

Atualizado com a Norma ISO/IEC 21838 Top-level ontologies

Foreword: Mathias Brochhausen

MAURICIO BARCELLOS ALMEIDA

# ONTOLOGIA EM CIÊNCIA DA INFORMAÇÃO

TECNOLOGIA E APLICAÇÕES

**COLEÇÃO** Representação do Conhecimento em  
Ciência da Informação – **VOLUME 02**





Mauricio Barcellos Almeida

# ONTOLOGIA EM CIÊNCIA DA INFORMAÇÃO: Tecnologia e Aplicações

Coleção Representação do conhecimento  
em Ciência da Informação

Volume 2

Editora CRV  
Curitiba – Brasil  
2021

Copyright © da Editora CRV Ltda.  
**Editor-chefe:** Railson Moura  
**Diagramação e Capa:** Designers da Editora CRV  
**Revisão:** Heliana Ribeiro de Mello

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)  
CATALOGAÇÃO NA FONTE  
Bibliotecária Responsável: Luzenira Alves dos Santos CRB9/1506

---

Al444

Almeida, Mauricio Barcellos.

Ontologia em Ciência da Informação: Tecnologia e Aplicações / Mauricio Barcellos Almeida  
– Curitiba: CRV: 2021.  
372 p. (Coleção: representação do conhecimento em Ciência da informação, v. 2)

Bibliografia

ISBN Coleção Digital 978-65-251-1473-6  
ISBN Coleção Físico 978-65-251-1475-0  
ISBN Volume Digital 978-65-251-1474-3  
ISBN Volume Físico 978-65-251-1477-4  
DOI 10.24824/978652511477.4

1. Ciência da informação 2. Ontologia 3. Representação do conhecimento 4. Organização do conhecimento 5. Tecnologia da informação 6. Web Ontology Language I. Título II. Série.

CDU 005.9

CDD 111  
005.9

---

Índice para catálogo sistemático  
1. Ciência da in formação - 111

ESTA OBRA TAMBÉM SE ENCONTRA DISPONÍVEL  
EM FORMATO DIGITAL.  
CONHEÇA E BAIXE NOSSO APLICATIVO!



2021

Foi feito o depósito legal conf. Lei 10.994 de 14/12/2004

Proibida a reprodução parcial ou total desta obra sem autorização da Editora CRV

Todos os direitos desta edição reservados pela: Editora CRV

Tel.: (41) 3039-6418 - E-mail: [sac@editoracrv.com.br](mailto:sac@editoracrv.com.br)

Conheça os nossos lançamentos: [www.editoracrv.com.br](http://www.editoracrv.com.br)

## Conselho Editorial:

Aldira Guimarães Duarte Domínguez (UNB)  
Andréia da Silva Quintanilha Sousa (UNIR/UFRN)  
Anselmo Alencar Colares (UFOPA)  
Antônio Pereira Gaio Júnior (UFRRJ)  
Carlos Alberto Vilar Estêvão (UMINHO – PT)  
Carlos Federico Dominguez Avila (Unicuro)  
Carmen Tereza Velanga (UNIR)  
Celso Conti (UFSCar)  
Cesar Gerónimo Tello (Univer. Nacional  
Três de Febrero – Argentina)  
Eduardo Fernandes Barbosa (UFMG)  
Elione Maria Nogueira Diogenes (UFAL)  
Elizeu Clementino de Souza (UNEB)  
Élsio José Corá (UFFS)  
Fernando Antônio Gonçalves Alcoforado (IPB)  
Francisco Carlos Duarte (PUC-PR)  
Gloria Fariñas León (Universidade  
de La Havana – Cuba)  
Guillermo Arias Beatón (Universidade  
de La Havana – Cuba)  
Helmuth Krüger (UCP)  
Jailson Alves dos Santos (UFRJ)  
João Adalberto Campato Junior (UNESP)  
Josania Portela (UFPI)  
Leonel Severo Rocha (UNISINOS)  
Lídia de Oliveira Xavier (UNIEURO)  
Lourdes Helena da Silva (UFV)  
Marcelo Paixão (UFRJ e UTexas – US)  
Maria Cristina dos Santos Bezerra (UFSCar)  
Maria de Lourdes Pinto de Almeida (UNOESC)  
Maria Lília Imbiriba Sousa Colares (UFOPA)  
Paulo Romualdo Hernandes (UNIFAL-MG)  
Renato Francisco dos Santos Paula (UFG)  
Rodrigo Pratte-Santos (UFES)  
Sérgio Nunes de Jesus (IFRO)  
Simone Rodrigues Pinto (UNB)  
Solange Helena Ximenes-Rocha (UFOPA)  
Sydione Santos (UEPG)  
Tadeu Oliver Gonçalves (UFPA)  
Tania Suely Azevedo Brasileiro (UFOPA)

## Comitê Científico:

Adilson Xavier da Silva (UFRJ)  
Claudia Pereira do Carmo Murta (UFES)  
Daniel Omar Perez (UNICAMP)  
Élsio José Corá (UFFS)  
Francisco Verardi Bocca (PUCPR)  
Jorge Augusto da Silva Santos  
(Bento Silva Santos) (UFES)  
José Euclimar Xavier de Menezes (UCSAL)  
José Manuel de Barros Dias  
(Universidade de Évora – Portugal)  
Josiane Cristina Bocchi (UNESP)  
Marcelo Martins Barreira (UFES)  
Maria José Moreira (Universidade  
NOVA de Lisboa – Portugal)  
Nilo César Batista da Silva (UFCA)  
Patrice Vermeren (UNIVERSITÉ – PARIS 8)  
Richard Theisen Simanke (UFJF)  
Selvino Antonio Malfatti (UNIFRAN/RS)

Este livro passou por avaliação e aprovação às cegas de dois ou mais pareceristas *ad hoc*.



*Em favor da honestidade intelectual:*

*Seja este o seu lema:  
que a mentira venha ao mundo.  
Que ela até triunfe.  
Mas não por meio de mim.*

*Aleksandr Solzhenitsyn (1918-2008)  
Dramaturgo e historiador russo  
Prêmio Nobel por “Arquipélago Gulag”*



# SUMÁRIO

FOREWORD .....	13
<i>Mathias Brochhausen</i>	

INTRODUÇÃO .....	15
------------------	----

## PARTE I – FRONTEIRAS DA WEB

### CAPÍTULO 1

A WEB É FEITA DE MARCAS .....	33
1.1 Contexto: dos anos 1960 aos anos 2000 .....	35
1.1.1 Os primórdios da marcação .....	35
1.1.2 Marcação na web .....	40
1.1.3 De marcas a padrões de metadados .....	45
1.2 A marcação no século XXI.....	50
1.2.1 O modelo de triplas .....	50
1.2.2 Como prover significado?.....	53
1.2.3 OWL: a última fronteira .....	58
1.3 Considerações finais .....	62
1.3.1 Breve sumário .....	62
1.3.2 Revisão .....	63

### CAPÍTULO 2

QUEM É QUEM NA WEB SEMÂNTICA?.....	67
2.1 <i>Linked Data, Knowledge Graphs, Ontologias etc...</i> .....	67
2.1.1 Linked data e ontologias são a mesma coisa? .....	68
2.1.2 O mapa das representações .....	73
2.2 Porque ontologias?.....	87
2.2.1 Formalidade e precisão.....	87
2.2.2 Raciocínio e inteligência.....	90
2.3 Considerações finais .....	92
2.3.1 Breve sumário .....	92
2.3.2 Revisão .....	93

## PARTE II – NO MUNDO DAS MÁQUINAS

### CAPÍTULO 3

REPRESENTAR O MUNDO PARA COMPUTADORES .....	99
---	----

3.1 Para conceber linguagens e modelos formais .....	99
3.1.1 Terminologia .....	101
3.1.2 Idealização .....	102
3.1.3 Formulação .....	104
3.1.4 Construção .....	105
3.2 Por onde começar? .....	109
3.2.1 Tarefas da etapa terminológica .....	110
3.2.2 Tarefas da etapa de idealização.....	114
3.2.3 Tarefas da formulação e construção .....	116
3.3 Diagramas ajudam .....	118
3.3.1 Conjuntos e diagramas .....	119
3.3.2 Notação de visualização .....	125
3.4 Do português para a OWL.....	128
3.4.1 Nomenclatura .....	128
3.4.2 Um primeiro exemplo .....	130
3.5 Considerações finais .....	136
3.5.1 Breve sumário .....	136
3.5.2 Revisão .....	136

## CAPÍTULO 4

PARA CONVERSAR EM OWL.....	139
4.1 Aspectos preliminares .....	139
4.1.1 Dados e metadados .....	140
4.1.2 Restrições e operadores .....	146
4.1.3 Classes anônimas .....	149
4.2 As classes são os blocos.....	153
4.2.1 Apresentar o código .....	153
4.2.2 Formas de enriquecer dados .....	158
4.3 Restrições também são classes.....	168
4.3.1 Restrições com equivalências e subclasses são diferentes? ....	168
4.3.2 Restrições com filtros .....	173
4.3.3 Restrições com quantificadores mais cardinalidade .....	177
4.3.4 Restrições com valor individual.....	180
4.4 As propriedades são a cola .....	183
4.4.1 Hierarquias, domínio e faixa em propriedades.....	183
4.4.2 Características das propriedades e cadeias .....	185
4.4.3 Propriedades de tipos de dados.....	190
4.5 Considerações finais .....	191
4.5.1 Breve sumário .....	191
4.5.2 Revisão .....	191

## CAPÍTULO 5

FERRAMENTAS AJUDAM NA CONVERSA.....	197
5.1 Noções do editor de ontologias.....	198
5.2 Classes, propriedades e indivíduos.....	207
5.2.1 Classes e indivíduos .....	207
5.2.2 Hierarquias de classes .....	208
5.2.4 Propriedades de objetos .....	214
5.2.5 Hierarquia de propriedades.....	216
5.2.6 Igualdade de indivíduos .....	219
5.2.7 Tipos de dados .....	221
5.3 Mais sobre propriedades de objetos .....	225
5.3.1 Classes complexas .....	225
5.3.2 Restrições de propriedades .....	230
5.3.3 Restrições de propriedades de cardinalidade.....	233
5.3.4 Enumeração de indivíduos.....	236
5.4 Ainda sobre propriedades de objeto e tipos de dados .....	238
5.4.1 Características de propriedades .....	238
5.4.2 Cadeias de propriedades e chaves.....	247
5.5 Informações sobre ontologia e anotações.....	252
5.5.1 Anotações para axiomas e entidades .....	252
5.5.2 Gestão da ontologia .....	255
5.5.3 Declarações de entidades.....	257
5.6 Considerações finais .....	258
5.6.1 Breve sumário .....	258
5.6.2 Revisão .....	258

## PARTE III – DE VOLTA AO MUNDO REAL

## CAPÍTULO 6

INFERÊNCIAS NA PRÁTICA.....	267
6.1 Um poucas palavras sobre o raciocínio humano.....	267
6.2 Dois tipos de raciocínio computacional .....	271
6.3 Inferências em ontologias OWL .....	275
6.3.1 Por que escolher ontologias para inferências? .....	275
6.3.2 Exemplos de inferências .....	277
6.4 Padrões de inferências válidas.....	288
6.5 Considerações finais .....	294
6.5.1 Breve sumário .....	294
6.5.2 Revisão .....	294

CAPÍTULO 7	
PARA OBTER BOAS ONTOLOGIAS .....	299
7.1 Evitar problemas conhecidos .....	299
7.1.1 Problemas de entendimento .....	301
7.1.2 Problemas lógicos .....	302
7.1.3 Problemas de modelagem .....	304
7.1.4 Problemas com a linguagem de representação .....	308
7.2 Adotar teoria ontológica bem fundamentada .....	310
7.2.1 Definir e validar com entidades BFO .....	311
7.2.2 Definir e validar através de relacionamentos BFO .....	317
7.3 Atuar em setores relevantes da sociedade .....	319
7.3.1 Medicina .....	321
7.3.2 Arquitetura da informação .....	328
7.3.3 Internet das Coisas .....	341
7.4 Considerações finais .....	349
7.4.1 Breve Sumário .....	349
7.4.2 Revisão .....	349
CAPÍTULO 8	
CONSIDERAÇÕES FINAIS .....	353
8.1 Limitações e justificativas .....	354
8.2 <i>Insights</i> úteis .....	357
8.3 Ontologia e sociedade .....	360
8.4 Recomendações sobre o percurso .....	361
ÍNDICE REMISSIVO .....	363

# FOREWORD

*Dr. Mathias Brochhausen*

---

An anecdote about a world-famous conductor claims that he once stepped out of an airport, jumped into a cab, and was asked by the cabbie: “Where to, Maestro”? He replied: “Doesn’t matter! They need me everywhere.”

As an expert in using Semantic Web Technologies, such as ontologies, to improve, integrate, and mine data in medicine, one cannot help but get the same feeling of being needed everywhere. Most of the data sources we encounter in daily practice would benefit from enhancement allowing for machine-interpretability and reasoning. Many projects aiming to collect and analyze data from heterogeneous sources are in dire need of semantic integration. Reasoning would allow filling existing data quality gaps. Furthermore, reasoning would allow inferring new knowledge from existing data without much human intervention. The latter could be especially fruitful for the medical field, where it could avoid duplicate testing.

In volume 1 of this important series of textbooks about the role of ontologies in information science, Mauricio Almeida introduced ontologies and best-practice approaches to create them. The book addresses ontology development for multiple domains. Volume 2 sets out to introduce methodologies and approaches to put these ontologies to work in information systems. It dives deeper into the world of Semantic Web Technologies, to introduce and explain a wide range of languages and tools for semantically enriched data management. Tackling the problem of how to put ontologies to work is of crucial importance in order to make use of the full range of benefits that employing ontologies can provide. Knowledge of these methodologies and tools allows data scientists to implement semantic data integration, reasoning over data, and data mining.

Over the last decades, there have been growing corpora of well-built and well-maintained ontologies, that are publicly available. For example, in biomedicine big projects like the OBO Foundry [Smith B, Ashburner M, Rosse C, et al. The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. *Nature Biotechnology* 25, 1251 - 1255 (2007)] and the NCBO BioPortal [Musen MA, Noy NF, Shah NH, et al. The National Center for Biomedical Ontology. *J Am Med Inform Assoc.* 2012 Mar-Apr;19(2):190-5] exist and provide libraries of ontologies inviting collaboration or usage. However, uptake of these artifacts is hindered by two factors: a) the small number of ontology curators that maintain ontologies

over their often largely unfunded lifespan; and b) the small number of experts in applying ontologies to real-world data problems. Especially, regarding the latter this volume provides a necessary bridge from state-of-the-art ontologies to semantically aware data management. This becomes even more important as the Artificial Intelligence-capabilities of ontologies appear to be greatly underutilized on domain-specific ontology use, where ontologies are often used merely as controlled vocabularies.

Barry Smith, in his foreword to the first volume, points out the wealth of contributions Mauricio Almeida has made to the field of Applied Ontologies. Mauricio has been an extremely active member of both the ontology development and the applied ontologies communities. Thus, his textbooks depict his intimate knowledge of both aspects of ontological engineering.

Having had the privilege to deliver a keynote at the Seminar on Ontology Research in Brazil (ONTOBRAS) 2019, I know that this book will contribute to a particularly rich and lively research community. I am looking forward to seeing the fruits this excellent introduction to the application of ontologies in information and data science is going to bring forth!

Little Rock, May 2021

# INTRODUÇÃO



A teoria da *Ontologia Aplicada* foi abordada no primeiro volume dessa coleção: o Volume 1 – Teoria e Método – forneceu diretrizes e fundamentos teóricos sobre o tema, ainda que não fosse possível esgotar assunto tão amplo. A distinção feita naquele volume remetia à dualidade da Ontologia Aplicada, a qual é essencial que seja entendida para que menções no presente volume façam sentido para leitor.

Ontologia, originalmente, um termo que nomeia uma disciplina da Filosofia, se ocupa da existência e das formas mais gerais daquilo que existe. Essa é a **ontologia como disciplina**.

Ontologia, no escopo das Ciências Aplicadas, se refere a um artefato de *software* que provê uma estrutura de mundo a máquinas, permitindo-lhes um tipo básico de raciocínio. Essa é a **ontologia como artefato**.

A **Ontologia Aplicada** é a combinação das teorias da disciplina e das técnicas do artefato.

Inicia-se aqui mais uma parte da jornada, agora para representar as coisas do mundo através da ontologia como artefato. A representação que se busca não é direcionada diretamente a pessoas, mas sim a máquinas. Assim

“alimentados”, os computadores podem se encarregar de tarefas intelectualmente mais simples, ainda que trabalhosas, liberando as pessoas para tarefas mais nobres.

Antes de iniciar a nova jornada, listam-se os volumes da Coleção Representação do Conhecimento em Ciência da Informação, da qual o presente é o Volume 2. Cada volume pode ser escolhido de acordo com os objetivos de momento do leitor:

- Volume 1 – Teoria e Método (2020)
- Volume 2 – Tecnologia e Aplicações (2021)
- Volume 3 – Estudos Avançados (2022)

## Contexto

Certos aplicativos são tão bem desenvolvidos que parecem “inteligentes”. A maioria deles contém milhares de linhas de código e vem sofrendo manutenção já há décadas. Esse enorme volume de código procedural é tratável pelos avanços do *hardware*, fazendo com que a máquina possa guardar quase todas as situações possíveis de um cenário, como em um jogo de xadrez. Isso confere a tais sistemas e a máquinas algo que lembra inteligência. Por outro lado, ao buscar inteligência, sistemas desenvolvidos no paradigma declarativo tornam explícito o conhecimento humano sobre o mundo. Isso quer dizer que se uma maçã é vermelha, a sentença “maçã é vermelha” é declarada no sistema de uma forma que tem correspondência com a linguagem humana.

Sistemas fazem uso de representações no contexto digital há mais de 50 anos. A abordagem ontológica, assunto do presente livro, se fundamenta no paradigma declarativo, ao invés do paradigma procedural utilizado em sistemas de uso rotineiro. Isso significa que o leitor, ao construir ontologias, não precisa programar – Java, C, C++, etc. – e criar *softwares* que executam tarefas.

Para os **proceduralistas**, sistemas inteligentes devem receber “conhecimento-como” – conhecimento sobre como realizar procedimentos – tanto quanto possível, através de códigos cada vez mais sofisticados.

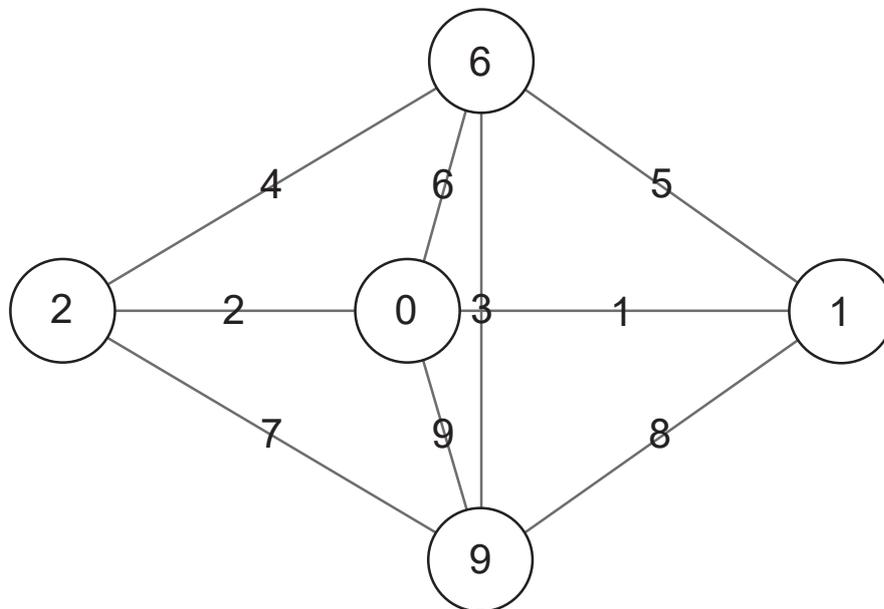
Para os **declarativistas**, sistemas inteligentes devem receber “conhecimento-que” – conhecimento sob a forma de representações do mundo – é aquele que torna inteligentes os sistemas.

Um desenvolvimento no âmbito da abordagem declarativa que desencadeou grandes avanços foi a *Lógica Descritiva* – também conhecida como *Linguagem Terminológica* ou *Linguagem Conceitual* – criada como uma

resposta a problemas dos formalismos de representação dos anos 1970. Tais formalismos, especialmente as redes semânticas e os sistemas baseados em *frames*, não possuíam semântica formal bem definida, o que é um requisito para o raciocínio automático. O raciocínio automático também é chamado de *inferência*.

A partir de meados dos anos 1980, a família das Lógicas Descritivas passou a fornecer semântica confiável e técnicas de inferência para sistemas baseados em conhecimento. Estes são os sistemas que armazenam dados nos projetos desenvolvidos no âmbito da Representação do Conhecimento. As Lógicas Descritivas permitem representar o mundo através de uma abstração usando *grafos diretos rotulados* (GDR).

**Figura I.1 – Um exemplo de grafo direto rotulado**



Fonte: *Wolfram Demonstrations Project*<sup>1</sup>.

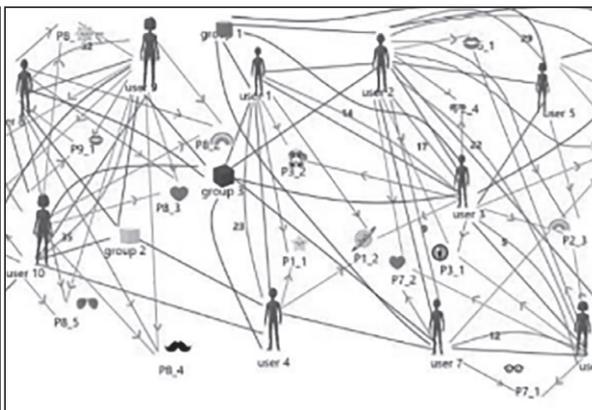
Nessa abstração, os nós (ou vértices) do GDR representam objetos do mundo e os rótulos dos nós representam características dos objetos. As arestas do GDR representam as relações entre pares de objetos e os rótulos de arestas representam os tipos de relações entre os objetos. Partes do mundo que podem ser representadas abstratamente em termos de um GDR são bons candidatos também para representação através de Lógicas Descritivas.

**Figura I.2 – Coisas do mundo real representadas por grafos**

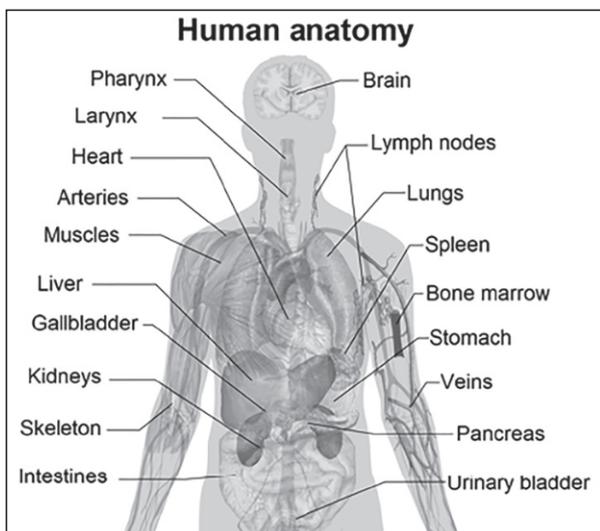
**a- linhas de metrô**



**b- conexões do Facebook**



**c- aspectos da anatomia humana**



**d- aspectos da classificação de documentos**



Fonte: Calvanese (2012, p. 21).

Paralelamente, no final dos anos 1980, o advento da Web trouxe muitas mudanças no ambiente de trabalho das organizações e na interação entre sistemas e pessoas. Os sistemas, em grande medida, migraram para o ambiente Web buscando as facilidades no uso e a interface padronizada do navegador. As linguagens de marcação, desenvolvidas desde os anos 1960 – sendo a mais popular o HTML<sup>2</sup> – evoluíram no escopo da Web Semântica para linguagens mais expressivas, passando pelo XML<sup>3</sup>, até a *Web Ontology Language* (OWL).

2

em: 15 fev. 2021.

3

mai. 2021.

## Por que ontologia é importante?

Pode-se executar hoje uma variedade de tarefas pela Web, desde algo simples como pedir um almoço ou escolher um filme, até procedimentos mais complexos como fazer uma matrícula na universidade ou pedir a expedição da carteira de motorista. Procedimentos e fluxos de trabalho imensos têm sido transferidos para a rede, sem contar a presença quase pervasiva das redes sociais. Ao usar todas essas facilidades, em dispositivos diversos, as pessoas estão lidando com softwares que fazem uso de ontologias. Isso ocorre mesmo em softwares que as pessoas usam no dia a dia e tomam como certo no seu planejamento.

Para aqueles que desconhecem os objetivos da Web Semântica, o próprio nome da linguagem OWL é um exemplo. Ao pesquisar por “owl” em um mecanismo de busca, a maioria das páginas recuperadas será sobre “coruja”. Em uma Web Semântica ideal seria possível especificar claramente o que é “*Web Ontology Language*” e o usuário não precisaria se preocupar com siglas ou sinônimos. Da mesma forma, o usuário não se incomodaria com homônimos enchendo os resultados da pesquisa de dados irrelevantes, como, por exemplo, dados sobre aves de rapina, tipos de corujas, seus hábitos noturnos, etc.

A OWL carregou qualidades das Lógicas Descritivas para o ambiente da Web: a abstração para representação via grafos e a semântica formal bem definida. A OWL, hoje na sua versão 2, é a principal linguagem em uso para construir ontologias como artefato e objetiva fornecer descrições da realidade para fins de inferência. A OWL possui três características que a distinguem das demais:

- i. É declarativa, pois contém conhecimento explicitamente definido<sup>4</sup>
- ii. É rigorosa e livre de ambiguidades, pois é baseada em uma família de lógicas
- iii. É capaz de reduzir decisões *ad hoc* por sua adesão à ontologia como disciplina
- iv. A família de lógicas a que pertence possui propriedades computacionalmente úteis

---

4 W3C (2012): “OWL 2 is not a programming language: OWL 2 is declarative i.e. it describes a state of affairs in a logical way”

## O que esse livro oferece?

Ao aprender português ou qualquer outro idioma, uma pessoa não estuda todas as possibilidades da língua para depois começar a usá-la. Isso seria impossível, dada a extensão e a complexidade da língua natural. A pessoa começa com um “vocabulário básico”, no passado, uma “cartilha”, que vai evoluindo à medida do uso e da interação.

Não é tão diferente quando se precisa aprender uma linguagem “artificial”, também complexa e extensa, e dotada de sintaxe e semântica específicas. Mas isso não parece ser levado em conta pela maioria dos livros sobre Web Semântica, os quais listam padrões técnicos repletos de símbolos desconhecidos (Figura I.3). Existem muitos recursos a considerar quando se fala em Web Semântica e a ontologia como artefato é apenas um desses recursos, ainda que o mais importante.

Um projeto de Web Semântica requer linguagens de programação, como *Python*, *Java*, etc.; tecnologias da Web, como protocolos diversos, *web-services*, linguagens de consulta, de apresentação, etc.; ferramentas de propósito geral, como *integrated development interfaces* (IDEs), *application programming interfaces* (APIs), dentre outros. Ainda que não se possa escapar de alguns desse requisitos, eles não são o objetivo aqui.

O presente volume busca uma abordagem diferenciada. Por isso não abrange todas as formas como os sistemas podem manipular o OWL, nem mesmo todos os seus elementos. Entretanto, explica e exemplifica um conjunto de construtos – para representar entidades, relacionamentos e valores – que engloba quase tudo o que é preciso para se criar ontologias como artefato usando OWL. Inferência é o que se espera de uma ontologia OWL e, por isso, o volume traz exemplos sumarizados em uma lista de inferências padrão. As inferências são “padrão” no sentido de que são esperadas de qualquer ontologia OWL.

**Figura I.3 – Fragmentos de sumários de livros sobre Web Semântica**

**a- Fragmento livro em Web Semântica    b- Fragmento livro em Web Semântica**

<p><b>2 Describing Web Resources: RDF</b></p> <p>2.1 Introduction . . . . .</p> <p>2.2 RDF: Data Model . . . . .</p> <p>2.3 RDF Syntaxes . . . . .</p> <p>2.4 RDFS: Adding Semantics . . . . .</p> <p>2.5 RDF Schema: The Language . . . . .</p> <p>2.6 RDF and RDF Schema in RDF Schema . . . . .</p> <p>2.7 An Axiomatic Semantics for RDF and RDF Schema . . . . .</p> <p>2.8 A Direct Inference System for RDF and RDFS . . . . .</p> <p>2.9 Summary . . . . .</p>	<p><b>CHAPTER 3</b> RDF—The Basis of the Semantic Web</p> <p>Distributing Data Across the Web</p> <p>Merging Data from Multiple Sources</p> <p>Namespaces, URIs, and Identity</p> <p>Expressing URIs in Print</p> <p>Standard Namespaces</p> <p>Identifiers in the RDF Namespace</p> <p>Challenge: RDF and Tabular Data</p> <p>Higher-Order Relationships</p> <p>Alternatives for Serialization</p> <p>N-Triples</p>
--	--

**c- Fragmento livro em lógica descritiva    d- Fragmento livro em inferências e lógica**

<p><b>3 Description Logics</b></p> <p>3.1 DL primer . . . . .</p> <p>3.1.1 Basic building blocks of DL ontologies . . . . .</p> <p>3.1.2 Constructors for concepts and roles . . . . .</p> <p>3.1.3 Description Logic semantics . . . . .</p> <p>3.2 Important DLs . . . . .</p> <p>3.2.1 A basic DL to start with: <i>ACC</i> . . . . .</p> <p>3.2.2 The DL <i>SROIQ</i> . . . . .</p> <p>3.2.3 Important fragments of <i>SROIQ</i> . . . . .</p> <p>3.3 Reasoning services . . . . .</p> <p>3.3.1 Standard reasoning services . . . . .</p> <p>3.3.2 Techniques: a tableau for <i>ACC</i> . . . . .</p>	<p><b>2 Mathematical Logic and Inference</b></p> <p>2.1 Representation and Logic . . . . .</p> <p>2.2 Propositional Logic . . . . .</p> <p>2.3 First-Order Logic . . . . .</p> <p>2.4 Sets . . . . .</p> <p>2.5 Description Logic . . . . .</p> <p>2.5.1 Description Language <i>AL</i> . . . . .</p> <p>2.5.2 Description Language <i>ACC</i> . . . . .</p> <p>2.5.3 Further Description Logic Constructors . . . . .</p> <p>2.6 Exercises and Further Reading . . . . .</p>
---	---

Fonte: busca simples no Google com *string* “livros web semântica”.

Procura-se trazer um texto acessível para não iniciados no tema, sem desvalorizar a tradição reflexiva da Ciência da Informação em troca de tecnicismos e intelectualismos. Isso não quer dizer que os assuntos dos livros técnicos não tenham valor, mas materiais técnicos sobre o assunto são facilmente obtidos na Web. Ao contrário, o presente volume aborda aspectos ausentes em outras obras e pouco disponíveis de forma acessível. Exemplos desses aspectos abordados com alguma exclusividade são:

- i. A integração com teoria bem fundamentada, testada e adotada em todo o mundo – disponível no Volume 1 – a qual ao final, é combinada com a tecnologia.
- ii. A explicação, decifrando o jargão matemático computacional, sobre para que servem, como se constroem e se utilizam linguagens formais.
- iii. Ênfase em explicações acessíveis, considerando-se que o maior impedimento ao lidar com ontologias é a falta de habilidades que modelos formais requerem
- iv. A inclusão de recursos didáticos intuitivos, como a Teoria dos Conjuntos, e que reduzem em muito a curva de aprendizado em OWL

O presente volume, assim como toda coleção, é também uma obra de Ciência da Informação e para a Ciência da Informação. É voltada para estudantes e pesquisadores da Ciência da Informação, cientistas da computação e analistas de sistemas, analistas de informação, linguistas, bibliotecários, bem como profissionais afins, como cientistas de dados, programadores, administradores de bancos de dados, dentre outros. É também de interesse para executivos e gerentes de tecnologia da informação que pretendem fomentar um salto qualitativo nas práticas de informação em ambientes profissionais.

### **Qual a motivação do livro?**

Uma das razões para criar o presente volume é compor uma coleção que engloba os principais assuntos da pesquisa em Ontologia Aplicada. Não existe livro brasileiro similar para ensino e pesquisa para a Ciência da Informação. Mantém-se a convicção, externada desde o Volume 1, de que a Ontologia Aplicada se constitui no tema a partir do qual a Ciência da Informação pode reencontrar suas origens no contexto digital do século XXI.

Um desafio que merece menção ao se lidar com ontologias é a crença equivocada de que criar um arquivo de extensão OWL é tudo o que é necessário. Essa é outra motivação para o livro, ou seja, explicar o porquê de essa crença estar longe da verdade. Objetiva-se deixar claro que, sem os princípios da Organização do Conhecimento e daqueles oferecidos pela Ontologia Aplicada, um arquivo OWL é apenas um arquivo texto, similar a um “.txt”. Acredita-se que os princípios teóricos para a criação de um vocabulário de representação, fundamentados em campos diversos – como Ciência da Informação, Linguística e Filosofia – são parte importante da criação de artefatos ontológicos.

Uma das consequências de se negligenciar tais princípios é colher inferências que até podem ser válidas, mas não são verdadeiras, não são condizentes com a realidade. De forma mais genérica, a negligência aos princípios teóricos pode levar ontologistas a violar regras do raciocínio dedutivo. Outra consequência indesejada é continuar carregando problemas dos sistemas de informação dos últimos 30 ou 40 anos, resultantes também, em grande medida, de negligência na modelagem.

Por outro lado, boas ontologias permitem obter um vocabulário que descreve parte do mundo, um domínio do conhecimento, mais as premissas sobre como o vocabulário deve ser interpretado. Uma vez construída a ontologia como artefato, baseada na ontologia como disciplina, é possível perguntar: que tipo de conclusão, de conhecimento novo e verdadeiro, pode se obter a partir das premissas aqui consideradas? Por “conhecimento novo” entende-se aquele que não estava disponível quando da formulação das premissas do argumento.

O objetivo geral do presente volume é ensinar a “falar algo em OWL”, de forma que se possa conversar com a máquina. O termo “conversar” é usado metaforicamente para explicar que se pode, muito mais facilmente que no passado recente, declarar algo e receber de volta conhecimento novo vindo de uma máquina. Sim, ontologias como artefatos fazem parte do conjunto de tecnologias da Inteligência Artificial.

Pode não parecer adequado, à primeira vista, escolher uma tecnologia específica, em vista da rápida evolução e obsolescência vivenciadas hoje. Entretanto, a obra oferece mais do que a tecnologia do momento: explica princípios que podem ser adotados com qualquer axiomatização que seja legível por máquina e que tenha semântica definida. Isso vale para OWL hoje, assim como para qualquer lógica descritiva indicada pelo W3C<sup>5</sup> como sucessora do OWL 2.

### Como o livro está organizado?

Cabe uma menção a cientistas e pesquisadores de áreas científicas das quais extratos foram tomados de empréstimo. Merecidos agradecimentos e desculpas antecipadas no caso de falta de rigor em interpretações ou argumentos. Agradecimentos ao Dr. *Barry Smith* da *State University of New York at Buffalo* (US), ao Conselho Nacional de Pesquisa e Desenvolvimento (CNPQ)<sup>6</sup>, aos integrantes do grupo de pesquisa Representação do Conhecimento, Ontologias e Linguagem (RECOL)<sup>7</sup> e a Dra. Heliana Ribeiro de Mello, Professora Titular de Linguística da Universidade Federal de Minas Gerais.

Para atingir seus objetivos, o restante do presente volume se organiza em capítulos distribuídos em três grandes partes, as quais são descritas a seguir:

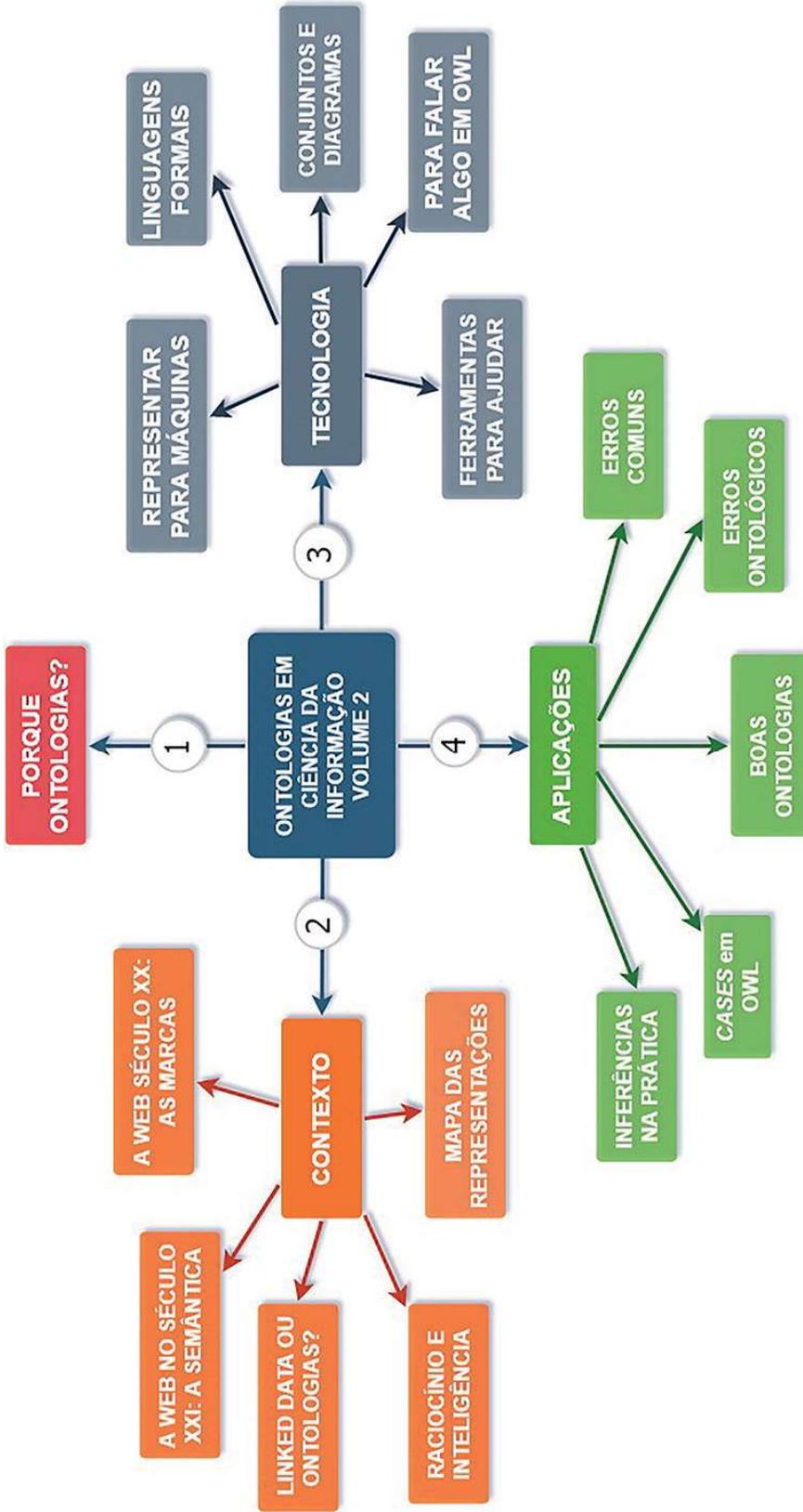
- PARTE I – Fronteiras da Web (Capítulos 1 e 2)
- PARTE II – No mundo das máquinas (Capítulos 3, 4 e 5)
- PARTE III – De volta ao mundo real (Capítulos 6, 7 e 8)

---

5 *The World Wide Web Consortium*

6 [chagas.cnpq.br/](http://chagas.cnpq.br/). Acesso em: 29 mar. 2021.

7



O **Capítulo 1** – *A Web é feita de marcas* – retorna às origens da Web para caracterizá-la como um tipo de artefato que atende à necessidade humana de fazer marcas em textos para estruturá-los. Aborda-se o contexto de criação e de evolução das linguagens de marcas até alcançar a *Web Ontology Language* (OWL). Explica-se evolução dessas linguagens em termos de representação e, em particular, da OWL, em prover semântica através de um modelo baseado em triplas no ambiente da Web Semântica.

O **Capítulo 2** – *Quem é quem na Web Semântica?* – apresenta tecnologias usadas para representação na Web Semântica – por exemplo, *Linked Data*, *Knowledge Graphs*, ontologias OWL – distinguindo suas capacidades e limitações. Um Mapa de Representações é apresentado para explicar como extrair o melhor de cada tecnologia e a importância de se considerar os tipos de representação de conhecimento ao planejar a arquitetura de sistemas. Discute as vantagens de se adotarem ontologias OWL, abordando suas habilidades de inferência automática e comparando-as com outras tecnologias.

O **Capítulo 3** – *Representar o mundo para computadores* – enfatiza a representação para máquinas introduzindo características das linguagens formais de representação e como criá-las. Indica, ao não iniciado, os primeiros passos para construir um modelo formal a partir de uma teoria formal, algo bastante similar ao que precisa ser feito em uma ontologia OWL. Traz notações diagramáticas e retoma aspectos da teoria dos conjuntos que reduzem a curva de aprendizado em OWL. Finalmente, apresenta os primeiros exemplos do processo de traduzir de português para OWL.

O **Capítulo 4** – *Para conversar em OWL* – segue a abordagem ao principal objetivo desse volume: “dizer algo em OWL”. Depois de recordar elementos dos primeiros exemplos, descrevem-se os “blocos” e a “cola” da Web Semântica – respectivamente, as classes e as propriedades OWL – e como são usados para enriquecer dados. Introduzem-se os principais construtos da linguagem OWL em diferentes sintaxes, enfatizando-se as alternativas para prover significado em exemplos simples. Os construtos considerados principais são aqueles usados na maioria dos casos mesmo por ontologistas experientes.

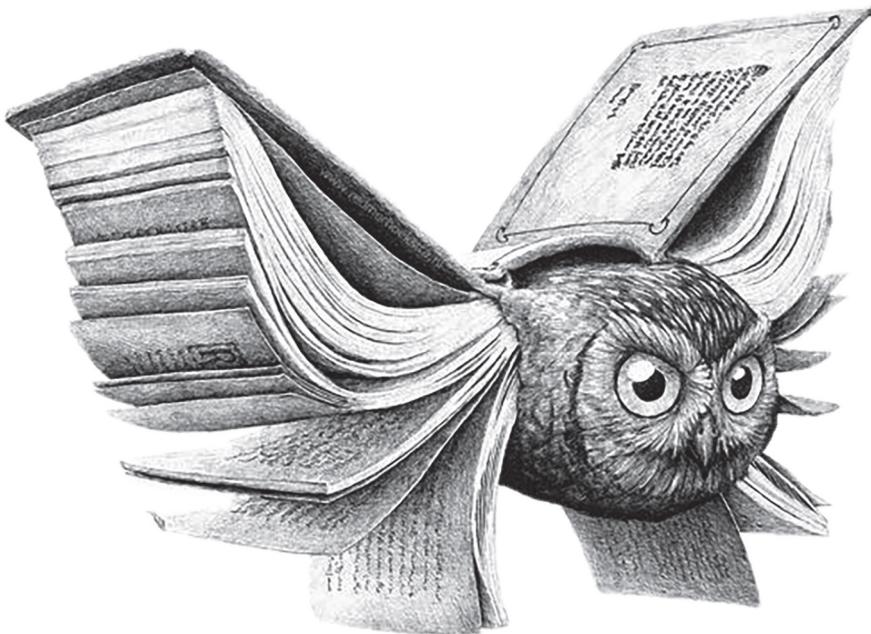
O **Capítulo 5** – *Ferramentas ajudam na conversa* – faz uma revisão dos construtos OWL já apresentados, introduzindo novos, gradativamente, num processo que se estende até o final do volume. Além de construtos, a principal novidade é a inclusão de um editor de ontologias como suporte à construção de ontologias OWL, com o objetivo de se evitar a complexidade inerente a linguagens computacionais. Noções de uso do editor são apresentadas junto a exemplos e a telas da ferramenta.

O **Capítulo 6** – *Inferências na prática* – aborda de maneira direta o principal motivo para se investir em ontologias OWL: o raciocínio ou, usando um termo mais técnico, as inferências automáticas. Um pequeno extrato teórico

distingue dois tipos de inferências populares oferecidos pela tecnologia da informação, seguido de exemplos práticos de como funcionam as inferências em OWL e como podem ser úteis. Uma lista de inferências OWL válidas mais comuns é então apresentada em formato padronizado.

O **Capítulo 7** – *Para obter boas ontologias* – advoga que construir ontologias e construir “boas ontologias” são processos distintos. Enquanto o primeiro caso exige apenas gerar arquivos OWL, o segundo caso requer considerar princípios ontológicos já descritos (veja Volume 1), sob pena de manter os mesmos problemas em outra roupagem. Apresenta-se uma lista de erros conhecidos no processo de desenvolvimento, classificados por tipos, mais uma descrição sobre o principal erro: a negligência aos princípios ontológicos.

O **Capítulo 8** – *Considerações finais* – traz um sumário do que foi apresentado no presente volume e destaca limitações e aspectos não incluídos, os quais são parte do volume seguinte (Volume 3). Após considerações sobre a importância de se adotar a abordagem da Ontologia Aplicada, encerra-se o presente volume de caráter tecnológico com digressões sobre a ontologia e a sociedade.



## Simbologia adotada

A nomenclatura e a terminologia utilizadas no volume anterior (Volume 1) se mantêm, em grande medida, mesmo que a ênfase aqui recaia sobre a tecnologia para se construir ontologias como artefatos. Isso requer um alinhamento com outra nomenclatura, que envolve aspectos matemáticos, lógicos e de tecnologias da Web. Sempre que necessário, definições e elucidacões do volume anterior são transcritas com o objetivo de prover ao volume o seu caráter individual, permitindo que possa ser lido separadamente dos outros.

### a-Texto

- No geral, o texto usa *itálico* ou **negrito** para diferenciar, na primeira ocorrência, palavras que podem ter significado diferente do uso cotidiano.
- Aspas são um dispositivo para indicar quando a linguagem é usada para teorizar ou para falar sobre propriedades da própria linguagem.
- Onde necessário, aspas duplas são aplicadas sobre aspas simples, aspas triplas são aplicadas sobre aspas duplas, e assim por diante.
- Em casos particulares e diferenciados, o próprio capítulo provê a convenção.

### b- Quadros de informação adicional

Quadros como esse, com linhas simples, sombreado claro, fonte Arial Narrow 10 podem conter:

- i. Exemplos ou sentenças às quais se quer dar destaque.
- ii. Material adicional para se entender um assunto novo introduzido.

### c- Fórmulas

As fórmulas são grafadas em fonte Agency FB 10 ou 12 como em:

- se  $X$  é semelhante a  $Y$ , então  $X$  representa  $Y$
- $F = x$  é feliz
- $\exists x [ Sx \wedge Ax ]$

## d- Lógica

Sentenças em lógica são evitadas, pois não fazem parte da formação usual em Ciência da Informação. A preferência é a explicação em linguagem natural, mas onde isso não for possível adota-se a notação básica:

Operador	Nome técnico	Significado	Exemplo
$\neg$	Negação	Não	$\neg P$
$\wedge$	Conjunção	E	$P \wedge Q$
$\vee$	Disjunção	Ou	$P \vee Q$
$\rightarrow$	Condicional	Se, então	$P \rightarrow Q$
$\leftrightarrow$	Bicondicional	Se e somente se	$P \leftrightarrow Q$
$\forall$	Quantificador universal	Para todo	$\forall x$ se lê “Para todo x...”
$\exists$	Quantificador existencial	Alguns	$\exists x$ se lê “Existe um x tal que...”

Além da notação básica da Lógica, a notação da *Web Ontology Language* segue sintaxes específicas que são explicadas à medida do necessário. Para uma referência inicial, sem entrar em aspectos técnicos, as formas de apresentar são:

- Elementos de linguagens Web no texto são mantidos entre marcas originais (< e >) e com a fonte Agency FB 10 ou 12, como em <owl:someValuesFrom> ou <rdfs:subClassOf>
- Equivalentes da sintaxe original em sintaxes mais avançadas e amigáveis são usadas em exemplos, também com a fonte Agency FB 10 ou 12, como em `some` e `class`
- O *namespace* fictício para todos os exemplos é a IRI<sup>9</sup> “exe:”, o qual corresponde ao endereço <https://www.ncor-brasil.org/livro-vol2/termos#>
- Termos compostos para classes OWL adotam a notação *CamelCase* com letras maiúsculas primeiro, como em “PacienteFeminino”
- Termos compostos para propriedades OWL adotam *CamelCase* com letras minúsculas primeiro, como em “temComponente”
- Um indivíduo OWL recebe destaque pelo símbolo de diamante, como em “◆Eva”

9 *Internationalized Resource Identifier* (IRI) internacionaliza o *Uniform Resource Locator* (URL) ao permitir o uso do *Universal Coded Character Set* (Unicode).

## REFERÊNCIAS

ALMEIDA, M. B. *Ontologia em Ciência da Informação: Teoria e Método*. Curitiba: CRV, 2020, 377 p. Coleção Representação do Conhecimento em Ciência da Informação: Volume I, ISBN 978-65-5578-679-8.

ALMEIDA, M. B. *Uma introdução ao XML, sua utilização na Internet e alguns conceitos complementares*. Ciência da Informação, volume 31, 2003.

CALVANESE, D. *Lectures notes. Knowledge Representation and Ontologies: Part 2, Description Logics*. Faculty of Computer Science Master of Science in Computer Science Free University of Bolzano, 2012.

GENESERETH, M. R.; NILSSON, J. *Logical foundations of artificial intelligence*. Morgan Kaufmann, San Francisco, 1988.

MCLAY, R. *Managing the rise of Artificial Intelligence*, 2018. Disponível em: <<https://tech.humanrights.gov.au/sites/default/files/inline-files/100%20-%20Ron%20McLay.pdf>>. Acesso em: 10 nov. 2019.

SMITH, B.; ALMEIDA, M. B.; BONA, J. *et al. Basic Formal Ontology 2.0: Specification and user's guide*. National Center for Ontological Research: Buffalo, USA.