



UNIVERSIDADE DE BRASÍLIA
FACULDADE DE CIÊNCIA DA INFORMAÇÃO
GRADUAÇÃO EM BIBLIOTECONOMIA

RAYRA SILVA DE ABREU

METADADOS

contribuições no campo da *web* semântica

RAYRA SILVA DE ABREU

METADADOS

contribuições no campo da *web* semântica

Monografia apresentada como requisito para a obtenção do título de Bacharel em Biblioteconomia pela Faculdade de Ciência da Informação (FCI), da Universidade de Brasília (UnB).

Orientador: Professor Dr. Márcio Bezerra da Silva

Brasília
2023

Dados Internacionais de Catalogação na Publicação (CIP)

A162 Abreu, Rayra Silva de.

Metadados : contribuições no campo da *web* semântica /
Rayra Silva de Abreu, orientador Márcio Bezerra da Silva. --
Brasília, 2023
60 f. : il. color.

Trabalho de Conclusão de Curso (Graduação em
Biblioteconomia) - Universidade de Brasília, Faculdade
de Ciência da Informação, Brasília, 2023.

1. Web semântica. 2. Metadados 3.
Interoperabilidade. 4. Mapa conceitual. I. Silva, Márcio
Bezerra da, orient. II. Título.

CDU 002:004

FOLHA DE APROVAÇÃO

Título: METADADOS: contribuições no campo da web semântica

Autor(a): Rayra Silva de Abreu

Monografia apresentada em **21 de julho de 2023** à Faculdade de Ciência da Informação da Universidade de Brasília, como parte dos requisitos para obtenção do grau de Bacharel em Biblioteconomia.

Orientador(a) (FCI/UnB): Dr. Márcio Bezerra da Silva
Membro Interno (FCI/UnB): Dra. Fernanda Farinelli
Membro Externo (FAC/UNB): Dra. Mônica Regina Peres

Em 20/10/2022.



Documento assinado eletronicamente por **Fernanda Farinelli, Professor(a) de Magistério Superior da Faculdade de Ciência da Informação**, em 25/07/2023, às 12:24, conforme horário oficial de Brasília, com fundamento na Instrução da Reitoria 0003/2016 da Universidade de Brasília.



Documento assinado eletronicamente por **Marcio Bezerra da Silva, Professor(a) de Magistério Superior da Faculdade de Ciência da Informação**, em 25/07/2023, às 12:26, conforme horário oficial de Brasília, com fundamento na Instrução da Reitoria 0003/2016 da Universidade de Brasília.



Documento assinado eletronicamente por **Monica Regina Peres, Técnico(a) em Assuntos Educacionais da Faculdade de Comunicação**, em 25/07/2023, às 13:53, conforme horário oficial de Brasília, com fundamento na Instrução da Reitoria 0003/2016 da Universidade de Brasília.



A autenticidade deste documento pode ser conferida no site http://sei.unb.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **10055217** e o código CRC **416F5650**.

Dedico este trabalho à tia Márcia, ao tio Marcos e ao meu sogro Marcos. Vocês se foram rápido demais, mas serão lembrados pra sempre com muito carinho.

“Metade de mim agora é assim
De um lado a poesia, o verbo, a saudade
Do outro, a luta, a força e a coragem pra
chegar no fim”

Teatro Mágico, O anjo mais velho

AGRADECIMENTOS

Quero começar agradecendo a minha família, a quem tornou o meu ingresso à Faculdade uma coisa possível. Obrigada mãe, Gizele, por sempre se esforçar tanto para que eu realizasse os meus sonhos, mesmo precisando mudar de Estado para que eles se tornassem realidade. Obrigada pai, Rangel, por todas as caronas ao “lonjão”, vulgo Bloco de Sala de Aula Norte. Obrigada irmã, Raíssa, por estar presente na minha vida nas horas boas e ruins. Obrigada prima, Isabor, por me apresentar a Biblioteconomia como um caminho possível. Obrigada primo, Juninho, por me levar pela primeira vez a uma biblioteca, e mostrar que o caminho dos livros e dos estudos é legal. Obrigada vó, Margarida, quem me presenteou com o meu primeiro computador, com muito esforço, para que eu pudesse desenvolver os meus estudos. E obrigada a todos, tios, tias, primos e primas, que de alguma forma me ajudaram a fazer essa caminhada.

Quero agradecer aos amigos que durante esses anos estiveram do meu lado, nos altos e baixos, e aos colegas que tornaram mais fácil e feliz a jornada da graduação. E um agradecimento mais especial a três pessoas a que devo essa conquista: meu orientador Márcio, minha psicóloga Denise e meu amor Ramón.

Obrigada, professor Márcio, por acreditar em mim, por todas as reuniões e e-mails trocados, por ser sempre tão prestativo e disponível. Eu definitivamente tive o melhor orientador possível!

Denise, obrigada por me mostrar que eu não preciso carregar o peso do mundo nas costas, e que sou mais do que capaz de alcançar os meus objetivos, os últimos anos não foram fáceis e eu não sei o que faria sem você.

E meu amor Ramón, obrigada por ser o parceiro mais incrível desse mundo, por aguentar todo o desespero e noites sem dormir, e por todas as palavras de carinho e incentivo. E por me mimar com todo o açúcar desse mundo, as barras de chocolate e os refrigerantes que viraram noites comigo foram muito úteis. Dividir a vida com você é um prazer, te amo!

RESUMO

Pesquisa que busca investigar, segundo a literatura da Ciência da Informação, as contribuições dos metadados na arquitetura da *web* semântica. Fundamenta-se a partir de três temáticas: estruturação da informação na *web*; abordando as *webs* de documentos, programável e de dados; *web* semântica, verificando as tecnologias que a compõem, e perpassando pelo conceito de interoperabilidade; e metadados, identificando seus atributos e aplicações. Caracteriza-se como uma investigação de método dedutivo, de natureza básica e de classificação descritiva, se propondo a descrever fenômenos específicos e suas relações segundo as pesquisas bibliográfica e documental. Ilustra os resultados inspirado na Teoria da Aprendizagem Significativa, vista nesta pesquisa como uma técnica de mapeamento conceitual. Destaca-se, entre os resultados, um conjunto de atributos e aplicações dos metadados, a contemplação de tecnologias que compõem a *web* semântica, incluindo os metadados, e a forma como os metadados contribuem para efetivar a *web* semântica. Conclui-se que a *web* semântica constitui-se de uma arquitetura específica, de um conjunto de tecnologias, de ontologias e de metadados, sendo que estes contribuem na representação de objetos digitais, na oferta de padrões, entre simples, estruturados e ricos, e na promoção da interoperabilidade.

Palavras-chave: *Web*. *Web* semântica. Metadados. Interoperabilidade. Mapa conceitual.

ABSTRACT

Research that seeks to investigate, according to Information Science literature, the contributions of metadata in semantic web architecture. It is based on three subjects: structuring information on the web; addressing document webs, programmable, and data web; semantic web, verifying the technologies that compose it, and going through the interoperability concept; and metadata, identifying its attributes and applications. It is characterized as an investigation of deductive method, basic nature and descriptive classification, proposing to describe specific phenomena and their relationships according to bibliographic and documentary research. It illustrates the outcomes inspired by the Theory of Meaningful Learning, seen in this research as a method of conceptual mapping. It highlights, among the results, a set of attributes and applications of metadata, the contemplation of technologies that make up the semantic web, including metadata, and how metadata contribute to make the semantic web effective. It is concluded that the semantic web is made of specific architecture, a set of technologies, ontologies and metadata, being those the ones that contribute in representing digital objects, in pattern offers, between simple, structured and rich, in promoting interoperability.

Keywords: Web. Semantic web. Metadata. Interoperability. Concept map.

LISTA DE FIGURAS

Figura 1	Diretório do <i>Yahoo</i> (1994)	22
Figura 2	Modelos de aplicações <i>web</i>	26
Figura 3	Arquitetura da <i>web</i> semântica	29
Figura 4	Exemplo de utilização da <i>tag</i> <meta>	34
Figura 5	Registro descrito em DC, codificado em XML	35
Figura 6	Trecho de um registro descrito em MARC 21	37
Figura 7	Triplas RDF	38
Figura 8	MC do tipo teia de aranha	41
Figura 9	MC que conceitua RD (características essenciais)	42
Figura 10	Legenda da lógica geométrica (MC)	43
Figura 11	MC relacionando recursos da <i>web</i> semântica	49

LISTA DE QUADROS

Quadro 1	Características da <i>web</i> 1.0	22
Quadro 2	Características da <i>web</i> 2.0	23
Quadro 3	Características da <i>web</i> 3.0	24
Quadro 4	Tipos de metadados	32
Quadro 5	Elementos DC	34
Quadro 6	Campos do MARC 21	36
Quadro 7	Tecnologias da <i>web</i> semântica	47
Quadro 8	Recursos da <i>web</i> semântica e suas contribuições	48

LISTA DE ABREVIATURAS E SIGLAS

AACR2	Anglo American Cataloguing Rules, second edition
AJAX	Asynchronous JavaScript and XML
API	Application Programming Interface
ARPANET	Advanced Research Projects Agency Network
CC	Ciência da Computação
CI	Ciência da Informação
CSS	Cascading Style Sheets
DC	Dublin Core
DCMI	Dublin Core Metadata Initiative
DOM	Document Object Model
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária
HTML	HyperText Markup Language
HTTP	HyperText Transfer Protocol
IA	Inteligência Artificial
IP	Internet Protocol
LC	Library of Congress
MARC	Machine Readable Cataloging
MC	Mapa conceitual
OCLC	OnLine Computer Library Center
OD	Objeto Digital
OWL	Web Ontology Language
PHP	Hypertext Preprocessor
RC	Representação do Conhecimento
RDF	Resource Description Framework
REST	Representational State Transfer
RSS	Really Simple Syndication
SOAP	Simple Object Access Protocol
SPARQL	Simple Protocol and RDF Query Language
TCC	Trabalho de Conclusão de Curso
TCP	Transmission Control Protocol
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
XHTML	eXtensible Hypertext Markup Language
XML	eXtensible Markup Language
XMLHttpRequest	eXtensible Markup Language HyperText Transfer Protocol Request
XSLT	eXtensible Stylesheet Language
WWW	World Wide Web
W3C	World Wide Web Consortium

SUMÁRIO

1 INTRODUÇÃO	14
1.1 PROBLEMATIZAÇÃO.....	16
1.2 JUSTIFICATIVA.....	17
1.3 OBJETIVOS.....	19
1.3.1 Geral.....	19
1.3.2 Específicos.....	19
2 FUNDAMENTAÇÃO TEÓRICA	19
2.1 ESTRUTURAÇÃO DA INFORMAÇÃO NA <i>WEB</i>	20
2.1.1 Arquitetura da <i>web</i>	24
2.2 <i>WEB</i> SEMÂNTICA.....	28
2.2.1 Camadas da <i>web</i> semântica.....	29
2.3 METADADOS.....	31
2.3.1 Padrões de metadados.....	33
3 METODOLOGIA	39
3.1 CARACTERÍSTICAS.....	39
3.2 APRESENTAÇÃO DOS RESULTADOS DA PESQUISA.....	40
3.2 ETAPAS DA PESQUISA.....	43
4 RESULTADOS	45
4.1 ATRIBUTOS E APLICAÇÕES DOS METADADOS.....	45
4.2 TECNOLOGIAS DA <i>WEB</i> SEMÂNTICA.....	46
4.3 CONTRIBUIÇÕES DE METADADOS NA <i>WEB</i> SEMÂNTICA.....	48
5 CONSIDERAÇÕES FINAIS	51
REFERÊNCIAS	53

1 INTRODUÇÃO

Ao longo do tempo, a informação se tornou objeto de estudo em diversas áreas do conhecimento, e devido ao seu caráter multidisciplinar, variadas são as definições que a conceituam, considerando as especificidades de onde se investiga e/ou aplica. Informação é uma “[...] mensagem inscrita que nasce na sociedade e a ela retorna [...]” (KOBASHI; TÁLAMO, 2003, p. 16). Nesta linha, de registro e comunicação, Le Coadic (1996, p. 5) conceitua informação da seguinte forma:

[...] é um conhecimento inscrito (gravado) sob a forma escrita (impressa ou numérica), oral ou audiovisual. [...] comporta um elemento de sentido. É um significado transmitido a um ser consciente [...] feita graças a um sistema de signos (linguagem).

Com o surgimento de tecnologias, como a invenção da prensa de Gutenberg em 1400, a produção de informação foi aumentando. No período após a Segunda Guerra Mundial, chegou-se a um ponto significativo na produção de informação, resultando em uma espécie de “explosão informacional”, marcada pelo processo de geração e consumo intenso de informação, especialmente a partir do advento da Internet, também conhecida como a rede mundial de computadores.

A Internet alcançou um nível global devido ao desenvolvimento da *World Wide Web* (WWW). Em seu primórdio, conhecido como *web* 1.0, a WWW caracterizou-se por ser um modelo unidirecional, formado por páginas estáticas com capacidade limitada de (qualquer) interação entre usuários e conteúdos em si, possibilitando apenas a leitura de informações inseridas pelos detentores dos *websites*, além de ser entendida apenas por humanos. Enquanto muitos apenas liam conteúdos, um pequeno conjunto de pessoas “[...] detinha poder de compra para custear as transações online e adquirir o software para criação e manutenção de sites” (COUTINHO; BOTTENTUIT JUNIOR, 2007, p. 199).

A segunda fase da *web* é marcada por uma maior dinamicidade, onde os usuários passaram a ser também produtores de conteúdo. A *web* 2.0 se caracterizava, e continua assim, pela possibilidade de interação dos usuários tanto entre si como com o ambiente digital, aspecto perceptível pela ascensão das redes

sociais, *blogs* e *tagueamentos*¹. Enquanto filosofia, a “[...] *web* 2.0 prima pela facilidade na publicação e rapidez no armazenamento [...]” (COUTINHO; BOTTENTUIT JUNIOR, 2007, p. 200), ocasionando em um exponencial fluxo de informação, visto o grande número de usuários criando e compartilhando conteúdos.

A *web* 3.0, conhecida como *web* semântica, objetiva tornar a WWW legível por máquinas e não apenas para humanos, com código aberto e maior transparência. A ideia é definir dados estruturados de forma que sejam analisados e integrados para obter novas informações, com contexto e significado. Tim Berners-Lee, James Hendler e Ora Lassila (2001), em um famoso artigo publicado na revista científica *Scientific American*², apresentaram suas idéias sobre o que seria a *web* semântica, e estabeleceu algumas condições para que ela se tornasse realidade, como representação do conhecimento (RC)³, *eXtensible Markup Language* (XML)⁴, *Resource Description Framework* (RDF)⁵ e ontologias⁶.

A base da *web* semântica é estabelecida pela RC, que se apoia nos metadados⁷, enquanto um conjunto de elementos padronizados, através de regras de inferência, permitindo, por um lado, uma melhor organização e tratamento dos objetos digitais (ODs)⁸ e proporcionando, pelo outro, diferentes mecanismos de busca e recuperação de informação na *web*.

Diante da contextualização introdutória, o presente Trabalho de Conclusão de Curso (TCC) se estrutura da seguinte forma: a primeira seção apresenta a

¹ Ação de definir palavras-chave na *web*, ou seja, “[...] ato de inserir tags no catálogo, de maneira livre, ou seja, ação que consiste na representação temática da informação realizada pelos usuários, neste caso na perspectiva digital [...]” (RIBEIRO; SILVA, 2015, p. 8).

² Acesso: <https://www.scientificamerican.com/>.

³ Entende-se que “[...] o processo de produção dos registros de conhecimento compreende a etapa de representação da coisa ou ser, gerando-se em decorrência um produto final, um conhecimento sobre a coisa [...]” (ALVARENGA, 2003, p. 22).

⁴ Linguagem de marcação derivada do antigo *Standard Generalized Markup Language* (SGML), que fornece uma sintaxe básica usada para representar e compartilhar informações estruturadas (WORLD WIDE WEB CONSORTIUM, 2016).

⁵ Formato que permite expressar informações sobre ODs de maneira que as máquinas consigam compreendê-las e interpretá-las, possibilitando a interoperabilidade e o processamento automático de dados (WORLD WIDE WEB CONSORTIUM, 2014)

⁶ Tomando como base o contexto da *web* semântica, “[...] ontologia é um documento ou arquivo que define formalmente relações entre termos” (ROBREDO, 2010, p. 26)

⁷ Metadados significa “[...] dados que descrevem outros dados em um sistema de informação, com o intuito de identificar de forma única uma entidade (recurso informacional) para posterior recuperação” (ALVES, 2010, p. 47).

⁸ Neste trabalho, OD é compreendido como um “[...] objeto de informação, de qualquer tipo de informação ou qualquer formato, que é expresso em formato digital” (THIBODEAU, 2002, p. 6, tradução nossa). Por exemplo, um *e-book*, um arquivo de um filme (vídeo) etc.

introdução, a problematização, a justificativa e os objetivos da pesquisa; a segunda seção fundamenta teoricamente o estudo, discorrendo sobre estruturação da informação na *web*, além de dissertar sobre *web* semântica e metadados; a terceira seção delinea a metodologia utilizada para alcançar os objetivos da pesquisa; a quarta seção expõe os resultados, incluindo uma representação imagética que ilustra as contribuições dos metadados na *web* semântica; e, por fim, a quinta seção apresenta as considerações finais e expectativas de estudos futuros.

1.1 PROBLEMATIZAÇÃO

A “explosão informacional” proporcionou mudanças de paradigmas em áreas que tratam da informação, sendo um objeto que constrói, edita e fortalece conhecimento que o indivíduo possui. Se um dia, o acesso era a dificuldade em pauta, diante do atual extenso volume de informação produzida, a recuperação e o uso (qualitativo) passaram a ser os pontos de preocupação, pois o simples acesso à informação não é mais suficiente.

A Ciência da Informação (CI) se enxerga presente tanto no debate sobre recuperação eficiente quanto no atendimento a necessidades informacionais, especialmente quando a área considera formas de preparar e disponibilizar informação como caminho. Trata-se de uma assertiva que passa a receber evidenciadas “[...] no período de crise e de transformações pós-guerra, [pois] a CI ganha destaque com o aumento do número de produção científica e com a importância sobre seu tratamento e disseminação das informações” (SILVA; SAMPAIO, 2017, p. 8).

Em sua essência, a CI tem suas origens atreladas a áreas que possuem a informação como um objeto comum de tratamento. No caso, o tratamento da informação, no contexto da terceira geração da WWW, exige certas condições para permitir o funcionamento da *web* semântica, e a RC e as ontologias são condições que a área se debruça há algum tempo. Enquanto a RC compreende “[...] uma questão que preocupa o mundo da documentação desde a sua origem” (VICKERY, 1986, p. 145), as ontologias se apresentam como uma expressão cooptada da

Filosofia e que foi incorporada pela CI, pela Ciência da Computação (CC)⁹, pela Inteligência Artificial (IA)¹⁰ e pela Linguística¹¹. Emygdio (et al., 2021) reforça que o interesse pelas ontologias possui um sentido amplo, ou seja, assenta-se na intenção de se prover sentido para os dados que estão disponíveis na *web*, pois trata-se de um espaço de caráter universal.

A importância da existência de sistemas de informação que atuem de maneira interoperável¹², a partir de uma RC eficaz e adoção de ontologias formais¹³, tornou-se evidente nos últimos tempos. Emygdio et al.(2021) apontou que, em 2020, com a pandemia que assolou o mundo, os problemas que a ciência vem enfrentando pela falta de interoperabilidade ficaram mais evidentes, especificamente nos registros e sistemas médicos.

Partindo-se do pressuposto de que a RC corporifica a *web* semântica, lançando mão de elementos de representação, como é o caso dos metadados, estudados por pesquisadores da CI, fomenta-se o seguinte questionamento: *conforme levantamento literário, de que maneira os metadados contribuem para a efetivação da web semântica?*

1.2 JUSTIFICATIVA

A ação de recuperação da informação pode ser considerada um relevante fator para o desenvolvimento da CI, sendo este um comportamento naturalmente identificável na área. Para Borko (1968), a CI se apresenta como uma disciplina que investiga as propriedades e o comportamento da informação.

⁹ Pode ser compreendida como “[...] a disciplina que busca construir uma base científica para tópicos como projeto e programação de computadores, processamento de informação, soluções algorítmicas de problemas e o próprio processo algorítmico” (BROOKSHEAR, 2013, p. 2).

¹⁰ Trata-se de “[...] uma área da ciência da computação que busca construir máquinas autônomas - que possam conduzir tarefas complexas sem intervenção humana” (BROOKSHEAR, 2013, p. 414).

¹¹ Integrante da semiologia, ou da semiótica, a Linguística “[...] estuda a principal modalidade dos sistemas sócio-sígnicos, as *línguas naturais*, que são a forma de comunicação mais altamente desenvolvida e de maior uso” (FIORIN, 2010, p. 19).

¹² Refere-se a capacidade dos sistemas de informação, como bases de dados e bibliotecas digitais, “[...] trocarem e compartilhar documentos, consultas e serviços, usando diferentes plataformas de hardware e software, estrutura de dados e interfaces [...]” (ALVES; SOUZA, 2007, p. 23).

¹³ Possuem axiomas lógicos que estabelecem as propriedades para cada termo e expressam relações entre entidades a fim de evitar qualquer ambiguidade (FARINELLI, ALMEIDA, 2014).

Pensar no desenvolvimento da *web* é também considerar a importância da informação (disponível) estar devidamente representada, compreensível ao público, e categorizada conforme atributos, com fins de recuperação. Observa-se, então, um cenário em que as “[...] ontologias têm sido um tema de interesse em ciência da informação [...]” (SOUZA; ALMEIDA; BARACHO, 2013, p. 169), assim como “o estudo dos metadados também pode ter sua importância reafirmada como essencial na sociedade da informação [...]” (LOURENÇO, 2007, p. 65).

A importância dos metadados pode ser observada em um estudo que realizou uma correspondência entre o formato MARC 21¹⁴ e o padrão *Dublin Core* (DC)¹⁵, especificamente em um sistema *web* gerenciado pela Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), chamado de Agência de Informação Embrapa, que organiza e possibilita o acesso à informação tecnológica e aos conhecimentos gerados pela empresa. Nesse estudo, discutiu-se a importância dos metadados na geração, integração e compartilhamento de dados entre sistemas, em relação a qualquer plataforma que use o formato MARC 21 como mecanismo de entrada de dados. No caso, ao considerar a qualidade dos dados gerados a partir de padrões de consenso internacional (ALVES; SOUZA, 2007), esse estudo foi um significativo estímulo para a escolha do tema do presente TCC.

O interesse pelo tema de pesquisa foi sendo alimentado durante a disciplina *Redes de Informação e Transferências de Dados* (2020.2), do Curso de Biblioteconomia, quando discutiu-se o supramencionado estudo e formas de interoperabilidade entre sistemas foram apresentadas, além da realização de atividades utilizando padrões de metadados e linguagens de marcação. É justamente nas características que os metadados possuem que cogitou-se o interesse em identificar benefícios desta forma de representação na *web* semântica.

¹⁴ O *Machine-Readable Cataloging* “[...] é datado da década de 1960 e foi desenvolvido pela *Library of Congress* (LC), com o objetivo de padronizar a descrição bibliográfica em meio eletrônico com o início da utilização de computadores para gerenciar o processo de catalogação” (CASTRO, 2008, p. 90).

¹⁵ O *Dublin Core Metadata* é um padrão de metadados desenvolvido pela *OnLine Computer Library Center* (OCLC) em 1995, “[...] criado para facilitar a descoberta do recurso informacional no ambiente *Web*, ou seja, abarcar uma representação mínima que proporciona a sua identificação bem como a sua localização” (CASTRO, 2008, p. 97).

1.3 OBJETIVOS

1.3.1 Geral

- Investigar, na literatura da CI, contribuições dos metadados na arquitetura da *web* semântica.

1.3.2 Específicos

- Verificar atributos e aplicações dos metadados;
- Identificar a presença de metadados entre as tecnologias da *web* semântica;
- Apontar a postura contributiva dos metadados na *web* semântica;
- Ilustrar as contribuições identificadas por meio de uma representação.

2 FUNDAMENTAÇÃO TEÓRICA

A Ciência vive em um ciclo de construção de novas verdades, e nessa busca (pela verdade), ela se apropria da realidade, construindo um conhecimento cumulativo, que pode ser inédito para uns ou consolidar o que já se compreendia por outros. Um caminho para alcançar verdades científicas é fundamentar as pesquisas, utilizando o conhecimento produzido na literatura sobre o tema investigado, o qual se assume como a base teórica do estudo.

Presente nas pesquisas científicas, tem-se a fundamentação teórica, que “[...] constitui o marco por meio do qual o problema da pesquisa é interpretado” (GASQUE, 2008, p. 108), ou seja, é onde o pesquisador trata sobre o tema proposto, com citações diretas e indiretas, evidenciando o domínio que assume-se sobre o tema, assim como ocorre neste TCC, que fundamenta-se em discussões sobre estruturação da informação na *web*, *web* semântica e metadados.

2.1 ESTRUTURAÇÃO DA INFORMAÇÃO NA WEB

A *Advanced Research Projects Agency Network* (ARPANET), precursora da Internet, foi uma rede de comutação de pacotes criada pelo Departamento de Defesa norte-americano em 1969, sendo composta por quatro pequenos computadores, conectados por linhas telefônicas dedicadas. Em 1974, a partir dessa proposta, de possibilitar a troca de informações, criou-se um protocolo, o qual viabilizou a comunicação entre computadores e pessoas dispersas localmente. A partir do *Transmission Control Protocol / Internet Protocol* (TCP/IP) foi possível ligar a ARPANET a várias outras redes heterogêneas, criando assim a Internet (WAZLAWICK, 2016), no formato que ela é mundialmente conhecida.

Enquanto a Internet foi viabilizada pelo protocolo TCP/IP, a *web* se tornou possível devido a criação e utilização do protocolo de comunicação de dados *Hypertext Transfer Protocol* (HTTP)¹⁶ e pela linguagem de computador HTML. A *web*

¹⁶ Trata-se de um “[...] protocolo de comunicação para acesso aos documentos. Esse protocolo estabelece as regras da comunicação entre duas aplicações, um cliente e um servidor, de forma a permitir a requisição de um conteúdo e o respectivo resultado retornado por essa requisição, por exemplo, uma página HTML” (LAUFER, 2015, p. 10-11). A linguagem de computador *Hypertext*

foi um sistema idealizado por Tim Berners-Lee em 1989, partindo da ideia de “[...] utilizar a Internet para a consulta e a utilização de itens de informação (documentos) organizados em uma estrutura hipertextual” (LAUFER, 2015, p. 10). Inevitavelmente, sendo uma característica inerente à Tecnologia da Informação (TI), em que há evolução/atualização, nesta pesquisa, apresenta-se e compreende-se a *web* a partir de três fases¹⁷ distintas, rotuladas como *web* 1.0, *web* 2.0 e *web* 3.0.

A ***web* 1.0, ou sintática**, teve início no começo da década de 1990, sendo a primeira fase da Internet, em sentido comercial. Essa fase foi marcada pelo crescimento caótico de páginas estáticas, pela comunicação unidirecional e pela presença de diretórios, os quais eram formados a partir da adoção de taxonomias¹⁸, ao passo que permitiam a navegação entre categorias pré-estabelecidas e busca de informações, ou seja, “[...] um servidor HTTP tem um diretório, uma pasta, que é designada como uma coleção de documentos, arquivos que podem ser enviados em resposta às requisições dos navegadores [...] identificados a partir de URLs¹⁹ [...]” (LAUFER, 2015, p. 12). O antigo diretório do *Yahoo* (Figura 1) é um exemplo de diretório que representava a fase da *web* em discussão, lançado em janeiro de 1994 e encerrado em dezembro de 2014.

Markup Language (HTML) é adotada “[...] para a descrição dos documentos. Ela foi inicialmente projetada como uma linguagem para a descrição de documentos científicos, mas adaptações efetuadas ao longo do tempo permitiram que ela seja utilizada para descrever diversos outros tipos de documentos” (LAUFER, 2015, p. 10).

¹⁷ Desde 2009, o conceito de *web* 4.0 vem sendo discutido por um número considerável de autores, porém sem chegar a um consenso. As abordagens a respeito da *web* 4.0 sugerem que ela coexiste com a *web* 3.0, e possui múltiplas dimensões que contam com tecnologias como Internet das Coisas, *Big Data* e IA para sua adoção e implementação (ALMEIDA, 2017).

¹⁸ Estrutura que “[...] surgiu como Ciência das leis da classificação de formas vivas e, por extensão, ciência das leis da classificação. No ambiente dos sistemas de classificação, das ontologias, da inteligência artificial, é entendida como classificação de elementos de variada natureza. Tradicionalmente, as taxonomias tiveram por função a classificação das espécies em botânica, e zoologia, adotando uma nomenclatura binária. Taxonomia é, por definição, classificação, sistemática” (CAMPOS; GOMES, 2007, p. 3).

¹⁹ A *Uniform Resource Locator* endereça um objeto, um documento presente na Internet, contendo “[...] o protocolo do recurso (por exemplo, <http://>, para hypertext transfer protocol, ou <ftp://> para file transfer protocol), o nome do domínio do recurso e o nome hierárquico do arquivo (endereço)” (SANTANA et al., 2001, p. 51).

Figura 1: Diretório do Yahoo (1994).



Fonte: Max (2008).

Em sua primeira fase, a *web* era vista como um meio de divulgação de informações no qual a comunicação partia de poucos produtores de conteúdos, como detentores de *websites* e gestores de diretórios, aspecto que definia o modelo chamado *top-down* (cima para baixo). A experiência dos usuários ficava, então, limitada, justamente pelos *websites* serem estáticos, além de outras características elencadas no Quadro 1.

Quadro 1: Características da *web* 1.0.

Conteúdo	Estático
Usuário	Consumidor
Dispositivo	<i>Desktop</i>
Navegação	Especulação de domínio
Monetização	<i>Page Views</i>
Indexação	Diretórios e taxonomia
Ferramentas	<i>DoubleClick, Ofoto, Britannica Online, Akamai etc.</i>

Fonte: Elaborado pela autora (2023), com base em O'Reilly (2005).

A segunda fase da *web*, chamada de **2.0 ou social**, eclode em meados de 2004 como uma proposta diferente da anterior, defendendo a ideia de ser um espaço interativo e colaborativo. Sendo marcada pela liberdade de taguear, os usuários contribuem na moldagem dos ambientes *web* ao representarem os conteúdos dispostos com palavras-chaves. Adotar *tags* é uma forma de facilitar a recuperação de informações por aqueles que comungam de palavras-chave semelhantes, que fazem parte de um mesmo grupo ou possuem gostos avizinados, gerando diretórios criados pelas próprias pessoas, ou seja, fomentando um modelo de organização do conhecimento²⁰ do tipo *bottom-up* (baixo para cima), intitulado folksonomia²¹. Segundo Tapscott e Williams (2007, p. 57-58), “[...] quando as pessoas ‘etiquetam’ (colocam tags) um conteúdo de maneira colaborativa, cria-se uma ‘folksonomia’ [...]”.

Ferramentas criadas na *web* 1.0 inovaram e se fizeram presentes na *web* 2.0 (Quadro 2), passando a estimular a colaboração de usuários na construção de conteúdos, como é o caso da *Wikipédia*²², e inaugurando o conceito de folksonomia, como implementado pelo *Flickr*²³ (O'REILLY, 2005). Ainda surgiram os *blogs*²⁴, os quais foram enriquecidos pela tecnologia *Really Simple Syndication* (RSS), permitindo que os usuários, a partir de uma assinatura, fossem notificados cada vez que houvesse mudança no conteúdo da página.

Quadro 2: Características da *web* 2.0.

Conteúdo	Dinâmico
Usuário	Consumidor participativo
Dispositivo	<i>Desktop</i> e dispositivos portáteis
Navegação	Motores de busca por palavras-chave
Monetização	Custo por clique/acesso

²⁰ A organização do conhecimento tem por base a análise do conceito para determinar a posição que ele ocupa num determinado domínio e as relações que ele possui com outros conceitos, a fim de construir representações do conhecimento (BRASCHER; CAFÉ, 2008).

²¹ Entende-se como sendo “[...] essencialmente uma taxonomia orgânica de baixo pra cima que organiza o conteúdo da web” (TAPSCOTT; WILLIAMS, 2007, p. 57-58).

²² Acesso: <https://pt.wikipedia.org/>.

²³ Acesso: <https://www.flickr.com/>.

²⁴ O blog (weblog) é um jornal eletrônico que combina texto, imagens e hiperlinks (ANDRADE et al., 2011, p. 33).

Indexação	Etiquetagem (<i>tagging</i>) e folksonomia
Ferramentas	<i>Google AdSense, Flickr, Bit Torrent, Wikipedia, blogs etc.</i>

Fonte: Elaborado pela autora (2023), com base em O'Reilly (2005).

A terceira fase, conhecida como **web 3.0 ou semântica** (Quadro 3), emerge a partir da experiência da antecessora, tendo o propósito de estruturar o conteúdo (significativo) das páginas *web*, a fim de que *softwares* possam percorrer estas páginas para executar solicitações dos usuários e trazer resultados específicos. Deve-se mencionar que o elemento “usuário” não é mencionado no Quadro 3, pois, enquanto nas primeiras fases o conteúdo era gerado e pensado com enfoque no usuário, o conteúdo na *web 3.0* é reconhecido pelos computadores, pelas máquinas.

Quadro 3: Características da *web 3.0*.

Conteúdo	Personalizado
Dispositivo	<i>Desktop</i> e dispositivos portáteis
Navegação	Busca semântica
Monetização	Engajamento do usuário
Indexação	Contexto/relevância e ontologia
Ferramentas	<i>Bitcoin, Ethereum, Brave, Metamask etc.</i>

Fonte: Elaborado pela autora (2023), com base em Magrani (2018).

As supracitadas fases da *web* podem ser entendidas para além das suas características, anteriormente elencadas, sendo compreendidas também por sua arquitetura. Sendo assim, seriam elas respectivamente, também em sentido de evolução: *web* de documentos; *web* programável; e *web* de dados.

2.1.1 Arquitetura da *web*

A *web 1.0*, ou do ponto de vista da sua arquitetura, ou seja, a **web de documentos**, é caracterizada por se organizar em uma estrutura de hipertexto²⁵, e baseada no conceito cliente-servidor, “[...] onde uma aplicação (um cliente Web)

²⁵ Hipertexto é um texto estruturado, composto por um conjunto interligado de itens de informação (nós) contendo, por exemplo, textos, imagens, vídeos, etc. (LAUFER, 2015, p. 10).

requisita um documento (um recurso) a uma outra aplicação (um servidor Web) informando a identificação desse documento” (LAUFER, 2015, p. 10).

Para possibilitar a *web* de documentos é fundamental a utilização de quatro elementos: URL, HTML, HTTP e *link*. A URL identifica um documento descrito em HTML, e essa identificação é fundamental para o servidor fazer a requisição desse documento, que é possível devido a comunicação entre o cliente e o servidor, viabilizada pelo protocolo HTTP. No caso do *link*, este proporciona o *hyperlink* em um hipertexto, visto que, utilizando uma *tag* (elemento específico da estrutura do *site*) específica da linguagem HTML, é possível incluir um *link* em uma página, interligando vários documentos e, assim, possibilitando a *web* de documentos. Laufer (2015) afirma que a estrutura mais comum nas aplicações *web* possui três camadas: **apresentação**, que seria um navegador, ou seja, a primeira camada; **lógica de negócios**, representada por um motor que utiliza alguma tecnologia de conteúdo dinâmico, isto é, a camada intermediária; e **armazenamento**, que ocorre em um banco de dados, considerado o terceiro nível.

A estrutura da *web* de documentos funciona da seguinte forma: o navegador (primeira camada) envia solicitações para o motor, que é uma aplicação (camada intermediária) que faz consultas e atualizações no banco de dados (terceiro nível), e retorna a resposta para o requisitante na forma de interface de usuário (LAUFER, 2015). Contudo, salienta-se que a cada solicitação é gerada uma nova requisição, e o processo então se repete, sendo necessário a atualização da página por inteiro.

A *web* 2.0, ou para fins da arquitetura, **web programável**, utiliza a estrutura da *web* de documentos de maneira aprimorada. As páginas de hipertexto estáticas ganham um conteúdo dinâmico, tornando possível várias requisições e transações, devido a um conjunto de recursos e de uma técnica: *Common Gateway Interface* (CGI)²⁶; linguagens de programação, como *JavaScript*, *Hypertext Preprocessor* (PHP) e *Python*; linguagens de marcação, como XML; e a técnica denominada *Asynchronous JavaScript and XML* (AJAX).

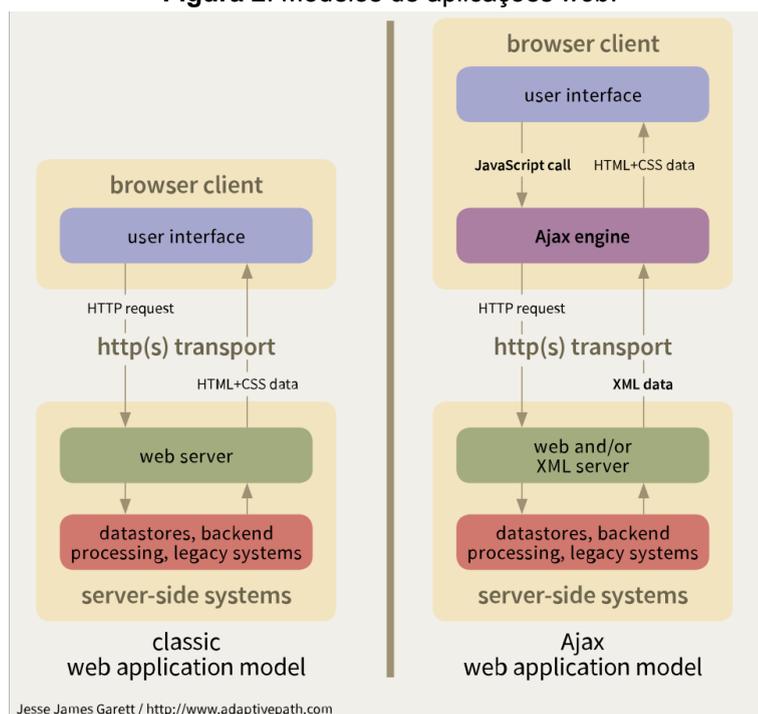
Garrett cunhou o termo AJAX em 2005, no artigo *Ajax: a new approach to web applications*, explicando que AJAX não é uma tecnologia, mas uma técnica que

²⁶ Entende-se CGI como “[...] um método padrão usado para gerar conteúdo dinâmico em páginas e aplicativos da Web. Ele fornece uma interface entre o servidor Web e programas que geram o conteúdo dinâmico da Web” (LAUFER, 2015, p. 12).

combina várias tecnologias: apresentação da interface baseada nos padrões *eXtensible Hypertext Markup Language* (XHTML)²⁷ e *Cascading Style Sheets* (CSS)²⁸; interface dinâmica e interação por meio do *Document Object Model* (DOM)²⁹; troca e manipulação de dados via XML e *eXtensible Stylesheet Language* (XSLT)³⁰; recuperação assíncrona de dados usando *XMLHttpRequest*³¹; e *JavaScript*, que liga todas as tecnologias supramencionadas. (GARRETT, 2005, tradução nossa).

A técnica AJAX permite que as solicitações de dados sejam feitas em segundo plano, não sendo mais necessário que a cada solicitação à página precise que ela seja recarregada por inteira, proporcionando, assim, uma experiência mais fluida ao usuário. No contexto das aplicações *web*, a Figura 2 apresenta uma comparação entre o modelo tradicional e o modelo que utiliza a técnica AJAX.

Figura 2: Modelos de aplicações *web*.



Fonte: Laufer (2015, p. 29).

²⁷ Trata-se da reformulação do HTML como uma aplicação de XML (WORLD WIDE WEB CONSORTIUM, 2010).

²⁸ É uma linguagem que adiciona estilo às páginas *web* sem alterar a semântica adicionada por outra linguagem utilizada para construí-las (WORLD WIDE WEB CONSORTIUM, 2023).

²⁹ É uma API para documentos HTML e XML que define a estrutura lógica dos documentos, assim como permite acesso e manipulação dos mesmos (WORLD WIDE WEB CONSORTIUM, 1998).

³⁰ Linguagem de transformação usada para manipular e converter documentos XML em outros formatos, como HTML ou texto, aplicando regras de transformação específicas (WORLD WIDE WEB CONSORTIUM, 2021).

³¹ Refere-se a uma API que permite fazer requisições assíncronas a um servidor, possibilitando a transferência de dados (WORLD WIDE WEB CONSORTIUM, 2016).

Nesse momento da *web* também surgem os *Web Services*³² e os *Web APIs*³³, devido à disseminação das aplicações *web* e às inúmeras linguagens de programação que surgiram. Utilizando o *Ajax*, as aplicações passaram a fornecer serviços que passariam a ser consumidos por outras aplicações, resultando em uma outra camada. Esses componentes *web* fornecem uma forma padrão de interoperação entre aplicações de *software* diferentes, sendo distinguidos segundo protocolos, visto que o *Web Service* utiliza o *Simple Object Access Protocol* (SOAP), enquanto a *Web APIs* utiliza o protocolo *Representational State Transfer* (REST).

No caso da *web 3.0*, chamada de **web de dados**, desponta trazendo à tona a ideia da *web* de documentos, porém não pensando nos documentos como um todo, mas nos dados que formam esses documentos, de forma que possibilite acessá-lo segundo um de seus dados específicos. Nesse sentido, os motores de busca³⁴ e os metamotores³⁵ são exemplos de aplicações que conseguem ser intermediários nesse processo, de extrair os dados que estão contidos em um documento para apresentá-los aos seus possíveis consumidores (LAUFER, 2015).

Considerar-se “[...] a Web de dados como uma camada de dados consumidos por uma camada de aplicações” (LAUFER, 2015, p. 30), em um ciclo onde aplicações geram dados, os quais, por sua vez, são utilizados como fontes de dados por outras aplicações. Dessa forma, a responsabilidade de interpretar os dados e compreender seus possíveis significados recai sobre as aplicações, que criam conexões a partir do vasto conjunto de informações disponíveis em diferentes níveis.

Buscando uma *web* de dados, na qual os ODs apresentam significados bem definidos, e para que possam ser entendidos e processados por aplicações, surge a *web semântica*, que procura utilizar técnicas e tecnologias específicas para organizar esses objetos, de modo a serem recuperados eficientemente.

³² São uma forma padronizada de permitir a comunicação e interação entre diferentes aplicações de *software*, ou seja, uma tecnologia que promove interoperabilidade em rede (LAUFER, 2015).

³³ *Web Application Programming Interface* é uma interface de programação de aplicações que garante interoperabilidade entre diferentes *softwares*, e funciona como o *Web Service*, porém utilizando um protocolo de comunicação diferente e uma formatação de dados menos restritiva (LAUFER, 2015).

³⁴ Realiza a indexação de *sites* e a criação de bancos de dados de forma automática, colecionando o maior número de recursos possível. Devido a essas características, suas bases são enormes, e a busca é feita por meio de palavras-chave, o que acarreta em baixa precisão (ALVES, 2005).

³⁵ Metamotores, também chamados de Multibuscadores ou Metapesquisadores, são ferramentas que buscam simultaneamente em vários mecanismos de busca, funcionando como um agente intermediário que compila os resultados obtidos, a fim de apresentar os resultados das buscas de maneira unificada (BRANSKI, 2000).

2.2 WEB SEMÂNTICA

Tim Berners-lee, em um *TED talk* de 2009³⁶, contou que no início dos anos 1990, estava frustrado com um projeto no qual estava trabalhando, juntamente com diversas pessoas, de inúmeros países, variados tipos de computadores e dados dos mais diferentes tipos de formatos (sendo todos incompatíveis). Dessa frustração, ele criou o conceito de *web*, e idealizou as tecnologias HTML, HTTP e URLs, buscando o que ele chamou de sistema de documentos virtuais, sistema esse que estaria disponível na Internet. Ele continuou explicando que, nesse primeiro momento, pediu para as pessoas disponibilizarem os seus documentos na *web*, considerando que dessa forma as informações poderiam ser compartilhadas, poderiam ser lidas. Em um segundo momento, o seu pedido passou a ser para que as pessoas compartilhassem seus dados, configurando-se o que ele chamou de *Linked Data*³⁷. O *Linked Data*, juntamente com outras tecnologias e ferramentas, permitiu a criação de um cenário onde a *web* semântica pudesse se tornar realidade (PALETTA; MUCHERONI, 2014).

A *web* semântica é um projeto idealizado por Berners-Lee, e que contou com a liderança do *World Wide Web Consortium (W3C)*³⁸, a qual ajudou, e continua ajudando, no desenvolvimento de tecnologias que darão suporte à *web* de dados. A *web* semântica não é uma nova *web*, mas uma extensão da *web* atual, que tem por objetivo estabelecer uma estrutura de dados organizada, a fim de que as informações contidas nesse espaço possam ser semanticamente descritas, ao passo que as máquinas, através de diversas ferramentas e tecnologias, possam identificar conceitos, processar informações e verificar possíveis relacionamentos. Conseqüentemente, busca-se obter respostas precisas às solicitações dos usuários, utilizando as mais diversas fontes de dados. O referido conjunto de ferramentas e tecnologias promovem-se em camadas, formando a arquitetura da *web* semântica.

³⁶ Acesso: https://www.ted.com/talks/tim_berniers_lee_the_next_web.

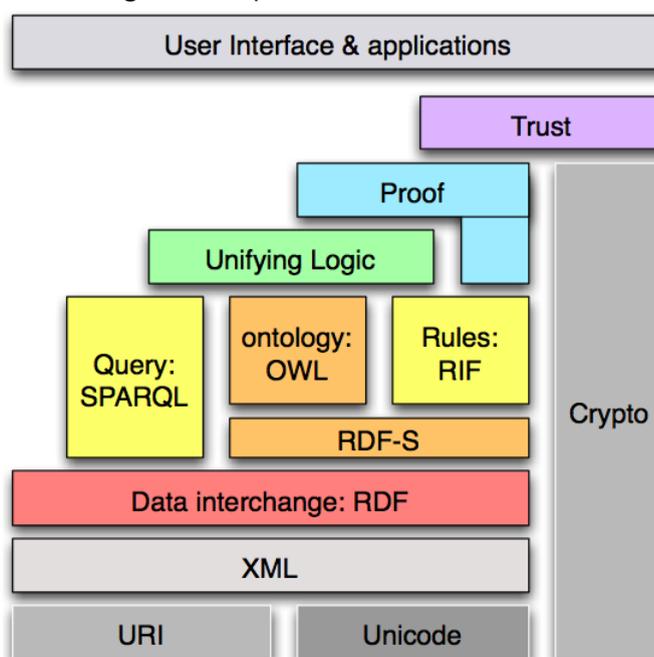
³⁷ *Linked Data*, em português Dados Vinculados ou Ligados, é uma abordagem para publicação de dados estruturados de modo que entidades sejam interligadas por URIs, e através do RDF são representadas com semântica, tendo as suas relações expressas. Assim, será possível compartilhar informações provenientes de diferentes fontes, de forma automática entre computadores, contribuindo para uma *web* mais conectada (PALETTA; MUCHERONI, 2014).

³⁸ Acesso: <https://www.w3c.br/>.

2.2.1 Camadas da *web* semântica

O W3C tem como missão desenvolver padrões e diretrizes que garantam o crescimento e desenvolvimento da *web*. Desta forma, o consórcio sugere que a arquitetura da *web semântica* seja organizada em camadas, cada uma fornecendo uma funcionalidade específica, ao passo que trabalham em conjunto para permitir a representação, inferência e consulta de informações de maneira semântica. São sete as camadas que constituem a *web* semântica (Figura 3), atuando de maneira consecutiva: faixa internacional; camada sintática; camada de dados; camada ontológica; camada lógica; camada de prova; e camada de validação.

Figura 3: Arquitetura da *web* semântica.



Fonte: Berners-Lee (2006).

A faixa **internacional** é a camada base da *web* semântica, sendo formada pelos elementos *Unicode*³⁹ e *Uniform Resource Identifier* (URI). O *Unicode* é um padrão universal que permite codificação técnica para textos que estejam em qualquer idioma, sendo adotado para o endereçamento, além de ser compatível com todos os tipos de *softwares*. No caso do URI, ele proporciona a localização de qualquer recurso na *web*, nomeando de forma única um OD.

³⁹ Acesso: <https://home.unicode.org/>.

Na sequência, têm-se a camada **sintática**, compreendida pela linguagem de marcação XML, que é responsável pela estrutura da informação e criação de novas *tags* para aplicações específicas, bem como identificadores contextuais. Em sentido funcional, “o XML serve para definir a estrutura dos dados e junto com os *namespaces*⁴⁰ separa o dado em si (conteúdo) do nome daquele dado (rótulo)” (MONTEIRO, 2013, p. 50). Portanto, essa camada facilita a interoperabilidade entre sistemas de informação.

A camada **de dados** vem em seguida, composta por RDF e suas extensões: RDF *Core* e RDF *Schema*. Essa camada é responsável por realizar a representação de metadados com significado codificado, para que um OD em específico seja referido, a fim de ser processado por máquinas. O RDF *Core* apresenta o modelo e a sintaxe para a codificação, enquanto o RDF *Schema* permite relacionar os ODs e criar hierarquias, classes e propriedades. Assim como os documentos se relacionam via *hiperlinks* entre páginas HTML, os dados se relacionam por meio de *links* RDF (PALETTA; MUCHERONI, 2014). Ainda, deve-se mencionar que nessa camada, enquanto o RDF (uma linguagem de estrutura) lida com os dados interoperacionais, o *SparQL*⁴¹ “[...] é uma linguagem de consulta que permite recuperar e manipular dados armazenados em RDF” (SILVA, 2013, p. 44).

Seguindo, a camada **ontológica** estabelece a semântica dos dados descritos e representados por metadados, definindo formalmente os conceitos a fim de evitar ambiguidades. Essa cama é composta principalmente pela *Web Ontology Language* (OWL), a linguagem de ontologia da *web* que verifica se a ontologia é logicamente consistente e se certo conceito se insere nela, tendo suas regras definidas pelo elemento *Rules*⁴². A OWL, na verdade, é uma extensão do modelo de dados RDF, e tem por propósito aumentar o seu poder de expressividade (WORLD WIDE WEB CONSORTIUM, 2004).

⁴⁰ Um *namespace* (NS) define um vocabulário controlado que identifica um conjunto de conceitos de forma única para que não haja ambigüidade [sic.] na sua interpretação. Os *namespaces* XML são conjuntos de tipos de elementos e atributos possíveis para cada tipo (SOUZA; ALVARENGA, 2004, p. 136).

⁴¹ *Simple Protocol and RDF Query Language* é uma “[...] linguagem que estrutura consulta semântica para RDF, que na Web Semântica tem como principal função a recuperação da informação por meio de queries, que extraem informações de fontes de dados baseado em triplas” (NUNES; MACULAN; ALMEIDA, 2020, p. 236).

⁴² Subcamada da camada ontológica que define as regras e especificações da linguagem OWL (MONTEIRO, 2013).

A camada **lógica** tem por objetivo possibilitar a combinação dos elementos das camadas anteriores, proporcionando uma interface lógica única, de forma que o computador consiga entender a semântica por meio de inferências, e realizar busca e recuperação mais eficientes. Por sua vez, a camada seguinte, **de prova**, é responsável por fazer as verificações necessárias para comprovar as inferências realizadas na camada lógica (MONTEIRO, 2013; SILVA, 2013).

Finalmente, a camada de **validação**, mais a assinatura digital e a criptografia, fornece meios para garantir a segurança, a autenticidade e a privacidade das informações, garantindo um certo grau de confiabilidade. Com a assinatura digital, por exemplo, é possível saber se a informação processada provém de uma fonte confiável (SILVA, 2013).

No contexto da arquitetura da *web* semântica é possível identificar a presença de metadados, os quais estão na terceira camada (de dados), contribuindo para a efetivação da *web* semântica ao fornecerem a estrutura e o contexto necessários para que os dados sobre ODs sejam compreendidos e utilizados de forma mais “inteligente” pelas máquinas.

2.3 METADADOS

As camadas bases da arquitetura da *web* semântica (internacional, sintática, de dados e ontológica) são essenciais para efetivar interoperabilidade entre sistemas. A interoperabilidade permite que múltiplos sistemas, com tipos de *hardware*, *software* e interfaces, possam comutar dados e informações entre si. Interoperabilidade pode ser compreendida como sendo

[...] a troca efetiva de conteúdo entre sistemas, depende de metadados que descrevem esse conteúdo para que os sistemas envolvidos possam efetivamente traçar o perfil do material recebido e combiná-lo com suas estruturas internas. (NATIONAL INFORMATION STANDARDS ORGANIZATION, 2017, p. 6, tradução nossa)

Entre os conceitos sobre metadados presentes na literatura está uma definição considerada clássica, literal, a qual se baseia na etimologia da própria palavra: metadados são "dados sobre dados". Em sentido mais amplo, metadados é

um “conjunto de elementos que descrevem as informações contidas em um recurso, com o objetivo de possibilitar sua busca e recuperação” (GRÁCIO, 2002, p. 23, grifo do autor). Assim, a utilização de metadados tem por objetivo descrever, identificar e definir um OD, a fim de garantir a sua organização, gestão, preservação e recuperação. Para tanto, pode-se considerar enriquecê-lo ou complementá-lo, aumentando o seu potencial informativo e visibilidade (ALVES, 2007; SILVA; SILVA, 2019).

Metadados são atributos que representam uma entidade (objeto do mundo real) em um sistema de informação. Em outras palavras, são elementos descritivos ou atributos referenciais codificados que representam características próprias ou atribuídas às entidades; são ainda dados que descrevem outros dados em um sistema de informação, com o intuito de identificar de forma única uma entidade (recurso informacional) para posterior recuperação. (ALVES, 2010, p. 47, grifo do autor)

Metadados possuem diferentes categorias de acordo com suas finalidades, variando em segmentos de aplicação. Neste sentido, o Quadro 4 apresenta os metadados nas perspectivas do tipo, função e uso, além de um breve conjunto de elementos (exemplos).

Quadro 4: Tipos de metadados.

Tipo		Função	Usos primários	Exemplos de elementos
Descritivos		Para encontrar ou entender um recurso	Descoberta; Exibição; Interoperabilidade.	Título; autor; assunto; gênero; data de publicação.
Administrativos	Técnicos	Para decodificar e compilar arquivos; Gerenciamento de arquivos a longo prazo; Direitos de propriedade intelectual associados ao conteúdo.	Interoperabilidade; Gerenciamento de ODs; Preservação.	Tipo de arquivo; Tamanho do arquivo; Criação data/hora; Esquema de compressão.
	De preservação		Interoperabilidade; Gestão de ODs; Preservação.	Verificação; Preservação.
	De direitos		Interoperabilidade; Gestão de ODs.	Direitos autorais; Licenças; Titular dos direitos.

Estruturais	Relacionamentos de partes de recursos entre si.	Navegação	Localização sequencial na hierarquia.
Linguagens de marcação	Integra metadados e sinalizadores ⁴³ para outros recursos estruturais ou semânticos no conteúdo.	Navegação; Interoperabilidade.	Parágrafo; Lista de cabeçalho; Autoridade; Data.

Fonte: Adaptado de Castro e Santos (2018, p. 744-745) e Riley (2017).

Buckland (2006) considera que há diferentes tipos de metadados, mas os qualifica como descritivos, pois defende que o propósito original dos metadados é descrever documentos. Ele afirma ainda que os metadados possuem dois papéis, ou seja, a descrição de ODs e a pesquisa e seleção. Ambos, no entanto, são inversamente proporcionais, pois, enquanto o primeiro parte do OD para a descrição, o segundo parte dos metadados para o OD.

Mas somente a utilização dos metadados não é suficiente para que as ações de descrição, pesquisa e seleção de ODs sejam efetivas. A interoperabilidade, como observa-se no Quadro 3, é uma iniciativa básica no contexto dos metadados em redes, a qual se faz presente em cinco, dos seis tipos de metadados apresentados.

E para que seja promovida essa “conversa” entre sistemas, deve-se considerar a forma como agentes de *software* entenderão a estrutura dos ODs representados. É neste sentido que evocam-se os padrões de metadados.

2.3.1 Padrões de metadados

A utilização apropriada de um padrão de metadados é um caminho para garantir uma RC de qualidade, contribuindo para a efetivação da interoperabilidade e recuperação de informação. É necessário não apenas conhecer as especificidades do ambiente informacional, mas entender para qual finalidade o padrão de metadados desejado foi desenvolvido (ALVES, 2010). Padrões de metadados podem ser entendidos como

⁴³ Elementos de marcação usados para identificar e delimitar partes específicas do conteúdo, sinalizando palavras com informações semânticas, ou informações sobre formatação do texto, como o itálico, por exemplo (NATIONAL INFORMATION STANDARDS ORGANIZATION, 2017).

[...] estruturas de descrição constituídas por um conjunto predeterminado de metadados (atributos codificados ou identificadores de uma entidade) metodologicamente construídos e padronizados. O objetivo do padrão de metadados é descrever uma entidade gerando uma representação unívoca e padronizada que possa ser utilizada para recuperação da mesma. (ALVES, 2010, p. 47-48, grifo do autor)

Padrões de metadados podem ser vistos quanto ao formato, variando entre simples, estruturados e formatos ricos (ALVES, 2010). O **formato simples** é composto por metadados não estruturados que são extraídos automaticamente por robôs, possuindo uma semântica limitada. Um exemplo desse tipo de formato é a utilização da *tag* <meta> na linguagem HTML, exemplificada na Figura 4. Essa *tag* possui dois atributos, possibilitando definir um nome e um conteúdo.

Figura 4: Exemplo de utilização da *tag* <meta>.

```
<meta name="color-scheme" content="dark">
<meta name="theme-color" content="#242526">
```

Fonte: Facebook⁴⁴ (2023).

O **formato estruturado** é constituído por metadados que propiciam uma descrição mínima do recurso, visando sua identificação, localização e recuperação. Um exemplo de formato que pertence a essa categoria é o padrão de metadados DC, criado em 1995 por profissionais de várias áreas do conhecimento, especificamente para a *web*, visando a interoperabilidade com base nos princípios da *web* semântica e *Linked Data*. O DC é composto por 15 elementos (Quadro 5), se destacando em ambientes mais genéricos e abrangentes, pois apresenta uma estrutura descritiva simples, que limita a representação de acervos mais específicos.

Quadro 5: Elementos DC.

Elemento		Definição
1	<i>Contributor</i>	Entidade responsável por fazer contribuições para o recurso.
2	<i>Coverage</i>	A aplicabilidade espacial do recurso ou a jurisdição sob a qual o recurso é relevante.
3	<i>Creator</i>	Entidade primariamente responsável por fazer o recurso.
4	<i>Date</i>	Período de tempo associado a um evento no ciclo de vida do recurso, como, por exemplo, data de publicação.

⁴⁴ Acesso: <https://www.facebook.com/>.

5	<i>Description</i>	Descrição do conteúdo do recurso, como um resumo, um sumário.
6	<i>Format</i>	O formato de arquivo, meio físico ou dimensões do recurso.
7	<i>Identifier</i>	Uma referência inequívoca ao recurso dentro de um determinado contexto, como URI e URL, por exemplo.
8	<i>Language</i>	A linguagem do recurso.
9	<i>Publisher</i>	Entidade responsável por disponibilizar, publicar o recurso.
10	<i>Relation</i>	Referência a um recurso relacionado.
11	<i>Rights</i>	Informações sobre os direitos sobre o recurso, como Direitos de Propriedade Intelectual e Direitos Autorais.
12	<i>Source</i>	Recurso do qual o recurso descrito é derivado.
13	<i>Subject</i>	O assunto do recurso, expresso por palavra-chave ou código de classificação (como a Classificação Decimal Universal - CDU ⁴⁵).
14	<i>Title</i>	Nome dado ao recurso.
15	<i>Type</i>	A natureza ou gênero do recurso, como imagem, som e texto.

Fonte: *Dublin Core Metadata Initiative* - DCMI (2012) e Hillmann (2005).

Na perspectiva das linguagens de marcação, o DC utiliza a linguagem XML para ser codificado, assim como esta pode ser inserida em páginas HTML. A Figura 5 apresenta um exemplo de código XML com DC, ilustrando a presença de três dos 15 elementos elencados no Quadro 4. Consecutivamente, os elementos são *description*, *title* e *publisher*.

Figura 5: Registro descrito em DC, codificado em XML.

```
<?xml version="1.0"?>
<dcx:descriptionSet
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcx="http://purl.org/dc/xml/">
  <dcx:description
    dcx:resourceURI="http://dublincore.org/pages/home">
    <dc:title>
      <dcx:valueString>DCMI Home Page</dcx:valueString>
    </dc:title>
    <dc:publisher>
      <dcx:valueString>Dublin Core™ Metadata Initiative</dcx:valueString>
    </dc:publisher>
  </dcx:description>
```

Fonte: Elaborado pela autora (2023), com base em DCMI (2006).

⁴⁵ Classificação Decimal Universal – Instrumento para a classificação de materiais bibliográficos, utilizado para a organização de bibliotecas, idealizado por Paul Otlet e Henry La Fontaine. Sua primeira edição foi publicada entre 1904 e 1907 (CARVALHO, 2009, p. 45).

Por fim, tem-se o **formato rico**, com padrões altamente estruturados e complexos, apresentando uma descrição formal e detalhada. Baseados em normas e códigos especializados de um domínio em particular, o formato MARC 21 é um significativo exemplo. O MARC 21 é um padrão de metadados desenvolvido nos anos 1960 pela *Library of Congress* (LC), particularmente, mas não somente, para descrever itens bibliográficos, de forma a intercambiar esses dados de maneira automatizada (ALVES, 2010). O MARC 21, tem sua estrutura (Quadro 6) baseada na semântica do *Anglo American Cataloguing Rules, second edition* (AACR2)⁴⁶, o qual apresenta uma estrutura complexa com campos, subcampos e indicadores⁴⁷.

Quadro 6: Campos do MARC 21.

Campos	Descrição
0XX	Informações de controle, números e códigos.
1XX	Entrada principal (autoria).
2XX	Títulos, edição, impressão etc.
3XX	Descrição física.
4XX	Série.
5XX	Notas.
6XX	Entradas de assunto.
7XX	Entradas secundárias (autoria e título).
8XX	Entrada secundária de série.
9XX	Uso local.

Fonte: Adaptado de Alves e Souza (2007).

⁴⁶ O AACR2 “constitui-se como um conjunto de regras e normas para o estabelecimento de uma padronização na representação de diversos recursos informacionais. Caracteriza-se por ser um código abrangente e detalhado, e devido a sua aceitação passou a ser utilizado no ensino de catalogação nos cursos de graduação em biblioteconomia brasileiros; além de ser considerado como um código internacional para a construção de formas de representação bibliográfica. O AACR2 comporta a descrição de qualquer tipo de informação independente do suporte, caracteriza-se por possuir um formalismo em sua estrutura de representação, pois estabelece, por meio de suas regras, uma relação semântica entre os elementos descritos, além disto, apresenta uma estrutura coerente, lógica e de fácil memorização” (ALVES, 2005, p. 48-49).

⁴⁷ Campos são unidades lógicas que representam uma informação bibliográfica, subcampos são elementos descritivos que especificam ainda mais a informação a ser representada e, por fim, indicadores são as duas primeiras posições de caracteres após os números que identificam o campo (ALVES, 2010).

Enquanto no Quadro 5 estão expostos os campos básicos do MARC 21, a Figura 6 apresenta um exemplo de registro completo, contendo campos (campo 245, relativo a título, edição e impressão, por exemplo), subcampos (letras “a”, referente ao título, e “b”, alusiva ao complemento do título, por exemplo) e indicadores (números “1”, que indica a entrada secundária de título, por exemplo). No caso, é possível observar a utilização de códigos ao invés de nomes de atributos.

Figura 6: Trecho de um registro descrito em MARC 21.

010	—	a 2011500780
020	—	a 9780199560288 (pbk.)
040	—	a DLC c DLC d DLC
042	—	a pcc
050	00	a P325 b .G65 2011
082	00	a 401/.43 2 23
100	1_	a Goddard, Cliff.
245	10	a Semantic analysis : b a practical introduction / c Cliff Goddard.
250	—	a 2nd ed.
260	—	a Oxford ; a New York : b Oxford University Press, c 2011.
300	—	a xvii, 490 p. : b ill., map ; c 25 cm.
490	0_	a Oxford textbooks in linguistics
500	—	a Previous ed.: 1998.
504	—	a Includes bibliographical references (p. [447]-475) and indexes.
650	_0	a Semantics.

Fonte: Goddard (2011).

Os formatos MARC e DC são padrões de metadados que, assim como outros, podem ser mapeados para utilizar a sintaxe e a semântica do RDF. Como apontado no capítulo anterior, o RDF é uma recomendação do W3C para ser implementada como parte da camada de dados da arquitetura da *web* semântica. O RDF permite fazer afirmações sobre ODs utilizando um esquema de triplas, ou seja, sujeito-predicado-objeto. Sujeito é entendido como uma entidade, objeto é uma outra entidade que se conecta ao sujeito e o predicado é o que descreve esse relacionamento. Essa tripla também é conhecida como recurso-propriedade-valor, e elas vão se relacionando a ponto de formar grafos complexos (LAUFER, 2015; NATIONAL INFORMATION STANDARDS ORGANIZATION, 2017; SOUZA; ALVARENGA, 2004). A Figura 7 apresenta um exemplo de tripla RDF, sendo o livro “*Dom Casmurro*” o sujeito (ou recurso), enquanto o predicado (ou propriedade) “foi escrito por” define a relação com o objeto (ou valor), que no caso é o autor “Machado de Assis”.

Figura 7: Triplas RDF

Fonte: Elaborada pela autora (2023), com base em Laufer (2015) e Riley (2017).

O RDF utiliza a linguagem XML para efetivar a interoperabilidade, através da utilização de *namespaces*, que serão especificados ao longo do código, garantindo que os descritores sejam utilizados de forma não-ambígua, podendo ser, inclusive, referenciados por meio de URI, para que a sintaxe e as propriedades semânticas sejam legíveis por máquinas.

3 METODOLOGIA

O ser humano, buscando compreender a si e ao mundo ao seu redor, recorre a diversos tipos de saberes, entre populares, teológicos, filosóficos, empíricos e científicos, e este último é construído por meio de pesquisas, as quais podem ser compreendidas como um “[...] procedimento racional e sistemático que tem como objetivo proporcionar respostas aos problemas que são propostos” (GIL, 2002, p. 17). Entre as partes que constituem uma pesquisa científica está a metodologia, ou seja, um conjunto de etapas que requer a definição do caminho pelo qual os resultados serão alcançados, isto é, “[...] métodos ou caminhos que são percorridos na busca do conhecimento” (ANDRADE, 2010, p. 117). A metodologia, então, tem a função de explicitar um conjunto de abordagens, técnicas e processos que são utilizados por pesquisadores para resolver problemas e gerar conhecimento.

3.1 CARACTERÍSTICAS

A presente pesquisa baseia-se no método dedutivo de raciocínio, pois parte de premissas verdadeiras e gerais, e por meio de pesquisa busca alcançar conclusões lógicas e específicas. Conforme Gil (2008), o método dedutivo foi proposto por racionalistas que defendiam razão como o caminho que levará ao conhecimento verdadeiro, baseando-se no silogismo que parte de duas premissas para gerar uma terceira, que se denomina conclusão.

Referente à natureza, trata-se de uma pesquisa básica, pois não possui aplicação prática, tendo o interesse de considerar verdades e reforçar compreensões identificadas na literatura. Trata-se de um tipo de natureza da pesquisa que “objetiva gerar conhecimentos novos e úteis para o avanço da ciência sem aplicação prática prevista” (SILVA; MENEZES, 2005, p. 20).

Quanto aos objetivos, a pesquisa qualifica-se como descritiva, pois se propôs descrever as “[...] características de determinada população ou fenômeno ou o estabelecimento de relações entre variáveis” (GIL, 2008, p. 28), no caso, identificar e descrever aspectos da *web* semântica, envolvendo as contribuições dos metadados.

Considerando os procedimentos técnicos, a pesquisa caracteriza-se como bibliográfica e documental. É uma pesquisa bibliográfica porque baseia-se em materiais como livros e artigos científicos, enquanto é documental porque utiliza materiais que não passaram por revisões de pares científicos, como *sites*, vídeos e documentações técnicas, por exemplo. Segundo Gil (2008), as pesquisas bibliográfica e documental se assemelham, diferenciando-se pela natureza das fontes, pois a primeira se vale da contribuição de diversos autores, enquanto a segunda baseia-se em materiais que não receberam tratamento analítico.

Do ponto de vista da abordagem do problema, a pesquisa é qualitativa, isto é, não faz uso de técnicas estatísticas, tratando os dados de forma subjetiva, voltada à interpretação. Para Silva e Menezes (2005, p. 20), “a interpretação dos fenômenos e a atribuição de significados são básicas no processo de pesquisa qualitativa”.

3.2 APRESENTAÇÃO DOS RESULTADOS DA PESQUISA

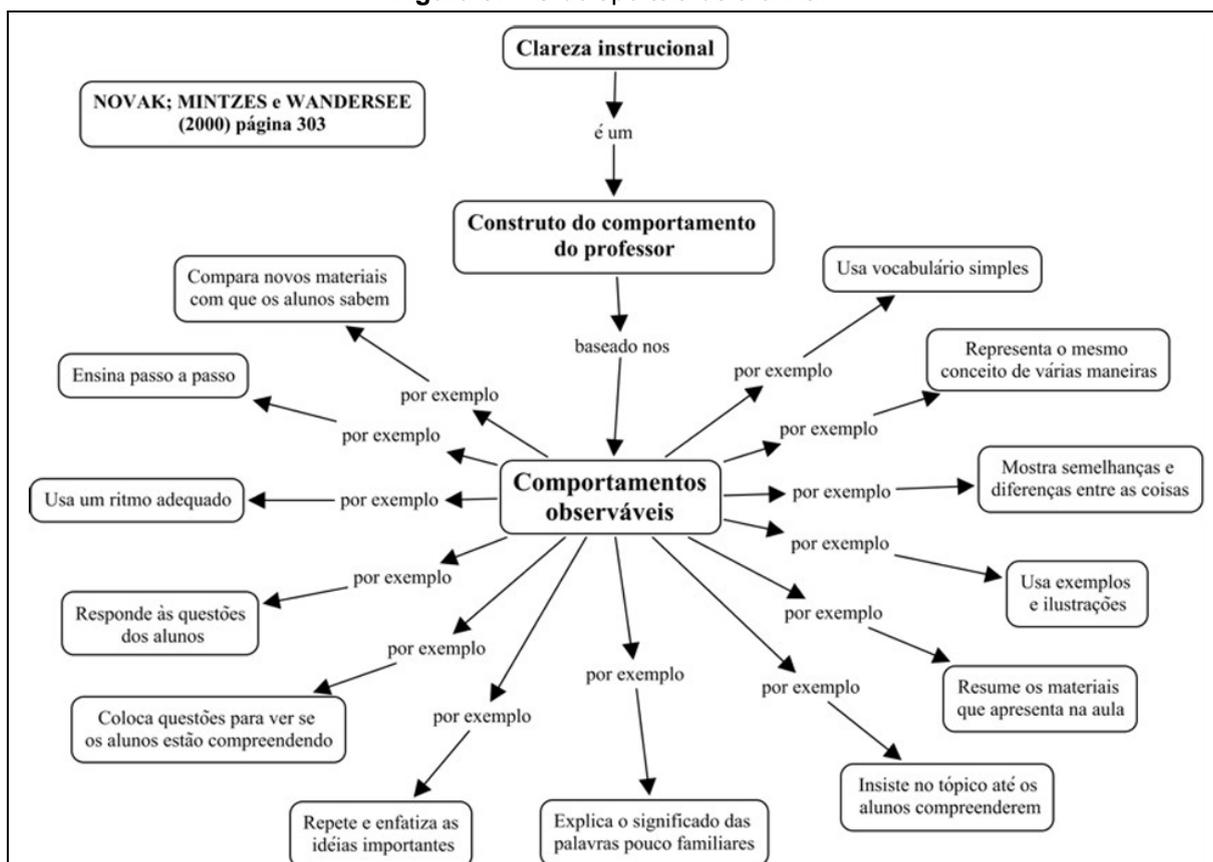
Para apresentação dos resultados, seguiu-se para a utilização de uma RC como forma de categorizar as contribuições dos metadados identificadas na literatura para a efetivação da *web* semântica, assim como outras características marginais. Como tipo de RC, optou-se pelos mapas conceituais (MCs), pois são diagramas que indicam “[...] relações entre conceitos, ou entre palavras que usamos para representar conceitos” (MOREIRA, 2010, p. 1). A ideia básica dos MCs está na Teoria da Aprendizagem Significativa, também conhecida como Teoria Cognitivista da Aprendizagem, partindo do pressuposto de que quando se possui uma informação nova, torna-se útil a ancoragem dessa nova informação há uma preexistente, servindo de base para a compreensão da nova informação. Foi justamente a partir dessa teoria que Joseph Novak⁴⁸ se inspirou para desenvolver a técnica de mapeamento conceitual (MOREIRA, 2010).

Existe uma variedade de MCs, que servem a diversos propósitos. Nesta pesquisa, adota-se um MC do tipo teia de aranha, devido à multiplicidade de

⁴⁸ Novak é um norte-americano, educador e professor da Universidade de Cornell (Estados Unidos da América), mundialmente conhecido pela criação da técnica de elaboração de MCs. A sua ideia intencionou a criação de uma ferramenta que atende-se ao lema o ato de “aprender a aprender” (ARAÚJO, 2019).

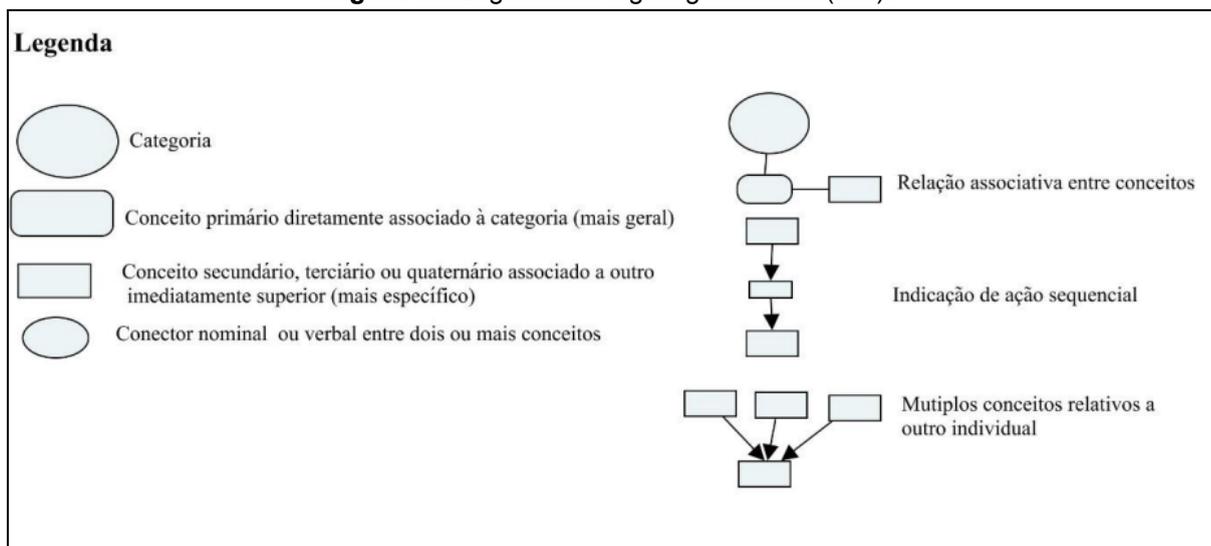
conceitos e a facilidade de estruturar as informações. A forma como o MC organiza-se ajuda na compreensão de um conceito, de um domínio. Conforme Tavares (2007), o conceito principal localiza-se no centro da representação, enquanto os demais vão se irradiando na medida que se afastam do centro.

Figura 8: MC do tipo teia de aranha.



Fonte: Tavares (2007).

Além da estrutura do MC do tipo teia de aranha, para facilitar a compreensão do conceito, e da consequente apresentação de suas características, adota-se a diferenciação das categorias por cores, assim como cada conceito tem um formato geométrico específico. Trata-se de um método adotado por Araújo (2019) e Araújo e Silva (2021) para conceituar repositório digital (RD), considerando um conjunto de características, ditas essenciais.

Figura 10: Legenda da lógica geométrica (MC).

Fonte: Araújo (2019) e Araújo e Silva (2021).

A utilização de um MC neste trabalho visa a representação visual, de forma clara e estruturada, das relações entre os conceitos identificados na literatura para explicar a contribuição de metadados na efetivação da *web* semântica. Para a construção do MC foi utilizado o *software* livre *CmapTools*, versão 6.04, devido a sua gratuidade e aos seus diversos recursos disponíveis (e intuitivos) para construir representações do tipo.

3.2 ETAPAS DA PESQUISA

Em suma, as fases desta pesquisa, desde a definição do tema, demarcação dos problemas e dos objetivos, até a apresentação dos resultados e considerações finais, ocorreram em 10 etapas:

1. Definição do tema, dos problemas da pesquisa e dos objetivos;
2. Realização do levantamento bibliográfico em livros, artigos científicos e trabalhos acadêmicos (monografias, teses e dissertações), localizados em periódicos científicos, bases de dados, repositórios e bibliotecas digitais;
3. Execução do levantamento documental em materiais como *sites* e documentações capitaneadas pelo W3C;
4. Produção da fundamentação teórica conforme os objetivos da pesquisa;

5. Realização de pesquisas no período de fevereiro de 2022 a abril de 2023, a partir de um conjunto de expressões de busca, entre outras: “*web semântica*” OR “*semantic web*”, “metadados” OR “metadata”, e “interoperabilidade”;
6. Definição das características que respondem ao problema da pesquisa;
7. Categorização das características identificadas;
8. Elaboração de um quadro contendo as características categorizadas;
9. Confeção do MC do tipo aranha;
10. Apresentação dos resultados.

4 RESULTADOS

Os resultados atingidos pela pesquisa contemplam as tecnologias que compõem a *web* semântica, os atributos e aplicações dos metadados e apontamentos sobre a forma como os metadados contribuem para a efetivação da *web* semântica. Os aspectos identificados na literatura levantada estão organizados em quadros, os quais atuam como instrumentos norteadores para o MC.

4.1 ATRIBUTOS E APLICAÇÕES DOS METADADOS

Os metadados são dados que descrevem outros dados. Eles fornecem contexto, detalhes e características sobre os dados, permitindo sua organização, preservação, busca e interpretação. Há variados tipos de metadados, os quais diferem de acordo com o recurso e as necessidades de representação.

Nesta pesquisa, entende-se que os metadados podem ser classificados por quatro segmentos. O primeiro são os metadados **descritivos**, que são pensados, assim como mencionado em seu nome, para descrever e identificar ODs, a fim de que estes possam ser buscados, descobertos. Também se menciona os metadados administrativos, que visam o gerenciamento de arquivos, a preservação (e a indicação de propriedade) e os direitos autorais. Os metadados **estruturais** relacionam partes de recursos entre si. Por fim, existem os metadados de **linguagem de marcação**, que visam a navegação, integrando metadados a outros recursos estruturais ou semânticos no conteúdo. Vale destacar que entre todos os elencados tipos de metadados estão os atributos, sendo comumente adotados no campo bibliográfico os seguintes: título; autor; assunto; gênero; data; formato; tipo e tamanho de arquivo; e direitos autorais.

Em tese, a utilização de metadados garante que os mais diversos *softwares* operacionalizem trocas de dados de maneira eficiente, sendo este um aspecto a ser considerado entre os padrões de metadados, pois, em sua essência, eles são estruturas e especificações estabelecidas que intencionam ofertar uniformidade na descrição e promover visibilidade dos ODs.

Os padrões de metadados podem variar em sua complexidade, sendo categorizados como: simples; estruturados; e ricos. Os formatos **simples** possuem dados semi-estruturados, segundo a recuperação realizada (automaticamente) por robôs, além de apresentarem semântica reduzida, justamente pela natureza básica dessas informações. Os formatos **estruturados** são baseados em normas emergentes, e possuem uma estrutura mais ampla de descrição, considerando características básicas, técnicas e de relacionamentos, além de proporcionar o armazenamento da informação em campos criados para este fim, o que facilita a recuperação. Finalmente, os formatos **ricos** são encarados como complexos, diante da sua ampla gama de elementos para descrever os ODs, além de serem baseados em normas especializadas e códigos específicos.

Os metadados na *web* semântica são encontrados nas camadas base, contextos onde se encontram padrões, assim como o RDF, enquanto uma tecnologia que tem a sua utilização recomendada pelo W3C, configurando-se, assim, um modelo semântico de dados. Por meio de triplas, o RDF oferece uma maneira padronizada de descrever ODs, e os relacionamentos entre eles, no caso, utilizando URIs. Quanto a sua capacidade, o W3C recomenda o RDF pela sua forma de expressão, ou seja, via linguagem XML, aspecto este que contribui para o compartilhamento de informações (estruturadas), de maneira padronizada e compreensível por máquinas. Como resultado, há uma facilitação na integração e interpretação de informações em diferentes sistemas na *web*.

Em suma, os metadados são aplicados de acordo com a necessidade do ambiente digital, e assim a sua utilização varia. Metadados podem ser utilizados para descoberta, no âmbito de busca e pesquisa, para gerenciamento e preservação de ODs e para viabilizar compartilhamento de informações, efetivando a interoperabilidade entre sistemas.

4.2 TECNOLOGIAS DA *WEB* SEMÂNTICA

A *web* semântica foi idealizada por Tim Berners-Lee, vista como uma extensão da *web* que deve apresentar informações enriquecidas com significado semântico, permitindo, assim, que máquinas compreendam, interpretem e

processem dados, de forma “inteligente” e automatizada. O projeto da *web* semântica essencialmente é a criação e implantação de normas e tecnologias que permitam o compartilhamento de dados entre dispositivos e sistemas de informação, com fins de efetiva e mais ampla interoperabilidade.

Como apontado no referencial teórico, a *web* semântica é frequentemente descrita em termos de camada, que desempenham papéis específicos na construção de uma *web* que privilegie o significado das coisas. No Quadro 7 é possível verificar as tecnologias que compõem as quatro primeiras camadas, juntamente com a sua correspondente descrição.

Quadro 7: Tecnologias da *web* semântica.

Camada	Tecnologia	Descrição
Faixa internacional	<i>Unicode</i>	Padrão de codificação que atribui um número único para cada caractere de texto, abrangendo um vasto conjunto de sistemas de escrita, de diferentes idiomas, permitindo a representação e o processamento de texto multilíngue de forma consistente.
	URI	Sequência de caracteres que identifica de forma única um recurso na <i>web</i> , permitindo sua localização e acesso por meio de um identificador único.
Camada sintática	XML	Linguagem de marcação que permite a criação de documentos estruturados, fornecendo uma forma flexível de representar e armazenar dados de forma legível tanto para humanos quanto para máquinas.
Camada de dados	RDF	Modelo de dados para representar informações estruturadas de forma flexível e interoperável, utilizando triplas, ou seja, sujeito-predicado-objeto.
	SPARQL	Linguagem de consulta para recuperar e manipular dados armazenados em formato RDF. Permite consultas complexas e expressivas em bancos de dados RDF.
Camada ontológica	OWL	Linguagem para criar ontologias, que são representações formais de conceitos, propriedades e relações entre entidades em um domínio específico.

Fonte: Da pesquisa (2023).

Apesar de serem quatro camadas descritas no Quadro 6, a *web* semântica é dividida em sete: faixa internacional; camada sintática; camada de dados; camada ontológica; camada lógica; camada de prova; e camada de validação. As três

últimas camadas não constam no Quadro 6 por serem compostas, basicamente, por normas e diretrizes, e não possuírem tecnologias específicas vinculadas a elas.

As quatro primeiras camadas, também chamadas de camadas de base, são vitais para a interoperabilidade que a *web* semântica tanto preza. As tecnologias que as permeiam são fundamentais para manipular os dados de forma que se obtenha um alto nível de padronização, através da utilização dos metadados. Neste sentido, a camada sintática, que é a camada que forma a estrutura da *web* semântica, é onde se localizam os metadados, estruturados pela linguagem XML, enquanto nas camadas seguintes, de dados e ontológica, os metadados recebem semântica a partir do modelo de dados RDF e da linguagem OWL.

4.3 CONTRIBUIÇÕES DE METADADOS NA WEB SEMÂNTICA

A *web* semântica, sendo uma *web* de dados, foi idealizada para tornar os conteúdos mais significativos para computadores, e para tal se faz necessário a utilização de vários recursos. O Quadro 8 é uma sintetização das contribuições de cada recurso que compõem a *web* semântica, identificadas no levantamento teórico em materiais da CI.

Quadro 8: Recursos da *web* semântica e suas contribuições.

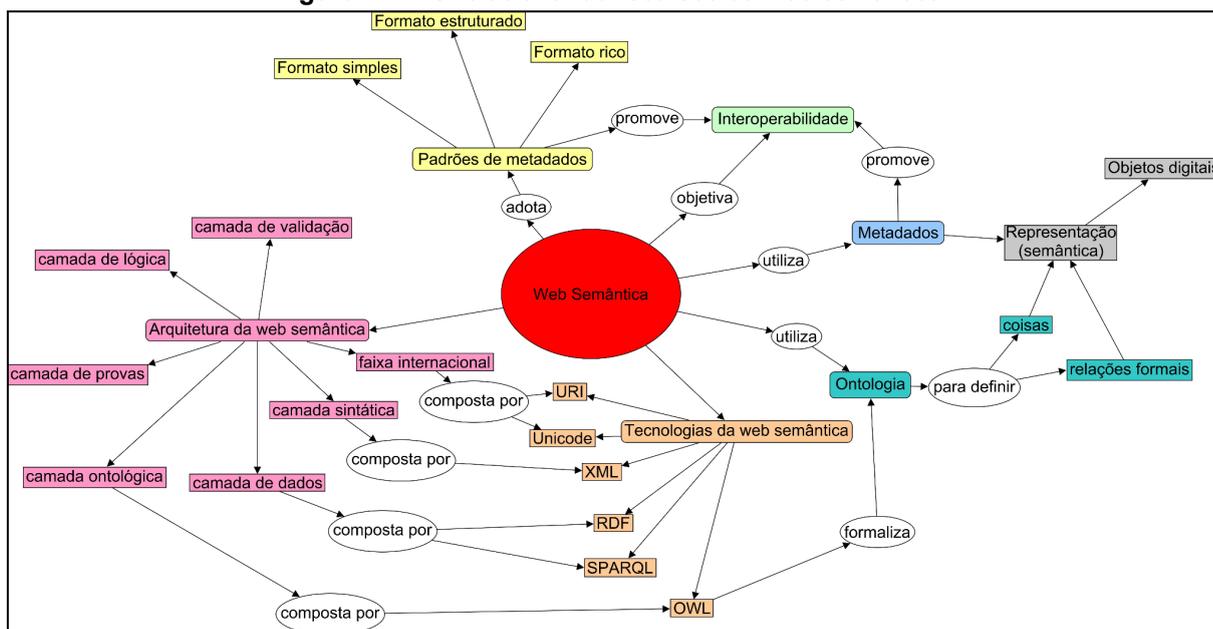
Recurso	Contribuição
Metadados	Os metadados desempenham no âmbito da <i>web</i> semântica o papel de garantir formas de representação de ODS , fornecendo informações que possibilitem troca de dados (ALVES, 2005; SILVA, 2013, grifo nosso).
Padrão de metadados	A <i>web</i> semântica adota padrões de metadados a fim de representar (com qualidade) as informações disponíveis na <i>web</i> e, assim, promover a interoperabilidade . Os padrões, ou formatos de metadados, são classificados entre simples, estruturados ou ricos (ALVES, 2005; SILVA, 2013; SOUZA; ALVARENGA, 2004, grifos nossos).
Arquitetura da <i>web</i> semântica	A arquitetura da <i>web</i> semântica é estruturada em sete camadas, ou seja, faixa internacional, camada sintática, camada de dados, camada ontológica, camada de lógica, camada de provas e camada de validação . Essas camadas são permeadas de diretrizes e tecnologias que juntas possibilitam a existência da <i>web</i> semântica (ALVES, 2005; MONTEIRO, 2013; SILVA, 2013, grifos nossos).

Tecnologias da <i>web</i> semântica	As tecnologias que visam embutir semântica na <i>web</i> , e são recomendadas pelo W3C, são: padrão de codificação Unicode , localizador de recursos URI , linguagem de marcação XML , modelo de dados RDF , linguagem de consulta SPARQL e linguagem ontológica OWL (ALVES, 2005; SILVA, 2013, grifos nossos).
Interoperabilidade	A <i>web</i> semântica promove a interoperabilidade, que é viabilizada pela utilização de metadados e padrões na descrição de ODs, permitindo que eles sejam compreendidos por outros programas ou sistemas (ALVES,2005; SILVA, 2013, grifos nossos).
Ontologia	As ontologias determinam semanticamente as coisas (como conceitos apresentados pelos metadados), fornecendo estrutura e vocabulário para representação de ODs, além de definir relações formais entre termos e conceitos, por meio de uma sintaxe estrutural clara. Na <i>web</i> semântica, a ontologia está formalizada via linguagem OWL (ALVES, 2005; SILVA, 2013, grifos nossos).

Fonte: Da pesquisa (2023).

Diante das contribuições identificadas, consideradas para fins de RC como características, as quais foram categorizadas em metadados, padrões de metadados, arquitetura da *web* semântica, tecnologias da *web* semântica, interoperabilidade e ontologia, o MC proposto (Figura 11) explicita a relação dos recursos que compõem a *web* semântica e como os metadados, neste contexto, se fazem presentes.

Figura 11: MC relacionando recursos da web semântica



Fonte: Da pesquisa (2023).

Os metadados estão presentes na base da arquitetura da *web* semântica, e através de tecnologias a semântica dos dados se torna compreensível por máquinas. Em um cenário prático, utilizando os metadados combinados com uma ontologia originada da linguagem OWL, é possível definir conceitos e relações formais na representação de ODs, o que acarreta em uma efetiva recuperação de informações, assim como defende Silva (2013, p. 47) “o emprego de metadados para otimizar a recuperação da informação na Web é a aplicação mais importante dos metadados [...]”.

No contexto da *web* semântica, de forma ampla, infere-se que os metadados “[...] fornecem informações adicionais sobre os dados, para ajudar desenvolvedores de aplicações e usuários finais a entender melhor o significado dos dados publicados, o seu conteúdo, a sua estrutura” (LAUFER, 2015, p. 20). Ratifica-se, portanto, a ideia de que padrões de metadados contribuem diretamente na operacionalização da interoperabilidade entre sistemas *web*, especialmente quando se fala na representação de ODs e recuperação de informações.

5 CONSIDERAÇÕES FINAIS

A “explosão informacional” trouxe consigo novas preocupações às áreas relacionadas à informação, como é o caso da Biblioteconomia e da CI. Esta preocupação evoca a importância dos metadados para a descrição, visibilidade e recuperação de recursos informacionais na *web* semântica. Ao fundamentar-se na RC, a *web* semântica utiliza metadados para estabelecer regras de inferência, as quais resultam em uma melhor organização e tratamento de ODs, incluindo o intercâmbio de dados entre diferentes agentes de *softwares*. Os metadados, no entanto, por si, não garantem a interoperabilidade, mas, em conjunto com padrões, diretrizes e tecnologias, permitam aos ODs, e suas relações, serem representados de maneira precisa e consistente.

Através da utilização das tecnologias que permeiam a arquitetura da *web* semântica, indicada pelo W3C, é possível garantir metadados padronizados e bem definidos, mas julga-se importante reforçar que as quatro primeiras camadas são a base, a partir de tecnologias que trabalham em conjunto para dar suporte à uma *web* mais significativa, compreensível não apenas pelos homens, mas entre máquinas.

Conforme a literatura levantada, apresenta-se um destaque à faixa internacional da *web* semântica, incluindo o padrão de codificação *Unicode*, ao permitir que computadores e dispositivos de diferentes sistemas operacionais se comuniquem, exibindo corretamente os caracteres, independente do idioma utilizado, e o localizador de recursos URI, que identifica cada OD de forma única, para que não ocorra ambiguidade. Por sua vez, na camada sintática, observa-se a presença da linguagem de marcação XML, empregada para representar dados estruturados, frequentemente utilizada em conjunto com o modelo de dados RDF, localizado na camada de dados, juntamente com a linguagem de consulta SPARQL, sendo esta importante para consultas complexas em bancos de dados RDF. Finalmente, a camada ontológica abrange a linguagem ontológica OWL, que define conceitos e relações entre ODs.

A interoperabilidade é mais um conceito chave na *web* semântica, identificado na literatura. Ela é a figura que atua diretamente na garantia da semântica na *web*, pois é a forma com que os agentes de *softwares* conseguem se comunicar.

Apointa-se, então, que os metadados, através de tecnologias e padrões desenvolvidos pelo W3C, e apontados durante a pesquisa, torna possível a interoperabilidade entre sistemas.

No contexto da *web* semântica, constituída de uma arquitetura específica, segundo camadas, entre tecnologias, linguagens e ontologias, na definição de relações formais entre coisas (conceitos), conclui-se que os metadados contribuem na representação de ODs, na oferta de padrões, entre simples, estruturados e ricos, e na promoção da interoperabilidade.

Espera-se que esta pesquisa fomente a discussão sobre metadados, especialmente no campo da CI, acompanhando as atualizações propostas pelo W3C em tecnologias e diretrizes para a efetivação da *web* semântica. Além disso, o breve contato com a literatura sobre *web* semântica fomenta o interesse em aprofundar os estudos nesta camada da Internet, observando a atuação do profissional da informação, como é o caso dos bibliotecários, frente a necessidade de tratar, organizar, permitir recuperação e preservar as informações na *web*.

REFERÊNCIAS

ALMEIDA, F. L. Concept and Dimensions of Web 4.0. **International Journal of Computers & Technology**, [S. l.], v. 16, n. 7, p. 7040–7046, 2017. Disponível em: <https://rajpub.com/index.php/ijct/article/view/6446>. Acesso em: 19 abr. 2023.

ALVARENGA, L. Representação do conhecimento na perspectiva da ciência da informação em tempo e espaço digitais. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 8, n. 15, p. 18-40, 2003. Disponível em: <https://brapci.inf.br/index.php/res/v/36962>. Acesso em: 25 jul. 2022.

ALVES, M. D. R.; SOUZA, M. I. F. Estudo de correspondência de elementos metadados: Dublin Core e MARC 21. **RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 5, n. 1, p. 20-38, 2007. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/2019>. Acesso em: 10 ago. 2022.

ALVES, R. C. V. **Metadados como elementos do processo de catalogação**. 2010. 134 f. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010. Disponível em: <https://repositorio.unesp.br/handle/11449/103361>. Acesso em: 28 jun. 2022.

ALVES, R. C. V. **Web semântica: uma análise focada no uso de metadados**. 2005. 180 f. Dissertação (Mestrado em Ciência da Informação) - Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, 2005. Disponível em: <http://hdl.handle.net/11449/93690>. Acesso em: 15 maio 2023.

ANDRADE, I. A. et al. Inteligência coletiva e ferramentas web 2.0: a busca da gestão da informação e do conhecimento em organizações. **Perspectivas em Gestão & Conhecimento (PG&C)**, João Pessoa, v. 1, p. 27-43, 2011. Disponível em: <https://periodicos.ufpb.br/ojs2/index.php/pgc/article/view/10385>. Acesso em: 22 abr. 2023.

ANDRADE, M. M. **Introdução à metodologia do trabalho científico**. 10. ed. São Paulo: Atlas, 2010. 158 p.

ARAÚJO, D. O. **Repositórios digitais: um estudo de características a partir de modelos categoriais**. 2019. 171 f. Trabalho de Conclusão de Curso (Bacharelado em Biblioteconomia) - Universidade de Brasília, Brasília, 2019. Disponível em: <https://bdm.unb.br/handle/10483/25859>. Acesso em: 12 jul. 2023.

ARAÚJO, D. O.; SILVA, M. B. Repositório digital: a delimitação de um conceito por meio de mapa conceitual. **Informação@Profissões**, Londrina, v. 10, n. 3, p. 18–33, 2021. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/infoprof/article/view/44878>. Acesso em: 13 jul. 2023.

BACHELARD, G. **A formação do espírito científico**. 3. ed. Rio de Janeiro: Contraponto, 2002. 316 p.

BERNERS-LEE, T. Artificial Intelligence and the Semantic Web. [S. l.], **W3C**, 2006. Disponível em: [http://www.w3.org/2006/Talks/0718-aaai-tbl/Overview.html#\(1\)](http://www.w3.org/2006/Talks/0718-aaai-tbl/Overview.html#(1)). Acesso em: 13 jul. 2023.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. 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, 2001, p. 28-37. Disponível em: https://www.researchgate.net/publication/225070375_The_Semantic_Web_A_New_Form_of_Web_Content_That_is_Meaningful_to_Computers_Will_Unleash_a_Revolution_of_New_Possibilities. Acesso em: 02 nov. 2021.

BORKO, H. Information science: what is it?. **American Documentation**, [S. l.], v.19, n.1, p. 3-5, Jan. 1968. Disponível em: <https://www.marilia.unesp.br/Home/Instituicao/Docentes/EdbertoFerneda/mri-01---information-science---what-is-it.pdf>. Acesso em: 04 nov. 2021.

BRANSKI, R. M. Localização de informações na internet: características e formas de funcionamento dos mecanismos de busca. **Transinformação**, Campinas, v. 12, n. 1, p. 11-19, 2000. Disponível em: <https://www.scielo.br/j/tinf/a/pwwhV5xN6FBtSYy8CJWJnvL/?lang=pt>. Acesso em: 10 jul. 2023.

BRASCHER, M; CAFÉ, L. Organização da Informação ou Organização do Conhecimento?. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 9., 2008, São Paulo. **Anais...** São Paulo: ENANCIB, 2008. Disponível em: <https://brapci.inf.br/index.php/res/v/176535>. Acesso em 11 mar. 2023.

BROOKSHEAR, J. G. **Ciência da Computação**: uma visão abrangente. 11. ed. Porto Alegre: Bookman, 2013. 501 p.

BUCKLAND, M. K. Descrição e pesquisa: metadados como infra-estrutura. **Brazilian Journal of Information Science: research trends**, [S. l.], p. 3–14, 2007. Disponível em: <https://revistas.marilia.unesp.br/index.php/bjis/article/view/26>. Acesso em: 14 jul. 2023.

CAFÉ, L.; SALES, R. Organização da informação: conceitos básicos e breve fundamentação teórica. In: ROBREDO, J.; BRÄSCHER, M. (Org.). **Passeios pelo bosque da informação: estudos sobre a representação e organização da informação e do conhecimento**. Brasília: IBICT, 2010. p. 115-129. Disponível em: <https://www2.senado.leg.br/bdsf/item/id/189812>. Acesso em: 14 ago. 2022.

CAMPOS, M. L. A; GOMES, H. E. Taxonomia e classificação: a categorização como princípio. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 8., 2007, Salvador, Bahia. **Anais...** Salvador: ENANCIB, 2007. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/172684>. Acesso em: 01 mar. 2023.

CARVALHO, R. A. **Perspectivas na web semântica para a ciência da informação**. 2009. 186 f. Dissertação (Mestrado em Ciência da Informação) - Pontifícia Universidade Católica de Campinas, Centro de Ciências Humanas e Sociais Aplicadas, 2009. Disponível em: <http://repositorio.sis.puc-campinas.edu.br/xmlui/handle/123456789/14788>. Acesso em: 20 jun. 2023.

CASTRO, F. F. **Padrões de representação e descrição de recursos informacionais em bibliotecas digitais na perspectiva da ciência da informação: uma abordagem do MarcOnt initiative na era da web semântica**. 2008. 201 f. Dissertação (Mestrado em Ciência da Informação) - Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, 2008. Disponível em: <http://hdl.handle.net/11449/93689>. Acesso em: 31 jan. 2023.

CASTRO, F. F.; SANTOS, P. L. V. A. C. Metadados em ciência da informação: trajetória científica no Brasil. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 19., 2018, Londrina, Paraná. **Anais...** Paraná: ENANCIB, 2018. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/102842>. Acesso em: 27 jun. 2023.

COUTINHO, C. P.; BOTTENTUIT JUNIOR, J. B.. Blog e wiki: os futuros professores e as ferramentas da web 2.0. In: SIMPÓSIO INTERNACIONAL DE INFORMÁTICA EDUCATIVA, 9., 2007, Porto, Portugal. **Anais...** Porto: ESSE-IPP, 2007. p. 199-204. Disponível em: <https://repositorium.sdum.uminho.pt/handle/1822/7358>. Acesso em: 01 nov. 2021.

DUBLIN CORE METADATA INITIATIVE. Dublin Core Metadata Element Set, Version 1.1: Reference Description. **Dublin Core Metadata Initiative (DCMI)**, [S. l.], 2012. Disponível em: <https://www.dublincore.org/specifications/dublin-core/dces/>. Acesso em: 03 jul. 2023.

EMYGDIO, J. L. *et al.* Ensaio sobre ontologia aplicada na recuperação da informação para a ciência da informação. **Ponto de Acesso**, Salvador, v. 15, n. 3, p.323-343, 2021. Disponível em: <https://periodicos.ufba.br/index.php/revistaici/article/view/47471>. Acesso em: 15 ago. 2022.

FACEBOOK. Facebook. **Meta**, Menlo Park, 2023. Disponível em: <https://www.facebook.com/>. Acesso em: 18 jul. 2023.

FARINELLI, F.; ALMEIDA, M.B. Interoperabilidade semântica em sistemas de informação de saúde por meio de ontologias formais e informais: um estudo da norma openehr. In: CONFERÊNCIA INTERNACIONAL ACESSO ABERTO, PRESERVAÇÃO DIGITAL, INTEROPERABILIDADE, VISIBILIDADE E DADOS CIENTÍFICOS, 2014, Porto Alegre. **Anais...** Porto Alegre, 2014. Disponível em: https://mba.eci.ufmg.br/downloads/Biredial2014_144_web.pdf. Acesso em: 14 abr. 2023.

FIORIN, J. L. **Introdução à linguística**. 6. ed. São Paulo: Contexto, 2010. 232p.

GARRETT, J. J. **Ajax: a new approach to web applications**. San Francisco: Adaptive Path, 2005. Disponível em: https://designftw.mit.edu/lectures/apis/ajax_adaptive_path.pdf. Acesso em: 29 maio 2023.

GIL, A. C. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002.

GIL, A. C. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Atlas, 2008.

GODDARD, C. **Semantic analysis: a practical introduction**. 2. ed. New York: Oxford University Press, 2011. Disponível em: <https://catalog.loc.gov/vwebv/holdingsInfo?searchId=4568&recPointer=8&recCount=25&bibId=16987636>. Acesso em: 11 jul. 2023.

GRÁCIO, J. C. A. **Metadados para a descrição de recursos da Internet: o padrão Dublin Core, aplicações e a questão da interoperabilidade**. 2002. 127 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2002. Disponível em: <https://repositorio.unesp.br/handle/11449/93722>. Acesso em: 03 nov. 2021.

HILLMANN, D. Using Dublin Core: the elements. **Dublin Core Metadata Initiative (DCMI)**, [S. l.], 2005. Disponível em: <https://www.dublincore.org/specifications/dublin-core/usageguide/elements>. Acesso em: 03 jul. 2023.

KOBASHI, N. Y.; TÁLAMO, M. F. G. M. Informação: fenômeno e objeto de estudo da sociedade contemporânea. **Transinformação**, Campinas, v. 15, n. 3, p. 7-21, set./dez. 2003. Edição Especial. Disponível em: <https://www.scielo.br/j/tinf/a/SZ5RXSnjnDcjhZ9ykPpCHHw/?lang=pt>. Acesso em: 01 nov. 2021.

LAUFER, C. **Guia de Web Semântica**. São Paulo: [s. n.], 2015. Disponível em: <https://ceweb.br/publicacao/guia-de-web-semantica/>. Acesso em: 18 jul. 2023.

LE COADIC, Y. F. **A ciência da informação**. Tradução Maria Yêda Filgueiras Gomes. Brasília: Brique de Lemos, 1996. 115 p.

LOURENÇO, C. A. Metadados: O grande desafio na organização da web. **Informação & Sociedade: Estudos**, João Pessoa, v. 17, n. 1, 2007. Disponível em: <https://periodicos.ufpb.br/ojs/index.php/ies/article/view/466>. Acesso em: 13 out. 2021.

MAX, W. A evolução da home page do Yahoo! **Google Discovery**, Internet, [S. l.], 2008. Disponível em: <https://googlediscovery.com/2008/06/13/a-evolucao-da-home-page-do-yahoo/>. Acesso em: 20 abr. 2023.

MAGRANI, E. **A internet das coisas**. Rio de Janeiro: FGV Editora, 2018. 192 p.

MONTEIRO, F. S. **Web semântica e repositórios digitais educacionais na área de saúde: uma modelagem com foco no objetivo de aprendizagem para refinar resultados de busca**. 2013. 189 f. Tese (Doutorado em Ciência da Informação)—Universidade de Brasília, Brasília, 2013. Disponível em: <https://repositorio.unb.br/handle/10482/13461>. Acesso em: 20 jun. 2023.

MOREIRA, M. A. **Mapas conceituais e aprendizagem significativa**. São Paulo: Centauro, 2010. 80 p.

NUNES, F. R. E.; MACULAN, B. C. M. S.; B.; ALMEIDA, M. B. Os fundamentos da Web Semântica como ferramentas de auxílio para as demandas da Sociedade da Informação. **Em Questão**, Porto Alegre, v. 26, n. 3, p. 224–249, 2020. DOI: 10.19132/1808-5245263.224-249. Disponível em: <https://seer.ufrgs.br/index.php/EmQuestao/article/view/92336>. Acesso em: 10 jul. 2023.

O'REILLY, T. **O que é Web 2.0: padrões de design e modelos de negócios para a nova geração de software**. [S. l.], [s. n.], 2005. Disponível em: https://edisciplinas.usp.br/pluginfile.php/114173/mod_resource/content/1/o-que-e-we-b-20_Tim%20O%C2%B4Reilly.pdf. Acesso em: 13 mar. 2023.

PALETTA, F. C.; MUCHERONI, M. L. O desenvolvimento da web 3.0: linked data e dbpedia. **Prisma.com**, Porto, n. 25, p. 73-90, 2014. Disponível em: <https://ojs.letras.up.pt/index.php/prismacom/article/view/1869>. Acesso em: 26 jun. 2023.

RIBEIRO, D. S.; SILVA, M. B. Sistema de Gerenciamento de Conteúdo: proposta de um catálogo bibliográfico 2.0 no Wordpress. **Revista Digital de Biblioteconomia e Ciência da Informação (RDBCI)**, Campinas, v. 14, n. 1, p. 144–163, 2015.

Disponível em:

<https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/8640820>. Acesso em: 27 jun. 2022.

RILEY, J. Understanding Metadata: what is metadata, and what is it for? **National Information Standards Organization (NISO)**, Baltimore, 2017. Disponível em: <http://www.niso.org/publications/press/UnderstandingMetadata.pdf>. Acesso em: 02 jul. 2021.

ROBREDO, J. Ciência da informação e web semântica: linhas convergentes ou linhas paralelas?. In: ROBREDO, J.; BRÄSCHER, M. (Org.). **Passeios pelo bosque da informação: estudos sobre a representação e organização da informação e do conhecimento**. Brasília: IBICT, 2010. p. 2-40. Disponível em: <https://repositorio.unb.br/handle/10482/7952>. Acesso em: 25 jul. 2022.

SANTANA, P. H. A. et al. Servidor de enlaces: motivação e metodologia. **Ciência da Informação**, Brasília, v. 30, n. 3, p. 48–55, set. 2001. Disponível em: <https://www.scielo.br/j/ci/a/QCgL7Jj88bhRpYLDmmMFRgP/?format=pdf&lang=pt>. Acesso em: 19 abr. 2023.

SILVA, E. L.; MENEZES, E. M. **Metodologia da Pesquisa e Elaboração de Dissertação**. 4. ed. Florianópolis: UFSC, 2015. Disponível em: https://tccbiblio.paginas.ufsc.br/files/2010/09/024_Metodologia_de_pesquisa_e_elaboracao_de_teses_e_dissertacoes1.pdf. Acesso em: 09 jul. 2023.

SILVA, E. B. F.; SAMPAIO, D. A. O boom informacional: a tecnologia e a gênese da ciência da informação. **Bibliocanto**, Natal, v. 3, n. 2, p. 3-16, 2017. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/120322>. Acesso em: 04 nov. 2021.

SILVA, R. E. **As tecnologias da Web Semântica no domínio bibliográfico**. 2013. 134 f. Dissertação (Mestrado em Ciência da Informação) - Universidade Estadual Paulista Júlio de Mesquita Filho, Faculdade de Filosofia e Ciências de Marília, 2013. Disponível em: <http://hdl.handle.net/11449/93653>. Acesso em: 20 jun. 2023.

SILVA, V. A.; SILVA, M. B. Metadados para preservação digital de dados abertos: um estudo de identificação. **Biblios**, [S. l.], n. 78, p. 44–60, 2021. Disponível em: <http://biblios.pitt.edu/ojs/biblios/article/view/793>. Acesso em: 11 jul. 2023.

SOUZA, R. R.; ALMEIDA, M. B.; BARACHO, R. M. A. Ciência da Informação em transformação: Big Data, Nuvens, Redes Sociais e Web Semântica. **Ciência da Informação**, Brasília, v. 42, n. 2, p. 159-173, maio/ago., 2013. Disponível em: <http://revista.ibict.br/ciinf/article/view/1379>. Acesso em: 4 nov. 2021.

SOUZA, R. R.; ALVARENGA, L. A Web Semântica e suas contribuições para a ciência da informação. **Ciência da Informação**, Brasília, v. 33, n. 1, p. 132-141, jan./abr. 2004. Disponível em: <https://www.scielo.br/j/ci/a/sp3XpmZhXw384H5Fw9H89YL/abstract/?lang=pt>. Acesso em: 02 nov. 2021.

TAPSCOTT, D.; WILLIAMS, A. D. **Wikinomics: como a colaboração em massa pode mudar seus negócios**. Rio de Janeiro: Nova Fronteira, 2007.

TAVARES, R. Construindo mapas conceituais. **Ciências & Cognição (Cien. Cogn.)**, Rio de Janeiro, v. 12, p. 72-85, 2007. Disponível em: <http://www.cienciasecognicao.org/revista/index.php/cec/article/view/641>. Acesso em: 09 jul. 2023.

THIBODEAU, K. Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years The State of Digital Preservation: An International Perspective. **Anais...** Washington: CLIR and Library of Congress, 2002. p. 4-31. Disponível em: <https://www.clir.org/pubs/reports/pub107/>. Acesso em: 11 ago. 2022.

VICKERY, B. C. Knowledge representation: a brief review. **Journal of Documentation**, [S. l.], v. 24, n. 3, p.145-159, 1986. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/eb026790/full/html>. Acesso em: 15 ago. 2022.

WAZLAWICK, R. **História da Computação**. Rio de Janeiro: Elsevier, 2016. E-book. ISBN 9788595156180. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788595156180/>. Acesso em: 24 fev. 2023.

WORLD WIDE WEB CONSORTIUM. CSS Snapshot 2023. **W3C**, [S. l.], 2023. Disponível em: <https://www.w3.org/TR/css-2023/#css-glossary>. Acesso em: 09 jul. 2023.

WORLD WIDE WEB CONSORTIUM. Level 1 Document Object Model Specification. **W3C**, [S. l.], 1998. Disponível em: <https://www.w3.org/TR/WD-DOM/introduction.html>. Acesso em: 09 jul. 2023.

WORLD WIDE WEB CONSORTIUM. OWL Web Ontology Language Overview. **W3C**, [S. l.], 2004. Disponível em: <https://www.w3.org/TR/2004/REC-owl-features-20040210/#s1.2>. Acesso em: 11 jul. 2023.

WORLD WIDE WEB CONSORTIUM. RDF 1.1 Primer. **W3C**, [S. l.], 2014. Disponível em: <https://www.w3.org/TR/rdf11-primer/>. Acesso em: 24 jul. 2022.

WORLD WIDE WEB CONSORTIUM. XHTML Modularization 1.1 Second Edition. **W3C**, [S. I.], 2010. Disponível em: https://www.w3.org/TR/xhtml-modularization/introduction.html#s_intro_whatixhtml. Acesso em: 09 jul. 2023.

WORLD WIDE WEB CONSORTIUM. Extensible Markup Language (XML). **W3C**, [S. I.], 2016. Disponível em: <https://www.w3.org/XML/>. Acesso em: 27 jun. 2022.

WORLD WIDE WEB CONSORTIUM. XMLHttpRequest Level 1. **W3C**, [S. I.], 2016. Disponível em: <https://www.w3.org/TR/XMLHttpRequest/>. Acesso em: 10 jul. 2023.

WORLD WIDE WEB CONSORTIUM. XSL Transformations (XSLT) Version 2.0 Second Edition. **W3C**, [S. I.] 2021,. Disponível em: <https://www.w3.org/TR/xslt20/#dt-xslt-namespace>. Acesso em: 10 jul. 2023.