



UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO DE CIÊNCIAS JURÍDICAS E ECONÔMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

VIVIANE BESSA LOPES ALVARENGA

**OBJETOS DE APRENDIZAGEM ORGANIZADOS COMO DADOS LIGADOS
INTEROPERÁVEIS NA WEB SEMÂNTICA: O REÚSO EM EVIDÊNCIA POR MEIO DE
UMA ONTOLOGIA OPERACIONAL**

VITÓRIA-ES

2023

VIVIANE BESSA LOPES ALVARENGA

**OBJETOS DE APRENDIZAGEM ORGANIZADOS COMO DADOS LIGADOS
INTEROPERÁVEIS NA WEB SEMÂNTICA: O REÚSO EM EVIDÊNCIA POR MEIO
DE UMA ONTOLOGIA OPERACIONAL**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal do Espírito Santo PPGCI-UFES como requisito para obtenção do título de Mestre em Ciência da Informação.

Orientador: Prof. Doutor Henrique Monteiro Cristovão

Linha de pesquisa 2: Memória, Representação e Informação

VITÓRIA-ES

2023

Ficha catalográfica disponibilizada pelo Sistema Integrado de
Bibliotecas - SIBI/UFES e elaborada pelo autor

A473o Alvarenga, Viviane Bessa Lopes, 1976-
Objetos de aprendizagem organizados como dados ligados
interoperáveis na Web semântica : o reúso em evidência por
meio de uma ontologia operacional / Viviane Bessa Lopes
Alvarenga. - 2023.
193 f. : il.

Orientador: Henrique Monteiro Cristovão.
Dissertação (Mestrado em Ciência da Informação) -
Universidade Federal do Espírito Santo, Centro de Ciências
Jurídicas e Econômicas.

1. Ciência da Informação. 2. Gestão do conhecimento. 3.
Normalização. 4. Ontologia. 5. Web semântica. 6. Educação à
distância. I. Cristovão, Henrique Monteiro. II. Universidade
Federal do Espírito Santo. Centro de Ciências Jurídicas e
Econômicas. III. Título.

CDU: 001

VIVIANE BESSA LOPES ALVARENGA

**OBJETOS DE APRENDIZAGEM ORGANIZADOS COMO
DADOS LIGADOS INTEROPERÁVEIS NA WEB SEMÂNTICA:
O REÚSO EM EVIDÊNCIA POR MEIO DE UMA ONTOLOGIA OPERACIONAL**

Dissertação apresentada ao Curso de Mestrado do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal do Espírito Santo (PPGCI/UFES) como requisito para a obtenção do título de Mestre em Ciência da Informação.

Linha de pesquisa 2: Memória, Representação e Informação.

Aprovada em 30 de março de 2023.

[assinatura digital]

Prof. Dr. Henrique Monteiro Cristóvão
Orientador

[assinatura digital]

Profa. Dra. Daniela Lucas da Silva Lemos
PPGCI/UFES

[assinatura digital]

Profa. Dra. Mariella Berger Andrade
IFES





Ata de Defesa e Folha de Aprovação - Viviane Bessa Lopes Alvarenga

Data e Hora de Criação: 18/04/2023 às 23:13:38

Documentos que originaram esse envelope:

- Ata de defesa - Viviane Bessa Lopes Alvarenga.pdf (Arquivo PDF) - 1 página(s)
- Folha de aprovação - Viviane Bessa Lopes Alvarenga.pdf (Arquivo PDF) - 1 página(s)



Hashs únicas referente à esse envelope de documentos

[SHA256]: e60692b08591a87bd344879e05695933b6a8e5c3180101a902322faee64c7f1c

[SHA512]: 14aca756305d3827baf2e927d11ddea67c99f50eae4c908b7f0dac83531e9623104eef276b845a98d0286b2262e60d8277654f751cff5298081b6636a2a70

Lista de assinaturas solicitadas e associadas à esse envelope



ASSINADO - Daniela Lucas da Silva Lemos (daniela.l.silva@ufes.br)

Data/Hora: 19/04/2023 - 21:19:00, IP: 187.36.169.81, Geolocalização: [-20.296746, -40.292655]

[SHA256]: 787ecdb18f42b77fb2e8050a01a4a2c6177e9edc1ad849ee9f5795c2f1d2e9c5



ASSINADO - Henrique Monteiro Cristovão (henrique.cristovao@ufes.br)

Data/Hora: 18/04/2023 - 23:27:22, IP: 177.104.235.45, Geolocalização: [-20.263252, -40.259875]

[SHA256]: 438a20e4aabe766577c66908aa3a0317a57555db9d166abf6de98a5a18961c9



ASSINADO - Mariella Berger Andrade (mariella.andrade@ufes.edu.br)

Data/Hora: 19/04/2023 - 07:05:13, IP: 200.137.82.60, Geolocalização: [-20.305717, -40.319062]

[SHA256]: c70a218f6c04a47e2a5abec911c536d2f55630f4b7efa4fb514392391af0a3b



ASSINADO - Viviane Bessa Lopes Alvarenga (viviane.alvarenga@edu.ufes.br)

Data/Hora: 19/04/2023 - 07:00:39, IP: 177.26.77.77, Geolocalização: [-20.318726, -40.378095]

[SHA256]: 51aa90e2e79a0eb5e02ce23b1623c1b7bed1c310753d554bcb531f7a87d2fefc

Histórico de eventos registrados neste envelope

19/04/2023 21:19:00 - Envelope finalizado por daniela.l.silva@ufes.br, IP 187.36.169.81

19/04/2023 21:19:00 - Assinatura realizada por daniela.l.silva@ufes.br, IP 187.36.169.81

19/04/2023 21:18:15 - Envelope visualizado por daniela.l.silva@ufes.br, IP 187.36.169.81

19/04/2023 07:05:13 - Assinatura realizada por mariella.andrade@ufes.edu.br, IP 200.137.82.60

19/04/2023 07:04:58 - Envelope visualizado por mariella.andrade@ufes.edu.br, IP 200.137.82.60

19/04/2023 07:00:39 - Assinatura realizada por viviane.alvarenga@edu.ufes.br, IP 177.26.77.77

19/04/2023 06:59:43 - Envelope visualizado por viviane.alvarenga@edu.ufes.br, IP 177.26.77.77

18/04/2023 23:27:22 - Assinatura realizada por henrique.cristovao@ufes.br, IP 177.104.235.45

18/04/2023 23:25:57 - Envelope visualizado por henrique.cristovao@ufes.br, IP 177.104.235.45

18/04/2023 23:22:31 - Envelope registrado na Blockchain por henrique.cristovao@ufes.br, IP 177.104.235.45

18/04/2023 23:22:30 - Envelope encaminhado para assinaturas por henrique.cristovao@ufes.br, IP 177.104.235.45

18/04/2023 23:13:40 - Envelope criado por henrique.cristovao@ufes.br, IP 177.104.235.45

Dedico este trabalho ao meu pai (in memorian), um Mestre, sem título ou diploma.

AGRADECIMENTOS

Primeiramente, a Deus, que me permitiu estar viva e com saúde para conseguir vencer esta etapa.

Agradeço aos meus pais que sempre me ensinaram o valor da educação e como ela é capaz de mudar a vida das pessoas.

Aos meus irmãos por me apoiarem em todos os meus projetos

Ao meu marido por seu apoio incondicional e por entender minhas ausências sempre que se fizeram necessárias nestes dois anos.

A todos os meus familiares e amigos por comemorarem minhas vitórias junto comigo.

Aos meus colegas e professores do PPGCI-UFES que souberam trocar suas experiências e se fizeram presentes mesmo numa situação de distanciamento imposto por uma Pandemia global que vivemos durante o período do curso.

A todos os professores da minha vida acadêmica que em cada etapa da minha vida deixaram um pouco deles em mim.

Aos membros internos e externos da Banca Examinadora, que desde qualificação se dispuseram e muito contribuíram com apontamentos importantes para o desenvolvimento desta pesquisa.

Neste momento em especial, quero agradecer meu orientador, Prof. Henrique Monteiro Cristovão que conduziu este trabalho com sabedoria, dedicação, gentileza e empatia.

Muito obrigada!

A ciência de hoje é a tecnologia de
amanhã.

Edward Teller

RESUMO

A quantidade de objetos de aprendizagem (OA) baseados em conteúdos desenvolvidos para uso na educação a distância têm crescido muito por conta do aumento da demanda por cursos nessa modalidade. Porém, o desenvolvimento desses OAs, em sua grande maioria são pensados para uma única plataforma, ou estão disponíveis em um repositório sem interoperabilidade, o que gera redundância de conteúdo, eleva desnecessariamente o custo de criação e manutenção, e indo de encontro à sua principal característica, que é a reusabilidade. Nesse sentido, tem sido investigado novas formas de organizar os OAs utilizando-se de processos padronizados de representação que viabilizem o acesso e o reúso. Um dos padrões elaborados com o objetivo de organizar os OAs é o LOM (Learning Object Metadata) da IEEE (Institute of Electrical and Electronics Engineers). Outro elemento indicado na representação de metadados e seus relacionamentos, com fortes características para reúso e acesso interoperável, é a ontologia operacional de domínio construída em um contexto de dados ligados na Web semântica. Dessa forma, o problema de pesquisa é: como utilizar o padrão LOM - IEEE para organizar o conhecimento sobre OAs com vistas a melhorar a capacidade de serem encontrados, acessíveis, interoperáveis e reutilizáveis em ambientes de dados abertos ligados? O objetivo geral da pesquisa é criar uma ontologia operacional de domínio para auxiliar a organização de OAs no padrão LOM-IEEE considerando-se o contexto da Web semântica. Com abordagem qualitativa, natureza aplicada e procedimentos de estudo de caso, a pesquisa foi desenvolvida de forma organizada em oito etapas procedimentais, sintonizadas com os objetivos específicos, que vão desde a investigação de repositórios de objetos de aprendizagem, até a construção da ontologia operacional e o seu teste com a criação de uma base de dados ligados e consultas como prova de conceito. Como resultados da pesquisa, destacam-se, a adequação do padrão LOM-IEEE para a representação de OAs; um modelo de design para a ontologia operacional de domínio; uma ontologia operacional de domínio implementada em Turtle; a indicação de elementos interoperáveis das camadas estrutural e semântica; um esquema de mapeamento para dados ligados RDF dos OAs; e testes de recuperação de informação por meio de consultas SPARQL sobre a base de OAs. Indica-se, como continuidade da pesquisa, a ampliação e publicação da base de dados ligados de OAs, englobando outros repositórios, e o desenvolvimento de uma interface com boa usabilidade para consultas aos OAs, bem como a sua validação com um corpo de usuários.

Palavras-chave: Objetos de aprendizagem. Organização do conhecimento. Padrão LOM-IEEE. Ontologia operacional de domínio. Dados Ligados. Web semântica.

ABSTRACT

The amount of learning objects (LO) based on content developed for use in distance education has grown a lot due to the increased demand for courses in this modality. However, the development of these LOs, for the most part, are designed for a single platform, or are available in a repository without interoperability, which generates content redundancy, unnecessarily increases the cost of creation and maintenance, and going against its main feature, which is reusability. In this sense, new ways of organizing LOs have been investigated using standardized representation processes that enable access and reuse. One of the standards developed with the objective of organizing the LOs is the LOM (Learning Object Metadata) of the IEEE (Institute of Electrical and Electronics Engineers). Another element indicated in the representation of metadata and their relationships, with strong characteristics for reuse and interoperable access, is the domain operational ontology built in a context of linked data in the Semantic Web. Thus, the research problem is: how to use the LOM - IEEE standard to organize knowledge about LOs in order to improve their ability to be found, accessible, interoperable and reusable in linked open data environments? The general objective of the research is to create a domain operational ontology to help the organization of LOs in the LOM-IEEE standard considering the context of the Semantic Web. With a qualitative approach, applied nature and case study procedures, the research was developed in an organized way in eight procedural stages, in tune with the specific objectives, ranging from the investigation of learning object repositories, to the construction of the operational ontology and your test with the creation of a linked database and queries as a proof of concept. As research results, the adequacy of the LOM-IEEE standard for representing LOs stands out; a design model for the domain operational ontology; a domain operational ontology implemented in Turtle; the indication of interoperable elements of the structural and semantic layers; a mapping scheme for OAs RDF linked data; and information retrieval tests through SPARQL queries on the LO base. It is indicated, as a continuation of the research, the expansion and publication of the linked database of LOs, encompassing other repositories, and the development of an interface with good usability for queries to LOs, as well as its validation with a body of users.

Keywords: Learning objects. Knowledge organization. LOM-IEEE standard. Domain operational ontology. Linked Data. Semantic Web.

LISTA DE FIGURAS

Figura 1 - Publicações sobre “Learning Object” Web of Science.....	25
Figura 2 - Uma visão geral das estruturas e funções do SOC.....	36
Figura 3 - Tipos de ontologias, de acordo com seu grau de dependência de uma determinada tarefa ou ponto de vista.....	43
Figura 4 - Processo de desenvolvimento de ontologia SABiO.....	45
Figura 5 - Dados FAIR.....	58
Figura 6 - Processo de FAIRificação dos dados.....	59
Figura 7 - Metáfora Lego.....	61
Figura 8 - Metáfora Átomo.....	62
Figura 9 - Tipo de OA - Imagem.....	64
Figura 10 - Tipo de OA - Áudio.....	65
Figura 11 - Tipo de OA – Vídeo.....	65
Figura 12 - Tipo de OA - Hipertexto.....	66
Figura 13 - Tipo de OA - Simulação.....	68
Figura 14 - Tipo de OA - Software.....	68
Figura 15 - Estágios do ciclo de vida de um OA.....	70
Figura 16 - Acontecimentos ligados ao movimento REA.....	72
Figura 17 - Esquema Base do LOM/IEEE-2020 - Categorias de 1 a 4.....	80
Figura 18 - Esquema Base do LOM/IEEE-2020 - Categorias de 5 a 9.....	81
Figura 19 - Camadas da Web semântica.....	84
Figura 20 - Representação em Triplas.....	87
Figura 21 - Padrões e recomendações que tratam de problemas de interoperabilidade.....	92
Figura 22 - Cobertura das recomendações da ISO 25964-2 para interoperabilidade entre tesouros e outros tipos de KOS.....	93
Figura 23 - Mapa conceitual fundamentação teórica.....	97
Figura24.....	110
Figura 25 - Mapa mental do esquema base do LOM/IEEE-2020.....	118
Figura 26 – Planilha com representação dos OA seguindo o padrão LOM.....	121
Figura 27 – Dados da amostra de OAs trabalhados no OpenRefine.....	122
Figura 28 – Modelo de design para a ontologia operacional de domínio do padrão LOM-IEEE - Categorias General, Life Cycle, Technical e Meta Metadata.....	123
Figura 29 – Modelo de design para a ontologia operacional de domínio Diagrama estrutural da Ontologia do padrão LOM-IEEE- Categoria Educational.....	125
Figura 30 - Modelo de design para a ontologia operacional de domínio Diagrama estrutural da Ontologia do padrão LOM-IEEE - Categorias Rigths, Relation, Annotation e Classification.....	126
Figura 31 - Namespaces usados na ontologia.....	127
Figura 32 - Meta-metadados da ontologia.....	128
Figura 33 - Recorte do código contendo a classe principal que constitui o OA e as três primeiras propriedades OWL.....	129
Figura 34 – Recorte do código da implementação dos três primeiros vocabulários da ontologia operacional.....	129
Figura 35 – Recorte do código da implementação dos três primeiros conceitos/termos dos vocabulários da ontologia operacional.....	130
Figura 36 – Recorte do código da implementação das três primeiras propriedades que conectam o OA aos valores literais.....	131

Figura 37 - Equivalência das classes ':LearningObject' e 'wdo:Q106261'.....	132
Figura 38 - Código Turtle com recorte de duas propriedades, ':generalIdentifierEntry' e ':generalTitle'.....	135
Figura 39 - Propriedade interoperáveis Dublin Core associadas ao indivíduo 'data:OA001'.....	136
Figura 40 – Mapeamento dos dados dos OAs.....	138
Figura 41 – Mapeamento dos dados dos OAs - Detalhe.....	139
Figura 42 – Mapeamento os dados dos OAs com xsd - Detalhe.....	139
Figura 43 – Mapeamento dos dados dos OAs - Idioma.....	140
Figura 44 - Recorte do mapeamento realizado com o GraphDB, registrado no formato JSON.....	141
Figura 45 - Dados ligados RDF gerados pelo mapeamento no GraphDB.....	142
Figura 46 - Características educacionais do objeto de aprendizagem 'data:OA001'.....	143
Figura 47 - Palavras-chave do OA 'data:OA001'.....	144
Figura 48 - Relação entre algumas propriedades dos objetos de aprendizagem 'data:OA001', 'data:OA002' e 'data:OA003'.....	145
Figura 49 – Consulta SPARQL com nome, tipo, formato e identificador.....	146
Figura 50 – Resultado da consulta SPARQL com nome, tipo, formato e identificador.....	146
Figura 51 – Consulta SPARQL dos OAs que trabalham com "gestão do conhecimento" e que possuam nível de interatividade baixo (Low) e nível de dificuldade fácil (Easy).....	147
Figura 52 – Resultado da consulta SPARQL dos OAs que trabalham com "gestão do conhecimento" e que possuam nível de interatividade baixo (Low) e nível de dificuldade fácil (Easy).....	147
Figura 53 – Consulta SPARQL dos OAs que tenham o termo 'saúde' no título ou nas palavras-chave.....	148
Figura 54 – Resultado da consulta SPARQL dos OAs que tenham o termo 'saúde' no título ou nas palavras chave.....	148
Figura 55 - Relacionamento entre as características educacionais dos OAs cujo usuário é aprendiz - consulta.....	149
Figura 56 - Relacionamento entre as características educacionais dos OAs cujo usuário é aprendiz - resultado.....	150

LISTA DE QUADROS

Quadro 1 - Paralelo entre os períodos históricos da humanidade e os períodos históricos da Catalogação.....	48
Quadro 2 – Principais diferenças entre os modelos da família FR e o IFLA LRM.....	52
Quadro 3 - Representação de um documento segundo o padrão de estrutura de metadados descritivo MARC e o padrão de conteúdo AACR2.....	54
Quadro 4 – Tipo de OA – Formas de Animação.....	67
Quadro 5 – Questões sob as perspectivas humana e técnica.....	71
Quadro 6 – Paralelo entre os princípios FAIR e as características técnicas dos OAs 95	
Quadro 7 – Resultados para os descritores nas bases de dados Google Scholar e Portal Capes.....	108
Quadro 8 – Categorias funcionais do LOM-IEEE segundo Barbedo, Corujo e Sant’ana (2011).....	119
Quadro 9 – Equivalências entre tipos LOM-IEEE e RDF.....	124
Quadro10 – Interoperabilidade da camada estrutural entre LOM-IEEE e DublinCore 132	
Quadro11 – Interoperabilidade da camada estrutural entre LOM-IEEE e Bibframe 136	

SUMÁRIO

1 INTRODUÇÃO.....	15
1.1 PROBLEMA DE PESQUISA.....	18
1.2 OBJETIVO GERAL.....	19
1.3 OBJETIVOS ESPECÍFICOS.....	19
1.4 JUSTIFICATIVA.....	20
2 ESTADO DA ARTE.....	24
2.1 OBJETOS DE APRENDIZAGEM COMO DADOS LIGADOS INTEROPERÁVEIS NA WEB SEMÂNTICA.....	24
2.2 USO DO PADRÃO LOM IEEE/2020 NA REPRESENTAÇÃO DE OBJETOS DE APRENDIZAGEM.....	27
2.3 ONTOLOGIA.....	30
3 FUNDAMENTAÇÃO TEÓRICA.....	31
3.1 ORGANIZAÇÃO DO CONHECIMENTO NA CIÊNCIA DA INFORMAÇÃO.....	31
3.1.1 Sistemas de Organização do Conhecimento.....	35
3.1.2 Metadados.....	38
3.1.3 Taxonomias.....	40
3.1.4 Ontologias.....	41
3.2 CATALOGAÇÃO.....	46
3.2.1 Fundamentos.....	46
3.2.2 Padrões de Conteúdo para Catalogação.....	49
3.2.3 Padrões de Estrutura de Metadados Descritivos para Catalogação.....	53
3.3 PRINCÍPIOS FAIR.....	56
3.4 OBJETOS DE APRENDIZAGEM.....	61
3.4.1 Tipos de Objetos de Aprendizagem.....	64
3.4.2 Características dos Objetos de Aprendizagem.....	69
3.4.2 Ciclo de Vida do Objeto de Aprendizagem.....	70
3.4.3 Recursos Educacionais Abertos - REA.....	72
3.4.4 Repositório de Objetos de Aprendizagem.....	73
3.4.5 Padrão LOM IEEE-2020.....	76
3.5 WEB SEMÂNTICA.....	83
3.5.1 Fundamentos.....	84

3.5.2 Dados Ligados.....	86
3.5.3 Interoperabilidade.....	91
3.6 CONSIDERAÇÕES FINAIS DA FUNDAMENTAÇÃO TEÓRICA.....	94
4 FUNDAMENTAÇÃO METODOLÓGICA.....	98
4.1 ORGANIZAÇÃO DA PESQUISA.....	98
4.1.1 Mapas Conceituais.....	98
4.1.2 Gerenciador Bibliográfico.....	99
4.2 DESENVOLVIMENTO DA PESQUISA.....	99
4.2.1 Software de Preparação de Dados.....	100
4.2.2 Software para Implementação de Ontologias Operacionais de Domínio.....	101
4.2.3 Sistemas de Banco de Dados Orientado a Grafos.....	103
4.2.4 Linguagem de Consultas em Bases de Dados Ligados.....	104
5 PROCEDIMENTOS METODOLÓGICOS.....	106
5.1 CLASSIFICAÇÃO DA PESQUISA E CENÁRIO DA PESQUISA.....	106
5.2 LEVANTAMENTO BIBLIOGRÁFICO.....	107
5.3 ETAPAS PROCEDIMENTAIS DO DESENVOLVIMENTO.....	109
5.3.1 Etapa 1: Investigação e seleção de um padrão de metadados para representar OAs.....	111
5.3.2 Etapa 2: Investigação de repositórios de objetos de aprendizagem.....	112
5.3.3 Etapa 3: Organização dos metadados necessários para a representação de OAs.....	113
5.3.4 Etapa 4: Elaboração do modelo de design para a ontologia operacional de domínio.....	114
5.3.5 Etapa 5: Implementação da ontologia operacional de domínio.....	115
5.3.6 Etapa 6: Investigação de propriedades e indivíduos equivalentes para a interoperabilidade das camadas estrutural e semântica.....	115
5.3.7 Etapa 7: Construção do esquema de mapeamento e geração da base de dados ligados RDF.....	116
5.3.8 Etapa 8: Elaboração de consultas SPARQL como prova de conceito da funcionalidade da ontologia operacional implementada.....	116
6 RESULTADOS E DISCUSSÃO.....	117
6.1 METADADOS PARA O PADRÃO LOM-IEEE.....	117
6.2 OBJETOS DE APRENDIZAGEM DO REPOSITÓRIO ARCA ORGANIZADOS CONFORME METADADOS DO PADRÃO LOM-IEEE.....	120

6.3 MODELO DE DESIGN PARA A ONTOLOGIA OPERACIONAL DE DOMÍNIO.....	122
6.4 ONTOLOGIA OPERACIONAL DE DOMÍNIO.....	127
6.5 INTEROPERABILIDADE DAS CAMADAS ESTRUTURAL E SEMÂNTICA.....	131
6.6 MAPEAMENTO E BASE DE DADOS LIGADOS RDF.....	137
6.7 CONSULTAS SPARQL.....	142
6.8 DISCUSSÕES GERAIS.....	151
7 CONSIDERAÇÕES FINAIS.....	156
REFERÊNCIA.....	158
ANEXO I - LISTA DE REPOSITÓRIOS DE OAs (Sbrogio e Valente, 2020).....	174
ANEXO II - TABLE 1 – LOMV1.0 BASE SCHEMA.....	176
ANEXO III - TABLE 2 – LANGSTRING.....	189
ANEXO IV - TABLE 3 – DATETIME.....	190
ANEXO V - TABLE 4 – DURATION.....	191
ANEXO VI - TABLE 5 – VOCABULARY.....	192
ANEXO VII - TABLE A.1 – MAPPING TO UNQUALIFIED DUBLIN CORE.....	193

1 INTRODUÇÃO

A Educação a Distância (EaD) foi criada com o objetivo de desenvolver o processo de ensino e aprendizagem mediado pelas tecnologias digitais da informação e comunicação (TDICs) solucionando o problema da exigência de tempo síncrono e, principalmente, da distância entre o estudante e a instituição. Moore e Kearsley (2007 p. 2), definem EaD como,

aprendizado planejado que ocorre normalmente em um lugar diferente do local do ensino, exigindo técnicas especiais de criação do curso e de instrução, comunicação por meio de várias tecnologias e disposições organizacionais e administrativas especiais (MOORE; KEARSLEY 2007, p. 2).

Esta modalidade de educação é pautada pelos seus aspectos tecnológicos, dentre eles, Ambientes Virtuais de Aprendizagem (AVAs), plataforma computacional onde as interações dos sujeitos envolvidos no processo acontecem. Behar (2006 apud Behar, 2009, p. 29), define AVA como,

...um espaço na internet formado pelos sujeitos e suas interações e formas de comunicação que se estabelecem por meio de uma plataforma, tendo como foco principal a aprendizagem. Entende-se por plataforma uma estrutura tecnológica composta pelas funcionalidades e interface gráfica que compõe o AVA.

A autora ressalta que neste novo espaço pedagógico surge também um novo modelo, que a mesma chama de “modelo pedagógico”, que está fortemente vinculado às TICs e, particularmente, aos Ambientes Virtuais de Aprendizagem (AVAs), utilizados para mediar e promover a educação a distância. Neste contexto, a autora ainda destaca que a produção de materiais educacionais digitais na forma de objetos de aprendizagem (OAs)

[...] recurso viável para enriquecer o espaço pedagógico” - tem sido uma opção mais dinâmica e interativa, na qual a utilização destes objetos remete a um novo tipo de aprendizagem onde o professor deixa de ser apenas um transmissor de informações para ser um mediador da aprendizagem. (BEHAR, 2009, p.66).

De acordo com Silva (2011), os OAs ajudam na compreensão do todo, no que diz respeito à matéria de um curso, “atomizando” o conteúdo, ou seja, decompondo-o em partes menores de forma a ilustrar para o aprendiz como cada elemento contribui para o funcionamento do todo. “É como um relojoeiro que desmonta um relógio para um visitante, explicando as propriedades e funções de cada peça, individualmente, e o funcionamento do relógio como um todo.” (SILVA, 2011, p. 17). Já Wiley (2000), cita

Wayne Hodgins que em 1996 usa a metáfora dos blocos de montar tipo Lego¹ para ilustrar o conceito de usabilidade dos OAs, onde nem todos os blocos combinam entre si, ou seja, nem todos os OAs podem ser combinados com qualquer outro OAs, tal como ocorre com esses blocos. Essas metáforas ressaltam a característica fundamental inerente aos OAs que é a de organizar o conteúdo educacional em pequenos segmentos combinados entre si, formando resultados gradativamente mais complexos.

Não somente, mas principalmente, no contexto da EaD, o que vem tornando o conceito de OA tão importante, é o suporte que as tecnologias têm oferecido de forma a possibilitar sua organização e catalogação, o que tem facilitado o seu acesso e o reuso. Porém, organizar as informações referentes a estes OAs de forma a torná-los acessíveis e reutilizáveis, traz à tona os desafios e a complexidade da gestão de grandes volumes de conteúdos educacionais digitais. Por isso, visando essa recuperação e reutilização, vários organismos internacionais como *IMS - Global standards*, *ADL - Advanced Distributed Learning Initiative* e *IEEE - Institute of Electrical and Electronics Engineers* vêm trabalhando colaborativamente para criar padrões de metadados para catalogação destes OAs.

Neste processo de padronização surgiu o *LOM - Learning Object Metadata (IEEE, 2020)*, um padrão internacional chancelado pelo *IEEE - Institute of Electrical and Electronics Engineers* (2009, p. 388) que "[...] possui blocos independentes dos conteúdos de aprendizagem que podem ter referências a outros objetos e ser combinados ou sequenciados visando construir grandes unidades educacionais" (tradução nossa).

Estes "blocos independentes" são metadados que existem para categorizar, indexar, recuperar, reutilizar e combinar diferentes OAs, onde esses padrões formam um conjunto mínimo de atributos que permitem a caracterização, localização e avaliação destes objetos.

Neste contexto, surgem os repositórios onde são armazenados os OAs e disponibilizados em seus diversos formatos (textos, vídeos, áudios, imagens etc.). Também existem os repositórios com acesso Web onde não se faz o armazenamento dos recursos

¹ O LEGO é um brinquedo produzido pelo LEGO Group, cujo conceito original se baseia em um sistema patenteado de peças de plástico que se encaixam, permitindo inúmeras combinações. O Grupo **LEGO** foi fundado no dia 10 de agosto de 1932, na Dinamarca, pelo carpinteiro Ole Kirk Kristiansen. Nascida como uma pequena oficina, a empresa era responsável pela fabricação de brinquedos artesanais de madeira. Disponível em: (<https://canaltech.com.br/empresa/lego/>).

propriamente ditos, mas dos metadados (catalogação extensa) que indicam quais são os repositórios que detêm recursos sobre determinado assunto, sendo um agregador de fonte de informação no ciberespaço.

O metadado de um OA descreve as características mais relevantes utilizadas na sua catalogação num repositório, permitindo assim, a sua recuperação por meio de sistemas de busca na Web ou mesmo num Ambiente Virtual de Aprendizagem.

A disponibilização destas informações na Web, utilizando padrões de metadados, permite a troca e reaproveitamento destes recursos. Mas, para que isso aconteça, é necessário que haja um processo de representação do conhecimento, utilizando-se vocabulários controlados, que segundo Kobashi (2008) é uma linguagem artificial constituída de termos/descriptores organizados numa estrutura relacional de forma a padronizar a troca de informações e ontologias que, na ciência da informação de acordo com a ideia defendida por Almeida (2013, p.1690-1691), “[...] é um tipo de vocabulário controlado na medida em que especifica o conjunto de predicados que podem ser usados para fazer declarações (ou seja, representações) sobre um recurso”. Desta forma, segundo o autor, uma ontologia é vista como “[...] um inventário das coisas no mundo e das relações entre elas em um domínio particular”. Então, seguindo essa premissa, pode-se inferir que as ontologias possibilitam a formação de relações entre vários conceitos/termos/descriptores e podem propiciar a representação de relações com um mesmo objeto como, por exemplo, objetos de aprendizagem. Desta forma, ao criar uma relação intencional de conceitos, essa rede conceitual cria relações de significação e de redes semânticas constituídas pelos termos que ali se encontram, de forma que não somente os seres humanos compreendam, mas, utilizando-se de ferramentas específicas, permite que as máquinas, na linguagem computacional, identifiquem estes conceitos na esfera da chamada Web semântica.

Segundo Berners-Lee, Hender e Lassila (2001, p. 2, tradução nossa), “A Web semântica não é uma Web separada, mas uma extensão da atual, na qual a informação recebe um significado bem definido, permitindo uma melhor cooperação no trabalho entre computadores e pessoas”, mas, para que isso aconteça, é importante utilizar os princípios “*Linked Data*”, ou “Dados Ligados”, que de acordo com Nhacuongue, Rozsa e Dutra (2018, p. 24) “[...] é um conjunto de dados publicados na Web, em formatos que sejam legíveis por máquinas, com estruturas semânticas bem definidas, que estejam ligados a

outros dados externos. É uma maneira de compartilhar, reutilizar e conectar recursos e usuários”. E, é neste contexto, que entram em cena os chamados princípios *FAIR*, um acrônimo para *Findable, Accessible, Interoperable e Reusable*, que estão presentes nas discussões e práticas contemporâneas da Ciência da Informação desde o início de 2014, fruto de uma conferência internacional denominada "*Jointly Designing a data FAIRPORT*" (*Jointly Designing a data FAIRPORT, 2014*) estes princípios enfatizam o aprimoramento da capacidade das máquinas de encontrar e usar automaticamente os dados, além de apoiar sua reutilização por indivíduos, criando o que se chama interoperabilidade semântica que, de acordo com a *Global Standards Harmonization*, citada por Dolin e Auschuler (2011, p. 99, tradução nossa, grifo nosso), é a “[...] capacidade dos dados compartilhados por sistemas serem compreendidos no nível de conceitos de domínio totalmente definidos”. Outra definição de interoperabilidade semântica diz que é “[...] a capacidade de dois ou mais sistemas de computador trocar informações e ter o significado dessa informação automaticamente interpretado pelo sistema receptor com precisão suficiente para produzir resultados úteis, conforme definido pelos usuários finais de ambos os sistemas.” (DOLIN; AUSCHULER, 2011, p. 99).

Neste processo de interoperabilidade semântica é importante que os dados sejam estruturados, ou seja, organizados e armazenados de uma maneira pré-definida, de forma a codificar o significado e preservar as relações entre diferentes pontos dos dados dentro de um conjunto. Desta forma, esses dados podem ser publicados na Web em ambientes de colaboração como a Wikidata².

1.1 PROBLEMA DE PESQUISA

A quantidade de OAs resultante da construção dos conteúdos que são disponibilizados nos ambientes virtuais de aprendizagem como vídeos, animações, textos, imagens, páginas Web, jogos, áudio, etc. têm se avolumado, por conta do aumento da demanda por cursos na modalidade EaD. Além disso, existe complexidade na produção, catalogação, armazenamento, recuperação e reutilização, pois estes objetos são pensados normalmente para uma única plataforma ou repositório, desperdiçando grande parte do potencial de uso. Os OAs são desenvolvidos de forma isolada o que gera redundância de conteúdo e eleva os custos de criação e manutenção durante todo o seu ciclo de vida.

² Wikidata é uma base de dados universal, editada colaborativamente, com intuito de oferecer suporte interoperável para diversos projetos. Disponível em: <https://www.wikidata.org/>.

Silveira (2014) aborda que, mesmo com avanços dos processos de padronização, alguns resultados de pesquisas empíricas relacionadas ao reúso efetivo de OAs têm mostrado baixos percentuais de reúso, algo em torno de 20%. Além disso, características como a granularidade - entendido aqui como seu nível de detalhe, ou seu “tamanho” - não é o único fator a ser considerado no potencial de reúso. O grau de liberdade do objeto, segundo o autor, é possivelmente o mais preponderante na maior parte das vezes, pois os OAs acabam ficando “presos” em AVAs ou repositórios institucionais sem serem vistos, acessados, sem interação com outros sistemas, e conseqüentemente, sem serem reutilizados.

Por isso, considerando os repositórios de OAs de Acesso Aberto como ponto de partida e analisando os OAs contidos nestes repositórios no contexto de “FAIRificação”, (termo comumente utilizado para aplicação dos princípios FAIR), segue-se com o intuito de responder à seguinte questão, sem perder de vista a interlocução deste processo com a Ciência da Informação: como organizar o conhecimento sobre OAs com vistas a melhorar a capacidade de serem encontrados, acessíveis, interoperáveis e reutilizáveis?

1.2 OBJETIVO GERAL

Criar uma ontologia operacional de domínio para auxiliar a organização de OAs no contexto de dados ligados interoperáveis na Web semântica.

1.3 OBJETIVOS ESPECÍFICOS

1. Investigar e selecionar um padrão de metadados para representar OAs;
2. Investigar e analisar as peculiaridades de alguns OAs disponíveis, em repositórios selecionados, fazendo um levantamento dos elementos necessários para a sua representação no padrão de metadados escolhido;
3. Organizar os metadados necessários para a representação de OAs em função do padrão selecionado no objetivo específico 1 e dos OAs selecionados no objetivo específico 2;
4. Elaborar um modelo de design para a ontologia operacional de domínio que represente o padrão selecionado no objetivo específico 1;

5. Implementar a ontologia operacional de domínio conforme o modelo desenvolvido no objetivo específico 4 com a incorporação de dados de alguns OAs;
6. Investigar propriedades e indivíduos equivalentes para a interoperabilidade das camadas estrutural e semântica;
7. Construir um esquema de mapeamento para receber a ontologia operacional criada no objetivo específico 5 juntamente com os elementos interoperáveis investigados no objetivo específico 6, os dados dos OAs selecionados no objetivo específico 2 e organizados com os metadados selecionados no objetivo específico 3, para criação de uma base de dados ligados RDF³;
8. Fazer consultas SPARQL⁴ sobre a base RDF gerada no objetivo específico 7 como prova de conceito da funcionalidade da ontologia operacional implementada.

1.4 JUSTIFICATIVA

A expansão da educação a distância, que vem ocorrendo gradativamente desde o início do segundo milênio, deve-se, principalmente, no interior do país, à implementação das políticas governamentais, como o sistema Universidade Aberta do Brasil (UAB), criado pelo Decreto nº 5.800, de 8 de julho de 2006, que em seu Artigo 1º diz que o Sistema UAB é um sistema: “voltado para o desenvolvimento da modalidade de educação a distância, com a finalidade de expandir e interiorizar a oferta de cursos e programas de educação superior no País” (BRASIL, 2006). Essa expansão tem aumentado a oferta de objetos de aprendizagem em vários repositórios, impondo desafios e complexidades não triviais ao desenvolvimento e gestão de grandes volumes de conteúdo educacional. Esses acervos de OAs das instituições são de grande importância, tornando-se recursos valiosos que cada vez mais precisam ser preservados e estarem prontamente disponíveis para recuperação e reutilização. Mas, nem sempre estas informações são facilmente encontradas, sendo que muitas delas ainda são esquecidas.

³ RDF é um modelo padrão para intercâmbio de dados na Web. O RDF possui recursos que facilitam a mesclagem de dados, mesmo que os esquemas subjacentes sejam diferentes, e suporta especificamente a evolução dos esquemas ao longo do tempo, sem exigir que todos os consumidores de dados sejam alterados. Disponível em: <https://www.w3.org/RDF/>

⁴ O SPARQL é um conjunto de especificações que fornecem linguagens e protocolos para consultar e manipular o conteúdo publicado em RDF na Web. (SANTARÉM SEGUNDO, 2014, p. 3871)

Alsubaie (2009) destaca que a falta de orientação para a produção e armazenamento de OAs, assim como a orientação para seu reúso correto, são elementos críticos para que a taxa de reutilização seja tão baixa. Assim, a proposição de um padrão a ser utilizado como base na representação de metadados dos OAs, como é o caso do LOM/IEEE, é apresentado como uma estratégia para resolver esta questão e propor uma orientação para a organização desses recursos, pois apesar de já estar disponível desde de 2002, um estudo conduzido por Reis e Ferneda (2016), onde os autores traçam um panorama sobre a utilização destes padrões em cinco repositórios nacionais, demonstrou que,

embora os padrões de metadados sejam extensos e complexos, na prática poucos atributos são utilizados para catalogar os recursos presentes nos repositórios, além da inserção de informações incorretas, ambíguas e, muitas vezes, imprecisas, o que impacta diretamente na redução das possibilidades de interoperabilidade, reusabilidade e recuperação dos objetos de aprendizagem (REIS; FERNEDA, 2016, p.1154).

E concluem dizendo que embora as pesquisas sobre o assunto tenham dado imensas contribuições, ainda existem muitas questões em aberto relativas à utilização dos objetos de aprendizagem, bem como uma carência por ferramentas que efetivamente possam suprir as necessidades informacionais. Assim, tendo como base os princípios FAIR, torna-se necessário propor mecanismos que ajudem na recuperação desses objetos e conseqüentemente, o seu reúso.

Isto posto, vislumbra-se que o conhecimento gerado por esta pesquisa possa ser de grande relevância tanto para a área de Educação quanto para as áreas da Ciência da Informação e Sistemas de Informação, visto que urge a necessidade de entendimento dos fenômenos informacionais que se relacionam a este processo que envolve a criação, a disponibilização e a recuperação de OAs no âmbito da EaD.

Esta pesquisa também trabalha o contexto informacional da interoperabilidade semântica, o que vai ao encontro da necessidade levantada no trabalho de Weiss (2021) em que a autora faz um levantamento das pesquisas sobre interoperabilidade semântica na Ciência da Informação e conclui que,

[...] os resultados alcançados nesse artigo apontam para a necessidade de pesquisas que adotem e explicitem claramente a perspectiva teórica que fundamenta o desenvolvimento dos processos e/ou produtos que tem como propósito final a interoperabilidade semântica. (WEISS, 2021, p. 447)

Segundo a autora, muitas vezes as pesquisas sobre este assunto, por serem exclusivamente tecnicistas, se limitam a descrever o desenvolvimento de processos e produtos sem se preocupar em apresentar os fundamentos que estão por trás dos mesmos. Por isso, esta pesquisa, além de ser uma pesquisa aplicada, passa pelo processo teórico da representação do conhecimento apresentando os fundamentos e conceitos envolvidos na interoperabilidade semântica que são intimamente ligados à Ciência da Informação.

Neste contexto, Nhacuongue, Rozsa e Dutra (2018) argumentam que há esforços das instituições no domínio da educação para publicar informações sobre ofertas de cursos, recursos educacionais abertos de forma acessível e reutilizável. Porém, mesmo enfatizando soluções no contexto de *Linked Data* na Web semântica, para transformar a educação, os autores admitem que estas ainda não são amplamente adotadas.

Assim, este trabalho se propõe a contribuir com a difusão do conhecimento que poderá gerar ações futuras em prol dos princípios de *Linked Data* na Web semântica e das iniciativas sobre os recursos educacionais abertos como um todo onde novos objetos podem ser interligados a partir daqueles já existentes, criando um ambiente altamente colaborativo e eficiente no que diz respeito ao reaproveitamento de recursos.

A dissertação foi estruturada da seguinte forma: na presente seção são apresentados o contexto da discussão, a temática, o problema de pesquisa, o objetivo geral e os objetivos específicos e a justificativa. Na seção dois, apresentam-se o Estado da arte num panorama sobre os trabalhos sobre o tema. Na seção três, disserta-se sobre os principais conceitos apresentados no trabalho divididos em 3.1 Organização do Conhecimento, neste contexto apresenta-se os Sistemas de Organização do Conhecimento, Metadados, Taxonomias e Ontologias. Na seção 3.2 trabalha-se os fundamentos da Catalogação e os padrões de conteúdo e de estrutura de metadados descritivos. Na seção 3.3 traz-se os princípios FAIR. Na seção 3.4 discorre-se sobre os Objetos de aprendizagem destacando os tipos, características ciclo de vida, recursos educacionais abertos, repositórios e apresentamos o padrão LOM-IEEE-2020. Na seção 3.5 recupera-se os fundamentos de Web semântica, dados ligados e interoperabilidade e fecha-se a seção com algumas considerações com o apanhado de tudo que foi trabalhado. A seção 4, fundamentação metodológica é onde mostra-se como a pesquisa foi organizada e desenvolvida e apresenta-se as principais ferramentas utilizadas. Na seção 5 trata-se os procedimento

metodológicos, o levantamento bibliográfico e apresenta-se as oito etapas procedimentais do desenvolvimento. Na seção 6 apresenta-se os resultados e as discussões com a bibliografia sobre o tema. Na seção 7, apresentam-se as considerações finais a partir dos resultados obtidos por esta pesquisa.

2 ESTADO DA ARTE

A presente pesquisa aborda diferentes áreas do conhecimento entre elas, Ciência da Informação, a Ciência da Computação, Sistemas de Informação e Educação, assim como envolve vários termos e assuntos, que, se tratados individualmente, podem confundir o escopo ou extrapolar o recorte que se pretende abordar. Desta forma, esta seção é dedicada a uma revisão de literatura na qual se procura levantar o que as pesquisas mais recentes estão discutindo sobre os assuntos abordados no presente trabalho.

Como há uma junção de diferentes temáticas na proposta do trabalho, esse capítulo foi organizado em duas seções, separando as temáticas em dois grupos, para que houvesse maior possibilidade de encontrar trabalhos correlatos considerando-se a relevância baseada na sua contemporaneidade e inovação. Dessa forma, as duas seções foram organizadas como se segue: (1) Objetos de aprendizagem como dados ligados interoperáveis na Web semântica e; (2) uso do padrão LOM IEEE/2020 na representação de objetos de aprendizagem.

Os métodos empregados para busca de literatura especializada de forma a satisfazer as temáticas do estado da arte são apresentados no capítulo 5, Procedimentos Metodológicos.

2.1 OBJETOS DE APRENDIZAGEM COMO DADOS LIGADOS INTEROPERÁVEIS NA WEB SEMÂNTICA

Com o desenvolvimento das TDICs e a aplicação dessas no contexto educacional, principalmente no que se refere à oferta de cursos na modalidade a distância, é possível observar que muito tem se pesquisado sobre OAs e a aplicabilidade dos mesmos em várias áreas do conhecimento. Ao fazer uma pesquisa simples utilizando o descritor “*LEARNING OBJECTS*” no Google Scholar⁵ foram encontrados mais de oitenta mil registros sobre o tema. Já no Portal de Periódicos da Capes⁶, a mesma busca recuperou mais de seis mil registros onde se pode observar um aumento da produção acadêmica a partir de 2003 e um ápice em 2017 de trabalhos sobre o tema. Ao fazer a mesma pesquisa na base de dados da Web Of Science⁷ são recuperadas cento e vinte quatro

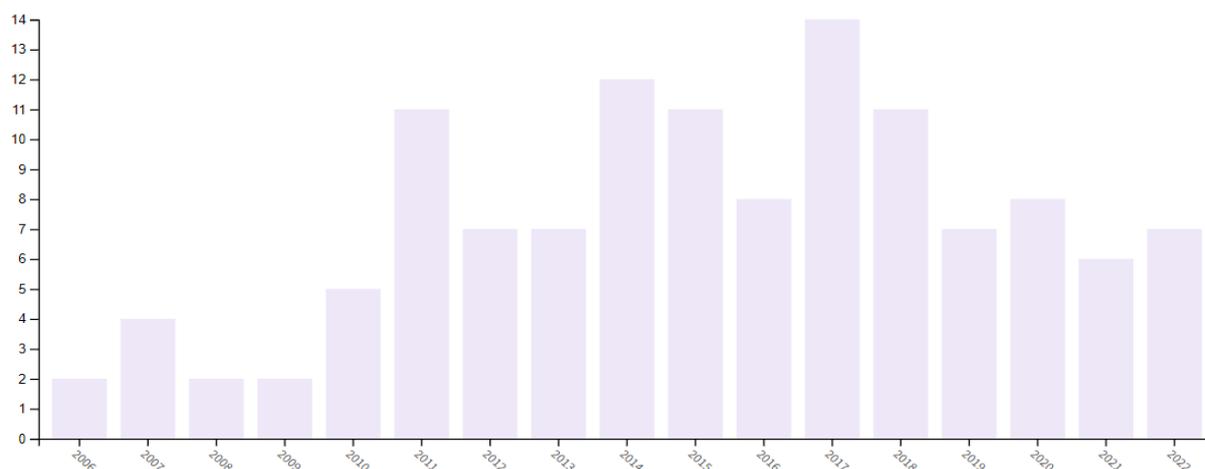
⁵Disponível em: <https://scholar.google.com.br/>

⁶Disponível em: www-periodicos-capes-gov-br

⁷Disponível em: <https://www-webofscience>

publicações que seguem o mesmo padrão com relação ao período de publicação conforme mostra a figura 1.

Figura 1 - Publicações sobre “Learning Object” Web of Science.



Fonte: Web of Science (2023)

Braga (2014) relata que, apesar deste número expressivo de trabalhos sobre OAs, estes estão espalhados em pequenos artigos escritos por grupos de pesquisadores ligados a algumas instituições acadêmicas pelo mundo afora. Torna-se difícil encontrar referências que tratem o tema “objetos de aprendizagem como dados ligados interoperáveis na Web semântica” de forma unificada e completa.

No contexto brasileiro, de acordo com Balan et al. (2010), em 2008 foi criado o Portal do Professor e o Banco Internacional de Objetos Educacionais - BIOE, projetos do Ministério da Educação da época, que tinham por objetivo, promover a inclusão digital dos professores de forma que pudessem explorar os recursos tecnológicos em sala de aula. A preocupação com a inserção das TDICs na formação docente fica evidente quando, em 2009, é instituída a Política Nacional de Formação de Profissionais do Magistério da Educação Básica (Decreto no 6.755, de 29 de janeiro de 2009), logo em seguida, o Plano Nacional de Formação dos Professores da Educação Básica (Portaria Normativa no 9, de 1o de julho de 2009). Acredita-se que estes eventos tenham motivado a realização de mais pesquisas, no que diz respeito às pesquisas brasileiras sobre OAs.

Em relação às pesquisas internacionais, Koutsomitropoulos, Solomou e Kalou (2017), propuseram uma estrutura e um sistema para lidar com a incapacidade de descobrir OAs novos e autênticos e sobre a falta de um único ponto de acesso para pesquisa e navegação de repositórios de objetos de aprendizagem na Web. Os autores desenvolvem

uma estrutura para expansão de consulta baseada em palavras-chave usando vocabulário *SKOS - Simple Knowledge Organization System* e implementaram um mecanismo de pesquisa federada integrando vários OAs diferentes em um sistema de gerenciamento de aprendizado. No artigo, os autores propõem uma estrutura e um serviço para mitigar os problemas de descoberta de OAs dentro de repositórios empregando a ideia de palavras-chave com expansão de consultas para enviar, expandir e refinar consultas de forma automática e transparente para repositórios, com base no conjunto de palavras-chave que descrevem originalmente um curso específico dentro de um ambiente virtual de aprendizagem - AVA. Para isso, o conhecimento terminológico especializado foi oferecido por dois tesouros de domínio, mudando o peso da competência dos pesquisadores para a experiência dos especialistas. Os tesouros foram implementados usando *SKOS* e *OWL - Ontology Web Language*, para poder aplicar operações de raciocínio durante a expansão. As consultas expandidas foram direcionadas para vários repositórios de forma federada. O mecanismo de pesquisa federado replicou as palavras-chave expandidas que foram distribuídas para os mecanismos de pesquisa dos repositórios e agregaram perfeitamente os resultados de volta ao AVA e o contexto do curso. Segundo os autores, os resultados mostraram que esta abordagem pode alcançar ganho de informação e que é possível encontrar um equilíbrio entre eficácia de busca, desvio de consulta e desempenho. Koutsomitropoulos é um autor grego que tem vários trabalhos neste mesmo tema.

Santarém (2018), reforça que boa parte das pesquisas e projetos desenvolvidos nos últimos anos dedicou-se a infraestrutura de organização e recuperação de dados em formato semântico, porém, o autor ressalta que não podemos esquecer que a Web semântica tem um papel social de extrema importância, sendo por meio dela que agentes computacionais podem realizar tarefas para facilitar a vida diária dos seres humanos.

Pouquíssimos estudos tratam os OAs como dados ligados interoperáveis na Web semântica. E se ainda for buscada a intersecção com a publicação destes como dados estruturados em portais colaborativos, na data desta pesquisa, nada foi encontrado. Não que este contexto de publicação de dados seja novidade, pois segundo Santarém (2018),

[...] há muitas teorias e pesquisas, algumas nem tão mais recentes, que tem pautado os estudos baseados em dados como um novo paradigma de pesquisa (quarto paradigma da ciência) e desenvolvimento, e também por isso é notável um grande interesse por vários segmentos de comunidades distintas em também publicar seus dados. (SANTARÉM, 2018, p. 120.)

Porém, o autor reforça que o processo de publicação de dados, que pode parecer simples, tem se tornado problemática para as equipes e pessoas que se propõe a realizá-lo, inclusive, ele inclusive, se propôs a oferecer uma disciplina chamada “Conceitos e tecnologias para publicação de dados abertos e semânticos seguindo as melhores práticas do *Linked Data*” no Programa de Pós-Graduação em Ciência da Informação da Unesp de Marília, que teve grande procura, devido ao grande interesse pela temática. E, ainda segundo o autor, publicar dados em formato aberto na Web semântica tem se tornado um desafio para as organizações e, por isso, propôs um fluxo com fases para publicação de dados em formato aberto na Web semântica seguindo as melhores práticas de dados ligado, inclusive, utilizando o conjunto de princípios para publicação de dados ligados ou Linked Data proposto por Berners-Lee.

2.2 USO DO PADRÃO LOM IEEE/2020 NA REPRESENTAÇÃO DE OBJETOS DE APRENDIZAGEM

De acordo com Alves e Santos (2013), como forma de organizar, promover o acesso e a recuperação de OAs, surgiram os repositórios digitais. Estes podem ser temáticos, quando armazenam recursos específicos de uma área de conhecimento, ou institucionais, quando estão relacionados a uma instituição específica. Possuem como principal objetivo, apoiar cursos presenciais ou à distância. Assim, para que seja possível a indexação e a recuperação deste recursos, os repositórios são estruturados com o auxílio de metadados⁸.

Numa aproximação da temática de pesquisa, a tese de doutorado de Souza (2020) do Programa de Pós-graduação em Ciência da Informação da Universidade da Paraíba, parte da problemática da inexistência de um modelo de representação da informação, no âmbito da Ciência da Informação, que considere a granularidade e a modularidade dos objetos de aprendizagem, de forma a permitir uma melhor recuperação de conteúdos educacionais. A pesquisadora trabalhou na hipótese de que a proposição de um modelo de representação de metadados que envolva a granularidade e a modularidade dos objetos de aprendizagem possibilita uma nova perspectiva de organização, acesso e recuperação desses objetos dentro de um repositório de objetos de aprendizagem, potencializando o reuso desses recursos digitais. Sob a luz dos autores pesquisados pela autora, considera-se, que: (1) A granularidade é a característica relacionada com o fato de arquivos ou ativos de conteúdo, estarem alinhados; (2) A modularidade é a característica

⁸ Esse conceito é abordado com mais detalhes na subseção 3.1.2 do capítulo do Referencial Teórico.

desses ativos serem reutilizados isoladamente, podendo assim se “descolar” do seu arquivo maior; (3) Nem todo ativo de conteúdo descrito para ser recuperado apresenta independência conceitual do objeto maior; (4) Um OA que apresenta granularidade não apresenta, necessariamente, a modularidade.

Souza (2020) utilizou o Acervo de Recursos Educacionais em Saúde, que por sua vez está vinculado ao Ministério da Saúde. Utilizou o método quadripolar, constituído pelos polos epistemológico, teórico, técnico e morfológico e chegou ao delineamento de um modelo operacional de objeto de aprendizagem para a Ciência da Informação que contempla três camadas distintas de granularidade: os grupos de objetos de aprendizagem, os objetos de aprendizagem e os ativos de informação. A modularidade, trabalhada na pesquisa da autora, foi proposta de uma forma distinta, sem contemplar categorias e sim sendo analisada a partir de sua presença ou não no conteúdo do objeto digital. Também foi apresentado um modelo de representação da informação que abarca essas três categorias de granularidade. Como resultado adicional da pesquisa, foi proposto um fluxo de trabalho como forma de visualizar em quais etapas da descrição dos objetos digitais estariam presentes as camadas do modelo de objeto de aprendizagem proposto.

A autora concluiu que o êxito da aplicação do Método Quadripolar (epistemológico, teórico, morfológico e técnico) contempla todos os polos e evidenciou que o fluxo entre eles não é sempre linear e o modelo operacional de objeto de aprendizagem, assim como o modelo de representação da informação e o fluxo de trabalho são produtos do polo morfológico. Desta forma, a autora propôs um modelo de representação de OAs por meio de metadados considerando a granularidade e modularidade - que são aspectos relacionados à diversidade de conteúdos dos objetos e, por isso, incidem diretamente na forma de descrição desses em um repositório - numa perspectiva de organização, acesso e recuperação dentro de um repositório de forma a potencializar o reuso desses objetos. Porém, o universo pesquisado foi os OAs da área da saúde que, e não foi usado o padrão LOM/IEEE, foi proposto outro modelo. A contribuição dessa autora para o presente trabalho é que a mesma trata da importância da representação dos OAs na perspectiva da organização, acesso e recuperação dos mesmos, o que é a essência do que aqui está sendo proposto.

No contexto educacional, de acordo com Ferlin (2009) surgiram várias iniciativas no sentido de desenvolvimento de padrões de metadados com o intuito de facilitar a reutilização e garantir a interoperabilidade de OAs nas diferentes plataformas, sem que haja perda de conteúdos e/ou de funcionalidades. Dentre estes padrões, destacam-se o *IEEE LOM - Learning Object Metadata* ou Metadados para Objetos de Aprendizagem, padrão internacional para objetos de aprendizagem e o OBAA - abreviação de Objetos de Aprendizagem Baseados em Agentes Autônomos, que é uma proposta brasileira baseada no IEEE LOM, com inclusão e expansão de elementos importantes para atender novas plataformas tecnológicas.

É importante ressaltar as pesquisas desenvolvidas a partir do projeto OBAA, dos pesquisadores da UFRGS em parceria com o FINEP, no qual foi criado um padrão de metadados para atender os requisitos de interoperabilidade e uso de recursos digitais no contexto educacional brasileiro. O padrão é compatível com padrões internacionais como o LOM-IEEE. Nesse projeto, destacam-se os trabalhos da equipe da professora Rosa Maria Vicari e de seu orientando André Behr. Esse último encontra-se, no momento da escrita deste trabalho, em Portugal desenvolvendo o projeto OBAA naquele país. Esse projeto gerou uma rica gama de trabalhos, cujo os temas, muito se aproximam da presente proposta, entre eles, Behr et al. (2021a) onde os autores apresentam uma revisão sobre o uso de ontologias em sistemas de repositórios de OAs com o objetivo de busca e sugestão, considerando a sua adoção num projeto chamado *seaThings* que visa a promoção da alfabetização oceânica. Além disso, os autores descrevem o uso do *Cognix* (OAs baseados em Sistema e em Agente) que usa o padrão de metadados OBAA. Sistema que inclui conceitos de inteligência artificial como agentes e ontologias que visam melhorar a busca e assim tornando o sistema mais responsivo. O referido artigo também sugere um modelo de implementação de uma ontologia, usando metadados em repositórios de OAs para fornecer aspectos relevantes, como interoperabilidade, reutilização e busca.

Outro trabalho do mesmo autor, Behr et al. (2021b) explora como os metadados de um OAs representados como dados vinculados, em um novo repositório, podem ser um facilitador para um catálogo mais completo e uma busca com recomendações de conteúdo. A abordagem proposta baseia-se no DBpedia Spotlight para anotação de texto não estruturado para fornecer recomendações na fase de catalogação de OAs e no

GEMET, um tesouro de domínio marinho, para expandir termos de pesquisa marinhos. Cada OA é descrito com metadados OBAA como um conjunto de triplas armazenadas no formato RDF para fornecer interoperabilidade e compatibilidade com dados vinculados. É uma proposta bem próxima ao presente trabalho, com a diferença de que na presente pesquisa serão trabalhados os metadados do padrão LOM-IEEE/2020.

2.3 ONTOLOGIA

Ao fazer uma pesquisa simples de trabalhos sobre o ontologia utilizando o descritor “*ONTOLOGY*” no Google Scholar foram encontrados mais de um milhão de registros sobre o tema e ao ordenar por relevância autores como Barry Smith, Nicola Guarino, Giancarlo Guizzardi aparecem com um número expressivo de citações sobre a temática. Ao pesquisar trabalhos sobre ontologias implementadas utilizando o padrão LOM-IEEE encontrou-se o trabalho de Gluz e Vicari (2012), “*An OWL Ontology for IEEE-LOM and OBAA Metadata*” que apresenta uma ontologia utilizando os metadados OBAA, que é uma ontologia *OWL* criada para representar todos os metadados do padrão IEEE-LOM e da proposta de metadados OBAA. Nessa ontologia os autores fornecem um vocabulário básico de termos linguísticos que os agentes podem usar para consultar e manipular informações dos metadados. O trabalho apresenta a estrutura principal da ontologia, e mostra um exemplo concreto de como representar objetos de aprendizagem utilizando-se da ontologia criada.

Buscou-se também na base do Linked Open Vocabulary⁹ - LOV para verificar a existência de alguma ontologia pronta e disponível utilizando o padrão LOM-IEEE onde foi encontrada apenas uma ontologia *OWL* de 2015, criada por Enayat Rajabi¹⁰ do *Information Engineering Research Unit (IERU) of University of Alcala*, segundo o autor, é uma ontologia e vocabulário usado para expor o IEEE LOM, um padrão de metadados para conteúdos educacionais, como dados ligados para ser uma ponte de ligação de metadados educacionais em Linked Open Data (LOD). Nessa ontologia foi utilizada a versão do LOM-IEEE anterior à última atualização que ocorreu em 2020 e como a forma de organização da ontologia em relação ao metadados foi feita de forma diferente ao que se pretende neste trabalho optou-se por não fazer o aproveitamento do arquivo disponibilizado.

⁹ LOV significa Vocabulários Abertos Vinculados. Este nome é derivado de LOD, que significa Linked Open Data. Disponível em: <https://lov.linkeddata.es/dataset/lov/about>.

¹⁰ Disponível em: http://data.opendiscovery.eu/lom_ontology_ods.owl

3 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo far-se-á uma abordagem teórica conceitual de forma a embasar os procedimentos metodológicos e o desenvolvimento da pesquisa. Visando a um encadeamento das ideias até aqui apresentadas, o capítulo foi organizado em sete seções, conforme apresentado a seguir. A Seção 3.1 trata da organização e representação do conhecimento na ciência da informação, com foco na representação descritiva e temática, apontando como as novas tecnologias têm mudado de forma significativa a representação, armazenamento e recuperação da informação, principalmente no que implica a utilização de Sistemas de Organização do Conhecimento. Na Seção 3.2 o enfoque maior é no processo de Catalogação, por entender que ela é a mediadora em um processo comunicativo e contribui para a organização e uso efetivo da informação por quem dela necessita.

Os princípios FAIR (*Findable, Accessible, Interoperable and Reusable*) são trabalhados na seção 3.3 pois estes têm sido adotados por instituições e considerados instrumentos norteador para a descoberta, acesso e interoperabilidade global, principalmente no que tange a adoção de padrões, metadados, vocabulários controlados, ontologias e identificadores persistentes. O principal objeto de estudo desta pesquisa é trabalhado na seção 3.4, os OAs - Objetos de Aprendizagem, nela trata-se os tipos, as características, seu ciclo de vida, a importância dos repositórios de OAs e como o padrão LOM/2020 da IEEE pode ser utilizado na descrição, busca, avaliação e uso dos OAs pelos estudantes, professores e buscadores automatizados.

Em seguida, na seção 3.5, a Web semântica entra em cena com a estrutura do conteúdo significativo das páginas da Web que ganha interoperabilidade com os dados ligados. A seção 3.6 apresenta as considerações finais do capítulo destacando relações importantes entre vários dos elementos tratados ao longo das seções da fundamentação teórica.

3.1 ORGANIZAÇÃO DO CONHECIMENTO NA CIÊNCIA DA INFORMAÇÃO

Pensando-se na questão conceitual, não é intenção desta pesquisa fazer um levantamento exaustivo sobre os vários conceitos que envolvem a organização do conhecimento na Ciência da Informação, e sim apenas trazer algumas reflexões que irão contribuir diretamente com a pesquisa.

Para Lima e Alvares (2012), organizar o conhecimento significa organizá-lo em assuntos e de forma sistemática. De acordo com Araújo (2018), a partir da década de 1970, as teorias cognitivistas na CI deram origem aos esforços de representação da informação com foco no usuário. Iniciativas que provaram que a representação não se limitava apenas ao documento, mas a quem iria utilizá-lo. Ainda segundo o autor, a teoria do conceito proposto por Dahlberg (1978) contribuiu também para trazer reflexões mais filosóficas à lógica de representação dos sistemas de informação, assim como ao desenvolvimento de tesouros e outros instrumentos de linguagem documentária.

Foi no final da década de 1980, com a criação da ISKO (Information Society for Knowledge Organization), que congregou pesquisadores dedicados a essa temática e passou a organizar eventos e publicações e início da década de 1990, que as inovações tecnológicas impulsionaram o surgimento de um processo mais colaborativo na representação da informação, a partir daí, deu-se também a denominação da área como “organização do conhecimento”. A informação passa a ser estudada como algo cognitivo, semântico ou subjetivo, e começa-se a considerar a articulação entre dado, informação e conhecimento onde,

[...] dados (os elementos presentes da realidade independente dos sujeitos) e o conhecimento (aquilo que os indivíduos sabem ou conhecem), sendo a informação a medida da alteração deste estado de conhecimento, ou, em outros termos, o produto da interação entre os dados e o conhecimento, no âmbito do indivíduo. (ARAÚJO, 2018, p.42)

Ainda de acordo com Araújo (2018), a partir do século XX, novas teorias surgem apontando para um modelo de conceito de informação mais pragmático, intersubjetivo e sociocultural onde a informação é vista como algo não apenas do objetivo ou do subjetivo mas também do coletivo, fruto de uma construção social. Estes aspectos e o incremento das tecnologias digitais fizeram com que as pesquisas em representação em termos de, descrição, classificação e organização da informação tivessem grande avanço nos últimos anos, destacando-se o surgimento de novos padrões de codificação, vocabulários controlados e ontologias (CURRÁS, 2010; USCHOLD, 1996; SCHIESSL; BRASCHER, 2012), estas últimas, se desenvolveram como novo modelo de representação da informação apresentando importantes especificidades em relação a modelos anteriores, como os tesouros e as taxonomias (GUARINO, 1998; CAMPOS; GOMES, 2017). Ainda relacionado às formas de representação da informação a partir das TDICs destacam-se os estudos em Web semântica, dados abertos (SANTARÉM SEGUNDO; CONEGLIAN;

LUCAS, 2016; JORENTE; PÁDUA; SANTARÉM SEGUNDO, 2017; SANTARÉM SEGUNDO, 2015) e o uso da perspectiva da classificação facetada.

Araújo (2018), ressalta ainda que fenômenos colaborativos como as ferramentas wiki e etiquetagem colaborativa colocam os processos de representação, na esteira do contexto da Web 2.0¹¹. Assim, os estudos na perspectiva da classificação facetada e as tecnologias digitais, passam a ter forte ligação com aplicações em metadados, modelagem de dados e sistemas de navegação e busca, entre outros. Neste contexto, desenvolve-se um novo conceito para a organização de recursos digitais na Web, as chamadas folksonomias, que etimologicamente em inglês significa folksonomy=folk(povo) +sonomy(taxonomia), cuja a cunhagem do termo é atribuída a Thomas Vander Wal, numa lista de distribuição em seu blog.

Folksonomia é o resultado da marcação pessoal gratuita de informações e objetos (qualquer coisa com uma URL) para a própria recuperação. A marcação é feita em um ambiente social (geralmente compartilhado e aberto a outras pessoas). A folksonomia é criada a partir do ato de marcação pela pessoa que consome a informação. (WAL, 2007, sem página)

Para Araújo (2018),

Trata-se da consideração da indexação livre, realizada pelos próprios usuários, no trabalho dos profissionais da informação, com o objetivo de proporcionar melhor recuperação da informação - trabalho este desenvolvido em ambiente aberto e de compartilhamento, portanto de construção conjunta. (ARAÚJO, 2018, p. 51)

Araújo (2018) traz uma linha de pesquisa inaugurada por Hjørland e Albrechtsen (1995) com outra tendência conceitual que é a construção de sistemas de classificação partindo da perspectiva sociocultural, os chamados estudos em “análise de domínio”, e aponta dois tipos:

“a descritiva (utilizada pelos pesquisadores para desenvolverem suas pesquisas em seus respectivos campos) e a instrumental (utilizada pelos desenvolvedores de sistemas de informação e sistemas de recuperação da informação), sendo que os domínios se caracterizam por ter determinado escopo e alcance, além de um contexto próprio.” ARAÚJO (2018, p. 53)

É importante que o profissional da informação desenvolva serviços e sistemas de informação que estejam adaptados a esses padrões coletivos, e o tratamento da informação para que atenda a estes padrões, precisa ser observado.

¹¹ O conceito de "Web 2.0" começou com uma sessão de brainstorming de conferência entre O'Reilly e Media Live International. Disponível em: <http://wellman.uni-trier.de/images/9/96/Web20.pdf>.

Alguns autores fazem distinção entre tratamento descritivo e temático da informação, para Dias (2013), o tratamento descritivo trata os aspectos mais objetivos para identificar, inequivocamente, um documento como o autor, o título, a editora e elementos similares. Já o tratamento temático, tem um forte viés subjetivo, pois caracteriza o documento do ponto de vista do seu conteúdo.

Para Ortega (2008), a representação temática atribuiu assuntos aos documentos por meio da classificação bibliográfica, indexação e elaboração de resumos. De acordo com Fujita (2003), sob o ponto de vista dos sistemas de informação, a indexação é reconhecida como uma parte importante do tratamento pois condiciona os resultados de uma estratégia de busca, ou seja, o bom ou mau desempenho da indexação impacta diretamente na recuperação da informação. De acordo com Lancaster (2004), os primeiros estudos sobre recuperação da informação buscaram medir o nível de eficácia dos diferentes sistemas, utilizando, para isso, os conceitos de revocação e precisão, sendo o primeiro termo relacionado à capacidade de o sistema recuperar documentos úteis e o segundo, com a capacidade de o sistema evitar documentos inúteis, descartando-os. A indexação diz respeito ao conteúdo do documento por meio da análise de assunto e da representação deste assunto utilizando certos conceitos. Esses conceitos são traduzidos em linguagem documentária no processo de intermediação entre o documento e o usuário no momento de recuperação da informação, utilizando-se catálogos, índices e bases de dados.

De acordo com Alvarenga (2013), desde o famoso *Traité de Documentation* de Paul Otlet, publicado em 1934, que definiu o documento como objeto de estudo, obra que representa a maturidade do seu pensamento sobre a organização e o acesso ao conhecimento, passando por Diderot, D'Alembert e Henry LaFontaine, estes pensadores e documentalistas sempre sonharam com a existência de coleções e catálogos que pudessem abarcar a totalidade dos conhecimentos de todos os tempos e espaços. Com o advento das novas tecnologias de informação que têm evoluído significativamente nas últimas décadas, têm surgido novos recursos que podem finalmente concretizar o sonhos daqueles documentalistas.

3.1.1 Sistemas de Organização do Conhecimento

Na perspectiva da Ciência da Informação, percebe-se que há significativas mudanças no que tange à representação, armazenamento e recuperação da informação, principalmente em Bibliotecas digitais, Repositórios, entre outros. E no processo de representação da informação, neste contexto da novas tecnologias, surgem os chamados SOC (Sistema de Organização do conhecimento)¹², termo que de acordo com Mazzocchi (2018), foi cunhado no primeiro Grupo de Trabalho de Sistemas de Organização de Conhecimento em Rede que ocorreu na Conferência *ACM Digital Libraries '98* em *Pittsburgh*. Mas já era utilizado em 2008 pelo grupo de trabalho criado pela *W3C*¹³, - o *SKOS (Simple Knowledge Organization System)*, que desenvolve especificações e padrões para apoiar o uso de sistemas de organização do conhecimento (KOS), como tesouros, esquemas de classificação, sistemas de cabeçalhos de assunto e taxonomias no âmbito da Web semântica. De acordo com Calan e Medeiros (2011), SOC são:

[...] sistemas conceituais semanticamente estruturados que contemplam termos, definições, relacionamentos e propriedades dos conceitos. Na organização e recuperação da informação, os SOC cumprem o papel de padronização terminológica para facilitar e orientar a indexação e os usuários. Quanto à estrutura variam de um esquema simples até o multidimensional, enquanto que suas funções incluem a eliminação da ambiguidade, controle de sinônimos ou equivalentes e estabelecimento de relacionamentos semânticos entre conceitos. (CALAN; MEDEIROS, 2011, p. 54)

Os SOC são estruturados de forma a criar uma padronização para evitar ambiguidade e recuperação equivocada de informação, por isso, precisam ser pensados como uma “ponte” entre a informação e quem dela necessita, conforme mostra Hodge 2000,

O termo sistemas de organização do conhecimento destina-se a abranger todos os tipos de esquemas para organizar a informação e promover a gestão do conhecimento [...] Os sistemas de organização do conhecimento são usados para organizar materiais para fins de recuperação e para gerenciar uma coleção. Um KOS serve como uma ponte entre a necessidade de informação do usuário e o material da coleção. Com ele, o usuário deve ser capaz de identificar um objeto de interesse sem conhecimento prévio de sua existência. Seja por meio de navegação ou pesquisa direta, seja por meio de temas em uma página da Web ou em um mecanismo de pesquisa de site, o KOS guia o usuário por meio de um processo de descoberta. Além disso, os KOSs permitem que os organizadores

¹² Knowledge Organization System” (KOS), proposto pelo Networked Knowledge Organization Systems Working Group na 1ª Conferência da ACM Digital Libraries em 1998, Pittsburgh, Pennsylvania. (MAZZOCCHI, 2018)

¹³ O World Wide Web Consortium (W3C) é uma comunidade internacional onde as organizações membros, uma equipe em tempo integral e o público trabalham juntos para desenvolver padrões da Web. Liderada pelo inventor e diretor da Web Tim Berners-Lee e pelo CEO Jeffrey Jaffe, a missão do W3C é levar a Web a todo o seu potencial. Disponível em: <https://www.w3c.br/Sobre/>

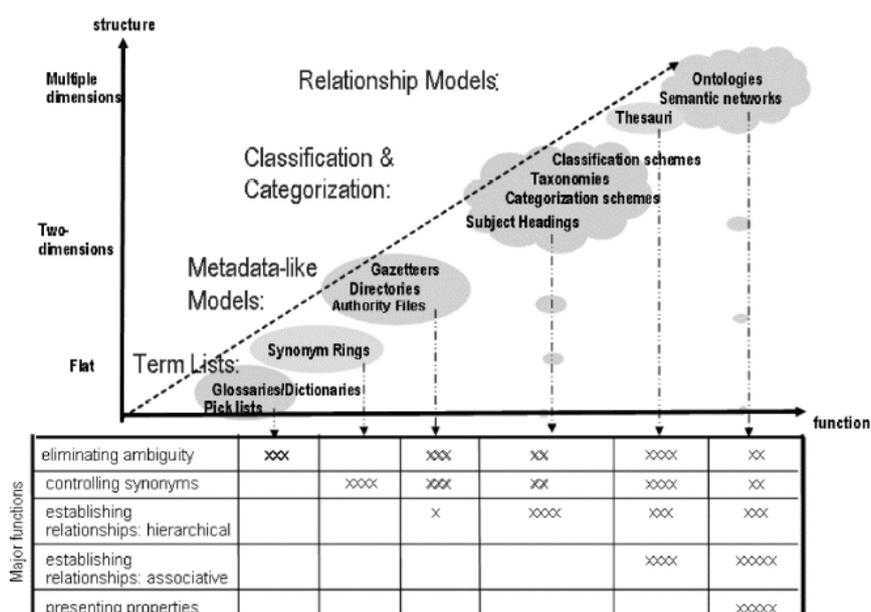
respondam a perguntas sobre o escopo de uma coleção e o que é necessário para completá-la. (HODGE, 2000, 3 tradução nossa)

Para isso, os SOC utilizam da linguagem documentária de uma determinada área de domínio, utilizando-se de conceitos semânticos para “rotular” a informação e facilitar a descoberta e a recuperação, conforme discorre Zeng (2008),

Esses sistemas modelam a estrutura semântica subjacente de um domínio e fornecem semântica, navegação e tradução por meio de rótulos, definições, digitação, relacionamentos e propriedades para conceitos [...] Incorporados como serviços (Web), eles facilitam a descoberta e recuperação de recursos por atuando como roteiros semânticos, possibilitando assim uma orientação comum para indexadores e futuros usuários, sejam humanos ou máquinas [...]. (ZENG, 2008 p.160 tradução nossa)

Quanto ao tipo, Zeng (2008) propôs uma estrutura para os classificar os SOC de acordo com quatro grandes grupos, do mais simples ao mais complexo, sendo: listas de termos, modelos semelhantes a metadados, classificação e categorização e modelos de relacionamento conforme mostra a Figura 2.

Figura 2 - Uma visão geral das estruturas e funções do SOC



Fonte: Zeng (2008, p. 161)

No modelo, Zeng utiliza as três grande categorias propostas por Hodge (2000), que são:

- listas de termos, que enfatizam listas de termos frequentemente com definições;
- classificações e categorias, que enfatizam a criação de conjuntos de assuntos;

- listas de relacionamentos, que enfatizam as conexões entre termos e conceitos. (HODGE, 2000, 1 tradução nossa)

Neste modelo, é possível observar que os Tesouros, as Taxonomias e as Ontologias são considerados pela a autora, os mais complexos, tendo em vista, as funções que desempenham no processo de organização e recuperação da informação, pois estabelecem relações hierárquicas, controlam sinônimos e eliminam ambiguidades.

De acordo com Fujita e Tolare (2019), apesar de a linguagem natural ser mais democrática e acessível e seu uso na indexação social, indiscutivelmente associado às redes sociais, blogs e sites diversos ser pertinente e necessário, em sistemas de recuperação da informação de universidades, por exemplo, a linguagem científica é inequivocamente associada ao vocabulário científico e tecnológico e a terminologia a ser utilizada como assunto em repositórios com uso de palavras-chave é mais especializada e mais precisa mesmo que seja da linguagem natural, segundo as autoras, a falta de controle de vocabulário afeta também a visibilidade das publicações arquivadas nos repositórios como a interação do sistema com o usuário final. Desta forma, o tesauro, como vocabulário controlado, passa a ter função importante no processo de mediação entre os termos utilizados no discurso e aqueles que funcionam efetivamente para a recuperação da informação, para isso, o usuário e o profissional da informação precisam aceitar um grau de artificialidade no vocabulário controlado para que sejam alcançados os benefícios no processo de recuperação.

No trabalho de Harper e Tillet (2009), os autores argumentam como os vocabulários controlados, esquemas de classificação e tesouros podem contribuir na construção da Web semântica compartilhando o controle de autoridade. Segundo os autores, esta contribuição partiria do uso efetivo das hierarquias, relações de equivalência e estruturas do conteúdo de vocabulários controlados contidos no registro de autoridade MARC¹⁴ (Machine Readable Cataloging) e que seu uso auxilia a busca na consulta de outros termos desconhecidos que fazem parte da estrutura hierárquica de um determinado termo ou que, também, seja equivalente ou associado.

Porém, a utilização de vocabulários em repositórios não é tarefa fácil e é preciso lançar mão de outros SOCs como as Ontologias e Taxonomias, que possuem maior intersecção

¹⁴ MARC é o nome de um sistema desenvolvido pela Library of Congress em parceria com a British Library. A sigla significa Machine Readable Cataloging, ou "catalogação legível para computadores"

com a presente pesquisa e, portanto, serão tratados com maior rigor em subseções desta seção.

3.1.2 Metadados

Metadado é um termo que pode ter várias definições em diferentes áreas do conhecimento. Pöttker, Ferneda e Moreiro-González (2018, p.29) postulam que na “Ciência da Informação e Ciência da Computação o metadado é considerado como uma descrição de dados que podem ser utilizados para identificar e descrever características comuns entre diferentes recursos informacionais com a finalidade de facilitar a interoperabilidade e a recuperação na Web”, ou de maneira simples, “dados sobre dados”.

Já o conceito proposto por Grácio (2012, p. 121) contempla as principais características e metodologias que garantem a preservação digital que é primordial quando se pensa em recuperação da informação, segundo o autor os metadados são “[...] um conjunto de dados, chamados ‘elementos’, cujo número varia de acordo com o padrão adotado, e que descreve o recurso, possibilitando a um usuário ou a um mecanismo de busca acessar e recuperar esse recurso”.

Barbedo, Corujo e Sant’Ana (2011) postulam algumas categorias funcionais de tipos de metadados, que são:

Metadados descritivos ou de identificação – O objetivo é a pesquisa, recuperação e identificação. Pode incluir elementos como o título, o assunto, autor e palavras-chave.

Metadados estruturais – Relaciona hierarquicamente os diferentes OD que fazem parte de um mesmo documento. O esquema METS (Metadata Encoding & Transmission Standard) permite agrupar metainformação descritiva, administrativa e estrutural sobre objetos guardados num repositório digital.

Metadados administrativos – O objetivo é a pesquisa, recuperação e identificação. Pode incluir elementos como o título, o assunto, autor e palavras-chave. O esquema EAD (Encoded Archival Description) é um bom exemplo, por ser um esquema muito completo e que respeita as normas de descrição arquivística;

Metadados técnicos – Descreve as características técnicas dos ficheiros e dos seus formatos, pelo que é a mais específica, por depender do formato do OD. Um exemplo deste tipo de MI é o esquema baseado na norma ANSI/NISO Z39.87 - Metadata for Images in XML schema, define um conjunto normalizado de elementos de metainformação para imagens digitais;

Metadados de preservação – Contém informação necessária para arquivar e preservar o OD. O esquema PREMIS (PREservation Metadata Implementation Strategies), produto da parceria entre a OCLC (Online Computer Library Center) e

a RLG (Research Libraries Group) é o mais conhecido. (BARBEDO, CORUJO E SANT'ANA, 2011, p.48)

Quanto à organização dos metadados, esta se dá por meio de estruturas formais chamadas de esquemas de metadados, que são conjuntos de elementos concebidos com objetivo específico. Encontram-se na literatura outras expressões como "formato", "sistema" e "conjuntos de elementos" para se referir a padrões de metadados (CHAN; ZENG, 2006). De acordo com Zeng e Qin (2008) um esquema de metadados é

Uma especificação possível de ser processada por máquinas que define a estrutura, a codificação de sintaxe, regras, e formatos para um conjunto de elementos de metadados em uma linguagem formal num esquema. Na literatura o termo "metadata schema" usualmente refere-se ao conjunto de elementos na sua totalidade, assim como a codificação dos elementos e a estrutura com uma linguagem de marcação. (ZENG; QIN, 2008, p.323, tradução nossa)

Assim, é possível entender "esquema" como uma entidade que contém componentes semânticos e de conteúdo "conjunto de elementos" e a codificação dos elementos com uma sintaxe ou linguagem de marcação, como Standard Generalized Markup Language (SGML) e Extensible Markup Language (XML). Desta forma é possível inferir que um conjunto de elementos de metadados é composto de dois componentes básicos: (1) Semântica – conceitos ou os significados dos elementos e seus refinamentos; e (2) Conteúdo – as instruções de quais valores e como devem ser atribuídos os elementos de dados. Um padrão de metadados geralmente especifica regras de conteúdo para cada elemento e define como o conteúdo deve ser criado ou incluído (como identificar o título principal), apresenta as regras de representação para o conteúdo (padrões de representação do tempo, por exemplo) e quais valores de conteúdo são admissíveis, ou seja, se os termos devem partir de um vocabulário controlado específico ou podem ser providos pelo autor, derivados do texto, ou adotados pelo trabalho de criadores de metadados sem uma lista de termos controlados. Ainda podem haver regras de sintaxe para codificação dos elementos e seu conteúdo (CHAN; ZENG, 2006; NATIONAL INFORMATION STANDARDS ORGANIZATION, 2004).

É possível identificar diversos padrões ou esquemas de metadados utilizados para a descrição de recursos em diferentes domínios, onde vários padrões de metadados vêm sendo construídos. Entretanto, os próprios conjuntos de elementos de metadados delimitam os esquemas, as particularidades e domínios de utilização.

3.1.3 Taxonomias

A palavra taxonomia deriva da biologia e trata da classificação científica dos seres vivos, sendo um trabalho do médico e botânico sueco Carolus Linnaeus. A etimologia do termo vem do grego onde taxis=ordem e onoma=nome, o que nos remete a ideia de ordenação, rotulagem, classificação e de sistema. Porém, mesmo tendo sido empregada primeiro na biologia, nos ambientes digitais, a palavra está relacionada com as formas automatizadas de organização da informação (VITAL; CAFÉ, 2011). Campos e Gomes (2008, p.3), dizem que taxonomia, “é, por definição, classificação, sistemática” e que se caracterizam por:

- Conter uma lista estruturada de conceitos/termos de um domínio.
- Incluir termos sem definição, somente com relações hierárquicas.
- Possibilitar a organização e recuperação de informação através de navegação.
- Permitir agregação de dados, diferentemente das taxonomias seminais, além de evidenciar um modelo conceitual do domínio.
- Ser um instrumento de organização intelectual, atuando como um mapa conceitual dos tópicos explorados em um Sistema de Recuperação de Informação.
- Ser um novo mecanismo de consulta em Portais institucionais, através de navegação. (CAMPOS; GOMES, 2008, p. 3)

De acordo com Martinez et al. (2004, p. 106) “a taxonomia, em um sentido amplo, é a criação da estrutura (ordem) e dos rótulos (nomes) que ajudam a localizar a informação relevante. Em um sentido mais específico, é o ordenamento e rotulação de metadados, que permite organizar sistematicamente a informação primária”.

Campos e Gomes (2008, p.3) colocam que a taxonomia possibilita, independentemente da área de conhecimento ser disciplinar ou multidisciplinar, uma organização de representação por classes de conceitos, por exemplo, coisas e seus tipos, processos e seus tipos etc, de forma a não priorizar nenhum deles, que se baseia no princípio de categorias conceituais representando os diversos aspectos (fenômenos, objetos etc) que ocorrem naquele campo de saber. Campos e Gomes (2008) ainda propuseram alguns princípios básicos de classificação adotados nas taxonomias, que são:

- Categorização, que fornece as bases para a apresentação sistemática;
- Cânones, para o trabalho no plano das ideias (princípios para a construção das classes);
- Princípios, para a ordenação das Classes e de seus elementos (CAMPOS; GOMES, 2008, p.5).

De acordo com a experiência vivenciada pelas autoras, elaborar taxonomias é uma atividade que requer da pessoa que pretende classificar a informação, um planejamento que deve ser iniciado com o dimensionamento das classes até as etapas de construção e alguns princípios são fundamentais para garantir a consistência da estrutura classificatória e conseguir boa recuperabilidade.

Então, como visto, as taxonomias podem ser utilizadas para a organização da informação em ambientes específicos, com vistas à recuperação eficaz e, para isso, estabelece parâmetros em todo o ciclo de produção da informação, o conceito, muitas vezes, é limitado a termos organizados em categorias hierárquicas, porém, quando trata de taxonomias em ambientes virtuais de aprendizagem deve-se levar em consideração os objetivos educacionais envolvidos.

Em 1948, com base no princípio e na importância de se utilizar o conceito de classificação como forma de se estruturar e organizar um processo, a Associação Norte Americana de Psicologia (American Psychological Association), solicitou a alguns de seus membros que montassem uma “força tarefa” para discutir, definir e criar uma taxonomia dos objetivos de processos educacionais, onde embora outros pesquisadores tenham colaborado significativamente no desenvolvimento dessa taxonomia, ela é conhecida como “Taxonomia de Bloom” (LOMENA, 2006). Segundo Bloom et al. (1977), vários pesquisadores utilizaram-se dessa terminologia conceitual baseada em classificações estruturadas e orientadas para definir algumas teorias instrucionais baseada em três domínios a saber: Cognitivo, Afetivo e Psicomotor. Voltando para Campos e Gomes (2008), é possível entender que uma taxonomia de domínio é voltada para a representação de um domínio do conhecimento, sendo policotômica, ou seja, onde a partir de um domínio ou tarefa, várias divisões são possíveis, seu recorte é feito por categorias e determinadas pelas características da organização ou área do conhecimento em que estão inseridos, possibilitando subdivisões. E se falando de OAs são estes domínios, principalmente o cognitivo, que precisam ser levados em consideração ao pensar uma estrutura para sua organização.

3.1.4 Ontologias

Foi na filosofia que surgiram as primeiras definições de ontologia, que se originaram na metafísica para designar “aquilo que existe” e etimologicamente deriva do grego, onde temos onto(ser) + logia (estudo), ou seja, ciência que estuda o ser. (SMITH, 2003). Porém, o sentido muda quando falamos de organização e representação da informação, onde a ontologia se propõe a categorizar e classificar as coisas sob a ótica do sujeito e da linguagem dominante.

Gruber (1996, p.1) utilizou a seguinte expressão para designar o termo: *“I use the term ontology to mean a specification of a conceptualization”*, numa tradução simples, “é uma especificação de uma conceituação”, para o autor, essa noção de conceituação é um grupo de relações extensionais que descrevem um “estado das coisas” de forma particular, sendo que normalmente o que se tem em mente é uma relação intencional, algo como uma rede conceitual onde se superpõe a vários possíveis “estados das coisas”. Neste aspecto de “relação intencional” esta rede conceitual cria relações de significação e de redes semânticas constituídas por termos que ali se encontram.

Caminhando nesta relação intencional, Gruber (1993, p.199, tradução nossa), define ontologia como “uma especificação formal e explícita de uma conceituação compartilhada”. De acordo com o autor essa “conceituação” refere-se a um modelo abstrato de algum fenômeno no mundo e identifica os conceitos relevantes deste fenômeno e “explícito” significa que tipo de conceitos são utilizados, e quais restrições sobre seu uso são explicitamente definidas. Studer, Benjamins e Fensel (1998), explicam que , no domínio da medicina, os conceitos são doenças e seus sintomas, e as relações entre eles são causais, neste caso, uma restrição seria que a uma doença não pode causar sintomas nela mesma. “Formal” refere-se ao fato de que a ontologia deve ser legível por máquina, o que exclui a linguagem natural. ‘Compartilhado’ reflete a noção de que uma ontologia capta o conhecimento consensual, ou seja, não é exclusivo de algum indivíduo, mas aceito por um grupo.

Assim, chegamos à ideia defendida por Almeida e Bax (2003, p. 17), de que as ontologias “[...] permitem formas de representação baseadas em lógica, o que possibilita o uso de mecanismos de inferência para criar novo conhecimento a partir do existente”. Então, seguindo esta premissa, pode-se inferir que as ontologias possibilitam a formação de relações entre vários conceitos e podem propiciar a representação de relações com um mesmo objeto. Essa representação é feita utilizando-se dos seguintes elementos:

- Indivíduos: são os objetos em sua unidade básica;
- Classes: são conjuntos, coleções ou os tipos de indivíduos agrupados;
- Atributos: são as propriedades dos indivíduos;

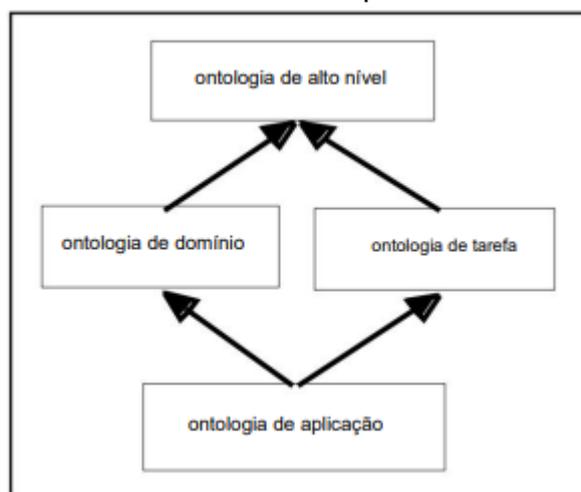
- **Relacionamentos:** é a forma como os indivíduos se relacionam.

Guarino (1998), faz uma distinção entre “Ontologia” (com “o” maiúsculo) e “ontologia” (com “o” minúsculo), segundo o autor o mesmo termo tem uma leitura incontável no primeiro caso e uma leitura contável no segundo. Enquanto a primeira leitura parece ser razoavelmente clara (como referindo-se a uma disciplina filosófica particular), dois sentidos diferentes são assumidos pela comunidade filosófica e pela comunidade de Inteligência Artificial (e, em geral, toda a comunidade da ciência da computação) para o último termo. O autor conceitua ontologia como,

Uma ontologia é uma teoria lógica responsável pelo significado pretendido de um vocabulário formal, ou seja, seu compromisso ontológico com uma conceituação particular do mundo. Os modelos pretendidos de uma linguagem lógica usando tal vocabulário são limitados por seu compromisso ontológico. Uma ontologia reflete indiretamente esse compromisso (e a conceituação subjacente) ao aproximar esses modos pretendidos. (GUARINO, 1998, p.7)

Guarino ainda sugere o desenvolvimento de diferentes tipos de ontologia de acordo com seu nível de generalidade, conforme mostra a figura 3.

Figura 3 - Tipos de ontologias, de acordo com seu grau de dependência de uma determinada tarefa ou ponto de vista



Fonte: Guarino (1998, p. 9)

De acordo com Guarino (1998),

Ontologias de alto nível: descrevem conceitos muito gerais como espaço, tempo, matéria, objeto, evento, ação, etc., que são independentes de um problema ou domínio

particular: parece, portanto, razoável, pelo menos em teoria, ter unificado ontologias para grandes comunidades de usuários.

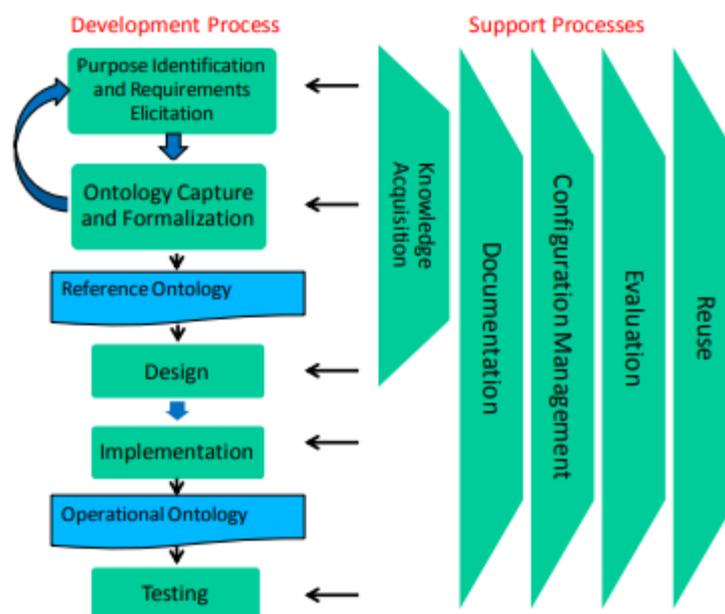
Ontologias de domínio e ontologias de tarefas: descrevem, respectivamente, o vocabulário relacionado a um domínio genérico (como medicina ou automóveis) ou uma tarefa ou atividade genérica (como diagnosticar ou vender), especializando os termos introduzidos na ontologia de nível superior.

Ontologias de aplicação: As ontologias de aplicação descrevem conceitos que dependem tanto de um domínio quanto de uma tarefa em particular, que muitas vezes são especializações de ambas as ontologias relacionadas. Esses conceitos geralmente correspondem a papéis desempenhados por entidades de domínio durante a execução de uma determinada atividade, como unidade substituível ou componente sobressalente.

Noy e McGuinness (2002) defendem que uma ontologia deve definir um vocabulário comum para pesquisadores que precisam compartilhar informações em um domínio. De acordo com os autores esta deve incluir definições interpretáveis por máquina de conceitos básicos no domínio e as relações entre eles.

Segundo Falbo (2014), uma vez que os usuários já concordaram com uma conceitualização comum, as ontologias operacionais (ontologias legíveis por máquina) podem ser implementadas. Ao contrário das ontologias de referência de domínio, as ontologias operacionais são projetadas com o foco em garantir propriedades computacionais desejáveis. A ideia por trás da análise ontológica é fornecer uma base sólida para modelar conceitos, caso assumamos que tais conceitos visam representar a realidade, entre os benefícios desta análise, o autor cita: "(i) a definição rigorosa de modelos, em termos de semântica do mundo real; (ii) a identificação de problemas na definição, interpretação ou uso de conceitos; e (iii) recomendações para melhorias na formalidade do modelo" (FALBO, 2014, p. 1). Desta forma, ao incorporar as experiências do autor no processo de software utilizando ontologias, assim como a importante distinção entre ontologias de referência e ontologias operacionais, o autor desenvolveu o SABiO, uma metodologia de criação de ontologia onde se incorporou as melhores práticas comumente adotadas em Engenharia de Software e Engenharia de Ontologias.

Figura 4 - Processo de desenvolvimento de ontologia SABiO



Fonte: Falbo (2014, p. 3)

O processo proposto pelo autor, conforme mostra a figura 4, compreende cinco fases principais, onde cada fase é composta por atividades, que devem ser executadas de acordo com determinadas funções, são elas:

- (i) especialista de domínio, que é um especialista na ontologia do negócio, e fornece o conhecimento que deve ser modelado e implementado na ontologia;
- (ii) usuário da ontologia, representando alguém que pretende utilizar a ontologia para um determinado fim;
- (iii) engenheiro de ontologia, responsável pela ontologia de referência e, portanto, responsável pelas fases iniciais do processo de desenvolvimento da ontologia;
- (iv) designer de ontologias, responsável pelo design de uma ontologia operacional;
- (v) programador de ontologias, responsável pela implementação de uma ontologia operacional;
- (vi) testador de ontologias, responsável por testar uma ontologia operacional. (FALBO, 2014, p.3, tradução nossa)

Ainda, segundo o autor, um mesmo profissional pode desempenhar diversas funções em um determinado projeto, onde geralmente as funções de designer de ontologias e programador de ontologias são desempenhadas pelo mesmo trabalhador, pois ambas as funções exigem grande conhecimento sobre o ambiente de implementação da ontologia.

Na metodologia SABiO as ontologias de domínio devem ser desenvolvidas com base numa ontologia de fundamentação, assim é possível organizar os níveis de ontologias desde o mais abstrato até o mais simples: Ontologia de fundamentação, Ontologia de referência de domínio e Ontologia operacional de domínio, onde de acordo com a metodologia, a ontologia operacional é oriunda de uma etapa de construção de um

modelo de design, seguida de uma implementação conforme mostra a figura 4. Assim, a SABiO sugere que conceitos e relações em uma ontologia de domínio de referência sejam analisados à luz de uma ontologia fundacional. Onde os conceitos e relações relevantes devem ser identificados e organizados. Neste processo, um modelo gráfico é considerado pelo autor, instrumento chave para apoiar a comunicação, significando negociação e estabelecimento de consenso com especialistas do domínio utilizando-se de linguagens altamente expressivas para criar ontologias fortemente axiomatizadas que se aproximem ao máximo da ontologia ideal do domínio. Segundo o autor, o foco nessas linguagens deve estar na adequação da representação, uma vez que as especificações resultantes são destinadas ao uso humano. Um exemplo de uma linguagem de representação ontológica que é adequada para ontologias de referência é OntoUML¹⁵, ela se expressa por meio de diagramas compostos por elementos gráficos que representam as entidades e as suas relações, é um perfil de diagrama de classe UML que incorpora importantes distinções fundamentais feitas pela *Unified Foundational Ontology (UFO)*¹⁶ representação ontológ

3.2 CATALOGAÇÃO

Essa seção aborda o processo de catalogação, processo que tem sofrido a influência das novas tecnologias, principalmente, no que tange a utilização de padrões para representação da informação de forma a viabilizar a recuperação e o acesso. Aborda também como estes padrões podem contribuir na organização e representação de OAs

3.2.1 Fundamentos

Na seção 3.1 fez-se um resgate histórico sobre a Organização do Conhecimento na Ciência da informação e os Sistemas que envolvem esta organização. Porém, como aqui trataremos da catalogação, é importante fazer um pequeno resgate conceitual do objeto da Ciência da Informação: a informação. De acordo com Le Coadic (2004, p. 4), “A informação é um conhecimento¹⁷ inscrito (registrado) em forma escrita (impressa ou

¹⁵ OntoUML é uma linguagem gráfica de modelagem orientada a ontologias baseada na UFO e na UML 2.0 (Unified Modeling Language). Disponível em: <https://ontouml.org/>.

¹⁶ Unified Foundational Ontology (UFO) tem sido o foco de um programa de pesquisa de longo prazo sobre fundamentos ontológicos para modelagem conceitual. O UFO (Guizzardi, 2005; Guizzardi et al., 2015b) foi desenvolvido reunindo consistentemente teorias da ontologia formal em filosofia, ciência cognitiva, linguística e lógica filosófica. (GUIZZARDI, 2021.p. 1)

¹⁷ Um conhecimento (um saber) é o resultado do ato de conhecer, ato pelo qual o espírito apreende um objeto. Conhecer é ser capaz de formar a ideia de alguma coisa; é tê-la presente no espírito. Isso pode ir da simples identificação (conhecimento comum) à compreensão exata e completa dos objetos (conhecimento científico). O saber designa um conjunto articulado e organizado de conhecimento a partir do qual uma

digital), oral ou audiovisual, em um suporte". Ela é vista como algo que pode gerar valor agregado e gerar conhecimentos inovadores, além de contribuir para o desenvolvimento da sociedade. Porém, para que isso aconteça, a informação precisa estar registrada, representada com vistas ao seu acesso e uso.

É por meio da catalogação, que informações como resultado de um conhecimento registrado, conforme defendido por Le Coadic (2004), podem ser recuperadas e acessadas. Desta forma, a catalogação passa a ser mediadora em um processo comunicativo, e contribui para a organização e uso efetivo da informação por quem dela necessita. Este processo vem sendo muito beneficiado pelos avanços tecnológicos, principalmente, da Internet, Web semântica e novas formas de recuperação de conteúdos digitais.

É nas tecnologias que a catalogação encontra importante suporte, de forma a criar um canal de comunicação entre o usuário e a informação. Este processo abrange todas as atividades do ciclo de representação da informação, desde o planejamento da estrutura do sistema, passando pela infraestrutura e definição de metadados, até o resultado final, que é a busca e recuperação de informação. Segundo Boar (2002).

A tecnologia da informação é a preparação, coleta, transporte, recuperação, armazenamento, acesso, apresentação e transformação de informações em todas as suas formas (voz, gráficos, texto, vídeo e imagem). A movimentação de informação pode ocorrer entre seres humanos, entre humanos e máquina e/ou entre máquinas. O gerenciamento da informação garante seleção, distribuição, administração, operação, manutenção e evolução dos bens de tecnologia da informação de forma coerente com as metas e objetivos da organização (BOAR, 2002, p. 2).

Nesse contexto, o processo de catalogação deve lançar mão das tecnologias disponíveis para realizar os principais processos dentre eles: a descrição, a importação, a transmissão, a organização, a indexação, o armazenamento, a proteção e a segurança, a localização, a recuperação, a visualização, a impressão e a preservação dos recursos informacionais. É no processo de descrição que a padronização auxilia na construção de representações mais completas e interoperáveis. A construção de estruturas eficientes e um processo de representação adequado é uma forma de tornar os recursos informacionais disponíveis, acessíveis, recuperáveis e únicos. Essa última característica, corrobora com o pensamento de Mey (1995), onde a autora defende que a Catalogação

tem como objetivo tornar o recurso informacional único, por meio de uma estrutura sucinta e padronizada dos dados e informações de um recurso informacional.

Quadro 1 - Paralelo entre os períodos históricos da humanidade e os períodos históricos da Catalogação

Períodos históricos da humanidade	Períodos históricos da Catalogação
Antiguidade	Período Pré-tipográfico
Idade Média	Período Pré-tipográfico
Idade Moderna	Período Tipográfico
Idade Contemporânea	Período Semimecanizado
	Período Mecanizado
	Período Automatizado
	Tecnologias de rede e Web Fonte

Fonte: Alves e Santos (2013, p. 25)

No quadro 1, Alves e Santos (2013) mostram as contribuições de cada período da história para construção da teoria da Catalogação, o desenvolvimento de padrões e o uso estratégico das tecnologias. É possível ver que foi na Idade Contemporânea que a Catalogação avançou significativamente. Nesse contexto, começaram a ser desenvolvidos alguns padrões e formatos de intercâmbio de dados que possibilitaram o compartilhamento de registros bibliográficos legíveis por máquinas.

Nos tempos atuais, o aumento da produção e a propagação de informação nos ambientes digitais tem exigido a adoção de padrões para representação da informação de forma que a recuperação e o acesso sejam possíveis de forma inequívoca, principalmente em sistemas cooperativos e centralizados.

Zafalon (2012) considera que, o que promove consistência, precisão e relevância aos resultados obtidos em resposta a uma consulta é a relação inerente entre representação e recuperação de documentos, levando-se em consideração as ferramentas de descrição, assim como, de estruturas de registros bibliográficos.

Segundo Foulonneau e Riley (2008), existem diferentes formas de categorizar os padrões de metadados. Todavia, para a presente pesquisa, apenas os padrões de conteúdo de metadados e padrões descritivos serão explorados, por alinharem-se à temática proposta.

3.2.2 Padrões de Conteúdo para Catalogação

Pretende-se aqui discorrer sobre os principais padrões de conteúdo, entre eles o AACR2, o seu sucessor, o RDA, e o caminho percorrido até a chegada do LRM, passando pelos modelos conceituais da família FR (FRBR, FRAD e FRSAD).

De acordo com Foulonneau e Riley (2008), os padrões de conteúdo fornecem regras para a sintaxe de uma entrada em um campo de metadado. Seu propósito é permitir uma melhor pesquisa e recuperação pelos usuários, sendo comumente usados em padrões de estrutura de metadados descritivos, mas também são aplicáveis a outros tipos, tais como padrões usados em datas, que podem ser usados como metadados de registros técnicos.

Os padrões de conteúdo usados em larga escala são o AACR2, o CCO (Cataloging cultural objects) e o DA:CS (Describing Archives: content standards), e os utilizados em pequena escala são o Archival Moving Image Materials: a cataloging manual (suplemento do AACR2 para catalogação de materiais de imagem em movimento), o W3CDTF (para codificação de sintaxe para datas), o DOI (para codificação de sintaxe para identificadores), o DCMI Type (para esquema de codificação de vocabulário para tipos de recursos) e o AAT (para esquema de codificação para assuntos de arte e arquitetura) (FOULONNEAU; RILEY, 2008, p.40).

O AACR é um padrão para catalogação, produto de um encontro de ideias de vários teóricos e de catalogadores de diversas partes do mundo.

Por sugestão de Melvil Dewey, em 1904, a American Library Association (ALA), nos Estados Unidos, e a Library Association, na Grã Bretanha, concordaram em cooperar para a elaboração de um código anglo-americano. O resultado da cooperação culminou na publicação, em 1908, do primeiro código internacional de catalogação: *Catalog Rules, Author and Title Entries*, com edição britânica e americana. Após separações causadas pela Segunda Guerra e a realização da *International Conference on Cataloguing Principles*, em 1961, foi publicada, em 1967, as *Anglo-American Cataloguing Rules (AACR)*, ainda em edições em cada país, que contou com considerações sobre os Princípios de Paris e com as ideias de Seymour Lubetzky. (MACHADO; ZAFALON,2020, p. 45)

De acordo com Machado e Zafalon (2020), o *International Standard Bibliographic Description (ISBD)* surgiu no *International Meeting of Cataloguing Experts*, realizado em Copenhague, em 1969.

É nesse contexto de discussões sobre a catalogação compartilhada e a automatização dos serviços de bibliotecas que nasce a *International Standard Bibliographic Description (ISBD)*, cujo objetivo era o de servir como um padrão voltado ao Controle Bibliográfico Universal, ou seja, tornar disponível, de modo internacionalmente aceito, os dados bibliográficos básicos de todos os recursos publicados em todos os países. (MACHADO; ZAFALON, 2020, p. 49)

A ISBD é considerado um dos principais marcos na história da catalogação mundial e umas das mais importantes contribuições para a padronização de descrição bibliográfica, considerado por alguns autores, o padrão de catalogação internacional mais bem-sucedido de todos os tempos. E foi em consequência dessa publicação, que o AACR passou por revisões para atender a essa padronização internacional que a partir de então foi exigida, foi aí que surgiu, no início da década de 1970, o AACR2. A partir de então, várias revisões aconteceram de forma a adequar o código, principalmente às mudanças tecnológicas de foram surgindo com o passar dos anos, uma delas foi a adequação em relação ao MARC, porém, os autores ressaltam que a publicação do AACR2 ocorreu às vésperas da grande mudança tecnológica: o computador e a rede eletrônica, o que evidenciou a necessidade de novas atualizações.

Porém, ante às discussões na comunidade de catalogação, passou-se a acreditar que um modelo teórico de informação bibliográfica precisava ser criado e usado pelos catalogadores no lugar das regras tradicionais. Neste contexto, surgem os chamados modelos conceituais, onde

[...] no universo bibliográfico foram propostos para servir de orientação à catalogação e colocam-se como posições teóricas no sentido de orientar a construção de padrões, normas, códigos que direcionam os registros bibliográficos usando de uniformidade, a despeito das diferenças culturais, linguísticas, tipos de conteúdo e sobre os seus suportes informacionais. (MACHADO; ZAFALON,2020, p. 54)

De acordo com Machado e Zafalon (2020), os modelos conceituais usados no domínio biblioteconômico, reconhecidos como família FR, são modelos abstratos, identificam as tarefas que os usuários devem empreender no processo de descobrimento dos recursos e como os diferentes tipos de dados bibliográficos e de autoridade servem de suporte à execução bem-sucedida dessas tarefas. Os requisitos funcionais propostos pela IFLA: *Functional Requirements for Bibliographic Records* (FRBR), *Functional Requirements for Authority Data* (FRAD) e *Functional Requirements for Subject Authority Data* (FRSAD) foram estabelecidos com a finalidade de orientar a construção de padrões, e não de tornarem-se padrões de catalogação em si mesmos. Neste contexto, surge o RDA - *RESOURCE DESCRIPTION AND ACCESS*, padrão que foi desenvolvido com o objetivo de atender às demandas oriundas do contexto tecnológico e que está alinhado com esses modelos conceituais.

De acordo com os autores, o RDA é resultado de um processo de desenvolvimento e revisão de algo que foi proposto como um padrão de catalogação no bojo da proposta teórica do FRBR e FRAD, mas está repleto de instruções advindas do AACR2. Assim, o RDA passa a ser um novo padrão para descrição e acesso ao recurso como resposta a um ambiente digital em desenvolvimento, no qual tanto a produção quanto a disseminação de recursos de informação e as tecnologias usadas para criar, armazenar e acessar dados descrevendo esses recursos, que também estão em transformação. Como é o caso da necessidade de descrição de novos recursos, novos tipos de mídias e conteúdos, como os catálogos agora disponíveis em ambiente Web. Para isso, o *Joint Steering Committee For Development of RDA* (2009) estabeleceu cinco princípios para orientar seu desenvolvimento:

- [1] generalização,
- [2] especificidade,
- [3] não redundância,
- [4] terminologia e
- [5] estrutura de referência;

Assim como outros oito princípios voltados à funcionalidade para os dados descritivos:

- [1] diferenciação,
- [2] suficiência,
- [3] relacionamentos,
- [4] representação,
- [5] precisão,
- [6] atribuição,
- [7] uso ou prática comum e
- [8] uniformidade.

(JOINT STEERING COMMITTEE FOR DEVELOPMENT OF RDA, 2009).

Ao observar estes princípios, é possível verificar que ao contrário do que acontece nos códigos de catalogação, comumente caracterizados pelo seu conjunto de regras, o RDA tem, desde sua construção, diretrizes e instruções, o que o configura, sob o ponto de vista de Machado e Zafalon (2020), como um padrão.

O *International Federation of Library Associations and Institutions* (2017), como resultado das discussões advindas do alinhamento do padrão dos três modelos conceituais da família FR (FRBR, FRAD e FRSAD) e, ao mesmo tempo, distinto dos mesmos, consolidou em um único modelo que originou o IFLA LRM (*IFLA Library Reference Model*), um modelo conceitual de referência para dados de bibliotecas. Para isso foram analisados comparativamente as tarefas de usuários, as entidades, os atributos e as relações dos três modelos da família FR, da seguinte forma segundo Machado e Zafalon (2020):

[1] exame das tarefas do usuário, o que proporcionou um foco e um alcance funcional para o resto das decisões da modelagem; [2] exame das entidades, e, depois, alternadamente, dos relacionamentos e atributos, processo realizado por meio de várias iterações, pois cada passagem revelou simplificações e refinamentos que, em seguida, precisavam ser aplicados de forma consistente em todo o modelo; e [3] redação das definições, notas de escopo e exemplos, o que fez com que houvesse a definição completa do modelo de forma consistente e integrada. (MACHADO; ZAFALON, 2020, p. 96)

Segundo os autores, no IFLA LRM não existe diferença entre dados armazenados nos registros bibliográficos ou de coleções e os dados armazenados em registros de autoridade de nome e de assunto, visto que todos esses dados são entendidos como informação bibliográfica. Assim como nos modelos da família FR, entidades, atributos, relacionamentos e tarefas de usuários são, predominantemente, as bases nas quais o IFLA LRM se organizou. Desta forma, o IFLA LRM não assume posição de código ou padrão de catalogação, serve apenas como base na implementação de sistemas bibliográficos.

Quadro 2 - Principais diferenças entre os modelos da família FR e o IFLA LRM.

Modelo Conceitual	FRBR (1998)	FRAD (2009)	FRSAD (2010)	IFLA LRM (2017)
Entidades	Obra Expressão Manifestação Item Pessoa Instituição Conceito Objeto Evento Lugar	Obra Expressão Manifestação Item Pessoa Instituição Conceito Objeto Evento Lugar Família Nome Identificador Pontos de acesso controlado Regras Agência	Thema Nomen	Res Expressão Manifestação Item Agente (pessoa e agente coletivo) Nomen Lugar Intervalo de tempo
Tarefas de usuário	Encontrar Identificar Selecionar Obter	Encontrar Identificar Contextualizar Justificar	Encontrar Identificar Selecionar Explorar	Encontrar Identificar Selecionar Obter Explorar

Fonte: Machado e Zafalon (2020, p. 100)

No quadro 2 é possível visualizar as entidades e as tarefas de usuários adotadas nos modelos conceituais da família FR e no IFLA LMR com destaque para as entidades novas propostas em relação ao modelo anterior.

Um componente importante nos modelos conceituais do tipo entidade-relacionamento são os relacionamentos entre as entidades, onde os relacionamentos exercem fundamental importância no IFLA LRM. Da mesma forma que ocorre com as entidades, os relacionamentos obedecem às hierarquias fazendo com que seja cumprida a função de restrição entre os relacionamentos.

Os autores colocam que mesmo que em épocas diferentes e contextos tecnológicos diversos, o que se observa no arcabouço teórico dos estudiosos da catalogação é a preocupação com o usuário. Num esforço para que o catálogo não sirva apenas como um canal de comunicação entre o acervo e seu público, mas que vá além, cumprindo o papel de democratização do acesso à informação.

3.2.3 Padrões de Estrutura de Metadados Descritivos para Catalogação

Essa subseção discorre sobre os principais padrões de estrutura de metadados descritivos, entre eles o ISBD, o formato MARC Bibliográfico e BibFrame.

Foulonneau e Riley (2008) postulam que padrões de estrutura de metadados descritivos são aqueles que listam elementos importantes no processo de descrição do recurso que inclui as características físicas e de conteúdo. Os padrões de estrutura de metadados descritivos mais gerais, indicados por Foulonneau e Riley (2008),

são o MARC Bibliográfico, o MARCXML, o MODS e o Dublin Core, e os mais especializados são o VRA Core (para recursos visuais), o CDWA lite (para arte e arquitetura), o GEM (para objetos de aprendizagem), o IMS Learning Resource Metadata (para objetos de aprendizagem), o ETD-MS (para dissertações e teses eletrônicas) e o DDI (para conjuntos de dados de ciências sociais e comportamentais) (FOULONNEAU; RILEY, 2008, p.40).

Zafalon (2012) também inclui a ISBD como padrão de estrutura de metadados descritivos, por entender que ela está voltada para os aspectos estruturais de um registro bibliográfico, este entendimento é tido a partir de estudos de Swanson (1973) e Langker (1974) visto que estes autores entendem que a ISBD especifica os elementos de descrição bibliográfica, descreve a ordem em que esses elementos devem ser apresentados e além disso, ela indica os sinais de pontuação que delimitam os elementos de descrição de

modo a precisar os elementos bibliográficos independentemente da linguagem, tanto para seres humanos quanto para máquinas. Desta forma, a autora destaca que a ISBD visa três objetivos: “fazer registros de diferentes fontes intercambiáveis, facilitar a sua interpretação acima das barreiras linguísticas e facilitar a conversão de tais registros à forma legível por máquina” (ZAFALON, 2012, p. 42).

Porém, quando se trata de descrição e recuperação de informações em meios digitais é necessário que em conjunto com às AACRs/RDA também seja utilizado um padrão de estrutura de metadados descritivos, neste caso, o formato MARC21 Bibliográfico seria o mais indicado, pois, abrange aspectos de leitura e de interpretação de dados disponíveis em registros bibliográficos, por meio de marcação desses registros, onde o processo de leitura e interpretação computacional de um registro bibliográfico em Formato MARC é facilitado.

A informação de um catálogo bibliográfico não pode ser simplesmente digitada em um computador para produção de um catálogo automatizado. O computador necessita de um meio para interpretar a informação contida em um registro bibliográfico. O registro MARC contém um guia para a interpretação destes dados antes de cada peça da informação bibliográfica (ZAFALON, 2008, p. 14).

A autora destaca que as convenções adotadas nas marcações de um registro bibliográfico, como padrão de estrutura de metadados descritivos, quer seja pela ISBD, quer seja pelo Formato MARC21 Bibliográfico, tornam o registro, o gerenciamento e a recuperação dos dados bibliográficos mais eficientes.

Quadro 3 - Representação de um documento segundo o padrão de estrutura de metadados descritivo MARC e o padrão de conteúdo AACR2

<i>Padrão de estrutura de metadados descritivos</i>				<i>Padrão de Conteúdo</i>
<i>Formato MARC21 Bibliográfico</i>				<i>AACR2r</i>
<i>TAG</i>	<i>Ind.1</i>	<i>Ind.2</i>	<i>Subcampo</i>	<i>Conteúdo do registro bibliográfico</i>
020	∅	∅	\$a	85-00-01493-9
090	∅	∅	\$a	925
			\$b	T49c
100	1	∅	\$a	Tiner, John Hudson
245	1	0	\$a	100 cientistas que mudaram a história do mundo
			\$c	John Hudson Tiner: Tradução de Marise Chinetti de Barros
260	∅	∅	\$a	Rio de Janeiro
			\$b	Ediouro
			\$c	c2002
300	∅	∅	\$a	259b.
490	0	∅	\$a	Colecção 100 mais
500	∅	∅	\$a	Tradução de: 100 scientists who shaped world history

650	1	4	\$a	CIENTISTAS
			\$b	BIOGRAFIA
650	2	4	\$a	CIÊNCIAS
			\$x	HISTÓRIA
740	0	Ø	\$a	Cem cientistas que mudaram a história do mundo

Fonte: Zafalon (2012 p.45)

De acordo com a autora, é possível observar no quadro 3 que um registro descrito no formato MARC recebe marcações para que o computador possa interpretá-lo, assim poderá ser recuperado pelos pontos de acesso que forem definidos.

De acordo com Espíndola (2018) o formato MARC possui dificuldade de descrever os relacionamentos existentes entre as entidades de dados bibliográficos. Desta forma, a Diretriz RDA com foco nos relacionamentos existentes entre as entidades, uma herança dos Modelos Conceituais FRBR e FRAD, pode não ser alcançada com o emprego do MARC. Apesar de sua ampla utilização na representação de dados bibliográficos, ainda em 2008 um relatório final de um grupo de trabalho da *Library of Congress* intitulado *Working Group on The Future of Bibliographic Control* concluiu o seguinte sobre o MARC:

O suporte de dados da comunidade de bibliotecas, MARC, é baseado em técnicas de quarenta anos para gerenciamento e está fora de sintonia com os estilos de programação de hoje. Nenhuma comunidade além da comunidade de bibliotecas usa este formato de registro, comprometendo severamente sua utilidade para outras comunidades como uma ferramenta de transmissão de dados. Aplicativos bibliográficos sendo desenvolvidos fora do ambiente da biblioteca não estão fazendo uso e podem não ser compatíveis com registros codificados em MARC. Usos novos e antecipados de dados bibliográficos requerem um formato que acomode e distinga metadados gerados por especialistas, automatizados e gerados pelo usuário, incluindo anotações (revisões, comentários) e dados de uso. O design flexível deve permitir o uso seletivo (modular) de metadados em diferentes ambientes (por exemplo, uso de vocabulários controlados apropriados para domínios) (LIBRARY OF CONGRESS, 2008, p. 24 tradução nossa).

Segundo Mitchell (2013), com base neste relatório, a *Library of Congress* lançou em maio de 2011, o modelo para representação de dados bibliográficos chamado Bibframe que visa a substituir o formato de intercâmbio bibliográfico MARC e foi criado com o propósito de melhor integrar a catalogação à Web semântica e atender as funcionalidades requeridas pelo Modelo Conceitual FRBR e seus associados. Modelo, este que tem por objetivo ser uma forma de representar os dados bibliográficos dos recursos informacionais na Web, a partir dos princípios *linked data*, ou dados ligados, que determinam as boas práticas para construir ligações semânticas entre dados de diferentes fontes, no contexto da Web semântica. De acordo com Espíndola (2018), um diferencial do Bibframe é o

emprego de identificadores únicos, como o *Uniform Resource Identifier - URI*¹⁸, visto que a catalogação realizada até o momento com o MARC, é focado em criar registros de catálogos que sejam independentemente compreensíveis utilizando, por exemplo, o nome do autor de uma obra, ao invés de adotar um identificador único para ele, como o *Open Researcher and Contributor ID (ORCID)* ou o identificador do Currículo Lattes, por exemplo. Desta forma, o Bibframe é visto como uma importante iniciativa de atualização das práticas e ferramentas de catalogação, de forma a integrar o processo e a prática de catalogação à Web assim como às necessidades dos usuários que utilizam este ambiente em busca de informações.

3.3 PRINCÍPIOS FAIR

De acordo com Wilkinson et al (2016), o bom gerenciamento de dados não é um objetivo em si, mas um caminho que leva à descoberta e à inovação do conhecimento, e conseqüentemente à integração e reutilização destes dados pela comunidade que deles necessita.

Os chamados princípios FAIR, acrônimo para *Findable, Accessible, Interoperable and Reusable*, estão presentes nas discussões da chamada ciência de dados desde de 2014, fruto de uma conferência internacional chamada *Jointly Designing a Data FAIRPORT*, mas que foram consolidados em 2017 quando houve a exigência, por parte da Comissão Europeia, de um plano de gestão de dados com base nestes princípios que desde então passaram a ser norteadores da descoberta, do acesso, da interoperabilidade, do compartilhamento e da reutilização dos dados de pesquisa. Tanto que em 2018 a *European Commission Expert Group on FAIR Data*¹⁹ publicou um instrumento com o intuito de ampliar a compreensão dos princípios chamado "*Turning FAIR into reality*", um instrumento para aplicação dos princípios em diferentes contextos. (JOINTLY DESIGNING A DATA FAIRPORT, 2014; EUROPEAN COMMISSION EXPERT GROUP ON FAIR DATA, 2018).

Wilkinson et al (2016), na primeira publicação formal dos princípios FAIR - (*The FAIR Guiding Principles for scientific data management and stewardship*), explica que existe

¹⁸Uniform Resource Identifier (URI) é um termo técnico (e anglicismo de tecnologia da informação) que foi traduzido para a língua portuguesa como um "identificador uniforme de recurso", é uma cadeia de caracteres compacta usada para identificar ou denominar um recurso na Internet. Disponível em: <https://pt.wikipedia.org/wiki/URI>.

¹⁹Disponível em: <https://ec.europa.eu/transparency/expert-groups-register/screen/expert-groups/consult?do=groupDetail.groupDetail&groupID=3464>

uma necessidade urgente de melhorar a infraestrutura que suporta a reutilização de dados acadêmicos. Por isso, as partes interessadas neste processo como a academia, indústria, agências de fomento e editores acadêmicos - se reuniram para projetar e endossar um conjunto conciso e mensurável de princípios que tem por objetivo, servir como uma diretriz para aqueles que desejam aumentar a reutilização de seus acervos de dados.

Diferentes de iniciativas semelhantes que se concentram no ser humano acadêmico, os princípios FAIR enfatizam especificamente o aprimoramento da capacidade das máquinas de encontrar e usar os dados, além de apoiar sua reutilização pelos indivíduos. A publicação dos autores, descreve a lógica por trás dos quatro princípios fundamentais do FAIR (Encontrabilidade, Acessibilidade, Interoperabilidade e Reutilização) que servem para orientar os produtores e editores de dados à medida que navegam por esses obstáculos, ajudando assim a maximizar o valor agregado adquirido pelos estudos acadêmicos formais contemporâneos. Os autores ressaltam que a intenção é que os princípios se apliquem, não apenas aos “dados” no sentido convencional, mas também aos algoritmos, ferramentas e fluxos de trabalho que levaram a esses dados, incluindo-se aqui, os metadados. Desta forma, todos os objetos se beneficiam da aplicação desses princípios, uma vez que todos os componentes do processo de pesquisa devem estar disponíveis para garantir transparência, reprodutibilidade e reutilização.

Hanning et al (2019) relatam que desde então os princípios FAIR vêm sendo adotados por instituições de pesquisa e considerados o instrumento norteador para a descoberta, o acesso, a interoperabilidade global e as discussões estão encaminhando no sentido de ampliar o reuso de dados científicos a partir da adoção de padrões, metadados, vocabulários controlados, ontologias e identificadores persistentes - exemplo DOI, ORCID, etc - que proporcionam significado preciso aos dados e aos demais objetos que a eles estejam vinculados.

Wilkinson et al (2016, tradução nossa), explica que são ao todo quinze princípios que estão divididos em quatro categorias:

Na categoria F – Findable temos: F1. Os (meta) dados devem ter identificadores globais, persistentes e identificáveis; F2. Os dados devem ser descritos com (Meta) dados ricos; F3. Os (meta) dados devem ser registrados ou indexados em recursos que ofereçam capacidades de busca; F4. Metadados devem especificar o identificador dos dados.

Na categoria A – Accessible temos: A1. (Meta) dados devem ser recuperáveis pelos seus identificadores usando protocolo de comunicação padronizado; A1.1 O protocolo deve ser aberto, gratuito e universalmente implementável; A1.2 O protocolo deve permitir procedimentos de autenticação e autorização, quando necessário; A2. (Meta) dados devem ser acessíveis, mesmo quando os dados não estão mais disponíveis.

Na categoria I – Interoperable temos: I1. (Meta) dados devem usar uma linguagem formal, acessível, compartilhada e amplamente aplicável para a representação do conhecimento; I2. (Meta) dados devem usar vocabulários que seguem os princípios FAIR; I3. (Meta) dados devem incluir referências qualificadas para outros (Meta) dados;

E por fim na categoria R – Reusable, temos R1. (Meta) dados devem ter atributos com pluralidade de precisão e serem relevantes, R1.1. (Meta) dados devem ser liberados com licenças de uso de dados claras e acessíveis, R1.2. (Meta) dados devem estar associados à sua proveniência e R1.3. (Meta) dados devem estar alinhados com padrões relevantes ao seu domínio.

É possível perceber que cada letra do FAIR faz menção a um conjunto de especificações que demandam domínios de conhecimentos técnicos e estruturais distintos. De acordo com Henning et al (2019), esse princípios não devem ser entendidos como um padrão, mas, como um guia para preparação dos dados visando o reuso, sob condições claramente descritas tanto por pessoas quanto por máquinas.

Figura 5 - Dados FAIR

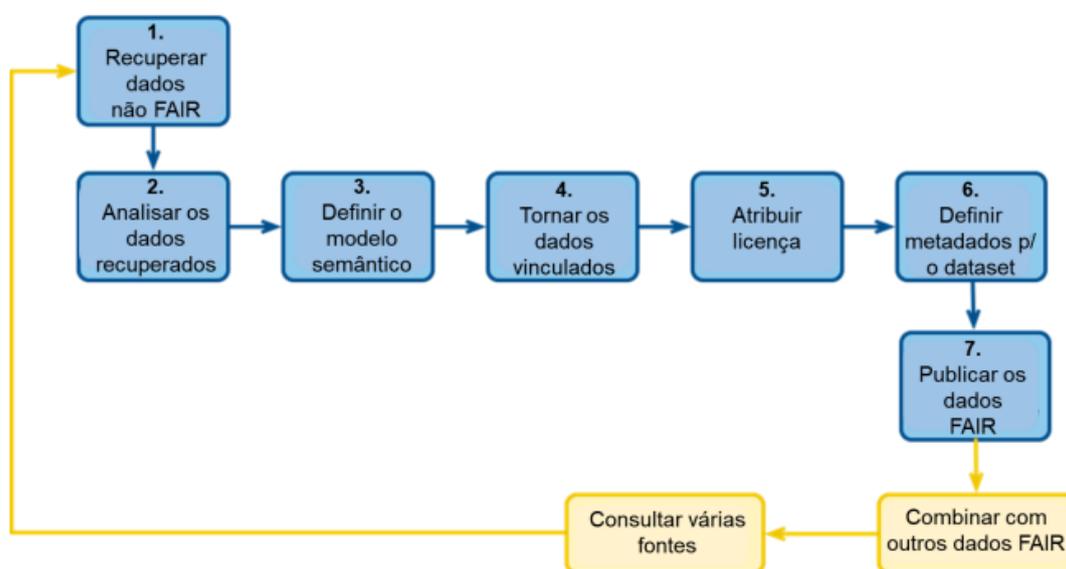


Fonte: Salles (2022)

De acordo com a Figura 5, Sales (2022) nos mostra que para que um dado/objeto possa atender aos princípios FAIR, ele deve ser:

- Encontrável: onde o dado/objeto precisa ter um identificador persistente único e deve ser registrado e indexado num mecanismo de busca;
- Acessível: pode ser baixado por meio de protocolos de comunicação padronizados;
- Interoperável: tratado como padrões abertos (formato, linguagem e vocabulário);
- Reusável: devido a condições de uso conhecidas, poderão ser reusados por outros;

Figura 6 - Processo de FAIRificação dos dados



Fonte: Salles (2022)

De acordo com a Figura 6, a Salles (2022) propõe que para FAIRificar os dados é necessário seguir os passos preconizados pela iniciativa GO FAIR²⁰ que incluem:

Para torná-lo **Encontrável**:

- Um repositório deve ser selecionado, e os formatos de dados e metadados; necessários devem ser verificados;
- Um identificador persistente precisa ser utilizado;

²⁰ GO FAIR é uma iniciativa de baixo para cima, orientada para as partes interessadas e autogovernada que visa implementar os princípios de dados FAIR, tornando os dados Localizáveis, Acessíveis, Interoperáveis e Reutilizáveis (FAIR). Oferece um ecossistema aberto e inclusivo para indivíduos, instituições e organizações que trabalham juntos por meio de Redes de Implementação (INs). Os INs atuam em três pilares de atuação: GO CHANGE, GO TRAIN e GO BUILD. Disponível em: <https://www.go-fair.org/go-fair-initiative/>

- Um catálogo ou mecanismo de busca deve ser selecionado.

Para torná-lo **Acessível**:

- A longevidade aos dados deve ser garantida por meio de certificação;
- Verifique as condições legais sob as quais os dados podem ser disponíveis e acessíveis;
- Estabeleça embargo quando necessário;
- Verifique se sua infraestrutura manterá o dado disponível no caso de defeito do equipamento ou erro humano.

Para torná-lo **Interoperável**:

- Selecione os formatos de dados mais usados;
- Selecione padrões mais usados;
- Selecione e aplique os vocabulários mais usados.

Para torná-lo **Reusável**

- Registre informação sobre a proveniência do dado;
- Selecione padrões gerais mínimos de metadados;
- Atribua uma licença específica aos dados;
- Relacione outros documentos e informações sobre o assunto do dado.

Porém, a autora esclarece que é muito difícil conseguir aplicar todos os princípios, e por isso, hoje se tem uma chamada escala de FAIRificação, que mede a qualidade dos dados, de acordo com a aplicação desses princípios. Segundo ela, quanto maior for o grau de FAIRificação dos dados, maior é a possibilidade desse dado ser encontrado, acessado, interoperável e reusável. Com base nessa premissa, entende-se que o mesmo ocorra quando se fala de OAs.

3.4 OBJETOS DE APRENDIZAGEM

O termo “*Learning Object*” foi cunhado por Wayne Hodgins em 1996, ao observar seu filho brincando com peças de Lego, quando percebeu o processo de *e-Learning* como blocos de construção interoperáveis - que pudessem ser executados em diferentes ambientes de aprendizagem - e permitissem a reutilização de conteúdos de aprendizagem (SILVEIRA, 2014). Porém, o conceito de objeto de aprendizagem só começou a ser amplamente estudado a partir da dissertação de Wiley (2000), “*Learning object design and sequencing theory*”, que a partir de então, passou a ser um referencial importante relacionado ao tema. No trabalho, o autor explica que o conceito de OAs não é novo e a ideia de usar pequenos segmentos de conteúdo educacional que possam ser reutilizados é bem antiga. (TAROUCO, 2012).

Já em 2003, Wiley faz críticas à metáfora do Lego de Hodgins pois, de acordo com o autor, essa metáfora apesar de conseguir explicar de maneira simples a ideia de OAs como “pedaços de instrução” que podem ser encaixados em outros de forma a criar uma estrutura maior ou reusados em outras estruturas, na visão do autor, é uma ideia bem limitada, pois existem propriedades nesses blocos de encaixe que não podem ser relacionadas com os OAs, como:

- Qualquer bloco de Lego é combinável com qualquer outro bloco de Lego.
- Os blocos de Lego podem ser unidos de várias maneiras.
- Os blocos LEGO são tão divertidos e simples que qualquer um pode montá-los (WILEY, 2003, p. 1).

Figura 7 - Metáfora Lego



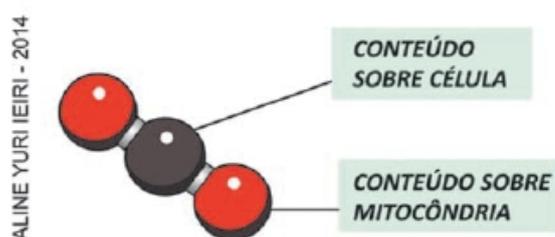
Fonte: Braga (2014, p. 44)

Na Figura 7, Braga (2014) utiliza a metáfora do Lego de Wayne Hodgins que mostra cada peça de lego como sendo um OA, e onde cada OA representa um conhecimento. Logo, unindo peças é possível criar novos conhecimentos. Porém, utilizando-se da crítica de Wiley a esta metáfora, a autora questiona: Um conhecimento pode ser unido a qualquer outro conhecimento? Um conhecimento pode ser unido da maneira que o professor quiser?

Assim como os blocos de Lego, os OAs são reusáveis, porém, a combinação com outros OAs não é automática, nem tão pouco simples. Por isso, o Wiley (2003) propõe outra metáfora, a do átomo, onde, um átomo, semelhante ao bloco do Lego, pode ser combinado com outros átomos, de forma a criar outras estruturas.

Figura 8 - Metáfora Átomo

Aula de Biologia Celular



Fonte: Braga (2014, p. 45)

De acordo com a Figura 8, a grande diferença é que nem todo átomo pode ser combinado com todos os outros átomos. Desta forma, as estruturas e características dos átomos são fatores determinantes para a combinação com outros átomos. Então, infere-se que é necessário certo nível de conhecimento sobre os átomos, ou seja, é importante conhecer seu conteúdo e suas especificidades para que se possa combiná-los.

Porém, o que essas metáforas mostram, é que a característica fundamental relacionada aos OAs é a de organizar o conteúdo educacional em pequenas unidades que se combinadas entre si formam outras unidades mais complexas.

Wiley (2003) afirma que a tecnologia é um grande agente de mudança de paradigma na maneira como as pessoas aprendem e a forma como os materiais educacionais são projetados, desenvolvidos e entregues para aqueles que desejam aprender, são consequência desta mudança. Neste contexto, devido ao seu potencial de reutilização,

capacidade de geração, adaptabilidade e escalabilidade o OA lidera outros candidatos como a tecnologia de escolha para a próxima geração de design instrucional. Segundo o autor, o conceito de OA deriva da ciência da computação, baseado na programação orientada a objetos, onde há a criação de componentes chamados de “objetos” que podem ser reutilizados em múltiplos contextos. Assim, a ideia fundamental desta tecnologia é que os designers instrucionais possam construir pequenos componentes instrucionais e estes serem reutilizados inúmeras vezes em diferentes contextos de aprendizagem. Além disso, os OAs são geralmente entendidos como entidades digitais que podem ser entregues pela Internet, o que significa que várias pessoas podem acessá-los e usá-los simultaneamente. Os indivíduos que incorporam OAs podem colaborar e se beneficiar imediatamente de novas versões, o que é um diferencial significativo entre estes e outras mídias existentes. Wiley aponta que essa reusabilidade é uma característica primordial dos OAs, pois permite a organização de uma grande justaposição de componentes básicos na ordem mais apropriada para cada necessidade educacional identificada. Os OAs são recursos utilizados para a enriquecer o espaço pedagógico nos AVAs e valorizar os conteúdos apresentados.

Os OAs ainda são considerados uma tecnologia relativamente recente, por isso, comumente encontra-se diferentes conceitos para os mesmos na literatura. O Comitê internacional de padrões para a tecnologia, denominado *IEEE Learning Technology Standards Committee (LTSC)*²¹, definiu um objeto de aprendizagem como: “qualquer entidade, digital ou não digital, que é usada para aprendizagem, educação ou treinamento.” (IEEE, 2020 p.13, tradução nossa). Já na a definição de Wiley (2000) um OA é “Qualquer recurso digital que possa ser reutilizado para apoiar a aprendizagem.” e Behar (2009) definiu um OA como,

[...] qualquer material digital como, por exemplo, textos, animação, vídeos, imagens, aplicações, página web de forma isolada ou em combinação, com fins educacionais. Tratam-se de recursos autônomos, que podem ser utilizados como módulos de um determinado conteúdo ou como um conteúdo completo. São destinados a situações de aprendizagem tanto na modalidade a distância quanto semipresencial ou presencial. Uma das principais características deste recurso é a reusabilidade, ou seja, a possibilidade de serem incorporados a múltiplos aplicativos (BEHAR, 2009, p. 67).

²¹ Comitê da IEEE, com a atribuição de desenvolver padrões, práticas recomendadas e guias para a tecnologia da aprendizagem que sejam aceitos internacionalmente. Disponível em: <https://sagroups.ieee.org/ltsc/>

Pode-se observar que a característica do uso e reuso é uma constante em todas essas conceituações apresentadas, o que mostra a importância deste processo quando se fala de OA. Em relação à reusabilidade, é importante ressaltar o trabalho do *IEEE Learning Technology Standards Committee (LTSC)* no estabelecimento de padrões de metadados para catalogação destes objetos em repositórios, onde o metadado de um OA descreve as características relevantes visando a sua recuperação por meio de sistemas de buscas na Web ou utilizados nos Ambientes Virtuais de Aprendizagem - AVA. (IEEE, 2020)

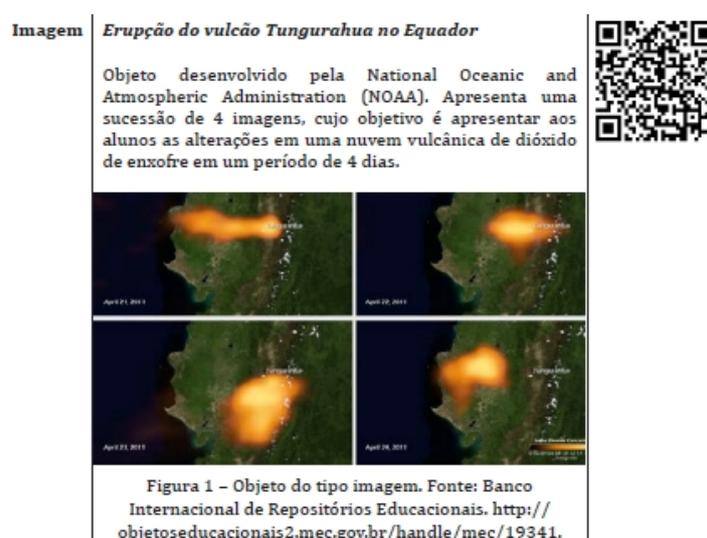
3.4.1 Tipos de Objetos de Aprendizagem

Como a definição de OA na literatura é muito ampla, para efeito deste trabalho é importante ressaltar que um AVA, apesar de ser um recurso digital, não é considerado como sendo um OA, pois voltando à analogia das peças de Lego de Wayne Hodgins, os OAs são as unidades que compõem o AVA no auxílio ao processo de ensino-aprendizagem.

De acordo com Braga e Meneses (2014), são considerados alguns tipos de OAs:

Imagem: Conforme modelo da Figura 9, imagem é a “[...] representação de uma pessoa ou coisa. Uma imagem digital pode ser utilizada para apoiar a aprendizagem e, devido a isso, é considerada como um tipo de OA” (BRAGA; MENESES, 2014, p. 22).

Figura 9 - Tipo de OA - Imagem



Fonte: Braga e Meneses (2014, p. 22)

Áudio: Conforme modelo da Figura 10, áudio “[...] é uma faixa do espectro reservada ao som, em contraposição ao vídeo. Pode atuar sozinho como um objeto de aprendizagem, desde que seja utilizado para aprendizagem” (BRAGA; MENESES, 2014, p. 22).

Figura 10 - Tipo de OA - Áudio

Áudio	<p><i>Gravidez na Adolescência - Parte I.</i></p> <p>Objeto desenvolvido na Universidade Estadual de Campinas (UNICAMP), que apresenta dados estatísticos sobre a gravidez na adolescência. http://objetoseducacionais2.mec.gov.br/handle/mec/20385</p>	
--------------	--	---

Fonte: Braga e Meneses (2014, p. 22)

Vídeo: Conforme Figura 11, Vídeo “[...] trata-se de uma gravação de imagens em movimento composta por fotos sequenciais que resultam em uma imagem animada (Stop--motion). Um vídeo utilizado para apoiar a aprendizagem é considerado um OA” (BRAGA; MENESES, 2014, p. 22).

Figura 11 - Tipo de OA – Vídeo

Vídeo	<p><i>Animais e Ambiente</i></p> <p>Objeto desenvolvido pela Universidade Estadual de Campinas (UNICAMP), que apresenta a diversidade de seres vivos e diferentes ambientes, para que o aluno entenda a relação existente entre as características de várias espécies de animais e o meio em que vivem. Possui tradução em Língua Brasileira de Sinais (Libras).</p>	
--------------	--	---



Figura 2 – OA do tipo vídeo. Fonte: Banco Internacional de Repositórios Educacionais. <http://objetoseducacionais2.mec.gov.br/handle/mec/21217>.

Fonte: Braga e Meneses (2014, p. 22)

Hipertexto: De acordo com a Figura 12, um hipertexto é uma “[...] organização da informação, no qual certas palavras de um documento estão ligadas a outros documentos, exibindo o texto quando a palavra é selecionada”. Um hipertexto é também

conhecido popularmente como páginas na Internet. Um hipertexto pode ser utilizado como apoio ao aprendizado, portanto, pode ser considerado um OA. Encaixam-se nessa categoria: aulas virtuais, cursos a distância (BRAGA; MENESES, 2014, p. 25).

Figura 12 - Tipo de OA - Hipertexto

Hipertexto

Fichamento: quando utilizar e como elaborar

Objeto desenvolvido pela Universidade de Brasília (UnB), que tem como objetivo prover referências para a elaboração e a utilização de fichamentos, apresentando sua definição, características e função.

Verifique se quer e se pode fazer parte do Fichamento de estudos e prepare os fichamentos com os seguintes:

Composição do texto	Fichamento para estudos e pesquisas	Fichamento científico ou acadêmico
Necessidade de definir o assunto do texto ou parte do texto	<input type="checkbox"/>	<input type="checkbox"/>
Indicar a referência do texto	<input type="checkbox"/>	<input type="checkbox"/>
Utilizar técnicas semelhantes a obtenção de resumos e resenhas	<input type="checkbox"/>	<input type="checkbox"/>
Prezará o código de arquivamento que pode ser numérico, alfabético ou alfanumérico	<input type="checkbox"/>	<input type="checkbox"/>
Necessidade de metodologia para organização	<input type="checkbox"/>	<input type="checkbox"/>

Estapas de elaboração de um fichamento: revisão

Composição do texto	Fichamento para estudos e pesquisas	Fichamento científico ou acadêmico
Prezará o arquivamento de determinado assunto para facilitar a busca	<input type="checkbox"/>	<input type="checkbox"/>
Não generalizar a identificação do texto	<input type="checkbox"/>	<input type="checkbox"/>
Indicar a localidade se encontrar o livro ou documento	<input type="checkbox"/>	<input type="checkbox"/>
Utilizar citação textual em sua resenha	<input type="checkbox"/>	<input type="checkbox"/>
Tratar o título de quem elaborou o fichamento	<input type="checkbox"/>	<input type="checkbox"/>

Tipos de fichamentos: diferenças

Analise se quer e se pode fazer parte do Fichamento de estudos e prepare os fichamentos com os seguintes:

Forma que compõem o Fichamento	Transcrição ou citação (de acordo com o conteúdo do Conteúdo ou referência)
1. Utiliza referências baseadas no texto (ZETAGINT 602)	<input type="checkbox"/>
2. Apresenta aspectos semelhantes ao do resumo	<input type="checkbox"/>
3. Utiliza apenas de início ao fim do texto	<input type="checkbox"/>
4. Analisa, com especial de quem elaborou o livro ou documento	<input type="checkbox"/>
5. Utiliza os mesmos procedimentos para a composição dos textos de resumo	<input type="checkbox"/>

Figura 4 – OA do tipo hipertexto. Fonte: Banco Internacional de Repositórios Educacionais. <http://objetoseducacionais2.mec.gov.br/handle/mec/16230>.

Fonte: Braga e Meneses (2014, p. 25)

Animações: animação significa “dar vida” a objetos estáticos, que podem ser imagens, textos etc. Segundo Braga e Meneses, existem três formas principais de animações com vários subtipos, que são classificados de acordo com a forma, os materiais envolvidos e a tecnologia utilizada para criar as imagens. A animação tradicional ou desenhada a mão. Isso inclui os muitos desenhos animados e filmes de longa-metragem feitos nos últimos 70 anos, o que é às vezes chamado de “animação cel.” Esse termo refere-se às folhas de acetato transparente onde os diagramas são desenhados e fotografados no filme, a fim de que possam ser mostrados rapidamente, para criar a ilusão de movimento. A animação stop-motion que consiste em tirar fotografias digitais fixas de objetos ou imagens depois de terem sido movidos manualmente para simular um movimento. Uma terceira forma de animação são aquelas realizadas por computador e têm imagens que são criadas digitalmente, com o uso de uma grande variedade de novas técnicas e softwares. (BRAGA; MENESES, 2014, p. 23). O quadro 4 mostra exemplos de OA de Animação.

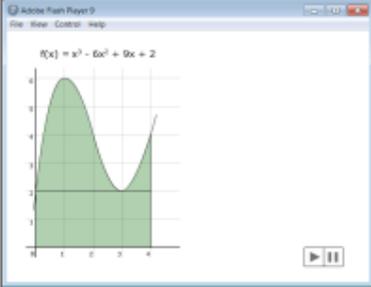
Quadro 4 - Tipo de OA - Formas de Animação

Forma de animação	Características	Tipos	Exemplos
Manual (<i>Hand-drawn</i>)	Imagens são desenhadas a mão e digitalizadas para o computador	Desenhos animados Personagens animados	A Turma da Mônica Flintstones Jetsons Rei Leão Desenhos da Disney
<i>Stop-motion</i>	Objetos, modelos ou imagens são criados e pequenos movimentos são feitos com a mão, e os modelos são fotografados individualmente.	Animação de argila Animação de cortes Animação de modelos Animação de objetos Animação de marionetes Animação de silhuetas	Wallace e Gromit Gumby A Fuga das Galinhas Os Muppets harvie krumpet
Animações computacionais	Imagens são criadas digitalmente e manipuladas pelo computador.	Animação 2-D e 3-D Animação com esqueleto Captura de movimento (ou <i>mocap</i>) Animação Morph (ou de metamorfose) Animação <i>Flash</i> Animação em Slides (ex.: Power Point)	Sherek Carros Procurando Nemo

Fonte: Braga e Meneses (2014, p. 24)

Simulação: As simulações são animações que representam um modelo da natureza e, devido a isso, podem ser muito utilizadas como objetos de aprendizagem. Sistemas computacionais para simulação auxiliam os desenvolvedores e pesquisadores, na medida em que permitem estudar o modelo em ambientes controlados, possibilitando a análise de itens tais como: a dinâmica do modelo, detalhes de sua estrutura, execução variada da simulação, alterando parâmetros de entrada para verificar os resultados obtidos etc. A Figura 13 mostra um exemplo de OA de simulação.

Figura 13 - Tipo de OA - Simulação

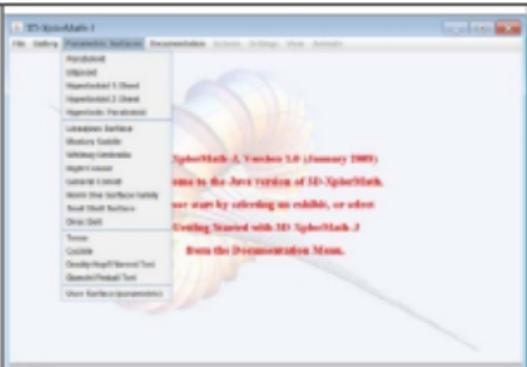
Simulação	<p>Análise Integral de Riemann</p> <p>Animação em flash desenvolvida pelo Instituto Superior de Formación y Recursos en Red para el Profesorado, que tem como objetivo compreender o cálculo de áreas, utilizando o conceito da integral de Riemann.</p>  <p>Figura 3 - OA do tipo simulação. Fonte: Banco Internacional de Repositórios Educacionais. http://objetoseducacionais2.mec.gov.br/handle/mec/17431.</p>	
-----------	---	--

Fonte: Braga e Meneses (2014, p. 27)

Softwares: são programas de computadores que permitem executar determinadas tarefas e resolver problemas de forma automática. Muitos softwares podem ser utilizados para apoiar a aprendizagem de maneira direta e, por isso, podem ser considerados objetos de aprendizagem. (BRAGA; MENESES, 2014, p. 25). A Figura 14 mostra um exemplo de OA de Software.

Figura 14 - Tipo de OA - Software

Software	<p>3D-XplorMath-J</p> <p>Objeto pertencente a The 3D-XplorMath Consortium, que permite a visualização de curvas, superfícies, fractais, dentre outros, possibilitando que os alunos tenham uma visão concreta das equações.</p>	
----------	--	---

	<p>Figura 5 - OA do tipo software educacional. Fonte: Banco Internacional de Repositórios Educacionais. http://objetoseducacionais2.mec.gov.br/handle/mec/14475.</p>
--	--

Fonte: Braga e Meneses (2014, p. 28)

De acordo com as autoras, o professor deve selecionar o tipo de OA que seja mais se adequa ao conteúdo e objetivos de aprendizagem. O OA tende a complementar o ensino, mas deve estar associado a uma estratégia pedagógica. Onde entra outra questão

importante sobre os OAs, que é o grau de interatividade, que é maior, quanto maior a capacidade de intervenção do aluno no conteúdo ensinado por esse OA. Um OA com alta interatividade possibilita a ação do aluno e o estabelecimento de uma relação de reciprocidade. Ou seja, quanto mais o OA permite que o aluno se aproprie de informações, reflita e seja ativo em seu processo de aprendizagem, mais interativo ele é.

3.4.2 Características dos Objetos de Aprendizagem

De acordo com Braga e Meneses (2014), os OAs possuem dois tipos de características: (1) Pedagógicas, que estão relacionadas ao trabalho de professores e alunos na busca por conhecimento, e esta se subdivide em:

- **Interatividade:** indica se há suporte às consolidações e ações mentais, requerendo que o aluno interaja com o conteúdo do OA de alguma forma, podendo ver, escutar ou responder algo.
- **Autonomia:** indica se os objetos de aprendizagem apoiam a iniciativa e tomada de decisão.
- **Cooperação:** indica se há suporte para os alunos trocarem opiniões e trabalhar coletivamente sobre o conceito apresentado.
- **Cognição:** refere-se às sobrecargas cognitivas alocadas na memória do aluno durante o processo de ensino-aprendizagem.
- **Afetividade:** refere-se aos sentimentos e motivações do aluno com sua aprendizagem e durante a interação com o OA (BRAGA; MENESES, 2014, p. 33).

E as características (2) Técnicas que se referem às questões tecnológicas associadas ao OA:

- **Disponibilidade:** indica se o objeto está disponível para ser utilizado.
- **Acessibilidade:** indica se o objeto pode ser acessado por diferentes tipos de usuários (ex: idosos, deficientes visuais etc), em diferentes lugares (ex.: lugares com acesso a Internet, lugares sem acesso a Internet etc.) e por diferentes tipos de dispositivos (ex.: computadores, celulares, tablets etc.).
- **Confiabilidade:** indica que o OA não possui defeitos técnicos ou problemas no conteúdo pedagógico.
- **Portabilidade:** indica se o OA pode ser transferido (ou instalado) para diferentes ambientes, como, por exemplo, diferentes tipos de AVAs ou sistemas operacionais.
- **Facilidade de instalação:** indica se o OA pode ser facilmente instalado caso ele exija esse recurso.
- **Interoperabilidade:** medida de esforço necessário para que os dados dos OAs possam ser integrados a vários sistemas.
- **Usabilidade:** indica a facilidade de utilização dos Oas por alunos e professores.
- **Manutenibilidade:** é a medida de esforço necessária para alterações do OA.
- **Granularidade:** de maneira geral, a palavra granularidade origina-se da palavra grão, sendo que quanto maior o número de grãos de um sistema maior a sua granularidade. Trazendo esse conceito para o âmbito dos objetos de aprendizagem, a granularidade é a extensão à qual um OA é composto por componentes menores e reutilizáveis.
- **Agregação:** indica se os componentes do OA (grãos) podem ser agrupados em conjuntos maiores de conteúdos como, por exemplo, as estruturas tradicionais de um curso.
- **Durabilidade:** indica se o OA se mantém intacto quando o repositório em que ele está armazenado muda ou sofre problemas técnicos.

- **Reusabilidade:** indica as possibilidades de reutilizar os OAs em diferentes contextos ou aplicações. Essa é a principal característica do OA e pode ser influenciada por todas as demais (BRAGA; MENESES, 2014, p. 34-35).

Em relação a essas características, as autoras ressaltam que os OAs não necessariamente irão apresentar todas essas que foram listadas, porém, quanto mais características ele apresentar, maior será a possibilidade de reutilização, pois são elas que servem como parâmetro para avaliar o nível de reuso de um OA.

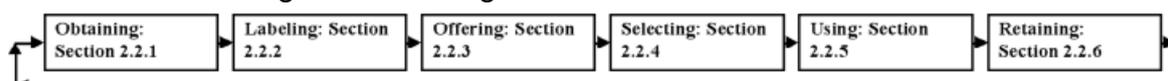
3.4.2 Ciclo de Vida do Objeto de Aprendizagem

Dalziel (2002) reforça que entre o processo de criação e armazenamento existem cinco atores diferentes envolvidos no ciclo de vida de um OA, que são:

- [1] Autoridade: responsável por prescrever os objetivos de aprendizagem e resultados;
- [2] Criador: o autor do objeto de aprendizagem e/ou responsável por submeter o mesmo para a publicação;
- [3] Organizador: responsável por projetar atividades de aprendizagem e revisar as licenças e direitos autorais e de uso;
- [4] Buscador de Informação (Infoseeker): tem o papel de buscar por recursos de acordo com os metadados fornecidos; e
- [5] Aprendiz: aquele que irá utilizar os OAs e realizar as avaliações (DALZIEL, 2002 apud CECHINEL, 2014, p.65-66).

Em relação, aos estágios do ciclo de vida, Collis e Strijker (2004) afirmam que um OA passa por seis estágios diferentes, conforme mostra a Figura 15.

Figura 15 - Estágios do ciclo de vida de um OA



Fonte: Collis; Strijker (2004, p.9)

Obtenção ou criação do OA: O material é obtido em formato digital para fácil distribuição e adaptabilidade. Também o desejo de qualidade em termos de comportamento profissional e consistência na apresentação desempenham um papel. O uso de templates fornece estrutura e pode ajudar os usuários para criar peças consistentes de material. Os modelos são uma ferramenta importante para obtenção de um novo objeto.

Etiquetamento: A forma mais arbitrária de rotulagem é apenas fornecer um nome de arquivo ou assunto para o objeto de aprendizagem. As ferramentas de marcação podem facilitar e apoiar a adição de metadados para os objetos.

Oferta: Quando um objeto de aprendizagem é obtido e rotulado, ele pode ser oferecido para seleção e uso eventual. Diferentes pessoas ou organizações podem oferecer objetos de aprendizagem, por exemplo, desenvolvedores de cursos, sociedades profissionais e instrutores, mas também fornecedores especializados na criação de material de curso.

Seleção: O processo de seleção é o quarto estágio no ciclo de vida do objeto de aprendizagem. As ferramentas podem apoiar desenvolvedores de cursos ou cursistas na seleção de material de repositórios que conter objetos de

aprendizagem. O processo de seleção também pode ser baseado nas necessidades de um desenvolvedor de curso. Outras ferramentas de suporte estão sendo desenvolvidas para ajudar os desenvolvedores de cursos a selecionar o material com base em conjuntos de critérios fornecidos.

Uso: O material pode ser usado de duas maneiras diferentes: Diretamente, o chamado uso “puro” do objeto, ou de forma adaptada. “Adaptado” significa aqui que o objeto é editado ou adaptado após a seleção para seu novo ambiente. O uso de um chamado puro objeto de aprendizagem é mais fácil do que um objeto adaptado porque a modificação da aprendizagem os objetos requerem ferramentas e habilidades específicas que podem não estar disponíveis para o desenvolvedor.

Retenção: Após ou durante o uso real de um determinado objeto de aprendizagem, esse objeto pode se tornar desatualizado e deve ser suprimido ou revisto. Decisões sobre a retenção de um objeto são influenciados por novos insights, experiências ou pesquisas de ambos, o desenvolvedor ou usuário do objeto. Novas instâncias ou versões podem ser criadas para rever o objeto original. Ferramentas de autoria podem ser usadas para revisar o aprendizado do objeto original. O controle de qualidade dos objetos de aprendizagem pode ser medido rastreando o uso. Também as ferramentas de classificação podem fornecer informações sobre a usabilidade e qualidade dos objetos de aprendizagem. As organizações precisam identificar quem é responsável pela manutenção, versionamento e controle de acesso de coleções de objetos de aprendizagem (COLLIS; STRIJKER, 2004, p. 10-15, tradução e adaptação nossa).

De acordo com os autores, cada um desses estágios se relaciona às questões, “Por quê?, O quê?, Como?, Quem? e Onde?” que devem ser feitas a partir de duas perspectivas, a humana, que lida com o porquê? e Quem?, e a perspectiva técnica que lida com o Como?, O quê? e Onde? (COLLIS; STRIJKER, 2004, p.9, tradução nossa). No Quadro 5 é possível observar estas questões levantadas pelos autores sob as duas perspectivas abordadas.

Quadro 5 - Questões sob as perspectivas humana e técnica

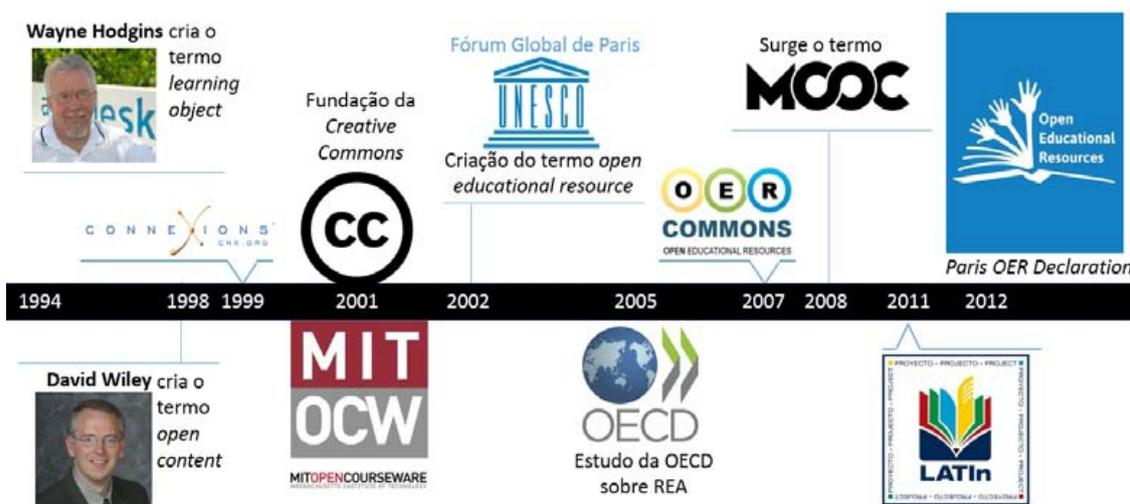
Perspectiva	Questão	Descrição
Humana	Porquê?	Quais os motivos de reúso? Por que o ser humano deve investir tempo e esforço durante os diferentes estágios do ciclo de vida do OA?
	Quem?	Quem são os envolvidos no processo de reúso? Quais funções/papéis podem ser identificados nos diferentes estágios do ciclo de vida do OA? Quais incentivos para cuidar dessas funções?
Técnica	O quê?	Que material é reusado? Qual é a granularidade e o tipo do material reusado?
	Como?	Como o material é atualmente reusado em termos de ferramentas? Que tipo de suporte técnico é possível no local e nos diferentes estágios do ciclo de vida do OA?
	Onde?	Onde se reusa em termos de sistema? Quais sistemas são avaliados para suportar o reúso e quais serviços são oferecidos pelo sistema nos diferentes estágios do ciclo de vida do OA?

Fonte: Collis e Strijker (2004, p.10, tradução nossa)

3.4.3 Recursos Educacionais Abertos - REA

No estágio de vida da oferta de um OA é importante que se atenha ao que Silveira (2014) chamou de “grau de liberdade”, pois este, segundo o autor, é um dos motivos para os baixos índices de reutilização destes recursos. Para minimizar esta questão, surgiu o movimento que se conhece hoje como REA – Recursos Educacionais Abertos (*OER – Open Educational Resources*). Os REA foram definidos pela UNESCO (2002, p. 24, tradução nossa) como “*The open provision of educational resources, enabled by information and communication technologies, for consultation, use and adaptation by a community of users for non-commercial purposes.*”, numa tradução simples, “quaisquer recursos educacionais baseados em tecnologia para consulta, utilização e adaptação por uma comunidade de usuários para fins não comerciais”.

Figura 16 - Acontecimentos ligados ao movimento REA



Fonte: Silveira (2014, p. 131)

Na Figura 16 é possível observar como a história do movimento REA se relaciona e se conecta com a história dos OAs, é possível observar que quatro anos após o termo *learning object* (objeto de aprendizagem) haver sido cunhado por Wayne Hodgins (1994), o pesquisador David Wiley (1998) traz a concepção de *open content* (conteúdo aberto), o que viria a inspirar e fundamentar todo o movimento de REA.

Ainda seguindo a linha do tempo de Silveira (2014), observa-se um grande marco para o movimento de REA que inclui os OAs, que é criação da iniciativa Creative Commons em 2001, que hoje em dia é responsável pela definição das licenças amplamente utilizadas no contexto dos REA, visto que, não há de se falar em reuso sem falar em licenças de

uso. Dez anos após o fórum global da UNESCO cunhar o termo REA, em 2012, novamente em Paris, a UNESCO emitiu a Declaração de Paris, conclamando os governos a investirem na iniciativa de REA.

Em 2005, o tema ganha importância acadêmica, política e econômica, visto que a Organisation for Economic Co-operation and Development (OECD) realiza um amplo estudo sobre o tema. Dois anos depois, em 2007, funda-se a OER Commons, portal de referência para busca de REA e, conseqüentemente, de OAs.

Em 2008, o termo MOOC (Massive Online Open Courses), surge pela primeira vez, termo que, atualmente, tem ganhado notoriedade, no que se refere aos estudos relacionados a Educação a Distância, inclusive com várias instituições inaugurando seus portais exclusivos para hospedar esses cursos abertos online que empregam, em boa parte, REA e OAs, porém, em muitos casos, nem tão abertos assim.

Não pode-se pensar em OAs como dados ligados interoperáveis na Web semântica sem o movimento REA, desta forma, Wiley (2014), propôs algumas características desejáveis, que até então era conhecido na literatura como “os 4R dos REA” e que o autor ampliou, passando a trabalhar com “os 5R dos REA” a saber:

- Reter – o direito de fazer, possuir e controlar cópias do conteúdo
- Reutilizar – o direito de usar o conteúdo de várias maneiras (por exemplo, em uma aula, em um grupo de estudo, em um site, em um vídeo)
- Revisar – o direito de adaptar, ajustar, modificar ou alterar o próprio conteúdo (por exemplo, traduzir o conteúdo para outro idioma)
- Remix – o direito de combinar o conteúdo original ou revisado com outro conteúdo aberto para criar algo novo (por exemplo, incorporar o conteúdo em um mashup)
- Redistribuir – o direito de compartilhar cópias do conteúdo original, suas revisões ou seus remixes com outras pessoas (por exemplo, dar uma cópia do conteúdo a um amigo) (WILEY, 2014, p.1, tradução nossa).

Os repositórios que guardam os OAs e que serão tratados na próxima seção, têm sido amplamente impactados pelo movimento REA.

3.4.4 Repositório de Objetos de Aprendizagem

Como visto na subseção anterior, na fase do ciclo de vida dos OAs definida por Collis e Strijker (2004) como oferta faz-se necessário que estes objetos sejam publicados em um local em que sejam facilmente encontrados e recuperados de forma que seja possível o uso e/ou reúso. Assim entram em cena os chamados Repositórios de Objetos de Aprendizagem - ROAs que de acordo com Heery e Anderson (2005) podem ser definidos

de forma simples como “coleções de objetos digitais”, onde os OAs são armazenados. Os autores salientam que o que diferencia os repositórios de outras coleções (catálogos, diretório, banco de dados, etc...) é que eles precisam ser qualificados quanto ao:

- Tipo: é um repositório institucional ou um repositório de assunto?
- Conteúdo: é de e-print, dados, objetos de aprendizagem?
- Propósito: é para preservação, acesso ou gerenciamento de dados?

Segundo os autores, os OAs ainda devem apresentar algumas características: [1] Permitir o depósito de conteúdo pelo criador, proprietário ou por um terceiro; [2] Possuir uma arquitetura que seja capaz de administrar os conteúdos e seus metadados; [3] Oferecer serviços para depósito, retirada e busca recursos, assim como controle de acesso aos recursos; [4] Serem confiáveis com bom suporte e boa administração.

Os repositórios são essenciais na economia dos OAs, pois são responsáveis pelo armazenamento e disponibilização destes em seus diversos formatos (textos, vídeos, áudios, imagens etc.). Também existem os repositórios com acesso Web onde não se faz o armazenamento dos recursos propriamente ditos, mas dos metadados (catalogação extensa) que indicam quais são os repositórios que detêm recursos sobre determinado assunto, sendo um agregador de fonte de informação no ciberespaço.

Uma definição de ROAs que vem ao encontro do presente trabalho é a de McGreal (2008), “bases de dados para armazenamento e/ou para possibilitar a interoperabilidade de OAs” e o autor os caracteriza-os da seguinte forma: [1] Nível do público-alvo: indica se o ROAs é direcionado a um nível de ensino específico ou não; [2] Granularidade do OAs: se são apenas para componentes curriculares, cursos completos, atividades ou recursos; [3] Tamanho do ROA em relação ao número de recursos armazenados; [4] Tipo: indica se no ROA existe um tipo predominante de OAs (vídeos, livros digitais, etc.) ou se estão em formatos variados; e por último, [5] Tipo do metadado utilizado para representação dos OAs: Dublin Core, IEEE-LOM, alguma taxonomia específica, ou nenhum.

Sbrogio e Valente (2020) fizeram um levantamento de alguns dos principais ROA que estão disponíveis na Web, o levantamento feito pelas autoras encontra-se na íntegra no Anexo 1, porém, dar-se-á destaque para alguns e outros serão incluídos nesta

seção por entender a importância de algumas informações destes para esta pesquisa, um pelo tipo, outro pelo conteúdo e o outro pelo propósito.

Pelo Tipo: ROA - Repositório de Recursos Educacionais Abertos para Educação Profissional e Tecnológica - PROEDU

O repositório de objetos educacionais para a educação profissional e tecnológica da Setec/MEC, chamado Proedu²², é uma ferramenta desenvolvida no âmbito da Rede e-Tec para reunir e disponibilizar recursos educacionais da educação profissional e tecnológica produzidos pelas instituições parceiras da Rede e-Tec e da Setec. (PROEDU, 2022)

Pelo Conteúdo: ROA - Banco Internacional de Objetos de Aprendizagem - BIOE

O Banco Internacional de Objetos Educacionais, doravante BIOE, é uma iniciativa da Secretaria de Estado de Educação a Distância (SEED) do Ministério da Educação (ME) em parceria com o Ministério da Ciência e Tecnologia (MCT), Rede Latinoamericana de Portais Educacionais (RELPE), Organização dos Estados Ibero-americanos (OEI) e outros, tendo, também, várias universidades como parceiras. Ele visa assessorar o professor e disponibilizar no banco recursos educacionais gratuitos em diversas mídias e idiomas (áudio, vídeo, animação/simulação, imagem, hipertexto, softwares educacionais) que atendem desde a educação básica até a superior, nas diversas áreas do conhecimento. Sua base é implementada no software livre DSpace²³. Baseados no esquema de metadados Dublin Core²⁴, os desenvolvedores incluíram campos que permitem, não apenas descrever melhor os objetos, mas apoiar na recuperação. Implementando campos obrigatórios como o campo título, ou optativos como o campo descrição, outros com vocabulários controlados e alguns de texto livre (SHINTAKU, 2014).

Pelo Propósito: ROA ARCA – Repositório Institucional da Fiocruz

²² Disponível em: <http://proedu.rnp.br/>

²³ O DSpace é um projeto inicial do Instituto de Tecnologia de Massachussets (MIT) com apoio da empresa Hewllet Packard (HP), que, atualmente é mantido pela organização DuraSpace e grande comunidade de usuários em todo o mundo. Disponível em: <https://www.dspace.com/en/pub/home.cfm>

²⁴ Dublin Core é um esquema de metadados que visa descrever objetos digitais, tais como, videos, sons, imagens, textos e sites na web. Aplicações de Dublin Core utilizam XML e o RDF. Disponível em: <https://www.dublincore.org/>.

Apesar de o ARCA não ser considerado um repositório apenas de OAs, e sim um Repositório Institucional da Fundação Oswaldo Cruz (Fiocruz) cuja função é reunir, hospedar, disponibilizar e dar visibilidade à produção intelectual da Instituição, assim como estimular a mais ampla circulação do conhecimento, fortalecendo o compromisso institucional com o livre acesso da informação em saúde, além de conferir transparência e incentivar a comunicação científica entre pesquisadores, educadores, acadêmicos, gestores, alunos de pós-graduação, bem como a sociedade civil. Este torna-se importante em vista do projeto Sudeste/RIAA (Rede Sudeste de Repositórios institucionais) capitaneado pela Fiocruz, esta rede faz parte da Rede Nacional de Repositórios coordenada pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), composta atualmente por 79 instituições de ensino e pesquisa, cujos principais objetivos são:

Criar, com o apoio dos profissionais de informação, repositórios digitais, institucionais e temáticos, observando os pressupostos do acesso aberto à informação científica; ✓ Incentivar o depósito da produção científica nos seus repositórios institucionais; ✓ Buscar o estabelecimento de uma política de ciência aberta e de funcionamento dos repositórios; ✓ Organizar ações contínuas de capacitação de usuários; ✓ Proporcionar ajuda mútua na implantação e gerenciamento de repositórios; ✓ Buscar novas instituições que possuam perfil para integrar a rede; ✓ Apoiar as ações dos grupos de trabalho criados no âmbito da rede; ✓ Promover o autoarquivamento das publicações produzidas nas Instituições nos RIs; ✓ Fortalecer a elaboração de Políticas Institucionais para os RIs nas Instituições da Rede Sudeste (QUEIROZ, 2021, p.6).

O Arca está organizado em comunidades que correspondem às unidades da Fiocruz. É mantido pelo Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT) e utiliza o software livre DSpace. Esta característica de foco no compartilhamento e incentivo aos dados abertos muito agrega a esta pesquisa.

3.4.5 Padrão LOM IEEE-2020

Conforme visto na subseção anterior, os ROA surgiram para facilitar o gerenciamento e promover o acesso e recuperação dos OAs de forma que seja possível o uso e/ou reúso. Porém, organizar as informações de forma a tornar estes OAs acessíveis e reutilizáveis traz à tona os desafios e a complexidade da gestão de grandes volumes de conteúdos educacionais digitais.

Com vistas a esta organização de OAs, os repositórios são estruturados com o auxílio de metadados. Reis e Ferneda (2016) trazem que várias iniciativas surgiram com o objetivo

de desenvolver padrões de metadados visando a reutilização e interoperabilidade de OAs nas diferentes plataformas. Porém, Duval e Verber (2018) argumentam que a abordagem do “evite os padrões não inventados aqui” fez com que se desenvolvessem vários padrões específicos, quando já havia padrões genéricos bastante apropriados já disponíveis ou em desenvolvimento, e que nos primórdios da padronização das tecnologias de aprendizagem, essa ressalva se aplicava menos, pois havia menos padrões relevantes ao redor.

Quando o IEEE LTSC iniciou o desenvolvimento do padrão LOM por volta de 1997, a Dublin Core Metadata Initiative (DCMI) também estava começando. Os autores apontam que, provavelmente, teria sido mais útil colaborar diretamente desde o início, mas nos primeiros anos, não estava claro qual iniciativa levaria a um padrão maduro e quando isso aconteceria. Assim, à medida que o DCMI e o LOM amadureceram, a colaboração se intensificou e nos últimos anos, o IEEE - Institute of Electrical and Electronics Engineers, os projetos Ariadne²⁵ (ARIADNE, 2022) e IMS (Instructional Management System)²⁶ (IMS, 2022) têm trabalhado conjuntamente na especificação de metadados objetivando a organização e documentação de cada tipo de material instrucional, entre eles, os metadados de OAs.

Assim surgiu o LOM - Learning Object Metadata (IEEE, 2020), um padrão criado pelo *Learning Technology Standards Committee (LTSC)*, um comitê criado pelo IEEE - Institute of Electrical and Electronics Engineers, cujas as especificações originais foram desenvolvidas pelo projeto Ariadne²⁷ e IMS para descrever os recursos de aprendizagem em formato XML (*extensible markup language*), este padrão facilita a descrição, busca, avaliação e uso dos OAs pelos estudantes, professores e buscadores automatizados.

O LOM é regido pela norma IEEE Std 1484.12.1™-2020 cuja primeira versão foi aprovada em 12 junho de 2002 e teve sua última atualização em 24 de setembro de 2020.

²⁵ Alliance of Remote Instructional Authoring and Distribution Networks for Europe. É uma proposta para integrar a infraestrutura de dados arqueológicos existentes em toda a Europa, para que os pesquisadores possam usar os vários conjuntos de dados distribuídos e novas tecnologias para explorar metodologias de pesquisa. Disponível em: <https://www.ariadne-eu.org/about-us/>.

²⁶ O IMS Global Learning Consortium (IMS Global/IMS) é uma organização membro altamente eficaz, sem fins lucrativos, comprometida com o avanço da tecnologia que pode escalar e melhorar a participação e o aproveitamento educacional de forma acessível. Disponível em: <https://www.imsglobal.org/aboutims.html>.

²⁷Inclusive nas páginas iniciais do LOM, o IEEE faz um agradecimento especial: “Este documento tem suas origens nos Projetos ARIADNE e IMS, sem os quais este documento não poderia ter sido criado. Este documento também se baseia no trabalho de metadados feito pelo grupo Dublin Core” (IEEE, 2020, p.6).

É interessante ressaltar que o primeiro grupo de trabalho do IEEE denominado *Learning Technology Standards Committee (LTSC)* que deu origem à primeira versão do LOM em 2002 foi presidida por Wayne Hodgins, que conforme visto na seção 3.4, cunhou a metáfora do Lego para OAs.

Logo no seu resumo, o LOM/IEEE-2020 nos apresenta um esboço do que se trata o documento: “Um esquema de dados conceitual que define a estrutura de uma instância de metadados para um objeto de aprendizagem é especificado neste padrão” (IEEE, 2020, resumo, tradução nossa).

Logo em seguida também traz uma conceituação do que, na visão da Learning Technology Standards Committee (LTSC) do IEEE, seja um OA: “qualquer entidade, digital ou não digital, que é usada para aprendizagem, educação ou treinamento” (IEEE, 2020, resumo, tradução nossa).

Continua a explanação informando que uma instância de metadados de um OA descreve as características que são relevantes para entender aquele objeto.

[...] uma instância de metadados para um objeto de aprendizagem descreve características relevantes do objeto de aprendizagem ao qual se aplica. Tais características podem ser reagrupadas em categorias gerais, ciclo de vida, meta-metadados, educacional, técnico, educacional, direitos, relação, anotação e classificação (IEEE, 2020, resumo, tradução nossa).

E finaliza informando, que a pretensão é que o padrão seja referenciado por outros padrões, onde propõe um esquema de dados conceitual que permite a diversidade linguística de modo que uma instância de um metadado de um OA possa ser utilizada por um sistema para gerenciar, localizar, avaliar ou trocar OA.

O esquema de dados conceitual definido neste padrão especifica os elementos de dados dos quais uma instância de metadados para um objeto de aprendizagem é composta e permite a diversidade linguística tanto dos objetos de aprendizagem quanto das instâncias de metadados que os descrevem. Pretende-se que este padrão seja referenciado por outros padrões que definirão as descrições de implementação do esquema de dados, de modo que uma instância de metadados para um objeto de aprendizagem possa ser usada por um sistema de tecnologia de aprendizado para gerenciar, localizar, avaliar ou trocar objetos de aprendizagem (IEEE, 2020, resumo, tradução nossa).

Os objetivos do LOM são:

1. Facilitar a busca, avaliação, aquisição e uso de objetos de aprendizagem, por exemplo, por alunos, instrutores ou processos de software automatizados.
2. Facilitar o compartilhamento e troca de objetos de aprendizagem, permitindo o desenvolvimento de catálogos e inventários, levando em

consideração a diversidade de contextos culturais e linguísticos em que os objetos de aprendizagem e seus metadados serão explorados (IEEE, 2020, p. 11, tradução nossa).

A ideia é que o LOM ajude a garantir que as ligações de Metadados de OAs levem a um alto grau de interoperabilidade semântica e como resultado, as transformações entre essas ligações sejam diretas.

O Esquema Base do LOM/IEEE-2020 consiste em nove categorias que descrevem os elementos de dados de um OA, são elas:

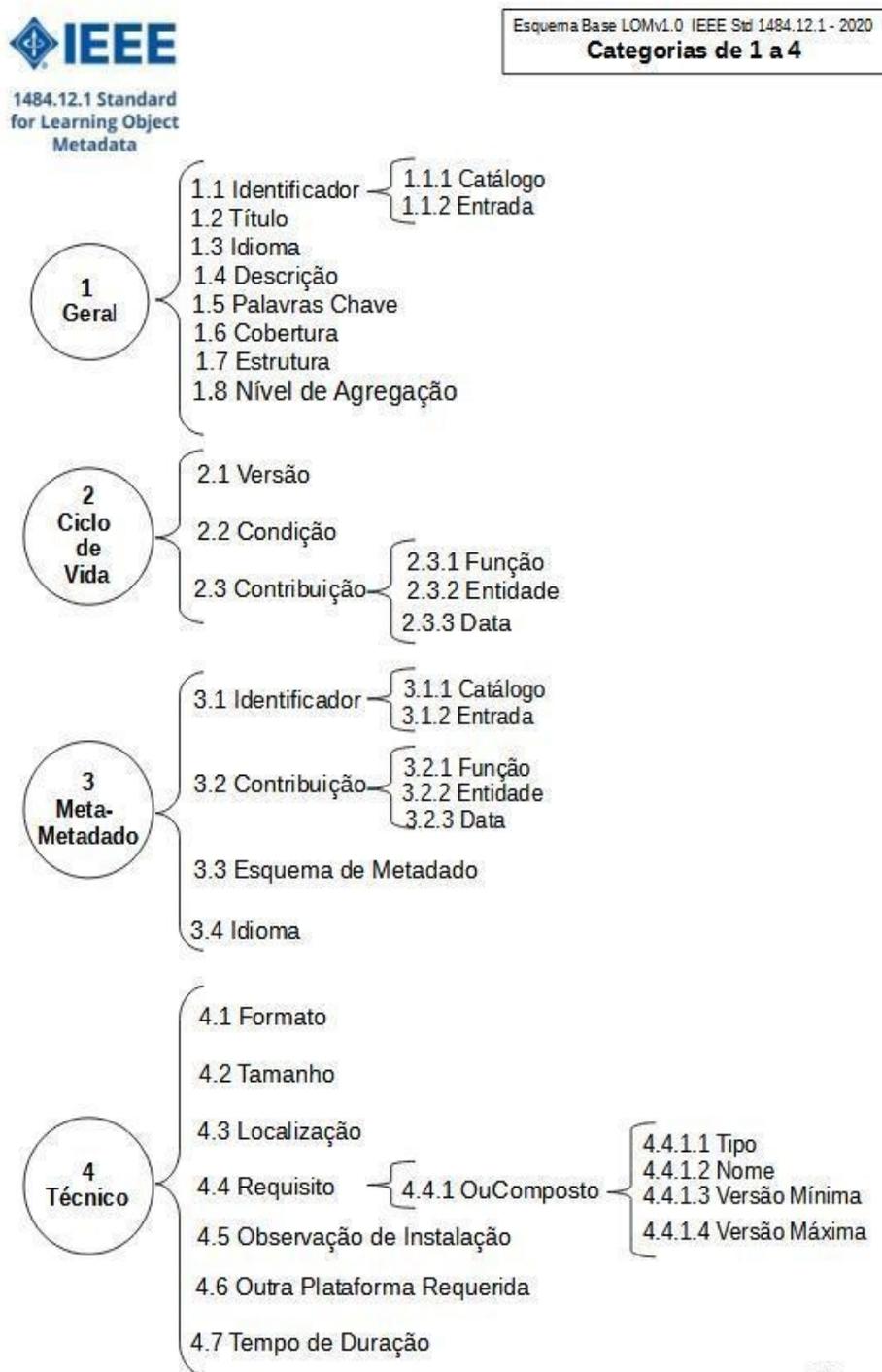
- a) **Geral:** agrupa as informações gerais que descrevem o objeto de aprendizagem como um todo.
- b) **Ciclo de Vida:** agrupa as características relacionadas ao histórico e estado atual deste objeto de aprendizagem e aquelas que afetaram este objeto de aprendizagem durante sua evolução.
- c) **Meta-Metadados:** agrupa informações sobre a própria instância de metadados (em vez do objeto de aprendizagem que a instância de metadados descreve).
- d) **Técnica:** agrupa os requisitos técnicos e as características técnicas do objeto de aprendizagem.
- e) **Educacional:** agrupa as características educativas e pedagógicas do objeto de aprendizagem.
- f) **Direitos:** agrupa os direitos de propriedade intelectual e as condições de uso do objeto de aprendizagem.
- g) **Relação:** agrupa características que definem a relação entre o objeto de aprendizagem e outros objetos de aprendizagem relacionados.
- h) **Anotação:** fornece comentários sobre o uso educacional do objeto de aprendizagem e fornece informações sobre quando e por quem os comentários foram criados.
- i) **Classificação:** descreve este objeto de aprendizagem em relação a um determinado sistema de classificação (IEE, 2020, p. 13-14, tradução nossa).

Todas essas categorias formam o esquema base do LOM, onde a categoria de “Classificação” é usada para fornecer certos tipos de extensões ao esquema base, pois qualquer sistema de classificação pode ser referenciado.

Cada uma das categorias é formada por elementos de dados que as compõem organizadas hierarquicamente, incluindo elementos de dados agregados e elementos de dados simples (como nós de folha da hierarquia)²⁸. Ao todo são 58 elementos de dados conforme mostram as Figuras 17 e 18.

²⁸ No desenvolvimento de tesouro e classificação, um método conhecido como nós folha tem sido usado no qual esquemas estendidos para subtópicos são apresentados como nós de uma estrutura de árvore em um vocabulário superior. Por outro lado, um novo vocabulário pode ser construído com base em mais de um vocabulário existente. Uma das principais tarefas dos desenvolvedores é não ser desnecessariamente redundante. Em vez disso, seu papel principal é estender a partir dos “nós” e cultivar “folhas” de vocabulário localizado (ZEN, 2019, sem página)

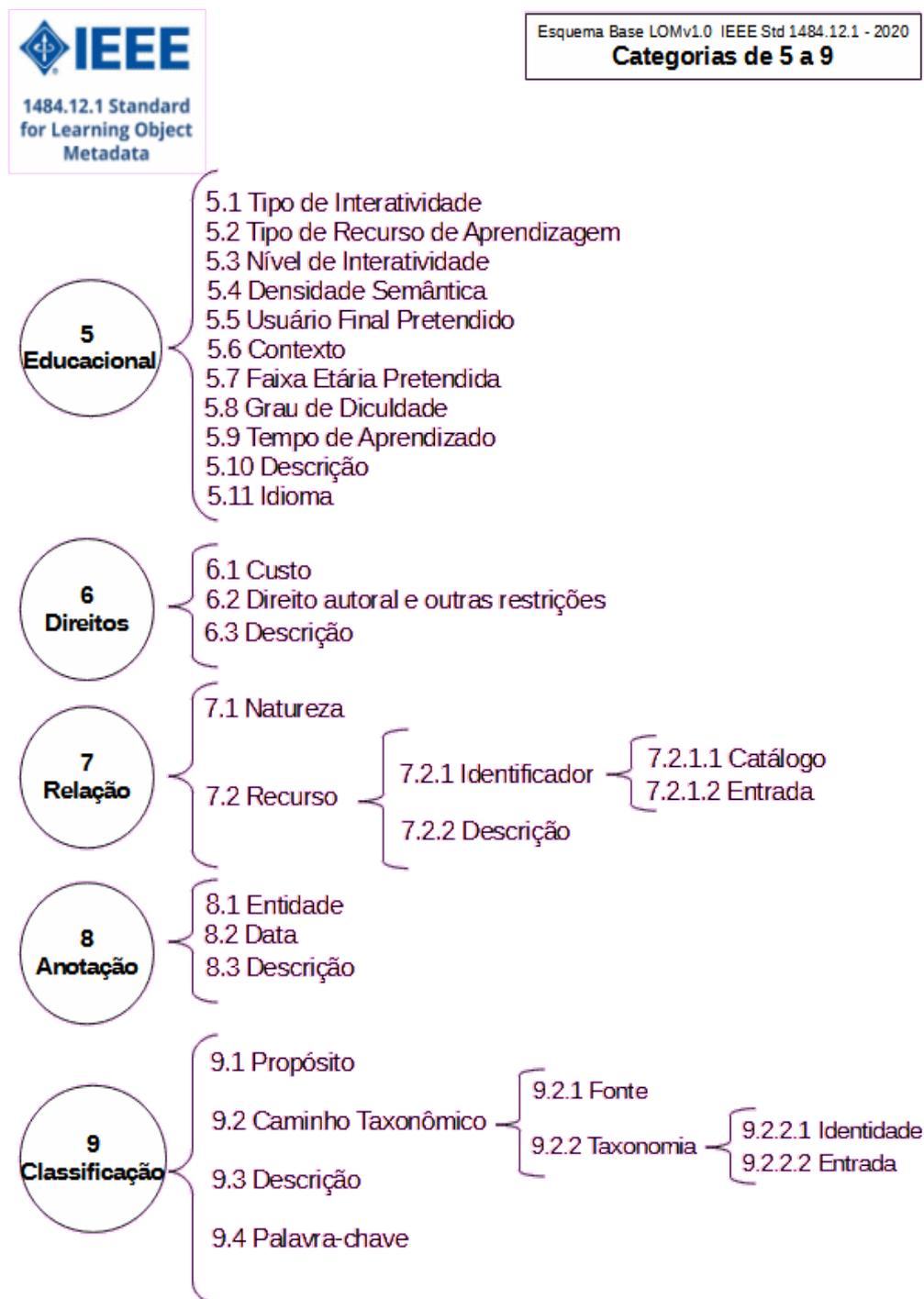
Figura 17 - Esquema Base do LOM/IEEE-2020 - Categorias de 1 a 4



1/2

Fonte: Mapa mental elaborado pela autora (2023) adaptado de IEEE (2020)

Figura 18 - Esquema Base do LOM/IEEE-2020 - Categorias de 5 a 9



As categorias do LOM foram traduzidas e organizadas em formato de mapa mental, Figuras 20 e 21, para facilitar o entendimento neste trabalho.

Os elementos de dados de cada categoria do LOM 2020 são organizados da seguinte forma:

- *Nr*: Número que define a estrutura hierárquica de cada categoria;
- *Name*: Nome do elemento de dado;
- *Explanation*: Explicação sobre cada elemento;
- *Size*: valor atribuído, tamanho máximo e mínimo de cada elemento;
- *Order*: ordem do valor do metadado, quando se aplica.
- *Value space*: um conjunto de valores permitidos para o elemento de dado, na forma de vocabulário, referência ou outro padrão de metadados;
- *Datatype*: indica o tipo do dado, com base em alguns padrões (*LangString*, *DataTime*, *Vocabulary* etc.)
- *Example*: um exemplo ilustrativo;

Uma característica interessante no padrão LOM é que todos os elementos de dados são opcionais, o padrão não define uma classificação única a respeito da obrigatoriedade, opcionalidade ou irrelevância de cada um dos metadados onde um atributo pode ser considerado opcional em um determinado contexto, e obrigatório em outro.

Essa característica tanto pode ser benéfica ou maléfica pois conforme saliente Vicari et al (2010, p.6) “[...] apesar de o IEEE LOM possuir uma grande variedade de documentação acessível e possuir elementos opcionais que possibilitam o preenchimento flexível para cada caso, a quantidade de metadados presentes no padrão pode ser tornar um desafio para os usuários iniciantes”.

Outra característica importante é que o LOM possui um mapeamento para o atender ao Dublin Core, onde faz as seguintes notas:

NOTA 1 - A Dublin Core Metadata Initiative também está desenvolvendo qualificadores de elementos de dados para refinar ainda mais a semântica dos elementos de dados Dublin Core.¹⁰ Um refinamento adicional do mapeamento na Tabela A.1 pode ser baseado nesses qualificadores.

NOTA 2 - O grupo de trabalho LOM está comprometido em trabalhar com a Dublin Core Metadata Initiative (DCMI) para desenvolver metadados interoperáveis,

conforme descrito no Memorando de Entendimento entre o IEEE LTSC LOM WG e o DCMI.11 (IEEE, 2020, p.48).

Este trabalho cooperativo aumenta a possibilidade de interoperabilidade ao trabalhar com o padrão LOM.

3.5 WEB SEMÂNTICA

Em 2001, Berners-Lee, Hendler e Lassila publicaram o primeiro artigo sobre a Web semântica, este texto tornou-se praticamente referência obrigatória sobre o tema desde então, nele, os autores propõem que a Web semântica trará estrutura ao conteúdo significativo das páginas da Web, mas ela não é uma Web separada, mas uma extensão da atual, onde a informação recebe um significado bem definido, permitindo assim que computadores e pessoas trabalhem em cooperação (NUNES, MACULAN, ALMEIDA, 2020). Os autores já tinham sinalizado que os primeiros passos para a Web semântica na estrutura da Web existente já estavam em andamento e em breve as novas tecnologias dariam início a novas funcionalidades significativas à medida que as máquinas se tornarem capazes de processar e "entender" os dados, para que a Web semântica funcionasse. Eles também alertaram que os computadores deveriam ter acesso a coleções estruturadas de informações e conjuntos de regras de inferência que pudessem usar para conduzir o raciocínio automatizado. É neste processo que entra a "Representação do Conhecimento" pois ela contém as sementes de aplicações importantes, mas para realizar todo o seu potencial deve ser vinculado a um único sistema global.

Hoje, depois duas décadas, percebe-se que os autores, Berners-Lee, Hendler e Lassila, estavam corretos quanto a indicação dos novos recursos e funcionalidades da Web, mas quanto ao tempo, previsto como "em breve", não aconteceu, mas vem ainda acontecendo em processo lento. Ainda em 2006, Shadbolt, Hall e Berners-Lee (2006) escreveram o texto *The Semantic Web Revisited (A Web semântica revisitada)*, onde relataram cinco anos após o primeiro trabalho de Berners-Lee, Hendler e Lassila em 2001, que as ferramentas da Web semântica se tornaram realidade, ao mesmo tempo que indicavam a necessidade de haver uma maior integração entre essas tecnologias.

Esperamos que os desenvolvimentos, metodologias, desafios e técnicas que discutimos aqui não apenas originem uma Web semântica, mas também contribuam para uma nova Ciência da Web - uma ciência que busca desenvolver, implantar e compreender sistemas e sistemas de informação distribuídos, de seres humanos e máquinas, operando em escala global. Inteligência Artificial - IA será

uma das disciplinas contribuintes. IA já nos forneceu métodos de programação funcional e lógica, formas de compreender sistemas distribuídos, ferramentas de detecção de padrões e de mineração de dados, abordagens de inferência, engenharia ontológica e representação de conhecimento. Tudo isso é fundamental para uma agenda da Web Science e para a realização da Web semântica (SHADBOLT; HALL; BERNERS-LEE, 2006, p. 101, tradução nossa).

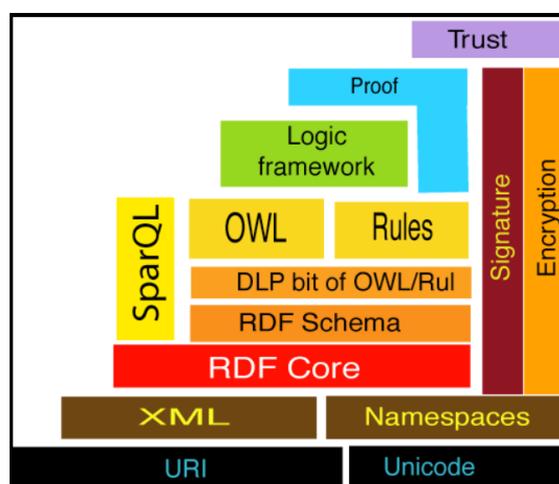
Alguns autores, como Santarem Segundo (2015), Santarem Segundo e Coneglian (2016) e Coneglian (2017) sinalizam que o *linked data* se mostra “[...] como a melhor forma de materialização dos conceitos e tecnologias da Web semântica [...]” e será tratado na subseção 3.5.2. Antes, a subseção 3.5.1 aborda fundamentos da Web semântica e, depois, a subseção 3.5.3 trata da interoperabilidade.

3.5.1 Fundamentos

De acordo com Lima e Carvalho (2004), os sistemas tradicionais de representação de conhecimento geralmente são centralizados, exigindo que todos compartilhem exatamente as mesmas definições e conceitos, o que, segundo os autores, pode ser um problema, tendo em vista o volume de dados da Web.

O desafio da Web semântica, portanto, é fornecer uma linguagem que expresse dados e regras para raciocinar sobre os dados e que permita que regras de qualquer sistema de representação de conhecimento existente sejam exportadas para a Web. Assim, propõem que para disponibilizar dados numa estrutura semântica é necessário pensar em partes do modelo, numa estrutura de camadas que apresenta a Web semântica como proposto por Berners-Lee (2004).

Figura 19 - Camadas da Web semântica



Fonte: Berners-Lee (2005)

Os princípios da Web semântica mostrados na Figura 19 são implementados em camadas de tecnologias e padrões disponíveis na Web e são explicados por Lima e Carvalho (2004), como segue:

Unicode e URI: estabelecem um conjunto de caracteres internacionais e fornece meios para a identificação de objetos na Web semântica.

XML: com definição de namespace e schema, estabelece que se pode integrar definições da Web semântica com outros padrões baseados em XML.

RDF + RDF Schema: possibilita estabelecer declarações sobre objetos com URIs e definir vocabulários que podem ser atribuídos por URIs. Esta é a camada onde se tem a capacidade de oferecer tipos para recursos e links.

Ontologia: suporta a evolução de vocabulários assim como pode definir relações entre conceitos diferentes.

Assinatura Digital: detecta as alterações em documentos. Esta é a camada que está sendo progressivamente padronizada por grupos de trabalho do W3C (LIMA; CARVALHO, 2004, p. 4).

Nunes et al (2020) explicam que os objetivos das camadas superiores Lógica, Prova e Validação são escritas de regras, onde:

A camada da Lógica: possibilita a escrita de regras.

A camada da Prova: executa estas regras e avalia, juntamente com a camada da Validação, os mecanismos que permitem às aplicações confiar ou não nas provas realizadas (LIMA; CARVALHO, p. 4).

Neste modelo destacam-se duas tecnologias importantes para o desenvolvimento da Web semântica, o eXtensible Markup Language (XML)²⁹ e o Resource Description Framework (RDF). O XML permite que todos criem suas próprias *tags* - (rótulos) que marcam páginas da Web, onde programas utilizam dessas tags, ou seja, o XML permite que os usuários adicionem estruturas arbitrárias a seus documentos, mas não diz nada sobre o que as estruturas significam. Nesse ponto, entra o RDF para dar significado, pois ele codifica em conjuntos de triplas, onde cada tripla é entendida como o sujeito, verbo e objeto de uma frase. Essas triplas podem ser escritas usando tags XML. Quando escrito em RDF, um documento afirma que uma coisa (pessoas, páginas da Web ou qualquer outra coisa) têm propriedades (como "é irmã de", "é o autor de") com determinados valores (outra pessoa, outra página da Web). Essa estrutura acaba sendo uma maneira natural de descrever a grande maioria dos dados processados pelas máquinas. O assunto e o objeto são identificados por um identificador universal de recursos (URI), exatamente como usado em um link em uma página da Web. (URLs, *Uniform Resource Locators*, são os documentos mais semânticos e o tipo mais comum de URI). (BERNERS-LEE; HENDLER; LASSILA, 2001)

²⁹ A XML (eXtensible Markup Language, ou Linguagem de Marcação Estendida) é um subconjunto da SGML (Standard Generalized Markup Language, ou Linguagem de Marcação Padrão Generalizada) que permite que uma marcação específica seja criada para especificar idéias e compartilhá-las na rede. Disponível em: <https://www.w3.org/XML/>

Mas é preciso deixar claro que a Web não “fala” e não “entende” a linguagem humana, por isso, o RDF usa URIs para codificar essas informações em um documento, os URIs garantem que os conceitos não sejam apenas palavras em um documento, mas estejam vinculados a uma definição exclusiva que todos podem encontrar na Web. (BERNERS-LEE;HENDLER;LASSILA, 2001)

3.5.2 Dados Ligados

Em 2006 Berners-Lee publicou um conjunto de princípios para publicação de dados usando as tecnologias da Web semântica, que chamou de *Linked Data* ou Dados Ligados. Bizer e Berners-Lee (2011) postulam que o *Linked Data* ou Dados ligados é um subcampo da Web semântica que aborda aspectos de boas práticas de como publicar e vincular conjuntos de dados na Web com a utilização do RDF para representar dados e codificá-los em triplas (sujeito, predicado e objeto) de forma a fornecer uma representação minimalista do conhecimento na Web.

Berners-Lee (2006) não se refere às boas práticas como regras, mas como um conjunto de expectativas de comportamento e que não segui-las faz com que se perca a oportunidade de tornar os dados interconectados. As regras ou boas práticas para publicação de dados são:

[1] Use URIs como nomes para coisas: bastante compreendida pela maioria das pessoas que fazem tecnologia da Web semântica. Se ele não usa o conjunto de símbolos URI universal, não o chamamos de Web semântica.

[2] Use URIs HTTP para que as pessoas possam pesquisar esses nomes: O único desvio tem sido, desde o início da web, uma tendência constante para as pessoas inventarem novos esquemas de URI, às vezes, isso tem a ver com não entender que URIs HTTP são nomes (não endereços) e que a pesquisa de nomes HTTP é um conjunto de padrões complexo, poderoso e em evolução.

[3] Quando alguém procura um URI, forneça informações úteis, usando os padrões (RDF*, SPARQL): bem seguida pela maioria das ontologias, mas, por algum motivo, não por alguns conjuntos de dados importantes. Pode-se, em geral, pesquisar as propriedades e classes encontradas nos dados e obter informações das ontologias RDF, RDFS e OWL, incluindo os relacionamentos entre os termos na ontologia.

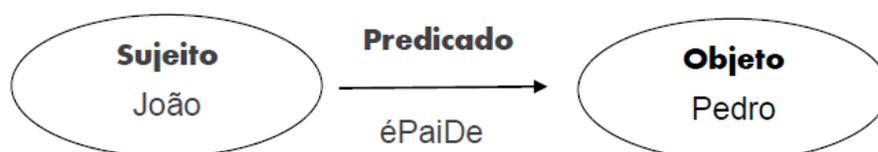
[4] Inclua links para outros URIs. para que possam descobrir mais coisas: para fazer links em outros lugares, é conectar os dados que temos em uma web, uma web séria e ilimitada na qual se pode encontrar todo tipo de coisa, assim como na web de hipertexto que conseguimos construir (BERNERS-LEE, 2006, p.1, Tradução nossa).

Nhacuongue, Rozsa e Dutra (2018) ressaltam que essas tecnologias citadas por Berners-Lee (URIs, RDF e SPARQL) viabilizam a publicação de dados na Web, onde as URIs servem como identificadores únicos dos recursos e podem ser utilizadas para acessar informações sobre um determinado recurso. Já o RDF é utilizado para descrever o

recurso e, por fim, o SPARQL é a ferramenta utilizada para a consulta do recurso no *dataset* (conjunto de dados). No contexto dos dados ligados, para que os recursos possam ser referenciados, de forma que se possa descrevê-los ou referenciá-los é necessário um método de identificação que não seja ambíguo, e são as URIs que trazem essa natureza de unicidade a esses recursos, sejam eles virtuais (páginas Web, vocabulários etc...) ou físicos (pessoas, lugares etc...).

Uma vez identificado de forma inequívoca o recurso, por meio do RDF é possível fazer a descrição destes recursos por meio de triplas.

Figura 20 - Representação em Triplas



Fonte: Nhacuongue, Rozsa e Dutra (2018, p. 25)

Conforme mostra a Figura 20, as triplas são compostas por três elementos: sujeito, predicado e objeto, onde o sujeito é o recurso do qual se fala, o predicado é o que se fala referente ao sujeito (uma propriedade do mesmo) e o objeto é o valor da propriedade, sendo que o objeto pode ser um valor literal (número, texto, data, etc.) ou a referência para outro recurso.

De acordo com Segundo (2014), uma base de dados que se destaca no processo de Dados Ligados é a DBpedia. A DBpedia é uma interface orientada a dados, constituída a partir de um esforço da comunidade para extrair informação estruturada da Wikipédia e tornar esta informação disponível na Web, na época do da escrita do referido artigo, a DBpedia descrevia

[...] mais de 4 milhão de itens, das quais 3,22 milhões são classificados de forma consistentes por ontologias, incluindo 832 mil pessoas, 63 mil lugares, além de músicas, filmes, games, espécies, doenças entre outras informações. Os dados da DBpedia apresentam estes itens em até 119 idiomas diferentes. A base de conhecimento DBpedia, em todos idiomas, consiste em aproximadamente 2,5 bilhões de declarações RDF, dos quais 470 milhões foram extraídos da edição Inglês da Wikipedia (SANTARÉM SEGUNDO, 2014, p. 3869).

Além DBpedia destacam-se ainda várias outras base de dados, como o *Geoname*³⁰, com mais de 10 milhões de nomes geográficos e composto por mais de 8 milhões de recursos exclusivos, entre outros:

Há ainda uma grande diversidade de bases de dados, que vão desde informações sobre Ciências da Vida (GeneID, PubMed, Geo Species, Gene Ontology, etc), Dados Geográficos (Aeroportos, Earth, Linked GeoData, etc), Dados de Uso Geral (Slideshare, Semantic Tweet, Delicious, Flickr, etc), Mídia (BBC, Music Brainz, New York Times, Last.FM, etc), Publicações (IEEE, ePrints, CiteSEER, theses.fr, etc) e Dados Governamentais (Patentsdata.gov.uk, researchdata.gov.uk, transportdata.gov.uk, etc) (SANTARÉM SEGUNDO, 2014, p. 3870).

Porém, de acordo com o autor, para que se desfrute de todo esse conteúdo informacional disponível nesses *datasets*³¹ é necessário transformar todo esse conjunto de informações em conhecimento útil e aplicável, de forma a mudar positivamente a vida das pessoas, conforme postulou Berners-Lee com a Web semântica. Ou seja, é necessário que haja a recuperação dessas informações, é aí que entra a consulta nestes bancos de dados.

Uma vez feita a descrição do objeto em triplas, é possível fazer uma consulta com o emprego do SPARQL que é uma linguagem de consulta e manipulação dos dados armazenados. Essa consulta possui os seguintes elementos, de acordo com (Santarém Segundo, 2014):

uma linguagem de consulta para RDF; uma especificação que define uma extensão do SPARQL Query Language para executar consultas distribuídas em diferentes terminais SPARQL; uma especificação que define a semântica de consultas SPARQL sob regimes de vinculação, como RDF Schema, OWL, ou RIF; um protocolo que define os meios para a transmissão de consultas SPARQL arbitrárias e solicitações de atualização para um serviço de SPARQL; uma especificação que define um método busca e descoberta e um vocabulário para descrever serviços SPARQL e um conjunto de testes, para avaliação da especificação SPARQL 1.1 (SANTARÉM SEGUNDO, 2014, p. 3870 - 3871)

Segundo Tim Berners-Lee, não dá para utilizar todo o potencial da Web semântica sem o SPARQL, para o autor isso seria o mesmo que tentar utilizar um banco de dados relacional sem usar a linguagem *SQL*³²(*Structured Query Language*).

³⁰ Disponível em <https://www.geonames.org/>.

³¹ Os datasets são bases de dados específicas, utilizadas para treinamentos de algoritmos de Inteligência Artificial (IA) ou para outros tipos de projetos de Data Science. Eles são geralmente estruturados em um formato de tabela, com linhas e colunas organizadas e preenchidas com informações claras e objetivas sobre sua finalidade. (DIGITALHOUSE, 2022) Disponível em: <https://www.digitalhouse.com/br/blog/dataset/>

³² SQL é uma linguagem que permite comunicar com banco de dados com o objetivo de manusear os dados que eles armazenam. (<https://harve.com.br/blog/analise-de-dados/o-que-e-sql/>)

De acordo com Santarém Segundo, entender como funciona o processo de consulta do SPARQL é importante para dominar o seu funcionamento, assim, uma consulta SPARQL é composta, em ordem, por:

- [1] Declarações de prefixos, para abreviar URIs • Definição do conjunto de dados, informando quais grafo(s) RDF estão sendo consultados
- [2] A cláusula de resultado, identificando que informação deve ser retornada a partir da consulta
- [3] O padrão de consulta, especificando o que consultar dentro do conjunto de dados
- [4] Modificadores de consulta, limites, ordenação, e outros que podem modificar o resultado final (SANTARÉM SEGUNDO, 2014, p. 3872)

Segundo o autor, o *Linked Data* é a principal referência quando se fala de Web semântica, o que tem motivado as organizações, instituições e setores a disponibilizarem os dados ligados de forma aberta. E o SPARQL é a base para a recuperação da informação nesses ambientes tratados semanticamente, porém, para que seja amplamente utilizado é preciso que haja o domínio de suas funcionalidades.

Só que nem sempre dois bancos de dados utilizarão os mesmos identificadores para um mesmo conceito, é neste momento que entra em cena o terceiro elo que forma a Web semântica que são as Ontologias, descrita na seção 3.1.1, a conceituação utilizada de Studer, Benjamins e Fensel (1998, p 184), de “uma especificação formal e explícita de uma conceituação compartilhada”, onde essa “conceituação compartilhada” vem ao encontro da necessidade que um programa que deseja comparar ou combinar informações entre os dois bancos de dados tem, pois precisa saber que esses dois termos estão sendo usados para significar a mesma coisa.

De acordo com Santarém Segundo (2014), há um grupo de vocabulários que são utilizados em larga escala nas principais ontologias conhecidas e também em grande parte dos exemplos de publicação de dados em formato semântico disponíveis na Internet. Dentre eles o autor destaca:

- Friend-of-a-Friend (FOAF) , o vocabulário para descrever pessoas.
- Dublin Core (DC) define os atributos gerais de metadados, baseado no padrão de metadados Dublin Core.
- Comunidades Online Semanticamente Interligados (SIOC) , vocabulário para representar as comunidades online.
- Descrição de um projeto (DOAP) , vocabulário para descrever projetos.
- Simple Sistema de Organização do Conhecimento (SKOS) , vocabulário para representar taxonomias, vocabulários controlados e conhecimento pouco estruturado.
- Music Ontology, prevê condições para descrever artistas, álbuns e faixas.
- Creative Commons (CC) , o vocabulário para descrever termos de licença (SANTARÉM SEGUNDO, 2014, p.3867-3868)

Normalmente as ontologias para Web ainda trazem consigo uma taxonomia que conforme visto na seção 3.1.2 é uma organização de representação por classes de conceitos, por exemplo, coisas e seus tipos, processos e seus tipos, etc, de forma a não priorizar nenhum deles, e se baseia no princípio de categorias conceituais representando os diversos aspectos (fenômenos, objetos etc) que ocorrem naquele campo de saber. E também trazem um conjunto de regras de inferência, onde uma ontologia pode expressar a regra "Se um código de Cidade está associado a um código de Estado e um endereço usa esse código de Cidade, então, **infere-se** que esse endereço tem o código de Estado associado" (BERNERS-LEE;HENDLER;LASSILA, 2001, grifo nosso).

Ainda segundo Berners-Lee, Hendler e Lassila (2001),

[...] o verdadeiro poder da Web semântica será percebido quando as pessoas criarem muitos programas que coletam conteúdo da Web de diversas fontes, processam as informações e trocam os resultados com outros programas. A eficácia de tais agentes de software aumentará exponencialmente à medida que mais conteúdo da Web legível por máquina e serviços automatizados (incluindo outros agentes) se tornarem disponíveis. (BERNERS-LEE;HENDLER;LASSILA, 2001, p.3, tradução nossa)

Em 2017 um grupo de trabalho da W3C publicou um documento com uma série de recomendações para publicação de dados na Web chamado “Melhores práticas para dados na Web” (Data on the Web Best Practices), este documento tem por objetivo descrever procedimentos que possam suportar um ecossistema de publicação de dados que seja autossustentável, onde os dados possam ser descobertos e compreensíveis por humanos e por máquinas. Nele são propostas 35 práticas que segundo seus proponentes podem trazer 8 benefícios aos dados que são: reuso, acesso, conexão, descoberta, processamento, confiança, interoperabilidade e compreensão. As primeiras boas práticas da lista são: “Prática recomendada 1 : forneça metadados”; “Prática recomendada 2 : forneça metadados descritivos”; “Prática recomendada 3: forneça metadados estruturais” (LÓSCIO; BURLE; CALEGARI, 2017).

É possível ver pelo documento proposto pela W3C que os metadados para o contexto da Web semântica são muito importantes, assim, pensando na “materialização da Web semântica” por meio de iniciativas de linked data como aspectos de boas práticas de como publicar e vincular conjuntos de dados na Web.

3.5.3 Interoperabilidade

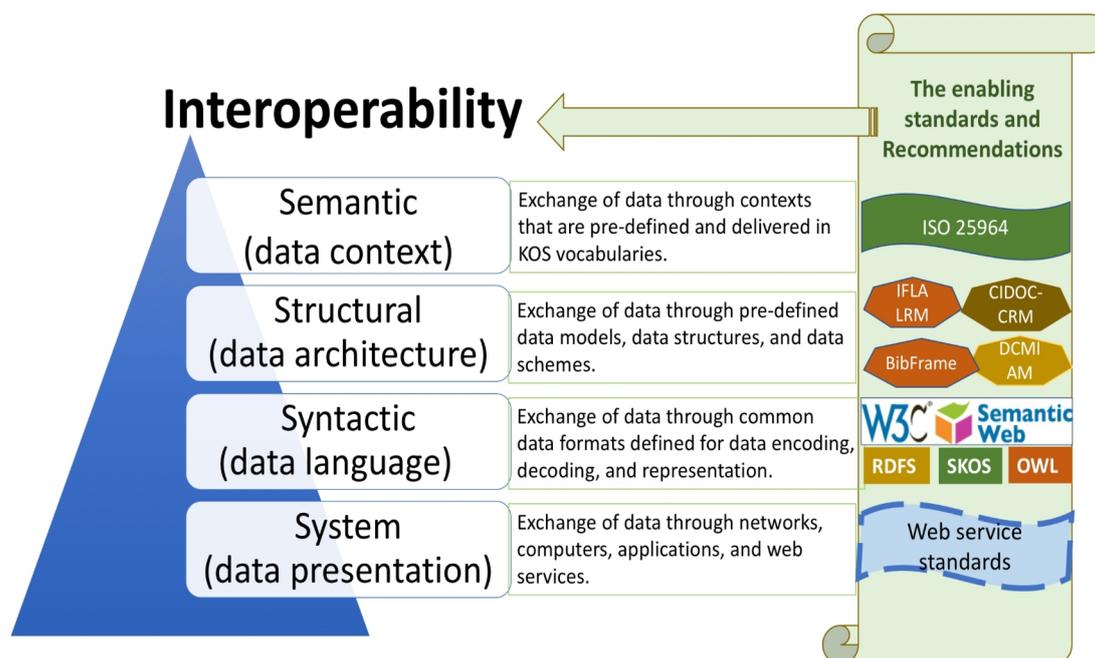
Segundo Berners-Lee, Hendler e Lassila (2001), a Web semântica promove sinergia onde mesmo agentes que não foram expressamente projetados para trabalhar juntos podem transferir dados entre si quando esses dados já vêm trabalhados semanticamente. É nesse processo de sinergia que ocorre a interoperabilidade que é definida por Zeng (2019, tradução nossa) como a “[...] capacidade de dois ou mais sistemas ou componentes de trocar informações e dados, e usar as informações e dados trocados sem esforço especial por qualquer sistema, ou sem qualquer manipulação especial”. De acordo com a autora, desde meados da década de 1980 alguns autores estabeleceram quatro tipos de heterogeneidade correspondentes a quatro tipos de possíveis problemas de interoperabilidade:

- Sistema: incompatibilidades entre hardware e sistemas operacionais.
- Sintática: diferenças nas codificações e na representação.
- Estruturais: variações em modelos de dados, estruturas de dados e esquemas.
- Semântica: inconsistências na terminologia e significados. (ZENG 2019, tradução nossa)

Porém em domínios relacionados aos KOS para a interoperabilidade normalmente são destacados em três desses tipos apresentados: sintático, estrutural e semântico. De acordo com Zeng (2019, tradução nossa),

Sem interoperabilidade sintática, dados e informações não podem ser tratados adequadamente em relação a formatos, codificações, propriedades, valores e tipos de dados; e, portanto, eles não podem ser fundidos nem trocados. Sem interoperabilidade semântica, o significado da linguagem, terminologia e valores de metadados usados não podem ser negociados ou compreendidos corretamente. Graus variados de expressividade semântica podem ser combinados com diferentes tipos: baixo em interoperabilidade sintática, médio em interoperabilidade estrutural e alto/muito alto em interoperabilidade semântica.

Figura 21 - Padrões e recomendações que tratam de problemas de interoperabilidade



Fonte: Zeng (2019)

Observando a Figura 21 a partir da base, segundo a autora, a interoperabilidade na camada do sistema aborda questões sobre incompatibilidades entre hardware e sistemas operacionais, para a troca técnica de dados por meio de redes, computadores, aplicativos e serviços da web. Exemplos citados pela autora para esta camada são o RDA e “Melhores práticas de dados na Web” da W3C³³ que apresenta as melhores práticas relacionadas à publicação e uso de dados na Web.

Na camada sintática, os problemas que afetam diretamente qualquer esforço de interoperabilidade são as diferenças na codificação, decodificação e representação de dados. Entre os padrões W3C amplamente aplicados em vocabulários KOS e trocas de dados, na Web semântica, a autora cita, o RDF e o OWL³⁴. Na camada estrutural, para permitir a troca de dados por meio de estruturas pré-definidas, modelos conceituais foram estabelecidos pelas comunidades LAM (biblioteca, arquivo e museu) na era digital, neste processo os modelos conceituais são independentes de qualquer sintaxe de codificação e

³³Documento produzido por um grupo de trabalho do W3C, publicado como recomendação W3C a partir de 31 de janeiro de 2017, chamado “Melhores práticas para dados na Web” (Data on the Web Best Practices), (resumidas na Seção 11 desta Recomendação do W3C). Ressalta também que, para promover a interoperabilidade entre conjuntos de dados, é importante adotar vocabulários e padrões de dados. Disponível em: <https://www.w3.org/TR/dwbp/>

³⁴Web Ontology Language (OWL), uma linguagem da Web semântica projetada para representar conhecimento rico e complexo sobre coisas, grupos de coisas e relações entre coisas. Os documentos OWL são conhecidos como ontologias (<https://www.w3.org/OWL/>).

sistemas de aplicativos específicos são os mais bem vistos, entre eles, a autora cita o IFLA-LRM, Dublin Core e o BibFrame. E na camada semântica, a interoperabilidade/integração semântica é basicamente impulsionada pela comunicação de propósito coerente de acordo com o contexto/domínio que inclui a interoperabilidade de tesouros com esquemas de classificação, taxonomias, esquemas de cabeçalho de assunto, ontologias, terminologias, listas de autoridade de nomes e anéis de sinônimos. A especificação cobre os detalhes dos recursos e funções de tesouros e outros tipos comuns de KOS com base na ISO 25964-2, citada por Zeng, conforme mostra a Figura 22.

Figura 22 - Cobertura das recomendações da ISO 25964-2 para interoperabilidade entre tesouros e outros tipos de KOS

Coverage of ISO 25964-2's recommendations for interoperability

Between thesauri and:	classification schemes [used by libraries (whether physical or electronic) or bibliographic databases]
	classification schemes used for records management (within business context, e.g., functional classification, file plan, business scheme)
	taxonomies (in networked environment, for organizing web pages and portals; a category label usually is not accompanied by notation.)
	subject heading schemes
	ontologies
	terminologies (E.g., glossaries and term bases)
	name authority lists
	synonym rings

Fonte: Zeng (2019)

De acordo com a autora, as instâncias individuais do KOS são chamadas de vocabulários KOS (para diferenciar dos vocabulários de metadados). Usando a terminologia das comunidades *Linked Open Data* (LOD), os vocabulários KOS são usados como vocabulários de valor (que se distinguem dos vocabulários de propriedade como conjuntos de elementos de metadados). Este termo refere-se ao seu uso nos modelos baseados em RDF.

Zeng (2019) ainda versa sobre os desafios especiais enfrentados pelos projetos para alcançar interoperabilidade entre os vocabulários KOS existentes, entre eles:

- **igualdade linguística** deve ser levada em consideração na análise e eventual seleção de uma solução para um problema específico;
- **diferentes culturas** há quatro requisitos básicos que devem ser harmonizados no trabalho terminológico: conceitos, sistemas conceituais, definições e termos;
- **microestruturas e macroestruturas** representam diferentes domínios de assunto ou têm escopo e abrangência diferentes; possuem diferenças semânticas causadas por variações na estruturação conceitual; seus graus de especificidade e uso da terminologia variam; e as características sintáticas (como a ordem das palavras dos termos e o uso de títulos invertidos) também são diferentes;
- **tipos de equivalências** também foram definidos por vários manuais e padrões importantes, identificados como: equivalência exata , equivalência inexata(ou quase equivalência), equivalência parcial , equivalência simples a múltipla e não equivalências;
- **requisitos e processos complexos** para correspondência de termos, que muitas vezes são imprecisos, podem ter um impacto significativo em vários aspectos do mapeamento de vocabulário: estrutura de navegação, exibição, profundidade, classes não tópicas e o equilíbrio entre consistência, precisão e usabilidade;
- muitos processos de mapeamento feitos no **nível sintático** (palavra, frase e contexto), **e não no nível semântico** . Os problemas de mapeamento incorreto de homógrafos para conceitos pertencentes a diferentes domínios podem ser encontrados nos serviços de mapeamento (ZENG, 2019, sem página, tradução nossa).

De acordo com a autora, a nova inteligência artificial com aprendizado de máquina apresenta grande potencial para reduzir todos esses conflitos e melhorar a interoperabilidade em todas as camadas, especialmente a interoperabilidade semântica.

E essa linguagem lógica unificadora abrirá caminho para o conhecimento e o funcionamento da humanidade para uma análise cheia de significado que será dada por agentes de software, fornecendo uma nova classe de ferramentas pelas quais pode-se viver, trabalhar e aprender juntos conforme preconizou Berners-Lee.

3.6 CONSIDERAÇÕES FINAIS DA FUNDAMENTAÇÃO TEÓRICA

Ao percorrer o caminho teórico conceitual dos elementos desta pesquisa é possível verificar várias conexões entre os termos aqui abordados. Ao propor a organização do conhecimento aplicada ao OA, foram tratados vários conceitos relacionados à Ciência da Informação, como o processo de representação descritiva e temática por meio da Catalogação que avançou significativamente com o uso da Tecnologias Digitais de Informação e Comunicação - TDICs e a utilização de vocabulários controlados, ontologias e taxonomias para armazenamento e recuperação da informação no que implica a utilização de Sistemas de Organização do Conhecimento. É possível traçar um paralelo dos princípios FAIR (*Findable, Accessible, Interoperable and Reusable*) com as características técnicas para os OAs propostos por Braga e Meneses (2014) conforme

mostra o Quadro 6, o que mostra como os princípios se aderem ao processo de representação destes objetos.

Quadro 6 - Paralelo entre os princípios FAIR e as características técnicas dos OAs

Princípios FAIR	Característica Técnica OA	Descrição da Característica
Encontrável	Disponibilidade	indica se o objeto está disponível para ser utilizado.
	Durabilidade	indica se o OA se mantém intacto quando o repositório em que ele está armazenado muda ou sofre problemas técnicos.
Acessível	Acessibilidade	indica se o objeto pode ser acessado por diferentes tipos de usuários (ex: idosos, deficientes visuais etc), em diferentes lugares (ex.: lugares com acesso a Internet, lugares sem acesso a Internet etc.) e por diferentes tipos de dispositivos (ex.: computadores, celulares, tablets etc.).
	Usabilidade	indica a facilidade de utilização dos OAs por alunos e professores.
Interoperável	Facilidade de instalação	indica se o OA pode ser facilmente instalado caso ele exija esse recurso.
	Portabilidade	indica se o OA pode ser transferido (ou instalado) para diferentes ambientes, como, por exemplo, diferentes tipos de AVAs ou sistemas operacionais.
	Interoperabilidade	medida de esforço necessário para que os dados dos OAs possam ser integrados a vários sistemas.
Reusável	Reusabilidade	indica as possibilidades de reutilizar os OAs em diferentes contextos ou aplicações. Essa é a principal característica do OA e pode ser influenciada por todas as demais.
	Manutenibilidade	é a medida de esforço necessária para alterações do OA.
	Granularidade	de maneira geral, a palavra granularidade origina-se da palavra grão, sendo que quanto maior o número de grãos de um sistema maior a sua granularidade. Trazendo esse conceito para o âmbito dos objetos de aprendizagem, a granularidade é a extensão à qual um OA é composto por componentes menores e reutilizáveis.
	Confiabilidade	indica que o OA não possui defeitos técnicos ou problemas no conteúdo pedagógico.
	Agregação	indica se os componentes do OA (grãos) podem ser agrupados em conjuntos maiores de conteúdos como, por exemplo, as estruturas tradicionais de um curso.

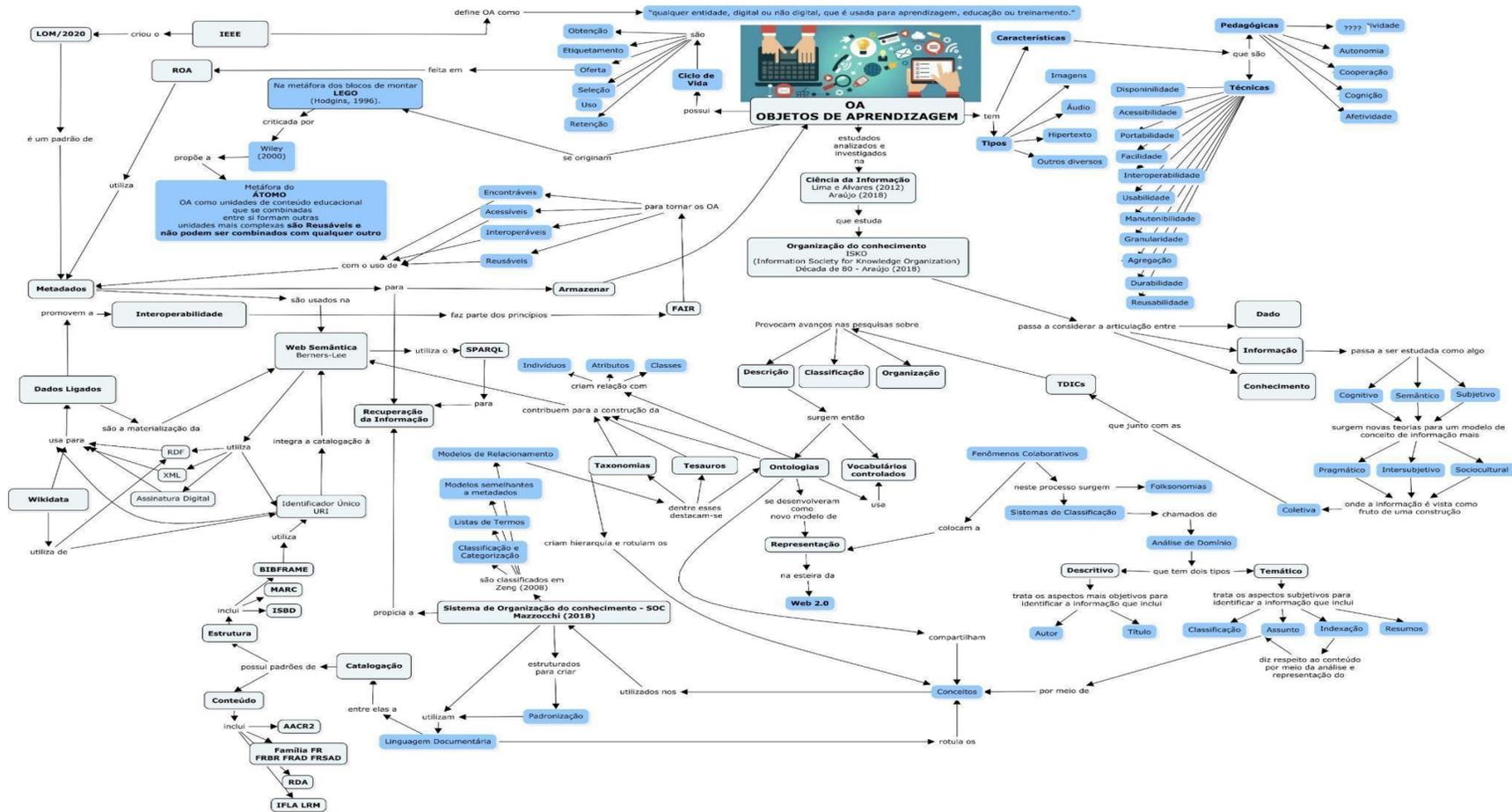
Fonte: elaborado pela autora (2023)

É possível observar também que as principais características esperada para o OA são as relacionadas à reusabilidade e a interoperabilidade o que também vai ao encontro do conceito proposto por Wiley (2014) para “os 5R dos Recursos Educacionais Abertos - REA” que são Reter, Reutilizar, Revisar, Remixar e Redistribuir. O que mostra a

importância da adoção de padrões, metadados e identificadores persistentes para tratar os OAs em seu estágio do ciclo de vida chamado oferta, onde os mesmos são armazenados e disponibilizados nos Repositórios - ROA.

Um desses padrões de metadados é o LOM/2020 da IEEE que pode facilitar a descrição, busca, avaliação e reuso dos OAs, esses padrões devem ser identificados por meio de identificadores universais de recursos (URI), exatamente como usado em um link em uma página da Web, onde a Web não “fala” e não “entende” a linguagem humana, por isso, o RDF usa URIs para codificar os OAs utilizando URIs que garantem que os conceitos não sejam apenas palavras em um documento, mas estejam vinculados a uma definição exclusiva que todos podem encontrar na Web desde que estejam publicados em redes colaborativas com dados estruturados de forma interoperável. O mapa conceitual da Figura 23 mostra diversas interconexões sobre conceitos trabalhados neste capítulo.

Figura 23 - Mapa conceitual fundamentação teórica



35 Fonte: Mapa conceitual elaborado pela autora (2023) por meio do CmapTools e disponível com melhor resolução em: <https://cmasccloud.ihmc.us:443/rid=1YB8RYT1P-20JQRHB-QB00/MAPA%20CONCEITUAL%20OA%20DISSERTA%C3%87%C3%83O.cmap>

4 FUNDAMENTAÇÃO METODOLÓGICA

Este capítulo trata dos conceitos e, principalmente, das ferramentas atinentes às etapas dos procedimentos metodológicos da pesquisa, tanto no processo de organização do conhecimento da pesquisa e a sua escrita, quanto na execução dos objetivos específicos. O entendimento de cada uma delas é necessário para a compreensão de como a pesquisa foi realizada, que será abordado no Capítulo 5 dos Procedimentos Metodológicos.

4.1 ORGANIZAÇÃO DA PESQUISA

Nesta seção far-se-á a apresentação dos sistemas utilizados para facilitar a organização e escrita da pesquisa.

4.1.1 Mapas Conceituais

De acordo com Sherrat e Schlabach (1990 apud Cristóvão, 2016) um mapa conceitual traz elementos capazes de envolver a identificação de conceitos ou ideias que pertencem a um assunto e descreve as relações existentes entre essas ideias de forma esquematizada, cuja finalidade é representar como um indivíduo compreende um corpo de conhecimento, onde o mapa ilustra a ideia que foi significativa para ele. Ou, simplesmente, como uma técnica para representação gráfica de conhecimento (LANZING, 1997 apud CRISTÓVÃO, 2016).

Assim, como forma de expressar muitas das ideias presentes neste trabalho, optou-se por utilizar mapas conceituais que serão escritos com a ajuda do software CmapTools³⁶, desenvolvido pelo IHMC³⁷, ferramenta que torna disponível uma grande variedade de funções para operação com mapas conceituais, desde a criação e edição de mapas com gravação do histórico de ações, até a formatação completa, o trabalho colaborativo e o compartilhamento dos mesmos por meio de servidores públicos ou do próprio IHMC (CMAPTOOLS, 2022).

³⁶ CmapTools é um software para construção e compartilhamento de mapas conceituais. Disponível em: <http://cmap.ihmc.us/>.

³⁷ IHMC (Florida Institute for Human & Machine Cognition), disponível em: <http://www.ihmc.us/>.

4.1.2 Gerenciador Bibliográfico

Os gerenciadores de referências bibliográficas são ferramentas que auxiliam na organização da bibliografia utilizada na pesquisa, geralmente softwares que permitem armazenar, gerenciar e citar as referências de acordo com normas específicas de cada periódico.

Os softwares de gerenciamento de referências permitem ao usuário criar a sua própria base de dados de referências bibliográficas. Ao adicionar à sua conta os documentos relevantes para sua pesquisa, o usuário tem um local para armazenar o conteúdo e controlar a própria leitura, podendo definir no software o que já leu ou referenciou. Estes softwares organizam os documentos por assunto e permitem pesquisar a partir de palavras-chave, facilitando o trabalho do pesquisador (PERIÓDICOS DE MINAS, 2022, sem página)

Entre eles, existem o Mendeley³⁸, EndNote³⁹ e o Zotero⁴⁰, este último de acordo com Yamakawa et al (2014), se destaca por ser gratuito, com recursos eficientes e permitir o armazenamento simultâneo dos arquivos no disco rígido do computador e na Web, além de buscas por palavras-chave, periódico, autores etc. Na presente pesquisa usa o software Zotero para a organização das referências bibliográficas e do método Zettelkasten⁴¹ que ajuda na organização de anotações e citações, o que contribui e agiliza o processo de escrita.

4.2 DESENVOLVIMENTO DA PESQUISA

Aqui pretende-se apresentar as ferramentas tecnológicas que foram usadas nas etapas procedimentais do desenvolvimento deste trabalho. Um dos principais critérios para escolha dessas ferramentas foi a gratuidade de uso e a sua boa usabilidade.

O detalhamento das ações que cada uma das ferramentas desempenhou é abordado no Capítulo 5 dos Procedimentos Metodológicos.

³⁸ Software da Elsevier disponível em: <https://www.mendeley.com/>

³⁹ Software da Clarivate disponível em: <https://endnote.com/>

⁴⁰ O Zotero é um gerenciador de referência bibliográfica de código aberto, arquivador de documentos, gerenciador de citações e uma ferramenta de colaboração (YAMAKAWA et al, 2014, p.173). Disponível em: <https://www.zotero.org/>

⁴¹ Zettelkasten significa “caixa de anotação”, em tradução direta, mas também pode significar aquelas caixinhas que continham fichas catalográficas (essas são as zettels) na biblioteca. Disponível em: <https://delchibruce.com/o-metodo-zettelkasten.html>.

As ferramentas apresentadas nesta seção são: software de preparação dos dados, software para implementação de ontologias operacionais de domínio, sistemas de bancos de dados orientados a grafos, linguagem consulta em bases de dados ligados.

4.2.1 Software de Preparação de Dados

De acordo com Guerra (2020), a etapa de preparação dos dados também conhecida por *Data Wrangling*⁴² consiste,

na coleta, limpeza e transformação de dados brutos para que sejam processados e qualificados. Geralmente, envolve a reformatação, a correção e a combinação de conjuntos de informações para que sejam enriquecidas e validadas.

Segundo o autor, é de muito útil quando se trabalha com dados confusos, inconsistentes ou sem padronização, quando é preciso combinar dados de várias fontes, quando os dados são extraídos de fontes não estruturadas, entre outros, pois garante que as informações geradas como resultado do processo sejam precisas e consistentes, o que permite uma análise válida e realista do fenômeno estudado,

Um dos softwares utilizado neste processo de preparação dos dados, é o OpenRefine⁴³ escolhido por ser um aplicativo de desktop de código aberto para limpeza e transformação de dados para outros formatos e de fácil manipulação. É semelhante a aplicativos de planilha e pode lidar com formatos de arquivo de planilha como CSV, mas se comporta mais como um banco de dados. Faz parte do CS&S - *Code for Science & Society*⁴⁴ (OPENREFINE, 2022).

O OpenRefine opera em linhas de dados que possuem células em colunas, semelhante à maneira como as tabelas de banco de dados relacional operam. Os projetos OpenRefine consistem em uma tabela, cujas linhas podem ser filtradas

42 Disponível em: <https://infochoice.com.br/site/index.php/data-wrangling-preparacao-de-dados/>

43 Aplicativo de desktop de código aberto para limpeza e transformação de dados para outros formatos. É semelhante a aplicativos de planilha e pode lidar com formatos de arquivo de planilha como CSV, mas se comporta mais como um banco de dados. Faz parte do CS&S - *Code for Science & Society* (OPENREFINE, 2022). Disponível em: <https://openrefine.org/>. Disponível em: <https://openrefine.org/>

44 Organização sem fins lucrativos sediada nos EUA que apoia a colaboração aberta em tecnologia de interesse público por meio de patrocínio fiscal e outros programas que apoiam o código aberto sustentável. Disponível em: <https://codeforscience.org/>

usando facetas que definem critérios (por exemplo, mostrando linhas onde uma determinada coluna não está vazia). Com o sistema é possível:

- **Limpar de dados confusos:** por exemplo, se estiver trabalhando com um arquivo de texto com alguns dados semi estruturados, ele pode ser editado usando transformações, facetas e clustering para tornar os dados estruturados de forma limpa.
- **Transformar de dados:** conversão de valores para outros formatos, normalização e desnormalização.
- **Analisar dados de sites:** OpenRefine tem um recurso de busca de URL e um analisador
- **Adicionar dados ao conjunto de dados:** buscando-os em serviços da web, por exemplo, pode ser usado para geocodificação de endereços para coordenadas geográficas.
- **Reconciliação de dados:** Alinhado ao Wikidata faz mapeamento de valores de string em células para entidades no Wikidata (OPENREFINE, 2022, tradução nossa).

Ainda de acordo com o site do OpenRefine, os dados podem ser importados e exportados em vários formatos incluindo o triplas RDF (formatos de serialização RDF/XML), consultas SQL e também aceita a definição de um modelo de exportação capaz de atender a necessidades diversas.

Na presente pesquisa, o software OpenRefine foi usado na Etapa 2, dos procedimentos do desenvolvimento, descrita na subseção 5.3.2.

4.2.2 Software para Implementação de Ontologias Operacionais de Domínio

A implementação de ontologias operacionais de domínio é um processo que visa tornar operacional uma ontologia de referência de domínio, isto é, transformá-la em uma linguagem que possa ser interpretada por máquina.

A metodologia de construção de ontologias de domínio apresentada por Falbo (2014), subseção 3.1.4 da Fundamentação Teórica, orienta que após a concepção da ontologia de referência de domínio, seja elaborado um modelo de design seguido da implementação da ontologia operacional. No caso da presente pesquisa foram usadas as informações detalhadas contidas no padrão LOM-IEEE como base conceitual para elaboração do modelo de design.

A ferramenta utilizada na elaboração do modelo de design foi o software Visual Paradigm⁴⁵. Ela se expressa por meio de diagramas compostos por elementos gráficos que representam as entidades e as suas relações. Para a diagramação foi utilizado o software Visual Paradigm que, apesar de possuir uma versão paga com vários recursos, possui também uma versão gratuita que atendeu bem às necessidades da pesquisa, por isso, foi a ferramenta de escolha.

Ele foi construído em formato de diagrama na linguagem OntoUML. Para suportar a linguagem OntoUML, foi adicionado o plugin gUFO ⁴⁶no Visual Paradigm.

Na presente pesquisa, o software Visual Paradigm foi usado na Etapa 4, dos procedimentos do desenvolvimento, descrita na subseção 5.3.4.

Entre as várias ferramentas de implementação ontologias, o VocBench⁴⁷ destaca-se por ser uma plataforma de desenvolvimento colaborativo, multilíngue e baseada na Web para gerenciamento de ontologias OWL, tesouros SKOS (XL), e conjuntos de dados RDF genéricos. Foi projetado para atender às necessidades de ambientes da Web semântica e Dados Ligados. Seu desenvolvimento foi impulsionado pelo feedback obtido de uma comunidade de usuários composta por organizações públicas, empresas e usuários independentes em busca de soluções de dados abertos para manutenção de suas ontologias, tesouros, códigos-listas e recursos de autoridade.

As camadas de acesso de negócios e dados VocBench são realizadas pelo ambiente Semantic Turkey, uma plataforma de código aberto para aquisição e gestão do conhecimento realizada pelo ART⁴⁸ *Research Group* da Universidade de Tor Vergata em Roma. A plataforma oferece um poderoso ambiente de edição, com facilidades para gerenciamento de ontologias OWL, tesouros SKOS/SKOS-XL,

⁴⁵ Visual Paradigm é um software para construção de modelos diversos. Disponível em: <https://www.visual-paradigm.com/>.

⁴⁶ gUFO é um plugin do Visual Paradigm para construção de modelos com a ontologia de fundamentação UFO utilizando-se da linguagem OntoUML. Disponível em: <https://github.com/OntoUML/ontouml-vp-plugin/>.

⁴⁷ Disponível em: <http://vocbench.uniroma2.it/>

⁴⁸ ART (Pesquisa de Inteligência Artificial no Tor Vergata) na Universidade de Tor Vergata é um grupo de pesquisa de média escala, incluindo professores, associados e estudantes de doutorado que trabalham em diferentes campos da Inteligência Artificial. Disponível em: <http://art.uniroma2.it/>

léxicos OntoLex⁴⁹ e qualquer tipo de conjunto de dados RDF. Ele visa estabelecer novos padrões de flexibilidade, abertura e poder expressivo como uma plataforma de modelagem RDF gratuita e de código aberto (VOCBENCH, 2022).

A linguagem utilizada para representar e serializar a ontologia operacional de domínio foi a Turtle RDF⁵⁰.

Na presente pesquisa, o software VocBench foi usado na Etapa 5, dos procedimentos do desenvolvimento, descrita na subseção 5.3.5.

4.2.3 Sistemas de Banco de Dados Orientado a Grafos

Os sistemas de bancos de dados orientados a grafos semânticos, também chamados de banco de dados orientados a grafos ou, simplesmente, *triplestores* RDF, trabalham com triplas e fornecem a infraestrutura central para soluções que agilizam a modelagem e integração de dados, exploração de relacionamento e publicação e consumo de dados.

Entre eles tem-se o GraphDB⁵¹ que é um banco de dados de grafos semântico compatível com os padrões W3C. É um banco de dados da família RDF, com excelente usabilidade, altamente eficiente, robusto e escalável. Ele agiliza a carga e o uso de dados de datasets de dados ligados, bem como seus próprios recursos. É compatível com vários padrões, como a especificação do protocolo W3C SPARQL e suporta todos os formatos de serialização RDF. É o banco de dados de escolha dos pequenos desenvolvedores independentes e grandes organizações corporativas por causa de seu suporte comercial e comunitário, além de excelentes recursos corporativos, como suporte a cluster e integração com aplicativos de pesquisa externos de alto desempenho (GRAPHDB, 2022).

O GraphDB é um dos poucos *triplestores* que podem realizar inferência semântica em escala, permitindo aos usuários derivar novos fatos semânticos de fatos

⁴⁹ OntoLex é o nome abreviado de um vocabulário para recursos lexicais na web de dados (OntoLex-Lemon) e o nome abreviado do grupo da comunidade W3C que o criou (W3C Ontology-Lexica Community Group).

⁵⁰ RDF Turtle é uma linguagem de marcação para representação de dados ligados RDF. Disponível em: <https://www.w3.org/TR/turtle/>.

⁵¹ Disponível em: <https://graphdb.ontotext.com/>

existentes. Ele lida com cargas massivas, consultas e inferências em tempo real. O “raciocínio” e a avaliação da consulta são executados em uma camada de armazenamento persistente, fazendo com que ocorram de forma extremamente rápida, mesmo em grandes ontologias e bases de conhecimento. De acordo com o *LDBC Semantic Publishing Benchmark*⁵², é um dos repositórios OWL mais escaláveis atualmente disponíveis (GRAPHDB, 2022).

O GraphDB também realiza mapeamentos a partir de uma ontologia operacional e uma base de dados, em diversos formatos, para uma base de dados ligados RDF, pelos motivos aqui elencados, foi o software de escolha para a etapa mapeamento e consulta de dados.

Na presente pesquisa o software GraphDB foi usado no mapeamento realizado na Etapa 7, dos procedimentos do desenvolvimento, descrito na subseção 5.3.7.

4.2.4 Linguagem de Consultas em Bases de Dados Ligados

Linguagem de consulta (*Query Language*) refere-se a qualquer linguagem de programação que solicita e recupera dados de bancos de dados e sistemas de informação por meio de consultas. Funciona em consultas baseadas em comandos de programação estruturada e formal feitas pelos usuários para localizar e extrair dados de bancos de dados. A linguagem de consulta também pode ser denominada linguagem de consulta de banco de dados (*database query language*) (SHENA, 2021).

Existem várias linguagens de consulta, porém, nesta pesquisa será usada a SPARQL (Protocol and RDF Query Language), pois é voltada às bases de dados ligados e trata-se de uma linguagem padronizada para a consulta de grafos RDF. Ela segue o padrão desenvolvido pelo RDF Data Access Working Group do W3C, o Grupo de Trabalho SPARQL da W3C que produziu uma “Recomendação”⁵³ para

⁵² O Semantic Publishing Benchmark (SPB) é um benchmark LDBC para testar o desempenho de mecanismos RDF inspirados na indústria de mídia/publicação. Em particular, a LDBC trabalhou com a British Broadcasting Corporation BBC para definir esse benchmark, para o qual a BBC doou cargas de trabalho, ontologias e dados. A indústria editorial é uma área onde a adoção significativa do RDF está ocorrendo. Disponível em: <https://ldbcouncil.org/benchmarks/spb/>

⁵³ A Recomendação do W3C é uma especificação ou conjunto de diretrizes que, após extensa construção de consenso, recebeu o endosso dos membros do W3C e do Diretor. O W3C recomenda

2008, cuja última atualização disponível foi do SPARQL 1.1 datada de 26 de março de 2013 (W3C, 2022).

Conforme já exposto, esta “recomendação” da W3C é um “[...] conjunto de especificações que fornecem linguagens e protocolos para consultar e manipular o conteúdo publicado em RDF na Web” (SANTAREM SEGUNDO, 2014, p. 3871). Estas especificações definem a sintaxe e a semântica da linguagem de consulta SPARQL para RDF. O SPARQL pode ser usado para expressar consultas em diversas fontes de dados, sejam os dados armazenados nativamente como RDF ou visualizados como RDF por meio de *middleware*⁵⁴. O SPARQL contém recursos para consultar padrões de gráfico obrigatórios e opcionais juntamente com suas conjunções e disjunções. O SPARQL também suporta testes de valor extensível e consultas de restrição por gráfico RDF de origem. Os resultados das consultas SPARQL podem ser conjuntos de resultados ou gráficos RDF (W3C, 2022).

Consultas escritas em SPARQL foram realizadas na Etapa 8, dos procedimentos do desenvolvimento, descrita na subseção 5.3.8.

a ampla implantação de suas Recomendações. Observação: as recomendações do W3C são semelhantes aos padrões publicados por outras organizações. Disponível em: <https://www.w3.org/2005/10/Process-20051014/tr.html#RecsW3C>

⁵⁴ Middleware é um software que fica entre um sistema operacional e os aplicativos executados nele. Funcionando essencialmente como uma camada de tradução oculta, o middleware permite a comunicação e o gerenciamento de dados para aplicativos distribuídos. Disponível em: <https://azure.microsoft.com/pt-br/resources/cloud-computing-dictionary/what-is-middleware/>

5 PROCEDIMENTOS METODOLÓGICOS

A proposta metodológica da presente pesquisa está organizada em três seções. A primeira seção classifica o trabalho de acordo com a literatura da área de metodologia de pesquisa. A segunda seção descreve o levantamento bibliográfico do estado da arte e da fundamentação teórica. A terceira seção apresenta as etapas procedimentais do desenvolvimento.

5.1 CLASSIFICAÇÃO DA PESQUISA E CENÁRIO DA PESQUISA

A pesquisa possui natureza aplicada, pois, segundo GIL, (2010), esse tipo de pesquisa abrange estudos elaborados com o objetivo de resolver problemas identificados no âmbito das sociedades em que os pesquisadores vivem, assim como, contribui para a ampliação do conhecimento científico e sugere novas questões a serem investigadas. No caso do presente trabalho, espera-se que a criação e a aplicação de uma ontologia operacional de domínio para representar OAs possa aproximar essa representação proposta aos princípios FAIR trazendo benefícios para o acesso e recuperação da informação. A abordagem da investigação é qualitativa, pois de acordo com Chizzotti (1991) os aspectos qualitativos na área de educação diz respeito a natureza das percepções que os pesquisadores trazem para a pesquisa que não podem ser quantificados, ou seja, trabalha com o universo de significados, de motivações, aspirações, crenças, valores e atitudes. Onde o pesquisador interage com o objeto de estudo numa interatividade que não é estática, pois na medida em que este sujeito interage com o objeto observado, cria-se uma rede de conectividade entre a realidade objetiva e os elementos subjetivos que constituem um sujeito. Essa abordagem orienta procedimentos de pesquisa sobre o objeto de estudo que requerem descrições e análises não numéricas de um determinado fenômeno.

Entende-se também que esta pesquisa possui procedimentos de estudo de caso, que segundo Yin (1994),

[...] as experiências estão completas quando um investigador pode manipular o comportamento directamente, precisamente e sistematicamente. Isto pode ocorrer num cenário laboratorial, no qual uma experiência pode concentrar-se numa ou duas variáveis isoladas (e

presume que o ambiente laboratorial pode "controlar" todas as restantes variáveis para além do campo de interesse)[...] (YIN, 1994)

Visto que a experiência se deu com a manipulação dos dados de três repositórios de OAs que trabalham com dados abertos, o primeiro escolhido, pelo critério do tipo, conforme apresentado na seção 3.4.2 da Fundamentação Teórica, foi o Repositório de Recursos Educacionais Abertos para Educação Profissional e Tecnológica - PROEDU. O segundo escolhido, pelo critério do conteúdo, foi o Banco Internacional de Objetos de Aprendizagem - BIOE. O terceiro foi escolhido pelo seu propósito institucional que foi o ARCA – Repositório Institucional da Fiocruz. Assim como o padrão escolhido para a representação das informações dos OAs foi o LOM da IEEE/2020 por entender que é o padrão internacionalmente mais utilizado e aceito no que diz respeito a objetos de aprendizagem.

A metodologia denominada de “prova de conceito”, aplicada na parte final do desenvolvimento do presente trabalho, refere-se, segundo Kendig (2016), a qualquer ideia que possa ser aplicada a uma classe de fenômenos onde a prova possa ser uma possibilidade de se obter com a prática experimental. A autora ainda complementa explicando que na prova de conceito existe uma conexão causal hipotética na estrutura proposta, onde a função sugerida ou a abordagem metodológica adotada na pesquisa são obtidas em pelo menos um caso real (caso de teste). Segundo Kendig, a prova de conceito é enquadrada em termos de um tipo particular de pesquisa que visa responder a uma questão, cuja resposta tem ampla aplicabilidade em áreas além daquela testada, ou seja, a pesquisa fornece justificativa na prática da transportabilidade potencial dessa pesquisa, por exemplo, a metodologia usada, processo descrito, o padrão instanciado, a aplicação prototípica, a estratégia de intervenção realizada, o modelo ou representação utilizada. Desta forma acredita-se que os resultados obtidos com a experimentação num pequeno recorte, podem ser replicados em outras instâncias maiores.

5.2 LEVANTAMENTO BIBLIOGRÁFICO

Esta pesquisa está diretamente relacionada com três áreas do conhecimento da Capes, sendo elas, 6.07.00.00-9 - Ciência da informação, 1.03.00.00-7 - Ciência da computação e 7.08.00.00-6 - Educação. Desta forma, ao fazer o levantamento de

trabalhos com temas similares, optou-se por bases de dados que abrangessem trabalhos das referidas áreas, mas não se limitando somente a elas, visto que os descritores utilizados também são recorrentes em outras áreas. Assim, optou-se pelas seguintes bases: Google Acadêmico⁵⁵: repositório que agrega materiais científicos como artigos, ensaios, livros, teses e dissertações. Nele, são indexados, conteúdos acadêmicos do Google Livros e bancos de materiais científicos com acesso público, entre eles, o Scientific Eletronic Library Online - SciELO, é considerado a maior fonte de arquivos científicos online. Portal de Periódicos da Capes⁵⁶: reúne e disponibiliza mais de 49 mil periódicos e 455 bases de dados de diversos conteúdos, referências, patentes, estatísticas, material audiovisual, normas técnicas, teses, dissertações, livros e obras.

Ao pesquisar os trabalhos com principais descritores aqui apresentados “Objetos de Aprendizagem”, “Web semântica”, “Dados Ligados” e “Interoperabilidade”, separadamente nas bases escolhidas, nos últimos cinco anos (para o Capítulo 2, Estado da Arte), houve uma recuperação muito exaustiva, com alta revocação, porém, com baixíssima precisão em relação ao tema principal da pesquisa, conforme demonstra a Quadro 7 abaixo.

Quadro 7 - Resultados para os descritores nas bases de dados Google Scholar e Portal Capes

Descritores	Google Acadêmico	Portal Capes
Objetos de Aprendizagem	9.790	218
Web semântica	15.400	2.381
Dados Ligados	15.900	1.464
Interoperabilidade	15.000	5.150

Fonte: elaborado pela autora (2023)

Apesar de encontrados muitos trabalhos com os descritores apresentados, conforme relatado, pesquisando-os separadamente, muitos dos trabalhos recuperados pouca ou nenhuma relação continha com a pesquisa ora apresentada. Desta forma, optou-se por pesquisar fazendo algumas combinações entre os termos, por exemplo:

⁵⁵ Disponível em: <https://scholar.google.com.br/?hl=pt> Acesso em: 15 jun. 2022.

⁵⁶ Disponível em: www-periodicos-capes-gov-br Acesso em: 15 jun. 2022.

“Objetos de Aprendizagem e Web semântica”, “Objetos de Aprendizagem e Dados Ligados” sempre em português e inglês com base nos últimos cinco anos e buscando sempre os mais recentes. Observou-se que a maioria das publicações recuperadas são artigos de revistas científicas, teses e dissertações. Também serviram como fontes de informação as bibliografias utilizadas nas disciplinas do Programa de Pós-graduação em Ciência da Informação da Universidade Federal do Espírito Santo - UFES, entre elas, “Epistemologia da Ciência da Informação”, “Informação e Cultura” e principalmente as disponibilizadas no gerenciador de referências Zotero do grupo da disciplina “Sistemas de Recuperação da Informação”, que foram muito importantes, pois as referências ali apresentadas serviram de base para encontrar outras a partir da leitura do material.

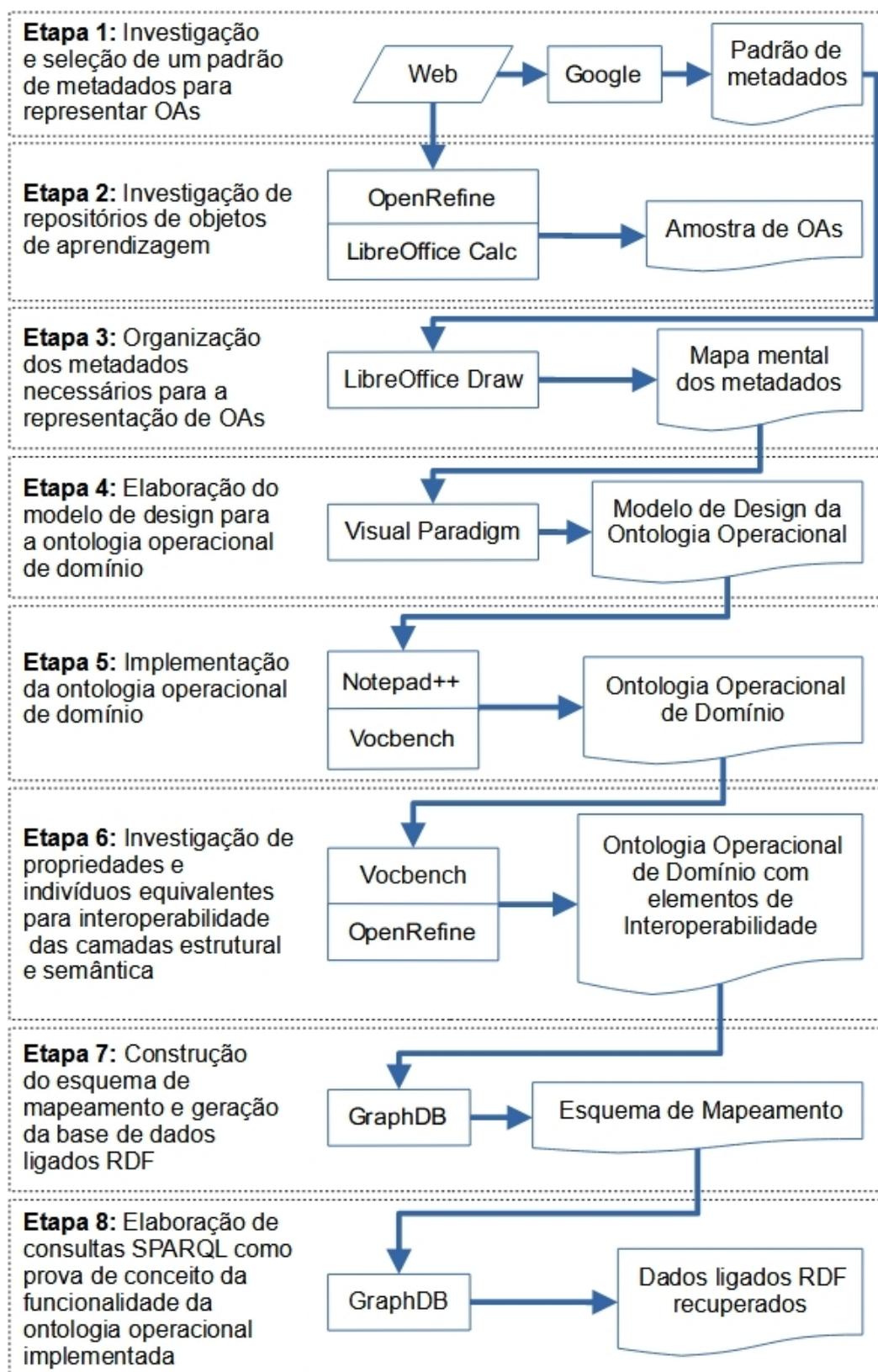
Todas as referências escolhidas foram organizadas na biblioteca pessoal do Zotero da autora utilizando-se de etiquetas com os descritores e utilizando-se o método Zettelkasten para anotações e citações.

Para este percurso de pesquisa, mais do que a quantidade de materiais recuperados para que aqui fossem apresentados, os critérios utilizados foram a aproximação mais estreita possível com os descritores desta pesquisa, onde os trabalhos encontrados foram apresentados na seção Estado da Arte, e a percepção da autora em relação aos autores mais recorrentes na literatura de acordo com os assuntos aqui tratados, onde uma obra encontrada acaba por trazer a outra, e assim sucessivamente, ou seja, pode-se dizer que a próprio desenvolvimento da pesquisa cooperou para a construção de seu referencial teórico.

5.3 ETAPAS PROCEDIMENTAIS DO DESENVOLVIMENTO

Esta seção apresenta as etapas procedimentais que foram executadas para obtenção dos objetivos da pesquisa. Ela foi organizada em oito etapas sequenciais que explicam como as ações foram desenvolvidas, principalmente com base nos processos e ferramentas apresentadas no Capítulo 4, Fundamentação Metodológica. Na Figura 24 mostra-se todas as etapas procedimentais apresentadas neste capítulo, assim como as ferramentas tecnológicas que auxiliaram em seu desenvolvimento e os resultados de cada uma delas.

Figura 24 – Fluxo das etapas procedimentais do desenvolvimento



Fonte: Elaborado pela autora (2023)

5.3.1 Etapa 1: Investigação e seleção de um padrão de metadados para representar OAs

Foram investigados padrões de metadados que pudessem melhor representar os OAs, primeiro buscou-se na literatura os conceitos relacionados à Metadados, “considerado como uma descrição de dados que podem ser utilizados para identificar e descrever características comuns entre diferentes recursos informacionais com a finalidade de facilitar a interoperabilidade e a recuperação na Web”, ou de maneira simples, “dados sobre dados” (PÖTTKER, FERNEDA E MOREIRO-GONZÁLEZ, 2018, p.29).

No processo de pesquisa, conforme já apresentado na Seção 3.2 da Fundamentação Teórica, descobriu-se que o uso de metadados é uma prática antiga da Biblioteconomia, fruto do desenvolvimento histórico de regras de catalogação que inclui padrões de conteúdo como AACR2, o seu sucessor, o RDA, e o caminho percorrido até a chegada do LRM, passando pelos modelos conceituais da família FR (FRBR, FRAD e FR SAD) e padrões de estrutura como os ISBD, o formato MARC Bibliográfico e BibFrame. Chegando-se aos mais atuais voltados para as novas Tecnologias Digitais de Informação e Comunicação - TDICs como DublinCore, linguagens eXtensible Markup Language - XML, Resource Description Framework - RDF e Ontology Web Language - OWL para representações de características de objetos digitais, especialmente na Web (FOULONNEAU; RILEY, 2008; ZAFALON, 2012; MACHADO; ZAFALON, 2020).

Assim, chegou-se ao Learning Object Metadata - LOM (IEEE, 2020) um padrão internacional que é regido pela norma IEEE Std 1484.12.1™, 2020, que foi apresentado na Seção 3.4.5 da Fundamentação Teórica. Há outros padrões para representação de objetos multimídias como o da W3C (World Wide Web Consortium) e da ISO/IEC (International Organization for Standardization/ International Electrotechnical Commission), que buscam soluções inteligentes para descrição de conteúdos multimídia processáveis por máquina e baseados em semântica como é o caso do padrão de metadados MPEG-7 ISO/IEC, comumente usado para descrição de conteúdo multimídia em rede, mas que

aqui não foi explorado, pois procurou-se por um padrão que apresentasse categorias específicas que descrevessem o objetivo educacional.

Por isso, apesar da existência desses outros padrões, optou-se pelo LOM-IEEE por ser um padrão com categorias específicas para catalogação de objetos de aprendizagem que se propõem a: [1] Facilitar a busca, avaliação, aquisição e uso de objetos de aprendizagem, por exemplo, por alunos, instrutores ou processos de software automatizados; [2] Facilitar o compartilhamento e troca de objetos de aprendizagem, permitindo o desenvolvimento de catálogos e inventários, levando em consideração a diversidade de contextos culturais e linguísticos em que os objetos de aprendizagem e seus metadados serão explorados (IEEE, 2020, p.11, tradução nossa).

Além disso, segundo IEEE (2020) o padrão traz em seu escopo a pretensão de que seja referenciado por outros padrões, onde propõe um esquema de dados conceitual que permite a diversidade linguística de modo que uma instância de um metadado de um OA possa ser utilizada por um sistema para gerenciar, localizar, avaliar ou trocar OA. A escolha do padrão LOM-IEEE objetiva garantir que as ligações de Metadados de OAs levem a um alto grau de interoperabilidade semântica e, como resultado, as transformações entre essas ligações sejam diretas.

5.3.2 Etapa 2: Investigação de repositórios de objetos de aprendizagem

Foram consultados os seguintes repositórios de acesso aberto: o Repositório de Recursos Educacionais Abertos para Educação Profissional e Tecnológica (PROEDU), o Banco Internacional de Objetos de Aprendizagem (BIOE) e o Repositório Institucional da Fiocruz (ARCA). Porém, em contato com os administradores dos repositórios PROEDU e BIOE, não foi possível a disponibilização dos arquivos de dados de OAs para continuidade da presente pesquisa. O único repositório cujo administrador se dispôs a fornecer os arquivos necessários para a coleta de dados foi o ARCA da Fiocruz que disponibilizou o arquivo CSV com todos os REA da plataforma.

É importante salientar que o ARCA não é considerado um repositório apenas de OA, e sim um Repositório Institucional da Fundação Oswaldo Cruz (Fiocruz) cuja função é reunir, hospedar, disponibilizar e dar visibilidade à produção intelectual da Instituição, assim como estimular a mais ampla circulação do conhecimento, fortalecendo o compromisso institucional com o livre acesso da informação em saúde, além de conferir transparência e incentivar a comunicação científica entre pesquisadores, educadores, acadêmicos, gestores, alunos de pós-graduação, bem como a sociedade civil. O projeto Sudeste/RIAA (Rede Sudeste de Repositórios institucionais) capitaneado pela Fiocruz faz parte da Rede Nacional de Repositórios coordenada pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), composta atualmente por 79 instituições de ensino e pesquisa. O ARCA está organizado em comunidades que correspondem às unidades da Fiocruz. É mantido pelo Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT) e utiliza o software livre DSpace. Esta característica de foco no compartilhamento e incentivo aos dados abertos muito agrega a esta pesquisa. (QUEIROZ, 2021).

Para a investigação dos dados contidos nos repositórios investigados, foram utilizados os softwares OpenRefine e LibreOffice Calc⁵⁷. Em seguida foram escolhidos dez OAs que serviram como amostra para o experimento de aplicação. A escolha desses OAs foi realizada levando-se em consideração o tipo, características e ciclo de vida dos objetos analisados conforme apresentado na seção 3.4.

5.3.3 Etapa 3: Organização dos metadados necessários para a representação de OAs

Todas as categorias do LOM IEEE foram analisadas de acordo com as tabelas de dados disponíveis no padrão (Anexos de II a VI) e desenhadas para facilitar a elaboração do modelo de design, nesta etapa utilizou-se o software livre LibreOffice Draw⁵⁸ para síntese do padrão e desenho do mapa mental, das categorias e subcategorias identificadas por variáveis nomeadas segundo o formato camelCase⁵⁹

⁵⁷ Calc é um software de planilha eletrônica pertencente ao pacote LibreOffice. Disponível em: <https://www.libreoffice.org/download/download-libreoffice/>

⁵⁸ Disponível em: <https://www.libreoffice.org/discover/libreoffice/>.

⁵⁹ O formato camelCase padroniza a terminologia do nome de variáveis para integrar o seu uso entre vários softwares e linguagens. Disponível em: <https://pt.wikipedia.org/wiki/CamelCase/>.

com a junção das categorias e subcategorias. Por exemplo, “generalIdentifierCatalog” é o nome do identificador do metadado “Catalog” que pertence à subcategoria “Identifier” que pertence à categoria “General”. Esta mesma lógica foi aplicada a todas as outras oito categorias e suas subcategorias e a Figura 25 com todas elas é apresentada na seção “Resultados e Discussão”.

Apesar de o idioma padrão do LOM-IEEE ser o inglês, alguns termos foram traduzidos para o português no decorrer do processo apenas para facilitar o entendimento.

5.3.4 Etapa 4: Elaboração do modelo de design para a ontologia operacional de domínio

Em direção da obtenção da ontologia operacional de domínio, segundo a metodologia de Falbo (2014), foi elaborado um modelo de design diagramado por meio da linguagem OntoUML que, apesar de ser voltada para modelagem conceitual orientada por ontologia, pode também auxiliar na construção do modelo de design para a ontologia operacional. Ele foi construído em formato de diagrama na linguagem OntoUML. Para suportar a linguagem OntoUML, foi adicionado o plugin gUFO no Visual Paradigm.

A especificação detalhada do padrão LOM-IEEE, conforme abordado na subseção 3.4.5 da Fundamentação Teórica, foi determinante para fornecer informações sobre cada propriedade e seus vocabulários. Considerando a riqueza dessa especificação na representação de OAs, foi possível partir diretamente para a elaboração do modelo de design para a ontologia operacional de domínio sem que fosse necessária a diagramação de uma modelagem ontológica conceitual, denominada de ontologia de referência de domínio conforme previsto na metodologia SABiO de Falbo (2014). Além disso, é importante considerar que o modelo de design para a ontologia operacional de domínio, conforme é apresentado na Seção 6.3 dos resultados, representa com grande riqueza de detalhes e proximidade da implementação o padrão LOM-IEEE.

5.3.5 Etapa 5: Implementação da ontologia operacional de domínio

Na ontologia operacional, segundo Falbo (2014), é uma das etapas da metodologia SABiO de desenvolvimento de ontologias de domínio, e tem foco na implementação da ontologia em uma linguagem específica que, no caso da presente pesquisa, foi escolhido o RDF Turtle por ser uma linguagem que, apesar de ser interpretada por máquina, ela é também de fácil compreensão por humanos, facilitando assim o processo de correção de erros de implementação.

O modelo de design descrito na Etapa 3 orientou a escrita do código de implementação para a ontologia operacional de domínio. Para a escrita foi utilizado o editor de texto Notepad++⁶⁰ associado a um plugin para reconhecimento de código em Turtle. Optou-se por um desenvolvimento híbrido com o uso de OWL e SKOS conforme recomendações e orientações de Bechhofer e Miles (2008). Após a escrita do código em Turtle, o mesmo foi carregado na plataforma VocBench para permitir a adição de indivíduos e realizar testes na ontologia operacional.

5.3.6 Etapa 6: Investigação de propriedades e indivíduos equivalentes para a interoperabilidade das camadas estrutural e semântica

Nesta etapa buscou-se propriedades e indivíduos equivalentes em bases de dados externas tanto para a camada estrutural quanto para a camada semântica, utilizou-se as informações do LOM-IEEE e pesquisas nas bases.

Foram usadas as seguintes ontologias e/ou esquemas de metadados para estabelecer equivalências com propriedades da ontologia operacional proposta:

- Wikidata
- Dublin Core
- Bibframe

Escolhidas devido a importância e características relatadas nas seções 3.2 e 3.5.2.

⁶⁰Notepad++ é um editor de texto e de código fonte de código aberto sob a licença GPL. Suporta várias linguagens de programação rodando sob o sistema Microsoft Windows (possível utilização no Linux via Wine). Disponível em: <https://notepad-plus-plus.org/downloads/>.

O resultado das equivalências encontradas foi incorporado no código da ontologia operacional implementada na Etapa 5. Essas equivalências foram adicionadas com apoio do software VocBench.

5.3.7 Etapa 7: Construção do esquema de mapeamento e geração da base de dados ligados RDF

Nesta etapa, foi elaborado um esquema de mapeamento sobre a ontologia operacional implementada, já com os elementos interoperáveis da Etapa 6 e juntamente com os dados dos OAs investigados na Etapa 2, para obter uma base de dados ligados interoperáveis na Web semântica.

Os softwares OpenRefine e o banco de dados orientado a grafo GraphDB foram utilizados neste processo de mapeamento. O GraphDB dispõe de uma linguagem visual e interativa para tal elaboração.

5.3.8 Etapa 8: Elaboração de consultas SPARQL como prova de conceito da funcionalidade da ontologia operacional implementada

Ao final foram feitas buscas escritas na linguagem SPARQL a fim de testar a recuperação de informação dos OAs da amostra. A recuperação serve como prova de conceito sobre a funcionalidade da ontologia implementada em conjunto com a base de dados ligados RDF dos OAs selecionados.

6 RESULTADOS E DISCUSSÃO

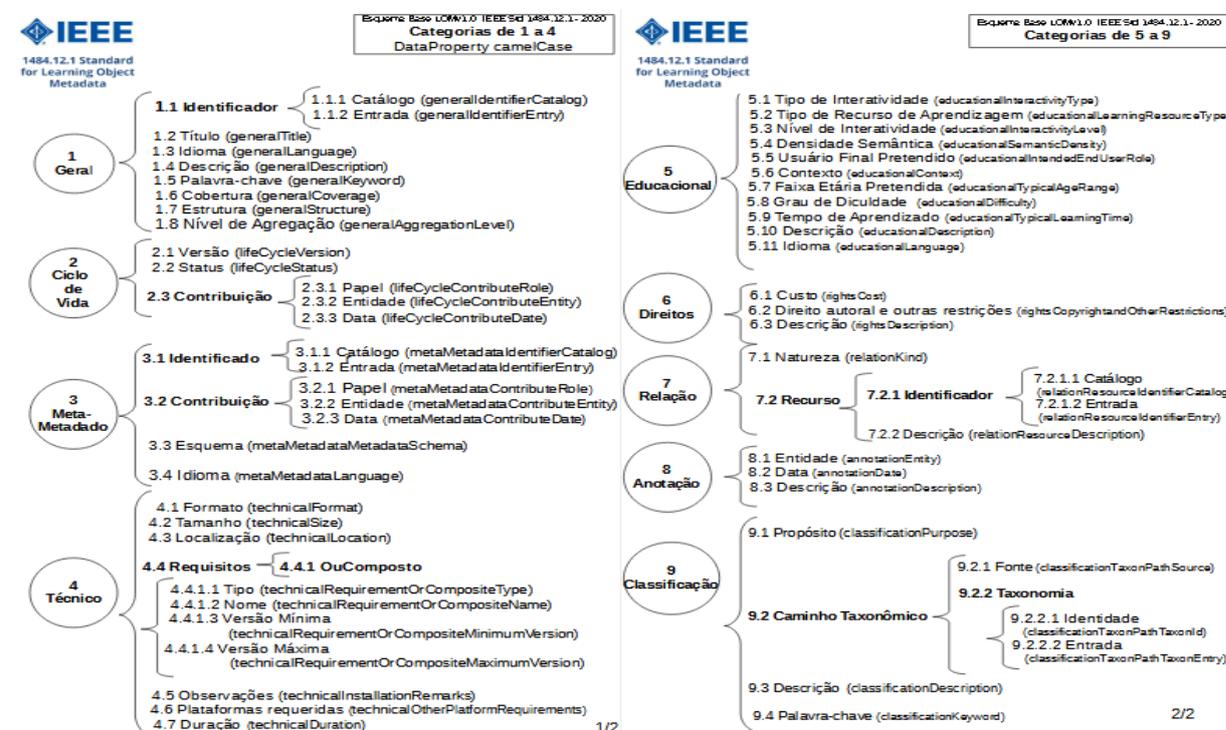
Esta seção apresenta os resultados da pesquisa conforme proposto nos objetivos específicos e os discute à luz da fundamentação teórica anteriormente apresentada. Os resultados são organizados em sete subseções e aparecem de forma sintonizada e na mesma sequência em que os objetivos específicos foram apresentados: metadados para o padrão LOM-IEEE; objetos de aprendizagem do repositório ARCA organizados conforme metadados do padrão LOM-IEEE; modelo de design para a ontologia operacional de domínio; ontologia operacional de domínio; interoperabilidade das camadas estrutural e semântica; mapeamento e base de dados ligados RDF; consultas SPARQL.

Discussões mais pontuais referentes aos resultados apresentados nas subseções foram realizadas diretamente nessas mesmas subseções. Outras discussões envolvendo elementos de diversas subseções, inclusive da Fundamentação Teórica, são apresentados na última subseção denominada de discussões gerais.

6.1 METADADOS PARA O PADRÃO LOM-IEEE

O padrão IEEE Std 1484.12.1™, 2020, foi selecionado conforme critérios apresentados na Seção 3.4.5 O LOM possui algumas tabelas que orientam o processo de interoperabilidade do padrão definindo também o esquema base da estrutura com nove categorias e suas respectivas subcategorias (Anexos de II a VI). A partir dessas informações, foi criado um mapa mental, Figura 25, contendo as nove categorias e seus itens, bem como o nome identificador de cada metadado, ou variável, criado na terminologia do formato camelCase.

Figura 25 - Mapa mental do esquema base do LOM/IEEE-2020



Fonte: elaborado pela autora (2023) a partir do padrão LOM-IEEE (IEEE, 2020)

Conforme já explicado na subseção 5.3.3 os nomes dos metadados, que aparecem entre parênteses na Figura 25, são o resultado da junção das categorias e no formato camelCase.

Conforme visto na subseção 3.1.2, Barbedo, Corujo e Sant'Ana (2011) postularam algumas categorias funcionais de tipos de metadados, que são: Metadados descritivos ou de identificação, Metadados estruturais, Metadados administrativos, Metadados técnicos e Metadados de preservação no Quadro 8 traçou-se um paralelo entre as categorias do LOM-IEEE e as categorias funcionais propostas pelo autores.

Quadro 8 - Categorias funcionais do LOM-IEEE segundo Barbedo, Corujo e Sant'ana (2011)

Categorias do LOM (camelCase)	Categorias funcionais de tipos de metadados segundo Barbedo, Corujo e Sant'Ana (2011)
generalIdentifierCatalog	descritivos/administrativos
generalIdentifierEntry	descritivos/administrativos
generalTitle	descritivos/administrativos
generalLanguage	descritivos/administrativos
generalDescription	descritivos/administrativos
generalKeyword	descritivos/administrativos
generalCoverage	descritivos/administrativos
generalStructure	descritivos/administrativos
generalAggregationLevel	descritivos/administrativos
lifeCycleVersion	descritivos/administrativos
lifeCycleStatus	descritivos/administrativos
lifeCycleContributeRole	descritivos/administrativos
lifeCycleContributeEntity	descritivos/administrativos
lifeCycleContributeDate	descritivos/administrativos
metaMetadataIdentifierCatalog	preservação
metaMetadataIdentifierEntry	preservação
metaMetadataContributeRole	preservação
metaMetadataContributeEntity	preservação
metaMetadataContributeDate	preservação
metaMetadataMetadataSchema	preservação
metaMetadataLanguage	preservação
technicalFormat	técnicos
technicalSize	técnicos
technicalLocation	técnicos
technicalRequirementOrCompositeType	técnicos
technicalRequirementOrCompositeName	técnicos
technicalRequirementOrCompositeMinimumVersion	técnicos
technicalRequirementOrCompositeMaximumVersion	técnicos
technicalInstallationRemarks	técnicos
technicalOtherPlatformRequirements	técnicos
technicalDuration	técnicos
educationalInteractivityType	estruturais
educationalLearningResourceType	estruturais
educationalInteractivityLevel	estruturais
educationalSemanticDensity	estruturais
educationalIntendedEndUserRole	estruturais
educationalContext	estruturais
educationalTypicalAgeRange	estruturais

educationalDifficulty	estruturais
educationalTypicalLearningTime	estruturais
educationalDescription	estruturais
educationalLanguage	estruturais
rightsCost	administrativos
rightsCopyrightandOtherRestrictions	administrativos
rightsDescription	administrativos
relationKind	estruturais
relationResourceIdentifierCatalog	estruturais
relationResourceIdentifierEntry	estruturais
relationResourceDescription	estruturais
annotationEntity	preservação
annotationDate	preservação
annotationDescription	preservação
classificationPurpose	estruturais
classificationTaxonPathSource	estruturais
classificationTaxonPathTaxonId	estruturais
classificationTaxonPathTaxonEntry	estruturais
classificationDescription	estruturais
classificationKeyword	estruturais

Fonte: elaborado pela autora (2023) a partir do padrão LOM-IEEE (2020) e Barbedo, Corujo e Sant'ana (2011)

É possível observar que itens de dados do LOM-IEEE transitam por todas as categorias funcionais, onde as nove categorias principais do padrão acabam por ditar se estes dados são descritivos ou de identificação, estruturais, administrativos, técnicos ou de preservação.

6.2 OBJETOS DE APRENDIZAGEM DO REPOSITÓRIO ARCA ORGANIZADOS CONFORME METADADOS DO PADRÃO LOM-IEEE

Os motivos da escolha do repositório ARCA foram apresentados na Seção 5.3.2. Após a limpeza e formatação dos dados de dez OAs escolhidos, foi criada uma planilha auxiliar para uma organização prévia deles, conforme mostra a Figura 26. Como o padrão de metadados utilizado pelo repositório ARCA é o *DublinCore*, foi fácil comparar as similaridades entre os dois padrões e identificar na base de dados, as categorias e subcategorias do padrão LOM.

Figura 26 – Planilha com representação dos OA seguindo o padrão LOM

A	1 General								
	1.1 Identifier		1.2 Title	1.3 Language	1.4 Description	1.5 Keyword	1.6 Coverage	1.7 Structure	1.8 Aggregation Level
	1.1.1 Catalog	1.1.2 Entry							
ID	generalIdentifierCatalog	generalIdentifierEntry	generalTitle	generalLanguage	generalDescription	generalKeyword	generalCoverage	generalStructure	generalAggregationLevel
OA001	"URI"	"https://www.arca.fiocruz.br/handle/arca/10001"	"Aprendizagem em rede"	"pt"	"Apresenta uma arquitetura de aprendizagem em rede"	"Aprendizagem em Rede", "Brasil"	"Brasil"	Atômico	1
OA002	"URI"	"https://www.arca.fiocruz.br/handle/arca/10002"	"A Pesquisa e a inovação"	"pt"	"Apresentação de Criação e a biodiversidade sob a ótica da inovação"	"Gestão de Ciência", "pt", "Brasil"	"Brasil"	Atômico	1
OA003	"URI"	"https://www.arca.fiocruz.br/handle/arca/10003"	"Museu mínimo do Rio de Janeiro"	"pt"	"Este material faz parte de um projeto de museu mínimo do Rio de Janeiro"	"Unidades de Terapia Intensiva", "pt", "Brasil"	"Brasil"	Atômico	1
OA004	"URI"	"https://www.arca.fiocruz.br/handle/arca/10004"	"Fiocruz no Ar - Os ritos da cidade"	"pt"	"Podcast que integra conteúdos de áudio e vídeo sobre a história da Fiocruz"	"Antibióticos", "pt", "medicina", "Brasil"	"Brasil"	Coleção	2
OA005	"URI"	"https://www.arca.fiocruz.br/handle/arca/10005"	"Fiocruz no Ar - Os ritos da cidade"	"pt"	"Vinheta de apresentação do podcast 'Fiocruz no Ar - Os ritos da cidade'"	"Podcast", "Fiocruz no ar", "pt", "Brasil"	"Brasil"	Coleção	2
OA006	"URI"	"https://www.arca.fiocruz.br/handle/arca/10006"	"Curso Profissional da Saúde"	"pt"	"o curso tem como principal objetivo o trabalho desenvolvido nos cursos de graduação em uma organização."	"repositório", "pt", "profissional", "Brasil"	"Brasil"	Coleção	4
OA007	"URI"	"https://www.arca.fiocruz.br/handle/arca/10007"	"Curso Profissional da Saúde"	"pt"	"Figura para divulgação do curso 'Curso Profissional da Saúde'"	"repositório", "pt", "profissional", "Brasil"	"Brasil"	Coleção	2
OA008	"URI"	"https://www.arca.fiocruz.br/handle/arca/10008"	"Desafio Agora 2030"	"pt"	"O Desafio Agora 2030 é um jogo que aborda temas como sustentabilidade, inovação e saúde"	"jogo", "pt", "sustentabilidade", "Brasil"	"Brasil"	Coleção	3
OA009	"URI"	"https://www.arca.fiocruz.br/handle/arca/10009"	"campus virtual de saúde"	"pt"	"Capítulo do E-book 'Campus Virtual de Saúde'"	"Gestão do Conhecimento", "pt", "Brasil"	"Brasil"	Atômico	1

Fonte: Dados de pesquisa (2023)

Na figura 26 observa-se um recorte dos dez OAs representados com os metadados da categoria General do LOM, onde cada um dos metadados recebeu um nome com a junção das categorias e subcategorias conforme mapa mental apresentado na Seção 6.1. Há de se ressaltar que este processo demandou bastante tempo, pois foram preenchidas as 58 categorias do LOM. Para conseguir catalogar apenas um OA, utilizando todas as categorias do padrão, é necessário que o profissional responsável, além das competências relacionadas à biblioteconomia, também tenha a habilidade para executar o objeto e entender suas características técnicas, assim como compreender os objetivos educacionais do mesmo para que todos os campos sejam preenchidos adequadamente.

Em seguida os OAs passaram por um processo de preparação e organização mais detalhada com apoio do software OpenRefine, cujo recorte do resultado pode ser visto na Figura 27. Diversas ações foram realizadas com o intuito, principalmente, em deixar os dados mais separados para que depois pudessem ser coletados no processo de mapeamento de forma mais fácil.

Figura 27 – Dados da amostra de OAs trabalhados no OpenRefine

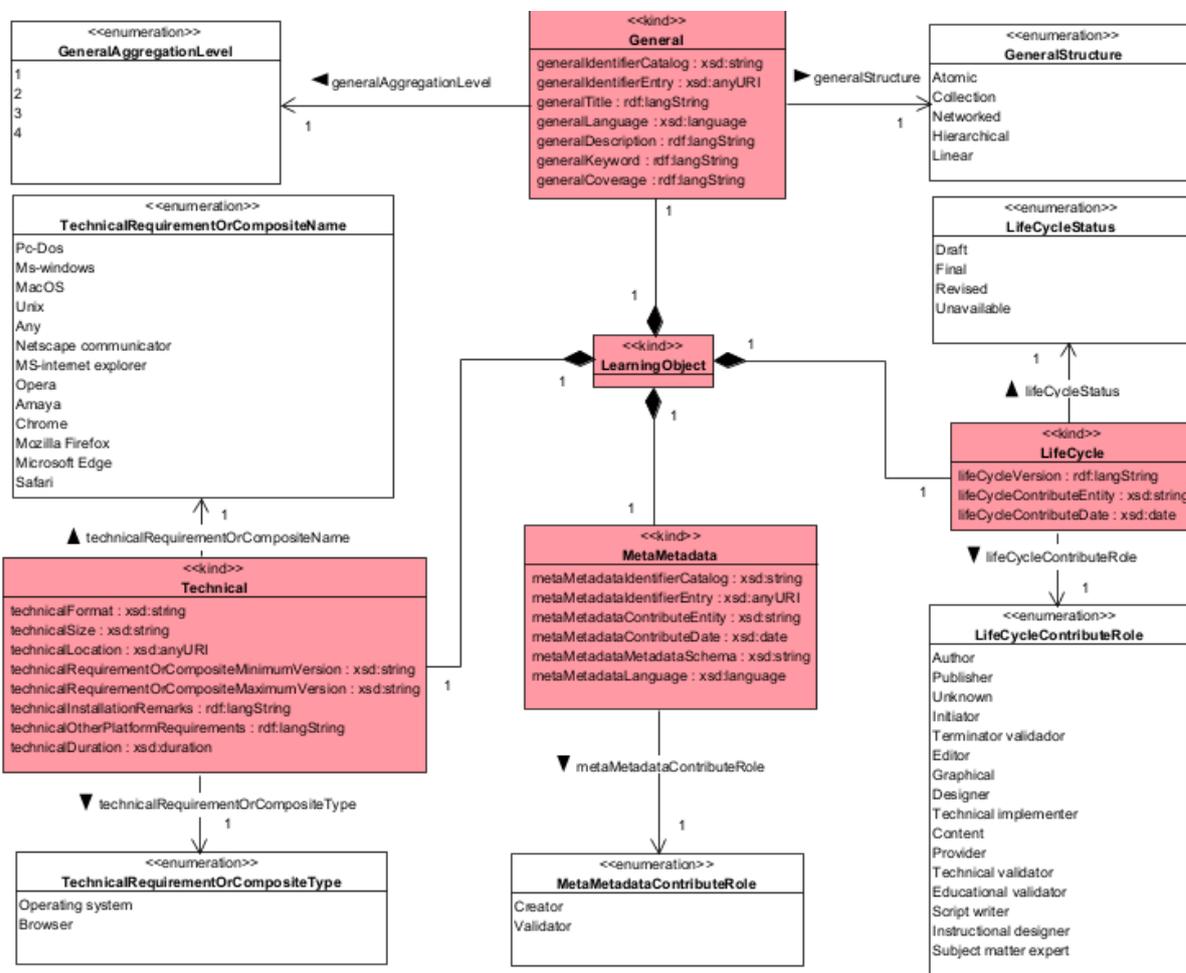
Column	generalIdentifierCatalog	generalIdentifierEntry	generalTitleLanguage	generalTitle	generalLanguage	generalDescriptionLanguage	generalDescription
1. OAO01	URI	https://www.arca.fiocruz.br/handle/icict/32122	pt-BR	Aprendizagem em rede	pt-BR	pt-BR	Apresenta uma arquitetura de aprendizagem em Rede, contemplando três dimensões: que interagem, se complementam e tem como objetivo o conhecimento aplicado, a geração de novas tecnologias e inovação para o desenvolvimento profissional e institucional.

Fonte: Dados de pesquisa (2023)

6.3 MODELO DE DESIGN PARA A ONTOLOGIA OPERACIONAL DE DOMÍNIO

O modelo de design proposto, baseado no padrão LOM-IEEE está representado nas Figuras 28, 29 e 30 diagramados com a linguagem OntoUML. Como o LOM tem muitas categorias e subcategorias, optou-se por dividir o modelo de design em três partes para melhor visualização e interpretação do seu conteúdo.

Figura 28 – Modelo de design para a ontologia operacional de domínio do padrão LOM-IEEE - Categorias General, Life Cycle, Technical e Meta Metadata.



Fonte: elaborado pela autora (2023) a partir do padrão LOM-IEEE (IEEE, 2020)

Na Figura 28 estão representadas as categorias General, Life Cycle, Technical e Meta Metadata. Cada uma das classes recebeu o estereótipo UFO “<<kind>>” para representar uma entidade forte. Além disso, foi criada uma classe para representar o OA, também com o estereótipo “<<kind>>”, onde foram associadas todas as nove classes a ela com a relação de “composição” (losango preto) e cardinalidade de 1 para 1, isto é, para cada OA existe exatamente um grupo de elementos de cada uma das categorias.

Acrescentou-se os *DatatypeProperty*, isto é, os atributos e tipos de dados, para cada uma das classes ligadas ao OA. Para cada atributo foi especificado o seu tipo de dado, conforme sugere o padrão LOM-IEEE. Contudo, os tipos foram adaptados

para uso em vocabulários controlados que seguem o padrão RDF para facilitar a posterior implementação da ontologia, conforme mostra o Quadro 9.

Quadro 9 - Equivalências entre tipos LOM-IEEE e RDF

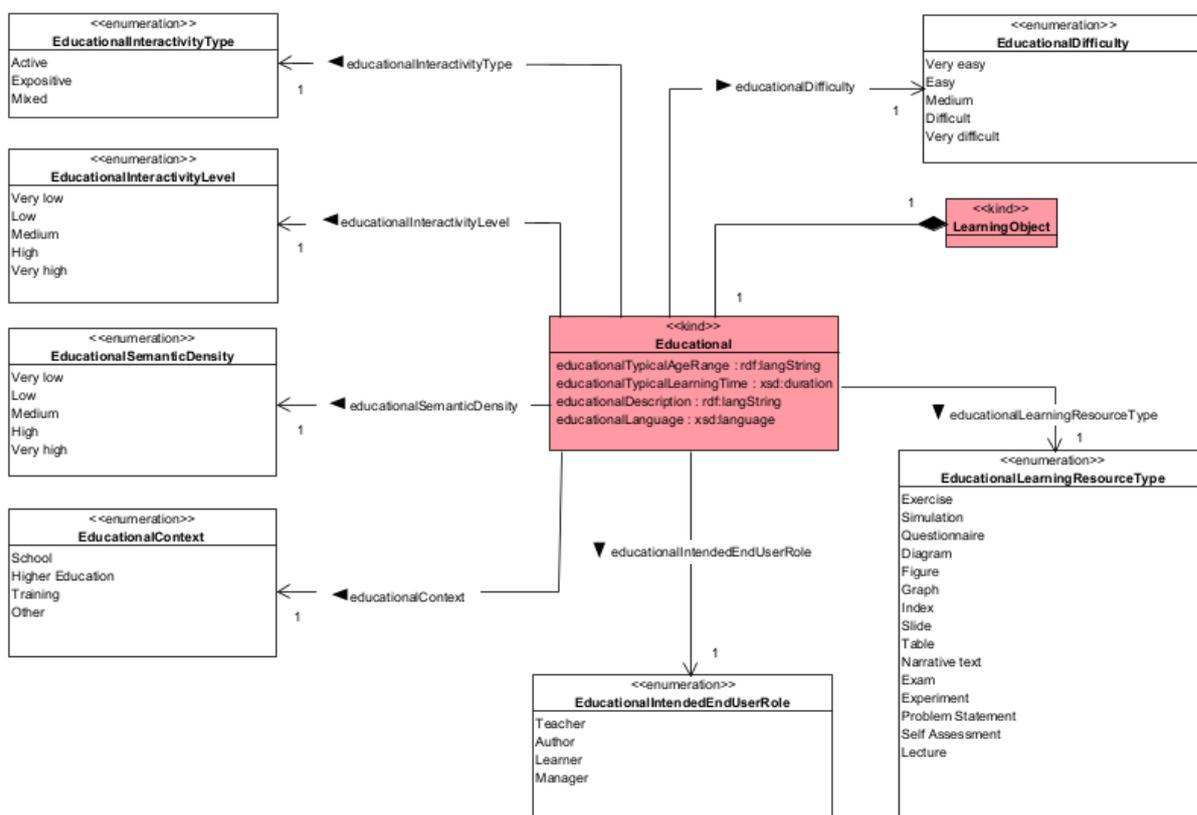
Tipo especificado no padrão LOM-IEEE	Tipo correspondente em RDF
CharacterString	xsd:string
CharacterString (para URIs)	xsd:anyURI
LangString	rdf:langString
DateTime	xsd:date
Duration	xsd:duration
Language	xsd:language

Fonte: elaborado pela autora (2023)

As classes que representam as categorias que apresentam vocabulários controlados têm um tratamento especial, pois representam uma lista de termos pertencentes a um vocabulário. Esses vocabulários foram tratados como *enumerations* nos diagramas para representar para cada uma das categorias do LOM. Isto foi feito observando-se a informação sobre “*value space*” disponível na norma IEEE (2020) para cada atributo.

Esse mesmo procedimento de organização das categorias com suas listas de termos de vocabulário, foi realizado para os outros dois diagramas, Figuras 32 e 33.

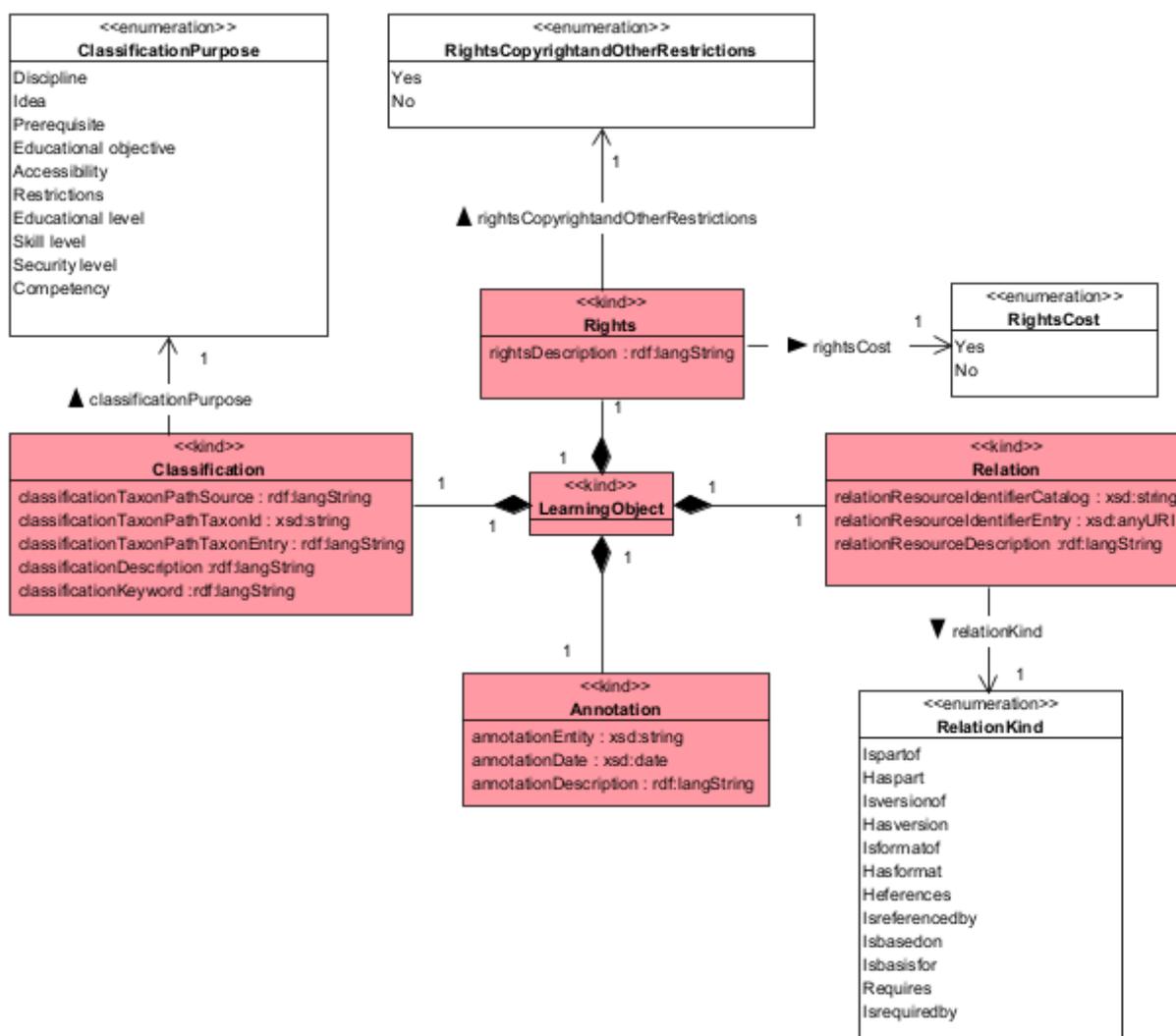
Figura 29 – Modelo de design para a ontologia operacional de domínio Diagrama estrutural da Ontologia do padrão LOM-IEEE- Categoria Educational



Fonte: elaborado pela autora (2023) a partir do padrão LOM-IEEE (IEEE, 2020)

A categoria “*Educational*”, Figura 29, é a que possui maior número de elementos de dados, o que enfatiza a característica e o objetivo educacional do padrão LOM. Na Figura 30 estão representadas as categorias Rights, Relation, Annotation e Classification.

Figura 30 - Modelo de design para a ontologia operacional de domínio Diagrama estrutural da Ontologia do padrão LOM-IEEE - Categorias Rigths, Relation, Annotation e Classification



Fonte: elaborado pela autora (2023) a partir do padrão LOM-IEEE (IEEE, 2020)

É importante destacar que a elaboração desse modelo de design, conforme a metodologia de Falbo (2014), necessitaria do desenvolvimento das etapas anteriores de identificação dos requisitos e a criação da ontologia de referência de domínio. Contudo isso não foi necessário, pois o padrão LOM-IEEE escolhido já formulou com detalhes todos os elementos que puderam servir de base para a criação do modelo de design apresentado nessa subseção.

6.4 ONTOLOGIA OPERACIONAL DE DOMÍNIO

Com base no modelo de design apresentado na seção anterior, foi escrito o código da ontologia operacional no formato de triplas RDF na linguagem Turtle. O código completo possui 1199 linhas e está disponível no link⁶¹. As Figuras 31, 32 e 33 destacam recortes importantes que serão explicados adiante.

A Figura 31 contém os namespaces usados na ontologia. São endereços para vocabulários e/ou ontologias usadas como referências externas na escrita do código. Dois namespaces são internos e foram especificados ainda como endereços fictícios, pois a ontologia não se encontra publicada, mas apenas na máquina local da pesquisadora. São os namespaces das linhas 1 e 12 que possuem, respectivamente, o local da ontologia e o local dos dados.

Figura 31 - Namespaces usados na ontologia

```

1 @prefix : <http://dados.ufes.br/lomieeee/onto#> .
2 @prefix grddl: <http://www.w3.org/2003/g/data-view#> .
3 @prefix dct: <http://purl.org/dc/terms/> .
4 @prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
5 @prefix owl: <http://www.w3.org/2002/07/owl#> .
6 @prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
7 @prefix skos: <http://www.w3.org/2004/02/skos/core#> .
8 @prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
9 @prefix dc: <http://purl.org/dc/elements/1.1/> .
10 @prefix cc: <http://creativecommons.org/ns#> .
11 @prefix wdo: <http://www.wikidata.org/ontology#> .
12 @prefix data: <http://dados.ufes.br/lomieeee/data#> .
13 @prefix gn: <https://sws.geonames.org/> .
14 @prefix wde: <http://www.wikidata.org/entity/> .

```

Fonte: Dados de pesquisa (2023)

A Figura 32 apresenta os meta-metadados da ontologia, desde o nome da ontologia e sua descrição, passando pela autoria e data de criação, até a licença '4.0' da Creative Commons.

⁶¹ Pasta contendo vários arquivos resultantes do desenvolvimento da pesquisa, entre eles, a ontologia operacional de domínio completa, cujo nome de identificação é 'Ontologia operacional - LOM IEE.ttl'. Disponível em: https://github.com/ViviLopesAlvarenga/Ontologia_LOM_IEEE/tree/main

Figura 32 - Meta-metadados da ontologia

```

18 <http://dados.ufes.br/lomieee#> a owl:Ontology;
19   rdfs:label "OntoLOM-IEEE";
20   rdfs:comment "Ontologia desenvolvida como parte da dissertação de Mestrado do
21   PPGCI/UFES da aluna Viviane Bessa Lopes Alvarenga"@pt-BR;
22   owl:versionInfo "1.0";
23   dct:title "Ontologia aplicada a Objetos de Aprendizagem com base no Padrão LOM da
24   IEEE/2020"@pt-BR;
25   dct:contributor "Henrique M. Cristovao";
26   dct:description "Ontologia operacional aplicada a Objetos de Aprendizagem com base
27   no Padrão LOM da IEEE/2020"@pt-BR;
28   dct:creator "Viviane Bessa Lopes Alvarenga";
29   dct:created "2023-02-01"^^xsd:date;
30   dct:modified "2023-03-07"^^xsd:date;
31   cc:licence "https://creativecommons.org/licenses/by-sa/4.0/"^^xsd:anyURI .

```

Fonte: Dados de pesquisa (2023)

A Figura 33 mostra um recorte do código contendo a classe principal que constitui o OA e as três primeiras propriedades OWL (ObjectProperty) que conectam o OA a um elemento de vocabulário SKOS (skos:Concept), ':generalStructure', ':generalAggregationLevel' e ':LifecycleStatus'. Todas as propriedades que conectam o OA a um vocabulário, especificadas nos três diagramas do modelo de design, foram ligadas diretamente à classe principal, ':LearningObject', conforme especificado no domínio mostrado nas linhas 655, 661 e 667 pela tripla 'rdfs:domain :LearningObject'. Na implementação da ontologia operacional observou-se que não foi necessário criar as classes intermediárias representadas nos diagramas da seção 6.3. Esse tipo de simplificação ou adaptação é bastante comum quando se implementa um código a partir de um modelo.

Figura 33 - Recorte do código contendo a classe principal que constitui o OA e as três primeiras propriedades OWL

```

30 :LearningObject a owl:Class;
31   rdfs:label "Learning Object"@en, "Objeto de Aprendizagem"@pt-BR;
32   owl:equivalentClass wdo:Q1062461 .

651 :generalStruture a owl:AsymmetricProperty, owl:ObjectProperty, owl:FunctionalProperty,
652   owl:IrreflexiveProperty;
653   rdfs:label "tem estrutura geral"@pt-BR, "has general structure"@en;
654   rdfs:range skos:Concept;
655   rdfs:domain :LearningObject .
656
657 :generalAggregationLevel a owl:AsymmetricProperty, owl:ObjectProperty, owl:
658   FunctionalProperty,
659   owl:IrreflexiveProperty;
660   rdfs:label "tem nível de granularidade"@pt-BR, "has level of granularity"@en;
661   rdfs:range skos:Concept;
662   rdfs:domain :LearningObject .
663
664 :lifeCycleStatus a owl:AsymmetricProperty, owl:ObjectProperty, owl:FunctionalProperty,
665   owl:IrreflexiveProperty;
666   rdfs:label "tem status do ciclo de vida"@pt-BR, "has life cycle status"@en;
667   rdfs:range skos:Concept;
668   rdfs:domain :LearningObject .

```

Fonte: Dados de pesquisa (2023)

Foram implementados os 18 vocabulários do padrão LOM-IEEE, conforme especificado como '*enumerations*' no modelo de design, e no código representados por esquemas conceituais do SKOS (skos:ConceptScheme). A Figura 34 mostra um recorte parcial dos três primeiros esquemas conceituais SKOS, ':GeneralStructure', ':GeneralAggregationLevel' e ':LifeCycleStatus'. Os vocabulários são documentados por meio de predicados do SKOS denominados de 'skos:prefLabel' e 'skos:note', sendo que a etiqueta preferencial (skos:prefLabel) foi identificado na língua portuguesa e na língua inglesa.

Figura 34 – Recorte do código da implementação dos três primeiros vocabulários da ontologia operacional

```

34 :GeneralStruture a skos:ConceptScheme;
35   skos:prefLabel "General Struture"@en, "Estrutura Geral"@pt-BR;
36   skos:note "Estrutura organizacional subjacente do objeto de aprendizagem"@pt-BR .
37
38 :GeneralAggregationLevel a skos:ConceptScheme;
39   skos:prefLabel "General Aggregation Level"@en, "Nível de Granularidade"@pt-BR;
40   skos:note "Granularidade funcional desse objeto de aprendizagem"@pt-BR .
41
42 :LifeCycleStatus a skos:ConceptScheme;
43   skos:prefLabel "Life Cycle Status"@en, "Status do Ciclo de Vida"@pt-BR;
44   skos:note "status ou condição de conclusão deste objeto de aprendizagem"@pt-BR .

```

Fonte: Dados de pesquisa (2023)

A Figura 35 mostra um recorte parcial do código que representa os dois primeiros conceitos que, neste caso, pertencem ao vocabulário ':GeneralStruture'. Todos os termos dos 18 vocabulários do padrão LOM-IEEE foram implementados por meio de conceitos do SKOS (skos:Concept).

Figura 35 – Recorte do código da implementação dos três primeiros conceitos/termos dos vocabulários da ontologia operacional

```
106 :Atomic a skos:Concept;  
107   skos:inScheme :GeneralStruture;  
108   skos:topConceptOf :GeneralStruture;  
109   skos:prefLabel "Atomic"@en, "Atômico"@pt-BR .  
110  
111 :Collection a skos:Concept;  
112   skos:inScheme :GeneralStruture;  
113   skos:topConceptOf :GeneralStruture;  
114   skos:prefLabel "Collection"@en, "Coleção"@pt-BR .  
115  
116 :Networked a skos:Concept;  
117   skos:inScheme :GeneralStruture;  
118   skos:topConceptOf :GeneralStruture;  
119   skos:prefLabel "Networked"@en, "Em Rede"@pt-BR .
```

Fonte: Dados de pesquisa (2023)

A Figura 36 apresenta um recorte parcial do código contendo as três primeiras propriedades OWL que conectam o OA aos valores literais (owl:DatatypeProperty), ':generalIdentifierCatalog', ':generalIdentifierEntry', e ':generalTitle'. Esse tipo de propriedade (owl:DatatypeProperty) possui como faixa de valores destino (range) um tipo correspondente a um literal, conforme mostrado no Quadro 9.

Figura 36 – Recorte do código da implementação das três primeiras propriedades que conectam o OA aos valores literais

```

757 :generalIdentifierCatalog a owl:DatatypeProperty, owl:FunctionalProperty;
758   rdfs:label "identificador do catálogo"@pt-BR, "catalog identifier"@en;
759   rdfs:comment "Nome ou designador do esquema de identificação ou catalogação"@pt-BR;
760   rdfs:range xsd:string;
761   rdfs:domain :LearningObject .
762
763 :generalIdentifierEntry a owl:DatatypeProperty, owl:FunctionalProperty;
764   rdfs:label "entrada do Identificador"@pt-BR, "identifier entry"@en;
765   rdfs:comment "O valor do identificador dentro do esquema de identificação ou
766   catalogação que designa ou identifica este OA. Uma string específica de namespace"
767   @pt-BR;
768   rdfs:range xsd:anyURI;
769   rdfs:domain :LearningObject;
770   owl:equivalentProperty dc:identifier .
771
772 :generalTitle a owl:DatatypeProperty, owl:FunctionalProperty;
773   rdfs:label "título"@pt-BR, "title"@en;
774   rdfs:comment "Nome dado a este objeto de aprendizagem"@pt-BR;
775   rdfs:range rdf:langString;
776   rdfs:domain :LearningObject;
777   owl:equivalentProperty dc:title .

```

Fonte: Dados de pesquisa (2023)

A opção pelo desenvolvimento híbrido, utilizando-se de elementos do OWL e do SKOS, conforme tratado nos procedimentos metodológicos, utiliza-se de potencialidades das duas linguagens. As duas linguagens são intercambiáveis e o SKOS possui vários elementos derivados de OWL (STELLATO, 2012). Enquanto o OWL possui uma poderosa linguagem de modelagem de dados, o SKOS é ótimo em representações simples como lista de termos em vocabulários (MILES, BECHHOFER; 2009). Além disso, a configuração dos termos dos vocabulários como conceitos do SKOS permite o estabelecimento de uma futura interoperabilidade com termos pertencentes a outros vocabulários externos à ontologia.

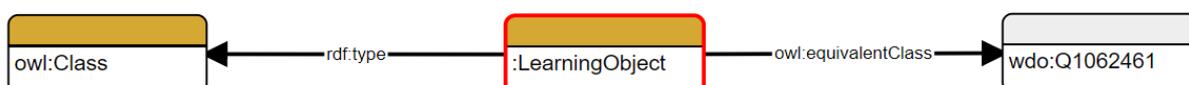
A ontologia operacional desenvolvida foi em seguida importada para a plataforma VocBench. Utilizando-se da própria interface do VocBench, alguns dados oriundos da relação de OAs mostrada na Figura 26, foram inseridos como teste. Os outros OAs foram inseridos por meio de mapeamento conforme descrito na seção 6.6.

6.5 INTEROPERABILIDADE DAS CAMADAS ESTRUTURAL E SEMÂNTICA

Nesta etapa buscou-se criar interoperabilidade para as camadas estrutural e semântica, segundo Zeng (2019), conforme abordado na seção 3.5 da Fundamentação Teórica, aplicada aos elementos da ontologia operacional. Para a

camada semântica, conseguiu-se apenas estabelecer equivalência entre o “:Learning Object” (owl:ClassObjet) e 'wdo:Q1062461', da Wikidata, que representa objetos de aprendizagem. Essa equivalência foi escrita com a propriedade OWL 'owl:equivalentClass', registrada na linha 32 do código da Figura 33. A Figura 37 apresenta via diagrama essa equivalência.

Figura 37 - Equivalência das classes ':LearningObject' e 'wdo:Q106261'



Fonte: elaboração própria, gerado com o software VocBench

Já na camada estrutural, utilizou-se a *Table A.1* do LOM-IEEE (IEEE, 2020) para mapear os 15 elementos de dados do *Dublin Core* correspondentes aos elementos do LOM-IEEE conforme mostra o Quadro 10.

Quadro 10 – Interoperabilidade da camada estrutural entre LOM-IEEE e *Dublin Core*

Elemento Dublin Core	Elemento LOM-IEEE	Relação Interoperável	Condição de Equivalência escrita em GREL no mapeamento
DC.Identifier	1.1.2:General.Identifier.Entry	dc:identifier ⇔ generalIdentifierEntry	-
DC.Title	1.2:General.Title	dc:title ⇔ generalTitle	-
DC.Language	1.3:General.Language	dc:language ⇔ generalLanguage	-
DC.Description	1.4:General.Description	dc:description ⇔ generalDescription	-
DC.Subject	1.5:General.Keyword or 9:Classification with 9.1:Classification	Equivalência Condicionada para generalKeyword	if(or(cells.classificationPurpose.value=="Discipline", cells.classificationPurpose.value=="Idea"), cells.generalKeyword1.val

	. Purpose equals "Discipline" or "Idea"		ue, null)
DC.Subject	1.5:General.Keyword or 9:Classification with 9.1:Classification . Purpose equals "Discipline" or "Idea"	Equivalência Condicionada para classificationKeyword	if(or(cells.classificationPurpose.value=="Discipline", cells.classificationPurpose.value=="Idea"), cells.classificationKeyword1.value, null)
DC.Coverage	1.6:General.Coverage	dc:coverage ⇔generalCoverage	-
DC.Type	5.2:Educational.LearningResourceType	dc:type ⇔educationalLearningResourceType	-
DC.Date	2.3.3:LifeCycle.Contribute.Date when 2.3.1:LifeCycle.Contribute.Role has a value of "Publisher".	Equivalência Condicionada	if(cells.lifeCycleContributeRole.value=="Publisher", cells.lifeCycleContributeDate.value, null)
DC.Creator	2.3.2:LifeCycle.Contribute.Entity when 2.3.1:LifeCycle.Contribute.Role has a value of "Author".	Equivalência Condicionada	if(cells.lifeCycleContributeRole.value=="Author", cells.lifeCycleContributeEntity.value, null)
DC.OtherContributor	2.3.2:LifeCycle.Contribute.Entity with the type of contribution specified in	Equivalência Condicionada	cells.lifeCycleContributeEntity.value + ", "+ cells.lifeCycleContributeRole.value

	2.3.1:LifeCycle. Contribute.Role.		
DC.Publisher	2.3.2:LifeCycle. Contribute.Entity when 2.3.1:LifeCycle. Contribute.Role has a value of “Publisher”.	Equivalência Condicionada	if(cells.lifeCycleContribute Role.value=="Publisher", cells.lifeCycleContributeEn tity.value, null)
DC.Format	4.1:Technical.Fo rmat	dc:format ⇔technicalForm at	-
DC.Rights	6.3:Rights.Descr iption	dc:rights ⇔ rightsDescription	-
DC.Relation	7.2.2:Relation.R esource.Descript ion	dc:relation ⇔relationResour ceDescription	-
DC.Source	7.2:Relation.Res ource when the value of 7.1:Relation.Kin d is “IsBasedOn.”	Equivalência Condicionada	if(cells.relationKind.value= ="IsBasedOn", cells.relationResource.valu e, Null)

Fonte: elaborado pela autora (2023) a partir do padrão LOM-IEEE (IEEE, 2020, p. 48)

Para se conseguir interoperar a ontologia feita no padrão LOM-IEEE para o Dublin Core, foi preciso, a partir do GraphDB criar o mapeamento tabular incluindo expressões GREL (General Refine Expression Language)⁶² para cada “*Predicate*” do Dublin Core, incluindo as expressões para condição equivalente no “*Object*” para cada elemento de dado do LOM-IEE conforme mostra a figura. Para os elementos sem condicionantes a relação foi feita diretamente nos elementos de dados sem necessidade de expressão GREL.

⁶²General Refine Expression Language (GREL). É uma linguagem de script similar ao javascript utilizada no ambiente do OpenRefine.

A Figura 38 mostra no código Turtle um recorte de duas propriedades, ':generalIdentifierEntry' e ':generalTitle' onde foi estabelecida a interoperabilidade com o Dublin Core, nas linhas 768 e 775, por meio do predicado 'owl:equivalentProperty'.

Figura 38 - Código Turtle com recorte de duas propriedades, ':generalIdentifierEntry' e ':generalTitle'

```

763 :generalIdentifierEntry a owl:DatatypeProperty, owl:FunctionalProperty;
764   rdfs:label "entrada do Identificador"@pt-BR, "identifier entry"@en;
765   rdfs:comment "O valor do identificador dentro do esquema de identificação ou
   catalogação que designa ou identifica este OA. Uma string específica de namespace"
   @pt-BR;
766   rdfs:range xsd:anyURI;
767   rdfs:domain :LearningObject;
768   owl:equivalentProperty dc:identifier .
769
770 :generalTitle a owl:DatatypeProperty, owl:FunctionalProperty;
771   rdfs:label "título"@pt-BR, "title"@en;
772   rdfs:comment "Nome dado a este objeto de aprendizagem"@pt-BR;
773   rdfs:range rdf:langString;
774   rdfs:domain :LearningObject;
775   owl:equivalentProperty dc:title .

```

Fonte: Dados da pesquisa (2023)

A Figura 39 apresenta, em formato de *knowledge graph*, o indivíduo 'data:OA001', referente ao primeiro objeto de aprendizagem cadastrado, com suas propriedades interoperáveis para o esquema de metadados Dublin Core.

generalTitle	bf:title
generalLanguage	bf:descriptionLanguage
generalDescription	bf:content
generalCoverage	bf:temporalCoverage
lifeCycleContributeEntity	bf:contribution
technicalFormat	bf:fileType
technicalSize	bf:fileSize
technicalLocation	bf:electronicLocator
annotationDate	bf:date
classificationTaxonPathSource	bf:classification

Fonte: elaborado pela autora (2023) a partir do padrão LOM-IEEE (IEEE, 2020) e Bibframe

Para se conseguir interoperar a ontologia feita no padrão LOM-IEEE para o Bibframe, foi preciso, a partir de alguns elementos de dados do LOM-IEE analisar a “*Property List*” (lista de propriedades) no Vocabulário BIBFRAME 2.0⁶³ e procurar por equivalências para ambos os padrões. Não foi intenção criar interoperabilidade com todos os elementos do LOM-IEEE, apenas alguns, de forma a demonstrar que a interoperabilidade entre os padrões é uma possibilidade.

6.6 MAPEAMENTO E BASE DE DADOS LIGADOS RDF

Nesta etapa todos os dados da amostra de OAs apresentados na Figura 26 foram mapeados no GraphDB tendo como base a ontologia operacional de domínio implementada para a criação de uma base de dados ligados RDF. O mapeamento foi feito utilizando-se os 58 elementos de dados constantes no LOM-IEEE conforme planilha de dados mostrados nas Figuras 26 e 27. A interface do sistema de mapeamento do GraphDB é mostrada na Figura 40.

⁶³ Disponível em: <https://id.loc.gov/ontologies/bibframe.html>

Figura 40 – Mapeamento dos dados dos OAs

The screenshot shows the OntoRefine web interface. At the top, there's a search bar and a dropdown menu set to 'LOM-IEEE_V01'. Below that, there are buttons for 'Save', 'Download JSON', 'Upload JSON', 'RDF', and 'SPARQL'. A 'Both' button is also visible. The main area contains a grid of property names: 'Column', 'gener ... talog', 'gener ... Entry', 'gener ... guage', 'generalTitle', 'generalLanguage', 'gener ... guage', 'gener ... ption', 'gener ... guage', 'generalKeyword1', 'gener ... guage', 'generalKeyword2', 'gener ... guage', 'generalKeyword3', 'gener ... guage', 'generalKeyword4', 'gener ... guage', 'generalKeyword5', 'gener ... guage', 'generalCoverage', and 'gener ... cture'. Below the grid is a URL field with 'http://example.com/base/'. Underneath, there's a section for repository prefixes with buttons for 'data', 'dc', 'rdf', and 'xsd'. The bottom part of the interface shows a table of mappings:

data: Column	<IRI>	a	<IRI>	: LearningObject	<IRI>
data: OA001		a		: LearningObject	
	dc: title	<IRI>		generalTitle	"Literal"
				general ... a	@Language

Fonte: elaboração própria com dados da pesquisa (2023) no software GraphDB

Na interface do sistema de mapeamento do GraphDB, cada elemento de dado do LOM-IEE assim como seus equivalentes no *Dublin Core* foram mapeados como *Predicates* e os dados dos OAs da amostra constantes no arquivo CSV gerado pelo OpenRefine entraram como *Objects* onde foi informado o *RDF Type (IRI)* ou (*Literal*), sendo *IRI* os que têm vocabulários e *Literal* os que são do tipo *string*, *langstring*, *data*, *anyURI*, *duration* e *language* conforme mostra o Quadro 9 de tipos equivalentes. Como mostra a Figura 41.

Figura 41 – Mapeamento dos dados dos OAs - Detalhe

Object RDF Value Mapping

RDF Type IRI **Literal** Unique Bnode Value Bnode

Source **Column** Constant Record ID Row Index GREL

Column Column

Literal attributes Datatype **Language**

Language source **Column** Constant GREL

Column Column

Fonte: elaboração própria com dados da pesquisa (2023) no software GraphDB

Foi informado o *Literal Attributes* como “constant” para *xsd:date*, *xsd:anyURI*, *xsd:language* para seguir todos os tipos colocados pelo VocBench no OA “OA001” no arquivo TTL da ontologia operacional conforme mostra a Figura 42.

Figura 42 – Mapeamento os dados dos OAs com xsd - Detalhe

Object RDF Value Mapping

Source **Column** Constant Record ID Row Index GREL

Column Column

Literal attributes **Datatype** Language

Datatype source Column **Constant** GREL

Constant Raw IRI Prefix Constant

Fonte: elaboração própria com dados da pesquisa (2023) no software GraphDB

E para os elementos que tem informação de idioma, foi criada uma coluna no OpenRefine com o mesmo nome do elemento acrescentando a palavra *Language* exemplo: `elementotechnicalInstallationRemarks` com outra `technicalInstallationRemarksLanguage` para informar “@pt-BR” conforme é mostrado na Figura 43.

Figura 43 – Mapeamento dos dados dos OAs - Idioma

Object RDF Value Mapping

Column	Column
	<u>technicalInstallationRemarks</u>
<hr/>	
Literal attributes	Datatype Language
Language source	Column Constant GREL
Column	Column
	<u>technicalInstallationRemarksLanguage</u>

Fonte: elaboração própria com dados da pesquisa (2023) no software GraphDB

O mapeamento realizado foi registrado em um arquivo do tipo JSON. A Figura 44 mostra um recorte desse arquivo para a situação do mapeamento do predicado 'dc:title' da ontologia operacional, linhas 14 a 23, para os dados da coluna 'generalTitle' do arquivo CSV gerado pelo OpenRefine, linhas 24 a 39.

Figura 44 - Recorte do mapeamento realizado com o GraphDB, registrado no formato JSON

```
10  "subjectMappings": [  
11    {  
12      "propertyMappings": [  
13        {  
14          "property": {  
15            "transformation": {  
16              "expression": "dc",  
17              "language": "prefix"  
18            },  
19            "valueSource": {  
20              "source": "constant",  
21              "constant": "title"  
22            }  
23          },  
24          "values": [  
25            {  
26              "valueSource": {  
27                "columnName": "generalTitle",  
28                "source": "column"  
29              },  
30              "valueType": {  
31                "type": "language_literal",  
32                "language": {  
33                  "valueSource": {  
34                    "columnName": "generalTitleLanguage",  
35                    "source": "column"  
36                  }  
37                }  
38              }  
39            ]  
40          }  
41    ],  
  ],  
}
```

Fonte: elaboração própria com dados da pesquisa (2023) no software GraphDB

A execução do mapeamento criou a base de dados ligados RDF, no formato Turtle, que conjuga a ontologia operacional de domínio com os dados dos OAs. Um recorte das primeiras triplas RDF dos dados, de um OA arbitrário, pode ser visto na Figura 45. Observa-se que as triplas desse OA partem todas do seu *Subject* identificado pela URI `data:OA003`, ou `http://dados.ufes.br/lomieeee/data#OA003`.

Figura 45 - Dados ligados RDF gerados pelo mapeamento no GraphDB

```

data:OA003 a :LearningObject;
  dc:title "Manuseio mínimo do recém-nascido"@pt-BR;
  dc:description "Este material faz parte do Portal de Boas Práticas em Saúde da Mulher, da Criança e do Adolescente do IFF/Fiocruz. O Portal oferece conteúdo sistematizado por especialistas e disponível em diversos formatos, como apresentação de slides e vídeos curtos, todos com links para as referências citadas. Esse material tem sido usado principalmente em atividades de educação permanente e de formação profissional em diferentes unidades do Sistema Único de Saúde (SUS), contribuindo para a revisão de processos clínicos norteados pelas boas práticas."@pt-BR;
  :generalStructure :Atomic;
  :generalAggregationLevel :Level1;
  :lifeCycleStatus :Final;
  :lifeCycleContributeRole :Author;
  :metaMetadataContributeRole :Creator;
  :technicalRequirementOrCompositeType :Browser;
  :technicalRequirementOrCompositeName :Any;
  :educationalInteractivityType :Expositive;
  :educationalLearningResourceType :Lecture;
  dc:type "Lecture";
  :educationalInteractivityLevel :Low;
  :educationalSemanticDensity :Medium;
  :educationalIntendedEndUserRole :Learner;
  :educationalContext :Training;
  :educationalDifficulty :Easy;
  :rightsCost :No;
  :rightsCopyrightandOtherRestrictions :Yes;
  :relationKind :Ispartof;
  :classificationPurpose :EducationalObjective;
  :generalIdentifierCatalog "URI";
  :generalIdentifierEntry "https://www.arca.fiocruz.br/handle/icict/30266"^^xsd:anyURI;
  dc:identifier "https://www.arca.fiocruz.br/handle/icict/30266"^^xsd:anyURI;
  :generalTitle "Manuseio mínimo do recém-nascido"@pt-BR;
  :generalLanguage "pt-BR"^^xsd:language;
  dc:language "pt-BR"^^dc:language;

```

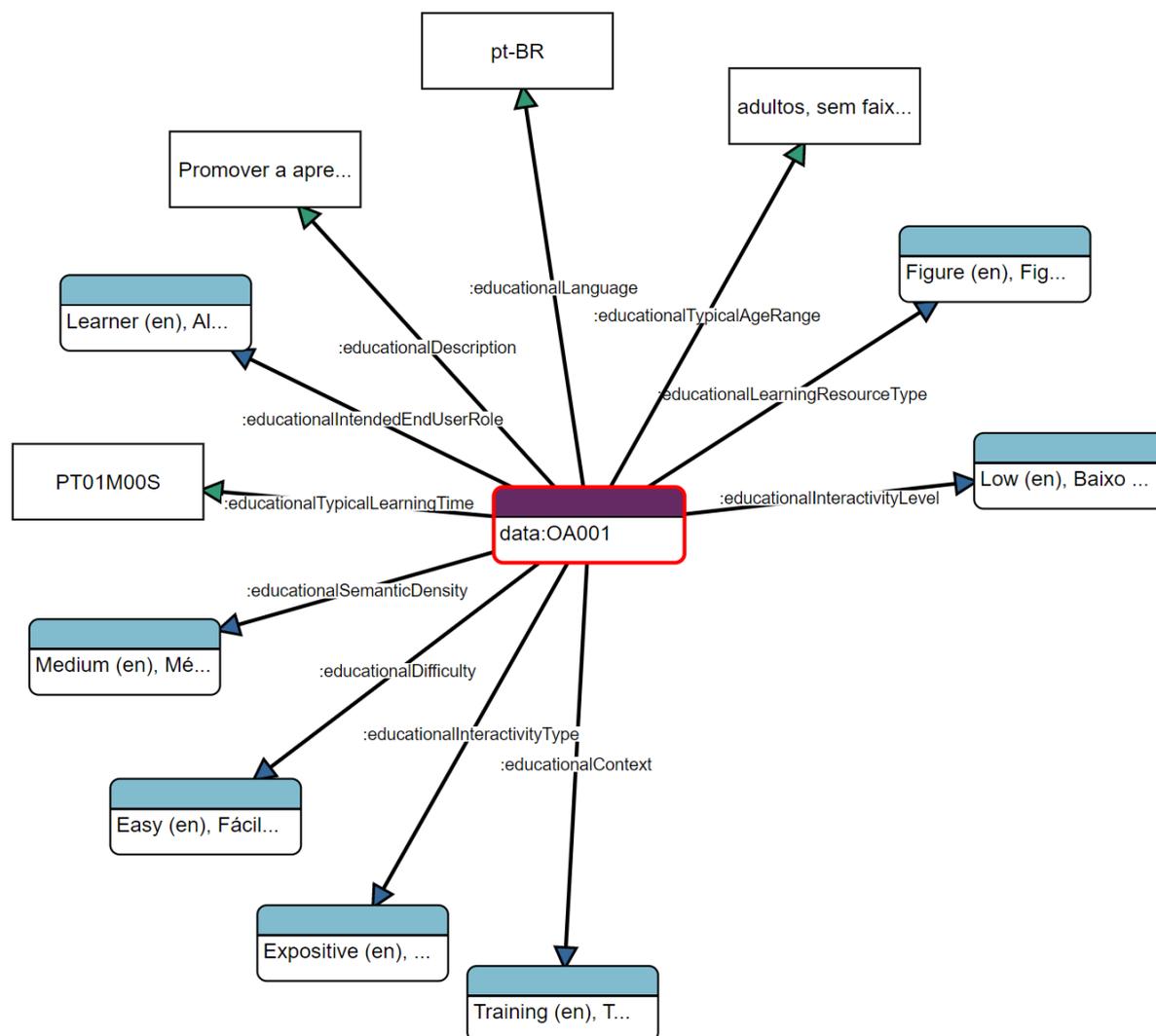
Fonte: Dados da pesquisa (2023)

6.7 CONSULTAS SPARQL

Após o processo de mapeamento, foram feitas algumas consultas inerentes ao OAs da amostra estudada como prova de conceito para a recuperação de informações. Apesar de existir uma infinidade de consultas que podem ser feitas, foram criadas algumas para explorar diversas situações de recuperação de informação sobre a base de dados ligados RDF.

Inicialmente foram realizadas algumas consultas por meio da interface gráfica do software VocBench para obtenção de resultados visuais por meio de *knowledge graphs*. São resultados mais fáceis de se fazer uma inspeção visual para obter uma visão mais geral ou bem específica de algumas situações da base.

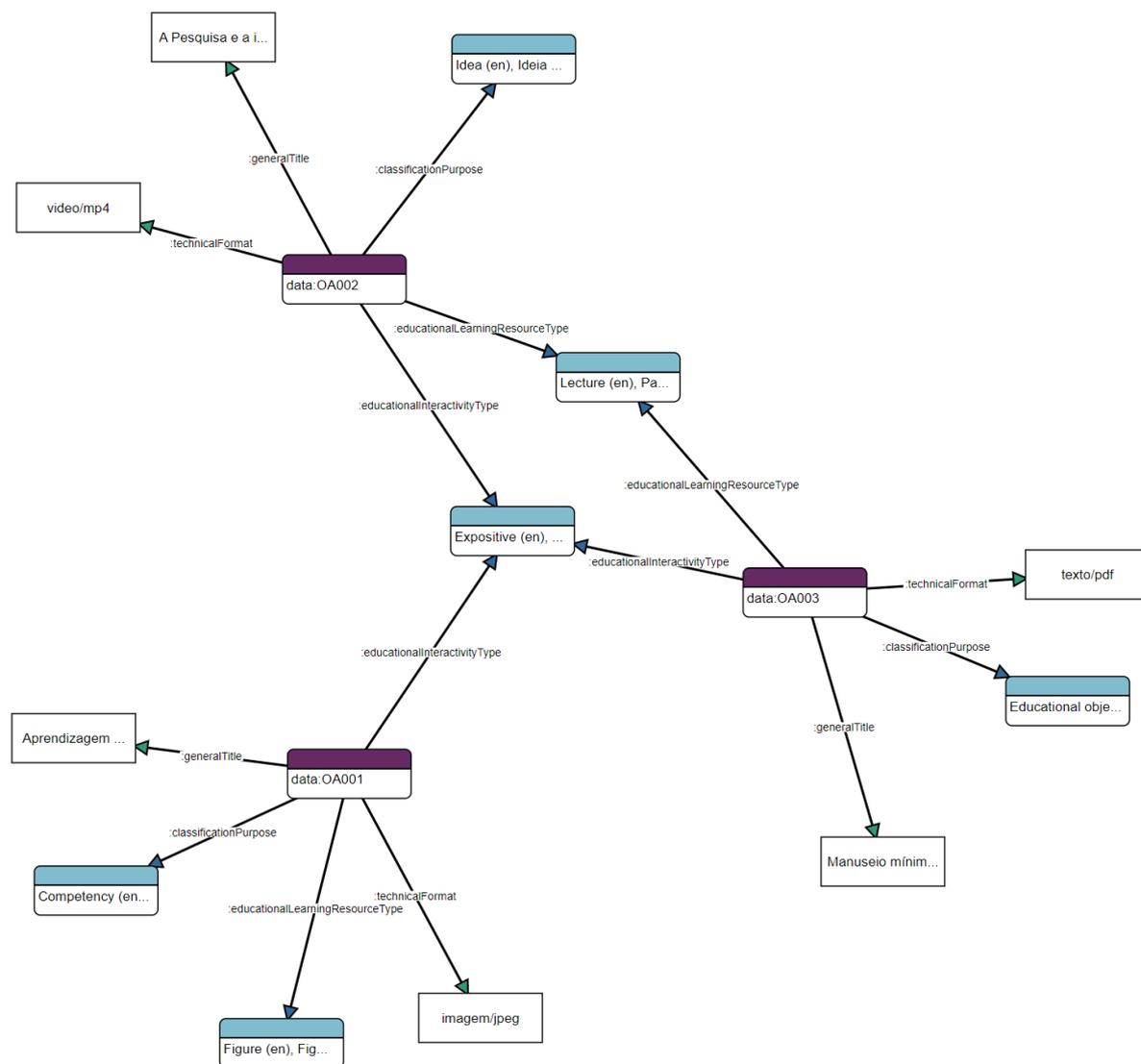
Figura 46 - Características educacionais do objeto de aprendizagem 'data:OA001'



Fonte: elaboração própria com dados da pesquisa (2023) no software VocBench

Na Figura 46 é possível verificar todas as características educacionais inerentes ao OA "OA001".

Figura 48 - Relação entre algumas propriedades dos objetos de aprendizagem 'data:OA001', 'data:OA002' e 'data:OA003'



Fonte: elaboração própria com dados da pesquisa (2023) no software VocBench

Figura 49 – Consulta SPARQL com nome, tipo, formato e identificador

```

1 # O nome, tipo, formato e identificador de todos os OAs
2 PREFIX : <http://dados.ufes.br/lomieeee/onto#>
3 PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
4 SELECT ?titulo ?tipo ?formato ?identificador
5 WHERE {
6     ?oa :generalTitle ?titulo ;
7         :educationalLearningResourceType ?recursoTipo ;
8         :technicalFormat ?formato ;
9         :generalIdentifierEntry ?identificador .
10    ?recursoTipo skos:prefLabel ?tipo .
11    FILTER(LANG(?tipo) = "pt")
12 }
13 LIMIT 100
14

```

Fonte: Dados da pesquisa (2023)

A Figura 49, mostra uma consulta feita para retornar os dados em formato tabular, essa foi feita com o intuito de retornar o nome, tipo, formato e identificador de todos os OAs cadastrados. O resultado é mostrado na Figura 50.

Figura 50 – Resultado da consulta SPARQL com nome, tipo, formato e identificador

	titulo	tipo	formato	identificador
1	"Aprendizagem em rede"@pt-BR	"Figura"@pt	"imagem/jpeg"	"https://www.arca.fiocruz.br/handle/ici ct/32122"xsd:anyURI
2	"A Pesquisa e a inovação da biodiversidade brasileira"@pt-BR	"Palestra"@pt	"video/mp4"	"https://www.arca.fiocruz.br/handle/ici ct/20416"xsd:anyURI
3	"Manuseio mínimo do recém-nascido"@pt-BR	"Palestra"@pt	"texto/pdf"	"https://www.arca.fiocruz.br/handle/ici ct/30266"xsd:anyURI
4	"Fiocruz no Ar - Os riscos do consumo de antibióticos sem receita médica"@pt-BR	"Entrevista"@pt	"áudio/mp3"	"https://www.arca.fiocruz.br/handle/ici ct/33461"xsd:anyURI
5	"Fiocruz no Ar - Os riscos do consumo de antibióticos sem receita médica"@pt-BR	"Figura"@pt	"imagem/png"	"https://www.arca.fiocruz.br/handle/ici ct/33461"xsd:anyURI
6	"Curso Profissional da Informação e Repositórios"@pt-BR	"Slide"@pt	"texto/pdf"	"https://www.arca.fiocruz.br/handle/ici ct/48701"xsd:anyURI
7	"Curso Profissional da Informação e Repositórios"@pt-BR	"Figura"@pt	"imagem/jpg"	"https://www.arca.fiocruz.br/handle/ici ct/48701"xsd:anyURI
8	"Desafio Ágora 2030"@pt-BR	"Jogo"@pt	"application/zip"	"https://www.arca.fiocruz.br/handle/ici ct/35768"xsd:anyURI

Fonte: Dados da pesquisa (2023)

É importante ressaltar que o filtro de idioma "LANG" fez toda diferença, pois como a ontologia foi implementada com o skos:prefLabel em português e inglês, sem este filtro o sistema retorna valores duplicados para o mesmo OA. Pode-se observar na

Figura 50 que os valores apresentados são apenas em idioma português brasileiro (@pt-BR). Caso fosse necessária a recuperação na língua inglesa, bastaria especificar isso no filtro da linha 11 da consulta mostrada na Figura 49.

Figura 51 – Consulta SPARQL dos OAs que trabalham com "gestão do conhecimento" e que possuam nível de interatividade baixo (Low) e nível de dificuldade fácil (Easy)

```

2 PREFIX : <http://dados.ufes.br/lomieeee/onto#>
3 PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
4 SELECT ?titulo ?tipo ?formato ?identificador
5 WHERE {
6   ?oa :generalTitle ?titulo ;
7       :educationalLearningResourceType ?recursoTipo ;
8       :technicalFormat ?formato ;
9       :generalIdentifierEntry ?identificador ;
10      :educationalInteractivityLevel :Low ;
11      :educationalDifficulty :Easy ;
12      :classificationTaxonPathTaxonEntry "gestão do conhecimento"@pt-BR .
13      ?recursoTipo skos:prefLabel ?tipo .
14      FILTER(LANG(?tipo) = "pt")
15 }
16 LIMIT 100
17

```

Fonte: Dados da pesquisa (2023)

Na consulta da Figura 51 buscou-se uma informação de um vocabulário controlado (SKOS) da ontologia para recuperar os OAs que trabalham com "gestão do conhecimento" e que possuam nível de interatividade baixo (Low) e nível de dificuldade fácil (Easy). A busca retornou apenas um OA que satisfaz o critério estabelecido conforme mostra a Figura 52.

Figura 52 – Resultado da consulta SPARQL dos OAs que trabalham com "gestão do conhecimento" e que possuam nível de interatividade baixo (Low) e nível de dificuldade fácil (Easy)

	titulo	tipo	formato	identificador
1	"Aprendizagem em rede"@pt-BR	"Figura"@pt	"imagem/jpeg"	"https://www.arca.fiocruz.br/handle/ici/ct/32122"@xsd:anyURI

Fonte: Dados da pesquisa (2023)

A consulta da Figura 53 foi elaborada com vistas a recuperação de uma informação que é muito buscada nos bancos de dados e muito útil em se tratando de OAs que é a busca por termos nas palavras-chave e título.

Figura 53 – Consulta SPARQL dos OAs que tenham o termo 'saúde' no título ou nas palavras-chave

```

1 # OAs que tenham o termo 'saúde' no título ou nas palavras chave
2 PREFIX : <http://dados.ufes.br/lomieee/onto#>
3 SELECT DISTINCT ?titulo
4 WHERE {
5     ?oa :generalTitle ?titulo ;
6         :generalKeyword ?generalKeyword ;
7         :classificationKeyword ?classificationKeyword .
8
9     FILTER(CONTAINS(?titulo, "saúde"@pt-BR) ||
10            CONTAINS(?generalKeyword, "saúde"@pt-BR) ||
11            CONTAINS(?classificationKeyword, "saúde"@pt-BR)) .
12 }
13 LIMIT 100
14
15

```

Fonte: Dados da pesquisa (2023)

Conforme visto na Figura 53, foi necessário informar dois elementos de dados do LOM-IEEE para formular a consulta, pois tanto o *generalKeyword* como o *classificationKeyword* guardam informações sobre as palavras-chave dos OAs, assim como a consulta também foi feita pelo *generalTitle*.

Figura 54 – Resultado da consulta SPARQL dos OAs que tenham o termo 'saúde' no título ou nas palavras chave

	titulo
1	"A Pesquisa e a inovação da biodiversidade brasileira"@pt-BR
2	"campus virtual de saúde pública (cvsp/opas): uma rede de conhecimento e formação em saúde pública"@pt-BR
3	"curso inteligência artificial documental: aumento de eficiência e qualidade"@pt-BR

Fonte: Dados da pesquisa (2023)

Na Figura 54 é possível verificar a recuperação de 3 OAs que atendem ao critério da consulta feita.

A Figura 55 apresenta uma consulta SPARQL um pouco mais elaborada para recuperar os relacionamentos entre as características educacionais dos OAs cadastrados, sem dar ênfase aos objetos de aprendizagem, mas nas relações existentes entre eles. O Resultado, por meio de um *knowledge graph*, é mostrado na Figura 56. Destacam-se os relacionamentos entre o tipo de recurso educacional, nível de interatividade, nível de dificuldade e densidade semântica.

Figura 55 - Relacionamento entre as características educacionais dos OAs cujo usuário é aprendiz - consulta

```
1 # Relações entre algumas características
educacionais dos OAs cujo papel do usuário final é
aprendiz
2 PREFIX : <http://dados.ufes.br/lomieee/onto#>
3 CONSTRUCT {
4     ?educ1 :tipoRecurso_densidadeSemantica ?educ3 .
5     ?educ1 :tipoRecurso_tipoInteratividade ?educ2 .
6     ?educ1 :tipoRecurso_dificuldade ?educ4 .
7     ?educ1 :tipoRecurso_nivelInteratividade ?educ5
8     .
9 }
10 WHERE {
11     ?oa :educationalIntendedEndUserRole :Learner;
12     :educationalLearningResourceType ?educ1;
13     :educationalInteractivityType ?educ2;
14     :educationalSemanticDensity ?educ3;
15     :educationalDifficulty ?educ4;
16     :educationalInteractivityLevel ?educ5 .
}
```

Fonte: Dados da pesquisa (2023)

Figura 56 - Relacionamento entre as características educacionais dos OAs cujo usuário é aprendiz - resultado



Fonte: elaboração própria, com dados da pesquisa (2023) no software GraphDB

Na Figura 56 é possível observar novamente que o tipo de OA (Exemplo: um jogo) de acordo com o seu tipo de interatividade (ativo, expositivo etc...) influencia no grau de dificuldade e de densidade semântica junto ao(s) usuário(s) principal(is) para o(s) qual(is) este OA foi projetado. E a análise destas relações podem interferir no tipo de OA de escolha para determinados objetivos educacionais.

Como visto, é possível fazer inúmeras consultas com esta pequena amostra de apenas 10 itens, o que poderia ser replicado para repositórios de OAs com milhares de itens. Com essa prova de conceito realizada sobre a base de dados ligados RDF organizada segundo o padrão LOM-IEEE é possível especular que a organização do conhecimento, em função dos elementos apresentados ao longo da pesquisa, e realizada sobre a amostra de OAs torna-os encontráveis, acessíveis, interoperáveis e, conseqüentemente, reutilizáveis.

6.8 DISCUSSÕES GERAIS

Zeng (2008), conforme apresentado na Fundamentação Teórica, estabelece um modelo para Sistemas de Organização do Conhecimento - SOCs, onde os Tesouros, as Taxonomias e as Ontologias são considerados, os mais complexos desses sistemas, tendo em vista, as funções que desempenham no processo de organização e recuperação da informação, pois estabelecem relações hierárquicas, controlam sinônimos e eliminam ambigüidades. Assim, ao propor a implementação de uma ontologia operacional de domínio entre os tipos de ontologia elencados por Guarino (1998), conforme a Figura 3, para auxiliar a organização de OAs no contexto da Web semântica, fez-se com vistas a diminuição e ou eliminação das “ambigüidades” citadas por Zeng (2008), onde acredita-se que o processo de organização do conhecimento relativo aos OAs possa trazer benefícios na recuperação da informação. O estudo do padrão LOM-IEEE (2020) ora apresentado vai ao encontro do conceito sobre o uso de metadados postulado por Pöttker, Ferneda e Moreiro-González (2018, p.29) [...] descrição de dados que podem ser utilizados para identificar e descrever características comuns entre diferentes recursos informacionais com a finalidade de facilitar a interoperabilidade e a recuperação na Web.” Ao resgatar o que foi discutido na seção 3.2.3, alinhado com o experimento da presente pesquisa, é possível inferir que utilizar padrões de estrutura de metadados descritivos para catalogação, de forma a representar as características físicas e de conteúdo, conforme exposto por Foulonneau e Riley (2008), aplicado aos OAs, torna-se de suma importância para facilitar a interpretação desses registros para além das barreiras linguísticas, convertendo-os à forma que seja legível por máquina. Para validar esta importância, utilizou-se aqui o modelo

para representação de dados bibliográficos chamado Bibframe, citado na referida seção como um formato de intercâmbio bibliográfico criado com o propósito de integrar a catalogação à Web semântica, assim como atender as funcionalidades requeridas pelo Modelo Conceitual FRBR e seus associados, como este modelo tem por objetivo ser uma forma de representar os dados bibliográficos dos recursos informacionais na Web, a partir dos princípios *linked data*, ou dados ligados, que determinam as boas práticas para construir ligações semânticas entre dados de diferentes fontes, no contexto da Web semântica, assim, como forma de atender essas boas práticas, utilizou-se de identificadores únicos, Uniform Resource Identifier – URI para criar interoperabilidade entre ele e algumas categorias do LOM-IEEE (2020) de forma a integrar o processo e a prática de catalogação à Web semântica e enriquecer os dados para que possam atender aos princípios FAIR.

Com relação ao processo de FAIRificação, descrito por Sales (2022), pode-se inferir que os princípios foram alcançados, pois grande parte dos requisitos estabelecidos por Wilkinson et al (2016) foram implementados:

Findable: os dados receberam identificadores globais, persistentes e identificáveis por meio dos URIs no formato de endereços Web; os dados foram ricamente descritos por meio predicados apropriados do próprio padrão LOM-IEEE; os dados foram registrados e indexados em recursos que oferecem capacidades de busca, feito por meio do formato de dados ligados RDF e em consonância com as boas práticas de Berners-Lee (2006); os metadados especificam o identificador dos dados.

Accessible: os dados podem ser recuperados pelos seus identificadores usando protocolo de comunicação padronizado, que no caso foram os URIs; o protocolo usado, isto é o padrão LOM-IEEE e os padrões RDF, são abertos, gratuitos e universalmente implementáveis. Inclusive o padrão LOM-IEEE contém várias categorias de metadados, inclusive de ordem técnica, que mostram quais ferramentas tecnológicas são necessárias para acessar o objeto e de direitos autorais para cientizar sob quais licenças aquele objeto foi disponibilizado.

Interoperable: os dados usaram uma linguagem formal, acessível, compartilhada e amplamente aplicável para a representação do conhecimento, suportada pelo padrão LOM-IEEE e as representações de dados ligados RDF; os dados usaram vocabulários que seguem os princípios FAIR, tal como o Dublin Core e o Bibframe; os metadados incluíram referências qualificadas para outros metadados, conforme estabelecido nas equivalências de propriedades e classes implementadas em OWL. Dessa forma, se conseguiu, de acordo com Zeng (2019), estabelecer equivalência aplicada aos elementos da ontologia operacional para as camadas semântica e estrutural conforme resultados apresentados na seção 6.5.

Reusable: os metadados possuem atributos relevantes, conforme estabelecido pelo padrão LOM-IEEE; os dados foram liberados com licenças de uso claras e acessíveis, estabelecidas também pelo padrão LOM-IEEE; os dados foram associados à sua proveniência e estão alinhados com padrões relevantes ao seu domínio, conforme estabelecem alguns atributos do padrão LOM-IEEE.

Na chamada escala de FAIRificação, de Sales (2022) que mede a qualidade dos dados, de acordo com a aplicação desses princípios onde quanto maior for o grau de FAIRificação dos dados, maior é a possibilidade desse dado ser encontrado, acessado, interoperável e reusável. Especula-se, diante o que já foi exposto, que um grau elevado de FAIRificação tenha sido alcançado para a amostra de OAs desta pesquisa.

É importante ressaltar que o processo de representação da amostra de OAs da presente pesquisa, com base nas categorias do LOM-IEEE (2020) trouxe à tona os aspectos discutidos na seção 3.4.2 sobre os tipos e características dos OAs citados por Braga e Meneses (2014), assim como do ciclo de vida de um OA proposto por Collis e Strijker (2004). Foi possível verificar as relações entre os tipos e o processo de interatividade envolvendo o usuário, que o nível de dificuldade também pode estar vinculado ao tipo de objeto, entre várias outras análises possíveis. O que mostra a especificidade do padrão em relação às informações de ordem educacional.

É difícil comparar os resultados aqui obtidos com os trabalhos apresentados no Estado da Arte por se tratarem de propostas bem distintas, mas alguns apontamentos serão feitos para efeito de discussão.

Koutsomitropoulos, Solomou e Kalou (2017) desenvolveram uma estrutura para expansão de consulta baseada em palavras-chave usando vocabulário SKOS - Simple Knowledge Organization System e implementaram um mecanismo de pesquisa federada integrando vários OAs diferentes em um sistema de gerenciamento de aprendizado. Assim como esta pesquisa, os autores também utilizaram vocabulário SKOS e OWL - Ontology Web Language para poder aplicar operações de raciocínio. As consultas expandidas foram direcionadas para vários repositórios de forma federada, pois a ideia, assim como a da presente pesquisa, também era recuperar a informação que estava dispersa nesses repositórios.

A presente pesquisa vai ao encontro da necessidade levantada por Santarém (2018), onde o autor ressalta que não podemos esquecer que a Web semântica tem um papel social de extrema importância, sendo por meio dela que agentes computacionais podem realizar tarefas para facilitar a vida diária dos seres humanos.

O trabalho de Souza (2020) é o que mais se aproxima das temáticas que aqui foram trabalhadas, pois a autora trabalhou a problemática da inexistência de um modelo de representação da informação, no âmbito da Ciência da Informação, que considere a granularidade (característica relacionada com o fato de arquivos ou ativos de conteúdo, estarem alinhados) e a modularidade (característica desses ativos serem reutilizados isoladamente, podendo assim se “descolar” do seu arquivo maior) dos OAs, de forma a permitir uma melhor recuperação de conteúdos educacionais. A partir do estudo do LOM IEEE (2020) é possível afirmar que o padrão trata a granularidade por meio da categoria 1.8 Aggregation Level e a modularidade por meio da categoria 7 Relation e suas subcategorias (7.1 Kind - que inclusive é baseada no Dublin Core, 7.2.1.1 Catalog, 7.2.1.2 Entry, 7.2.2 Description) todas elas descrevem a relação do OA com outros OAs, desta forma o LOM IEEE (2020)

pode ser considerado um modelo de representação da informação que contempla essas informações problematizadas pela autora.

Em Behr et al. (2021a) os autores apresentam uma revisão sobre o uso de ontologias em sistemas de repositórios de OAs com o objetivo de busca e sugestão, considerando a sua adoção num projeto chamado seaThings que visa a promoção da alfabetização oceânica e utiliza o padrão OBAA que foi adaptado do LOM IEEE e muito se aproxima da presente pesquisa, com a diferença de que foram desenvolvidos Sistemas para dar suporte ao processo de catalogação dos OAs visando armazenar o conjunto de triplas no formato RDF para fornecer interoperabilidade e compatibilidade com dados vinculados.

É possível observar que o ponto de convergência entre a presente pesquisa e os trabalhos avaliados é a preocupação em facilitar a recuperação da informação para quem dela necessita, pois ela só poderá ser reutilizada se for encontrada e para isso, um processo organização do conhecimento eficiente que leve em consideração as especificidades do tipo de informação que se deseja recuperar, se faz necessário.

7 CONSIDERAÇÕES FINAIS

Foram discutidas as razões para a seleção e escolha do padrão LOM IEEE para representação de OAs. Mostrou-se como se deu a escolha do Repositório ARCA para análise das peculiaridades de alguns OAs disponíveis, estudando-se os elementos necessários para a sua representação no padrão LOM IEEE. Apresentou-se as ferramentas selecionadas para desenvolvimento das etapas metodológicas e os metadados necessários para a representação dos OAs selecionados no repositório, organizados de acordo com o padrão LOM IEEE. Elaborou-se um modelo de design para a ontologia operacional de domínio representando todas as categorias e subcategorias do LOM IEEE e implementou-se a ontologia conforme o modelo desenvolvido com vistas a auxiliar a organização de OAs no contexto da Web semântica.

Procurou-se por propriedades e indivíduos equivalentes para a interoperabilidade das camadas estrutural e semântica utilizando-se, por exemplo, de elementos do DublinCore, Wikidata, Bibframe entre outros.

Foi realizado o mapeamento para carregar os dados dos OAs para o formato RDF; onde se realizou consultas SPARQL para testar a recuperação de informação aplicada aos OAs da amostra. Observa-se ainda que a recuperação de informação ainda vai carecer de uma interface ou camada de software apropriada, pois a linguagem proposta, SPARQL, é de difícil domínio por um usuário final.

Todos os objetivos aqui propostos foram atingidos conforme mostrado na seção 6.8 das discussões gerais. Dessa forma, é possível inferir que, com base nos resultados apresentados, a utilização de um padrão para a representação das informações inerentes ao OAs facilita o processo de localização, acesso, interoperabilidade e recuperação dos mesmos. Também foi possível observar que a utilização de ferramentas tecnológicas e apropriadas na implementação de ontologias podem auxiliar muito neste processo. Porém, é importante que os profissionais (bibliotecários, professores, designers etc.) envolvidos neste processo conheçam as ferramentas envolvidas.

Espera-se que esta pesquisa colabore com o amadurecimento e a formalização de um modelo de organização e representação da informação em repositórios de OAs com vistas a interoperabilidade fazendo com que se tornem acessíveis, auxiliando os vários agentes envolvidos no processo ensino-aprendizagem (professores, alunos, designers instrucionais etc.) trazendo também, maior visibilidade ao trabalho riquíssimo desenvolvido nos ambientes virtuais de aprendizagem. Principalmente visando atender ao princípio da reutilização/reusabilidade, apontada na literatura como a característica mais importante de um OA.

O uso do padrão LOM-IEEE foi feito com vistas a derrubar as barreiras que impedem a recuperação dos mesmos de forma integrada a outras bases. Além disso, é uma forma de colocá-los de forma padronizada e interoperável na Web, sintonizando-os ao mundo informacional altamente conectado, trazendo economia de recursos de mão de obra, tecnologia, tempo, e expandindo os horizontes para tornar a EaD mais efetiva pelo acesso e reuso facilitado de OAs.

REFERÊNCIA

ABED (Org.). **Censo EAD.BR**: relatório analítico da aprendizagem a distância no Brasil 2019. ABED – Associação Brasileira de Educação a Distância. Tradução de Camila Rosa. Curitiba: InterSaberes, 2019. Disponível em: http://abed.org.br/arquivos/CENSO_DIGITAL_EAD_2018_PORTUGUES.pdf. Acesso em: 11 jul. 2022.

ALMEIDA, M. B.; BAX, M. P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Ciência da Informação**, Brasília, v. 32, n. 3, p. 7-20, set./dez. 2003. Disponível em: <https://www.scielo.br/j/ci/a/LR68syZsPSSmwvPHrNXmC8N/?format=pdf&lang=pt>. Acesso em: 28 jul. 2021.

ALMEIDA, M. B. Revisiting ontologies: A necessary clarification. **Journal of the American Society for Information Science and Technology**, 64(8), 1682–1693. ago. 2013. Disponível em: https://www.researchgate.net/publication/256441144_Revisiting_Ontologies_A_Necessary_Clarification. Acesso em: 20 dez.2022.

ALSUBAIE, M.; ALSHAWI, M. Reusable objects: Learning object creation cycle. **Proceedings - International Conference on Developments in eSystems Engineering, DeSE 2009**, pp. 321-325.

ALVARENGA, L. Representação do conhecimento na perspectiva da ciência da informação em tempo e espaço digitais. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, v. 8, n. 15, p. 18-40, 2003. Disponível em: 10.5007/1518-2924.2003v8n15p18. Acesso em: 29 jun. 2022.

ARIADNE. Foundation. Disponível em: <https://www.ariadne-eu.org/>. Acesso em: 20 jun. 2022.

ALVES, R. C. V.; SANTOS, P. L. V. A. C. **Metadados no domínio bibliográfico**. Rio de Janeiro: Intertexto, 2013.

ARAÚJO, Carlos Alberto Ávila. **O que é Ciência da informação?** Belo Horizonte: KMA, 2018.

BALAN, A. M. O. A; SCHULUNZEN, E. T. M.; CHULUNZEN, K.; MELQUES, P. M. Banco internacional de objetos educacionais: uma ferramenta para auxiliar no processo de ensino-aprendizagem por meio do uso das Tecnologias de Informação e Comunicação (TIC). **Etic-encontro de iniciação científica**, v.6, n. 6, 2010. Disponível em: <http://intertemas.unitoledo.br/revista/index.php/ETIC/article/viewArticle/2609>. Acesso em: 20 jan. 2022.

BALSTER, Kevin. BIBFRAMEing for non-BIBFRAMers: an introduction to current and future cataloging practices. **The Serials Librarian**, [S. l.], v. 74, n. 1–4, p. 151–155, 2018. Disponível em: <https://doi.org/10.1080/0361526X.2018.1428474>. Acesso em: 20 jun. 2022.

BARBEDO, F.; CORUJO, L.; SANT'ANA, M. Recomendações para a produção de planos de preservação digital. Lisboa: **Direção-Geral do Livro, dos Arquivos e das Bibliotecas (DGLAB)**, nov. 2011. 111 p. Disponível em: <http://arquivos.dglab.gov.pt/wp-content/uploads/sites/16/2014/02/Recomend_producao_PPD_V2.1.pdf>. Acesso em: 9 nov. 2022.

BARUQUE. Lúcia Blondet. **Design instrucional e objetos de aprendizagem**. Tópico 2: Learning Objects (Objetos de Aprendizagem). Disponível em: https://extensao.cecierj.edu.br/material_didatico/biotecava/dilo/creditos.html. Acesso em: 20 jun. 2022.

BECHHOFER, Sean; MILES, Alistair. **Using OWL and SKOS**. W3C. Web page. 2008. Disponível em: <https://www.w3.org/2006/07/SWD/SKOS/skos-and-owl/master.html>. Acesso em: 1 fev. 2023.

BEHAR, Patrícia Alejandra (Orgs.). **Modelos pedagógicos em educação a distância**. Porto Alegre: Artimed, 2009.

BEHR, André et al. Recommending Metadata Contents for Learning Objects Through Linked Data. c **Springer Nature Switzerland AG** 2021 F. De La Prieta et al. (Eds.): PAAMS Workshops 2021, CCIS 1472, pp. 115–126, 2021b. Disponível em: https://doi.org/10.1007/978-3-030-85710-3_10. Acesso em: 10 jan. 2022.

BEHR, André. et al. Enhancing Learning Object Repositories with Ontologies. c The Author(s), under exclusive license to **Springer Nature Switzerland AG** 2021 A. Rocha et al. (Eds.): WorldCIST 2021, AISC 1368, pp. 463–472, 2021a. Disponível em: https://doi.org/10.1007/978-3-030-72654-6_44. Acesso em: 10 jan. 2022.

BERNERS-LEE, Tim. **Linked Data**. 2006. Disponível em: <https://www.w3.org/DesignIssues/LinkedData.html>. Acesso em: 24 mar. 2022.

BERNERS-LEE, Tim. **Web for real people**. Web page: W3C - World Wide Web Consortium, 2005. Disponível em: <https://www.w3.org/2005/Talks/0511-keynote-tbl/>. Acesso em: 24 mar. 2022.

BERNERS-LEE, Tim; HENDER, James; LASSILA, Ora. The Semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. **Scientific American**, New York, v. 284, n. 5, - Maio 2001. Disponível em: https://www-sop.inria.fr/acacia/cours/essi2006/Scientific%20American_

%20Feature%20Article_%20The%20Semantic%20Web_%20May%202001.pdf
Acesso em: 17 mar. 2022.

BIZER, C., HEATH, T., BERNERS-LEE, T.: **Linked data**: the story so far. In: Semantic Services, Interoperability and Web Applications: Emerging Concepts, pp. 205–227. IGI global. 2011.

BLOOM, B. S.; ENGELHART, M. D.; FURST, E. J.; HILL, W. H.; KRATHWOHL, D. R. **Taxonomia de objetivos educacionais** : domínio cognitivo. Globo: Porto Alegre – RS. 1977.

BOAR, B. **Tecnologia da informação**: a arte do planejamento estratégico. 2.ed. São Paulo: Berkeley, 2002.

BRAGA, Katia Soares. Aspectos relevantes para a seleção de metodologia adequada à pesquisa social em Ciência da Informação. In: MUELLER, SUZANA P. M. **Métodos para a pesquisa em Ciência da Informação**. Brasília: Tesaurus, 2007. p. 17–38.

BRAGA. Juliana Cristina; MENESES. Lilian. Introdução aos objetos de aprendizagem. In: BRAGA. Juliana Cristina (Org). **Objetos de aprendizagem**: introdução e fundamentos. v.1. Santo André : Editora da UFABC, 2014. p. 18-40. Disponível em:
<https://pesquisa.ufabc.edu.br/intera/wp-content/uploads/2015/12/objetos-de-aprendizagem-v1.pdf>. Acesso em: 20 jun. 2022.

BRASIL. **Decreto n. 5.800, de 8 de junho de 2006**. Dispõe sobre o Sistema Universidade Aberta do Brasil. Disponível em:
http://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2006/Decreto/D5800.htm. Acesso em: 31 mar. 2022.

CAMPOS, M. L. A.; GOMES, H. E. Ontology as knowledge organization system: role of definitions and relations in a domain conceptual modeling. **Knowledge Organization**, v. 44, p. 178-186, 2017. Disponível em:
<https://www.sciencedirect.com/science/article/abs/pii/S1042814383710083>. Acesso em: 30 jan. 2022.

CAMPOS, M. L. A.; GOMES, H. E. Taxonomia e classificação: o princípio de categorização. **DataGramZero**, v. 9, n. 4, 2008. Disponível em:
<http://hdl.handle.net/20.500.11959/brapci/6615>. Acesso em: 26 jul. 2021.

CAMPOS, M. L. de A.; GOMES, H. E. Taxonomia e classificação: o princípio de categorização. **DataGram Zero**, v. 9, n. 4, p. 2008. Disponível em:
http://www.datagramzero.org.br/ago08/F_I_art.htm. Acesso em: 22 jun. 2022.

CALAN, Eliana; MEDEIROS, Marisa Bräscher Basílio. Sistemas de Organização do Conhecimento na visão da Ciência da Informação. **RICI**: R.Ibero-amer. Ci. Inf., ISSN 1983-5213, Brasília, v. 4, n.2, p. 53-73, ago./dez. 2011.

CECHINEL, Cristian; DA SILVA CAMARGO, Sandro. **Localization of Learning Objects**. Analysis of State-of-the-Art Solutions for Personalised Learning Support, p. 12.

CHAN, L. M.; ZENG, M. L. Metadata interoperability and standardization: a study of methodology part i: achieving interoperability at the schema level. **D-Lib Magazine**, v. 12, n. 6, June 2006. Disponível em: <<http://www.dlib.org/dlib/june06/chan/06chan.html> >. Acesso em: 9 jul. 2022.

CHIZZOTTI, A. **Pesquisa em ciências humanas e sociais**. São Paulo: Cortez, 1991.

COLLIS, B. and STRIJKER, A. **Technology and human issues in reusing learning objects**. Journal of Interactive Media in Education, 2004(1), p.Art. 4. Disponível em: <https://jime.open.ac.uk/articles/10.5334/2004-4-collis/>. Acesso em: 30 jun. 2022.

CMAPTOOLS. **Construct, Navigate, Share and Criticize**. Disponível em: <https://cmap.ihmc.us/cmaptools/>. Acesso em: 20 jul. 2022.

CONEGLIAN, C. S. **Modelo computacional de recuperação da informação para repositórios digitais utilizando ontologias**. Dissertação (Mestrado em Ciência da Informação). Faculdade de Filosofia e Ciências. Universidade Estadual Paulista. Marília. 2017. Disponível em: https://repositorio.unesp.br/bitstream/handle/11449/148996/coneglian_cs_me_mar.pdf?sequence=3&isAllowed=y. Acesso em: 13 mar. 2022.

CONEGLIAN, C. S.; SANTAREM SEGUNDO, J. E. Materialização da Web semântica: um modelo de construção dinâmica de consultas baseados em mapeamento de ontologias. **Perspect. ciênc. inf.**, Belo Horizonte, v. 23, n. 2, p. 33-49, June 2018. Disponível em: <https://www.scielo.br/j/pci/a/W45CkKWQw3rtXPwwXDkcSnf/?format=pdf&lang=pt>. Acesso em: 04 jul. 2022.

CONGRESSO DE INICIAÇÃO CIENTÍFICA DA UFPE, 21., 2013, Recife. **Anais eletrônicos...** Recife: UFPE, 2013. Disponível em: <https://www3.ufpe.br/conic/images/stories/anais/2013/ANAIS.html>. Acesso em: 14 jun. 2021.

CRISTÓVÃO, Henrique Monteiro. **Um modelo híbrido de recuperação de informação e conhecimento baseado na síntese de mapas conceituais obtidos por operações de transformação de redes complexas orientadas por busca de relacionamentos entre termos de consulta em bases de dados ligados**. Tese de Doutorado. Programa de Pós-graduação em Ciência da Informação. Faculdade de Ciência da Informação. Universidade de Brasília. Brasília - DF. 315f. 2016. Disponível em: <https://repositorio.unb.br/handle/10482/22284>. Acesso em: 20 mar. 2022.

CURRÁS, E. **Ontologias, taxonomia e tesouros em teoria de sistemas e sistemática**. Brasília: Thesaurus, 2010.

DALZIEL, James et al. Reflections on the colis (collaborative online learning and information systems) demonstrator project and the "learning object lifecycle". In: **ASCILITE**. 2002. p. 159-166. Disponível em: <<https://www.ascilite.org/conferences/auckland02/proceedings/papers/207.pdf>>. Acesso em: 20 jun. 2022.

DIAS, E. W.; NAVES, M. M. **Análise de assunto: teoria e prática**. 2. ed. rev. Brasília: Briquet de Lemos / Livros, 2013.

DOLIN, R. H.; ALSCHULER, L. Approaching semantic interoperability in Health Level Seven. **Journal of the American Medical Informatics Association**, [s.l.], v. 18, n. 1, p. 99-103, jan. 2011. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/21106995/>. Acesso em: 20 maio 2022.

DUVAL, Erik., VERBER, Katrien. On the role of technical standards for learning technologies. **IEEE transactions on learning technologies**, vol. 1, no. 4, october-december 2008. Disponível em: <https://ieeexplore.ieee.org/document/4760130>>. Acesso em: 20 jun. 2022.

ESPÍNDOLA, Priscilla Lüdtkke. **A influência do Bibframe para visibilidade dos dados bibliográficos**. Dissertação. Programa de Pós-Graduação em Gestão da Informação, Universidade do Estado de Santa Catarina. Florianópolis, SC. 2018. 242f. Disponível em: https://www.udesc.br/arquivos/faed/id_cpmenu/1439/a_influencia_do_bibframe_para_1568900489299_1439.pdf. Acesso em: 22 jun. 2022.

EUROPEAN COMMISSION EXPERT GROUP ON FAIR DATA. Turning FAIR data into reality: final report and action plan from the European Commission Expert Group on FAIR Data. Brussels. 76 p. Disponível em: https://ec.europa.eu/info/sites/default/files/turning_fair_into_reality_1.pdf. Acesso em: 20 jun. 2022

FALBO, Ricardo de Almeida. SABiO: Systematic approach for building ontologies. Em: 2014, Rio de Janeiro, RJ. **Anais [...]**. Em: 1st Joint Workshop ONTO.COM/ODISE on Ontologies in Conceptual Modeling and Information Systems Engineering co-located with 8th International Conference on Formal Ontology in Information Systems. Rio de Janeiro, RJ: CEUR Workshop Proceedings, 2014. p. 14. Disponível em: http://ceur-ws.org/Vol-1301/ontocomodise2014_2.pdf. Acesso em: 20 jan. 2023.

FERLIN, J. **Repositório de objetos de aprendizagem para a área de Informática**. 2009. 118 f. Trabalho de Conclusão de Curso (Graduação em Ciência da Computação) - Universidade Estadual de Santa Catarina (UDESC), 2009.

FORMENTON, D.; GRACIOSO, L. de S. Padrões de metadados no arquivamento da Web: recursos tecnológicos para a garantia da preservação digital de websites arquivados. **Rdbci**: Revista digital de biblioteconomia e ciência da informação, Campinas, sp, v. 20, n. 00, p. e022001, 2022. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/8666263>. Acesso em: 7 jul. 2022.

FOULONNEAU, M.; RILEY, J. **Metadata for digital resources**: implementation, systems design and interoperability. Oxford: Chandos, 2008.

FUJITA, M. S. L. A identificação de conceitos no processo de análise de assunto para indexação. **Rdbci**: Revista digital de biblioteconomia e ciência da informação, Campinas, v. 1, n. 1, p. 60-90, jul./dez. 2003. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/2089> . Acesso em: 14 jul. 2021.

FUJITA, Mariângela Spotti Lopes; TOLARE, Jessica Beatriz. Vocabulários controlados na representação e recuperação da informação em repositórios brasileiros. **Informação & Informação**, [S.l.], v. 24, n. 2, p. 93-125, nov. 2019. ISSN 1981-8920. Disponível em: <<https://www.uel.br/revistas/uel/index.php/informacao/article/view/37985>>. Acesso em: 24 mar. 2022.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 5. ed. São Paulo: Atlas, 2010.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. São Paulo: Atlas, 2002.

GRÁCIO, J. C. A. Preservação digital na gestão da informação: um modelo processual para as instituições de ensino superior. São Paulo, SP: Cultura Acadêmica, 2012. 214 p. Disponível em: <http://www.culturaacademica.com.br/_img/arquivos/Preservacao_digital_na_gestao_da_informacao-WEB_v2.pdf>. Acesso em: 9 nov. 2022.

GRAPHDB. **About GraphDB**. Disponível em: <https://graphdb.ontotext.com/documentation/10.0/about-graphdb.html>. Acesso em: 20 jul. 2022.

GRUBER, Thomas R. **What is an ontology?** 1996. Disponível em: <http://www-ksl.stanford.edu/kst/what-is-an-ontology-1997.html>. Acesso em: 30 jul. 2021.

GRUBER, Thomas R. A translation approach to portable ontology specifications. **Knowledge Acquisition**, [S. l.], v. 5, n. 2, p. 199–220, 1993. DOI: 10.1006/knac.1993.1008. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1042814383710083>. Acesso em: 16 set. 2021.

GUARINO, N. Formal ontology and information systems. In: Formal ontology in information systems - FOIS, 1, 1998, Trento. **Proceedings...** Trento: IOS Press, 1998. Disponível em:
https://www.researchgate.net/publication/272169039_Formal_Ontologies_and_Information_Systems. Acesso em: 19 ago. 2022.

GUERRA, Bruno. **O que é preparação de dados e qual sua importância?**. Inteligência de Negócios. 2020. Disponível em: <https://blog.in1.com.br/o-que-e-preparacao-de-dados-e-qual-sua-importancia>. Acesso em: 05 jul. 2022.

GUIZZARDI, Giancarlo ; Benevides, Alessander Bottes ; Fonseca, Claudemir M. ; Almeida, João Paulo A. ; Sales, Tiago Prince & Porello, Daniele (2022). UFO: Unified Foundational Ontology. **Applied ontology** 1 (17):167-21. Disponível em:
<https://philarchive.org/rec/PORUUF>. Acesso em: 01 mar. 2023.

HARPER, C. A., TILLET, B. B. Library of Congress controlled vocabularies and their application to the Semantic Web. **Cataloging & Classification Quarterly**, v.43, n.34, p.47-68, 2009. Disponível em:
 <https://www.tandfonline.com/doi/abs/10.1300/J104v43n03_03> Acesso em: 22 jun. 2022.

HEERY, Rachel; ANDERSON, Sheila. Digital repositories review. **Joint Information**

HENNING, Patricia Corrêa. et al. Desmistificando os princípios FAIR: conceitos, métricas, tecnologias e aplicações inseridas no ecossistema dos dados FAIR. **Pesquisa brasileira em ciência da informação e biblioteconomia**. V. 14 Issue 3. 2019. Disponível em : <https://periodicos.ufpb.br/index.php/abcib/article/view/46969>. Acesso em: 04 nov. 2021.

HJORLAND, B. Semantics and knowledge organization, **Annual Review of Information Science and Technology**, 41: 367-405. 2007. Disponível em:
<https://asistdl-onlinelibrary-wiley.ez120.periodicos.capes.gov.br/doi/epdf/10.1002/aris.2007.1440410115>

HODGE, Gail. **Systems of knowledge organization for digital libraries**: Beyond Traditional Authority Files. Washington, DC: Council on Library and Information Resources. Available at 2000. Disponível em:
<http://www.clir.org/pubs/reports/pub91/contents.html>. Acesso em: 20 jun. 2022.

IEEE - INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS.
IEEE1484.12.1-2020 - IEEE Standard for Learning Object Metadata - LOM.2020. Disponível em <https://ieeexplore.ieee.org/servlet/opac?punumber=9262116>. Acesso em: 20 jul. 2022.

IMS. Intructional Management Systems. Disponível em: <https://www.1edtech.org/> Acesso em: 20 jun. 2022.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. **IFLA Library Reference Model**: a conceptual model for bibliographic information. Netherlands: IFLA, 2017.

JOINT STEERING COMMITTEE FOR DEVELOPMENT OF RDA. **RDA — Resource Description and Access**: objectives and principles. 2009. Disponível em: <http://www.rda-rsc.org/archivedsite/docs/5rda-objectivesrev3.pdf>. Acesso em: 22 jun. 2022.

Jointly Designing a Data FAIRPORT, 2014, Workshop: 13 - 16 January 2014, Leiden, the Netherlands. Disponível em: <https://www.lorentzcenter.nl/jointly-designing-a-data-fairport.html>. Acesso em: 17 Mar. 2022.

JORENTE, M. J.; PADUA, M. C.; SANTAREM SEGUNDO, J. E. Criação de padrões na Web semântica: perspectivas e desafios. **Em Questão**, v. 23, n. 3, p. 157-178, 2017. Disponível em <http://seer.ufrgs.br/index.php/EmQuestao/article/view/70466>. Acesso em: 30 jan. 2022.

KENDIG, C. E O. What is proof of concept research and how does it generate epistemic and ethical categories for future scientific practice? **Sci Eng Ethics** 22, 735–753 (2016). Disponível em: <https://link.springer.com/article/10.1007/s11948-015-9654-0>. Acesso em: 18 jun. 2022.

KOBASHI, N. Y. **Vocabulário controlado**: estrutura e utilização. Brasília, Escola Nacional de Administração Pública, 2008. Disponível em: <https://repositorio.enap.gov.br/bitstream/1/1289/41/Vocabul%C3%A1rio%20controlado%20-%20estrutura%20e%20utiliza%C3%A7%C3%A3o.pdf> . Acesso em: 24 mar. 2022.

KOIVUNEN, Marja-Riitta ; MILLER Eric. **W3C Semantic Web Activity**. In: Semantic Web Kick-Off in Finland: Vision, Technologies, Research, and Applications, Finlândia, 2001.

LANCASTER, F. W. **Indexação e resumos**: teoria e prática. Brasília, DF: Briquet de Lemos, 2004.

LE COADIC, Yves-François. **A ciência da informação**. 2. rev. e ampl. Brasília, DF: Briquet de Lemos, 2004.

LIBRARY OF CONGRESS. Working group on the future of bibliographic control. On the record: **report of the library of congress working group on the future of bibliographic control**. Washington, DC: LC, 2008. Disponível em: <https://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf> . Acesso em: 20 jun. 2022.

LIMA, J. L.O.; ALVARES, L. Organização e representação da informação e do conhecimento. In: ALVARES, Lillian (Org.). **Organização e representação da informação e do conhecimento**: conceitos, subsídios, interdisciplinaridades e aplicações. São Paulo: B4, 2012. cap. 1. LONGMIRE, W. A primer on learning objects. American Society for Training & Development: Virginia, 2001

LIMA, Júnio César de; CARVALHO Cedric L. de. **Uma visão da Web semântica**. Relatório Técnico March - 2004 - Março. Instituto de Informática Universidade Federal de Goiás. Disponível em: https://ww2.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF_001-04.pdf. Acesso em: 20 jun. 2022.

LOMENA, M. **Benjamin Bloom**. Disponível em: <https://everything2.com/title/Benjamin+Bloom>. Acesso em: 28 Jul. 2021.

LÓSCIO, B. F.; BURLE, C.; CALEGARI, N. **Data on the Web Best Practices**: challenges and benefits. W3C Recommendation, 2017. Disponível em: <https://bit.ly/2FG1EoK>. Acesso em: 10 jul. 2022.

MACHADO, Raildo de Souza; I. ZAFALON, Zaira Regina. **Catálogo**: dos princípios e teorias ao RDA e IFLA LRM. João Pessoa: Editora UFPB, 2020.

MARTINEZ, A. et al. Las categorías o facetas fundamentales: una metodología para el diseño de taxonomías corporativas de sitios Web argentinos. **Ciência da Informação**, Brasília, v. 33, n. 2, p. 106-111, maio/ago. 2004. Disponível em: https://www.researchgate.net/publication/242446103_Las_categorias_o_facetas_fundamentales_una_metodologia_para_el_diseno_de_taxonomias_corporativas_de_sitios_Web_argentinos. Acesso em: 25 jul. 2021.

MAZZOCCHI, Fulvio. Knowledge organization system (KOS). **Knowledge organization**, [S. l.], v. 45, n. 1, p. 54-78, 2018. Disponível em: <https://www.isko.org/cyclo/kos.htm>. Acesso em: 20 ago. 2021.

MCGREAL, Rory. A typology of learning object repositories. In: Handbook on information technologies for education and training. Springer Berlin Heidelberg 2008 p. 5-28.

MITCHELL, Erik. Three case studies in linked, open data. Library Technology Reports, [S. l.], v. 49, n. 5, p. 26-43. jul. 2013. **Library linked data**: research and adoption. Disponível em: <https://journals.ala.org/index.php/ltr/article/view/4693/5586>. Acesso em: 11 jun. 2022.

MOORE. Michael G. , KEARSLEY. Greg. **Educação a distância** : uma visão integrada. São Paulo : Thomson Learning, 2007.

MOREIRA, F. M.; ZAFALON, Z. R.; SANTOS, P. L. V. A. C.; SANTANA, R. C. G. Metadados para descrição de datasets e recursos informacionais do portal brasileiro de dados abertos. **Perspectivas em Ciência da Informação**, v. 22, n. 3, p. 158-185, 2017. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/31970>. Acesso em: 19 jun. 2022.

NATIONAL INFORMATION STANDARDS ORGANIZATION (NISO). Understanding metadata. Bethesda: **NISO Press**, 2004. 16 p. Disponível em: <<http://www.niso.org/publications/press/UnderstandingMetadata.pdf> >. Acesso em: 9 jul. 2022.

NHACUONGUE, Januário Albino; ROZSA, Vitor; DUTRA, Moisés Lima. Linked Data e Ciência da Informação: diretrizes para a publicação de datasets institucionais abertos. **Biblios**, nº 73 (outubro): 20–34. 2018. Disponível em: <http://www.scielo.org.pe/scielo.php?script=sci_abstract&pid=S1562-47302018000400002&lng=es&nrm=iso&tlng=es>. Acesso em: 24 mar. 2022.

NOY, N.F. ; MCGUINNESS, D.L. **Ontology Development 101**: a guide to creating your first ontology, Stanford Knowledge System Laboratory Technical Report, [s/l], p1-25, june 2002.

NUNES, F. R. E.; MACULAN, B. C. M. dos Santos; ALMEIDA, M. B.. Os fundamentos da Web semântica como ferramentas de auxílio para as demandas da Sociedade da Informação. **Em Questão**. Universidade Federal do Rio Grande do Sul. vol. 26, núm. 3, 2020. Disponível em: <https://www.redalyc.org/articulo.oa?id=465664724010> Acesso em: 19 jun. 2022.

OCHOA, X.; DUVAL, E. **Measuring learning object reuse**. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 5192 LNCS, 2008, pp. 322-325.

OLIVER, Chris. **Introdução à RDA**: um guia básico. Brasília, DF: Briquet de Lemos. 2011.

OPENREFINE. **Introduction to OpenRefine**. Disponível em: <https://openrefine.org/>. Acesso em: 20 jul. 2022.

O'REILLY, Tim. **What is Web 2.0**. 2005. Disponível em: <<http://wellman.uni-trier.de/images/9/96/Web20.pdf>> Acesso em 14 jun. 2022.

ORTEGA, C. D. Fundamentos da organização da informação frente à produção de documentos. **Transinformação**, v. 20, n. 1, p. 7-15, jan./abr., 2008. Disponível em: <https://brapci.inf.br/index.php/res/v/115965> . Acesso em: 16 jun. 2022.

PERIÓDICOS DE MINAS. Entenda. **O que é um gerenciador de referências?**. Disponível em: <https://www.periodicosdeminas.ufmg.br/o-que-e-um-gerenciador-de-referencias/>. Acesso em: 10 jul. 2022.

PÖTTKER, Luciana Maria Vieira; FERNEDA, Edberto; MOREIRO-GONZÁLEZ, José Antonio. Mapeamento relacional entre padrões de metadados educacionais. **Perspectivas em Ciência da Informação**, v.23, n.3, p.25-38, jul./set. 2018. Disponível em: <http://dx.doi.org/10.1590/1981-5344/2843>. Acesso em: 20 jun. 2022.

QUEIROZ, Claudete Fernandes de. **Rede Sudeste de Repositórios Institucionais: retrospectiva 2021**. In: Reunião rede sudeste de repositórios institucionais, 25., 2021. Rio de Janeiro: Fiocruz/Icict, 2021. 77 p. Disponível em: <https://www.arca.fiocruz.br/handle/icict/50477>. Acesso em: 20 jun. 2022.

REIS, Marcia Cristina dos; FERNEDA, Edberto. Panorama sobre a utilização dos padrões de metadados IEEE LOM e OBAA em repositórios educacionais brasileiros. **Anais...** Seminário em Ciência da Informação - Secin. VI. Londrina. Paraná. 3 a 5 de Agosto. 2016. Disponível em: <http://www.uel.br/eventos/cinf/index.php/secin2016/secin2016/paper/viewFile/312/201>. Acesso em 30 jun. 2022.

SALES, Luana Farias. **Curso princípios FAIR aplicados à repositórios**. Rio de Janeiro: Sudeste/RIAA; IBICT/PPGCI, 2021. Disponível em: https://www.arca.fiocruz.br/bitstream/icict/50571/2/RIAA_Curso_Princ%C3%ADpios_FAIR_20_10_2021.pdf. Acesso em: 30 jun. 2022.

SALES, Luana Farias. **Princípios FAIR**. Em entrevista à comunicação social do Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict). 2022. Disponível em: <https://www.gov.br/ibict/pt-br/central-de-conteudos/noticias/2022/marco-2022/principios-go-fair> Acesso em: 30 jun. 2022.

SANTAREM SEGUNDO, J. E. Web semântica, dados ligados e dados abertos: uma visão dos desafios do Brasil frente às iniciativas internacionais. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 8, n. 2, p. 219-239, 2015. Disponível em: <http://www.ufpb.br/evento/index.php/enancib2015/enancib2015/paper/viewFile/3149/1193>. Acesso em: 30 jan. 2022.

SANTAREM SEGUNDO, J. E. Web semântica, dados ligados e dados abertos: uma visão dos desafios do Brasil frente às iniciativas. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 8, n. 2, p.219-239, 2015. Disponível em: <http://inseer.ibict.br/ancib/index.php/tpbci/article/view/207>. Acesso em: 21 fev. 2022.

SANTAREM SEGUNDO, J. E.; CONEGLIAN, C. S. Web semântica e ontologias: um estudo sobre construção de axiomas e uso de inferências. **Informação & Informação**, v. 21, n. 2, p. 217-244, 2016. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/viewFile/26417/20131>>. Acesso em: 23 mar. 2022.

SANTAREM SEGUNDO, J. E.; CONEGLIAN, C. S. Web semântica e ontologias: um estudo sobre construção de axiomas e uso de inferências. **Informação & Informação**, v. 21, n. 2, p. 217-244, 2016. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/26417>. Acesso em: 30 jan. 2022.

SANTAREM SEGUNDO, J.E.; CONEGLIAN, C.; LUCAS, E. Conceitos e tecnologias da Web semântica no contexto da colaboração acadêmico-científica: um estudo da plataforma Vivo. **Transinformação**, v. 29, n. 3, p. 297-309, 2017. Disponível em: http://www.scielo.br/scielo.php?script=sci_abstract&pid=S0103-37862017000300297&lng=en&nrm=iso&tlng=-pt. Acesso em: 30 jan. 2022.

SANTAREM SEGUNDO, José Eduardo. Web semântica: Introdução a recuperação de dados usando SPARQL. In: ENCONTRO NACIONAL DE PESQUISAS EM CIÊNCIA DA INFORMAÇÃO (ENANCIB), 14., 2014, Belo Horizonte. **Além das nuvens: expandindo as fronteiras da ciência da informação**. Belo Horizonte: UFMG, 2014. p. 3242 - 3261. Disponível em: https://www.researchgate.net/publication/304216019_WEB_SEMANTICA_INTRODUCAO_A_RECUPERACAO_DE_DADOS_USANDO_SPARQL. Acesso em: 21 maio 2022.

SBROGIO, Renata VALENTE, Vania. **Repositórios e Referatórios de Objetos de Aprendizagem: Recursos Livres e Gratuitos**. ResearchGate. 1. 1. 2020. Disponível em : https://www.researchgate.net/publication/346952531_REPOSITARIOS_E_REFERATORIOS_DE_OBJETOS_DE_APRENDIZAGEM_RECURSOS_LIVRES_E_GRATUITOS/related. Acesso em: 22 jun. 2022.

SCHIESSL, M.; BRASCHER, M. Do texto às ontologias: uma perspectiva para a ciência da informação. **Ciência da Informação**, v. 40, n. 2, set. 2012. Disponível em: <http://revista.ibict.br/ciinf/article/view/1318>. Acesso em: 30 jan. 2022.

SHADBOLT, Nigel; BERNERS-LEE, Tim; HALL, Wendy. The Semantic Web. **IEEE Intelligent Systems**, vol. 21, no. 03, pp. 96-101, 2006. Disponível em: <https://www.computer.org/csdl/magazine/ex/2006/03/x3096/13rUwwaKpw>. Acesso em: 25 jun. 2022.

SHENA. **Definição - o que significa a linguagem de consulta?** Definirtec. 2021. Disponível em: <https://definirtec.com/linguagem-de-consulta/>. Acesso em: 10 jun. 2022.

SHINTAKU, Milton. et al. Banco Internacional de Objetos Educacionais – BIOE. **Revista Intercâmbio das humanidades**. v.1. n.421. 2014. Disponível em: <https://2014.revistainterambio.net.br/24h/pessoa/temp/anexo/1/421/749.pdf>. Acesso em: 20 jun. 2022.

SILVA, Edna Lúcia da; CAFÉ, Lígia; CATAPAN, Araci Hack. Os objetos educacionais, os metadados e os repositórios na sociedade da informação. **Ciência da Informação** [online]. 2010, v. 39, n. 3. p. 93-104. Disponível em: <https://doi.org/10.1590/S0100-19652010000300008>. Acesso em: 25 Jun. 2022.

SILVA, Robson Santos da. **Objetos de aprendizagem para a educação a distância**. São Paulo: Novatec, 2011.

SILVEIRA, ISMAR FRANGO. Rumo ao reúso: recursos educacionais abertos. In: BRAGA, Juliana Cristina (Org). **Objetos de aprendizagem: introdução e fundamentos**. v.1. Santo André : Editora da UFABC, 2014. p. 127-149. Disponível em: <https://pesquisa.ufabc.edu.br/intera/wp-content/uploads/2015/12/objetos-de-aprendizagem-v1.pdf>. Acesso em: 20 jun. 2022.

SMITH, B. Ontology. In: FLORIDI, L. **Blackwell guide to the philosophy of computing and information**. Oxford: Blackwell, 2003. p.155-166. Disponível em: <http://ontology.buffalo.edu/smith/articles/>. Acesso em: 15 jul. 2021.

SOUZA, R. M. F. de. **Representação da informação de objetos de aprendizagem por meio de metadados**: considerações sobre granularidade e modularidade. Tese de Doutorado. Departamento de ciências sociais aplicadas. Programa de Pós-graduação em ciência da informação. Universidade da Paraíba. João Pessoa-PB. 212 f. 2020

STELLATO, Armando. Dictionary, thesaurus or ontology? Disentangling our choices in the semantic Web jungle. **Journal of Integrative Agriculture**, [S. l.], v. 11, n. 5, p. 710–719, 2012. DOI: 10.1016/S2095-3119(12)60060-4. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S2095311912600604>. Acesso em: 5 mar. 2023.

STUDER, R., Benjamins, V.R. and Fensel, D. Knowledge Engineering: Principles and Methods. **Data and Knowledge Engineering**, 25, 61-197. 1998. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0169023X97000566>. Acesso em: 20 jun. 2022.

Systems Committee: UKOLN. 2005. Disponível em: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.211.2993&rep=rep1&type=pdf> . Acesso em: 30 jun. 2022.

TAROUCO, Liane Margarida Rockenbach. Objetos de Aprendizagem e a EAD In: LITTO, Fredric M.; FORMIGA, Marcos. **Educação a distância: o estado da arte**. v. 2 . São Paulo: Pearson Education do Brasil. 2012.

UNESCO. **Forum on the impact of open courseware for higher education in developing countries**. Final Report, 2002. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000128515>. Acesso em: 22 jun. 2022.

USCHOLD, M. Building ontologies; towards a unified methodology. In: Annual conference of the british computer society specialist group on expert systems, 16, 1996, Cambridge. **Anais...** Cambridge, 1996. p. 16-18.

VAZ, M. F. R. Os padrões internacionais para a construção de material educativo on-line. In: LITTO, Fredric M. ; FORMIGA, Marcos. (Orgs). **Educação a distância: o estado da arte**. São Paulo: Pearson Education do Brasil, 2009 p. 386 - 394.

VICARI, Rosa Maria et al. Proposta brasileira de metadados para objetos de aprendizagem baseados em agentes (OBAA). **Revista Novas Tecnologias na Educação**, Porto Alegre, v. 8, n. 2, 2010. Disponível em: <<https://seer.ufrgs.br/index.php/renote/article/view/15257/9015>>. Acesso em: 01 jun.2022.

VICARI, Rosa Maria et al. **Relatório Técnico RT-OBAA-01**: proposta de padrão para metadados de objetos de aprendizagem multiplataforma. Porto Alegre: UFRGS, 2009 a. 78 p. Disponível em:https://www.researchgate.net/publication/236669555_Proposta_de_Padrao_para_Metadados_de_Objeto_de_Aprendizagem_Multiplataforma . Acesso em: 15 out. 2021.

VITAL, Luciane Paula; CAFÉ, Lígia Maria Arruda. Ontologias e taxonomias: diferenças. **Perspectivas em Ciência da Informação**, v.16, n.2, p.115-130, abr./jun. 2011. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/200/927>. Acesso em: 28 jul.2021.

VOCBENCH. **VocBench**. Disponível em: <http://vocbench.uniroma2.it/>. Acesso em: 20 jul. 2022.

W3C. **SPARQL Query Language for RDF**. 2008. Disponível em: <https://www.w3.org/TR/rdf-sparql-query/>. Acesso em: 20 jul.2022.

WAL,Thomas Vander. **Folksonomy**. 2007. Disponível em: <http://vanderwal.net/folksonomy.html>. Acesso em: 02 jul. 2022.

WEISS, Leila Cristina. **Interoperabilidade semântica: uma análise sob a perspectiva da abordagem ontológica de Willard Van Orman Quine**. Tese (doutorado) - Universidade Federal de Santa Catarina, Centro de Ciências da Educação, Programa de Pós-Graduação em Ciência da Informação, Florianópolis, 2019. Disponível em: <https://repositorio.ufsc.br/bitstream/handle/123456789/194269/PCIN0191-T.pdf?sequence=1&isAllowed=y> . Acesso em: 24 mar. 2022.

WEISS, Leila Cristina. Interoperabilidade semântica: uma análise das perspectivas teóricas dos estudos desenvolvidos na área de Ciência da Informação. **Em Questão** 27 (3): 431–57. 2021. Disponível em: <https://seer.ufrgs.br/EmQuestao/article/view/107807>. Acesso em: 20 mar. 2022.

WILEY, D. A. **Connecting learning objects to instructional design theory: a definition, a metaphor, and a taxonomy.** In: D. A. Wiley (ed.), *The Instructional Use of Learning Objects: Online Version*, 2000. Disponível em: <http://reusability.org/read/chapters/wiley.doc>. Acesso em: 22 nov. 2021.

WILEY, D. A. **The Access Compromise and the 5th R.** 2014. Disponível em: <http://opencontent.org/blog/archives/3221>. Acesso em: 2 dez. 2021.

WILEY, D. A. **The Post-LEGO Learning Object.** 2003. Disponível em: <http://davidwiley.org/docs/post-lego.pdf>. Acesso em: 22 jun. 2022.

WILEY, D. A. **Learning object design and sequencing theory.** Brigham Young University, 2000. Disponível em: <https://opencontent.org/docs/dissertation.pdf>. Acesso em: 01 jan. 2022.

WILEY, D. A. **The instructional use of learning objects.** Agency for Instructional Technology; Association for Educational Communications Technology. Bloomington, Indiana: 2002. Disponível: <https://members.aect.org/publications/InstructionalUseofLearningObjects.pdf>. Acesso em: 01 jan. 2022.

WILKINSON, M. D., et al. **A design framework and exemplar metrics for FAIRness.** November, 2017. Disponível em: <https://www.biorxiv.org/content/biorxiv/early/2017/11/29/225490.full.pdf> Acesso em: 30 Jun. 2022.

WILKINSON, Mark D. et al. **The FAIR guiding principles for scientific data management and stewardship.** *Scientific Data*, [S. l.], v. 3, n. 1, p. 160018, 2016. DOI: 10.1038/sdata.2016.18. Disponível em: <https://www.nature.com/articles/sdata201618>. Acesso em: 15 jun. 2022.

YAMAKAWA, Eduardo Kazumi et al. Comparativo dos softwares de gerenciamento de referências bibliográficas: Mendeley, EndNote e Zotero. **Transinformação** [online]. 2014, v. 26, n. 2, pp. 167-176. Disponível em: <https://doi.org/10.1590/0103-37862014000200006>. Acesso em 24 Jun 2022.

YIN, R. K. **Pesquisa Estudo de Caso: desenho e métodos**(2 ed.). Porto Alegre: Bookman. 1994.

ZAFALON, Z. R. **Formato MARC21 Bibliográfico: estudos e aplicações para livros, folhetos, folhas impressas e manuscritos.** São Carlos: EdUFSCar, 2008.

ZAFALON, Z. R. **Scan for MARC: princípios sintáticos e semânticos de registros bibliográficos aplicados à conversão de dados analógicos para o Formato MARC21 Bibliográfico.** 2012. Tese (Doutorado) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2012.

ZENG, Marcia L. Knowledge Organization Systems (KOS). **Knowledge Organization**, 35, nos. 2-3: 160–182, 2008. Disponível em: <https://www.researchgate.net/publication/297530633_Knowledge_Organization_Systems_KOS>. Acesso em: 20 jun. 2022.

ZENG, Marcia L. Interoperability. **Knowledge Organization** 46, n.2 .2019. 122–146. Disponível em: <<http://www.isko.org/cyclo/interoperability>>. Acesso em: 25 jun. 2022.

ZENG, Marcia L.; QIN, J. **Metadata**. New York: Neal-Schuman Publishers, 2008. 365 p.

ANEXOS

ANEXO I - LISTA DE REPOSITÓRIOS DE OAs (Sbrogio e Valente, 2020)

Nome do ROA	Descrição e endereço eletrônico
Acervo Multimeios	ROA com fotos vídeos e áudios de temas curriculares e transversais diversos. Nele, é possível compartilhar, copiar, distribuir e transmitir as obras sob algumas condições: atribuição de crédito, não comercial e sem derivação. http://multimeios.seed.pr.gov.br/resourcespace-seed/pages/home.php
Biblioteca Digital do Centro de Trabalho Indigenista	ROA com registros de documentos, fotos, vídeos, livros, objetos de cultura material, entre outros tipos de AO, sobre povos indígenas: http://bd.trabalhoindigenista.org.br/
BNDigital: Biblioteca Nacional Digital	ROA com artigos, dossiês, exposições, acervo digital e hemeroteca, faz parte da Fundação Biblioteca Nacional: http://bndigital.bn.gov.br/
BIOE: Banco Internacional de Objetos Educacionais	Objetos educacionais de todos os níveis de ensino, inclusive para Educação Profissional, de acesso livre: http://objetoseducacionais.mec.gov.br/#/inicio
Casa das Ciências	Portal colaborativo com recursos digitais de matemática e ciências para diferentes níveis de ensino https://www.casadasciencias.org/edulog.html
CESTA: Coletânea de Entidades de Suporte ao uso de Tecnologia na Aprendizagem	Projeto elaborado pelo CINTED - Centro Interdisciplinar de Novas Tecnologias na Educação da UFRGS, que visa criar um repositório de objetos educacionais para fins de reuso: http://www.cinted.ufrgs.br/CESTA/cestadescr.html
Co-laborando	Plataforma desenvolvida e disponibilizada pela FCS Comercial Importadora e Exportadora Ltda., apoiada e patrocinada pela Faber-Castell com AO para Educação Infantil, Ensino Fundamental I e II e Ensino Médio: http://www.colaborando.net.br/
Currículo Digital da Cidade de São Paulo	Sequências de atividades curriculares para Ensino Fundamental, Médio e EJA: https://curriculo.sme.prefeitura.sp.gov.br/
Currículo +	Recursos digitais articulados com o Currículo do Estado de São Paulo, com atividades de apoio ao trabalho docente no uso de objetos digitais de aprendizagem: http://curriculomais.educacao.sp.gov.br/
Educopédia²	Organização da rede municipal de Ensino do Estado do Rio de Janeiro, plataforma on-line colaborativa de aulas digitais, onde alunos e professores podem acessar atividades: http://www.educopedia.com.br/
Eukatu	Rede de incentivo para a troca de conhecimentos e práticas de consumo consciente com vídeos, áudios, imagens, textos e gráficos: https://edukatu.org.br/
LABVIRT	Laboratório Didático Virtual da USP com objetos de aprendizagem de química e física como: simulações; links para simulações e <i>link</i> para <i>sites</i> de interesse: http://www.labvirtq.fe.usp.br/indice.asp

LUME Repositório Digital (UFRGS)	Biblioteca digital de recursos educacionais produzidos na UFRGS e/ou por autores da UFRGS: https://lume.ufrgs.br/
Escola Digital	Plataforma gratuita com recursos digitais de aprendizagem para interatividade, com OA, planos de estudo, cursos e ferramentas de autoria (criação): http://www.escoladigital.org.br/
Portal do Professor	Acesso e compartilhamento de planos de aula, conteúdos educativos na forma de mídias, leitura de notícias entre outros. Iniciativa do MEC (Ministério da Educação e Cultura do Brasil): http://portaldoprofessor.mec.gov.br/index.html
Portal Domínio Público	Obras literárias, artísticas e científicas na forma de textos, imagens, sons e vídeos, disponibilizadas em domínio público ou com divulgação previamente autorizada, constituintes do patrimônio cultural brasileiro e/ou universal. http://www.dominiopublico.gov.br/pesquisa/PesquisaObraForm.jsp
Portal MAIS	Repositório com recursos educacionais de Matemática, desenvolvidos em parceria de diversas empresas, para Ens. Fundamental e Médio: http://www.mais.mat.br/wiki/P%C3%A1gina_principal
Portal Unicamp: GGTE	Ambientes virtuais de aprendizagem e conteúdos educacionais de acesso livre: https://ggte.unicamp.br/wp/?page_id=647
Porto OCW: Porto OpenCourseWare	Conteúdos educacionais digitais do Colégio Visconde de Porto Seguro: http://oec.portoseguro.org.br/
PROEDU: Acervo de Recursos Educacionais para Educação Profissional e Tecnológica	ROA para compartilhamento de Recursos Educacionais Abertos (REA) validados para os Cursos de Formação Profissional de nível técnico, Formação Inicial Continuada (FIC) e formações técnico profissionalizantes em geral: http://proedu.rnp.br/
REA Dante	Recursos Educacionais Abertos do Colégio Dante Alighieri com materiais didáticos em formato digital, sob licença Creative Commons: https://sistemas.colegiodante.com.br/auxiliar/rea/
REALPTL: Recursos Educacionais Abertos para Leitura e Produção de Textos nas Licenciaturas	Recursos Educacionais Abertos para leitura e produção de textos na forma de jogos, tutoriais, tabelas, entre outros: http://realptl.letras.ufmg.br/realptl/
RIVED: Rede Internacional Virtual de Educação	Conteúdos digitais interativos para Educação Básica: http://www.dmm.im.ufrj.br/projeto/rived/index.html
SaberCom: Repositório de Objetos Educacionais Digitais – Universidade Federal do Rio Grande (FURG)	Objetos de aprendizagem dos cursos de graduação, especialização, aperfeiçoamento e extensão na modalidade EaD da Universidade Federal do Rio Grande (FURG) para fomentar o uso, reuso e a criação de novos materiais. http://www.sabercom.furg.br/
SOPLAR (nome originado da junção das palavras: Sócrates, Platão e Aristóteles)	Rede social para relacionamentos e estudos coletivos de ensino e aprendizagem com ferramentas como: chat, sala de estudos, área para recados e criação de textos, procurar amigos, entre outros: http://www.soplaar.com/index.php?idioma=1
TATU repositório digital	Repositório digital com objetos iconográficos, cartilhas, livretos, revistas entre outros do Grupos de Estudos em Educação, História e Narrativas (UNIPAMPA): http://sistemas.bage.unipampa.edu.br/tatu/

ANEXO II - TABLE 1 – LOMV1.0 BASE SCHEMA

Table 1—LOMv1.0 Base Schema

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
1	General	This category groups the general information that describes this learning object as a whole.	1	Unspecified	—	—	—
1.1	Identifier	A globally unique label that identifies this learning object.	Smallest permitted maximum: 10 items	Unspecified	—	—	—
1.1.1	Catalog	The name or designator of the identification or cataloging scheme for this entry. A namespace scheme.	1	Unspecified	Repertoire of ISO/IEC 10646-1:2000	CharacterString (smallest permitted maximum: 1000 char)	"ISBN", "ARIADNE", "URI"
1.1.2	Entry	The value of the identifier within the identification or cataloging scheme that designates or identifies this learning object. A namespace specific string.	1	Unspecified	Repertoire of ISO/IEC 10646-1:2000	CharacterString (smallest permitted maximum: 1000 char)	"2-7342-0318", "LEAO875", "https://www.ieee.org/documents/1234"
1.2	Title	Name given to this learning object.	1	Unspecified	—	LangString (smallest permitted maximum: 1000 char)	("en", "The life and works of Leonardo da Vinci")

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
1.3	Language	The primary human language or languages used within this learning object to communicate to the intended user. NOTE 1— An indexation or cataloging tool could provide a useful default. NOTE 2— If the learning object had no lingual content (as in the case of a picture of the Mona Lisa, for example), then the appropriate value for this data element would be "none." NOTE 3— This data element concerns the language of the learning object. Data element 3.4:Meta-Metadata. Language concerns the language of the metadata instance.	smallest permitted maximum: 10 items	Unordered	LanguageID = Langcode ["-Subcode]* with Langcode a language code as defined by the code set ISO 639:1988 and Subcode (which can occur an arbitrary number of times) a country code from the code set ISO 3166-1:1997. NOTE 4— This value space is also defined by RFC1766:1995 and is harmonized with that of the xml:lang attribute. NOTE 5— ISO 639:1988 also includes "ancient" languages, like Greek and Latin. The language code should be given in lower case and the country code (if any) in upper case. However, the values are case insensitive. "none" is also an acceptable value.	CharacterString (smallest permitted maximum: 100 char)	"en", "en-GB", "de", "fr-CA", "it", "grc" (ancient greek, until 1453) "en-US-philadelphia" "eng-GB-cockney" "map-PG-buin" (Austronesian - Papua New Guinea - buin) "gem-US-pennsylvania"

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
1.4	Description	A textual description of the content of this learning object. NOTE 6— This description need not be in language and terms appropriate for the users of the learning object being described. The description should be in language and terms appropriate for those that decide whether or not the learning object being described is appropriate and relevant for the users.	smallest permitted maximum: 10 items	Unordered	—	LangString (smallest permitted maximum: 2000 char)	("en", "In this video clip, the life and works of Leonardo da Vinci are briefly presented. The focus is on his artistic production, most notably the Mona Lisa.")
1.5	Keyword	A keyword or phrase describing the topic of this learning object. This data element should not be used for characteristics that can be described by other data elements.	smallest permitted maximum: 10 items	Unordered	—	LangString (smallest permitted maximum: 1000 char)	("en", "Mona Lisa")
1.6	Coverage	The time, culture, geography or region to which this learning object applies. The extent or scope of the content of the learning object. Coverage will typically include spatial location (a place name or geographic coordinates), temporal period (a period label, date, or date range) or jurisdiction (such as a named administrative entity). Recommended best practice is to select a value from a controlled vocabulary (for example, the Thesaurus of Geographic Names [TGN]) and that, where appropriate, named places or time periods be used in preference to numeric identifiers such as sets of coordinates or date ranges. NOTE 7— This is the definition from the Dublin Core Metadata Element Set. ⁴	smallest permitted maximum: 10 items	Unordered	—	LangString (smallest permitted maximum: 1000 char)	("en", "16th century France") NOTE 8— A learning object could be about farming in 16th century France: in that case, its subject can be described with 1.5:General.Keyword = ("en", "farming") and its 1.6:General.Coverage can be ("en", "16th century France").

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
1.7	Structure	Underlying organizational structure of this learning object.	1	Unspecified	atomic: an object that is indivisible (in this context). collection: a set of objects with no specified relationship between them. networked: a set of objects with relationships that are unspecified. hierarchical: a set of objects whose relationships can be represented by a tree structure. linear: a set of objects that are fully ordered. Example: A set of objects that are connected by "previous" and "next" relationships.	Vocabulary (State)	NOTE 9— A learning object with Structure = "atomic" will typically have 1.8:General.AggregationLevel = 1. A learning object with Structure = "collection", "linear", "hierarchical" or "networked" will typically have 1.8:General.AggregationLevel = 2, 3 or 4.

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
1.8	Aggregation Level	The functional granularity of this learning object.	1	Unspecified	1: the smallest level of aggregation, e.g., raw media data or fragments. 2: a collection of level 1 learning objects, e.g., a lesson. 3: a collection of level 2 learning objects, e.g., a course. 4: the largest level of granularity, e.g., a set of courses that lead to a certificate. NOTE 10— Level 4 objects can contain level 3 objects, or can recursively contain other level 4 objects.	Vocabulary (Enumerated)	If the learning object is a digital picture of the Mona Lisa, 1.7:General.Structure = Atomic and 1.8:General.AggregationLevel = 1. If the learning object is a lesson with the digital picture of the Mona Lisa, 1.7:General.Structure = Collection or Networked (since there are two descriptions of the same type of Structure) and 1.8:General.AggregationLevel = 2. If the learning object is a course on the Mona Lisa, 1.7:General.Structure = Linear if the documents are intended to be viewed linearly and 1.8:General.AggregationLevel = 3. If the learning object is a collection of lessons from different sources, 1.7:General.Structure = Collection and 1.8:General.AggregationLevel = 3. Lastly if the learning object is a set of courses with a full history, description, interpretation, etc. of the Mona Lisa, 1.7:General.Structure = Linear or Hierarchical and 1.8:General.AggregationLevel = 4. NOTE 11— A learning object with AggregationLevel = 1 will typically have 1.7:General.Structure = "atomic". A learning object with AggregationLevel = 2, 3 or 4 will typically have 1.7:General.Structure = "collection", "linear", "hierarchical" or "networked".

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
2	Life Cycle	This category describes the history and current state of this learning object and those entities that have affected this learning object during its evolution.	1	Unspecified	—	—	—
2.1	Version	The edition of this learning object.	1	Unspecified	—	LangString (smallest permitted maximum: 50 char)	("en", "1.2.alpha"), ("nl", "voorlopige versie")
2.2	Status	The completion status or condition of this learning object.	1	Unspecified	draft final revised unavailable NOTE 12— When the status is "unavailable" it means that the learning object itself is not available.	Vocabulary (State)	—
2.3	Contribute	Those entities (i.e., people, organizations) that have contributed to the state of this learning object during its life cycle (e.g., creation, edits, publication). NOTE 13— This data element is different from 3.3:Meta-Metadata. Contribute. NOTE 14— Contributions should be considered in a very broad sense here, as all actions that affect the state of the learning object.	smallest permitted maximum: 30 items	Ordered	—	—	—

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
2.3.1	Role	Kind of contribution. NOTE 15— Minimally, the Author(s) of the learning object should be described.	1	Unspecified	author publisher unknown initiator terminator validator editor graphical designer technical implementer content provider technical validator educational validator script writer instructional designer subject matter expert NOTE 16— “terminator” is the entity that made the learning object unavailable.	Vocabulary (State)	—
2.3.2	Entity	The identification of and information about entities (i.e., people, organizations) contributing to this learning object. The entities shall be ordered as most relevant first.	Smallest permitted maximum: 40 items	Ordered	vCard, as defined by IMC vCard 3.0 (RFC 2425, RFC 2426).	CharacterString (smallest permitted maximum: 1000 char)	“BEGIN:VCARD\nFN:Joe Friday\nnTEL:+1-919-555-7878\nnTITLE:Area Administrator\nAssistant\nEMAIL;TYPE=INTERN\nnET:jfriday@host.com\nnEND:VCARD\n”
2.3.3	Date	The date of the contribution.	1	Unspecified	—	DateTime	“2001-08-23”

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
3	Meta-Metadata	This category describes this metadata record itself (rather than the learning object that this record describes). This category describes how the metadata instance can be identified, who created this metadata instance, how, when, and with what references. NOTE 17— This is not the information that describes the learning object itself.	1	Unspecified	—	—	—
3.1	Identifier	A globally unique label that identifies this metadata record.	smallest permitted maximum: 10 items	Unspecified	—	—	—
3.1.1	Catalog	The name or designator of the identification or cataloging scheme for this entry. A namespace scheme.	1	Unspecified	Repertoire of ISO/IEC 10646-1:2000	CharacterString (smallest permitted maximum: 1000 char)	“Ariadne”, “URI”
3.1.2	Entry	The value of the identifier within the identification or cataloging scheme that designates or identifies this metadata record. A namespace specific string.	1	Unspecified	Repertoire of ISO/IEC 10646-1:2000	CharacterString (smallest permitted maximum: 1000 char)	“KUL532”, “https://www.ieee.org/descriptions/1234”
3.2	Contribute	Those entities (i.e., people or organizations) that have affected the state of this metadata instance during its life cycle (e.g., creation, validation). NOTE 18— This data element is concerned with contributions to the metadata. Data element 2.3.Lifecycle. Contribute is concerned with contributions to the learning object.	smallest permitted maximum: 10 items	Ordered	—	—	—
3.2.1	Role	Kind of contribution. Exactly one instance of this data element with value “creator” should exist.	1	Unspecified	creator validator	Vocabulary (State)	—

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
3.2.2	Entity	The identification of and information about entities (i.e., people, organizations) contributing to this metadata instance. The entities shall be ordered as most relevant first.	smallest permitted maximum: 10 items	Ordered	vCard, as defined by IBC vCard 3.0 (RFC 2425, RFC 2426)	CharacterString (smallest permitted maximum: 1000 char)	"BEGIN:VCARD\nFN:Joe Friday\nTEL:+1-919-555-7878\nTITLE:Area Administrator\nAssistant\nEMAIL;TYPE=INTERNET:jfriday@host.com\nEND:VCARD\n"
3.2.3	Date	The date of the contribution.	1	Unspecified	—	DateTime	"2001-08-23"
3.3	Metadata Schema	The name and version of the authoritative specification used to create this metadata instance. If multiple values are provided, then the metadata instance shall conform to multiple metadata schemas.	smallest permitted maximum: 10 items	Unordered	Repertoire of ISO/IEC 10646-1:2000	CharacterString (smallest permitted maximum: 30 char)	"LOMv1.0"

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
3.4	Language	Language of this metadata instance. This is the default language for all LangString values in this metadata instance. If a value for this data element is not present in a metadata instance, then there is no default language for LangString values. NOTE 19— This data element concerns the language of the metadata instance. Data element 1.3:General.Language concerns the language of the learning object.	1	Unspecified	See 1.3:General.Language. For this data element, "none" shall not be an acceptable value. NOTE 20—"none" is unacceptable, because the metadata instance is in one or more human languages. "none" is acceptable for 1.3:General.Language, as the learning object itself could be in no particular human language. For example, a picture of the Mona Lisa has "none" for 1.3:General.Language. If its description (i.e., metadata instance) is in Swedish, then 3.4:Metadata.Language has value "sv."	CharacterString (smallest permitted maximum: 100 char)	"en"
4	Technical	This category describes the technical requirements and characteristics of this learning object.	1	Unspecified	—	—	—
4.1	Format	Technical datatype(s) of (all the components of) this learning object. This data element shall be used to identify the software needed to access the learning object.	Smallest permitted maximum: 40 items	Unordered	MIME types based on IANA registration (IETF RFC 2048:1996) or "non-digital"	CharacterString (smallest permitted maximum: 500 char)	"video/mpeg", "application/x-toolbook", "text/html"

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
4.2	Size	The size of the digital learning object in bytes (octets). The size is represented as a decimal value (radix 10). Consequently, only the digits "0" through "9" should be used. The unit is bytes, not Mbytes, GB, etc. This data element shall refer to the actual size of this learning object. If the learning object is compressed, then this data element shall refer to the uncompressed size.	1	Unspecified	ISO/IEC 646:1991, but only the digits "0".."9"	CharacterString (smallest permitted maximum: 30 char)	"4200"
4.3	Location	A string that is used to access this learning object. It could represent a location (e.g., Universal Resource Locator), or a method that resolves to a location (e.g., Universal Resource Identifier). The first element of this list shall be the preferable location. NOTE 21— This is where the learning object described by this metadata instance is physically located.	Smallest permitted maximum: 10 items	Ordered	Repertoire of ISO/IEC 10646-1:2000	CharacterString (smallest permitted maximum: 1000 char)	"http://host/id"
4.4	Requirement	The technical capabilities necessary for using this learning object. If there are multiple requirements, then all are required, i.e., the logical connector is AND.	Smallest permitted maximum: 40 items	Unordered	—	—	—
4.4.1	OrComposite	Grouping of multiple requirements. The composite requirement is satisfied when one of the component requirements is satisfied, i.e., the logical connector is OR.	Smallest permitted maximum: 40 items	Unordered	—	—	—
4.4.1.1	Type	The technology required to use this learning object, e.g., hardware, software, network, etc.	1	Unspecified	operating system browser	Vocabulary (State)	—

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
4.4.1.2	Name	Name of the required technology to use this learning object. NOTE 22— The value for this data element is allowed to be derived from 4.1:Technical.Format automatically, e.g., "video/mpeg" implies "multi-os." NOTE 23— This vocabulary includes most values in common use at the time that this standard was approved.	1	Unspecified	if Type = "operating system", then: pc-dos ms-windows macos unix multi-os none if Type = "browser" then: any netscape communicator ms-internet explorer opera amaya	Vocabulary (State)	—
4.4.1.3	Minimum Version	Lowest possible version of the required technology to use this learning object.	1	Unspecified	Repertoire of ISO/IEC 10646-1:2000	CharacterString (smallest permitted maximum: 30 char)	"4.2"
4.4.1.4	Maximum Version	Highest possible version of the required technology to use this learning object.	1	Unspecified	Repertoire of ISO/IEC 10646-1:2000	CharacterString (smallest permitted maximum: 30 char)	"6.2"
4.5	Installation Remarks	Description of how to install this learning object.	1	Unspecified	—	LangString (smallest permitted maximum: 1000 char)	("en", "Unzip the zip file and launch index.html in your web browser.")
4.6	Other Platform Requirements	Information about other software and hardware requirements. NOTE 24— This element is intended for descriptions of requirements that cannot be expressed by data element 4.4:Technical.Requirement.	1	Unspecified	—	LangString (smallest permitted maximum: 1000 char)	("en", "sound card"), ("en", "runtime X")

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
4.7	Duration	Time a continuous learning object takes when played at intended speed. NOTE 25— This data element is especially useful for sounds, movies or animations.	1	Unspecified	—	Duration	"PT1H30M", "PT1M45S"
5	Educational	This category describes the key educational or pedagogic characteristics of this learning object. NOTE 26— This is the pedagogical information essential to those involved in achieving a quality learning experience. The audience for this metadata includes teachers, managers, authors, and learners.	Smallest permitted maximum: 100 items	Unspecified	—	—	—

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
5.1	Interactivity Type	Predominant mode of learning supported by this learning object. "Active" learning (e.g., learning by doing) is supported by content that directly induces productive action by the learner. An active learning object prompts the learner for semantically meaningful input or for some other kind of productive action or decision, not necessarily performed within the learning object's framework. Active documents include simulations, questionnaires, and exercises. "Expositive" learning (e.g., passive learning) occurs when the learner's job mainly consists of absorbing the content exposed to him (generally through text, images or sound). An expositive learning object displays information but does not prompt the learner for any semantically meaningful input. Expositive documents include essays, video clips, all kinds of graphical material, and hypertext documents. When a learning object blends the active and expositive interactivity types, then its interactivity type is mixed. NOTE 27— Activating links to navigate in hypertext documents is not considered to be a productive action.	1	Unspecified	active expositive mixed	Vocabulary (State)	Active documents (with learner's action): simulation (manipulates, controls or enters data or parameters); questionnaire (chooses or writes answers); exercise (finds solution); problem statement (writes solution). expositive documents (with learner's action): hypertext document (reads, navigates); video (views, rewinds, starts, stops); graphical material (views); audio material (listens, rewinds, starts, stops). mixed document: hypermedia document with embedded simulation applet.

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
5.2	Learning Resource Type	Specific kind of learning object. The most dominant kind shall be first. NOTE 28— The vocabulary terms are defined as in the OED:1989 and as used by educational communities of practice.	Smallest permitted maximum: 10 items	Ordered	exercise simulation questionnaire diagram figure graph index slide table narrative text exam experiment problem statement self assessment lecture	Vocabulary (State)	—
5.3	Interactivity Level	The degree of interactivity characterizing this learning object. Interactivity in this context refers to the degree to which the learner can influence the aspect or behavior of the learning object. NOTE 29— Inherently, this scale is meaningful within the context of a community of practice.	1	Unspecified	very low low medium high very high	Vocabulary (Enumerated)	NOTE 30— Learning objects with 5.1:Educational.InteractivityType = "active" may have a high interactivity level (e.g., a simulation environment endowed with many controls) or a low interactivity level (e.g., a written set of instructions that solicit an activity). Learning objects with 5.1:Educational.InteractivityType = "expositive" are defined as having a low interactivity level (e.g., a piece of linear, narrative text produced with a standard word processor) or a medium to high interactivity level (e.g., a sophisticated hyperdocument, with many internal links and views).

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
5.4	Semantic Density	The degree of conciseness of a learning object. The semantic density of a learning object could be estimated in terms of its size, span, or—in the case of self-timed resources such as audio or video— duration. The semantic density of a learning object is independent of its difficulty. It is best illustrated with examples of expositive material, although it can be used with active resources as well. NOTE 31— Inherently, this scale is meaningful within the context of a community of practice.	1	Unspecified	very low low medium high very high	Vocabulary (Enumerated)	<i>Active documents:</i> user interface of a simulation Low semantic density: a screen filled up with explanatory text, a picture of a combustion engine, and a single button labeled "Click here to continue" High semantic density: screen with short text, same picture, and three buttons labeled "Change compression ratio", "Change octane index", "Change ignition point advance" <i>Expositive documents:</i> Medium difficulty text document Medium semantic density: "The class of Marsupial animals comprises a number of relatively primitive mammals. They are endowed with a short placentation, after which they give birth to a larva. The larva thereafter takes refuge in the mother's marsupium, where it settles to finish its complete development." High semantic density: "Marsupials are primitive mammals, with short placentation followed by the birth of larva, which thereafter takes refuge in the marsupium to finish its development."

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
							<p>easy video document</p> <p>low semantic density: The full recorded footage of a conversation between two experts on the differences between Asian and African elephants; 30 min duration.</p> <p>high semantic density: An expertly edited abstract of the same conversation; 5 min duration</p> <p>difficult mathematical notation</p> <p>medium semantic density: The text representation of the theorem: For any given set ϕ, it is always possible to define another set ψ, which is a superset of ϕ.</p> <p>very high semantic density: The symbolic representation (formula) of the theorem ($\forall\phi \exists\psi: \psi \supset \phi$)</p>
5.5	Intended End User Role	<p>Principal user(s) for which this learning object was designed, most dominant first.</p> <p>NOTE 32— A learner works with a learning object in order to learn something. An author creates or publishes a learning object. A manager manages the delivery of this learning object, e.g., a university or college. The document for a manager is typically a curriculum.</p> <p>NOTE 33— In order to describe the intended end user role through the skills the user is intended to master, or the tasks he or she is intended to be able to accomplish, the category 9.Classification can be used.</p>	Smallest permitted maximum: 10 items		teacher author learner manager	Vocabulary (State)	An authoring tool that produces pedagogical material is a typical example of a learning object whose intended end user is an author

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
5.6	Context	The principal environment within which the learning and use of this learning object is intended to take place. NOTE 34— Suggested good practice is to use one of the values of the value space and to use an additional instance of this data element for further refinement, as in (“LOMv1.0”, “higher education”) and (“ http://www.ond.vlaanderen.be/onderwijsinvlaanderen/Default.htm ”, “kandidatuursonderwijs”)	Smallest permitted maximum: 10 items	Unordered	school higher education training other	Vocabulary (State)	—
5.7	Typical Age Range	Age of the typical intended user. This data element shall refer to developmental age, if that would be different from chronological age. NOTE 35— The age of the learner is important for finding learning objects, especially for school age learners and their teachers. When applicable, the string should be formatted as minimum age-maximum age or minimum age-. (note that this is a compromise between adding three component elements (minimum age, maximum age, and description) and having just a free text field.) NOTE 36— Alternative schemes for what this data element tries to cover (such as various reading age or reading level schemes, IQ’s or developmental age measures) should be represented through the 9:Classification category.	Smallest permitted maximum: 5 items	Unordered	—	LangString (smallest permitted maximum: 1000 char)	“7-9”, “0-5”, “15”, “18-”, (“en”, “suitable for children over 7”), (“en”, “adults only”)

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
5.8	Difficulty	How hard it is to work with or through this learning object for the typical intended target audience. NOTE 37— The “typical target audience” can be characterized by data elements 5.6:Educational.Context and 5.7:Educational.TypicalAgeRange.	1	Unspecified	very easy easy medium difficult very difficult	Vocabulary (Enumerated)	—
5.9	Typical Learning Time	Approximate or typical time it takes to work with or through this learning object for the typical intended target audience. NOTE 38— The “typical target audience” can be characterized by data elements 5.6:Educational.Context and 5.7:Educational.TypicalAgeRange.	1	Unspecified	—	Duration	“PT1H30M”, “PT1M45S”
5.10	Description	Comments on how this learning object is to be used.	smallest permitted maximum: 10 items	Unspecified	—	LangString (smallest permitted maximum: 1000 char)	(“en”, “Teacher guidelines that come with a textbook.”)
5.11	Language	The human language used by the typical intended user of this learning object.	smallest permitted maximum: 10 items	Unordered	See 1.3:General.Language	CharacterString (smallest permitted maximum: 100 char)	“en”, “en-GB”, “de”, “fr-CA”, “it” NOTE 39— As an example, for a learning object in French, intended for English-speaking students, the value of 1.3:General.Language will be French, and the value of 5.11:Educational.Language will be English.

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
6	Rights	This category describes the intellectual property rights and conditions of use for this learning object. NOTE 40—The intent is to reuse results of ongoing work in the Intellectual Property Rights and e-commerce communities. This category currently provides the absolute minimum level of detail only.	1	Unspecified	—	—	—
6.1	Cost	Whether use of this learning object requires payment.	1	Unspecified	yes no	Vocabulary (State)	—
6.2	Copyright and Other Restrictions	Whether copyright or other restrictions apply to the use of this learning object.	1	Unspecified	yes no	Vocabulary (State)	—
6.3	Description	Comments on the conditions of use of this learning object.	1	Unspecified	—	LangString (smallest permitted maximum: 1000 char)	("en", "Use of this learning object is only permitted after a donation has been made to Amnesty International.")
7	Relation	This category defines the relationship between this learning object and other learning objects, if any. To define multiple relationships, there may be multiple instances of this category. If there is more than one target learning object, then each target shall have a new relationship instance.	Smallest permitted maximum: 100 items	Unordered	—	—	—

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
7.1	Kind	Nature of the relationship between this learning object and the target learning object, identified by 7.2:Relation.Resource.	1	Unspecified	Based on Dublin Core: ispartof: is part of haspart: has part isversionof: is version of hasversion: has version isformatof: is format of hasformat: has format references: references isreferencedby: is referenced by isbasedon: is based on isbasisfor: is basis for requires: requires isrequiredby: is required by	Vocabulary (State)	—
7.2	Resource	The target learning object that this relationship references.	1	Unspecified	—	—	—
7.2.1	Identifier	A globally unique label that identifies the target learning object.	Smallest permitted maximum: 10 items	Unspecified	—	—	—
7.2.1.1	Catalog	The name or designator of the identification or cataloging scheme for this entry. A namespace scheme.	1	Unspecified	Repertoire of ISO/IEC 10646-1:2000	CharacterString (smallest permitted maximum: 1000 char)	"ISBN", "ARIADNE", "URI"

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
7.2.1.2	Entry	The value of the identifier within the identification or cataloging scheme that designates or identifies the target learning object. A namespace specific string.	1	Unspecified	Repertoire of ISO/IEC 10646-1:2000	CharacterString (smallest permitted maximum: 1000 char)	"2-7342-0318", "LEAO875", "https://www.ieee.org/"
7.2.2	Description	Description of the target learning object.	Smallest permitted maximum: 10 items	Unspecified	—	LangString (smallest permitted maximum: 1000 char)	("en", "The QuickTime movie of the Mona Lisa on the website of the Louvre museum.")
8	Annotation	This category provides comments on the educational use of this learning object, and information on when and by whom the comments were created. This category enables educators to share their assessments of learning objects, suggestions for use, etc.	Smallest permitted maximum: 30 items	Unordered	—	—	—
8.1	Entity	Entity (i.e., people, organization) that created this annotation.	1	Unspecified	vCard, as defined by IMC vCard 3.0 (RFC 2425, RFC 2426).	CharacterString (smallest permitted maximum: 1000 char)	"BEGIN:VCARD\nFN:Joe Friday\nnTEL:+1-919-555-7878\nTITLE:Area Administrator, Assistant\nEMAIL,TYPE=INTERN\nnET:jfriday@host.com\nEND:VCARD"
8.2	Date	Date that this annotation was created.	1	Unspecified	—	DateTime	"2001-08-23"
8.3	Description	The content of this annotation.	1	Unspecified	—	LangString (smallest permitted maximum: 1000 char)	("en", "I have used this video clip with my students. They really enjoy being able to zoom in on specific features of the painting. Make sure they have a broadband connection or the experience becomes too cumbersome to be educationally interesting.")

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
9	Classification	This category describes where this learning object falls within a particular classification system. To define multiple classifications, there may be multiple instances of this category.	Smallest permitted maximum: 40 items	Unordered	—	—	—
9.1	Purpose	The purpose of classifying this learning object.	1	Unspecified	discipline idea prerequisite educational objective accessibility restrictions educational level skill level security level competency	Vocabulary (State)	—
9.2	Taxon Path	A taxonomic path in a specific classification system. Each succeeding level is a refinement in the definition of the preceding level. Different paths, in the same or different classifications, which describe the same characteristic, are allowed.	Smallest permitted maximum: 15 items	Unordered	—	—	—
9.2.1	Source	The name of the classification system. NOTE 41—An indexation, cataloging or query tool could provide the top-level entries of a well-established classification, such as the Library of Congress Classification (LOC), Universal Decimal Classification (UDC), Dewey Decimal Classification (DDC), etc.	1	Unspecified	Repertoire of ISO/IEC 10646-1:2000	LangString (smallest permitted maximum: 1000 char)	("en", "ACM"), ("en", "MESH"), ("en", "ARIADNE")

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
9.2.2	Taxon	A particular term within a taxonomy. A taxon is a node that has a defined label or term. A taxon may also have an alphanumeric designation or identifier for standardized reference. The label and/or the entry may be used to designate a particular taxon. An ordered list of taxons creates a taxonomic path, i.e., "taxonomic stairway"; this is a path from a more general to more specific entry in a classification.	Smallest permitted maximum: 15 items	Ordered	—	—	{["12",("en","Physics")], ["23",("en","Acoustics")], ["34",("en","Instruments")], ["45",("en","Stethoscope")]} A 2nd taxon path for the same learning object could be: {"56",("en","Medicine")}, {"67",("en","Diagnostics")}, {"34",("en","Instruments")}, {"45",("en","Stethoscope")}]
9.2.2	Taxon	A particular term within a taxonomy. A taxon is a node that has a defined label or term. A taxon may also have an alphanumeric designation or identifier for standardized reference. The label and/or the entry may be used to designate a particular taxon. An ordered list of taxons creates a taxonomic path, i.e., "taxonomic stairway"; this is a path from a more general to more specific entry in a classification.	Smallest permitted maximum: 15 items	Ordered	—	—	{["12",("en","Physics")], ["23",("en","Acoustics")], ["34",("en","Instruments")], ["45",("en","Stethoscope")]} A 2nd taxon path for the same learning object could be: {"56",("en","Medicine")}, {"67",("en","Diagnostics")}, {"34",("en","Instruments")}, {"45",("en","Stethoscope")}]
9.2.2.1	Id	The identifier of the taxon, such as a number or letter combination provided by the source of the taxonomy.	1	Unspecified	Repertoire of ISO/IEC 10646-1:2000	CharacterString (smallest permitted maximum: 100 char)	"320", "4.3.2", "BF180"
9.2.2.2	Entry	The textual label of the taxon	1	Unspecified	—	LangString (smallest permitted maximum: 500 char)	("en", "Medical Sciences")

Table continues

Table 1—LOMv1.0 Base Schema (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
9.3	Description	Description of the learning object relative to the stated 9.1:Classification. Purpose of this specific classification, such as discipline, idea, skill level, educational objective, etc	1	Unspecified	—	LangString (smallest permitted maximum: 2000 char)	("en", "A medical instrument for listening called a stethoscope.")
9.4	Keyword	Keywords and phrases descriptive of the learning object relative to the stated 9.1:Classification. Purpose of this specific classification, such as accessibility, security level, etc., most relevant first.	Smallest permitted maximum: 40 items	Ordered	—	LangString (smallest permitted maximum: 1000 char)	("en", "diagnostic instrument")

*<http://www.dublincore.org/documents/dces/>

ANEXO III - TABLE 2 – LANGSTRING

Table 2—LangString

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
1	LangString	A datatype that represents one or more character strings. A LangString value is allowed to include multiple semantically equivalent character strings, such as translations or alternative descriptions.	Smallest permitted maximum: 10 items	Unordered	—	—	—
1.1	Language	Human language of the character string. NOTE—Indexation or cataloging tools can provide a useful default, for instance, the language of the user interface.	1	Unspecified	See 1.3:General.Language If no Language is specified, then LangString.String should be interpreted as a string in 3.4:Meta-Metadata.Language. If no language is specified, and a value for 3.4:Meta-Metadata is not present in the metadata instance, then the language for the LangString value is undefined.	CharacterString (smallest permitted maximum: 100 char)	“en”, “en-GB”, “de”, “fr-CA”, “it”
1.2	String	Actual character string.	1	Unspecified	Repertoire of ISO/IEC 10646-1:2000 excluding the NUL-character (UCS character U00000000)	CharacterString	“A picture of the Mona Lisa”

ANEXO IV - TABLE 3 – DATETIME

Table 3—Date Time

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
1	DateTime	A point in time with accuracy at least as small as one second.	1	Unspecified	<p>YYYY[-MM][-DD[Thh[:mm[:ss[.s[TZD]]]]]] where: YYYY = four-digit year (>= 0001) MM = two-digit month (01 through 12 where 01 = January, etc.) DD = two-digit day of month (01 through 31, depending on value of month and year) hh = two digits of hour (00 through 23) (am/pm NOT allowed) mm = two digits of minute (00 through 59) ss = two digits of second (00 through 59) s = one or more digits representing a decimal fraction of a second TZD = time zone designator ("Z" for UTC or +hh:mm or -hh:mm)</p> <p>At least the four digit year must be present. If additional parts of the DateTime are included, the character literals "-", "T", ":", and ".", are part of the character lexical representation for the datetime.</p> <p>If the time portion is present, but the time zone designator is not present, the time zone is interpreted as being UTC.</p> <p>NOTE 1—This value space is based on ISO8601:2000.*</p> <p>NOTE 2—The date portion only represents dates in the Common Era (CE). The date portion follows the Gregorian calendar for dates after October 15, 1582, and the Julian calendar for dates prior to October 05, 1582, independent of locale. Dates Before Common Era (BCE) and other cases should be represented using the "Description" data item.</p> <p>NOTE 3—The square bracket meta characters ("["; "]") indicate optional elements that appear zero or one time in the character lexical representation of the DateTime. These meta characters do not appear in the result; only the associated values described appear, e.g., "DD" is replaced by the corresponding 2 digit value for day of month.</p>	CharacterString (smallest permitted maximum: 200 char)	"1999-01-11" (January 11th, 1999) "1997-07-16T19:20:30+01:00" (July 16th, 1997, 30 s past 7.20 p.m. with a time offset of 1 h with respect to UTC)
2	Description	Description of the date.	1	Unspecified	—	LangString (smallest permitted maximum: 1000 char)	("en", "circa 1300 BCE"),

*See also <http://www.w3.org/TR/NOTE-datetime-970915.html>.

ANEXO V - TABLE 4 – DURATION

Table 4—Duration

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
1	Duration	An interval in time with accuracy at least as small as one second.	1	Unspecified	<p>P[yY][mM][dD][T][hH][nM][s[.s]S] where:</p> <p>y = number of years (integer, > 0, not restricted)</p> <p>m = number of months (integer, > 0, not restricted, e.g., > 12 is acceptable)</p> <p>d = number of days (integer, > 0, not restricted, e.g., > 31 is acceptable)</p> <p>h = number of hours (integer, > 0, not restricted, e.g., > 23 is acceptable)</p> <p>n = number of minutes (integer, > 0, not restricted, e.g., > 59 is acceptable)</p> <p>s = number of seconds or fraction of seconds (integer, > 0, not restricted, e.g., > 59 is acceptable)</p> <p>The character literal designators "P", "Y", "M", "D", "T", "H", "M", "S" must appear if the corresponding nonzero value is present.</p> <p>If the value of years, months, days, hours, minutes or seconds is zero, the value and corresponding designation (e.g., "M") may be omitted, but at least one designator and value must always be present. The designator "P" is always present. The designator "T" shall be omitted if all of the time (hours/minutes/seconds) are zero. Negative durations are not supported.</p> <p>NOTE 1— This value space is based on ISO8601:2000.*</p> <p>NOTE 2— The value is designated in the Gregorian calendar.</p> <p>NOTE 3— The ordering of durations may be indeterminate (e.g., 1 month may be 28, 29, 30 or 31 d).</p> <p>NOTE 4— For durations that apply only while the learning object is in use, but not when its use is suspended, it is recommended that only hours and smaller units of duration be used. Examples: PT43H, PT5M35S. For durations that express a time span, regardless of whether the learning object is actually used continuously during that time, days and larger units of duration are allowed to be used. Examples: P1Y6M, P20D.</p> <p>NOTE 5— The square bracket meta characters ("["", "]") indicate optional elements that appear zero or one time in the character lexical representation of the Duration. These meta characters do not appear in the result; only the associated values described appear, e.g., "dD" is replaced by the corresponding value for the number of days in the duration and is followed by the character literal designator "D."</p>	CharacterString (smallest permitted maximum: 200 char)	"PT1H30M" (1 h and 30 min)

Table continues

Table 4—Duration (continued)

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
2	Description	Description of the duration.	1	Unspecified	—	LangString (smallest permitted maximum: 1000 char)	("en", "Fall Semester 1999")

*See also: <http://www.w3.org/TR/xmlschema-2/#duration>.

ANEXO VI - TABLE 5 – VOCABULARY

Table 5—Vocabulary

Nr	Name	Explanation	Size	Order	Value space	Datatype	Example
1	Source	"LOMv1.0", or an indication of the source of the value, for instance through a URI (see 4.4).	1	Unspecified	—	CharacterString (smallest permitted maximum: 1000 char)	"LOMv1.0" " http://www.vocabularies.org/OSList "
2	Value	The actual value. If the source is "LOMv1.0", then the value shall come from the list defined in the LOMv1.0 Base Schema for the data element. In ISO 11404:1996 terminology, when there is implied ordering in the value list, then the values are "enumerated"; when there is no implied ordering, then they are "state."	1	Unspecified	—	CharacterString (smallest permitted maximum: 1000 char)	"MacOS"

ANEXO VII - TABLE A.1 – MAPPING TO UNQUALIFIED DUBLIN CORE

Table A.1—Mapping to unqualified Dublin Core Metadata Element Set

DC.Identifier	1.1.2:General.Identifier.Entry
DC.Title	1.2:General.Title
DC.Language	1.3:General.Language
DC.Description	1.4:General.Description
DC.Subject	1.5:General.Keyword or 9:Classification with 9.1:Classification. Purpose equals “Discipline” or “Idea”.
DC.Coverage	1.6:General.Coverage
DC.Type	5.2:Educational.LearningResourceType
DC.Date	2.3.3:LifeCycle.Contribute.Date when 2.3.1:LifeCycle.Contribute.Role has a value of “Publisher”.
DC.Creator	2.3.2:LifeCycle.Contribute.Entity when 2.3.1:LifeCycle.Contribute.Role has a value of “Author”.
DC.OtherContributor	2.3.2:LifeCycle.Contribute.Entity with the type of contribution specified in 2.3.1:LifeCycle.Contribute.Role.
DC.Publisher	2.3.2:LifeCycle.Contribute.Entity when 2.3.1:LifeCycle.Contribute.Role has a value of “Publisher”.
DC.Format	4.1:Technical.Format
DC.Rights	6.3:Rights.Description
DC.Relation	7.2.2:Relation.Resource.Description
DC.Source	7.2:Relation.Resource when the value of 7.1:Relation.Kind is “IsBasedOn.”

NOTE 1—The Dublin Core Metadata Initiative is also developing data element qualifiers to further refine the semantics of the Dublin Core data elements.¹⁰ A further refinement of the mapping in [Table A.1](#) can be based on these qualifiers.

NOTE 2—The LOM working group is committed to working with the Dublin Core Metadata Initiative (DCMI) to develop interoperable metadata, as outlined in the Memorandum of Understanding between the IEEE LTSC LOM WG and the DCMI.¹¹