específica, como o diagnóstico; e 3 - ontologia de aplicação: incluem conceitos que são especializações habituais de ambos as ontologias de domínio e de tarefa e fornecem estruturação de conceitos no nível do aplicativo (Figura 3). No campo da recuperação da informação é de interesse a ontologia de domínio ou tarefa (AROUA; MOURAD, 2017; GUARINO, 1998; PAIVA; COSTA; SILVA, 2014; SILVELLO *et al.*, 2017).

Figura 3 - Tipos de ontologia.



Fonte: Adaptada de Guarino (1998).

A ontologia de domínio vem sendo aplicada para a recuperação da informação na Web em diferentes áreas como no comércio eletrônico (BAHAFID *et al.*, 2015), na medicina (CLUNIS, 2018, TAHMASEBI *et al.*, 2019), na pesquisa de imagens e objetos (ASPURA; NOAH, 2017); na nanotecnologia (SHARMA, 2016); na tomada de decisão organizacional baseada em banco de dados (GIRASE; PATANAIK; PATIL, 2016).

No campo da recuperação da informação a ontologia assemelha-se a uma taxonomia ampliada que é a linguagem documentária usada para representação do

conhecimento (SALES; CAFÉ, 2008). A taxonomia apresenta um vocabulário controlado formado por termos relacionados que pode ser utilizado para recuperar informações com base em palavras-chave (SALES; CAFÉ, 2008). Entretanto, na taxonomia, os conceitos partem de relações hierárquicas simples, mas que não incluem a definição de termos.

O uso de palavras-chave livres para fornecer uma primeira ideia aproximada do conteúdo semântico de um documento pode ser considerado como uma técnica simples ou básica de anotação semântica/conceitual. Uma evolução desta abordagem básica diz respeito ao uso de dicionários de tesouros especializados que incluem um vocabulário fechado e parcialmente estruturado, próprio de domínios de aplicação específicos e que ainda pode conter limitações (SILVEIRA *et al.*, 2021; ZARRI, 2014).

Em um contexto de anotação semântica, a tendência consiste em passar de uma simples abordagem de “palavra-chave” para uma abordagem de “metadado”. Os metadados podem ser definidos em geral como coleções estruturadas e compreensíveis por máquina de itens lexicais a serem usados para expressar afirmações sobre a organização e o conteúdo de alguns conjuntos de documentos digitais ou não digitais (ZARRI, 2014). Existe uma grande variedade de formatos de metadados para recursos digitais, como o básico Dublin Core, Iniciativa de Codificação de Texto (TEI) e a Catalogação de Leitura por Máquinas (MARC) (DUMER; ALBUQUERQUE, 2021; ZARRI, 2014;).

Para fazer a análise semântica da palavra-chave em uma pesquisa ambígua inserida pelos usuários é necessária uma base de conhecimento ontológica construída com base em metadados refinados (LEE *et al.*, 2018). A técnica de análise semântica de palavras-chave reflete as intenções de pesquisa do usuário a partir da base de conhecimento bem estabelecida e que pode gerar resultados de pesquisa mais precisos. Na pesquisa semântica a palavra-chave é convertida automaticamente em linguagem de consulta estruturada, interpretando as palavras-chave com base em ontologias. A consulta da palavra-chave digitada pelo usuário é separada e para cada termo é identificada uma ontologia. O sistema procura por relacionamentos entre objetos pela navegação facetada em ontologias com base no termo identificado, oferecendo resultados de pesquisa precisos (LEE *et al.*, 2018).

A navegação facetada é a forma de navegação pelas facetadas, ou seja, pelas categorias fundamentais de um assunto, em cada campo de conhecimento é

representado pelos componentes de uma ontologia (LEE *et al.*, 2018; SALES; CAFÉ, 2008).

A recuperação da informação na Web é, portanto, o maior campo de aplicação de ontologia formal, segundo Telnov (2015), já que esta proporciona a melhoria na precisão das pesquisas *online* adicionando informação semântica que é útil para reduzir o espaço de busca da informação. A ontologia pode ser usada para resolver vários problemas de heterogeneidade da Web, enriquecendo os metadados com etiquetas semânticas significativas, que auxiliam os usuários com relação às suas necessidades de informação (TELNOV, 2015)

Uma ontologia formal pode ser escrita para diferentes tarefas uma vez que muitos domínios podem precisar de uma representação específica e formal do conhecimento. Na integração de dados, por exemplo, o objetivo é integrar sistemas de informação heterogêneos. Muitas vezes, um mesmo tipo de informação é mantido em diferentes bancos de dados, com diferentes padrões de modelagem de dados. Uma ontologia pode ser usada como mediador entre os esquemas de banco de dados, o que permite integrar a informação em diferentes padrões tendo a ontologia como um intérprete entre os dados de duas fontes diferentes (TELNOV, 2015).

O compartilhamento de informações faz parte das interações humanas desde a era primitiva até o advento do WORLD WIDE WEB - W3C (MALHOTRA; NAIR, 2015). Destarte, o grande volume de informações armazenadas em diferentes domínios levou a mudanças tecnológicas para facilitar a recuperação de informações relevantes em diferentes tipos de sistemas, como os sistemas especialistas e os sistemas baseados em conhecimento. Os sistemas especialistas foram idealizados para executarem tarefas complexas em menor tempo do que um especialista humano; já os sistemas baseados em conhecimento fornecem uma arquitetura estruturada para representar o conhecimento por meio de classes, subclasses e instâncias a partir do uso de ontologia formal (MALHOTRA; NAIR, 2015).

Para que os sistemas baseados em conhecimento disponibilizem em sua interface pesquisas baseadas em ontologias, faz-se necessário dispor de um conjunto de tecnologias da Web Semântica – difundidas por Tim Berners-Lee em 2001. Na Seção seguinte aborda-se a representação semântica da informação juntamente com as estruturas e tecnologias apregoadas pela Web Semântica.

2.2.2 Representação semântica da informação.

A representação semântica da informação pode ser obtida a partir dos preceitos da Web Semântica que inclui uma arquitetura computacional com o objetivo de inserir significado nos documentos para que as máquinas também possam interpretá-los e, desta forma, aumentar a eficiência das buscas e facilitar a recuperação de informações relevantes para o usuário da Web (BERNERS-LEE; HENDLER; LASSILA, 2001).

Com a explosão das informações na Web, o algoritmo tradicional de mapeamento de classificação das páginas na Internet não consegue lidar com a classificação complexa dos *sites* da Web. Alguns grandes *sites* de portal de notícias e empresas de mídia social vêm utilizando abordagens otimizadas para classificação das páginas da Web e de recuperação da informação a partir do uso de redes semânticas usadas em *sites* da Web. Esta prática tem por objetivo equalizar o interesse dos usuários e o conteúdo da Web (LI *et al.*, 2017; MALGAONKAR; DEVALE, 2016).

A Web Semântica, ou WEB. 3.0, é conhecida como uma Web com significado, cujas informações são legíveis por máquina, o que possibilita a oferta de serviços automatizados (SIRISHA; SUBBARAO; KAVITHA, 2015). O objetivo da arquitetura da Web Semântica é fornecer representação do conhecimento de dados ligados de modo a permitir o processamento por máquina em uma escala global. Para tanto, é necessário dispor de componentes semânticos da Web, ferramentas e tecnologias específicas (SIRISHA; SUBBARAO; KAVITHA, 2015).

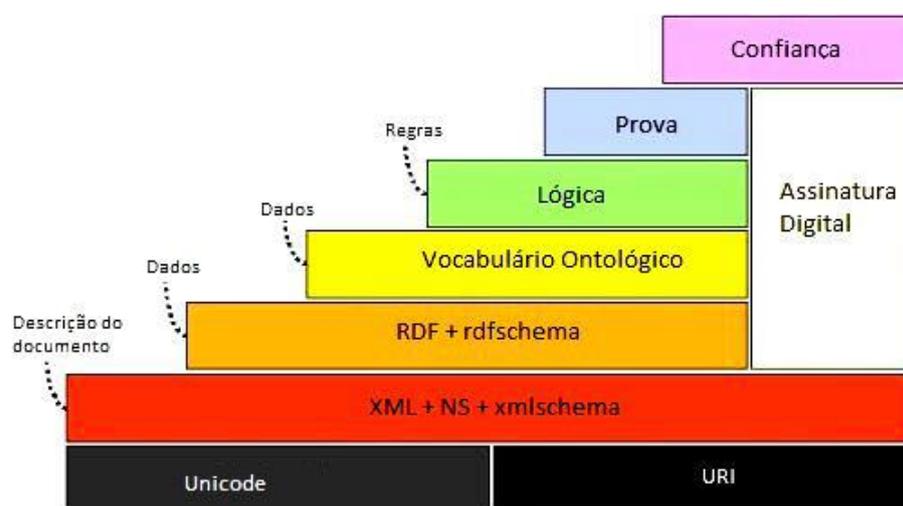
As tecnologias da Web Semântica fornecem interoperabilidade semântica definida pelos padrões do WORLD WIDE WEB - W3C (HAWKE *et al.*, 2013). Assim, forma-se uma grande rede semântica que fornece um quadro comum que permite que dados sejam compartilhados e reutilizados em diversas aplicações em diferentes campos a fim de auxiliar a tomada de decisão (HAWKE *et al.*, 2013).

A representação semântica é marcada pela ação de agentes inteligentes, os *softwares*, que por meio de uma infraestrutura de dados, utilizam técnicas de inteligência artificial que permitem a troca de informações, automatizando tarefas de rotina dos usuários (BERNERS-LEE; HENDLER; LASSILA, 2001). O conteúdo distribuído na Web é anotado utilizando-se diferentes ontologias permitindo que

agentes inteligentes possam “compreender” e utilizar a informação para auxiliar os usuários nas pesquisas (BERNERS-LEE; HENDLER; LASSILA, 2001).

A Web Semântica apresenta uma estrutura dividida em camadas de linguagens, propostas por Koivunen e Miller (2001). Esta estrutura é composta por tecnologias de base, ou seja, sistema operacional, banco de dados, servidor Web e sistema de gerenciamento de conteúdo. Inclui também tecnologias específicas e necessárias para que os conteúdos das páginas possam ser compreendidos por computadores (Figura 4).

Figura 4 - Arquitetura da Web Semântica.



Fonte: Adaptado de Koivunen e Miller (2001, p. 34).

Segundo Koivunen e Miller (2001), compõem a Web semântica as seguintes camadas:

- a) Camada Unicode /Uniform Resource Identifier (URI) - define uma forma única para identificar e localizar páginas na Web.
- b) Camada XML/NS/XMLschema - a linguagem Extensible Markup Language (XML) possibilita a inclusão de *tags* (etiquetas ou metadados) para descrever o conteúdo do documento, fornecendo sintaxe básica para criar documentos com dados organizados de forma hierárquica. O Name Space (NS) é um padrão para prover um nome único de elemento e de atributo em uma instância XML, eliminando assim, a ambiguidade entre nomes idênticos de elementos ou atributos. O XML Schema possibilita a definição do formato válido de documentos XML, incluindo quais os elementos e os atributos são permitidos e suas

- localizações, o número de ocorrências de cada elemento e outras características, sendo possível definir gramáticas para correção de documentos XML.
- c) Camada RDF e RDF Schema – o modelo de dados Resource Description Framework (RDF) é considerado o fundamento da Web Semântica que está relacionado com a representação descritiva dos recursos da Web que expressam significados que são codificados em um conjunto de três elementos – triplas, cada uma contendo sujeito, verbo e objeto de uma sentença elementar. Essas triplas podem ser escritas em *tags* XML, descrevendo relações entre recursos em termos de propriedades e valores. Além do modelo e da sintaxe, a camada RDF inclui também a linguagem RDF Schema que permite a definição de vocabulário para representação de ontologias simples e descrição de classes e propriedades.
 - d) Camada de ontologia – composta pela metaontologia Web Ontology Language (OWL) que possibilita ampliar o vocabulário da RDF Schema para a inclusão de elementos com maior poder de expressividade e inferência. Esta linguagem é composta por outras três sublinguagens: (i) - OWL Lite - permite definir uma ontologia de classes, propriedades e instâncias, possibilitando uma rápida migração de tesouros e taxonomias; (ii) - OWL DL (Description Logic ou Lógica de Descrição) - permite definir uma ontologia de classes e propriedades e instâncias, possibilitando uma rápida migração de tesouros e taxonomias; (iii) - OWL Full – permite o tratamento simultâneo de classes como coleções e instâncias. Cada uma destas sublinguagens é uma extensão da sua precedente, sendo que as linguagens menos expressivas (OWL Lite e DL) estão contidas nas mais expressivas (OWL DL e Full).
 - e) Camada de lógica - possibilita descrever regras para expressar relações sobre os conceitos de uma ontologia, facilitando a construção de inferências e possibilitando que os agentes inteligentes possam relacionar e processar informações. Na descrição de regras semânticas podem ser utilizadas linguagens como Rule Markup Language (RuleML) e Semantic Web Rule Language (SWRL) e o Rule Interchange Format (RIF) para fornecer suporte ao intercâmbio das diversas tecnologias baseadas em regras.
 - f) Camadas de prova (*Proof*), confiança (*Trust*) e assinatura digital - as camadas de Prova e Confiança fornecem o suporte para a execução das regras, além de avaliar a correção e a confiabilidade dessa execução; possibilita provar uma hipótese a partir da representação semântica das informações contidas nos

documentos e que já foram convertidas nas camadas anteriores. A partir disto é possível inferir respostas sobre o documento. A assinatura digital visa garantir a procedência de um documento, o que é fundamental para decidir se a informação é confiável ou não. Essa assinatura contém um pequeno código que pode ser usado para verificar a autoria de um documento e que é aplicado em várias fases da Web Semântica. Após a assinatura digital é possível atribuir o grau de confiança da informação.

A Web Semântica é uma grande rede semântica que pode fornecer um quadro comum para que dados sejam compartilhados e reutilizados em diferentes áreas do conhecimento. Modelos que incluem ontologia para representação semântica com utilização de tecnologias de busca da Web Semântica têm sido utilizados para melhorar a resposta das buscas em sistemas complexos da Web (AMATO *et al.*, 2017; BENABDERRAHMANE *et al.* 2017; LI *et al.*, 2017; SILVELLO *et al.*, 2017). A título de exemplo tem-se:

- Aroua e Mourad (2017) utilizaram modelagem de ontologia baseada no perfil de usuário, aplicada em um sistema de transporte público para recuperar rotas mais adequadas. O modelo foi composto por preferências explícitas e implícitas do usuário, estabelecidas mediante relações hierárquicas e semânticas em conjunto com propriedades relacionadas. A fase de extração das preferências e propriedades foi obtida pelas regras da linguagem semântica SWRL, estendidas por meio da técnica de ativação *Spreading*, que possibilita ampliar, tanto quanto possível, novos conhecimentos a partir dos termos de busca e o enriquecimento de novos conceitos e relações dentro de uma ontologia de domínio específico. Este método pode facilitar a descoberta de interesses implícitos e antecipar novos resultados de busca em pesquisas na Web.
- Moura *et al.* (2015) utilizaram uma abordagem de integração semântica e dados vinculados para publicar e recuperar dados sobre ecologia na Web. O modelo usado pelos autores supramencionados incluiu dados produzidos e coletados de diferentes recursos de mídia, armazenados em bancos de dados relacionais distintos e publicados como triplos RDF. Uma ontologia de aplicação foi utilizada para compartilhar vocabulário dos conjuntos de dados e facilitar a recuperação da informação contida nos dados vinculados.

Segundo Ta e Thi (2015a), a ontologia vem desempenhando um papel importante nos últimos anos e suas aplicações são mais populares e variadas. São utilizadas em diferentes áreas relacionadas à tecnologia da informação, à biologia e à medicina, especialmente em recuperação de informação, extração de informação e resposta a perguntas.

Para Ta e Thi (2015b), o objetivo da ontologia específica de domínio é eliminar a confusão conceitual e terminológica. Isto é realizado especificando um conjunto de conceitos genéricos que caracteriza o domínio, bem como suas definições e inter-relações. Existem algumas linguagens para representar ontologias, como RDF e OWL, conforme já citado. No entanto, essas linguagens são adequadas apenas com ontologias ou com poucos dados. Para representar ontologias em grandes bancos de dados é necessário o uso de abordagens para pesquisa de palavras-chave baseada em semântica sobre bancos de dados relacionais.

Seneviratne, Fernando e Karunaratne (2019) destacaram que a despeito de esforços para classificar documentos em grupos de acordo com o domínio do assunto ou de acordo com os outros atributos, ainda prevalece um grande desafio, a extração de informações devido a grande quantidade de informações da Web. Esses autores utilizaram o modelo Markov, baseado em lógica e regras para classificação de documentos na Web, com uma combinação de técnicas de extração de relações que considera os pesos ideais para cada regra, o que contribuiu para a medida de desempenho das regras na classificação do documento, mantendo-se apenas as relações relevantes.

Uma grande quantidade de dados estão disponíveis na Web e a utilização destes requer um sistema de recuperação baseado em semântica (REYES-ÁLVAREZ; ROLDÁN-GARCÍA; ALDANA-MONTES, 2019). Na Seção seguinte discute-se sobre dados vinculados e abertos em ambientes semânticos com foco na organização e acesso a dados na Web.

2.2.3 Dados vinculados e dados abertos vinculados

Um conjunto de dados estruturados logicamente forma um banco de dados que contém informações de um determinado domínio (DATE, 2004). Segundo a Open Knowledge Internacional (2019), dados abertos se relacionam ao livre acesso de um dado a qualquer pessoa que deseje utilizar, modificar e compartilhar para qualquer finalidade. Dados abertos por extensão constituem blocos de conhecimento aberto quando estes se tornam úteis, utilizáveis, usáveis e usados (OPEN KNOWLEDGE INTERNACIONAL, 2019).

Dentre as principais características da abertura de um dado incluem-se (OPEN KNOWLEDGE INTERNACIONAL, 2019):

- a) disponibilidade e acesso: os dados devem estar disponíveis como um todo de forma conveniente e modificável para acesso via Internet;
- b) reutilização e redistribuição: os dados devem ser legíveis por máquina de modo que possibilitem a reutilização e a redistribuição;
- c) participação universal: os dados devem estar disponíveis para uso, reutilização e redistribuição sem qualquer restrição de área de atuação, pessoa ou grupos. Para isso, a publicação de dados em formato aberto deve ser feita mediante uma licença aberta, ou seja, sem controle exclusivo, sem regulações de direitos autorais, marcas, patentes ou segredo industrial.

Dados vinculados ou interligados se referem à conexão de dados a outros conjuntos de dados que, quando unidos a dados abertos, possibilitam a recuperação de documentos e informações relacionadas que explicam e descrevem o conteúdo, seus significados e a relação entre as informações mostradas (MOURA *et al.*, 2015).

Segundo Reyes-Álvarez, Roldán-García e Aldana-Montes (2019), dados vinculados referem-se a conjuntos de dados semanticamente bem estruturados, interconectados e sintaticamente interoperáveis que são distribuídos entre vários repositórios e descritos em Resource Description Framework (RDF) como o formato comum. Além dos dados RDF, existem linguagens de ontologia, como o Web Ontology Language (OWL), que podem fornecer vocabulários para estruturar domínios de conhecimento dos dados.

Na Web clássica os *hiperlinks* conectam documentos em um único espaço global de informações (HEATH; BIZER, 2011). Com o advento da Web semântica, um

link pode ser definido entre itens em diferentes fontes de dados e, assim, conectar essas fontes em um único local. O uso de padrões da Web e um modelo de dados comum possibilitam a implementação de aplicativos que operam em todo o espaço de dados. Essa é a essência dos dados vinculados (HEATH; BIZER, 2011).

A tecnologia de dados vinculados torna legíveis os dados de rede em todas as plataformas de modo que esses sejam compartilháveis e vinculáveis a conjuntos de dados externos. Como consequência, essa abordagem tem o potencial de suportar redes abertas e dinâmicas, redes que podem ser cada vez mais enriquecidas com camadas de semântica adicional de fontes internas e externas (BERNERS-LEE; HENDLER; LASSILA, 2001).

Thangsupachai, Niwatthanakul e Chamnongsri (2015) destacaram que a transição da Web de dados e de conteúdo compartilhável para o padrão aberto, como os dados vinculados abertos, exige nova estruturação para os dados tradicionais. A interoperabilidade dos sistemas de informação é alcançada pelo uso do formato RDF que proporciona heterogeneidade semântica. Esses autores apresentaram uma solução baseada em padrões da Web Semântica no qual os bancos de dados relacionais são diretamente mapeados para RDF e OWL. Dado um esquema de banco de dados relacional e suas restrições de integridade, esse mapeamento direto produz uma ontologia de OWL, que fornece a base para gerar instâncias de RDF. A semântica desse mapeamento é definida usando o modelo de mapeamento de metadados. Duas propriedades fundamentais incluem a preservação de informações e a preservação de consultas, obtidas a partir do uso de Application Programming Interface – API – Jena, pela linguagem de consulta Open Knowledge Base Connectivity – OKBC e por algoritmo de pesquisa baseado em técnicas de pontuação de similaridade semântica.

O banco de dados usado na Web Semântica é formado por metadados que descrevem o sujeito, o predicado e o objeto de um documento, sendo possível correlacionar esses metadados, a partir da linguagem RDF e OWL, para recuperar um contexto informacional com mais exatidão (BIZER; HEATH; BERNERS-LEE, 2009; KOIVUNEN; MILLER 2001).

Na integração de dados, a interoperabilidade, ou seja, a capacidade de um sistema se comunicar de forma transparente com outro sistema é considerada um grande desafio para a recuperação da informação, já que em um sistema é essencial suportar consultas que forneçam uma visão de dados consolidada a partir de diferentes repositórios heterogêneos. Neste contexto, o uso de abstrações

conceituais fornecidas por ontologias, juntamente a abordagem de dados vinculados, ajudaram a superar esse desafio. Quando usadas em conjunto, ontologias e dados vinculados, se tem uma boa estratégia para integrar dados (HEATH; BIZER, 2011).

Estudos descrevem a aplicação de recursos semânticos no acesso de dados abertos (BRANDT; VIDOTTI; SANTAREM SEGUNDO, 2018; MOURA *et al.*, 2015; PATTUELLI; MILLER, 2015; SILVELLO *et al.*, 2017), como por exemplo:

- Pattuelli e Miller (2015) desenvolveram uma ferramenta para examinar dados de diferentes fontes primárias da comunidade do *jazz* digitadas em redes sociais. O modelo incluiu a combinação de redes semânticas e princípio dos dados abertos vinculados de forma integrada e interativa. Com o uso de tecnologias semânticas da Web, incluindo a estruturação RDF, aplicado à construção e aprimoramento da rede das fontes primárias e do conjunto de dados subjacente, foi possível tornar os dados legíveis em todas as plataformas e vinculáveis a conjuntos de dados externos. Essa abordagem teve o potencial de tornar as redes sociais de *jazz* dinâmicas e abertas à integração com fontes de dados externos.
- Silvello *et al.* (2017) propuseram um modelo semântico para recuperação da informação de dados abertos na Web. O modelo incluiu o RDF como base para o processamento de metadados, proposto pela Web Semântica, com a finalidade de aprimorar a capacidade de descoberta e facilitar conexões dos dados científicos com a produção científica relacionada. Com o modelo seria possível fornecer uma metodologia para enriquecer automaticamente os dados, explorando dados externos relevantes da nuvem de dados abertos vinculados. A modelagem semântica contribui em tarefas de anotar e interligar dados científicos experimentais, melhorando a interpretação, compartilhamento e reutilização destes.
- Um modelo de dados abertos conectados foi proposto por Brandt, Vidotti e Santarem Segundo (2018), envolvendo um conjunto legislativo da Câmara dos Deputados do Brasil, que contém informações sobre partido político, unidade federativa, endereço eletrônico, legislatura, entre outras, relacionando com os parlamentares. A estruturação dos dados foi feita mediante o recurso RDF e uso de vocabulários e padrões já estabelecidos na Web Semântica como Dublin Core, Friend of a Friend (FOAF), RDF e RDF Schema, além de vocabulários de áreas correlatas, como a Ontologia da Câmara dos Deputados

Italiana e a da Assembleia Nacional Francesa. Os recursos foram relacionados também a outros conjuntos de dados abertos conectados para enriquecimento semântico, como as bases Geonames e DBpedia. Em conclusão esses autores consideraram que a disponibilização dos dados legislativos, de acordo com as recomendações do WORLD WIDE WEB - W3C e integrados a outros dados legislativos da Web, podem ampliar as possibilidades de reuso e aplicações dos dados em ações de transparência e fiscalização, aproximando os cidadãos do Congresso e de seus representantes.

- Reyes-Álvarez, Roldán-García e Aldana-Montes (2019) propuseram um modelo otimizado para armazenar ontologias OWL classificadas em um banco de dados orientado por coluna, de código aberto, como um passo inicial para o desenvolvimento de um modelo de raciocínio dado, ligado, escalável e incremental. A técnica permitiria o armazenamento distribuído dos dados RDF, explorando a distribuição inerente dos nós do banco de dados NoSQL, fornecendo vocabulários para estruturar domínios de conhecimento dos dados para facilitar a consulta de dados.

O grande volume de dados abertos disponíveis na Web forma sistemas complexos que serão descritos a seguir.

2.2.4 Sistemas complexos e *big data*

De acordo com Wilson (2017), o termo *big data* se refere à quantidade total de dados gerados e armazenados eletronicamente, incluindo conjuntos de dados específicos que são grandes em tamanho e em complexidade para os quais são necessárias novas técnicas algorítmicas para extrair destes dados e informações úteis.

O *big data* apresenta cinco características (5Vs) (NIMBIX, 2021), descritas a partir de 2012, a saber: 1- Volume – alta quantidade dados que precisam ser processados em um determinado momento; 2- Variedade – diversidade de informações e dados produzidos, formatos e fontes de dados heterogêneas que precisam ser identificados e normalizados antes que a computação possa ocorrer; 3- Velocidade – agilidade em que os dados são coletados, processados, analisados e utilizados; 4- Veracidade – qualidade dos dados e a confiabilidade sobre as fontes dos

dados; 5- Valor – importância ou significado de um dado para uma determinada solução (FAM; NORISZURA; SHINYIE; 2019). Mais recentemente, outras três características foram arroladas: 6- Viabilidade - refere-se à capacidade de um modelo de representar a realidade; 7-Vocabulário - remete a comunicação do provedor e o usuário e também a busca semântica a partir de ontologias; 8- Vaguidade - imprecisão na interpretação dos resultados que retornam das buscas

Os sistemas complexos, a exemplo do *big data*, vêm sendo alvo da área de recuperação da informação uma vez que estes oferecem uma variedade de aplicações para a indústria e a sociedade tais como; pesquisa médica, pesquisa empresarial, propriedade intelectual, pesquisa de patentes e muitos outros (SILVELLO *et al.*, 2017). Os portais do conhecimento de universidades também disponibilizam grande quantidade de dados do domínio acadêmico que podem ser reutilizados, a exemplo dos sistemas complexos.

A crescente demanda pelo desenvolvimento de sistemas de informação complexos que reduzam custos e garantam a confiabilidade levam à adoção de diferentes paradigmas com abordagens que incluem compartilhamento de conhecimento, interoperabilidade, integridade, qualidade e reutilização de dados (GARGIULO *et al.*, 2014).

Considerando que as técnicas tradicionais de recuperação da informação são ineficazes para tratar com o advento do *big data*, o desafio consiste em projetar e desenvolver esses sistemas complexos e facetados, por meio de metodologias que possibilitem mensurar o desempenho alcançado, avaliar o impacto de alternativas e novas ideias e garantir os níveis de eficácia necessários para atender às expectativas dos usuários por informações relevantes (SILVELLO *et al.*, 2017).

A recuperação semântica da informação, fornecida pelo uso de ontologias constitui uma opção que fornece alta precisão e resultados mais rápidos em sistemas complexos ou *big data*. As informações semânticas podem expressar as intenções de pesquisa dos usuários e essa abordagem se tornou uma ferramenta importante no campo da recuperação de informações (GUO *et al.*, 2018).

Alomran (2014), combinou técnicas de mineração de dados e da Web semântica para extrair informações relevantes da Web em um ambiente de aprendizagem virtual. A mineração de dados consiste na análise dos dados contidos no texto de linguagem natural e muitas vezes inconsistente com dados não estruturados, o que torna difícil a recuperação da informação. A ontologia auxilia na

extração de metadados contidos nos documentos minerados (títulos, resumos, data de criação, autores) e também para extrair informação semântica que é deduzida a partir do material recuperado. A arquitetura combinando mineração de dados e semântica pode melhorar os resultados de pesquisa na Web.

Segundo Ni (2015), com grande quantidade de informações disponíveis na era do *big data*, o conhecimento está surgindo em várias formas de dados não estruturados, exigindo novas metodologias de construção de recursos significativos para facilitar a aprendizagem do conhecimento. A maioria dos métodos baseados em dados existentes, tal como a mineração de dados, serve apenas ao objetivo de melhorar o poder discriminativo das características, enquanto subestima a importância da interpretabilidade. Esse autor propôs uma estrutura de construção de características baseada em ontologias para a descoberta e o aprendizado de novas hipóteses e conhecimentos de forma significativa e compreensível a partir de dados não estruturados. A metodologia proposta foi aplicada em dados da área médica para construir características interpretáveis e altamente preditivas para a previsão de hospitalização. De forma interativa, especialistas de domínio da área médica foram envolvidos para auxiliar na evolução da ontologia e avaliar a significância das características construídas e assim, explorar novas hipóteses com o auxílio de algoritmos heurísticos baseados em dados.

Já no estudo proposto por Malgaonkar e Devale (2016) foram utilizados princípios de rede semântica de recuperação de imagens na Web. Os mecanismos de busca são baseados no modelo de classificação de palavras ou sistema de relevância por reclassificação. A relevância é um fator importante que influencia o desempenho do sistema de recuperação de informações com precisão. A partir da consulta do usuário, agentes inteligentes filtram os termos de busca e os compara semanticamente com o padrão de *links* de redes de informações disponíveis na Web, selecionando melhores resultados.

Benabderrahmane *et al.* (2017) propuseram, igualmente, um sistema de recuperação da informação com uma arquitetura que combina informações na área de recrutamento de pessoal. O sistema é composto de dois módulos: um módulo de recomendação temporal, baseado nas buscas por ofertas de trabalho em determinado período e outro módulo de recuperação semântica, baseada na análise textual das ofertas de trabalho mediante uso de vocabulário controlado. Os módulos foram

implementados em uma arquitetura de *big data*, levando em consideração o grande volume de dados complexos.

Amato *et al.* (2017) desenvolveram um sistema de recuperação de Informação baseada no conhecimento de coleções de arte para consultar e analisar conteúdos digitais culturais a partir de um conjunto de repositórios multimídia distribuídos e heterogêneos. O sistema fornece funcionalidades de recuperação e apresentação necessárias para busca de informação, capazes de acessar de forma eficaz e eficiente os dados desejados, explorando ao mesmo tempo características e descrição de metadados, juntamente com a semântica relacionada aos objetos culturais a fim de que o resultado das buscas corresponda às necessidades e preferências do usuário.

Para Nadal *et al.* (2019), as arquiteturas do *big data* permitem armazenar e processar de forma flexível dados heterogêneos, a partir de múltiplas fontes, em seu formato original. A estrutura desses dados, normalmente fornecida por meio de APIs, está em constante evolução e os analistas de dados precisam adaptar seus processos após lançada cada API. Para lidar com a evolução sintática nas fontes de dados, foi apresentado um algoritmo que adapta de forma semiautomática a uma ontologia chamada Big Data Integration que permite uma abordagem sistemática para a governança do sistema de integração de dados em tarefas de consulta e análise de dados. A ontologia é estruturada com base em vocabulário RDF que permite modelar e integrar dados em evolução de vários provedores e anotar semanticamente o processo de integração de dados.

Barba-González *et al.* (2019) propuseram uma ontologia chamada BIGOWL, que abrange um amplo vocabulário de termos relacionados aos fluxos de trabalho de análise de *big data*, incluindo seus componentes e como eles estão conectados, desde as fontes de dados até a visualização analítica. Também leva em consideração aspectos como parâmetros, restrições e formatos.

Consequentemente, a ontologia define não apenas as relações taxonômicas entre os diferentes conceitos, mas também possibilita anotar semanticamente todos os metadados que fluem de várias fontes de dados. Assim, é possível representar e consolidar o conhecimento que envolve análise semântica das fontes e dos resultados do *big data* e, ainda, identificar o caminho correto que os dados seguem e como eles são modificados para obter valor agregado, para um dado domínio de conhecimento. Então, esgotada as questões de recuperação e de tratamento da informação, que se

julgam pertinentes, no Capítulo que se segue serão aprofundados os Portais Semânticos detalhando-se sua estrutura e tecnologias.

3 PORTAIS SEMÂNTICOS

Neste Capítulo serão abordados o conceito de portais – dos tradicionais aos semânticos –, aprofundando a descrição da sua arquitetura e das tecnologias semânticas usadas na sua estruturação. Finalizando o Capítulo serão tratados os indicadores de avaliação do potencial semântico.

3.1 CONCEITO DE PORTAIS

Os portais são definidos como pontos de entrada para apresentação e intercâmbio de informações na Web para diferentes comunidades, com estrutura computacional que favoreça o acesso às informações de forma rápida e precisa (LAUSEN *et al.*, 2005).

Um portal pode ter diferentes designações tais como, portal da Web, portal corporativo, portal do conhecimento e portal semântico (HAUKE; OWOC; PONDEL, 2015). O portal da Web consiste em um serviço de informação *online* estendido a uma variedade de funções da Internet, disponível a partir de um único endereço desta; contém informações de interesse do público em geral e pode oferecer serviços gratuitos, como *e-mail* e hospedagem na Web (HAUKE; OWOC; PONDEL, 2015).

O portal corporativo, também conhecido como portal de negócios, portal de informações corporativas, portal de informações empresariais (DIAS, 2001), constitui-se de um ponto único de acesso a todos os recursos e aplicativos utilizados para troca de informações, gestão do conhecimento na empresa e realização de diversas transações comerciais (HAUKE; OWOC; PONDEL, 2015).

De acordo com Shilakes e Tylmann (1998), os portais corporativos apresentam as seguintes características: integração e categorização de dados heterogêneos, estruturados e não estruturados da organização e seu ambiente; publicação e distribuição de informações específicas e para conhecimento dos colaboradores; integração de aplicações heterogêneas; estes favorecem a comunicação, a colaboração e o compartilhamento de informações, sendo uma interface da organização adaptada às necessidades individuais.

Ainda segundo Hauke, Owoc e Pondel (2015), o portal do conhecimento é uma versão avançada do portal corporativo que apoia e estimula a criação e o compartilhamento de conhecimento, o armazenamento e a recuperação deste,

favorecendo sua criação, integração e aplicação em processos de Gestão de Conhecimento, fornecendo acesso a relevantes artefatos de conhecimento. Esse tipo de portal disponibiliza acesso e pesquisa a repositórios por meio de sistema de organização do conhecimento, aplicativos e serviços *online*, com informações confiáveis sobre um fragmento específico da realidade que pode ser usado no desenvolvimento posterior de uma área ou vinculá-lo a outra área do conhecimento (HAUKE; OWOC; PONDEL, 2015). Um ponto fundamental desses portais é o acesso ao conhecimento expresso em muitas de suas formas, incluindo para sua construção teorias próprias, subjacentes à criação, ao reuso e ao compartilhamento de conhecimento envolvendo diferentes áreas da atuação humana (HAUKE; OWOC; PONDEL, 2015).

Com o advento da Web Semântica, os portais disponibilizados até então passaram a ser considerados portais tradicionais por não incluírem arquitetura favorável para recuperar informações relevantes diante do acúmulo de informações na Web (CORREA, 2012; REYNOLDS; SHABAJEE; CAYZER, 2004).

O portal semântico surge como uma alternativa para superar as limitações dos portais tradicionais (CORREA, 2012; REYNOLDS; SHABAJEE; CAYZER, 2004), constituindo uma extensão natural destes, mas com características distintas quanto à navegação, à atualização, ao processamento, ao acesso e publicação da informação (BERNERS-LEE; HENDLER; LASSILA, 2001; LAUSEN *et al.*, 2005).

Um portal semântico pode ser definido como aquele que utiliza tecnologias semânticas para ofertar serviços de armazenamento, troca de informações e estruturação de recursos de informação mediante o uso de ontologias de domínios que expressam conhecimentos específicos de uma determinada área (GAVRILOVA, 2011; HUAN; MAIDMENT; TIAN, 2011).

A utilização de ontologia em portais semânticos possibilita uma hierarquia flexível de classificação com incorporação de homônimos e sinônimos de conceitos que melhoram os resultados obtidos (WIMMER; YOON; RADA, 2013). Assim, a semântica, fornecida pelo uso de ontologia, é uma alternativa para melhorar a precisão da busca nos portais, pois possibilita uma melhor compreensão do significado contextual de termos digitados nos campos de pesquisa. Nesse tipo de pesquisa o motor de busca não só decodifica, compara e coteja as relações hierárquicas dos termos, mas também as relações complexas que ocorrem entre os termos definidos dentro de ontologias (PAIVA; COSTA; SILVA, 2014).

Além da ontologia, a construção de portais semânticos requer um conjunto de tecnologias integradas que fornecem a base para o funcionamento adequado do portal com uma arquitetura específica, conforme será descrito na Seção seguinte.

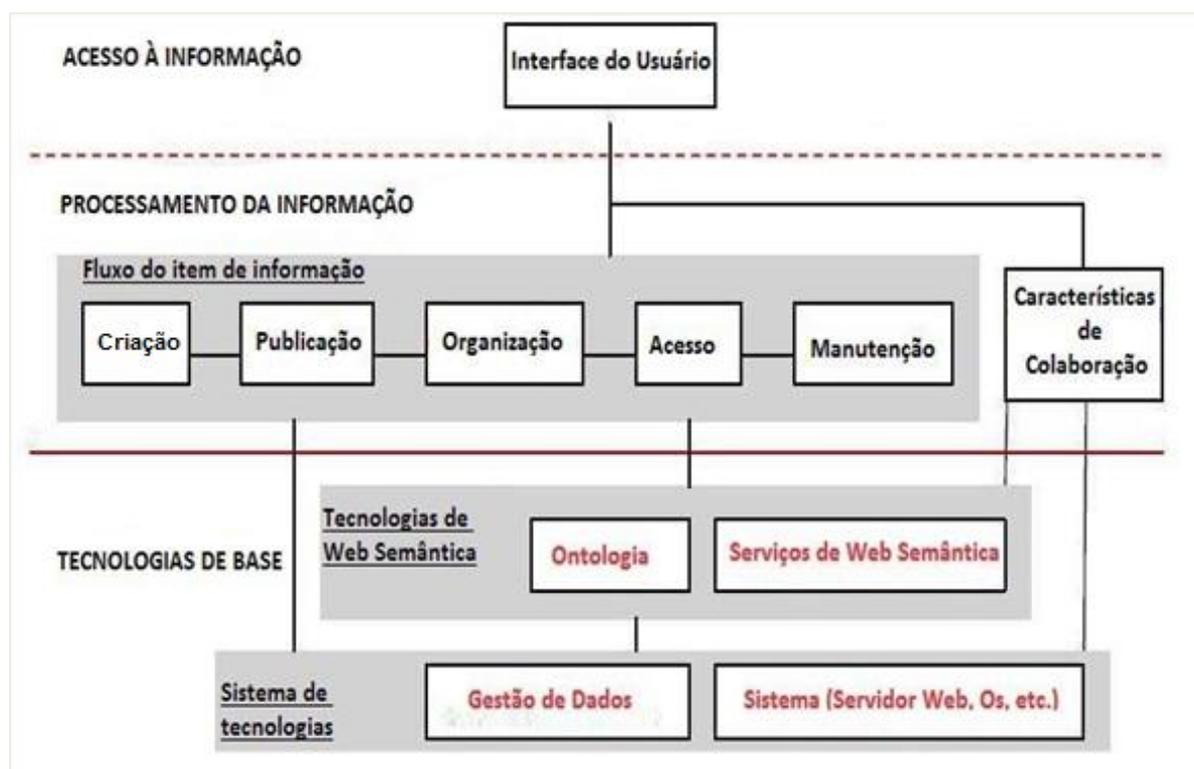
3.2 ARQUITETURA DE PORTAIS SEMÂNTICOS

De acordo com Maedche *et al.* (2001), basicamente a estrutura de um portal semântico inclui um repositório de dados e um mecanismo de inferência que funciona como um mediador entre diferentes fontes de informação. Em um modelo semântico ocorre a interação entre os agentes inteligentes (agentes de *software*), a comunidade de usuários e os usuários em geral que se comunicam com o sistema por meio do servidor Web. Nessa interação, os agentes inteligentes com motores de busca baseados em linguagem RDF acessam e coletam dados e informações armazenados nos sítios em questão – *sites*. Os usuários navegam pelo portal explorando *hyperlinks* de documentos e fazendo buscas de informação. Os usuários da comunidade podem fornecer dados para serem inseridos no portal de forma semiautomática por meio de formulários HTML e, posteriormente os dados são armazenados nos repositórios.

Maedche e Staab (2001) propuseram uma arquitetura para construção de portais semânticos formada por quatro componentes: 1) OntoEdit – Engenharia de Ontologia Workbench; 2) Gestão do componente; 3) Processamento de Recursos e 4) Biblioteca de algoritmo. Segundo esses autores, o engenheiro de ontologia interage apenas por meio das interfaces gráficas, que compreendem o módulo OntoEdit e da Gestão do componente. O OntoEdit oferece, segundo Maedche e Staab (2001), uma interface gráfica para modelagem manual e para o refinamento final da ontologia. A interface do utilizador propicia diferentes pontos de vista, visando um nível epistemológico ao invés de uma linguagem de representação particular. No entanto, o usuário pode exportar as estruturas ontológicas para as linguagens da Web Semântica como o OIL e DAML-ONT, bem com as extensões F-Logic baseadas em modelo RDF (MAEDCHE; STAAB, 2001). Além disso, os usuários podem gerar e acessar representações executáveis para restrição de verificação e aplicação de depuração a partir do Simple Logic-based RDF Interpreter – SilRi, além de usar o motor de inferência F-Logic motor, que se conecta diretamente ao OntoEdit (MAEDCHE; STAAB, 2001).

Lausen *et al.* (2005) utilizaram um roteiro para avaliação de portais semânticos composto por três partes (Figura 5). A primeira parte envolve a camada de acesso à informação e as características de interação entre o usuário e o sistema, tal como a usabilidade de acesso do usuário ao portal como uma aplicação da Web. A segunda parte inclui o processamento de informação que abrange o caminho da informação processada pelo portal, composto pelas etapas de criação, de publicação, de organização, de acesso e de manutenção. Também estão incluídos nessa parte os recursos de colaboração. A terceira parte engloba as tecnologias de suporte do sistema e as tecnologias de Web Semântica. Dentre as tecnologias de base figuram as ferramentas para gestão de dados: sistema operacional; sistema de banco de dados; servidor Web; sistema de segurança. Já as tecnologias semânticas envolvem a ontologia e outras ferramentas para ofertar serviços semânticos tais como: gerenciador de conteúdo, mecanismo de inferência, linguagem XML, modelagem de dados RDF e metaontologia OWL (LAWAN; RAKIB, 2019).

Figura 5 - Roteiro para avaliação de portal semântico.



Fonte: Adaptado de Lausen *et al.* (2005, p. 3).

Rolim *et al.* (2019) conduziram um estudo de caso envolvendo o SemanticSefaz, um portal semântico idealizado para integração entre bases heterogêneas voltadas para o domínio da contratação pública. A arquitetura desse portal apresenta as seguintes camadas: a) Fontes de dados disponibilizadas pelo portal; b) extração e anotação das fontes de dados; c) ontologia de domínio, para anotação semântica de dados a partir do uso de metaontologia OWL; d) Integração semântica das múltiplas fontes de dados; e) acesso a fontes de dados a partir da modelagem de dados RDF e por serviços de consulta SPARQL; f) Aplicação - a interface de pesquisa dos usuários e aplicações que fazem uso dos serviços de integração dos dados.

Rantala *et al.* (2020) apresentaram uma infraestrutura para construção de portais semânticos e dados vinculados usando ontologia no domínio do patrimônio cultural. O portal foi composto por três componentes: 1) modelo para a criação e publicação de dados vinculados heterogêneos e distribuídos com base em estrutura de ontologia compartilhada; os metadados sobre conteúdos provenientes de vários provedores são automaticamente vinculados e enriquecidos formando um gráfico de conhecimento representado por modelo RDF; 2) desenho de interface em que os dados podem ser reutilizados e acessados independentemente de várias perspectivas de aplicativo, enquanto os dados residem em um único terminal SPARQL; 3) modelo em duas etapas para acesso e análise dados, a primeira composta pela seleção do foco de interesse a partir de pesquisa semântica facetada, seguida da visualização ou análise de dados por ferramentas de humanidades digitais que são disponibilizadas para uso do portal. A estrutura proposta é a base do modelo intitulado Sampo, utilizado para a construção de portais semânticos na área das humanidades.

A arquitetura dos portais semânticos é composta, portanto, por tecnologias específicas que incluem linguagem de representação semântica, ferramentas de gerenciamento e programação de ontologias, a serem descritas na Seção seguinte.

3.3 TECNOLOGIAS SEMÂNTICAS

As tecnologias semânticas são usadas para estruturar semanticamente os dados bem como para oferecer serviços de Web Semântica. Fornecem integração inteligente de dados e acesso mais ágil a diferentes tipos de dados por meio de linguagens específicas (STANESCU, 2018).

De acordo com Guedes (2015), a Web semântica prevê a utilização das seguintes tecnologias: a) sistemas de base de padrão aberto; b) gerenciador de conteúdo de código aberto; c) editor de ontologia; d) mecanismos de Inferência; e) linguagens de programação de marcação semântica.

A Plataforma SEALS de acesso aberto e sustentável, tem por objetivo difundir e avaliar as tecnologias semânticas, suas linguagens e ferramentas fornecendo acesso a um conjunto de serviços e de melhores práticas que apoiam todo o ciclo de vida de avaliação dessas tecnologias e que orientam e atualizam os desenvolvedores de portais acerca da sua evolução (STANESCU, 2018).

3.3.1 Linguagem de representação semântica e ferramentas de Ontologia

O modelo de dados RDF e a metaontologia OWL são as principais tecnologias para representar ontologias. O modelo RDF ou triplas (sujeito-objeto-predicado) é o formato adequado para armazenar os dados em bases de conhecimento, que por sua vez estão apoiados em um conjunto de padrões universais estabelecidos pela Web Semântica (WORLD WIDE WEB - W3C, 2021). A base do conhecimento é descrita como uma ontologia geral (terminologia) enriquecida pelo conhecimento extensível individual de uma área específica (BIZER; HEATH; BERNERS-LEE, 2009; TELNOV, 2015).

O XML é a linguagem de representação comumente usada para que o modelo conceitual de dados RDF seja processado pelos computadores, para que, a partir de um vocabulário de domínio específico, seja possível fazer inferências e relações semânticas entre os conteúdos (BIZER; HEATH; BERNERS-LEE, 2009).

A metaontologia OWL é o padrão atualmente recomendado pelo WORLD WIDE WEB - W3C para representar ontologias e inclui ferramentas para comparação de propriedades e classes: identidade, equivalência, oposição, cardinalidade, simetria, transitividade, disjunção, entre outros (MARTINEZ-CRUZ; BLANCO; VILA, 2012). A metaontologia OWL é uma extensão do vocabulário de RDF usada para enriquecer o modelo de dados RDF, definindo um vocabulário vasto para descrever ontologias complexas (EL HAJI; AZMANI; EL HARZLI, 2014; LAWAN; RAKIB, 2019).

O OWL e o RDF, associados às linguagens de descrição lógica SWRL, RIF e RuleML, caracterizam um modelo de dados que representam o conhecimento rico e complexo sobre hierarquias de coisas e as relações entre as coisas, refletindo uma

ontologia de domínio de uma área do conhecimento (LAWAN; RAKIB, 2019; TELNOV, 2015).

Dois tipos de ferramentas suportam tarefas de engenharia de ontologia: 1- editores de ontologia, que são orientados ao usuário e permitem criar e manter ontologias principalmente através de interfaces de usuário; e 2- interfaces de programação de gerenciamento de ontologias, que são orientadas ao desenvolvedor e permitem a criação e a manutenção de ontologias mediante programação em interfaces.

Considerando as diferentes linguagens ontológicas, cada uma com diferentes capacidades de expressividade e raciocínio, é fundamental observar as características de conformidade e interoperabilidade das tecnologias semânticas em relação às especificidades da linguagem ontológica (WRIGLY, GARCIA-CASTRO, NIXON, 2012).

3.3.2 Interfaces de programação de gerenciamento de ontologias

A Web semântica prevê o uso de sistemas de padrão aberto, ou seja, aqueles que estão publicamente disponíveis e não são controlados por corporações ou governos e independem de taxas ou discriminação de uso. Essas tecnologias trazem benefícios a todos, potencializando a interoperabilidade (colaboração) entre todos os envolvidos no processo de comunicação (NGUYEN *et al.*, 2020; OPEN KNOWLEDGE INTERNACIONAL, 2019). Assim, para que um sistema ou *software* seja considerado de padrão aberto, quatro condições devem ser atendidas: 1) as especificações do padrão devem estar disponíveis para uso de qualquer pessoa ou organização; 2) ausência de critérios, favorecimento ou discriminação por parte de seus desenvolvedores; 3) devem permitir extensões; 4) não deve ser cobrado nenhum tipo de taxa para sua implementação (NGUYEN *et al.*, 2020; OPEN KNOWLEDGE INTERNACIONAL, 2019).

São exemplos de sistemas básicos de padrão aberto utilizados em estruturas semânticas: a) Linux - sistema operacional; b) Oracle, MySQL, Postgrate - banco de dados; c) Apache, Nginx - servidor Web; d) Nagios São exemplos de sistemas básicos de padrão aberto utilizados em estruturas semânticas: a) Linux – sistema operacional; b) Oracle, MySQL, Postgrate – banco de dados; c) Apache, Nginx – servidor Web; d) Nagios Core, Snort, OpenStack, KeePassXC – segurança de rede e nuvem; e)

Wordpress, Joomla, Drupal, Textpattern, Radiant – gerenciadores de conteúdo (ELMASRI; NAVATHE, 2011; MASNER *et al.*, 2019; ROSENFELD, MORVILLE; ARANGO, 2015; STALLINGS, 2015; WORLD WIDE WEB - W3C, 2019).

Gerenciadores de conteúdo de código aberto para Web tais como Drupal, Plone e WordPress são compatíveis para uso em aplicações de Web semântica. Com o uso de gerenciadores de conteúdo é possível realizar operações automatizadas de criação, de seleção, de filtro e de categorização dos dados armazenados para apresentá-los ao usuário. Sistemas de gestão de conteúdo baseados preponderantemente em *softwares* e linguagens de programação com padrão aberto possibilitam também trabalho simultâneo de muitos editores em uma variedade de documentos e o controle de versão de todos os documentos, com manutenção consistente da qualidade da informação (MASNER *et al.*, 2019; WORLD WIDE WEB – W3C, 2021).

As ferramentas de raciocínio ontológico baseiam-se em lógicas de descrição (DLs) que são uma família de formalismos de representação de conhecimento baseados em lógica projetados para representar e raciocinar sobre o conhecimento de um domínio de aplicação de maneira estruturada e bem compreendida. Além das linguagens formais de representação do conhecimento, as DLs também fornecem serviços de inferência. A finalidade de tais serviços é extrair novas informações implícitas das informações explicitamente declaradas. Toda linguagem de representação do conhecimento geralmente oferece um conjunto diferente de serviços de inferência. Um serviço de inferência satisfatório deve incluir: satisfabilidade de classe, classificação, implicação lógica e satisfabilidade da ontologia (WRIGLY, GARCIA-CASTRO, NIXON, 2012).

Mecanismos de inferência como o SPARQL Engine e XQuery de padrão aberto são utilizados para otimização das respostas para buscas semânticas. Nos portais semânticos, a linguagem SPARQL é utilizada para consultas sobre uma base de dados RDF, relacionando, por meio das triplas RDF, as informações a serem procuradas com aquelas já definidas como ponto de partida para a busca (DUARTE; HARA, 2018; EL HAJI; AZMANI; EL HARZLI, 2014).

Os motores de buscas tradicionais utilizam habitualmente buscas heurísticas ou critérios de relevância durante a apresentação dos resultados aos usuários. Por outro lado, a busca facetada, possível com o uso do SPARQL, se contrapõe a esses mecanismos, pois possibilita ao usuário navegar nas hierarquias das classificações,

generalizando ou especificando o nível das informações apresentadas e permite ao usuário uma navegação expressivamente interativa com os resultados apresentados, podendo verificar as relações existentes entre as informações pesquisadas (BÖRNER *et al.*, 2012).

O motor de busca XQuery é baseado na linguagem XML, que foi projetada para suportar otimizações poderosas e pré-compilação, levando a pesquisas muito eficientes em grandes quantidades de dados, incluindo os chamados bancos de dados nativos XML que leem e escrevem XML. Esta linguagem possibilita ainda consulta e desenvolvimento de aplicativos, capazes de processar o conteúdo de informações de diversas fontes de dados, incluindo documentos estruturados e semiestruturados, bancos de dados relacionais e bases de dados em árvore (WORLD WIDE WEB – W3C, 2019).

Para a edição de ontologias a ferramenta ontológica mais difundida e utilizada, segundo Stanescu (2018) é o editor Protégé (<http://protege.stanford.edu/>), desenvolvido pelo Stanford Research Centre for Biomedical Informática. Foi especialmente concebida para o desenvolvimento e o apoio de ontologias, permitindo aos cientistas reutilizar ontologias (STANESCU, 2018).

No âmbito nacional, em 2007 foi desenvolvida a ferramenta OntoKEM, no Laboratório de Engenharia do Conhecimento do Programa de Engenharia e Gestão do Conhecimento da Universidade Federal de Santa Catarina, para apoiar atividades de especificação e aquisição do conhecimento no domínio do setor elétrico (RAUTENBERG *et al.*, 2010). A OntoKEM foi baseada em processos de desenvolvimento da metodologia 101, da On-to-Knowledge e da METHONTOLOGY (RAUTENBERG *et al.*, 2010).

As tecnologias da Web Semântica são usadas na estruturação de portais semânticos para busca em diferentes áreas; os portais semânticos e sua caracterização são o objeto da Seção seguinte.

3.4 PORTAIS SEMÂNTICOS DE BUSCA – OBJETIVOS, CARACTERÍSTICAS, INDICADORES DE AVALIAÇÃO E EVOLUÇÕES

Delineadas as tecnologias semânticas de estruturação de portais, apresenta-se na Seção seguinte os propósitos e os traços – alguns dos aspectos que marcam os portais semânticos.

3.4.1 Objetivos, características e evolução dos portais semânticos

Os portais semânticos de busca visam o gerenciamento de ontologias desenvolvidas a fim de melhorar a busca e a recuperação destas na Web Semântica (FAHAD *et al.*, 2011). Alguns exemplos destes portais são apresentados a seguir:

- O portal semântico DataFiland, foi proposto por Frosterus, Hyvonen e Laitio (2011) com o objetivo de disponibilizar um modelo de criação de conteúdo distribuído com ferramentas para anotação e publicação de metadados sobre dados abertos e o sistema tem sido utilizado por desenvolvedores para reuso de dados governamentais da Finlândia. O esquema de metadados do portal foi baseado em uma modificação da versão do vocabulário void usado para descrição de conjunto de dados RDF. As anotações são feitas usando o editor de metadados em linha SAHA, conectado ao banco de ontologias ONKI, que forneceu um conjunto controlado de conceitos de anotação. Em uma pesquisa, o conteúdo é publicado imediatamente e a navegação é feita pelo motor de busca HAKO com interface para o usuário e também por um arquivo SPARQL para leitura pelas máquinas.
- Paiva, Costa e Silva (2014) propuseram uma abordagem híbrida usando sistemas de recomendação baseado no perfil do usuário combinado com o uso de ontologia de domínio. Um protótipo de portal foi concebido e implementado para o domínio do mercado comprador *online* de veículos de uma concessionária. O portal foi concebido para recomendar os melhores anúncios disponíveis na Web incluindo preço, modelo, marca, índice de depreciação e preço do seguro, de acordo com o perfil de interesse do usuário. Foi construída uma ontologia para anúncio de veículos, sendo reaproveitados conceitos ontológicos de outros modelos da área já existentes. Foram avaliadas a precisão e a recuperação de informações no portal, apontando índices de 86,66% e 68,42%, respectivamente. Os resultados mostraram que houve melhora nas pesquisas, particularmente no que diz respeito a precisão dos resultados fornecidos aos usuários do portal.
- O Semantic Web Environmental Directory – SWED é um portal semântico desenvolvido pelo projeto europeu Semantic Web Advanced Development – SWAD Europe, com o objetivo de criar um sistema que seja escalável para manutenção, permitindo o desenvolvimento de portais semânticos (SWAD,

2021). O SWED permite que cada organização seja responsável pela inclusão e manutenção de suas informações de forma distribuída e posteriormente agregadas. As informações são disponibilizadas pelas organizações em seus servidores e podem ser geradas de diferentes maneiras: a partir de um formulário Web para a geração de um arquivo RDF, de um editor de texto ou a partir de um Banco de Dados. Com isso, os robôs do portal percorrem esses servidores em busca de novas informações. Posteriormente, essas informações são recuperadas e analisadas pelo SWED. Como cada organização tem interesse em manter suas informações atualizadas, o portal permanece sempre atualizado e sem intervenção humana. O portal SWED é composto por uma ontologia chamada Operational Area Ontology, que tem como objetivo descrever propriedades das organizações e projetos utilizados pelo SWED. Esta ontologia define as áreas que estão contidas dentro de outras áreas, usadas para gerar a faceta de navegação da área operacional do SWED. Tem como suporte um dicionário de palavras e sinônimos em um determinado domínio (SWAD, 2021).

- O OntoSearch é um portal que faz buscas semânticas, busca por estruturas e instâncias de ontologias. O objetivo do OntoSearch é facilitar a reutilização do conhecimento disponibilizando quantidade significativa de arquivos com estruturas e instâncias de ontologia disponíveis na Web Semântica para consulta rápida e eficaz a partir de uma ferramenta *online*. Os resultados da pesquisa conterão ontologias e outros documentos semânticos para permitir que engenheiros do conhecimento possam navegar por documentos relevantes e encontrar informações suplementares sobre um assunto, se uma ontologia completa não estiver disponível (ZHANG; VASCONCELOS; SLEEMAN, 2005). O OntoSearch é baseado em Java, JSP, Jena e tecnologias JBoss. Pode ser utilizado combinado com motores de busca, para encontrar as ontologias. O OntoSearch permite pesquisa de ontologias disponíveis em linguagens RDF, OWL e DAML (+ OIL), possibilitando ao usuário realizar pesquisas com palavras-chave em determinados tipos de arquivos e inspecionar visualmente os arquivos para verificar a sua relevância (ZHANG; VASCONCELOS; SLEEMAN, 2005).
- O portal Swoogle foi desenvolvido pela Universidade de Maryland, segundo Nazário (2013) e possibilita buscas realizadas nas ontologias disponíveis na

Web, em instâncias de ontologias, em URIs nos vocabulários da Web Semântica, além de cruzar ontologias. Consiste em um mecanismo de indexação e recuperação de documentos de Web Semântica que indexa documentos RDF e OWL em vez de HTML puros. Extrai metadados para cada documento descoberto e computa relações, incluindo similaridades, entre documentos, segundo Nazário (2013). Apresenta serviços de pesquisa que suporta restrições em URLs de documentos de Web Semântica e classes/propriedades usadas ou definidas por elas. Indexa classes e propriedades definidas pelos documentos e ainda apresenta dados estatísticos baseados nos metadados de documentos semânticos descobertos (NAZÁRIO, 2013).

- O portal OntClassfire foi desenvolvido por Fahad *et al.* (2011) com o objetivo de propor uma ferramenta de gestão de ontologia para classificar as ontologias disponíveis. O portal utiliza o conhecimento semântico “escondido” dentro das ontologias e não apenas algoritmos tradicionais de texto simples ou páginas em HTML para classificar ontologias (FAHAD *et al.*, 2011). Cada categoria predefinida é considerada como um domínio para classificar objetos, documentos, dados da Web e das próprias ontologias. A classificação ontológica aumenta a precisão da classificação e, por conseguinte, a localização e acesso às ontologias.
- Cruz *et al.* (2019) apresentaram um protótipo do SemanticSUS, um portal semântico para acesso, análise e visualização de grande quantidade de dados do Sistema Único de Saúde (SUS). O Portal SemanticSUS foi estruturado em camadas: a) fontes de dados de diferentes tipos e proveniência; b) extração e anotação de metadados que descrevem o formato das fontes de dados; c) ontologia de domínio, com vocabulário formal e explícito compartilhado para a anotação semântica das fontes de dados, utilizando-se para modelagem a linguagem OWL; d) integração semântica da ontologia de domínio, de ontologias locais e a base de dados; e) integração de dados com a disponibilização de serviços de acesso às fontes de dados locais por meio de visão ontológica. A pesquisa no portal é efetuada a partir da estrutura facetada da ontologia.
- Hyvönen *et al.* (2021) apresentam um portal semântico intitulado *Mapping Manuscript Migrations* (MMM) que possibilita a modelagem, a agregação, a

publicação e o estudo de bancos de dados de manuscritos na área de humanidades digitais. O portal está vinculado ao Banco de Dados de Manuscritos Schoenberg do Instituto de Estudos de Manuscritos Schoenberg (Universidade da Pensilvânia) e aos Manuscritos Medievais das Bibliotecas de Oxford e Bibale do Institut de recherche et d'histoire des textes em Paris. A estrutura do portal segue o Modelo Sampo, criado para uma série de portais na área de humanidades digitais e que fornece aos editores de manuscritos uma forma colaborativa de enriquecer conteúdos com dados relacionados de outros fornecedores por meio SPARQL, mecanismo de raciocínio lógico também utilizado na pesquisa semântica facetada para facilitar a exploração dos dados pelo usuário final.

Com a difusão do conceito semântico para estruturação computacional e para recuperação de informações, também os portais de diversas áreas passaram a utilizar modelos ontológicos em sua estruturação. A seguir destacam-se os portais semânticos da área de educação, objeto maior de interesse desta pesquisa:

- O OntoWeb Portal, segundo Spyns *et al.* (2002), é um portal da comunidade acadêmica voltado para companhias e institutos de pesquisa da União Europeia que compartilham interesses na Web Semântica. Ele foi criado como parte do projeto OntoWeb delineado com aplicações do servidor ZOPE que fornece funcionalidades básicas de gerenciamento de conteúdo (SPYNS *et al.*, 2002). Um servidor remoto foi usado para armazenar informações utilizadas por aplicações distribuídas em um servidor central, possibilitando a reutilização de informações em diferentes portais semânticos que usam a tecnologia do portal OntoWeb (SPYNS *et al.*, 2002). A ontologia usada no OntoWeb representa uma ontologia de aplicação, o que significa que fornece todos os aspectos ontológicos que são necessários para fornecer a terminologia, a fim de executar o portal. O portal não se encontra disponível para consulta neste momento.
- O Portal UNI-POINT foi desenvolvido por Jorge (2005) para a comunidade ligada à Universidade Presidente Antônio Carlos – UNIPAC, localizada no estado de Minas Gerais e com vários *campi*. O objetivo do portal foi colocar os atores em contato com seus pares via Web, para troca e compartilhamento de informações e conhecimento. Foi utilizada a ontologia ONTO-POINT que

modela o ambiente, suas unidades, os cursos oferecidos, o projeto pedagógico de cada um desses cursos, além de classificar os atores envolvidos no dia a dia da Universidade (JORGE, 2005). A ontologia foi construída pela ferramenta Protégé, desenvolvida pela Universidade de Stanford. O portal UNI-POINT oferece os seguintes serviços: montar perfil de usuário e grupos de usuários; encontrar usuário e grupos de usuários; permitir colaboração entre os usuários; publicar conteúdo; recuperar conteúdo; navegar semanticamente. O portal não se encontrava disponível para consulta, quando da pesquisa em tela.

- O Portal Esperanto foi proposto como um estudo de caso de um grupo de pesquisadores da Faculdade de Informática da Universidade Politécnica de Madrid (LAUSEN *et al.*, 2005). O projeto do portal incorporou uma arquitetura clássica de três camadas e incluiu um banco de dados Oracle, conectado a um servidor que executa aplicações de gerenciamento de serviços e de acesso ao repositório da ontologia WebODE. Os documentos são armazenados em um servidor de arquivos e cinco diferentes ontologias de domínio foram desenvolvidas, incluindo projeto, documentação, pessoa, organização e reunião. Estas ontologias descrevem projetos de pesquisa e desenvolvimento que ficam disponíveis para reutilização. As ontologias são conectadas por meio de várias relações e apresentam profundidade de até cinco níveis, cada um deles comportando de dois a oito conceitos. Todas as instâncias podem ser modificadas, usando a interface do portal, embora apenas os administradores e membros registrados com permissão apropriada possam executar modificações. As buscas no portal eram feitas por palavras-chave, mas o usuário poderia navegar pela ontologia e recuperar todas as instâncias para construção de conceitos e subconceitos em cada etapa de navegação (LAUSEN *et al.*, 2005). O portal não se encontra disponível para consulta, no momento.
- O portal educacional semântico POSEDU, foi desenvolvido por Lachtim *et al.* (2009) e gerado a partir do sistema SiGePoS que foi baseado em uma arquitetura genérica, criada para a geração de portais semânticos. O portal foi fruto de uma dissertação de mestrado e não se encontra mais disponível na Web. A organização do conteúdo do POSEDU foi efetuada a partir de dois módulos funcionais do sistema SiGePoS: 1) módulo que permite a publicação de conteúdo semântico dinâmico, e 2) módulo que possibilita a publicação de

conteúdo dinâmico com base na descoberta de conhecimento em banco de dados Knowledge Discovery in Database. O primeiro módulo enfatizava a instanciação de conteúdo, isto é, fornece e organiza conteúdos para o Portal recuperados a partir da Web Semântica por meio de uma ontologia educacional. O segundo módulo usa a mesma ontologia para classificar e agrupar os documentos recuperados da Web tradicional, a partir de técnicas de mineração de dados, para a sua exposição no portal. A partir do POSEDU, informações oriundas de outros portais educacionais na Web são integradas via mapeamento de ontologias. Evidencia-se a organização e a instanciação de conteúdos tanto por meio de documentos semânticos oriundos da Web Semântica quanto pela Web aberta, pela classificação automática de documentos disponíveis. O Portal POSEDU, segundo Lachtim *et al.* (2009), foi baseado em uma ontologia leve chamada BOEDU, que por sua vez, foi baseada em uma análise exaustiva de cursos de graduação e pós-graduação brasileiros e de outros países responsáveis por atividades acadêmicas, envolvendo também grupos de pesquisa, projetos e financiamento. Esta ontologia base incorporou também a funcionalidade do Swoogle dos motores de busca Yahoo para obter informações. O portal não se encontra disponível para consulta.

- O Portal Educational Information Intelligent Search – EIIS foi projetado por Huang *et al.* (2011) com o objetivo de fornecer uma máquina de busca semântica inteligente para recuperação de informações e oferta de serviços no contexto educacional. O sistema de EIIS foi desenvolvido a partir de ferramenta de acesso aberto J2EE e funciona nos ambientes JDK 1.6.0_05 e Tomcat 5.5.26. A plataforma semântica KIM é usada para anotação de pesquisas heterogêneas. A base integrada de ontologia é desenvolvida pelo sistema Protégé 3.3.1 e o algoritmo EIISReasoning é remodelado com base em Jena (HUANG *et al.*, 2011). O portal é baseado no modelo SWIUSM, desenvolvido pelos autores supracitados e que inclui o algoritmo EIISRanking que contém regras relevantes para alinhar conceitos semânticos. Além disso, o algoritmo EIISRanking classifica conceitos em dois níveis a partir de um fator de semelhança semântica e mesma relevância que é aplicado para diferenciar termos importantes do campo educacional (HUANG *et al.*, 2011). A ontologia foi construída manualmente, uma vez que não havia uma ontologia aplicável

no domínio do assunto de tecnologia educacional. Para compartilhar extensamente vários conceitos deste assunto, todo o vocabulário especializado básico da tecnologia educacional foi extraído e considerados de primeiro nível; os conceitos considerados de alta frequência de uso (estabelecidos mediante análise de aplicações na Internet e presentes em trabalhos originais especializados) foram coletados e classificados em nível secundário; foram considerados ainda conceitos relacionados por docentes e peritos da área educacional. Todos estes conceitos foram categorizados durante a construção da ontologia. O portal não se encontra disponível para consulta.

- Um protótipo de portal educacional semântico russo foi proposto por Telnov (2015), baseado no modelo RDF, ontologia e mecanismo de consulta SPARQL. O portal inclui módulos com informações sobre departamentos educacionais e gráficos do conhecimento organizados pelo Ontotext Graph DB. A interface do usuário permite navegar através dos gráficos de conhecimento disponíveis no portal que são complementados por bases de conhecimento internacionais como o Dbpedia, wikidata e outros repositórios de conhecimento especializados. As ferramentas do portal permitem que um aluno ou professor acessem facilmente lugares específicos do gráfico de conhecimento para encontrar objetos de aprendizagem necessários. O princípio de pesquisa é muito semelhante aos de motores de busca populares (Google, Yandex, entre outros). Quando o usuário digita as letras da palavra-chave na linha de entrada, o sistema lança uma lista de conceitos relevantes do gráfico de conhecimento, sendo possível escolher o conceito mais adequado e aceder diretamente para a área desejada do gráfico. O uso do navegador RDF possibilita que o usuário navegue nos gráficos a partir de seus nós e ao se concentrar em um nó específico do gráfico, é possível obter ligações de metadados de texto, conteúdo de mídia e hipertexto que estão associados a esse nó. O portal não se encontra disponível para consulta.
- Terziyan, Golovianko e Shevchenko (2015), lançaram um portal semântico envolvendo informações educacionais de quatro universidades ucranianas. A estrutura do portal inclui uma base de conhecimento estruturada a partir de ontologia, RDF, OWL e um mecanismo de raciocínio lógico. Além de disseminar conteúdo do domínio educacional foi projetado um mecanismo para avaliação da qualidade da educação ucraniana, a partir de avaliações feitas pela

comunidade usuária do portal. O mecanismo de avaliação é baseado em tecnologias de comunicação dos usuários com a instituição e cada usuário recebe uma pontuação associada à sua reputação, de acordo com sua interatividade no portal. O portal é gerenciado com base na atividade social dos usuários, a partir de avaliações de qualidade dos conteúdos disponibilizados e também dos aspectos relacionados de qualidade da educação. O portal não se encontra disponível para consulta no momento.

- Borges e Silveira (2019) apresentaram o portal semântico educacional AvaOne, projetado para recuperação de vídeos educacionais por alunos e professores a partir de cadastro com *login* e senha. A arquitetura do portal é formada por três bases: a) banco de dados semântico (Metadados, OWL, RDF) que armazena os metadados dos descritores segmentados em RDFs com base em ontologias, o que permite a pesquisa e recuperação dos vídeos; b) banco de dados multimídia, com armazenamento de arquivos em servidor; c) banco de dados de usuários incluindo perfis e outras informações. O portal possibilita a segmentação e anotação de vídeos, a partir de um processador de vídeo usado para acesso aos dados armazenados no banco de multimídia. A pesquisa semântica é efetuada por meio de ontologia e mecanismo de inferência SPARQL. O portal não se encontra disponível para consulta.

Comparativamente, em termos de recuperação da informação, conforme já citado anteriormente, nos portais tradicionais a busca por informações recupera documentos a partir de relações sintáticas, ou seja, da sintaxe das palavras-chaves sem considerar o significado destas. Já nos portais semânticos a recuperação da informação é feita a partir de relações semânticas, resgatando o significado de termos de buscas e associados ao conteúdo dos documentos. Os portais semânticos apresentam, contudo, características diferenciadas em termos de tecnologias, sistema de busca, forma de gestão de conteúdo, navegação, processamento, publicação e atualização da informação, como se depreende dos exemplos apresentados.

Considerando-se que a educação é fundamental para o desenvolvimento humano, os portais educacionais semânticos, no entanto, devem ser ambientes que forneçam dados, informações e objetos de aprendizagem e que impulsionem devem ser ambientes que forneçam dados, informações e objetos de aprendizagem e que

impulsionem o conhecimento de forma facilitada e amplamente disseminada (BORGES; SILVEIRA, 2019; SCHLEICHER, 2006; TERZIYAN; GOLOVIANKO; SHEVCHENKO, 2015).

Sintetizando-se, então, apresenta-se uma comparação das principais características dos portais semânticos frente aos portais tradicionais concebidos de acordo com os preceitos da Web Semântica propostos inicialmente por Bernes-Lee, Hendler e Lassila (2001), Lausen *et al.* (2005) e atualizados via W3C (2019) - Quadro 1.

Quadro 1 - Características de portais tradicionais e de portais semânticos – comparativo.

CARACTERÍSTICAS	PORTAL TRADICIONAL	PORTAL SEMÂNTICO
TECNOLOGIAS	Tecnologias convencionais	Tecnologias de base de padrão aberto para facilitar a interoperabilidade, ou seja, a colaboração entre os envolvidos no processo de comunicação. Uso de tecnologias da Web Semântica
BUSCA	Pesquisa por texto livre e hierarquia estável de classificação	Pesquisa multidimensional por significado de uma ontologia rica de domínio
NAVEGAÇÃO	Hierarquia fixa de classificação e manutenção centralizada	Multidimensional com facetas, informações semiestruturadas e extensíveis, com alterações descentralizadas
ATUALIZAÇÃO	Adição de informações anotações	Adição de novas classificações pela comunidade de usuários para expansão da estrutura da informação
GESTÃO DE CONTEÚDO	Gestão centralizada	Gestão descentralizada com possibilidade de múltiplas agregações e visões do mesmo dado
PROCESSAMENTO/ ACESSO	Processamento manual	Processamento automatizado para acesso de agentes Inteligentes e facilitar integração entre portais
PUBLICAÇÃO	Manual ou via formulários específicos e cada portal é suprido e mantido separadamente	Automática integração de ontologias; provedores publicam dados em formulários reutilizados pela organização em múltiplos portais.

Fonte: Adaptado de Reynolds, Shabajee e Cayzer (2004, p. 290) e Correia (2012, p.36).

Todas essas características semânticas associadas aos portais semânticos atendem as normativas da Web Semântica (BERNERS-LEE; HENDLER; LASSILA, 2001) e estão associadas ao potencial semântico do portal, conforme conceito introduzido nesta Tese, inferido a partir da literatura (BERNERS-LEE; HENDLER; LASSILA, 2001; WORLD WIDE WEB - W3C, 2021, 2019). Delineada a estruturação dos portais semânticos com suas tecnologias e ferramentas, apresenta-se na próxima

Seção, a definição e as funções dos indicadores, com ênfase no uso destes para avaliar o potencial semântico de portais.

3.4.2 Indicadores de avaliação do potencial semântico

Os indicadores são instrumentos usados para identificar e medir aspectos relacionados a um determinado conceito, fenômeno, problema ou resultado de uma intervenção na realidade (BRASIL, 2010). Os indicadores devem traduzir, de forma mensurável, determinado aspecto de uma realidade dada ou construída, tornando operacional a sua observação e avaliação (BRASIL, 2010). De acordo com a Organização para a Cooperação e o Desenvolvimento Económico - OCDE (2021), os indicadores fornecem informações que ajudam a simplificar a realidade, auxiliam a compreensão de fenômenos complexos, detectam padrões de comportamento e tendências de futuro, propiciando a tomada de decisões de forma consciente.

Os indicadores possuem duas funções básicas a saber (BONNEFOY; ARMIJO 2005): (i) função descritiva: arrolar informação sobre uma determinada realidade empírica, situação social ou ação pública; (ii) função valorativa ou avaliativa: implica em agregar informação de juízo de valor à situação em foco ou verificar a adequação de desempenho de um programa.

Os indicadores de qualidade ou de desempenho são usados para medir ou avaliar na prática a *performance* de um programa ou sistema (AMBROZEWICZ, 2015). Na construção de indicadores qualitativos, objeto desta tese, a partir de estratégias quantitativas busca-se mensurar valores, opiniões, relações e vivências intersubjetivas (MINAYO, 2009). Os indicadores qualitativos podem ser validados por diferentes escalas como a de Thurstone (1928), que se baseia em questionários para coleta de opiniões de concordâncias ou discordâncias de certas afirmações cujos resultados serão submetidos a especialistas no tema estudado e que definirão as tendências comportamentais.

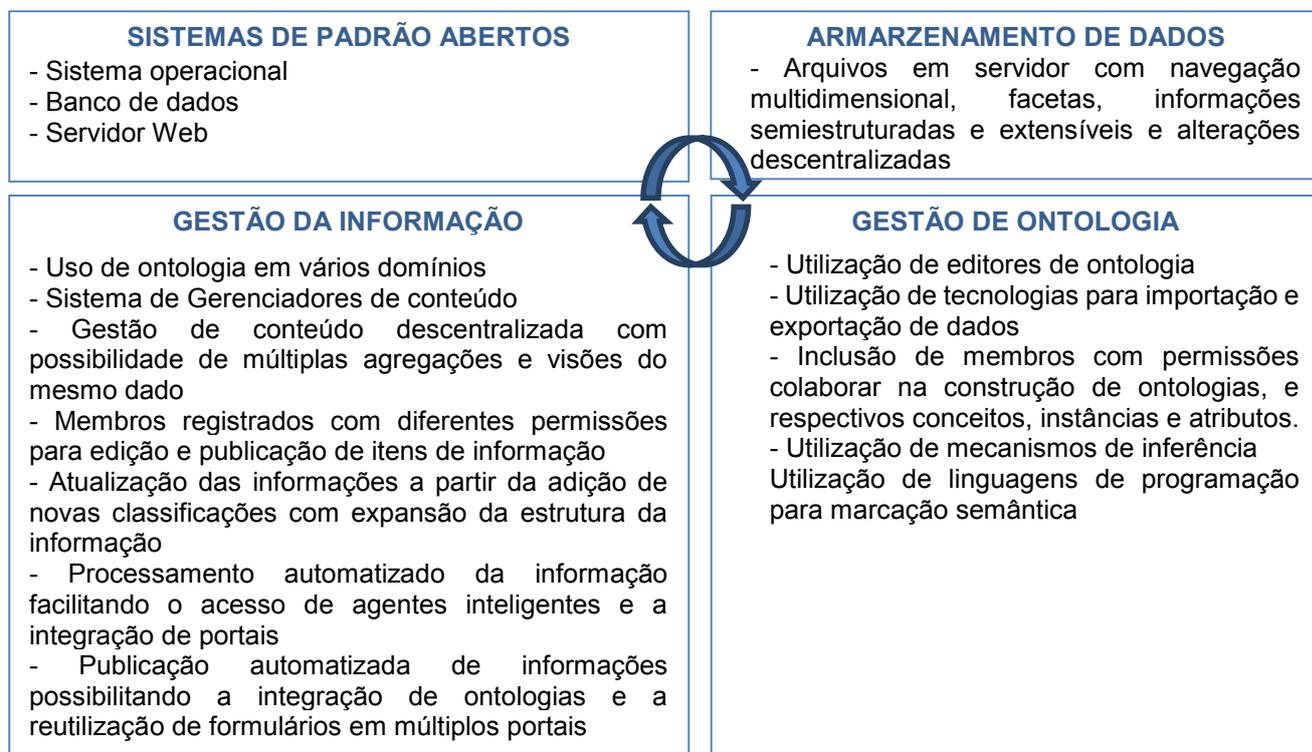
A partir da análise crítica do recorte da revisão da literatura objeto deste estudo, e da sua revisão sistemática, iniciada em Guedes (2015) e aqui aprofundada, postula-se que o potencial semântico está correlacionado a qualidade de resposta obtida quando da busca de informações em portais. O potencial é resultado do uso de ferramentas e de tecnologias específicas o que retornaria serviços aos usuários compatíveis com a estruturação da Web Semântica.

O modelo semântico de recuperação da informação é, conforme mencionado recursivamente, fundamentado no uso da ontologia formal (GUARINO, 1997; GRUBER, 1996) e das ferramentas técnicas propostas pela Web Semântica, podendo conferir aos portais da Web a qualidade semântica requerida. Neste escopo os portais semânticos do conhecimento de universidades podem melhor atender as necessidades e interesses dos seus usuários finais (BERNERS-LEE; HENDLER; LASSILA, 2001; MAEDCHE *et al.*, 2001; TELNOV, 2015).

Com base nos estudos da Web semântica de Berners-Lee, Hendler e Lassila, (2001) e do modelo de avaliação de portais semânticos descrito por Lausen *et al.*, (2005), Guedes (2015) e Guedes e Strauhs (2016), propuseram um quadro referencial de condições mínimas para estabelecimento de potencial semântico dos portais do conhecimento de universidades, conforme Figura 6.

O quadro referencial compõe-se de quatro dimensões a serem consideradas: (i) sistemas abertos – incluem tecnologias básicas de padrão aberto a serem utilizadas; (ii) armazenamento de dados – inclui forma de armazenamento e de organização da informação; (iii) gestão da informação – inclui o uso de ontologia, forma de gestão de conteúdo e publicação, atualização e processamento de informações; (iv) gestão da ontologia – inclui editor de ontologia, tecnologias para marcação semântica e ferramentas para controle e colaboração na construção de ontologias.

Figura 6 - Condições mínimas para estabelecimento de potencial semântico de portais.



Fonte: Guedes e Strauhs (2016).

Sob a ótica sociotécnica, um portal semântico pode ser considerado um artefato complexo construído por atores humanos e não humanos com o objetivo de disponibilizar informações acadêmicas para a comunidade universitária e a sociedade em geral. (CALLON; LATOUR, 1981; BURNES, 1986). Assim, considerando as dimensões e as respectivas características elencadas por Guedes e Strauhs (2016), no quadro referencial citado anteriormente, e adicionando os aspectos sociais envolvidos, apresenta-se na Figura 7, potenciais indicadores qualitativos para análise de potencial semântico de portais.

Figura 7 - Potenciais indicadores qualitativos semânticos.

POTENCIAIS INDICADORES SEMÂNTICOS			
ARMAZENAMENTO DE DADOS	GESTÃO DA INFORMAÇÃO	GESTÃO DE ONTOLOGIA	SISTEMAS SOCIOTÉCNICOS
Tecnologias de padrão aberto	Ontologia Navegação por facetas Mecanismo de inferência RDF - OWL Gestão de conteúdo Atualização da informação	Colaboração Comunicação Editor de ontologia	Equipe técnica Usuários

Fonte: A autoria própria (2021).

Os potenciais indicadores semânticos estão relacionados às condições necessárias que favorecem a construção ou manutenção de portais que disponibilizem informações atualizadas e pertinentes do domínio, com recuperação rápida e precisa, além de oferecer outros serviços e facilidades de colaboração e comunicação entre os diferentes atores envolvidos.

Considerando-se então estabelecido um quadro referencial propício ao desenvolvimento da pesquisa em tela, no Capítulo seguinte apresenta-se o alinhamento conceitual das teorias tratadas no tema em estudo e de como elas estão inter-relacionadas para embasar e justificar a consecução desta pesquisa.

4 ALINHAMENTO CONCEITUAL

Este Capítulo apresenta conceitos e características dos sistemas técnicos e sistemas complexos concebidos para atender diferentes demandas da sociedade. Na visão da tecnologia social, apresenta-se os portais do conhecimento de universidades como artefatos sociotécnicos já que agregam pessoas, tecnologias e processos para disseminação da informação no domínio da educação.

4.1 SISTEMAS SOCIOTÉCNICOS

Os sistemas técnicos, de acordo com Martínez-García e Hernández-Lemus (2013), constituem sistemas complexos com características de modularidade e robustez. São formados por estruturas hierarquicamente organizadas, incluindo sistemas compostos e subsistemas que abarcam uma coleção de constituintes individuais (ou agentes) que se articulam por meio de interações em que um agente pode afetar as condições em que outros agentes se comportam. Por isso, a abordagem reducionista que tem sido a base da maioria das ciências ocidentais apresenta limitações já que não se pode simplesmente decompor um sistema em uma coleção de subsistemas simples e de fácil análise (MARTÍNEZ-GARCÍA; HERNÁNDEZ-LEMUS, 2013).

Segundo Hughes (1989) em um sistema sociotécnico, as relações são quase sempre permeadas por conflitos, cujas resoluções são sempre o resultado de negociações. Assim, as controvérsias científicas são aspectos interligados na construção do sistema juntamente com aquelas advindas da sociedade. Esse autor discorreu também sobre a complexidade dos sistemas, destacando duas grandes situações. A primeira trata das falhas ou obstáculos que surgem durante o desenvolvimento do sistema e que necessitam de uma ação coletiva e concentrada para corrigi-las, a fim de que o processo de expansão do sistema tenha continuidade. A segunda refere-se à aceitação do desenvolvimento de uma tecnologia pelos indivíduos e instituições relacionados a ela e quando o contexto é favorável. A partir desse momento, os sistemas técnicos expandem-se rapidamente, adquirindo uma espécie de autonomia.

Revela-se, portanto a necessidade do desenvolvimento, e do uso de ferramentas tecnológicas para resolver problemas sociais, com oferta de produtos voltados para a necessidade da sociedade (DAGNINO, 2014), a prevalência, portanto, das tecnologias dita sociais. Nesse contexto, os sistemas sociotécnicos estão permeados de normativas voltadas para o uso benéfico e igualitário da tecnologia (BURNS, 2006).

De acordo com Burns (2006), a tecnologia é entendida como um conjunto complexo de artefatos físicos, juntamente com sistemas de regras empregadas pelos atores sociais para utilizar e gerir esses artefatos. A tecnologia tem uma face cultural/institucional e outra face material em que algumas regras integram um conjunto de instruções para que a tecnologia funcione e seja gerenciada de forma eficaz. Essas regras têm um caráter prático e podem ser distinguidas de outras regras que envolvem arranjos culturais e institucionais inerentes aos sistemas sociotécnicos em que a tecnologia está inserida (BURNS, 2006). Esses sistemas de regras incluem leis e princípios normativos, tais como o uso legítimo ou aceitável da tecnologia, os proprietários e operadores apropriados ou legítimos, os locais e tempos de uso, as perdas e os ônus da aplicação das tecnologias (BURNS, 2006).

Os elementos técnicos e sociais imbricados caracterizam a tecnologia como uma interação de escolhas humanas, ações, histórias sociais e contextos institucionais, refletidos em produtos físicos e em resultados voltados toda a sociedade (DAGNINO, 2010; LATOUR, 1994b). Nesse contexto, tecnologia social pressupõe a resolução de problemas sociais, a partir do uso de tecnologia, e está estritamente relacionada e indicada para aplicações que envolvem os sistemas sociotécnicos (DAGNINO, 2010).

Para Burns e Dietz (2001), as diversas estruturas técnicas e físicas que fazem parte de sistemas sociotécnicos são geridas e organizadas por diferentes atores com diversos profissionais, compondo uma variedade de grupos, de redes sociais e de entidades envolvidas na construção, na operação e na manutenção de sistemas sociotécnicos. Os atores humanos, incluindo indivíduos, grupos, organizações, comunidades e outras coletividades são os produtores, transportadores e reformadores de sistemas de regras sociais (BURNS; DIETZ, 2001). Eles interpretam, adaptam, implementam e transformam as regras de forma cautelosa, e às vezes, até radicalmente.

O papel dos atores sociais, segundo Burns (2006), é o de compreender e de prever o que acontece em um determinado contexto social, a fim de justificar, explicar e criticar uma ação e/ou suas consequências. Para tanto, esses atores estabelecem e utilizam normas e sistemas de regras apropriados para cada situação e, ao longo do tempo adquirem experiências e habilidades para adaptar ou reformar regras em ambientes de interação, a partir de políticas situacionais para administrar processos. Se houver barreiras ou distorções nesses encadeamentos, falhas podem ocorrer já que o fator humano seria responsável pelo bom desempenho do sistema de forma que tais falhas podem estar relacionadas às características organizacionais e de comunicação que são difíceis de analisar e de compreender. Assim define-se o cenário para a luta social, o exercício do poder de impor ou resistir às regras, e a negociação para mudanças nas regras.

Para Iyamu e Mgudlwa (2018), as teorias técnicas sociais são usadas geralmente em estudos de sistemas de informação e de tecnologia para analisar os sistemas complexos. Os sistemas complexos envolvem tecnologias, pessoas, processos, registro de dados e atividades que processam dados e informações para realização de tarefas em diferentes ambientes organizacionais (IYAMU; MGUDLWA, 2018). Os sistemas complexos são, portanto, considerados sistemas sociotécnicos que consistem em pessoas, tecnologias e processos atuando juntos para a realização dos objetivos dos sistemas (IYAMU; MGUDLWA, 2018).

Os sistemas sociotécnicos complexos incluem tarefas associadas ao aprendizado de máquina, inteligência artificial, algoritmos, *big data*, tecnologia de automação e robótica, aplicáveis a uma ampla gama de campos e os projetistas precisam entender as implicações desses desenvolvimentos para a prática de um desenho que envolva humanos e não-humanos (FORLANO, 2017). As considerações sobre o não-humano, coisas e o mundo artificial, exigem novas formas de conhecimento e abrem novos problemas, questões, oportunidades e soluções para o campo do desenho dos objetos (FORLANO, 2017).

As teorias sociotécnicas são usadas, portanto, para guiar e sustentar os estudos de um sistema de informações e baseiam-se na interconexão entre os componentes em que é necessária para uma compreensão mais aprofundada do impacto da implantação, do uso e do gerenciamento das tecnologias (IYAMU; MGUDLWA, 2018). Rose e Scheepers (2001, p. 217) destacaram a importância da visão social da tecnologia na formação dos profissionais da área de TI, afirmando que:

[...] “a teoria social tem um papel substancial a desempenhar no desenvolvimento da disciplina de TI principalmente ajudando a entender e interagir com os contextos sociais”.

Nos contextos aludidos as universidades são fundamentalmente a principal fonte criadora e disseminadora do conhecimento e da experiência social (SCHLEICHER, 2006), tendo os seus portais como mecanismo de difusão de informações. Um portal universitário semântico disponibiliza dados em constante atualização, impulsionado por conhecimentos em constante evolução e estruturado em bases que envolvem princípios de rede social (TERZIYAN; GOLOVIANKO; SHEVCHENKO, 2015). A função social dos portais semânticos é marcada pela ação de humanos interagindo com as tecnologias, conforme depreendido da Teoria Ator-Rede (CALLON; LATOUR, 1981; LAW, 1992; IYAMU; MGUDLWA, 2018). Assim os portais semânticos de universidades constituem artefatos sociotécnicos, na medida em que todos os atores da rede, especialmente os usuários, participam de forma colaborativa na manutenção e na evolução do portal como uma aplicação da Web Semântica (BURNS, 2006; FORLANO, 2017; WORLD WIDE WEB - W3C, 2021).

A próxima Seção aborda, então, a Teoria Ator-Rede que realça o efeito relacional e a interação entre humanos e não-humanos em redes heterogêneas, possibilitando uma análise social dos portais do conhecimento de universidades como artefatos sociotécnicos.

4.2 TEORIA ATOR-REDE

O desenvolvimento de novas metodologias está alinhado às práticas emergentes de desenhos técnicos que enfatizam as inter-relações entre atores humanos e não-humanos conforme apregoa a TAR.

A TAR baseia-se em preceitos defendidos por autores como Callon e Latour (1981) e Law (1992) e pressupõe que atores humanos e não-humanos possuem a mesma importância em um sistema, em razão da sua capacidade de agência. Desta perspectiva os seres humanos não são vítimas do processo ou protagonistas autônomos em ação, bem como também os objetos têm agência o que significa “ser associado de tal maneira que faz outros atores fazerem coisas” (LATOUR, 2012, p. 158), sendo regidos também pelo princípio da simetria generalizada que descarta sobreposições de um ator sobre o outro (LATOUR, 1994b).

Law (1992, p. 3) descreveu a Teoria Ator-Rede como teoria relacional e sociológica orientada a processos, que trata agentes, organizações e dispositivos como efeitos interativos. A sociedade é formada por redes padronizadas de materiais heterogêneos, “compostas não só de pessoas, mas também de máquinas, animais, textos, dinheiro, arquiteturas” enfim, de outro qualquer material (LAW, 1992, p. 3).

Segundo Iyamu e Mgodlwa (2018), a TAR é considerada uma teoria altamente influente dentro da Sociologia da Ciência que procura explicar e interpretar o desenvolvimento e a mudança sociotécnica. Na TAR, os atores são entidades humanas e não humanas, que têm a capacidade de mudar o estado das coisas, sendo esses atores vistos de forma igualitária, sem prioridade dada à tecnologia ou a questões sociais que a envolve. Pessoas, tecnologia e processos são considerados de igual importância e valor e, independentemente de o ator ser humano ou não-humano, ambos são igualmente ponderados, oferecendo a mesma contribuição para uma rede (IYAMU; MGUDLWA, 2018).

Com base na TAR, um modelo e uma estrutura de portal semântico podem ser desenvolvidos, criando-se uma plataforma que destaca a interação entre não humanos – objetos, artefatos, políticas – e humanos, a partir da qual evidencia-se que os elementos não humanos influenciam o comportamento humano e podem alterar as relações e as ações esperadas, modificando o desenho dos objetos técnicos (CALLON, 1986c; LATOUR, 2001).

Gao (2005) destacou que a TAR possibilita analisar dados qualitativos em processos de formulação de estratégias, examinando os componentes e os elementos que juntos formam os contextos sociais e tecnológicos e como eles estão dinamicamente conectados dentro de um ambiente. Shaanika e Iyamu (2015), deste prisma, usaram a TAR para desvendar fatores técnicos e não técnicos dentro de um sistema, examinando as ações individuais, os relacionamentos e as interações dentro de uma comunidade de arquitetura corporativa.

Para Latour (2012), não existe diferenciação entre homem e objeto, mas sim um híbrido homem-objeto em uma composição na qual o objeto é mais do que meramente uma ferramenta, antecedente ou estágio, nos quais os atores humanos desempenham seus papéis principais. Latour (2012) rejeita o próprio conceito de ator como substância, que é algo que preexiste – está acabado e perdura – e apregoa que os atores são eventos que não podem ser diferenciados de suas manifestações e relações.

A TAR se origina no campo da Sociologia, mas há muitos anos tem sido adotada por outros campos para estudar e interpretar diferentes tipos de fenômenos, como nos estudos de sistemas de tecnologia e informação (SHIM; SHIN, 2016; WALSHAM, 1997). De acordo com Latour (2000), a TAR aborda os elementos da ciência e da tecnologia em formação em oposição à ciência e à tecnologia prontas, descompactando e analisando os componentes elementares tais como atores, processos, política, ambiente, políticas e redes, que estão frequentemente associados ao desenvolvimento e implementação de artefatos tecnológicos.

A arquitetura da TAR, segundo Latour (2017), é baseada em três partes principais. A primeira é a compreensão semiótica da construção de entidades, sendo possível ver todos os atores como iguais, independentemente de suas características ou contextos. Estendendo a semiótica às coisas permite-se usar um quadro metodológico vazio que pode ajudar a seguir qualquer tipo de ator e desvendar cadeias de associações e permanecer entre formas descritivas e explicativas, que é a essência da segunda parte da arquitetura da TAR. E, finalmente, a terceira parte consiste no caráter ontológico de redes e atores, já anteriormente enfatizado por Latour (2001).

O conceito de rede remete, portanto, à ideia de alianças, fluxos e mediações resultantes da conexão entre elementos heterogêneos que podem ser materiais e humanos em constante interação (MAIA; SERAFIM, 2011). A reordenação da rede desmistifica o conceito de caixa-preta que denota a ideia da cibernética, de algo acabado, estabilizado, consagrado e inquestionável (CALLON, 1986). Para Latour (2001, p. 353), a caixa-preta, está associada à “maneira como o trabalho científico e técnico torna-se invisível decorrente de seu próprio êxito [...] paradoxalmente, quanto mais a ciência e a tecnologia obtêm sucesso, mais opacas e obscuras se tornam”.

A solidez de um fato depende do movimento de todos aqueles que o mantêm em movimento, os aliados (PEDRO, 2010) e, assim, as sucessivas caixas-pretas formam uma estrutura com macro e microatores. Todo actante é uma caixa-preta, que pode e deve ser aberta para revelar conexões, articulações e redes, como também destacou Latour (2012).

Erofeeva (2017) argumenta que a TAR considera os objetos tecnológicos híbridos (quase-objetos) que produzem outros tipos de objetos (naturais e sociais) e, a partir do padrão relacional da TAR, esses objetos podem ser descritos por diferentes ações dos atores que a eles estão conectados.

Segundo Latour (2001, p. 222), objetos “não-humanos podem entrelaçar-se com os humanos graças aos processos-chave da translação, articulação, delegação, deslocamento para fora e para baixo”. Para explicar os deslocamentos, esse autor traz os conceitos da semiótica, “quando o leitor é enviado de um plano de referência para outro, dá-se a isso o nome de deslocamento para fora; quando é trazido para o plano de referência original, deslocamento para dentro; quando o material expressivo é inteiramente modificado, deslocamento para baixo” (LATOURE, 2001, p. 348).

A TAR se fundamenta, à vista disso, em conceitos de tradução e deslocamento que refletem os interesses dos atores (CALLON,1986). Na tradução se definem interesses, responsabilidades, necessidades, mecanismos de ação e de regulação. O deslocamento é o movimento literal necessário para solidificar os atores da rede e que pode envolver “deslocamento de objetivos e interesses, deslocamentos de dispositivos e de seres humanos” (CALLON, 1986b, p.72).

O alinhamento dos interesses dos atores é feito por intermédio de um processo de negociação contínua em que ocorre a reinterpretação, a representação ou a apropriação dos interesses de um ator, de forma que estes possam ser seguidos pelos outros atores na rede (CALLON; LATOUR, 1981).

Na tradução ocorrem processos de identificar, mobilizar, atribuir papéis, além de influenciar atores formando uma aliança entre eles para que todos apoiem a construção de uma rede (LATOURE, 1992). Assim, as traduções constituem processos de análise de interesses comuns ou divergentes e como estes interesses são considerados nas estratégias para solucionar problemas (CALLON, 1986 a, b, c).

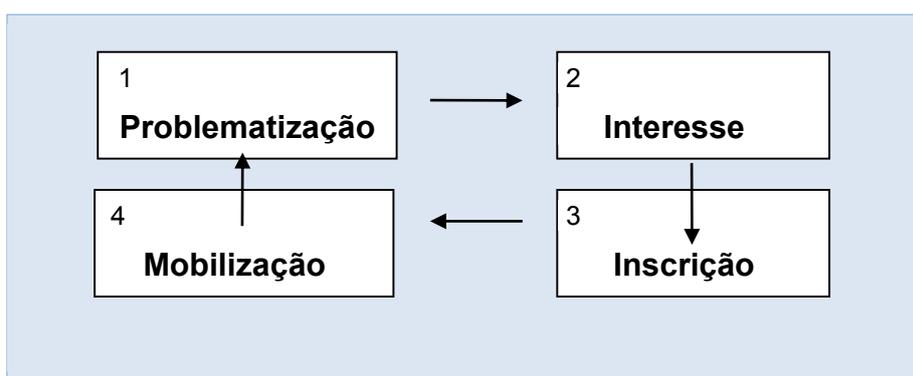
Segundo Prado e Baranauskas (2012), o processo de tradução consiste na construção de associações entre atores no qual há desejo de mudar um determinado estado de coisas, estimular os interesses dos atores a se unirem e definir papéis para garantir o cumprimento das responsabilidades. Na tradução uma atividade se transforma ou recebe um novo significado.

Elbanna (2012) enfatizou que a tradução é o mecanismo pelo qual o construtor de redes recruta atores e assegura suas associações fiéis. No entanto, a tradução não é determinista, pois o que os atores fazem quando se reúnem em uma rede é imprevisível. Cordella (2010) esclareceu que os atores são influenciados pela rede e, por sua vez, influenciam a rede negociando suas forças por meio da tradução.

Callon (1986a, p. 18) descreveu o movimento de tradução e deslocamento em que cada ator tem definido o seu papel dentro da rede, sendo que um destes assume

o papel de “porta-voz”, que interpreta os interesses de todos, propondo um caminho único a seguir, estabelecendo-se os “pontos de passagem obrigatórios” para os atores que desejam permanecer e se desenvolver na rede e solucionar problemas. Esse autor explicou os processos de tradução a partir de quatro fases: problematização, interesse, inscrição e mobilização, conforme apresentado na Figura 8.

Figura 8 - Fases da Tradução.



Fonte: Autoria própria, conforme Callon (1986).

A Problematização é a primeira etapa da tradução que lida com a identificação de um problema, que se destina a criar uma solução. De acordo com Callon (1986a), problematização é o processo em que um ator focal define um problema de uma maneira que permite que outros atores o reconheçam como seu próprio problema, convencendo-os a aderir à rede.

Interesse é o segundo estágio, no qual atores interessados se juntam a uma rede de problemas que foram propostos. Hu (2011) inferiu que essa fase da tradução constitui uma série de processos pelos quais os atores buscam bloquear outros atores nos papéis propostos para eles em um determinado programa.

Na fase de Inscrição alguns atores interessados dentro de uma rede começam a participar na busca de solução para o problema. Segundo Macome (2008), apenas atores que aceitam os papéis definidos para eles na etapa anterior participam da construção da rede.

A Mobilização é o último estágio da tradução, o que significa que todos os atores foram inscritos com sucesso na rede. Conforme Prado e Baranauskas (2012), quando os atores se conectam, as consequências do sucesso ou do fracasso se espalham pela rede, de modo que há interesse mútuo de que todos tenham sucesso. Callon (1986, p. 209) ressalta que é também nesta fase que um “porta-voz” consegue

falar previsivelmente em nome dos outros, que conseguem traduzir com sucesso os interesses, funções e relações de toda a rede. Esta etapa se traduz no estabelecimento bem-sucedido de uma rede de atores, com interesses alinhados.

De acordo com Latour (2017, p. 310), as redes são o ponto final da redução; não há nada além de redes que são todas contingentes na natureza. Esse autor explica sua abordagem baseando-se nos seguintes temas principais: a) remover a oposição entre distância e proximidade; b) substituir a metáfora da escala pela metáfora da associação, acabando por provocar o colapso da dicotomia “micromacro” (LATOURE, 2017, p. 371) ao remover a divisão entre o interior e o exterior e a substituição dessa divisão pela imanência das redes. Além desta topologia, a abordagem baseia-se na afirmação ontológica de que uma rede é um trabalho que é feito por atores que estão agindo, ou entidades que são acionadas.

No contexto da análise sociotécnica, Bijker, Hugh e Pinch (1989) analisaram a evolução histórica das redes constituídas por componentes complexos de interconexão que se ligam para formar infraestruturas como a ferroviária, a elétrica e, mais recentemente, a Internet. Eles destacaram que esses sistemas são moldados por forças sociais, e as tecnologias de rede são desenvolvidas dentro de contextos sociais, históricos e institucionais particulares.

Aprofundando, as redes, segundo Granovetter (1983), são constituídas por laços fortes e laços fracos. Nas redes de laços fortes há uma identidade comum, em que se procuram referências para a tomada de decisão. Os laços fortes são caracterizados por relações com alto nível de credibilidade e influência, mas as dinâmicas geradas nessas interações não se estendem além dos *clusters* e os indivíduos que compartilham esses laços fortes comumente participando de um mesmo círculo social. Já as redes de laços fracos são importantes porque conectam vários grupos, rompendo a configuração de *clusters* e funcionando como pontes pelas quais surgem e circulam as inovações. Sem as conexões de laços fracos, a tendência é o isolamento da rede ficando confinada em seus grupos. Nesse sentido, os laços fracos são vitais para a integração dos sistemas sociais, pois permitem que novas ideias se espalhem mais rapidamente e inovações surjam.

Uma rede científica ficará em desvantagem se tiver poucos laços fracos, com dificuldades e caminhos fechados para inovar. Por outro lado, as inovações precisam ser avaliadas por todos os membros da comunidade, o que remete ao papel dos laços fortes (GRANOVETTER, 1983). Assim, laços fracos possibilitam o acesso à

informação que é à base da inovação, mas essa depende também dos laços fortes para se concretizar (GRANOVETTER, 1983).

Infere-se, a partir deste escopo conceitual que todo o arcabouço da teoria social e da TAR pode ser aplicada na construção de portais do conhecimento, especificamente os portais do conhecimento de universidades, foco deste estudo. O quadro 2 apresenta o alinhamento das teorias consubstanciadas pela literatura.

Quadro 2 - Arcabouço conceitual.

CONCEITOS	AUTORES
Construção do portal do conhecimento de universidades como um artefato sociotécnico	Bijker; Hugh; Pinch, 1989; Callon, 1986a; Callon; Latour, 1981; Elbanna, 2012; Erofeeva, 2017; Granovetter, 1983; Hu, 2011; Hughes, 1989; Latour, 1992; Latour, 200; Latour, 2012; Latour, 2017; Macome, 2008; Prado; Baranauskas, 2012; Shim; Shin, 2016; Iyamu; Mgudlwa, 2018
Recuperação semântica da informação para atender os interesses dos usuários, considerando a grande quantidade de informação disponíveis na web.	Alomran, 2014; Amato <i>et al.</i> , 2017, Barba-González <i>et al.</i> , 2019; Benabderrahmane <i>et al.</i> , 2017; Brandt, Vidotti ; Santarem Segundo, 2018; Gargiulo <i>et al.</i> , 2014; Heath; Bizer, 2011; Li <i>et al.</i> , 2017; Malgaonkar; Devale, 2016; Malhotra; Nair, 2015; Moura <i>et al.</i> , 2015; Nadal <i>et al.</i> , 2019; Pattuelli; Miller, 2015; Raza <i>et al.</i> , 2018; Reyes-Álvarez; Roldán-García; Aldana-Montes, 2019; Silvello <i>et al.</i> , 2017; Thangsupachai; Niwatthanakul; Chamnongsri, 2015; Wilson, 2017
Uso de ontologia formal como modelo de recuperação semântica da informação. uso tecnologias da web torna o portal semântico	Berners-Lee; Hendler; Lassila, 2001; Correa, 2012; Fahad <i>et al.</i> , 2011; Frosterus; Hyvonen; Laitio, 2011; Gavrilova, 2011; Guarino, 1997; Gruber, 1996; Huan; Maidment; Tian, 2011; Huang <i>et al.</i> , 2011; Jorge, 2005; Lausen <i>et al.</i> , 2005; Lachtim <i>et al.</i> , 2009; Maedche <i>et al.</i> , 2001; Nazário, 2013; Paiva; Costa; Silva, 2014; Reynolds; Shabajee; Cayzer, 2004; Spyns <i>et al.</i> , 2002; Telnov, 2015; Wimmer, Yoon; Rada, 2013; Zhang; Vasconcelos; Sleeman, 2005.

Fonte: Autoria própria (2021).

Isto posto, infere-se que os portais do conhecimento de universidades como artefatos tecnológicos podem ser considerados um sistema sociotécnico, já que são construídos por dispositivos complexos de interação social (BURNS, 2006; HUGHES, 1989). No desenho dos portais, além das regras e das ferramentas técnicas, os princípios organizadores sociais são marcantes e regulam as atividades dos atores que operam e gerenciam a tecnologia quer sejam, administradores, desenvolvedores e usuários finais (BURNS, 2006; CALLON, 1986c; LATOUR, 2001).

4.3 PORTAIS SEMÂNTICOS COMO ARTEFATOS SOCIOTÉCNICOS

Um artefato, segundo Walsham (1997), é um fato ou produto criado por seres humanos e que geralmente se refere a um desenho tecnológico. Assim, os portais semânticos são artefatos sociotécnicos já que são concebidos pelos atores humanos socialmente com bases em regras técnicas.

Os artefatos sociotécnicos compõem sistemas complexos com diferentes tipos de interações que podem interferir, prever, gerenciar ou controlar o resultado do funcionamento do sistema (MARTÍNEZ-GARCÍA, HERNÁNDEZ-LEMUS, 2013). Em uma estrutura complexa, ocorrem dinâmicas intrincadas que além da conectividade estrutural estão associadas a fortes interações com o meio ambiente já que sistemas complexos são sistemas abertos (MARTÍNEZ-GARCÍA, HERNÁNDEZ-LEMUS, 2013).

Neste contexto dos sistemas complexos os portais semânticos podem constituir redes semânticas cujo desenvolvimento requer uma análise aprofundada de todas as relações que envolvem pessoas e tecnologias. Os gestores definem os objetivos e o público-alvo e a equipe de TI estabelece as tecnologias de estruturação, bem como os recursos e as facilidades de acesso ao portal (NIELSEN, 2000; NIELSEN; LORANDER, 2007). Os conceitos complexos das estruturas semânticas e que envolvem as ontologias e experiência do usuário, são absorvidos e aplicados pelos designers e desenvolvedores, com auxílio de tecnologias semânticas que modulam os interesses do usuário, o contexto e o conteúdo do portal (ROSENFELD; MORVILLE; ARANGO, 2015)

Walsham (1997) destacou que no processo de implementação e desenho de uma tecnologia a TAR pode ser utilizada para investigar o processo de alinhamento de interesses de diferentes elementos humanos e não humanos em um arranjo ou artefato social e tecnológico.

Considerando a análise de fatores técnicos e não técnicos e como eles se articulam, Shaanika e Iyamu (2015) desenvolveram uma arquitetura de portal governamental na Namíbia. Verificaram que os fatores não técnicos foram completamente ignorados ou às vezes considerados menos significativos pelos técnicos de TI. Como fatores que afetaram a implantação do portal, esses autores destacaram: o processo de negócios governamentais, a comunicação, a falta de consciência, as funções e responsabilidades, o relacionamento entre atores,

políticas e pessoas. Concluíram que se os fatores técnicos e não técnicos não forem considerados na implementação de sistemas e tecnologias, a disponibilização de serviços no portal seria prejudicada.

Iyamu e Mgudlwa (2018) utilizaram a TAR como uma lente para orientar a análise de *big data* da área de saúde. O objetivo da estrutura proposta foi orientar análises que poderiam traduzir e transformar os dados em um recurso mais útil e significativo para os profissionais de saúde, melhorando assim os serviços prestados aos pacientes. Para os autores, supramencionados, a TAR propiciou uma análise relacional, moldada pela interação entre humanos e não-humanos e suas redes heterogêneas.

Gao (2005) utilizou a TAR na análise da construção sociotécnica de estratégias governamentais da China na área de telecomunicações. O mercado de telecomunicações foi considerado o ator não humano e os atores humanos incluíram o público e a sociedade, o Estado e as operadoras, representando os interesses sociais no setor de telecomunicações. Concluiu-se que a partir das aplicações da TAR foi possível conectar dinamicamente contextos sociais e tecnológicos, com a disputa dos atores para inscrever seus interesses na formulação de estratégias governamentais para o setor de telecomunicações.

Logo, os sistemas complexos são concebidos como redes em que os agentes ou atores interagem por meio de um relacionamento simétrico (MARTÍNEZ-GARCÍA, HERNÁNDEZ-LEMUS, 2013).

Callon e Latour (1981), estudando a macro-ordenação da sociedade e suas relações de poder, destacam que os atores formam alianças e aliciam outros atores para fortalecer estas alianças, criando as redes heterogêneas feitas de humanos e de artefatos não humanos em que as associações ocorrem mutualmente. Maia e Serafim (2011), corroboram com Callon e Latour (1981), ressaltando a formação de fluxos e de mediações em constante interação na rede de atores.

Nessa ótica, os portais semânticos, enquanto artefatos tecnológicos complexos também podem ser analisados pela abordagem de rede, já que são formados por uma estrutura que inclui diferentes agentes que se articulam entre si. Esta análise passa pelo estudo das instruções técnicas e do conjunto de normas que regem os sistemas sociotécnicos, mas também pela análise de todos os atores da rede, considerando seus traços fortes e fracos (BIJKER; HUGHS; PINCH, 1989; BURNS, 2006;

GRANOVETTER, 1983; HUGHES, 1989; MARTÍNEZ-GARCÍA; HERNÁNDEZ-LEMUS, 2013).

A interação entre elementos humanos e não-humanos fica, portanto, destacada na TAR, de modo que é possível guiar a análise de dados com foco em efeitos relacionais que modificam o desenho dos objetos técnicos (CALLON, 1986a, b, c); IYAMU; MGUDLWA, 2018). Infere-se, então, que a TAR pode ser usada, portanto, não apenas como elemento conceitual de entendimento dos atores, mas também como metodologia para o resgate do seu estabelecimento e desenho.

Desta abordagem a perspectiva sociotécnica de construção dos portais alinha-se fortemente com a TAR, teoria que destaca a interação do homem com os objetos e como estes atores se interconectam para que um artefato tecnológico agregue valor à sociedade. Assim, os elementos não humanos, que influenciam o comportamento humano, podem alterar as relações e as ações esperadas, modificando o desenho dos objetos técnicos, como os portais do conhecimento de universidades (BIJKER; HUGH; PINCH, 1989; BURNS, 2006, BURNS; DIETZ, 2001; HUGHES, 1989), conforme exaustivamente citado.

A recuperação semântica da informação surge então como uma alternativa para tornar mais eficiente o resultado de buscas na Web considerando a grande quantidade de informações disponíveis e o interesse dos usuários (AMATO *et al.*, 2017; LI *et al.*, 2017).

A partir dos estudos alinhados com esta pesquisa, infere-se, portanto, que é possível estabelecer indicadores para classificar o potencial semântico dos portais do conhecimento de universidades que é o objetivo central deste estudo. Um conjunto de indicadores neste escopo, cerne da proposta desta tese, constitui-se então como uma ferramenta de gestão e de desenho de portais, com vistas a atender os interesses de todos os atores, sobretudo dos usuários finais do portal, que é a recuperação rápida e precisa da informação.

Estabelecidos fundamentos e cenários, no próximo Capítulo, detalham-se os procedimentos metodológicos, com a caracterização e o fluxo de desenvolvimento da pesquisa, seus protocolos, a análise da literatura, a coleta e o tratamento dos dados, que permitiram elencar e propor indicadores qualitativos para a avaliação do potencial semântico de portais do conhecimento de universidades.

5 PROCEDIMENTOS METODOLÓGICOS DA PESQUISA

Nesse Capítulo serão apresentados os procedimentos metodológicos utilizados para o desenvolvimento desta pesquisa. São abordados os aspectos relativos à classificação da pesquisa, incluindo sua natureza e propósito, a descrição do seu planejamento e desenvolvimento com as técnicas e os procedimentos de coleta e de tratamento de dados, a operacionalização da pesquisa e da análise e a validação dos resultados auferidos.

5.1 CLASSIFICAÇÃO, DESENVOLVIMENTO E PROTOCOLO DE PESQUISA

Um conjunto de técnicas aliado a procedimentos intelectuais são necessários para atingir os objetivos do estudo e estes, são alcançados mediante preparação, pesquisa, execução e relatos da pesquisa. Para tanto, faz-se necessária a adoção de métodos que forneçam as bases lógicas à investigação (GIL, 2010; MARCONI; LAKATOS, 2010a).

Quanto ao delineamento metodológico, portanto, o estudo foi predominantemente bibliográfico, mas com pesquisa de campo do tipo diagnóstico a partir de múltiplas unidades experimentais, utilizando questionário como instrumento metodológico (GIL, 2010).

Quanto aos propósitos da pesquisadora, esta pesquisa caracteriza-se por ser exploratória e explicativa correlacional. O caráter exploratório tem por objetivo modificar e clarificar conceitos, aplicando-se procedimentos sistemáticos para obtenção de observações empíricas ou para análise de dados, permitindo o uso de descrições quantitativas e qualitativas sobre o objeto em estudo com procedimentos de amostragem flexíveis (MARCONI; LAKATOS, 2010b). Na vertente explicativa correlacional é possível clarificar o problema e, de forma estruturada, quantificar as relações entre variáveis (SAMPIERI; COLLADO; LUCIO, 2010).

Esta pesquisa do ponto vista dos procedimentos técnicos, para tratamento dos dados, é considerada mista já que envolve dados qualitativos e quantitativos, conforme acentuam Creswell e Clark (2013), no entanto a análise final será fundamentalmente qualitativa.

O delineamento do desenho metodológico da pesquisa é apresentado na Matriz de Consistência – Quadro 3 – a partir da qual recupera-se desde a questão de

pesquisa, já apresentada no Capítulo 1, passando pelos objetivos, população e amostra, tipos e métodos de pesquisa.

Quadro 3 - Matriz de consistência.

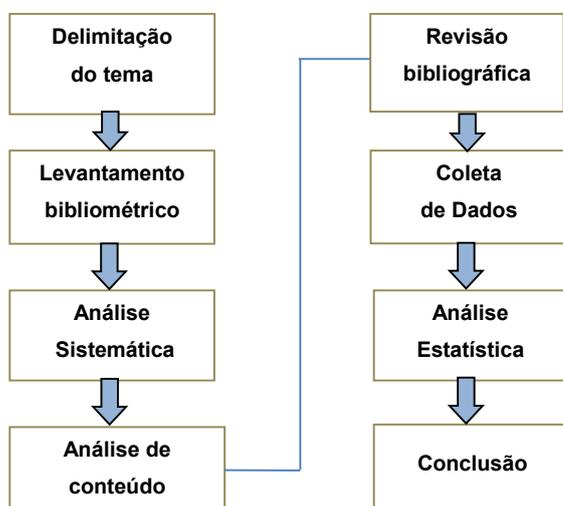
QUESTÃO DE PESQUISA	POPULAÇÃO E AMOSTRA
Quais indicadores qualitativos podem ser usados na avaliação do potencial semântico de portais do conhecimento de universidades?	Portais das 69 universidades federais brasileiras, arroladas com base no cenário de junho de 2020, representados pelos responsáveis da TI destas universidades
OBJETIVOS	AMOSTRA
<p>Geral Propor indicadores qualitativos contendo parâmetros para a avaliação do potencial semântico de portais do conhecimento de universidades, de acordo com os conceitos da TAR.</p> <p>Específicos</p> <ol style="list-style-type: none"> Levantar as condições básicas para o estabelecimento do potencial semântico de portais, como um artefato sociotécnico. Mapear os atores humanos e não humanos que compõem a rede estrutural de portais semânticos e suas funções. Elencar indicadores qualitativos iniciais, considerando os atores envolvidos, seus interesses, funções e características. Validar os indicadores elencados a partir de um levantamento (<i>survey</i>) em universidades e das percepções de um painel de especialistas Apresentar indicadores qualitativos para avaliação de portais. 	<p>Dados de questionário enviados aos 69 coordenadores de TI responsáveis pelos portais das universidades federais brasileiras para redistribuição interna e coleta de respostas.</p>
	TIPOS DE PESQUISA
	Aplicada, exploratória e explicativa correlacional, quantitativa e qualitativa
	MÉTODO DE PESQUISA
	<p>Os métodos de pesquisas estão delineados no fluxograma apresentado na Figura 1 - Capítulo 1 e incluem:</p> <ol style="list-style-type: none"> Levantamento bibliométrico e bibliográfico Aplicação de questionário Mapeamento da rede de portais Análise de conteúdo Análise estatística descritiva

Fonte: Autoria própria (2021).

O desenvolvimento da pesquisa iniciou-se com a definição do tema, o levantamento bibliométrico, a análise sistemática e a revisão bibliográfica. A seguir foi desenhada a rede de portais do conhecimento de universidades e foi feita a coleta de

dados, a análise estatística descritiva e a descrição das conclusões do estudo. A Figura 9 apresenta as fases da pesquisa.

Figura 9 - Fases de desenvolvimento da pesquisa.



Fonte: Autoria própria (2020).

A seguir serão descritas as etapas da pesquisa, detalhando-se os instrumentos de coleta, tratamento e operacionalização de dados. Foi realizado um levantamento bibliométrico para subsidiar a revisão da literatura e a composição do questionário de pesquisa. Sequencialmente foi efetuado o tratamento dos dados mediante análise de conteúdo, seguida pelo tratamento estatístico dos dados.

5.1.1 Levantamento bibliométrico

O objetivo do levantamento bibliométrico é fornecer indicadores para composição da revisão da literatura e para a análise de conteúdo, provisionando esta com recursos como a frequência de palavras em artigos de determinado periódico e o número de citações e cocitações, conforme preceitos biblioteconômicos (SOARES; PICOLI; CASAGRANDE, 2018).

Os indicadores bibliométricos, utilizados nas seções de metodologia, e os resultados devem conter uma descrição detalhada das bases de dados, dos bancos de teses e dissertações e das demais fontes de informação utilizadas, assim como a sua abrangência e cobertura; as justificativas para a escolha dessas bases, também

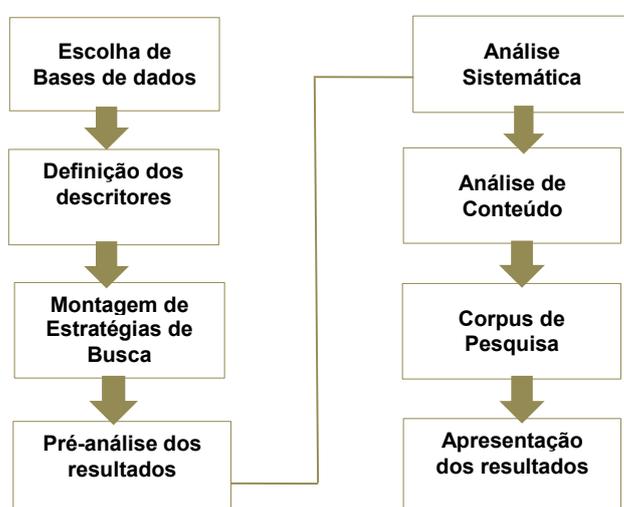
devem ser incluídas devendo relevar a escolha do recorte temporal (SOARES; PICOLI; CASAGRANDE, 2018).

Seguindo esse roteiro, neste estudo, o levantamento bibliométrico teve início com a seleção das bases de dados para pesquisa, optando-se pelas bases que indexam matérias das áreas de Ciências Sociais Aplicadas e Informática, com foco nos Sistemas de Informação, além das bases multidisciplinares de maior credibilidade em nível internacional.

A seguir foram definidos as palavras chaves e/ou descritores de assunto relacionados ao tema e posteriormente, formou-se a estratégia de busca para consulta nas bases elencadas.

Os resultados encontrados foram exportados para o aplicativo de gerenciamento bibliográfico EndNote, para então, prosseguir com as demais fases de execução da pesquisa quais sejam: a) análise e seleção dos materiais; b) análise sistemática; c) definição do público-alvo; d) coleta de dados primários: questionário (elaboração/aplicação) (Figura 10).

Figura 10 - Fases do levantamento bibliométrico



Fonte: Autoria própria (2021).

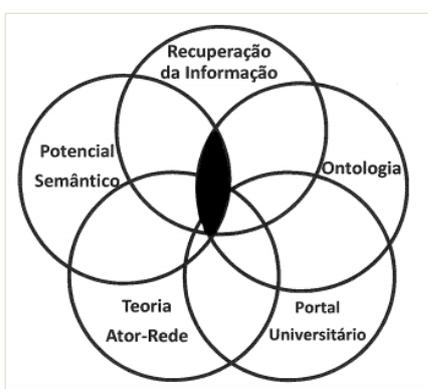
Tradicionalmente, salienta-se, os estudos bibliométricos são desenvolvidos a partir das informações obtidas em grandes bases de dados como Web of Science e Scopus, embora alguns estudos nacionais utilizem o Google Scholar com indicador bibliométrico (SOARES; PICOLI; CASAGRANDE, 2018).

Neste estudo as bases de dados de escolha para a pesquisa foram : Web of Science – Coleção Principal, Proquest (Information Science & Technology Abstracts, Computer and Information Systems Abstracts, Applied Social Sciences Index & Abstracts, Social Services Abstracts; Technology Collection, Materials Science & Engineering Database information), Scopus e SciELO Citation Index (Web of Science). Essas bases foram elencadas pela abrangência e disponibilização dos principais periódicos relacionados ao tema em estudo. O espaço temporal de 5 anos foi eleito por recuperar informações mais atualizadas que tratam de Sistemas de Informação e Tecnologia, áreas que possuem um grande volume de publicação e seus desdobramentos temáticos – subtemas.

Os critérios de elegibilidade dos itens de informação para o estudo incluíram: artigos de periódicos, de evento científico e livros eletrônicos em língua inglesa, em espanhol e em língua nacional, avaliados por pares com conteúdo associado aos temas e subtemas em estudo, no intervalo temporal indicado.

O resultado do levantamento bibliométrico pressupõe a intercessão de todos os termos de busca, bem como dos critérios de inclusão/exclusão de documentos citados, indicando que esses devem aparecer em todos os artigos recuperados. Para este estudo foram delineados 5 principais termos que envolvem os subtemas em estudo e representam suas palavras-chaves: 1- Ontologia; 2- Recuperação da informação, 3- Portal universitário; 4- Potencial semântico e, 5- Teoria ator-rede. A interface dos termos está ilustrada no diagrama da Figura 11.

Figura 11 – Diagrama representativo da combinação de palavras-chaves usadas na pesquisa nas bases de dados e a interface de resultado (central em preto).



Fonte: Autoria própria (2021).

A busca torna-se mais precisa quando se utiliza termos associados em estratégias de busca. Desta forma, os descritores foram conectados a operadores booleanos, formando uma estratégia única de busca para pesquisa em todas as bases de dados elencadas. A busca com todos os descritores juntos constitui uma técnica de revisão sistemática da literatura que pode ser reproduzível, ou seja, a mesma estratégia de busca pode ser utilizada para checar os resultados encontrados. Assim, para este estudo foram reunidos os descritores relacionados ao tema e subtemas e seus sinônimos, formando a estratégia delineada no Quadro 4.

Quadro 4 - Estratégia única de pesquisa

```
(((("retrieval information" AND ("university portals" OR "academic portals" OR "educational portals") AND ("semantic potential" OR "semantic evaluation") AND Ontology AND "actor-network theory" AND "semantic index" AND evaluation AND portals))))
```

Fonte: Autoria própria (2021).

Verificou-se, entretanto, que não houve resultado válido para a estratégia única em nenhuma das bases de dados estabelecidas denotando a originalidade do estudo, com o enfoque que se pretende dar nesta pesquisa. Um dos propósitos deste estudo é validar as condições básicas para estabelecimento de potencial semântico de portais já elencadas por Guedes (2015) a fim de propor indicadores qualitativos contendo parâmetros para a avaliação do potencial semântico de portais do conhecimento de universidades, de acordo com os conceitos da TAR.

Adicionalmente, para verificar a existência de indicadores para avaliação de potencial de portais, foi utilizada a expressão de busca "*semantic indicators*" AND *evaluation AND portals*, mas essa busca não retornou resultados nas bases Scielo, Web of Science e Scopus. Já na base Proquest foram encontrados 12 resultados, entretanto, verificou-se que apenas alguns termos da busca apareciam nos resumos das obras. Isto foi comprovado com a substituição do operador booleano AND pelo NEAR/5 que aproxima os termos da expressão de busca e evidencia a relação entre os termos em um mesmo documento. O operador NEAR é usado para consultas de conteúdo em que a proximidade de termos identifica a relação entre eles, isto é, um documento corresponderá à consulta se as palavras estiverem a uma distância de até 50 palavras uma da outra no documento (WEB OF SCIENCE, 2020). Quanto mais próximas as palavras estiverem, mais elevada será a classificação atribuída ao documento no conjunto de resultado (WEB OF SCIENCE, 2020). Assim, utilizando-se

NEAR/5, para aproximar os termos de busca a uma distância de cinco palavras, não foi recuperado, igualmente, nenhum documento. Desse modo, comprovou-se a inexistência de indicadores formatados para avaliação semântica de portais, nas bases utilizadas.

Diante do resultado negativo da estratégia de busca unificada, os descritores foram arranjados formando 12 conjuntos de estratégias de busca que foram novamente utilizadas para pesquisa nas bases de dados. Desses conjuntos, sete estratégias de busca, foram consideradas válidas, isto é, aquelas que recuperaram pelo menos um item de informação independente da base (Tabela 3).

Tabela 3 - Resultados do Levantamento bibliométrico.

ESTRATÉGIAS DE BUSCA	RESULTADOS
<i>("retrieval information") AND "actor-network theory"</i>	2
<i>("retrieval information") AND ontology</i>	170
<i>("retrieval information" AND ("university portals" OR "academic portals" OR "educational portals"))</i>	1
<i>("retrieval information" AND ("semantic potential" OR "semantic evaluation"))</i>	13
<i>("semantic potential" OR "semantic evaluation") AND Ontology</i>	118
<i>("university portals" OR "academic portals" OR "educational portals") AND ontology</i>	17
<i>ontology AND "actor-network theory"</i>	558
TOTAL	879

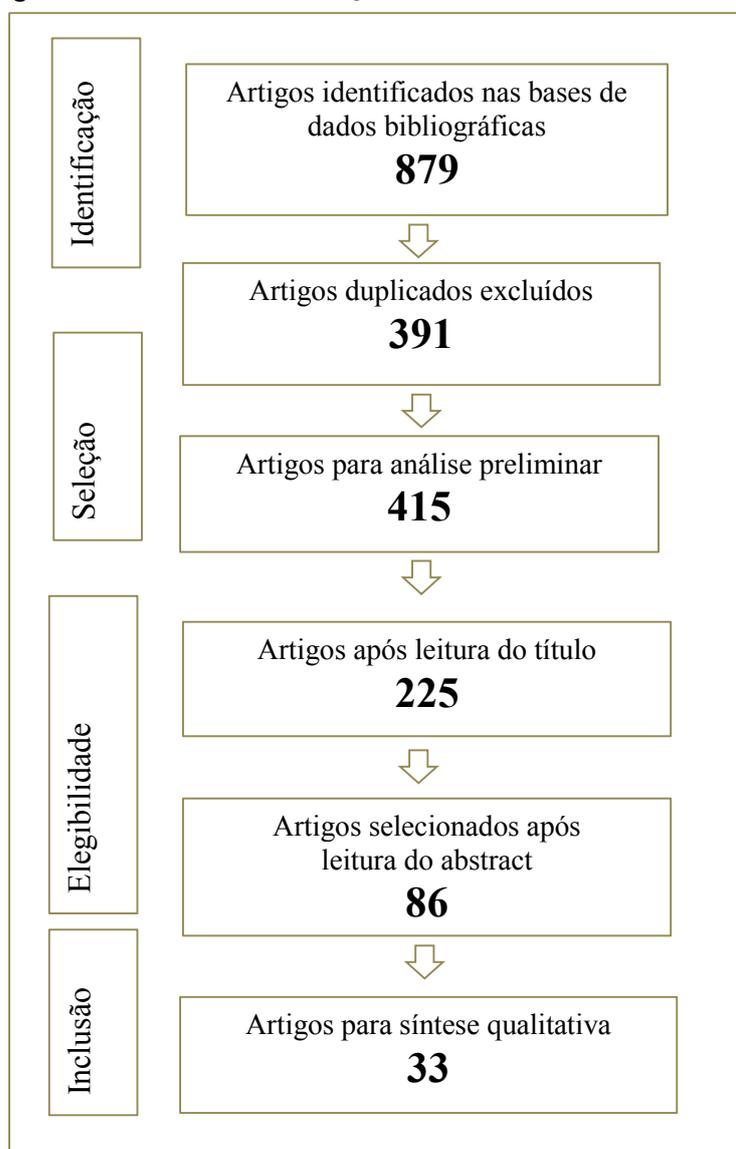
Fonte: Aatoria própria (2021).

A seleção dos itens para revisão sistemática da literatura foi fundamentada na Recomendação PRISMA que tem por objetivo auxiliar os pesquisadores no relato de revisões sistemáticas que variam de acordo com a habilidade e o entendimento do leitor de analisar e avaliar os pontos fortes e fracos (MOHER *et al.*, 2009).

Após extrair os itens duplicados do total de 879 resultados recuperados, restaram para análise preliminar 415 itens de informação. Na seleção pelo título foram selecionados 225 potenciais itens de interesse. Após a leitura dos resumos/*abstract* foram selecionados 86 itens para leitura do texto na íntegra. Ao final desta etapa o *corpus* de pesquisa ficou com 33 documentos – na forma de artigos –, uma vez que foram excluídos aqueles artigos não alinhados ou cujos textos completos não estavam

disponíveis para acesso (Figura 12). Estes 33 documentos fazem parte do *corpus* dinâmico de pesquisa e estão associados aos temas em estudo (MARICATO, 2011).

Figura 12 - Fluxo da informação com as fases da revisão sistemática.

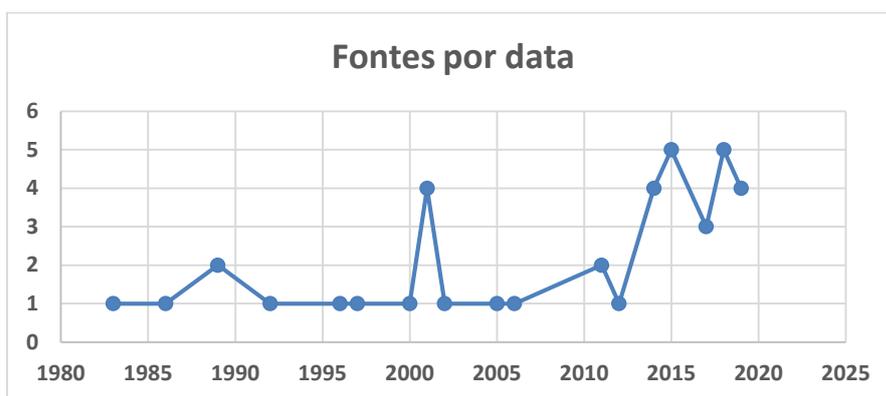


Fonte: Adaptado de Scielo (2021).

Adicionalmente, uma pesquisa com a mesma estratégia unificada efetuada na Biblioteca Digital de Teses e Dissertações (BDTD) e no Catálogo de Teses e Dissertações (CAPES) não retornou nenhum estudo. Pesquisando ainda pelo conjunto de estratégias listadas na Tabela 3, filtrando as áreas de conhecimento associadas a Sociais Aplicadas, Computação e Administração, recuperou-se 11 estudos na BDTD e 77 na CAPES, entretanto, da análise dos títulos e resumos resultaram dois trabalhos, um deles a dissertação de mestrado de Guedes (2015).

Os 33 artigos foram analisados detalhadamente, sendo extraídas informações para compor a Revisão de Literatura e para discussão de resultados. Na distribuição das fontes por data, verifica-se que há um pico nos anos 2000 que coincide com o início dos estudos da Web Semântica, seguido por uma concentração maior de fontes entre 2015 e 2019, cobrindo os últimos 5 anos, período em que ocorre o avanço da área de recuperação semântica da informação (Gráfico 1).

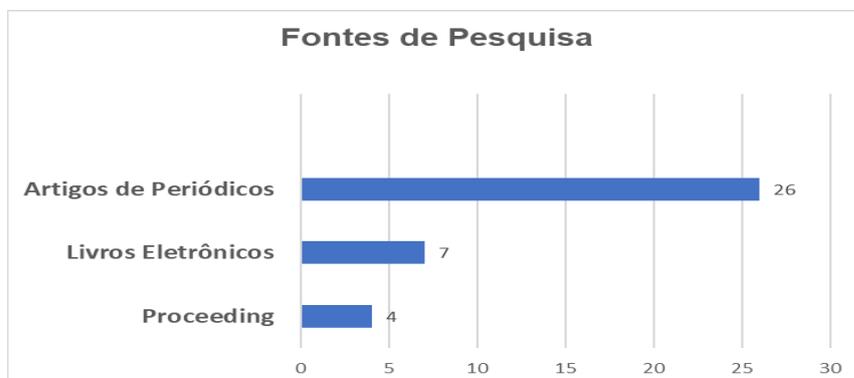
Gráfico 1 - Distribuição das 33 fontes de estudo por data de publicação.



Fonte: Autoria própria (2021).

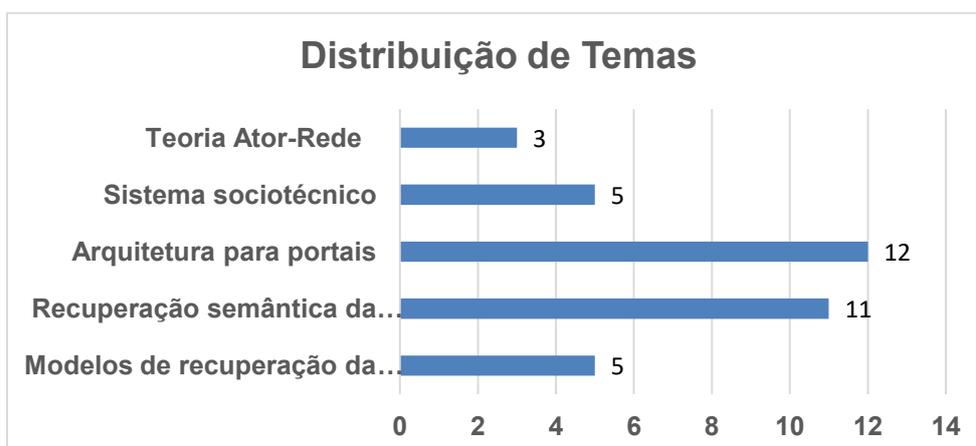
Dentre os tipos de fontes de informação utilizadas, em sua maioria foram artigos de periódicos – 26 (Gráfico 2), seguido por livros eletrônicos e *proceedings* (materiais publicados em eventos técnicos). Quanto à distribuição de temas, verifica-se maior quantidade de fontes para a Recuperação da Informação (11 fontes) e para a Arquitetura de Portais (12 fontes) (Gráfico 3).

Gráfico 2 - Fontes de informação utilizadas no estudo.



Fonte: Autoria própria (2021).

Gráfico 3 – Distribuição de fontes de informação por tema abordado



Fonte: Autoria própria (2021).

A pesquisa incluiu também um *corpus* estático previamente conhecido pela pesquisadora e que complementa o conjunto de dados e materiais teóricos que fundamenta a pesquisa (MARICATO, 2011).

A fim de ratificar a importância dos estudos desse *corpus*, foi efetuado uma análise das citações de alguns artigos nas bases de dados. Segundo Borgman e Furner (2002), a análise de citações é a abordagem bibliométrica mais conhecida e que estabelece relações semânticas entre os artigos, o que justifica sua importância em trabalhos científicos.

Já para Alvarenga, (1998), as citações propiciam o mapeamento da rede de relações estabelecidas entre autores e textos, evidenciando ângulos específicos do processo de produção de conhecimento. À luz desses preceitos pesquisou-se na base Web of Science e na Scopus o número de citações dos documentos seccionados.

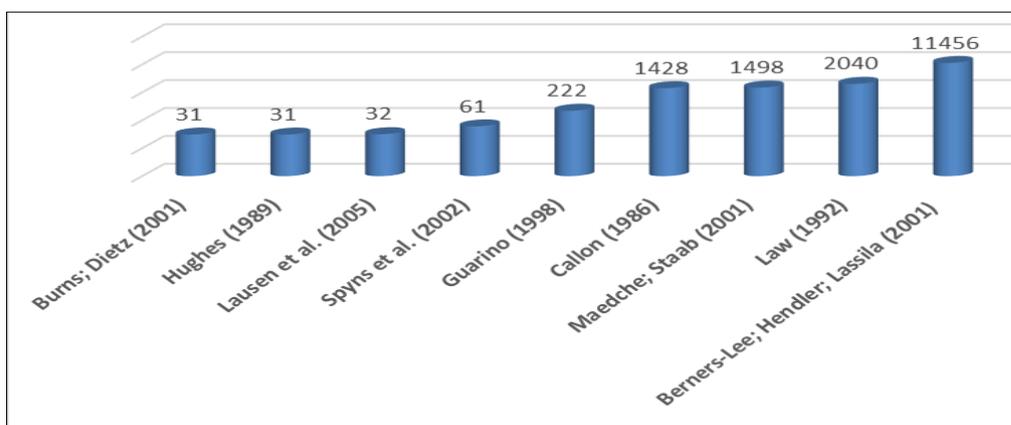
O Quadro 5 apresenta os trabalhos que tiveram mais de 10 citações em pelo menos uma das bases, já que este foi o número mínimo de citações encontrado para um autor nas bases selecionadas para esta análise.

Quadro 5 - Número de citações por autor e base de referência.

AUTOR	WEB OF SCIENCE	SCOPUS
Berners-Lee, Hendler e Lassila (2001)	7.757	3.699
Burns e Dietz (2001)	17	14
Callon (1986)	416	1.012
Law (1992)	1.155	885
Guarino (1998)	-	222
Hughes (1989)	31	-
Lausen <i>et al.</i> (2005)	32	-
Maedche e Staab (2001)	939	559
Spyns <i>et al.</i> (2002)	33	30

Fonte: Autoria própria (2021).

Verifica-se que, dentre os trabalhos mais citados, pode ser feito uma correlação entre os subtemas do presente estudo (Gráfico 4). Para a área de Web Semântica o estudo de Berners-Lee, Hendler e Lassila (2001) teve um número expressivo de citações, inferindo-se que estes são autores de importância no assunto. Na área de estudos sobre TAR e sistemas sociotécnicos, os estudos de Law (1992) e Callon (1986a, b, c) tiveram grande destaque, o que aponta estes trabalhos como sendo obras fundamentais no tema. O estudo de Maedche e Staab (2001) fundamentam a construção de portais conforme preceitos da Web Semântica. Já o estudo de Guarino (1998) denota expressiva quantidade de citações, sendo este autor considerado um expoente da área de Ontologia Formal. Evidentemente, como já citado, não se descuidou do *corpus* dinâmico de pesquisa fundamentado na Revisão de Literatura no espaço temporal de 5 anos.

Gráfico 4 - Número de citações por autor.

Fonte: Autoria própria (2021).

Considerando o número expressivo de citações, depreende-se que estes autores constituem as teorias de base do presente estudo. Os trabalhos citados nos artigos foram também verificados, buscando-se a contextualização das citações com os temas correlatos deste estudo, de acordo com a metodologia *Snowball* ou Bola de Neve, que possibilita a formação de uma cadeia de referências ou informantes sobre determinado tema (BALDIN; MUNHOZ, 2011; BIERNACKI; WALDORF, 1981; MOREIRA; CALEFFE, 2006). Evidentemente, como já citado, não se descuidou do *corpus* dinâmico de pesquisa fundamentado na Revisão de Literatura no espaço temporal de 5 anos.

O levantamento bibliométrico, com seus 879 documentos resultantes foi a base para análise sistemática e análise de conteúdo efetuada a partir dos preceitos de Meireles e Cedón (2010) e Bardin (2011).

5.2 ANÁLISE SISTEMÁTICA E ANÁLISE DE CONTEÚDO

Os métodos de análises sistemática e de conteúdo, que serão descritos nesta seção, auxiliam a organização dos itens de informação recuperados pelo levantamento bibliométrico, possibilitando a articulação destes, de forma a cumprir os objetivos deste estudo (BARDIN, 2011; GUERTIN; BERNHARDT, 2005; MEIRELES; CEDÓN, 2010).

5.2.1 Análise sistemática

A partir dos resultados do levantamento bibliométrico foi efetuada a sistematização de conteúdos existentes nos documentos por meio da leitura dos textos na íntegra a fim de obter informações relevantes para a revisão bibliográfica, conforme destacaram Meireles e Cedón (2010). A análise sistemática foi realizada para a construção da Revisão de Literatura e também para a análise de resultados.

Com a análise sistemática dos 33 trabalhos selecionados foi possível verificar a aderência destes ao tema em estudo, identificando-se abordagens semelhantes e complementares. Para cada item de informação foram identificados a teoria de base utilizada, os objetivos, a metodologia e os principais resultados. O quadro 6 apresenta

um recorte da análise sistemática incluindo os trabalhos mais citados e os demais estudos seguem listados no Apêndice 1.

Quadro 6 - Recorte da análise sistemática.

TÍTULO / AUTOR / ANO	TEORIA DE BASE USADA	OBJETIVO	MÉTODOS E PROCEDIMENTOS DE PESQUISA	PRINCIPAIS RESULTADOS
Notes on the theory of the actor network. (LAW, 1992)	Teoria Ator-Rede	Destaca as relações sociais, incluindo poder e organização, como efeitos de rede heterogênea	Pesquisa qualitativa e documental	Descreve a teoria ator-rede como teoria relacional e sociologia orientada a processos que tratam agentes, organizações e dispositivos como efeitos interativos. A sociedade é formada por redes padronizadas de materiais heterogêneos, compostas não só de pessoas, mas também de máquinas, animais, textos, dinheiro, arquiteturas enfim e de outro qualquer material.
Understanding, building and using ontologism. (GUARINO, 1997)	Ontologia Representação do conhecimento	Definição, construção e reuso de ontologias	Qualitativa e quantitativa. Revisão de Literatura	Conceitua a ontologia como sendo um artefato de engenharia para recuperação de informação; Apresenta classificação da ontologia de acordo com o tipo de estrutura e assunto da conceitualização. Sugere a construção de ontologias a partir de teorias, metodologias e ferramentas que induzam e organizem o conhecimento de forma reutilizável e transparente
Ontology learning for the Semantic Web. (MAEDCHE; STAAB, 2001)	Ontologia Web Semântica	Construção de ontologia para Web Semântica e portais semânticos	Qualitativa e quantitativa Revisão de Literatura e apresentação de modelos	Apresenta o editor de ontologias OntoEdit para construção de ontologia em quatro estágios: importação e reuso extração, poda e refinamento. Apresenta editor específico cujas características precisam ser comparadas com outros mais recentes
The Semantic Web... BERNERS-LEE; HANDLER; LASSILLA (2001)	Ontologia Web Semântica	Estruturação da Web para inserção de significado nos documentos visando à leitura destes por agentes computacionais inteligentes	Qualitativos Apresentação de modelo de estruturação da Web	A Web Semântica possibilita aumentar a eficiência das buscas e facilitar a recuperação de informações relevantes para o usuário; Usa ontologia formal como sistema de recuperação de informação A estrutura de Web Semântica necessita tecnologias específicas. As últimas camadas da Web Semântica e devem estar bem desenvolvidas e consolidadas e ainda estão em desenvolvimento.

Fonte: Autoria própria (2021).

Após a análise sistemática procedeu-se a análise de conteúdo, detalhada na seção seguinte, a fim de descrever exhaustivamente os itens de informação relacionados aos temas em estudo.

5.2.2 Análise de conteúdo

A análise de conteúdo, de acordo com Bardin (2011), tem por um dos objetivos a descrição de conhecimentos relacionados aos textos selecionados para pesquisa de forma sistemática. Essa análise inclui três etapas: 1) consiste na pré-análise que envolve os procedimentos que antecedem a análise propriamente dita, incluindo o processo de escolha de documentos, de formulação de hipóteses, dos objetivos da análise e da elaboração dos indicadores que fundamentam a interpretação final; 2) envolve a exploração do material, alocando os dados brutos sistematicamente em unidades que possibilitem a descrição das características dos documentos; 3) inclui os processos de inferência e a interpretação obtidos a partir do tratamento dos resultados.

5.2.2.1 Pré-Análise

Nessa fase é feita a organização do material a ser analisado, estabelecendo os objetivos, hipóteses e indicadores para a interpretação final do conteúdo (BARDIN, 2011). A pré-análise envolveu a leitura e a seleção dos itens de informação a serem analisados e a definição dos objetivos da análise de conteúdo (BARDIN, 2011). A seguir foram definidas as categorias de análise para responder aos objetivos de pesquisa e contribuir para o estabelecimento de instrumento de gestão de portais do conhecimento de universidades pautados na recuperação semântica da informação. As etapas da análise de conteúdo, para construção da Revisão de Literatura, incluíram:

- a) Identificação prévia dos atores envolvidos na construção de portais do conhecimento semânticos de universidades que foi complementada com a Análise dos Questionários.
- b) Verificação das condições básicas para o estabelecimento do potencial semântico de portais;

Salienta-se que foi feita uma segunda análise de conteúdo para a identificação de parâmetros para construção de indicadores qualitativos para portais semânticos. O processo de análise de conteúdo, com a exploração do material, é descrito a seguir.

5.2.2.2 Exploração do Material

Nesta fase foram realizadas atividades de codificação dos conteúdos, que compreendem a transformação dos dados, agrupando-se as características gerais e delimitando-se categorias de contexto, de análise e de registro (BARDIN, 2011).

As unidades de registro podem ser palavras temas, expressões, objetos, personagens, acontecimentos, documentos que abrangem significado codificado que pode ser categorizado e identificado sequencialmente (BARDIN, 2011). As categorias de contexto explicam ou se relacionam com as unidades de registro. Já as categorias de análise agrupam em um mesmo código os conteúdos analisados

A partir da revisão da literatura foram definidas as categorias de contexto em três grandes grupos, a saber: (i) Recuperação da Informação, (ii) Portal Semântico e (iii) Sistemas Sociotécnicos. Estas categorias foram determinadas a partir da análise heurística com a geração de imagens de nuvens de palavras, mediante uso da ferramenta Tag Crowd.

A análise heurística, resultante em nuvens de palavras formadas com base nos artigos analisados para cada categoria de contexto, é apresentada nas figuras 13 a 15. Ressalta-se que as palavras com maior destaque gráfico remetem as palavras mais citadas e representativas dos temas em estudo, por outro lado, as de menos destaque representam oportunidades de pesquisa.

A categoria de contexto Recuperação da Informação tem a análise ilustrada na nuvem de palavras da Figura 13, sendo as mais citadas: *ontology*, *search*, *semantic*, *retrieval*, *data*, *query*, *results*.

Figura 13 - Nuvem de palavras gerada para a categoria de contexto “Recuperação da Informação”.



Fonte: Autoria própria, utilizando a ferramenta Tag Crowd (2021).

Para a categoria de contexto Portal Semântico, foram destacadas na nuvem de palavras da Figura 14, tendo como palavras em destaque: *portals*, *ontology*, *semantic*, *technologies*, *search*, *information*.

Figura 14 - Nuvem de palavras gerada para a categoria de contexto “Portal Semântico”.



Fonte: Autoria própria, utilizando a ferramenta Tag Crowd (2021).

Para a categoria de contexto Sistemas Sociotécnicos, ilustrada na Figura 15, os destaques foram para as palavras: *science, social, systems, technology, ties, weak*.

Figura 15 - Nuvem de palavras geradas para a categoria de contexto “Sistemas Sociotécnicos”



Fonte: Autoria própria, utilizando a ferramenta Tag Crowd (2021).

As categorias de análise são uma redução, ou subdivisão, das categorias de contexto, buscando-se agrupar sob um mesmo código os conteúdos analisados (BARDIN, 2011). Foram estabelecidas três categorias de análise a saber: (i) Recuperação semântica da informação – relacionados à primeira categoria de contexto; (ii) Arquitetura dos portais semânticos – relacionada à segunda categoria de contexto e (iii) Artefatos sociotécnicos - relacionados à terceira categoria de contexto

Para cada categoria de contexto foram estabelecidas unidades de contexto e a partir desta foram elencadas as unidades de registro. As unidades de registro foram definidas de modo a correlacionar elementos e características de busca semântica em portais à luz dos preceitos sociotécnicos.

Assim, para cada uma das categorias de análise baseadas nas unidades de contextos, foram identificadas unidades de registro para facilitar a análise de conteúdo. Por exemplo, para a categoria de análise “Arquitetura de portais semânticos” foram identificadas as unidades de registro: Web semântica; Tecnologias semânticas; Motor de busca; Tecnologias de padrão aberto; Tecnologias de base e

Portais semânticos. O Quadro 7 apresenta uma síntese das categorias de análise deste estudo.

Quadro 7 - Categorias de Contexto, de análise e Unidades de Contexto e de Registro para a Análise de Conteúdo.

CATEGORIA DE CONTEXTO	CATEGORIA DE ANÁLISE	UNIDADE DE CONTEXTO	UNIDADE DE REGISTRO
Recuperação da informação	Recuperação semântica da informação	No campo da recuperação da informação a ontologia assemelha-se a uma taxonomia ampliada que é a linguagem documentária usada para representação do conhecimento (SALES; CAFÉ, 2008)	Ontologia
Portal semântico	Arquitetura dos Portais semânticos	A Web Semântica , ou WEB. 3.0, é conhecida como uma Web com significado, cujas informações são legíveis por máquina, o que possibilita a oferta de serviços automatizados (SIRISHA; SUBBARAO; KAVITHA, 2015).	Web Semântica
		As tecnologias semânticas são usadas para estruturar semanticamente os dados bem como para oferecer serviços de Web Semântica. Fornecem integração inteligente de dados e acesso mais ágil a diferentes tipos de dados por meio de linguagens específicas (STANESCU, 2018).	Tecnologias semânticas
		Na pesquisa semântica o motor de busca não só entende as relações hierárquicas dos termos, mas também as relações complexas que ocorrem entre os termos definidos dentro de ontologias (PAIVA; COSTA; SILVA, 2014).	Motor de busca
		Mecanismos de Inferência como o SPARQL Engine e XQuery de padrão aberto são utilizados para otimização das respostas para buscas semânticas. São exemplos de sistemas básicos de padrão aberto utilizados em estruturas semânticas: Linux, Oracle, MySQL, Apache, Nginx, Drupal, Plone e WordPress (WORLD WIDE WEB - W3C, 2019; MASNER <i>et al.</i> , 2019).	Tecnologias de padrão aberto
		As tecnologias de base são aquelas convencionalmente usadas na estruturação de portal da Web quer sejam sistema operacional, banco de dados, servidor Web, todas de padrão aberto (BERNERS-LEE; HENDLER; LASSILA, 2001)	Tecnologias de base
		Os portais semânticos de busca visam o gerenciamento de ontologias desenvolvidas a fim de melhorar a busca e a recuperação destas na Web Semântica (FAHAD <i>et al.</i> , 2011).	Portais Semânticos
Sistemas sociotécnicos	Artefatos sociotécnicos	Os sistemas técnicos, constituem sistemas complexos com características de modularidade e robustez. de acordo com (MARTÍNEZ-GARCÍA; HERNÁNDEZ-LEMUS, 2013).	Sistemas complexos
		A Teoria Ator-Rede e pressupõe que atores humanos e não-humanos possuem a mesma importância em um sistema, em razão da sua capacidade de agência (CALLON; LATOUR, 1981; LAW, 1992).	Teoria Ator-Rede

Fonte: Autoria própria (2021).

A terceira e última etapa da análise de conteúdo se refere ao tratamento dos dados, à interpretação dos resultados (assertivas e conclusões) e à inferência. Esta etapa permite, em primeira instância, a construção do instrumento de coleta de dados, detalhado na seção seguinte e que, posteriormente, foi retomada na análise dos dados coletados.

5.3 COLETA, TRATAMENTO, INTERPRETAÇÃO DOS DADOS E INFERÊNCIA

Esta Seção foi composta pela definição da população e amostra, da forma de coleta e de tratamento dos dados, delineando-se também o protocolo adotado para a coleta de dados e suas inferências.

5.3.1. População e Amostra

A população investigada, o universo de pesquisa, foi composta por 69 universidades federais brasileiras, quantificadas de acordo com o cenário das instituições federais de ensino superior em junho de 2020 (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, 2021). O público-alvo para coleta de dados foi composto por técnicos da área de Tecnologia da Informação (TI), ou por elementos da área da Comunicação ou da Administração Institucional, relevando que estes devem ser responsáveis pela criação do portal da instituição. Definiu-se que, independentemente do número de elementos envolvidos, a resposta seria unitária para cada universidade.

Considerando, portanto, uma população finita, o cálculo amostral foi efetuado de acordo com a técnica de amostragem probabilística simples, aleatória e sem repetição, com grau de confiança de 90% e margem de erro de 10 % (TRIOLA, 2017), de acordo com a fórmula:

$$\text{Tamanho da amostra} = \frac{\frac{z^2 \times p(1-p)}{e^2}}{1 + \left(\frac{z^2 \times p(1-p)}{e^2 N} \right)}$$

N = tamanho da amostra
E = margem de erro
Z = escores z

Salienta-se, no entanto, que apesar de se usar o cálculo amostral dentro da orientação da amostragem probabilística simples aleatória não foi feita uma codificação para seguimento e verificação dos retornos, pois estes redundaram e foram influenciados pelas respostas das universidades.

A partir do cálculo feito, a amostra mínima deveria ser de 41 respondentes para o questionário de coleta de dados, mas como se obteve, na data de 29/09/2021, 42 respostas, o erro recalculado final foi de 9,53% com margem de confiança de 95%, atendendo e superando os limites estatísticos previamente estabelecidos. A amostra foi, então, considerada representativa para pesquisas mistas, lembrando, no entanto, que esta é uma pesquisa com análise preponderantemente qualitativa; o percentual de retorno de 60,86% da população estudada com o envio de questionários respondidos obteve taxa acima da média de retorno para questionários sugerida por Marconi e Lakatos (2010) que seria de 25%.

As questões formuladas para o questionário de coleta de dados e sua fundamentação na literatura seguem delineadas na Seção seguinte.

5.3.2 Instrumento de coleta de dados

O questionário foi escolhido como ferramenta para a coleta de dados e foi elaborado a partir da revisão da literatura que por sua vez, foi delineada pelas categorias de contexto, já citadas: Recuperação da Informação e Portal Semântico com delineamento dos requisitos para a criação de portais semânticos e Sistemas Sociotécnicos, com a caracterização sociotécnica do portal semântico, conforme detalhado no Quadro 8. A escolha pelo questionário deu-se devido à localização geográfica das universidades e a dificuldade de acesso aos potenciais respondentes, já que esse instrumento é de natureza impessoal e possibilita uma estruturação ordenada de perguntas para serem respondidas sem a presença do entrevistador (MARCONI; LAKATOS, 2010b).

Um pré-teste do questionário foi elaborado para verificar a pertinência e a adequação de perguntas do instrumento. O pré-teste foi enviado para 4 instituições de ensino de mesmas características da população em estudo (2 estaduais e 2 particulares), obtendo-se 3 respostas. A partir do questionário piloto verificou-se a necessidade de excluir a opção de resposta “Não se aplica” pela dualidade dessa

resposta para a opção " não", já que o objetivo das questões foi saber se era utilizada ou não tecnologias e/ou condições indicadas nas perguntas.

Visando a precisão e a rapidez no retorno de dados, o questionário foi eletronicamente construído com o auxílio de ferramenta do Google Docs e foi encaminhado via *e-mail* para cada um dos responsáveis pelos portais das 69 universidades federais em funcionamento no momento da pesquisa – universo este na data de início do envio⁵. A identificação dos responsáveis pelos portais foi feita mediante lista de contatos disponível pela Agência de Tecnologia da Universidade Federal do Paraná (2021), instituição a qual a autora pertence

As questões utilizadas no questionário criado foram baseadas no quadro referencial de condições mínimas para estabelecimento de portais semânticos, proposto por Guedes (2015), em pesquisa de dissertação de mestrado. A pesquisa foi desenvolvida em 11 universidades federais da Região Sul do Brasil, partindo-se de um estudo comparativo entre as características da Web Semântica e as ferramentas e tecnologias usadas nas instituições pesquisadas. As questões respondidas pelos técnicos de TI, na época, incluíram o uso de tecnologias utilizadas, a forma de gestão da informação e de gestão de ontologia adotada nos portais.

No presente estudo, foi considerada como população o total das 69 universidades federais brasileiras, de acordo com o cenário das instituições instaladas até junho de 2020, conforme já citado, com amostra ampliada e significativa a fim de propor indicadores qualitativos contendo parâmetros para a avaliação do potencial semântico de portais do conhecimento de universidades, de acordo com os conceitos da TAR, mediante a identificação de ferramentas e tecnologias usadas na estruturação do portal e na descrição da rede de atores envolvida na construção deste.

O Quadro 8 apresenta as questões elaboradas para coleta de dados, a fim de elencar os requisitos técnicos semânticos que caracterizam a recuperação da informação em portais semânticos e as condições que evidenciam a relação entre portais e artefatos sociotécnicos, incluindo o número de questões feitas e a quais objetivos específicos atendem.

⁵ Os questionários foram reenviados sucessivamente até se chegar em 42 respostas, que foram consideradas para análise qualitativa e quantitativa.

Quadro 8 - Estrutura do questionário.

QUESTÕES	DETALHAMENTO	Nº. DE QUESTÕES	OBJETIVOS ESPECÍFICOS
REQUISITOS PARA CRIAÇÃO DE PORTAL SEMÂNTICO	Uso de tecnologias de padrão aberto e formas de armazenamento de dados	6	2, 4, 5
	Formas de gestão da informação	10	2, 4, 5
	Gestão de Ontologia	8	2, 4, 5
CARACTERIZAÇÃO SOCIOTÉCNICA DO PORTAL	Atores envolvidos e avaliação do portal	7	1, 3, 4, 5

Fonte: Autoria própria (2021).

Detalhando as questões e correlacionando-as com a Revisão de Literatura tem-se, sequencialmente, quadros específicos para os grupamentos de temas que levaram à construção do Questionário.

As 6 questões elaboradas para verificar o uso de tecnologias de padrão aberto estão apresentadas no Quadro 9. O conceito de padrão aberto implica em que um determinado dado esteja disponível para todos, sem restrições (BRANDT; VIDOTTI; SANTAREM SEGUNDO, 2018; OPEN KNOWLEDGE INTERNACIONAL, 2019). Sistemas de informação, baseados em padrões abertos facilitam a manipulação de dados por *softwares* diversos. No ambiente Web, o uso de padrões abertos internacionalmente debatidos e aceitos, facilita a participação de todos os interessados no desenvolvimento da Web, evitando-se a dependência de fornecedores e de tecnologias proprietárias, que restringem o acesso aos dados e informações (OPEN KNOWLEDGE INTERNACIONAL, 2019).

Quadro 9 - Questões sobre armazenamento de dados e tecnologias de padrão aberto.

QUESTÕES SOBRE ARMAZENAMENTO DE DADOS E SUA GESTÃO	
Questão	Referências
Q1: Utiliza alguma tecnologia de padrão aberto na estruturação do portal	Open Knowledge International (2019)
Q2: Sistema usado para repositório de documentos e navegação pelos dados	Reyes-Álvarez; Roldán-García; Ldana-Montes (2019)
Q3: Sistema operacional utilizado	Thangsupachai; Niwatthanakul; Chamnongsri (2015)
Q4: Banco de dados utilizado	Azani (2009), Lausen <i>et al.</i> (2005);
Q5: Servidor Web utilizado	Elmasri; Navathe (2011); Masner <i>et al.</i> (2019);
Q6: Sistema de segurança para rede e nuvem	Rosenfeld; Morville; Arango (2015); Stallings (2015); W3C (2019)

Fonte: Autoria própria (2021).

Para analisar as formas de Gestão da Informação foram elaboradas 10 questões listadas no Quadro 10.

Quadro 10 - Questões sobre formas e gestão da informação.

QUESTÕES SOBRE FORMAS DE GESTÃO DA INFORMAÇÃO	
Questão	Referências
Q:7 Sistema usado para pesquisa no portal	Hawke et al. (2013)
Q:8 Forma de gestão do conteúdo	Lausen et al. (2005).
Q:9 Como é feita a edição de itens de informação	W3C (2019)
Q:10 Forma de atualização das informações	Masner et al. (2019)
Q:11 Forma de processamento de informações	Börner et al. (2012)
Q:12 Forma de publicação de informações	Stanescu (2018)
Q:13 Sistema de gerenciador de conteúdo utilizado	Berners-Lee; Hendler; Lassila (2001)
Q:14 Mecanismos de inferência utilizado	Koivunen; Miller (2001)
Q:15 Linguagens de programação e modelo de dados usados	
Q:16 A disponibilização da informação no portal segue os requisitos da Lei Geral de Proteção de Dados Pessoais (LGPD)?	

Fonte: Autoria própria (2021).

Para analisar as formas de gestão de Ontologia foram elaboradas 8 questões listadas no Quadro 11.

Quadro 11 - Questões sobre gestão de ontologia

QUESTÕES SOBRE FORMAS DE GESTÃO DE ONTOLOGIA	
Questão	Referências
Q:17: Utiliza editor de ontologia?	Raza et al. (2018)
Q:18 Utiliza tecnologias para importação e exportação de dados com vistas ao reuso de dados e informações?	Malhotra; Nair (2015)
Q:19 Possibilita a inclusão de membros com permissões para escrever conceitos, instâncias, atributos e colaborar na construção de ontologias?	Berners-Lee; Hendler; Lassila (2001)
Q:120 Apresenta descrição de processamento semântico ou estrutura da ontologia com materiais para consulta?	Koivunen; Miller (2001)
Q:21 Apresenta informação semanticamente completa e consistente?	Amato et al. (2017)
Q:22 Apresenta profundidade de cobertura das informações?	Lausen et al. (2005)
Q: 23- Há controle no uso de vocabulários /, evitando-se termos ambíguos que possuem o mesmo significado?	Benabderrahmane et al. (2017)
Q:24 Há maturidade de implantação/manutenção do portal sem erros ou <i>link</i> quebrado?	Li et al. (2017)
	Silvello et al. (2017)
	Barba-González et al. (2019)
	Reyes-Álvarez; Roldán-García;
	Aldana-Montes (2019)
	Paiva; Costa; Silva (2014)
	Maedche et al. (2001)
	Maedche; Staab (2001)
	Stanescu (2018)
	El Haji; Azmani; El Harzli (2014)

Fonte: Autoria própria (2021).

As 7 questões elaboradas para avaliação do acesso à informação e tecnologias usadas no portal segundo a percepção dos sujeitos pesquisados estão listadas no Quadro 12.

Quadro 12 – Questões sobre gestão e avaliação do portal.

QUESTÕES SOBRE CARACTERIZAÇÃO SOCIOTÉCNICA DO PORTAL	
Questão	Referências
Q:25 Identifique os integrantes da equipe de administradores do portal	Gavrilova (2011) Huan; Maidment; Tian (2011)
Q:26 Identifique os integrantes da equipe de desenvolvimento/manutenção do portal:	Lachtim; Moura; Cavalcanti (2009) Lausen <i>et al.</i> (2005)
Q:27 Indique a participação dos usuários na construção/melhoramento do portal	Ngo; Cao (2018) Callon; Latour (1981)
Q:28 Os usuários dispõem de informações relevantes do domínio específico da comunidade acadêmica?	Callon (1986) Law (1992) Latour (2012)
Q:29 Os usuários dispõem de mapas de tópicos pessoais para melhoria da gestão de informação?	Burns (2006) Burns; Dietz (2001)
Q:30 Que recursos para colaboração, comunicação e compartilhamento de informação os usuários dispõem?	Hughes (1989) Iyamu; Mgudlwa (2018)
Q:31 Os usuários dispõem de mecanismos de ajuda ou mapa do site com funcionalidades do portal?	Granovetter (1983)

Fonte: Autoria própria (2021).

O questionário completo encontra-se no Apêndice 2. As respostas a estas questões contribuíram na construção de indicadores de potencial semântico para portais conforme pressuposto já citado na seção de Introdução deste estudo:

- Um conjunto de indicadores para avaliação do potencial semântico de portais constituirá em uma ferramenta de gestão para construção de portais voltados para ações de criação, compartilhamento e difusão do conhecimento dentro ou fora do âmbito acadêmico.

Salienta-se que se utilizou da análise estatística descritiva para organizar, resumir e descrever os dados coletados a partir dos questionários recebidos (TRIOLA, 2017).

As questões do Questionário, mas não apenas estas, permitiram o atendimento de cada um dos objetivos propostos nesta pesquisa; na Seção seguinte descreve-se,

a título de ordenamento o protocolo estabelecido para atendimento aos objetivos de pesquisa.

5.3.3 Protocolo de Coleta de Dados

No protocolo de orientação de coleta de dados destacam-se os procedimentos relativos ao cumprimento do objetivo geral e dos objetivos específicos, com delineamento do instrumento de coleta, consubstanciado na revisão da literatura e na análise de conteúdo. É importante destacar que este protocolo norteou as ações da pesquisadora e que as perguntas do questionário foram adaptadas à linguagem dos respondentes.

Os protocolos de pesquisa deste estudo são apresentados nos Quadros 14 a 18, arrolando os objetivos específicos associados às perguntas que fizeram parte do questionário para coleta de dados.

Para atender o objetivo 1, foi elaborada uma análise dos requisitos básicos levantados por Guedes (2015) e em observância aos preceitos da Web Semântica. Adicionalmente foi destacada a associação entre portais semânticos e redes sociotécnicas, seguidos ainda, pela pesquisa junto às universidades federais brasileiras.

Guedes (2015) agrupou os requisitos básicos em 4 categorias a saber: (i) - sistemas de padrão aberto; (ii) - armazenamento de dados, (iii) - gestão da informação; e (iv) - gestão de ontologia. Berners-Lee, Hendler e Lassila, (2001) e Lausen *et al.* (2005) apresentaram um conjunto de técnicas e preceitos necessários para a estruturação semânticas de portais.

Neste estudo, as dimensões propostas por Guedes (2015) foram mantidas, mas com reagrupamento de perguntas em 3 dimensões – (i) Armazenamento de dados e sua gestão, (ii) Gestão da Informação e (iii) Gestão de Ontologia – acrescentando-se mais uma dimensão para associar questões referentes à Rede de Atores e à caracterização sociotécnica.

A partir da análise dos estudos de Burns (2006), Hughes (1989) e Thomas (2008) depreende-se que os portais do conhecimento semânticos de universidades estão sujeitos a normas e a regras que envolvem atores humanos e não humanos, caracterizando, assim, redes sociotécnicas que têm por objetivo atender as necessidades informacionais da comunidade acadêmica e da sociedade em geral.

Assim, os interesses de todos os atores, humanos e não humanos precisam ser equalizados na rede de portais semânticos.

As questões que nortearam a busca no levantamento junto às universidades encontram-se no Quadro 13, a seguir.

Quadro 13 - Protocolo de coleta de dados – Objetivo específico 1.

Objetivo geral: Propor indicadores qualitativos contendo parâmetros para a avaliação do potencial semântico de portais do conhecimento de universidades, de acordo com os conceitos da TAR.		
Objetivo específico 1: Levantar as condições básicas para o estabelecimento do potencial semântico de portais, como um artefato sociotécnico		
Dimensão	Questões Norteadoras	Referências
Armazenamento de dados e sua gestão	Quais as tecnologias usadas para armazenamento de dados em um portal semântico?	Aksoy (2016); Bellandi <i>et al.</i> (2012); Fernández <i>et al.</i> (2011); Ferneda (2003); Gavrilova (2011); Lausen <i>et al.</i> (2005); Lee <i>et al.</i> (2018); Abdullar; Zamil (2018); Malhotra; Nair (2015); Ngo; Cao (2018); Zerarga; Djouadi (2018)
Gestão da informação	Como é feita a gestão da informação em portais semântico, quanto a: tipo de sistema de recuperação da informação, gerenciamento do conteúdo; edição, atualização, processamento e publicação da informação; mecanismos de inferência?	Lausen <i>et al.</i> (2005); Abdullar; Zamil (2018); Lee <i>et al.</i> (2018); Ngo; Cao (2018); Aksoy (2016); Malhotra; Nair (2015); Ferneda (2003); Gavrilova (2011); Zerarga; Djouadi (2018); Fernández <i>et al.</i> (2011); Bellandi <i>et al.</i> (2012)
Gestão da ontologia	De que forma é feita a gestão da ontologia nos portais semânticos quanto a: ferramentas e tecnologias, serviços e facilidades?	Guarino (1998); Gruber (1996); Telnov (2015); Zarri (2014); Lee <i>et al.</i> (2018); Malhotra; Nair (2015); Shirisha; Subbarao; Kavitha, (2015); Hawke <i>et al.</i> (2013); Amato <i>et al.</i> (2017); <i>et al.</i> (2017); Li <i>et al.</i> (2017); Silvello <i>et al.</i> (2017); Guo <i>et al.</i> (2018); Aroua; Mourad (2017); Moura <i>et al.</i> (2015); Ta; Thi (2015); Berners-Lee (2006); Lausen <i>et al.</i> (2005); Heath; Bizer (2011); Pattuelli; Miller (2015); Silvello <i>et al.</i> (2017); Wimmer; Yoon; Rada (2013); Paiva; Costa; Silva (2014)
Rede de atores e caracterização sociotécnica	Porque os portais semânticos podem ser considerados artefatos sociotécnicos?	Li <i>et al.</i> , 2017; Malgaonkar e Devale. (2016); Burns (2006); Burns; Dietz (2001); Martínez-García; Hernández-Lemus (2013); Hughes (1989); Iyamu; Mgudlwa (2018); Forlano (2017); Rose; Scheepers (2001); Law (1992); Latour (2017); Latour (2012); Erofeeva (2017); Callon (1986a, b, c); Bijker; Hugh; Pinch (1989); Granovetter (1983)

Fonte: Autoria própria (2021).

Para o objetivo específico 2 desta pesquisa, que trata de mapeamento de atores, foi realizada análise conceitual, seguida do levantamento junto às universidades federais brasileiras.

De acordo com Martínez-García e Hernández-Lemus (2013), a análise de sistemas complexos, como por exemplo, os sistemas tecnológicos, tem sido comumente efetuada por meio da abordagem de rede, derivada da teoria dos grafos, que estuda as relações entre os objetos de um determinado conjunto. Assim, os sistemas complexos foram concebidos como gráficos ou redes nos quais os agentes ou constituintes são nós (ou vértices) e a interação entre eles são *links* (ou arestas). Estes gráficos serão inicialmente levantados, de acordo com a literatura.

De outra parte, a Teoria Ator-Rede - TAR, como metodologia de estudo, tem sido utilizada para acompanhar e avaliar desenhos e aparatos técnicos. De acordo com Prado e Baranauskas (2012), o foco da TAR está nas associações formadas entre os atores, resultando em redes e sua transformação. Na prática, acentua Walsham (1997), a TAR é aplicada em estudos de sistema de informação para tentar rastrear e explicar os processos pelos quais redes relativamente estáveis de interesses alinhados são criadas e mantidas durante um período de tempo, e para examinar porque tais redes se estabelecem ou não se estabelecem. Para Macome (2008), a estabilidade das redes depende da integração entre seus atores humanos e não humanos. A TAR, portanto, orientará a identificação dos rastros dos atores.

As questões que orientaram a busca no levantamento junto às universidades encontram-se no Quadro 14, a seguir.

Quadro 14 – Protocolo de coleta de dados – Objetivo específico 2.

Objetivo geral: Propor indicadores qualitativos contendo parâmetros para a avaliação do potencial semântico de portais do conhecimento de universidades, de acordo com os conceitos da TAR.		
Objetivo específico 2: Mapear os atores que compõe a rede estrutural de portais semânticos		
Dimensão	Questões Norteadoras	Referências
Humanos	Quais são os atores humanos envolvidos no portal semântico? Quais são os interesses dos atores humanos na rede de portais semânticos?	Rosenfeld; Morville; Arango (2015); Maniega-Legarda; Pérez-Salmerón; Guerrero-Torres (2008); Elmasri; Navathe (2011); Stallings (2015); Masner <i>et al.</i> (2019)
Não-Humanos	Quais são os atores não-humanos envolvidos no portal semântico? Como os interesses dos atores não-humanos são traduzidos na rede de portais semânticos?	Spyns <i>et al.</i> (2002); Open Knowledge International (2019) Nguyen <i>et al.</i> (2020); Bizer, Heath e Berners-Lee (2009); Koivunen; Miller (2001); Lausen <i>et al.</i> (2005); Berners-Lee, Hendler e Lassila (2001); ; Börner <i>et al.</i> , 2012); Lawan; Rakib (2019); Azani (2009); Gruber (1996); Guarino (1997); Malhotra; Nair (2015)

Fonte: Autoria própria (2021).

O atendimento do objetivo 3 foi efetuado mediante análise dos interesses dos atores humanos e não humanos à luz da TAR. A partir da literatura, caracterizou-se as tecnologias e suas funções, entendidas como interesses dos atores não-humanos; já os atores humanos e suas ações foram canalizadas em interesses em prol da evolução do portal como artefato sociotécnico.

Callon (1986 a, b, c), Law (1992) e Latour (2001) destacam que os atores da rede deixam traços que permitem identificar as lacunas, interpretar necessidades e apontar negociações. Esses elementos permitem realinhar conceitos e indicar caminhos em prol da construção ou melhoramento de portais do conhecimento de universidades. Nos traços da rede estão presentes as variáveis de interesse deste estudo (MALHOTRA, 2019), sendo possível, a partir da análise relacional aliada à análise de dados, verificar as condições da estrutura técnica e de pessoal dos portais, auxiliando o estabelecimento de indicadores de potencial semântico para avaliação destes portais.

As questões que identificaram os interesses dos atores dos portais do conhecimento de universidades foram delineadas no Quadro 15. As variáveis de interesse que estão relacionadas aos atores humanos incluíram os usuários, a equipe de administradores e de desenvolvedores do portal. Já as variáveis de interesse dos atores não humanos abarcaram as tecnologias, os mecanismos e as funcionalidades de um portal semântico.

Quadro 15 – Protocolo de coleta de dados – Objetivo específico 3.

Objetivo geral: Propor indicadores qualitativos contendo parâmetros para a avaliação do potencial semântico de portais do conhecimento de universidades, de acordo com os conceitos da TAR.		
Objetivo específico 3: Elencar os indicadores qualitativos iniciais, considerando os atores envolvidos, seus interesses, funções e características		
Dimensão	Questões Norteadoras	Referências
Rede de atores e caracterização sociotécnica	Quais são os interesses dos atores humanos na construção de portais semânticos?	Callon; Latour (1981); Callon (1986); Law (1992); Latour (2012); Shim; Shin (2016); Latour (2017); Prado; Baranauskas (2012); Iyamu;
	Quais são os interesses dos atores não-humanos na construção de portais semânticos?	Mgudlwa (2018); Elbanna (2012); Cordella (2010); Hu (2011); Macome (2008)

Fonte: Autoria própria (2021).

Para atender o objetivo específico 4 desta pesquisa, que trata da validação dos indicadores, foi encaminhado um novo questionário para 22 dos especialistas, aqueles respondentes do questionário de pesquisa que manifestaram explicitamente o

interesse em receber os resultados finais da tese. Esse segundo questionário – ver Apêndice 3 - incluiu o conjunto de indicadores levantados com base na literatura revisada e na confrontação com os requisitos propostos, seguido pela análise do levantamento feito junto às universidades, tendo sido recebidas 11 respostas. O detalhamento do processo de validação encontra-se na Seção 5.3.3.

Estudos clássicos de Koivunen e Miller (2001), Berners-Lee Hendler e Lassila (2001), Lausen *et al.* (2005) e Maedche e Staab (2001) preveem o uso de tecnologias apropriadas e condições técnicas específicas para notação semântica de portais. Nguyen *et al.* (2020) e Masner *et al.* (2019) ratificaram a importância do uso de tecnologias de padrão aberto. Estudos de Li *et al.*, 2017; Silvello *et al.* (2017); Stanescu (2018) e Lee *et al.*, (2018), reafirmam a necessidade do uso agentes inteligentes, modelagem ontológica e específica de dados para otimização de buscas nos portais. Tais diretrizes foram empregadas para formulação de questões que figuram no Quadro 16.

Quadro 16 – Protocolo de coleta de dados – Objetivo específico 4.

Objetivo geral: Propor indicadores qualitativos contendo parâmetros para a avaliação do potencial semântico de portais do conhecimento de universidades, de acordo com os conceitos da TAR.		
Objetivo específico 4: Validar os indicadores elencados a partir de um levantamento (survey) em universidades e das percepções de um painel de especialistas		
Dimensão	Questões Norteadoras	Referências
Armazenamento de dados e sua gestão	Utiliza alguma tecnologia de padrão aberto na estruturação do portal? Qual o sistema usado para repositório de documentos e navegação pelos dados Qual o sistema operacional utilizado? Qual o banco de dados utilizado? Qual o servidor Web utilizado Qual o sistema de segurança para rede e nuvem?	Open knowledge international (2019); Thangsupachai, Niwatthanakul; Chamnongsri (2015); Reyes-Álvarez; Roldán-García; Ldana-Montes (2019); Nguyen <i>et al.</i> (2020); Lausen <i>et al.</i> (2005)
Gestão da informação	Qual o sistema usado para pesquisa no portal? Qual a forma de gestão do conteúdo? Como é feita a edição de itens de informação? Qual a forma de atualização das informações? Qual a forma de processamento de informações? Qual a forma de publicação de informações? Quais os mecanismos de inferência utilizados?	Hawke <i>et al.</i> (2013); Lausen <i>et al.</i> (2005); W3C (2019); Masner <i>et al.</i> (2019); Börner <i>et al.</i> (2012); Stanescu (2018); Berners-Lee; Hendler; Lassila (2001); Koivunen; Miller (2001); Li <i>et al.</i> (2017); Silvello <i>et al.</i> (2017); Maedche e Staab (2001); Maedche <i>et al.</i> (2001)

(continua)

Cont. do quadro 17, Objetivo específico 4: Validar os indicadores elencados a partir de um levantamento (<i>survey</i>) em universidades e das percepções de um painel de especialistas		
Dimensão	Questões Norteadoras	Referências
Gestão da ontologia	<p>O portal utiliza ontologia em sua estruturação?</p> <p>O portal utiliza editor de ontologia?</p> <p>O portal utiliza tecnologias para importação e exportação de dados?</p> <p>Possibilita a inclusão de membros com permissões para escrever conceitos, instâncias, atributos e colaborar na construção de ontologias</p> <p>Quais as linguagens de programação são usadas que favorecem a semântica?</p> <p>Apresenta descrição de processamento semântico ou estrutura da ontologia com materiais para consulta?</p> <p>O portal apresenta informação semanticamente completa e consistente?</p>	<p>Raza <i>et al.</i> (2018); Malhotra; Nair (2015) ; Berners-Lee; Hendler; Lassila (2001); Koivunen; Miller (2001); Amato <i>et al.</i>, (2017); Lausen <i>et al.</i> (2005); <i>et al.</i> (2017); Li <i>et al.</i> (2017); Silvello <i>et al.</i> (2017); Barba-González <i>et al.</i> (2019); Reyes-Álvarez; Roldán-García; Aldana-Montes (2019); Paiva; Costa; Silva (2014); Maedche <i>et al.</i> (2001); Maedche; Staab (2001) ; Stanescu (2018); El Haji; Azmani; El Harzli (2014); Stanescu (2018)</p>
Rede de atores e caracterização sociotécnica	<p>Quais profissionais administram o conteúdo do portal?</p> <p>Quais profissionais de TI que manipulam o portal?</p> <p>Os são os usuários do portal?</p> <p>Os usuários participaram da construção do portal?</p> <p>O portal apresenta profundidade de cobertura das informações?</p> <p>O portal apresenta informações relevantes do domínio específico da comunidade acadêmica?</p> <p>Há maturidade de implantação do portal sem erros ou links quebrados?</p> <p>O portal dispõe de mapas de tópicos pessoais para melhoria da gestão de informação?</p> <p>O portal disponibiliza recursos para colaboração, comunicação e compartilhamento de informação tais como boletins eletrônicos, lista de endereços, lista de discussão, chat, fórum)?</p> <p>Há controle no uso de vocabulários e ações ambíguas que significam os mesmos conceitos?</p> <p>Dispõe de mecanismos de ajuda ou mapa do site com funcionalidades do portal?</p>	<p>Gavrilova (2011); Huan; Maidment; Tian (2011);</p> <p>Lachtim; Moura; Cavalcanti (2009); Lausen <i>et al.</i> (2005); Ngo; Cao (2018); Callon; Latour (1981); Callon (1986 a, b, c); Law (1992); Latour (2012);</p> <p>Burns (2006); Burns; Dietz (2001); Hughes (1989); Iyamu; Mgodlwa (2018); Granovetter (1983).</p>

Fonte: Autoria própria (2021).

Para contemplar o objetivo 5, que trata da apresentação dos indicadores qualitativos para avaliação de portais, a formulação destes foi embasada sobretudo

nos preceitos semânticos de estruturação de portais como delineado por Guedes (2015) cujo estudo tem por normativas as diretrizes da Web semântica.

A partir do estudo de Andrade, Vaitsman e Farias (2010), depreende-se que os indicadores podem ser propostos para traduzir as dimensões da realidade dos portais do conhecimento de universidades, traduzindo-as empiricamente em valores investigados. Sendo assim, o Quadro 17 registra a pergunta que será respondida pelos achados das demais questões formuladas.

Quadro 17 – Protocolo de coleta de dados – Objetivo específico 5.

Objetivo geral: Propor indicadores qualitativos contendo parâmetros para a avaliação do potencial semântico dos portais do conhecimento de universidades, de acordo com os conceitos da TAR		
Objetivo específico 5: Apresentar indicadores qualitativos para avaliação de portais		
Variável	Pergunta	Referências
Indicadores	É possível estabelecer indicadores qualitativos para avaliação de portais semânticos?	Guedes e Strauhs (2016); Andrade; Vaitsman; Farias (2010); Last, 1998; Ambrozewicz, 2015; Zatsiorski (1989); Van Bellen (2005)

Fonte: Autoria própria (2021).

A partir do mapeamento dos atores da rede de portais semânticos foi possível descrever os elementos humanos e não humanos e suas inter-relações (objetivo específico 1), tendo como referencial as condições básicas elencadas para a construção de um portal semântico (objetivo específico 2). A partir dessas etapas, o interesse de cada ator foi descrito e confrontado com os demais atores (objetivos específicos 3 e 4) para estabelecer indicadores qualitativos de potencial semântico (objetivo 5).

Na próxima Seção será detalhada a forma de tratamento dos dados que envolve a descrição dos instrumentos de análise, as inferências e a interpretação dos resultados.

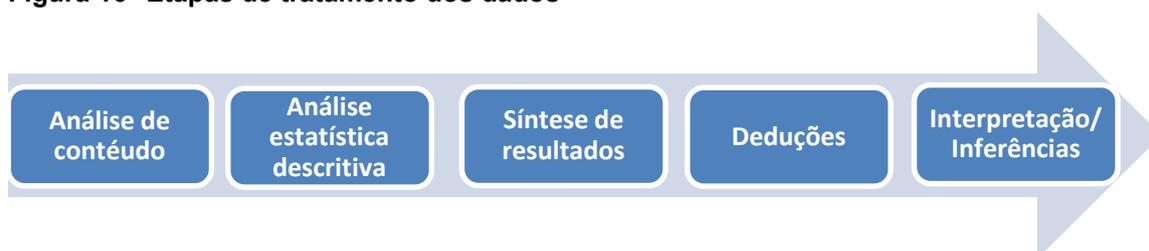
5.3.3 Tratamento e validação dos dados

Os conteúdos relacionados ao tema em estudo, categorizados e codificados, representam dados brutos que precisam ser tratados para se tornarem significativos. O tratamento desses dados será efetuado mediante análise de conteúdo e análise estatística descritiva para, à luz da literatura, embasar discussões e inferências.

Segundo Bardin (2011) a dedução lógica, que usa aspectos de inferência, é uma fase intermediária entre a descrição dos dados e o significado concedido aos

dados. A inferência enquanto operação lógica, possibilita ligações de uma proposição com outras proposições consideradas como verdadeiras. Desta forma, o tratamento de dados qualitativos e quantitativos possibilita a consecução dos objetivos traçados no estudo. A Figura 16 ilustra as etapas do tratamento dos dados.

Figura 16 - Etapas de tratamento dos dados



Fonte: Autoria própria (2021).

A análise de conteúdo indicará, a partir das unidades de contexto e de registro, os principais parâmetros associados ao tema em estudo, sendo a seguir efetuada a análise estatística descritiva dos dados coletados pelos questionários. A partir dessa análise foi produzida uma síntese dos resultados e a partir de deduções, foi possível interpretar e realizar inferências que auxiliaram na construção dos indicadores de potencial semântico para portais do conhecimento semântico de universidades.

Considerando que as variáveis analisadas neste estudo são de caráter qualitativo, o tratamento dos dados pautou-se na média simples, com foco nas variáveis de interesse e que refletem percepções e a compreensão do contexto do problema de pesquisa delineado, conforme destacado por Malhotra (2019).

Após a análise qualitativa das respostas, confrontadas com a literatura, foi efetuada também uma pesquisa direta nos portais por esta pesquisadora como profissional da Informação, para fazer um contraponto de algumas respostas, a fim de identificar as características, sobretudo, aquelas associadas à busca e a recuperação da informação.

A partir da análise das variáveis de interesse, foi estabelecido o conjunto inicial de indicadores que na sequência foram validados por um painel de especialistas, considerando que esta etapa é importante no desenvolvimento de novas medidas associadas a conceitos abstratos (LAST, 1998). Para cada questão do questionário foi identificado a variável de interesse associada a uma alternativa de resposta que por sua vez foi fundamentada na literatura, a partir das teorias da Web Semântica, da Recuperação da Informação e da Teoria Ator-Rede.

Salienta-se que na literatura não há um consenso sobre o número de um painel de especialistas para um processo de validação de um conteúdo. Alguns autores sinalizam o uso de 2 a 20 integrantes (RUBIO *et. al.*, 2003; WALZ; STRICKLAND; LENZ, 1991). Zatsiorski (1989) pontua um número ideal de especialistas entre 15 e 30. Davis (1992) ressalta que o número de especialistas pode ser dependente do nível desejado de experiência e da diversidade de conhecimento. Crespo Borges (2007) destaca a importância da representatividade, ou seja, o painel de especialistas não deve ser muito grande ou muito pequeno e sugeriu um cálculo de margem de erro obtido com respostas de até 30 especialistas respondentes. Aplicando esse cálculo neste estudo, considerando as respostas dos 22 especialistas determinados, a margem de erro assumida para a resposta da validação seria de 2%, dentro de parâmetros estatísticos aceitáveis.

Do aspecto de aceitação do indicador, no processo de validação pelos especialistas, para ser considerado válido o indicador deve receber julgamentos de opiniões convergentes, com nível favorável mínimo de 75%, seguindo as indicações de Donabedian (1985), parâmetro adotado neste estudo.

Estabelecido o desenho metodológico da pesquisa, a Figura 17 ilustra a estrutura de apresentação dos dados que será feita no Capítulo subsequente, incluindo a apresentação do universo da pesquisa, o desenho da rede de atores do portal delineado e a validação de dados.

Figura 17 – Estrutura do capítulo 6.

APRESENTAÇÃO DE DADOS	REDE DE ATORES DOS PORTAIS DO CONHECIMENTO DE UNIVERSIDADES	ANÁLISE E VALIDAÇÃO DOS DADOS
Caracterização do universo da pesquisa Instrumento de coleta	Atores Humanos da Rede de Portais do Conhecimento de Universidades Atores Não-Humanos do Conhecimento de Universidades	Variáveis de interesse Proposição de indicadores Validação dos indicadores

Fonte: Autoria própria (2021).

Finalizado o detalhamento metodológico, apresenta-se no próximo Capítulo a análise dos dados coletados frente às variáveis de interesse e às discussões consubstanciadas na literatura consultada.

6 APRESENTAÇÃO DOS DADOS E DISCUSSÃO DOS RESULTADOS DA PESQUISA

Neste Capítulo serão apresentados os dados resultantes das diversas fases de pesquisa com prevalência para a análise qualitativa das variáveis de interesse. A rede de atores nos portais, os serviços, as ferramentas, as tecnologias utilizadas nestes e os indicadores qualitativos de potencial semântico para portais do conhecimento de universidades auferidos, igualmente serão apresentados.

6.1 APRESENTAÇÃO DOS DADOS DA PESQUISA

Neste estudo, conforme já mencionado, foi utilizado o questionário como um dos instrumentos de coleta de dados de campo; os questionários foram enviados para os coordenadores da área de Tecnologia e Informação responsáveis pelos portais em cada universidade federal brasileira, após a adequação nas opções de perguntas, consonante ao indicado no pré-teste – o estudo piloto – descrito na Seção 5.3.2.

O intuito do questionário foi identificar as tecnologias, os recursos humanos e os serviços utilizados nos portais e confrontá-los com a estruturação técnica apontada na literatura e que fundamentam as condições necessárias para estabelecer o potencial semânticos de portais. Os dados coletados foram usados para validar os parâmetros elencados por Guedes (2015) e fundamentaram a construção de indicadores qualitativos semânticos, considerando os atores envolvidos, seus interesses, funções e características sociotécnicas.

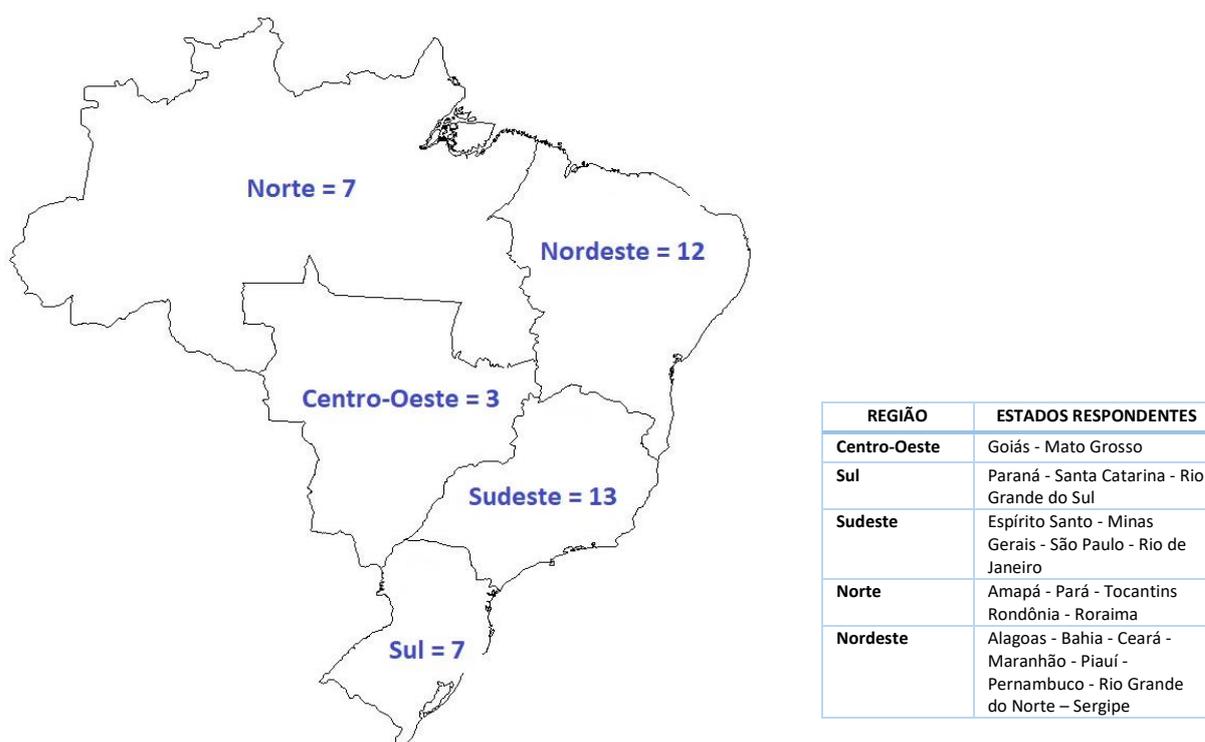
Ressalta-se que as universidades federais brasileiras estão distribuídas nas cinco regiões do país, totalizando 69 instituições, sendo que as regiões Sudeste (n=19) e Nordeste (n=20), concentram um maior número de universidades para atender estas localidades que são as mais populosas do país (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, 2021), seguidas pela região Sul (n=11), Norte (n=11) e Centro-Oeste (n= 8).

Dentre as universidades mais antigas, criadas a partir de 1912, rememora-se, destacam-se a Universidade do Paraná, a Universidade do Amazonas, a Universidade do Rio de Janeiro e Universidade São Paulo (SOUZA; MIRANDA; SOUZA; 2019; UNIVERSIDADE FEDERAL DO PARANÁ, 2021). Em 2019, foram criadas as cinco

mais novas universidades: a Universidade Federal de Jataí, a Universidade Federal do Agreste de Pernambuco; a Universidade Federal de Rondonópolis, Universidade Federal do Delta do Parnaíba e a Universidade Federal de Catalão (MASSALLI, 2019).

Da totalidade das universidades pesquisadas foram obtidas 42 respostas que atenderam à amostra mínima conforme delineado na Seção 5.3.1. A Figura 18 ilustra a localização das universidades por região e que responderam ao questionário de pesquisa, observando-se uma concentração maior do número respostas nas regiões Sudeste e Nordeste, seguindo o padrão de concentração das universidades no país. De outra parte, salienta-se, obteve-se respostas de todas as regiões brasileiras.

Figura 18 - Distribuição das universidades federais nas regiões do país com indicação do número de respostas recebidas ao Questionário.



Fonte: Autoria própria (2021).

Ao receber-se resultados de todas as regiões brasileiras em número significativo, constatou-se que se atendeu, de forma inopinada, aos critérios de amostragem probabilística estratificada, em que a população total pode ser dividida em grupos e subgrupos a fim de se ter uma amostra representativa (TRIOLA, 2017).

Sendo assim, os dados desta pesquisa refletiram com maior precisão o panorama dos portais das universidades federais brasileiras (Tabela 4).

Tabela 4 - Respostas do questionário por estratos da população.

REGIÃO	POPULAÇÃO	%	AMOSTRA	%
Norte	11	15,94	7	16,67
Nordeste	20	28,99	12	28,57
Centro-Oeste	8	11,59	3	7,14
Sudeste	19	27,54	13	30,95
Sul	11	15,94	7	16,67

Fonte: Autoria própria (2022).

Desta forma, e conforme os propósitos de pesquisa, fundamentando-se nos dados coletados e analisados, nos preceitos da TAR, dos sistemas sociotécnicos e da Web semântica e suas tecnologias (BERNERS-LEE; HENDLER; LASSILA, 2001; CALLON; LATOUR, 1981; GUEDES; STRAUHS 2016; LAUSEN *et al.*, 2005; MAEDCHE; STAAB, 2001), foi possível desenhar, entre outros elementos, a rede conceitual de atores que contribuem para construção de portais semânticos e estabelecer a inter-relação entre eles caracterizando-os como uma rede sociotécnica, conforme delineado no Capítulo 4. A rede de atores será apresentada na Seção seguinte.

6.2 REDE DE ATORES DOS PORTAIS DO CONHECIMENTO DE UNIVERSIDADES

Nesta seção apresenta-se o desenho da rede de portais de universidades, considerando a estruturação semântica e os preceitos sociotécnicos apreendidos da literatura e da coleta de dados. Seguem delineados os atores humanos – os administradores, desenvolvedores e usuários e os atores não-humanos – as tecnologias de estruturação dos portais.

6.2.1 Atores humanos da rede de portais do conhecimento de universidades

As tecnologias da Web Semântica, se aplicadas aos portais possibilitam a formação de redes abertas, dinâmicas complexas, enriquecidas de camadas de semântica provenientes de fontes internas e externas (BERNERS-LEE; HENDLER;

LASSILA, 2001). Para Iyamu e Mgudlwa (2018), as redes dinâmicas formam um sistema de informação complexo que exige análise criteriosa dos processos, das tecnologias, dos papéis e das responsabilidades.

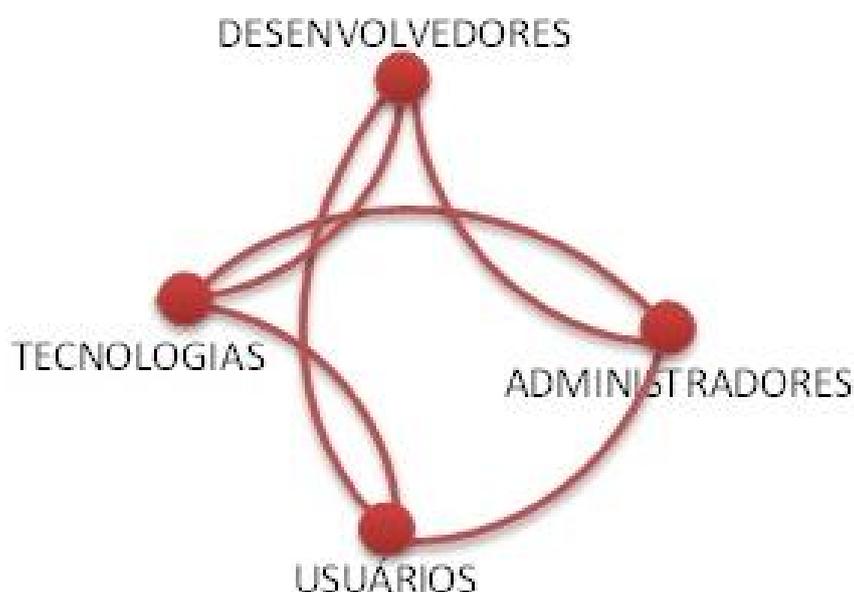
De acordo com Callon e Latour (1981), na Teoria Ator-Rede – TAR, os atores formam alianças e aliciam outros atores para fortalecer estas alianças, criando as redes heterogêneas feitas de actantes humanos e não humanos em que as associações ocorrem mutualmente. Logo, o mapeamento da rede de atores envolvidos na construção de um portal universitário semântico possibilita compreender os procedimentos necessários para disponibilizar uma interface de pesquisa que tem por finalidade promover o acesso a informações acadêmicas com rapidez e precisão.

No presente estudo, a TAR foi aplicada metodologicamente para a construção do que se entende como a rede de atores dos portais do conhecimento de universidades, bem como para orientar o seu desenho, especialmente com base nos processos de tradução já que a partir das diferentes etapas das traduções aprimora-se a compreensão de como as redes surgem, evoluem e se estabilizam (WALSHAM, 1997). Na tradução, um ator, o porta voz, interpreta os interesses de outros atores, sejam humanos ou não humanos, e os realinha, incorporando ou adequando papéis e influenciando os atores de modo que estes formem uma aliança para construir e manter uma rede (CALLON; LATOUR, 1981; LATOUR, 1992).

Então, infere-se que a rede de um portal universitário é composta pelos seguintes elementos básicos – atores, à luz da TAR: 1) Administradores que fazem o gerenciamento global do portal; 2) Desenvolvedores que estruturam e coordenam as ferramentas e os mecanismos para funcionamento do portal; 3) Usuários, que são atores que efetuam buscas no portal; 4) Tecnologias usadas na construção do portal.

Tomando como base a estruturação do portal semântico proposta por Guedes (2015), o processo de tradução da rede do portal envolve fundamentalmente os interesses dos administradores, desenvolvedores, tecnologias e usuários, conforme a Figura 19.

Figura 19 - Rede simples de tradução do portal universitário.



Fonte: Autoria própria (2021).

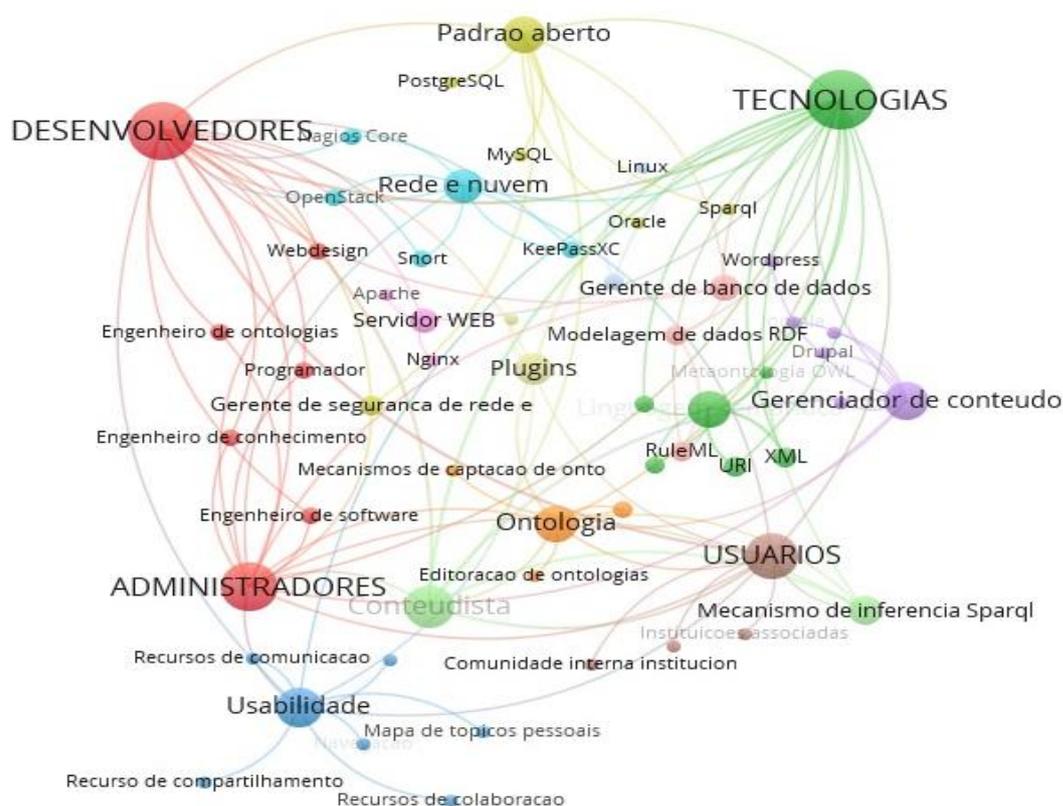
Na rede de portais tem-se como atores humanos a administração do portal, os desenvolvedores e os usuários, que utilizam artefatos não humanos que são as tecnologias para o desenvolvimento e utilização dos portais (ELBANNA, 2012; CORDELLA, 2010; IYAMU; MGUDLWA (2018). Nesse sentido, constitui-se um fluxo em movimentação no qual as práticas são executadas por atores humanos e não humanos. Essa ordenação se torna incorporada em uma rede de práticas.

Latour (2001) afirma que considerar humanos e não humanos tem implicações metodológicas quando da análise de redes, por tratar-se da movimentação ou da descrição do que as redes transportam, sendo preciso estudá-las simetricamente. Então, com base nos estudos da Web Semântica de Berners-Lee, Hendler e Lassila, (2001), da estruturação semântica para portais descrita por Lausen *et al.*, (2005) e do quadro referencial de condições mínimas para estabelecimento de potencial semântico proposto por Guedes (2015) e Guedes e Strauhs (2016), foi possível desdobrar os processos de tradução da rede de um portal, de modo a especificar melhor os traços da rede de um portal universitário semântico, estabelecendo-se as relações entre atores humanos e não humanos.

Os traços da rede - relações entre humanos e não humanos - neste estudo - aparecem entre desenvolvedores e tecnologias, desenvolvedores e administradores

e desenvolvedores e usuários. Os macroatores - tecnologias, desenvolvedores, administradores e usuários - se relacionam entre si e com todos os atores da rede – programadores, mecanismo de inferência, comunidade interna, entre outros recursos, conforme ilustra a Figura 20. Cabe ressaltar que também podem figurar na rede outros atores intangíveis tais como políticas, legislação e cultura organizacional, dentre inúmeros outros elementos, que não foram aqui considerados, pois fogem do escopo desta pesquisa.

Figura 20 - Rede de atores do portal universitário semântico.



Fonte: Autoria própria (2021), com suporte da ferramenta Vosviewer.

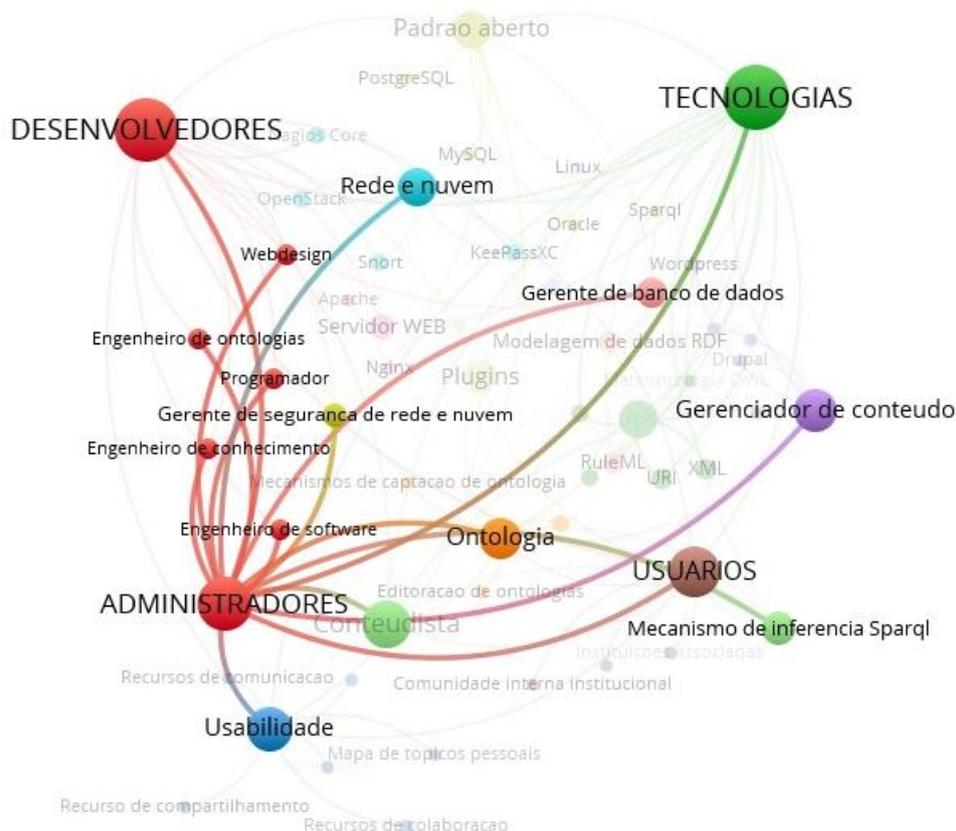
Na rede de portais ocorrem processos de identificação, de mobilização e de atribuição de papéis, além de se influenciar atores formando uma aliança entre eles para que todos apoiem a construção da rede (CALLON; LATOUR, 1981; LATOUR, 1992). O alinhamento dos interesses dos atores é feito por intermédio de um processo de negociação contínuo em que ocorre a reinterpretação, a representação ou a apropriação dos interesses de um ator, de forma que estes possam ser seguidos pelos outros atores na rede (CALLON; LATOUR, 1981). Assim, os atores influenciam a rede e também são influenciados por ela (CORDELLA, 2010; ELBANNA, 2012; PRADO;

BARANAUSKAS, 2012). Exemplificando, para que a pesquisa semântica seja possível, é necessário que a ontologia - elemento não humano - esteja presente na rede para que possa desempenhar suas funções e para tal, os engenheiros de ontologia - atores humanos - precisam dominar as regras de estruturação da ontologia para aplicação na prática. Nesse processo, outros atores devem ser alistados a participar na rede tais como a metaontologia OLW, o modelo de dados RDF, o mecanismo de inferência, necessários para a ontologia, bem como os engenheiros devem também exercer suas funções de forma plena dentro da rede.

Na rede de portais os desenvolvedores, os administradores, as tecnologias e os usuários constituem pontos de passagem obrigatórios. Nesses pontos, de acordo com Callon (1986a), os atores alteram suas preferências a fim de superar barreiras para o alcance do objetivo inicial que é a estruturação semântica dos portais. Em cada ponto de passagem os atores humanos e não humanos se conectam para fortalecer e caracterizar a rede, dotando-as de elementos heterogêneos que apresentam funções específicas dentro da rede.

Os administradores são responsáveis pela definição dos objetivos e do público-alvo, bem como pelo estabelecimento das diretrizes de governança do portal (NIELSEN; HORANDER, 2007; ROSENFELD; MORVILLE; ARANGO, 2015). Os administradores constituem um macroator que se subdivide em microatores: os contedistas, o gerente de conteúdo, o gerente de banco de dados, o gerente de segurança de rede e nuvem. A Figura 21 exemplifica os principais relacionamentos dos administradores no portal semântico levantados na literatura (MANIEGA-LEGARDA; PÉREZ-SALMERÓN; GUERRERO-TORRES, 2008).

Figura 21 - Administradores na rede do portal semântico.



Fonte: Autoria própria (2021), com suporte da ferramenta Vosviewer.

Os conteudistas são responsáveis pela indicação de textos e de recursos informacionais para disponibilização no portal, que contém dados e informações que podem ser utilizados pelos usuários, conforme descrito por Nielsen (2000). O gerente de conteúdo é responsável pelo recebimento das demandas advindas dos conteudistas, pela aprovação e atualização de conteúdos, bem como pela exclusão de conteúdo obsoleto, traduzindo assim os interesses dos conteudistas e dos usuários (MANIEGA-LEGARDA; PÉREZ-SALMERÓN; GUERRERO-TORRES, 2008). Os gerenciadores de conteúdo são mecanismos que traduzem os interesses do gestor de conteúdo e dos conteudistas, sobretudo aqueles que são configurados com padrão aberto tais como: Wordpress, Joomla, Drupal, Textpattern, Radiant, entre outros (WORLD WIDE WEB - W3C, 2019).

O gerente de banco de dados é o responsável por definições acerca da organização de dados distribuídos em repositórios informacionais, esquemas, tabelas, relatórios entre outros objetos que são organizados para modelar aspectos da

realidade, de uma forma que suporte os processos de pedidos de informação (ELMASRI; NAVATHE, 2011).

Neste cenário, os portais convencionais necessitam de um sistema de gerenciamento de dados para ações de criação, consulta, atualização e administração do banco de dados utilizado (ELMASRI; NAVATHE, 2011). Já no portal semântico o banco de dados é constituído pelos dados nativos RDF (*triple stores*), o que constitui uma quebra de paradigma e se configura como uma abertura da caixa-preta dos portais convencionais, já que se trata de um nova modelagem e uma nova forma de estruturar a informação nos portais (BERNERS-LEE; HENDLER; LASSILA, 2001; BIZER; HEATH, BERNERS-LEE, 2009; CALLON, 1986 a, b, c).

A caixa preta dos portais, representada pelo uso de tecnologias convencionais, é aberta para os desenvolvedores e administradores, pela incorporação das tecnologias semânticas na interface dos portais. No entanto, os usuários são beneficiados pela evolução dos portais decorrente da incorporação das novas tecnologias que facilitam o acesso à informação, sendo este um benefício indireto da abertura da caixa preta para os desenvolvedores e administradores.

Na rede de portais semânticos, à luz da Web Semântica e da TAR, novamente ocorre uma reordenação de atores com interesses no uso de bancos de dados gratuitos e de código aberto, tais como: Oracle, MySQL, PostgreSQL, tendo o Linux como sistema operacional (OPEN KNOWLEDGE INTERNACIONAL, 2019; NGUYEN *et al.*, 2020). O banco de dados liga-se com os bancos nativos RDF da modelagem semântica de dados.

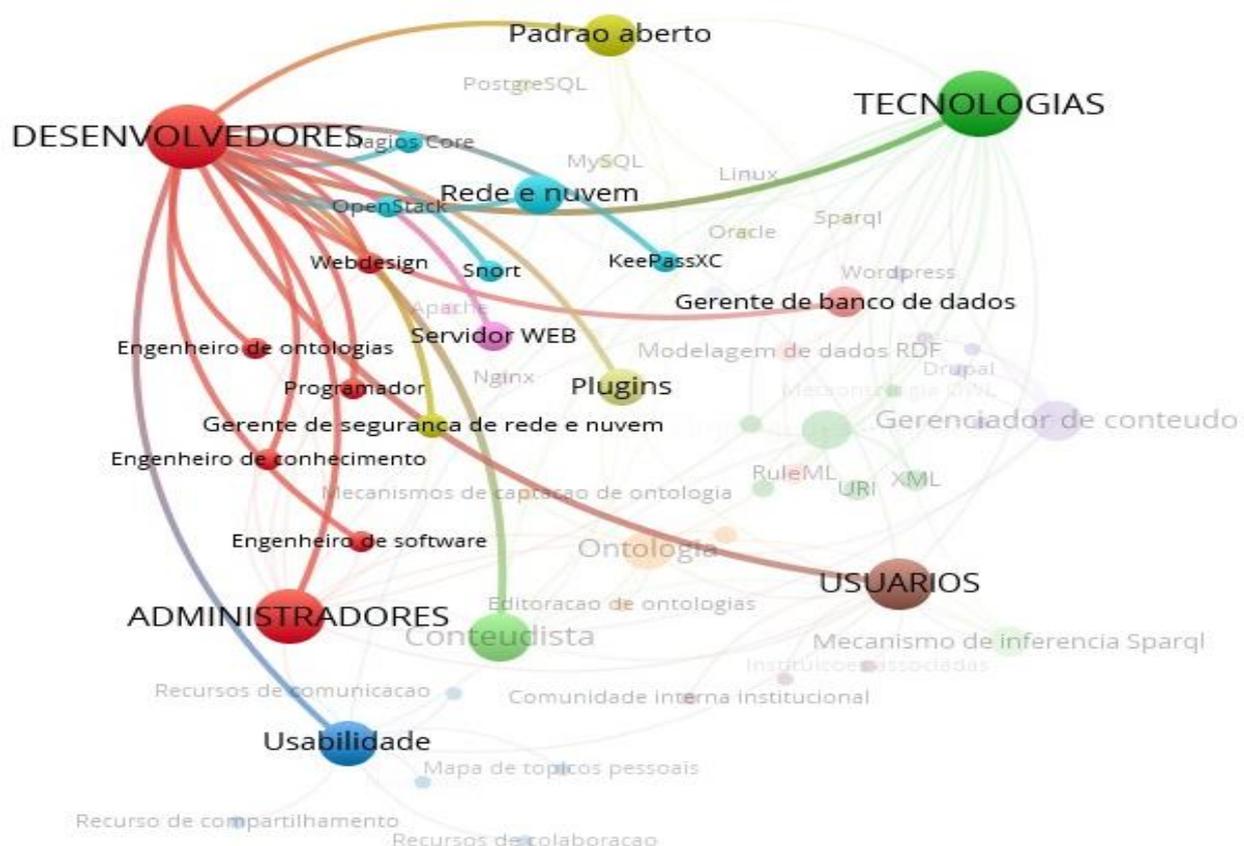
O gerente de segurança local e de nuvem responde pela definição de regras de uso e de controle de credenciais de acesso ao banco de dado, garantindo sigilo de informações (MASNER *et al.*, 2019; ROSENFELD; MORVILLE; ARANGO, 2015; STALLINGS, 2015). Na rede do portal semântico este ator traduz a segurança do banco de dados nativos RDF e dos serviços de nuvem. O convencimento de outros atores da rede de portais permeia o uso de ferramentas para segurança de rede ou nuvem de código aberto a exemplo das ferramentas Nagios Core, Snort, OpenStack, KeePassXC.

Considerando que a construção de portais do conhecimento de universidades surge como uma demanda interna das instituições de ensino, coordenada pelos seus gestores, a tradução dos interesses dos Administradores do portal semântico envolve

a construção do portal a fim de disponibilizar informações relativas ao contexto universitário e que são de interesses da comunidade acadêmica e do público em geral. A estruturação tecnológica do portal deve permitir a entrega de serviços de qualidade para sua comunidade. Esses interesses devem ser alinhados na rede do portal, por meio de processos de negociação que envolvem a agência dos desenvolvedores, das tecnologias e dos usuários do portal semântico (LAW,1992; CALLON; LATOUR, 1981).

Os desenvolvedores constituem outro macroator da rede dos portais semânticos, que se subdividem, igualmente, em microatores compostos por profissionais da área de tecnologia que atuam em tarefas distintas como: programação, engenharia de *software*, engenharia de conhecimento, engenharia de ontologias, *webdesigner* para desenho e usabilidade do portal (NIELSEN; LOHANGER, 2007; ROSENFELD; MORVILLE; ARANGO, 2015) A Figura 22 destaca os relacionamentos dos desenvolvedores na rede.

Figura 22 – Desenvolvedores na rede do portal semântico.



Fonte: Autoria própria (2021), com suporte da ferramenta Vosviewer.

Os desenvolvedores são atores mobilizados pelos interesses dos administradores e dos usuários, assumindo papéis e mobilizando-se na rede para apoiar a construção, a manutenção e a atualização do portal, negociando também a escolha de tecnologias que expressam semântica. Em portais convencionais o *designer* projeta a interface (*layout*) do portal, modulando também questões envolvidas com a usabilidade, mediante testes e aplicação de regras que envolvem a experiência do usuário (NIELSEN; LORANGER, 2007).

Na interface semântica do portal, controvérsias são moduladas de modo que as tarefas de usabilidade vão além das regras básicas para também expressarem acessibilidade, por meio de outros atores, como a modelagem específica de dados, a ontologia e os recursos advindos das ferramentas semânticas, que rastreiam as informações nos portais por meio dos metadados, modificando as normas estabelecidas (CALLON; LATOUR, 1981; ROSENFELD; MORVILLE; ARANGO, 2015). A usabilidade e a acessibilidade traduzem os interesses dos usuários do portal semântico.

Os programadores ou engenheiros de *software* são responsáveis pela modulação e pela manutenção de motores de busca usados na recuperação de informações no portal (NIELSEN; LOHANGER, 2007). Os especialistas de domínio e os engenheiros do conhecimento estão envolvidos na gestão de ontologias e conectam-se ao vocabulário semântico compartilhado e aos mecanismos de edição e de captação de ontologia que possibilitam o compartilhamento automático e integrado de informações. Esses mecanismos permitem a extração, a interpretação e o processamento de informações, de modo que a publicação em um portal semântico pode ser efetuada modificando-se ou ampliando-se os conceitos do vocabulário semântico utilizado (LACHTIM; MOURA; CAVALCANTI, 2009). Os principais atores humanos e não humanos da rede de portais semânticos, levantados na literatura – seções 6.2.1 e 6.2.2 são listados no Quadro 18.

Quadro 18 - Atores da rede de portais semânticos

ADMINISTRADORES	DESENVOLVEDORES	TECNOLOGIAS	USUÁRIOS
Conteudistas Gerente de Conteúdo Gerente de Banco de Dados Gerente de Segurança e Rede	Programadores <i>Designer</i> Engenheiro do Conhecimento Engenheiro de Ontologia	Ontologia Modelo de RDF Metaontologia OWL Gerenciador de conteúdo Mecanismo de Inferência Tecnologias semânticas	Usuários diretos Comunidade acadêmica, professores, alunos, servidores da instituição Usuários indiretos Instituições educacionais, Instituições e organizações parceiras e sociedade em geral

Fonte: Autoria Própria (2021), fundamentada em Nielsen (2000), Bernes-Lee, Handler e Lassila (2001) e Iyamu e Mgudlwa, 2018.

Delineados os administradores e os desenvolvedores, na próxima Seção caracteriza-se as tecnologias e os usuários juntamente com os relacionamentos destes com os demais macroatores da rede de portais semânticos.

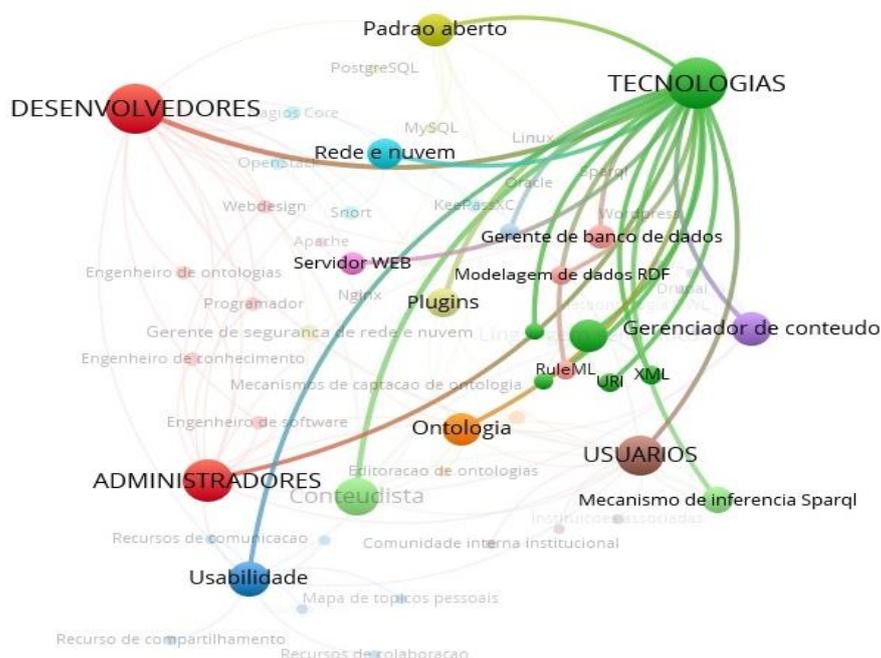
6.2.2 Atores não-humanos e usuários da rede de portais do conhecimento de universidades

Na rede de portais semânticos, as tecnologias provocam abertura de caixas-pretas, com a reordenação de atores que revelam novas possibilidades de arranjos técnicos (LAW, 1992; IYAMU; MGUDLWA, 2018). As tecnologias fornecem a estrutura técnica necessária para que o portal funcione como um aplicativo da Web. Em termos de recuperação da informação, os portais convencionais disponibilizam pesquisa sintática com resultados numerosos e pouca relevância, conforme já tratado no Capítulo 2. Este cenário de tecnologias estruturantes de portais constituiu uma caixa-preta da Tecnologia da Informação, pois o uso dessas ferramentas ficou estabilizado até o início do ano de 2000 com o advento da Web Semântica. A partir desse período as tecnologias semânticas rompem as barreiras da pesquisa sintática para disponibilizar busca semântica que permite o refinamento de documentos mais relevantes pelo uso de ontologias.

Da estrutura semântica dos portais, desvenda-se a Tecnologia com um macroator que se associa com diferentes microatores tais como o servidor Web, o

gerenciador de conteúdo, as linguagens de programação semântica, o RDF, o OWL. As tecnologias também estão diretamente associadas aos administradores, desenvolvedores e usuários, pois todas as ações emanadas por esses atores é que vão caracterizar o portal como um artefato sociotécnico semântico. Na Figura 23 foram mapeados os laços que ligam as tecnologias aos demais atores da rede, conforme levantado na literatura.

Figura 23 – Tecnologias na rede do portal semântico.



Fonte: Autoria própria (2021), com suporte da ferramenta Vosviewer.

As tecnologias utilizadas na estruturação do portal são definidas pelos administradores e desenvolvedores (NIELSEN, 2000; NIELSEN; LORANDER, 2007; (ROSENFELD; MORVILLE; ARANGO, 2015) e revelam a estrutura técnica necessária para que o portal ofereça pesquisa semântica, atendendo aos interesses dos usuários da informação. As tecnologias de código aberto (banco de dados, sistema operacional e servidor Web) usadas no portal semântico traduzem interesses de colaboração, de intercâmbio e o reuso de dados, a comunicabilidade, a publicização, a autoregulação, a auto-organização e democratizam o acesso às informações, atendendo preceitos da Web semântica (BERNERS-LEE; HENDLER; LASSILA, 2001; MAEDCHE; STAAB, 2001; NGUYEN *et al.*, 2020).

De acordo com Gavrilova (2011), as tecnologias semânticas possibilitam a estruturação de portais com características semânticas cujo objetivo é melhorar a recuperação da informação contida nos portais para os usuários. Essas são compostas pelas linguagens de programação semântica, *plugins*, mecanismos de inferência, gerenciador de conteúdo e ontologia.

O RDF e OWL são ferramentas essenciais que fornecem suporte para o intercâmbio das diversas tecnologias baseadas em regras para a estruturação semântica do portal (WORLD WIDE WEB - W3C, 2019). Estas tecnologias desvendaram a caixa-preta de estruturação dos portais convencionais, modificando a estruturação técnica desses artefatos (LATOUR, 2001; LAW, 1992; EL HAJI; AZMANI; EL HARZLI, 2014; LAWAN; RAKIB, 2019; STANESCU, 2018; WRIGLY, GARCIA-CASTRO, NIXON, 2012).

O modelo de dados Resource Description Framework (RDF) possibilita a representação descritiva dos recursos da Web que expressam significados (BIZER; HEATH; BERNERS-LEE, 2009). Esta modelagem constitui um avanço técnico dos portais convencionais e proporciona uma recuperação da informação mais eficiente (LAWAN; RAKIB, 2019; ROSENFELD; MORVILLE; ARANGO, 2015). O RDF liga-se com o RDF Schema, definindo o vocabulário para representação de ontologias simples e de descrição de classes e propriedades (KOIVUNEN; MILLER 2001).

A metaontologia OWL é também uma tecnologia importante para implementação do portal semântico. Possibilita, por meio da criação de ontologias, o processamento do conteúdo da informação e não somente a apresentação desta aos atores humanos, como acontece nos portais convencionais. (LAUSEN *et al.*, 2005; BERNERS-LEE; HENDLER; LASSILA, 2001). SWRL, RIF e RuleML configuram linguagens de regras semânticas que fornecem expressividade para inclusão eficiente nas metaontologias OWL (LAWAN; RAKIB, 2019).

A linguagem de consulta SPARQL, é um padrão fundamental dos portais semânticos, pois fornece recuperação rápida e precisa da informação por meio de motor de busca com adição de raciocínio lógico automatizado (BELLANDI *et al.*, 2012; BERNERS-LEE; HENDLER; LASSILA, 2001; DUARTE; HARA, 2018; KOIVUNEN; MILLER, 2001; MAATOUK 2021). Este ator liga-se com o banco de dados nativos RDF, OWL, traduzindo interesses de desenvolvedores, administradores e usuários.

Para Wimmer, Yoon e Rada (2013), a ontologia possui papel fundamental na estruturação semântica já que melhora a precisão da busca nos portais por fornecer

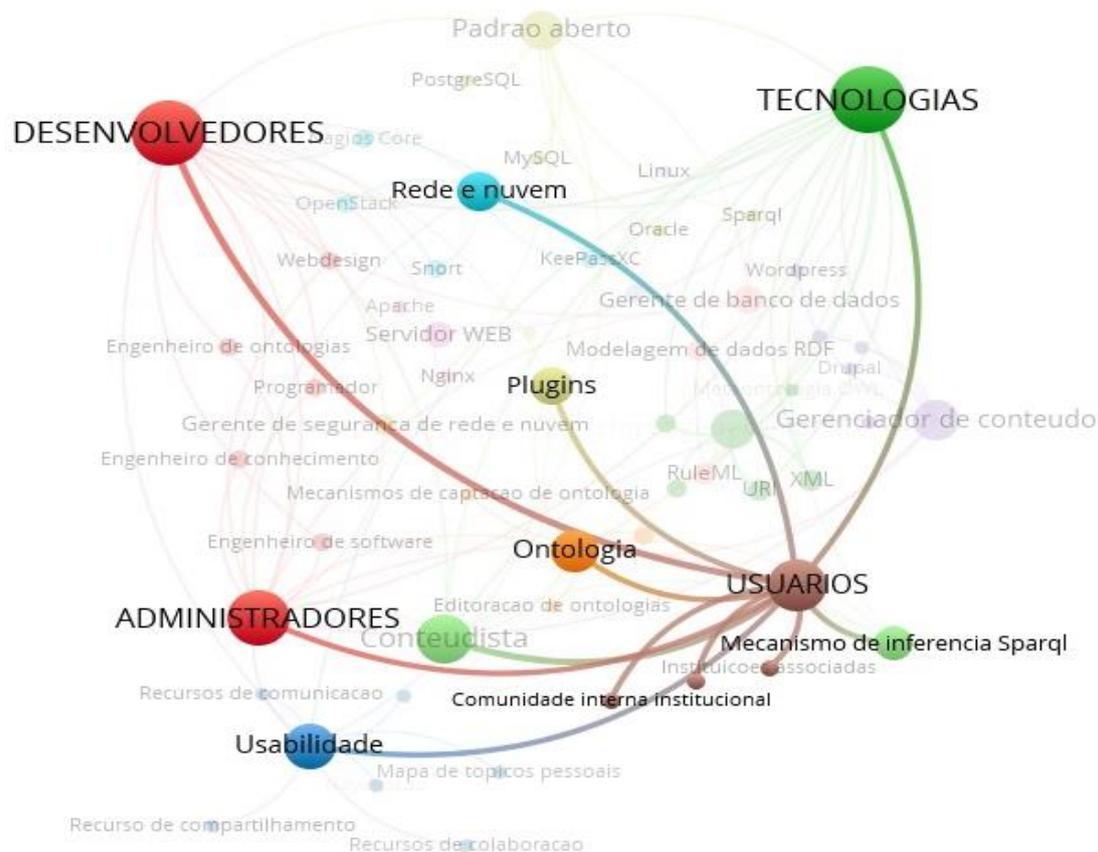
um vocabulário semântico que possibilita a interpretação do significado contextual de termos digitados nos campos de pesquisa. Essa tarefa, segundo Paiva, Costa e Silva (2014) é facilitada pelos mecanismos de busca RDF/SPARQL que entendem as relações hierárquicas dos termos e as relações complexas que ocorrem entre os termos definidos dentro de ontologias.

A ontologia de domínio mobiliza atores não-humanos representados pelas tecnologias semânticas para caracterizar o portal semântico, conectando-se com as tecnologias envolvendo banco de dados, banco de dados nativos de RDF, sistema operacional, servidor Web e mecanismos de inferência, alinhando-se aos interesses dos atores humanos representados por administradores, desenvolvedores e usuários (BIZER; HEATH, BERNERS-LEE, 2009; CALLON, 1986c; GUARINO, 1998; KOIVUNEN; MILLER, 2001). A ontologia se conecta ainda aos *plug-ins* e mecanismos que possibilitam a utilização de recursos de colaboração, de comunicação, de compartilhamento e de personalização, ampliando as funcionalidades do portal semântico (LAUSEN *et al.*, 2005; LACHTIM; MOURA; CAVALCANTI, 2009).

À ontologia conectam-se ainda, o vocabulário semântico compartilhado, os mecanismos de edição e captação de ontologia, a metaontologia OWL, as linguagens de regras semânticas, o gerenciador de conteúdo e os plugins. De acordo com Lachtim, Moura e Cavalcanti (2009), os mecanismos de edição e de captação possibilitam o compartilhamento automático e integrado de informações que permitem a extração, a interpretação e o processamento de informações, de modo que a publicação em um portal semântico pode ser efetuada modificando-se ou ampliando-se os conceitos do vocabulário semântico utilizado.

Os usuários constituem outro macroator do portal semântico, que se subdividem-se em usuários diretos e indiretos. Os usuários estão ligados a todos os demais atores da rede de portais, tendo seus interesses traduzidos pelos administradores, desenvolvedores e tecnologias semânticas. São usuários diretos aqueles que mantêm vínculo com a comunidade acadêmica de cada universidade, incluindo os docentes, os discentes e os servidores técnico-administrativos. Dentre os usuários indiretos podem figurar as instituições associadas, os patrocinadores e outros *stakeholders*/atores com necessidades comuns, a recuperação rápida e precisa da informação. A interconexão dos usuários na rede do portal é ilustrada na Figura 24.

Figura 24 – Usuários na rede do portal semântico.



Fonte: Autoria própria (2021), com suporte da ferramenta Vosviewer.

A semântica é o ponto de passagem dos usuários, pois fornece a qualidade na recuperação de informações. Assim, o vocabulário formal e controlado fornecido pela ontologia, a modelagem de dados RDF, o mecanismo de inferência SPARQL, a metaontologia OWL e as tecnologias que expressam regras semânticas traduzem os interesses dos usuários. A aliança desses atores na rede heterogênea composta por humanos e artefatos não humanos configuram os portais semânticos como artefatos sociotécnicos, socialmente construídos (CALLON; LATOUR, 1981).

Uma síntese das principais tecnologias utilizadas em portais semânticos é apresentada no Quadro 19.

Quadro 19 – Tecnologias para estruturação semântica de portais

TECNOLOGIAS	
Ontologia	→ Vocabulário formal - precisão nos resultados
Modelo de RDF	→ Estrutura <i>Triple Stores</i> / sujeito, objeto, predicado
Metaontologia OWL	→ Amplia o vocabulário RDF – fornece expressividade e inferência
Gerenciador de Conteúdo	→ Para criar, editar, gerenciar e publicar conteúdo
Mecanismo de Inferência	→ Motores de busca RDF/SPARQL para otimização das respostas de buscas
Tecnologias semânticas	→ Linguagens para expressar regras semânticas (SWRL, RIF, RuleML)
Tecnologias de Base	→ Banco de dados, Servidor Web, Segurança de Nuvem e Rede - Padrão aberto

Fonte: Autoria Própria (2021).

Apresentada a rede de atores, com os elementos humanos e as tecnologias e o detalhamento de funções e de responsabilidades – e os elementos não humanos, na próxima Seção apresenta-se a análise qualitativa dos dados coletados, a formulação e a validação dos indicadores e por fim, a apresentação do conjunto de indicadores qualitativos para avaliação do potencial semântico de portais do conhecimento de universidades.

6.3 ANÁLISE E VALIDAÇÃO DE DADOS

O questionário de pesquisa, após os ajustes derivados do pré-teste, foi encaminhado, eletronicamente, no dia 04 de junho de 2021 para os 69 coordenadores da área de TI, responsáveis pelos portais das universidades federais brasileiras. Para facilitar o preenchimento e o retorno dos dados, o questionário e o termo de autorização para participação na pesquisa foram enviados pela Plataforma D4sign. Até a data de 30/06/2021 foram obtidas 07 respostas, razão pela qual foram efetuados novos contatos por telefone com os coordenadores e procedeu-se o reenvio do formulário, por pelo menos mais duas vezes. Na data de 29/09/2021 foram obtidas 42 respostas, que perfizeram 60,86% em relação ao total da população estudada, atingindo o número de respostas mínimo calculado para uma amostra aleatória

simples, também acima do estimado para estudos qualitativos, de acordo com Marconi e Lakatos (2010), sendo este o viés principal de análise desta Tese.

Após a análise dos dados coletados, fez-se uma proposição de indicadores iniciais, posteriormente validada por um painel de especialistas formado pelos próprios respondentes dos questionários. Para validação dos dados foi encaminhado um segundo formulário para 22 universidades que responderam ao questionário, sendo obtidas 11 respostas dos coordenadores ou técnicos que atuam nos portais (Apêndice 3), perfazendo uma margem de erro de aproximadamente 9 % conforme indicado por Crespo Borges (2007), ainda dentro de padrões aceitáveis nos cálculos estatísticos

Na Seção seguinte inicia-se pela apresentação e discussão das respostas obtidas nos questionários e que embasaram a formulação dos indicadores qualitativos semânticos.

6.3.1 Dados coletados

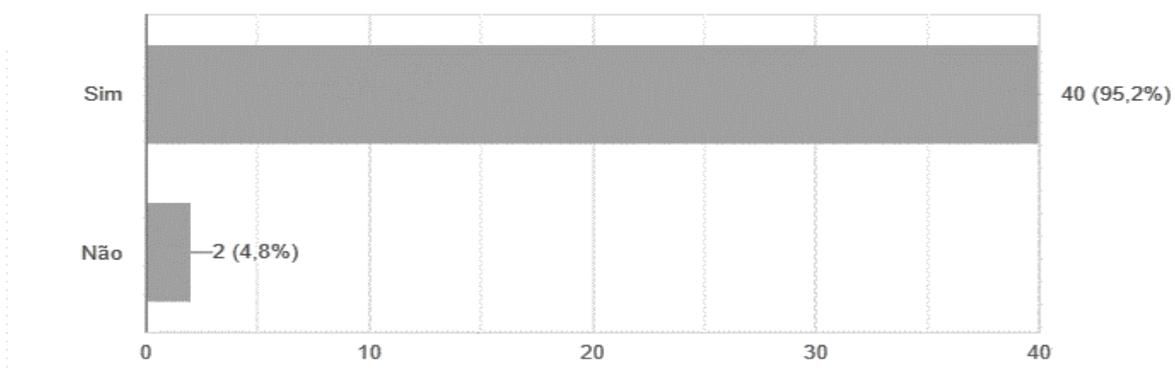
O questionário foi estruturado a partir de 4 dimensões concebidas pelo estudo de Guedes (2015), com questões norteadoras conforme o Quadro 16 do Protocolo de dados, incluindo: i) Armazenamento de dados e sua gestão; ii) Gestão da Informação; iii) Gestão de Ontologia e iv) Rede de Atores e caracterização sociotécnica. Desta forma, cada questão, dentre as 31 questões propostas, identificou a estruturação dos portais em termos de recursos humanos e de tecnologias. As subseções seguintes estão relacionadas às quatro dimensões utilizadas nesta pesquisa.

6.3.1.1 Armazenamento de dados e sua gestão

Nesta Seção apresenta-se as respostas para as questões 1 e de 3 a 5 do questionário. A primeira questão foi formulada com o intuito de verificar o uso de tecnologia de padrão aberto na estruturação do portal, condição que estava presente na maioria dos portais (95,2%), não sendo utilizada em apenas dois portais (4,8%) (Gráfico 5).

As tecnologias de padrão aberto são essenciais em portais semânticos já que não são controladas por corporações ou governos e podem ser utilizadas sem restrição (NGUYEN *et al.*, 2020; OPEN KNOWLEDGE INTERNACIONAL, 2019).

Gráfico 5 - Utilização de tecnologias de padrão aberto.

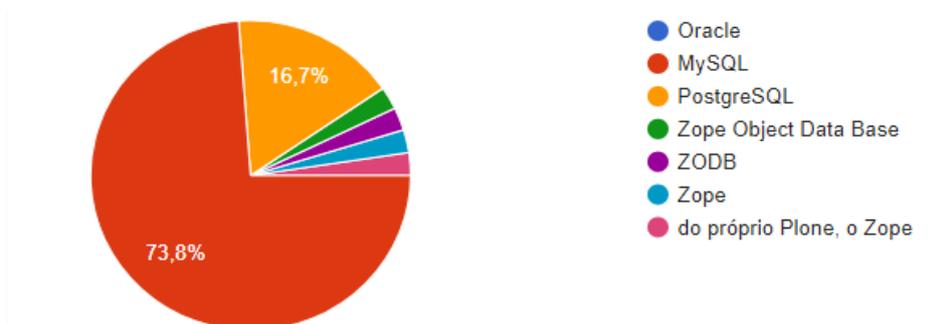


Fonte: Autoria própria (2021).

As especificações das tecnologias de código aberto estão disponíveis, sem custo, para uso de qualquer pessoa ou organização, além de ser possível adicionar extensões e facilitar a colaboração (interoperabilidade) e a comunicação no contexto do uso dessas tecnologias (NGUYEN *et al.*, 2020; OPEN KNOWLEDGE INTERNACIONAL, 2019).

Nas questões de número 3 a 5, procurou-se identificar as tecnologias de estruturação básica dos portais. Foi observado que o sistema operacional Linux foi utilizado em todos os portais. Para banco de dados, verificou-se uma predominância para o *software* MySQL (73,8%) seguido pelo PostgreSQL (16,7%) e Zope – e suas variações (9,6%) (Gráfico 6).

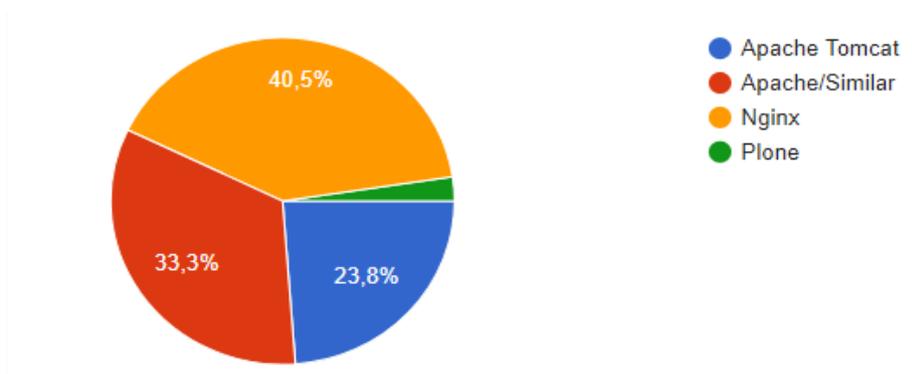
Gráfico 6 - Sistema de banco de dados utilizado



Fonte: Autoria própria (2021).

Para servidor Web o mais utilizado foi o Apache - e suas variações (57,1%), seguido pelo Nginx (40,5%) e Plone (2,4%) (Gráfico 7).

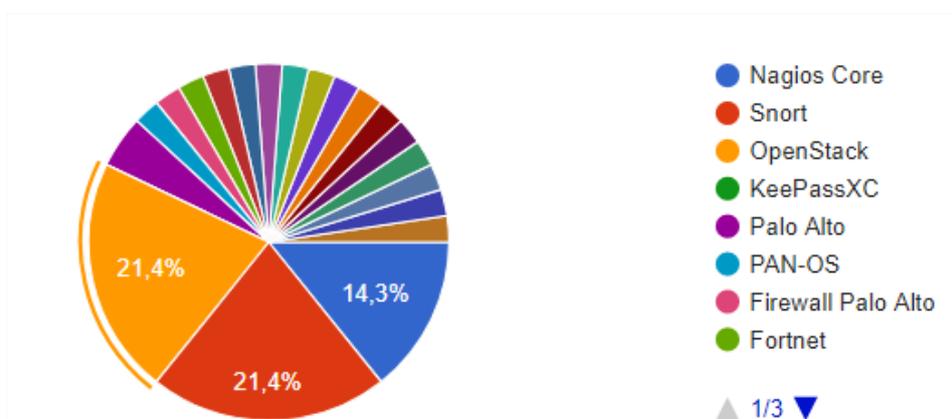
Gráfico 7 - Servidor Web utilizado.



Fonte: Autoria própria (2021).

Para segurança de rede e nuvem, os *softwares* mais citados foram o Snort e o OpenStack que tiveram o mesmo percentual de uso (21,4%), seguidos pelo Nagios Core (14,3 %) (Gráfico 8), marcando o uso de *software* de padrão aberto. Verificou-se o uso de *software* de segurança proprietário em 8 portais, sendo em um desses utilizado a combinação de proprietário e padrão aberto. Para outros dois portais não houve resposta direta, um deles indicando que a aplicação ainda está em desenvolvimento e outro respondeu que “não se aplica”.

Gráfico 8 - Sistema de segurança para nuvem e rede



Fonte: Autoria própria (2021).

O uso marcante de tecnologias de padrão aberto observado na estruturação de base dos portais é compatível com a estruturação semântica por favorecer a colaboração e a comunicação entre os atores do portal (REYNOLDS; SHABAJEE; CAYZER, 2004, LAUSEN *et al.*, 2005; CORREIA, 2012). Nesse quesito, as universidades cumprem as normativas da Web Semântica.

Considerando-se o Armazenamento de dados e sua gestão, verificou-se que os portais utilizam tecnologias básicas de padrão aberto para sistema de banco de dados, sistema servidor Web, sistema de segurança para nuvem e rede.

6.3.1.2 Gestão da Informação

A dimensão que trata da Gestão da Informação incluiu a questão de número 2 e de 6 a 15. Na questão número dois, o propósito foi a identificação do tipo de armazenamento e da navegação pelos dados nos portais (Gráfico 9). Todos os portais utilizam sistema de arquivo em servidor. Em 54,8% dos portais utiliza-se hierarquia fixa de classificação que desfavorece a inclusão de informações novas e não permite participação do usuário do portal (CORREIA, 2012; LAUSEN *et al.*, 2005; REYNOLDS; SHABAJEE; CAYZER, 2004). Em 33,3% dos casos utiliza-se arquivos com hierarquia de classificação flexível e o usuário pode incluir informações e anotações na estrutura do portal (CORREIA, 2012; LAUSEN *et al.*, 2005; REYNOLDS; SHABAJEE; CAYZER, 2004).

Gráfico 9 - Sistema usado para repositório de documentos e navegação pelos dados.

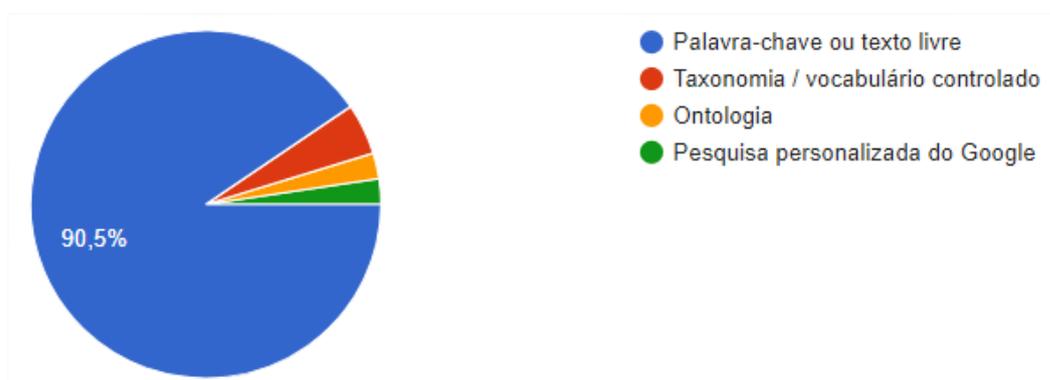


Fonte: Autoria própria (2021).

Para 11,9% dos portais foi identificado o uso de navegação multidimensional, com facetas, informações semiestruturadas e extensíveis. Esta estruturação está de acordo com os requisitos dos portais semânticos, já que o usuário pode navegar pelas facetas e incluir na busca novas classificações, estendendo a estrutura da informação, sendo possível agregar múltiplas visões sobre um mesmo dado (CORREIA, 2012; BÖRNER *et al.*, 2012; LAUSEN *et al.*, 2005; REYNOLDS; SHABAJEE; CAYZER, 2004). A navegação multidimensional, a partir de hierarquia das classificações associadas ao conteúdo pesquisado, possibilita ao usuário refinar a busca para obter resultados pertinentes. Considerando que a hierarquia fixa de classificação e a manutenção centralizada de documentos foi vista na maioria dos portais (54,8%), percebe-se que neste quesito, os portais universitários distanciam-se das prerrogativas semânticas de estruturação.

Na questão 6, foi possível verificar o sistema utilizado para pesquisa no portal. Na maioria dos portais (90,5%), foi utilizada a palavra-chave ou texto livre para recuperar informações no portal (Gráfico 10).

Gráfico 10 - Sistema utilizado para pesquisa no portal



Fonte: Autoria própria (2021).

Na pesquisa por palavra-chave são considerados apenas os aspectos sintáticos ou formais das palavras e não as relações entre os conceitos que estão associados aos documentos. Assim, este sistema é incompatível com a Web Semântica, pois não possui recurso para estabelecer as especificidades dos assuntos pesquisados, diminuindo a pertinência da pesquisa com recuperação de documentos irrelevantes para o usuário do portal (AKSOY, 2016; LI *et al.*, 2017; MALGAONKAR; DEVALE, 2016; RAZA *et al.*, 2018). Tais limitações podem ser minimizadas com o uso de ontologia formal de domínio na estruturação do portal que explora as relações

semânticas dos termos de busca e os documentos (BARBA-GONZÁLEZ *et al.*, 2019; GRUBER, 1996; GUARINO, 1997; GUO *et al.*, 2018; TELNOV, 2015). Deste modo, na pesquisa semântica os termos digitados na busca são convertidos automaticamente em linguagem de consulta estruturada, interpretando as palavras-chave com base em ontologias (LEE *et al.*, 2018).

Percebe-se, no gráfico 10 que o uso de taxonomia/vocabulário controlado foi indicado em 2 portais (5,6%). Entretanto, ao pesquisar-se direto nesses portais, para uma verificação prática, com um termo de busca que é composto por mais de uma palavra-chave (comitê de ética) verificou-se a marcação de palavras chaves no título dos documentos, mas em alguns casos sem relação direta com o termo, ou recuperando isoladamente cada termo, inclusive a preposição (de), o que gerou dúvidas sobre o efetivo uso de vocabulário controlado.

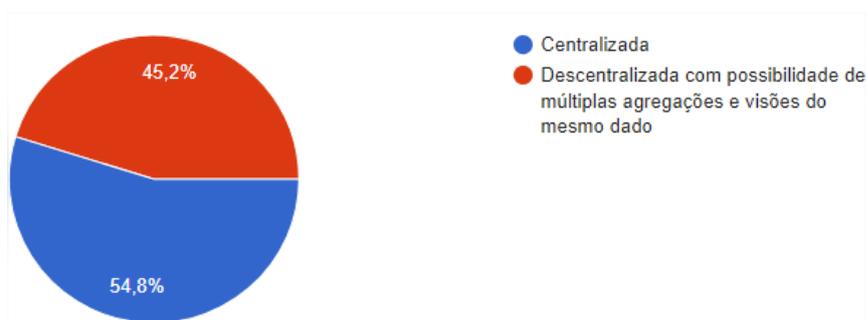
O uso de ontologia como sistema de recuperação de informações foi apontado em um portal (2,8%), mas ao pesquisar-se na interface desse portal verificou-se que é utilizado o mecanismo de pesquisa do Google como sistema de busca. O mecanismo inclui técnicas de rastreamento, indexação e ranqueamento e é usado para indexar as notícias publicadas no portal possibilitando uma recuperação melhorada de conteúdo, mas não foi possível identificar o uso de ontologia de domínio da área educacional. Diante disso, conclui-se que nenhum portal pesquisado usa a ontologia em sua estruturação.

Muito embora em apenas um portal (2,8%) tenha sido sinalizado o uso da pesquisa personalizada do Google acoplada, observou-se em 10 outros portais da amostra a utilização dessa mesma ferramenta. Contudo, de acordo com os preceitos semânticos, o uso da ontologia de domínio na estrutura dos portais do conhecimento de universidades, potencializa a qualidade dos resultados das buscas, pela expansão da consulta a partir de regras e de classificação hierárquica de assuntos que melhoram a recuperação de informações com resultados mais precisos (MALHOTRA; NAIR, 2015; NGO; CAO, 2018).

Na questão 7, ficou registrada a forma de gestão do conteúdo nos portais. A gestão centralizada de conteúdo foi observada na maioria dos casos (54,8 %), mas a forma descentralizada foi usada também em boa parcela dos casos (45,2,9%). (Gráfico 11). A gestão centralizada pressupõe a existência de um órgão ou setor que controla todo o processo de Gestão da Informação. Por outro lado, na gestão descentralizada permite-se que os dados disponíveis possam ser compartilhados e

manipulados por organizações provedoras e pelos usuários do portal, o que é compatível com estruturação semântica dos portais.

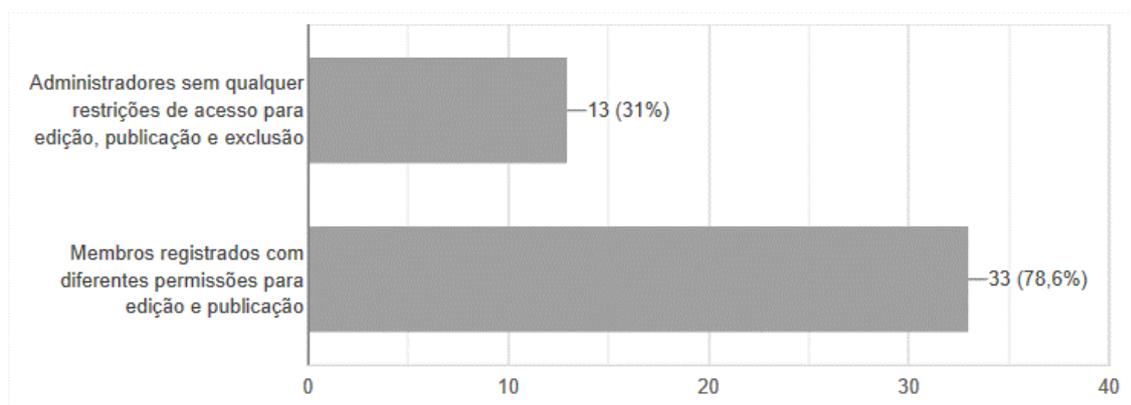
Gráfico 11 - Forma de gestão de conteúdo adotada.



Fonte: Autoria própria (2021).

A forma de edição de itens de informação foi investigada na questão 8 (Gráfico 12.). Na maioria dos portais (78,6%), observou-se que membros registrados com diferentes permissões podem editar e publicar conteúdos no portal. Já em oito portais (31%) a edição, a publicação e a exclusão de conteúdo são feitas por administradores sem quaisquer restrições de acesso. O registro de membros e suas permissões está de acordo com protocolos de segurança (NIELSEN, 2000), mas sobretudo, em um portal semântico, é possível extrair, interpretar e processar informações mediante a ampliação de conceitos do vocabulário semântico utilizado (LACHTIM; MOURA; CAVALCANTI, 2009), o que contribui para a precisão do resultado de busca.

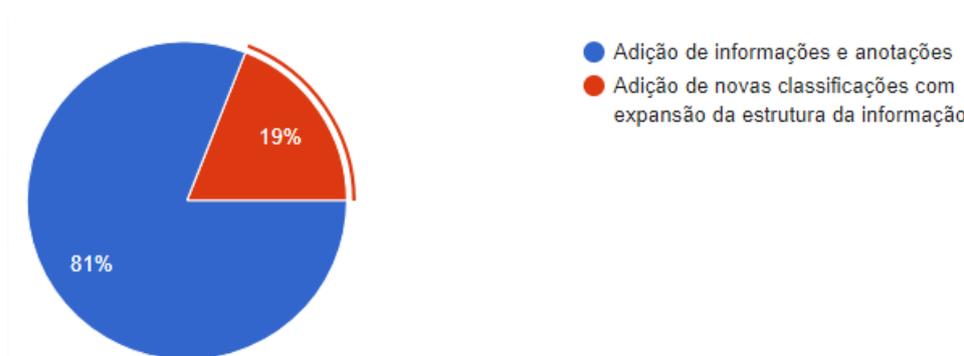
Gráfico 12 - Forma de edição de itens de informação.



Fonte: Autoria própria (2021).

Na questão 9, investigou-se a forma de atualização das informações, sendo que para 81,3% dos portais a atualização é feita a partir da adição de informações e anotações (Gráfico 13). Já em 19% dos portais as informações são atualizadas por meio da adição de novas classificações com expansão da estrutura da informação, sendo esta tarefa desejável em portais semânticos. Isto porque a inclusão de classes e subclasses com ênfase nas propriedades, nas facetas e nos axiomas possibilita mais qualidade de resultados da pesquisa (GRUBER, 1996; NOY; GUINNESS, 2005).

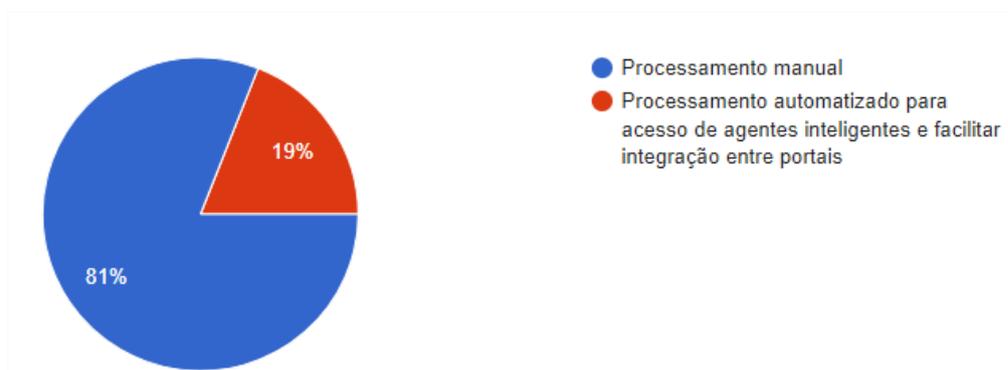
Gráfico 13 - Forma de atualização das informações.



Fonte: Autoria própria (2021).

Na questão 10, evidenciou-se a forma de processamento de informações e, para a maioria dos casos (81%), o processo é feito de forma manual se comparado com o processamento automatizado (19%), este desejável em portais semânticos por permitir o acesso de agentes inteligentes e facilitar a integração entre portais (CORREIA, 2012; JORGE, 2005; LAUSEN *et al.*, 2005; REYNOLDS; SHABAJEE; CAYZER, 2004) (Gráfico 14).

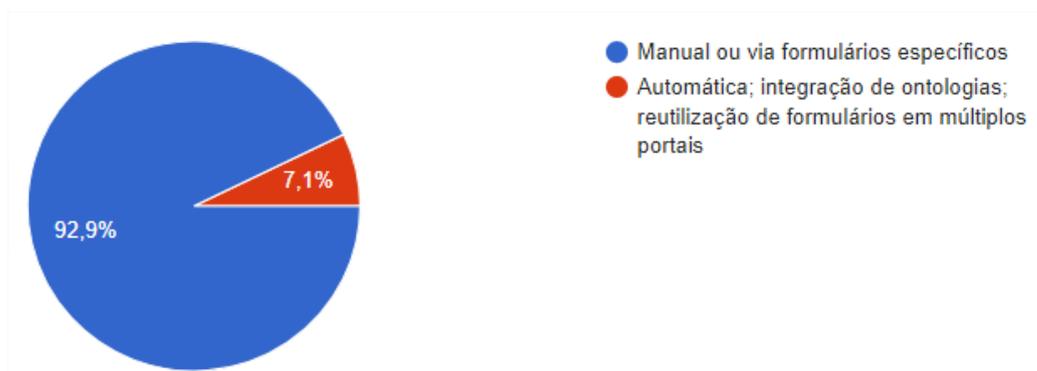
Gráfico 14 - Forma de processamento da informação



Fonte: Autoria própria (2021).

Quanto à forma de publicação de informações, investigada na questão 11, também a maioria dos portais (92,9%) adota publicação manual ou mediante formulário específico do portal. Apenas três portais (7,1%) adotam a publicação automatizada que possibilita a integração de ontologias e a reutilização de formulários em múltiplos portais semânticos (LAUSEN *et al.*, 2005; REYNOLDS; SHABAJEE; CAYZER, 2004) (Gráfico 15).

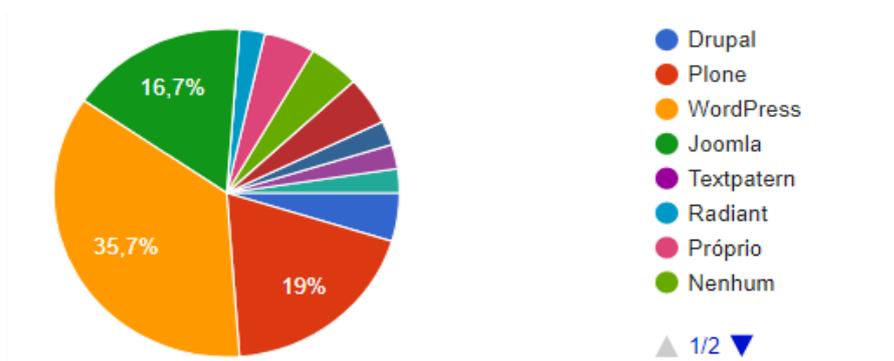
Gráfico 15 - Forma de publicação de informações



Fonte: Autoria própria (2021).

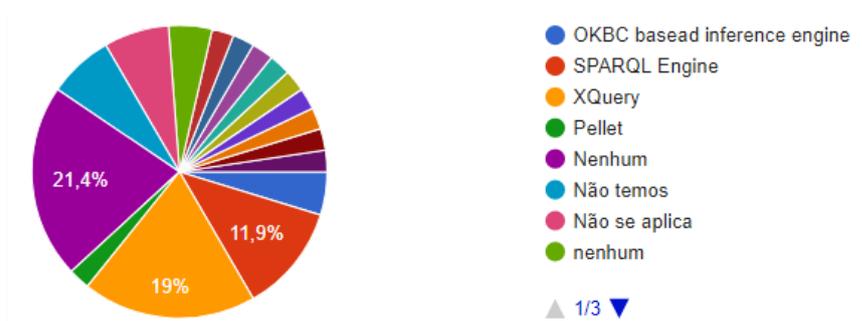
A predominância de processamento e de publicação manual de informações distancia os portais universitários da estruturação semântica, sendo necessário automatizar esses procedimentos.

Na questão 12 foi investigado o uso de sistema gerenciador de conteúdo, sendo observado que o WordPress figurou como o sistema mais utilizado (35,7%), seguido pelo Plone (19%) e Joomla (16,7%) (Gráfico 16). Em quatro portais utiliza-se sistema próprio e em cinco outros não é utilizado gerenciador de conteúdo. Observa-se que a maioria dos portais utiliza gerenciador de conteúdo de código aberto, compatíveis para uso em aplicações de Web Semântica e esta ferramenta possibilita automatizar a criação, a seleção e a categorização de dados, o que contribui para que resultados de pesquisa relevantes sejam apresentados ao usuário do portal (MASNER *et al.*, 2019; WORLD WIDE WEB - W3C, 2019). Além disso, o processamento automatizado de conteúdos favorece a manipulação de vários documentos por vários editores, reforçando o trabalho colaborativo (MASNER *et al.*, 2019; WORLD WIDE WEB - W3C, 2019).

Gráfico 16 - Gerenciador de conteúdo usado

Fonte: Autoria própria (2021).

O uso de mecanismo de inferência foi investigado na questão 13, sendo os mais citados XQuery (19%), SPARQL (11,9%) seguido pelo OKBC (4,8%); Drupal (2,4%) e Pellet (2,4%) (Gráfico 17). Porém observou-se que esta ferramenta não é usada em 59,5% dos portais, o que inviabiliza a pesquisa semântica. Por outro lado, o uso de mecanismo de inferência, especialmente o SPARQL, possibilita o raciocínio automatizado que aumenta a eficiência da busca nos portais semânticos (BELLANDI et al., 2012; BERNERS-LEE; HENDLER; LASSILA, 2001; KOIVUNEN; MILLER, 2001; MAATOUK, 2021).

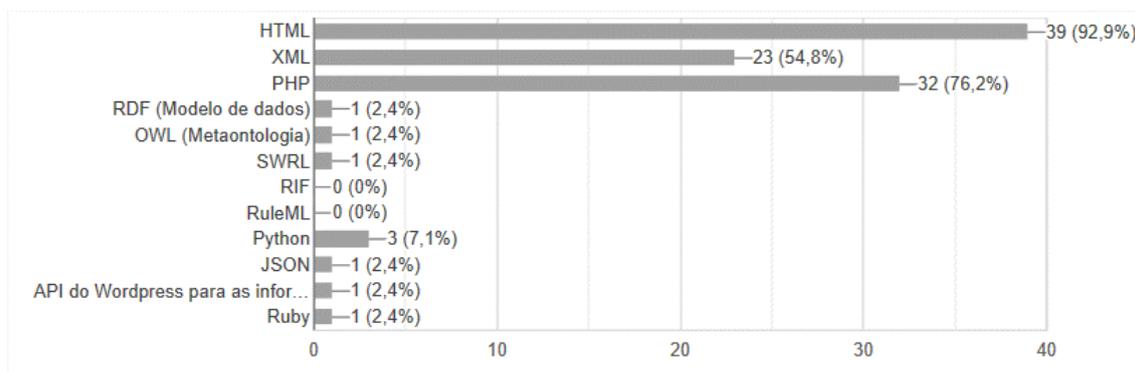
Gráfico 17 Mecanismo de inferência usado

Fonte: Autoria própria (2021).

A partir da questão 14, foi possível identificar as linguagens de programação e os modelos de dados usados nos portais. As três linguagens mais citadas foram HTML (91,7%), PHP (77,8%) e XML (55,6%) (Gráfico 18). Dessas linguagens, o XML figura no contexto da pesquisa semântica para que seja possível fazer inferências e relações semânticas entre os conteúdos (BIZER; HEATH; BERNERS-LEE, 2009) e, juntamente com o XML Schema, possibilita uma abordagem intuitiva e poderosa para descrição dos elementos e atributos de um documento XML, constituindo assim, a estrutura

padrão de representação que torna possível relacionamentos entre as linguagens semânticas (BIZER; HEATH; BERNERS-LEE, 2009; KOIVUNEN; MILLER, 2001; WORLD WIDE WEB – W3C 2019).

Gráfico 18 - Linguagens de programação e modelo de dados usados



Fonte: Autoria própria (2021).

Quanto à modelagem de dados, em apenas um portal foi sinalizado o uso do RDF (modelo de dados), juntamente com a OWL (Metaontologia) e a linguagem de programação SRWL. Estas tecnologias estão associadas ainda a outras linguagens de descrição lógica quais sejam RIF e RuleML para representar um modelo de dados para construção de hierarquias e de relacionamentos entre as coisas que refletem ontologias de domínio (LAWAN; RAKIB, 2019; TELNOV, 2015). Entretanto, nesse portal não foi sinalizado o uso de ontologia e de gerenciador de ontologia que são requisitos essenciais para, junto com o mecanismo de inferência e as linguagens semânticas, seja possível criar ou evoluir o portal para o contexto semântico. Assim, observa-se que nenhum portal utiliza linguagens de programação necessárias para evoluir os portais e torná-los semânticos.

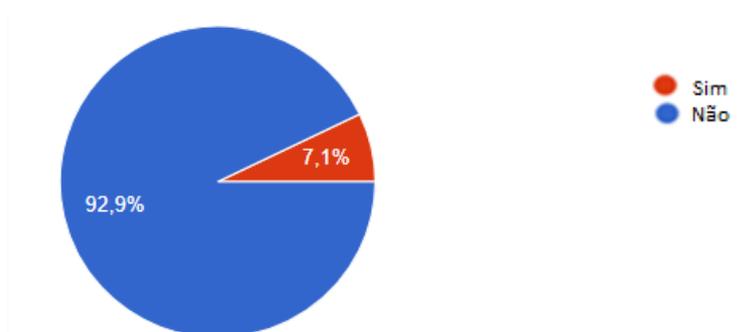
O resultado levantado pela questão 15, mostrou que 59,6% dos portais seguem a Lei Geral de Proteção de Dados Pessoais (LGPD) na disponibilização da informação no portal, o que denota a preocupação com proteção de direitos fundamentais de liberdade e de privacidade (BRASIL, 2018).

De maneira geral, no quesito Gestão da Informação, observou-se também o uso de gerenciador de conteúdo e de mecanismo de inferência de padrão aberto, mas o processamento e a publicação de informações ainda são manuais. Em nenhum portal usa-se ontologia e, tampouco, modelagem de dados RDF e metaontologia OWL, requisitos básicos para estruturação semântica.

6.3.1.3 Gestão de Ontologia

As questões 16 a 23 identificaram as características da dimensão Gestão de Ontologia. Na questão 16 investigou-se o uso de editor de ontologia. O uso dessa ferramenta foi sinalizado em três portais (7,1%), muito embora nesses portais não tenha sido apontado o uso de ontologia como sistema de busca ou linguagens semânticas o que justificaria o uso de editor para construção e reuso de ontologias disponíveis na Web (Gráfico 19).

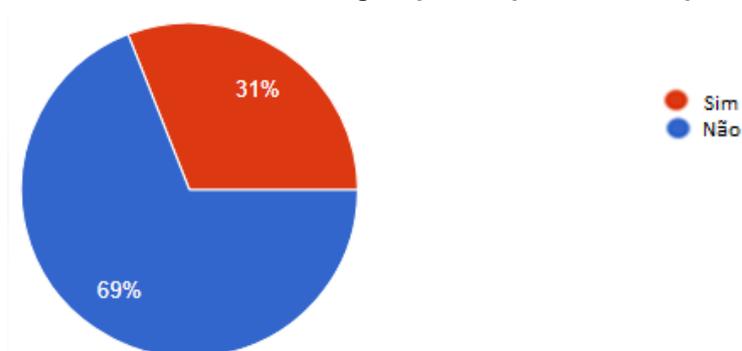
Gráfico 19 - Uso de editor de ontologia



Fonte: Autoria própria (2021).

Quanto aos serviços que proveem interoperabilidade e troca de informações entre portais semânticos, investigados nas questões 17 a 19, verificou-se que a maioria dos portais (69%) não utiliza tecnologias para importação e exportação de ontologias e reuso de informações (Gráfico 20).

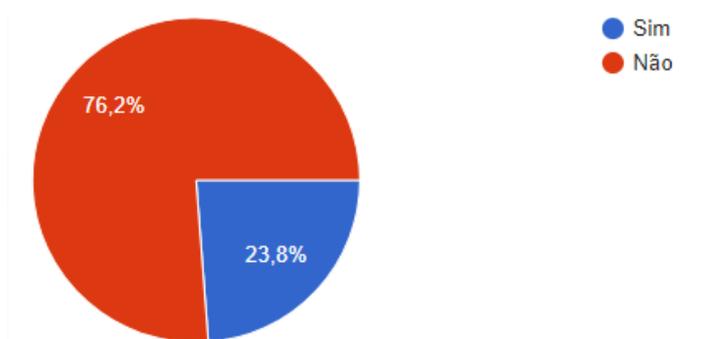
Gráfico 20 - Uso de tecnologias para importação e exportação.



Fonte: Autoria própria (2021).

De forma semelhante, a grande parte dos portais (76,2%) não possibilita a inclusão de membros para escrever conceitos, instâncias, atributos e colaborar na construção de ontologias (Gráfico 21).

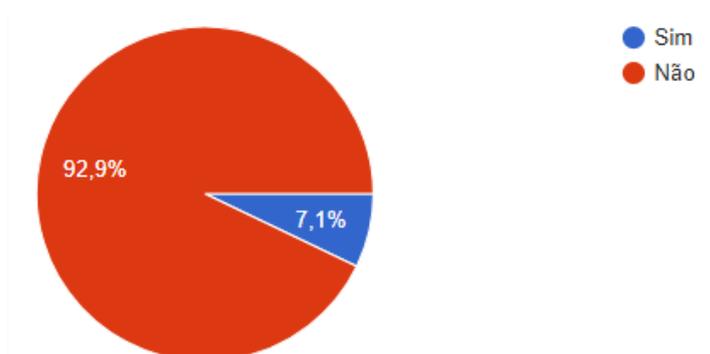
Gráfico 21 - Membros com permissões para escrever conceitos, instâncias, atributos e colaborar na construção de ontologias



Fonte: Autoria própria (2021).

Quanto à disponibilidade de materiais para consulta sobre a estrutura da ontologia e processamento semântico, este serviço também não estava disponível na maioria dos portais (92,9%) (Gráfico 22). Este panorama está congruente com a resposta negativa para uso de ontologia e linguagens semânticas na grande maioria dos portais.

Gráfico 22 - Processamento semântico ou estrutura da ontologia com materiais para consulta

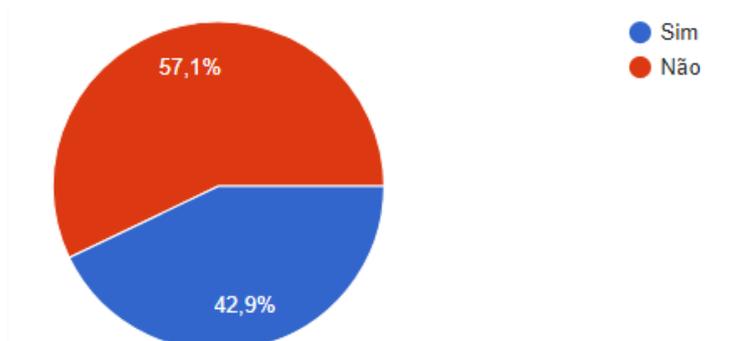


Fonte: Autoria própria (2021).

Sobre a completude e a consistência das informações, investigadas na questão 20, boa parte dos portais (42,9%) indicaram que estas condições estavam presentes,

embora a maioria dos portais tenha sinalizado a ausência dessa condição (Gráfico 23).

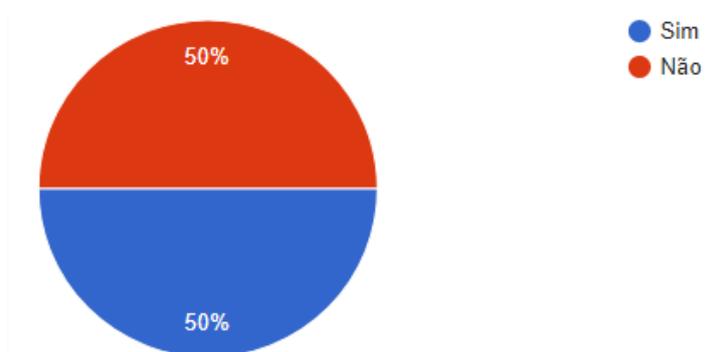
Gráfico 23 - Informação semanticamente completa e consistente



Fonte: Autoria própria (2021).

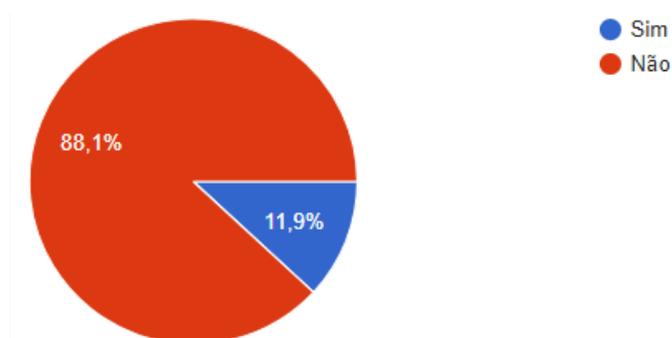
Já, quanto à profundidade de cobertura das informações investigada na questão 21, houve igual proporcionalidade entre respostas positivas e negativas (Gráfico 24). A qualidade no retorno das buscas em um portal semântico está diretamente associada à cobertura, à completude e à consistência das informações, o que é possível pela recuperação de dados da ontologia a partir do uso de mecanismo de raciocínio lógico e tecnologias semânticas (AZMANI; EL HARZLI, 2014; BÖRNER *et al.*, 2012; MAATOUK, 2021).

Gráfico 24 - Profundidade de cobertura das informações



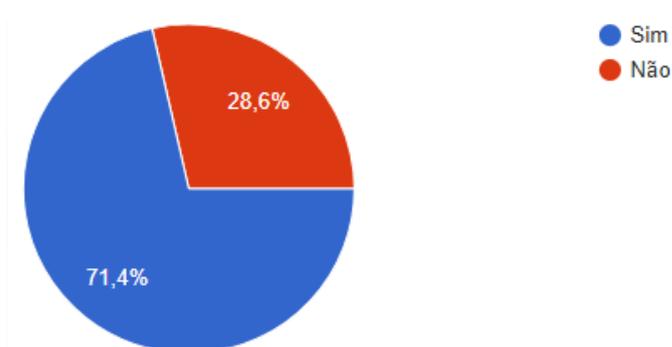
Fonte: Autoria própria (2021).

Na questão 22, investigou-se o uso de vocabulários para evitar termos ambíguos que possuem o mesmo significado, sendo observado que a grande maioria dos portais (88,1%) não faz uso ou controle de vocabulários o que pode tornar o resultado da busca menos relevante (Gráfico 25).

Gráfico 25 - Gráfico - Uso de vocabulários

Fonte: Autoria própria (2021).

Já quanto à maturidade de implantação do portal, observada pelas respostas da questão 23, na maioria dos portais (71,4%) esta condição foi sinalizada positivamente indicando a ausência de erros ou *link* quebrado, sendo o que o usuário da informação espera na apresentação destas (Gráfico 26).

Gráfico 26 - Maturidade de implantação/manutenção do portal

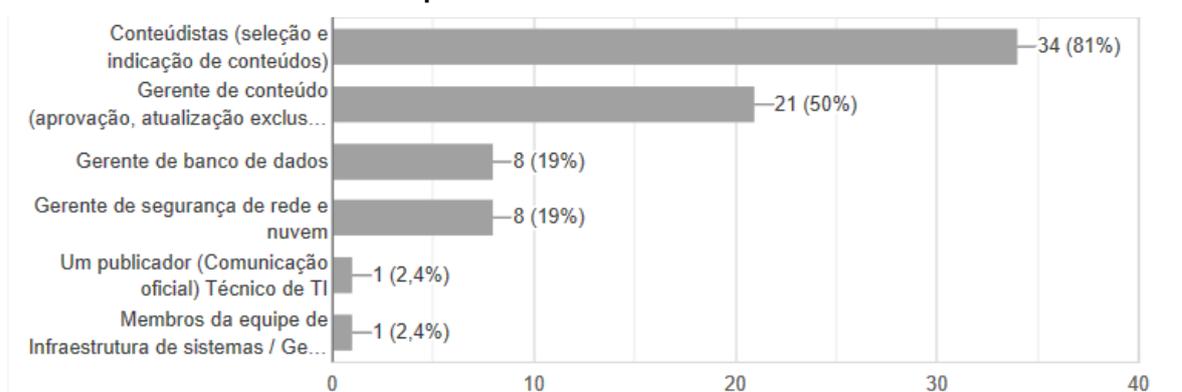
Fonte: Autoria própria (2021).

Os achados desta dimensão indicam que nenhum portal usa editor de ontologia e a maioria deles não utiliza tecnologias de importação e exportação, manuais de consulta de ontologia e vocabulários controlados, o que distancia estes portais das estruturas semânticas. Por outro lado, verificou-se razoável profundidade de cobertura de informações e ausência de erros nos portais.

6.3.1.4 Rede de Atores - Caracterização Sociotécnica

Dados associados à dimensão Rede de Atores – caracterização sociotécnica – foram coletados a partir das questões 24 a 31. A questão 24 apontou os integrantes da equipe de administradores dos portais, com o conteudista figurando em 81% dos portais, o gerente de conteúdo em 50%, o gerente de banco de dados 19% e o gerente de segurança de rede e nuvem também em 19%. Observa-se que o conteudista é o profissional que está mais presente na rede de portais com a tarefa de selecionar e indicar conteúdos, seguido pelo gerente de conteúdo que aprova e exclui conteúdos e, com menor participação figuram o gerente de banco de dados e o gerente de segurança de rede e nuvem. Adicionalmente, em dois portais, foram apontados a presença de publicador e de técnico de TI, associados ao setor de comunicação com participação de membros da equipe de infraestrutura/ desenvolvimento (Gráfico 27).

Gráfico 27 - Administradores do portal

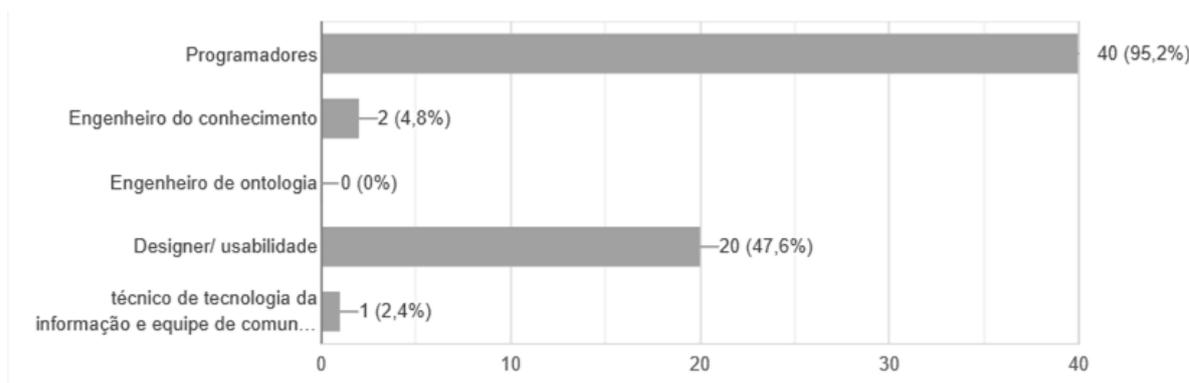


Fonte: Autoria própria (2021).

A questão 25 foi formulada para identificar os integrantes da equipe de desenvolvimento/manutenção do portal (Gráfico 28). Na maioria dos portais (97,2%) os programadores são responsáveis pelo desenvolvimento e manutenção dos portais. O *designer*, encarregado das tarefas de desenho e usabilidade dos portais esteve presente na metade dos portais analisados (50%). Por outro lado, em nenhum portal figuram os engenheiros do conhecimento e de ontologia, responsáveis pelo desempenho correto de ferramentas de raciocínio lógico e compartilhamento automático e integrado de informações (NIELSEN; LOHANGER, 2007; LACHTIM; MOURA; CAVALCANTI, 2009). Considerando a estrutura de pessoal das universidades públicas federais, a ausência de profissionais especializados pode estar

relacionada à insuficiência do quadro de servidores especializados para acompanhar exclusivamente as atividades relacionadas aos portais, sobretudo nas questões de segurança e organização de dados.

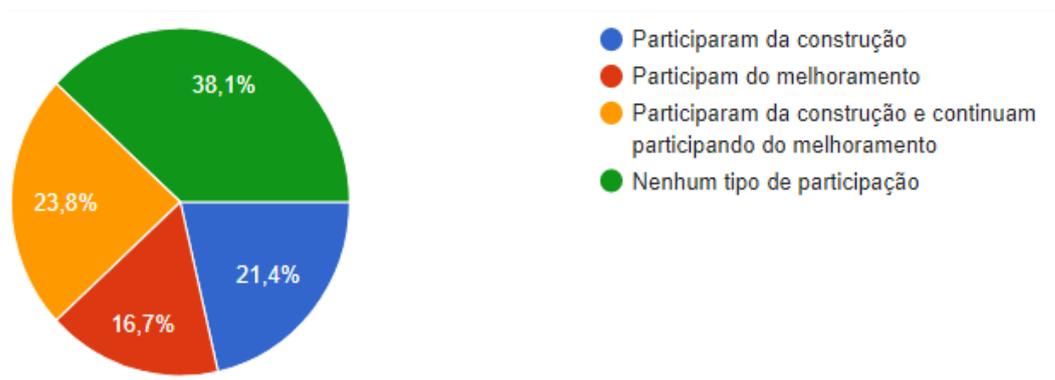
Gráfico 28 - Desenvolvedores do portal.



Fonte: Autoria própria (2021).

A participação dos usuários na construção ou melhoramento do portal foi investigada a partir da questão 26 (Gráfico 29). Para 38,1% não houve nenhum tipo de participação do usuário tanto na construção como no melhoramento dos portais. Por outro lado, em 61,9% dos portais houve algum tipo de participação do usuário o que denota a preocupação dos gestores quanto à avaliação do portal, já que este é projetado para atender as necessidades do usuário da comunidade acadêmica ou consumidor da informação de modo geral. Houve uma pequena diferença nos percentuais em que os usuários participaram apenas da construção dos portais (21,4%) comparado com aqueles que participaram da construção e continuam participando do melhoramento dos portais (23,8%). Uma parcela menor (16,7%) denota a participação apenas no melhoramento do portal.

Gráfico 29 - Participação dos usuários na construção/ melhoria do portal.



Fonte: Autoria própria (2021).

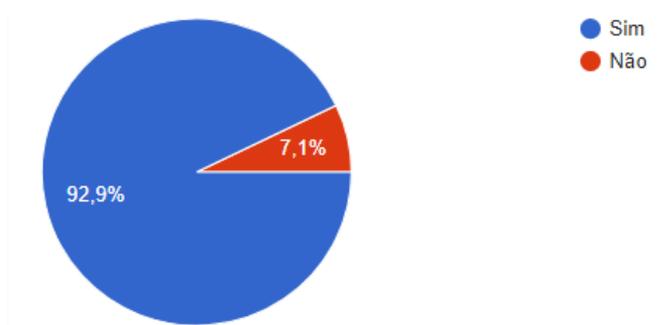
No contexto da tecnologia social, os artefatos técnicos, incluindo-se aqui os portais do conhecimento de universidades, devem ser desenvolvidos para facilitar ação humana nas suas diversas atividades e interesses (BIJKER; HUGH; PINCH, 1989; CALLON; LATOUR, 1981; HUGHES, 1989; IYAMU; MGUDLWA, 2018; LAW, 1992; LI *et al.*, 2017; MALGAONKAR; DEVALE, 2016; MALHOTRA; NAIR, 2015; MARTÍNEZ-GARCÍA; HERNÁNDEZ-LEMUS, 2013; PRADO; BARANAUSKAS, 2012; SHIM; SHIN, 2016) e, portanto, não se concebe o desenvolvimento de tecnologias, em todas as suas fases, quer seja do desenvolvimento à evolução, sem o contínuo diálogo com os usuários finais (LINSINGEN; BAZZO; PEREIRA, 2003; MARCUSE, 1998; VACCAREZZA, 2011). Sendo assim, a participação dos usuários no projeto dos portais semânticos, contribuindo também com melhorias, caracterizaria os portais como artefatos sociotécnicos, tornando-se importante ampliar a abertura de espaço e de mecanismos que favoreçam a comunicação constante dos usuários e os demais atores da rede de portais.

A baixa participação dos usuários de forma contínua, no desenho e na manutenção dos portais analisados, é uma condição que precisa ser revertida, afinal, é para os usuários que o portal é implantado, e portanto, suas contribuições são de fundamental importância para a evolução dos portais enquanto artefatos sociotécnicos.

Por meio da questão 27, observou-se que para a maioria dos respondentes (92,9%) os usuários dispõem de informações relevantes do domínio específico da comunidade acadêmica (Gráfico 30). Tais resultados mostram a preocupação dos técnicos de TI com o conteúdo disponibilizado, entretanto, denota-se a necessidade de também obter a avaliação dos usuários finais, considerando que uma parcela

significativa não teve nenhuma participação na construção ou melhoramento do portal de sua instituição.

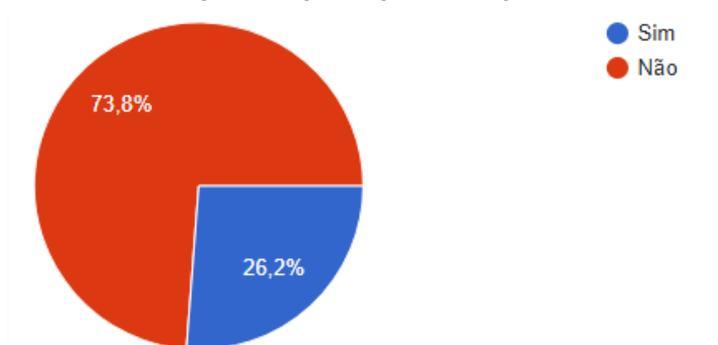
Gráfico 30 - Informações relevantes do domínio específico da comunidade acadêmica.



Fonte: Autoria própria (2021).

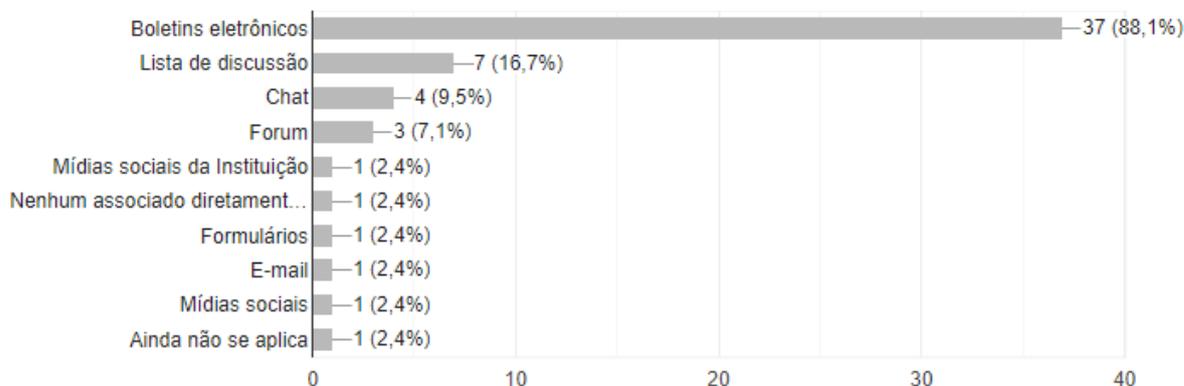
A partir das questões 28 a 30, investigou-se os recursos e os serviços ofertados para os usuários dos portais. Na maioria dos portais (73,8%) não é disponibilizado para o usuário mapas de tópicos pessoais para melhorar a gestão de informação pessoal e a eficácia do portal (Gráfico 31).

Gráfico 31 - Mapa de tópicos pessoais para usuários



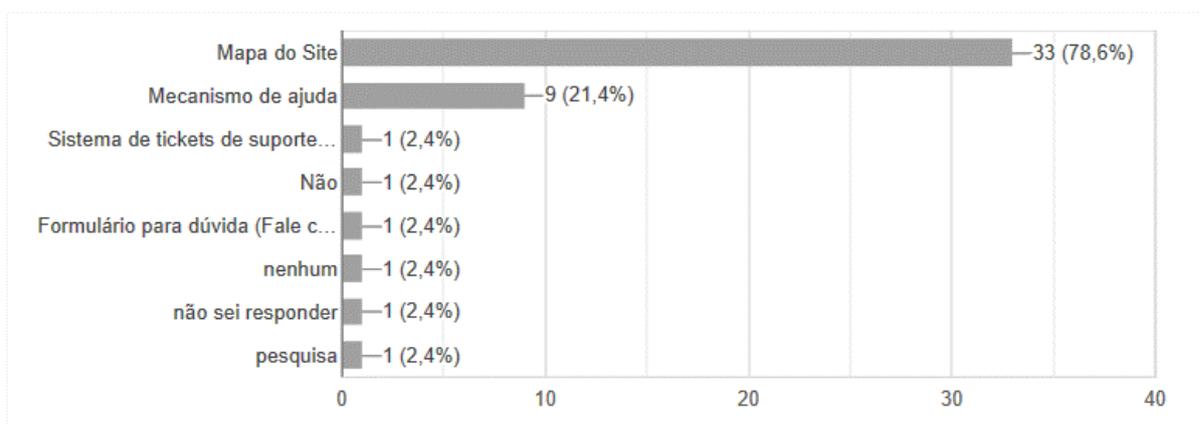
Fonte: Autoria própria (2021).

Quanto aos recursos para colaboração, comunicação e compartilhamento de informação, os usuários dispõem de boletim eletrônico na grande maioria dos portais (86,1%); lista de discussão, *chat* e fórum tiveram menor destaque (16,7%, 9,5% e 7,1%, respectivamente), figurando também o acesso a mídias sociais, formulários e *e-mail* da instituição, embora com pouco destaque (1,2% para cada modalidade). (Gráfico 32).

Gráfico 32 - Recurso de colaboração, comunicação e compartilhamento

Fonte: Autoria própria (2021).

Quanto a mecanismos de ajuda, foi sinalizada a disponibilidade de mapa do *site* na maioria dos portais (78,6%) e o mecanismo de ajuda em menor proporção (21,4 %) (Gráfico 33). Ainda foram apontados como serviços disponíveis, mas com baixa frequência, mecanismos de suporte/sugestões (2,4%) e formulário para Fale Conosco (2,4%).

Gráfico 33 - Mecanismos de ajuda

Fonte: Autoria própria (2021).

A disponibilização ao usuário de serviços personalizados é de suma importância no desenvolvimento de um sistema interativo, como os portais do conhecimento de universidades. No entanto, antecipar o interesse do usuário continua sendo um desafio a ser superado (AROUA; MOURAD, 2017) e, portanto, canais de comunicação variados e abertos são imprescindíveis. Assim, ressaltando novamente a finalidade primordial em se desenvolver aparatos técnicos socialmente construídos,

os portais semânticos devem disponibilizar serviços de informação diversificados para os usuários, a fim de atender as expectativas dos usuários forma abrangente.

Dos dados analisados para a Gestão de Dados, é marcante o uso de tecnologias de padrão aberto, em conformidade com os preceitos semânticos; na Gestão da Informação, notou-se a falta do uso de ontologias e de tecnologias semânticas, requisitos fundamentais para a pesquisa semântica; na Gestão de Ontologia, como a própria ontologia estava ausente, justifica-se o desuso do editor de ontologia, mas a disponibilização de poucos serviços personalizados prejudicou a comunicação e colaboração entre os usuários; e, por fim, na Rede de Atores ficou nítida a ausência de especialistas para administrar e desenvolver os portais e a baixa participação dos usuários.

Da revisão de literatura, perpassando pelas teorias que embasam este estudo e pela análise de dados coletados foram elencadas variáveis de interesse que seguem descritas na próxima Seção.

6.3.2 Proposta de Indicadores

Finalizada a análise dos dados correlacionados com a revisão da literatura que envolveu as teorias da Web Semântica, dos Sistemas Sociotécnicos e da Teoria Ator-Rede, apresenta-se na Tabela 5 as variáveis de interesse levantadas para este estudo, com o percentual de respostas levantado a partir do questionário.

As variáveis de interesse corresponderam às respostas do questionário que estão de acordo com a estruturação semântica associada e que também puderam ser relacionadas às condições sociotécnicas para que o portal possa ser entendido como um artefato social. Os percentuais correspondem a totalização de respostas para cada variável obtida na análise do questionário.

Tabela 5 - Variáveis de interesse

VARIÁVEIS DE INTERESSE	RESPOSTAS (%)
Uso de tecnologias de padrão aberto	95%
Banco de dados (Padrão aberto)	100%
Servidor Web (Padrão aberto)	100%
Sistema operacional (Padrão aberto)	100%
Sistema de segurança de rede e nuvem	90%
Arquivos em servidor com navegação multidimensional com facetas, informações semi estruturadas e extensíveis, alterações descentralizadas	12%
Ontologia usada como sistema de recuperação da informação	2%
Gestão descentralizada de conteúdos	45%
Membros registrados com diferentes permissões para edição e publicação	79%
Atualização da informação por adição de novas classificações com expansão da estrutura da informação	19%
Processamento automatizado da informação para acesso de agentes inteligentes e facilitar a integração entre portais	19%
Publicação automatizada de informações que possibilita a integração de ontologias e reutilização de formulários em múltiplos portais	7%
Uso de gerenciador de conteúdo	95%
Uso de mecanismo de inferência	41%
Uso de linguagens de programação semântica, modelagem de dados RDF e metaontologia OWL	2%
Uso de editor de ontologia	7%
Uso de tecnologia de importação e exportação para colaboração e reuso de informações	3%
Apresenta descrição de processamento semântico ou estrutura da ontologia com materiais para consulta	7%
Possibilita a inclusão de membros com permissões para escrever conceitos, instâncias, atributos e colaborar na construção de ontologias	24%
Apresenta informação semanticamente completa e consistente	43%
Apresenta profundidade de cobertura das informações	50%
Há controle no uso de vocabulários, evitando-se termos ambíguos que possuem o mesmo significado?	12%
Maturidade de implantação/manutenção do portal, ou seja, sem erros ou <i>link</i> quebrado?	71%
Conteudistas (seleção e indicação de conteúdo)	81%
Gerente de conteúdo (aprovação, atualização exclusão de conteúdo)	50%
Gerente de banco de dados	19%
Gerente de segurança de rede e nuvem	19%
Programadores	95%
Engenheiro do conhecimento	48%
Engenheiro de ontologia	0%
Designer	48%
Usuários participaram da construção e continuam participando do melhoramento do portal	24%
Usuários dispõem de informações relevantes do domínio específico da comunidade acadêmica	93%

(Continua)

VARIÁVEIS DE INTERESSE	RESPOSTAS (%)
Usuários dispõem de mapas de tópicos pessoais para melhorar a gestão de informação pessoal e eficácia do portal	26%
Usuários dispõem de boletins eletrônicos	37%
Usuários dispõem de lista de discussão	17%
Usuário dispõem de chat	10%
Usuário dispõem de fórum	7%
Usuário dispõem de mecanismo de ajuda	21%
Usuário dispõem de mapa do site	79%

Fonte: Autoria própria (2021).

Foram elencadas 40 variáveis de interesse que estavam associadas às condições estruturantes de semântica nos portais e que embasaram o conteúdo para determinação dos indicadores de potencial semântico. Algumas variáveis foram aglutinadas em um único indicador por tratar-se de assuntos afins ou complementares.

A partir do exposto, apresenta-se, no Quadro 20, o conjunto de indicadores qualitativos formulados para estabelecer o potencial semântico dos portais do conhecimento de universidades. Este conjunto é o indicado, relevando os principais elementos tais como o uso da Ontologia, linguagens de programação semântica e o motor de busca semântico (DUARTE; HARA, 2018; LAWAN; RAKIB, 2019; TELNOV, 2015). Os indicadores podem ser usados para evolução dos portais tornando-os semânticos e também para orientar a construção de novos portais semânticos.

Quadro 20 – Conjunto de indicadores qualitativos de potencial semântico.



Fonte: Autoria Própria (2021).

Considerou-se que o rol de indicadores qualitativos e seu uso como instrumento de gestão requer, no entanto, validação para que estes possam ser utilizados na

prática diária. Para atender esta diretiva, os indicadores iniciais propostos neste estudo, foram submetidos à validação de um conjunto de especialistas, conforme descrito na Seção seguinte.

6.3.3 Validação dos indicadores propostos

Com base na revisão de literatura combinada com a análise das variáveis de interesse foram determinados, portanto, 18 indicadores qualitativos para avaliação de potencial semântico de portais do conhecimento de universidades. Os indicadores foram validados por um conjunto de especialistas, composto pelos técnicos de TI que responderam ao questionário de pesquisa, a fim de coletar a percepção destes sobre os indicadores propostos. Obteve-se 11 respostas, todas com registro de aceitação acima de 89 % para cada um dos indicadores, embasando a proposição destes como ferramenta para gestão dos portais (Tabela 6), lembrando-se que se estipulou como válido um percentual igual ou superior a 75% (DONABEDIAN, 1985).

Tabela 6 - Validação dos dados.

INDICADORES PROPOSTOS	VALIDAÇÃO (%)
Uso de tecnologias de padrão aberto	91%
Armazenamento de documentos em servidor, com informações organizadas de forma semiestruturada e extensível, com possibilidade alterações descentralizadas	100%
Uso de ontologia como sistema de recuperação da informação	91%
Navegação por facetas	91%
Gestão descentralizada de conteúdos	91%
Membros registrados com diferentes permissões para manipulação de dados e informações	100%
Atualização de informações por adição de novas classificações e expansão da estrutura da informação	100%
Processamento e publicação automática de informações	100%
Uso de gerenciador de conteúdo	100%
Uso de mecanismo de inferência	100%
Uso de linguagens de programação semântica, modelagem de dados RDF e metaontologia OWL	91%
Uso de editor de ontologia	89%
Disponibilização de materiais sobre descrição do processamento semântico ou estrutura da ontologia usada	89%
Disponibilização de recursos para importação, exportação e colaboração para reuso de informações	89%
Cobertura, precisão, completude e consistência das informações	100%
Equipe especializada para gestão do portal	100%
Participação contínua dos usuários na construção e melhoramento do portal	89%
Disponibilização de recursos e serviços de suporte de comunicação e compartilhamento	100%

Fonte: Autoria própria (2021).

Analisando as variáveis de interesse frente a validação dos especialistas, verifica-se que mesmo aquelas que receberam menores percentuais de resposta no questionário da pesquisa foram todas validadas pelos especialistas (Tabela 7), destacando-se o uso de linguagens de programação semântica e uso de ontologia como sistema de recuperação da informação, ambos com 2% de respostas iniciais, e a disponibilização de recursos para importação, exportação e colaboração para reuso de informações, com 3% de respostas no questionário.

Tabela 7 - Variáveis de interesse e validação de indicadores.

INDICADORES	% VARIÁVEIS	% VALIDAÇÃO
Uso de tecnologias de padrão aberto	95%	91%
Armazenamento de documentos em servidor, com informações organizadas de forma semiestruturada e extensível, com possibilidade alterações descentralizadas	12%	100%
Uso de ontologia como sistema de recuperação da informação	2%	91%
Navegação por facetas	12%	91%
Gestão descentralizada de conteúdos	45%	91%
Membros registrados com diferentes permissões para manipulação de dados e informações	79%	100%
Atualização de informações por adição de novas classificações e expansão da estrutura da informação	19%	100%
Processamento e publicação automática de informações	13%	100%
Uso de gerenciador de conteúdo	95%	100%
Uso de mecanismo de inferência	41%	100%
Uso de linguagens de programação semântica, modelagem de dados RDF e metaontologia OWL	2%	91%
Uso de editor de ontologia	7%	89%
Disponibilização de materiais sobre descrição do processamento semântico ou estrutura da ontologia usada	7%	89%
Disponibilização de recursos para importação, exportação e colaboração para reuso de informações	3%	89%
Cobertura, precisão, completude e consistência das informações	64%	100%
Equipe especializada para gestão do portal	51%	100%
Participação contínua dos usuários na construção e melhoramento do portal	24%	89%
Disponibilização de recursos e serviços de suporte de comunicação e compartilhamento	28%	100%

Fonte: Autoria própria (2021).

Verificou-se que na maioria dos portais (95%) usa-se tecnologias de base de padrão aberto (sistema operacional, sistema de banco de dados, servidor Web e

sistema de segurança) para estruturação do portal, sendo este tipo de tecnologia validada por 91% dos especialistas.

A variável de interesse “Arquivos em servidor com navegação multidimensional com facetas, informações semiestruturadas e extensíveis, alterações descentralizadas” foi condição presente em apenas 12% dos portais. No entanto, para o painel de especialistas, esta modalidade de organização da informação foi validada com 100% de aceitação e a “Navegação por facetas” da ontologia recebeu 91% de validação.

A ontologia usada como sistema de recuperação da informação foi sinalizada apenas em 2% dos portais, mas foi validada pelos especialistas com 91% de aceitação. Observou-se que a despeito da prática diária, houve uma convergência de opiniões favoráveis dos especialistas frente a literatura que enfatiza a importância do uso da ontologia na estruturação dos portais para recuperação semântica da informação.

A variável “Gestão descentralizada de conteúdos” esteve presente em 45% dos portais e foi validada por 91% dos especialistas, confirmando-se, à luz da literatura, que o conteúdo armazenado e manipulado por uma rede descentralizada de organizações provedoras e indivíduos pode tornar possível múltiplas agregações e visões de um mesmo dado. A presença de “Membros registrados com diferentes permissões para edição e publicação” foi relatada por 79% dos portais e validada em 100% para o controle destas tarefas. A “Atualização de informações por adição de novas classificações e expansão da estrutura da informação” foi relatada por apenas 19% dos portais e validada em 100%, denotando que esta modalidade de atualização pode facilitar o acesso a novas informações disponibilizadas.

O processamento automatizado da informação foi descrito como processo adotado em 19% dos portais; já a publicação automatizada de informações foi usada em 7% dos portais. Pelo painel de especialistas ambas as condições, no entanto, foram 100% validadas já que estas favorecem o acesso de agentes inteligentes, facilitam a integração entre portais e possibilitam a integração de ontologias e a reutilização de formulários para publicação de informações em múltiplos portais.

O uso de gerenciador de conteúdo foi verificado em 95% dos portais e validado por 100% dos especialistas para automatizar tarefas de criação, seleção, filtro e categorização dos dados armazenados. Já o uso de mecanismo de inferência foi relatado por 41% dos portais e validado também por 100% dos especialistas, o que

reforça a fundamental importância deste mecanismo para otimização das respostas de busca, mediante navegação interativa por facetas, com visualização das relações existentes entre as informações traduzindo-as para o usuário.

O uso de linguagens de programação semântica, modelagem de dados RDF e metaontologia OWL foi relatado em apenas 2% dos portais, porém foi validado por 91% dos especialistas, confirmando-se que estas tecnologias são essenciais para o funcionamento do portal semântico.

Nos portais apenas 7% utiliza editor de ontologia, mas esse parâmetro foi validado por 89% dos especialistas. Por se tratar de programa para edição de ontologia ele só terá função se a ontologia fizer parte da estrutura técnica do portal. Esse programa se inter-relaciona com o engenheiro de ontologia para atualizar, incluir e organizar as ontologias de domínio acadêmico para que estas possibilitem a pesquisa semântica com resultados de busca que atendam os interesses dos usuários.

As tecnologias de importação e exportação estavam presentes em 3% dos portais, mas esses recursos foram validados por 89% dos especialistas e, assim, as funções de gestão da ontologia e interesses do engenheiro de ontologia, contribuem para que a recuperação das informações disponibilizadas possa ocorrer de forma mais eficiente, com resultados mais relevantes para os usuários. De forma semelhante, a descrição do processamento semântico ou estrutura da ontologia foi validada por 89% dos especialistas, muito embora tenha sido indicado em 7% dos portais a disponibilização de manuais sobre a estruturação da ontologia. A forma como se dá o processamento semântico se traduz em regras e relacionamentos entre as facetas da ontologia, sendo importante que tanto o engenheiro como outros atores possam trabalhar de forma colaborativa para contribuir com a evolução da ontologia usada no portal.

Ainda, sobre manipulação de ontologias, 24% dos portais registraram a inclusão de membros com permissões para escrever conceitos, instâncias, atributos e colaborar na construção de ontologias e 89% registraram o controle de permissões para edição e publicação de informações. O painel de especialista validou em 100% o acesso de membros registrados com diferentes permissões para manipulação de dados e informações. O acesso e o controle de tarefas no portal são, para além das ações de segurança de dados e informações gerenciadas pelo engenheiro de segurança e de rede, uma forma de promover a colaboração entre os atores da rede

de portais que se dá mediante o alinhamento de interesses quer sejam dos atores humanos - usuários e membros da equipe técnica e as tecnologias - os atores não-humanos, estruturados por seus conjuntos de regras para exercerem ações planejadas.

Em 43% dos portais a informação disponibilizada no portal é semanticamente completa e consistente; em 50% dos portais a informação tem profundidade de cobertura; em 71% dos portais há maturidade de implantação, sem erros; e em (93%) são disponibilizadas informações relevantes do domínio da comunidade acadêmica. Os especialistas validaram essas condições, com 100% de aceitação para a cobertura, a precisão, a completude e a consistência das informações disponibilizadas. Tais condições são essenciais para os usuários, já que estão diretamente relacionadas à qualidade e à relevância das informações disponibilizadas.

Quanto à equipe de administradores e de desenvolvedores, tiveram mais destaque os conteudistas, os programadores, o engenheiro do conhecimento e o *designer* com percentuais de 81%, 95%, 48% e 48%, respectivamente. Entretanto, os especialistas validaram em 100% que o portal mantenha uma equipe especializada incluindo também o gerente de conteúdo, o gerente de banco de dados, o gerente de segurança e nuvem e o engenheiro de ontologia.

A participação do usuário tanto na construção como no melhoramento do portal foi sinalizada por 24% dos respondentes, já os especialistas validaram em 89% a participação contínua dos usuários. A participação ativa dos usuários desde a criação até a manutenção e atualização do portal constitui uma ação fundamental que envolve colaboração, uma das prerrogativas da Web Semântica, propiciando interatividade entre administradores, desenvolvedores e usuários para tornar o portal dinâmico e eficiente.

As facilidades para os usuários do portal, incluindo a disponibilização de mapas de tópicos pessoas (26%), boletins eletrônicos (37%), lista de discussão (17%), *chat* (10%), mecanismo de ajuda (21%), e mapa do *site* (79%) foram citadas como funcionalidades presentes na interface dos portais. Apesar dos diferentes percentuais dessas funcionalidades, todos esses recursos e serviços foram validados em 100% pelos especialistas, o que denota a preocupação em disponibilizar um portal dinâmico e colaborativo, com facilidades de comunicação e colaboração.

Do arcabouço teórico aos dados empíricos validados observou-se que os portais do conhecimento de universidades disponibilizam informações associadas à

prática de suas ações institucionais. Das características elencadas a partir do levantamento da estruturação sociotécnica, então, adaptou-se a ideia dos laços fortes e fracos da rede de Granovetter (1983), para analisar como se estabelecem tais laços em relação às variáveis levantadas. Todos os laços, sejam fortes ou fracos, traçam a visão geral da estruturação da rede, salientando que laços fracos em uma rede, aparentemente um elemento negativo, podem ser vistos como oportunidade de inovação por envolverem vários grupos, possibilitando a troca de informações e de conhecimento. Os laços fortes da rede são importantes por estes trazem confiança, influência e predisposição para trabalho em rede, embora sejam limitados ao mesmo círculo social (GRANOVETTER,1983). Estabeleceu-se parâmetros para a identificação de laços fortes e fracos – Quadro 21.

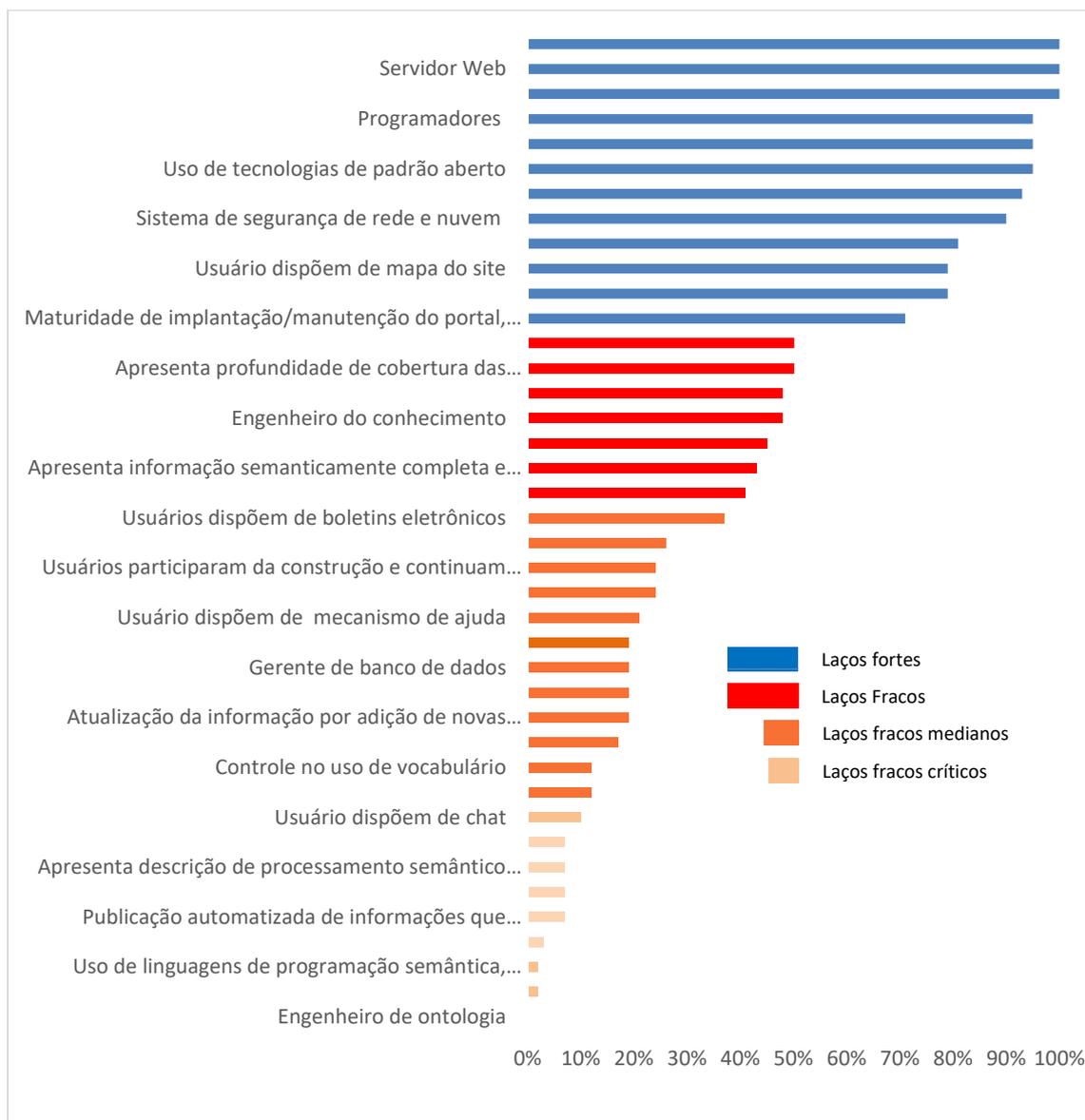
Quadro 21 - Parâmetros para laços fortes e fracos

LAÇO	PARAMETRO
Laços fortes	Acima de 70%
Laços Fracos	Abaixo de 70%
Laços fracos medianos	40% a 50%
Laços muito fracos	0 a 37%
Laços criticamente fracos	Abaixo de 10%

Fonte: Autoria Própria (2021).

Apresenta-se no Gráfico 35 estas relações.

Gráfico 34 - Laços fortes e fracos da rede de portais de universidades.



Fonte: Autoria Própria (2021).

A validação de todas as variáveis pelo painel de especialistas, independente da presença marcante ou mais discreta das características elencadas, confirma a importância dos laços fortes e fracos da rede de portais. Os laços fortes ficaram evidenciados pelas variáveis que apresentaram percentuais acima de 70% e denotam o alto nível de credibilidade na adoção de 12 características da estruturação dos portais a saber: (i) maturidade de implantação/manutenção do portal, sem erros; (ii) membros registrados com diferentes permissões para edição e publicação; (iii) participação do contendedor e do programador; (iv) disponibilização de informações

relevantes e (v) mapa do site; (vi) uso de tecnologias de padrão aberto; (vii) uso de gerenciador de conteúdo; (viii) uso de banco de dados, (ix) servidor Web, (x) sistema operacional, (xi) sistema de segurança e (xii) gerenciador de conteúdo de padrão aberto.

Todas as demais características se aproximam mais de laços fracos, algumas mais acentuadamente fracas do que outras. Em uma situação mediana, com percentuais entre de 40% a 50% figuraram 7 laços fracos: (i) uso de mecanismo de inferência; (ii) disponibilização de informação semanticamente completa e consistente; (iii) profundidade de cobertura das informações; (iv) gestão descentralizada de conteúdo; (v) presença de engenheiro do conhecimento, (vi) designer e (vii) gerente de conteúdo.

Laços muito fracos, partindo de 0 a 37% foram constatadas em 12 condições estruturais dos portais: (i) arquivos em servidor com navegação multidimensional com facetadas, informações semi estruturadas e extensíveis, alterações descentralizadas; (ii) controle no uso de vocabulário; (iii) disponibilização de lista de discussão, (iv) mecanismo de ajuda, (v) boletins eletrônicos e (vi) mapas de tópicos pessoais; (vii) atualização da informação por adição de novas classificações; (viii) processamento automatizado da informação; (ix) participação do gerente de banco de dados e (x) gerente de segurança de rede e nuvem; (xi) inclusão de membros com diferentes permissões; (xii) usuários participaram da construção e continuam participando do melhoramento do portal.

A participação do usuário como laço muito fraco marca o baixo envolvimento desses nos portais o que merece atenção cuidadosa, pois, esses atores também são demandantes de melhorias uma vez que são consumidores de serviços e produtos informacionais disponibilizados.

Laços criticamente fracos, até 10%, envolveram 9 condições: (i) presença de engenheiro de ontologia; (ii) uso de ontologia; (iii) uso de linguagens de programação semântica, (iv) modelagem de dados RDF e (v) metaontologia OWL; (vi) uso de tecnologia de importação de informações; (vii) publicação automatizada de informações; (viii) uso de editor de ontologia; (ix) disponibilização de materiais para consulta de processamento semântico, fórum e chat.

Na rede de portais todos os laços fortes e fracos foram representados no conteúdo dos indicadores qualitativos semânticos validados, já que representam os portais em um cenário de evolução sociotécnica.

Nesse contexto, o conjunto de indicadores qualitativos pode ser utilizado para avaliação do potencial semântico dos portais do conhecimento de universidades e implementação de melhorias, a partir de estruturação evoluída, marcadamente na recuperação de informações relevantes para toda a comunidade usuária. No Quadro 22, sintetiza-se, então, os indicadores finalizados e inferidos como parâmetros mínimos para a usabilidade semântica e sociotécnica.

Quadro 22 - Indicadores qualitativos com parâmetros semânticos e sociotécnicos

DIMENSÃO	INDICADOR	PARÂMETRO	
		SEMÂNTICO	SOCIOTÉCNICO
ARMAZENAMENTO DE DADOS	USO DE TECNOLOGIAS DE PADRÃO ABERTO PARA TECNOLOGIAS DE BASE E TECNOLOGIAS SEMÂNTICAS	✓	
GESTÃO DA INFORMAÇÃO	DOCUMENTOS ARMAZENADOS EM SERVIDOR, COM INFORMAÇÕES ORGANIZADAS DE FORMA SEMI-ESTRUTURA E DESCENTRALIZADAS	✓	
	ONTOLOGIA COMO SISTEMA DE RECUPERAÇÃO DA INFORMAÇÃO	✓	✓
	NAVEGAÇÃO POR FACETAS	✓	
	GESTÃO DESCENTRALIZADA DE CONTEÚDOS	✓	
	MEMBROS REGISTRADOS COM DIFERENTES PERMISSÕES PARA MANIPULAÇÃO DE DADOS E INFORMAÇÕES	✓	✓
	ATUALIZAÇÃO DE INFORMAÇÕES POR ADIÇÃO DE NOVAS CLASSIFICAÇÕES E EXPANSÃO DA ESTRUTURA DA INFORMAÇÃO	✓	✓
	PROCESSAMENTO E PUBLICAÇÃO AUTOMÁTICA DE INFORMAÇÕES	✓	✓
	USO DE MECANISMO DE INFERÊNCIA	✓	✓
	USO DE LINGUAGENS DE PROGRAMAÇÃO SEMÂNTICA, MODELAGEM DE DADOS RDF E METAONTOLOGIA OWL	✓	
GESTÃO DE ONTOLOGIA	USO DE EDITOR DE ONTOLOGIA	✓	
	RECURSOS PARA IMPORTAÇÃO, EXPORTAÇÃO E COLABORAÇÃO DO USUÁRIO NA ESTRUTURA DA ONTOLOGIA	✓	✓
	COBERTURA, PRECISÃO, COMPLETEDE E CONSISTÊNCIA DAS INFORMAÇÕES	✓	✓
	DISPONIBILIZAÇÃO DE MATERIAIS SOBRE DESCRIÇÃO DO PROCESSAMENTO SEMÂNTICO OU ESTRUTURA DA ONTOLOGIA USADA	✓	✓
REDE DE ATORES	EQUIPE ESPECIALIZADA PARA GESTÃO DO PORTAL	✓	✓
	PARTICIPAÇÃO CONTÍNUA DOS USUÁRIOS NA CONSTRUÇÃO E MELHORAMENTO DO PORTAL	✓	✓
	DISPONIBILIZAR RECURSOS E SERVIÇOS DE SUPORTE DE COMUNICAÇÃO E COMPARTILHAMENTO	✓	✓

Fonte: Aatoria Própria (2021).

Finalizadas a análise de dados e a proposição dos indicadores qualitativos semânticos, apresentam-se as Considerações Finais delineadas no Capítulo 7, retomando-se a pergunta de pesquisa, os objetivos propostos e a análise final da pesquisadora.

7 CONSIDERAÇÕES FINAIS

Neste Capítulo apresenta-se os principais achados de Tese e responde-se à pergunta de pesquisa e aos objetivos traçados. Sequencialmente a essas pontuações discorre-se sobre a delimitação da pesquisa e as sugestões de estudos futuros.

7.1 ATENDIMENTO AOS OBJETIVOS E À PERGUNTA DE PESQUISA

Os portais do conhecimento de universidades constituem uma ferramenta dinâmica e importante para disseminar informação organizada, relevante e geradora de conhecimento. Para tanto, se faz necessário estruturar esses portais com recursos humanos e tecnologias adequadas para que seja possível disponibilizar serviços informacionais que correspondam aos interesses da comunidade acadêmica e demais usuários da informação.

A recuperação da informação nos portais, a partir de modelos tradicionais que recuperam documentos baseados em palavras-chave sem considerar as relações semânticas entre elas, fica limitada já que não é possível interpretar o contexto da consulta com precisão. A ontologia de domínio surge, então, como uma forma de minimizar tais limitações pela possibilidade de expandir a consulta incorporando conceitos significativos. Assim, com a ontologia fazendo parte da estrutura dos portais os resultados da busca são mais precisos e relevantes para o usuário.

Para que os portais do conhecimento de universidades se tornem semânticos é preciso utilizar tecnologias e condições propostas pela Web Semântica. Na estruturação e gestão desses portais, as atividades que envolvem planejamento se tornam mais facilitadas, pressupõe-se, quando se tem disponíveis indicadores que possam embasar decisões sobre a adoção de tecnologias e a provisão de outras condições que se façam necessárias para o pleno funcionamento dos portais dentro de uma abordagem semântica.

A questão norteadora desta pesquisa, neste contexto, foi relacionada a **quais indicadores qualitativos podem ser usados na avaliação do potencial semântico de portais do conhecimento de universidades**, entendendo que a semântica é o elemento fundamental para a recuperação rápida e precisa da informação nos portais.

A resposta a questão de pesquisa foi possível, pois, foram elaborados **indicadores qualitativos para avaliar o potencial semântico de um portal, como**

uma estratégia válida para melhorar a recuperação de informação, e pode ser usado como ferramenta prática para uma possível implantação e gestão de portais semânticos, favorecendo o aprendizado, o compartilhamento e a aplicação de conhecimento específico e interdisciplinar.

Os preceitos semânticos fundamentados na Web de significado e de colaboração alinham-se aos postulados dos sistemas sociotécnicos e da Teoria Ator-Rede que enfatizam as relações sociais entre diferentes atores. Tendo estes norteadores teóricos aliados a dados empíricos foi possível atender aos objetivos deste estudo.

O objetivo geral **Propor indicadores qualitativos contendo parâmetros para a avaliação do potencial semântico de portais do conhecimento de universidades, de acordo com os conceitos da TAR**, foi cumprido, conforme Figura 25 – Seção 6.3.1 e pelo estudo e análise do cenário dos portais das universidades federais brasileiras à luz do entrelaçamento das teorias estudadas e da pesquisa de campo de análise majoritariamente qualitativa.

Quanto aos objetivos específicos traçados também todos foram atendidos, conforme detalhados a seguir:

- **Objetivo específico 1** - Levantar as condições básicas para o estabelecimento do potencial semântico de portais, como um artefato sociotécnico. Foi atendido mediante a análise das condições mínimas para elaboração de estabelecimento de potencial semântico dos portais do conhecimento de universidades levantados por Guedes (2015), correlacionando-as com as teorias dos sistemas sociotécnicos e da Teoria Ator-Rede, descritas nas Seções 4.1 e 4.2, respectivamente. Foi estabelecido o alinhamento dos preceitos semânticos e sociotécnicos que apontam para a recuperação da informação e para a comunicação e a cooperação a partir da interface dos portais, o que é obtido a mediante aplicação de regras técnicas complexas envolvendo os atores sociais.
- **Objetivo específico 2** - Mapear os atores humanos e não humanos que compõem a rede estrutural de portais semânticos e suas funções. Atendido nas seções 6.2.1 e 6.2.2, sintetizados no Quadros 19 e 20. Os atores humanos compreendem a equipe de administradores, desenvolvedores e usuários do portal; já os atores não-humanos envolvem as tecnologias de estruturação básica e as tecnologias semânticas.

- **Objetivo específico 3** - Elencar indicadores qualitativos, considerando os atores envolvidos, seus interesses, funções e características técnicas. Foi atendido pelo mapeamento da rede nas Seções 6.2.1, 6.2.2 e 6.3, seguindo os traços deixados por atores humanos e humanos, suas funções, ações e responsabilidades.
- **Objetivo específico 4** - Validar os indicadores qualitativos iniciais propostos a partir de um levantamento (*survey*) em universidades e de um painel de especialistas. Atendido pela elaboração, aplicação e análise do questionário que está no Apêndice 2. Os parâmetros, validados pela literatura e também por um painel de especialistas formado pelos técnicos de TI das universidades que responderam ao questionário de pesquisa. Deste modo, ficou ratificada a importância dos parâmetros estabelecidos cujo conteúdo integram os indicadores qualitativos semânticos propostos.
- **Objetivo específico 5** - Apresentar os indicadores qualitativos para avaliação de portais. Atendido conforme a Figura 25, cujo conteúdo reflete uma estruturação evoluída dos portais para análise e adequação a partir dos preceitos semânticos e sociotécnicos.

A análise dos traços deixados pelos atores da rede evidenciou laços fortes e fracos, adaptando-se o conceito de Granovetter (1983). Os laços fracos que oportunizam a inovação, envolveram mais variáveis o que denota a importância do fortalecimento destes para a evolução dos portais.

A partir dos resultados deste estudo foi possível concluir que, neste momento, nenhum portal de universidade federal usa ontologia e tecnologias semânticas em sua estruturação, o que inviabiliza a pesquisa semântica e o retorno de informação relevantes nas buscas. A baixa participação do usuário e a oferta de poucos serviços de comunicação e compartilhamento dificultam o uso do portal como uma ferramenta agregadora de informação e conhecimento.

O conjunto de indicadores validados podem servir como ferramenta de gestão para adequação dos portais do conhecimento de universidades como sistemas sociotécnicos, em sintonia com a evolução tecnológica. Com o uso destes indicadores é possível estabelecer estruturação semântica, partindo de processos de negociação e equalização de interesses dos atores da rede de portais, especialmente, os interesses dos usuários traduzidos pela recuperação semântica da informação que agrega valor no processo de construção do conhecimento. Assim, os portais das

universidades federais podem tornar-se mais dinâmicos, correspondendo às expectativas e interesses da comunidade universitária e demais usuários da informação.

De outra parte, este relato de Tese pode tornar-se um referente teórico para outras pesquisas sobre o tema, posto que levanta um conjunto de indicadores dentro de quatro dimensões – (i) Gestão de Dados; (ii) Gestão de Informação; (iii) Gestão de Ontologia e (iv) Redes de Atores e Sistemas Sociotécnicos – não trabalhados nesta combinação na literatura pertinente.

7.2. LIMITAÇÕES DA PESQUISA E TRABALHOS FUTUROS

A pesquisa teve com objeto central a proposição de indicadores qualitativos para avaliação do potencial semântico de portais do conhecimento de universidades. A construção dos indicadores foi pautada na análise de conteúdo da literatura, no mapeamento da rede de atores e na análise de dados empíricos coletados no ambiente dos portais do conhecimento de universidades nacionais.

Trata-se, portanto, de um **recorte de análise de sistemas de recuperação** da informação com poucas unidades, o que configura uma limitação da pesquisa, muito embora o retorno dos questionários tenha ficado acima do patamar definido nos procedimentos metodológicos adotados e chegado de maneira inopinada a uma amostra estratificada do território explorado. Por outro lado, a metodologia, as reflexões derivadas do ambiente analisado, especialmente o uso de ontologia e tecnologias de estruturação semântica, podem servir como ponto de partida para estudos futuros.

A avaliação dos portais unicamente feita pelos técnicos de TI é outra limitação do estudo, já que pode não refletir a visão dos usuários em termos do funcionamento do sistema de recuperação da informação e das condições e facilidades associadas aos serviços disponibilizados pelos portais, bem como as políticas institucionais para estes quesitos. Assim, estudos envolvendo a percepção da comunidade usuária pode oferecer outros parâmetros de avaliação, bem como sensibilizar a alta gestão acerca da importância da disponibilização de portais do conhecimento de universidades estruturados em normativas semânticas e sociotécnicas.

A pesquisa nos portais revelou o uso do mecanismo de pesquisa do Google como sistema de busca, o que parece ter adicionado algumas melhorias na

recuperação da informação. Deste modo, recomenda-se estudos mais acurados sobre a contribuição deste mecanismo nas pesquisas em portais.

Os indicadores propostos retratam tecnologias semânticas utilizadas na atualidade, mas o avanço do conhecimento pode colocar em evidência outras tecnologias emergentes, assim sendo, sugere-se novos estudos em ambientes sociotécnicos relacionados aos portais do conhecimento de universidades, bem como uma ampla avaliação da pertinência destes, envolvendo estudos interdisciplinares. Lembra-se que estes estudos, no entanto, foram deflagrados por Bernes-Lee, em 2001, mas na prática ainda há um longo caminho a percorrer para que os portais se tornem semânticos. Nesta perspectiva, estudos de como estruturar tecnicamente portais, com protocolos detalhados, podem auxiliar a evolução dos portais do conhecimento de universidades.

Com este estudo procurou-se enfatizar a importância da recuperação semântica da informação em portais, propondo um conjunto de indicadores qualitativos que podem ser usados para estabelecer o potencial semântico dos portais do conhecimento de universidades. Os indicadores qualitativos semânticos refletem uma estrutura necessária para que os portais forneçam pesquisa semântica, como um artefato sociotécnico, e facilitem a busca de informações que possam contribuir para a criação e o compartilhamento do conhecimento.

A tecnologia social como ferramenta transformadora agrega conhecimento à vida prática diária, os portais semânticos e sociotécnicos impulsionam esta engrenagem pela facilitação do acesso à informação e à interação com a sociedade, rumo a sustentabilidade das organizações no contexto educacional.

Salienta-se que esta pesquisa foi concluída no período da pandemia da Covid-19 que assolou o mundo, com vidas perdidas, sequelas, reclusão e recolhimento social, o que também trouxe dificuldades na fase da coleta de dados deste estudo. Por outro lado, a Ciência não parou, muito pelo contrário, e como nunca, a informação foi crucial no campo da Saúde, assim como o é em todas as áreas do conhecimento, abrindo caixas-pretas e dando passagem para a inovação.

REFERÊNCIAS

- ABDULLAH, Maher; ZAMIL, Mohammed AL. The effectiveness of classification on Information Retrieval System (Case Study). **ArXiv.org**, Ithaca, N.Y., Apr 2, 2018. Disponível em; <https://arxiv.org/abs/1804.00566>. Acesso em: 18 nov. 2018.
- AGÊNCIA DE TECNOLOGIA DA INFORMAÇÃO E COMUNICAÇÃO - UFPR. Disponível em: <https://www.agtic.ufpr.br/portal/a-agtic/>. Acesso em 07 out. 2021.
- AKSOY, Cem. **Semantics and result disambiguation for keyword search on tree data**. New Jersey: Institute of Technology, ProQuest Dissertations Publishing, 2016.
- AL BALUSHI, Taiseera Hazeem; SAMPAIO, Pedro Rafael Falcone; LOUCOPOULOS, Pericles. Eliciting and prioritizing quality requirements supported by ontologies: a case study using the ElicitO framework and tool. **Expert Systems**, New York, v. 30, n. 2, p. 129-151, May 2013. Disponível em: <https://doi.org/10.1111/j.1468-0394.2012.00625.x>. Acesso em: 20 fev. 2020.
- ALBAGLI, Sarita. Território e territorialidade. *In*: LAGES, Vinícius; BRAGA, Christiano; MORELLI, Gustavo. **Territórios em movimento: cultura e identidade como estratégia de inserção competitiva**. Rio de Janeiro: RelumeDumará; Brasília, DF: SEBRAE, 2004, p. 23-69.
- ALOMRAN, Hamad Ibrahim. Text mining-based semantic Web architecture (TMSWA) for e-Learning systems. **International Journal of Machine Learning and Computing**, Denver, v.4, n. 4, p. 333-338, Aug. 2014. Disponível em: <https://doi.org/10.7763/IJMLC.2014.V4.433>. Acesso em: 20 fev. 2020.
- ALVARENGA, Lídia. Bibliometria e arqueologia do saber de Michel Foucault: traços de identidade teórico-metodológica. **Ciência da Informação**, Brasília, v. 27, n. 3, set. 1998.
- AMATO, Flora. *et al.* Big data meets digital cultural heritage: design and implementation of SCRABS, a smart context-aware browsing assistant for cultural environment. **Journal on Computing and Cultural Heritage**, New York, n. 6, April 2017. Disponível em: <https://doi.org/10.1145/3012286>. Acesso em: 15 abril 2021.
- AMBROZEWICZ, Paulo Henrique Laporte. **Gestão da qualidade na administração pública: histórico, pbqp, conceitos, indicadores, estratégia, implantação e auditoria**. São Paulo: Atlas, 2015.
- ANDRADE, Gabriela Rieveres Borges de; VAITSMAN, Jeni; FARIAS, Luis Otávio. Metodologia de elaboração do índice de responsividade do serviço (IRS). **Cad. Saúde Pública**, Rio de Janeiro, v. 26, n.3, p. 523-534, mar, 2010. Disponível em: <https://doi.org/10.1590/S0102-311X2010000300010>. Acesso em: 15 fev. 2021.

AROUA, Essayeh; MOURAD, Aabed. An ontology-based framework for enhancing personalized content and retrieval information. **INTERNATIONAL CONFERENCE ON RESEARCH CHALLENGES IN INFORMATION SCIENCE**, 2017. **Proceedings**...Brighton, 2017. Disponível em: <https://doi.org/10.1109/RCIS.2017.7956547>. Acesso em: 15 fev. 2020.

ASPURA, Yanti Idaya Mohd Khalid; NOAH, Shahrul Azman Mohd. **Semantic text-based image retrieval with multi-modality ontology and DBpedia**, Oxford, v. 35, n. 6, p. 1191-1214, 2017.

AZANI, Cyrus. An open systems approach to system of systems engineering. *In*: JASMSHIDI, M. (ed.) **System of systems engineering: innovations for the 21st century**. p. 21-43. New Jersey: John Wiley & Sons, 2009. Disponível em: <https://doi.org/10.1002/9780470403501.ch2>. Acesso em: 16 nov. 2021.

BAEZA-YATES, Ricardo; RIBEIRO-NETO, Berthier. **Modern Information Retrieval**, New York: ACM Press, 1999.

BAHAFID, Abdelhadi *et al.* Toward a Semantic search engine for E-Commerce. **International Journal of Advanced Studies in Computers, Science and Engineering**, India, v. 4, n. 10, p. 26-34, 2015.

BALDIN, Nelma; MUNHOZ, Elzira Maria Bagatin. Snowball (bola de neve): uma técnica metodológica para pesquisa em educação ambiental comunitária. *In*: CONGRESSO NACIONAL DE EDUCAÇÃO – EDUCERE, 5, Curitiba, 2011. **Anais**... PUC, Curitiba, 2011.

BARBA-GONZÁLEZ, Cristóbal *et al.* BIGOWL: Knowledge centered big data analytics. **Expert Systems with Applications**, New York, v. 115, p. 543, Jan, 2019.

BARDIN, L. **Análise de conteúdo**. 3. ed. Lisboa: Edições 70, 2011.

BELLANDI, Andrea *et al.* Assisted knowledge base generation, management and competence retrieval. **International Journal of Software Engineering and Knowledge Engineering**, Singapur, v.22, n.8, p.1007-1037, 2012. Disponível em: <https://doi.org/10.1142/S021819401240013X>. Acesso em: 15 nov. 2020.

BENABDERRAHMANE, Sidahmed *et al.* Smart4Job: a big data framework for intelligent job offers broadcasting using time series forecasting and semantic classification. **Big data Research**, Netherlands, v.7, p. 16-30, March, 2017. Disponível em: <https://doi.org/10.1016/j.bdr.2016.11.001>. Acesso em: 10 nov. 2020.

BERNERS-LEE, Tim; HENDLER, James; LASSILA, Ora. The semantic *Web*: a new form of *Web* content that is meaningful to computers will unleash a revolution of new possibilities. **Scientific American**, New York, v. 284 n. 5, May, 2001.

BIERNACKI, Patrick; WALDORF, Dan. Snowball sampling: problems and techniques of chain referral sampling. **Sociological Methods & Research**, Thousand Oaks, California, n. 1, p. 141-163, Nov. 1981. Disponível em: <https://doi.org/10.1177/004912418101000205>. Acesso em: 28 out. 2021.

BIJKER, Wiebe Eco; HUGHES, Thomas Park; PINCH, Trevor. **The social construction of technological systems**. Cambridge: The MIT Press, 1989.

BIREME. **Biblioteca Virtual em Saúde - tutorial de pesquisa bibliográfica versão reduzida**. 2009. Disponível em https://bvsmis.saude.gov.br/bvs/publicacoes/apostila_biblioteca_virtual_saude_reduzida.pdf. Acesso em: 20 jan. 2022.

BIZER, Christian; HEATH, Tom; BERNERS-LEE, Tim. Linked data-the story so far. **International Journal on Semantic Web and Information Systems**, Hershey, Pennsylvania, EUA v. 5, p.1-22, 2009. Disponível em: <http://dx.doi.org/10.4018/jswis.2009081901>. Acesso em: 17 out. 2021.

BONNEFOY, C.; ARMIJO, M. **Indicadores de desempeño en el sector público**. Santiago do Chile: ILPES, 2005.

BORGMAN, Christine Louise; FURNER, Jonathan. Scholarly Communication and Bibliometrics. **Annual Review of Information Science and Technology**, Leesburg, Virginia, v. 36, n. 1, p. 2-72, 2002. Disponível em: <https://doi.org/10.1002/aris.1440360102>. Acesso em: 17 nov.2021.

BORGES, Paulo Renato Soares; SILVEIRA, Ismar Franco. **Adding and segmenting educational videos: experiences of teacher users in an educational portal**. 2019. Disponível em: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8746107>. Acesso em 20 set. 2021.

BÖRNER, Katy *et al.* VIVO: A semantic approach to scholarly networking and discovery. **Synthesis Lectures on the Semantic Web: Theory and Technology**, San Rafael, California, v. 7, n. 1, p. 1-178, 2012. Disponível em: <https://doi.org/10.2200/S00428ED1V01Y201207WBE002>. Acesso em: 20 jul. 2021.

BRAGA, Luis Paulo Vieira. **Introdução à mineração de dados**. 2. ed. rev. e ampl. Rio de Janeiro: E-papers, 2005.

BRANDT, Mariana Baptista; VIDOTTI, Silvana Aparecida Borsetti Gregorio; SANTAREM SEGUNDO, José Eduardo. Modelo de dados abertos conectados para informação legislativa. **Informação & Sociedade: Estudos**, João Pessoa, v. 28, n. 2, p. 149-161, maio/ago. 2018.

BRASIL. Ministério do Planejamento, Orçamento e Gestão. **Indicadores de programas: guia metodológico**. Brasília, 2010.

BRASIL. Presidência da República. Secretaria Geral - Subchefia para Assuntos Jurídicos. **Lei nº 13.709, de 14 de agosto de 2018**. Brasília, 2018.

BURNS, Tom R. The sociology of complex systems: an overview of actor-system-dynamics theory. **World Futures: the Journal of New Paradigm Research**, Milton, United Kingdom, v. 62, n. 6, 2006, p. 411-440. Disponível em: <https://doi.org/10.1080/02604020600798619>. Acesso em: 22 set. 2021.

BURNS, Tom R.; DIETZ, Thomas. Revolution: an evolutionary perspective. **International Sociology**, London, United Kingdom, v.16, n. 4, p. 531–555, 2001. Disponível em: <https://doi.org/10.1177/0268580901016004001>. Acesso em: 13 maio 2021.

CALLON, Michel. **La Science et ses réseaux**: genèse et circulation des faits scientifiques. Paris: La Découverte, 1989.

CALLON, Michel. Society in the making: the study of technology as a tool for sociological analysis. *In*: Wiebe Eco Bijker *et al.* (ed.). **The social construction of technological systems**: new directions in the sociology and history of technology. Cambridge, Mass: Mit Press, 1986. p. 83-103. (c)

CALLON, Michel. Some elements of a sociology of translation: domestication of the scallops and the fishermen of St Brieuc Bay. *In*: LAW, John (ed.). **Power, action and belief**: a new sociology of Knowledge? London: Routledge Kegan Paul, 1986, p. 196-233. (b)

CALLON, Michel. The sociology of an actor-network: the case of the electric vehicle. *In*: CALLON, Michel; LAW, John; RIP, Arie (ed.). **Mapping the dynamics of science and technology**. Houndmills: Macmillan Press, 1986, p. 19-34. (a)

CALLON, Michel; LATOUR, Bruno. Unscrewing the big leviathan: how actors macro-structure reality and how sociologist help them to do so. *In*: KNORR-CETINA, Karin; CICOUREL, Aaron Victor (ed.). **Advances in social theory and methodology**: towards an integration of micro and macro-sociology. Boston: Routledge, 1981. p. 277-303.

CALLON, Michel; LAW, John; RIP, Arie (ed.). **Mapping the dynamics of science and technology**. Houndmills: Macmillan Press, 1986. p. 19-34.

CASADEVALL, Arturo; FANG, Ferric C. *et al.* Rigorous science: a how-to guide. **mBio** [online], Washington, v. 07, n. 06, 2016. e01902-16. Disponível em <https://journals.asm.org/doi/epub/10.1128/mBio.01902-16>. Acesso em 05 jan. 2021.

CLUNIS, Juliane. Designing an ontology for managing the diets of hypertensive individuals. **International Journal on Digital Libraries**, Basingstoke, England, v. 20, n. 3, p. 269-284, September 2019. Disponível em: <https://doi.org/10.1007/s00799-018-0253-y>. Acesso em: 18 abr. 2021.

CORDELLA, Antonio. Information infrastructure: an actor-network perspective. **International Journal of Actor-Network Theory and Technological Innovation**, Hershey, Pennsylvania, v.2, n.1, p. 27-53, 2010.

CORREA, Débora Alvernaz. **Uma abordagem para extração de conteúdos baseada em características estruturais e navegacionais de portais Web**. 2012. 110 fls. Dissertação (mestrado) - Instituto Militar de Engenharia, Rio de Janeiro, 2012.

CRESPO BORGES, Thomás. **Respuestas a 16 preguntas sobre el empleo de expertos encruz la investigación pedagógica**. Lima: Editorial San Marcos, 2007.

CRESWELL, John W.; CLARK, Vicki L. P. **Pesquisa de métodos mistos**. 2. ed. Porto Alegre: Penso, 2013.

CRUZ, Matheus Mayron Lima da *et al.* SemanticSUS: um portal semântico baseado em ontologias e dados interligados para acesso, integração e visualização de dados do SUS. In: WORKSHOP DE FERRAMENTAS E APLICAÇÕES - SIMPÓSIO BRASILEIRO DE COMPUTAÇÃO APLICADA À SAÚDE (SBCAS), 19, 2019, Niterói. **Anais ...** Porto Alegre: Sociedade Brasileira de Computação, 2019. p. 13-18.

DAGNINO, Renato. **Tecnologia social**: contribuições conceituais e metodológicas. Campina Grande: EDUEPB, 2014.

DATE, C. J. **Introdução a sistemas de bancos de dados**. Rio de Janeiro: Campus, 2004.

DAVIS, Linda L. Instrument review: getting tyhe most from your panel of experts. **Applied Nursing Research**, New York, v. 5, p. 194-197, 1992.

DIAS, Cláudia A. Portal corporativo: conceitos e características. **Ciência da Informação**, Brasília, v. 30, n. 1, p. 50-60, jan. /abr. 2001.

DONABEDIAN, Avedis. **The methods and findings of quality assessment and monitoring: an ilustrated analysis**. Ann Arbor, Michigan: Health Administration Press, 1985.

DUARTE, Mariana Machado Garcez; HARA, Carmem S. Otimização do mapeamento de consultas SPARQL para SQL. In: ESCOLA REGIONAL DE BANCO DE DADOS, 24, 2018, Porto Alegre. **Anais ...**, Porto Alegre, SBC, 2018.

DUMER, Luciana; ALBUQUERQUE, Maria Elizabeth Baltar Carneiro de. **O MARC 21 e outros formatos de intercâmbio bibliográfico**. João Pessoa: UFPB, 2021.

EL HAJI, Essaid; AZMANI, Abdellah.; El Harzli, Mohamed. Multi-expert system design for educational and career guidance: an approach based on a multi-agent system and ontology. **International Journal of Computer Science Issues (IJCSI)**, Mauritius, v.11, n. 5, p. 46-52, Sep 2014.

ELBANNA, Amany. Applying actor network theory and managing controversy. **Information Systems Theory**. New York: Springer, 2012. p.117-129. Disponível em: https://link.springer.com/chapter/10.1007%2F978-1-4419-9707-4_7. Acesso em: 18 nov. 2021.

ELMASRI, Ramez; NAVATHE, Shamkant B. **Database systems**. Boston, MA: Pearson Education, 2011.

EMERSON, Ralph Waldo. Disponível em https://www.pensador.com/autor/ralph_waldo_emerson/. Acesso em 14. Jul 2021.

EROFEEVA, Maria. **Actor-network theory: an object-oriented sociology without objects?** *Logos* (Russian Federation), v. 27, n. 3, p. 83-112, 2017.

FAHAD, Muhammad *et al.* Towards classification of *Web* ontologies for the emerging Semantic Web. **Journal of Universal Computer Science**, Sofi, Bulgária, v. 17, n. 7, p. 1021-1042, 2011. Disponível em; https://www.jucs.org/jucs_17_7/towards_classification_of_web/jucs_17_07_1021_1042_fahad.pdf. Acesso em 14. jul 2021.

FAM, Soo-Fen; NORISZURA, Ismail; SHINYIE, Wendy L. THE magnitude of big data 5vs in business macroclimate. **International Journal of Recent Technology and Engineering (IJRTE)**, Malásia, v. 8, n. 1S5, p. 497-503, June 2019. Disponível em: https://www.researchgate.net/publication/335028796_The_Magnitude_of_Big_Data_5vs_in_Business_Macroclimate. Acesso em: 15 nov. 2021.

FEENBERG, Andrew. **Critical theory of technology**. New York: Oxford University Press, 1992.

FERNÁNDEZ, Miriam. *et.al.* Semantically enhanced information retrieval: an ontology-based approach. **Web semantics: Science and agents on the World Wide Web**, Oxford, v.9, p.434-52, 2011.

FERNEDA, Edberto. **Recuperação de informação: análise sobre a contribuição da Ciência da Computação para a Ciência da Informação**. 2003. 147 fls. Tese (doutorado) – Escola de Comunicação e Artes, Universidade de São Paulo, 2003

FERNEDA, Edberto. Redes neurais e sua aplicação em sistemas de recuperação de informação. **Ciência da Informação**, Brasília, v.35, n.1, p. 26-30, 2006. Disponível: <https://www.scielo.br/j/ci/a/SQ9myjZWLxnyXfstXMgCdch/?lang=pt>. Acesso em: 18 nov 2021.

FORLANO, Laura. Post humanism and design. **SheJi**, Shanghai, v. 3, n. 1, p. 16-29, 2017. Disponível em; <https://www.sciencedirect.com/science/article/pii/S2405872616300971?via%3Dihub>. Acesso em: 18 nov. 2021.

FROSTERUS, Matias.; HYVÖNEN, Eero.; LAITIO, Joonas. DataFinland: a semantic portal for open and linked datasets. **The Semantic Web: research and applications**. ESWC: Extended Semantic Web Conference, Heraklion, 29 May 2011. Disponível em: <https://link.springer.com/book/10.1007/978-3-642-21064-8>. Acesso em: 18 nov 2021.

GAO, Ping. Using actor-network theory to analyze strategy formulation. **Information Systems Journal**, Oxford, v.15, n. 3, p. 255-275, 2005. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1111/j.1365-2575.2005.00197.x>. Acesso em: 18 nov 2021.

GARGIULO, Francesco *et al.* Aerospace information system based on semantic technologies and ontology management: A Web portal for semantic search and document categorization. In: INTERNATIONAL CONFERENCE ON DATA MANAGEMENT TECHNOLOGIES AND APPLICATIONS, 3, Vienna, 2014, **Proceedings** ... Vienna, 2014 p.341-348. Disponível em: <https://www.scitepress.org/Link.aspx?doi=10.5220/0004994703410348>. Acesso em; 18 nov 2021.

GAVRILOVA, E. A. Implementation features a university data portal using the Semantic *Web* technology. **Programming and Computer Software**, New York, v. 37, n. 1, p. 48-55, Jan 2011. Disponível em; <https://link.springer.com/content/pdf/10.1134/S0361768811010026.pdf>. Acesso em: 17 nov. 2021.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 5. ed. São Paulo: Atlas, 2010.

GIRASE, Ashutosh V., PATNAIK, Girish K. ; PATIL, Sandip S. Ontology driven approach for effective decision making. **International Journal of Computer Applications**; New York, v. 148, n .7, p.15-21, 2016.

GRANOVETTER, Mark. The strength of weak ties: a network theory revisited. In: **Sociological Theory**. San Francisco: Ed. Randall Collins, 1983. v.1. p. 2001-2233.

GRUBER, Tom. **What is an ontology?** 1996. Disponível em: <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>. Acesso em: 01 maio. 2021.

GRUBER, Thomas R. A Translation approach to portable ontology specifications. **Knowledge Acquisition**, Stanford, Ca, v. 5, n. 2, p. 199-220,1993.

GUARINO, Nicola. **Formal ontology in information systems**. Proceedings of FOIS'98, Trento, Italy, 6-8 June 1998. Amsterdam: IOS Press, p. 3-15, 1998.

GUARINO, Nicola. Understanding, building and using ontologies. **International Journal of Human and Computer Studies**, London, v. 45, n. 2/3, p. 2, 1997. Disponível em: <http://ksi.cpsc.ucalgary.ca/KAW/KAW96/guarino/guarino.html>. Acesso em: 10 Jan. 2020.

GUEDES, Josefina A. S. **Portais do conhecimento de universidades**: proposta de um quadro referencial para estabelecimento do potencial semântico. 2015. 169 f. Dissertação (Mestrado em Tecnologia) – Programa de Pós-Graduação em Tecnologia, Universidade Tecnológica Federal do Paraná. Curitiba, 2015.

GUEDES, Josefina. A. S.; STRAUHS, Faimara do Rocio. Portais do conhecimento de universidades: um quadro referencial para avaliação de potencial semântico. **Liinc em Revista**, Rio de Janeiro, v.12, n.1, p. 166-179, maio 2016. Disponível em: <http://revista.ibict.br/liinc/article/view/3667/3114>. Acesso em 17 nov. 2021.

GUERTIN, H el ene. BERNHARD, Paulette. **Les 6  tapes d'un projet de recherche d'information**.  cole de biblioth conomie et des sciences de l'information (EBSI), Universit  de Montr al, Qu bec, 2005. *Online*. Dispon vel em: <http://www.ebsi.umontreal.ca/jetrouve/projet/index.htm>. Acesso em: 02 jan. 2021.

GUO, Kehua *et al.* SOR: An optimized semantic ontology retrieval algorithm for heterogeneous multimedia big data, **Journal Computational Science Education**, Durham v. 28, p. 455-465, 2018. Dispon vel em: <https://www.sciencedirect.com/science/article/pii/S187775031730176X?via%3Dihub>. Acesso em: 17 nov. 2021.

HAUKE, Krzysztof; OWOC, Mieczslaw L.; PONDEL, Maciej. Usability of knowledge portals for exclusives in local governments. In: AI4KM: ARTIFICIAL INTELLIGENCE FOR KNOWLEDGE MANAGEMENT; IFIP ADVANCES IN INFORMATION AND COMMUNICATION TECHNOLOGY, Warsaw, 469., 2014. Cham: Springer, 2015. Dispon vel em: <https://link.springer.com/book/10.1007/978-3-319-28868-0>. Acesso em: 06 out. 2021.

HAWKE, Sandro. *et al.* **W3C Semantic Web Activity**, 2013, retrieved from <http://www.w3.org/2001/sw/>. Acesso em 20 dez 2021.

HEATH, Tom; BIZER, Christian. Linked data: evolving the Web into a global data space. 1st ed. In: **Synthesis lectures on the semantic Web: theory and technology**. California: Morgan & Claypool, 2011. p.1-136. Dispon vel em: <https://www.morganclaypool.com/doi/abs/10.2200/S00334ED1V01Y201102WBE001>. Acesso em: 17 nov. 2021.

HU, Danchao. **Using actor-network theory to understand inter-organizational network aspects for strategic information systems planning**. Master's thesis, University of Twente, 2011. Dispon vel em: http://essay.utwente.nl/63005/1/master_thesis_s0211494_final_version.pdf. Acesso em 17 nov. 2021.

HUANG, Chang-Qin, *et al.* EIS: An educational information intelligent search engine supported by semantic services. **International Journal of Distance Education Technologies**, Hershey, PA, v.9, n.1, p.21-43, Jan-March, 2011. Dispon vel em: https://go.gale.com/ps/i.do?id=GALE%7CA275852527&sid=googleScholar&v=2.1&it=r&linkaccess=abs&issn=15393100&p=AONE&sw=w&userGroupName=ufpr_br. Acesso em: 17 nov. 2021.

HUANG, Mutao; MAIDMENT, David R.; TIAN, Yong. Using SOA and RIAs for Water Data Discovery and Retrieval. **Environmental Modelling & Software**, Oxford, v. 26, p. 1309-1324, 2011. Dispon vel em: <http://dx.doi.org/10.1016/j.envsoft.2011.05.008>. Acesso em: 17 nov. 2021.

HUGHES, Thomas. The Evolution of large technological systems. In: BIJKER, W., HUGHES, T. e PINCH, T. **The social construction of technological systems**. Cambridge: The MIT Press, 1989, p. 51-82.

HYVÖNEN, Eero *et al.* Mapping manuscript migrations on the semantic web: a semantic portal and linked open data service for premodern manuscript research. In: Hotho, A. *et al.* (eds) **The Semantic Web - ISWC 2021**. New York: Springer, 2021. Disponível em https://doi-org.ez22.periodicos.capes.gov.br/10.1007/978-3-030-88361-4_36. Acesso em 20 set. 2021

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE.

Estimativas da população 2021. Disponível em:

https://ftp.ibge.gov.br/Estimativas_de_Populacao/Estimativas_2021/estimativa_dou_2021.pdf. Acesso em 24 out 2021

IYAMU, Tiko; MGUDLWA, Sibulela. Transformation of healthcare big data through the lens of actor network theory. **International Journal of Healthcare Management**, London, v.11, n.3, p. 182-192, 2018.

JORGE, Marcelo Alvin. A. **Ontologias no suporte a portais semânticos**. 2005. 119 f. Dissertação (mestrado) - Universidade Federal de Minas Gerais. Minas Gerais, 2005.

KOIVUNEN, Marja-Riitta; MILLER, Eric. W3C: Semantic *Web* activity. In: SEMANTIC WEB KICK-OFF SEMINAR IN FINLAND, 2004. **Proceedings ...** Finland: Helsinki Institute for Information Technology, Nov. 2001.

LACHTIM, Fernanda A. *et al.* POSEDU: A Semantic Educational Portal. **IEEE Multidisciplinary Engineering Education Magazine**, Boston, v. 4, n. 3, Sep. 2009.

LACHTIM, Fernanda Aparecida; MOURA, Ana Maria C; CAVALCANTI, Maria Cláudia. Ontology matching for dynamic publication in semantic portals. **Journal of the Brazilian Computer Society**, Rio de Janeiro, v. 15, n. 1, p. 27-43, 2009. Disponível em: <https://www.scielo.br/j/jbcos/a/JFX7KG8S7DbwMwhmyjDx4KG/?lang=en>. Acesso em: 17 nov.2021.

LAST, John M. **A dictionary of epidemiology**. 2nd ed. New York: Oxford University Press, 1998.

LATOUR, Bruno. **Ciência em ação**: como seguir cientistas e engenheiros sociedade afora. São Paulo: Editora UNESP, 2001.

LATOUR, B. **Jamais fomos modernos**. Rio de Janeiro: Ed. 34, 1994. (b)

LATOUR, Bruno. On actor-network theory: a few clarifications, plus more than a few complications. **Logos**, Russian Federation, v. 27, n. 1, p. 173-200, 2017.

LATOUR, Bruno. On technical mediation. **Common Knowledge**, New York, v. 3, n. 2, p. 29-64, 1994. (a)

LATOUR, Bruno. **Reagregando o social**: uma introdução à teoria do Ator-Rede. São Paulo: Edusc, 2012, 301p.

LATOURE, Bruno. **Petite réflexion sur le culte moderne des dieux faitiches**. Paris: Les Empêcheurs de Penser en Rond, 1996.

LATOURE, Bruno. Where are the missing masses? The sociology of a few mundane artifacts,” in **Shaping technology/building society: studies in sociotechnical change**. Cambridge: MIT Press, 1992.

LAUSEN, Holger *et al.* Semantic Web portals – state of the art survey. **Journal of Knowledge Management**, West Yorkshire, v. 9, n. 5, p. 40-49, 2005. Disponível em:

<https://www.emerald.com/insight/content/doi/10.1108/13673270510622447/full/html>. Acesso em: 17 nov. 2021.

LAW, John. **After method: mess in social Science research**. London: Routledge, 2010, 188p.

LAW, John. **Notes on the theory of the actor network: ordering, strategy and heterogeneity**. Lancaster, Lancaster University, 1992.

LAWAN, Abba; RAKIB, Abdur. The semantic Web rule language expressiveness extensions-a survey. **Arkiv.org**, Malásia, arXiv preprint arXiv:1903.11723, 2019.

LEE, Jinhyur *et al.* Ranking paragraphs for improving answer recall in open-domain question answering. In: CONFERENCE ON EMPIRICAL METHODS IN NATURAL LANGUAGE PROCESSING, 2018. **Proceedings ...** Belgium, October 31 - November 4, 2018 Pro Oct / Nov, 2018. Disponível em: arXiv.org > cs > arXiv:1810.00494. Acesso em: 17 nov. 2021.

LI, Huankang *et al.* An optimized approach for massive web page classification using entity similarity based on semantic network. **Future Generation Computer Systems**, Amsterdam, v. 76, p. 510–518, 2017. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0167739X17303321?via%3Dihub>. Acesso em: 17 nov. 2021.

LINSINGEN, Irlan von; BAZZO, Walter A.; PEREIRA, Luiz T. V. O que é ciência, tecnologia e sociedade? In: **Introdução aos estudos CTS: ciência, tecnologia e sociedade**. Espanha: OEI, 2003. p. 119-156 (Cadernos de Ibero-América).

LOPES, Ilza Leite. Estratégia de busca na recuperação da informação: revisão da literatura. **Ci. Inf.**, Brasília, v. 31, n. 2, p. 60-71, maio/ago. 2002.

MAATOUK, Yasser. Building AIPedia ontology to evaluate research impact in artificial intelligence area. **Academia Letters**, August, 2021. Disponível em: <https://doi.org/10.20935/AL2781>, 2021. Acesso em 20 set. 2021.

MACOME, Esselina. On implementation of an information system in the mozambican context: the EDM case viewed through ANT lenses. **Information Technology for Development**, Amsterdam, v. 14, n. 2, p.154-170, 2008.

MAEDCHE, Alexander. *et al.* SEAL - a framework for developing semantic portals. In BRITISH NATIONAL CONFERENCE ON DATABASES, 18,

Proceedings... Oxford, UK, LNCS Springer, July, 2001. Disponível em: <https://dl.acm.org/doi/10.1145/500737.500762>. Acesso em: 18 nov. 2021.

MAEDCHE, Alexander; STAAB, Steffen. Ontology learning for the Semantic Web. **IEEE intelligent systems & their applications**, Los Alamitos, CA, v. 16, n. 2, p. 72-79, Mar-Apr, 2001. Disponível em: <https://ieeexplore.ieee.org/document/920602>. Acesso em 18 nov. 2021.

MAIA, Alessandra; SERAFIM, Sabrina. Análise da Teoria Ator-Rede (TAR) e sua relação com os paradigmas de relações públicas. **Contemporânea**, Rio de Janeiro, v. 9, n. 1, p. 123-137, 2011.

MALGAONKAR, Rohit; DEVALE, Prakash. Optimizing web image search with alternative optimization technique. **International Journal of Application or Innovation in Engineering & Management**, India, v. 5, n. 6, p. 1-10, June 2016.

MALHOTRA, Meenakshi; NAIR, T. R. Gopalakrisnan. Evolution of knowledge representation and retrieval techniques. **International Journal of Intelligent Systems and Applications**, Hong Kong, v. 7, n. 7, p. 18-28, Jun 2015.

MALHOTRA, Naresh K. **Pesquisa de marketing**: uma orientação aplicada. 7. ed. Porto Alegre: Bookman, 2019.

MANIEGA-LEGARDA, David; PÉREZ-SALMERÓN, Gloria; GUERRERO-TORRES, Yolanda. El Portal de Bibliotecas de la Generalitat de Catalunya: un proyecto para los ciudadanos y los profesionales de la biblioteca pública. **El profesional de la información**, Spain, v. 17, n. 2, p. 183-187, 2008.

MARCONI, Maria Andrade.; LAKATOS, Eva Maria. **Técnicas de pesquisa**. 7. ed. São Paulo: Atlas, 2010b.

MARCONI, Maria Andrade; LAKATOS, Eva Maria. **Fundamentos de metodologia científica**. 7. ed. São Paulo: Atlas, 2010a.

MARCUSE, Herbert. Algumas implicações sociais da tecnologia moderna. In: **Tecnologia, guerra e fascismo**. São Paulo: Ed. Unesp, 1998. p.730104.

MARICATO, João de Melo. Procedimentos metodológicos em estudos bibliométricos e cientométricos: opções e reflexões no contexto dos processos de recuperação e organização da informação. In: COSTA, R. L. M. **Estudos contemporâneos em comunicação e artes: melhores teses e dissertações da ECA/USP**, 2010. São Paulo: ECA/USP, 2011.

MAROUF, Laila N.; KHALIL, Omar E. M. The influence of individual characteristics on knowledge sharing practices, enablers, and barriers in a project management context. **International Journal of Knowledge Management**, Hershey, v.1, n. 1, p. 1-27, 2015.

MARTINEZ-CRUZ, Carmen; BLANCO, Ignacio J.; VILA, M. Amparo Health systems as complex systems. **American Journal of Operations Research**, Irvine, Calif, v.3, p. 113-126, 2013.

MARTÍNEZ-GARCÍA, Mireya; HERNÁNDEZ-LEMUS, Enrique. Health Systems as Complex Systems. **American Journal of Operations Research**, Irvine, Calif., v.3, p. 113-126, 2013.

MARTINS, Mariana Thamires; SOARES, Maria Sueny Barbosa; FRANCELIN, Marivalde Moacir. Questões preliminares sobre os espaços de significação e a organização do conhecimento: conceitos, linguagens e conhecimentos marginais. **Informe: Estudos em Biblioteconomia e Gestão da Informação**, Recife, v. 1, n. 1, p. 78-86, 2012.

MASNER, Jan *et al.* Creation, storage and presentation of information content – semantics, sharing, presentation, and archiving. **Agris On-line Papers in Economics and Informatics**, Czechia, v.11, n. 1, p. 75-82, March 2019.

MASSALLI, Fábio. Ministério da Educação cria 5 novas universidades e empossa reitores. **Agência Brasil**, Brasília, publicado em 12/12/2019. Disponível em: <https://agenciabrasil.ebc.com.br/educacao/noticia/2019-12/ministerio-da-educacao-cria-5-novas-universidades-e-empossa-reitores>. Acesso em 24 out 2021.

MEIRELES, Magali Rezende Gouvêa; CEDÓN, Beatriz Valadares. Aplicações práticas dos processos de análise de conteúdo e de análise de citações em artigos relacionados a redes neurais artificiais. **Informação & Informação**, Londrina, v. 15, n. 2, p.77–93, 2010.

MINAYO, Maria Cecília de Souza. Construção de indicadores qualitativos para avaliação de mudanças. **Revista Brasileira de Educação Médica**, Brasília, v. 33, n. 1, supl. 1, p. 83-91, 2009.

MIOR, Luiz Carlos. Desenvolvimento rural: a contribuição da abordagem das redes sociais e sociotécnicas. In: SCHNEIDER, S. **A diversidade da agricultura familiar**. Porto Alegre: Editora UFRGS, 2006. p. 268-293.

MOHER, David. *et al.* **The PRISMA** Group preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. **PLoS Med**, San Francisco, v. 6, n. 7, e1000097, 2009. Disponível em: <https://doi.org/10.1371/journal.pmed.1000097>. Acesso em: 17 nov. 2021.

MORIN, Edgar. **Ciência com consciência**. 16. ed. Rio de Janeiro: Bertrand, 2014.

MOREIRA, H.; CALEFFE, L. G. **Metodologia da pesquisa para o professor pesquisador**. Rio de Janeiro: DP&A, 2006.

MORRISON, Donald R. PATRICIA - practical algorithm to retrieve information coded in alphanumeric. **Journal of the ACM (JACM)**, San Francisco, v. 15, n. 4, p. 514-534, 1968.

MOURA, Ana Maria de *et al.* A semantic integration approach to publish and retrieve ecological data. **International Journal of Web Information Systems**, Leichester, U.K, v. 1, n.1, p.87-119, 2015.

NADAL, Sergi *et al.* An integration-orient ontology to govern evolution in Big Data ecosystems. *Stijn. Information Systems*, Oxford, v. 79, Jan 2019. Disponível em: http://ceur-ws.org/Vol-1810/DOLAP_paper_09.pdf. Acesso em: 17 nov. 2021.

NAZÁRIO, Débora Cabral. Semantic portals as a support for knowledge management in organizations. INTERNATIONAL CONFERENCE ON INFORMATION SYSTEMS AND TECHNOLOGY MANAGEMENT, 10. **Proceedings...** São Paulo: CONTECSI, 2013. p. 3100-3119.

NGO, Vuong M.; CAO, Tru. H. Ontology-based query expansion with latently related named entities for semantic text search. **ArXiv.org**, Ithaca, N.Y, v.1, p. 1-12, 2018. Disponível em: <https://arxiv.org/ftp/arxiv/papers/1807/1807.05579.pdf>. Acesso em: 17 nov. 2021.

NGUYEN, Phuong T. et al. An automated approach to assess the similarity of GitHub repositories. **Software Quality Journal**, Andover, U.K., v. 28, n. 3, p. 1-37, 2020.

NI, Weizeng. **Ontology-based feature construction on non-structured data**. Cincinnati, ProQuest Dissertations Publishing, 2015. Disponível em: <https://www.proquest.com/openview/a8e847452f6a09c3080981701307a598/1?pq-origsite=gscholar&cbl=18750>. Acesso em: 17 nov. 2021.

NIELSEN, J. **Projetando web sites**. Rio de Janeiro, Editora Campus. 2000.

NIELSEN, Jacob.; LORANDER, Hoa. **Usabilidade na Web: projetando websites com qualidade**. Rio de Janeiro: Elsevier, 2007.

NIMBIX. **The eight v's of supercomputing and big data**. Disponível em: <https://www.nimbix.net/eight-vs-supercomputing-big-data>. Acesso em 14 out. 2021.

NONAKA, Ikujiro; TAKEUCHI, Hirotaka. **Criação de conhecimento na empresa: como as empresas japonesas geram a dinâmica da inovação**. Rio de Janeiro: Campus, 1997.

NOVAES, Henrique; DAGNINO, Renato. O fetiche da tecnologia. **ORG & DEMO**, Marília v.5, n.2, p.189-210, 2004.

NOY, Natalya F.; McGUINNESS, Deborah L. **Desarrollo de ontologías – 101: guía para crear tu primera ontología**. 2005. 29 p. Disponível em <<http://goo.gl/OMNr7n>>. Acesso em: 10 abr. 2007.

OCDE. **Indicadores e instrumento**. Disponível em: <https://www.ovtt.org/pt/guias/guia-de-indicadores-de-ciencia-e-tecnologia/#categorias-de-indicadores>. Acesso em 06. out. 2021.

OPEN KNOWLEDGE INTERNACIONAL. Disponível em <https://okfn.org/opendata/how-to-open-data/>. Acesso em: 23 mar 2020.

PAIVA, Fabio A. P. de; COSTA, José Alfredo F.; SILVA, Cláudio. R. M. An ontology-based recommender system architecture for semantic searches in vehicles sales portals. In: INTERNATIONAL CONFERENCE ON HYBRID ARTIFICIAL INTELLIGENCE SYSTEMS, HAIS, 2014. Cham: Springer, 2014. p. 537-548.

PATTUELLI, M. Cristina; MILLER, Matthew. Semantic network edges: a human-machine approach to represent typed relations in social networks. **Journal of Knowledge Management**, Hershey, Pa, v. 19, n. 1, p. 71-81, 2015. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/JKM-11-2014-0453/full/html>. Acesso em: 17 nov. 2020.

PEDRO, Rosa. Sobre redes e controvérsias: ferramentas para compor cartografias psicossociais. In: FERREIRA, Arthur *et al.* **Teoria ator-rede e psicologia**. Rio de Janeiro: Nau, 2010. p. 78-96.

PINHEIRO, Wallace Anacleto. **Busca em portais semânticos: uma abordagem baseada em ontologias**. 2004. 170 f. Dissertação (mestrado) - Instituto Militar de Engenharia, Rio de Janeiro, 2004.

POLANYI, M. **The tacit dimension**. Gloucester: Peter Smith, 1996.

PRADO, Alysson Bolognesi; BARANAUSKAS, Maria Cecilia Calani. Representing scientific associations through the lens of actor-network theory, 2012. INTERNATIONAL CONFERENCE ON COMPUTATIONAL ASPECTS OF SOCIAL NETWORKS (CASoN), 4, São Carlos. **Proceedings ...** São Carlos, 2012. p. 87-92, 2012. Disponível em: <https://tnr.nied.unicamp.br/ecoweb/nied/ecoweb/publicacoes/artigos-em-conferencias/representing-scientific-associations-through-the-lens-of-actor-network-theory/view.html>. Acesso em 17 fev. 2020.

PPGTE. Disponível em: <http://www.utfpr.edu.br/cursos/coordenacoes/stricto-sensu/ppgte/sobre/tecnologia-e-desenvolvimento>. Acesso em 15 dez 2021.

RANTALA, Heikki. *et al.* WarVictimSampo 1914–1922: a semantic portal and linked data service for Digital Humanities Research on War History. In: Harth A. *et al.* (eds) **The Semantic Web: ESWC 2020 Satellite Events**, Creta, v. 12124, maio 2020. Disponível em: https://link.springer.com/chapter/10.1007%2F978-3-030-62327-2_33. Acesso em: 06. out. 2021.

RAUTENBERG, Sandro. *et al.* Ferramenta ontoKEM: uma contribuição à Ciência da Informação para o desenvolvimento de ontologias. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.15, n.1, p. 239-258, abr. 2010.

RAUTENBERG, Sandro.; TODESCO, José Leomar.; STEIL, Andrea Valéria. Ontologias de domínio no mapeamento de instrumentos da gestão do conhecimento e de agentes computacionais da engenharia do conhecimento: o estado da arte. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 15, n. 2, p. 163-182, maio/ago. 2010.

RAZA, Muhammad Ahsan *et al.* Query expansion using conceptual knowledge in computer science. **Advanced Science Letters**, Stevenson Ranch, v. 24, n. 10, p. 7490-7493, Oct. 2018.

REYES-ÁLVAREZ, Ludmila; ROLDÁN-GARCÍA, Maria del Mar; ALDANA-MONTES, José. F. Tool for materializing OWL ontologies in a column-oriented database. **Software: Practice & Experience**, Chichester, Sussex, v. 49, n.1, p. 100-119, Jan, 2019. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/spe.2645>. Acesso em: 02 jan. 2021.

REYNOLDS, Dave; SHABAJEE, Paul; CAYZER, Steve. **Semantic Information Portals**. New York, ACM, 2004. p.290-291. Disponível em: <https://dl.acm.org/doi/abs/10.1145/1013367.1013440>. Acesso em 08 maio. 2020.

ROLIM, Tulio Vidal *et al.* SemanticSefaz: an ontology-based semantic portal for the government spending. SIMPÓSIO BRASILEIRO DE MULTIMÍDIA E WEB, WEBMEDIA (Webmedia 2019), 25., 2019, Rio de Janeiro. **Anais ...** Rio de Janeiro, Sociedade Brasileira de Computação, 2019.

RODRIGUES, Sara Firmino. **As redes sociais como canal de distribuição de notícias**. 2020. Dissertação - Universidade Lusófona de Humanidades e Tecnologias. Lisboa, 2020

ROSE, Jeremy; SCHEEPERS, Rose. Structuration theory and information system development - frameworks for practice. **European Conference on Information Systems**, Bled, Slovenia, v. 9, 27-29, June, 2001.

ROSENFELD, Louis; MORVILLE, Peter; ARANGO, Jorge. **Information architecture: for the Web and beyond**. 4th ed. Sebastopol, CA: O'Reilly Media, Inc., 2015.

RUBIO, Doris M. *et al.* Objectifying content validity: conducting a content validity study in social work research. **Social Work Research**, Washington, v. 27, n. 2, p. 94 – 111, 2003.

SACK, Robert D. **Human territoriality: a theory**. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8306.1983.tb01396.x>. Acesso em 02 jan. 2020.

SALES, Rodrigo; CAFÉ, Lúcia. Semelhanças e diferenças entre tesouros e ontologias. **DataGramZero**, Rio de Janeiro, v. 9, n. 4, p. 00, ago. 2008.

SALTON, Gerard. **Automatic information organization and retrieval**. New York: McGraw-Hill, New, 1968.

SAMPIERI, Roberto H.; COLLADO, Carlos F.; LUCIO, María del Pilar B. **Metodologia de Pesquisa**. 5. ed. Mexico: McGraw-Hill, 2010.

SCHLEICHER, Andreas. The economics of knowledge: weducation is key for Europe's success. Lisboa: Lisbon Council, 2006. Disponível em: <http://www.oecd.org/dataoecd/43/11/36278531.pdf>. Acesso em 24 out, 2021

SENEVIRATNE, Malage D. S.; FERNANDO, K. S. D.; KARUNARATNE, D. D. Generation of relation-extraction-rules based on Markov logic network for document classification. **International Journal of Advanced Computer Research**, Bhopal, p. 94- 111, Mar 2019. Disponível em: <https://www.accentjournals.org/PaperDirectory/Journal/IJACR/2019/3/2.pdf>. Acesso em: 29 out. 2021.

SHAANIKA, Ilrja; IYAMU, Tiko. Deployment of enterprise architecture in the Namibian Government: the use of activity theory to examine the influencing factors. **EJISDC**, Hong Kong, v. 71, n. 6, p. 1-21, 2015. Disponível em: <https://www.readcube.com/articles/10.1002%2Fj.1681-4835.2015.tb00515.x>. Acesso em: 29 out. 2021.

SHARMA, Abhishek. Nanotechnology ontology: semantic access to information in the nano world. **Journal of Library & Information Technology**, Delhi, v. 36, n. 1, p. 9-29, Jan 2016. Disponível em: <https://publications.drdo.gov.in/ojs/index.php/djlit/article/view/9198>. Acesso em: Acesso em 29 out. 2021.

SHILAKES, Christopher C.; TYLMAN, Julie. **Enterprise information portals**. New York, NY: Merrill Lynch, 16 Nov. 1998. Disponível em [\http://www.sagemaker.com/home.asp?id=500&file=Company/WhitePapers/lynch.htm. Acesso em 27 out. 2021.

SHIM, Yonwoon; SHIN, Dong-Hee. Analyzing china's fintech industry from the perspective of actor-network theory". **Telecommunications Policy**, London, v. 40, n. 2/3, p. 168-181, 2016. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0308596115001810>. Acesso em: 29 out. 2021.

SILVA, Renata; SANTOS, Plácida, FERNEDA, Edberto. Modelos de recuperação de informação e Web Semântica: a questão da relevância. **Inf. Inf.**, Londrina, v. 18, n. 3, p. 27 – 44, set. / dez. 2013. Disponível em: <https://www.uel.br/revistas/uel/index.php/informacao/article/view/12822>. Acesso em: 29 out. 2021.

SILVEIRA, Lucia da et al. Ciência aberta na perspectiva de especialistas brasileiros: proposta de taxonomia. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, Florianópolis, v. 26, p. 1-27, 2021. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/79646>. Acesso em: 16 dez. 2021.

SILVELLO, Gianmaria *et al.* Semantic representation and enrichment of information retrieval experimental data. **International Journal Digital Library Systems**, Hershey, PA, v. 18, p. 145–172, 2017. Disponível em: https://scholar.google.com/citations?view_op=view_citation&hl=en&user=pY24h-YAAAAJ&citation_for_view=pY24h-YAAAAJ:U4n9YNQMCAIC. Acesso em: Acesso em 29 out. 2021.

SIMÕES, Marcelo Godoy; SHAW, Ian S. **Controle e modelagem fuzzy**. São Paulo: Editora Blucher, 2007.

SIRISHA, Jagannadham; SUBBARAO, B. V.; KAVITHA, Donepudi. A. Cram on Semantic Web Components. **International Journal of Advanced Research in Computer Science**, Udaipur, India, v. 6, n. 3, p. 62-67, May 2015. Disponível em: <http://ijarcs.info/index.php/ljarcs/article/view/2471>. Acesso em: 29.10 2021.

SOARES, Sandro Vieira; PICOLLI, Icaro Roberto A.; CASAGRANDE, Jacir. Leonir. Pesquisa bibliográfica, pesquisa bibliométrica, artigo de revisão e ensaio teórico em administração e contabilidade. **Administração: Ensino & Pesquisa**, Rio de Janeiro, v. 19, n. 2, p. 308-339, 2018.

SOUZA, Dominique G. de; MIRANDA, Jean Carlos; SOUZA, Fabiano, dos Santos. Breve histórico acerca da criação das universidades no Brasil. **Revista Educação Pública**, Rio de Janeiro v.9, n.5, 2019

SPYNS, Peter *et al.* OntoWeb: a semantic Web community portal. In: INTERNATIONAL CONFERENCE ON PRACTICAL ASPECTS OF KNOWLEDGE MANAGEMENT, 4, 2002, Austria. **Proceedings ... Austria**, 2002. p. 298-302.

STALLINGS, William. **Arquitetura e organização de computadores**. São Paulo: Pearson, 2015.

STANESCU, Nicolae G. A smart approach for e-learning domain: the use of business models and semantic technologies. **Informática Econômica**, Bucharest, v.22, n.2, p. 68-80, 2018. Disponível em: <https://www.proquest.com/docview/2067316838/fulltext/A778C2CF1C154A46PQ/1?accountid=146694>. Acesso em: 29 out. 2021.

STRAUHS, Faimara. R. *et al.* **Gestão do conhecimento nas organizações**. Curitiba: Aymara, 2012. 128 p

SWAD - SEMANTIC WEB ADVANCED DEVELOPMENT. 2004. Disponível em <https://www.w3.org/2001/sw/Europe/>. Acesso em: 29 out. 2021.

TA, Chen D. C.; THI, Tuoi Phan. Automatic evaluation of the computing domain ontology. In: FUTURE DATA AND SECURITY ENGINEERING - FDSE, 2., Cham: Springer, 2015. v.9446. p.285-295. Disponível em: https://link.springer.com/chapter/10.1007/978-3-319-26135-5_21. Acesso em: 16 nov. 2021.(a)

TA, Chen D. C.; THI, Tuoi Phan. Improving the algorithm for mapping to relational database schema. In: PERNER, P. (Ed.). **Machine learning and data mining in pattern recognition**. Cham: Springer, 2015, v.9166. Disponível em: https://link.springer.com/chapter/10.1007/978-3-319-21024-7_9. Acesso em: 16 nov. 2021.(b).

TAG Crowd. Disponível em: <https://tagcrowd.com/>. Acesso em: 20 abr. 2021

TAHMASEBI, Amir M. *et al.* Automatic normalization of anatomical phrases in radiology reports using unsupervised learning. **Journal of Digital Imaging**, Feb 2019. Disponível em; <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6382634/>. Acesso 16 nov. 2021.

TATTO, L.; BORDIN, R. A. Filosofia e gestão do conhecimento: um estudo do conhecimento na perspectiva de Nonaka e Takeuchi. **Cadernos EBAPE.BR**, Rio de Janeiro, v. 14, n. 2, p. 340-350, 2016.

TELNOV, V. P. Semantic Web and search agents for Russian higher education. A pilot project. CEUR Workshop **Proceedings**, Moscow, 2015. p.195-204.

TERZIYAN, Vagan; GOLOVIANKO, M.; SHEVCHENKO, O. Semantic portal as a tool for structural reform of the ukrainian educational system, **Information Technology for Development**, Oxford, v. 21, n.3, p.381-402, 2015.

THANGSUPACHAI, Nopol; NIWATTHANAKUL, Supakrit; CHAMNONGSRI, Nisashal. Mapping learning object resource metadata for retrieval via linked open data. **International Information Institute (Tokyo) Information**; Koganei, v. 18, n.12, p. 5129-5142, Dec 2015.

THOMAS, Herman. Estructuras cerradas vs. Procesos dinámicos: trayectorias y estilos de innovación y cambio tecnológico. In: **Actos, actores y artefactos. Sociología de la Tecnología**, UNQ, Bernal, 2008, pp. 217-262.

THURSTONE LL. Attitudes can be measured. **American Journal of Sociology**, Washington, v. 33, p. 529-44, 1928. Disponível em: https://brocku.ca/MeadProject/Thurstone/Thurstone_1928a.html. Acesso em 06 out. 2021.

TRIOLA, M. F. **Introdução à estatística**. 12. ed. Rio de Janeiro: LTC, 2017.

TRIVIÑOS, Augusto Nivaldo Silva. Três enfoques na pesquisa em ciências sociais: o positivismo, a fenomenologia e o marxismo. In: **Introdução à pesquisa em ciências sociais**. São Paulo: Atlas, 1987.

UNIVERSIDADE FEDERAL DO PARANÁ - UFPR. **A mais antiga do Brasil**, Disponível em: <http://www.ufpr.br/portalufpr/a-mais-antiga-do-brasil/>. Acesso em: 24 out. 2021

VACCAREZZA, Leonardo Silvio. Ciencia, tecnología y sociedad: el estado de la cuestión em America Latina. **Revist@ do Observatorio do Movimento pela Tecnologia Social da América Latina**, Belo Horizonte, v.1, n,1, p.42-64, jul., 2011.

VAN BELLEN, Hans Michael. **Indicadores de Sustentabilidade: uma análise comparativa**. 2. ed. Rio de Janeiro: FGV Editora, 2007.

VERGARA, S. C. **Projetos e relatórios de pesquisa em administração**. 12. ed. São Paulo: Atlas, 2010.

VITAL, Luciene Paula; CAFÉ, Lígia Maria Arruda. Ontologias e taxonomias: diferenças. **Perspectivas de Ciência da Informação**, Belo Horizonte, v. 16, n. 2, p. 115-130, 2011. Disponível em: <https://www.scielo.br/j/pci/a/VxhxGfW6zfHpfTswwPM5wXP/abstract/?lang=pt>. Acesso em: 14 nov. 2021.

WALSHAM, Geoff. Actor-network theory and IS research: current status and future prospects. In: LEE, A.; LIEBENAU, J.; DEGROSS, J. (eds.) **Information systems and qualitative research**. Boston, MA: IFIP – The International Federation for Information Processing, 1997. DOI:10.1007/978-0-387-35309-8_23. Disponível em: <https://www.semanticscholar.org/paper/Actor-network-theory-and-IS-research%3A-current-and-Walsham/1e963eb52b2d5f65003093944c77080a6e632ba8>. Acesso em: 02 out. 2020.

WALTZ, Carolyn Feher; STRICKLAND, Ora L.; LENZ, Elizabeth R. **Measurement in Nursing Research**. 2nd ed. Philadelphia: F.A. Davis Company, 1991.

WEB OF SCIENCE, 2020. Disponível em: http://images.Webofknowledge.com/WOKRS535R52/help/pt_BR/WOS/hp_search.html. Acesso em 02 fev. 2020.

WILSON, T.D. Review of big data: a very short introduction. **IR Information Research**, Boras, v. 22, n. 4, 2017.

WIMMER, Hayden; YOON, V.; RADA, Roy. Integrating Knowledge sources: an ontological approach. **International Journal of Knowledge Management**, Hershey, PA, v. 9, n. 1, p. 60-75, 2013.

WORLD WIDE WEB- W3C. **Standards**. Disponível em: <http://www.w3.org/standards/>. Acesso em 13. mar. 2021.

WRIGLY, Stuart N.; GARCIA-CASTRO, Raúl; NIXON, Lyndon. Semantic Evaluation At Large Scale (SEALS). OntoSearch: an ontology search engine. WWW 2012 – European Projects Track. **Lyon**, France, p. 16-20, April 2012. Disponível em: https://www.researchgate.net/publication/262390777_Semantic_evaluation_at_large_scale_SEALS. Acesso em: 14 out. 2021.

ZARRI, Gian Piero. Conceptual and content-based annotation of (multimedia) documents. **Multimedia Tools and Applications**, Dordrecht, v. 72, n. 3, p. 2359-2391, Oct 2014. Disponível em: <https://dl.acm.org/doi/10.1007/s11042-013-1463-3>. Acesso em: 14 out. 2020.

ZERARGA, Louffi; DJOUADI, Yassine. A many-sorted theory proposal for information retrieval: axiomatization and semantics. **Knowledge and Information Systems**; London, v. 55, n. 1, p.113-139, Apr 2018. Disponível em: <https://link.springer.com/article/10.1007/s10115-017-1074-9>. Acesso em: 14 out. 2020.

ZATSIORSKI. Vladimir. **Metrología deportiva**. Habana: Editorial Pueblo y Educación, 1989.

ZHANG, Yi; VASCONCELOS, Wamberto; SLEEMAN, Derek. OntoSearch: an ontology search engine. In: RESEARCH AND DEVELOPMENT IN INTELLIGENT SYSTEMS 21., Springer, 2005. **Proceedings...** Springer, 2005. p. 58-69. Disponível em: https://link.springer.com/chapter/10.1007/1-84628-102-4_5. Acesso em: 14 out. 2020.

APÊNDICE 1 – ANÁLISE SISTEMÁTICA DA LITERATURA

TÍTULO / AUTOR / ANO	TEORIA DE BASE USADA	OBJETIVO	MÉTODOS E PROCEDIMENTOS DE PESQUISA	PRINCIPAIS RESULTADOS
Notes on the theory of the actor network. (LAW, 1992)	Teoria ator-rede	Destacar as relações sociais, incluindo poder e organização, como efeitos de rede heterogêneas	Pesquisa qualitativa e documental	Descreve a teoria ator-rede como teoria relacional e sociologia orientada a processos que trata agentes, organizações e dispositivos como efeitos interativos. Sociedade é formada por redes padronizadas de materiais heterogêneos, compostas não só de pessoas, mas também de máquinas, animais, textos, dinheiro, arquiteturas enfim, de outro qualquer matéria
Understanding, building and using ontologism. (GUARINO, 1997)	Ontologia Representação do conhecimento	Definição, construção e reuso de ontologias	Qualitativa e quantitativa. Revisão de literatura	Conceitua a ontologia como sendo um artefato de engenharia para recuperação de informação; Apresenta classificação da ontologia de acordo com o tipo de estrutura e assunto da conceitualização. Sugere a construção de ontologias a partir de teorias, metodologias e ferramentas que induzam e organizem o conhecimento de forma reutilizável e transparente
Ontology learning for the Semantic Web. (MAEDCHE; STAAB, 2001)	Ontologia Web Semântica	Construção de ontologia para Web Semântica e portais semânticos	Qualitativa e quantitativa Revisão de literatura e apresentação de modelos	Apresenta o editor de ontologias OntoEdit para construção de ontologia em quatro estágios: importação e reuso extração, poda e refinamento. Apresenta editor específico cujas características precisam ser comparadas com outros mais recentes
The Semantic Web... BERNERS-LEE; HANDLER; LASSILLA (2001)	Ontologia Web Semântica	Estruturação da Web para inserção de significado nos documentos visando à leitura destes por agentes computacionais inteligentes	Qualitativos Apresentação de modelo de estruturação da Web	A Web Semântica possibilita aumentar a eficiência das buscas e facilitar a recuperação de informações relevantes para o usuário; Usa ontologia formal como sistema de recuperação de informação A estrutura de Web Semântica necessita tecnologias específicas. As últimas camadas da Web Semântica estejam bem desenvolvidas e consolidadas ainda estão em desenvolvimento.
Text mining-based semantic Web ... (ALOMRAN, 2014).	Ontologia Web semântica	Construção de um sistema de um sistema de recuperação da informação a partir de ontologia para itens de e-learning para portal acadêmico.	Qualitativos Apresentação de modelo de estruturação semântico da Web	A mineração de texto é usada com o modelo proposto para melhorar o processamento de dados não estruturados disponíveis nos formatos XML e RDF.
Ontology-based feature construction (NI, 2015)	Ontologia Web Semântica	Construção de características baseada em ontologias para a descoberta e o aprendizado de novas hipóteses e conhecimentos.	Qualitativos. Apresentação de modelo de estruturação semântico da Web	Aplicação de da área médica para construir características interpretáveis e altamente preditivas para a previsão de hospitalização.
Query expansion using conceptual ... (RAZA <i>et al.</i> , 2018)	Ontologia Recuperação da informação	Proposta de framework de expansão de consultas baseado em ontologia no domínio da ciência da computação.	Qualitativos. Apresentação de modelo de consulta baseado em ontologia	Pesquisa em sistemas baseados na correspondência em palavras-chave simples é incapaz de fornecer resultados de acordo com as necessidades de informações do usuário. O protótipo do sistema mostrou precisão média de 89,2% sobre a pesquisa baseada em palavras-chave
Ontology-based query expansion ... (NGO; CAO, 2018)	Ontologia Recuperação da informação	Explorar recursos ontológicos de entidades nomeadas e latentes para revelar a semântica de documentos e consultas	Qualitativos. Modelo de espaço vetorial generalizado baseado em ontologias para a busca de texto semântico	O modelo apresentou boa qualidade no retorno de pesquisas a partir de experiências em um conjunto de dados de referência

(Continua)

TÍTULO / AUTOR / ANO	TEORIA DE BASE USADA	OBJETIVO	MÉTODOS E PROCEDIMENTOS DE PESQUISA	PRINCIPAIS RESULTADOS
Semantics and result disambiguation for keyword ... (Aksoy, 2016)	Recuperação da informação	Utiliza heurística para quantificar a relevância e dissimilaridade dos padrões de resultados de pesquisa com palavras-chave.	Qualitativos Apresenta modelo heurístico para recuperação da informação por palavra-chave	Vantagens e desvantagens da pesquisa por palavra-chave. Análise heurística
Evolution of Knowledge Representation. ... (MALHOTRA; NAIR, 2015)	Ontologia e Recuperação da informação	Fornecer uma referência para futuras pesquisas no campo da representação do conhecimento para fornecer técnicas aprimoradas para a recuperação do conhecimento.	Qualitativo Artigo de revisão sobre representação do conhecimento inicial e técnicas de recuperação	A ontologia como técnicas de recuperação de conhecimento melhora os sistemas de conhecimento existentes
A many-sorted theory proposal for information retrieval ... (ZERARGA; DJOUADI, 2018)	Recuperação da informação	Propuseram uma teoria axiomática recuperar informações	Qualitativos Apresenta modelo de recuperação da informação	Inovação dos sistemas de recuperação da informação pela utilização conjunta de lógica booleana, a fuzzy e a vetorial
KIRA: a system for knowledge-based ... (AMATO et al., 2017)	Ontologia e Recuperação da informação	Apresentaram sistema KIRA - Informação Baseada em conhecimento para recuperação de conteúdo digitais de coleções de arte.	Qualitativos Apresenta modelo de recuperação da informação	Modelos que incluem ontologia para representação semântica com utilização de tecnologias de busca da Web Semântica melhoraram a resposta das buscas em sistemas complexos da Web
Linked data: evolving the Web ... (HEATH; BIZER, 2011)	Dados vinculados Web semântica	Apresentam visão geral dos aplicativos de dados vinculados existentes e as arquiteturas usadas para vincular dados na Web	Qualitativos Revisão sobre dados vinculados e arquiteturas para utilização desses dados	O uso de padrões da Web e um modelo de dados comum possibilitam a implementação de aplicativos que operam em todo o espaço de dados. Essa é a essência dos dados vinculados
Mapping learning object resource metadata ... (THANGSUPACHAI, NIWATTHANAKUL e CHAMNONGSRi, 2015)	Ontologia Web semântica Dados vinculados	Presentaram uma solução baseada em padrões da Web Semântica no qual os bancos de dados relacionais são diretamente mapeados para RDF e OWL	Qualitativos Modelo de mapeamento de dados vinculados e arquiteturas para utilização desses dados	Utilização de ontologia e de tecnologias semânticas RDF, OWL, Sparql para maior precisão semântica da informação de dados vinculados
Semantic representation and enrichment ... (SILVELLO et al, 2017)	Ontologia Web semântica Dados vinculados	Proposta de modelo semântico para recuperação da informação de dados abertos na Web	Qualitativos Modelo semântico para recuperação da informação	A modelagem semântica contribui em tarefas de anotar e interligar dados científicos, melhorando a interpretação, compartilhamento e reutilização dos mesmos

(Continua)

TÍTULO / AUTOR / ANO	TEORIA DE BASE USADA	OBJETIVO	MÉTODOS E PROCEDIMENTOS DE PESQUISA	PRINCIPAIS RESULTADOS
Smart4Job: a big data ... (BENABDERRAH MANE <i>et al.</i> , 2017)	Ontologia Web semântica Big Data	Proposta do sistema Smart4Job para recuperação da informação de Big Data com arquitetura que combina informações na área de recrutamento de pessoal	Qualitativos Modelo semântico para recuperação da informação	Descreve os princípios básicos de dados vinculados incluindo cobertura de aspectos relevantes da arquitetura da web.
Paiva, Costa e Silva (2014)	Portais semânticos Ontologia	Um protótipo de portal concebido e implementado para o domínio do mercado comprador de veículos de concessionária <i>online</i> .	Qualitativos Protótipo de portal semântico	A abordagem híbrida usando sistemas de recomendação baseado no perfil do usuário combinado com o uso de ontologia de domínio. Precisão e recuperação de informações de 86,66% e 68,42%, respectivamente, com melhorias nas pesquisas para os usuários
Semantic Educational Web P (TELNOV, 2015)	Portais semânticos educacionais	Apresenta projeto piloto dedicado a aplicação de gráficos do conhecimento em atividades educacionais das universidades da Rússia	Qualitativos Protótipo de portal semântico	A ontologia do currículo e dos cursos de formação, bem como os dados de autoria, foram usadas para enriquecimento e adaptação dos objetos de aprendizagem. Uso de modelo RDF para navegação visual de dados e consulta SPARQL e serviços RESTful.
Posthumanism and Design. (FORLADNO, 2017)	Sistemas sociotécnicos Teoria Ator-Rede Ontologia	Discussões emergentes sobre pós-humanismo em uma série de disciplinas e perspectivas, e considera exemplos de design emergente práticas que enfatizam as inter-relações entre humanos e não humanos	Qualitativos Revisão sobre design de Design. sistemas sociotécnicos	As considerações sobre o não-humano, coisas e o mundo artificial, exigem novas formas de conhecimento e abrem novos problemas, questões, oportunidades e soluções para o campo de desenho dos objetos
Actor-network theory and IS research ... (WALSHAM, 1997)	De informação e Teoria Ator-Rede	Discorre sobre a teoria interpretativa e a teoria construtivista da TAR	Qualitativos Revisão sobre sistemas de informação e TAR	Na estruturação de sistemas de informação deve se levar em conta a teoria construtivista da TAR que enfatiza as a construção de interações entre humanos e não-humanos
Unscrewing the big leviathan (CALLON, LATOUR, 1981)	Teoria Ator-Rede	Apresenta a macro-ordenação da sociedade e suas relações de poder	Qualitativos Bases da teoria social, para explicar a TAR	Os macroatores têm sucesso traduzindo os interesses de outros atores um único interesse que é extremamente poderoso devido as forças A micro e macro-ordenação de atores os atores definem sua própria escala de crescimento que é refletido no crescimento da ciência e da tecnologia. Os atores formam alianças e aliciam outros atores para fortalecer estas alianças, criando as redes heterogêneas feitas de humanos e de artefatos não humanos em que as associações ocorrem mutuamente.

(Continua)

TÍTULO / AUTOR / ANO	TEORIA DE BASE USADA	OBJETIVO	MÉTODOS E PROCEDIMENTOS DE PESQUISA	PRINCIPAIS RESULTADOS
The sociology of an actor-network (CALLON, 1986a)	Teoria Ator-Rede	Introdução do termo "ator-rede" como uma ferramenta conceitual para "descrever as dinâmicas e estruturas internas de atores-mundos	Qualitativos Bases da Teoria-Ator rede	A tradução inclui quatro fases: 1- Problematização - define-se o problema e possíveis soluções; 2- Interessamento - envolve estratégias e mecanismos para atrair os atores para que aceitem participar da proposta da resolução da questão; 3- Inscrição - estabelece estratégias específicas para convencer os atores a associarem-se aos demais atores da rede (pessoas, tecnologias e normas, dentre outros); 4- mobilização, utiliza métodos para avaliar e garantir que os atores aderiram à proposta sugerida para resolver a controvérsia e se manterão firme no propósito de atingir os interesses da rede.
Iyamu e Mgodlwa (2018)	Teoria Ator-Rede e Big Data Sistemas complexos	Análise de sistemas complexo como os sistemas de informação e de tecnologia a partir da teorias sociais	Qualitativos Bases da Teoria-Ator rede	Um sistema de informação complexo envolve diferentes tecnologias, pessoas, processos, registro de dados e atividades que processam dados. Um sistema de informação é um sistema complexo.
WarVictimSamp o 1914–1922. (RANTALA et al. 2020)	Portal semântico	Apresentaram uma infraestrutura para construção de portais semânticos e dados vinculados usando ontologia no domínio do patrimônio cultural.	Qualitativos Web Semântica e Portais semântico	A estrutura do portal inclui: 1) modelo para a criação e publicação de dados vinculados heterogêneos e distribuídos com base em estrutura de ontologia compartilhada e conhecimento representado por modelo RDF; 2- desenho de interface que possibilita a reutilizados de dados com acesso independente de várias perspectivas de aplicativos mas os dados residem em um único terminal SPARQL; 3) acesso e análise dados a partir de pesquisa semântica facetada, seguida da visualização ou análise de dados por ferramentas disponíveis para uso no portal.
On actor-network theory...(LATOR, (2017)	Teoria Ator-Rede	Apresenta propriedades das redes e clarifica sobre o uso indevido da rede e as conotações que eles implicam	Qualitativos Bases da Teoria-Ator rede	As redes são o ponto final da redução; não há nada além de redes, que são todas contingentes na natureza. Esse autor explica sua abordagem baseando-se nos seguintes temas principais: a) remover a oposição entre distância e proximidade; b) substituir a metáfora da escala pela metáfora da associação, acabando por provocar o colapso da dicotomia micromacro; c) remover a divisão entre o interior e o exterior e a substituição dessa divisão pela imanência das redes. A partir da afirmação ontológica uma rede é um trabalho que é feito por atores que estão agindo, ou entidades que são acionadas.
Actor-network theory ... (EROFEEVA, 2017)	Teoria Ator-Rede	Compara diferentes formas de conceituar o objeto na teoria ator-rede	Qualitativos Bases da Teoria-Ator rede	Os objetos são instituições e projetos em diferentes formas de existência de um objeto, criando-se uma composição instituição-objeto, ou seja, tecnologia embutida na vida coletiva. A instituição e o projeto representam as dimensões actancial e temporal, respectivamente, ou seja, a capacidade de agir enquanto estiver ausente. Os objetos tecnológicos são híbridos (quase-objetos) que produzem outros tipos de objetos (naturais e sociais) e, a partir do padrão relacional da TAR, esses objetos podem ser descritos por diferentes ações dos atores que a eles estão conectados.

(Continua)

TÍTULO / AUTOR / ANO	TEORIA DE BASE USADA	OBJETIVO	MÉTODOS E PROCEDIMENTOS DE PESQUISA	PRINCIPAIS RESULTADOS
The strength of weak ties (GRANOVETTE R, 1983)	Teoria Ator-Rede	Conceitua e explica os laços fortes e fracos das redes	Qualitativos Bases da Teoria-Ator rede	Os laços fortes indicam relações com alto nível de credibilidade e influência nas relações de grupos de indivíduos que participam de um mesmo círculo social. Os laços fracos conectam vários grupos, rompendo a configuração de <i>clusters</i> e funcionando como pontes pelas quais circulam as inovações.
The sociology of complex systems (BURNS, 2006)	Teoria de sistemas aplicados a ciências sociais	Propuseram uma arquitetura básica para portais semânticos	Qualitativos Bases da Teoria de Sistemas	Os sistemas sociotécnicos são caracterizados por incluírem objetos técnicos complexos que interagem e são regulados no campo social. Dependem de estruturas administrativas, investimentos em ciência e tecnologia e são temporariamente estáveis. Os sistemas sociotécnicos são compostos por normativas técnicas e que seguem regras e preceitos como os princípios organizadores sociais destinados a regular as atividades dos atores que operam e gerenciam a tecnologia.
SEAL - a framework for developing semantic portals (MAEDCHE <i>et al.</i> , 2001)	Web Semântica Portais semânticos	Proposta de arquitetura básica para portais semânticos	Qualitativos Modelo de portais semânticos	Plataforma composta por um repositório de dados e um mecanismo de inferência (motores de busca RDF), a partir da qual seria possível a interação entre os agentes de software e usuários em geral que se comunicariam com o sistema por meio de servidor Web.
An ontology-based framework ... (AROUA, E.; MOURAD, 2017)	Web semântica Ontologia	Proposta de método para construção de ontologias baseadas no perfil dos usuários de transporte público	Qualitativos Construção de ontologia	Na construção de ontologias baseadas nas preferências do usuário de sistema de transporte recupera rotas mais adequadas. As preferências explícitas e implícitas do usuário, foram traduzidas em relações hierárquicas e semânticas em conjunto com propriedades relacionadas. O método pode facilitar a descoberta de interesses implícitos e antecipar novos resultados de busca em pesquisas na Web.
Tool for materializing OWL ontologies (REYES-ÁLVAREZ; ROLDÁN-GARCÍA; ALDANA-MONTES, 2019)	Ontologia	Armazenamento de ontologias OWL	Qualitativos Construção e manipulação de ontologia	O armazenamento distribuído dos dados RDF e estruturação de vocabulários facilita a construção de ontologias de domínio e a consulta em banco de dados.
Adding and segmenting educational videos (BORGES; SILVEIRA, 2019)	Web Semântica Portal semântico	Projeto de portal educacional com busca semântica	Qualitativos Web Semântica e Portais semântico	A estrutura de portal para disponibilização de vídeos educacionais possibilitou a pesquisa semântica e anotação de informações pelos usuários.
EIIS: An Educational Information ... HUANG, Chang-Qin, <i>et al</i>	Web Semântica Portal semântico	Projeto de portal educacional com busca semântica	Qualitativos Web Semântica e Portais semântico	A ontologia construída pelo Protegé para o domínio de tecnologia educacional e mecanismo de busca semântica delineado podem ser utilizados em outros portais da área educacional

Fonte: Autoria própria (2021).

APÊNDICE 2 – QUESTIONÁRIO DA PESQUISA

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ



Programa de Pós-Graduação em
Tecnologia e Sociedade

Seção 1 de 7

Pesquisa sobre o Portal da Universidade ✕ ⋮

Este questionário tem como objetivo mapear o uso de tecnologias semânticas e aspectos de administração geral dos portais de universidades federais brasileiras, com ênfase nas áreas gestão da informação, gestão do conhecimento e Web Semântica. Este levantamento faz parte de uma tese do Programa de Pós-Graduação em Tecnologia e Sociedade (PPGTE) da UTFPR.

Título da pesquisa:
Indicadores qualitativos para avaliação de potencial semântico de portais do conhecimento de universidades

O projeto foi aprovado pelo Comitê de Ética, conforme processo CAAE 40011220.8.0000.5547.

Pesquisadores responsáveis:
Josefina Aparecida Soares Guedes (doutoranda)
Profa. Dra. Faimara do Rocio Strauhs (professora orientadora)
E-mail para contato: josefinaguedes@yahoo.com.br

Desde já agradecemos a sua participação!

Após a seção 1 ▼ Continuar para a próxima seção

Seção 2 de 7

Título da seção (opcional) ✕ ⋮

Descrição (opcional)

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO - (TCLE)

Seção 3 de 7

Identificação da Universidade ✕ ⋮

Descrição (opcional)

AMAZENAMENTO DE DADOS - E SUA GESTÃO



Descrição (opcional)

Utiliza alguma tecnologia de padrão aberto na estruturação do portal? *

Sim

Não

Sistema usado para repositório de documentos e navegação pelos dados: *

Arquivos em servidor com hierarquia fixa de classificação e manutenção centralizada

Arquivos em servidor com hierarquia de classificação flexível

Arquivos em servidor com navegação multidimensional com facetar, informações semi estruturadas e ex...

Sistema operacional usado: *

Windows

Linux

Outros...

Banco de dados utilizado: *

Na opção outra é obrigatório especificar.

Oracle

MySQL

PostgreSQL

Outros...

Servidor web utilizado: *

- Apache Tomcat
- Apache/Similar
- Nginx
- Outros...

Sistema de segurança para rede e nuvem *

- Nagios Core
- Snort
- OpenStack
- KeePassXC
- Outros...

Após a seção 4 Continuar para a próxima seção ▼

Seção 5 de 7

GESTÃO DA INFORMAÇÃO



Descrição (opcional)

Sistema usado para pesquisa no portal: *

- Palavra-chave ou texto livre
- Taxonomia / vocabulário controlado
- Ontologia
- Outros...

Forma de gestão do conteúdo: *

- Centralizada
- Descentralizada com possibilidade de múltiplas agregações e visões do mesmo dado
-

Como é feita a edição de itens de informação: *

- Administradores sem qualquer restrições de acesso para edição, publicação e exclusão
- Membros registrados com diferentes permissões para edição e publicação
-

Forma de atualização das informações *

- Adição de informações e anotações
- Adição de novas classificações com expansão da estrutura da informação
-

Forma de processamento de informações *

- Processamento manual
- Processamento automatizado para acesso de agentes inteligentes e facilitar integração entre portais
-

Forma de publicação de informações *

- Manual ou via formulários específicos
- Automática; integração de ontologias; reutilização de formulários em múltiplos portais
-

Sistema de gerenciador de conteúdo *

- Drupal
- Plone
- WordPress
- Joomla
- Textpattern

Outros...

Mecanismo de inferência utilizado *

OKBC basead inference engine

SPARQL Engine

XQuery

Pellet

Outros...

Linguagens de programação e modelo de dados usados *

HTML

XML

PHP

RDF (Modelo de dados)

OWL (Metaontologia)

SWRL

RIF

RuleML

Outros...

A disponibilização da informação no portal segue os requisitos da Lei Geral de Proteção de Dados Pessoais (LGPD)? *

Sim

Não

GESTÃO DE ONTOLOGIA



Descrição (opcional)

Utiliza editor de ontologia *

Não

Sim

Utiliza tecnologias para importação e exportação de dados, com vistas ao reúso de dados e informações? *

Não

Sim.

Possibilita a inclusão de membros com permissões para escrever conceitos, instâncias, atributos e colaborar na construção de ontologias *

Sim

Não

Apresenta descrição de processamento semântico ou estrutura da ontologia com materiais para consulta: *

Sim

Não

Apresenta informação semanticamente completa e consistente *

Sim

Não

Apresenta profundidade de cobertura das informações *

- Sim
- Não

Há controle no uso de vocabulários, evitando-se termos ambíguos que possuem o mesmo significado? *

- Sim
- Não

Há maturidade de implantação/manutenção do portal, ou seja, sem erros ou link quebrado? *

- Sim
- Não

Após a seção 6 Continuar para a próxima seção

Seção 7 de 7

REDE DE ATORES - CARACTERIZAÇÃO SOCIOTÉCNICA



Descrição (opcional)

Identifique os integrantes da equipe de administradores do portal: *

- Conteúdistas (seleção e indicação de conteúdos)
- Gerente de conteúdo (aprovação, atualização exclusão de conteúdos)
- Gerente de banco de dados
- Gerente de segurança de rede e nuvem
- Outros...

Identifique os integrantes da equipe de desenvolvimento/manutenção do portal: *

- Programadores
- Engenheiro do conhecimento
- Engenheiro de ontologia
- Designer/ usabilidade
- Outros...

Indique a participação dos usuários na construção ou melhoramento do portal *

- Participaram da construção
- Participam do melhoramento
- Participaram da construção e continuam participando do melhoramento
- Nenhum tipo de participação

Os usuários dispõem de informações relevantes do domínio específico da comunidade acadêmica: *

- Sim
- Não

Os usuários dispõem de mapas de tópicos pessoais para melhorar a gestão de informação pessoal e eficácia do portal: *

- Sim
- Não

Que recursos para colaboração, comunicação e compartilhamento de informação os usuários dispõem *

Boletins eletrônicos

Lista de discussão

Chat

Forum

Outros...

Os usuários dispõem de mecanismos de ajuda ou mapa do site *

Mapa do Site

Mecanismo de ajuda

Outros...

APÊNDICE 3 – QUESTIONARIO PARA VALIDAÇÃO DE DADOS

 UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

Programa de Pós-Graduação em
Tecnologia e Sociedade

Validação de Indicadores de potencial semântico para portais universitários

Olá.
Para você que participou da pesquisa indicando as ferramentas e tecnologias utilizadas no portal de sua instituição, peço a gentileza de indicar se considera aceitável ou não o conjunto de indicadores propostos

USO DE TECNOLOGIAS DE PADRÃO ABERTO - tecnologias para sistema operacional, servidor Web, banco de dados, sistema de segurança, mecanismo de inferência e gerenciador de conteúdo *

Aceitável

Não aceitável

ARMAZENAMENTO DE DOCUMENTOS EM SERVIDOR, COM INFORMAÇÕES ORGANIZADAS DE FORMA SEMI-ESTRUTURADA E EXTENSÍVEL, COM POSSIBILIDADE ALTERAÇÕES DESCENTRALIZADAS *

Aceitável

Não aceitável

USO DE ONTOLOGIA COMO SISTEMA DE RECUPERAÇÃO DA INFORMAÇÃO - Mediante vocabulário controlado e formal, possibilita explicitar e enriquecer os dados com semântica, facilitando a interpretação/entendimento sobre dados/informações e maior precisão nos resultados de busca *

Aceitável

Não aceitável

NAVEGAÇÃO POR FACETAS - Navegação pelas categorias fundamentais de um assunto representado pelos componentes da ontologia *

- Aceitável
- Não aceitável

GESTÃO DESCENTRALIZADA DE CONTEÚDOS - Conteúdo armazenado e manipulado por uma rede descentralizada de organizações provedoras e indivíduos, o que torna possível múltiplas agregações e visões de um mesmo dado *

- Aceitável
- Não aceitável

MEMBROS REGISTRADOS COM DIFERENTES PERMISSÕES PARA MANIPULAÇÃO DE DADOS E INFORMAÇÕES *

- Aceitável
- Não aceitável

ATUALIZAÇÃO DE INFORMAÇÕES POR ADIÇÃO DE NOVAS CLASSIFICAÇÕES E EXPANSÃO DA ESTRUTURA DA INFORMAÇÃO *

- Aceitável
- Não aceitável

PROCESSAMENTO E PUBLICAÇÃO AUTOMÁTICA DE INFORMAÇÕES - Para acesso de agentes inteligentes, integração de portais, integração e o reuso de informações e ontologias *

- Aceitável
- Não aceitável

...

USO DE GERENCIADOR DE CONTEÚDO - Para automatizar tarefas de criação, seleção, filtro e categorização dos dados armazenados *

- Aceitável
- Não aceitável

USO DE MECANISMO DE INFERÊNCIA - Para otimizar das respostas de busca, mediante navegação interativa por facetar, com visualização das relações existentes entre as informações *

- Aceitável
- Não aceitável

USO DE LINGUAGENS DE PROGRAMAÇÃO SEMÂNTICA, MODELAGEM DE DADOS RDF E METAONTOLOGIA OWL - Ferramentas e modelagem de dados para disponibilização de pesquisa semântica *

- Aceitável
- Não aceitável

USO DE EDITOR DE ONTOLOGIA - Para manipulação e atualização da ontologia utilizada *

- Aceitável
- Não aceitável

DISPONIBILIZAÇÃO DE MATERIAIS SOBRE DESCRIÇÃO DO PROCESSAMENTO SEMÂNTICO OU ESTRUTURA DA ONTOLOGIA USADA *

- Aceitável
- Não aceitável

DISPONIBILIZAÇÃO DE RECURSOS PARA IMPORTAÇÃO, EXPORTAÇÃO E COLABORAÇÃO PARA REUSO DE INFORMAÇÕES *

- Aceitável
- Não aceitável

COBERTURA, PRECISÃO, COMPLETUDE E CONSISTÊNCIA DAS INFORMAÇÕES *

- Aceitável
- Não aceitável

EQUIPE ESPECIALIZADA PARA GESTÃO DO PORTAL - Conteudista, gerente de conteúdo, gerente de banco de dados, gerente de segurança de dados, programadores, engenheiro de ontologia, engenheiro do conhecimento *

- Aceitável
- Não aceitável
-

PARTICIPAÇÃO CONTÍNUA DOS USUÁRIOS NA CONSTRUÇÃO E MELHORAMENTO DO PORTAL *

- Aceitável
- Não aceitável
-

DISPONIBILIZAÇÃO DE RECURSOS E SERVIÇOS DE SUPORTE DE COMUNICAÇÃO E COMPARTILHAMENTO (mecanismo de ajuda, mapa do site, boletins eletrônicos, chat, fórum, lista de discussão e outros) *

- Aceitável
- Não aceitável

APÊNDICE 4 – ÍNDICE ONOSMÁTICO

A

Abdullar e Zamil (2018), 60, 143
 ABDULLAR; ZAMIL, 2018, 35, 41, 55
 Aksoy (2016), 56, 143, 175, 226
 AKSOY, 2016, 174
 AL BALUSHI; SAMPAIO; LOUCOPOULOS,
 2013, 35
 Albagli (2004), 33
 ALBAGLI, 2004, 48, 49
 Alomran (2014), 76
 Alomran, 2014, 111
 Alvarenga, (1998), 125
 Amato *et al.* (2017), 77, 140, 143
 AMATO *et al.* 2017, 37
 AMATO *et al.*, 2017, 32, 36, 39, 69, 111, 114
 AMBROZEWICZ, 2015, 98
 Andrade, Vaitsman e Farias (2010), 148
 Andrade; Vaitsman; Farias (2010), 149
 Andrade; Vaitsman; Farias (2010), 149
 ANDRADE; VAITSMAN; FARIAS, 2010, 32, 38
 Aroua e Mourad (2017), 69
 AROUA; MOURAD, 2017, 63, 190
 ASPURA; NOAH, 2017, 63
 Azani (2009), 138, 145
 AZANI, 2009, 36, 165
 AZMANI; EL HARZLI, 2014, 184

B

BAEZA-YATES; RIBEIRO-NETO, 1999, 59, 60
 BAHAFID *et al.*, 2015, 63
 BALDIN; MUNHOZ, 2011, 127
 Barba-González *et al.* (2019), 78, 140, 147
 Barba-González *et al.*, 2019, 111
 BARBA-GONZÁLEZ *et al.*, 2019, 56, 62, 174
 Bardin (1977), 127
 Bardin (2011), 50, 52, 129, 149
 Bardin (2011),, 129, 149

BARDIN, 1977, 127
 BARDIN, 2011, 129, 130, 132
 Bellandi *et al.* (2012), 143, 145
 BELLANDI *et al.*, 2012, 36, 37, 38, 41, 42, 43,
 179
 BELLANDI *et al.*, 2012, 37
 BENABDERRAHMANE *et al.* 2017, 35
 Benabderrahmane *et al.* (2017), 77, 11, 140,
 147, 227
 BENABDERRAHMANE *et al.* 2017, 39, 69
 Berners-Lee, Hendler e Lassila (2001), 33, 49,
 97 99, 126, 142, 145, 146, 155, 157,
 BERNERS-LEE; HENDLER; LASSILA, 2001,
 30, 37, 38, 41, 43, 52, 66, 67, 80, 99, 111,
 134, 139, 140, 147, 154, 160, 165, 166,
 179
 BIERNACKI; WALDORF, 1981, 127
 Bijker, Hugh e Pinch (1989), 110
 Bijker, Hugh; Pinch (1989), 144
 BIJKER; HUGH; PINCH, 1989, 37, 114, 188
 Bijker; Hughes; Pinch (1989), 48
 BIJKER; HUGHS; PINCH, 1989, 113
 BIREME, 2022, 46, 58
 BIZER, HEATH; BERNERS-LEE, 2009, 85,
 180
 BIZER; HEATH, BERNERS-LEE, 2009, 160
 BIZER; HEATH; BERNERS-LEE, 2009, 73,
 165
 BONNEFOY, 2005, 98
 Borges e Silveira (2019), 96, 229
 BORGES; SILVEIRA, 2019, 97
 Borgman e Furner (2002), 125
 Börner *et al.* (2012), 139
 Börner *et al.* (2012);, 147
 BÖRNER *et al.*, 2012, 87, 174, 184
 Börner *et al.*, 2012), 145
 BRAGA, 2005, 61
 Brandt, Vidotti ; Santarem Segundo, 2018, 111

Brandt, Vidotti e Santarem Segundo (2018), 74
 BRANDT; VIDOTTI; SANTAREM SEGUNDO,
 2018, 73, 138
 BRASIL, 2010, 98
 Bruno Latour (2012), 31
 Burns (2006), 34, 102, 103, 141, 144, 148
 BURNS (2006), 50
 Burns e Dietz (2001), 102, 103
 Burns(2006), 143
 BURNS, 2006, 30, 37, 39, 42, 43, 48, 53, 102,
 104, 111, 113
 Burns; Dietz (2001), 141, 144, 148

C

Callon (1986 a, b, c), 34, 50
 Callon (1986 c), 48
 Callon (1986), 109, 126, 141, 146, 158
 Callon (1986a), 109
 Callon (1986a, b, c), 144
 Callon (1986a,b,c), 126, 145, 148
 Callon e Latour (1981), 104, 105, 155, 227
 CALLON, 1986, 30, 37, 42, 43, 44, 53, 105,
 107, 111, 166
 CALLON, 1986a,b,c, 113
 CALLON, 1986c, 30, 44
 CALLON, 1989, 32
 CALLON,1986, 107, 160
 CALLON,1986a,b,c, 160
 Callon; Latour (1981), 146, 148
 CALLON; LATOUR, 1981, 108, 134, 154, 155,
 161, 162, 168, 188
 CALLON; LATOUR,1981, 158
 CALLON; LAW; RIP, 1986, 52
 Callon e Latour (1981),, 113
 Campos; Gomes (2008), 143
 CAO, 2005, 108
 CLUNIS, 2018, 63
 Cordella (2010), 30, 108, 146
 CORDELLA, 2010, 158
 Correa (2012), 50

Correa, 2012, 111
 CORREA, 2012, 36
 CORREIA, 2012, 80, 173, 174, 177
 Costa e Silva (2014), 89, 166, 227
 COSTA; SILVA, 2014, 37
 Crespo Borges (2007), 151, 170
 Creswell e Clark (2013), 50, 51, 53, 116
 Creswell e Clark (2013),, 50, 51, 53, 116
 Cruz et al. (2019), 91

D

DATE, 2004, 71
 Davis (1992), 151
 DONABEDIAN, 1985, 151
 DUARTE, HARA, 2018, 166
 DUARTE; HARA, 2018, 87
 DUMER; ALBUQUERQUE, 2021, 64

E

El Hajj; Azmani; El Harzli (2014), 140
 EL HAJI; AZMANI; EL HARZLI, 2014, 85, 87,
 165
 Elbanna (2012), 30, 108, 146
 ELBANNA, 2012, 156, 158
 Elmasri; Navathe (2011), 138
 ELMASRI; NAVATHE, 2011, 86, 160
 Erofeeva (2017), 30, 107, 144, 228

F

Fahad *et al.* (2011), 91
 Fahad *et al.*, 2011, 111
 FAHAD *et al.*, 2011, 89, 91, 134
 FAM; NORISZURA; SHINYIE; 2019, 75
 FEENBERG, 1999, 44
 Fernández *et al.* (2011), 143
 FERNÁNDEZ *et al.*, 2011, 29, 31, 32, 36, 37,
 42, 53, 59
 Ferneda (2003), 143
 Ferneda (2006), 61
 FERNEDA, 2003, 57, 58

FERNEDA, 2013, 61
 Forlano (2017), 144
 FORLANO, 2017, 28, 33, 34, 53, 103, 104
 Frosterus; Hyvonen e Laitio (2011), 89
 Frosterus; Hyvonen; Laitio, 2011, 111

G

Gao (2005), 42, 106, 113, 228
 GARGIULO *et al.*, 2014, 29, 35
 Gargiulo *et al.*, 2014, 111
 GARGIULO *et al.*, 2014, 30, 76
 Gavrilova (2011), 141, 143, 148, 165
 Gavrilova, 2011, 111
 GAVRILOVA, 2011, 37, 41, 53, 57, 80
 Gil (2010), 50, 51
 GIL, 2010, 50, 51, 116
 GIRASE; PATANAİK; PATIL, 2016, 63
 Granovetter (1983), 49, 110, 141, 144, 148, 200
 GRANOVETTER, 1983, 110, 113
 Gruber (1993), 33, 49, 143, 145
 GRUBER, 1993, 52, 56, 62, 99, 175
 Gruber, 1996, 111
 GRUBER, 1996, 53, 62, 174, 177
 Guarino (1997), 33, 49, 145
 Guarino (1998), 63, 126, 143
 GUARINO, 1997, 37, 52, 53, 62, 99, 128, 174, 225
 GUARINO, 1998, 63
 Guedes (2015), 38, 40, 50, 84, 121, 123, 137, 142, 147, 203
 Guedes e Strauhs (2016), 99, 100, 148, 156, 157
 Guedes e Strauhs (2016), 99, 156
 Guedes e Strauhs (2016), 149
 GUEDES, 2015, 39
 GUEDES; STRAUHS 2016, 154
 GUERTIN; BERNHARDT, 2005, 127
 Guo *et al.*, 2018, 111

GUO *et al.*, 2018, 37, 38, 41, 55, 56, 60, 62, 76, 174

H

Hauke, Owoc e Pondel (2015), 29, 79, 80
 HAUKE, OWOC; PONDEL, 2015, 79
 HAWKE *et al.*, 2013, 67
 Hawke *et al.* (2013), 139, 143, 147
 HAWKE *et al.*, 2013, 38, 41
 HAWKE *et al.*, 2013, 30, 31, 66
 Heath; Bizer (2011), 143
 Heath; Bizer, 2011, 111
 HEATH; BIZER, 2011, 72
 Hu (2011), 109, 146
 HU, 2011, 30
 Huan; Maidment; Tian (2011), 141, 148
 Huan; Maidment; Tian, 2011, 111
 HUAN; MAIDMENT; TIAN, 2011, 80
 Huang *et al.* (2011), 94
 Huang *et al.*, 2011, 111
 HUANG *et al.*, 2011, 94, 95
 HUANG; MAIDMENT; TIAN, 2011, 29
 Hughes (1989), 34, 102, 114, 126, 141, 143, 144, 148
 HUGHES, 1989, 42, 111, 113, 188
 Hyvönen *et al.* (2021), 52, 91

I

IBGE, 2021, 153
 Iyamu e Mgudlwa (2018), 103, 105, 112, 155
 Iyamu; Mgudlwa (2018), 141, 144, 146, 148
 IYAMU; MGUDLWA, 2018, 28, 33, 53, 103, 104, 105, 114, 163, 188
 IYAMU; MGUDLWA, 2018), 104
 IYAMU; MGUDLWA, 2018)., 28, 188

J

John Law (2010), 31
 Jorge (2005), 92
 JORGE, 2005, 93, 177

Jorge, 2005, 111

K

Koivunen e Miller (2001), 49, 67, 146

Koivunen; Miller (2001), 139, 147

Koivunen; Miller (2001), 145

KOIVUNEN; MILLER 2001, 73, 165

KOIVUNEN; MILLER, 2001, 179, 180

L

Lachtim *et al.* (2009), 93, 94

Lachtim; Moura e Cavalcanti (2009),, 167

Lachtim; Moura; Cavalcanti (2009), 141, 148

LACHTIM; MOURA; CAVALCANTI, 2009, 162,
166, 176, 186

BERNERS-LEE, HENDLER; LASSILA, 2001;, 63

LAST, 1998, 149, 150

LAST, 1998),, 150

Latour (1996), 32

Latour (2000), 30, 34, 43, 48, 103, 106, 145,
156

Latour (2000),, 48, 106

Latour (2001), 107

Latour (2001), 31, 50, 107

Latour (2001), 107

Latour (2012), 106, 107, 141, 144, 146, 148,
228

Latour (2017), 106, 109, 144, 146

LATOUR, 1992, 108, 158

LATOUR, 1994, 44, 53

LATOUR, 2000, 31, 42, 43, 105, 111, 165

LATOUR, 2001, 33, 34, 107

LATOUR, 2012, 44, 105

LAUSEN *et al.*, 2005, 97

Lausen et al. (2005), 50, 82, 126, 138, 139,
140, 141, 142, 143, 145, 146, 147, 148

Lausen *et al.*, (2005), 99, 157

LAUSEN *et al.*, 2005, 31, 42, 43, 79, 80, 93,
154, 166, 173, 174, 176, 177, 178

Lausen *et al.*, 2005;, 111

Law (1992), 50, 105, 126, 141, 144, 145, 146,
148

LAW, 1992, 128, 163, 225

LAW, 1992, 28, 104, 134, 188

LAW, 2010, 37, 44

(LAW, 1992, 128, 163, 225

LAW,1992, 161

Lawan; Rakib (2019, 145

LAWAN; RAKIB, 2019, 52, 82, 85, 166, 180

Lee *et al.* (2018), 143

Lee *et al.*, (2018), 146

LEE *et al.*, 2018, 37, 41, 53, 55, 64, 65, 175

LEE *et al.*, 2018, 55, 64

LEE, *et al.*, 2018, 35, 41

Li *et al.* (2017), 147

Li *et al.* (2017), 140, 143, 147

Li *et al.*, 2017, 111

LI *et al.*, 2017, 31, 33, 35, 36, 38, 39, 41, 42,
61, 66, 69, 114, 174, 188

Li *et al.*, 2017, 146

LINSINGEN; BAZZO; PEREIRA, 2003, 28,
188

LOPES, 2002, 58

M

MAATOUK, 2021, 33, 34, 179, 184

Macome (2008), 109, 144, 146

MACOME, 2008, 30

Maedche e Staab (2001), 49, 81, 126, 146,
147

Maedche *et al.* (2001), 81, 140, 147

Maedche *et al.*, 111

MAEDCHE *et al.*, 2001, 99

Maedche; Staab (2001), 140, 147

MAEDCHE; STAAB, 2001, 53, 82, 154, 165

MAIA; SERAFIM, 2011, 107

MALGAONKA *et al.*, 2016, 33, 35, 38

MALGAONKA; DEVALE, 2016, 41, 66

Malgaonkar e Devale (2016), 77

MALGAONKAR; DEVALE , 2016, 36, 188
 Malgaonkar; Devale, 2016, 111
 MALGAONKAR; DEVALE, 2016, 31, 42, 61,
 174
 Malhotra (2019), 150
 Malhotra; Nair (2015), 143
 Malhotra; Nair (2015), 140, 145, 147
 MALHOTRA; NAIR, 2015, 39, 57, 65, 175,
 188, 226
 Maniega-Legarda, Pérez-Salmerón e
 Guerrero-Torres (2008), 55, 145
 MANIEGA-LEGARDA; PÉREZ-SALMERÓN;
 GUERRERO-TORRES, 2008, 158, 159
 Marconi e Lakatos (2010, 50, 135, 170
 Marconi e Lakatos (2010b), 116
 MARCONI; LAKATOS, 2010, 50
 MARCONI; LAKATOS, 2010a, 116
 MARCONI; LAKATOS, 2010b, 136
 Marcuse (1998), 48
 MARCUSE, 1998, 28, 188
 MARCUSSE, 1998, 44
 MARICATO, 2011, 122, 125
 MAROUF; KHALIL, 2015, 30
 MARTINEZ-CRUZ; BLANCO; VILA, 2012, 85
 Martínez-García e Hernández-Lemus (2013),
 101, 144
 MARTÍNEZ-GARCÍA; HERNÁNDEZ-LEMUS,
 2013, 47, 102, 113, 116, 134, 188
 MARTINS; SOARES; FRANCELIN, 2012, 37
 MARTINS; SORAES; FRANCELIN, 2012, 36
 Masner *et al.* (2019), 138, 139, 145, 146, 147
 Masner *et al.* (2019), 138
 MASNER *et al.*, 2019, 86, 87, 134, 160, 179
 MASSALLI, 2019, 153
 MASSALLI, 2019;, 34
 Meireles e Cedón (2010), 127, 128
 MEIRELES; CEDÓN 2010, 127
 Michel Callon (1986), 31
 MINAYO, 2009, 98
 MIOR, 2006, 49

MOHER *et al.*, 2009, 122
 MOREIRA; CALEFE, 2006, 50
 MOREIRA; CALEFFE, 2006, 127
 MOREIRA; CALEFFE, 2006), 52
 MORIN, 2014, 36
 MORRISON, 1968, 36
 Moura *et al.* (2015), 70, 143
 MOURA *et al.* 2015, 36
 MOURA *et al.*, 2015, 72, 73

N

Nadal *et al.* (2019), 77
 Nadal *et al.*, 2019, 111
 Nazário (2013), 91
 Nazário, 2013, 111
 NAZÁRIO, 2013, 91
 Ngo e Cao (2018), 55, 59, 175
 NGO, CAO, 2018, 60
 Ngo; Cao (2018), 141, 143
 Nguyen *et al.* (2020), 145, 146, 147
 NGUYEN *et al.*, 2020, 86, 160, 165, 170, 171
 Ni (2015), 76, 225
 NIELSEN, 2000, 112, 164
 NIELSEN; HORANDER, 2007, 158
 NIELSEN; LOHANGER, 2007, 161, 162
 NIELSEN; LORANDER, 2007, 112, 164, 162,
 187
 NIMBIX, 2021, 75
 NONAKA; TAKEUCHI, 1997, 28, 30
 NOVAES; DAGNINO, 2004, 28
 NOY; GUINNESS, 2005, 62, 177
 NOY; MCGUINNESS (2005), 70

O

OCDE - Organização para a Cooperação e o
 Desenvolvimento Económico (2021), 98
 Open Knowledge Internacional (2019), 71
 OPEN KNOWLEDGE INTERNACIONAL,
 2019, 71, 86, 138, 160, 170, 171
 Open knowledge international (2019), 147

Open Knowledge International (2019), 138,
145

P

Paiva; Costa; Silva (2014), 140, 147

PAIVA; COSTA; SILVA, 2014, 43, 81, 133

Pattueli e Miller (2015), 73

Pattueli; Miller, 2015, 111

PATTUELLI; MILLER; 2015, 73

PEDRO, 2010, 30, 42, 107

PINHEIRO, 2004, 59

POLANYI, 1966, 30

Prado e Baranauskas (2012), 108, 109, 144

Prado; Baranauskas (2012), 146

PRADO; BARANAUSKAS, 2012, 156, 158,
188

PROGRAMA DE PÓS-GRADUAÇÃO EM
TECNOLOGIA E SOCIEDADE, 2021, 47

R

Rantala *et al.* (2020), 83

RAUTENBERG *et al.*, 2010, 88

RAUTENBERG; TODESCO; STEIL, 2010, 52

Raza *et al.* (2018), 56, 140, 147

RAZA *et al.* 2018, 60

Raza *et al.*, 2018, 111

RAZA *et al.*, 2018, 55, 56, 57

Reyes-Álvarez, Roldán-García e Aldana-
Montes (2019), 72

REYES-ÁLVAREZ, ROLDÁN-GARCÍA;
ALDANA-MONTES, 2019, 71

Reyes-Álvarez; Roldán-García, 140

Reyes-Álvarez; Roldán-García, e Aldana-
Montes (2019), 74, 229

Reyes-Álvarez; Roldán-García; Aldana-Montes
(2019), 147

Reyes-Álvarez; Roldán-García; Aldana-
Montes, 2019, 111

Reyes-Álvarez; Roldán-García; Ldana-Montes
(2019), 138, 147

Reynolds, Shabajee e Cayzer (2004),, 49

Reynolds; Shabajee; Cayzer, 2004, 111

REYNOLDS; SHABAJEE; CAYZER, 2004, 80,
173, 174

REYNOLDS; SHABAJEE; CAYZER, 2004), 80

RODRIGUES, 2020, 55, 60

Rolim *et al.* (2019), 83

Rose e Scheepers (2001, p. 217), 104

Rose; Scheepers (2001), 144

ROSENFELD, MORVILLE; ARANGO, 2015,
86

Rosenfeld; Morville; Arango (2015), 138, 145

ROSENFELD; MORVILLE; ARANGO, 2015,
112, 158, 160, 161, 162

RUBIO *et al.*, 2003, 151

S

SACK, 1986, 49

SALES; CAFÉ, 2008, 63, 64, 65, 133

SALTON, 1968, 57

Sampieri, Collado e Lucio (2010), 36, 39,99

SCHLEICHER, 2006, 97, 104

Seneviratne, Fernando e Karunaratne (2019),
70

Shaanika e Iyamu (2015), 106, 112

SHARMA, 2016, 63

SHAW; SIMÕES, 1999, 60

Shilakes e Tylmann (1998),, 79

Shim; Shin (2016);, 146

SHIM; SHIN, 2016, 42, 106, 188

Shirisha; Subbarao; Kavitha, (2015), 143

SHIRISHA; SUBBARAO; KAVITHA, 2015, 31,
37, 39

PAIVA, COSTA, 63

SILVA; SANTOS; FERNEDA, 2013, 33, 58,
59, 60

Silvello *et al.* (2017), 74, 140, 143, 147

SILVELLO *et al.*, 2017, 31, 33, 35, 36, 37, 38,
39, 41, 43, 55, 69, 73, 75, 76

SILVELLO *et al.*, 2017), 31, 33, 36, 37, 43, 55,
63, 69, 73, 75
Silvello *et al.*,(2017), 146
Silvello *et al.*, 2017, 111
SIRISHA; SUBBARAO; KAVITHA, 2015, 66,
133
SOARES; PICOLI; CASAGRANDE, 2018, 118,
119
SOUZA, MIRANDA; SOUZA; 2019, 34, 153
Spyns *et al.* (2002),, 92
SPYNS *et al.*, 2002, 92
Stallings (2015), 138
STALLINGS, 2015, 86, 160
Stanescu (2018), 140, 146, 148
STANESCU, 2018, 52, 84, 133, 165
Staniscu (2018), 139, 147
STANISCU, 2018, 88
STRAUHS *et al.*, 2012, 29, 30
SWAD, 2021, 89, 90

T

Ta e Thi (2015a), 70
Ta e Thi (2015b), 70
TA; THI, 2015, 62
TA; THI, 2015a, 62
TAHMASEBI *et al.*, 2019, 63
TATTO; BORDIN, 2016, 30
Telnov (2015), 95, 143
Telnov (2015),, 65, 95
Telnov, 2015, 111
TELNOV, 2015, 37, 56, 62, 65, 85, 99, 175,
181, 227
TELNOV, 2015;, 56
Terziyan, Golovianko e Shevchenko (2015), 95
TERZIYAN; GOLOVIANKO; SHEVCHENKO,
2015, 104
TERZIYAN; GOLOVIANKO; SHEVCHENKO,
2015, 97
Thangsupachai, Niwatthanakul e Chamnongsri
(2015), 72

Thangsupachai, Niwatthanakul; Chamnongsri
(2015), 147
Thangsupachai; Niwatthanakul; Chamnongsri
(2015), 138
Thangsupachai; Niwatthanakul; Chamnongsri,
2015, 111
Thomas (2008), 143
THOMAS (2008), 50
Thomas (2008),, 34
THOMAS, 2008, 39, 42, 53
Thurstone (1928), 98
TRIOLA, 1999, 135
TRIOLA, 2017, 142
Triviños (1987), 51

U

UFPR, 2021, 34, 153
Universidade Tecnológica Federal do Paraná,
12

V

VACCAREZZA, 2011, 28, 188
Van Bellen (2005), 38, 149
VERGARA, 2010, 135
VITAL; CAFÉ, 2011, 32

W

W3C (2019), 138
W3C, 2019, 29, 97, 159, 165, 180
W3W, 2019, 97
Walsham (1997), 144
WALSHAM, 1997, 106, 155
WALSHAM,1997, 47
WALZ, STRICKLAND, LENZ, 1991, 151
WEB OF SCIENCE, 2020, 46, 121, 223
Wilson (2017), 75
Wimmer, Yoon; Rada, 2013, 111
Winner, Yoon e Rada (2013), 166
WINNER; YOON; RADA, 2013, 29, 38, 41, 80

WORLD WIDE WEB - W3C, 29, 35, 65, 66,
74, 85, 86, 87, 88, 97, 104, 134, 159, 165,
178

World Wide Web (2019), 139, 147

WORLD WIDE WEB, 2019, 35, 86, 88, 134,
178, 179

WRIGLY, GARCIA-CASTRO, NIXON, 2012,
86, 87, 165

Z

Zarri (2014), 143

ZARRI, 2014, 64

Zatsiorski (1989), 149, 151

Zerarga e Djouadi (2018), 60, 61

Zerarga; Djouadi (2018), 143

ZERARGA; DJOUADI, 2018, 57

ZERARGA; DJOUADI, 2018, p.15, 57

Zhang; Vasconcelos; Sleeman, 2005, 111

ZHANG; VASCONCELOS; SLEEMAN, 2005,
90