

Universidade de Brasília (UnB)

Faculdade de Ciência da Informação (FCI)

Programa de Pós-Graduação em Ciência da Informação (PPGCInf)

Web semântica e repositórios digitais educacionais na área de saúde: uma modelagem com foco no objetivo de aprendizagem para refinar resultados de busca

Fernanda de Sousa Monteiro

Brasília – DF

Março de 2013

FERNANDA DE SOUSA MONTEIRO

Web semântica e repositórios digitais educacionais na área de saúde: uma modelagem com foco no objetivo de aprendizagem para refinar resultados de busca

Tese apresentada ao Programa de Pós-graduação em Ciência da Informação da Universidade de Brasília como requisito parcial para obtenção do título de Doutor em Ciência da Informação.

Orientador: Prof. Dr. Cláudio Gottschalg Duque

Brasília – DF

Março de 2013

Universidade de Brasília (UnB)

Faculdade de Ciência da Informação (FCI)

Programa de Pós-Graduação em Ciência da Informação (PPGCInf)

Doutorado em Ciência da Informação

Diretora da Faculdade: Profa. Dra. Elmira Luzia Melo Soares Simeão

Coordenadora da Pós-Graduação: Profa. Dra. Lillian Maria A. R. Alvares

Banca examinadora composta por:

Prof. Dr. Prof. Dr. Cláudio Gottschalg Duque (UnB) – Orientador.

Profa. Dra. Kelley Cristine Gonçalves Dias Gasque (UnB) – Membro.

Profa. Dra. Alessandra Dahmer (UFCSPA) – Membro.

Profa. Dra. Ana Estela Haddad (USP) – Membro.

Profa. Dra. Janaina Ferreira Fialho (UFG) – Membro.

Prof. Dr. André Porto Ancona Lopez (UnB) – Suplente.

M775w

MONTEIRO, Fernanda de Souza.

Web semântica e repositórios digitais educacionais na área de saúde: uma modelagem com foco no objetivo de aprendizagem para refinar resultados de busca, 2013.

202 f.

Tese (Doutorado)- Universidade de Brasília,
Brasília, 2012. Orientador: Prof. Dr. Cláudio Gottschalg
Duque.

1. Web semântica. 2. Repositórios digitais educacionais. I. Título

Endereço: Universidade de Brasília. Campus Universitário Darcy Ribeiro – Asa Norte.
Brasília – DF – Brasil. CEP 70910-900.

Site: <<http://www.cid.unb.br>>.

LISTA DE FIGURAS

Figura 1: Anatomy of an eLearning Object.....	36
Figura 2: Modelo de conteúdo molecular.....	38
Figura 3: Ciclo de vida do objeto de aprendizagem.....	40
Figura 4: Reutilização de dados interligados.....	46
Figura 5: Evolução Web atual para WS.....	47
Figura 6: Dados definidos e interligados com uma relação explícita e formal.....	47
Figura 7: Arquitetura da WS versão 2005.....	50
Figura 8: Triplas e grafos RDF.....	54
Figura 9: Construção de sentenças com o RDF.....	55
Figura 10: Metadados de um formulário de dados bibliográficos.....	73
Figura 11: Metadados embutidos em um documento.....	74
Figura 12: Elemento title, do Dublin Core.....	79
Figura 13: Entidades do Dublin Core como recursos.....	82
Figura 9: Representação esquemática da hierarquia dos elementos do LOM.....	84
Figura 15: Abrangência da modelagem proposta.....	96
Figura 16: Primeiros níveis da modelagem de conceitos.....	102
Figura 17: Ontologia de Carsten Ullrich (ULLRICH, 2004).....	106
Figura 18: Dengue Fever Ontology.....	111
Figura 19: Vaccine Ontology.....	111
Figura 20: International Classification of Functioning, Disability and Health (ICF).....	112
Figura 21: Higher Education Reference Ontology.....	114
Figura 22: Propriedades da The Cognitive Characteristics Ontology 0.2.....	115
Figura 22: Ontologia para ambientes interativos de aprendizagem.....	116
Figura 24: Health Insurance Portability and Accountability.....	117

Figura 25: Health Care Provider Taxonomy.	118
Figura 26: Classificação Brasileira de Ocupações – Família Médicos Clínicos.....	119
Figura 27: Ontologia de objetivos de aprendizagem.	120
Figura 28: Propriedades da classe “Objetivo de aprendizagem”.....	121
Figura 29: Exemplo de valores da propriedade “taxonomia” da classe “Assunto”..	124
Figura 30: Público-alvo e os conceitos relacionados.....	127
Figura 31: Recorte do código fonte da modelagem proposta.....	157
Figura 32: Diagrama de caso de uso – contexto e requisitos da modelagem proposta.	159
Figura 33: Diagrama de atividades – fluxo de disponibilização.....	160
Figura 34: Metadados da modelagem proposta.	163
Figura 35: Arquitetura da modelagem proposta.	165
Figura 36: Acervo de Recursos Educacionais em Saúde (ARES).	168

LISTA DE QUADROS

Quadro 1: Componentes correspondentes do DC, LOM e especificações para WS88	
Quadro 2: Conceitos que especificam o objetivo de aprendizagem.....	129
Quadro 3: Resumo de termos do vocabulário de metadados DC - Objetivo de aprendizagem.....	134
Quadro 4: Resumo de termos do vocabulário de metadados DC - Assunto	136
Quadro 5: Resumo de termos do vocabulário de metadados DC - Público-alvo....	136
Quadro 6: Resumo de elementos LOM - Objetivo de aprendizagem.	140
Quadro 7: Resumo de elementos LOM - Assunto.....	142
Quadro 8: Resumo de elementos LOM – Público-alvo	143
Quadro 9: Correlação DC e LOM para os conceitos que especificam o objetivo de aprendizagem.....	145
Quadro 10: Integração de metadados e valores propostos pela modelagem.	150
Quadro 11: Metadados ARES, metadados da modelagem proposta e recomendações.....	170

LISTA DE SIGLAS E REDUÇÕES

ADL	Advanced Distributed Learning Initiative
AGLS-MES	AGLS Metadata Element Set
AICC	Aviation Industry Aviation Industry Computer-Based Training Committee
ARES	Acervo de Recursos Educacionais em Saúde
ARIADNE	Aliance of Remote Instructional Authoring e Distribution
AVA	Ambiente Virtual de Aprendizagem
BIOE	Banco Internacional de Objetos Educacionais
BVS	Biblioteca Virtual de Saúde
CanCore	Canadian Core
CETL	Reusable Learning Objects
CSS	Cascading Style Sheets
CVSP	Campus Virtual de Saúde Pública
DAML + OIL	DARPA Agent Markup Language acrescida da Ontology Inference Layer ou Ontology Interchange Language
DC	Dublin Core
DCAM	Modelo Abstrato Dublin Core
DCMI	Dublin Core Metadata Initiative
DLP	Description Logic Programs
DLP	Description Logic Programs
EAD	Educação à Distância
EdShare	repositório da University of Southampton
Fiocruz	Fundação Oswaldo Cruz
F-Logic	Programação Lógica
GEM	Gateway to Educational Materials
HEAL	Health Education Digital Assets

HTML	HyperText Markup Language
IEEE LTSC	Institute of Electrical and Electronic Engineers Learning Technology Standards Committee
IMS	IMS Global Learning Consortium
IMS-LRM	IMS Learning Resource Metadata
JVM	Java Virtual Machine
LMS	Learning Management System
LOM	Learning Object Metadata
LOM	Standard for Learning Object Metadata
MARC 21	Machine-Readable Cataloging 21
MCT	Ministério da Ciência e Tecnologia
MDIS	Meta Data Interchange Specification
MDR	Information technology Metadata Registries
MEC	Ministério da Educação
Merlot	Multimedia Educacional Resource for Learning and Online Teaching
MET	Metadadta Ecoding e Transmition Standart
MLR-ISO 19788-2	Metadata Standard for Learning Resources
MMMI	Model for Metadata for Multimedia Information
MPEG-7	Multimedia Content Description Interface
MS	Ministério da Saúde
NDLR	The National Digital Learning Resources
AO	Objeto de Aprendizagem
OAI-PMH	Open Archives Initiative - Protocol for Metadata Harvesting
OBO	Open Biomedical Ontologies
OPAS	Organização Pan-Americana da Saúde
OWL	Web Ontology Language
PML	Proof Markup Language

RD	Repositório Digital
RDE	Repositório Digital Educacional
RDF	Resource Description Framework
REA Rede/OER	Rede Regional de Recursos Educacionais Abertos
RIO	Reusable Information Object
ROAR	Registry of Open Access Repositories
SCO	Sharable Content Object
SCORM	Shareable Content Reference Model
SEED	Secretaria de Educação a Distancia
SGML	Standard Generalized Markup Language
SGTES	Secretaria de Gestão do Trabalho e da Educação na Saúde
SparQL	RDF Query Language and Protocol
SUO	IEEE Standard Upper Ontology
SUS	Sistema Único de Saúde
TICs	Tecnologias da Informação e Comunicação
TM4L	Topics Maps for eLearning
UML	Unified Modelling Language
UMLS	Unified Medical Language System
UNA-SUS	Universidade Aberta do Sistema Único de Saúde
URI	Uniform Resource Identifier
VC	Vocabulário Controlado
W3C	World Wide Web Consortium
Web	Word Wide Web
Wisc-online	Wisconsin Technical College System
XML	eXtensible Markup Language
XSL	eXtensible Style Language

SUMÁRIO

1	INTRODUÇÃO	15
2	PROBLEMA	19
3	OBJETIVO GERAL	21
3.1	OBJETIVOS ESPECÍFICOS	21
4	JUSTIFICATIVA	22
5	REVISÃO DE LITERATURA	24
5.1	REPOSITÓRIO DIGITAL EDUCACIONAL	24
5.1.1	Educação à Distância e a importância do Repositório Digital Educacional	26
5.1.2	Diretrizes e padrões para Repositórios Digitais Educacionais	28
5.1.3	Exemplos de Repositórios Digitais Educacionais	31
5.1.4	Objeto de Aprendizagem disponibilizado em Repositórios digitais educacionais	35
5.1.5	Objeto de Aprendizagem e objetivo de aprendizagem	39
5.2	WEB SEMÂNTICA	42
5.2.1	Arquitetura da Web semântica	49
5.2.2	Estrutura e sintaxe da Web semântica	52
5.2.3	Lógica e semântica da Web semântica	57
5.3	METADADOS	72
5.3.1	Padrões de metadados	77
5.3.2	Interoperabilidade entre Dublin Core e Learning Object Metadata	86
6	METODOLOGIA	90
6.1	PESQUISA DOCUMENTAL	92
6.2	MODELAGEM BASEADA EM WEB SEMÂNTICA	94
7	DESENVOLVIMENTO DA MODELAGEM	95
7.1	TECNOLOGIAS DA WEB SEMÂNTICA PARA A RECUPERAÇÃO DA INFORMAÇÃO EM REPOSITÓRIOS DIGITAIS EDUCACIONAIS	96
7.1.1	Tecnologias para estrutura e sintaxe	96
7.1.2	Tecnologias para lógica e semântica	97
7.2	ESTRUTURA DA INFORMAÇÃO DESCRITIVA DOS OBJETOS DE APRENDIZAGEM	98
7.2.1	Modelagem de conceitos que especificam o objetivo de aprendizagem	99

7.2.2 Ontologias que formalizam os conceitos que especificam o objetivo de aprendizagem	103
7.2.3 Metadados que descrevem os conceitos que especificam o objetivo de aprendizagem	128
8 RESULTADOS	154
8.1 SISTEMATIZAÇÃO DA MODELAGEM PROPOSTA	154
8.2 SIMULAÇÃO DO RESULTADO DE BUSCA EM RDE NA ÁREA DE SAÚDE COM A IMPLEMENTAÇÃO DA MODELAGEM PROPOSTA	165
8.3 CONSIDERAÇÕES FINAIS	176
8.3.1 Utilização da ferramenta Protégé	176
8.3.2 A modelagem proposta	179
8.3.3 Ontologias e a WS	181
8.3.4 RDE, metadados e a WS	184
9 CONCLUSÃO	187
10 SUGESTÕES E PESQUISAS FUTURAS	189
REFERÊNCIAS	190

Dedicatória

*A memória da minha vizinha
Elília que não teve a oportunidade de
estudar, mas sabia de muitas coisas.*

Resumo

Apresenta modelagem baseada em Web Semântica para repositórios digitais educacionais na área de saúde, que disponibilizam objetos de aprendizagem com aspectos educacionais descritos em formulários de metadados. Foram identificadas características dos objetos de aprendizagem, com foco no objetivo de aprendizagem; agregadas tecnologias da Web Semântica à arquitetura desses repositórios e definida estrutura de informação adequada. Isso viabilizou inferências automáticas e contribuiu com o refinamento dos resultados de busca do usuário, aprimorando a recuperação da informação. Aborda repositórios digitais educacionais, as características e peculiaridades que influenciam sua utilização; Web Semântica e suas tecnologias para estrutura e sintaxe, lógica e semântica; e também são feitas considerações sobre metadados utilizados na descrição da informação dos objetos de aprendizagem. Como resultado, obtem-se modelagem que considera o processo de descrição e recuperação de objetos de aprendizagem, segundo seus aspectos educacionais; informações sobre a implementação; quadro com metadados LOM e Dublin Core correlacionados; e a simulação do resultado de busca em repositório digital educacional na área de saúde, com a implementação da modelagem proposta.

Palavras-chave: Objetos de aprendizagem. Repositórios digitais educacionais. Web Semântica. Recuperação da informação. Educação na saúde.

Abstract

The study presents modeling based on Semantic Web for educational digital repositories in healthcare, which provide learning objects with educational aspects described in metadata forms. Characteristics of learning objects have been identified, focusing on the learning objective; Semantic Web technologies have been aggregated to the architecture of these repositories; and adequate information structure has been defined. These enabled automatic inferences and contributed to the refinement of user's search results, improving information retrieval. The introduction discusses educational digital repositories and the specific characteristics that influence their use; then it discusses Semantic Web technologies to structure and syntax, logic and semantics; and metadata used to describe the information of learning objects. The resulting modeling considers the processes of description and retrieval of learning objects according to its educational aspects and information on its implementation. More over, the study obtained a framework with correlated LOM and Dublin Core metadata, and the simulation of search results in digital library education in healthcare, with the implementation of the proposed modeling.

Keywords: Learning objects. Educational digital repositories. Semantic Web. Information retrieval. Healthcare education.

1 Introdução

O Objeto de Aprendizagem (OA) é um recurso estruturado para disponibilizar conteúdo para ensino-aprendizagem. Nessa perspectiva, os OAs são adequados para desenvolver conteúdos para Educação a distância (EAD), especialmente com a aplicação de modelos de referencia que especificam diretrizes para reutilização de conteúdos como o Sharable Content Object Reference Model (SCORM), Institute of Electrical and Electronics Engineers (IEEE), entre outros.

Esses OAs, à medida que são disponibilizados na Web, contribuem com a redução de gastos com sua produção, além de permitir constante aprimoramento. Para viabilizar essa disponibilização são constituídas coleções no Learning Management System (LMS) para que os OAs sejam acessados no próprio ambiente dos cursos ou, mais recentemente, são desenvolvidos Repositórios Digitais Educacionais (RDE) onde um número cada vez maior de objetos é compartilhado. Para isso, o RDE deve apresentar uma arquitetura adequada e dispor de informação descritiva suficiente à seleção criteriosa de OAs, especialmente relacionadas aos seus aspectos educacionais. Além disso, é recomendável que o RDE possa interagir com o LMS (MIRANDA, 2004).

Ao mencionar a importância da informação que descreve um OA e a necessidade de arquitetura adequada, cabe acrescentar que esses são pontos intensamente discutidos no desenvolvimento de RDEs. Existem diversos padrões de metadados para a descrição dos objetos e essa descrição, muitas vezes, é realizada pelo autor ou desenhista instrucional. Esses metadados são utilizados na busca de OAs para atender uma ação educacional, isso demanda a análise de diferentes aspectos que subsidiam a escolha daqueles mais adequados. No que se refere à arquitetura, a estrutura navegacional, as opções de busca e a apresentação de resultados nesses RDEs, observam-se características comuns a outros tipos de repositórios, ou seja, nem sempre é relevada sua peculiaridade educacional.

Esse contexto motiva o desenvolvimento dessa pesquisa que pretende agregar à arquitetura dos RDEs tecnologias capazes de refinar o resultado de busca de OAs, levando em consideração seu enfoque educacional. Para tanto, as contribuições da Web Semântica (WS) são o ponto crucial, uma vez que a

evolução de suas tecnologias está direcionada a estruturação e atribuição formal de semântica aos conteúdos disponibilizados na Web. A WS é uma evolução da Web que destaca a estruturação de conteúdos, descrição e atribuição de semântica aos dados a partir da implementação de padrões (TIM BERNERS-LEE, 2000). Considerando que o OA é um documento digital e a descrição da sua informação é necessária para sua busca em RDE, questões inerentes à sintaxe e semântica formais são fundamentais e podem ser favorecidas com soluções propostas pela WS. A WS visa facilitar e melhorar a recuperação de conteúdos relevantes, permite que aplicações computacionais façam associações e inferências automáticas para minimizar esforços humanos no momento da seleção de resultados de busca. DUQUE (2005) sugere que o uso da lingüística computacional, conjuntamente com a geração semi-automática de ontologias, auxilia e melhora a recuperação de informação relevante a partir de uma base de dados textual em língua portuguesa. Além disso, destaca a comunicação baseada em usuários, notadamente considerando o papel das redes sociais, fator relevante para o sucesso dos RDEs (TEIXEIRA; DUQUE, 2009).

A informação que descreve um OA é representada em metadados e a relação semântica, centrada no objetivo de aprendizagem, contribui com a pertinência do resultado de busca. O objetivo de aprendizagem é aquilo [resultado] que o professor [autor] espera do aluno [usuário] ao final do processo de aprendizagem (DRISCOLL, 1998). Um modelo capaz de estabelecer quais tecnologias podem ser implementadas para estruturar a representação da informação de OAs, ressaltando as características educacionais, e formalizar a semântica, contribui para maior satisfação do usuário que possui interesses pautados não apenas no assunto do documento digital, mas na valoração de um OA frente a um interesse de aprendizagem específico, com contexto e público-alvo definidos.

Essa pesquisa é a confluência de investigações anteriores que elencaram temas como “Representação da informação”, “Construção de ontologias de domínio”, “Arquitetura da informação para repositórios”, além de experiências com repositórios educacionais digitais. A citar o capítulo de livro “Organização da informação: proposta de elementos de arquitetura da informação para repositórios digitais institucionais, baseados na descrição física e descrição temática” (MONTEIRO, 2010); e os artigos “Plataformas de software livres para o armazenamento e descrição de objetos educacionais digitais: estudo inicial para customização do Banco Internacional de Objetos Educacionais digitais” (MONTEIRO; DUTRA, 2008); “Uma política para

desenvolver o acervo de recursos educacionais em saúde” (MONTEIRO; JACOB, 2011); e “Life Cycle Assessment Ontology” (BRASCHER; MONTEIRO, 2007).

A Ciência da Informação abarca essas temáticas ao ter a informação como seu objeto de pesquisa. Segundo Le Coadic (2004), para a Ciência da Informação, a informação comporta um elemento de sentido. É um significado transmitido a um ser consciente por meio de uma mensagem inscrita em um suporte espaço-temporal: impresso, sinal-elétrico, onda sonora, etc. Capurro (2003) completa que a área de Ciência da Informação se atém aos aspectos contextuais, semânticos, sociais e culturais da informação.

Ter a WS como base para o refinamento dos resultados de busca nos RDEs aproxima seu conceito dos problemas investigados pela Ciência da Informação. O conceito de WS surgiu no âmbito da Ciência da computação para descrever uma série de pesquisas que têm como objetivo principal possibilitar melhor aproveitamento das potencialidades do ambiente Web por meio do uso de linguagens computacionais e metadados (RAMALHO; VIDOTTI; FUJITA, 2007). Como afirma Ramalho, Vidotti e Fujita (2007):

Os conceitos e tecnologias subjacentes ao projeto Web Semântica podem ser considerados como uma renovação ou desdobramento dos tradicionais métodos representação, organização e recuperação de informações, apontando a possibilidade de contribuições da área de Ciência de Informação, devido ao seu embasamento teórico referente a formas de representação e as práticas profissionais identificadas em seu campo de atuação. (RAMALHO; VIDOTTI; FUJITA, 2007, p. 1).

Assim, pesquisas com temática semelhante foram encontradas na Ciência da Informação e, também, na Ciência da Computação e Educação. Essas pesquisas abordam, em sua maioria, o impacto da WS nos resultados de busca em ambientes informacionais diversos; a emergência dos OAs no ensino-aprendizagem, evidenciando a necessidade de soluções para melhor descrevê-los; e como os OAs podem ser compartilhados. A investigação de trabalhos anteriores foi fundamental para identificar os avanços ocorridos na área e certificar a efetiva contribuição desse estudo a partir do desenvolvimento de objetivos ainda não explorados. Foram realizadas pesquisas em teses e dissertações no Brasil e exterior, além de artigos científicos e *sites* de grandes grupos responsáveis por diretrizes e

recomendações, tanto para *e-learning*¹ quanto para WS.

Os objetivos propostos nesta pesquisa revelam seu ineditismo, bem como, os aspectos a serem estudados e relacionados. Ainda assim, alguns estudos com propósitos similares podem ser destacados, tais como, o EDUCO que tem como objetivo a representação do conteúdo de OA para possibilitar buscas mais adequadas e promover seu reuso a partir da exploração de navegações entre seus relacionamentos semânticos. Esse estudo aborda o impacto da WS nos sistemas de *e-learning* que disponibilizam conteúdos educativos e recomenda o uso de metadados, ontologias e a mapas de conceitos ou mapas de tópicos para orientar os agentes de software na localização, recuperação e reutilização de objetos de aprendizagem (GONÇALVES, 2007). Já a proposta de Wanderson Rigo, Renato Fileto e Roberto Willrich (2010) é a construção de sistema para apoiar a anotação e a navegação semântica sobre o conteúdo de repositórios na Web, utilizando conhecimento específico de domínio e técnicas de visualização. O Learning content management system Using Innovative Semantic web services Architecture (LUIA, 2012)² é um projeto fruto de consórcio formado por países da União Europeia, incluindo pesquisadores e instituições com a missão de enriquecer e tornar mais flexíveis os processos de especificação, busca e utilização de OAs em LMSs e REDs. O LUIA utiliza WS para fornecer semântica necessária aos sistemas computacionais e automatizar tarefas.

¹ O termo *e-learning*, cuja definição mais adequada é “educação a distância mediada pela Web”, será adotado na redação desta pesquisa devido sua recorrente utilização no meio científico e de desenvolvimento.

² Link: <<http://www.luisa-project.eu/dvd/motivation.php>>.

2 Problema

Repositórios digitais têm sido amplamente utilizados para disponibilizar conteúdos informacionais em diferentes áreas do conhecimento. A Educação, com o propósito de tornar acessíveis conteúdos de apoio ao ensino-aprendizagem, tem desenvolvido Repositórios Digitais Educacionais (RDE). No âmbito da Educação a Distância (EAD), vertente da Educação amplamente utilizada na capacitação e formação profissional em larga escala, o interesse na busca em RDEs vai além do conteúdo expresso, representado pelo assunto. Buscam-se materiais informacionais sobre um assunto, mas que possam ser incorporados a estratégias educacionais e, conseqüentemente, atendam o objetivo de aprendizagem.

A produção em EAD na área de saúde têm se destacado, especialmente com iniciativas governamentais que oferecem oportunidades de formação e qualificação ao profissional de forma integrada ao seu processo de trabalho (PORTAL DA SAÚDE, 2012). A multidisciplinaridade das equipes envolvidas nessa produção é comum e advém da ampla utilização de tecnologias para localizar, aprimorar e desenvolver materiais informacionais. É fundamental que os membros das equipes dialoguem entre si e compreendam as reais necessidades de informação.

Assim, diversos tipos de materiais informacionais podem ser buscados e, atualmente, cresce o interesse pelo Objeto de Aprendizagem (OA) (DRISCOLL, 1998). O termo OA, adotado nesta pesquisa sem fazer menções às influências pedagógicas, generaliza o material informacional que possui conteúdo, atividade e avaliação; apresenta explicitamente objetivo de aprendizagem e é disponibilizado em RDEs. Também é encontrado na literatura com os sinônimos material instrucional (MOULIN; PEREIRA, 1999), material de aprendizagem online (MERLOT, 2011), entre outros.

No RDE, os OAs são classificados por tema e/ou tipos, com possibilidade de busca por palavras-chave. O resultado de busca, mesmo que apresentado a partir de estratégia de busca que represente o objetivo de aprendizagem, não considera a relação entre o assunto do OA e o objetivo de aprendizagem, que por sua vez, relaciona-se com o público-alvo e o contexto educacional. A dissociação desses aspectos descritivos compromete a seleção

adequada dos OAs recuperados. O RDE deve apresentar estrutura apropriada e dispor de informação descritiva suficiente à seleção de OAs, especialmente relacionadas aos seus aspectos educacionais (MIRANDA, 2004).

A descrição do OA em RDEs é realizada a partir do preenchimento de metadados como título, autor, assunto, objetivo de aprendizagem, público-alvo, contexto educacional, entre outros. Esses metadados, independente do padrão adotado, agregam informação, ou mesmo atribuem sentido educacional ao OA em si. Entretanto, a busca de quaisquer termos registrados nesses metadados retorna uma lista de resultados que não considera a relação semântica entre os aspectos descritivos que indicam o uso do objeto no ensino-aprendizagem. E, ainda, algum OA pode ser desconsiderado simplesmente por não apresentar os termos utilizados na composição da estratégia de busca.

Tendo em vista a representação da informação e o refinamento dos resultados de busca em RDEs na área de saúde, destaca-se a necessidade da modelagem baseada em WS, com foco no objetivo de aprendizagem e suas relações semânticas. Ao relevar as tecnologias da WS para este trabalho será possível estruturar e formalizar aspectos descritivos dos OAs, considerando o contexto de ensino-aprendizagem. Assim, a questão que se pretende responder com esta pesquisa é:

Quais elementos uma modelagem baseada em Web Semântica para repositórios digitais educacionais na área de saúde deve ter para contribuir com o refinamento dos resultados de busca do usuário?

3 Objetivo geral

Propor modelagem baseada em Web Semântica, para repositórios digitais educacionais na área de saúde, com foco no objetivo de aprendizagem.

3.1 Objetivos específicos

- 1) Identificar as características dos objetos de aprendizagem disponíveis nos repositórios digitais educacionais na área de saúde;
- 2) Verificar tecnologias da Web Semântica adequadas para a estruturação da informação descritiva dos objetos de aprendizagem dos repositórios digitais educacionais na área de saúde;
- 3) Definir modelagem que contribua para o refinamento do resultado da busca dos objetos de aprendizagem disponíveis em repositórios digitais educacionais na área de saúde.

4 Justificativa

A produção de cursos de EAD assistidos pela Web aumenta à medida que é reconhecido seu potencial para atingir um maior número de pessoas com custos cada vez menores. Esses cursos são formados por OAs organizados em ambientes virtuais de aprendizagem. Para isso, os OAs possuem características como flexibilidade para reutilização, concepção modular (modularidade) com respeito às características das partes que compõem um todo, interoperabilidade e possibilidade de busca a partir da sua indexação com uso de metadados. Esses metadados apresentam informação descritiva do OA, que além de identificá-lo individualmente, com seu título e autor, por exemplo, agregam informação educacional, como objetivo de aprendizagem e contexto educacional. Dessa forma, OAs são disponibilizados em REDs para usuários com necessidades de busca específicas. As buscas podem ter resultados mais precisos ao serem identificadas às relações existentes entre os metadados que descrevem o OA. A autora Margaret Driscoll (1998) destaca que o crescimento na produção de EAD reforça a importância do aprimoramento do processo de busca de OAs.

Propor uma modelagem baseada em WS para repositórios digitais educacionais na área de saúde, com foco no objetivo de aprendizagem, irá contribuir para resultados de buscas mais precisos. Na WS, a especificação do objetivo de aprendizagem pode compor sentenças com sujeito, predicado e objeto, nas quais a relação entre o sujeito e objeto é estabelecida pelo predicado. Cada um dos termos que compõem sentenças pode ter conceitos determinados, categorizados, relacionados entre si e formalizados. As relações estabelecidas com termos que descrevem o público-alvo e contexto para o qual o OA foi pensado permitem, entre outras coisas, apresentar resultados de busca que considerem que o objeto de aprendizagem é para um *“aluno de química de ensino médio identificar a composição química”* e, ainda, que outro objeto de aprendizagem é para um *“aluno de química de ensino superior ter capacidade de diferenciar a composição química dos ácidos”*. Assim, a busca pelo assunto *“composição química”* pode ter seu resultado refinado a partir do público-alvo do OA, com nível de abstração expresso em seu objetivo de aprendizagem, observando, por exemplo, que o processo cognitivo de *“diferenciar”* requer o processo cognitivo de *“identificar”*.

Esses termos utilizados para descrever OAs são apresentados em metadados, contudo, o estabelecimento de suas relações e a estrutura dos RDEs que irão contê-los demandam estudos. As tecnologias para representar o conteúdo dos OAs de maneira formal são fundamentais para modelagem a ser proposta. Isto permitirá maior satisfação das necessidades de busca do usuário em questão. Assim, justifica-se o interesse da pesquisa de identificar as características dos objetos de aprendizagem, verificar tecnologias da WS adequadas para a estruturação da sua informação descritiva e definir a modelagem. Para esse trabalho, as tecnologias da WS viabilizam o compartilhamento dos conceitos e relações que definem objetivos de aprendizagem. Além disso, elimina ambiguidades no momento da seleção dos resultados de busca.

Gonçalves (2007), ao discutir a busca de OAs, ressalta que as estratégias atuais de indexação e recuperação de informação mostram-se insatisfatórias e ineficazes, pois a demanda por capacitação é crescente, ao passo que a descoberta de conteúdos de aprendizagem torna-se problemática. O autor afirma que isso acontece porque a busca por palavras-chave desprovidas de significado e contexto faz com que esses conteúdos permaneçam perdidos, dependentes da verificação ou interpretação humana. Logo, expectativa paralela, relacionada aos resultados da pesquisa, é aumentar a qualidade dos metadados atribuídos à descrição de OAs, pois os termos a serem utilizados podem ter significado e contexto agregados, quando baseados na estrutura que representa a semântica dos objetivos de aprendizagem.

5 Revisão de literatura

Apresenta-se neste capítulo o referencial teórico da pesquisa. Isto se dá a partir de uma abordagem sobre repositórios digitais educacionais, as características e peculiaridades que influenciam sua utilização. Seguindo a perspectiva, é realizado estudo da Web Semântica e sua arquitetura é percorrida segundo as tecnologias para estrutura e sintaxe, lógica e semântica. Assim, evidenciam-se as contribuições dessa área para representação do conteúdo dos objetos de aprendizagem e sua recuperação. E diante da proposta de definir uma modelagem, com foco no objetivo de aprendizagem, que contribua para o refinamento do resultado da busca em repositórios digitais educacionais, são feitas considerações sobre metadados e como estes são utilizados na descrição da informação, ou mais especificamente, dos objetos de aprendizagem a serem buscados.

5.1 Repositório digital educacional

De maneira geral, a *Word Wide Web* (Web) tem favorecido a publicação e disseminação de conteúdos informacionais. Esses conteúdos podem ser acessados em blogs, sites, portais, entre outras fontes de informação. Nesse contexto, crescem as soluções para o agrupamento de documentos que apresentam conteúdos com objetivos e/ou temas comuns. Uma dessas soluções é o Repositório Digital (RD), que permite além do armazenamento, a descrição e busca de informação em formatos digitais.

O RD surge no contexto acadêmico, a partir da Iniciativa dos arquivos abertos, com a pretensão de disponibilizar uma variedade de documentos não publicados tradicionalmente por editoras e revistas científicas (HARNAD, 2001). Vale citar a iniciativa de Paul Ginsparg ao criar, em 1991, um repositório mantido pelo Laboratório Nacional de Los Alamos (Novo México) (LAGOZE; SOMPEL, 2000). Esses RDs possuem uma concepção mais ampla que as bibliotecas digitais no sentido de permitir a disseminação de documentos em versões preliminares para apreciação de pares e podem ter diferentes definições que variam de acordo com seu contexto de desenvolvimento e

objetivo. A definição do Digital Repositories JISC³ (HAYES, 2005) é ampla e adequada aos propósitos dessa pesquisa ao afirmar que um repositório digital é aquele onde os conteúdos informacionais digitais podem ser pesquisados e recuperados para uso posterior.

Os RDs apresentam estrutura e características próprias, capazes de permitir a um usuário comum a recuperação e o acesso a documentos de seu interesse. Geralmente estão sob a responsabilidade técnica e administrativa de uma instituição ou organismo (CAFÉ, 2003). Em um RD destacam-se as seguintes características gerais (HEERY; ANDERSON, 2005 apud MARTINS; RODRIGUES; NUNES, 2008):

- Suporta a submissão de documentos pelo autor, profissional de informação ou responsável pelo repositório;
- Disponibiliza documentos acrescidos de seus metadados descritivos;
- Oferta serviços de busca, localização e acesso aos documentos;
- Garante a atualização e disponibilidade do conteúdo a partir de uma gestão sustentável.

Um RD com conteúdo e funcionalidades focadas na educação é chamado Repositório Digital Educacional (RDE), diferenciando-o dos demais. A preocupação de seus mantenedores não é disponibilizar documentos que sirvam apenas de referência para pesquisas, levantamentos bibliográficos e subsídios teóricos, como a maioria dos RDs acadêmicos. Um RDE deve disponibilizar documentos que possam ser incorporados e utilizados diretamente no ensino-aprendizagem.

Para isso, além das características gerais elencadas para o RD, o RDE deve apresentar uma arquitetura adequada, dispor de informações descritivas suficientes à seleção criteriosa de documentos, especialmente relacionadas aos aspectos educacionais. Também é recomendável que o RDE possa interagir com ambientes virtuais de aprendizagem possibilitando, por exemplo, a exportação e importação de seus documentos (MIRANDA, 2004; NASCIMENTO, 2010; MCLAREN, 2004).

³ Repositório digital educacional. Disponível em: <<http://www.jisc.ac.uk/>>.

5.1.1 Educação à Distância e a importância do Repositório Digital Educacional

A Educação à Distância (EAD) é uma vertente da Educação que tem sido amplamente utilizada na capacitação e formação profissional. Isso se deve ao potencial de alcance, que ultrapassa as fronteiras geográficas, e a redução de custos, com um investimento inicial robusto, mas que diminui gradativamente a cada oferta de curso (SLOEP, 2003). Logo, a disponibilização de conteúdos em RDEs para acesso e utilização por um número cada vez maior de interessados é uma alternativa muito pertinente.

Vale lembrar, que a EAD não é unicamente assistida pela Web, o que pode ser observado com exemplos clássicos de envio de materiais impressos via correio, gravação e distribuição de CD-ROM e mesmo aulas e cursos inteiros televisionados (SIMONSON et. al., 2000). Porém, para esta pesquisa, a EAD assistida pela Web interessa especialmente por ser a principal “consumidora” dos conteúdos informacionais disponibilizados nos RDEs. Como já foi dito anteriormente, o termo *e-learning* será adotado no decorrer do texto para nomear a Educação à distância mediada pela Web. O *e-learning* utiliza Tecnologias da Informação e Comunicação (TICs) e novas estratégias de aprendizagem (SILVA, 2007).

Entre os diversos conceitos de EAD apresentados na literatura, destaca-se o de Kaplan-Leiserson (2005) que afirma que EAD é o processo de aprendizagem sem contato regular e presencial entre o estudante e o professor que se encontram separados temporal e/ou geograficamente. Desse modo, os conteúdos informacionais para a aprendizagem são transmitidos via meios assíncronos, como a correspondência escrita, e síncronos, como plataformas *online* que permitem interações em tempo real.

A transposição do ensino presencial para ensino mediado por plataformas *online* é uma tarefa complexa que envolve recursos e profissionais distintos. O professor ou especialista assume o papel de conteudista e elaborador das estratégias educacionais planejadas para o cumprimento de objetivos de aprendizagem – essa segunda tarefa também pode ser desempenhada por um desenhista instrucional. Demais conhecimentos necessários para implementação do *e-learning* são aplicados por *designers*, profissionais da computação, entre outros. Em meio a essa atuação multidisciplinar, está o emprego da informática que, como afirma Kratz *et al.* (2005), cumpre o papel de auxiliar no processo de ensino-aprendizagem.

Na composição de cursos, módulos de cursos, unidades ou atividades de ensino-aprendizagem pode ser utilizado um Ambiente Virtual de Aprendizagem (AVA) com diferentes funcionalidades capazes de estruturar e promover a estratégia educacional. Essas funcionalidades permitem o uso de conteúdos como textos, vídeos, animações, simulações, etc., chamados genericamente de objeto de aprendizagem (OA). Os AVAs, ou Learning Management System (LMS), são ambientes criados especialmente para permitir a interação entre o aluno, o OA e o professor ou tutor. Ou seja, ocorre a mediação da aprendizagem, implementação da estratégia educacional e utilização de conteúdos informacionais, definidos como OA. Alguns exemplos de AVA são o Moodle⁴, TeleEduc⁵, E-proinfo⁶, Blackboard Learn⁷ e TopClass⁸

Os OAs são disponibilizados no RDE e podem ser combinados e recombinados compondo novos e mais complexos OAs. A organização e tratamento informacional adequado para esses objetos são primordiais para garantir sua localização e utilização no momento desejado. Dada a diversidade dos OAs e dos seus usuários, os RDEs assumem fundamental importância na sua recuperação. Não é apenas um documento de referência o que se busca, é um objeto que condensa, aplica, demonstra, instrui, apresenta um conteúdo. Assim, o RDE possui um papel que transcende o de fonte de informação sobre determinado assunto e passa a fornecer conteúdos capazes de viabilizar a produção de uma estratégia educacional, permitindo a recuperação de OAs que podem, por exemplo, ser diretamente inseridos em um curso. O OA, dependendo de sua complexidade, assume parte substancial na composição de um curso, módulo de curso, unidade ou atividade de aprendizagem, podendo, até mesmo, ser um curso completo.

O crescimento na produção de *e-learning* reforça a importância da localização e utilização de OAs como mostra a autora Margaret Driscoll ainda em 1998 (DRISCOLL, 1998). O RDE favorece a formação de coleções específicas, a organização, o compartilhamento, a recuperação e reutilização de OAs. Isso reduz custos e amplia o público-alvo, propostas intrínsecas à EAD. Assim, é necessário compreender não apenas a finalidade do RDE, mas a importância da sua alimentação, as diretrizes de desenvolvimento e a definição de padrões para submissão e acesso de OAs.

⁴ Link: <<http://www.moodlebrasil.net/moodle/>>

⁵ Link: <<http://www.teleduc.org.br/>>.

⁶ Link: <<http://eproinfo.mec.gov.br/>>.

⁷ Link: <<http://www.blackboard.com/Platforms/Learn/Overview.aspx>>.

⁸ Link: <<http://www.wbtsystems.com/product/#tooltip10>>.

5.1.2 Diretrizes e padrões para Repositórios Digitais Educacionais

Frequentemente o volume de informações disponíveis faz com que os usuários de RDEs tenham expectativas frustradas com um resultado de busca e assim criem novos OAs, com dispendioso esforço que poderia ser evitado caso obtivesse resultados mais precisos (MIRANDA, 2004). O cumprimento de diretrizes e padrões faz com que os RDEs sejam mais efetivos na disseminação de OAs, como destaca Rafael Relvão (2006):

Na cadeia do processo que vai desde a recolha da informação até a sua disponibilização em formato adequado, há uma etapa que se reveste da máxima importância: a organização dos conteúdos de aprendizagem. A organização de conteúdos é especialmente importante quando eles são partilhados ou armazenados e, mais tarde, recuperados para reutilização. Para que a partilha, reutilização ou difusão seja feita por diferentes entidades, em Plataformas diferentes - situação muito generalizada - é essencial que a organização dos conteúdos seja feita segundo alguma estrutura, normas técnicas ou leis consensuais (RELVÃO, 2006, p. 2).

A alimentação dos RDEs, muitas vezes, adota o auto-arquivamento, ou seja, requer que os próprios autores depositem os OAs e gerenciem versões e atualizações, bem como descrição com o uso de metadados. Contudo, as diretrizes do repositório devem ser previamente definidas, inclusive o fluxo de trabalho que fornece critérios como permissões para a submissão, políticas para avaliação, edição de versões e o acesso por meio de pesquisas e *downloads*. Entretanto, o auto-arquivamento é uma responsabilidade que recai sobre o autor e gera discussões, pois modifica o modo de divulgação da produção que passa a não ter mais intermediários. Alguns problemas têm sido identificados nesse contexto como o fato de que qualquer autor pode submeter OAs sem que a avaliação e revisão sigam critérios rígidos, comprometendo a qualidade dos conteúdos disponíveis. Outro problema, talvez ainda maior, está no preenchimento dos metadados pelos próprios autores, considerando que estes podem não ter conhecimento suficiente sobre como os OAs devem ser descritos. Uma das possíveis soluções para minimizar os problemas gerados pelo auto-arquivamento é a qualificação do que é disponibilizado. Nesse sentido, a padronização e estabelecimento de políticas que discorram sobre a avaliação do conteúdo, dos aspectos midiático-pedagógicos e, ainda, o estabelecimento de regras para a descrição, são soluções adequadas.

No tocante à estrutura do conteúdo a ser armazenado, é possível observar que a formação de um RDE depende da instituição mantenedora, dos seus objetivos e do seu escopo. Este tipo de repositório agrupa os OAs para produção e oferta de *e-learning* e é preciso fazer definições que delimitem e expressem o conteúdo a ser coberto. As primeiras definições dizem respeito aos tipos de OAs abrangidos, granularidade e temas abordados. Algumas definições podem ser ampliadas ao longo do desenvolvimento do repositório, porém, vale lembrar que elas impactam diretamente na especificação dos metadados.

Quanto ao processo de submissão dos OAs, na maioria das vezes, condicionado à ferramenta escolhida para o desenvolvimento do repositório, é preciso prever níveis de acesso, responsáveis e permissões. Essas ferramentas possuem interface de submissão baseada na Web, que permite o depósito (*upload*) do documento e o preenchimento de metadados descritivos. Uma vez depositados, os OAs são submetidos ao controle de qualidade, ou seja, a avaliação antes de serem disponibilizados para busca e recuperação no repositório. Esse controle pode variar, contudo, é possível verificar se o indivíduo que submeteu tem autoridade para isso, se o documento está de acordo com as definições do repositório e se os metadados associados são adequados (HUNTER; DAY, 2005).

O acesso ao conteúdo dos RDEs pode ser viabilizado aos usuários internos ou externos à instituição mantenedora. O conteúdo é suscetível a barreiras ou níveis diferenciados de acesso, pois pode haver razões para não estar disponível publicamente ou possuir restrições estabelecidas pelo regime de direitos autorais, por exemplo. É comum que um OA disponível em um RDE tenha os direitos patrimoniais cedidos pelos autores à instituição mantenedora, de acordo com a Lei 9.610 de 19 de fevereiro de 1998 (BRASIL, 1998), ou conforme licenças específicas. Isso se deve a necessidade de armazenar, copiar, reproduzir, exibir, executar, distribuir, transferir e criar obras derivadas – novos objetos. Essas diretrizes devem ser claras e, em geral, a transferência é concedida em caráter parcial e não exclusivo, não havendo impedimento para que os autores utilizem os OAs como desejarem, inclusive comercialmente. A referência aos nomes dos autores, que constitui um direito moral, deve ser respeitada sempre que os objetos forem veiculados ou utilizados.

Uma vez respeitado o regime de direitos autorais, combinar ou reutilizar diferentes OAs são as principais preocupações do estabelecimento de padrões. Para isso, é possível adotar padrões para criar conteúdos; estruturar

e reunir objetos considerando suas características individuais; e definir metadados que permitam sua recuperação (OLSEN, 2012). Atualmente, muitas organizações, como a World Wide Web Consortium (W3C), IMS Global Learning Consortium (IMS), Aviation Industry Aviation Industry Computer-Based Training Committee (AICC), Institute of Electrical and Electronic Engineers Learning Technology Standards Committee (IEEE LTSC) e Advanced Distributed Learning Initiative (ADL), têm se preocupado em desenvolver especificações para prover a reutilização de OAs e a integração entre plataformas, tanto LMS quanto RDE. Entre as iniciativas das organizações encontradas na literatura, cabe destaque para o Shareable Content Reference Model (SCORM), que será discutido nessa seção, e o Learning Object Metadata (LOM), discutido na Seção 5.3.1.2 Learning Object Metadata IEEE LTSC, p.71

O SCORM apresenta um conjunto unificado de padrões e especificações para conteúdos, tecnologias e serviços para *e-learning* que, conforme Rouyet e Martín (2011) pontuam, permite ao AVA encontrar, importar, compartilhar, reusar e exportar OAs de maneira normalizada. A proposta do SCORM é ser um padrão amplamente utilizado, desenvolvido a partir de outros padrões, como o AICC, o IMS e o Alliance of Remote Instructional Authoring e Distribution (ARIADNE) (GOMEZ *et al.*, 2005) (KRATZ *et al.*, 2005). Este padrão pretende garantir a interoperabilidade, durabilidade e acessibilidade de OAs a partir da composição de manuais técnicos, a saber (SCORM, 2011):

- a) SCORM Overview Book: apresenta os principais conceitos e informação;
- b) SCORM Content Aggregation Model: modelo de agregação de conteúdos que especifica como combinar, agregar, descrever e sequenciar OAs utilizando metadados;
- c) SCORM Run-Time Environment: especifica o ambiente de execução dos conteúdos e como registrar o percurso do aluno na utilização de OAs e LMSs;
- d) SCORM Sequencing and Navigation: modelo de sequenciamento e navegação que descreve como os conteúdos podem ser ordenados para o aluno.

Com o uso do SCORM, OAs podem ser desenvolvidos, armazenados e distribuídos independentemente das plataformas, favorecendo a criação de RDEs (RELVÃO, 2006). Este padrão traz considerações importantes sobre

interoperabilidade, requisito amplamente discutido no âmbito dos repositórios, que envolve a adoção e uso de protocolos comunicação e transferência de informação.

5.1.3 Exemplos de Repositórios Digitais Educacionais

A ampla utilização do OA é assegurada a partir da sua disponibilização em RDEs (ANTA, 2003 apud SALVE, 2010). É crucial a definição de como os usuários acessaram os OAs, em qual sistema e a partir de quais serviços (BARRITTI; ALDERMAN, 2004). Assim, os RDEs viabilizam a busca e recuperação de OAs (DOWNES, 2001). A descrição dos objetos nesses repositórios permite que o usuário os encontre a partir da sua necessidade de busca, considerando aspectos como assunto, formato, data, etc.

Com o crescimento e valorização da EAD, muitos repositórios têm surgido para atender necessidades específicas, fortalecer e oferecer subsídios a esse tipo de ensino. Nessa perspectiva, os RDEs apresentam classificações distintas, arquiteturas, metadados, e características próprias de disponibilização de conteúdos. A seguir são apresentados alguns exemplos.

Multimedia Educational Resource for Learning and Online Teaching (Merlot)⁹: é um repositório de acesso livre que apresenta diversos *links* para conteúdos relacionados ao ensino-aprendizagem assistidos pela Web, tais como, redes sociais, sites, objetos de aprendizagem, etc. Além disso, mantém uma comunidade colaborativa que contribui com informação, documentos, comentários e experiências de ensino-aprendizagem. Compartilha dicas e conselhos sobre educação e ainda permite o reconhecimento e qualificação de todos os contribuidores (MERLOT, 2011). No Merlot, os OAs estão classificados por tema em uma estrutura navegacional e há possibilidade de busca por palavras-chave. O resultado de busca apresenta para o usuário um formulário com informação descritiva e o *link* para acesso ao conteúdo. Também há uma estrutura navegacional organizada por tipos de OAs que apresenta tutoriais, animações, simulações, etc. O site recomenda que para utilizar o OA o usuário vá até o site de origem e se informe dos direitos autorais e condições para uso.

Wisc-online¹⁰: é um repositório de OAs produzidos pela Wisconsin Technical College System para serem utilizados nos seus cursos ou aplicações *online* diversas.

⁹ *Link*: <<http://www.merlot.org/merlot/index.htm>>.

¹⁰ *Link*: <<http://www.wisc-online.com/Default.aspx>>.

Os OAs são mantidos como alternativas para o aprendizado com recursos midiáticos que adicionam flexibilidade ao ensino-aprendizagem, permitindo aplicação de conceitos, simulações e instruções que ampliam o conhecimento transmitido em sala de aula (WISC, 2011). Os OAs estão classificados por tema em uma estrutura navegacional e há possibilidade de busca por palavras-chave.

Reusable Learning Objects (CETL)¹¹: projeto financiado pelo Conselho de Ensino Superior da Inglaterra (HEFCE) que possui RDE mantido por um centro de excelência focado na concepção, desenvolvimento e uso de OAs. As instituições envolvidas são a London Metropolitan University, University of Cambridge e University of Nottingham. O CETL desenvolve e avalia os OAs contando com a participação de comunidades de especialistas em workshops e oficinas. Docentes, estudantes e desenvolvedores multimídia trabalharam em equipe para criar, utilizar e avaliar pedagogicamente objetos. Os OAs estão classificados por tema em uma estrutura navegacional.

The National Digital Learning Resources (NDLR)¹²: é uma iniciativa irlandesa para promover o compartilhamento de conteúdos de aprendizagem e experiência de ensino com a participação de universidades, institutos e faculdades de tecnologia (NDLR, 2011). Os OAs estão classificados por tema em uma estrutura navegacional e há possibilidade de busca por palavras-chave. Também há uma estrutura navegacional em que os OAs podem ser consultados por autor, título e data de submissão.

EdShare¹³: é o repositório da University of Southampton. Foi elaborado como uma plataforma *online* para ensino-aprendizagem, em que os usuários podem gerenciar seus conteúdos com o uso de *tags* e criação de coleções em seu próprio perfil. Fornece estatísticas de uso de OAs, os disponibiliza classificados por tema em uma estrutura navegacional e há possibilidade de busca por palavras-chave. Também há uma estrutura navegacional em que os OAs podem ser consultados por autor, título e data de submissão, estrutura da universidade e curso. Este RDE implementa recursos da Web 2.0 permitindo, por exemplo, a indexação realizada pelos próprios usuários e a sugestão de novos temas.

Banco Internacional de Objetos Educacionais (BIOE)¹⁴: o maior RDE brasileiro, é uma iniciativa da Secretaria de Educação a Distância (SEED) do Ministério da Educação (MEC) em conjunto com o Ministério da Ciência e Tecnologia (MCT), além de várias universidades parceiras. Criado em 2008,

¹¹ *Link*: <<http://www.rlo-cetl.ac.uk/whatwedo/rlos/completedrlos.php>>.

¹² *Link*: <<https://dspace.ndlr.ie/jspui/>>.

¹³ *Link*: <<http://www.edshare.soton.ac.uk/>>.

¹⁴ *Link*: <<http://objetoseducacionais2.mec.gov.br/>>.

disponibiliza documentos em diversos formatos sobre temas da educação infantil, ensino fundamental e médio, educação profissional e superior e modalidades de ensino brasileiras. Assim, os objetos estão classificados por tema em uma estrutura navegacional e há possibilidade de busca por palavras-chave e busca avançada. Também há uma estrutura navegacional em que os OAs podem ser consultados por níveis de ensino, tipos de recursos, título dos objetos, título das coleções de objetos, autor, assunto, tema e data de submissão. Ademais, é possível realizar buscas filtradas por país, idioma, tipo de recurso, nível de ensino e tema. A avaliação dos OAs que integram o BIOE é composta de duas etapas, uma avaliação realizada pelas universidades parceiras que fazem a submissão e uma avaliação de pares, ou especialistas no assunto (BIOE, 2010).

Esses são exemplos localizados a partir de consultas no próprio Merlot e no cadastro do Registry of Open Access Repositories (ROAR)¹⁵. Existem vários outros RDEs, pagos, abertos ou restritos a comunidades específicas. É importante acrescentar que os objetos encontrados nesses repositórios nem sempre atendem as características ideais para um AO, com conteúdo, atividade, avaliação e objetivo de aprendizagem definido; e não adotam amplamente, diretrizes e padrões para torná-los interoperáveis. Como esta pesquisa é delimitada pelos objetos de aprendizagem disponíveis nos RDEs na área de saúde, as peculiaridades desses repositórios serão apresentadas na subseção seguinte.

5.1.3.1 Repositórios digitais educacionais na área de saúde

Especificamente na área de saúde, a conscientização quanto à necessidade de disseminação e acesso a informação tem ganhando espaço (CARVALHO; GOMES, 2011). No âmbito nacional observa-se:

Hoje, no Sistema Único de Saúde (SUS), uma série de iniciativas voltadas à Educação na Saúde, colocam na agenda dos centros de formação e das instituições acadêmicas questões como a reorientação do ensino, a reorganização curricular, a revisão de modalidades de oferta de cursos, de práticas pedagógicas e de conteúdos (CAVALCANTE; VASCONCELLOS, 2006).

¹⁵ *Link:* <<http://roar.eprints.org/index.php>>.

A comunicação, educação e a informação são componentes essenciais para se alcançar equidade, qualidade e humanização dos serviços de saúde e o controle social no SUS. Em consonância com a crescente ênfase dada a este tema, a Fundação Oswaldo Cruz (Fiocruz)¹⁶, por exemplo, tem ampliado a sua oferta de produtos e serviços. Bibliotecas e repositórios digitais desenvolvidos ou dos quais essa instituição faz parte, conjugam informação técnico-científico para a comunicação entre pares e constituem segmentos destinados às atividades didáticas (FIOCRUZ, 2002). Outros exemplos são as iniciativas do MS como a Política de Educação Permanente, a Formação de Formadores e de Formuladores de Políticas e a Rede de Ensino para a Gestão Estratégica do SUS.

Assim, os repositórios educacionais na área de saúde têm sido utilizados para facilitar o acesso à informação específica. Da mesma forma que outros repositórios, esses buscam diretrizes e padrões para disseminar conteúdos amplamente. No Brasil a Rede Regional de Recursos Educacionais Abertos (REA Rede/OER), um projeto da Organização Pan-Americana da Saúde (Opas) sob a liderança do Campus Virtual de Saúde Pública (CVSP) e da Biblioteca Virtual de Saúde (BVS), pretende trocar experiências sobre a integração de repositórios das instituições pertencentes ao CVSP. O projeto busca a interoperabilidade entre o CVSP e a BVS e promove a publicação, busca e recuperação de OAs na região das Américas. O Acervo de Recursos Educacionais em Saúde (ARES), da Universidade Aberta do Sistema Único de Saúde (UNA-SUS), é um repositório de destaque em conformidade com a proposta de integração da REA Rede/OER (ENSP, 2012).

Internacionalmente, repositórios disponibilizam conteúdos na área de saúde, tanto para educadores como para alunos. O Health Education Digital Assets (HEAL) é mantido de maneira colaborativa e seu controle de qualidade é assegurado por um processo de revisão por pares (EDUCATIONAL TECHNOLOGY FOR HEALTHCARE EDUCATION, 2011). Cavalcante e Vasconcellos (2006) destacam o Consórcio MedBiquitous fundado pela Johns Hopkins Medicine e por sociedades médicas, incluindo organizações internacionais e universidades que, baseado nas recomendações do IEEE LTSC, desenvolveu o Healthcare Learning Object Metadata para descrever atividades e conteúdos educacionais em saúde.

¹⁶ Instituição para promover a saúde e o desenvolvimento social, gerar e difundir conhecimento científico e tecnológico, vinculada ao Ministério da Saúde (MS).

Muitos desses repositórios são compostos por documentos, como vídeos, imagens, animações, etc. que podem ser utilizados no contexto educacional ou mesmo incluídos como conteúdo e/ou atividade de um objeto de aprendizagem. Neste momento, faz-se necessária uma explanação sobre os documentos disponibilizados por esses repositórios, uma abordagem sobre OA em si.

5.1.4 Objeto de Aprendizagem disponibilizado em Repositórios digitais educacionais

Na literatura, observa-se que o OA também é chamado de objeto educacional, material instrucional (MOULIN; PERDEIRA 1999), material de aprendizagem online (MERLOT, 2011), entre outros. Sem fazer menções às influências pedagógicas, o termo objeto de aprendizagem, para esta pesquisa, generaliza os documentos acessados em RDEs que possuem a premissa de oferecer conteúdo e funcionalidades focadas na educação. Um AO reúne um conjunto de itens de informação para o cumprimento de um objetivo de aprendizagem específico. Esses itens são ordenados, ligados uns aos outros, para que o aprendiz possa seguir as etapas definidas na estratégia de aprendizagem e atingir o objetivo proposto (HORN, 1993). A composição de um OA com conteúdo, atividade e avaliação, para cumprimento de objetivos de aprendizagem, como mencionado anteriormente, também é reforçada pelo modelo conceitual de objetos de informação reutilizáveis – Reusable Information Object (RIO) – da Cisco Systems (CISCO SYSTEMS, 2003). Os OAs podem ser entendidos como entidades digitais, distribuídas pela Web para que diversas pessoas possam acessar e utilizar (WILEY, 2000). A figura a seguir ilustra sua composição ou “Anatomy”, como sugere Tanya Heins e Frances Himes (2002):

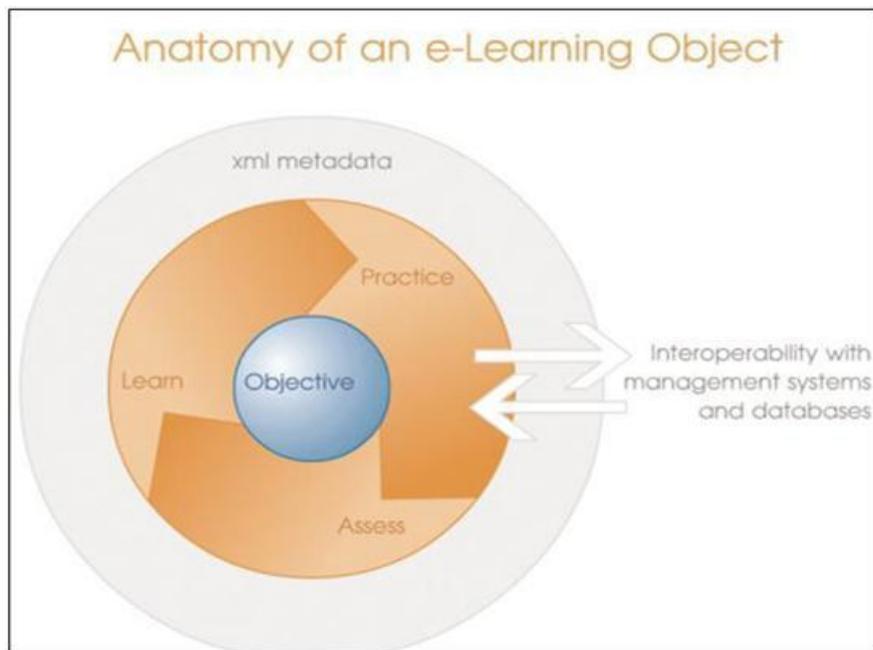


Figura 1: Anatomy of an eLearning Object.

Fonte: (HEINS; HIMES, 2002, p. 2).

A figura ilustra como um OA é formado. Seus autores afirmam que objetos de aprendizagem reutilizáveis podem ser usados sozinhos, para o rápido aprendizado, ou “*just-in-time learning*”, ou podem ser inseridos em estratégias mais complexas, que demandam mais tempo para o cumprimento de objetivos amplos. O OA é composto de três elementos fundamentais: conteúdo (*learn*), atividade (*practice*) e avaliação (*assess*). Do ponto de vista instrucional, um conteúdo pode conter uma oportunidade para a prática, simulação ou interação colaborativa e, ainda, uma avaliação a cerca do seu objetivo de aprendizagem. Os metadados descrevem os atributos do OA e permitem que este seja buscado e compreendido. A interoperabilidade permite a comunicação entre o OA e os sistemas de gestão de aprendizado, bases de dados e aplicações Web.

O OA pode ser usado, reusado ou referenciado durante o ensino-aprendizagem, com auxílio da tecnologia (IEEE; LTSC, 2008), e passa a ter maior visibilidade justamente a partir do interesse na reutilização, momento em que diretrizes e padrões, como aqueles citados outrora, são adotados. Recomenda-se que sejam projetados para serem pequenos, flexíveis, portáteis e adaptáveis a diversas aplicações Web. Os OAs podem preencher uma variedade de situações de ensino e estilo de aprendizagem, visto que, oferecem uma forma eficiente de adaptar atividades para diferentes tipos de alunos (MIRANDA, 2004).

Algumas características recomendáveis aos OAs, segundo o especialista em e- *learning* Warren Longmire (2000), são:

- Flexibilidade para reutilização sem a necessidade de excessivo esforço tecnológico;
- Possibilidade de customização e adequação ao contexto de aprendizagem, o que permite a readequação da sua forma de apresentação e estrutura;
- Facilidade para atualização, uma vez que apresentam características descritas em metadados e são reutilizados;
- Modularidade na sua concepção com respeito às características das partes que compõe um todo, permitindo o reuso sem causar entendimentos fragmentados ao usuário;
- Interoperabilidade devido a sua adequação aos padrões para criar conteúdos, estruturar e reunir objetos, permitindo sua utilização em diferentes plataformas de *software*, *browsers*, *hardwares*, etc.;
- Facilidade de busca e recuperação viabilizada pela sua indexação com o uso de metadados padrões e regras de descrição dos RDEs que os disponibilizam.

No que se refere a EAD, o OA é um elemento que compõe uma unidade de instrução, módulo de curso ou curso completo (GOMEZ, 2005), denominações aqui apresentadas para ressaltar a estrutura educacional (EDTech, 2005 apud KRATZ; PINTO; BARBOSA, 2007). Assim, Allert, Richter e NejdI (2002) e Longmire (2000) destacam alguns dos seus valores pedagógicos, tais como: atenção às teorias de aprendizagem; estabelecimento de paradigmas educacionais, como construtivismo e cognitivismo; cumprimento de princípios instrucionais; construção de estruturas conceituais; e aprimoramento de estratégias educacionais.

Segundo o SCORM (SCORM, 2011) deve ser seguido um Modelo de Conteúdos, baseado na forma de agregação dos itens de informação ou conteúdos que compõem um AO. Este Modelo de Conteúdos estabelece níveis de agregação que definem granularidades – de 1 (um) a 4 (quatro) – que aumentam conforme a complexidade das partes agregadas. Os RDEs podem disseminar conteúdos com diferentes granularidades, contudo, aqueles com menor granularidade, denominados *Assets*, são mais comumente disponibilizados. Os *Assets* não são compreendidos como objetos de aprendizagem, de acordo com sua definição literal.

- *Assets*: conteúdo de aprendizagem em formato básico, tal como,

texto, imagem, áudio, vídeo, animação, simulação, etc. Granularidade 1 (um);

- *Sharable Content Object* (SCO): conjunto de conteúdos compartilháveis formado por *Assets*, com exclusiva especificação de regras de comunicação com o AVA. Granularidade 2 (dois) ou 3 (três);
- *Content Package*: pacote de conteúdo formado a partir do processo de agregação (*Content Aggregation*) de SCOs e *Assets*, na composição de uma unidade de instrução, módulo de curso ou curso completo. Granularidade 3 (três) ou 4 (quatro) (SCORM, 2011).

Os níveis de agregação estabelecem relação entre os OAs, criando uma hierarquia de objetos complexos aos mais simples. Pesquisadores acreditam que os OAs são um núcleo de conceito para a criação, manutenção e gestão da aprendizagem. Desse modo, independentemente do nível de agregação, esses têm a pretensão de condensar informação necessária para um indivíduo atingir um objetivo de aprendizagem. Esse objetivo de aprendizagem está relacionado a um público-alvo e um contexto educacional, considerados na elaboração de uma estratégia educacional significativa (WAGNER, 2002).

A figura a seguir, baseada em Wayne Hodgins apud Wagner (2002), mostra como um OA pode ser dividido em partes menores e mais simples, capazes de serem construídas e reconstruídas para atender necessidades específicas. Quanto mais complexo for um objeto, melhor a sua contextualização e aderência à proposta para o qual foi desenvolvido, porém, menor a sua reutilização em outras propostas.

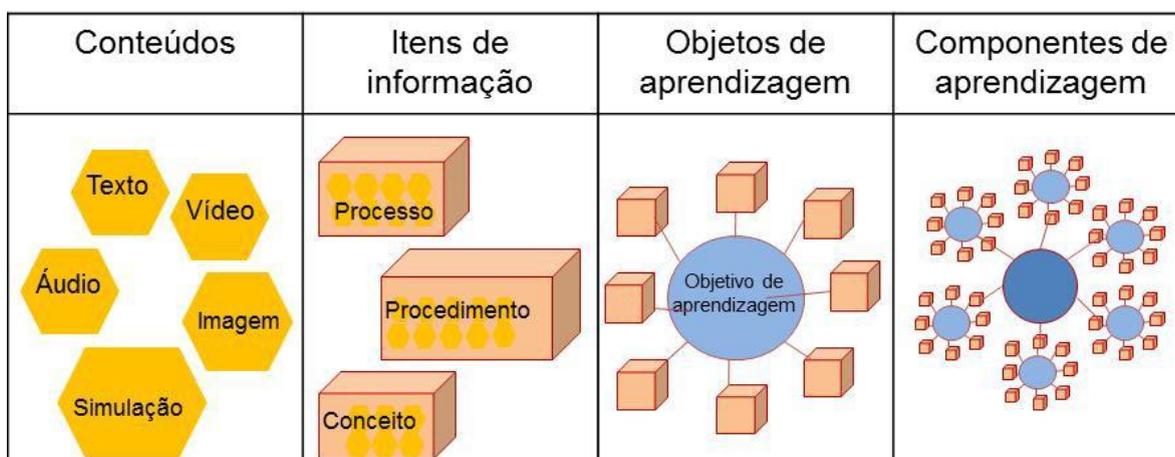


Figura 2: Modelo de conteúdo molecular.

Fonte: (WAGNER, 2002).

Conforme o modelo apresentado na figura (Figura 2), os conteúdos podem ser apresentados em diversos formatos. Estes serão utilizados na composição de itens de informação a serem agregados na composição de OA com objetivo de aprendizagem definido, capaz de ser tomado isoladamente para “*Compreender um conceito*” ou “*Diferenciar um procedimento certo de um procedimento errado*”, por exemplo. Um conjunto de OAs podem formar um OA ainda mais complexo, compondo um curso com amplo objetivo, como por exemplo “*Diagnosticar uma doença*”. A seguir discorre-se sobre o objetivo de aprendizagem de um OA.

5.1.5 Objeto de Aprendizagem e objetivo de aprendizagem

O OA deve apresentar um objetivo de aprendizagem para ser definido como tal, afinal, um conjunto de vídeos e imagens é simplesmente um conjunto de vídeos e imagens se não possui um objetivo de aprendizagem para ser utilizado no ensino. Como já mencionado, o OA pode ter conteúdos diversos e é possível ter abordagens e pontos de vista distintos sobre um mesmo objetivo de aprendizagem. O objetivo de aprendizagem é aquilo [resultado] que o professor [autor] espera do aluno [usuário] ao final de um processo de aprendizagem (DRISCOLL, 1998).

Segundo Cisco Systems (2003), o OA deve ser baseado em um único objetivo de aprendizagem e, a partir da sua estrutura granular, pode proporcionar experiência de aprendizagem mais ou menos ampla, por exemplo: o OA módulo de um curso tem como objetivo de aprendizagem “*Compreender os princípios de Mecânica dos fluidos*” e o OA, parte integrante desse primeiro, tem como objetivo de aprendizagem “*Identificar a diferença entre a equação de Bernoulli e a equação de Navier-Stoke*¹⁷”. Essa experiência permite a estimulação do pensamento (SRDEB-SCRDE, 2007 apud SALVE, 2010.). Assim, o objetivo de aprendizagem é o centro do OA, composto por conteúdos estruturados (WAGNER, 2002).

O objetivo de aprendizagem é uma informação descritiva, registrada nos metadados, que indica sua utilização e permite sua reutilização em contextos distintos, mediante adequações (WILEY, 2005). O LOM, entre outros padrões de metadados, estabelece critérios para permitir a descrição do objetivo de

¹⁷ Exemplo de duas equações utilizadas na análise da Mecânica dos fluidos.

aprendizagem que são úteis à recuperação dos OAs em RDEs. Esse assunto será abordado especificamente na Seção 5.3 Metadados, p. 61.

De acordo com Relvão (2006) o conceito de objetivo de aprendizagem é reconhecido no *design* instrucional e deriva dos trabalhos de Robert Frank Mager (MAGER, 1997), Robert Gagne (GAGNE, 1985), Walter Dick e Lou Carey (DICK; CAREY, 1996), entre outros, citados por Robson (2004). Para esses autores o objetivo de aprendizagem é passível e verificação e avaliação ao passo que são apresentados conteúdos e atividades ao aprendiz (aluno) e é definido um resultado esperado. O objetivo pode abranger o domínio psicomotor, afetivo ou cognitivo, seja no nível do conhecimento, da compreensão, síntese ou avaliação, segundo a taxonomia de Bloom (BLOOM, 1956).

O objetivo de aprendizagem tem relação com o resultado de uma busca de OAs, como pode ser observado no ciclo de vida do objeto (ANTA 2003 apud SALVE, 2010). O ciclo de vida (Figura 3) vincula o resultado de busca ao interesse de atingir um objetivo de aprendizagem. Indo além, o autor demonstra que um resultado não satisfatório para essa busca indica a necessidade de uma nova produção de OA. A busca é motivada pelo cumprimento da proposição inicial, contudo, as bases conceituais e o próprio objetivo de aprendizagem representam a necessidade latente de informação, que deve ser estabelecida antes da formulação da estratégia de busca no RDE.

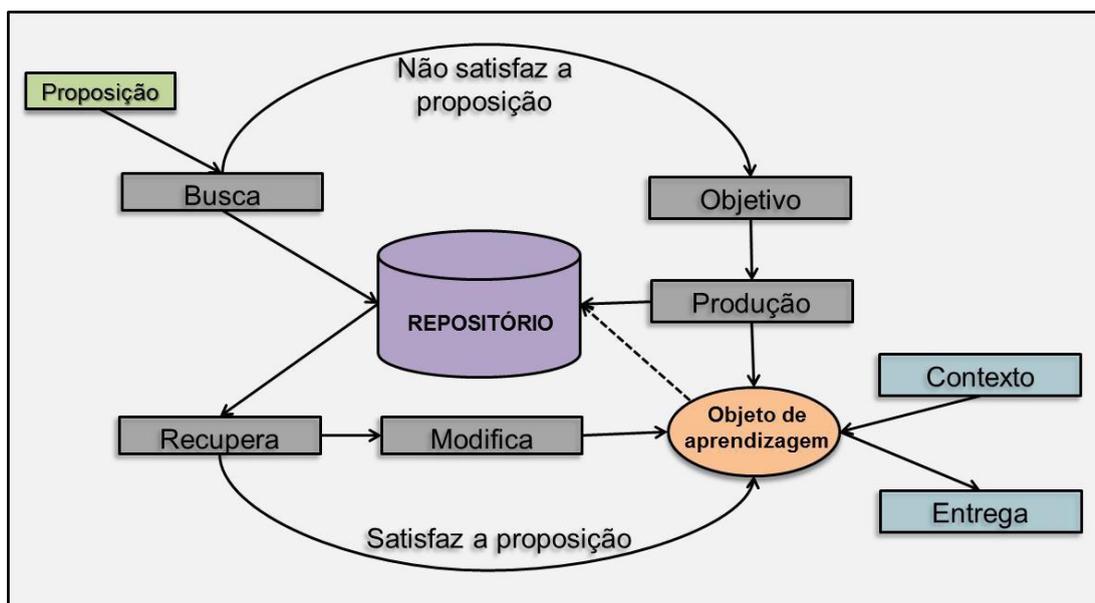


Figura 3: Ciclo de vida do objeto de aprendizagem

Fonte: adaptado de (ANTA 2003 apud SALVE, 2010).

A busca por palavra-chave faz referência ao assunto do OA e, inclusive, este é o principal parâmetro de classificação encontrado nos repositórios pesquisados (Seção 5.1.3 Exemplos de Repositórios Digitais Educacionais, p. 24). Contudo, para sua avaliação e uso efetivo, a descrição de um OA requer, necessariamente, a definição de assunto, geralmente baseado na área de conhecimento ou conceito abordado; e objetivo de aprendizagem (WAGNER, 2002). A dissociação dessas partes na formulação da estratégia de busca afeta a precisão dos OAs recuperados. Logo, a busca de OAs por palavra-chave carece de elementos que melhorem a precisão dos resultados apresentados em RDEs.

Diante do interesse do usuário de um RDE, o objetivo de aprendizagem é parte fundamental do processo de seleção e utilização do OA, como mencionado. Entretanto, o objetivo de aprendizagem relaciona-se com o público-alvo e o contexto educacional. Tudo isso evidencia o interesse da pesquisa em desenvolver uma modelagem baseada em Web Semântica, para repositórios digitais educacionais na área de saúde, com foco no objetivo de aprendizagem.

A explanação sobre repositórios digitais educacionais situa o contexto deste trabalho, ressalta o interesse na representação descritiva dos objetos, especialmente na afirmação de que o assunto é um aspecto importante na busca de OA, contudo, o objetivo de aprendizagem é um aspecto determinante na seleção de objetos relevantes em um RDE. A hipótese da pesquisa é de que uma modelagem baseada em WS para RDEs na área de saúde pode contribuir com o refinamento dos resultados de busca do usuário, considerando o contexto de ensino-aprendizagem. Assim, a seção seguinte irá abordar WS e as suas contribuições para esta pesquisa.

5.2 Web Semântica

Para manter a coesão e a coerência da abordagem sobre WS apresentada nessa sessão, é importante esclarecer alguns termos e conceitos usados, uma vez que podem ser adotadas acepções diferentes pelos autores. Isso irá facilitar a compreensão do texto e a visualização da relação existente entre os conceitos adotados para discorrer sobre o assunto. Assim, os termos a seguir serão utilizados ao longo dessa pesquisa de acordo com os conceitos apresentados.

Embora “informação” seja um termo amplamente utilizado para generalizações, como qualquer mensagem que se deseja comunicar, e para especializações, como um arquivo com registro dessa mensagem, o termo “**informação**” pode ser compreendido como conhecimento registrado em um suporte para sua comunicação (LE COADIC, 2004). A informação comporta um elemento de sentido, é um significado transmitido a um ser consciente por meio de uma mensagem inscrita em suporte material-temporal (impresso, sinal elétrico, onda sonora, etc). “**Documento**” é o registro da informação em um determinado formato, como papel, página na Web ou um OA dentro de um RDE, ou seja, o documento é o suporte da informação (SVENONIUS, 2001). Ao longo do texto, quando se fala em informação, geralmente esta se falando da informação em formato digital, registrada em documentos acessíveis pela Web.

O “**dado**” está na composição da informação e muitas vezes pode ser utilizado para comunicar algo a cerca da informação, como seus atributos. Se o “dado” é referido como parte da informação ele é um conjunto de caracteres com sentido semântico dentro do contexto em que está inserido. Esse contexto é definido pela estrutura e conteúdo do documento. O “**conteúdo**” é aquilo que se deseja comunicar, expresso pela informação em um ou mais documentos. O termo “dado” quando faz referencia a comunicação de algo a cerca da informação, indica que ele não compõe o conteúdo, mas pode ser integrado a este. Em outras palavras, o dado acrescenta algo sobre a informação. O “**metadado**” é o termo utilizado para descrever o dado que descreve outro dado que compõe uma informação. Dessa forma, o ato de comunicar algo a cerca da informação com o uso de dados específicos, ou metadados, é chamado “**descrição da informação**” ou simplesmente descrição.

De maneira geral, o “**conceito**” é formado pelo termo que o representa acrescido da sua definição. Um conceito pode, por exemplo, ser parte do conteúdo de um documento ou pode ser representado por um metadado quando se adota um padrão para melhor comunicar algo a cerca da informação. O padrão de metadado é exatamente o que traz sua definição, fazendo com que haja consenso entre o termo que representa o metadado e a sua definição para fins de utilização na descrição da informação. A “**relação entre dados**” permite a definição de um conceito e a “**relação entre conceitos**” permite representações semânticas mais apuradas como generalização, especialização e a associação. O estabelecimento de uma relação se dá com a “**interligação**” entre as partes e essas devem possuir significados individuais claros para compreensão da relação.

Uma vez definidos alguns dos importantes termos usados para falar de WS, pode-se relembrar como começou o processo de disponibilização de informação digital com a Web para contextualizar esse assunto. Desde seu surgimento a World Wide Web, ou simplesmente Web, passa por constantes evoluções. Em princípio, permitia apenas a troca de dados entre computadores com páginas HTML estáticas. Em um segundo momento, priorizou a apresentação da informação em páginas dinâmicas com conteúdos extraídos de bases de dados. Atualmente a Web caminha para estruturação dos seus conteúdos, descrição e atribuição de semântica aos seus dados, a chamada WS. Com isso pretende-se minimizar as limitações dos motores de busca e apoiar o usuário na realização de tarefas mais sofisticadas, sem demandar tanto tempo na busca e interpretação do grande volume de informação que se tem disponível (GOÑI *et al.*, 2002). A WS tem o objetivo de identificar e interpretar os conteúdos disponibilizados na Web de maneira automática, a partir da implementação de padrões (TIM BERNERS-LEE, 2000).

Segundo Decker *et al.* (2000) e Berners-Lee *et al.* (2002), ainda nos primórdios da WS, este era apenas o nome de um projeto capitaneado pelo W3C, que pretendia embutir alguma inteligência e contexto às linguagens utilizadas para o desenvolvimento de páginas Web, para melhorar a forma com que aplicações de computador interagem com estas páginas. O projeto, em andamento, passou a criar e implantar padrões para compartilhar, armazenar e descrever dados para serem utilizados por humanos e máquinas, de maneira automática e não ambígua.

O amplo acesso e utilização da Web pela sociedade em geral, faz com que essa ganhe dimensão e seja construída por qualquer indivíduo ou grupo de indivíduos que tenham interesse. As pessoas envolvidas na produção de conteúdos contribuem para que qualquer pergunta possa encontrar uma resposta na Web, basta ter um pouco de paciência e habilidade. Contudo, esse conjunto imenso de informação disponível tem pouco significado expresso e o computador é incapaz de interpretá-lo minimizando o esforço do usuário na seleção do que é relevante. A WS pretende resolver este problema e sua ideia pode resumir-se à questão: *Como fazer com que os computadores compreendam o conteúdo da Web?* (W3C.ORG, 2012). Para tanto, o primeiro passo é organizar e estruturar a informação e o segundo passo é adicionar semântica (GOÑI *et al.*, 2002).

O conteúdo acessível na Web é composto de textos, imagens, sons, etc. Que são escritos em linguagens próprias, lidas, interpretadas e/ou compiladas para serem exibidas em *browsers*. Este conteúdo pode ser estático ou dinâmico, de acordo com a apresentação de dados armazenados em bases de dados específicas responsáveis pela alteração do que é exibido conforme a solicitação de um usuário. Com a iniciativa da WS, esse conteúdo passa a agregar dados, também em linguagem própria, sobre o que é e qual o significado daquilo que será exibido ao usuário. Para compreender melhor, é possível analisar um RED: o repositório tem uma aplicação para busca que, ao ser utilizada por um usuário qualquer, irá retornar uma resposta de acordo com o interesse e aquilo que se tem disponível. A aplicação computacional é construída em uma linguagem de programação e irá se comportar de acordo com os requisitos definidos, por exemplo, verificando se um termo de busca do usuário está presente nos metadados descritivos de um volume de documentos presentes em sua base.

Uma infraestrutura adequada é capaz de fazer com que a informação disponível na Web, dentro e fora de um RDE, tenha relações semânticas, facilitando sua contextualização e conseqüente utilização. Serviços disponíveis em *sites* diferentes podem ser associados, grandes bases de dados podem ser compartilhadas, a atualização de informação dispersa sobre uma mesma temática pode ser facilitada. O comportamento inteligente da Web depende de aplicações computacionais inteligentes e integradas (ALLEMANG; HENDLER, 2011).

Hoje a infraestrutura da Web baseia-se essencialmente em *links* entre documentos, com significado e contexto dependente da interpretação humana.

Com a proposta de aprimorar essa infraestrutura, a WS passa a explicitar a relação estabelecida por um *link*, permitindo que os computadores sejam capazes de estabelecer seu contexto e inferir seu significado. Segundo a visão da WS, um usuário navegando em um diretório ou realizando uma busca em um *site*, irá indicar seu assunto de interesse e ser auxiliado por agentes inteligentes. Os agentes percorrerão páginas Web anotadas e repositórios com metadados para recuperar documentos sobre o assunto em questão, com contexto e significado pertinente, em vez de simplesmente apresentar todos os documentos que contenham as palavras-chave indicadas. Ou seja, já não será a sintaxe a determinar o resultado do processo de recuperação de informação, mas sim a semântica (GONÇALVES, 2007).

A partir da iniciativa da W3C, a infraestrutura da WS passa a formar uma rede de dados interconectados com a implementação de um conjunto de tecnologias, como RDF, OWL, entre outras, que fornecem um ambiente onde uma aplicação Web pode consultar os dados, fazer inferências, usar vocabulários de conceitos, etc. No entanto, para essa evolução da Web atual é importante que os dados estejam disponíveis em um formato padrão, acessível e gerenciável pelas ferramentas da WS. Ademais, a WS precisa de acesso não apenas aos dados, mas às relações entre eles. Esta coleção de dados e suas relações tem sido referida como *Linked Data* (ALLEMANG; HENDLER, 2011). Para alcançar e criar a *Linked Data*, o uso de tecnologias deve permitir fazer conversões dinâmicas com o acesso às bases de dados existentes, que podem ser bases relacionais, XML, HTML, etc. Também é importante a configuração de terminais de consulta para acessar os dados da forma adequada (BERNERS-LEE, 2009) e (W3C.ORG, 2012).

A figura a seguir (Figura 4) ilustra como atender algumas regras para interligar dados na Web garante que esses dados não tenham sua reutilização limitada. A reutilização dos dados de maneiras não previstas é justamente o que agrega valor a WS (BERNERS-LEE, 2006):

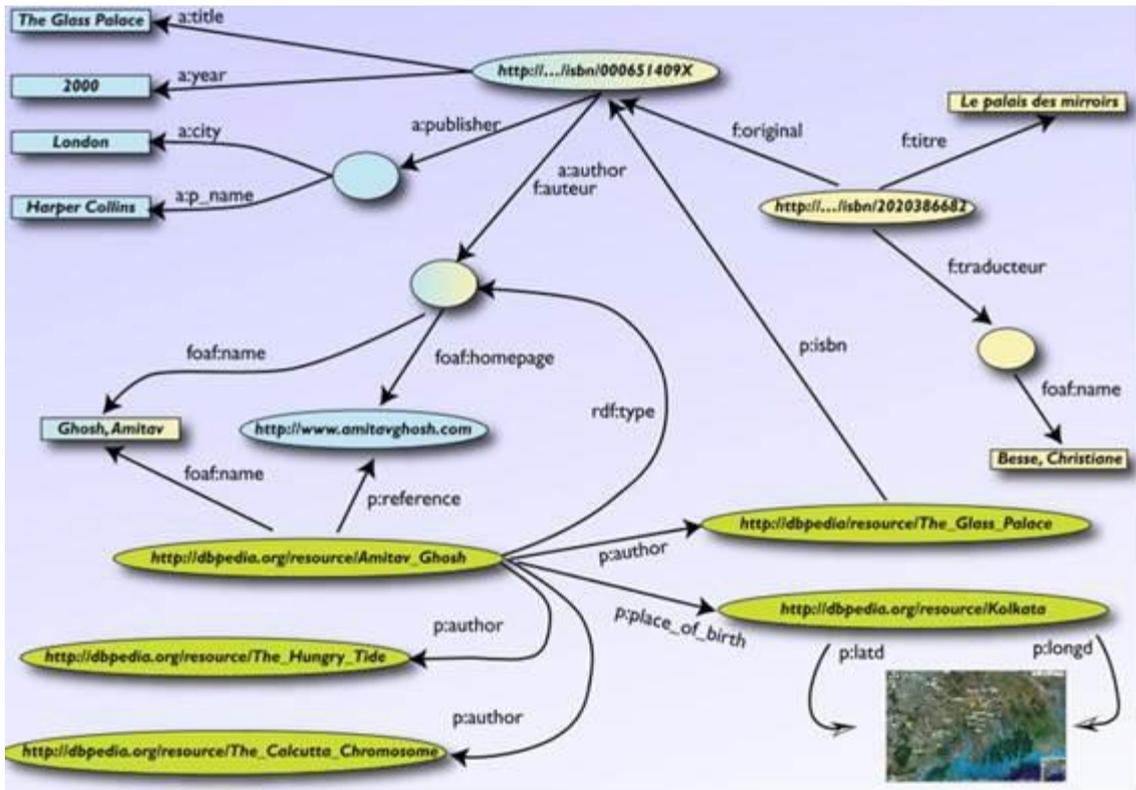


Figura 4: Reutilização de dados interligados

Fonte: (HERMAN, 2009).

A Figura 5 apresenta dados interligados e nomeados, ilustrando a evolução Web atual para WS: dados que descrevem um livro estão relacionados, destacados na cor azul, tais como, *title* (The Glass Palace), *year* (2000), *publisher city* (London) e *publisher name* (Harper Collins), *author name* (Ghosh, Amitav), etc. Um outro livro, a tradução do primeiro livro descrito, também possui seus dados descritos e relacionados, destacados na cor bege, tais como, *titre* (Les Palais des Mirroirs), *traducteur name* (Besse, Christiane), etc. Os dados são interligados a partir do entendimento que são “original” e a sua “tradução”, *linkando* a URL que indica o ISBN da tradução a URL que indica o ISBN do original. Isto permite realizar inferências e responder questões que antes necessitariam de interpretação humana, como por exemplo, se o livro “The Glass Palace” tem uma tradução. Ou seja, o computador só infere que “Les Palais des Mirroirs” é uma tradução a partir do momento que esses dados (*title*, *titre*, ISBN) são definidos e interligados com uma relação explícita e formal (Figura 6).

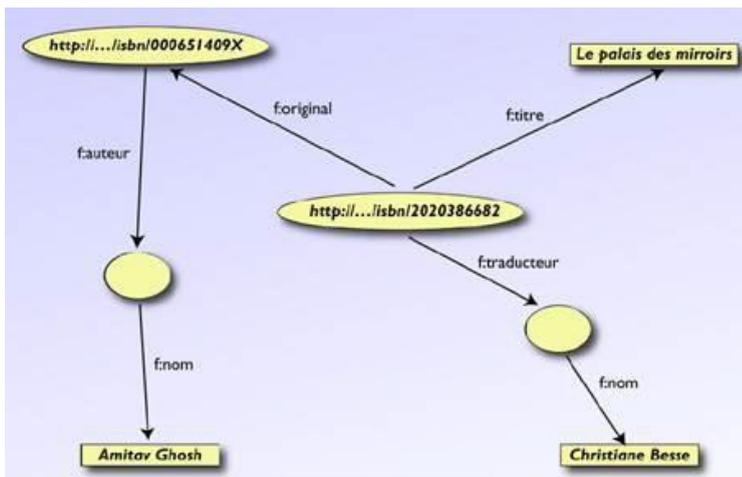


Figura 5: Evolução Web atual para WS.

Fonte: (HERMAN, 2009).

A WS também permite que informação adicional possa ser agregada: se alguns dados possuem as mesmas características, podem ser agrupados em uma mesma categoria, como no caso de “*author name* (Ghosh, Amitav)” e “*traducteur name* (Besse, Christiane)”, perfeitamente adequados a uma categoria “Pearson”, com definição consensual na comunidade que a está adotando: “entidade que possui nome próprio e homepage”. Com essas novas relações mais inferências podem ser feitas automaticamente. Pode-se inferir “qual a *homepage* do autor original do livro “Les Palais des Miroirs”.

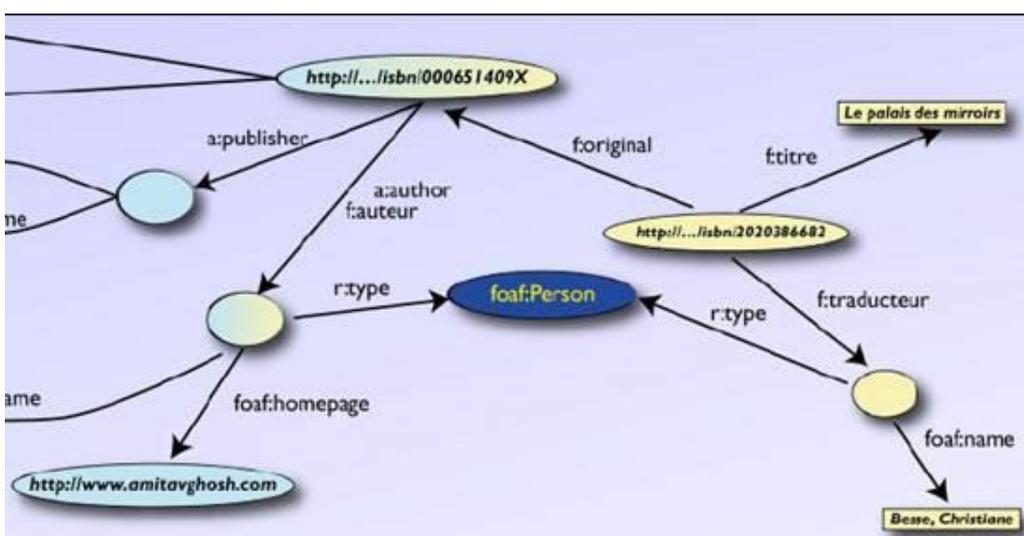


Figura 6: Dados definidos e interligados com uma relação explícita e formal.

Fonte: (HERMAN, 2009).

Usando, por exemplo, uma categoria "Pearson", o conjunto de dados que está sendo relacionado pode ser interligado com dados relacionados de outras fontes, como a Wikipedia. Os dados da Wikipedia podem ser interligados utilizando ferramentas específicas¹⁸. Assim, é possível apresentar, sem mais esforço no momento da busca, outras obras do autor original, seu local de nascimento, etc., como destacado na cor amarelo na parte inferior da Figura 6. Inferências relacionadas a idioma também podem ser facilitadas uma vez que fica claro que "*title*" e "*titre*" têm o mesmo significado, assim como "*author*" e "*auteur*". Tudo isso é feito por humanos com facilidade e um pouco de tempo, porém, para que o computador faça e poupe o tempo do usuário é necessário um pouco de rigor (HERMAN, 2009).

Devido a preocupação com a explicitação de significados e relações entre os dados, algumas vezes a WS é entendida como um esforço para que todos os usuários e desenvolvedores tenham a mesma opinião e conceito sobre o que está sendo disponibilizado. Isto não é verdade como afirma Allemang e Hendler (2011) em seu recente livro "Semantic web for the working ontologist: effective modeling in RDFS and OWL". Ao descrever os dados acessados na Web de maneira que os computadores compreendam, como mostrado anteriormente, tem-se uma Web mais inteligente, capaz de entender os dados como entidades que podem ser vistas sobre diferentes pontos de vista e em diferentes contextos. Essas descrições dos dados são abstrações ou modelos que ocultam alguns detalhes e destacam outros.

Os modelos auxiliam a interligar dados e estruturar o conteúdo de três maneiras essenciais (ALLEMANG; HENDLER, 2011):

- Na comunicação entre as pessoas: descreve um contexto de determinada maneira para que outras pessoas possam compreender;
- Ao explicar e fazer previsões: relaciona fenômenos simples a fenômenos mais complexos, fornecendo explicações e previsões sobre como estes podem ocorrer;
- Ao mediar múltiplos pontos de vista: representam pontos em comum a cerca de diferentes fenômenos, permitindo que duas pessoas que discordem completamente sobre eles explorem suas diferenças.

¹⁸ DBpedia.

A modelagem é o processo de organização da informação para comunidade utilização. No contexto da WS esse é um processo contínuo (ALLEMANG. HENDLER, 2011). Isso quer dizer que os padrões da WS permitem a representação do contexto, a previsão de interpretações e o compartilhamento de pontos de vista sobre uma informação disponibilizada na Web. Em outras palavras, modelos estabelecem uma comunicação a partir de um padrão capaz de causar consenso à interpretação; proporciona o contexto para o qual as inferências são validas; e assim, permite o compartilhamento de significados. As pessoas podem colaborar tanto com conteúdos como com os modelos que interligam seus dados e lhe fornecem estrutura.

A WS trará melhoria no índice de precisão das buscas realizadas hoje, como afirma Souza e Alvarenga (2004), pois os modelos criados e disponibilizados permitirão que sistemas de recuperação evitem a polissemia e sinonímia, entre outros problemas de origem semântica. Faz necessário compreender como que a WS, com a utilização de tecnologias específicas e modelos, pode ser utilizada em REDs, para isso sua arquitetura será abordada na seção seguinte.

5.2.1 Arquitetura da Web semântica

As tecnologias e especificações técnicas para a WS são definidas com o objetivo de estruturar, definir a sintaxe, atribuir semântica e permitir inferências lógicas. Desde a concepção da arquitetura da WS no ano de 2000 (W3C, 2000) essas tecnologias e especificações foram organizadas em camadas sobrepostas. Essas camadas são complementares e compatíveis, além de tornarem-se mais complexas na medida em que são aprofundadas.

Camadas inferiores são independentes das superiores facilitando sua evolução e utilização. Essa característica ainda permite a progressão da implementação da WS, ou seja, *sites* e aplicações da Web podem cumprir com os requisitos de uma camada baixa e ter resultados mais simples quanto à estruturação de dados. De acordo com as possibilidades e interesses outras camadas podem ser implementadas e novos resultados, cada vez mais elaborados, serão atingidos.

De maneira geral, as camadas provêm linguagens de modelagem que diferem em seu nível de expressividade e fornecem sintaxe para representar

metadados e os vocabulários para exprimir esses metadados, ou seja, as ontologias (HARMELEN, 2004). Essas linguagens têm sido testadas e seus recursos aprimorados ao longo dos anos. A versão atual das camadas da WS é ilustrada na Figura 7 com suas linguagens e demais elementos (BERNERS-LEE, 2006).

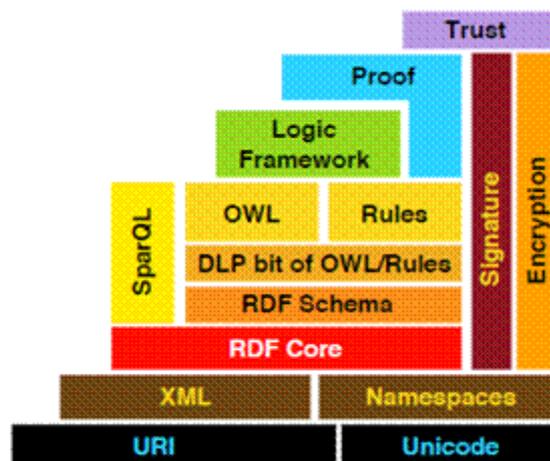


Figura 7: Arquitetura da WS versão 2005

A Figura 7 mostra como as camadas da WS estão sobrepostas e organizadas. Uma visão ampla permite destacar:

- Endereçamentos em URI são utilizados para nomear de forma única documentos, entidades, conceitos, relações, etc. O Unicode é padrão universal adotado para o endereçamento e garante sua legibilidade;
- O XML serve para definir a estrutura dos dados e junto com os Namespaces separa o dado em si (conteúdo) do nome daquele dado (rótulo). O XML Schema é o esquema a ser utilizado em um documento XML;
- O RDF descreve os recursos que integram os dados estruturados em XML. O RDF Core apresenta o modelo e a sintaxe para essa descrição e o RDF Schema é o esquema adotado para as definições. A partir da camada RDF Schema é possível realizar algumas inferências;
- A DPL bit of OWL/Rules é a camada inserida com o avanço do uso de ontologias na WS e seu objetivo é facilitar a comunicação entre RDF e OWL;
- A OWL e suas regras definidas na subcamada Rules compõe a linguagem recomendada para definição de ontologias para descrever

- semanticamente os dados;
- Logic framework é a camada que guarda modelos lógicos a cerca dos dados e sua semântica permitindo a realização de inferências;
 - Proof e Trust são responsáveis pelas verificações necessárias para comprovar e exibir uma inferência;
 - Signature e Encryption foram inseridas recentemente, de maneira perpendicular, com o objetivo de dar suporte a assinatura, ou identificação e criptografia, das descrições e definições que implicam diretamente na realização de inferências. Abarcam desde o RDF Core a camada Proof;
 - A camada SparQL é a definição da linguagem utilizada na pergunta, ou *query*, realizada durante as inferências viabilizadas pela WS. Por esse motivo a SparQL é perpendicular a RDF Schema, DPL bit of OWL/Rules e OWL/Rules.

Poucos meses após a publicação dessa arquitetura foram apresentadas críticas e sugestões de alterações para a mesma. Segundo Horrocks, *et al.* (2005), algumas das linguagens apresentadas nesta arquitetura não são semanticamente compatíveis de forma direta, como a DLP e a linguagem de desenvolvimento de ontologias, OWL, de modo que tais tecnologias deveriam ser apresentadas paralelamente ao invés de sobrepostas. Independentemente, essas camadas e seu conjunto de tecnologias, aqui apresentado com várias siglas, serão discutidos a seguir. Para facilitar a compreensão, as tecnologias e requisitos técnicos estão divididos em duas sessões.

A primeira seção (Seção 5.2.2 Estrutura e sintaxe da Web semântica, p. 43) irá abordar as camadas utilizadas para estruturar os dados e descreve-los de maneira não ambígua, de acordo com padrões universais e linguagens consolidadas, capazes de viabilizar seu intercambio e compartilhamento com aplicações Web distintas. A segunda seção (Seção 5.2.3 Lógica e semântica da Web semântica, p. 49), por sua vez, irá abordar como o significado desses dados pode ser formalizado gerando consenso na sua interpretação por aplicações e computadores. Não há preocupação em esgotar os assuntos, mas sim em abordar como cada camada pode ser implementada e destacar como isto se aplica no contexto dos RDEs.

Como afirma Gonçalves (2007), um repositório está estruturado, organizado com informações descritas em metadados. O que facilita a implementação da WS. Esses metadados devem ser preenchidos com conteúdos que seguem um mínimo de padronização uma vez que seu

significado é expresso. A semântica desses conteúdos deve ser formalizada com o uso de ontologias para permitir que computadores identifiquem os metadados e interpretem seus conteúdos. Esses dois assuntos, metadados e ontologias, serão abordados com maiores detalhes.

5.2.2 Estrutura e sintaxe da Web semântica

Os conteúdos da Web são escritos em linguagem própria para sua apresentação em *browsers*, como a HyperText Markup Language (HTML). Porém, com a premissa da WS, linguagens capazes de indicar como o conteúdo deve ser apresentado e ainda informar o que é esse conteúdo, tornam-se fundamentais. A estrutura e sintaxe da WS têm a preocupação justamente de estabelecer como os conteúdos da Web devem ser escritos para permitir seu acesso e interpretação consensual. Nessa seção serão apresentados padrões para nomear os dados da Web, como o URI; linguagem com sintaxe e estrutura adequada para escrever os conteúdos, o XML; e recomendações para explicitar as características dos dados, o RDF. A partir do uso dessas tecnologias, pode-se garantir a interoperabilidade estrutural e sintática da WS. Com isso, diferentes aplicações computacionais podem acessar a informação e identificar o seu significado.

5.2.2.1 URI e Unicode

Os dados na WS devem estar relacionados para que uma aplicação Web, ao encontrar parte do conteúdo, possa encontrar outras partes referenciadas. Essas relações são estabelecidas utilizando linguagens apropriadas para compor o documento. O que efetiva essa relação é a interligação entre dados nomeadamente identificados com o Uniform Resource Identifier (URI). Os URIs são identificadores únicos e requerem um padrão adequado para nomear os dados, as relações estabelecidas entre eles e os próprios conteúdos da Web. O padrão recomendado é o Unicode, que define a composição de um URI a partir de uma sequência de caracteres com interpretação padronizada por meio de codificação numérica universal.

O uso de URIs é o primeiro passo para a construção de uma rede de conteúdos, pois as suas especificações permitem interligar e nomear

genericamente tudo aquilo que for definido com as tecnologias acrescentadas pelas outras camadas da WS. Além da sua adoção, é importante deixar os URIs disponíveis para as aplicações computacionais e assim as relações podem incluir links (URIs) de outras fontes. Isso conseqüentemente irá potencializar a realização de inferências e a descoberta de conteúdos (BERNERS- LEE, 2009).

5.2.2.2 XML e Namespaces

Enquanto a HTML é uma linguagem para confecção de documentos para Web que tem preocupação com a apresentação do conteúdo, a eXtensible Markup Language (XML) preocupa-se com a marcação desse conteúdo. Essa marcação se dá com o uso de *tags* para declarar que tipo de dado compõe um conteúdo de forma separada do conteúdo em si. A HTML é uma linguagem pré-definida que não permite o acréscimo de *tags* na sintaxe do documento. Já a XML permite que sejam criadas *tags* para marcar o conteúdo de acordo com a necessidade, facilitando seu processamento por computadores. As duas linguagens derivam da Standard Generalized Markup Language (SGML).

As *tags* da XML delimitam o início e o fim do conteúdo ou de parte do conteúdo de um documento. Este conteúdo delimitado é tratado como um elemento de informação, considerado um objeto ou entidade a qual se pode atribuir características específicas (OLIVEIRA, 2002). Para facilitar o intercâmbio de informação na Web, podem ser adotados entidades e atributos padronizados, ou Namespaces (W3C, 2012).

Uma vez que o conjunto de *tags* pode ser estabelecido pelo usuário para definir o conteúdo, indicar seu significado e sua apresentação, devem ser utilizados os esquemas de XML, ou XML Schema para orientar a marcação de um conteúdo, sua estrutura e formatos válidos (W3C, 2006). Assim, a estrutura XML é formada basicamente por três arquivos principais: o XML Schema, com a declaração de quais *tags* são usadas e como são usadas; o Cascading Style Sheets (CSS) ou eXtensible Style Language (XSL), que especificam como apresentar o conteúdo, indicando os elementos e quais estilos aplicar, além de associar as partes do documento; e o próprio documento com seu conteúdo em XML (OLIVEIRA, 2002). Os arquivos XML são codificados em

Unicode e isso permite sua leitura, interpretação e exibição em diferentes plataformas Web (W3C.ORG, 2012).

Essa camada da WS permite a definição da sintaxe do conteúdo a ser disponibilizado na Web e especifica sua estrutura, porém essa estrutura não possui indicação semântica. A XML e o XML Schema viabilizam a interoperabilidade sintática e interoperabilidade estrutural (OLIVEIRA, 2002). Com o uso do XML é possível definir um texto na Web e seu conteúdo, sua forma de apresentação e os dados sobre esse conteúdo, como por exemplo, o seu “autor”. Não é possível inferir se esse “autor” é uma pessoa ou instituição, ou mesmo relacionar este texto a outros textos do mesmo “autor”. Para isso são necessárias outras tecnologias a serem discutidas.

5.2.2.3 RDF, RDF Core e RDF Schema

O Resource Description Framework (RDF) é uma tecnologia para representar as informações na Web, considerada como a base da WS. Para isso, possui uma sintaxe abstrata com grafos baseados em modelos de dados. A estrutura básica de qualquer expressão em RDF é um conjunto de triplas que compõem os grafos e são formadas por um sujeito, um predicado e um objeto. A figura a seguir ilustra as triplas e grafos RDF (Figura 8).

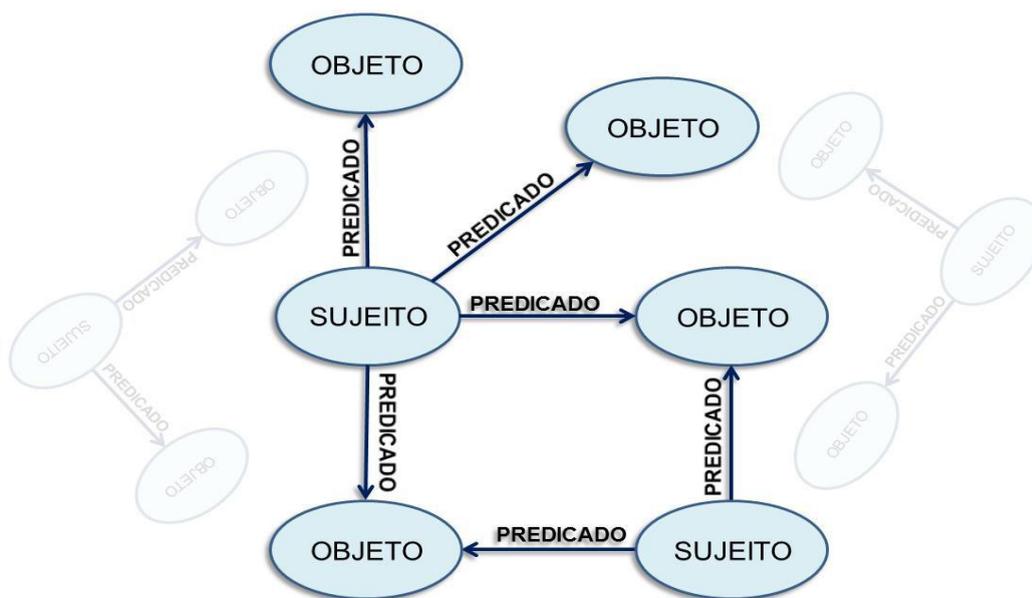


Figura 8: Triplas e grafos RDF.

O conteúdo de um documento definido e estruturado com o uso da XML pode ser representado a partir da composição de grafos RDF. Os grafos irão representar declarações simples sobre os dados que descrevem o conteúdo, estabelecendo a propriedade desse dado e atribuindo um valor para mesma. Cada tripla do grafo é uma relação entre seu sujeito e objeto, interligados pelo predicado. Um predicado, também chamado de propriedade, é o que define a relação.

Pode-se afirmar que o documento para Web definido em XML, com sintaxe que expressa os dados sobre seu conteúdo, como por exemplo, o seu “autor”, pode agregar uma semântica formalizada aos dados que descrevem seu conteúdo a partir da utilização do RDF. Isto porque, o dado que expressa o “autor” do documento passa a ter seus atributos formalizados em triplas, tais como: o *autor* (sujeito) *possui* (predicado) *CPF* (objeto); *CPF* (sujeito) *pertence a* (predicado) *pessoa* (objeto); e *pessoa* (sujeito) *possui* (predicado) *data de nascimento* (objeto). A construção dessas sentenças com o RDF irá ampliar a capacidade de realizar inferências (KLYNE; CARROLL, 2004), como observando de maneira simplificada na Figura 9.

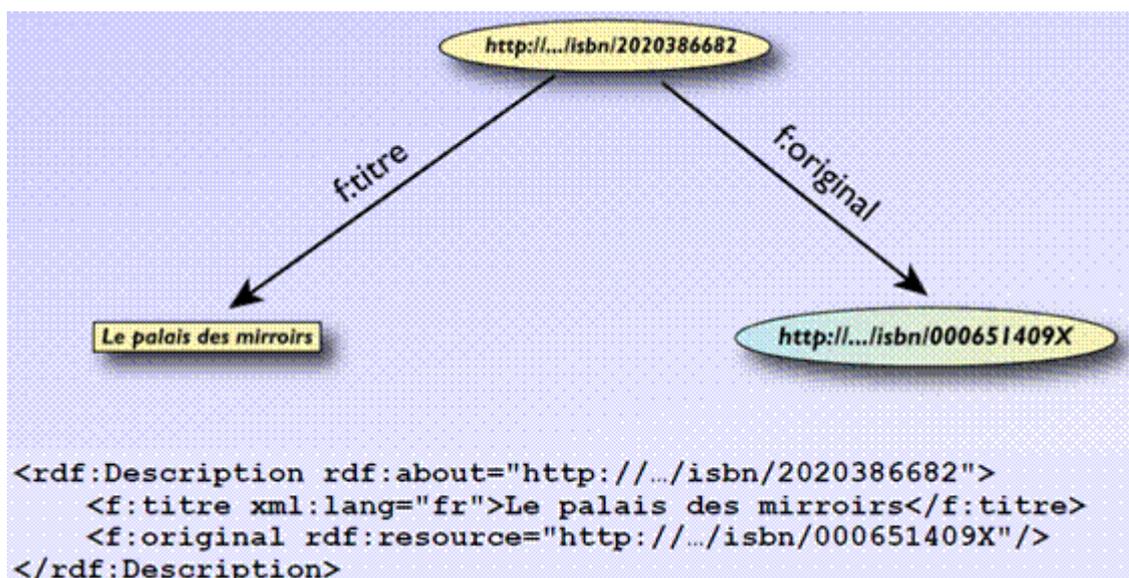


Figura 9: Construção de sentenças com o RDF.

Fonte: (HERMAN, 2009).

Figura 9: o autor do texto é delimitado por *tags* XML. As *tags* utilizadas e seus atributos são declarados no XML Schema que pode conter, por exemplo, a informação que o valor das *tags* `<author>` e `</author>` é um conjunto de caracteres alfabético. A sua forma de apresentação, por exemplo, em negrito, é definida pelo XSL. Já o RDF estabelece que o “autor” possui um

CPF. Uma vez que essas afirmações e atributos são nomeados por URIs e processados por uma aplicação Web, pode-se inferir automaticamente que *“se o autor do documento possui CPF também possui uma data de nascimento. Este autor é uma pessoa, não uma instituição”*.

Os grafos do RDF irão compor metadados, como no exemplo da Figura 8, a partir de dados e também unidades de informação delimitadas por *tags*. De maneira geral, ambos são chamados de recursos, tal como a interligação estabelecida entre quaisquer deles. Os recursos podem estar contidos em classes e, ao mesmo tempo, as classes também são recursos. Assim, as classes são coleções de recursos com propriedades e/ou valores semelhantes que denotam sua generalização. As interligações das triplas ocorrem entre as classes e recursos, e um recurso pode pertencer a mais de uma classe. Ora a classe pode ser tratada como um conjunto de recursos e ora como um recurso específico interligado a outro recurso ou pertencente a uma nova classe. Isso tudo permite concluir que no RDF não são tratados tipos de dados, mas sim as propriedades dos dados ou unidades de informação. Dessa forma, dá-se um passo a frente da sintaxe e estruturação dos documentos disponibilizados na Web. Este avanço está no campo da semântica, porém ainda não permite a conceituação e desambiguação dos termos.

Na arquitetura da WS o RDF foi dividido em RDF Core e RDF Schema (BECHHOFFER, 2004). Genericamente, o RDF Core (termo utilizado para distinguir o núcleo RDF do esquema RDF Schema) determina como descrever as propriedades e os valores dos recursos, enquanto que o RDF Schema define quais as classes, subclasses, propriedades e restrições podem ser utilizadas. Souza e Alvarenga (2004) afirmam que essa tecnologia estabelece na verdade um padrão de metadados e a indicação de seu uso para o processamento adequado. Esses metadados devem ser embutidos na XML, considerados como Namespaces definidos no RDF Schema. Metadados, dada sua relevância para esta pesquisa, serão abordado na Seção 5.3 Metadados, observadas as especificidades da descrição dos OAs disponibilizados em RDEs.

De acordo com Graham e Carroll (2004) o RDF foi concebido, entre outras coisas, vislumbrando a descrição dos conteúdos da Web com o uso de metadados capazes de estabelecer o consenso na interpretação de um documento e ainda combinar seus dados com dados de outros documentos, disponíveis em fontes distintas ou não. Atualmente, é uma tecnologia recomendada pela W3C com os seguintes objetivos:

- Oferecer um modelo de dados simples para facilitar suas aplicações computacionais e seu processamento;
- Permitir a formalização semântica e as inferências;
- Utilizar um padrão de nomeação para os recursos e suas interligações baseado em URI;
- Utilizar uma sintaxe baseada em XML para codificar o modelo de dados e facilitar a troca de informação entre as aplicações computacionais;
- Suportar o uso de XML Schema e seus tipos de dados;
- Permitir a representação de quaisquer declarações sobre quaisquer recursos.

Os objetivos da RDF ressaltam sua complementariedade com as outras camadas da WS, tais como URI e Unicode, XML e Namespaces. Recomenda-se que um recurso do grafo RDF seja identificado por um URI e que sua sintaxe utilize a XML e seu XML Schema. O RDF Query Language and Protocol (SparQL), camada que perpassa do RDF Schema a OWL, propicia linguagens e protocolos para consultar e manipular grafos RDF possibilitando diversas inferências. Para isso, o SparQL é capaz de limitar o número de resultados retornados e os classifica-los, facilitar a especificação de várias fontes (via URIs) dentro de uma mesma consulta, entre outras possibilidades (HERMAN, 2009). A fim de tornar os resultados das consultas reutilizáveis e interoperáveis, a SparQL atende a XML, entre outros formatos (SPARQL, 2012).

Conforme discutido anteriormente, as inferências possíveis com o uso da RDF advêm da relação entre as entidades que formam as triplas. O RDF Schema Plus é uma recomendação que amplia essas inferências ao classificar e caracterizar essas entidades. Entretanto, para conceituar uma entidade de maneira consensual e não ambígua, permitindo que as aplicações computacionais compreendam a semântica da informação disponibilizada na Web (ALLEMANG; HENDLER, 2011), são utilizadas ontologias. Esse assunto será abordado na próxima seção.

5.2.3 Lógica e semântica da Web semântica

A seção anterior abordou tecnologias e recomendações para nomear, compor a estrutura sintática e expressar o significado dos conteúdos da WS. O

RDF foi a recomendação apresentada capaz de agregar alguma semântica a esses conteúdos. Para compreender melhor a contribuição dessa tecnologia para a WS e, conseqüentemente, para o modelo a ser proposto por essa pesquisa, é importante estabelecer o seguinte raciocínio (DAN BRICKLEY, R.V. GUHA, 2004) (KLYNE; CARROLL, 2004):

- Um documento na Web é formado por conteúdo. Esse conteúdo pode ser descrito para facilitar a compreensão do seu significado (É um livro?, Possui autor?, Possui título?), a sua contextualização (É direcionado para especialistas? Trata do assunto didática em sala de aula?) e a sua localização por sistemas de busca entre outras aplicações;
- A descrição do conteúdo pode ser feita com metadados. Por exemplo, é possível descrever o autor de um documento usando o metadado *dc:author*. Esse metadado tem uma identificação única e universal definida pelo seu padrão (*dc:author* padrão *Dublin Core*) e um valor que corresponde ao nome do autor do referido conteúdo;
- A descrição do conteúdo também pode ser feita com a delimitação da informação contida no conteúdo, associada a um atributo para ganhar significado. Por exemplo, uma unidade de informação é delimitada pelas *tags* em um documento: “Eu, *<author>Fulano de Tal</author>* ofereço esse poema aos meus filhos”. Essa unidade de informação ganha significado quando é estabelecida a relação Fulano de Tal (sujeito) é (predicado) autor (objeto). Isso, na verdade, retoma o conceito de metadado;
- A sintaxe XML é utilizada para delimitar unidades de informação no conteúdo de um documento. Além disso, permite o uso de metadados na estrutura de um documento;
- O RDF estabelece a definição de triplas para relacionar sujeito, predicado e objeto;
- O RDF Schema formaliza o esquema de triplas utilizadas, agrupando as entidades relacionadas em classes e subclasses, definindo o domínio, as propriedades, os valores de propriedades apropriados e as instancias;
- O RDF Schema permite a composição e o manuseio de um padrão de metadados formado por um conjunto de entidades que criam uma rede de informação, assim, é possível interpretar o conteúdo de um documento, descrito com o uso dessas entidades.

O padrão URI, discutido na Seção 5.2.2.1 URI e Unicode, permite identificar cada entidade não apenas com palavras, mas com combinação de caracteres que é única e possui entendimento universal. Logo, se um recurso (sujeito), atributo (predicado) e o valor do atributo (objeto) representam uma conceituação, como por exemplo, “*aluno é pessoa*”, ao utilizar URIs, tem-se uma identificação única para o conceito. Entretanto, é fundamental a utilização de URIs diferentes para conceitos diferentes, e URIs iguais para conceitos iguais, o que se torna problemático com a quantidade de documentos disponibilizados na Web. Outro problema não resolvido com a XML e o RDF é a necessidade de estabelecer relações mais complexas para expressar um mesmo conceito como, “*aluno é sinônimo de estudante*” ou “*aluno não é professor*”. As ontologias poderão ser utilizadas na WS para resolver problemas como estes, de ordem conceitual e semântica. Seu uso permitirá a formalização dos diferentes conceitos de um termo, a composição de um vocabulário comum e o entendimento consensual sobre os termos desse vocabulário.

Uma vez que as ontologias tratam de conceitos que representam a semântica dos conteúdos na WS, a subseção seguinte discorre sobre modelagem conceitual e o estabelecimento de relações entre conceitos. Considerações necessárias para o uso de ontologia são apresentadas na Seção 5.2.3.2 Ontologia e a linguagem recomendada para sua construção, denominada OWL, é tratada na Seção 5.2.3.3 OWL, juntamente com a discussão da sua adequação e complementariedade com RDF Schema. A realização de inferências a partir de modelos lógicos que utilizam a semântica dos dados que descrevem os conteúdos da WS é abordada na Seção 5.2.3.4 Realização de inferências na Web Semântica.

5.2.3.1 Modelagem de conceitos

Os documentos disponíveis na Web precisam ser localizados e interpretados para terem utilidade. Quando uma pessoa busca um OA para atender uma necessidade específica ela determina alguns dos atributos que esse AO deve ter. Uma necessidade de busca muitas vezes está centrada no assunto dos documentos, mas pode ser refinada a partir de outros dos seus atributos, como por exemplo, a data de publicação. Independente do atributo que determina a necessidade de busca, este é representado por um conceito

que deve ser adotado da mesma forma no conteúdo do documento e pelo usuário para que haja satisfação da sua necessidade.

O conceito é definido como a reunião e compilação de enunciados verdadeiros a respeito de determinado objeto, fixada por um símbolo linguístico. Objeto, por sua vez, é tudo aquilo que circunda o homem e é designado pelo mesmo. Estes podem ser, individuais, caracterizados por tempo e espaço, ou gerais, independentes destas duas características. Um conceito possui *características* referenciadas em cada um dos seus enunciados constituindo os elementos do conceito, obtidos através da análise do conceito e síntese das características (método analítico-sintético). A soma das características de um conceito é a sua intensão, a extensão *do conceito* é a classe dos conceitos para os quais a intensão é verdadeira. Contudo, definição é a delimitação ou fixação do conteúdo de um conceito (conteúdo de conceito = intensão) (DAHLBERG, 1978).

Um modelo é a “representação simplificada e abstrata de fenômeno ou situação concreta, e que serve de referência para a observação, estudo ou análise” (FERREIRA, 2004, p. 1345). Segundo a Teoria do Sistema Geral, também denominada Teoria da Modelização de Le Moigne (1977 *apud* CAMPOS, 2004, p. 23), conhecer é modelizar, ou seja, o processo de conhecer equivale à construção de modelos que nos permitem descrever o que se observa.

Modelar conceitos é o processo abstrato que possibilita a criação de estruturas semânticas apropriadas para descrever a informação a partir da relação entre os conceitos do conteúdo. Os modelos devem ser compostos de explicações baseadas em princípios e formalizadas para permitir a inferência de conhecimento. A formalização é o que garante a compreensão do significado expresso tanto por humanos como por computadores. A descrição da informação com o uso de estruturas semânticas permite gerenciar conteúdos semelhantes e conteúdos diferentes, e isso é um aspecto essencial para WS (ALLEMANG. HENDLER, 2011).

Uma etapa fundamental para modelar conceitos é a própria definição dos conceitos dentro do seu domínio. Essa etapa precede a formulação de axiomas a partir dos conceitos que originarão outros conceitos. A Teoria do Conceito (DAHLBERG, 1978) esquematiza a formação de conceitos ao passo que define, estabelece sua composição e seus relacionamentos. Os princípios desta teoria procuram responder o que é o conceito, quais suas partes e para que este serve. Segundo esta autora, existem conceitos

individuais e gerais, relativos aos diferentes tipos de objetos. “Os elementos contidos nos conceitos gerais encontram-se também nos conceitos individuais, sendo, portanto, possível reduzir os conceitos individuais aos gerais e ordená-los de acordo com os conceitos gerais” (DAHLBERG, 1978, p. 102). A característica de um conceito (ou atributo, em nível de objeto) apresentada em um enunciado pode ser sujeito de um novo enunciado até o estabelecimento de uma característica generalista considerada categoria. Leska (1981, apud CAMPOS, 1995), aponta na Terminologia princípios para expandir as referências lexicais e tornar os conceitos mais precisos e bem definidos, qualificando-os de acordo com seu relacionamento com outros conceitos.

A estrutura semântica do domínio a ser modelado deve ser entendida como a exteriorização do conhecimento organizado conforme as inferências e abstrações pretendidas. A organização hierárquica de conceitos pode ser favorecida com a definição de categorias capazes de identificar os conceitos de acordo com sua natureza, predeterminando a perspectiva a ser analisada. Ranganathan, no âmbito da Ciência da Informação, sugere perspectivas em sua Teoria da Classificação Facetada (RANGANATHAN, 1967):

- Matéria – característica ou propriedade do que esta sendo descrito;
- Energia – referente aos processos envolvidos;
- Espaço – aspecto de localização;
- Tempo – aspecto temporal;
- Personalidade – entidades relacionadas à definição e demais atributos.

A forma como os conceitos são dispostos e relacionados no modelo constitui um todo coeso com definições especificadas até o ponto em que podem ser estabelecidos exemplos. A relação entre os conceitos é fundamental para sua compreensão. Alguns tipos de relações que podem ser estabelecidas entre os conceitos são (CAMPOS, 2004):

- Relação categorial – reúne, em um primeiro grande agrupamento, os conceitos por sua natureza, ou seja, entidades, processos, entre outros. Constitui uma classe de maior amplitude e possibilita diminuir erros lógicos no estabelecimento de outras relações, pois determina a natureza do conceito;
- Relação hierárquica – permite relacionar conceitos de uma mesma natureza ou não, dependendo da abordagem realizada. É uma das principais relações em qualquer estrutura semântica e estabelece uma abstração de gênero e espécie, formando cadeias de conceitos;

- Relação partitiva – recai sobre a constituição do objeto, ou seja, na definição de um todo e suas partes;
- Relação entre categorias – dá-se a partir da relação prescritiva entre conceitos de naturezas distintas. Pode estabelecer um encadeamento, que relaciona contiguidades no tempo; ou uma causalidade, que estabelece um elo sucessivo de causas;
- Relação associativa – estabelecida a partir de uma propriedade específica, comum ou não entre dois conceitos, mas que é capaz de complementar suas definições. A associação facilita a descoberta de conhecimento por meio da indicação de conceitos que agregam significado àquele buscado;
- Relação de equivalência – estabelecida no âmbito da denominação dos conceitos que podem ser representados por mais de um termo. Nessa relação deve ser indicado o termo preferencial ou adotado no domínio específico.

Existem diferentes formas de representar o conhecimento resultante do processo de modelagem de conceitos. A representação possui funcionalidades distintas que demandam instrumentos capazes de contemplar seus objetivos. Brachman (1979, apud CAMPOS, 2004, p. 24) classifica a representação do conhecimento de acordo com os tipos de primitivas oferecidas ao usuário, em quatro níveis:

O nível lógico é o nível da formalização. Não existe, entretanto, preocupação com a semântica em termos dos conceitos e de suas relações; [...] No nível epistemológico, a noção genérica de um conceito é introduzida como uma primitiva de estruturação de conhecimento; ele é o nível da estruturação. O nível ontológico tem por objetivo restringir o número de possibilidades de interpretação do conceito dentro de um dado contexto, a partir de um formalismo que pretende representar o conteúdo do conceito. No nível conceitual, independentemente de um formalismo, os conceitos possuem, a priori, uma interpretação definida. [...] (BRACHMAN, 1979, apud CAMPOS, 2004, p. 24)

A partir desses níveis, alguns exemplos de instrumentos são indicados a seguir:

- Classificação – conjunto de conceitos organizados de acordo com um critério específico (ISO TR 14177, 1994 apud TRISTÃO; FACHIN; ALARCON, 2004) que produz o encadeamento de termos. Os termos

são identificados com notações que podem evidenciar grupos de termos afins;

- Taxonomia – organização hierárquica de termos, ou seja, disposição de conceitos em categorias e subcategorias que facilitam sua compreensão por generalização ou especialização (SILVA, 2004a);
- Tesouro – “vocabulário de termos relacionados genérica e semanticamente sobre determinada área do conhecimento” (MOTTA, 1987). É um sistema de conceitos no qual todos os termos relevantes devem encontrar seu lugar apropriado a partir de relações hierárquicas, associativas ou de equivalência (LESKA, 1981 apud CAMPOS, 1995);
- Ontologia – estrutura de conceitos que define as regras que regulam a combinação e a relação entre termos para viabilizar o entendimento comum e compartilhado de um domínio de conhecimento, de forma que o mesmo possa ser compreendido e explorado por pessoas e computadores (ALMEIDA; BAX, 2003).

Na WS, o interesse em instrumentos que possibilitam a representação do conhecimento, é pautado pela necessidade de descrever a semântica do conteúdo utilizando conceitos relacionados entre si e formalizados. Com isto, os conceitos serão usados com metadados ou delimitações de conteúdo que terão sua semântica inferida por aplicações de computadores. A ontologia preenche os requisitos que circundam a abstração dos conceitos e permite a representação formal do conhecimento (GUARINO, 1997), este assunto será abordado na próxima seção.

5.2.3.2 Ontologia

O uso de ontologia na WS viabiliza a formalização da semântica do conteúdo, bem como a sua relação e contextualização com um conjunto de documentos. Ao implementar as tecnologias apresentadas nas camadas inferiores da WS temos documentos com sintaxe XML e conteúdos representados de acordo a recomendações RDF e seu RDFS. A partir dessa etapa pode-se adotar uma ontologia para agregar semântica ao conjunto de *tags* usadas nessa representação de conteúdos, descrevendo formalmente seu significado. Dessa forma, haverá consenso e compartilhamento dos conceitos relacionados ao documento e inferências semânticas poderão ser realizadas pelas aplicações computacionais.

Na filosofia, ontologia é o estudo de entidades e suas relações. A questão em torno da sua definição é “Que tipo de coisas existem ou pode existir no mundo, e que tipo de relações essas coisas podem ter uma com a outra?” (Semantic Web, 2011). Existem diversas definições de ontologia encontradas na literatura e decorrentes contradições. Uma das definições mais conhecidas é proposta por Gruber (1992) que diz que ontologia é a “especificação formal de uma conceitualização compartilhada”.

A definição proposta por Gruber (1992) é discutida por Guarino & Giaretta (1995, apud MONTEIRO, 2006):

[...] um ponto inicial nesse esforço de tornar claro o termo será uma análise da interpretação adotada por Gruber. O principal problema com tal interpretação é que ela é baseada na noção de conceitualização, a qual não corresponde à nossa intuição. [...] Uma conceitualização é um grupo de relações extensionais descrevendo um ‘estado das coisas’ particular, enquanto a noção que temos em mente é uma relação intencional, nomeando algo como uma rede conceitual a qual se superpõe a vários possíveis ‘estados das coisas. (GUARINO & GIARETTA, 1995 apud MONTEIRO, 2006).

A contraposição dessas definições evidenciam algumas características importantes da ontologia. Para facilitar a compreensão do que se pretende dizer com o uso de ontologias na WS é importante destacar alguns conceitos presentes na definição de Gruber (1996) analisada por Guarino & Giaretta (1995, apud MONTEIRO, 2006). “Formal” significa processado por computadores; “especificação explícita” diz respeito a conceitos, propriedades, relações, funções, restrições e axiomas claramente definidos; “compartilhado” quer dizer conhecimento consensual; e “conceitualização” diz respeito a um modelo abstrato de algum fenômeno do mundo real.

Uma ontologia define os termos, relacionamentos e demais elementos usados para descrever e representar conceitos. Também formaliza o conhecimento do domínio e o que pode ser interpretado sobre o mesmo. Assim, podemos enumerar algumas características para que a ontologia possibilite realizar inferências:

- Definição e estruturação dos termos;
- Estabelecimento de propriedades inerentes ao conceito representado por um termo;
- Povoamento da estrutura através de exemplos que satisfaçam um

conceito e as suas propriedades;

- Estabelecimento de relações entre os conceitos;
- Elaboração de sentenças para restringir inferências de conhecimento baseadas na estrutura.

De acordo com Gruber (1992), alguns componentes básicos e comuns a uma ontologia podem ser enumerados. São estes:

- Classes – organizadas em uma taxonomia;
- Relações – representam o tipo de interação entre os conceitos do domínio;
- Axiomas – usados para modelar sentenças sempre verdadeiras.
- Instâncias – utilizadas para representar elementos específicos, ou seja, os próprios dados.

O uso de ontologias na WS permite criar ‘compromissos ontológicos’ para realizar consultas e fazer afirmações, ou ainda, determinar um vocabulário comum entre as aplicações (Semantic Web, 2011). Em outras palavras, a WS pode ser vista como um conjunto de pequenos componentes ontológicos que apontam entre si e que facilitarão o intercâmbio e processamento da informação entre aplicações Web (HENDLER, 2001).

Diferentes iniciativas tais como, W3C, WebOnt Working Group, Semantic Web Deployment Working Group, Semantic Web for Health Care and Life Sciences Interest Group e Semantic Web Interest Group, concentram esforços para desenvolver metodologias, ferramentas e linguagens para a construção e integração de ontologias. A linguagem OWL é baseada em XML e RDF e por isso esta de acordo com a WS. É a linguagem recomendada pelo W3C e conta com uma camada da WS específica para sua comunicação com RDF, a camada DLP bit of OWL/Rules. Suas regras e especificações são dadas na camada OWL e Rules. Essas tecnologias serão abordadas na seção seguinte.

5.2.3.3 OWL

A XML fornece uma sintaxe para os documentos, mas não impõe restrições semânticas sobre o conteúdo desses documentos. O XMLS define a estrutura dos documentos XML e estabelece tipos de dados. O RDF é uma recomendação para modelar os recursos e as relações entre eles, que pode

ser escrita com uma sintaxe XML. O RDFS, também considerado um vocabulário, descreve classes, subclasses e suas instancias; oferecendo um conjunto de primitivas para construir ontologias simples.

O RDFS não permite caracterizar as propriedades das classes, identificar entidades diferentes, estabelecer disjunções, equivalências e cardinalidade. Para tanto, são necessárias linguagens expressivas para construção de ontologias, como o Web Ontology Language (OWL) (MCGUINNESS; HARMELEN, 2004). A OWL é projetada para o uso por aplicações computacionais que precisam processar o conteúdo da informação em vez de apenas apresentá-lo para os seres humanos. Essa linguagem amplia a capacidade de inferências subsidiadas por XML, RDF e RDFS, fornecendo um vocabulário adicional juntamente com uma semântica formal.

A OWL pode ser usada para representar explicitamente o significado de termos e as relações entre estes termos, ou seja, na construção de ontologias. Esta linguagem é uma revisão da DAML + OIL, DARPA Agent Markup Language acrescida da Ontology Inference Layer ou Ontology Interchange Language, que é uma linguagem de marcação semântica para recursos da Web baseada em RDF e RDFS (CONNOLLY *et al.*, 2001). Possui três sublinguagens em níveis crescentes de expressividade, sendo uma a extensão da sua antecessora, como se segue (MCGUINNESS; HARMELEN, 2004):

- OWL Lite – permite a representação de hierarquias de classes e restrições simples. Tem uma menor complexidade formal;
- OWL DL – permite uma representação com alta expressividade, mantendo completude computacional (*computational completeness*) e a decidibilidade (*decidability*), ou seja, todas as conclusões são computáveis e seu processamento termina em tempo finito. OWL DL inclui todas as construções da linguagem OWL, mas com algumas restrições, como por exemplo, enquanto uma classe pode ser uma subclasse de muitas classes, uma classe não pode ser uma instância de outra classe.
- OWL Full – permite uma representação máxima expressividade e liberdade sintática, mas não possui garantias computacionais. Por exemplo, em OWL Full uma classe pode ser tratada simultaneamente como classe e como instância.

A escolha entre as sublinguagens OWL Lite e OWL DL depende do quanto a expressividade é necessária para representar os conceitos de um domínio. Algumas precauções devem ser tomadas para garantir que aquilo que

foi definido na camada anterior, RDF e seu RDFS, esteja em conformidade com as restrições adicionais impostas por OWL DL e OWL Lite. Já a escolha da OWL Full conta com uso de sintaxe amplamente adequada a definições de RDFS já existentes, mas o apoio à inferências é menor. O uso de pequenas ontologias genéricas e menos expressivas é recomendado para aplicações WS satisfatórias (HERMAN, 2009).

Entre as camadas OWL e RDFS, esta a DLP bit of OWL/Rules que acrescenta a Description Logic Programs (DLP) como uma linguagem para integrar RDFS e OWL (GROSOF *et al.*, 2003). Esta camada baseia-se na lógica descritiva da OWL DL e na Programação Lógica (F-Logic) e tem sido discutida por pesquisadores, não havendo recomendações do W3C ainda (HORROCKS *et al.*, 2005).

Posicionando a OWL na estrutura da WS destaca-se sua contribuição para o cumprimento do objetivo da pesquisa sem a preocupação de esgotar o assunto. O W3C disponibiliza um conjunto de documentos com detalhes adequados a sua aplicação, tais como, o OWL Guide que demonstra o uso da linguagem OWL e apresenta um glossário da terminologia utilizada (MCGUINNESS; HARMELEN, 2004); OWL Reference que sistematiza e descreve as primitivas de modelagem da linguagem (SMITH; WELTY; MCGUINNESS, 2004) e o OWL Semantics and Abstract Syntax que é a definição normativa da linguagem (PATEL-SCHNEIDER; HAYES; HORROCKS, 2004).

5.2.3.4 Realização de inferências na Web Semântica

Como foi discutido, a proposta da WS é viabilizada a partir da implementação de tecnologias dispostas em camadas que devem estar integradas. A ontologia desempenha um papel crucial ao agregar semântica aos conteúdos por meio da definição de conceitos utilizados na sua descrição. Uma vez que os documentos da WS estão estruturados e sua semântica é representada, são necessárias regras de inferência e definições lógicas a serem usadas na descoberta de conhecimento. A descoberta de conhecimento diz respeito à identificação dos dados, seus conceitos e relacionamentos que permitem respostas a diversas consultas a cerca dos documentos.

Segundo a W3C.ORG (2011), uma inferência pode ser caracterizada pela descoberta de novos relacionamentos entre os conteúdos disponibilizados.

"Inference" significa que os processos automáticos das aplicações computacionais podem gerar novas relações com base nos dados e em alguma informação adicional, como por exemplo, um conjunto de regras. Ambas as abordagens recorrem à representação do conhecimento. As novas relações são formalmente acrescentadas ao conjunto de dados, ou devolvidas no momento da consulta.

As regras de inferência são introduzidas na WS desde o momento que predicados e valores são associados aos dados, ou seja, com o RDF. As questões lógicas a serem interpretadas e respondidas por uma aplicação são apresentadas na camada Logic Framework. Essa camada acrescenta regras de inferência mais abrangentes, capazes de gerar conhecimento mediante um processo racional. Para comprovar a coerência dos resultados de uma inferência é preciso traduzir a racionalidade interna em uma linguagem de prova, acrescentada pela camada Proof. A camada Proof possibilita a verificação da coerência lógica para garantir que os aspectos semânticos estejam descritos adequadamente, de acordo com os requisitos das camadas inferiores (SILVA; MCGUINNESS; FIKES, 2006).

Trust é a camada que agrega confiabilidade ao resultado de uma inferência. Essa camada é importante, pois muitas vezes os recursos são relacionados entre fontes distintas, por exemplo, agregando um Projeto X a um Projeto Y que utiliza uma ontologia OWL construída por pessoas desconhecidas. Dessa forma, o resultado deve ser acompanhado por uma justificativa para o usuário que assegure que a fonte é correta, que não sofreu alteração e que não possa ser negada por seu emissor. Linguagens como a Proof Markup Language (PML) são adotadas nessa camada (SILVA; MCGUINNESS; FIKES, 2006).

Por fim, as camadas Signature e a Encryption preveem a assinatura e autenticação do responsável por um documento, mensagem ou registro na WS. A criptografia da assinatura deve garantir sua integridade e dificultar a reprodução por terceiros (BARTEL; BOYER; FOX; LAMACCHIA; SIMON, 2008).

Todas essas camadas, Logic Framework, Proof, Trust, Signature e Encryption, estão em discussão e ainda não possuem recomendações por parte do W3C, motivo pelo qual não foram detalhadas. A OWL, camada que as antecede, ainda não possui informação suficiente sobre sua integração às camadas inferiores, mesmo que a camada DLP bit of OWL/Rules tenha sido introduzida com essa finalidade. Isto é necessário para o desenvolvimento de

camadas superiores. Como pode ser percebido, a WS esta em constante desenvolvimento e sua arquitetura é sempre revista por especialistas. Entretanto, os objetivos que norteiam sua utilização tendem a permanecer estáveis e permeiam alguns projetos em andamento, inclusive, na área de Ciência da informação, Educação e Saúde, temáticas dessa pesquisa, como apresentado a seguir.

5.2.3.5 Web Semântica e Repositórios Digitais Educacionais

A investigação sobre as tecnologia da WS procurou destacar sua contribuição para o acesso aos conteúdos disponibilizados, com foco na sua descrição semântica. Isto porque, a construção do modelo proposto no objetivo dessa pesquisa requer soluções para o problema de refinamento dos resultados de busca em RDEs. O refinamento desses resultados deve ser pautado pelas relações semânticas existentes entre seus conteúdos, mas que atualmente requer a análise e interpretação dos usuários para serem estabelecidas.

A WS propõe minimizar o esforço do usuário na avaliação e seleção de documentos pertinentes ao passo que formaliza a semântica encontrada nos conteúdos e estabelece interpretações consensuais e computáveis. Um modelo inovador para descrição de documentos em RDEs em consonância com as tecnologias da WS permitirá ainda que outras iniciativas nas áreas de interesse do RDE possam ser agregadas, gerando novos conhecimentos.

Os RDEs apresentam a descrição dos OAs utilizando padrões de metadados, oportunamente discutidos na seção seguinte. O limite desse estudo é estabelecido pela necessidade de estruturar sintática e semanticamente esses metadados; relacioná-los e definir um vocabulário adequado, que garanta a interpretação semântica dos OAs. Também é importante acrescentar que o modelo a ser proposto deve atender ao refinamento de buscas em repositórios na área de saúde, em que a necessidade de busca extrapola o assunto do OA, podendo ser qualificada pelo objetivo de aprendizagem. Dessa forma, as camadas XML, XMLS, RDF, RDFS e OWL são as mais relevantes para o modelo.

Estudos anteriores já identificaram as contribuições da WS para repositórios e devem ser considerados diante do objetivo da pesquisa. Esses estudos, na maior parte das vezes, destacam a necessidade de explorar

metadados e ontologias, contudo, as principais contribuições divulgadas estão no estabelecimento de perfis de usuários. O foco nas características dos objetos não tem sido evidenciado. Araújo, (2004a), Mohan (2003), Bernes-lee (2007), Botelho (2008), Franciscato (2009b) entre outros autores citados por Relvão (2006), têm falado sobre a necessidade de modificar a forma como os OAs são descritos e disponibilizados na Web e destacam a contribuição da WS para favorecer esse processo. A intenção é justamente permitir a interpretação do conteúdo semântico desses OAs e melhorar sua recuperação. Stojanovic, Staab e Studer (2001, p. 4) lembram que se os OAs forem relacionados com a partir de uma ontologia permitirão a expansão da busca de tópicos de interesse durante o estudo. O IEEE Standard Upper Ontology (SUO), vem trabalhando na definição de ontologias para aplicações na área de educação (IEEE.ORG).

No cenário internacional, existe uma forte comunidade realizando pesquisas de WS aplicada à educação. Tal comunidade organiza desde 2002 uma série de workshops, como o International Workshop on Applications of Semantic Web Technologies for E- Learning. No trabalho apresentado por Mizoguchi et al. (2007) nesse workshop é proposto um conjunto de ontologias para representar o planejamento instrucional baseado em diferentes teorias instrucionais e de aprendizagem. Através do uso destas ontologias, os autores desenvolveram um sistema inteligente de suporte ao professor que o auxilia durante o planejamento de atividades de ensino.

Em outro trabalho, Dicheva e Dichev (2006) apresentam o Topics Maps for eLearning (TM4L), uma interface baseada na WS que auxilia professores e alunos a procurar, usar (reusar) e compartilhar documentos com o uso de mapas conceituais. Gonçalves (2007) especifica uma arquitetura para um sistema de recuperação de OAs baseado nas tecnologias para a WS. Sua pesquisa, citada na introdução desta tese, desenvolve um agente para descobrir OAs e cursos de formação armazenados em sistemas de e-learning, repositórios e outros servidores Web. Diferentemente do modelo a ser proposto, o TM4L e a pesquisa de Gonçalves (2007) oferecem ferramentas de busca distribuída e não necessariamente propiciam o refinamento dos resultados.

Isotani et al. (2009) ressalta que os novos modelos de busca para conteúdo educacional na Web devem levar em consideração diversos parâmetros como, por exemplo, aspectos culturais, estilos de aprendizagem e estratégias pedagógicas. Alguns desses parâmetros podem estar presentes nos metadados dos OAs. Ao relacioná-los a semântica de um termo que figure

explicitamente nesses metadados pode contribuir com o refinamento da busca. O modelo a ser proposto irá considerar os parâmetros dos OAs que descrevem seus aspectos educacionais para aumentar satisfação diante dos resultados de busca de um RDE na área de saúde.

5.3 Metadados

Amplamente denominados “dados sobre dados” devido a sua etimologia e análise do prefixo “meta”, os metadados são elementos que compõem estruturas para tratar a informação, tendo em vista o seu gerenciamento (ALVES, 2005). Para os profissionais da Ciência da Informação, o termo está relacionado, mais especificamente, com as formas de representação de um documento para fins de descrição, identificação, busca e recuperação, ou seja, dados bibliográficos e catalográficos que servem para organizar e tornar a informação acessível (CASTRO; SANTOS, 2007). Essa interpretação pode ser ampliada para o uso do termo no contexto da WS.

Para a WS os metadados são compreendidos como a representação, processada por aplicações de computadores e compreendida por pessoas, de quaisquer características de um conteúdo a ser disponibilizado. A representação requer um rótulo, que indica ou sinaliza o que se pretende informar sobre o conteúdo; um dado, que é o conteúdo em si ou uma referência ao conteúdo; e um padrão lógico que irá fornecer o significado e a formalização necessária para o metadado. A Figura 10 mostra metadados de um formulário de dados bibliográficos e a Figura 11 metadados embutidos em um documento.

Formulário de um RDE

Título:	Divisão celular (com libras)
Tipo do recurso:	Vídeo
Objetivo:	Ilustrar as etapas da mitose e os principais acontecimentos em cada uma de suas fases
Descrição do recurso:	Vídeo da série Viagem à Célula, ilustra as etapas da mitose e os principais acontecimentos em cada uma de suas fases
Observação:	Este material faz parte de uma série de conteúdos digitais voltados ao ensino de Biologia, produzidos pelo Projeto EMBRIO, da Universidade Estadual de Campinas com recursos do FNDE, MCT e MEC
Componente Curricular:	Ensino Médio::Biologia
Tema:	Educação Básica::Ensino Médio::Biologia::Identidade dos seres vivos
Autor(es):	Galembeck, Eduardo; Santos, Erica Rodrigues dos; Lourenço, Víctor Toni; Projeto Condigital MEC - MCT; Milani, Renato; Vicente, Cristina Pontes; Universidade Estadual de Campinas - Unicamp - Biologia
Idioma:	Português (pt)
País:	Brasil (br)
Data de publicação:	2012-05-09

Metadados

Campo Dublin Core	Valor
dc.contributor.author	Galembeck, Eduardo
dc.contributor.author	Santos, Erica Rodrigues dos
dc.contributor.author	Lourenço, Víctor Toni
dc.contributor.author	Projeto Condigital MEC - MCT
dc.contributor.author	Milani, Renato
dc.contributor.author	Vicente, Cristina Pontes
dc.contributor.author	Universidade Estadual de Campinas - Unicamp - Biologia
dc.relation.requires	VTS_28_Libras.flv
dc.identifier.uri	http://objetoseducacionais2.mec.gov.br/handle/mec/21216
dc.date.issued	2012-05-09T14:53:59Z
dc.date.available	2012-05-09T14:53:59Z
dc.date.accessioned	2012-05-09T14:53:59Z
dc.date.submitted	2012-05-09T14:53:59Z

Figura 10: Metadados de um formulário de dados bibliográficos

Fonte: Disponível em: <<http://objetoseducacionais2.mec.gov.br/handle/mec/21216?show=full>>.

Conteúdo WS



Allegorie op de ondergang van de Spaanse Armada en de ontdekking van het buskruitverraad, 1588-1605 ()

Creator: [prentmaker: anoniem](#) | [vermeld op object drucker: anoniem](#) | [Contributor: Ward, Samuel \(naar ontwerp van\)](#)

Date: [eerste kwart 17e eeuw](#) ; 1621 - 1621

Geographic coverage: [Noordelijke Nederlanden](#) ; [Amsterdam](#) ; [Ipswich](#)

Type: [prent](#) | [propagandaprent](#) | [Subject: Filips II, koning van Spanje](#) | [House of Parliament \(Londen\)](#) | [Iconclass code: 11P9](#) | [Iconclass code: 45C17\(GUNPOWDER\)](#) | [Description: Politieke allegorie op de ondergang van de Spaanse Armada in 1588 en de ontdekking van het buskruitverraad in 1605. Links de Spaanse vloot op zee, in het midden een tafel met baldakijn waaraan de duivel, Filips II, de paus en jezueïet en andere Roomse geestelijken zitten, rechts Guy Fawkes met een lantaarn bij de kelder van de Houses of Parliament op 5 november 1605. In de hemel windkopjes.](#)

[See more](#) >

Data provider: [Rijksmuseum](#) | [Provider: Rijksmuseum](#) | [Netherlands](#) | >

Metadados embutidos

```
xmlns="http://www.w3.org/1999/xhtml"
xmlns:og="http://ogp.me/ns#"
xmlns:fb="http://www.facebook.com/2008/fbml"
xmlns:dcterms="http://purl.org/dc/terms/"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:europa="http://www.europeana.eu/schemas/ese/"
xmlns:cc="http://creativecommons.org/ns#"
xmlns:marcrel="http://id.loc.gov/vocabulary/relators/"
xmlns:foaf="http://xmlns.com/foaf/0.1/"
xmlns:enrichment="http://www.europeana.eu/schemas/ese/enrichment/"
xmlns:dc="http://purl.org/dc/elements/1.1/"
<meta about="http://www.europeana.eu/resolve/record/90402/185215BD42FE734547BCF8FE16DB1722E690E5D0"
property="dcterms:spatial" content="Noordelijke Nederlanden"/>
<meta about="http://www.europeana.eu/resolve/record/90402/185215BD42FE734547BCF8FE16DB1722E690E5D0"
property="dcterms:spatial" content="Amsterdam"/>
<meta about="http://www.europeana.eu/resolve/record/90402/185215BD42FE734547BCF8FE16DB1722E690E5D0"
property="dcterms:spatial" content="Ipswich"/>
<meta about="http://www.europeana.eu/resolve/record/90402/185215BD42FE734547BCF8FE16DB1722E690E5D0"
property="dc:contributor" content="Ward, Samuel (naar ontwerp van)"/>
<meta about="http://www.europeana.eu/resolve/record/90402/185215BD42FE734547BCF8FE16DB1722E690E5D0"
property="dc:creator cc:attributionName" content="prentmaker: anoniem"/>
<meta about="http://www.europeana.eu/resolve/record/90402/185215BD42FE734547BCF8FE16DB1722E690E5D0"
property="dc:creator cc:attributionName" content="vermeld op object drucker: anoniem"/>
<meta about="http://www.europeana.eu/resolve/record/90402/185215BD42FE734547BCF8FE16DB1722E690E5D0"
```

Figura 11: Metadados embutidos em um documento

Fonte: Disponível em: <view- ource:http://www.europeana.eu/portal/record/90402/185215BD42FE734547BCF8FE16DB1722E690E5D0.html>

Conforme mencionado, o conceito de metadado pode ser diretamente relacionado ao propósito dos RDEs e seu uso é requisito para implementação e cumprimento dos objetivos da WS. Em repositórios, seja qual for sua plataforma ou arquitetura de software, os metadados são apresentados em formulários para o registro de dados descritivos a cerca dos documentos disponibilizados. É por meio de metadados que são compostas listas e coleções de documentos afins para facilitar sua gestão e utilização. No caso dos documentos específicos disponibilizados em RDEs, os OAs, é comum também que os metadados estejam na sua composição – estrutura e sintaxe, descrevendo características bibliográficas

e, principalmente, educacionais para facilitar seu compartilhamento, uso, reuso e agregação.

Na WS os metadados são parte do conteúdo, sinalizados com *tags* específicas; ou relacionados ao documento, com a adoção formal de um padrão. São os metadados os responsáveis pela caracterização e interligação dos dados. Para ampliar a compreensão de sua semântica são introduzidas as ontologias. As ontologias podem ser usadas para conceituar e acrescentar regras de inferência estabelecendo a semântica dos dados que descrevem outros dados e dos próprios dados. Em outras palavras, uma ontologia pode ser adotada para descrever o significado e delimitar a interpretação de um metadado em si, como “dc:subject” sinônimo de “lom:subject”; e também pode ser utilizada para conceituar e classificar o conteúdo desse metadado, como “saúde coletiva” – hierarquicamente relacionado a “saúde”; estabelecendo um vocabulário próprio com definições, relações conceituais, etc. Alves (2005) afirma que as ontologias funcionam como qualificadores dos metadados e que não haverá definição da semântica sem a representação por metadados.

Um dos objetivos do uso de metadados no contexto da Web e dos repositórios digitais é permitir que computadores identifiquem os metadados associados aos documentos, para que esses sejam recuperados e pessoas possam avaliar sua relevância. A WS permitirá que essa avaliação seja apoiada por aplicações computacionais específicas, com a realização de inferências subsidiadas por RDF e OWL, por exemplo. Dessa forma, os metadados promovem a mediação entre a informação em ambiente digital e o usuário (ALVES, 2005). Relvão (2006) define metadado como uma informação estruturada que resume, enriquece ou complementa os objetos ou serviços referenciados [documentos], enfatizando como se dá essa mediação. Para Modesto (2005) metadados são meios de descobrir quais são os recursos existentes e como podem ser acessados, evitando a ambiguidade.

As diferentes definições do termo devem-se à amplitude de perspectivas e abordagens de uso, parametrizadas em diversos campos teóricos e aplicados (RODRIGUES; TAGA; VIEIRA, 2011). No processamento de dados, os metadados fornecem informação sobre outros dados dentro de uma aplicação ou de um ambiente. Nesse contexto, são utilizados para documentar elementos (definição de atributos como nome, tamanho, tipo), estruturas (definição de campos, colunas) ou dados (definição de localização, associação, propriedade). Podem ainda conter anotações sobre o contexto, a qualidade, a condição ou características dos dados (RELVÃO, 2006).

O emprego de metadados pode ser definido sob o ponto de vista do usuário, fornecendo o título de um documento, por exemplo; ou sob o ponto de vista da aplicação, com informação para o intercâmbio e interoperabilidade, como endereçamentos URIs. Observando os metadados usados em aplicações para busca e recuperação de documentos, Márdero Arrellano (2009) e Taylor e Jourdre (2009) definem seus tipos:

- Administrativos – definem a forma de armazenamento e processamento dos dados. Ex. Open Archives Initiative - Protocol for Metadata Harvesting (OAI- PMH);
- Estruturais – marcam e estabelecem a forma de apresentação dos dados. Ex. HTML e XML;
- Descritivos – representam e identificam o conteúdo com descritores e respectivas descrições. Ex. <dc:title> e <lom:title>.

Como o objetivo da pesquisa está relacionado aos REDs, pode-se afirmar que a descrição dos documentos – os OAs – se dá de maneira estruturada, com uso de formulários de metadados associados aos documentos, ainda que os OAs tenham metadados embutidos. Esses metadados apresentados em formulários são chamados extrínsecos e seu conteúdo é preenchido durante o processo de submissão aos repositórios. Os REDs têm a preocupação de organizar e tornar a informação acessível e durante sua implementação são estabelecidos padrões para os metadados adotados. Essa característica faz com que sejam propensos aos objetivos da WS.

Recomenda-se que os OAs sejam disponibilizados acompanhados por um registo de metadados, informando, no mínimo, “Título”, “Autor”, “Assunto”, “Identificador”, tal como encontramos na ficha catalográfica de um livro. Um registo mais completo deveria descrever seu contexto educacional, público-alvo, etc. (RELVÃO, 2006). Esses metadados que compõem o próprio documento são denominados intrínsecos e são empacotados de acordo com padrões específicos como o SCORM. Com a correta utilização de metadados intrínsecos o preenchimento de formulários de metadados em RDEs seria facilitado, muitas vezes automatizado, e a interligação entre OAs seria favorecida, enriquecendo a experiência educacional iniciada com a busca. Pesquisas citadas por Relvão (2006), como Friesen (2004, apud) e Najjar *et al.* (2003, apud), indicam que poucos elementos dos padrões de metadados têm sido utilizados para descrever OAs durante sua concepção, dificultando sua localização e reduzindo o reuso.

Os metadados para OAs são uma evolução da marcação de conteúdo feita durante a sua elaboração para permitir o reuso e interoperabilidade. Quanto maior

seu reuso ou o nível de granularidade, maior o detalhamento com metadados. Nesse contexto, seu objetivo inicial não é a organização de conteúdos, mas sim a descrição, que posteriormente pode ser utilizada para a organização (RELVÃO, 2006).

Padrões para descrever os OAs, seja com metadados intrínsecos ou extrínsecos, favorecem as perspectivas pedagógicas e o uso dos objetos. Entretanto, o interesse nos seus benefícios faz com que novos padrões de metadados sejam criados ou padrões consolidados sejam estendidos, dificultando sua adoção. Para que a padronização promova a interoperabilidade de modo que os benefícios obtidos sejam significativos, são necessárias normas de utilização bem definidas e amplamente divulgadas (RODRIGUES; TAGA; VIEIRA, 2011). Os padrões de metadados mais comuns e com normas claras para permitir seu correto uso na descrição de OAs serão detalhados na subseção seguinte.

5.3.1 Padrões de metadados

A definição de um padrão de metadados tem como propósito a descrição e entendimento comum da informação a cerca de um documento, neste caso um OA. No caso dos OAs disponibilizados em RDEs, os padrões mais utilizados em formulários para descrição – metadados descritivos extrínsecos – são o Dublin Core (DC) e o Standard for Learning Object Metadata (LOM). Outros padrões também são citados: Metadadta Ecoding e Transmition Standart (MET), Machine-Readable Cataloging 21 (MARC 21), IMS Learning Resource Metadata (IMS-LRM), AGLS Metadata Element Set (AGLS-MES), Information technology Metadata Registries (MDR), Model for Metadata for Multimedia Information (MMMI), Meta Data Interchange Specification (MDIS), Multimedia Content Description Interface (MPEG-7), Metadata Standard for Learning Resources (MLR-ISO 19788-2), entre outros (RELVÃO, 2006; RODRIGUES; TAGA; VIEIRA, 2011). Esses padrões são adequados tanto para descrições em formulários, metadados extrínsecos ao OA, como para descrições embutidas nos objetos, metadados intrínsecos no OA. Os metadados descritivos extrínsecos, estruturados em formulários para submissão de OAs em RDEs são o foco desta pesquisa.

A padronização de metadados permite que os dados sejam eficiente e consistentemente armazenados, indexados, recuperados, processados e compartilhados por diversos sistemas (BERTOLETTI-DE-MARCHI; COSTA, 2004, apud RODRIGUES; TAGA; VIEIRA, 2011). Para refinar os resultados de busca em

RDEs, a adoção de padrões de metadados descritivos é fundamental, pois viabilizará modelagem dos conceitos utilizados na representação dos conteúdos do OA com relações e definições formalizadas em ontologias OWL. Além disso, vocabulários controlados, como as taxonomias e ontologias, podem ser utilizadas no preenchimento desses metadados descritivos.

Nas subseções seguintes serão abordados os padrões DC e LOM, uma vez que estes padrões serão utilizados para definir os elementos de metadados que descrevem os OAs no escopo desta pesquisa. O DC é o padrão de metadados mais popular, enquanto que o LOM é o mais utilizado na área de Educação. O RDF também é uma proposta de padronização de metadados e é capaz de expressar elementos de metadados DC e LOM, promovendo seu processamento em aplicações WS.

5.3.1.1 Dublin Core

O Dublin Core (DC) é mantido sob a responsabilidade da Dublin Core Metadata Initiative (DCMI), uma organização aberta que apoia o desenvolvimento de metadados e suas melhores práticas. As atividades da DCMI incluem discussões, colaborações em grupos de trabalho específicos, realização de conferências, reuniões e oficinas, além de esforços para promover a aceitação generalizada de padrões de metadados (DCMIa, 2012).

A ideia que se tornou popular desde o início da divulgação desse padrão, em 1995, é a importância de um núcleo de metadados simples e genérico para descrição de recursos na Web. A partir disso, foram estabelecidos os 15 elementos do DC – ou Dublin Core Metadata Element Set, ratificados pelo informe IETF RFC 5013 do The Internet Engineering Task Force¹⁹, ANSI/NISO Standard Z39.85-2007²⁰ e ISO Standard 15836:2009²¹ (DCMIb, 2012).

Os 15 elementos do Dublin Core Metadata Element Set são: *contributor, coverage, creator, date, description, format, identifier, language, publisher, relation, rights, source, subject, title, type*. Suas definições e comentários sobre o uso são apresentadas no *site* da DCMI²². Os elementos podem ser associados a qualificadores de refinamento de elemento e/ou de esquema de codificação. Os

¹⁹ Link: <<http://www.ietf.org/>>.

²⁰ Link: <http://www.niso.org/apps/group_public/project/details.php?project_id=57>.

²¹ Link: <http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=52142>.

²² Link: <<http://dublincore.org/documents/dces/>>.

qualificadores de refinamento de elemento são termos que especificam com mais precisão o significado desse elemento. Um elemento refinado possui o mesmo significado do elemento não qualificado, porém tem sua amplitude restrita, por exemplo, o elemento *date* tem como qualificadores de refinamento de elemento os termos *Issued*, *Created*, *Available*, *Modified* e *Valid*. Qualificadores de esquema de codificação identificam esquemas que ajudam na interpretação do valor de um elemento, por exemplo, o mesmo elemento *date* tem como qualificadores de esquema de codificação os esquemas W3C-DTF6 e DCMI7 Period.

Esses elementos compõem um vocabulário que faz parte de um conjunto amplo de metadados e especificações técnicas. O conjunto completo apresenta todos os termos usados na definição de metadados (DCMI-TERMS)²³. Cada termo é especificado pelos atributos *Nome do termo* (nome no idioma inglês atribuído de acordo com o padrão), *URI* (identificador único do termo), *Rótulo* (etiqueta em qualquer idioma que melhor representa o termo), *Definição* (afirmação que representa o conceito do termo) e *Tipo de termo* (indicação do valor que pode ser atribuído ao termo). E, quando aplicável, por exemplo, na qualificação, outros atributos são usados para fornecer informação adicional sobre o termo, como: comentário sobre sua aplicação; indicação de documentação oficial relacionada ao termo; definição de suas instâncias, propriedades, classes e domínio; conjunto de valores aceitos. Ademais, são apresentados esquemas de codificação do vocabulário e esquemas de codificação de sintaxe (DCMIc, 2012).

O elemento *title*, com seus atributos, pode ser observado na figura 12. O atributo *Refines* indica a propriedade da qual o termo descrito é uma sub-propriedade. *Version* apresenta uma descrição anteriormente adotada para o termo. *Has range* é a classe da qual o valor descrito pelo termo é uma instância.

Term Name:	title
URI:	http://purl.org/dc/terms/title
Label:	Title
Definition:	A name given to the resource.
Type of Term:	Property
Refines:	http://purl.org/dc/elements/1.1/title
Version:	http://dublincore.org/usage/terms/history/#titleT-002
Has Range:	http://www.w3.org/2000/01/rdf-schema#Literal

Figura 12: Elemento *title*, do Dublin Core.

Fonte: (DCMIc, 2012).

²³ Link: <<http://dublincore.org/documents/2012/06/14/dcmi-terms/>>.

A utilização do DC pode cumprir quatro diferentes níveis de interoperabilidade, de acordo com a perspectiva da DCMI. Os níveis apresentados a seguir decorrem da necessidade de integrar ou relacionar diferentes iniciativas que usam o DC e especificam opções, custos e benefícios das ações para aumento da interoperabilidade (DCMI, 2012).

- Nível 1 – Compartilhamento de termos e definições: os elementos DC possuem definições em linguagem natural. Deixando de lado a semântica formal que permite o processamento por computadores, iniciativas podem fornecer vocabulários para compartilhar o significado de um termo, em seu contexto de utilização, entre pessoas. Por exemplo, o mapeamento dos termos e definições DCMI para conceber o LOM pode ser considerado uma “conformidade Informal” com os elementos DC que compartilha termos e definições;
- Nível 2 – Interoperabilidade semântica formal: a interoperabilidade "semântica" é atingida com formalização propiciada pelas triplas RDF, baseados em vocabulários com termos DCMI. "Semântica", neste caso, se refere a relações formalmente estabelecidas entre os termos e as regras para a utilização de tais declarações para tirar conclusões automáticas (inferências lógicas). Isso inclui o uso de URIs em conformidade com os domínios formalmente especificados para os metadados, as faixas de termos correspondentes e as propriedades relacionadas. Este nível corresponde a aplicação da semântica do RDF aos termos DCMI. Assim, qualquer utilização dos termos deve ser mais precisa, em conformidade com o grafo RDF;
- Nível 3 – Interoperabilidade sintática do conjunto de termos da descrição: acima das especificações dos grafos RDF, a camada com o modelo abstrato DCMI, ou DCMI Abstract Model, fornece noções de sintaxe e limitam o conjunto de termos da descrição, garantindo uma base para a validação e troca de registros de metadado. Este nível corresponde ao uso explícito do modelo abstrato DCMI;
- Nível 4 – Interoperabilidade do perfil do conjunto de termos da descrição: a especificação Description Set Profiles: a constraint language for Dublin Core Application Profiles [DC-DSP] é uma linguagem genérica, com um modelo de informação e expressões XML, para estabelecer as restrições estruturais do conjunto de termos da descrição. O DSP limita os recursos que podem ser descritos, as propriedades que podem ser utilizadas, e o modo como um valor podem ser determinado. Neste nível, uma aplicação é

interoperável quando oferece restrições formais compatíveis com a especificação.

A compreensão da complexidade e do potencial de cada um desses níveis é fundamental no uso do DC em aplicações WS, pois quanto maior o nível, maior a capacidade de estabelecer uma semântica formal e regras de inferências. A interoperabilidade de fato só inicia-se no Nível 2, que permite a formalização com RDF. Os Níveis 3 e 4 já são capazes de viabilizar a troca de informação automática entre diferentes iniciativas (DCMIa, 2012).

O DCMI Abstract Model²⁴, utilizado para atingir o Nível 3 de interoperabilidade, foi projetado para atender tanto o contexto da WS, ou Linked Data, com dados relacionados em grafos RDF, como o contexto tradicional de validação de registros de metadados com protocolos como OAI-PMH (DCMIb, 2012). O objetivo do modelo abstrato DCMI é especificar os componentes e constructos dos metadados DC, independente de qualquer sintaxe de codificação. Neste modelo é definida a natureza dos componentes utilizados e como estes podem ser combinados para criar estruturas de informação. Com o DCMI Abstract Model obtém-se uma melhor compreensão dos tipos de descrições codificadas com DC, facilitando seu uso, adequação e, conseqüentemente, o estabelecimento de relações entre metadados, como previsto para a definição do modelo dessa pesquisa (DCMIe, 2012).

O DCMI Abstract Model é baseado no trabalho do W3C de especificação do RDF. O modelo possui uma sintaxe abstrata de acordo com essa especificação e se tornou um dos vocabulários mais usados com RDF. A descrição de um termo DCMI, entendido como um recurso no modelo abstrato, é estabelecida por um ou mais pares de propriedades e respectivos valores. O valor é a entidade física, digital ou conceitual, literalmente informado (valor literal) ou associado com a propriedade (valor não literal). Por exemplo, um valor associado à propriedade do Dublin Core Creator é uma pessoa, organização ou serviço (uma entidade física). Um valor associado à propriedade do Dublin Core Subject é um conceito (uma entidade conceitual) (DCMIe, 2012). Cada uma destas entidades também é um recurso (Figura 13).

²⁴ Link: <<http://dublincore.org/documents/abstract-model/>>.

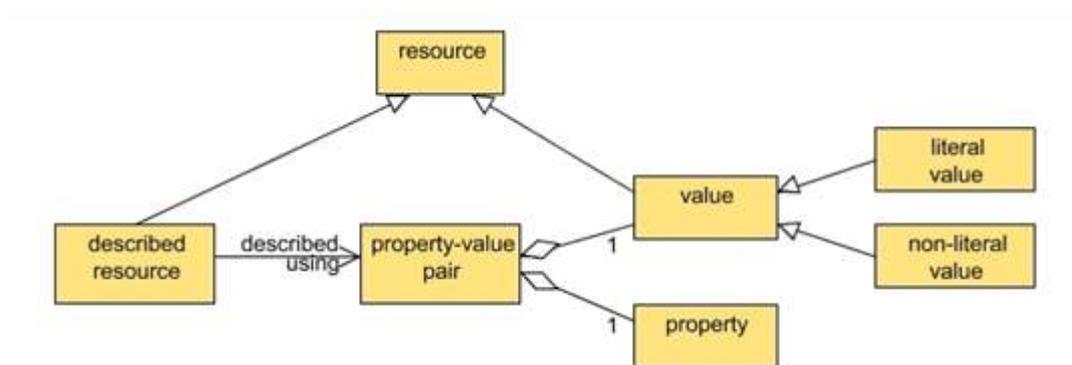


Figura 13: Entidades do Dublin Core como recursos.

Fonte: Disponível em: <<http://dublincore.org/documents/abstract-model/#RDFMT>>.

Os termos DCMI podem ser combinados com outros vocabulários compatíveis com o seu modelo abstrato DCMI Abstract Model, ampliando a possibilidade de aplicação do DC, independente do contexto. O modelo abstrato ainda fornece a noção de um "conjunto de descrição" como uma entidade delimitada com identidade específica. Por exemplo, em um RDE cada "registro" é um conjunto de descrições que pode ser gerenciado (DCMIe, 2012). Não se pretende definir explicitamente uma semântica formal com o DCMI Abstract Model. A intenção é que a semânticas formais possam ser definidas com RDF e por isso há equivalência entre alguns conceitos utilizados por ambos, como *resource*, *class* e *has range*.

Publicações recentes complementam o uso do DC em projetos WS. Uma versão de manutenção de termos de metadados DCMI inclui marcação HTML descrevendo todas as suas propriedades, classes, tipos de dados e esquemas de codificação de vocabulário legível por RDF de acordo com o novo RDFa Lite 1.1, uma sintaxe para a incorporação de dados estruturados em páginas da Web com detalhes necessários para aplicações da WS. Termos de metadados DCMI disponíveis em vários formatos, seguindo as diretrizes do W3C para publicação de vocabulários em RDF, podem agora ser solicitados por navegadores da Web e aplicações de software. Por exemplo, uma descrição RDF do "title" DCMI pode ser solicitado como um arquivo em RDF/XML, via HTTP. O Singapore Framework for Dublin Core Application Profiles é um *framework* para criação e documentação de aplicações com metadados para garantir o máximo de interoperabilidade e reutilização. O *framework* define um conjunto de elementos descritivos e como estes se relacionam com as recomendações da WS (DCMIf, 2012).

A peculiaridade dos OAs levou a necessidade de descritores específicos, principalmente para atender seus aspectos educacionais. Algumas adaptações do DC foram criadas e iniciou-se uma investigação para criar extensões do conjunto de metadados para atender esse contexto. Essa investigação contribuiu para a definição de metadados específicos para objetos de aprendizagem, o padrão Learning Objects Metadata (LOM) (RELVÃO, 2006), comumente adotado nesse contexto, será abordado na subseção seguinte.

5.3.1.2 Learning Object Metadata IEEE LTSC

O Learning Objects Metadata (LOM) surge da necessidade de descobrir, gerenciar e usar objetos de aprendizagem, produzidos em quantidades cada vez maiores. O LOM tem sua origem no ARIADNE²⁵ e no IMS²⁶, além de basear-se no trabalho do DCMI, apresentando, inclusive, uma correlação entre seus elementos e os 15 elementos DC. A intenção dos responsáveis pelos padrões é desenvolver metadados interoperáveis.

Este padrão foi projetado para realizar descrições bem estruturadas de OAs e assim facilitar sua descoberta, localização, avaliação e aquisição por alunos, professores ou aplicações computacionais. O uso do LOM permite compartilhar essas descrições reduzindo o custo de serviços com base na descrição de objetos de alta qualidade; adaptar as descrições para atender a necessidades específicas a partir da escolha adequada vocabulários controlados; “marcar” os próprios OAs com sua descrição (BARKER, 2005).

É um padrão desenvolvido sob a responsabilidade do IEEE LTSC que especifica um esquema de dados conceituais para estabelecer instâncias metadados para OAs, viabilizando a interoperabilidade semântica. A instância de metadados de um OA possui elementos de dados que descrevem suas características (IEEE, 2002).

O LOM é um modelo de dados hierárquico (Figura 14), normalmente codificados em XML, que inclui elementos de dados agregados ou simples. Em seu esquema, apenas os elementos simples possuem valores associados e tipo de dados definidos. Os elementos de dados usados na descrição de OAs são agrupados nas seguintes categorias:

²⁵ Link: <<http://www.ariadne-eu.org/>>.

²⁶ Link: <<http://www.imsproject.org/>>.

- Geral (*General*) – informação geral que descreve o OA como um todo;
- Ciclo de vida (*Lifecycle*) – características relacionadas com o histórico e o estado atual do OA e como essas têm influenciado sua evolução;
- Meta-Metadados (*Meta-Metadados*) – caracteriza a instância de metadados em si e não o objeto;
- Requisitos técnicos (*Technical*) – requisitos e características técnicas do objeto;
- Educacional (*Educational*) – características educacionais e pedagógicas;
- Direitos de propriedade e uso (*Rights*) – especifica os direitos de propriedade intelectual e condições de uso do AO;
- Relação entre objetos (*Relation*) – define quais relações podem ser estabelecidas entre os OAs;
- Uso educacional (*Annotation*) – fornece comentários sobre o uso educacional do OA e informa quando e por quem os comentários foram criados;
- Classificação (*Classification*) – descreve um OA em relação a um sistema de classificação específico e pode ser usada para fornecer extensões ao esquema.

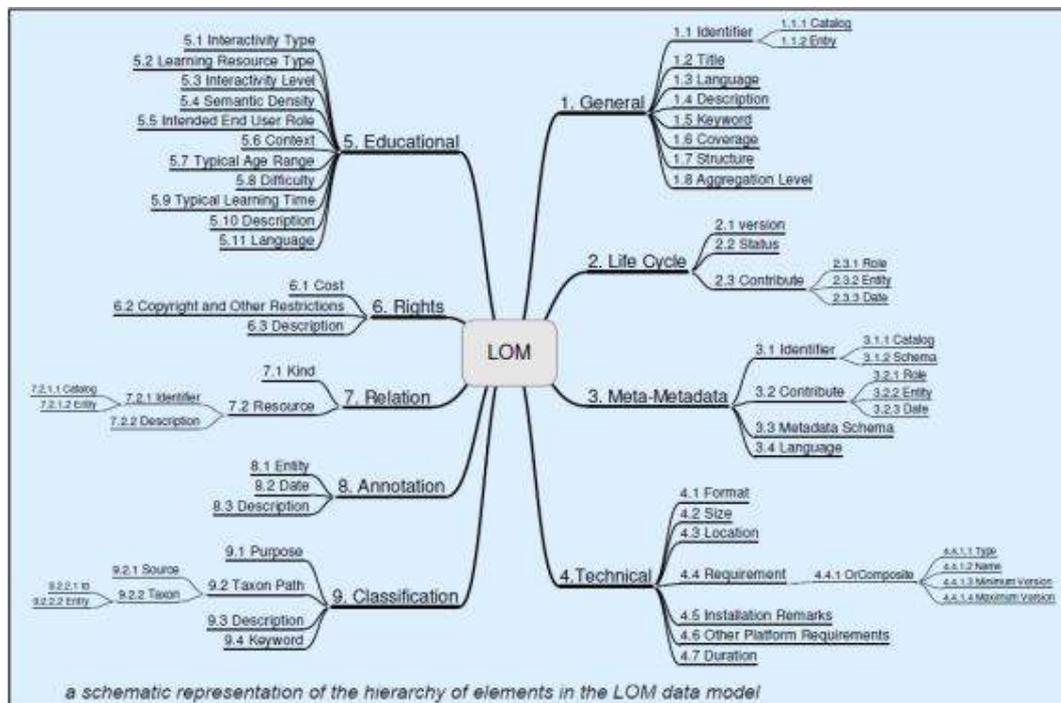


Figura 14: Representação esquemática da hierarquia dos elementos do LOM.

Fonte: (BARKER, 2005).

Os primeiros elementos da hierarquia compõem nove categorias, cada uma

dessas com seus próprios elementos. Estes elementos podem ser simples, armazenando dados, ou podem ser de agregação, concentrando outros sub-elementos. A semântica de um elemento é determinada pelo seu contexto, pelas características dos elementos hierarquicamente acima (*pai*) e por aqueles que compõem a mesma categoria. Por exemplo, os diversos elementos “Description” 1.4, 5.10, 6.3, 7.2.2, 8.3 e 9.3 têm significados e utilizações distintas herdadas de seus elementos *pai*. Além disso, o elemento 9.3 é contextualizado a partir da definição dos elementos 9.1 e 9.2 oriundos da categoria 9, exemplo que se aplica a todas as demais categorias (BARKER, 2005).

Cada elemento de dado possui os atributos: Nome (nome pelo qual o elemento de dado é referenciado), Explicação (definição do elemento de dado), Tamanho (número de valores permitidos), Ordem (indicação se ordem dos valores é significativa) e Exemplo (ilustração do uso). Para os elementos de dados simples, também são definidos os atributos: Espaço de valor (conjunto de valores permitidos para o elemento de dados), Tipo de dados (indica se os valores são *LangString* - qualquer sequência de caracteres Unicode, *DateTime* - data ou a duração em formato legível por máquina, *Vocabulary* - opção de uma lista declarada, ou seja, um vocabulário controlado) (IEEE, 2002).

A IEEE 1484 (IEEE, 2002) é um dos documentos do padrão e descreve o modelo de dados LOM, seus metadados e significados (semântica), as relações entre os elementos (estrutura), o tipo e o tamanho dos dados a serem preenchidos nos elementos (tipo de dados) (ROBSON, 2000). Esta norma especifica quais os aspectos de um objeto de aprendizagem deve ser descritos e que vocabulários podem ser usados para essa descrição, além disso, define como este modelo de dados pode ser alterado.

Os elementos LOM são todos opcionais, mas segundo (Godby, 2004) Relvão (2006), a utilização do padrão determina níveis de descrição:

- Nível 1 – semelhante a uma descrição bibliográfica simples, requer metadados da categoria *Geral*, como título, autor, identificador único, idioma e palavras-chave. Esse nível de descrição pode ser ampliado adicionando os elementos *Relação entre objetos*, *Direitos de propriedade e uso* e *Classificação*;
- Nível 2 – contempla elementos que descrevem o contexto social e técnico do OA. Os elementos das categorias *Educacional* e *Uso educacional* representam detalhes da experiência de aprendizagem como, Para qual público foi planejado? Que tipo de interação é requerida? Qual o grau de dificuldade?

- Nível 3 – reconhece o conjunto de metadados que descreve um OA como um item com propriedade intelectual, que deve ser mantido separadamente, porque uma descrição tem o seu próprio autor, idioma, vocabulário controlado, todos elementos da categoria *Meta-Metadata*.

Embora o LOM possua metadados para descrever informações educacionais, não é incluída informação para especificar modelos e teorias instrucionais, mantendo o foco no que vai ser ensinado e não em como ensinar, caracterizando um nível básico de abstração. Allert (2002, apud RELVÃO 2006) propõe um *framework* para ampliar a dimensão pedagógica do LOM, o Basic Instrucional Framework. Sua proposta é dividida em quatro camadas de abstração que se inicia com os aspectos educacionais representados pelo LOM e estende a descrição de modelos pedagógicos, tipos de aprendizagem (baseada em problemas, casos, etc.) e, por último, informação sobre a teoria do conhecimento, epistemologia.

Iniciou-se a elaboração de documentos para definir associações do modelo de dados LOM, ou seja, definir como registros LOM devem ser representados em XML e RDF, como as recomendações IEEE 1484/12/03 e IEEE 1484/12/04 (BARKER, 2005). Contudo, em reunião do seu grupo de trabalho, foi decidido que a associação entre o LOM e o RDF LOM seria interrompida. O motivo foi a formação de uma força-tarefa entre o LTSC IEEE e o Dublin Core Metadata Initiative, que produzirá uma recomendação conjunta para a utilização de elementos de metadados LOM de acordo com o modelo abstrato Dublin Core (DCAM). Ao usar o DCAM e sua interligação com o RDF, elementos LOM podem ser expressos adequadamente para aplicações da WS. Isso conduzirá uma significativa interoperabilidade entre os padrões (Joint DCMI/IEEE LTSC Taskforce, 2012).

5.3.2 Interoperabilidade entre Dublin Core e Learning Object Metadata

Uma vez especificados os padrões de metadados DC e LOM, torna-se evidente o avanço do modelo abstrato Dublin Core, bem como, a consolidação de suas recomendações, inclusive para atender a WS. Entretanto, o DC mostra-se insuficiente para descrição de OAs, prerrogativa inerente ao LOM (KOUTSOMITROPOULOS, 2010). Diante disso, a associação entre os dois padrões, levando em consideração a melhor contribuição de cada para RDEs, é fundamental para uma descrição eficiente em nível pedagógico, de domínio e documento. Em nível pedagógico explicitam características e atividades necessárias para o êxito no processo de ensino-aprendizagem. O nível de domínio diz

respeito ao conteúdo, sua temática. O nível de documento diz respeito a estrutura, com informação sobre título, autor, etc. (AZOAOU; DESMOULINS, 2005).

A adoção de um esquema de metadados único pode ser inadequada. Como solução para este problema, é proposto a utilização de perfis de aplicação. Um perfil de aplicação é uma agregação de elementos de metadados selecionados entre um ou mais esquemas de metadados para composição de um novo. O objetivo é atender as necessidades e manter a interoperabilidade com seu esquema base. Alguns perfis de aplicação na área educacional são o Canadian Core (CanCore) Learning Resource Metadata Application Profile, o Gateway to Educational Materials (GEM), Metadata Element Set, UK Learning Object Metadata Core (KOUTSOMITROPOULOS, 2010). Esses perfis tanto podem basear-se em DC ou LOM, como podem mesclar os dois, enriquecendo a adoção do DC com metadados LOM, por exemplo.

O crescente interesse em interoperar os dois padrões levou a criação do grupo de trabalho Joint DCMI/IEEE LTSC Taskforce (Joint DCMI/IEEE LTSC TASKFORCE, 2012). Esse grupo tem discutido e avançado na definição de uma recomendação para utilização conjunta do LOM e DC. A proposta do grupo é desenvolver um padrão RDF para os elementos de dados LOM (Standard for Resource Description Framework (RDF) Vocabulary for IEEE Learning Object Metadata (LOM) Data Elements) e uma recomendação para expressar metadados LOM, ou instâncias de objetos, utilizando o modelo abstrato Dublin Core (Recommended Practice for Expressing IEEE Learning Object Metadata Instances Using the Dublin Core Abstract Model). Um *draft* das recomendações já está publicado para consulta na *Wiki*²⁷ do grupo.

O IEEE P1484.12.4™/D1Draft Recommended Practice for Expressing IEEE Learning Object Metadata Instances Using the Dublin Core Abstract Model, disponibilizado para consulta em dezembro de 2008 (IEEE DRAFT, 2008), é um documento ainda não aprovado que deve sofrer alterações antes da sua publicação como padrão, mas já apresenta considerações sobre a interoperabilidade entre DC e LOM, além de descrever o longo processo para alinhar os padrões, analisando suas incompatibilidades fundamentais. A recomendação indica como construir instâncias LOM (um conjunto de elementos LOM para descrever um OA em conformidade com o documento IEEE 1484.12.1-2002) usando o modelo abstrato DC. Isso representa uma solução parcial para o problema de interoperabilidade de metadados em contextos de aprendizagem, educação e formação.

Com essa recomendação em desenvolvimento, destaca-se que o primeiro passo para atingir a interoperabilidade entre os padrões é a compreensão e o mapeamento dos seus elementos. Este mapeamento deve buscar equivalências entre os elementos e seus

²⁷ Link: <<http://dublincore.org/educationwiki/DCMIIEEELTSCTaskforce>>.

valores predefinidos. Instâncias LOM que não estão estritamente de acordo com o padrão, ou seja, que apresentam elementos de metadados que não são definidos no IEEE 1484.12.1-2002, devem ser mapeados para atingir esta conformidade. A categoria Educacional, não pode ser mapeada diretamente. O documento Analysis of each of the LOM Elements²⁸ apresenta uma análise inicial dos elementos LOM e seu mapeamento para DCAM. O documento An example DCLOM description set²⁹ é um exemplo definido com base no mapeamento proposto.

Nilsson, Johnston, Naeve e Powell (2006) propõem um modelo para servir como um guia para a compreensão das relações conceituais entre as estruturas dos diversos padrões de metadados. Isso é ilustrado com um framework de comparação entre DC e LOM. Os autores destacam que o DCAM especifica os conceitos utilizados no framework, a natureza dos elementos e como eles se relacionam para formar uma estrutura de informação. O conjunto de metadados DC deve ser interpretado como um conjunto de afirmações sobre o objeto descrito e o modelo abstrato é a chave para aplicações específicas realizarem essa interpretação.

Para a WS a representação dos elementos e suas relações deve ser processável por máquina. É importante que o valor preenchido nesses elementos também seja processável, de forma que a informação a cerca do OA, registrada em metadados, seja acessada e compreendida por diferentes aplicações. O DC tem histórico de uso do RDF como base para descrever tanto elementos quanto valores de preenchimento. Ao comparar DC, LOM e especificações para WS, Mikael Nilsson, Pete Johnston, Ambjörn Naeve, Andy Powell (2006) identificam componentes correspondentes e sintetizam no quadro a seguir (Quadro 1).

Quadro 1: Componentes correspondentes do DC, LOM e especificações para WS

<i>Framework concept</i>	<i>Dublin Core framework</i>	<i>LOM framework</i>	<i>Semantic Web framework</i>
Abstract Model	DCMI Abstract Model	Implicit in LOM Data Model	RDF Concepts and Abstract Syntax
Metadata Formats	XML, RDF and HTML bindings	XML binding	RDF/XML syntax, N-triples, etc.
Metadata Element Vocabularies	DCMES, large set of external properties and encoding schemes	LOM Data Model includes element vocabulary, various extensions to LOM	Many external element vocabularies
Metadata Value Vocabularies	DCMIType vocabulary. Many external value vocabularies	LOM Data Model includes several basic value vocabularies, many external vocabularies	Many external value vocabularies
Vocabulary Model	Not formalized, but see Baker (2003)	Not formalized	RDF Vocabulary Description Language
Application Profiles	Some published by DCMI, many external application profiles	LOM Data Model includes basic application profile, many external application profiles.	Many in the form of ontologies
Profile Model	Not formalized, but cf "Guidelines".	Not formalized	Possibly OWL, the Web Ontology Language

Fonte: (NILSSON; JOHNSTON; NAEVE; POWELL, 2006)

²⁸ Link: <<http://dublincore.org/educationwiki/DCMIIEEEELTSCTaskforce/LomDCAMAnalysis>>.

²⁹ Link: <<http://dublincore.org/educationwiki/DCMIIEEEELTSCTaskforce/Examples>>.

O quadro traz especificações necessárias para implementação e interoperabilidade, contudo, nem todas estão formalizadas. Os responsáveis pelo DC e o LOM têm feito esforços nesse sentido. O vocabulário para definição de metadados DC é mais maduro que o vocabulário para definição de metadados LOM e não suporta a identificação de elementos URI (NILSSON *et al.*, 2006). Os padrões são motivados a incorporar URIs aos seus elementos para que sejam mais facilmente mapeados e comparados com outros. Nilsson *et al.* (2006) mostra que isso não minimiza diferenças entre os construtos de seus modelos abstratos. O maior avanço está nos vocabulários para valores de preenchimento, em que muitas fontes externas devem ser incorporadas.

Embora haja correspondência entre os componentes dos padrões apresentados no quadro, existem diferenças significativas entre eles. Por exemplo, as estruturas de informação conceituais que eles descrevem, bem como a natureza dos termos utilizados nestas estruturas, são bastante diferentes. No caso do DC e das especificações para WS, os dois modelos abstratos são compatíveis, e isso é refletido na utilização do RDF para descrever metadados DC (NILSSON; JOHNSTON; NAEVE; POWELL, 2006).

6 Metodologia

Com o objetivo de propor uma modelagem, baseada em Web Semântica, para repositórios digitais educacionais na área de saúde, com foco no objetivo de aprendizagem, foram identificadas as características dos objetos de aprendizagem disponíveis nesses repositórios. Dessa forma, foi possível compreender as peculiaridades dos documentos disponibilizados e do seu público-alvo. Em um segundo momento, abordou-se as tecnologias da WS adequadas para a estruturação da informação descritiva dos objetos de aprendizagem. Metadados, recurso utilizado para descrever essa informação, também foram discutidos.

Após estudo exploratório realizado por meio de pesquisas bibliográficas em diversas fontes citadas, será definida uma metodologia para desenvolver a modelagem e destacar sua contribuição para o refinamento dos resultados das buscas em RDEs na área de saúde. A metodologia é o conjunto de processos ou operações mentais empregadas na pesquisa científica (GIL, 1999; LAKATOS; MARCONI, 2006). Para Gil (1999, p.42), a pesquisa tem caráter pragmático, é um “processo formal e sistemático de desenvolvimento do método científico. O objetivo fundamental da pesquisa é descobrir respostas para problemas mediante o emprego de procedimentos científicos”. Assim, a metodologia proposta tem a preocupação de minimizar o caráter subjetivo da pesquisa, pois sistematiza esta ação, primeiramente pautada pela interpretação, e evita impasses na extensão dos resultados a outros contextos.

A contínua evolução da Internet, meio em que se encontram os repositórios, podem tornar esta pesquisa trabalhosa e ao mesmo tempo ressaltar a importância de estudos sobre organização da informação no ambiente digital, assim como este, inserido na linha de pesquisa de Arquitetura da Informação da Pós-graduação em Ciência da Informação da Faculdade de Ciência da Informação da Universidade de Brasília (PPGCInf – UnB). Portanto, vale ressaltar que a Ciência da Informação possui abordagens e conceitos fundamentais para esta pesquisa e ainda, profissionais adequados para elaboração e execução deste tipo de estudo, conforme completa Tillman (2003) ao apontar os profissionais da informação como os mais preparados para determinar e expandir critérios para a informação disponível na Internet.

Com essência interdisciplinar, a Ciência da Informação e seus problemas apropriam-se de metodologias de diferentes áreas do conhecimento, embora haja direcionamento para os instrumentos metodológicos das Ciências Sociais, qualitativos ou quantitativos. Por seu objetivo, essa pesquisa utiliza abordagem qualitativa fundamentada pela interpretação da literatura para formulação teórica, a partir da pesquisa documental. A pesquisa qualitativa considera que há um vínculo entre o mundo objetivo e a subjetividade do sujeito que não pode ser traduzido em números. A interpretação dos fenômenos e a atribuição de significados são básicas no processo de pesquisa qualitativa, que é tipicamente descritiva. O ambiente em que se desenvolve é a fonte direta para coleta de dados, enquanto o pesquisador é o instrumento-chave que tende a analisar seus dados indutivamente (LAKATOS; MARCONI, 2006).

A percepção do sujeito, obtida com a realização de Grupo Focal, apresentada na qualificação desta tese como técnica para verificar a contribuição da modelagem proposta, foi descartada após considerações da banca que ressaltaram a pesquisa documental e sistematização da modelagem como resultados suficientes à elucidação dos objetivos. O Grupo Focal é um conjunto de pessoas selecionadas e reunidas por pesquisadores para discutir um item a partir de sua experiência pessoal (GATTI, 2005), logo, estudos futuros poderão acolher ideias e explorar a opinião de usuários de RDEs no aprimoramento da modelagem (GASKELL, 2005). Ademais, o produto a ser verificado pelo Grupo Focal, compreendido como exemplos de refinamento do resultado de busca em RDEs com a modelagem proposta implementada, extrapola a pretensão de constituir um modelo teórico. Para visualizar a contribuição da modelagem, será apresentado o cenário da implementação e discutidas simulações de resultados de busca refinados.

A revisão bibliográfica sobre o tema mostra a ausência de pesquisas ou reflexões que tenham contemplado esta abordagem, destacando a importância da fundamentação teórica do estudo. Assim, é possível classificar a pesquisa de acordo com seus objetivos, como pesquisa documental, pois estudou o assunto tratado de maneira ainda incipiente na literatura, facilitando o seu entendimento. Quanto aos procedimentos técnicos, trata-se de uma pesquisa bibliográfica (MARCONI; LAKATOS, 2006), pois utiliza como fonte de consulta, material publicado sobre os temas investigados. Desta forma, apresenta-se a seguir, as diretrizes metodológicas para o andamento e realização da pesquisa, organizadas em duas etapas distintas, cada qual com respectivo detalhamento.

6.1 Pesquisa documental

Esta etapa é fundamental para caracterizar os repositórios digitais educacionais, os objetos de aprendizagem disponibilizados nesses e o interesse de busca dos seus usuários. Além disso, apresenta abordagem da Web Semântica identificando suas tecnologias, dispostas em camadas, que potencialmente contribuem para o refinamento da busca em RDE. A investigação sobre metadados destaca como a informação descritiva dos objetos de aprendizagem pode ser representada. Foi realizado o levantamento bibliográfico que respalda o estudo e guia escolhas que melhor atendem seu objetivo. A pesquisa documental constitui-se da pesquisa, leitura e interpretação de documentos científicos (artigos, teses, dissertações, livros, etc.) nos idiomas português, inglês e espanhol, com preferência às publicações mais atuais. Os resultados dessa etapa antecedem as demais, pois subsidiam a compreensão do contexto da pesquisa e ressaltam como os assuntos abordados contribuem para o desenvolvimento do modelo a ser proposto. Os resultados desta etapa são apresentados na Revisão bibliográfica (Seção 5), com o seguinte destaque:

- Repositórios digitais educacionais: assunto abordado no âmbito da educação à distância, ressaltando suas diretrizes e padrões. São apresentados exemplos de repositórios com detalhamento para aqueles na área de saúde, tipo específico para o qual o modelo será proposto. É discutido o objeto de aprendizagem disponibilizado nesses repositórios, bem como, objetivo de aprendizagem;
- Web Semântica: investigada arquitetura, subdividida em “Estrutura e sintaxe” e “Lógica e semântica”. Para aprofundar sobre sua contribuição na representação semântica da informação descritiva do objeto de aprendizagem são feitas considerações sobre modelagem de conceitos e ontologia, além de destacar a realização de inferências na Web Semântica e sua aplicação em repositórios digitais educacionais;
- Metadados: como sendo um recurso adotado em repositórios para descrever informação sobre objetos de aprendizagem, são abordados padrões adequados ao contexto da pesquisa, Dublin Core e Learning Object Metadata e, ainda, a interoperabilidade entre eles.

Como já mencionado, a revisão bibliográfica é necessária para compreender o problema da pesquisa e consolidar sua proposta, para tanto, as seguintes fontes de informação foram pesquisadas:

- Bibliotecas com acervo de livros impressos e anais de congresso:
 - Biblioteca Central da Universidade de Brasília;
 - Biblioteca do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT).

- Bancos de teses e dissertações:
 - Banco de Teses da CAPES <<http://www.capes.gov.br/capes/portal>>;
 - Banco de Teses e Dissertações da UnB <<http://www.bce.unb.br/>>;
 - Biblioteca Digital de Teses e Dissertações (BDTD) <<http://www.ibict.br/>>.

- Bases de dados, Repositórios, Bibliotecas digitais e *sites*:
 - Aléplus <<http://www.ibict.br/aléplus/>>;
 - E-prints in Library and Information Science <<http://eprints.rclis.org/>>;
 - Google <<http://www.google.com.br>>;
 - Library and Information Science Abstracts (LISA) <<http://www.csa.com/factsheets/lisa-set-c.php>>;
 - Repositório <<http://repositorium.sdum.uminho.pt/>>;
 - OAIster <<http://oaister.umd.umich.edu/o/oaister/>>;
 - Scielo - Scientific Electronic Library Online <<http://www.scielo.br/>>;
 - Science Direct <<http://www.sciencedirect.com/>>;
 - Web of Science <<http://www.isinet.com/products/citation/wos/>>;
 - W3C - World Wide Web Consortium <<http://http://www.w3.org/>>.

- Periódicos:
 - Ciência da Informação;
 - DataGramaZero - Revista de Ciência da Informação;
 - Information and Organization;
 - Information Systems Research;
 - Journal of Information Science;
 - Journal of the American Society of Information Science and Technology.

Os documentos recuperados foram classificados segundo os assuntos da pesquisa, Repositórios digitais educacionais, Web Semântica e Metadados. Também foi pesquisado literatura de apoio à contextualização do estudo, com foco na Ciência da informação, e à metodologia, sobre pesquisas documentais com objetivos similares.

6.2 Modelagem baseada em Web Semântica

Adquirida a fundamentação teórica, nessa etapa será consolidada a modelagem proposta. São definidos os passos para verificar tecnologias da WS adequadas para a estruturação da informação descritiva dos OAs dos repositórios digitais educacionais na área de saúde, com foco no objetivo de aprendizagem, como se segue:

1. Estabelecer tecnologias da Web Semântica adequadas à recuperação da informação de repositórios digitais educacionais na área de saúde;
 - 1.1. Recomendar tecnologias para estrutura e sintaxe
 - 1.2. Recomendar tecnologias para lógica e semântica;
2. Estruturar a informação descritiva dos objetos de aprendizagem;
 - 2.1. Modelar os conceitos que especificam objetivo de aprendizagem;
 - 2.2. Identificar ontologias que formalizam os conceitos que especificam o objetivo de aprendizagem;
 - 2.3. Identificar metadados que descrevem os conceitos que especificam o objetivo de aprendizagem.
3. Sistematizar tecnologias a serem implementadas e a estrutura da informação descritiva que compõem a modelagem;
4. Simular resultado de busca em repositório digital educacional na área de saúde com a implementação da modelagem.

7 Desenvolvimento da Modelagem

Este capítulo discorre sobre o desenvolvimento da modelagem para repositórios digitais educacionais na área de saúde, considerando que o referencial teórico foi abordado no capítulo que trata da revisão bibliográfica (Capítulo 5) e as etapas definidas na metodologia da pesquisa (Capítulo 6). São sistematizadas a seguir tecnologias da WS adequadas para a estruturação da informação descritiva dos objetos de aprendizagem. Para tanto, levou-se em consideração o processo de descrição e o processo de recuperação, ou seja, o fluxo de disponibilização de OAs em RDEs.

No que diz respeito à descrição dos OAs, os RDEs contam com um formulário *online* de entrada de dados, a etapa de *upload* de arquivos e a disponibilização para busca. O formulário possui metadados específicos para representação da informação. O processo de recuperação prevê a formulação de uma estratégia de busca pelo usuário, a consulta a base de dados e a apresentação de resultados. A estratégia de busca é composta por termos que representam a necessidade de busca do usuário. Não é objeto dessa pesquisa a discussão sobre as funcionalidades e particularidades de cada sistema ou plataforma de *software*. Assim, a abrangência da modelagem proposta é ilustrada na Figura 15.

O limite desse estudo é estabelecido pela necessidade de estruturar sintática e semanticamente esses metadados; relacioná-los e definir um vocabulário adequado, que garanta a interpretação semântica dos OAs. Também é importante acrescentar que o modelo a ser proposto deve atender ao refinamento de buscas em repositórios na área de saúde, em que a necessidade de busca extrapola o assunto do OA, podendo ser qualificada pelo objetivo de aprendizagem. Dessa forma, as camadas XML, XMLS, RDF, RDFS e OWL são as mais relevantes para o modelo.

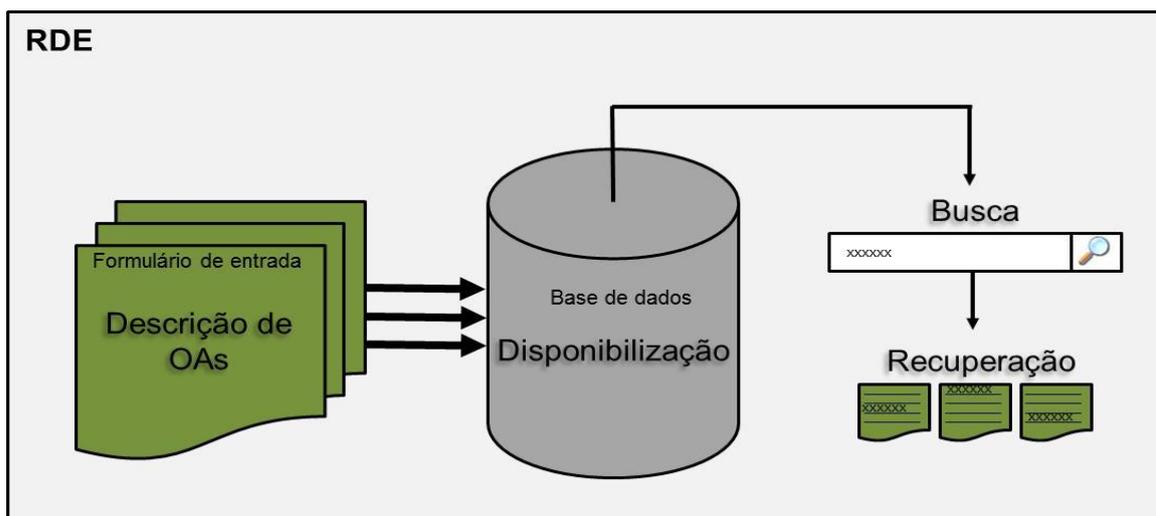


Figura 15: Abrangência da modelagem proposta

7.1 Tecnologias da Web Semântica para a recuperação da informação em repositórios digitais educacionais

Nesta seção é definida a implementação de tecnologias da Web Semântica com o objetivo de refinar os resultados de busca em RDEs na área de saúde, ou seja, favorecer a recuperação da informação. As tecnologias recomendadas são relacionadas ao processo de descrição de OAs, uma vez que este consolida a entrada da informação a ser buscada no processo de recuperação. Essa descrição requer a definição de metadados e o preenchimento dos mesmos com informação a cerca do objeto. O processo de recuperação é compreendido como uma das finalidades do processo de descrição, logo, o refinamento do resultado de busca será consequência das tecnologias implementadas. Para tanto, o RDE deve possuir plataforma de *software* própria que permita a disponibilização de OAs em formato digital, descritos em formulário de entrada com metadados, para serem recuperados pelos usuários. O formulário de entrada de um RDE exibe metadados e campos de preenchimento correspondente, além disso, pode ser acompanhado por um documento com a definição dos metadados, indicando sua descrição, suas relações com outros metadados, suas restrições e regras para o preenchimento.

7.1.1 Tecnologias para estrutura e sintaxe

As tecnologias recomendadas à estruturação e sintaxe da informação que descreve o OA em RDEs contribuem com a definição dos metadados.

Os metadados do RDE devem ser nomeados com URIs no padrão Unicode, identificados de maneira única, tal qual suas características, respeitando o conceito que cada metadado representa e sua relação com os demais.

A descrição do metadado, suas relações com outros metadados, suas restrições e regras para o preenchimento, devem ter sintaxe baseada em XML e XML Schema. Vale lembrar que o conteúdo marcado pelas *tags* XML com informação referente ao metadado é tratado como um elemento de informação, considerado um objeto ou entidade a qual se pode atribuir características específicas (OLIVEIRA, 2002). Para facilitar o intercâmbio de informação na Web, devem ser adotados entidades e atributos padronizados, ou Namespaces (W3C, 2012).

Para permitir inferências no momento da recuperação da informação, a RDF com seu RDF Core e RDF Schema, deverá ser implementada. Para isso devem ser propostos grafos baseados em modelos de dados, ou um conjunto de triplas formadas por um sujeito, um predicado e um objeto, observado a definição, as relações, as restrições e regras para o preenchimento de metadados. Os grafos representam declarações simples sobre os metadados, estabelecendo suas propriedades e atribuindo valores para as mesmas. Os grafos do RDF irão compor metadados, suas triplas e interligações são recursos que devem estar contidos em classes, sempre que houver propriedades e/ou valores semelhantes que denotam sua generalização.

7.1.2 Tecnologias para lógica e semântica

As tecnologias recomendadas para representação da lógica e da semântica da informação que descreve o OA em RDEs contribuem com a definição e com o preenchimento dos metadados.

Os metadados do RDE, com estrutura e sintaxe já definidas, possuem formalização semântica a partir da adoção do RDF. Para caracterizar as propriedades das classes, identificar entidades diferentes, estabelecer disjunções, equivalências e cardinalidade, ou seja, agregar semântica e permitir inferências que contribuam com o refinamento da busca, devem ser adotadas ontologias. Essas ontologias ampliam a semântica dos metadados, aprofundando sua conceituação, e, quando utilizadas no preenchimento dos metadados, ampliam a semântica das instancias descritas, ou seja, dos OAs. A ontologia é um vocabulário adicional.

A OWL é linguagem adotada para a construção de ontologias projetadas para o uso por aplicações que precisam processar o conteúdo da informação em vez de apenas

apresenta-lo para os seres humanos. Essa deve ser usada para representar explicitamente o significado de termos e as relações entre estes termos. A sublinguagens recomendada é a OWL Lite, que permite a representação de hierarquias de classes e restrições simples, ou seja, tem uma menor complexidade formal.

O RDF Query Language and Protocol (SparQL), camada que perpassa do RDF Schema a OWL, propicia linguagens e protocolos para consultar e manipular grafos RDF possibilitando diversas inferências. O SparQL apoiará a delimitação do número de resultados retornados e sua classificação.

Nesta modelagem são adotadas tecnologias até a camada OWL uma vez que a preocupação é atender a descrição e recuperação de OAs em RDEs na área de saúde, um universo específico que não demanda esforços de camadas superiores para garantir a inserção de regras de inferência, prova e criptografia. Além disso, as tecnologias mais discutidas e aplicadas na WS são mais adequadas. Para ampliar os resultados e permitir a interligação com outras fontes de informação recomenda-se a implementação dessas camadas.

7.2 Estrutura da informação descritiva dos objetos de aprendizagem

A informação descritiva dos OAs disponibilizados nos RDEs na área de saúde abordados nessa pesquisa refere-se ao objetivo de aprendizagem e suas relações para favorecer o refinamento da busca. Dessa forma, representam características importantes para seleção, análise e utilização dos OAs no ensino-aprendizagem. Essa informação está estruturada da seguinte forma:

- Conceitos: características, elementos de definição do objetivo de aprendizagem;
- Metadados: descrição de metadados para representar os conceitos que especificam o objetivo de aprendizagem, suas relações com outros metadados, suas restrições, regras para o preenchimento, formulário com conjunto de metadados;
- Preenchimento dos metadados: conteúdo preenchido a cerca do objetivo de aprendizagem; semântica do conteúdo com estabelecimento de classes, relações, instâncias, etc. que podem ser formalizadas.

Para o refinamento do resultado de busca são definidas relações semânticas entre metadados. O preenchimento dos metadados pode ter sua semântica formalizada, dependendo de cada caso e da utilização de vocabulários controlados. As tecnologias para estrutura, sintaxe, lógica e semântica são utilizadas para expressar essa estrutura de informação.

Para a modelagem proposta na pesquisa são considerados metadados e preenchimento de metadados que descrevem aspectos educacionais dos OAs na área de saúde. Assim, são modelados conceitos que especificam o objetivo de aprendizagem, propostas ontologias que formalizam a intensão e extensão desses conceitos e, então, estabelecidos metadados que representam os conceitos. Cada uma dessas abordagens, apresentadas nas subseções seguintes, possuem uma proposta genérica, com as principais definições. É facultada a possibilidade de expandir a estrutura da informação descritiva, abordando outros conceitos que representam diferentes aspectos do OA. Também podem ser constituídas estruturas específicas para necessidades de informação pré-definidas com a realização de inferências robustas.

7.2.1 Modelagem de conceitos que especificam o objetivo de aprendizagem

Nessa etapa são modelados os principais conceitos que especificam o objetivo de aprendizagem dos OAs em RDEs na área de saúde e, para tanto, observa-se o referencial teórico da pesquisa documental (Capítulo 5). A modelagem requer a leitura interpretativa dos conceitos e a identificação de definições, atributos e relações entre termos. Para facilitar seu desenvolvimento é utilizada ferramenta Protégé³⁰, versão 3.4.8, que também dispõe de interface gráfica para edição de ontologias e uma arquitetura para a criação de bases em conhecimento. Vale destacar que a modelagem de conceitos é parte do desenvolvimento de ontologias, assunto abordado na etapa seguinte do desenvolvimento da pesquisa (Seção 7.2.2).

A Protégé foi desenvolvida pelo Stanford Medical Informatics na escola de medicina da Universidade de Stanford (Califórnia, EUA) com o apoio de diversos colaboradores. Pode ser usada tanto por desenvolvedores de sistema como por especialistas em domínio para criar bases de conhecimento, permitindo representar facilmente a multidimensionalidade do conhecimento de uma área. A ferramenta é capaz de tratar classes com definições e exemplos elaborados simultaneamente

³⁰ Link: < <http://protege.stanford.edu/doc/users.html> >.

(PROTÉGÉ, 2012). A escolha da Protégé para a modelagem é favorecida por ser uma ferramenta gratuita e de código aberto; com interface gráfica interativa e amigável; suporte a diferentes formatos de armazenamento, tais como, OWL, RDF e XML; além de facilitar a constituição e agregação de ontologias já existentes.

A leitura interpretativa traz considerações fundamentais sobre o objetivo de aprendizagem para o desenvolvimento da modelagem de conceitos que favorecem o refinamento do resultado de busca, como destacado a seguir:

- O OA é composto de três elementos fundamentais: conteúdo (*learn*), atividade (*practice*) e avaliação (*assess*). O conteúdo advém de um domínio do conhecimento específico e pode ser apresentado em diferentes formas e formatos. A atividade é o elemento proposto para a apreensão do conteúdo. A avaliação mensura a apreensão do conteúdo. O que indica o domínio do conhecimento, o tipo de atividade e a medida da apreensão do conteúdo é o objetivo de aprendizagem (HEINS; HIMES, 2002).
- Como destaca Driscoll (1998), o objetivo de aprendizagem é o resultado esperado de um indivíduo ao final do processo de ensino-aprendizagem. O OA deve ser baseado em um único objetivo de aprendizagem, contudo, o mesmo objetivo de aprendizagem pode abranger conteúdos diferentes, quando abordado por OAs diferentes.
- O nível de agregação dos OAs com objetivos de aprendizagem similares estabelece uma relação entre eles, criando uma hierarquia de objetos complexos aos mais simples dentro de um mesmo domínio. Independente disso, o OA condensa informação necessária para um indivíduo atingir um objetivo de aprendizagem. Logo, o nível de agregação hierarquiza OAs e, por conseguinte, hierarquiza o nível de abstração do público-alvo, definido pela exigência de requisitos para realização de processos cognitivos demandados no objetivo de aprendizagem.
- O objetivo de aprendizagem é determinado para um público-alvo, com contexto educacional específico. Para isso, é constituído a partir da definição de um verbo que indica a ação do indivíduo para com o conteúdo. O OA é sobre um assunto a ser abordado nas atividades. O assunto é parte de um domínio, representado por uma hierarquia de conceitos com intenção, extensão, sinônimos, etc. O domínio influencia diretamente as relações estabelecidas entre assuntos.

Como já mencionado, o objetivo de aprendizagem, assim como o assunto e outros aspectos que caracterizam o OA, são representados por conceitos. A modelagem desses conceitos é a criação de uma estrutura semântica da informação descritiva do OA que se refere ao objetivo de aprendizagem. A modelagem é formalizada com a utilização da ferramenta Protégé para definição dos conceitos.

O desenvolvimento da modelagem constitui um conjunto de decisões tomadas e consequentes atividades realizadas na ferramenta, com base nos passos para Modelagem conceitual apresentada em Monteiro (2006) e nas instruções contidas no *Ontology Development 101: a guide to creating your first ontology* (NOY; MCGUINNESS, 2001), guia recomendado no *site* da ferramenta Protégé. As especificações que se seguem elucidam a modelagem de conceitos.

7.2.1.1 Especificações da modelagem de conceitos

O **objetivo da modelagem** é definir e relacionar conceitos que caracterizam o objetivo de aprendizagem dos OAs disponibilizados em RDEs. A modelagem identifica relações semânticas que contribuem com o refinamento dos resultados em RDEs na área de saúde.

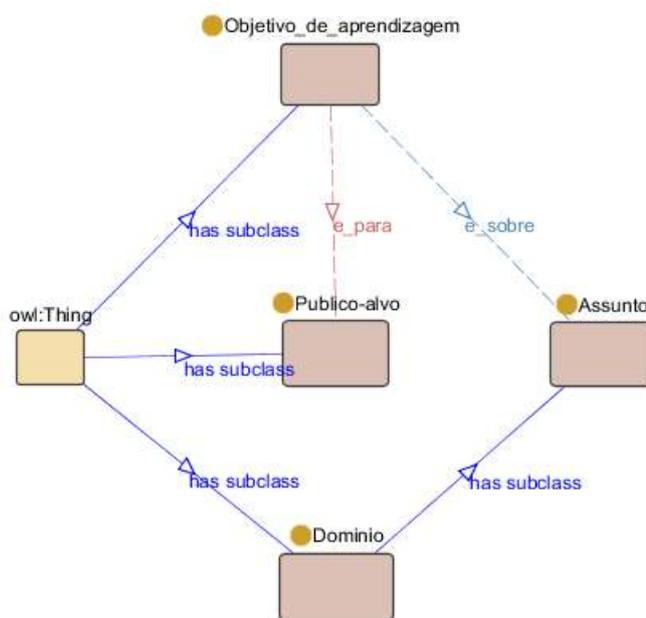
O **domínio** dessa modelagem é “a informação descritiva do OA, disponibilizado em RDE, que caracteriza seu objetivo de aprendizagem”. O **escopo** prevê o entendimento do domínio e uma visão geral de definições e termos. A compreensão semântica dessa informação será mais clara e objetiva, restringindo o número de interpretações dos conceitos e viabilizando inferências.

Os **termos que compõem o domínio** foram retirados da pesquisa documental (Capítulo 5). A partir da análise da informação pesquisada, constitui-se uma lista dos principais termos que serão modelados:

- Objetivo de aprendizagem;
- Domínio;
- Assunto;
- Público-alvo.

Os termos “objetivo de aprendizagem”, “domínio” e “público-alvo”, dadas as suas características individualizantes, são definidos como **classes** ou conceitos (*Class*). São definidas **subclasses** (*Subclass*) a partir do agrupamento e hierarquias de conceitos, como entre a superclasse “domínio” e as subclasses “assunto” e “conteúdo”.

As características ou **propriedades** (*Properties*) dos conceitos são definidas, como por exemplo, o “contexto educacional” para o “público-alvo”. Os valores das propriedades também são definidos a partir do estabelecimento de restrições, como na definição do “tipo: literal” (*type: string*) para o preenchimento da propriedade



“contexto educacional”. **Instâncias** (*Individuals*), ou exemplos de conceitos com respectivas propriedades, são preenchidas para criar uma base de conhecimento. O arranjo de classes e propriedades apresentado a seguir, ilustra os primeiros níveis da modelagem (Figura 16):

Figura 16: Primeiros níveis da modelagem de conceitos.

Como pode ser observado, “objetivo de aprendizagem” possui propriedades extrínsecas que fazem referência aos conceitos de “público-alvo” e “assunto”. Para completar o seu significado também é definida a propriedade intrínseca “especificação” que indica a redação do objetivo em si. O “assunto”, subclasse de “domínio”, possui as propriedades intrínsecas “texto livre”, para indicação de assuntos em linguagem natural, e “taxonomia”, para indicação de assuntos a partir de um vocabulário controlado (VC)³¹. “Público-alvo” possui as propriedades intrínsecas “contexto educacional”, para redação do contexto em que o público-alvo está inserido; “nível educacional” para indicação de níveis educacionais pré-definidos de acordo com a estrutura educacional de ensino-aprendizagem; e “pré-requisito” para indicar a exigência de requisitos para o público-alvo.

³¹ O termo *vocabulário controlado* é adotado para generalizar ontologias, taxonomias e quaisquer outros vocabulários controlados com controle terminológico e semântico.

As etapas seguintes indicam novos conceitos, propriedades, VCs e instâncias que irão ampliar a modelagem de conceitos.

7.2.2 Ontologias que formalizam os conceitos que especificam o objetivo de aprendizagem

Após a modelagem dos principais conceitos são definidos VCs, ou ontologias, que aumentam a compreensão e formalização semântica do objetivo de aprendizagem. Para isso, são consideradas as principais questões que a modelagem deve responder, a reutilização de ontologias e a proposição de novas ontologias de acordo com o interesse da pesquisa. Não há pretensão de exaurir todas as possibilidades, mas sim de identificar vocabulários adequados e demonstrar como são agregados a modelagem proposta.

Principais questões que a modelagem deve responder

As questões que a modelagem deve responder são definidas de acordo com o interesse de refinar os resultados de busca de RDEs na área de saúde. As ontologias devem dispor de informação suficiente às respostas dessas questões.

Tendo o objetivo de aprendizagem como o conceito central para realização de inferências, as seguintes questões podem ser estabelecidas:

- Qual a classificação dos objetivos de aprendizagem de acordo com processo cognitivo, níveis de abstração ou de aprofundamento de conteúdo?
- Um público-alvo específico tem relação associativa com objetivo de aprendizagem específico?
- Um assunto, disposto em uma hierarquia, tem relação associativa com o objetivo de aprendizagem, disposto em uma classificação?
- Qual o assunto relacionado ao objetivo de aprendizagem?
- Qual o público-alvo relacionado ao objetivo de aprendizagem?

Com essas questões sendo respondidas pretende-se refinar a busca com resultados expandidos por inferências, de acordo com as considerações exemplificadas a seguir:

- O resultado de busca pode ser expandido com OAs que possuem objetivos de aprendizagem em um mesmo nível de classificação, para o mesmo público-alvo (fulano definir X, fulano definir Z);
- O resultado de busca pode ser expandido com OAs de acordo com a classificação do objetivo de aprendizagem, levando em consideração o assunto.

Essas questões e as considerações sobre as inferências guiam a definição das ontologias que ampliam a modelagem de conceitos. Outras questões podem surgir e outras inferências poderão ser realizadas de acordo com a informação disposta nas ontologias.

7.2.2.1 Reutilização de ontologias

Como já foi abordado, ontologia é a “especificação formal de uma conceitualização compartilhada” (GRUBER, 1992). As ontologias investigadas e/ou agregadas à modelagem devem definir os termos, relacionamentos e demais elementos usados para descrever e representar conceitos de “objetivo de aprendizagem”, do assunto “saúde” e do “público-alvo dos OAs na área de saúde”. Cabe destacar que também deverão ser atendidas as especificações tecnológicas recomendadas, logo, são adequadas ontologias na linguagem OWL, baseada em XML e RDF.

Foram feitas pesquisas com o intuito de encontrar uma ontologia ou, quando necessário, outro tipo de esquema de representação do conhecimento, como uma taxonomia, sobre o assunto e os conceitos inerentes ao domínio estudado. A busca se deu na Internet, em repositórios de ontologias, projetos vinculados ao W3C e a ferramenta Protégé. Iniciativas identificadas na revisão bibliográfica também foram consultadas.

As ontologias vinculadas a ferramenta Protégé estão disponibilizadas na sua Wiki³², em lista alfabética de projetos cadastrados. O Ontology Repositories³³, disponível na Wiki da W3C, indica outros repositórios de ontologias que foram consultados, tais como, Cupboard³⁴, DERI Vocabularies³⁵, Knoodl³⁶ e o BioPortal³⁷, que é específico para a área de saúde (medicina e biomedicina) e possui mais de

³² *Link:* <http://protegewiki.stanford.edu/wiki/Protege_Ontology_Library>

³³ *Link:* <http://www.w3.org/wiki/Ontology_repositories>.

³⁴ *Link:* <<http://cupboard.open.ac.uk:8081/cupboard-search/>>.

³⁵ *Link:* <<http://vocab.deri.ie/>>.

³⁶ *Link:* <<http://knoodl.com/>>.

³⁷ *Link:* <<http://bioportal.bioontology.org/>>.

300 ontologias que podem ser consultadas e manuseadas em ferramentas próprias, inclusive fazendo comparativos de termos apresentados nas diferentes ontologias.

Também foram realizadas buscas no Google pelos termos que representam o domínio de interesse, em português e em inglês, e posteriormente verificada autenticidade do responsável pela disponibilização das ontologias encontradas. Neste momento, manteve-se o foco em ontologias mantidas por instituições reconhecidas, atualizadas e aplicadas em projetos de WS.

Ontologias com conceitos de objetivo de aprendizagem

Conceitos de **objetivo de aprendizagem** foram identificados em ontologias sobre OA, metadados para OAs e teoria de aprendizagem. Algumas ontologias específicas são consideradas para reutilização de acordo com os propósitos dessa pesquisa, como será discutido.

O LOCO-Analyst é uma ferramenta educacional para fornecer informação ao professor sobre o processo de ensino-aprendizagem que utiliza tecnologias de WS e um framework baseado em ontologias (Learning Object Context Ontologies) para representação formal dos dados. É resultado do trabalho de Jelena Jovanović, Dragan Gašević e Carlo Torniai no âmbito da Canadian International Development Agency e disponibiliza as ontologias em *site* próprio.

A “Learning Context ontology” relaciona conceitos do contexto de aprendizagem que ampliam a compreensão do objetivo de aprendizagem, porém não apresenta conceitos específicos do domínio. A aprendizagem é compreendida como um “evento” com características específicas como atividade (classe “activity”) que um aluno ou professor realizou ao interagir com um conteúdo de aprendizagem. A classe “contente” faz referência ao conteúdo de aprendizagem utilizado ou produzido no ensino-aprendizagem, “description” é a propriedade que descreve conteúdo, “domaintopic” é a propriedade aborda o tema do conteúdo. “Interagir de forma assíncrona” com outros alunos e “ler” um documento são exemplos de atividade (JOVANOVIĆ, 2012).

Carsten Ullrich propõe uma ontologia desenvolvida no âmbito do EC Project LeActiveMath do German Research Center for Artificial Intelligence que apresenta conceitos sobre a função de um OA. OA é a classe principal da ontologia e as propriedades possuem URIs e metadados LOM. A classe “concept” apresenta informação do que está sendo ensinado no OA. A subclasse “definition”, por exemplo, descreve OAs que apresentam o significado de uma palavra, frase ou símbolo; já a subclasse “theorem” descreve OAs que apresentam uma ideia demonstrada como verdadeira (ULLRICH, 2004).

As propriedades da classe “Concept” e de suas subclasses criam relações que estabelecem condições ou circunstâncias que um OA deve preencher para que ela possa fazer parte da classe. A classe “Satellite” apresenta os principais elementos constitutivos do domínio, definindo essas propriedades. A propriedade “interactivity”, por exemplo, caracteriza OA que oferece aspectos interativos. Uma “interactivity” é usada para desenvolver ou treinar uma habilidade relativa a um conceito e é mais geral que os diferentes tipos de exercício (“application_exercise”, “analysis_exercise”, “synthesis_exercise”, “evaluation_exercise”), pois não possui necessariamente a conclusão de um objetivo. A dificuldade de uma “interactivity” é representada numa propriedade de mesmo nome. Em geral, o modo com a atividade do OA é realizada, por exemplo, uma questão múltipla escolha, é independente da sua função (ULLRICH, 2004).

Essa ontologia tem o foco no ensino e não na aprendizagem, embora esses processos não estejam dissociados. Dessa forma, apresenta conceitos importantes para a compreensão do objetivo de aprendizagem, pois dadas as funções do OA, são esperadas ações, ou objetivos, para a conclusão da aprendizagem. Na Figura 17 podem ser visualizadas classes dessa ontologia.

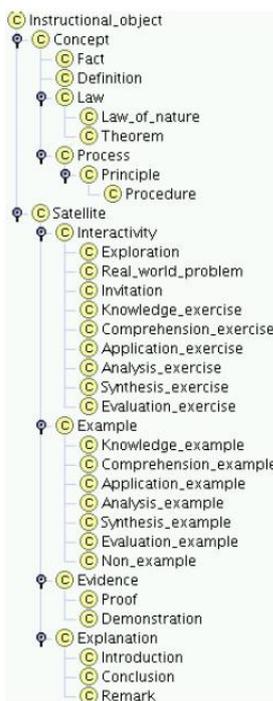


Figura 17: Ontologia de Carsten Ullrich (ULLRICH, 2004).

Lai Ng (NG, 2005), em sua dissertação de mestrado defendida na School of Interactive Arts and Technology da Simon Fraser University (Canadá), desenvolveu uma ontologia para ser aplicada na personalização da aprendizagem com OAs de acordo com o objetivo de aprendizagem. O domínio da ontologia é a representação descritiva do objetivo de aprendizagem e suas instâncias são os objetivos de aprendizagem em si. Para a autora,

objetivo de aprendizagem é uma declaração precisa de uma capacidade que pode ser observada como resultado da experiência educacional a partir de um desempenho pré-definido. Esse objetivo é composto por “comportamento para a aprendizagem”, “contexto da aprendizagem”, “público-alvo” e “nível de desempenho”.

Com a ontologia pretende-se responder “O que o aprendiz deve fazer?”, “Com quais condições ele deve fazer isso?” e “Como isso deve ser feito?”. Para isso, o objetivo de aprendizagem é caracterizado com conceitos das classes “descrição”, “desempenho” e “referência de conhecimento”. Um objetivo requer a discriminação, identificação e aplicação de conceitos definidos na classe “referência de conhecimento” que pode ser estendida com outra ontologia. A partir das classes, são definidas as propriedades “título”; “descrição” de características como, desempenho requerido, contexto, etc.; “referência de conhecimento”; e “metadado” (NG, 2005).

O domínio, a que se refere o conceito da “referência de conhecimento” pertence a classe “taxonomia”. OA é instância da classe “taxonomia” e da classe “objetivo de aprendizagem”, logo, OAs com objetivos comuns são referenciados na mesma classe. Um conceito pode ser composto por conceitos mais específicos que indicam a relação *todo-parte*, o sequenciamento de informação ou o estabelecimento de pré-requisito. Por analogia, a relação entre objetivos de aprendizagem que abordam conceitos pode indicar ordenação, hierarquia ou dependência entre objetivos de aprendizagem (NG, 2005).

Um objetivo de aprendizagem pode ser amplo, chamado de objetivo de aprendizagem terminal, e pode ser composto por vários outros objetivos de aprendizagem, chamados de objetivos de aprendizagem intermediários. Além disso, pode ter objetivos que constituem pré-requisitos. Assim, o objetivo de aprendizagem define um caminho para a aquisição de conhecimento sobre certo assunto e não pode ser expresso com uma sentença simples (NG, 2005).

Essa ontologia é direcionada para entendimento comum de aprendizes, desenhistas instrucionais, catalogadores, não apenas por apresentar e controlar definições, mas também por permitir a expansão com outras ontologias com propósitos similares. Suas classes e propriedades serão consideradas na proposição de ontologias para a modelagem da pesquisa.

Além das ontologias citadas, Aroyo e Dicheva (2002), Wilson (2004), Qin e Hernández (2004), Finneran (2002), Brase e Nejdí (2003) também apresentam ontologias no contexto educacional, porém genéricas, com enfoque no aluno, estrutura de aprendizagem, etc. Como as ontologias não atendem completamente a proposta da pesquisa, outros esquemas de representação de objetivos de aprendizagem são verificados.

7.2.2.2 Esquemas de representação de objetivos de aprendizagem

Foram pesquisados outros esquemas que representam objetivos de aprendizagem e oportunamente podem ser considerados na proposição de ontologias para este trabalho. Um esquema, com definições e relações estabelecidas, é capaz de direcionar a criação de classes e propriedades necessárias. Foram considerados esquemas mencionados nas pesquisas com objetivos similares, citadas anteriormente, e esquemas amplamente discutidos e utilizados no contexto dos OAs:

- Novak - pesquisa centrada na aprendizagem humana, estudos educacionais e representação de conhecimento. O autor desenvolveu uma teoria da educação para orientar a investigação e instrução, publicada pela primeira vez em 1977 na Universidade de Cornell (Estados Unidos). Seu mapeamento de conceitos representa o conhecimento de estudantes de ciência e foi usado como uma ferramenta para aumentar a aprendizagem significativa de acordo com o conhecimento prévio do aprendiz e seus objetivos de aprendizagem. O trabalho é baseado na teoria cognitiva de David Ausubel (NOVAK; GOWIN 1984);
- Gagné - teoria que estipula que existem vários tipos e níveis de aprendizagem e para cada um desses tipos e níveis há uma instrução adaptada para atender às necessidades do aluno. Suas cinco categorias de aprendizagem são *habilidades intelectuais* (criar competência individual e capacidade de responder aos estímulos), *estratégias cognitivas* (capacidade de aprender, pensar e lembrar), *informação verbal* (memorização de nomes, rostos, datas, números de telefone, etc.), *habilidades motoras* (capacidade de aprender a dirigir, andar de bicicleta, desenhar uma linha reta, etc.) e *atitudes* (pré-conceito em direção diferentes ideias, pessoas, situação). Cada categoria exige métodos diferentes para que a habilidade definida seja aprendida. Seus oito métodos para aprender são: aprendizagem por sinal, estímulo-resposta, encadeamento, associação verbal, aprendizagem por discriminação, aprendizagem por conceito, regra de aprendizagem, solução de problemas (GAGNÉ, 1985);
- Bloom - a taxonomia dos objetivos educacionais ou Taxonomia de Bloom (Bloom's Taxonomy) auxilia o planejamento, organização e gestão dos objetivos de aprendizagem. É uma terminologia conceitual baseada em classificações orientadas para definir algumas teorias de aprendizagem. Aborda domínios específicos de desenvolvimento cognitivo, afetivo e psicomotor, com objetivos agrupados em categorias dispostas em hierarquia

de complexidade e dependência. O Domínio Cognitivo está relacionado a aprender, dominar um conhecimento. Inclui o reconhecimento de fatos específicos, procedimentos padrões e conceitos que estimulam o desenvolvimento intelectual. As categorias desse domínio são *conhecimento, compreensão, aplicação, análise, síntese e avaliação*. O Domínio Afetivo relaciona sentimentos e posturas. Envolve categorias ligadas ao desenvolvimento da área emocional e afetiva, que incluem comportamento, atitude, responsabilidade, respeito, emoção e valores. As categorias desse domínio são *receptividade, resposta, valorização, organização e caracterização*. O Domínio Psicomotor é relacionado a habilidades físicas específicas. Bloom e sua equipe não chegaram a definir categorias para esse domínio. Outros autores definem categorias que incluem ideias ligadas a reflexos, percepção, habilidades físicas, movimentos aperfeiçoados e comunicação não verbal. As categorias desse domínio são *imitação, manipulação, articulação e naturalização*. Para ascender a uma nova categoria de qualquer Domínio é preciso ter obtido um desempenho adequado na anterior, pois utilizam capacidades adquiridas nos níveis anteriores (BLOOM; KARTHWOHL, 1984);

- Merrill - simplificação da taxonomia de Bloom (BLOOM; KARTHWOHL, 1984) que propõe uma relação de pré-requisito entre os componentes de conhecimento. Os componentes de conhecimento são *fatos, conceitos, procedimentos e princípios* pensados para serem adequados e precisos para descrever certos tipos de conteúdo cognitivo (MERRILL, 2000).

Dos esquemas apresentados, a Taxonomia de Bloom será utilizada e as considerações sobre sua estrutura serão abordadas oportunamente. A escolha se deu por atender diretamente ao propósito de definir e classificar objetivos de aprendizagem. Os outros esquemas não incidem diretamente sobre esse domínio, ainda que integrem os processos cognitivos que facultam o respectivo cumprimento de objetivos. Porém, não está excluída a consulta e utilização de conceitos abordados por outros esquemas. A taxonomia de Bloom (BLOOM; KARTHWOHL, 1984) é amplamente difundida no contexto de ensino-aprendizagem e seu uso não exigirá demais esforços de levantamento bibliográfico e análise de aplicação.

Ontologias com conceitos da área de saúde

Ontologias com conceitos da área de “saúde” são utilizadas na definição do assunto sobre o qual o objetivo de aprendizagem propõe um processo de ensino-aprendizagem. O assunto é o elemento mais utilizado para apoiar a busca de OAs, justamente por que descreve seu tema. A descrição temática de um documento representa seu conteúdo e a profundidade da abordagem. Para tanto, é possível utilizar informação extraída do próprio documento ou de instrumentos capazes de sintetizar o assunto.

A descrição temática fornece a síntese do conteúdo dos documentos e estabelece categorias para seu armazenamento (GUINCHAT; MENO, 1994). Na descrição temática de um documento com base na sua própria informação é utilizada a *linguagem natural* ou linguagem do discurso técnico-científico. A descrição com o uso de esquemas específicos utiliza a *linguagem controlada* ou vocabulário controlado que seleciona da *linguagem natural* um conjunto restrito de palavras e frases, tratadas semanticamente para estabelecer termos significativos (SVENONIOUS, 2000). A *linguagem controlada* tem o objetivo de evitar sinonímia, polissemia e homografia, enquanto produz entidades com afirmações sobre as propriedades que identificam um documento ou suas relações com outros documentos. Se a representação do assunto for com *linguagem natural*, sem restrições, o resultado das buscas pode ser pouco preciso, ambíguo e com ruídos, logo, a adoção de uma ontologia minimiza esses problemas.

As ontologias da área de saúde localizadas e disponíveis para *download* estão no idioma inglês, na linguagem OWL, Open Biomedical Ontologies (OBO)³⁸ ou Unified Medical Language System (UMLS)³⁹. OBO e UMLS são iniciativas de padronização para informação em saúde com linguagens próprias para o desenvolvimento de sistemas de informática. O OBO possui uma plataforma para registro de ontologias que evoluiu para biblioteca digital de ontologias médicas e biomédicas. A OBO dissemina a linguagem patenteada em que a maior parte das suas ontologias estão codificadas. O UMLS constitui um metatesauro com terminologias, tesouros, sistemas de classificação, e ontologias biomédicas em linguagem própria. Além de facilitar o acesso transparente às fontes de informação em saúde, o metatesauro UMLS estabelece ligações entre conceitos de diferentes VCs, categorizados por tipos semânticos da UMLS Semantic Network (FREITAS, F.; SCHULZ, S.; MORAES, E., 2009).

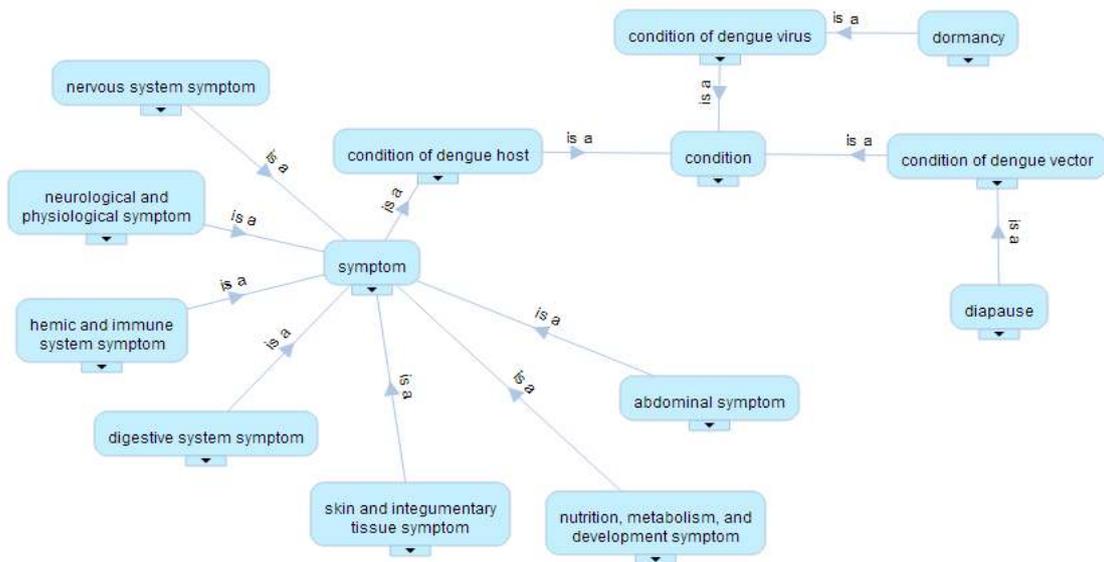
O domínio “saúde” é bastante amplo e, por esse motivo, buscaram-se por VCs que abordem conceitos tratados pelas áreas prioritárias do Ministério da Saúde do Brasil⁴⁰, tais

³⁸ *Link:* <<http://obofoundry.org/>>.

³⁹ *Link:* <<https://uts.nlm.nih.gov/home.html>>.

⁴⁰ *Link:* <<http://portal.saude.gov.br/portal/saude>>.

como doenças crônicas, atenção primária, prevenção de doenças e urgência e emergência. Não se pretende avaliar ou readequar uma ontologia do domínio da saúde, pois demandaria trabalho de especialistas. A ontologia escolhida tem caráter demonstrativo, contudo diversas ontologias encontradas podem ser utilizadas para futuras aplicações. Alguns exemplos de ontologias que podem ser adotadas são:



Dengue Fever Ontology⁴¹ - 4.904 classes e 51 propriedades sobre dengue.

Figura 18: Dengue Fever Ontology

Vaccine Ontology⁴² - 5.150 classes e 203 propriedades sobre vacinas e imunologia.

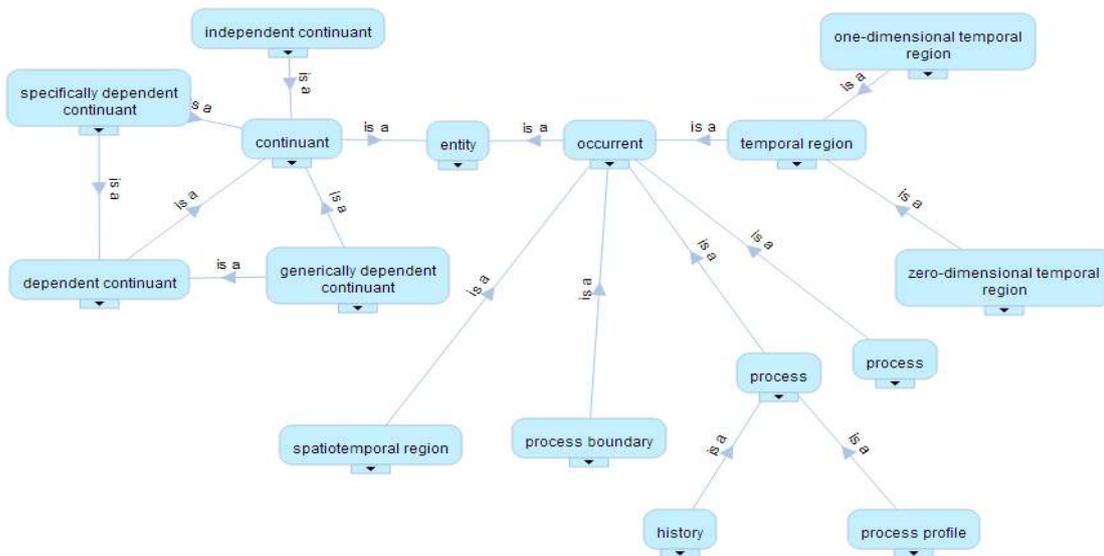


Figura 19: Vaccine Ontology

⁴¹ Link: <<http://bioportal.bioontology.org/ontologies/3174>>.

⁴² Link: <<http://bioportal.bioontology.org/ontologies/1172>>.

International Classification of Functioning, Disability and Health (ICF)⁴³ - 1.594 classes e 67 propriedades da Classificação Internacional de Funcionalidade, Incapacidade e Saúde.

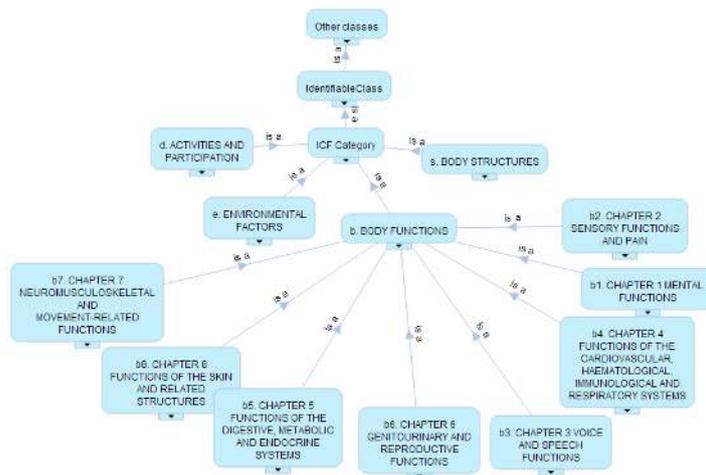


Figura 20: International Classification of Functioning, Disability and Health (ICF)

Esquemas de representação na área de saúde

A área de saúde é caracterizada por uma ampla terminologia constituída para disciplinar definições, recuperar documentos, registrar estatísticas de mortalidade e morbidade, ou faturamento de serviços de saúde. Essas terminologias definem os termos a partir da linguagem natural e compõem hierarquias que indicam termos preferenciais e relações entre conceitos para sua melhor compreensão e utilização.

O progresso dos esquemas de representação na área de saúde está relacionado ao estabelecimento de vocabulários e sistemas de classificação, como a Classificação Internacional de Doenças e o Index Medicus, impulsionados pelos interesses da saúde pública, epidemiologia, com apoio da Biblioteconomia; e pelas pesquisas de sistemas de suporte a tomada de decisões médicas, com apoio da Inteligência Artificial (FREITAS, F.; SCHULZ, S.; MORAES, E., 2009).

Nessa perspectiva, além das ontologias, vários outros esquemas definem conceitos desse domínio, tais como:

- North American Nursing Diagnosis Association (NANDA)⁴⁴: diagnósticos em enfermagem;

⁴³ Link: <<http://bioportal.bioontology.org/ontologies/1411>>.

- Nursing Outcomes Classification (CNC)⁴⁵: classificação de intervenções em enfermagem;
- International Classification of Diseases (ICD10)⁴⁶: classificação internacional de doenças e problemas relacionados à saúde que inclui sinais e sintomas, diagnósticos e procedimentos;
- Systematized NOMenclature of MEDicine (SNOMED)⁴⁷: nomenclatura sistematizada de humanos e medicina;
- Logical Observation Identifiers Names and Codes (LOINC)⁴⁸: base de dados com códigos, nomes e sinônimos de exames laboratoriais;
- Descritores em Ciências da Saúde (DeCS)⁴⁹: descritores baseados no Medical Subject Headings (MeSH);
- International Classification of Primary Care (ICPC)⁵⁰: sistema de classificação para os cuidados primários de saúde.

Além disso, a área de saúde possui padrões internacionais desenvolvidos para interoperabilidade de informação, como o Health Level Seven (HL7)⁵¹, um protocolo internacional para intercâmbio de dados eletrônicos em ambientes da área da saúde, que integra informação de natureza clínica e administrativa; e o Open Electronic Health Record (openEHR)⁵², uma plataforma que permite interoperabilidade entre sistemas de saúde. O TC 251⁵³ do Comitê Europeu de Normalização é o grupo de trabalho dentro da União Europeia para normalização de Tecnologia da Informação e Comunicação (TIC) da Saúde.

Ontologias com conceitos de público-alvo dos OAs na área de saúde

Ontologias que tratam o “público-alvo dos OAs na área de saúde” devem representar indivíduos que potencialmente vão aprender com os OAs, ou seja, são o foco do objetivo de aprendizagem. Esses indivíduos têm características generalistas, tal como “pessoas” ou “aprendizes”; e também têm características específicas, como indivíduos com necessidades de informação na área de saúde. Assim, faz-se a delimitação do público-alvo ao grupo de aprendizes da área de saúde. Uma vez que o conhecimento especializado na área de saúde

⁴⁴ *Link:* <<http://www.nanda.org/>>.

⁴⁵ *Link:* <<http://www.nursing.uiowa.edu/cncce/nursing-outcomes-classification-overview>>.

⁴⁶ *Link:* <<http://apps.who.int/classifications/icd10/browse/2010/en>>.

⁴⁷ *Link:* <<http://www.ihtsdo.org/snomed-ct/>>.

⁴⁸ *Link:* <<http://loinc.org/>>.

⁴⁹ *Link:* <<http://decs.bvs.br/>>.

⁵⁰ *Link:* <<http://icpc2.danielpinto.net/K/1/>>.

⁵¹ *Link:* <<http://www.hl7brazil.org/>>.

⁵² <http://openehr.codeplex.com/>

⁵³ *Link:* <<http://www.cen.eu/CEN/Sectors/TechnicalCommitteesWorkshops/CENTechnicalCommittees/Pages/Standards.aspx?param=6232&title=CEN/TC+251>>.

na estrutura de ensino brasileira é abordado a partir do nível técnico, esse aprendiz pode estar em fase de formação, especialização ou capacitação. A partir da conclusão da formação ele também pode ser um profissional na área de saúde.

Diante disso, foram buscadas ontologias para representar características gerais do aprendiz, com o intuito de posicionar o público-alvo no processo de ensino-aprendizagem; e ontologias que representem a formação profissional na área de saúde. A formação profissional, ainda que em andamento, tem relação com assuntos de interesse a serem buscados e com o processo cognitivo demandado pelo objetivo de aprendizagem.

Algumas ontologias abordam amplamente o contexto de ensino-aprendizagem, mas não aprofundam aspectos que qualificam o público-alvo. Como exemplo, a Higher Education Reference Ontology que é ontologia de referência, com cobertura do domínio “universidade”, abordando aspectos da sua estrutura organizacional, administração, pessoal, etc. (Figura 21). Seu objetivo é descrever qualquer universidade, dado seu contexto de ensino-aprendizagem (HERO, 2012).

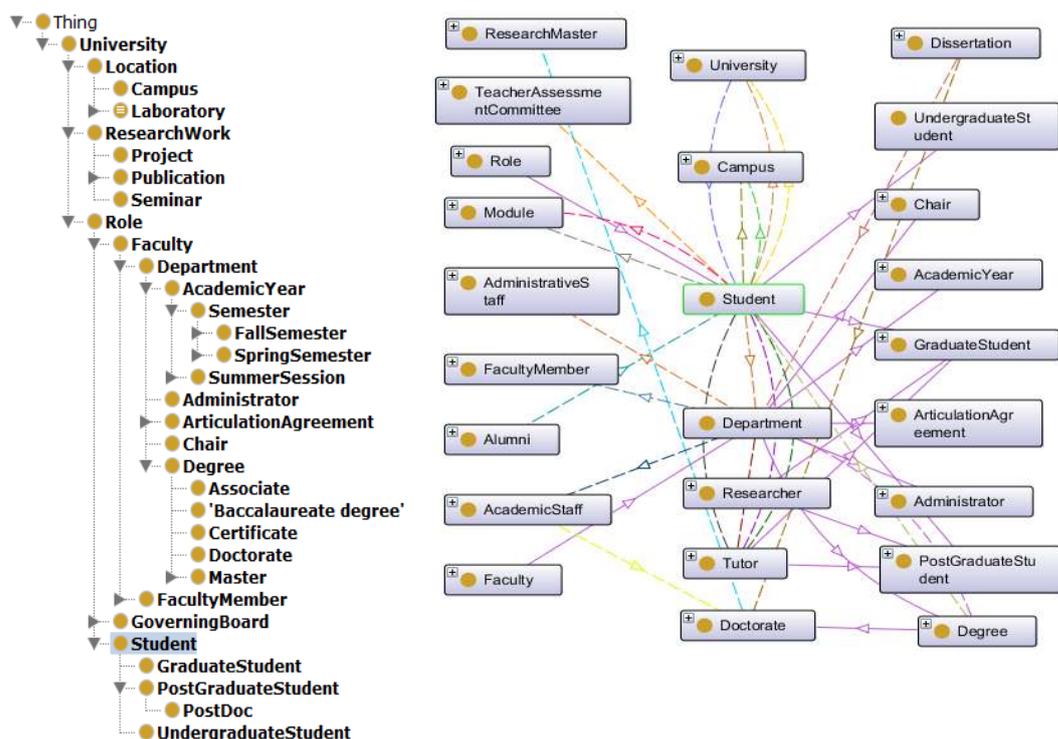


Figura 21: Higher Education Reference Ontology

Fonte: (HERO, 2012)

A Cognitive Characteristics Ontology 0.2 aborda diferentes características cognitivas que um indivíduo possui para permitir a personalização da aprendizagem. Apresenta

conceitos que constituem perfis de usuários com dados, características pessoais e preferências. A ontologia apresenta duas opções para modelar padrões cognitivos. A primeira é a representação de características cognitivas dos usuários relacionadas semanticamente aos padrões cognitivos. A segunda é a definição de características cognitivas de um contexto para descrever padrões cognitivos para um usuário específico ou grupo de usuários. Na categoria "CognitiveCharacteristic" as propriedades que expressam padrões cognitivos são interesse (cco:interest), que indica um certo interesse ou preferência, com valores especificados em assuntos de interesse (foaf:topic_interest); competência (cco:competence), que indica capacidade para fazer algo ou saber algo (cco:setting) relacionado a um ambiente específico, uma situação ou aplicação (BRICKLEY; *et. al*, 2010).

Competência (cco:competence) possui propriedades específicas como habilidade para fazer alguma coisa, por exemplo, andar, tocar piano ou trabalhar em equipe; conhecimento (cco:skill) ou experiência (cco:expertise) sobre um certo domínio ou de um tema específico, por exemplo, futebol, linguagens de programação ou música; convicção (cco:belief) para a representação de competência, o que significa crenças ou opiniões, que também pode ser equívocos. Tudo isso de acordo com as cinco categorias de aprendizagem de Gagné (GAGNÉ, 1985), citadas anteriormente (BRICKLEY; *et. al*, 2010).

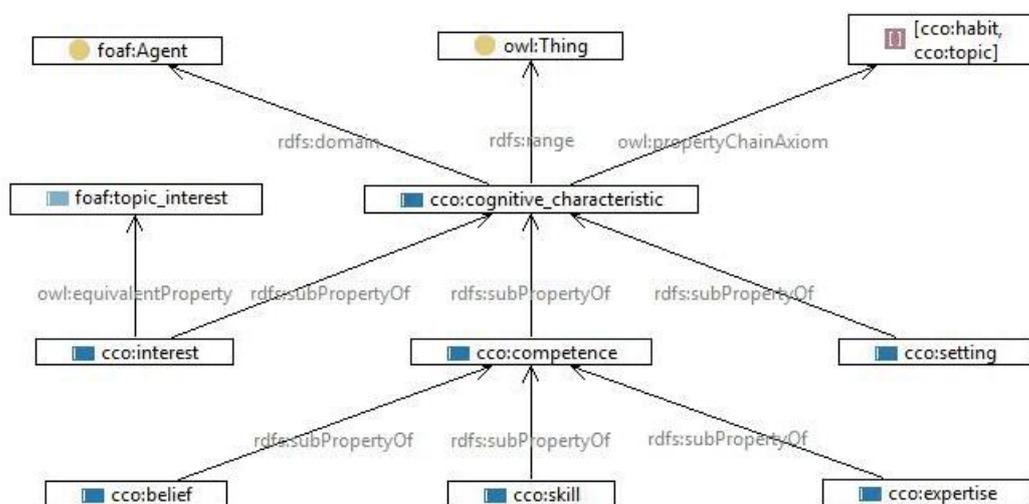


Figura 22: Propriedades da The Cognitive Characteristics Ontology 0.2

Fonte: (BRICKLEY; *et. al*, 2010).

Holanda; *et. al* (2007) propõe ontologia para ambientes interativos de aprendizagem que aborda conceitos do domínio "estudante" e "colaboração pedagógica". "Estudante" relaciona termos sobre quem será ensinado, tais como, informação estática integrada a definição do perfil do estudante (proficiência, estilo de aprendizagem, informação pessoal); e informação dinâmica que é alterada a partir da interação do estudante com o ambiente de

aprendizagem (informação cognitiva, afetiva, meta-cognitiva, motivacional) (HOLANDA; et al, 2007).

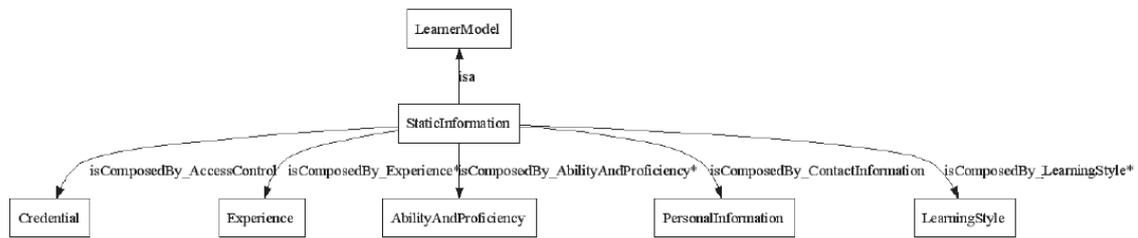


Figura 23: Ontologia para ambientes interativos de aprendizagem

Fonte: (HOLANDA; et al, 2007).

Para caracterizar o indivíduo formando ou formado em alguma especialidade na área de saúde levou-se em consideração as profissões relacionadas às Ciências da Saúde no Brasil, de acordo com a Resolução nº 287 de 08 de outubro de 1998, do Conselho Nacional de Saúde: médicos, enfermeiros, fisioterapeutas, nutricionistas, profissionais de educação física, serviço social, fonoaudiólogos, dentistas, terapeutas ocupacionais, psicólogos, biomédicos, farmacêuticos, veterinário, biólogo e assistente social, além das especialidades técnicas em enfermagem, saneamento, nutrição e dietética, odontologia, prótese, ortoprotésico, laboratório de análises clínicas, administração hospitalar, optometria, ortóptica, radiologia médica, higiene dental, patologia clínica e histologia, reabilitação, regulação em saúde. De forma que uma ontologia com tais termos atenda a caracterização da atuação profissional, área de formação ou capacitação do público-alvo para o qual o OA foi elaborado.

Não foram encontradas ontologias específicas nesse domínio, contudo, ontologias bastante amplas, que agregam mais de um VC na área de saúde, possuem classes e instâncias que abordam as especialidades na área de saúde. Como exemplo, a ontologia do Health Level Seven que agrega a taxonomia do Health Insurance Portability and Accountability (HIPAA) e apresenta especialidades em saúde reguladas para planos de saúde nos Estados Unidos.

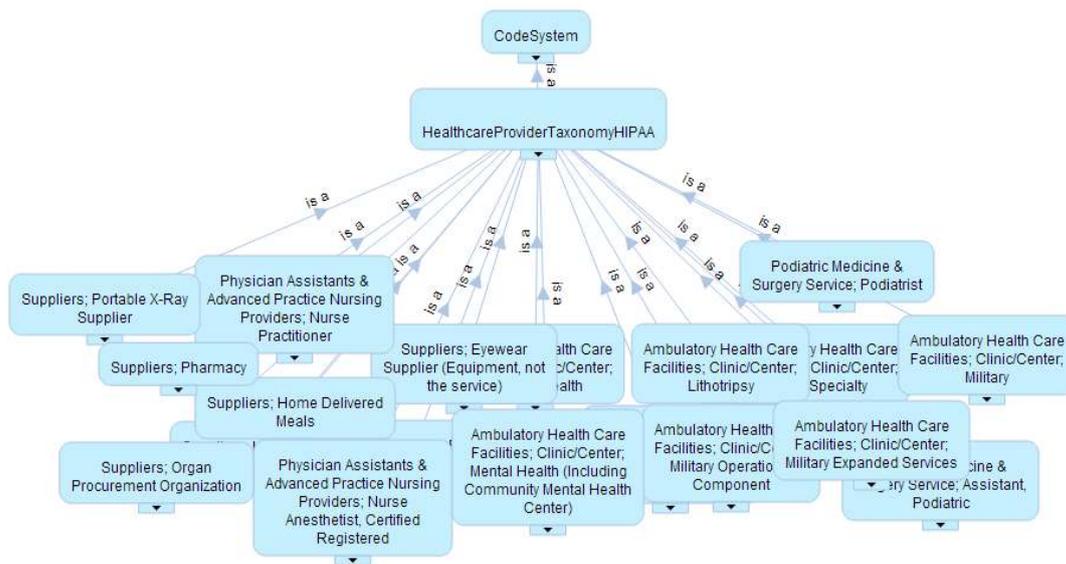


Figura 24: Health Insurance Portability and Accountability.

Esquemas de representação para público-alvo dos OAs na área de saúde

Foram encontradas taxonomias e classificações que conferem estrutura, definição e legitimidade a atuação profissional na área de saúde. Esses esquemas são validados em âmbito nacional e não podem ser diretamente aplicados ou estendidos a qualquer contexto. Uma vez reconhecidas atuações profissionais, podem ser estabelecidas competências e habilidades a serem observadas na definição de objetivos de aprendizagem.

Como exemplo, o norte-americano Health Care Provider Taxonomy (HCPT)⁵⁴ que codifica especialidades em saúde (Figura 25).

⁵⁴ Link: <http://www.nucc.org/index.php?option=com_wrapper&view=wrapper&Itemid=126>.

<ul style="list-style-type: none"> • Ophthalmology - 207W00000X [definition] • Oral & Maxillofacial Surgery - 204E00000X [definition] ⊕ Orthopaedic Surgery - 207X00000X [definition] <ul style="list-style-type: none"> • Adult Reconstructive Orthopaedic Surgery - 207XS0114X [definition] • Foot and Ankle Surgery - 207XX0004X [definition] • Hand Surgery - 207XS0106X [definition] • Orthopaedic Surgery of the Spine - 207XS0117X [definition] • Orthopaedic Trauma - 207XX0801X [definition] • Pediatric Orthopaedic Surgery - 207XP3100X [definition] • Sports Medicine - 207XX0005X [definition] ⊕ Otolaryngology - 207Y00000X [definition] <ul style="list-style-type: none"> • Facial Plastic Surgery - 207YS0123X [definition] • Otolaryngic Allergy - 207YX0602X [definition] • Otolaryngology/Facial Plastic Surgery - 207YX0905X [definition] • Otolaryngology & Neurotology - 207YX0901X [definition] • Pediatric Otolaryngology - 207YP0228X [definition] • Plastic Surgery within the Head & Neck - 207YX0007X [definition] • Sleep Medicine - 207YS0012X [definition] ⊕ Pathology - <ul style="list-style-type: none"> • Anatomic Pathology - 207ZP0101X [definition] • Anatomic Pathology & Clinical Pathology - 207ZP0102X [definition] • Blood Banking & Transfusion Medicine - 207ZB0001X [definition] • Chemical Pathology - 207ZP0104X [definition] • Clinical Pathology - 207ZC0006X [definition] • Clinical Pathology/Laboratory Medicine - 207ZP0105X [definition] • Cytopathology - 207ZC0500X [definition] • Dermatopathology - 207ZD0900X [definition] • Forensic Pathology - 207ZF0201X [definition] • Hematology - 207ZH0000X [definition] 	<p>Status Active</p> <p>Code 207XS0114X</p> <p>Type Level III Area of Specialization</p> <p>Adult Reconstructive Orthopaedic Surgery Recognized by several state medical boards as a fellowship subspecialty program of orthopaedic surgery, adult reconstructive orthopaedic surgeons deal with reconstructive procedures such as joint arthroplasty (i.e., hip and knee), osteotomy, arthroscopy, soft-tissue reconstruction, and a variety of other adult reconstructive surgical procedures.</p> <p><i>Source: American Board of Medical Specialties, 2007. www.abms.org [7/1/2007: definiton added, source added]</i></p> <p>Additional Resources: American Board of Orthopaedic Surgery, 2007. http://www.abos.org/.</p> <p>Separate board certification is not currently offered.</p>
---	---

Figura 25: Health Care Provider Taxonomy.

No Brasil, a Classificação Brasileira de Ocupações (CBO)⁵⁵, elaborada pelo Ministério do Trabalho e Emprego do Brasil, tem a finalidade de identificar as ocupações no mercado de trabalho para fins classificatórios junto aos registros administrativos e domiciliares (Figura 26). A CBO é o documento normalizador do reconhecimento, da nomeação e da codificação dos títulos e conteúdos das ocupações do mercado de trabalho brasileiro. É ao mesmo tempo uma classificação enumerativa, que codifica empregos e outras situações de trabalho (inclui códigos, títulos ocupacionais e descrição sumária); e uma classificação descritiva, que detalha as atividades realizadas no trabalho, os requisitos de formação e experiência profissionais e as condições de trabalho (BRASIL MTE, 2012).

⁵⁵ Link: <<http://www.mteco.gov.br/cbsite/pages/pesquisas/BuscaPorTituloResultado.jsf>>.

2251 :: Médicos clínicos

Títulos
2251-03 - Médico infectologista Infectologista, Médico de doenças infecciosas e parasitárias
2251-05 - Médico acupunturista
2251-06 - Médico legista Legista
2251-09 - Médico nefrologista Nefrologista
2251-10 - Médico alergista e imunologista Alergista, Alergista e imunologista, Alergologista, Médico alergista, Médico imunologista
2251-12 - Médico neurologista Médico neuropediatra, Neurologista, Neuropediatra
2251-15 - Médico angiologista Angiologista

Figura 26: Classificação Brasileira de Ocupações – Família Médicos Clínicos.

As ontologias observadas nesta seção serão avaliadas quanto à equivalência e adequação dos termos, ou ainda com relação adoção das definições apresentadas para os conceitos que interessam a essa modelagem. A seção seguinte trata da proposição de ontologias para atender o objetivo dessa pesquisa. Quando as ontologias citadas forem consultadas e utilizadas serão feitas as considerações necessárias.

7.2.2.3 Proposição de ontologias

Nesta etapa discorre-se sobre a proposta de desenvolvimento de ontologias ou readequação das ontologias apresentadas na Seção 7.2.2.1. Espera-se que as ontologias venham contribuir com o refinamento dos resultados de busca de OAs em RDEs na área de saúde. Na etapa anterior, além de ontologias, foram discutidos esquemas de representação dos domínios de interesse para pesquisa que serão considerados no levantamento de conceitos e propriedades.

Property	Cardinality	Type
e_para	Multiple	Publico-alvo
e_sobre	Multiple	Assunto
especificacao	Multiple	string
identificador	Single	string
objetivo_de_aprendizagem_posterior	Multiple	Objetivo_de_aprendizagem
objetivo_de_aprendizagem_previo	Multiple	Objetivo_de_aprendizagem
objetivo_de_aprendizagem_requerido	Multiple	Objetivo_de_aprendizagem
processo_cognitivo	Multiple	Processo_cognitivo

Figura 28: Propriedades da classe “Objetivo de aprendizagem”.

A ontologia proposta por Carsten Ullrich (ULLRICH, 2004) futuramente poderá ser integrada a modelagem proposta pela pesquisa, com a abordagem de conceitos que definem o OA, indicando suas funções. Isso permitirá uma visão geral do OA e a correlação entre sua função educacional e o objetivo de aprendizagem para o qual foi pensado. Neste momento, essa possibilidade não contribui com o objetivo do trabalho.

Ontologia de saúde

Anderson (2001), aluno de Bloom que revisa sua taxonomia de objetivos de aprendizagem, destaca que, além do processo cognitivo, o ensino-aprendizagem se desenvolve a cerca de “*algo ou alguma coisa*” que é “*o que saber*”. Isso é chamado de “dimensão do conhecimento”, com a classificação da abordagem do conhecimento como *factual, conceitual, procedimental e meta-cognitivo*. O conhecimento factual inclui elementos isolados de informação, como definições de vocabulário e conhecimento de detalhes específicos. O conhecimento conceitual consiste em sistemas de informação, como classificações e categorias. O conhecimento procedimental, ou “saber como fazer”, inclui algoritmos, heurística ou método empírico, técnicas e métodos, bem como o conhecimento sobre quando usar esses procedimentos. O conhecimento meta-cognitivo, ou reflexão sobre o que se sabe, refere-se ao conhecimento dos processos cognitivos e da informação sobre como manipular esses processos de forma eficaz.

Não se pode negar a forte relação entre o objetivo de aprendizagem e o assunto ou domínio do conteúdo abordado pelo OA. Domínios de conteúdo representam as áreas de conhecimento como, matemática, informática, medicina, economia, direito etc. Todo domínio de conteúdo tem sua própria estruturação de conhecimento, o que torna bastante trabalhoso definir um modelo conceitual para cada domínio. O domínio que especifica o assunto a que se refere o objetivo de aprendizagem na modelagem proposta será classificado de acordo

com a dimensão do conhecimento (ANDERSON, 2001) e a hierarquia de termos ou descritores será baseada em uma ontologia da área da saúde.

O “Assunto” nessa modelagem é uma classe, ou conceito, e o VC que o representa é uma propriedade com valores estabelecidos pela ontologia da área de saúde agregada. “Assunto” também possui a propriedade “texto livre” que pode ser preenchida por qualquer valor alfanumérico que indique o assunto a que se refere o objetivo de aprendizagem em linguagem natural.

A ontologia escolhida tem o propósito de demonstrar como deve ser a agregação de uma ontologia de domínio a modelagem proposta para estender e controlar a representação do assunto. Quaisquer ontologias podem ser utilizadas, desde que sejam desenvolvidas em linguagem apropriada, neste caso a OWL, e atendam a conceitualização esperada. Ademais, como o assunto é uma descrição temática, também podem ser utilizadas na modelagem proposta taxonomia e outros esquemas de representação, ainda que demandem esforços para formalização dos conceitos de acordo com as tecnologias da WS.

Para facilitar o manuseio e verificação do cumprimento do objetivo da pesquisa foi agregada a modelagem a ontologia International Classification of Functioning, Disability and Health (ICF Ontology). Diversos fatores motivaram essa escolha, tais como:

- Está em constante desenvolvimento e possui documentação ampla sobre as decisões tomadas;
- É desenvolvida sob a responsabilidade de uma entidade reconhecida na área de saúde, a Family of International Classifications da World Health Organization (WHO)⁵⁶;
- É um VC reconhecido e utilizado em contextos da área de saúde, por isso apresenta termos e definições explícitas;
- Está disponível para *download* na linguagem OWL no BioPortal e apresenta termos mapeados de acordo com conceitos utilizados em outros VCs na área de saúde, como Computer Retrieval of Information on Scientific Projects (CRISP Thesaurus), Logical Observation Identifier Names and Codes, Medical Subject Headings (MeSH), National Cancer Institute (NCI Thesaurus), Radiology terms for practice, education, and research (RadLex) e SNOMED Clinical Terms (BIOPORTAL, 2012);
- Possui importação facilitada a partir da opção *import classes of BioPortal* da ferramenta Protégé;
- Atende as áreas prioritárias do Ministério da Saúde do Brasil, pois contempla conceitos a cerca da prevenção de doenças de uma população específica,

⁵⁶ Link: <<http://www.who.int/classifications/icf/en/>>.

responsabilidade do Sistema Único de Saúde do Brasil (SUS), estabelecida pela Política Nacional de Saúde da Pessoa Portadora de Deficiência (BRASIL, 2008), com propósito de reabilitar a pessoa portadora de deficiência na sua capacidade funcional e no seu desempenho humano, proteger a saúde, bem como prevenir agravos que determinem o aparecimento de deficiências.

A ICF Ontology possui 1.594 classes, com até 8 níveis hierárquicos e 819 definições. Além de 67 propriedades, 3.250 instâncias. É uma ontologia agregada a diversas outras ontologias que realizam a expansão dos seus conceitos e pode ser visualizada em partes separadas, como ICF-d1 Learning and applying knowledge, Icf-d5 Self-care, Icf-d810-d839 Education e Icf-d840-d859 Work and employment. O projeto OntoCat⁵⁷ utiliza a ICF Ontology para realizar inferências baseadas em ontologias e disponibilizar *webservices* para reutilização e integração de termos em aplicações Web na área de saúde. O projeto OntoMaton⁵⁸ também utiliza a ICF Ontology para disponibilizar aplicativo para o Google que permite buscas baseadas em ontologia e marcação semântica de documentos diretamente no Google Docs (BIOPORTAL, 2012).

A Classificação Internacional de Funcionalidade, Incapacidade e Saúde (CIF) é uma classificação de domínios da área de saúde e suas relações. Estes domínios são classificados a partir de perspectivas do corpo, individuais e sociais, por meio da lista de funções e estrutura do corpo (classes “b. BODY FUNCTIONS” e “s. BODY STRUCTURES”), e da lista de atividade e participação (classe “d. ACTIVITIES AND PARTICIPATION”). A partir do funcionamento do indivíduo e da deficiência que ocorre em um contexto, também dispõe lista de fatores ambientais (classe “e. ENVIRONMENTAL FACTORS”) (WHO, 2012). Na CIF, o termo funcionalidade refere-se a funções do corpo, atividades e participação, enquanto o termo incapacidade refere-se a deficiências, limitações e atividades com restrição de participação.

As relações semânticas deste domínio ajudam a descrever mudanças na função e estrutura corporal, o que a pessoa com certa condição de saúde pode fazer num ambiente normal (seu nível de capacidade), bem como, o que ela realmente faz em seu ambiente habitual (seu nível de desempenho) (classe “ICF Qualifier” e subclasses “Capacity” e “Performance”). Essas subclasses possuem valores pré-definidos que constituem propriedades das classes “b. BODY FUNCTIONS” e “s. BODY STRUCTURES”.

⁵⁷ Link: <<http://www.ontocat.org/>>.

⁵⁸ Link: <<http://isa-tools.org/>>.

área de saúde, como o DeCs. A opção de utilizar outros esquemas de representação requer ações para disponibilizá-los em linguagem apropriada para a modelagem proposta, ou seja, OWL. Isso pode ser feito de maneira automática ou semiautomática, dependendo do formato de disponibilização do esquema. Nesta pesquisa não foram dispendidos esforços nesse sentido devido ao fato da agregação da ontologia que representa o assunto a que se refere o objetivo de aprendizagem ter caráter demonstrativo das relações semânticas existentes, conforme a modelagem proposta. A implementação dessa modelagem requer medidas para contornar questões de idioma, especialmente se isso constitui um problema para os usuários.

Ontologia de público-alvo dos OAs na área de saúde

O público-alvo é delimitado pelo aprendiz para o qual o objetivo de aprendizagem dos OAs na área de saúde é elaborado. A partir dessa afirmação, a ontologia deve tipificar o público-alvo, de acordo com sua área de interesse específica, e fornecer informação sobre seu perfil. Isso contribui com o agrupamento do público-alvo de acordo com características comuns para posterior refinamento de resultados de busca de OAs.

Brickley, *et. al* (2010) e Holanda, *et. al* (2007) auxiliaram na identificação de propriedades que compõem o perfil do público-alvo. Assim, foi inserida a propriedade “assunto de interesse”, com valores da classe “ICF ontology”, que permite definir assuntos mais genéricos ou mais específicos para expandir os resultados de busca. A propriedade “nível educacional” teve valores pré-definidos de acordo com a estrutura de formação educacional na área de saúde (Graduação, Mestrado, Doutorado, Residência médica, etc.). Estilos de aprendizagem, como propõe Holanda, *et. al* (2007), não foram estabelecidos por ser uma informação muito pessoal, não adequada a generalizações de público-alvo. A propriedade “contexto educacional”, que especifica o ambiente em que o público-alvo deve estar inserido, permanece com valor de preenchimento alfanumérico.

O tipo de público-alvo foi definido de acordo as profissões relacionadas às Ciências da Saúde no Brasil, citadas anteriormente. Para isso recorreu-se a Classificação Brasileira de Ocupações (CBO). Foi criada a propriedade “formação profissional” cujos valores são determinados pelas subclasses da classe “CBO”. A formação do público-alvo (indicada pelo seu campo profissional) é aquela para qual o objetivo de aprendizagem é proposto, assim como o nível educacional, ou seja, não são pré-requisitos. “Pré-requisito” é uma propriedade do objetivo de aprendizagem e pode ser preenchida por nível educacional, objetivo de aprendizagem ou um valor alfanumérico.

Para a CBO, ocupação é um conceito artificial, o que existe de concreto são as atividades desempenhadas pela pessoa em um emprego ou outro tipo de relação de trabalho. Ocupação é a agregação de empregos ou situações de trabalho similares quanto às atividades realizadas. Emprego ou situação de trabalho é um conjunto de atividades desempenhadas, para as quais são mobilizadas competências, que podem ser adquiridas em processo de ensino-aprendizagem. O conceito de competência tem duas dimensões: 1) Nível de competência: função da complexidade, amplitude e responsabilidade das atividades desenvolvidas; e 2) Domínio da competência: características do contexto do trabalho como área de conhecimento, função, etc., que identificam o tipo de profissão ou ocupação (BRASIL MTE, 2012).

A estrutura proposta agrega os empregos por habilidades cognitivas comuns exigidas no exercício de um campo de trabalho, composto por um conjunto de empregos similares que vai se constituir em um campo profissional do domínio. Assim, um campo profissional agrega profissionais diferentes, mas com habilidades cognitivas e interesses comuns, identificados por processos, funções ou ramos de atividades. Os campos profissionais são denominados Subgrupos e, assim como a ocupação, são categorias sintéticas, construtos. Família é a unidade do sistema de classificação e representa empregos ou conjunto de atividades. O Índice Analítico indica o código e o nome do Subgrupo de uma Família (BRASIL MTE, 2012).

Dessa forma, a classe “CBO” possui subclasses que representam os campos profissionais, tais como “Profissionais da educação física” e “Profissionais da medicina”; as famílias, tais como “Médicos clínicos”, “Médicos em especialidades cirúrgicas” e “Médicos em medicina diagnóstica e terapêutica” (Famílias definidas no Subgrupo Profissionais da medicina), que permite categorizar o público-alvo segundo habilidades cognitivas e interesses comuns; e ocupações, tais como “Médico em medicina intensiva”, “Médico psiquiatra” e “Médico da estratégia de saúde da família” (Empregos definidos nas Famílias), que agrupa o público-alvo de acordo com a similaridade das atividades desempenhadas. Os sinônimos das ocupações são representados pela propriedade “sinônimo”.

Consideração feita pela CBO é que os Grandes Grupos são agregados por nível de competência e similaridade nas atividades executadas, porém, por falta de indicador mais adequado, usou como nível de competência a escolaridade. Assim, são divididos os profissionais de nível superior (Profissionais das ciências e das artes, por exemplo) dos profissionais de nível técnico (Técnicos de nível médio, por exemplo). A compreensão mais atualizada do conceito de "competência" requer a classificação da complexidade das atividades exercidas e não simplesmente do nível de escolaridade (BRASIL MTE, 2012).

formalizada, com o uso de tecnologias da WS adequadas, a semântica dos conceitos utilizados. Como o objetivo da pesquisa é refinar os resultados de busca de OAs em RDEs na área de saúde, na seção seguinte são identificados e discutidos os metadados que descrevem os conceitos que especificam o objetivo de aprendizagem, modelados anteriormente.

7.2.3 Metadados que descrevem os conceitos que especificam o objetivo de aprendizagem

Metadados disponibilizados em formulários para a descrição de OAs em RDEs devem descrever os conceitos que especificam o objetivo de aprendizagem para que o resultado de busca possa ser refinado de acordo com a proposta dessa pesquisa. Para tanto, foram verificados os dois principais padrões utilizados em RDEs, DC e LOM, e identificados metadados adequados. Após essa etapa, foi realizada a correlação entre os dois padrões para que a modelagem possa ser utilizada em RDEs com ambos. Por fim, são feitas considerações sobre a integração dos metadados a modelagem da pesquisa.

Os metadados são fundamentais para a busca em RDEs e a adoção de um padrão implica na definição de conceitos que descrevem os OAs. O preenchimento dos metadados é a informação a ser indexada e recuperada posteriormente, logo, o uso de VCs minimiza ambiguidades no resultado da busca. Com o conjunto de metadados relacionados semanticamente com a modelagem será possível fazer inferências para expandir os resultados de busca de acordo com conceitos que especificam o objetivo de aprendizagem.

Os conceitos que especificam o objetivo de aprendizagem, definidos na modelagem proposta, são especificados no quadro a seguir (Quadro 2):

Quadro 2: Conceitos que especificam o objetivo de aprendizagem

Conceito	Propriedade / Relação	Valor
● Objetivo de aprendizagem	■ é para	● Público-alvo
	■ é sobre	● Assunto
	■ processo cognitivo	● Processo cognitivo
	■ especificação	◆ <i>String</i>
	■ pré-requisito	◆ <i>String</i>
	■ objetivo de aprendizagem requerido	● Objetivo de aprendizagem
	■ objetivo de aprendizagem prévio	● Objetivo de aprendizagem
	■ objetivo de aprendizagem posterior	● Objetivo de aprendizagem
● Assunto	■ classificação	● Dimensão do conhecimento
	■ taxonomia	● ICF Ontology
	texto livre	◆ <i>String</i>
● Público-alvo	■ formação profissional	● CBO
	■ nível educacional	◆ <i>String</i>
	■ contexto educacional	◆ <i>String</i>
	■ assunto de interesse	● ICF Ontology
	■ atividades	● Competências
● Processo cognitivo	■ especificação	◆ <i>String</i>
● CBO	■ sinônimo	◆ <i>String</i>
● ICF Ontology	-	-
● Dimensão do conhecimento	-	-

Legenda: ● Conceito: classe e subclasse. ■ Propriedade Objeto: estabelece relações entre conceitos. ■ Propriedade RDF: define características sobre conceitos.

Os metadados adequados à modelagem devem ter definições que coincidem com os conceitos apresentados (como por exemplo, “Assunto” e “Público-alvo”). Assim, são relacionados de acordo com as propriedades (como por exemplo, “é sobre” e “é para”) e os valores de preenchimento são pré-definidos com a adoção de ontologias que expandem os conceitos (como por exemplo, “ICF Ontology” e “CBO”) ou são valores de texto livre, ou *Strings* (como por exemplo, “texto livre” e “contexto

educacional”). Os valores de texto livre, para esta pesquisa, não constituem insumos para o refinamento dos resultados de busca.

O objetivo de aprendizagem é uma informação descritiva, registrada nos metadados. As propriedades desse conceito constituem a sua redação com o controle de alguns aspectos (“Assunto” e “Público-alvo”, por exemplo), que por sua vez, possuem semântica formalizada pelos respectivos VCs adotados. Na descrição de um OA no RDE, ainda que o preenchimento do metadado “objetivo de aprendizagem” não faça menção aos conceitos relacionados e as propriedades citadas, a implementação da modelagem permitirá o estabelecimento e formalização semântica dessa informação. Necessariamente, o conjunto de metadados do RDE deve apresentar esses conceitos e VCs para descrição dos OAs.

Para a WS, metadados são a representação, processada por aplicações de computadores e compreendida por pessoas, de quaisquer características de um conteúdo a ser disponibilizado. A representação requer um rótulo, que indica ou sinaliza o que se pretende informar sobre o conteúdo; um dado, que é a informação em si ou uma referência ao conteúdo; e um padrão lógico que irá fornecer o significado e a formalização necessária para o metadado. Assim, alguns campos do formulário de metadados irão apresentar opções de dados para preenchimento, pré-definidos de acordo com ontologias e esquemas de representação para que a semântica possa ser garantida na realização de inferências para o refinamento do resultado de busca.

Como observado, os metadados são usados na caracterização e interligação dos dados, e para ampliar a compreensão de sua semântica são introduzidas ontologias. A modelagem conceitua o objetivo de aprendizagem no contexto dos OAs na área de saúde e as ontologias acrescentam regras de inferência estabelecendo a semântica dos dados que descrevem outros dados. Por exemplo, na descrição de um OA:

O objetivo de aprendizagem que requer “Comparar”⁵⁹ a estrutura dos lobos corticais com a estrutura do mesencéfalo, é sobre o assunto “Estrutura do cérebro”,

⁵⁹ “Comparar” advém do processo cognitivo “Entender”, que está relacionado a estabelecer uma conexão entre o novo e o conhecimento previamente adquirido de acordo com Anderson (2001).

um tipo de “Estrutura do corpo”⁶⁰, classificado como “Conhecimento conceitual” de “Teorias, modelos e estruturas”⁶¹.

O dado “Estrutura do cérebro” é um assunto específico da categoria “Estrutura específica do sistema nervoso”, que permite inferir sua semelhança (relação *parte-todo*) com outras estruturas que compõem o sistema nervoso como, “Medula espinhal e estruturas relacionadas”, “Estrutura de meninges”, “Estrutura do sistema nervoso simpático” e “Estrutura do sistema nervoso parassimpático”, de acordo com a ICP Ontology. O objetivo de aprendizagem é sobre (propriedade) um Assunto (classe), que possui *classificação* (propriedade) de acordo com a *Dimensão do conhecimento* (classe) e valor especificado pela *taxonomia* (propriedade) ICP Ontology (classe) ou *texto livre* (propriedade).

Ao avançar as possibilidades de inferência nesse contexto, outro OA com objetivo de aprendizagem que requer “Comparar” e é sobre um assunto classificado como “Conhecimento conceitual” de “Teorias, modelos e estruturas” pode ser relacionado ao primeiro e compor resultado de busca satisfatório, especialmente se for considerada a hierarquia do assunto.

O uso de metadados nos REDs na área de saúde consolida a descrição do conteúdo e a organização da informação, ou seja, são metadados descritivos. Esses são associados aos OAs e, ao representarem conceitos inerentes ao objetivo de aprendizagem, possibilitam que o resultado de busca considere inferências e seja refinado automaticamente. Os conceitos da modelagem qualificam os metadados, formalizam a semântica, delimitam a interpretação e especificam regras para o preenchimento. A seção seguinte (Seção 7.2.3.1) identifica quais metadados adequados à representação desses conceitos e como são integrados a modelagem da pesquisa.

7.2.3.1 Metadados DC e LOM identificados

Os metadados descritivos extrínsecos, estruturados em formulários para submissão de OAs em RDEs, são o foco desta pesquisa. Dublin Core (DC) e

⁶⁰ Segundo a ICP Ontology: “Structure of brain” is specific structure of the nervous system, part of Body structures. Body structures are anatomical parts of the body such as organs, limbs and their components. Em tradução livre: “Estrutura de cérebro” é a estrutura específica do sistema nervoso, que faz parte da estrutura de corpo. Estruturas do corpo são as partes anatômicas do corpo, tais como, órgãos, membros e seus componentes.

⁶¹ De acordo com a Dimensão do conhecimento, o Conhecimento conceitual aborda as relações entre as partes de uma estrutura maior que as fazem funcionar em conjunto (ANDERSON, 2001).

Learning Objects Metadata (LOM) são padrões de metadados amplamente utilizados na descrição de OAs, como já foi mencionado. A definição de um padrão de metadados tem como propósito a descrição e entendimento comum da informação a cerca do OA, além de permitir que os dados sejam eficiente e consistentemente armazenados, indexados, recuperados, processados e compartilhados por diversos sistemas (RODRIGUES; TAGA; VIEIRA, 2011).

Nesta etapa são identificados metadados para os conceitos que especificam o objetivo de aprendizagem, de acordo com os padrões DC e LOM. Não se pretende optar pela adoção de um dos padrões, mas sim destacar como podem ser adaptados ou estendidos para atender aos propósitos da pesquisa. Para tanto, foram verificadas documentações referentes aos padrões já mencionadas na pesquisa documental sobre o tema.

Metadados DC para os conceitos que especificam o objetivo de aprendizagem

O DC destaca 15 elementos (Dublin Core Metadata Element Set) que fazem parte de um conjunto amplo de vocabulário de metadados e especificações técnicas. Esse conjunto, o DCMI Terms⁶², apresenta todos os termos e definições (Properties in the / terms/ namespace), classes de termos DCMI (Classes), vocabulário de tipos de recursos (DCMI Type Vocabulary), esquemas de codificação de vocabulário (Vocabulary Encoding Schemes), esquemas de codificação de sintaxe (Syntax Encoding Schemes) e termos relacionados ao DCMI Abstract Model (Terms related to the DCMI Abstract Model) (DCMIId, 2012).

O DCMI Abstract Model⁶³ especifica os componentes e constructos dos metadados, independente de qualquer sintaxe de codificação. Neste modelo é definida a natureza dos componentes utilizados e como estes podem ser combinados para criar estruturas de informação (DCMIe, 2012).

Os termos DC podem ser utilizados combinados com outros termos de vocabulários de metadados compatíveis, observado o contexto de aplicação e o modelo abstrato DCMI Abstract Model. Além disso, estão de acordo com a Política de Namespace DCMI que descreve como atribuir URIs, estabelece limites para alterações de rótulos, apresenta definições e comentários de uso.

⁶² Link: <<http://dublincore.org/documents/2012/06/14/dcmi-terms/>>.

⁶³ Link: <<http://dublincore.org/documents/abstract-model/>>.

Dessa forma, o DCMI Terms e o DCMI Abstract Model foram consultados para identificar os metadados. O DCMI Abstract Model permite melhor compreensão dos tipos de descrições codificadas com DC e facilitou o uso, adequação e estabelecimento de relações entre metadados. Já o DCMI Terms é a referência recomendada que acompanha alterações do DC, agrega os 15 elementos do Dublin Core Metadata Element Set e é mais preciso semanticamente do que esse, pois atribui domínios formais e faixa de valores às propriedades (elementos).

De acordo com o DCMI Abstract Model (DCMIe, 2012), o vocabulário utilizado na definição de metadados DC identifica a relação dos aspectos que podem descrever um recurso e são representados com metadados. *Recurso* é qualquer coisa que possa ser identificada, como por exemplo, um documento, um OA ou uma coleção de outros recursos. O modelo teórico de recursos descritos estabelece que *descrição* é uma ou mais declarações sobre um, e somente um, *recurso*. Cada *recurso* é descrito usando um ou mais pares de propriedade e valor. Assim, destaca-se:

- Um vocabulário é um conjunto de um ou mais termos. Cada termo é um membro de um ou mais vocabulários;
- Um termo é uma propriedade (elemento), classe, esquemas de codificação de vocabulário ou esquemas de codificação de sintaxe;
- Cada propriedade pode estar relacionada a uma ou mais classes com as relações “possui domínio” ou “possui valor”;
- Cada recurso pode ser uma ocorrência de uma ou mais classes;
- Cada recurso pode ser membro de um ou mais esquemas de codificação de vocabulário;
- Cada classe pode estar relacionada a uma ou mais classes em uma hierarquia, constituindo classes e subclasses;
- Cada propriedade pode estar relacionada a uma ou mais outras propriedades em uma hierarquia, constituindo propriedades e subpropriedades;
- Cada esquema de codificação de sintaxe é uma classe.

Diante disso, foram verificadas propriedades, classes, esquemas de codificação de vocabulário ou esquemas de codificação de sintaxe do DCMI Terms para estabelecer quais são adequados à proposta da pesquisa. Para cada termo que constitui o vocabulário de metadados foram especificados os seguintes atributos (DCMIId, 2012):

- Nome (Name): nome relacionado a URI do Namespace DCMI, ou seja, a codificação no padrão DC;
- Rótulo (Label): rótulo em português que identifica o termo;
- Definição (Definition): declaração que representa o conceito e a natureza do termo;

- Tipo de termo (Type of Term): tipo de termo de acordo com o DCMI Abstract Model.

Quanto adequado, também são indicadas propriedade (Refines) da subpropriedade, relações hierárquicas da classe (Broader Than e Narrower Than), classe que constitui domínio do termo (Has Domain), classe que constitui valor para o termo (Has Range), esquema de codificação de vocabulário do qual o termo é membro (Member Of) e classe da qual o termo descrito é uma instância (Instance Of). Consequentemente, são feitas considerações sobre a aplicação dos termos do vocabulário de metadados DC para a descrição dos conceitos que especificam o objetivo de aprendizagem dos OAs na área de saúde.

Para facilitar a compreensão, propriedades são sublinhadas e **classes** são formatadas em negrito, de acordo com a modelagem proposta. “Termos” do DC são apresentados entre aspas.

O **objetivo de aprendizagem** pode ser especificado com a propriedade “instructionalMethod” da classe “MethodOfInstruction”, inclusive com informação do pré-requisito, quando for o caso. Como um objetivo de aprendizagem é estabelecido para um OA e OA é o recurso descrito, a relação entre os recursos descritos pode representar a relação entre objetivos de aprendizagem. Assim, **objetivo de aprendizagem prévio** e **objetivo de aprendizagem posterior** podem ser relacionados ao **objetivo de aprendizagem** de um OA com o uso da propriedade “relation” para indicação das URLs dos OAs descritos que apresentam essas relações. Seguindo a mesma analogia, o **objetivo de aprendizagem requerido** ao qual o OA se refere pode ser relacionado a outro OA que possui um **objetivo de aprendizagem** que requer esse primeiro com as propriedades “requires” e “isRequiredBy”, estabelecendo, inclusive, relações inversamente proporcionais. O quadro resumo DC com essas propriedades e classes é apresentado a seguir (Quadro 3).

Quadro 3: Resumo de termos do vocabulário de metadados DC - Objetivo de aprendizagem.

Nome: instructionalMethod	
Rótulo	Método de ensino
Definição	Processo para gerar conhecimentos, atitudes e habilidades, para o qual o recurso descrito é projetado. Incluem todos os aspectos do ensino-aprendizagem.
Tipo de termo	Propriedade
Has range	Classe MethodOfInstruction: processo usado para gerar conhecimentos, atitudes e habilidades.
Nome: relation	

Rótulo	Relação
Definição	Recurso relacionado ao recurso descrito.
Tipo de termo	Propriedade
Nome: requires	
Rótulo	Requer
Definição	Recurso relacionado que é exigido pelo recurso descrito para apoiar a sua função, utilização ou coerência.
Tipo de termo	Propriedade
Refines	Propriedade relation: recurso relacionado ao recurso descrito.
Nome: isRequiredBy	
Rótulo	É requerido por.
Definição	Recurso relacionado que exige o recurso descrito para apoiar a sua função, utilização ou coerência.
Tipo de termo	Propriedade
Refines	Propriedade relation: recurso relacionado ao recurso descrito.

Para formalizar que **objetivo de aprendizagem** é sobre um **assunto** e é para um **público-alvo**, a modelagem proposta relaciona semanticamente os metadados apropriados, ainda que essa informação possa ser contemplada em “instructionalMethod”.

Já **processo cognitivo**, que qualifica o objetivo de aprendizagem e escabece categorias de verbos para sua redação, não pode ser descrito de maneira controlada nas propriedades e classes DC, embora também possa ser contemplado em “instructionalMethod”. Alternativa é definir um esquema de codificação de vocabulário para a classe “MethodOfInstruction”, ou seja, uma subclasse com os valores indicados na modelagem. Depois estender a propriedade “instructionalMethod” e estabelecer valores dessa subclasse (indicação de “possui valor”).

O **assunto** é representado com a propriedade “subject” e pode ser descrito usando texto livre ou taxonomias, como a **ICP Ontology**. O uso de VC é recomendado e para isso deve ser definida uma classe, ou esquema de codificação de vocabulário, que estabelece os valores para a propriedade “subject”. Alguns esquemas de codificação de vocabulário do DC são o MESH e a National Library of Medicine Classification (NLM).

A **dimensão do conhecimento**, que classifica o **assunto** de acordo com sua abordagem, não pode ser descrita de maneira controlada pelas propriedades e classes DC. Sua informação em texto livre no “subject”, por ser muito genérica, pode gerar resultados de busca pouco satisfatórios e por isso não é recomendada. Classificar a **ICP Ontology**, esquema de codificação de vocabulário adotado para o **assunto**, demanda muito tempo e

requer decisões arbitrárias que podem não atender o consenso. Alternativa é adotar outro termo para representar a **dimensão do conhecimento** com valores estabelecidos em esquema de codificação de vocabulário. A opção mais coerente é a propriedade “description”, que possibilita um relato sobre o recurso descrito. Também é possível qualificar a propriedade “subject” para atender esse caso específico, o que é pertinente, porém, diminui a interoperabilidade do esquema de metadados.

Quadro 4: Resumo de termos do vocabulário de metadados DC - Assunto

Nome: subject	
Rótulo	Assunto
Definição	Assunto abordado no recurso.
Tipo de termo	Propriedade
Nome: description	
Rótulo	Descrição
Definição	Relato sobre o recurso.
Tipo de termo	Propriedade

O **público-alvo** é representado pela propriedade “audience” da classe “AgentClass” que permite a adoção de um esquema de codificação de vocabulário, como a **CBO**, para indicar a formação profissional e atividades de acordo com as **competências**. Os sinônimos das ocupações da CBO são tratados no próprio esquema. O nível educacional é descrito com a propriedade “educationLevel”, subpropriedade de “audience”. Os outros conceitos que descrevem o **público-alvo**, contexto educacional e **assunto de interesse**, demandam a qualificação da propriedade “audience”, a definição de uma nova propriedade para a classe “AgentClass”, ou podem simplesmente serem considerados com a adoção da ontologia para público-alvo. O contexto educacional pode ser informado em texto livre, caso não seja adotado um esquema.

Quadro 5: Resumo de termos do vocabulário de metadados DC - Público-alvo.

Nome: audience	
Rótulo	Público-alvo
Definição	Classe de entidades para a qual o recurso se destina ou é útil.
Tipo de termo	Propriedade
Has Range	Classe AgentClass: grupos de estudantes, instituições, etc. Instância da classe Agent (Instance Of AgentClass).

Nome: educationLevel	
Rótulo	Nível educacional do público-alvo
Definição	Classe de entidades, definida em termos da progressão através de um contexto educacional ou de formação, para o qual o recurso descrito se destina.
Tipo de termo	Propriedade
Has Range	Classe AgentClass: grupos de estudantes, instituições, etc. Instância da classe Agent (Instance Of AgentClass).
Refines	audience

Como observado, a maior parte dos conceitos da modelagem são representados pelo DC, mesmo sendo elaborado para descrições gerais e não específicas de recursos de ensino-aprendizagem. Apenas **dimensão do conhecimento** e **processo cognitivo** não têm especificação clara e necessitam de alternativas para atender a necessidade. Em todos os casos, as propriedades podem ser informadas, ainda que sua redação ocorra em um elemento mais amplo e não permita a adoção de VCs, como na descrição do processo cognitivo; ou sua redação demande a qualificação de termos DC, como no caso do contexto educacional. Cabe destacar, que os valores dos elementos de metadados permitem o estabelecimento de relações entre estes e a adoção dos VCs propostos.

Metadados LOM para os conceitos que especificam o objetivo de aprendizagem

O LOM é um modelo de dados hierárquico com elementos de dados simples e de agregação. Os elementos simples armazenam dados e os elementos de agregação concentram outros elementos, formando categorias. Em seu esquema, apenas os elementos simples possuem valores associados e tipo de dados definidos. Os elementos de dados usados na descrição constituem instâncias, logo, a instância de metadados de um OA possui elementos de dados que descrevem suas características (BARKER, 2005).

A semântica de um elemento é determinada pelo seu contexto, pelas características dos elementos hierarquicamente acima (*pai*) e por aqueles que compõem a mesma categoria. A norma IEEE 1484 apresenta o esquema conceitual de dados, elementos e significados (semântica), relações entre elementos (estrutura), e requisitos para o preenchimento dos elementos (BARKER, 2005).

De acordo com o modelo de dados LOM, categoria é um grupo de elementos de dados relacionados. Essas categorias são: Geral (1 *General*), Ciclo de vida (2 *Lifecycle*),

Meta-Metadados (3 *Meta-Metadados*), Requisitos técnicos (4 *Technical*), Educacional (5 *Educational*), Direitos de propriedade e uso (6 *Rights*), Relação entre objetos (7 *Relation*), Uso educacional (8 *Annotation*) e Classificação (9 *Classification*). Outras definições fundamentais para a utilização do padrão são:

- Elemento de dado é um elemento para o qual são definidos nome, explanação, tipo de dado, valor determinado e relevância da ordem dos valores;
- Tipo de dado é uma propriedade com valores específicos (*LangString* - qualquer sequência de caracteres Unicode, *Vocabulary* - lista de opções, ou seja, VC, etc.);
- Valor determinado é o conjunto de valores permitidos para o elemento de dados;
- Elemento de dado estendido é um elemento de outra estrutura de dados.

A partir da verificação do esquema conceitual e do modelo de dados (BARKER, 2005), foram identificados elementos LOM adequados para a representação dos conceitos. Vale destacar que extensões do esquema devem manter o tipo de dado e o valor determinado para os elementos de dados, além disso, esses atributos não podem ser atribuídos aos elementos de dados agregados já especificados. Para cada elemento são definidos:

- Nome (Name): numeração e nome que referencia o elemento no padrão. A numeração indica a posição na hierarquia de elementos de dados agregados e a categoria;
- Rótulo (Label): rótulo em português que identifica o elemento;
- Explanação (Explanation): definição do elemento.

A numeração dos elementos diz respeito à agregação daqueles com características e significado comuns, imposta pelo padrão. Ou seja, todos os elementos pertencentes à mesma hierarquia devem ser incluídos na instância de metadados do OA, mesmo que apenas um seja preenchido na descrição. Assim, os elementos agregados são informados quando constituem a hierarquia dos elementos de dados adequados a proposta da pesquisa. De acordo com a necessidade, também são informados tipo de dado (*Datatype*), valor determinado (*Value space*) e se a ordem dos valores do elemento é significativa (*Order*). Por fim, são feitas considerações sobre a aplicação dos elementos LOM para a descrição dos conceitos que especificam o objetivo de aprendizagem dos OAs na área de saúde.

Assim como nas considerações sobre o DC, propriedades são sublinhadas e **classes** são formatadas em negrito, de acordo com a modelagem proposta. “Elementos” do LOM são apresentados entre aspas.

O **objetivo de aprendizagem** pode ser especificado com o elemento “5.10 Description”, e também o pré-requisito, quando for o caso. Esse elemento é genérico, pois permite qualquer informação sobre o uso educacional do OA, mas sua categoria (5 *Educational*) estabelece elementos para descrever informação educacional ou características pedagógicas essenciais para a qualidade da experiência de ensino-aprendizagem, o que justifica sua utilização de acordo com a proposta da pesquisa. Também é possível especificar o **objetivo de aprendizagem** e pré-requisito com os elementos “9.1 Purpose – valor *objetivo de aprendizagem*” e “9.3 Description”, se essa informação indicar uma classificação dos OAs. O **processo cognitivo**, que qualifica o objetivo de aprendizagem e escabece categorias de verbos para sua redação, pode ser representado com os elementos “9.1 Purpose – valor *objetivo de aprendizagem*” e “9.4 Keyword” elementos LOM, embora também possa ser contemplado em “5.10 Description”. O elemento

Outros elementos definidos na categoria Educacional (5 *Educational*) são recomendados para caracterizar (ou qualificar) a aprendizagem proposta pelo objetivo de aprendizagem para o OA, mas não atendem diretamente o interesse da pesquisa, tais como, “5.1: Educational.InteractivityType”, ou Tipo de interatividade, classifica a abordagem da proposta de aprendizagem como "ativa" ou "expositiva"; “5.3 Interactivity Level”, ou Nível de interatividade, que define quanto o aprendiz pode influenciar o aspecto ou o comportamento do OA; e “5.4 Semantic Density”, ou Densidade semântica, indica níveis de concisão da aprendizagem (baixo, médio, alto, muito alto).

Como um objetivo de aprendizagem é estabelecido para um OA e OA é o recurso descrito, a relação entre os recursos descritos pode representar a relação entre os objetivos de aprendizagem, da mesma forma como verificado no DC. Assim, **objetivo de aprendizagem prévio** e **objetivo de aprendizagem posterior** podem ser relacionados ao **objetivo de aprendizagem**, tal qual, **objetivo de aprendizagem requerido** pode ser relacionado a outro OA que possui **objetivo de aprendizagem** que requer esse primeiro, com o uso do elemento “7.1 Kind” para indicação da natureza da relação (reference, isreferencedby, requires, isrequiredby). O OA é relacionado a outro com seu respectivo identificador no elemento “7.2.1.2 Entry” e descrição no “7.2.2 Description” (caso necessária). O tipo de identificador adotado, elemento “7.2.1 Identifier”, é a URL do OA no RDE. O esquema de catalogação, elemento “7.2.1.1 Catalog”, não precisa ser preenchido pelo fato das relações se darem no mesmo contexto (RDE), utilizando os mesmos padrões. O quadro resumo de elementos LOM é apresentado a seguir (Quadro 6).

Quadro 6: Resumo de elementos LOM - Objetivo de aprendizagem.

Nome: 5.1 Interactivity Type	
Rótulo	Tipo de interatividade
Explicação	Modo predominante da estratégia de ensino-aprendizagem proposta pelo OA.
Datatype	Vocabulário controlado (vocabulary)
Value space	Aprendizagem "ativa" (active): aprender fazendo. O conteúdo propõe uma ação produtiva por parte do aluno. Aprendizagem "expositiva" (expositive): aprendizagem passiva. Consiste principalmente em absorver o conteúdo exposto. Aprendizagem "mista" (mixed): combina aprendizagem ativa e expositiva.
Nome: 5.10 Description	
Rótulo	Descrição
Explicação	Comentário sobre o OA.
Nome: 7.1 Kind	
Rótulo	Tipo de relação
Explicação	Natureza da relação entre o OA descrito e o OA identificado pelo elemento 7.2 Relation.Resource.
Datatype	Vocabulário controlado (vocabulary)
Value space	Segue as relações especificadas pelo DC: ispartof, haspart, isversionof, hasversion, isformatof, hasformat, reference, isreferencedby, isbasedon, sbasisfor, requires, isrequiredby.
Nome: 7.2 Resource – Elemento agregador	
Rótulo	Recurso relacionado
Explicação	OA relacionado ao OA descrito.
Nome: 7.2.1 Identifier – Elemento agregado	
Rótulo	Identificador do OA relacionado
Explicação	Identificador do OA relacionado ao OA descrito.
Nome: 7.2.1.1 Catalog – Elemento agregado	
Rótulo	Identificador do OA relacionado
Explicação	Nome ou designação do identificador ou esquema de catalogação do OA relacionado ao OA descrito.
Datatype	Letras, números e símbolos (CharacterString)
Value space	Repertoire of ISO/IEC 10646-1:2000
Nome: 7.2.1.2 Entry – Elemento agregado	
Rótulo	Entrada do OA relacionado
Explicação	Valor do identificador de acordo com a identificação ou esquema de

	catalogação do OA relacionado ao OA descrito.
Datatype	Letras, números e símbolos (CharacterString)
Value space	Repertoire of ISO/IEC 10646-1:2000
Nome: 7.2.2 Description – Elemento agregado	
Rótulo	Descrição do OA relacionado
Explicação	Descrição do OA relacionado ao OA descrito.
Datatype	Letras, números (LangString)
Value space	Repertoire of ISO/IEC 10646-1:2000
Nome: 9.1 Purpose	
Rótulo	Objetivo
Explicação	Característica para classificação do OA.
Datatype	Vocabulário controlado (vocabulary)
Value space	Disciplina, ideia, requisito prévio, objetivo de aprendizagem, acessibilidade, restrições, nível educacional, nível de habilidade, nível de segurança, competência.
Nome: 9.3 Description	
Rótulo	Descrição relativa a 9.1 Purpose
Explicação	Descrição do OA, relativa ao valor escolhido no elemento 9.1 Purpose.
Nome: 9.4 Keyword	
Rótulo	Palavra-chave relativa a 9.1 Purpose
Explicação	Palavra-chave ou frase que descreve o OA, relativa ao valor escolhido no elemento 9.1 Purpose.
Ordered	A ordem das palavras-chave é significativa.

Outros elementos que caracterizam o OA podem indicar o direcionamento do objetivo de aprendizagem, porém, essa relação é demasiado subjetiva e não será considerada na modelagem proposta. O elemento “1.7 Structure”, ou Estrutura, por exemplo, estabelece a estrutura organizacional do OA, que pode ser Atômica (objeto indivisível), Coleção (conjunto de objetos sem relação especificada entre eles), Rede (conjunto de objetos com relacionamentos indeterminado), Hierárquica (conjunto de objetos cujas relações podem ser representadas por uma estrutura de árvore) e Linear (conjunto de objetos totalmente ordenado). Para cada estrutura pode ser previsto um conjunto típico de objetivos de aprendizagem como, reconhecer e lembrar para um OA com estrutura Atômica.

O **assunto** é representado com o elemento “1.5 Keyword”, caso seja um texto livre. Para descrever o **assunto** com a **ICP Ontology** (valor definido pela propriedade taxonomia), deve ser usado o elemento “9.2.1 Source” para indicar sua adoção, elemento “9.2.2 Taxon” para informar termos que representam categorias, elemento “9.2.2.1 Id” para informar identificador do termo e o elemento “9.2.2.2 Entry” para informar o termo.

O LOM, assim como o DC, não apresenta elemento para descrever a **dimensão do conhecimento**, que classifica o **assunto**. Ela pode ser informada em texto livre no elemento “1.5 Keyword”, ou em uma extensão do mesmo. Como alternativa, é possível utilizar os elementos já citados “9.1 Purpose – valor *disciplina*” e “9.4 Keyword”.

Quadro 7: Resumo de elementos LOM - Assunto

Nome: 1.5 Keyword	
Rótulo	Assunto
Explicação	Palavra-chave ou frase que descreve o tema do OA.
Nome: 9.2 Taxon Path – Elemento agregador	
Rótulo	Taxonomia indicada
Explicação	Indicação de uma taxonomia ou classificação específica para descrever o tema do OA.
Nome: 9.2.1 Source – Elemento agregado	
Rótulo	Fonte
Explicação	Nome da taxonomia ou classificação indicada para descrever o tema do OA.
Datatype	Letras, números (LangString)
Value space	Repertoire of ISO/IEC 10646-1:2000.
Nome: 9.2.2 Taxon – Elemento agregado	
Rótulo	Taxonomia
Explicação	Termo que representa uma categoria da taxonomia ou classificação indicada para descrever o tema do OA.
Ordered	A ordem dos termos é significativa.
Nome: 9.2.2.1 Id – Elemento agregado	
Rótulo	Identificador do termo
Explicação	Identificador do termo da taxonomia ou classificação indicada para descrever o tema do OA.
Datatype	Letras, números e símbolos (CharacterString)
Value space	Repertoire of ISO/IEC 10646-1:2000.
Nome: 9.2.2.2 Entry – Elemento agregado	
Rótulo	Termo

Explicação	Termo da taxonomia ou classificação indicada para descrever o tema do OA.
Nome: 9.1 Purpose	
Rótulo	Objetivo
Explicação	Característica para classificação do OA.
Datatype	Vocabulário controlado (vocabulary)
Value space	Disciplina, ideia, requisito prévio, objetivo de aprendizagem, acessibilidade, restrições, nível educacional, nível de habilidade, nível de segurança, competência.
Nome: 9.4 Keyword	
Rótulo	Palavra-chave relativa a 9.1 Purpose
Explicação	Palavra-chave ou frase que descreve o OA, relativa ao valor escolhido no elemento 9.1 Purpose.
Ordered	A ordem das palavras-chave é significativa.

O **público-alvo** não pode ser representado com a **CBO** para indicar formação profissional e atividades de acordo com as **competências**. O elemento “5.5 Intended End User Role” apenas define tipos de público-alvo e deve ser estendido para atender a necessidade. O nível educacional também não pode ser representado adequadamente, pois o elemento “5.6 Context” apenas define níveis básicos que não contemplam o público-alvo dos OA de RDEs na área de saúde. O contexto educacional não é atendido pelo elemento “5.6 Context”, assim como o **assunto de interesse**. O elemento “5.6 Context” precisa ser estendido para representar nível educacional, contexto educacional e **assunto de interesse**.

Quadro 8: Resumo de elementos LOM – Público-alvo

Nome: 5.5 Intended End User Role	
Rótulo	Tipo de público-alvo
Explicação	Indicação do principal tipo de público-alvo para qual o OA foi elaborado.
datatype	Vocabulário controlado (vocabulary)
value space	Professor, autor, aprendiz, gerente.
Nome: 5.6 Context	
Rótulo	Contexto
Explicação	Contexto para qual o OA foi elaborado.
datatype	Vocabulário controlado (vocabulary)
value space	Educação básica, Ensino superior, Treinamento, outro.

Os elementos LOM não representam alguns conceitos, ainda que o padrão seja direcionado ao contexto de ensino-aprendizagem. O **processo cognitivo** e a **dimensão do conhecimento** não são contemplados por um elemento específico, mas são apresentadas alternativas adequadas. O **público-alvo** não pode ser especificado de acordo com a proposta, pois o LOM restringe o valor dos elementos que representam esse conceito, tornando-se necessária a extensão.

Embora haja a formação de força-tarefa entre o LTSC IEEE e a DCMI (Joint DCMI/IEEE LTSC Taskforce, 2012) para elaborar padrão RDF para os elementos de dados LOM e recomendar como esses elementos podem ser expressos utilizando o modelo abstrato DC, é importante a correlação dos padrões. Essa correlação, apresentada na seção seguinte (Seção 7.2.3.2), estabelece interoperabilidade semântica entre termos DC e elementos LOM, utilizados para representar os conceitos que especificam o objetivo de aprendizagem.

7.2.3.2 Correlação DC e LOM

A correlação entre DC e LOM contribui com o objetivo da pesquisa, pois o modelo abstrato DC oferece interligação com RDF e viabiliza expressão de metadados de maneira adequada para aplicações da WS (NILSSON; *et al.*, 2006) e (DCMIe, 2012). Logo, a interoperabilidade semântica entre os padrões permite que RDEs na área de saúde possam implementar a modelagem proposta, independentemente de adotarem o LOM. Além disso, a correlação dos padrões destacou a melhor contribuição de cada para descrição eficiente em nível pedagógico, ou seja, características necessárias para o êxito no processo de ensino-aprendizagem; e em nível de domínio, que diz respeito ao conteúdo, a temática.

De acordo com recomendações do IEEE P1484.12.4™/D1Draft Recommended Practice for Expressing IEEE Learning Object Metadata Instances Using the Dublin Core Abstract Model (IEEE DRAFT, 2008), o primeiro passo para atingir a interoperabilidade entre os padrões é a compreensão e o mapeamento entre termos DC e elementos LOM, como realizado na correlação. Este mapeamento buscou equivalências semânticas entre termos e elementos a partir das definições apresentadas pelos padrões. Para tanto, além dos documentos já citados, foram consultados *Analysis of each of the LOM Elements*⁶⁴, *An example*

⁶⁴ Link: <<http://dublincore.org/educationwiki/DCMIIEEELTSCTaskforce/LomDCAMAnalysis>>.

*DCLOM description set*⁶⁵ e a ontologia de metadados DC, disponibilizada na lista de projetos que utilizam o Protégé⁶⁶.

Os metadados adequados para a pesquisa foram identificados na seção anterior (Seção 7.2.3.1) e isto evidenciou que os padrões verificados utilizam terminologias diferentes. No DC, o recurso é descrito por pares de propriedade e valor, essas também chamadas de elementos. Propriedades, classes, valores e esquemas de codificação de vocabulário ou codificação de sintaxe, são termos. Os termos constituem o vocabulário dos metadados DC. Já no LOM, elementos de dados simples armazenam dados e são reunidos por elementos de dados de agregação, conforme suas características e significados. Os elementos constituem o esquema de metadados LOM. Diante disso, uma propriedade DC e um elemento de dados LOM, denominados metadados, são correlacionados no quadro a seguir (Quadro 9), assumindo-se que podem ser adotados em RDEs na área de saúde para descrever a mesma informação a cerca do OA. Valores DC e dados LOM, que compõem a representação do metadado, constituem informação descritiva do OA ou indicam relações entre metadados.

Assim, para os conceitos que especificam o objetivo de aprendizagem, são definidas propriedades e classes DC, elementos de dados e elementos de agregação LOM. Caso as propriedades e elementos já definidos não atendam plenamente os requisitos da pesquisa, são sugeridas alternativas a partir da qualificação dos metadados existentes ou extensão do padrão.

Quadro 9: Correlação DC e LOM para os conceitos que especificam o objetivo de aprendizagem.

Conceito	DC	LOM
Objetivo de aprendizagem	Propriedade: “instructionalMethod” Classe: “MethodOfInstruction”	Elemento de dado: “5.10 Description”
Objetivo de aprendizagem requerido	Propriedades: “requires” e “isRequiredBy” (Identificador do OA no RDE)	Elemento de dado: “7.1 Kind” (Natureza da relação) Elemento de agregação: “7.2 Resource” Elemento de dado: “7.2.1 Identifier” (Tipo de identificador)

⁶⁵ Link: <<http://dublincore.org/educationwiki/DCMIIEEELTSCSTaskforce/Examples>>.

⁶⁶ Link: <<http://protege.cim3.net/cgi-bin/wiki.pl?DublinCoreProtegeOntology>>.

		Elemento de dado: “7.2.1.2 Entry” (Identificador do OA no RDE)
Objetivo de aprendizagem prévio	Propriedade: “relation” (Identificador do OA no RDE)	Elemento de dado: “7.1 Kind” (Natureza da relação) Elemento de agregação: “7.2 Resource” Elemento de dado: “7.2.1 Identifier” (Tipo de identificador) Elemento de dado: “7.2.1.2 Entry” (Identificador do OA no RDE)
Objetivo de aprendizagem posterior	Propriedade: “relation” (Identificador do OA no RDE)	Elemento de dado: “7.1 Kind” (Natureza da relação) Elemento de agregação: “7.2 Resource” Elemento de dado: “7.2.1 Identifier” (Tipo de identificador) Elemento de dado: “7.2.1.2 Entry” (Identificador do OA no RDE)
Processo cognitivo	Classe: “MethodOfInstruction” Subclasse estendida: “ProcessoCognitivo” Propriedade estendida: “processoCognitivo” (Propriedade “possui valor” da Subclasse).	Elemento de dado: “9.1 Purpose – valor <i>objetivo de aprendizagem</i> ” Elemento de dado: “9.4 Keyword”
Assunto (Taxonomia)	Classe estendida: “SubjectTaxonomia” Propriedade: “subject” (Propriedade “possui valor” da Classe)	Elemento de agregação: 9.2 Taxon Path Elemento de dado: “9.2.1 Source” (Indicar taxonomia) Elemento de dado: “9.2.2 Taxon”

		(Categorias da taxonomia) Elemento de dado: “9.2.2.1 Id” (Identificador do termo) Elemento de dado: “9.2.2.2 Entry” (Termo)
Assunto (Texto livre)	Propriedade: “subject”	Elemento de dado: “1.5 Keyword”
Dimensão do conhecimento	Classe estendida: “DimensaoConhecimento” Propriedade qualificada: “subjectDimCon” (Propriedade “possui valor” da Classe)	Elemento de dado: “9.1 Purpose – valor <i>disciplina</i> ” Elemento de dado: “9.4 Keyword”
Público-alvo	Classe: “AgentClass” Propriedade: “audience”	Elemento de dado estendido: “5.12 Publico Alvo”
Nível educacional	Classe: “AgentClass” Propriedade: “audience” Subpropriedade: “educationLevel”	Elemento de dado qualificado: “5.6.1 Context Nivel Educacional”
Contexto educacional	Classe: “AgentClass” Propriedade qualificada: “contextoEducacional”	Elemento de dado estendido: “5.13 Contexto Educacional”
Assunto de interesse	Classe: “AgentClass” Propriedade: “audience” Propriedade estendida: “assuntoInteresse”	Elemento de dado estendido: “5.14 Assunto Interesse”

De acordo com o DC, mais especificamente com os níveis de interoperabilidade definidos pela DCMI (DCMIId, 2012), a correlação apresentada no “Quadro 9: Correlação DC e LOM para os conceitos que especificam o objetivo de aprendizagem” fornece vocabulário para compartilhar o significado da informação descrita, de acordo com o contexto de utilização (Nível 1 – Compartilhamento de

termos e definições). A interoperabilidade semântica também é atingida com formalização da relação entre termos, propiciada pelas triplas RDF, baseadas em vocabulários com termos DCMI. Ou seja, a semântica do RDF é aplicada aos termos DCMI, assim, a utilização dos termos é mais precisa e permite a realização de inferências (Nível 2 – Interoperabilidade semântica formal). A discussão sobre interoperabilidade entre os padrões foram apresentadas na Seção 5.3.1.

Dada a correção entre DC e LOM, a seção seguinte (Seção 7.2.3.3) discorre sobre a integração dos metadados a modelagem, ou seja, como os conceitos, relações, propriedades e valores estabelecidos para especificar o objetivo de aprendizagem são representados com metadados dispostos em RDEs na área de saúde para descrever OAs.

7.2.3.3 Integração dos metadados a modelagem da pesquisa

Nesta seção discute-se como a informação descritiva dos OAs, a cerca do objetivo de aprendizagem, é representada no formulário de RDEs que reúne o conjunto de metadados utilizados para este fim. Para tanto, são definidos valores de preenchimento para metadados que especificam o objetivo de aprendizagem, com a recomendação de VC, quando adequado à formalização e realização de inferências. As relações entre metadados, que agregam semântica de acordo com a proposta da pesquisa, também são evidenciadas.

A relação entre metadados é orientada pelas especificações do modelo abstrato DC, DCMI Abstract Model. Isto se justifica, pois esse modelo promove o processamento dos metadados em aplicações WS; existem esforços para a definição de instâncias LOM usando DCMI Abstract Model, como discutido anteriormente; e ainda, define a natureza dos constituintes de vocabulários de metadados e descreve como são combinados para criar estruturas de informação. O DCMI Abstract Model é um modelo de informação independente de qualquer sintaxe de codificação particular (DCMIe, 2012).

De acordo com o DCMI Abstract Model, recurso é qualquer coisa que possa ser identificada, como por exemplo, um OA, serviço ou coleção de outros recursos. O modelo teórico de recursos descritos (por descrições) especifica:

- Descrição é uma ou mais declarações sobre um, e somente um recurso;

- Cada recurso descrito é descrito usando um ou mais pares de propriedade/valor (property/value pair);
- Um par de propriedade/valor é a combinação de uma propriedade e um valor, usado para descrever uma característica;
- Propriedade é um aspecto específico, característica, atributo, a relação usada para descrever recursos;
- Valor é a entidade física, entidade conceitual ou literal (um recurso), que está associada a propriedade quando um par de propriedade/valor é usado para descrever recurso;
- Cada valor é também um recurso (físico, digital, entidade conceitual ou literal) que está associado a propriedade. Portanto, cada valor é um valor literal ou valor não-literal;
- Valor literal é uma entidade que utiliza uma sequência de caracteres Unicode como uma forma lexical para denotar um recurso;
- Valor não-literal é uma entidade física, digital ou conceitual.

Assim, OA é um recurso descrito com declarações sobre suas características, especificadas por propriedades e valores. Quando o valor da propriedade denota relação entre metadados ele é um valor não-literal, quando apresenta a informação descritiva em si é um valor literal ou valor não-literal, para casos em que são adotados esquemas de codificação de vocabulário.

A partir de então, o DCMI Abstract Model (DCMIe, 2012) define o modelo teórico para o conjunto de metadados que compõem a descrição de recurso, cujas principais considerações são:

- Cada declaração instancia um par propriedade/valor e é composta de uma propriedade (identificada com URI) e valor preenchido;
- Valor preenchido é um valor literal ou não-literal;
- Valor preenchido literal é composto de exatamente uma cadeia de valor. A cadeia de valor é um literal que codifica o valor literal;
- Valor preenchido não-literal é composto de zero ou valor URI (URI identifica o valor não-literal associado a propriedade), zero ou esquema de codificação de vocabulário URI (URI identifica o esquema do qual o valor não-literal é membro) e zero ou cadeias de valor. Cada cadeia de valor é um literal que representa o valor não-literal;
- Cadeia de valor é um valor pré-definido ou digitado.

Para a pesquisa, as relações entre os metadados do conjunto que descreve o OA, conforme a abordagem DCMI Abstract Model (DCMIe, 2012), são especificadas

pelas declarações a cerca do objetivo de aprendizagem, com os pares de propriedade/valor:

- *É sobre* Assunto/*string* (valor literal), Assunto/ICF Ontology (valor não-litera);
- *É para* Público-alvo/CBO (valor não-litera);
- *Possui* Processo cognitivo/processo cognitivo (valor não-litera); Especificação/*string* (valor literal), Pré-requisito/*string* (valor literal);
- *É relacionado a* Objetivo de aprendizagem posterior/objetivo de aprendizagem (valor não-litera), Objetivo de aprendizagem prévio/objetivo de aprendizagem (valor não-litera), Objetivo de aprendizagem requerido/objetivo de aprendizagem (valor não-litera).

As propriedade com valor não-litera indicam relações semânticas formalizadas com tecnologias da WS, discutidas anteriormente. Como por exemplo, para declarar formalmente que **objetivo de aprendizagem é sobre** um **assunto** e **é para** um **público-alvo**. O Quadro 10 destaca os valores de preenchimento literais e valores de preenchimento não-literais que ampliam a semântica com a definição de relações ou adoção de VC. Em outras palavras, são apresentados conceitos definidos pela modelagem proposta, sua representação com os padrões DC e LOM, e valores que são a informação descritiva em si ou uma referência a ontologias ou outro metadado.

Quadro 10: Integração de metadados e valores propostos pela modelagem.

Conceito	DC	LOM	Valor
Objetivo de aprendizagem	Propriedade: "instructionalMethod" Classe: "MethodOfInstruction"	Elemento de dado: "5.10 Description"	<i>String</i>
Objetivo de aprendizagem requerido	Propriedades: "requires" e "isRequiredBy" (Identificador do OA no RDE)	Elemento de dado: "7.1 Kind" (Natureza da relação) Elemento de agregação: "7.2 Resource" Elemento de dado: "7.2.1 Identifier" (Tipo de identificador) Elemento de dado: "7.2.1.2 Entry"	Objetivo de aprendizagem

		(Identificador do OA no RDE)	
Objetivo de aprendizagem prévio	Propriedade: “relation” (Identificador do OA no RDE)	Elemento de dado: “7.1 Kind” (Natureza da relação) Elemento de agregação: “7.2 Resource” Elemento de dado: “7.2.1 Identifier” (Tipo de identificador) Elemento de dado: “7.2.1.2 Entry” (Identificador do OA no RDE)	Objetivo de aprendizagem
Objetivo de aprendizagem posterior	Propriedade: “relation” (Identificador do OA no RDE)	Elemento de dado: “7.1 Kind” (Natureza da relação) Elemento de agregação: “7.2 Resource” Elemento de dado: “7.2.1 Identifier” (Tipo de identificador) Elemento de dado: “7.2.1.2 Entry” (Identificador do OA no RDE)	Objetivo de aprendizagem
Processo cognitivo	Classe: “MethodOfInstruction” Subclasse estendida: “ProcessoCognitivo” Propriedade estendida: “processoCognitivo” (Propriedade “possui valor” da Subclasse).	Elemento de dado: “9.1 Purpose – valor <i>objetivo de aprendizagem</i> ” Elemento de dado: “9.4 Keyword”	Processo cognitivo
Assunto (Taxonomia)	Classe estendida: “SubjectTaxonomia” Propriedade: “subject” (Propriedade “possui valor” da Classe)	Elemento de agregação: 9.2 Taxon Path Elemento de dado: “9.2.1 Source” (Indicar taxonomia) Elemento de dado:	ICF Ontology

		<p>“9.2.2 Taxon” (Categorias da taxonomia) Elemento de dado: “9.2.2.1 Id” (Identificador do termo) Elemento de dado: “9.2.2.2 Entry” (Termo)</p>	
Assunto (Texto livre)	Propriedade: “subject”	Elemento de dado: “1.5 Keyword”	<i>String</i>
Dimensão do conhecimento	Classe estendida: “DimensaoConhecimento” Propriedade qualificada: “subjectDimCon” (Propriedade “possui valor” da Classe)	Elemento de dado: “9.1 Purpose – valor <i>disciplina</i> ” Elemento de dado: “9.4 Keyword”	Dimensão do conhecimento
Público-alvo	Classe: “AgentClass” Propriedade: “audience”	Elemento de dado estendido: “5.12 Publico Alvo”	CBO
Nível educacional	Classe: “AgentClass” Propriedade: “audience” Subpropriedade: “educationLevel”	Elemento de dado qualificado: “5.6.1 Context Nivel Educacional”	<i>String</i>
Contexto educacional	Classe: “AgentClass” Propriedade qualificada: “contextoEducacional”	Elemento de dado estendido: “5.13 Contexto Educacional”	<i>String</i>
Assunto de interesse	Classe: “AgentClass” Propriedade: “audience” Propriedade estendida: “assuntoInteresse”	Elemento de dado estendido: “5.14 Assunto Interesse”	ICF Ontology

É importante destacar que conforme os níveis de interoperabilidade definidos pela DCMI, o DCMI Abstract Model fornece noções de sintaxe e delimita o conjunto

de termos da descrição do OA, garantindo base para a validação de metadados (Nível 3 – Interoperabilidade sintática do conjunto de termos da descrição). Isto aumenta a capacidade de estabelecer semântica formal e regras de inferências.

Para a modelagem proposta, integrada com metadados, estes representam conceitos que especificam o objetivo de aprendizagem. Propriedades da modelagem que qualificam os conceitos são fornecidas pelas ontologias, adotadas como valores de preenchimento de metadados para diminuir ambiguidades, possibilitar generalizações e especificações. Propriedades da modelagem que relacionam conceitos são formalizadas nas camadas da WS para, assim, refinar resultados de busca.

8 Resultados

Esse capítulo discorre sobre os resultados atingidos pela pesquisa. A hipótese da pesquisa é que uma modelagem baseada em WS para RDEs na área de saúde pode contribuir com o refinamento dos resultados de busca do usuário, considerando o contexto de ensino-aprendizagem. Assim, para demonstrar que esta hipótese foi comprovada, é realizada a sistematização da modelagem proposta com devidas observações. As contribuições da modelagem para o refinamento dos resultados de busca são discutidas com a simulação da sua implementação. Por fim, são apresentadas considerações identificadas durante o desenvolvimento da pesquisa, pertinentes ao estudo.

8.1 Sistematização da modelagem proposta

Nesta seção são sistematizadas as tecnologias da WS e a estrutura da informação descritiva que constituem a modelagem proposta. Evidencia-se a relação entre as partes da modelagem e o RDE, para facilitar sua utilização. Aqui são apresentados requisitos para o refinamento dos resultados de busca de acordo com o objetivo da pesquisa.

A modelagem proposta tem foco no objetivo de aprendizagem de OAs disponibilizados em RDEs na área de saúde, dessa forma, o resultado de busca considera a relação existente entre os conceitos que especificam o objetivo de aprendizagem e, com uso das tecnologias da WS, são realizadas inferências para refinar esse resultado. As inferências são previstas a partir da formalização semântica dos aspectos que impactam o resultado de busca nesse contexto. O usuário, aprendiz ou educador com necessidades de buscas, é favorecido, pois o assunto e/ou outros aspectos que compõem a sua estratégia de busca são relacionados semanticamente para melhor representar seu interesse. Nesse sentido, considera-se que documentos para atender demandas do processo de ensino-aprendizagem não são estritamente representados por palavras-chave que indicam o assunto, como tradicionalmente observado em RDEs.

As inferências são realizadas automaticamente a partir da verificação da informação descritiva contida em metadados, propiciando a recuperação de OAs de acordo com objetivo de aprendizagem, público-alvo, processo cognitivos, etc. Embora a definição do objetivo de aprendizagem do OA não explicita a forma como a aprendizagem é realizada, indica o contexto e o que se espera do aprendiz ao final do processo de ensino-aprendizagem. Assim, o objetivo de aprendizagem é composto de uma ou mais declarações sobre o OA que destacam suas características educacionais, pertinentes para o usuário que realiza busca nesse contexto.

Como exemplo, um professor, ao elaborar estratégias de ensino-aprendizagem para EAD, não busca simplesmente termos expressos no título ou o assunto dos documentos, busca OAs capazes de viabilizar o cumprimento de objetivos de aprendizagem. Ademais, o interesse de busca dos usuários de RDEs na área de saúde pode estar unicamente registrado nos metadados do OA, fazendo com que a qualificação desses metadados esteja proporcionalmente vinculada à satisfação com a recuperação da informação. A modelagem proposta procurou justamente qualificar esses metadados em favor do refinamento dos resultados de busca, considerando aspectos semânticos relevantes, de acordo com a pesquisa documental.

A modelagem é uma especificação abstrata (modelo teórico) que permite formalizar a semântica inerente à definição do objetivo de aprendizagem com tecnologias da WS. A especificação de metadados, bem como, as funcionalidades, processos e atividades que o RDE deve realizar para refinar os resultados de busca, corroboram com sua utilização. No âmbito do desenvolvimento da modelagem tem-se a ontologia geral da modelagem que concentra desde a modelagem de conceitos que representam o objetivo de aprendizagem às ontologias dos domínios objetivo de aprendizagem, assunto e público-alvo. Essa ontologia está em XML/RDF e incorpora definições OWL para formalizar a semântica dos VCs adotados. Logo, as ontologias de domínio constituem ontologias separadas, em OWL, oportunamente integradas à ontologia geral, como pode ser observado no recorte do código fonte da Figura 31. A disponibilização da modelagem está relacionada à sua aplicação, assim, seus arquivos contêm sintaxe e semântica dos principais conceitos e ontologias, conforme especificações das camadas da WS adotadas. Para permitir a implementação em RDEs na área de saúde, o quadro Integração de metadados e valores propostos pela modelagem (Quadro 10) relaciona metadados DC e LOM aos conceitos, além de indicar seus valores de preenchimento.

```

<?xml version="1.0"?>
<!DOCTYPE rdf:RDF [
  <rdf:RDF xmlns="http://www.owl-ontologies.com/Ontology1356747637.owl#"
    xml:base="http://www.owl-ontologies.com/Ontology1356747637.owl"
    xmlns:sqwrl="http://sqwrl.stanford.edu/ontologies/built-ins/3.4/sqwrl.owl#"
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
    xmlns:owl="http://www.w3.org/2002/07/owl#"
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  /]
  <owl:Ontology rdf:about="">
    <owl:imports
rdf:resource="http://swrl.stanford.edu/ontologies/3.3/swrla.owl"/>
      <owl:imports      rdf:resource="http://sqwrl.stanford.edu/ontologies/built-
ins/3.4/sqwrl.owl"/>
    </owl:Ontology>
    <rdfs:Class rdf:ID="Assunto">
      <rdfs:subClassOf rdf:resource="#Dominio"/>
    </rdfs:Class>
    <owl:ObjectProperty rdf:ID="taxonomia">
      <rdfs:domain rdf:resource="#Assunto"/>
      <rdfs:range rdf:resource="#ICF_Ontology"/>
    </owl:ObjectProperty>
    <owl:ObjectProperty rdf:ID="e_sobre">
      <rdfs:domain rdf:resource="#Objetivo_de_aprendizagem"/>
      <rdfs:range rdf:resource="#Assunto"/>
    </owl:ObjectProperty>
    <owl:Class rdf:about="http://who.int/icf#Attention_functions">
      <rdfs:subClassOf
rdf:resource="http://who.int/icf#Specific_mental_functions"/>

```

```

    <rdfs:label rdf:datatype="&xsd:string"
      >Attention functions</rdfs:label>
  </owl:Class>
  <rdfs:Class rdf:ID="ICF_Ontology">
    <rdfs:subClassOf rdf:resource="#Dominio"/>
  </rdfs:Class>
  <owl:ObjectProperty rdf:ID="classificacao">
    <rdfs:domain rdf:resource="#Assunto"/>
    <rdfs:range rdf:resource="#Dimensao_do_conhecimento"/>
  </owl:ObjectProperty>
  <rdfs:Class rdf:ID="Classificacoes_e_categorias">
    <rdfs:subClassOf rdf:resource="#Conhecimento_conceitual"/>
  </rdfs:Class>
  <rdfs:Class rdf:ID="Classificar">
    <rdfs:subClassOf rdf:resource="#Entender"/>
  </rdfs:Class>
  <rdfs:Class rdf:ID="Comparar">
    <rdfs:subClassOf rdf:resource="#Entender"/>
  </rdfs:Class>
  <rdfs:Class rdf:ID="Concluir">
    <rdfs:subClassOf rdf:resource="#Entender"/>
  </rdfs:Class>

```

Figura 31: Recorte do código fonte da modelagem proposta

A implementação da modelagem é delimitada pelo processo de descrição e recuperação de OAs. A seguir são elencados os principais componentes da modelagem e respectivas funcionalidades do RDE. Para isso, destaca-se o contexto e requisitos que determinam onde e como a modelagem está inserida no repositório, o comportamento das suas funcionalidades, ou seja, registro da informação descritiva dos OAs em formulário de metadados, busca a partir da formulação de

estratégia em campo livre, apresentação de resultados conforme a indexação dos metadados. A proposição da modelagem vislumbra melhorias para essas funcionalidades e cria condições para que a descrição, busca e recuperação sejam mais efetivas. Assim, a arquitetura do RDE com a modelagem proporciona a visão geral do refinamento dos resultados de busca, conforme o objetivo da pesquisa.

Para especificar, visualizar e documentar o contexto, comportamento e arquitetura de implementação da modelagem são elaborados diagramas, seguindo recomendações da Unified Modelling Language (UML). A UML fornece linguagem e notação gráfica para modelagem de sistemas, amplamente utilizada pela engenharia de *software*. É adaptável e atende diferentes necessidades, tais como a modelagem de aplicações e serviços baseados na Web. Basicamente, a UML permite que desenvolvedores visualizem alterações ou sistemas completos (como o RDE) em diagramas padronizados, ou seja, representação gráfica da informação (KON, 2006).

A pesquisa utilizou *diagrama de casos de uso* para representar o contexto e requisitos do RDE, de acordo com a perspectiva do seu usuário. Já o *diagrama de atividade* ilustra o comportamento do RDE, dada a implementação da modelagem proposta pela pesquisa. Por fim, o *diagrama de componentes* apresenta a arquitetura do RDE com a modelagem e especificação de tecnologias da WS adotadas (KON, 2006). Em síntese, as relações e associações entre as partes dos diagramas são:

- Seta pontilhada - dependência da parte com a ponta da seta;
- Seta em linha - generalização da parte com a ponta da seta;
- Ligação sem seta - interconexão, associação entre partes;
- Nome da associação indica a natureza e direção.
- Seta com losango vazio - agregação das partes de um todo;
- Seta com losango cheio - composição de partes de um todo, agregadas e inseparáveis.

Contexto e requisitos

O contexto para o qual a modelagem é prevista foi detalhado na pesquisa documental e considera os OAs disponibilizados em RDEs na área de saúde, descritos em formulário de metadados (processo de descrição) para atender a necessidade de busca do usuário (processo de recuperação) que tem interesse em conteúdos dessa área para planejar ou desempenhar processos de ensino-aprendizagem, ou seja, são educadores ou aprendizes.

O processo de descrição requer a definição de um conjunto de metadados e o preenchimento dos mesmos com informação a cerca do OA. O processo de recuperação é composto pela interrogação ao RDE, por meio de estratégia de busca que representa a necessidade do usuário, e apresentação de resultados. Esses processos contam com a interação de dois tipos de atores, Usuário e RDE. O Usuário interage com o RDE na descrição do OA no formulário de metadados e também interage com o RDE na recuperação de OAs com a composição de estratégia de busca. São ações distintas, mas ambas realizadas através de interface que facilita a interação do usuário. O RDE, por sua vez, armazena instancias de metadados que descrevem o OA e também os indexa para verificação e apresentação de resultados de busca durante a recuperação da informação. As funcionalidades que envolvem esses processos demandam o armazenamento e consulta ao banco de dados do RDE.

No desenvolvimento da modelagem (Capítulo 7) foram recomendadas tecnologias da WS, discutidos conceitos que especificam o objetivo de aprendizagem, ontologias e metadados que representam esses conceitos. Como a modelagem prevê o refinamento dos resultados de busca do RDE, as etapas que culminaram na proposta foram cruciais e especificam requisitos para tanto. Um requisito corresponde à funcionalidade ou condição que o RDE satisfaz com a implementação da modelagem. O *diagrama de caso de uso* apresentado a seguir (Figura 32) indica as partes da modelagem propostas, os processos que delimitam seu escopo e os atores interagem no contexto dos RDEs.

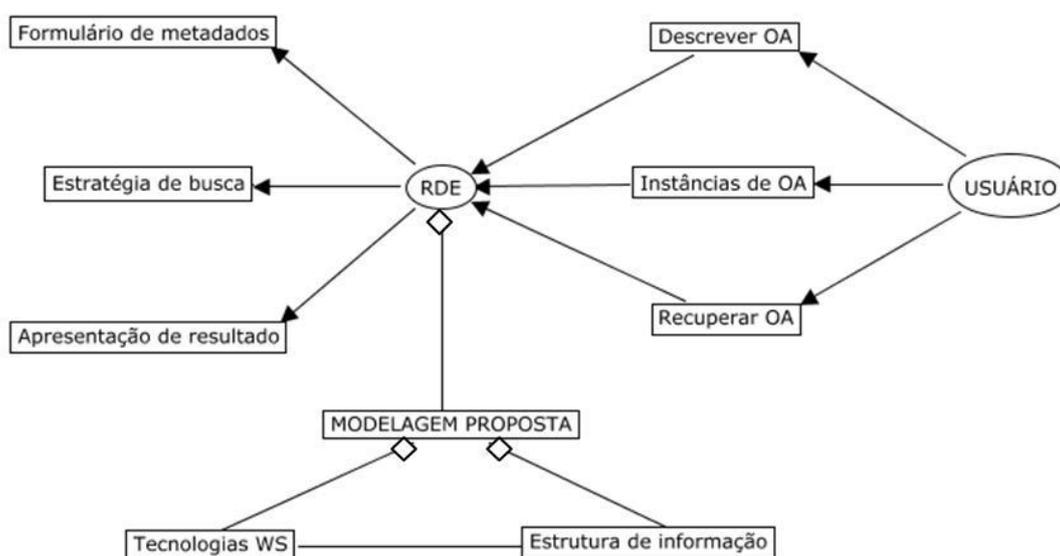


Figura 32: Diagrama de caso de uso – contexto e requisitos da modelagem proposta.

O *diagrama de caso de uso* (Figura 32) indica a interação entre os atores, usuário e RDE, através de ações que permitem a descrição e recuperação de OAs. O refinamento do

resultado de busca demanda o cumprimento de requisitos que estabelecem condições para que a recuperação da informação seja favorecida com a realização de inferências semânticas. A estruturação da informação descritiva e as tecnologias da WS definem essas condições e são agregadas ao fluxo de disponibilização de OAs em RDEs. O comportamento desse fluxo é explicitado a seguir.

Comportamento

O comportamento do fluxo de disponibilização, composto pelos processos de descrição e recuperação da informação, é balizado pela necessidade de estruturar sintática e semanticamente os metadados, garantindo a interpretação automática da informação descritiva dos OAs. O *diagrama de atividades* apresentado a seguir ressalta as atividades que constituem esse fluxo. A atividade é realizada por um ator (Usuário) que interage com outro (RDE) para consecução de ações que culminam no êxito dos processos (Figura 33).

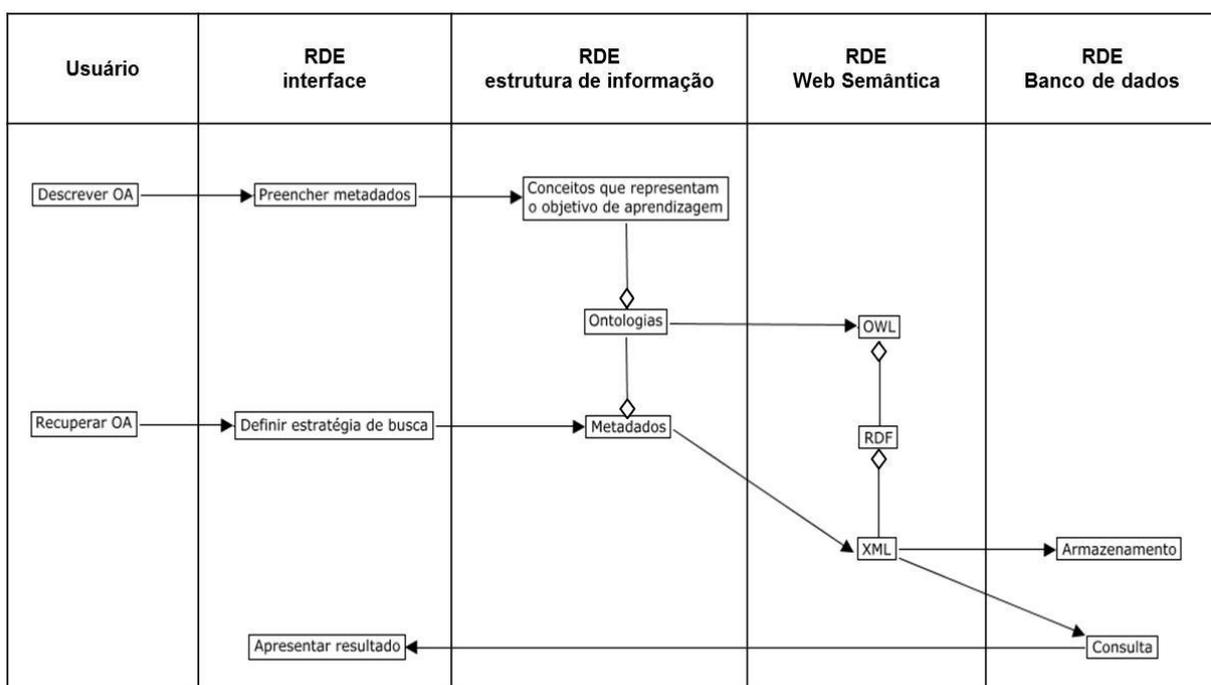


Figura 33: Diagrama de atividades – fluxo de disponibilização.

O fluxo de disponibilização determina a interação do Usuário com o RDE a partir da sua interface. O RDE possui diferentes partes com componentes que fornecem estrutura sintática e semântica aos metadados. As principais funções desses componentes são resumidas a seguir:

1. Apresentar formulário para preenchimento de metadados de acordo com a estrutura de informação proposta e semântica formalizada pelas tecnologias da WS;
2. Armazenar metadados no banco de dados;
3. Indexar metadados que especificam o objetivo de aprendizagem;
4. Consultar metadados indexados de acordo com a estratégia de busca do usuário;
5. Realizar inferências de acordo com a modelagem proposta;
6. Refinar resultado de busca de acordo com a consulta de metadados e as inferências realizadas;
7. Apresentar resultado de busca.

A recuperação de OAs, como proposto pela pesquisa, requer a formalização semântica dos metadados que especificam o objetivo de aprendizagem. Isto permite relacionar OAs semanticamente e apresentar resultados de busca pertinentes. Para tanto, a estrutura da informação descritiva e as tecnologias da WS adequadas são apresentadas na arquitetura de implementação da modelagem em RDEs.

Arquitetura

A arquitetura demonstra como a modelagem proposta deve ser implementada em RDEs, observado o contexto e o comportamento do fluxo de disponibilização de OAs, já apresentados. Descreve-se genericamente como os componentes da modelagem contribuem para que o usuário do RDE possa ser satisfeito ao buscar um OA, considerando a semântica do objetivo de aprendizagem, a partir da informação descritiva registrada em metadados. O *diagrama de componentes* apresenta metadados, tecnologias para defini-los sintaticamente e para formalizar sua semântica, possibilitando o armazenamento e consulta ao banco de dados do RDE.

A arquitetura está dividida em dois níveis diferentes, mas complementares, *Nível de descrição* e *Nível de recuperação de OA*, cada um com devidas especificações das tecnologias da WS adotadas. É importante lembrar que foram modelados conceitos que especificam o objetivo de aprendizagem, propostas ontologias que formalizam a intensão e extensão desses conceitos e, então, estabelecidos metadados que representam os conceitos. Além disso, os metadados foram relacionados semanticamente e determinados valores de preenchimento, assumindo posição fundamental para a proposta da pesquisa. Por esse motivo, a arquitetura destaca metadados utilizados.

No *Nível de descrição* e *Nível de recuperação de OA* são consideradas as especificações abordadas na seção *Integração dos metadados a modelagem da pesquisa* (Seção 7.2.3.3), que ainda fornece a correlação entre DC e LOM. A relação entre metadados é orientada pelo modelo abstrato DC, DCMI Abstract Model. Assim, OA é o recurso descrito com declarações sobre suas características, especificadas por propriedades e valores. Cada declaração instancia um par propriedade/valor. Quando o valor da propriedade denota relação entre metadados ele é valor não-literal, quando apresenta a informação descritiva em si é valor literal ou valor não-literal, para casos em que são adotados esquemas de codificação de vocabulário.

O *Nível de descrição de OA* indica como a informação descritiva a cerca do objetivo de aprendizagem é representada com as tecnologias da WS. Em síntese, a arquitetura deste nível constitui:

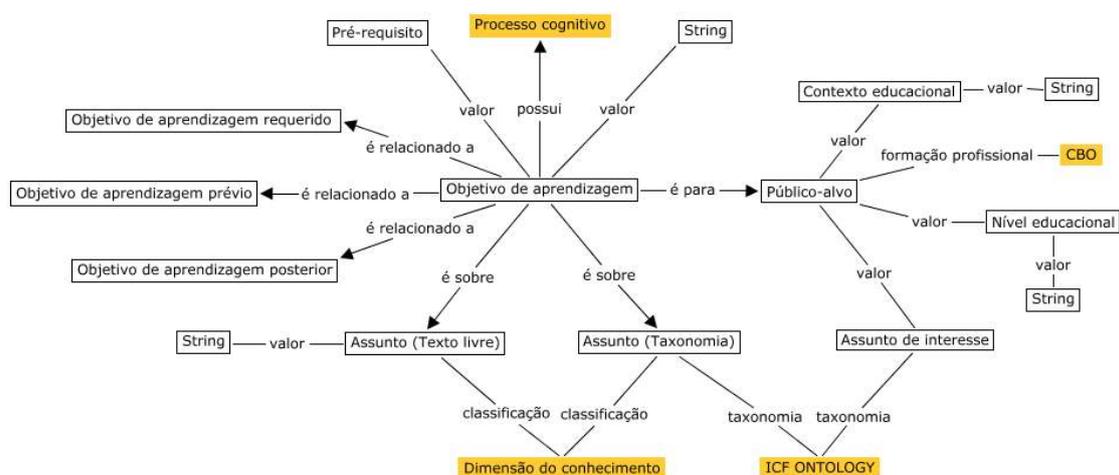
- Definição de metadados - sintaxe dos conceitos que especificam o objetivo de aprendizagem e são representados por metadados no RDE;
- Semântica de metadados - formalização da relação semântica entre os metadados que possibilitam inferências na recuperação da informação;
- Preenchimento de metadados - delimitação de valores de preenchimento de metadados e armazenamento de instâncias que descrevem OAs.

A definição de metadados se dá com o uso do XML e XML Schema. Na composição da instância de metadados que descreve OA todos os metadados utilizados são relacionados sintaticamente. Os metadados são nomeados com URIs no padrão Unicode, estabelecidos Namespaces de acordo com o DCMI Abstract Model para aqueles que podem ser generalizados em uma classe para processamento computacional comum. Assim, são agrupados na *Classe Objetivo*: processo cognitivo, objetivo de aprendizagem, objetivo de aprendizagem requerido, objetivo de aprendizagem prévio e objetivo de aprendizagem posterior. Na *Classe Assunto*: dimensão do conhecimento, assunto (taxonomia), assunto (texto livre) e assunto de interesse. E na *Classe Público-alvo*: público-alvo, nível educacional e contexto educacional.

A semântica de metadados utiliza RDF e RDF Schema. Dessa forma, são compostos grafos com propriedades e valores que especificam os metadados, relacionados com predicados, como exemplo, a tripla da *Classe Público-alvo* “Público-alvo (sujeito) possui formação profissional (predicado) estabelecida pela CBO (objeto)”. Os conceitos representados por metadados possuem propriedades extrínsecas que fazem referência a outros metadados e propriedades intrínsecas que indicam características do próprio metadado. As propriedades extrínsecas são observadas em metadados que representam

conceitos mais gerais. Para ilustrar, as propriedades extrínsecas do objetivo de aprendizagem são: público-alvo, assunto, processo cognitivo, objetivo de aprendizagem requerido, objetivo de aprendizagem prévio e objetivo de aprendizagem posterior; e as propriedades intrínsecas do público-alvo são: formação profissional, nível educacional, contexto educacional e assunto de interesse.

A figura a seguir ilustra metadados da proposta (Figura 34), com propriedades extrínsecas indicadas por seta em linha, propriedades intrínsecas por ligação sem seta, predicados informados textualmente e adoção de VC destacada por retângulos na cor



amarelo ocre.

Figura 34: Metadados da modelagem proposta.

A delimitação dos valores de preenchimento de metadados, quando prevê a adoção de VC, utiliza OWL para estabelecer disjunções, equivalências e cardinalidade. Assim, ICF Ontology e CBO são expressas em OWL e dispõem termos para o preenchimento dos metadados assunto e formação profissional, respectivamente. Isso permite observar que a OWL é adotada para representar explicitamente o significado de termos e as relações entre estes termos para diminuir ambiguidades, possibilitar generalizações e especificações.

O armazenamento de instâncias de metadados na base de dados do RDE é em formato XML e RDF. Vale lembrar que o conteúdo marcado pelas tags XML com informação descritiva referente ao metadado é tratado como elemento de informação, considerado um objeto ao qual se pode atribuir características específicas, assim, o valor de preenchimento do metadado pode ser também propriedade, como um termo da ICF Ontology (valor do assunto) visto como propriedade com dimensão do conhecimento estabelecida como valor.

O *Nível de recuperação de OA* indica como a informação descritiva a cerca do objetivo de aprendizagem é recuperada para apresentar resultados de buscas refinados, conforme o objetivo da pesquisa. Em síntese, a arquitetura deste nível constitui:

- Indexação - valores de preenchimento de metadados são indexados para possibilitar a busca;
- Inferência - identificação e interpretação da semântica da informação descritiva do OA que relaciona estes de acordo com o objetivo de aprendizagem;
- Resultado - verificação dos termos da estratégia de busca do usuário nos valores de preenchimento de metadados indexados e realização de inferências para apresentar resultados refinados.

Os metadados indexados são insumos para a busca, ou seja, são os valores preenchidos e armazenados no banco de dados, consultados durante a recuperação da informação. Mais especificamente, a busca inicia-se com o estabelecimento da conexão com o banco de dados, arquivos de metadados e ontologias. A estratégia de busca do usuário constitui parâmetros de consulta ao banco de dados. Os parâmetros recebidos são checados nos campos das tabelas de metadados. Após a consulta ao banco de dados são verificadas regras de inferência estabelecidas pela modelagem. Algumas inferências são viabilizadas unicamente com a adoção da RDF outras, com mais expressividade, recorrem a OWL.

A recuperação de OAs requer a comunicação entre as partes da modelagem implementadas no RDE. Assim, resultados podem ser expandidos com OAs semanticamente relacionados, mas que não apresentam os termos buscados pelo usuário em sua instância de metadados, ou resultados podem ser classificados conforme coincidências semânticas da sua informação descritiva. As duas opções são consideradas e o refinamento do resultado de busca elenca OAs com termos buscados, classificados de acordo com coincidências semânticas e OAs sem termos buscados, mas com relações semânticas, classificados de acordo com coincidências semânticas. A apresentação do resultado de busca com sintaxe XML é formatada com folhas de estilo em linguagem XSL (eXtensible Style Language).

Ao buscar e organizar semanticamente os resultados tem-se a recuperação semântica de OAs. A recuperação semântica considera inferências RDF e OWL e demanda a consulta SPARQL. A SPARQL é uma camada da WS que perpassa do RDF Schema a OWL, dispõe de linguagens e protocolos para consultar e manipular grafos RDF, além de apoiar a delimitação do número de resultados e sua classificação. Para implementação

recomenda-se a RDF API, que inclui motor de consulta SPARQL para realizar consultas baseadas em ontologias e arquivos RDF/XML. Complementarmente, a recuperação semântica pode ser aperfeiçoada através de uma interface baseada no servidor SPARQL Server Joseki, que suporta o protocolo SPARQL e a linguagem SPARQL RDF Query Language.

A figura a seguir (Figura 35) ilustra a arquitetura a partir do *diagrama de componentes* baseados nas tecnologias da WS.

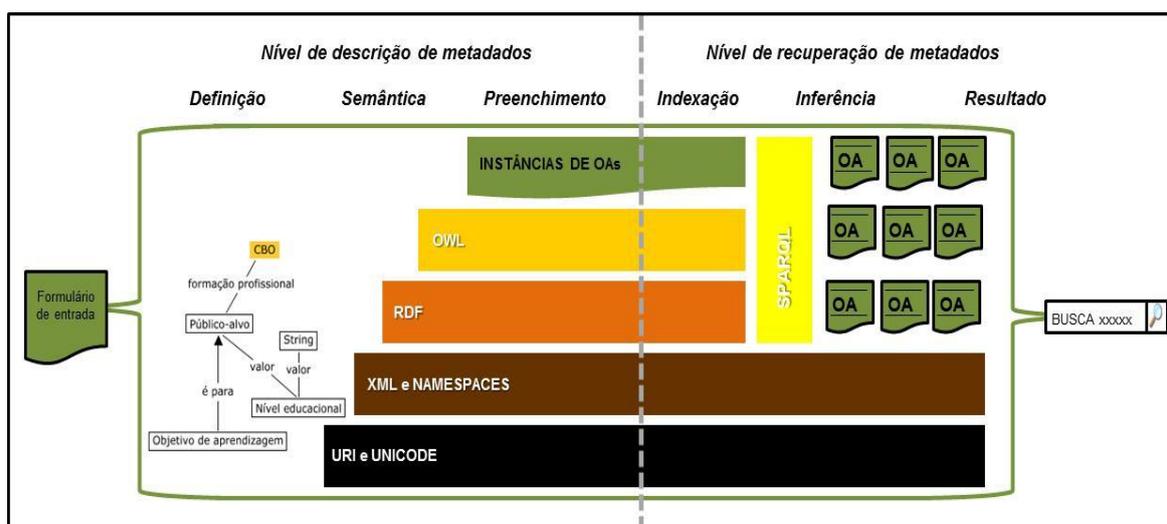


Figura 35: Arquitetura da modelagem proposta.

O *diagrama de casos de uso*, *diagrama de atividade* e *diagrama de componentes* fornecem a perspectiva da implementação da modelagem em RDEs e demonstram sua aplicabilidade. A sistematização realizada nesta seção e o capítulo que discorre sobre o desenvolvimento da modelagem (Capítulo 7), constituem documentação a ser consultada para implementação. Além disso, as considerações finais (**Seção X**) abordam constatações importantes para o êxito na recuperação semântica de OAs. Para elucidar contribuições da modelagem, a seção seguinte (**Seção X**) apresenta o cenário de implementação e simulações do resultado de busca.

8.2 Simulação do resultado de busca em RDE na área de saúde com a implementação da modelagem proposta

A simulação do resultado de busca em RDE na área de saúde considera a implementação da modelagem desenvolvida e sistematizada. Esta seção discute contribuições da pesquisa para recuperação de OAs, de acordo com a hipótese comprovada. São ressaltadas as principais características do RDE e investigadas

possíveis adaptações, com base no cenário de implementação da modelagem em RDE condizente com o contexto. Por fim, são comentadas inferências para refinamento dos resultados de busca e apresentados exemplos. Para facilitar a compreensão do texto, **metadados** são transcritos em negrito, padrão de “metadados propostos” entre aspas, padrão de metadados do RDE utilizado como cenário da implementação sublinhado e *valores de preenchimento* de metadados em itálico.

Principais características do RDE

O RDE possui documentos e funcionalidades focadas no ensino-aprendizagem. Essas funcionalidades subsidiam a disponibilização de OAs com uso de metadados padronizados. RDEs oferecem serviços de busca que permitem acesso aos OAs e, para isso, dispõem de informação descritiva suficiente à seleção, especialmente relacionadas aos aspectos educacionais.

A alimentação dos RDEs muitas vezes adota o auto-arquivamento, ou seja, os próprios autores fazem a submissão de seus OAs com o preenchimento de metadados e *upload* de arquivos. Nesse sentido, recomenda-se o estabelecimento de regras para a descrição, inclusive com adoção de VCs. A submissão dos OAs se dá a partir da interface do repositório, assim como a busca.

Na área de saúde, tem crescido a quantidade de repositórios, especialmente os educacionais, para promover equidade e capacitação aos profissionais com a disseminação de materiais educacionais.

Cenário de implementação

Com intuito de identificar possíveis adaptações, necessárias à implementação da modelagem, foi escolhido um repositório com as características citadas e analisada sua estrutura de informação. Assim, constitui-se o cenário de implementação com detalhamento do fluxo de disponibilização dos OAs neste RDE, para discutir sua situação atual e projetar melhorias. O Acervo de Recursos Educacionais em Saúde (ARES), já investigado na pesquisa documental, foi escolhido, pois possui objetivos, documentos e metadados condizentes com a proposta (Figura 36).

The screenshot shows the ARES website interface. At the top, there is a blue navigation bar with the 'Acervo' logo on the left and 'Universidade Aberta do SUS' in the center. To the right of the navigation bar is a search bar with the placeholder text '<<Escolha um serviço;>>'. Below the navigation bar, the main header area features the 'UNA-SUS' logo and the title 'Acervo de Recursos Educacionais em Saúde'. A secondary navigation bar includes a 'Página Inicial' link. The main content area is divided into two columns. The left column, under the heading 'Submissões recentes', lists three educational resources with their titles, authors, and types. The right column contains three menu sections: 'Visualizar' with options like 'Todo o ARES', 'Coleções', 'Autor', 'Título', 'Data', 'Palavras-chave DeCS', 'Público-Alvo', 'Área temática', and 'Tipo de recurso'; 'Minha conta' with an 'Entrar' button; and 'Filtros' with options like 'Coleção', 'Objetos Simples (195)', 'Objetos de Aprendizagem (152)', 'Modelos de Cursos (30)', and 'Área temática'.

Figura 36: Acervo de Recursos Educacionais em Saúde (ARES).

O ARES é iniciativa da Secretaria de Gestão do Trabalho e da Educação na Saúde (SGTES) do Ministério da Saúde do Brasil (MS) e foi desenvolvido no âmbito da Universidade Aberta do Sistema Único de Saúde (UNA-SUS) para disseminar conteúdos produzidos e utilizados por suas instituições parceiras em ofertas de capacitação. Esses conteúdos são chamados de recurso educacional, entendido como documento digital utilizado no ensino-aprendizagem. Todo recurso educacional, ao ser incorporado ao acervo, deve ter definido seu objetivo educacional ou de aprendizagem (BRASIL, 2011). Os tipos de recursos educacionais contemplados, de acordo com a forma de apresentação do conteúdo, são: animação, áudio, diagrama, gráfico, imagem e material interativo.

A abrangência temática do ARES é delimitada pelo escopo da educação dos profissionais em saúde, que define categorias para o armazenamento dos recursos educacionais. Diante disso, o assunto principal desse acervo é Saúde e os temas atendem áreas prioritárias do MS, tais como: Atenção Primária / Saúde da Família, Doenças crônicas, Doenças infecciosas, Enfermagem, Saúde Bucal e Saúde da Pessoa Portadora de Deficiência (BRASIL, 2011).

Os recursos educacionais estão divididos em coleções específicas de acordo com suas características. A Coleção Modelos de Cursos possui cursos completos ou módulos de cursos. Módulo é a primeira subdivisão do curso, independente da nomenclatura originalmente utilizada, como eixo, parte, etc., desde que, apresente sentido educacional quando tomado isoladamente. A Coleção Objetos de Aprendizagem possui partes constituintes dos módulos que possuem objetivos de aprendizagem, conteúdo, atividade e avaliação. A Coleção Objetos Educacionais Simples possui recursos educacionais isolados ou sem agregação, tais como conteúdos e atividades, com objetivo educacional

estabelecido. Os objetos dessa Coleção podem ser utilizados para compor objetos de aprendizagem, módulos ou cursos. A Coleção Itens de avaliação contempla itens de avaliação formativa e são tipos específicos de objetos simples ou objetos de aprendizagem (BRASIL, 2011).

A disponibilização de recursos educacionais e armazenamento de metadados no banco de dados compreendem o fluxo geral do ARES, desenvolvido na plataforma de software DSpace⁶⁷. Com este fluxo se dá a formação do acervo:

“Os recursos são selecionados em bibliotecas, repositórios ou acervos pessoais; readequados ou produzidos por instituições parceiras. Em todos os casos os direitos autorais devem ser verificados e o regime adotado deve ser descrito, garantindo-se o cumprimento das regras para sua disponibilização. Já no sistema do repositório, o responsável pela submissão ou responsável pela validação, após fazer o *login*, deve catalogar e fazer o carregamento do(s) arquivo(s) do recurso educacional a partir de um processo de *upload* para armazenamento físico dos dados. [...] Após a avaliação dos aspectos conceituais, educacionais e descritivos, os recursos educacionais são publicados no ARES para busca e recuperação” (BRASIL, 2011, p. 6).

O formulário de metadados do ARES permite o tratamento da informação dos recursos educacionais disponibilizados. Na etapa de catalogação do fluxo é realizada descrição, conforme diretrizes estabelecidas no Manual de descrição e preenchimento de metadados (BRASIL, 2011). Os metadados utilizados estão no padrão DC, estendido e qualificado com base em especificações LOM, agrupados em: Classificação, Descrição geral, Descrição educacional, Descrição temática, Direitos autorais e Submissão. Além disso, o ARES adota VC para alguns metadados, como DeCS para **Assunto** e CBO para **Público-alvo** (BRASIL, 2011).

O ARES dispõe de busca simples de todos metadados indexados e também permite realizar busca avançada, em Coleções específicas e/ou combinando metadados com operadores *booleanos* (OU, E, NÃO). Os resultados de busca podem ser exibidos em ordem alfabética, crescente ou decrescente, e em ordem cronológica, do mais antigo para o mais recente ou do mais recente para o mais antigo, de acordo com a data de publicação (BRASIL, 2011).

Diante disso, a implementação da modelagem proposta requer adaptações, discutidas de acordo com a Arquitetura apresentada.

⁶⁷ Link: <<http://www.dspace.org/>>.

Em *Nível de descrição de OA*, o ARES adota sintaxe XML e Namespaces do DC para definição de metadados. A semântica desses metadados precisa ser formalizada, bem como suas relações. Valores de preenchimento de metadados com VC utilizam dicionários em tabelas do banco de dados e *webservices*⁶⁸ que compõem listas de opções, fazendo-se necessário a adoção da OWL para representação semântica adequada.

A arquitetura atual do ARES permite apenas a verificação sintática de hierarquias de **Assunto** e **Público-alvo**, por exemplo. A lista de opções composta pela CBO utiliza o formato NNNNXX para Famílias, preservando sua codificação (*2251XX::Médicos clínicos* e *225122::Médico cancerologista pediátrico*, por exemplo). Seguindo essa orientação, o DeCS está no formato Código hierárquico::Descrição (*C01.252.410.040.552.846::Tuberculose* e *C01.252.410.040.552.386::Hanseníase*, por exemplo). Esses formatos favorecem a implementação da modelagem e garantem que os recursos educacionais já descritos possam ser representados semanticamente sem retrabalho. O armazenamento de instâncias que descrevem recursos educacionais é em XML e deverá atender também RDF, após a implementação.

Como já discutido amplamente, a modelagem proposta estabelece conceitos que especificam o objetivo de aprendizagem, representados por metadados. Assim, além das adaptações citadas, é importante certificar a adoção dos metadados propostos, bem como, valores de preenchimento previstos. O quadro a seguir (Quadro 11) sintetiza metadados adotados no ARES e recomenda possíveis adaptações para atender a modelagem, quando for o caso.

Quadro 11: Metadados ARES, metadados da modelagem proposta e recomendações.

Conceito	DC ARES	Recomendações
Objetivo de aprendizagem	unasus.educacionalDescription	Adotar metadado previsto na modelagem.
Objetivo de aprendizagem requerido	dc.relation.{relation} Qualificador relation: requires isrequiredby	-
Objetivo de aprendizagem prévio	dc.relation.{relation} Qualificador relation: haspart ispartof	-
Objetivo de	dc.relation.{relation}	-

⁶⁸ *Webservice* é utilizado na integração de sistemas e na comunicação entre aplicações diferentes. São componentes que permitem às aplicações enviar e receber dados em formato XML (Wikipédia, 2012).

aprendizagem posterior	Qualificador relation: haspart ispartof	
Processo cognitivo	-	Adotar metadado previsto na modelagem.
Assunto (Taxonomia)	unasus.subject.decs Lista de opções: DeCS	Adotar ontologia.
Assunto (Texto livre)	dc.subject.keyword	Substitui metadado previsto na modelagem.
Dimensão do conhecimento	-	Adotar metadado previsto na modelagem.
Público-alvo	unasus.audience.occupation Lista de opções: Subconjunto CBO	Substitui metadado previsto na modelagem e adotar ontologia.
Nível educacional	unasus.learningContext Lista de opções: Estrutura de ensino brasileira	Substitui metadado previsto na modelagem.
Contexto educacional	unasus.audience.context	Substitui metadado previsto na modelagem.
Assunto de interesse	-	Adotar metadado previsto na modelagem e adotar ontologia.

A definição dos metadados ARES utiliza o Dublin Core Metadata Element Set e não o DCMI Terms, por isso alguns metadados que poderiam ser contemplados pelo DC, como o **Objetivo de aprendizagem** (“instructionalMethod”), adotam qualificações específicas para aplicação (unasus.educacionalDescription). O metadado instructionalMethod é usado para especificar **Recomendações de uso educacional** no ARES, mas recomenda-se sua adoção conforme a proposta (BRASIL, 2011).

Na Descrição temática também é usado metadado unasus.classification para especificar **Área temática**, isso atende a necessidade de categorizar **Assuntos** especificados em unasus.subject.decs e dc.subject.keyword, o que pode ser substituído pela adoção de ontologias para **Assunto (Taxonomia)**. Os valores de preenchimento dos metadados **Assunto (Taxonomia)** e **Público-alvo** no ARES já preveem VCs, mas não adotam ontologias OWL para formalização semântica recomendada, tal como deve ocorrer no metadado não contemplado **Assunto de interesse**.

Os metadados que especificam relações entre OAs e, conseqüentemente, entre objetivos de aprendizagem, atendem a proposta. Uma relação não contemplada na

modelagem é definida para indicar **versão** no ARES, ou seja, recursos educacionais com mesmo objetivo de aprendizagem, mas com conteúdos diferentes. Essa relação pode ser útil para proposta.

A maior parte dos metadados do ARES atendem a modelagem e estão agrupados em Descrição educacional e temática. Os metadados adotados para **Assunto (Texto livre)** (dc.subject.keyword), **Público-alvo** (unasus.audience.occupation), **Nível educacional** (unasus.learningContext) e **Contexto educacional** (unasus.audience.context) são utilizados para representar os mesmos conceitos previstos na modelagem e podem substituir metadados proposto sem perda semântica. Os metadados **Processo cognitivo** (“processoCognitivo”), **Dimensão do conhecimento** (“subjectDimCon”) e **Assunto de interesse** (“assuntoInteresse”) precisam ser definidos conforme a modelagem proposta, pois não são contemplados pelo ARES.

Em *Nível de recuperação de OA*, o ARES indexa metadados e possibilita recuperação da informação a partir da estratégia de busca do usuário. É realizada consulta ao banco de dados, porém, não há inferências para refinar o resultado de busca, o que será viabilizado pela implementação da modelagem. Assim, recomendações e possíveis adaptações são importantes. Após essas observações e dada a implementação da modelagem proposta no ARES, são comentadas inferências.

Refinamento dos resultados de busca

Considerando o cenário que discute a implementação da modelagem proposta no ARES, algumas inferências para refinamento dos resultados de busca são comentadas a partir de exemplos⁶⁹ de recuperação da informação. As inferências são afirmações baseadas na semântica dos conceitos que especificam o objetivo de aprendizagem dos OAs. Para tanto, a semântica é formalizada com a estrutura da informação e tecnologias da WS propostas.

No desenvolvimento da modelagem (Capítulo 7) foram consideradas questões a serem respondidas para a realização de inferências. Dessa forma, a implementação da modelagem no ARES permite respostas a questionamentos referentes ao seu acervo, que contribuem com o refinamento dos resultados de busca.

Ao analisar o atendimento das questões propostas, o ARES apresentará recursos educacionais classificados de acordo com processo cognitivo, o que permite ordenação da complexidade do que se espera do público-alvo ao fim do processo de ensino-aprendizagem

⁶⁹ Os exemplos apresentados são meramente ilustrativos, baseados em valores de preenchimento dos VCs propostos, com livre tradução dos termos da ICF Ontology. O ARES não dispõe de recursos educacionais suficientes para a comparação de resultados de busca antes e depois da implementação da modelagem proposta.

(Exemplo 1). O público-alvo será associado ao objetivo de aprendizagem, conforme as famílias de ocupação da CBO e suas atividades. Ademais, as famílias de ocupação podem apresentar assuntos de interesse diretamente relacionados às atividades a serem desempenhadas. Sendo esses assuntos de interesse definidos pela mesma ontologia do assunto, conteúdos específicos são pertinentes para certo público-alvo (Exemplo 2). O nível educacional e o contexto educacional são informação a mais considerada nas inferências sobre o perfil do público-alvo.

Exemplo 1:

Título	Recurso educacional 001
Objetivo de aprendizagem	Diferenciar funções sensoriais relacionadas a visão
Processo cognitivo	Analisar, diferenciar
Público-alvo	Médico clínico

Título	Recurso educacional 002
Objetivo de aprendizagem	Executar procedimento cirúrgico para reabilitação das glândulas lacrimais
Processo cognitivo	Aplicar, executar
Público-alvo	Médico oftalmologista

Exemplo 2:

Título	Recurso educacional 003
Objetivo de aprendizagem	Classificar níveis de sensibilidade à luz
Assunto	Sensibilidade à luz, qualidade de visão
Público-alvo	Médico oftalmologista

Título	Recurso educacional 004
Objetivo de aprendizagem	Classificar níveis de sensibilidade ao contraste
Assunto	Sensibilidade ao contraste, qualidade de visão
Público-alvo	Médico oftalmologista

A busca por assunto poderá ser expandida com termos mais genéricos e mais específicos, além de controlar sinônimos, traduções e distinguir significados, tais como, *corda vocal* (Estrutura do corpo envolvendo voz e fala) e *corda vocal* (qualidade de Função do corpo envolvendo voz e fala). O assunto considerará que o objetivo de aprendizagem é sobre este e será classificado de acordo com a dimensão do conhecimento. Assim, assuntos não representados na estratégia de busca do usuário poderão ser considerados observando a dimensão do conhecimento e a hierarquia de termos (Exemplo 3).

Exemplo 3:

Título	Recurso educacional 005
Objetivo de aprendizagem	Implementar critérios para estímulo do contato físico entre pacientes com deficiência psicomotora
Assunto	Interações interpessoais, contato físico
Dimensão do conhecimento	Critérios para usar procedimentos

Título	Recurso educacional 006
Objetivo de aprendizagem	Implementar critérios para estímulo do respeito e cordialidade nas relações interpessoais de pacientes com deficiência psicomotora
Assunto	Interações interpessoais, respeito e cordialidade nas relações interpessoais
Dimensão do conhecimento	Critérios para usar procedimentos

O resultado de busca apresenta recursos educacionais descritos com termos buscados e recursos educacionais sem termos buscados na sua descrição, mas com relações semânticas. Esse resultado é classificado de acordo com coincidências semânticas. Para identificar coincidências são verificadas definições de metadados da instância. Para ponderar coincidências considera-se a quantidade de relações semânticas entre recursos educacionais, verificado, respectivamente, os metadados das classes Objetivo, Assunto e Público-alvo. Dessa forma, recursos educacionais com o mesmo processo cognitivo e objetivo de aprendizagem requerido são ordenados antes de recursos educacionais com o mesmo público-alvo e nível educacional (Exemplo 4). Quando a estratégia de busca do usuário representar o assunto, os primeiros recursos educacionais listados serão sobre o assunto, ordenados conforme a relevância das classes de metadados indicada anteriormente.

Exemplo 4:

Título	Recurso educacional 007
Objetivo de aprendizagem	Comparar funções urinárias da mulher gestante e da mulher não gestante
Processo cognitivo	Entender, comparar
Objetivo de aprendizagem requerido	Recurso educacional 020
Público-alvo	Médico urologista
Nível educacional	Especialização

Título	Recurso educacional 008
Objetivo de aprendizagem	Comparar filtração de urina de pacientes com diferentes graus de diabetes
Processo cognitivo	Entender, comparar
Objetivo de aprendizagem requerido	Recurso educacional 020
Público-alvo	Médico clínico
Nível educacional	Graduação

Título	Recurso educacional 009
Objetivo de aprendizagem	Verificar incidência de incontinência urinária causada por redução da função renal
Processo cognitivo	Avaliar, verificar
Objetivo de aprendizagem requerido	Recurso educacional 033
Público-alvo	Médico urologista
Nível educacional	Especialização

Como discutido, a recuperação da informação irá considerar a semântica dos metadados e suas relações para apresentar resultados que abrangem a especificação do objetivo de aprendizagem. A seguir, abordam-se considerações finais, visto que os objetivos específicos da pesquisa foram cumpridos e resultados apresentados.

8.3 Considerações finais

Esta seção apresenta considerações observadas ao longo do desenvolvimento da pesquisa, que contribuem com a elucidação dos objetivos iniciais. Destaca-se que a metodologia adotada foi pertinente e os resultados obtidos com sucesso, assim, são discutidos nas subseções seguintes detalhes da utilização da ferramenta Protégé, aspectos relativos à modelagem proposta e ponderações sobre ontologias, metadados, RDE e WS.

8.3.1 Utilização da ferramenta Protégé

A ferramenta Protégé foi utilizada em todo desenvolvimento da modelagem proposta, desde a modelagem dos conceitos principais até a agregação de ontologias de objetivo de aprendizagem, assunto e público-alvo. A ontologia de objetivo de aprendizagem ampliou conceitos e relações semânticas essenciais para pesquisa. As ontologias de assunto e público-alvo expandiram a definição de conceitos e o estabelecimento de instâncias. Assim, a formalização dos conceitos que especificam objetivo de aprendizagem, permitiu a representação estrutural, sintática, lógica e semântica dos metadados a serem adotados em RDEs para descrição de OAs. Outras ferramentas também seriam aplicáveis, embora não tenham sido testadas com afinco, entre estas a OntoEdit⁷⁰.

A ferramenta facilitou a representação semântica com sua interface gráfica para desenvolvimento de ontologias. Seu código aberto e arquitetura modular permitem a inserção de novos recursos em estudos futuros. Apesar das vantagens proporcionadas, observaram-se limitações para manusear grandes quantidades de termos, especialmente em demonstrações visuais de conceitos relacionados. Isso foi contornado com “recortes” da modelagem em arquivos específicos. Além disso, sendo RDF e OWL tecnologias robustas para formalização semântica, tornou-se necessário a declaração de quaisquer definições importantes para o êxito da modelagem, como por exemplo, a ordenação gradativa dos processos cognitivos.

A instalação e configuração da Protégé foram a partir do *download* gratuito no *site*⁷¹ oficial. Para tanto, foi feito registro do projeto de utilização da ferramenta;

⁷⁰ Link: < <http://www.daml.org/tools/#OntoEdit>>.

⁷¹ Link: < <http://protege.stanford.edu/>>.

download completo da versão de interesse, acrescido de *plug-ins* opcionais e do Java Virtual Machine (JVM); instalação do arquivo executável e configuração dos componentes desejados; determinação de pasta no disco rígido para salvar arquivos; e criação do novo projeto com a escolha do formato adequado (XML/RDF).

Para especificar classes na Protégé, baseou-se na identificação dos conceitos gerais e, a partir desses, agrupadas especificidades com a definição de subclasses. Observou-se neste momento que o relacionamento expresso pela hierarquia de classes é do tipo *gênero-espécie*, com herança de atributos por parte das subclasses. Ao passo que se criava uma nova classe, identificava-se quais seriam seus atributos e valores, de modo que cada conceito foi trabalhado intensamente antes de dar início à representação de outro.

A determinação de atributos para classes compôs sua definição e relacionamentos. A definição dos atributos deu-se pela indicação de propriedades RDF e propriedades Objeto. De maneira geral, a propriedade RDF permitiu a criação de listas de valores de preenchimento (*allowed values*) pré-definidas (propriedade *nível educacional*, valores *graduação*, *mestrado*, *doutorado*) e a propriedade Objeto permitiu o relacionamento entre classes, com especificação de uma classe como valor de preenchimento (*range*) de outra classe (propriedade *formação profissional*, valor *CBO*). Também foram definidos domínios do atributo (*domain*), ou seja, em qual classe são contemplados; e tipo de valor (*value type*), como por exemplo, texto livre (*string*), classe, etc.

Além disso, atributos constituíram campos do formulário de entrada de dados para criação de instâncias. O preenchimento dos atributos da classe são exemplos, ou seja, a intensão do conceito representado. Observou-se que o atributo deve ser preenchido com um único valor, pois o detalhamento da propriedade implica na declaração de classe. Com isso, atributos com definições que extrapolam a delimitação do seu valor, foram definidos como classes. Isso pode ser constatado na definição da classe *processo cognitivo* que também é atributo da classe *objetivo de aprendizagem*, mas possui atributos próprios.

Quando identificado atributo comum ao conjunto de subclasses, esse foi definido na respectiva superclasse. Dessa forma, definiu-se a propriedade *atividade* para classe *competências de médicos clínicos* (família da CBO) e essa foi herdada por *médico dermatologista*, *médico do trabalho*, etc. (ocupações da família de médicos clínicos da CBO).

Os relacionamentos entre conceitos formaram a estrutura necessária para recuperação semântica de OAs. A pesquisa documental vislumbrou relacionamentos

possíveis e a investigação da estrutura de informação e tecnologias da WS complementaram a definição das relações semânticas. Para tanto, observou-se a importância de conhecer recursos da Protégé para representação de relacionamentos.

Diversos relacionamentos foram traçados, sendo o primeiro deles a própria relação categorial, estabelecida no momento em que os termos foram listados e discernidos entre conceitos e propriedades. Essa divisão inicial diminuiu erros lógicos no estabelecimento de outras relações, pois determinou a natureza do objeto a ser representado.

A hierarquização das classes delineou a estrutura de conceitos, pois a relação hierárquica estabelece o primeiro elemento da definição e permite relacionar objetos da mesma natureza, ou seja, classes e subclasses relacionadas de acordo com características comuns, expressas por atributos. Conforme a Teoria do Conceito de Dalhberg (1978), que define a relação hierárquica com base na lógica de implicação, os conceitos hierarquizados formaram cadeias de conceitos. Vale lembrar, que as classes são elemento agregador da ontologia, estabelecidas para facilitar o entendimento do conceito e não simplesmente classificar o domínio.

O relacionamento com instâncias permitiu que o conceito fosse completado com informação que não está apenas no atributo, mas sim nos exemplos da classe que, por sua vez, possui atributos próprios. O atributo *público-alvo* da classe *objetivo de aprendizagem*, por exemplo, foi definido com instâncias que apresentam *formação profissional, nível educacional e contexto educacional*, atributos da classe *público-alvo*.

Além de expressar relacionamentos, a criação de instâncias povoa a ontologia para compor base de conhecimento sobre o domínio. Porém, este estudo não teve foco na criação de instâncias e o povoamento apenas ilustrou a representação dos conceitos para melhorar sua compreensão. As instâncias foram criadas à medida da definição dos conceitos, tornando possível a revisão de classes, atributos e relacionamentos.

Destaca-se ainda, que o desenvolvimento da modelagem com a ferramenta Protégé viabilizou a definição de restrições para inferências (axiomas), afinal, representou adequadamente a sintaxe e semântica dos conceitos, para que aplicações computacionais sejam capazes de interpretá-las. Por exemplo, a definição de que o *assunto* é classificado de acordo com a *dimensão conhecimento*, restringe sua classificação. A classe *dimensão do conhecimento* indica valores permitidos para essa classificação.

8.3.2 A modelagem proposta

Com o desenvolvimento da pesquisa constatou-se a necessidade de reavaliar a modelagem proposta para identificar possíveis revisões de seus componentes. Isto ocorreu ao longo do estudo e configurou o retorno a algumas definições já estabelecidas para devidos ajustes. O esforço de elaborar figuras para ilustrar conceitos favoreceu a revisão e a eliminação de incoerências. No início da modelagem, foi adotada abordagem *bottom up* (NOY, 2001), partindo-se dos conceitos e propriedades que especificam o objetivo de aprendizagem, para posteriormente definir metadados representativos. Quando um conjunto maior de conceitos foi identificado para a estrutura de informação, valeu-se da abordagem *top down* (NOY, 2001) para criar hierarquias de valores de preenchimento de metadados.

No contexto deste estudo, a modelagem conceitual, passo inicial para desenvolvimento da modelagem, mostrou-se adequada à organização e estruturação dos principais conceitos. O modelo conceitual constituiu representação simplificada e abstrata para descrever e fornecer informação de referência. Após esse passo, foram incorporadas tecnologias da WS, garantindo que a ampliação do modelo corroborasse com a versão final da modelagem proposta.

A premissa de contribuir com o refinamento do resultado de busca em RDEs na área de saúde guiou o desenvolvimento da pesquisa. O estudo exploratório forneceu embasamento teórico para respaldar a estrutura de informação e a adoção de tecnologias da WS. No que diz respeito à estrutura de informação, foram identificadas opções para atender o objetivo e realizadas escolhas baseadas na disponibilidade de documentação, ampla utilização e melhor adequação aos propósitos do trabalho. Quanto às tecnologias da WS, procurou-se atender recomendações da W3C e as considerações de experiências anteriores.

Outra observação importante refere-se à sistematização da modelagem proposta, que forneceu visão geral da implementação em RDEs na área de saúde. Além disso, a simulação de resultados de busca refinados foi determinante para compreensão da recuperação semântica de OAs. Analisando o fato do interesse de busca do usuário não ser unicamente representado pelo assunto ou qualquer termo isolado, o aumento de RDEs e OAs disponibilizados e a autonomia na submissão e recuperação da informação na Web, o resultado deste trabalho torna-se ainda mais relevante.

Para além da recuperação semântica de OAs, a apresentação dos resultados de busca também colabora com a satisfação do usuário. Isso permitiu intuir, que árvores hiperbólicas e mapas com *links* para navegação pelos resultados, podem facilitar a visualização de suas relações semânticas por parte do usuário e ainda facilitar a redução ou aumento do nível de exaustividade (aumento de revocação). Com essa perspectiva, refinar os resultados de busca a partir dos objetivos de aprendizagem, permitirá ao usuário buscar um assunto e selecionar o tipo de relação semântica entre os resultados apresentados que melhor atende sua necessidade (aumento de precisão).

De maneira geral, a investigação de metadados que especificam o objetivo de aprendizagem possibilitou ampla discussão sobre a semântica dos conceitos que representam. Isso ressaltou observações para correta definição de metadados para modelagem, como destacado a seguir:

- Objetivo de aprendizagem remete ao que se vai aprender. “O que” é definido com processo cognitivo demandado ao aprendiz. “Como aprender” é a estratégia de aprendizagem. A ação, atividade ou comportamento possibilitado com a aprendizagem, tem forte relação com o objetivo de aprendizagem, podem ter a mesma natureza e exigir o mesmo processo cognitivo, mas não são, necessariamente, a mesma coisa;
- O processo cognitivo é uma ação que envolve conteúdo e atividades do OA. É relacionado com a competência a ser apreendida, e a mesma competência pode ser adquirida com processos cognitivos sucessivos e encadeados (diferentes objetivos de aprendizagem). A atividade ou habilidade que constitui a competência, demanda processos cognitivos adequados para maior efetividade do ensino-aprendizagem. Por conseguinte, a competência tem relação com a dimensão do conhecimento;
- O público-alvo é definido por sua formação profissional, com respectivas competências. Determinados objetivos de aprendizagem são para públicos-alvo específicos adquirirem competências necessárias para o exercício da profissão. Competências delimitam atividades profissionais e podem ser formalizadas em documentos próprios e VCs;
- O nível educacional do público alvo pode ser herdado da sua formação profissional. Entretanto, o público-alvo pode ter formação profissional XYZ, mas lhe é exigido nível educacional maior. Por exemplo, é um

profissional graduado, mas exige-se doutorado para o objetivo de aprendizagem. O nível educacional apresenta vocabulário controlado, mas o que gera insumos para o refinamento de resultados de busca é a formação profissional. Ainda assim, essa informação pode facilitar a ordenação, após categorização por formação profissional.

O desenvolvimento da modelagem proposta gerou indagações sobre o cumprimento dos objetivos da pesquisa, contudo, a comprovação da hipótese levantada no início da tese elucidou respostas. Resultado já esperado no cumprimento das etapas da metodologia que contribuiu fortemente para boa parte dessas indagações foi a compreensão semântica de conceitos do contexto em que se insere o estudo, restringindo interpretações dos termos citados nos documentos consultados. Em outras palavras, o entendimento do domínio propiciou a visão geral de definições reconhecidas e foi fundamental na tomada de decisões que permearam o desenvolvimento.

Não foi efetuada avaliação da modelagem com especialistas no domínio, mas foram realizadas simulações e recorrentes investigações do que tem sido utilizado na prática, procurando, assim, obter como resultado um modelo teórico, mas que reflete a realidade. Vale ressaltar que esse esforço não exime a modelagem proposta de falhas, uma vez que qualquer alteração, avanço tecnológico ou mudança de perspectiva implicam adequações, o que não decresce sua credibilidade.

8.3.3 Ontologias e a WS

O desenvolvimento da modelagem com a concepção da construção de uma ontologia na ferramenta Protégé, foi fundamental para a adoção das tecnologias da WS adequadas. A função da ontologia é representar formalmente a semântica de domínios de conhecimento para restringir interpretação dos conceitos em um dado contexto. Logo, a inferência de OAs que satisfazem a necessidade de busca do usuário considerará a informação descritiva relevante para o ensino-aprendizagem.

A ontologia, além de organizar semanticamente a estrutura de informação, definiu regras para regular a combinação e relação entre metadados, que representam conceitos e propriedades, a serem adotados pelo RDE para viabilizar o entendimento comum e compartilhado da descrição de OAs. Assim, a instância de metadados do OA poderá ser compreendida e explorada por pessoas e computadores.

Para identificar conceitos e propriedades diretamente relacionados com objetivo de aprendizagem e outros capazes de especificá-lo, foi discutido a reutilização de ontologias já existentes, para sua posterior consideração ou agregação na ontologia que constituiu a modelagem proposta. As ontologias investigadas utilizam a linguagem OWL e representam conceitos de “objetivo de aprendizagem”, do assunto “saúde” e do “público-alvo dos OAs na área de saúde”. Para complementar, foram pesquisados outros esquemas de representação do conhecimento.

Apenas a ontologia proposta por Lai Ng, com conceitos de objetivo de aprendizagem, atendeu necessariamente a pesquisa. O domínio dessa ontologia é a representação descritiva do objetivo de aprendizagem e suas instâncias são os objetivos de aprendizagem em si. Lai Ng caracteriza o objetivo de aprendizagem com conceitos das classes “descrição”, “desempenho” e “referência de conhecimento” (NG, 2005). A Taxonomia de Bloom e sua atualização (BLOOM; KARTHWOHL, 1984) também foram consideradas na proposição da modelagem.

Ontologias pesquisadas com conceitos da área de “saúde” para descrever o assunto sobre o qual o objetivo de aprendizagem trata, não apresentaram termos em português. Isso não foi observado como empecilho à utilização, pois a agregação dessas ontologias tem caráter demonstrativo, facultando o desenvolvimento ou tradução de ontologias às demandas para implementação da modelagem. Um dos VCs investigados com termos na área de saúde, o DeCS, é adotado pelo ARES, RDE discutido no cenário de implementação. Sua utilização diante da proposta da pesquisa requer a representação de conceitos em OWL. Recomenda-se estudos sobre a geração automática de ontologias e outras tecnologias para garantir a cobertura de assuntos e não condicionar a implementação da modelagem à disponibilidade de ontologias prontas em OWL.

Ontologias à cerca do “público-alvo dos OAs na área de saúde” não foram verificadas, porém compreendeu-se a representação desse público-alvo específico, dividida entre conceitos e propriedades que definem características gerais dos aprendizes e conceitos e propriedades que especificam sua formação profissional. As ontologias verificadas suscitaram conceitos apenas para características gerais dos aprendizes. Para representar formação profissional recorreu-se a Classificação Brasileira de Ocupações (CBO), documento que reconhece ocupações do mercado de trabalho brasileiro.

Após identificação de propriedades e conceitos, sucedeu-se a proposição de ontologias para objetivo de aprendizagem, assuntos da área de saúde e público-alvo. Ao longo desse processo, foram consultadas aplicações com as ontologias e

esquemas de representação do conhecimento investigados que trouxeram constatações importantes para o trabalho.

O uso de ontologias em consonância com as tecnologias da WS apresentou vantagens à compreensão de conceitos e ainda à disponibilização da modelagem proposta para aplicações computacionais específicas, permitindo, por exemplo, a correlação dos OAs do RDE na área de saúde com outras fontes de informação que adotam estrutura de informação ou tecnologias compatíveis. Assim, contribuiu-se com a nomeada Linked Web, viabilizando, por exemplo, que a Wikepedia possa fornecer informação adicional ao usuário que busca determinado assunto no RDE, mas também se interessa por ações de capacitação referentes a esse assunto, promovidas por instituições e divulgadas em sua página na Wikepedia. Outra possibilidade vislumbrada é a relação do OA com documento de referência utilizado na sua concepção e disponibilizado em bibliotecas digitais adaptadas a proposta da Linked Web.

Devido à formalização dos conceitos, a ontologia elimina ambiguidades e inconsistências na recuperação da informação, logo, compõe vocabulário que representa o conhecimento do domínio de forma explícita e com alto nível de abstração. Entretanto, alguns problemas com uso de ontologias foram identificados, como por exemplo, o tempo demandado para representar conceitos consistentemente, já que a ontologia não é totalmente flexível às abstrações feitas por indivíduos, frente ao domínio.

Ontologias não são necessariamente estacionárias, isto é, necessitam evoluir com a inserção de conhecimento consensual. Assim, pode ser considerada a ampliação de instâncias da ontologia à medida que a modelagem é implementada e são descritos OAs no repositório, constituindo exemplos de público-alvo, objetivo de aprendizagem, etc. Para tanto, são necessários esforços adicionais para viabilizar a edição colaborativa de ontologias. Uma solução adequada é a utilização do *framework* Jena, que oferece compatibilidade com RDF e suporte à OWL.

Ao avaliar a ontologia frente à definição de Gruber (1996), analisada por Guarino & Giaretta (1995, apud MONTEIRO, 2006), de que "ontologia é uma especificação formal e explícita de uma conceitualização compartilhada", o produto desta pesquisa, é uma ontologia na qual sua conceitualização traz o modelo abstrato de um fenômeno do mundo real, definido com relações intencionais que descrevem o 'estado particular das coisas', para criar uma rede conceitual que superpõe vários possíveis 'estados das coisas. Ademais, o desenvolvimento da modelagem compartilha definições de especialistas que escreveram sobre o domínio, considerando as referências consultadas.

A ontologia foi especificada com termos em linguagem natural e adotadas tecnologias da WS para definição de conceitos, sendo assim, é explícita e formal. No que diz respeito aos componentes básicos das ontologias, está organizada em classes com definições claras de seus atributos, relacionando e especificando conceitos dentro do domínio. A ontologia que representa a modelagem proposta foi desenvolvida de maneira precisa, frente às fontes de informação consultadas, porém, constatou-se que o resultado é influenciado pela interpretação do modelador e sua visão do domínio representado.

8.3.4 RDE, metadados e a WS

Metadados são fundamentais para o cumprimento do objetivo da pesquisa, tal como para WS. Isto porque, é a definição de metadados que viabiliza a identificação e processamento da informação que descreve conteúdos disponíveis na Web. Para o estudo, metadados são dispostos em formulários de entrada de dados em RDEs na área de saúde e utilizam padrões específicos. Para WS de maneira geral, metadados podem ser definidos com RDF no próprio conteúdo do documento.

A pesquisa definiu metadados para representar os conceitos que especificam o objetivo de aprendizagem. A formalização semântica, constituída na concepção de ontologias, é retratada nos metadados do RDE para descrição de OAs. Diante disso, possibilitou-se que, independentemente da informação preenchida no metadado “objetivo de aprendizagem”, as relações semânticas desse metadado com outros sejam estabelecidas para promover inferências. Também foram definidos valores de preenchimento de metadados, com adoção de VCs para alguns casos, que contribuem para diminuição de ambiguidades e erros no momento da descrição de OAs.

A investigação de padrões de metadados, DC e LOM, destacou a complementariedade de ambos, um com relação à documentação e instrumentalização para o uso (DC), outro com relação à perspectiva da informação descritiva no contexto de ensino-aprendizagem (LOM). Diferente do esperado, o LOM não contemplou boa parte dos conceitos e propriedades propostos. A correlação desses padrões mostrou-se fundamental para implementação da modelagem em RDEs com metadados baseados em qualquer um dos dois e, por conseguinte, contribuiu com a interoperabilidade semântica do DC e LOM.

O DC apresentou ampla possibilidade de representação semântica dos metadados e, inclusive, seu modelo abstrato (DCMI Abstract Model) foi utilizado para

tratar da integração de metadados à modelagem proposta. O LOM suscitou a definição de outros conceitos e propriedades adequados para a modelagem, com base na categoria Educacional (5 *Educational*). Os dois padrões demandaram extensões e qualificações de metadados para atender os requisitos da pesquisa.

Uma vez definidos todos metadados adequados para pesquisa, algumas considerações observadas trouxeram reflexões sobre metadados e a modelagem proposta. Fazendo menção apenas do conceito representado pelo metadado, independente do padrão adotado, destaca-se:

- Mesmo que haja a relação entre o metadado *objetivo de aprendizagem* e o metadado *assunto (Texto livre)*, este último não foi considerado para inferências, pois seu preenchimento é livre, de difícil controle. Nesse caso, apenas o metadado *assunto (Taxonomia)* é considerado para inferências. Essa consideração foi adotada para demais metadados com valores de preenchimento não controlado;
- O processo cognitivo pode ser transcrito na redação do objetivo de aprendizagem, contudo, o metadado *processo cognitivo* com valores de preenchimento determinados por VC apropriado, foi definido e permite a categorização de OAs;
- A dimensão do conhecimento não figura na ontologia adotada para descrever assunto e, portanto, o metadado *dimensão do conhecimento* com valores de preenchimento determinados por VC apropriado, permite a categorização de assuntos;
- *Objetivo de aprendizagem prévio, requerido e posterior* assumem valores não literais, pois preveem a indicação da URL de OAs relacionados;
- O *pré-requisito* descrito para o objetivo de aprendizagem do OA e *objetivo de aprendizagem requerido*, podem coexistir para indicar requisito que o RDE não fornece em termos de disponibilidade de OAs ou qualquer outro tipo de requisito não contemplado com o relacionamento de OAs;
- O assunto de interesse do *público-alvo* deve ser pré-determinado pela CBO para generalizar o interesse com base na formação profissional do aprendiz. Não é viável detalhá-lo no preenchimento, pois trata preferências pessoais, com ampla margem de erro na descrição;
- Metadados como *nível educacional*, com valores de preenchimento pré-determinados, mas sem muita expressividade semântica, são úteis

para o refinamento dos resultados de busca, porém podem constituir filtro simples para recuperação de OAs;

- A obrigatoriedade (preenchimento do metadado como condição para a finalização da descrição) e repetitividade (permissão para preenchimento do metadado com mais de um valor) dos metadados podem garantir a recuperação semântica, uma vez que a descrição completa do OA subsidia a indexação de metadados e o estabelecimento de relações entre eles.

O uso de metadados semanticamente definidos com a modelagem proposta, trouxe benefícios importantes para o RDE na área de saúde, tais como, orientações para criar e manter vocabulários de metadados personalizados, interoperabilidade entre RDEs com padrões de metadados diferentes, diretrizes genéricas para definição e relacionamento de metadados, modelo único para utilização, armazenamento e consulta de metadados de acordo com o propósito da pesquisa.

9 Conclusão

A educação em saúde, especialmente a capacitação de profissionais da área de saúde com estratégias de educação a distância, tem ampliado a disseminação de objetos de aprendizagem. O OA, disponibilizado em repositórios digitais educacionais, tem seus aspectos educacionais descritos em formulários de metadados. Nesse contexto, o cumprimento dos objetivos da pesquisa viabilizou o desenvolvimento de modelagem baseada em Web Semântica para RDEs na área de saúde, com foco no objetivo de aprendizagem. Para isso, foram identificadas características dos OAs e tecnologias da WS. A modelagem proposta agrega essas tecnologias à arquitetura dos repositórios, além de definir a estrutura de informação adequada. Isso viabilizou inferências automáticas para refinar os resultados de busca e minimizar esforços humanos no momento da seleção de objetos de interesse. Assim, conclui-se que a hipótese da pesquisa de que a modelagem baseada em WS para RDEs na área de saúde contribui com o refinamento dos resultados de busca do usuário é confirmada.

A informação descritiva do OA, registrada nos metadados, teve sua semântica representada pela definição dos conceitos que especificam o objetivo de aprendizagem. O objetivo de aprendizagem é aquilo [resultado] que o professor [autor] espera do aluno [usuário] ao final do processo de aprendizagem (DRISCOLL, 1998), logo, a proposta aumentou a pertinência do resultado de busca e a satisfação do usuário que possui interesses pautados não apenas no assunto, mas em características que indicam o uso do objeto no ensino-aprendizagem. Assim o usuário pode tomar decisões não apenas com base no tema, mas com relação aquilo que deseja aprender. A pesquisa levou em consideração o processo de descrição e o processo de recuperação, ou seja, o fluxo de disponibilização de OAs em RDEs, que delinearam a necessidade de estruturar sintática e semanticamente os metadados, valendo-se das camadas XML, XMLS, RDF, RDFS e OWL da WS. XML forneceu sintaxe e estrutura para os dados, RDF permitiu expressar a semântica dos metadados e OWL aumentou a expressividade semântica da informação.

O preenchimento de metadados como assunto, objetivo de aprendizagem, público-alvo, contexto educacional, entre outros, no padrão DC ou LOM, agregam informação, ou mesmo atribuem sentido educacional ao OA. Consequentemente, a busca de quaisquer termos registrados nesses metadados em RDE na área de saúde com a modelagem implementada, retornará OAs com termos buscados, classificados de acordo com coincidências semânticas, e OAs sem termos buscados, mas com relações semânticas, classificados de acordo com coincidências semânticas. Vocabulários controlados, como a Taxonomia de Bloom, ICF Ontology e CBO, foram utilizados na definição de valores de preenchimento com semântica formalizada, aumentando a possibilidade de realizar inferências e relacionar OAs.

Reflexões e considerações sobre WS, RDE e OA foram suficientemente expostas ao longo dos capítulos desta tese. Por fim, a pesquisa cumpriu a expectativa de ampliar definições e restringir a interpretação de conceitos importantes para o ensino-aprendizagem. E, ainda, os padrões de metadados mais utilizados em RDEs, DC e LOM, foram discutidos e correlacionados, mostrando avanços na interoperabilidade semântica.

10 Sugestões e pesquisas futuras

O desenvolvimento da pesquisa vislumbrou sugestões para avançar o estudo, tais como, a utilização da modelagem proposta em RDEs diversos a partir da investigação de adaptações. Além disso, ferramentas de autoria para elaboração de OAs, sistemas de recuperação de informação, integração com LMS, ou quaisquer outras aplicações computacionais dentro do contexto, que adotem a modelagem, são favorecidas. A consulta de novas fontes de informação poderá enriquecer a estrutura de conceitos. Já a revisão das definições e estrutura semântica por especialistas da área, possibilitará o aprimoramento. Por fim, a pesquisa motivou novos estudos, apresentados como propostas para trabalhos futuros:

- Construir interface com API Jena para que usuários do RDE possam manipular e acrescentar conceitos nas ontologias adotadas;
- Definir e relacionar um conjunto maior de metadados, observando elementos da Classe Educacional LOM;
- Classificar OAs e, mediante cadastro do perfil de usuário, recomendar conteúdos em ofertas de capacitação no LMS;
- Incorporar definições especificadas pela modelagem em agentes de *softwares* para realizar buscas em fontes sem implementação da modelagem, mas com concepções do ser contexto;
- Gerar ontologias automaticamente para garantir a cobertura de assuntos dos OAs;
- Indexar automaticamente texto livre com a ponderação de termos relevantes estabelecida pela modelagem proposta;
- Medir precisão e revocação da recuperação da informação em RDEs com e sem implementação da modelagem para estudos comparativos.

REFERÊNCIAS

- ALLEMANG, Dean; HENDLER, James. **Semantic web for the working ontologist: effective modeling in RDFS and OWL**. 2. ed. USA: Morgan Kaufman, 2011.
- ALLERT, H.; HADHAMI Dhaief; WOLFGANG, Nejd. **How are learning objects used in learning processes?** Instructional roles of learning objects in LOM. 2002. Disponível em: <<http://projekte.learninglab.uni-hannover.de/pub/bscw.cgi/d6193/How%20are%20Learning%20Objects%20Used%20in%20Learning%20Processes.%20PDF-file>>. Acesso em: 08 dez. 2012.
- ALMEIDA, Mauricio B.; BAX, Marcello P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Ci. Inf.**, Brasília, v. 32, n. 3, p. 7-20, set./dez. 2003.
- ALVES, R. C. V. **Web semântica: uma análise focada no uso de metadados**. 2005. 180 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2005.
- ANDERSON, L. W. et. al. **A taxonomy for learning, teaching and assessing: a revision of Bloom's Taxonomy of Educational Objectives**. Nova York: Addison Wesley Longman, 2001. 336 p.
- ANDERSON, L. W.; et. al. A taxonomy for learning, teaching and assessing: a revision of Bloom's Taxonomy of Educational Objectives. Nova York: Addison Wesley Longman, 2001. 336 p.
- AZOUAOU, Faïçal; DESMOULINS, Cyrille. Semantic annotation for the teacher: models for a computerized memory tool. In: INTERNATIONAL WORKSHOP ON APPLICATIONS OF SEMANTIC WEB TECHNOLOGIES FOR E-LEARNING (AIED), 2005, Amsterdam. **Proceedings...** Amsterdam: AIED, 2005. p. 47-56.
- BARKER, Phil. **What is... What is IEEE Learning Object Metadata / IMS Learning Resource Metadata?** In: CETIS, Standards Briefing Series, 2005.
- BARRITT, C.; ALDERMAN, F.L. **Creating a reusable learning objects strategy: leveraging information and learning in a knowledge economy**. [S.l.]: John Wiley & Sons, 2004.
- BARTEL, Mark; et. al. **XML Signature Syntax and Processing**. 2.ed. [S.l.]: W3C Recommendation, 2008. Disponível em: <<http://www.w3.org/TR/2008/REC-xmldsig-core-20080610/#def-Integrity>>. Acesso em: ago. 2012.
- BECHHOFFER, Sean; et. al. **OWL Web Ontology Language: reference**. [S.l.]: W3C Recommendation, 2004. Disponível em: <<http://www.w3.org/TR/owl-ref/>>. Acesso em: ago. 2010.

- BERNERS-LEE, T. **Linked Data**. 2006. Disponível em: <<http://www.w3.org/DesignIssues/LinkedData.html>>. Acesso em: set. 2012.
- BERNERS-LEE, T.; et al. **The semantic toolbox: building semantics on top of XML-RDF**. 1999. Disponível em: <<http://www.w3.org/DesignIssues/Toolbox.html>>. Acesso em: jun. 2003.
- BERNERS-LEE, Tim. **Semantic Web: XML2000**. [2000]. Disponível em: <<http://www.w3.org/2000/Talks/1206-xml2k-tbl/Overview.html>>. Acesso em: jul. 2012.
- BERNERS-LEE, Tim; HENDLER, James; LASSILA, Ora. The semantic web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. In: SCIENTIFIC AMERICAN SPECIAL ONLINE ISSUE. **Proceedings...** New York: Scientific American, 2002, p. 24-30. Disponível em: <<http://www.med.nyu.edu/research/pdf/mainim01-1484312.pdf>>. Acesso em: set. 2012.
- BIOE. Banco Internacional de Objetos Educacionais. 2009. Disponível em: <<http://objetoseducacionais2.mec.gov.br/>>. Acesso em: mar. 2009.
- BLOOM, B. S. **Taxonomy of educational objectives, handbook: the cognitive domain**. New York: David McKay Co Inc., 1956.
- BLOOM, B. S.; KARTHWOHL, D. R. **Taxonomy of Educational Objectives, Handbook 1: Cognitive Domain**. New York: Addison-Wesley and Co, 1984.
- BRASCHER, M.; MONTEIRO, Fernanda. Life cycle assessment ontology. In: BRAVO, Blanca Rodríguez; ALVITE DÍEZ, Maria Luisa (Org.). **La interdisciplinariedad y la transdisciplinariedad en la organización del conocimiento científico**. León: Universidad de León, 2007. v. 1. p. 169-179.
- BRASIL. Lei nº 9.610, de 19 de fevereiro de 1998. Altera, atualiza e consolida a legislação de direitos autorais e dá outras providências. **Diário Oficial [da] República Federativa do Brasil**, 20 fev. 1998. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/L9610.htm>. Acesso em: out. 2012.
- BRASIL. Ministério da Saúde. Secretaria de Gestão do Trabalho e da Educação na Saúde. Universidade Aberta do Sistema Único de Saúde.
- Manual de descrição e preenchimento de metadados**. Responsabilidade técnica: Aline Santos Jacob e Fernanda de Souza Monteiro – Brasília : UNA-SUS, 2011.
- BRASIL. Ministério da Saúde. Secretaria de Gestão do Trabalho e da Educação na Saúde. Universidade Aberta do Sistema Único de Saúde. **Política do acervo de recursos educacionais em saúde**. Responsabilidade técnica: Fernanda de Souza Monteiro e Aline Santos Jacob. – Brasília : UNA-SUS, 2011.

BRASIL. Ministério da Saúde. Secretaria de Atenção à Saúde. **Política Nacional de Saúde da Pessoa Portadora de Deficiência**. Brasília: Ed. MS, 2008. Disponível em: <http://bvsmms.saude.gov.br/bvs/publicacoes/politica_nacional_saude_pessoa_deficiencia.pdf>. Acesso em: nov. 2012.

BRASIL MTE. **Site**. Disponível em:< <http://www.mtecbo.gov.br/cbosite/pages/informacoesGerais.jsf;jsessionid=A5E20FC9E6EF0DFCF6EDCF7CFC3BCA17.Ibroutef121p005>>. Acesso em: nov. 2012.

BRICKLEY, Dan; et. al. **Cognitive Characteristics Ontology 0.2. 2010**. Disponível em: <http://smiy.sourceforge.net/cco/spec/cognitive_characteristics.html>. Acesso em: out. 2011.

BRICKLEY, Dan; GUHA, R. V. (eds.). **RDF Vocabulary Description Language 1.0: RDF Schema**. 2004. Disponível em: <<http://www.w3.org/TR/2004/REC-rdf-schema-20040210/>>. Acesso em: fev. 2012.

CAFÉ, Lúcia. et al. Repositórios institucionais: nova estratégia para publicação científica na Rede. In: CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO, 26, 2003, Belo Horizonte. **Anais...** Belo Horizonte: INTERCOM, 2003. Disponível em: <http://www.intercom.org.br/papers/nacionais/2003/www/pdf/2003_ENDOCOM_TRABALHO_cafe.pdf >. Acesso em: out. 2011.

CAMPOS, Maria Luiza Almeida. Modelização de domínios de conhecimento: uma investigação de princípios fundamentais. **Ci. Inf.**, Brasília, v. 33, n. 1, p. 22-32, jan./abr. 2004.

CAMPOS, Maria Luiza Almeida. Modelos de abstração: perspectivas para o estudo da área de representação da informação. **Ci. Inf.**, Brasília, v. 25, n. 2, p. 22-32, jan./abr. 1995.

CAPURRO, R. Epistemologia e ciência da informação. IN: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB), 5, 2003, Belo Horizonte. **Anais...** Belo Horizonte: UFMG, 2003.

CARVALHO, Maria da Conceição Rodrigues de; GOMES, Sandra Lúcia Rebel. Repositório institucional no campo da saúde: um estudo exploratório de reuso da informação técnico-científica para a pós-graduação do ICICT/Fiocruz. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB), 12, 2011, Brasília. **Anais...** Brasília: UnB, 2011.

CASTRO, Fabiano Ferreira de; SANTOS, Plácida Leopoldina V. A. da Costa. Os metadados como instrumentos tecnológicos na padronização e potencialização dos recursos informacionais no âmbito das bibliotecas digitais na era da web semântica. **Inf. & Soc.:** Est., João Pessoa, v.17, n.2, p.12-21, maio/ago. 2007.

CAVALCANTE, Maria Tereza Leal; VASCONCELLOS, **Miguel Murat. Base tecnológica para a educação na saúde:** objetos de aprendizagem e padrões. Rio de Janeiro: ENSP/FIOCRUZ, 2006.

- CETL. Reusable Learning Objects. Disponível em: <<http://www.rlo-cetl.ac.uk/whatwedo/rlos/completedrlos.php>>. Acesso em: jun. 2011.
- CISCO SYSTEMS. **Reusable learning object strategy**: designing and developing learning object for multiple learning approaches. [S.l.]: White paper, 2003.
- CONNOLLY, D. et al. DAML+OIL (March 2001) Reference Description. 2001. Disponível em: <<http://www.w3.org/TR/daml+oil-reference>>. Acesso em: jul. 2011.
- DAHLBERG, I. Teoria do conceito. **Ci. Inf.**, Rio de Janeiro, v.7, n.2, p.101-7, 1978.
- DECKER, S. et al. The semantic web: the roles of XML and RDF. **IEEE Expert**, v. 15, n. 3, oct. 2000. Disponível em: <<http://www.cs.vu.nl/~frankh/postscript/IEEE-IC00.pdf>>. Acesso em: jan. 2010.
- DICHEVA, D.; DICHEV, C. TM4L: Creating and browsing educational topic maps. **British Journal of Educational Technology**, v. 37, n. 3, p. 391-404, 2006.
- DOWNES, S. Learning objects: learning object repositories. 2001. Disponível em: <http://www.atl.ualberta.ca/downes/naweb/column000523_4.htm>. Acesso em: jul. 2003.
- DRISCOLL, M. M. **Web-based training**: using technology to design adult learning experiences. San Francisco: Jossey-Bass, 1998.
- DUBLIN CORE METADATA INITIATIVE (DCMI). DCMI Abstract Model. 2007. Disponível em: <<http://dublincore.org/documents/abstract-model/>>. Acesso em: jan. 2010.
- DUBLIN CORE METADATA INITIATIVE (DCMI). DCMI Metadata Terms. 2012. Disponível em: <<http://dublincore.org/documents/2012/06/14/dcmi-terms/#H1>>. Acesso em: jun. 2012.
- DUBLIN CORE METADATA INITIATIVE (DCMI). Metadata basics. [20--]. Disponível em: <<http://dublincore.org/metadata-basics/>>. Acesso em: ago. 2012.
- DUBLIN CORE METADATA INITIATIVE (DCMI). **Home**. Disponível em: <<http://dublincore.org/>>. Acesso em: jan. 2010.
- DUBLIN CORE METADATA INITIATIVE (DCMI). **Interoperability Levels for Dublin Core Metadata**. 2009. Disponível em: <<http://dublincore.org/documents/interoperability-levels/>>. Acesso em: jan. 2010.
- DUBLIN CORE METADATA INITIATIVE (DCMI). **The Singapore Framework for Dublin Core Application Profiles**. 2008. Disponível em: <<http://dublincore.org/documents/singapore-framework/>>. Acesso em: jul. 2011.

DUQUE, Claudio Gottschalg. **SiRILiCO**: uma proposta para um Sistema de Recuperação de Informação baseado em Teorias da Lingüística Computacional e Ontologia. 2005. Tese (Doutorado Sanduíche em Ciência da Informação e Angewandte Sprachwissenschaft und Computerlinguist - Universität Giessenem) – Universidade Federal de Minas Gerais, 2005.

EDUCATIONAL TECHNOLOGY FOR HEALTHCARE EDUCATION. **Site**. Disponível em: <http://sonet.nottingham.ac.uk/rlos/rlo_repositories.html>. Acesso em: fev. 2012.

ESCOLA NACIONAL DE SAÚDE PÚBLICA SÉRGIO AROUCA (ENSP). **Site**. Disponível em: <<http://micro010.ensp.fiocruz.br/~ensp/acessoaberto/>>. Acesso em: ago. 2012.

FARIA, Carla Gomes de; GIRARDI, Rosário. **Uma análise da web semântica e suas implicações no acesso à informação**. Maranhão: Universidade Federal do Maranhão, 2001.

FERREIRA, Aurélio Buarque de Holanda. **Novo dicionário Aurélio da língua portuguesa**. 3. ed. rev. e amp. Curitiba: Positivo, 2004. 2120 p.

FIOCRUZ. **Programa integrado de informação e comunicação da Fiocruz**. Rio de Janeiro: FIOCRUZ, 2002.

FRANCISCATO, Fábio Teixeira. **ROAD**: Repositório Semântico de objetos de aprendizagem para dispositivos móveis. 2010. Dissertação (Mestrado em Informática) - Universidade Federal de Santa Maria, 2010.

FREITAS, Fred ; SCHULZ, S.; MORAES, Eduardo. Pesquisa de terminologias e ontologias atuais em biologia e medicina. **RECIIS. Revista Eletrônica de Comunicação, Informação & Inovação em Saúde**, v. 3, p. 1-13, 2009.

Disponível em: <<http://www.reciis.cict.fiocruz.br/index.php/reciis/article/viewArticle/239>>. Acesso em: jul. 2010.

GAGNÉ, R. M. **The Conditions of Learning**. New York: Holt, Rinehart and Winston, 1985.

GATTI, B. A. **Grupo focal na pesquisa em ciências sociais e humanas**. Brasília: Líber Livro, 2005.

GIL, Antonio Carlos. Métodos e técnicas de pesquisa social. São Paulo: Atlas, 1999.

GONÇALVES, Vitor. **A web semântica no contexto educativo**: um sistema para a recuperação de objectos de aprendizagem baseado nas tecnologias para a web semântica, para o e-learning e para os agentes. 2007. 380 f. Tese (Doutorado em Engenharia Electrotécnica e de Computadores) - Universidade do Porto, Faculdade de Engenharia, Porto, 2007.

GOÑI, J.; FERNANDES, C.; LUCENA, C. E-learning e a Web Semântica. CONGRESSO DA RIBIE, 6, 2002, Pontifícia Universidade Católica do Rio de Janeiro. **Anais...** Rio de Janeiro: PUC, 2002. Disponível em: <<http://lsm.dei.uc.pt/ribie/docfiles/txt2003729185022paper-212.pdf>>. Acesso em: 01 abr. 2004.

GROSOFF, B. et al. **Description logic programs**: combining logic programs with description logics. 2003. Disponível em: <<http://www.cs.man.ac.uk/~horrocks/Publications/download/2003/p117-grosoff.pdf>>. Acesso em: 16 set. 2005.

GRUBER, T. R. A translation approach to portable ontologies. **Knowledge Acquisition**, v. 5, n. 2, p.199-220, 1993. Disponível em: <<http://tomgruber.org/writing/ontologia-kaj-1993.pdf>>. Acesso em: ago. 2010.

GUARINO, N. Understanding, building, and using ontologies: a commentary to using explicit ontologies in KBS development. **International Journal of Human and Computer Studies**, v. 46, n. 2/3, p. 293-310, 1997.

HARMELEN, F. **The semantic web**: what, why, how, and when. IEEE Intelligent Systems, v. 5, issue 3, 2004. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1285880>. Acesso em: maio 2011.

HARNAD, Stevan. et al. **The access/impact problem and the green and gold roads to open access**. 2001. Disponível em: <<http://www.ecs.soton.ac.uk/~harnad/Temp/impact.html>>. Acesso em: fev. 2011

HAYES, Helen. **Digital Repositories**: helping universities and colleges. JISC Briefing Paper – Higher Education Sector. Helen Hayes (Ed.). England: August 2005. Disponível em: <<http://archive.excellencegateway.org.uk/page.aspx?o=117227>>. Acesso em: fev. 2010.

HEINS, Tanya; HIMES, Frances. **Creating learning objects with macromedia flash MX. Macromedia, Inc. A Macromedia white paper**. San Francisco: Released April, 2002. Disponível em: <<http://www.macromedia.com/learning>>. Acesso em: nov. 2011.

HERMAN, Ivan. Introduction to the Semantic Web (tutorial). In: SEMANTIC TECHNOLOGY CONFERENCE SAN JOSE, 2009, California, USA. **Proceedings...** California: [s.n], 2009. Disponível em: <<http://ilps.science.uva.nl/Teaching/II0607/twiki/pub/Main/CourseSchedule/nice-ii-0607-week15-Ivan-Herman-Semantic-web-8up.pdf>>. Acesso em: out. 2010.

HOLANDA, Raphael; et. al. **Uma Ontologia para Ambientes Interativos de Aprendizagem**. Maceió: Universidade Federal de Alagoas (UFAL), 2007.

HORROCKS, I. et al. Semantic Web Architecture: stack or two towers? In: FAGES, François; SOLIMAN, Sylvain. **PPSWR**: principles and practice of

semantic web reasoning, Third International Workshop, [S.l.]: Springer; 2005. p. 37-41. Disponível em: <<http://ect.belllabs.com/who/pfps/publications/stack.pdf>>.

HUNTER, Philip; DAY, Michael. Institutional repositories, aggregator services and collection development, 2005. Disponível: <<http://eprints-uk.rdn.ac.uk/project/docs/studies/colldevelopment/coll-development.pdf>>. Acesso em: nov. 2011.

IEEE. **Draft Recommended Practice for Expressing IEEE Learning Object Metadata Instances Using the Dublin Core Abstract Model**, 2008. New York: IEEE, 2008.

IEEE. **Draft Standard for Learning Object Metadata**. New York: IEEE, 2002.

IEEE. **Standard Upper Ontology Working Group (SUO WG)**. Site. Disponível em: <<http://suo.ieee.org/>>. Acesso em: dez. 2011.

INTERNATIONAL CLASSIFICATION OF FUNCTIONING, DISABILITY AND HEALTH (ICF). Site. Disponível em:<<http://bioportal.bioontology.org/ontologies/47404/?p=summary>>. Acesso em: out. 2012.

ISOTANI, Seiji. et al. Estado da arte em web semântica e web 2.0: potencialidades e tendências da nova geração de ambientes de ensino na internet. **Revista Brasileira de Informática na Educação**, v. 17, n. 1, 2009. Disponível em: <http://www.ei.sanken.osakau.ac.jp/~isotani/artigos/estado_da_arte_sw.pdf>. Acesso em: jan. 2012.

JOVANOVIĆ. **Site**. Disponível em:< <http://jelenajovanovic.net/LOCO-Analyst/index.html>>. Acesso em: out. 2012.

KAPLAN-LEISERSON, Eva. **Trend**: Podcasting in Academic and Corporate Learning. 2005. Disponível em: <http://www.learningcircuits.org/2005/jun2005/0506_trends>. Acesso em: fev. 2009.

KLYNE, Graham; CARROL, Jeremy J. (eds.). **Resource Description Framework (RDF): concepts and abstract syntax**. USA: W3C, 2004. Disponível em: <<http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/>>. Acesso em: maio 2010.

KOUTSOMITROPOULOS, Dimitrios A.; et. al. The use of metadata for educational resources in digital repositories: practices and perspectives. **D-Lib Magazine**, v. 16, n. 1/2, Jan./Feb. 2010. Disponível em: <<http://www.dlib.org/dlib/january10/kout/01kout.html>>. Acesso em: nov. 2011.

KRATZ, R; PINTO, S.; SCOPEL, M. **An architecture for courseware adequacy to re-useable learning objects (scorm)**. IEEE INTERNATIONAL CONFERENCE ON NEXT GENERATION WEB SERVICES PRACTICES (NWESP'05).

KRATZ, Ricardo de Andrade; PINTO, Sérgio Crespo C. S.; BARBOSA, Jorge. Fábrica de adequação de objetos de aprendizagem. **Revista Brasileira de**

Informática na Educação, v. 15, n. 3, 2007. Disponível em: <<http://www.br-ie.org/pub/index.php/rbie/article/view/23>>. Acesso em: out. 2010.

LAGOZE, Carl; SOMPEL, Herbert van den. The Santa Fe Convention of the Open Archives Initiative. **D-Lib Magazine**, v. 6, n. 2, fev. de 2000. Disponível em: <<http://www.dlib.org/dlib/february00/vandesompel-oai/02vandesompel-oai.html> > Acesso em: nov. 2005.

LE COADIC, Yves-François. **A ciência da informação**. Brasília: Briquet de Lemos, 2004. 124 p.

LONGMIRE, W. **A primer on learning objects**. Learning Circuits Webzine. 2000. Disponível em: <<http://www.learningcircuits.com/mar2000/primer.html>>. Acesso em: maio 2003.

MARCONI, M. A.; LAKATOS, M.E. **Técnicas de Pesquisa**. 6. ed. São Paulo: Atlas, 2006.

MÁRDERO ARELLANO, M. A. Repositórios, acesso livre, preservação digital: questões para hoje e amanhã. In: SEMINÁRIO DE REPOSITÓRIOS INSTITUCIONAIS E DE OBJETOS VIRTUAIS DE APRENDIZAGEM, 2009, Florianópolis. **Socializando as melhores práticas**. Florianópolis: UFSC, 2009. p. 1-68. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/13706/12570>>. Acesso em: 16 set. 2010.

MARTINS, Ana Bela; RODRIGUES, Eloy; NUNES, Manuela Barreto. Repositórios de informação e ambientes de aprendizagem: criação de espaços virtuais para a promoção da literacia e da responsabilidade social. **Rede de bibliotecas escolares**, newsletter n. 3, 2008.

MCGUINNESS, Deborah L.; HARMELEN, Frank Van. **OWL Web Ontology Language: overview**. USA: W3C, 2004. Disponível em: <<http://www.w3.org/TR/2004/REC-owl-features-20040210/>>. Acesso em: abr. 2009.

MCLAREN, I. New trends in web-based learning: objects, repositories and learner engagement. **European Journal of Engineering Education**, v. 29, n. 1, p. 65-71, 2004.

MERLOT. **Site**. Disponível em: <<http://www.merlot.org/merlot/index.htm>>. Acesso em: ago. 2011.

MIRANDA, Raquel Melo de. **Groa**: um gerenciador de repositórios de objetos de aprendizagem. 2004. 80 f. Dissertação (Mestrado em Ciência da Computação) - Universidade Federal do Rio Grande do Sul, Porto alegre, 2004.

MERRILL, M. D. Knowledge Objects and Mental-Models. In: **WILEY, D.** (Ed.). The Instructional Use of Learning Objects. 2000. Disponível em: <<http://www.id2.usu.edu/Papers/KOMM.PDF>>. Acesso em: out. 2004.

MODESTO, F. **Metadados**: introdução básica. São Paulo: Depto. de Biblioteconomia e Documentação da Escola de Comunicações e Artes da

Universidade de São Paulo. 2005. p. 6-35. Disponível em:
<<http://www.eca.usp.br/prof/fmodesto/textos/livrometadados.pdf>>. Acesso em: 9 set. 2010.

MONTEIRO, Fernanda de Souza. **Modelagem conceitual: a construção de uma ontologia sobre Avaliação do Ciclo de Vida (ACV) para fomentar a disseminação de seus conceitos**. 2006. 132 f. Monografia (Graduação em Biblioteconomia) – Faculdade de Ciência da Informação. Universidade de Brasília, Brasília, 2006.

MONTEIRO, Fernanda; DUTRA, Jacqueline. Plataformas de software livres para o armazenamento e descrição de objetos educacionais digitais: estudo inicial para customização do Banco Internacional de Objetos Educacionais digitais. In: ENCONTRO NACIONAL DE ENSINO E PESQUISA EM INFORMAÇÃO, 2008, Salvador. **Anais...** Salvador, UFBA, 2008.

MONTEIRO, Fernanda; JACOB, A. Uma política para desenvolver o acervo de recursos educacionais em saúde. In: ENCONTRO NACIONAL DE ENSINO E PESQUISA EM INFORMAÇÃO - CIFORM, 10. 2011, Salvador. Responsabilidade social na representação, preservação e disseminação de conteúdos. **Anais...** Salvador: UFBA, 2011.

MOULIN, N.; PEREIRA, V. Material impresso para educação a distância: formas e funções. **Caderno de Estudos e Pesquisas**, ano 3, n. 5, ago. 1999.

NASCIMENTO, Anna Christina de Azevedo. Repositórios no contexto da EaD: o Banco Internacional de Objetos Educacionais. **Enc. Bibli: R. Eletr. Bibliotecon. Ci. Inf.**, Florianópolis, v. 15, n. 29, 2010.

NDLR National Digital Learning Resources. **Site**. Disponível em:
<<https://dspace.ndlr.ie/jspui/>>. Acesso em: jul. 2012.

NG, Lai Kuen April. **Ontological model for representation of learning objectives**. 2005. 94 f. Thesis (Master of Science) – School of Interactive Arts and Technology, Simon Fraser University, Burnaby, 2005. Disponível em:
<<http://summit.sfu.ca/item/8007>>. Acesso em: out. 2012.

NILSSON, M.; JOHNSTON, P.; NAEVE, A.; POWELL, A. The future of learning object metadata interoperability. 2006. Disponível em: <<http://kmr.nada.kth.se/papers/SemanticWeb/FutureOfLOMI.pdf>>. Acesso em: out. 2010.

NILSSON, Mikael; JOHNSTON, Pete; NAEVE, Ambjörn; POWELL, Andy. Towards an interoperability framework for metadata standards. In: INTERNATIONAL CONFERENCE ON DUBLIN CORE AND METADATA APPLICATIONS: metadata for knowledge and learning, 2006, Mexico. **Proceedings...** Mexico: Universidad de Colima, 2006. p. 13-23.

NOVAK, Joseph D.; GOWIN, D. Bob. **Learning How to Learn**. Cambridge University Press, 1984c.

NOY, Natalya F.; MCGUINNESS, Deborah L. **Ontology Development 101: A Guide to Creating Your First Ontology**. 2001. Disponível em: <<http://www.ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness-abstract.html>>. Acesso em: nov. 2012.

OLIVEIRA, D. H. de. Introdução ao XML e suas aplicações. 2002. Disponível em: <http://www.xml.com.br/docs/intro_xml_apli.pdf>. Acesso em: set. 2011.

OLSEN, G. S. Stand ready? Emerging e-learning standards in a pedagogical perspective. In: WEB EDUCATION SYSTEMS IN EUROPE. **Fern Universität, Germany: ZIFF Papiere 118**. Disponível em: <<http://www.nettskolen.com/forskning/standards.pdf>>. Acesso em: mar. 2012.

PATEL-SCHNEIDER, P.; HAYES, P.; HORROCKS, I. **OWL Web Ontology Language Semantics and Abstract Syntax**. 2004. Disponível em: <<http://www.w3.org/TR/owl-semantics/>>. Acesso em: jul. 2011.

PORTAL da Saúde. **Site**. Disponível em: <http://portal.saude.gov.br/portal/saude/profissional/area.cfm?id_area=933>. Acesso em: mar. 2012.

RAMALHO, Rogério Aparecido Sá; VIDOTTI, Silvana Aparecida Borsetti Gregorio; FUJITA, Mariângela Spotti Lopes. Web semântica: uma investigação sob o olhar da Ciência da Informação. **DataGramaZero Revista de Ciência da Informação**, v.8, n.6, dez. 2007.

RANGANATHAN, S. R. **Prolegomena to library classification**. Bombay: Asia Publ.House, 1967. 640 p.

RELVÃO, **Rafael**. **Estudo sobre a utilização e interoperabilidade entre conteúdos de aprendizagem com diferentes granularidades**. 2006. 185 f. Dissertação (Mestrado em Sistemas de Informação) – Escola de Engenharia. Universidade do Minho, Guimarães, 2006.

ROBREDO, J. ; BRASCHER, M. ; MONTEIRO, Fernanda. Organização da informação: proposta de elementos de arquitetura da informação para repositórios digitais institucionais, baseados na descrição física e descrição temática. In: ROBREDO, Jaime; BRÄSCHER, Marisa. (Org.). **Passeios pelo bosque da informação: estudos sobre a representação e organização da informação e do conhecimento eroic**. Brasília: Ibict, 2010. 335 p. ISBN:

ROBSON, R. **Reusable learning**: enabling the reuse of learning content. 2004. Disponível em: <<http://www.reusablelearning.org/index.asp?id=28>>. Acesso em: jan. 2010.

RODRIGUES, Rosângela Schwarz; TAGA, Vitor; VIEIRA, Eleonora Milano Falcão. Repositórios educacionais: estudos preliminares para a Universidade Aberta do Brasil. **Perspectivas em Ciência da Informação**, v.16, n.3, p.181-207, jul./set. 2011. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/13706/12570>>. Acesso em: maio 2012.

ROUYET, J. I.; MARTÍN, V. **A comparative study of the metadata in SCORM and Dublin Core**. Disponível em: <http://www.cc.uah.es/spdece/papers/Rouyet_Final.pdf>. Acesso em: jan. 2011.

SALVE, G. B. **Modelo de planejamento para repositório de aprendizagem em organizações educacionais (MOPROA)**. 2010. 227 f. Tese (Doutorado em Engenharia de Produção) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2010.

SCORM - Sharable Content Object Reference Model. In: **Educational Standards Portal**. Disponível em: <<http://www.adlnet.gov/capabilities/scorm#tab-learn>>. Acesso em: mar. 2011.

SEMANTIC WEB. Ontology. **Site**. Disponível em: <<http://semanticweb.org/wiki/Ontology>>. Acesso em: mar. 2011.

SILVA, Nalva Neila Alves da. **A utilização da linguagem owl na definição de uma ontologia para o currículo lattes**. 2004. 112 f. Monografia (requisito da disciplina Prática de Sistemas de Informação I do curso de Sistemas de Informação) – ULBRA, Palmas, 2004.

SILVA, Patric Ferreira da. **Uso de rede de Kohonen para a clusterização de objetos de aprendizagem**. 2007. 114 f. Dissertação (Mestrado em Engenharia Elétrica)-Universidade Presbiteriana Mackenzie, São Paulo, 2007.

SILVA, Paulo Pinheiro da; MCGUINNESS, Deborah L.; FIKES, Richard. A proof markup language for Semantic Web services. **Information Systems**, v. 31, n. 4–5, p. 381–395, jun./jul 2006. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0306437905000281>>. Acesso em: fev. 2011.

SIMÕES, L.; GOUVEIA, L. Geração Net, Web 2.0 e ensino superior. 2008. In: FREITAS, E.; TUNA, S. (Orgs.). **Novos média, novas gerações, novas formas de comunicar**. [S.l.]: Universidade Fernando Pessoa, 2009. p. 21-32. Disponível em: <http://homepage.ufp.pt/lmbg/com/ls_cem6_09.pdf>. Acesso em: out. 2011.

SIMONSON, M.; SMALDINO, S.; MICHAEL, A.; ZVACEK, S. **Teaching and learning at distance: foundations of distance education**. 2. ed. New Jersey: Prentice Hall, 2002.

SLOEP, P. B. Learning objects: the answer to the knowledge economy's predicament? In: JOCHEMS, W.; KOPER, R.; MERRIËNBOER, J. VAN. (Eds.). **Integrated E-Learning**. London: Routledge/Falmer, 2003. p. 139-15.

SMITH, M.; WELTY, C.; MCGUINNESS, D. **OWL Web Ontology Language: guide**. USA: W3C, 2004. Disponível em: <<http://www.w3.org/TR/owl-guide/>>. Acesso em: out. 2011.

SOUZA, Renato Rocha; ALVARENGA, Lídia. A Web Semântica e suas contribuições para a ciência da informação. **Ci. Inf.**, Brasília, v. 33, n. 1, p. 132-141, jan./abril 2004.

SPARQL 1.1 Overview. 2012. Disponível em: <<http://www.w3.org/TR/2012/WD-sparql11-overview-20120501/>>. Acesso em: out. 2012.

STOJANOVIC, Ljiljana; STAAB, Steffen; STUDER, Rudi. **eLearning based on the Semantic Web**. [2001]. Disponível em: <<http://www.aifb.uni-karlsruhe.de/~sst/Research/Publications/WebNet2001eLearningintheSemanticWeb.pdf>>. Acesso em: jul 2007.

SVENONIUS, Elaine. **The intellectual foundation of information organization**. Cambridge: MIT press, 2001.

TAYLOR, A. G.; JOUDREY, D. N. **The organization of information**. 3. ed. Westport: Libraries Unlimited, 2009.

TEIXEIRA, F.A.G. ; DUQUE, C. G. . Colaboração na Web como ferramenta para a Recuperação da Informação. In: **Simpósio Brasileiro de Sistemas de Informação - Workshop de Teses e Dissertações**, 2009, Brasília. . Anais do Simpósio Brasileiro de Sistemas de Informação. Porto Alegre: Sociedade Brasileira de Computação, 2009. v. 5. p. 288-293.

TILLMAN, Hope N. **Evaluating Quality on the Net**. 2003. Disponível em: <<http://www.hopetillman.com/findqual.html>>. Acesso: ago. 2003.

TRISTÃO, Ana Maria Delazari; FACHIN, Gleisy Regina Bóries; ALARCON, Orestes Estevam. Sistema de classificação facetada e tesouros: instrumentos para organização do conhecimento. **Ci. Inf.**, Brasília, v. 33, n. 2, p. 161-171, maio / ago. 2004.

ULLRICH, C. Description of an Instructional Ontology and its Application in Web Services for Education, In: International Semantic Web Conference, 3., 2004. **Proceedings...** Hiroshima: ISWC2004, 2004., p. 93-94.

UMLS. **Site**. Disponível em: <<http://www.nlm.nih.gov/research/umls/>>. Acesso em: fev. 2009.

W3C. What is Linked Data? **Site**. Disponível em: <<http://www.w3.org/standards/semanticWeb/data>>. Acesso em: maio 2012.

DALY, Janet; FORGUE, Marie-Claire; HIRAKAWA, Yasuyuki. World Wide Web Consortium Issues RDF and OWL Recommendations: Semantic Web emerges as commercial-grade infrastructure for sharing data on the Web. [2004]. Disponível em: <<http://www.w3.org/2004/01/sws-pressrelease.html.en>>. Acesso em: ago. 2012.

WAGNER, E. D. **The new frontier of learning object design**. 2002. Disponível em: <<http://www.elearningguild.com/pdf/2/061802dst-h.pdf>>. Acesso em: 20 Out. 2003.

WHO. **Classifications**. 2012. Disponível em: <<http://www.who.int/classifications/icf/en/>>. Acesso em: out. 2012.

WIKIPEDIA. **Unicode**. Disponível em: <<http://en.wikipedia.org/wiki/Unicode>>.

WILEY, D. A. Connecting learning objects to instructional design theory: a definition, a metaphor, and a taxonomy. 2005. In: WILEY, D. A. (Ed.). **The instructional use of learning objects**: online version.[S.l.]: Agency for Instructional Technology, 2002. Disponível em: <<http://reusability.org/read/chapters/wiley.doc>>. Acesso em: out. 2004.

WILEY, D. A. **The instructional use of learning objects**. Bloomington: D. A. Wiley, 2000.

WISC ONLINE.**Site**. Disponível em:<<http://www.wisc-online.com/Default.aspx>>. Acesso em: nov. 2011.