

Alison Carlos Filgueiras

Um Modelo Semântico, Aberto e Conectado para Acervos Digitais de Patrimônio Cultural



Universidade Fernando Pessoa  
Porto, 2022



Alison Carlos Filgueiras

Um Modelo Semântico, Aberto e Conectado para Acervos Digitais de Patrimônio  
Cultural

Universidade Fernando Pessoa  
Porto, 2022





Alison Carlos Filgueiras

Um Modelo Semântico, Aberto e Conectado para Acervos Digitais de Patrimônio Cultural

Tese apresentada à Universidade Fernando Pessoa como parte dos requisitos para obtenção do grau de Doutor em Ciências da Informação, sob a orientação do Prof. Doutor Feliz Alberto Ribeiro Gouveia.



[...]

*Mas a índia explicou que o mais temível da doença da insônia não era a impossibilidade de dormir, pois o corpo não sentia cansaço, mas sim a sua inexorável evolução para uma manifestação mais crítica: o esquecimento. Queria dizer que quando o doente se acostumava ao seu estado de vigília, começavam a apagar-se da sua memória as lembranças da infância, em seguida o nome e a noção das coisas, e por último a identidade das pessoas e ainda a consciência do próprio ser, até se afundar numa espécie de idiotice sem passado.*

[...]

*Assim, continuaram vivendo numa realidade escorregadia momentaneamente capturada pelas palavras, mas que de fugir sem remédio quando esquecessem os valores da letra escrita.*

*Gabriel Garcia Marquez – Cem Anos de Solidão*



## RESUMO

ALISON CARLOS FILGUEIRAS: Um Modelo Semântico, Aberto e Conectado para Acervos Digitais de Patrimônio Cultural

(Sob orientação do Prof. Doutor Feliz Alberto Ribeiro Gouveia)

Essa tese parte da seguinte questão de investigação: “como integrar, de forma semântica, aberta e conectada, os acervos digitais de patrimônio cultural dos espaços de memória da Secretaria de Estado da Cultura do Estado de Goiás?”. O patrimônio cultural, segundo a UNESCO, é uma fonte insubstituível de vida e inspiração. Além de monumentos e lugares, esse patrimônio envolve nossas expressões e tradições enquanto povo. Nossos conhecimentos, técnicas e outros elementos, tangíveis ou não, precisam ser preservados e acessíveis às gerações futuras. No que diz respeito à preservação e acesso ao patrimônio cultural, existem atualmente várias ações dedicadas à organização, à recuperação e à digitalização dos diversos acervos presentes nos GLAMs (Galerias, Bibliotecas, Arquivos e Museus). Essas iniciativas de preservação e acesso geralmente esbarram em vários problemas relacionados à colaboração e à reutilização de informações de seus diversos e multiformes acervos digitais. O objetivo desta tese é apresentar um modelo de Sistema de Informação para permitir a interoperabilidade semântica entre os repositórios de patrimônio cultural através das ferramentas da Web Semântica. A metodologia utilizada foi a Pesquisa-Ação e Engenharia de Software e a amostragem é composta pelos espaços de memória da Secretaria de Estado da Cultura de Goiás. Como resultado apresenta o Guará, um protótipo de Sistema de Informação semântico que contribui para a organização e acesso à informação cultural em GLAMs para que curadores digitais interliguem os acervos destes.

**Palavras-chave:** sistemas de informação; interoperabilidade semântica; patrimônio cultural; acervos digitais; Estado de Goiás.

## ABSTRACT

ALISON CARLOS FILGUEIRAS: A semantic, open and connected model for digital collections of cultural heritage

(Under orientation of Professor Feliz Alberto Ribeiro Gouveia)

This thesis starts from the following research question: "how to integrate, in a semantic, open and connected way, the digital collections of cultural heritage from the memory centers of Secretary of Culture of the State of Goiás". Cultural heritage, according to UNESCO, is an irreplaceable source of life and inspiration. In addition to monuments and places, this heritage involves our expressions and traditions as a people. Our knowledge, techniques and other elements, tangible or not, need to be preserved and accessible to future generations. With regard to the preservation and access to cultural heritage, there are currently several actions dedicated to the organization, recovery and digitalization of the diverse collections present in the GLAMs (Galleries, Libraries, Archives and Museums). Such initiatives generally come up against several problems related to the collaboration and reuse of information from its diverse and multiform digital collections. The objective of this thesis is to present an Information System model to allow semantic interoperability between cultural heritage repositories through the tools of the Semantic Web. The methodology used was Action Research and Software Engineering and the sample is composed of the memory spaces of the Goiás State Secretariat for Culture. As a result, it presents Guar, a prototype of a Semantic Information System that contributes to the organization and access to cultural information in GLAMs so that digital curators can interconnect their collections.

**Keywords:** information systems; semantic interoperability; cultural heritage; digital collections; Goias State.

## RÉSUMÉ

ALISON CARLOS FILGUEIRAS: Un modèle sémantique, ouvert et connecté pour les collections numériques du patrimoine culturel

(Sous la direction du professeur Feliz Alberto Ribeiro Gouveia)

Cette thèse part de la question de recherche suivante: "comment intégrer, de manière sémantique, ouverte et connectée, les collections numériques du patrimoine culturel des espaces de mémoire du Secrétariat d'Etat à la Culture de l'Etat de Goiás ". Le patrimoine culturel, selon l'UNESCO, est une source irremplaçable de vie et d'inspiration. Outre les monuments et les lieux, ce patrimoine implique nos expressions et nos traditions en tant que peuple. Nos connaissances, techniques et autres éléments, tangibles ou non, doivent être préservés et accessibles aux générations futures. En matière de préservation et d'accès au patrimoine culturel, il existe actuellement plusieurs actions dédiées à l'organisation, la récupération et la numérisation des diverses collections présentes dans les GLAM (Galleries, Bibliothèques, Archives et Musées). De telles initiatives se heurtent généralement à plusieurs problèmes liés à la collaboration et à la réutilisation des informations issues de ses collections numériques diverses et multiformes. L'objectif de cette thèse est de présenter un modèle de système d'information pour permettre l'interopérabilité sémantique entre les référentiels du patrimoine culturel à travers les outils du Web sémantique. La méthodologie utilisée était la recherche-action et le génie logiciel et l'échantillon se compose des espaces de mémoire du Secrétariat d'État à la culture de Goiás. En conséquence, il présente un prototype qui contribue à l'organisation et à l'accès à l'information culturelle dans les GLAM afin que les conservateurs numériques s'interconnectent leurs collections.

**Mots clés:** systèmes d'information; interopérabilité sémantique; héritage culturel; collections numériques; État de Goiás

## **DEDICATÓRIA**

Aos que lutam pelo patrimônio cultural do Estado de Goiás. Os curadores, entusiastas, professores, estudantes e demais pessoas sensíveis à nossa memória enquanto povo, que atuam, cada um à sua maneira, contribuindo para a preservação e para a construção contínua da nossa identidade e consciência social.

## **AGRADECIMENTOS**

Há muito a agradecer aos tantos familiares, amigos, mestres e colegas que contribuíram direta ou indiretamente nessa trajetória de pesquisa:

Agradeço de forma muito especial os meus colegas nessa jornada, sobretudo, aos amigos Fabrizio, Paula e Paulo Henrique.

Agradeço à Juliana, companheira de vida e também colega de doutoramento, que para além do companheirismo, contribuiu com muitas ideias e debates durante toda a pesquisa.

Agradeço aos professores da UFP, em especial ao Prof. Feliz Gouveia, que foi um orientador cuidadoso e presente, e ao Prof. Luis Borges Gouveia, que sempre se colocou à disposição para tantas dúvidas que surgiram nesse percurso.

Agradeço, profundamente, ao Prof. Dalton (UnB) que, sempre receptivo, abriu as portas do grupo de pesquisa que coordena, nos atendeu pacientemente e apontou muitos caminhos para a área de pesquisa.

Aos professores Tathiana, Keley e Francisco da UEG, meus agradecimentos pelas leituras e sugestões.

Aos colegas da Gerência de Tecnologia da UEG, em especial Danyllo, Ibanez, Letícia, Stellio e Reinaldo, muito obrigado por suas sugestões e contribuições tecnológicas.

Agradeço também ao professor Renato Bulcão, da UFG, que contribuiu com sugestões de livros e artigos sobre ontologias e Web Semântica.

Aos curadores, servidores dos espaços culturais e gestores da SECULT-GO especialmente à Marília – que foi uma grande parceira nessa pesquisa – meus sinceros agradecimentos.

Pela leitura, colaborações e correção textual, agradeço ao amigo e professor Prof. Roberto Almeida.

Por fim, agradeço a generosidade, simpatia e acolhimento da Doutora Silvia Ribeiro, e em sua pessoa, estendo os agradecimentos a todos os servidores da UFP que foram sempre amigáveis e prestativos nesses cinco anos de jornada.



## SUMÁRIO

<b>ÍNDICE DE FIGURAS</b> .....	<b>XIX</b>
<b>ÍNDICE DE QUADROS</b> .....	<b>XXI</b>
<b>ÍNDICE DE ALGORITMOS</b> .....	<b>XXII</b>
<b>ÍNDICE DE SERIALIZAÇÕES</b> .....	<b>XXIII</b>
<b>CAPÍTULO I. INTRODUÇÃO</b> .....	<b>25</b>
1.1 VISÃO GERAL .....	25
1.2 MOTIVAÇÃO PARA O TRABALHO .....	26
1.3 CONTEXTO E RELEVÂNCIA .....	27
1.4 PROBLEMA E DESAFIOS.....	31
1.5 QUESTÃO DE PARTIDA .....	35
1.6 HIPÓTESE .....	36
1.7 OBJETIVOS .....	37
1.8 OBJETO DE ESTUDO E LIMITES DA PESQUISA .....	37
1.9 ESTRUTURA DA TESE .....	38
<b>CAPÍTULO II. PRESSUPOSTOS, LINGUAGENS E ESTRATÉGIAS DA CIÊNCIA DA INFORMAÇÃO.....</b>	<b>40</b>
2.1 INTRODUÇÃO .....	40
2.2 PRESSUPOSTOS DA CIÊNCIA DA INFORMAÇÃO: A INFORMAÇÃO DIGITAL E O PATRIMÔNIO CULTURAL .....	41
2.3 UMA NOVA REDE? A WEB, O PATRIMÔNIO CULTURAL E O PROBLEMA DA INTEROPERABILIDADE .....	46
2.4 LINGUAGENS E FERRAMENTAS DA WEB .....	51
2.5 MODELOS DE REFERÊNCIA .....	69
2.6 O CENTRO DE MEMÓRIA DE GOIÁS .....	76
2.7 SUMÁRIO DO CAPÍTULO.....	78
<b>CAPÍTULO III. INTEROPERABILIDADE SEMÂNTICA DE PATRIMÔNIO CULTURAL: UMA REVISÃO SISTEMÁTICA .....</b>	<b>79</b>
3.1 INTRODUÇÃO .....	79
3.2 SELEÇÃO DOS TRABALHOS .....	80
3.3 RESULTADOS DA BUSCA .....	81
3.4 TRATAMENTO DOS REGISTROS RETORNADOS .....	82
3.5 SÍNTESE DOS TRABALHOS .....	89
3.6 ARTIGOS EXCLUÍDOS APÓS LEITURA .....	118
3.7 SUMÁRIO DO CAPÍTULO.....	119
<b>CAPÍTULO IV. UM MODELO SEMÂNTICO, ABERTO E CONECTADO PARA O PATRIMÔNIO CULTURAL .....</b>	<b>121</b>
4.1 INTRODUÇÃO .....	121
4.2 ESTRATÉGIA UTILIZADA .....	121
4.3 PONTO DE PARTIDA: ARQUITETURA GENÉRICA .....	123
4.4 CONCEPÇÃO DO MODELO .....	126
4.5 PRIMEIRA CAMADA: SISTEMA WEB .....	130
4.6 SEGUNDA CAMADA: ARMAZENAMENTO (RDF STORE) .....	132
4.7 TERCEIRA CAMADA: SEMÂNTICA .....	133
4.8 QUARTA CAMADA: INTEROPERABILIDADE.....	143
4.9 UM ARCABOUÇO TECNOLÓGICO PARA OPERACIONALIZAÇÃO DA PROPOSTA .....	145
4.10 OPERACIONALIZAÇÃO DA PROPOSTA .....	150
4.11 SUMÁRIO DO CAPÍTULO .....	151

<b>CAPÍTULO V. PERCURSO METODOLÓGICO DA INVESTIGAÇÃO .....</b>	<b>152</b>
5.1 INTRODUÇÃO .....	152
5.2 INSTRUMENTOS E ESTRATÉGIAS .....	153
5.3 DIAGNÓSTICO INICIAL .....	154
5.4 ANÁLISE DOS DADOS .....	159
5.5 REVISÃO BIBLIOGRÁFICA E ENQUADRAMENTO TEÓRICO .....	160
5.6 ENGENHARIA DE REQUISITOS .....	161
5.7 CONSTRUÇÃO DE PROTÓTIPOS .....	163
5.8 SUMÁRIO DO CAPÍTULO.....	163
<b>CAPÍTULO VI. DA PESQUISA-AÇÃO À ENGENHARIA DE SOFTWARE:</b>	
<b>MODELAGEM DO PROTÓTIPO .....</b>	<b>164</b>
6.1 INTRODUÇÃO .....	164
6.2 RESULTADOS DO INQUÉRITO SOBRE OS ESPAÇOS DE MEMÓRIA DA SECRETARIA DE ESTADO DA CULTURA .....	164
6.3 RESULTADO DAS ENTREVISTAS PARA REFINAMENTO E LEVANTAMENTO DAS TECNOLOGIAS ENVOLVIDAS NO ACERVO .....	172
6.4 MODELAGEM DO PROTÓTIPO DE SISTEMA DE INFORMAÇÃO PARA OS ESPAÇOS DE MEMÓRIA.....	175
6.4.1 <i>DEFINIR ACERVOS</i> .....	178
6.4.2 <i>GERENCIAR OBJETOS FÍSICOS</i> .....	184
6.4.3 <i>GERENCIAR OBJETOS DIMENSIONAIS</i> .....	193
6.4.4 <i>ADICIONAR/REMOVER PROPRIEDADES</i> .....	199
6.4.5 <i>COLETAR LIGAÇÕES</i> .....	203
6.4.6 <i>ADMINISTRAR REPOSITÓRIO</i> .....	205
6.5 SUMÁRIO DO CAPÍTULO.....	208
<b>CAPÍTULO VII. GUARÁ: UM PROTÓTIPO DE SISTEMA DE INFORMAÇÃO SEMÂNTICO PARA O PATRIMÔNIO CULTURAL .....</b>	<b>209</b>
7.1 INTRODUÇÃO .....	209
7.2 ARQUITETURA E DESENVOLVIMENTO DO PROTÓTIPO .....	209
7.3 EXEMPLOS DE INTERAÇÃO.....	224
7.3.1 <i>CENÁRIO 1: ADICIONANDO UMA CLASSE NO ACERVO</i> .....	226
7.3.2 <i>CENÁRIO 2: ADICIONANDO UM OBJETO FÍSICO</i> .....	230
7.3.3 <i>CENÁRIO 3: ADICIONANDO UM OBJETO DIMENSIONAL</i> .....	235
7.3.4 <i>CENÁRIO 4: ADICIONANDO UMA NOVA RELAÇÃO</i> .....	239
7.3.5 <i>CENÁRIO 5: COLETAR LIGAÇÕES ENTRE REPOSITÓRIOS</i> .....	243
7.4 SUMÁRIO DO CAPÍTULO.....	246
<b>CAPÍTULO VIII. DOS ESPAÇOS DE MEMÓRIA AO SISTEMA DE INFORMAÇÃO: UMA ANÁLISE DOS RESULTADOS .....</b>	<b>247</b>
8.1 INTRODUÇÃO .....	247
8.2 OS ESPAÇOS DE MEMÓRIA DA SECRETARIA DE ESTADO DE CULTURA-GO .....	247
8.3 MODELOS CONCEITUAIS DE REPRESENTAÇÃO DA INFORMAÇÃO DE PATRIMÔNIO CULTURAL E TECNOLOGIAS DA WEB-SEMÂNTICA .....	256
8.4 UM MODELO DE SISTEMA DE INFORMAÇÃO SEMÂNTICO .....	258
8.5 PROTÓTIPO DO SISTEMA DE INFORMAÇÃO GUARÁ .....	261
8.6 O AMBIENTE SEMÂNTICO E A COMPLEXIDADE PARA OS CURADORES .....	265
8.7 SUMÁRIO DO CAPÍTULO .....	267
<b>CAPÍTULO IX. CONCLUSÃO E TRABALHOS FUTUROS .....</b>	<b>268</b>
9.1 INTRODUÇÃO .....	268
9.2 SUMÁRIO DA INVESTIGAÇÃO .....	268
9.3 QUESTÃO DE PESQUISA, QUESTÕES ACESSÓRIAS E HIPÓTESE .....	269
9.4 OBJETIVOS DO TRABALHO .....	271

9.5 CONTRIBUTOS DO TRABALHO .....	272
9.6 LIMITAÇÕES DO TRABALHO.....	273
9.7 TRABALHOS FUTUROS .....	273
9.8 RECOMENDAÇÕES.....	274
9.9 SUMÁRIO DO CAPÍTULO .....	275
<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>276</b>

## ÍNDICE DE FIGURAS

Figura 1 – Níveis de Interoperabilidade .....	48
Figura 2 – Esquema de Camadas da Web Semântica .....	51
Figura 3 – Representação básica de uma tripla RDF .....	53
Figura 4 – Exemplo de grafo RDF .....	55
Figura 5 – Estrutura básica de uma ontologia OWL .....	64
Figura 6 – Modelo 5 estrelas dos Arquivos Abertos Conectados.....	69
Figura 7 – Conceitos fundamentais do CIDOC-CRM.....	71
Figura 8 – Classes e propriedades básicas do CIDOC-CRM .....	73
Figura 9 – Exemplo de fato em CIDOC-CRM .....	74
Figura 10 – Classes do modelo EDM .....	75
Figura 11 – Esquema dos núcleos do CMG .....	78
Figura 12 – Critério de avaliação dos artigos .....	120
Figura 13 – Arquitetura geral da proposta.....	123
Figura 14 – Esquema de atividades associadas à arquitetura geral.....	125
Figura 15 – Exemplo de Repositórios.....	127
Figura 16 – Camadas do Modelo Genérico .....	128
Figura 17 – Visão dos fluxos de informação das camadas .....	129
Figura 18 – Esquema de Funcionamento dos Repositórios .....	130
Figura 19 – Estrutura Interna do Sistema Web.....	131
Figura 20 – Funcionamento interno da Camada de Armazenamento.....	132
Figura 21 – Dimensões do Centro de Memória .....	134
Figura 22 – Esquema das Ontologias da camada semântica .....	135
Figura 23 – Esquema da Ontologia de Definições de Classes .....	136
Figura 24 – Exemplo de organização do acervo de um cartório.....	137
Figura 25 – Esquema de organização inicial dos objetos.....	138
Figura 26 – Organização inicial dos conceitos.....	141
Figura 27 – Visão ampliada dos objetos.....	142
Figura 28 – Esquema da Camada de Interoperabilidade .....	143
Figura 29 – Esquema do arcabouço tecnológico.....	145
Figura 30 – Esquema de funcionamento geral do Apache Jena .....	146
Figura 31 – Exemplo de modelo RDF em <i>Apache Jena</i> .....	147
Figura 32 – Interface Web do Apache Jena Fuseki .....	149
Figura 33 – Uma visão do Protégé 5. ....	150
Figura 34 – Exemplo de Catálogo Físico .....	173
Figura 35 – Exemplos de planilhas de catálogos.....	174
Figura 36 – Banco de dados em Access do Museu da Imagem e do Som de Goiás .....	174
Figura 37 – Gerenciar Acervos Digitais.....	176
Figura 38 – Diagrama de Caso de Uso de Gerenciar Acervos .....	177
Figura 39 – Diagrama de Caso de uso de Definir Acervo .....	178
Figura 40 – Diagrama de Atividades de Definir Acervo.....	180
Figura 41 – Esquema Estrutural de Definir Acervo .....	181
Figura 42 – IHC – Pesquisar Classes .....	182
Figura 43 – IHC – Modificar Classe.....	183
Figura 44 – IHC – Destruir Classe .....	183
Figura 45 – IHC – Adicionar Classes.....	184

Figura 46 – Caso de Uso Gerenciar Objetos Físicos .....	185
Figura 47 – Diagrama de Atividades de Gerenciar Objetos Físicos.....	187
Figura 48 – Diagrama de Classes Gerenciar Objetos Físicos .....	188
Figura 49 – IHC – Adicionar Novo Objeto Físico.....	189
Figura 50 – IHC – Pesquisar Objetos Físicos .....	190
Figura 51 – IHC – Modificar Objeto Físico .....	191
Figura 52 – IHC – Destruir Objeto Físico.....	191
Figura 53 – IHC – Visualizar Propriedades.....	192
Figura 54 – Diagrama de Caso de Uso Gerenciar Objetos Dimensionais.....	193
Figura 55 – Diagrama de Atividades Gerenciar Objetos Digitais .....	195
Figura 56 – IHC – Adicionar Objetos Dimensionais .....	196
Figura 57 – IHC – Pesquisar Objetos Dimensionais.....	197
Figura 58 – IHC – Modificar Objeto Dimensional.....	197
Figura 59 – IHC – Destruir Objeto Dimensional .....	198
Figura 60 – Diagrama de Caso de Uso de Adicionar/remover Propriedades.....	199
Figura 61 – Diagrama de Atividades Adicionar/remover Propriedades .....	201
Figura 62 – IHC – Adicionar/Remover Propriedades.....	202
Figura 63 – Diagrama de Casos de uso Coletar Ligações .....	203
Figura 64 – Diagrama de Atividades Coletar Ligações .....	205
Figura 65 – Caso de Uso Administrar Repositório.....	206
Figura 66 – Diagrama de Atividades Administrar Repositório .....	207
Figura 67 – IHC – Administrar Repositório.....	208
Figura 68 – Arquitetura modular do protótipo .....	210
Figura 69 – Esquema das Páginas Web do Protótipo.....	211
Figura 70 – Página Inicial do Protótipo .....	212
Figura 71 – Esquema Funcional do Gerenciador de Eventos.....	215
Figura 72 – Esquema da Camada de Persistência .....	216
Figura 73 – Esquema de Armazenamento RDF (RDF Store).....	221
Figura 74 – Estrutura dos Acervos dos Repositórios de Exemplo .....	225
Figura 75 – Tela da Ação no Cenário 1.....	228
Figura 76 – Tela do resultado da ação no Cenário 1 .....	228
Figura 77 – Ficha de batismo de Pedro Ludovico Teixeira .....	230
Figura 78 – Ações do Curador no Exemplo de Interação 2 .....	231
Figura 79 – Resultados da Ação do Curador 2 na Interação 2 .....	234
Figura 80 – Ação do Curador 3 ao Adicionar Novo Objeto Dimensional .....	236
Figura 81 – Resultado da Ação do Curador 3 na Interação 3.....	239
Figura 82 – Ações do Curador no Cenário 3 – Adicionar Nova Propriedade .....	240
Figura 83 – Resultado da Ação do Curador 3 no Cenário 4.....	242
Figura 84 – Ilustração do Resultado da Ação de Interoperabilidade no Cenário 5.....	245
Figura 85 – Gibiteca Jorge Braga .....	249
Figura 86 – Imagens do Acervo do Arquivo Histórico do Estado de Goiás .....	250
Figura 87 – Objetos do Acervo do Museu Pedro Ludovico .....	251
Figura 88 – Imagem do Tour Virtual de Museus – SECULT-GO.....	252

## ÍNDICE DE QUADROS

Quadro 1 – Classes RDF/RDF-S.....	57
Quadro 2 – Propriedades RDF/RDF-S.....	57
Quadro 3 – Palavras-Chave por Base Pesquisada .....	80
Quadro 4 – Critério C4 - possuir técnicas ou estratégias de integração e/ou interoperação entre repositórios ou tecnologias.....	118
Quadro 5 – Descrição das Propriedades da Ontologia 1 .....	136
Quadro 6 – Descrição das propriedades da Ontologia 2 .....	139
Quadro 7 – Descrição das Propriedades da Ontologia 3 .....	142
Quadro 8 – Detalhamento dos métodos utilizados quanto aos objetivos .....	153
Quadro 9 – Sítios Eletrônicos dos Espaços de Memória .....	156
Quadro 10 – Base de Dados Utilizadas na Pesquisa.....	160
Quadro 11 – Sistemas de Informação presentes nos Acervos.....	172
Quadro 12 – Tabela de Requisitos da funcionalidade Gerenciar Acervos Digitais .....	176
Quadro 13 – Especificação do Caso de Uso Definir Acervo .....	179
Quadro 14 – Especificação do Caso de Uso Gerenciar Objetos Físicos .....	185
Quadro 15 – Especificação do Caso de Uso Gerenciar Objetos Dimensionais .....	193
Quadro 16 – Especificação do Caso de Uso Adicionar/Remover Propriedades .....	200
Quadro 17 – Especificação do Caso de Uso Coletar Ligações .....	204
Quadro 18 – Especificação do Caso de Uso Administrar Repositório.....	206
Quadro 19 – Estado dos Repositórios no Cenário 5.....	243
Quadro 20 – Objetos Retornados pelo Motor de Interoperabilidade no Cenário 5.....	245

## ÍNDICE DE ALGORITMOS

Algoritmo 1 – Exemplo de Consulta SPARQL .....	65
Algoritmo 2 – Exemplo de SPARQL em um evento de inserção de classes .....	133
Algoritmo 3 – Algoritmo para Interoperabilidade .....	144
Algoritmo 4 – Parte do código fonte da classe PIC .....	213
Algoritmo 5 – Código simplificado de Classe para Conexão com o TDB.....	216
Algoritmo 6 – Código simplificado de Conexão Remota com o TDB .....	217
Algoritmo 7 – Pesquisar Objetos Físicos por Palavra-chave .....	219
Algoritmo 8 – Pesquisar Objetos Dimensionais por Palavra-chave .....	219
Algoritmo 9 – Pesquisar Objetos Dimensionais por Relação .....	220
Algoritmo 10 – pesquisar Objetos Físicos por Dimensão.....	221
Algoritmo 11 – SPARQL do Cenário 1 – Inserindo uma Nova Classe. ....	229
Algoritmo 12 – SPARQL do Cenário 2 – Inserindo uma Novo Objeto Físico. ....	232
Algoritmo 13 – SPARQL do Cenário 3 – Inserindo uma Novo Objeto Dimensional. ....	237
Algoritmo 14 – SPARQL do Cenário 2 – Inserindo uma Nova Relação ao Objeto.....	241

## ÍNDICE DE SERIALIZAÇÕES

Serialização 1 – Exemplo de serialização RDF .....	55
Serialização 2 - Exemplo de Serialização RDF/XML da classificação de um acervo .....	137
Serialização 3 - Exemplo de Serialização de um objeto .....	140
Serialização 4 – Parte do arquivo pom.xml de um projeto Jena/Fuseki .....	148
Serialização 5 - RDF/XML da ontologia para Classificação dos acervos .....	222
Serialização 6 - Serialização em RDF/XML da ontologia de Definições de Objetos .....	222
Serialização 7 - Serialização da ontologia de definição das dimensões .....	223
Serialização 8 - Situação Inicial do TDB de classes do Acervo no Cenário 1. ....	227
Serialização 9 – Retorno em JSON da operação de Inserir Nova Classe .....	229
Serialização 10 – Resultado em JSON da operação de Inserir Novo Objeto Físico .....	233
Serialização 11 – Resultado em JSON da Operação de Visualizar Propriedades .....	233
Serialização 12 – Resultado em JSON da operação de Inserir Novo Objeto Físico .....	237
Serialização 13 – Resultado em JSON da Operação de Visualizar Propriedades .....	238
Serialização 14 – Resultado em JSON das relações manipuladas no Cenário 4.....	241

## LISTA DE ABREVIATURAS

ABNT	Associação Brasileira de Normas Técnicas
ACM	American Computer Machinery
API	Application Programming Interface
CIDOC-CRM	Modelo de Referência Conceitual CIDOC
DAO	Data Access Object (Objeto de Acesso a Dados)
DIS	Sistema de Ingestão de Documentos
EDM	Europeana Data Model
GLAM	Galerias, Bibliotecas, Arquivos e Museus
GO	Estado de Goiás
GUI	Graphical User Interface
HTML	Hyper Text Markup Language
HTTP	HyperText Transfer Protocol
ICCU	Instituto Central pelo Catálogo Único da Biblioteca Italiana
ICOM	International Council of Museums
IEEE	Institute of Electrical and Electronic Engineers
IRI	Internationalized Resource Identifier
JDBC	<i>Java Database Connectivity</i>
JSON	JavaScript Object Notation
JSON-LD	JavaScript Object Notation Linked Data
LIMS	Sistema de Informação para Laboratórios
LOD	<i>Linked Open Data</i>
N3	<i>Notation 3</i>
OWL	<i>Web Ontology Language</i>
PIC	<i>Plain Initial Class</i> (Classe Inicial Simples)
PIDO	<i>Plain Initial Data Object</i> (Objeto de Dados Inicial Simples)
PPV	<i>Plain Property Value</i> (Propriedade – Valor simples)
RDF	<i>Resource Description Framework</i>
RDF-S	<i>Resource Description Framework Schema</i>
SECULT	Secretaria de Estado da Cultura
SI	Sistema de Informação
SPARQL	<i>SPARQL Protocol and RDF Query Language</i>
SWRL	<i>Semantic Web Rule Language</i>
TDB	<i>Triple Database</i> - Banco de Dados em Tripla
TI	Tecnologia da Informação
Turtle	<i>Terse RDF Triple Language</i>
UIID	<i>Universally Unique Identify</i>
UML	<i>Unified Modeling Language</i>
UNESCO	Organização das Nações Unidas para Educação, Ciência e Cultura.
URI	<i>Uniform Resource Identifier</i>
URL	<i>Uniform Resource Locator</i>
W3C	<i>World Wide Web Consortium</i>
XHTML	<i>Extensible HyperText Markup Language</i>
XML	<i>Extensible Markup Language</i>

## CAPÍTULO I. INTRODUÇÃO

### 1.1 Visão geral

Essa tese aborda a memória e a cultura em uma perspectiva da Ciência da Informação que envolve organização, guarda, busca e disseminação de fragmentos, aqui denominados objetos digitais, em um recorte sobre o patrimônio cultural digitalizado ou digitalizável, aqui denominado acervo digital de patrimônio cultural. E para fazê-lo, utiliza-se dos conceitos e métodos da Web Semântica e, conseqüentemente, dos Dados Abertos Conectados (*Linked Open Data - LOD*).

O problema central abordado é o da interoperabilidade semântica de patrimônio cultural (*Cultural Heritage*), especificamente no que se refere ao patrimônio material e imaterial do Estado de Goiás, e, portanto, tem como objeto os espaços de memória da Secretaria de Estado da Cultura (SECULT-GO), mais precisamente os museus, bibliotecas, arquivos e galerias (GLAMs) da Superintendência de Patrimônio desta secretaria. Diante desse cenário específico da SECULT-GO e dados os problemas já conhecidos quanto à sistematização, estratégias e serviços para compartilhamento, descoberta, acesso e reuso de dados de patrimônio cultural conforme Doerr (2009;2019) e Meghini *et al.* (2017), e as ferramentas da Web Semântica (Berners-Lee *et al.*, 2001), a investigação pretendeu, através de uma pesquisa-ação realizada nos espaços de memória da SECULT-GO, contribuir com um modelo de sistemas de informação que possa auxiliar nas ações de acesso e preservação deste patrimônio nos diversos processos de digitalização dos fenômenos e coisas presentes e circundantes naqueles espaços. Com um olhar mais pontual ao que concerne à interligação entre repositórios em um nível de contexto que permita que os curadores possam ampliar as informações digitais dos objetos sob guarda de tais centros de memória. O modelo proposto é constante de um conjunto de práticas metodológicas, tecnológicas e conceituais para a interoperabilidade semântica em 4 dimensões: Pessoas, Tempo, Local e Fenômeno aplicados ao patrimônio cultural do Estado de Goiás, especificamente nesses espaços de memória do recorte da investigação.

O estado de Goiás é um dos 26 estados do Brasil localizado na região Centro-Oeste, possui aproximadamente 7 milhões de habitantes, segundo dados do IBGE em 2019, perfazendo uma extensão de aproximadamente 340 mil km<sup>2</sup> e uma ampla diversidade cultural latente de ações de preservação.

O primeiro espaço de memória do Estado de Goiás, o Museu Zoroastro Artiaga, em Goiânia, data da década de 40, é um dos exemplos de instituição de memória que apresenta diversos elementos do patrimônio material e imaterial do povo goiano. Há outros espaços com o mesmo objetivo sob jurisdição da Secretaria de Cultura do Estado de Goiás. Também muitos outros espalhados pelas diversas cidades do estado de Goiás que, de uma maneira ou de outra, contribuem para a preservação do patrimônio e possuem acervos potencialmente viáveis para serem digitalizados e expostos sob diversas perspectivas. É também objetivo deste trabalho trazer contribuições na relação potencial entre estes espaços de memória, entendendo de forma preliminar, que existe uma forte motivação social, científica e tecnológica em cuja relevância, no contexto da Sociedade em Rede, parece ser urgente em vários sentidos, em especial, auxiliar na preservação e no acesso aos bens culturais.

Este capítulo apresenta os aspectos introdutórios e iniciais da pesquisa, bem como seu *locus* em torno do problema abordado, passando pela motivação, contexto e relevância, desafios, questões de pesquisa, objetivos, hipóteses, justificação, objetos e limites, e, por fim, a estrutura da tese.

## **1.2 Motivação para o trabalho**

Há como ponto de partida uma inquietação pessoal e acadêmica durante a atuação no projeto de extensão universitária denominado Centro de Memória de Goiás, que desde sua gênese tem a pretensão insólita de contribuir para que as informações de patrimônio cultural do povo goiano possam ser, de forma livre e colaborativa, armazenadas e disponibilizadas para as comunidades. Pretende envolver professores, estudantes e parceiros externos em ações de digitalização, armazenamento, pesquisa e exposição, dentre outras.

Pessoal e academicamente, a motivação da investigação está, desde o início, associada a esse projeto — que é melhor detalhado na Seção 2.6 - p. 76 — e que se iniciou em forma de ação de extensão institucional vinculada à Pró-Reitoria de Extensão, Cultura e Assuntos Estudantis da Universidade Estadual de Goiás (UEG), cujas frentes de atuação direcionam-se à preservação e valorização do patrimônio material e imaterial do povo Goiano, conforme afirmação em (Chagas *et al.*, 2018, p. 4):

Comumente, um centro de memória se caracteriza pelos trabalhos de organização e gestão de documentos de determinada instituição, com um caráter

meramente instrumental. Na prática, temos vivenciado um fenômeno onde o termo “memória”, presente no nome desses locais, adquire um significado muito mais amplo, no sentido de rememorar e perpetuar aspectos essenciais da história de um local ou um povo, adquirindo *status* de centros de difusão do saber e espaços de formação culturais (Chagas *et al.*, 2018, p. 4).

No entanto, esse e outros projetos relacionados às humanidades digitais tendem a enfrentar duas dificuldades básicas. A primeira, relacionada a limitações tecnológicas em vários níveis. A segunda, está relacionada à quantidade de informações geradas nos processos de digitalização que têm contribuído para o aumento massivo dos dados na Web, mas em direção contrária, não ao seu aproveitamento. Parte disso porque em termos de organização, armazenamento e distribuição, o processo digital como um todo depende de arquiteturas e ferramentas que estão além do esforço da transformação dos acervos em formato digital.

Também é fator motivador o estado de Goiás possuir um grande número de oportunidades de digitalização nos GLAMs, a começar por acervos quase inexplorados presentes nas cidades mais antigas desconhecidos de grande parte da população com difícil acesso a pesquisadores. Há ainda registros históricos da memória oral e imagem e muitos outros artefatos com grande potencial para produção de conhecimento sobre a cultura do povo goiano. Nesse contexto, portanto, “as coisas estão para serem feitas”.

### **1.3 Contexto e relevância**

Os bens materiais e imateriais representam a cultura de um povo e pode se dizer que há um consenso dentro da comunidade científica que devem ser preservados. Guardar, colecionar, proteger e transmitir os vestígios da produção intelectual, material, cultural ou social é característica tão humana quanto a necessidade de defesa e preservação da própria espécie (Menezes, 2006).

Isso vem ao encontro a Vannevar Bush naquilo que pode ser considerado como um pensamento precursor da era digital, em que, no seu texto mais clássico - “*As we may think*” -, este autor afirma que “se um documento é importante para a ciência, deve ser preservado, armazenado e, principalmente, consultado” (Bush, 1945 s.p.).

Segundo a UNESCO (Unesco, 2020), o patrimônio cultural se trata de uma fonte insubstituível de vida e inspiração. Para além dos monumentos e lugares, este patrimônio envolve nossas expressões e tradições enquanto povo e vê-se em uma

emergente necessidade de reconhecimento, valorização e proteção em um momento em que a informação e o conhecimento são vistos como elementos fundamentais de uma sociedade em rede.

Esta sociedade se desenvolve freneticamente e torna-se cada vez mais dependente das ferramentas tecnológicas, sobretudo da Internet que, em 2021, viu o número de usuários ultrapassar a casa dos 4,66 bilhões, com uma taxa de penetração de 59,5% da população mundial<sup>1</sup>.

Assim, como a própria Internet, diferentes objetos e métodos foram historicamente utilizados para a preservação dos diversos bens culturais e contribuíram, de acordo com o tempo, para a disseminação destes. As pinturas rupestres, a história oral e a prensa nos permitem visualizar diferentes formas de uma constante luta em favor da memória dos bens culturais.

Estas tecnologias, quase sempre envolvidas de símbolos, cores, misturas e mistérios — seja qual o formato e meio utilizados —, serviram para carregar pelas eras o que pode ser visto hoje como nosso conhecimento, e, portanto, nossa cultura — sem qualquer pretensão de entrar na seara dos questionamento sobre o que é conhecimento ou que é cultura quem, conforme Peter Burke (Burke, 2003, p. 19), “é uma pergunta tão capciosa como *o que é verdade?*”.

Na perspectiva das tecnologias da informação e comunicação digitais, em especial com os Sistemas de Informação, há várias décadas em que é experimentado um aumento considerável na relação “produtos e consumo” do digital através da construção de acervos digitais de forma que possam ser organizados, protegidos e disseminados em vários formatos, especialmente pela Internet.

Há também que reconhecer que houve (e ainda há em grande parte do mundo, e em menor quantidade, no Brasil) um conjunto de políticas públicas que, aliadas a iniciativas privadas com ou sem fins lucrativos, tem resultado em um significativo número de projetos nacionais e internacionais a buscar a valorização da cultura e memória, muitas vezes de forma localizada, evidenciando um contraste à crescente exposição dos fenômenos globalizados com a massificação dessas novas tecnologias da informação e comunicação.

---

<sup>1</sup> <https://datareportal.com/reports/digital-2021-global-overview-report>

Alguns desses projetos foram destinados a expor estruturas de dados, conjuntos de dados integrados, vocabulários e ontologias para apoiar outras iniciativas no desenvolvimento de aplicativos na área de patrimônio digital, alguns outros simplesmente representam implementações de aplicativos ou soluções localizadas (Calvanese *et al.*, 2016).

A Rede e suas tecnologias, por si só, apresentam desafios que são experimentados e acrescidos a estas à medida em que se desenvolvem. Enquanto fenômeno tecnológico, a Internet permite um aumento significativo de repositórios digitais desde as últimas décadas, e esse número cresce a cada dia, de forma que há uma vastidão de dados sobre praticamente todo conhecimento humano. A maior parte desses dados, no entanto, estão organizados nas redes para a recuperação através do contato de usuários humanos com o conteúdo.

Para além do que está facilmente acessível na Web, há uma grande quantidade de dados difíceis ou totalmente inalcançáveis em formatos diversos, como arquivos de texto, planilhas e principalmente em bancos de dados relacionais, os quais são gerados incrementalmente à uma velocidade espantosa a contribuir para este fenômeno conhecido como *Big Data*.

O *Big Data*, segundo Olivier (2015), pode ser considerado como um fenômeno concreto que “está atraindo todos os tipos de empresas que enfrentam questões estratégicas e decisórias, e também campos científicos interessados em entender observações complexas que podem levar a importantes descobertas” (Curé & Blin, 2015, p. 1).

Com o advento da Web Semântica, diversas iniciativas e estratégias de organização da informação foram propostas pela comunidade científica, de forma a permitir uma nova abordagem que facilitasse o intercâmbio de informações.

Kiryakov (2006) afirma que repositórios semânticos permitem o armazenamento, a consulta e o gerenciamento de dados estruturados, podendo substituir os gerenciamentos de banco de dados tradicionais.

Isso porque os repositórios permitem mais facilidade de integração e mais poder analítico, ou seja, “um repositório semântico pode interpretar dinamicamente esquemas e ontologias de metadados que determinam a estrutura e a semântica dos dados e das consultas nestes” (Kiryakov, 2006, p. 116).

De forma acessória à diversidade de fontes de dados, há que se levar em conta que a *Web* foi construída como uma plataforma para garantir que usuários pudessem saltar nos conteúdos através dos hipertextos e apresentou para isso dois elementos básicos, o protocolo HTTP (*Hyper Text Transfer Protocol*) que permite requisições e respostas na rede, e a linguagem HTML (*Hyper Text Markup Language*) que permite, por meio de um navegador, que os usuários humanos possam visualizar o conteúdo dos hipertextos.

Do ponto de vista da interoperação, isso representa uma limitação. Com base nessa limitação, Berners-Lee propôs a Web Semântica, cujo maior objetivo é melhorar as informações de forma a prover significados através de um conjunto de padrões e protocolos que têm como intenção expandir essa interação também para os algoritmos e outras máquinas na Internet, possibilitando assim, que novos conhecimentos sejam gerados. (Berners-Lee *et al.*, 2001).

Para além das ligações comumente conhecidas no mundo HTTP (*Links* e documentos), nessa extensão da *Web*, o contexto dos objetos digitais é altamente relevante e a recuperação de informação assume uma nova e mais emergente complexidade (Berners-Lee *et al.*, 2001).

A Web Semântica tem se configurado como uma plataforma atual de compartilhamento de dados e, de forma interdisciplinar, que se aplica com os pesquisadores da Ciência da Informação (Souza *et al.*, 2018). Associada à iniciativa dos dados abertos, há grandes oportunidades para construção de sistemas de informação semânticos, abertos e conectados.

Dados abertos, segundo a *Open Knowledge International* (OKF, 2020), são dados que podem ser livremente acessados, utilizados, modificados e compartilhados por quem acessá-los para qualquer finalidade. Esse acesso está sujeito a, no máximo, exigências que visem preservar sua proveniência e sua abertura. Já os Dados Abertos Conectados, segundo a W3C (W3C, 2021), são um conjunto de iniciativas que garantem uma visão de dados vinculados e acessíveis globalmente na Internet com base na Web Semântica.

Nesta visão, a *Web* é abordada como um grande banco de dados virtual, que qualquer pessoa (ou sistema) pode acessar qualquer dado que lhe esteja autorizado e também pode, sob determinadas condições, adicionar dados complementares sem perturbar a fonte dos dados original. Através de um ambiente aberto, os dados podem ser

criados, conectados e consumidos, gerando mais valor. Isto é, os dados serão mais significativos se puderem ser conectados a outros dados (Berners-Lee, 2006).

Para além de uma visão tecnicista dos elementos da Web Semântica, busca-se focalizar o fenômeno social que assume, perante a Sociedade em Rede, uma abordagem interdisciplinar, partindo da premissa da cultura enquanto direito fundamental, sendo do patrimônio cultural o conjunto de todas as coisas referentes à herança cultural enquanto povo, e da visão de que repositórios de patrimônio cultural como dados abertos pode contribuir para a geração de ações de preservação, governança, educação e outras relacionadas à cultura.

A ciência, com a preocupação essencial em responder a questões fundamentais da sociedade, também é impactada pelos fenômenos informacionais, vê-se aqui uma possibilidade concreta de contribuição, conforme explicitam os autores (Carvalho Junior *et al.*, 2019, p. 13).

A ciência, obviamente, é também impactada diretamente pela explosão da informação digital e, coloca-se em questão sua capacidade em responder aos complexos dilemas informacionais da sociedade contemporânea.

O professor, o estudante, o Curador, o gestor, ou qualquer outro usuário da informação do século XXI não querem simplesmente consumir a informação como em outros tempos o fazia. Estes querem participação. Almejam a possibilidade de produzir, indexar, classificar, contestar, interferir; em outras palavras, querem ser sujeitos ativos em todos os processos da informação (Araújo, 2017), e um dos caminhos possíveis é a construção de uma base de processos da informação digital constituída à luz dos dados abertos conectados.

#### **1.4 Problema e desafios**

Como evidenciado na primeira parte introdutória desta tese, há inúmeros desafios a serem enfrentados pela comunidade científica nas temáticas que relacionam os sistemas de informação, o patrimônio cultural, a Rede, seus atores e fenômenos. No entanto, pretende-se aqui fechar o foco em uma parte fundamental que está relacionada, especificamente, ao problema de integração e colaboração entre fontes heterogêneas de informação que, de maneira incidental, envolve organização e recuperação da informação e gestão da informação.

Os espaços de memória, de uma maneira ou de outra, possuem acervos e

repositórios (ativos ou potenciais) de documentos digitais. Ao observar os tantos desafios enfrentados por estes espaços (justificados e detalhados no CAPÍTULO V), optou-se por abordar um problema científico presente de forma transversal nas tecnologias que envolvem o GLAMs através de abordagem sistemática em um problema também técnico que envolve os espaços de memória e os sistemas de informação para o patrimônio cultural. Na literatura há uma vasta lista de estratégias para tentar responder os desafios impostos pelo problema da interoperabilidade semântica. A criação de agregadores como na Europeia (Europeana, 2017; N Freire *et al.*, 2018), extensões de linguagens como em Bannour *et al.* (2018) (Bannour *et al.*, 2018), mapeamento e harmonização de conceitos (Carrasco *et al.*, 2017; Daquino *et al.*, 2017), dentre outras.

Sem abandonar a perspectiva sociocultural, o problema abordado nesta tese é o da interoperabilidade semântica entre repositórios de patrimônio cultural, e, para tanto parte de uma visão deste patrimônio em repositórios digitais que são sistemas de informação que devem estar na Web. A questão da interoperabilidade é um dos problemas clássicos dos sistemas de informação e está relacionada à necessidade de integração de sistemas com determinados graus de heterogeneidade.

A integração de sistemas de informação só é possível garantindo de alguma forma e em algum nível a interoperabilidade. Os sistemas são formados, no geral, por fontes de dados diversas tecnológica, semântica e sintaticamente. Se fazem necessárias estratégias que permitam uma troca qualificada de informações para que estas sejam mutuamente úteis.

Davies *et al.* (2006) afirmam que a integração é um desafio importante para os gestores de Tecnologia da Informação (TI), uma vez que os custos dos processos de integração, interna e externamente, são um componente significativo de seu orçamento (Davies *et al.*, 2006).

Essa necessidade de troca de informações se esbarra, inicialmente, em dois desafios importantes, um sintático e outro semântico. Sobre essa barreira (Hebeler *et al.*, 2009, p. 65) apresenta o seguinte paralelo:

Considere o caso de duas pessoas que estão tentando se comunicar. Existem muitas maneiras de se comunicar e, dependendo das circunstâncias, elas podem escolher uma abordagem ou outra. Eles podiam falar ou escrever um para o outro, mas também podiam usar Código Morse, sinalizadores de semáforo ou até algo mais exótico, como sinais de fumaça. Encontrar um meio comum para

comunicação é análogo ao aspecto sintático do compartilhamento de dados. É um primeiro passo muito importante no problema de compartilhamento de informações.

Nesse paralelo, aborda-se a questão sintática que envolve a comunicação, que é escolher um meio adequado na leitura desta. Uma vez que esse meio é escolhido, as duas pessoas devem concordar com um idioma franco compartilhado, que se trata de uma codificação para suas ideias, e é a partir daí que se apresenta o problema do compartilhamento semântico, conforme Hebler *et al.* (2009).

Uma linguagem é uma maneira de mapear as ideias em nossas cabeças de uma forma que possa ser comunicada com outras pessoas. Para se comunicar, as pessoas devem concordar com um idioma médio e comum. Se duas pessoas falam idiomas diferentes, elas não serão capazes de se entender, mesmo que tenham decidido se comunicar por meio de, por exemplo, palavras faladas. Encontrar uma codificação mutuamente inteligível de conceitos dentro do meio comum é análogo ao problema de compartilhamento de dados semântico (Hebler *et al.*, 2009, p. 65) .

Este problema, que é conhecido como o problema da interoperabilidade semântica, está epistemologicamente ancorado no campo de estudos da representação e recuperação da informação (Araújo, 2014) e tem suas raízes nos primórdios da Web.

Conforme Lima *et al.* (2016), a Web nasceu descentralizada, funcionando como um grande depósito de documentos, por isso, há imensa dificuldade para a recuperação das informações, tornando impossível que agentes automatizados compreendam o sentido que os dados possuem (Lima *et al.*, 2016).

Os sistemas de informação do patrimônio cultural, que tratam geralmente sobre o domínio, a guarda, os processos de aquisição (digitalização) e exposição de objetos digitais, também vivenciam em seus repositórios os desafios da interoperabilidade. Doerr (2009) afirma que é comum as pessoas imaginarem o patrimônio cultural como um conjunto de objetos de coleções de arte cuja descrição e indexação estão mais relacionadas ao campo da reflexão acadêmica do que como uma questão de lógica útil para sua organização.

Para Doerr (2009), o patrimônio cultural é muito mais que um domínio. “Compreende um amplo espectro de funções sobre o estudo e preservação de evidências físicas do passado de todos os tipos de atividade humana” (Doerr, 2009 p. 2).

No sentido estrito, pode-se considerar o patrimônio cultural como as coisas a serem preservadas pelas instituições de memória, museus, bibliotecas, arquivos, galerias e outros espaços (GLAMs), as organizações guarda-chuva, por manter padrões e protocolos sobre esses domínios específicos, a rede (Web) e seus diversos usuários nas mais diversas possibilidades de uso como componentes desse problema.

GLAM representa um acrônimo em inglês para “galerias, bibliotecas, arquivos e museus”, que muitas vezes são chamados espaços ou instituições de memória, como pode ser encontrado em (Carvalho Junior *et al.*, 2019, p. 13)

[...] o termo “Instituição de Memória” tem ganhado proeminência quando nos referimos a museus, bibliotecas, arquivos, cinematecas, centros de documentação cultural e instituições similares. Outra referência que representa a agregação natural do campo na era da cultura digital é o termo “GLAM” - um acrônimo de “galerias, bibliotecas, arquivos e museus”, e refere-se a instituições culturais que têm o acesso ao conhecimento como missão.

Os sistemas de informação para esse fim têm que frequentemente permitir a manipulação de informações com uma terminologia muito rica e em diferentes idiomas. Tais terminologias são geralmente muito específicas e apresentam, para além de questões técnicas, desafios intelectuais na aproximação e harmonização de conceitos por meio de deduções lógicas que requerem estratégias complexas de semântica (Doerr, 2009; Doerr *et al.*, 2007).

Estes autores abordam a complexidade dos sistemas de patrimônio cultural e apresentam aquilo que seria uma das facetas do problema da interoperabilidade semântica:

A maioria dos sistemas profissionais mencionados é baseada em esquemas de banco de dados bastante complexos. Por exemplo, o CIDOC propôs até 1995 um esquema relacional padrão para museus com mais de 400 tabelas (Doerr *et al.*, 2007, p. 6).

Como descrito acima, as informações sobre o patrimônio cultural são distribuídas em muitos sistemas diferentes que se complementam. Uma fonte pode relacionar nomes romanos às inscrições romanas, outras inscrições romanas às pedras, outras pedras ao local da descoberta e outro local às coordenadas. Ainda assim, porém, a maioria dos esforços para integrar informações patrimoniais concentra-se em encontrar

elementos mínimos de descrição comum para objetos, como auxílio em vez de documentação (Doerr, 2009).

Tradicionalmente, padrões de metadados têm sido utilizados para permitir interoperabilidade sintática entre sistemas de informação. O compartilhamento sintático de informações é extremamente útil, nem sempre, entretanto, consegue abranger a maior parte das questões semânticas envolvidas.

Essa diferença é fundamental, uma vez que o aspecto sintático do problema envolve obter acesso aos dados compartilhados, enquanto o aspecto semântico envolve a incorporação dessas informações nas estruturas de dados dos sistemas envolvidos (Hebeler *et al.*, 2009).

Os Sistemas de Informação de patrimônio cultural, segundo Doerr (2009), destinam-se à Gestão de acervos, Conservação, Pesquisa e Apresentação (publicação) do patrimônio cultural. O autor afirma que essas informações são advindas de fontes de dados diversificadas com estratégias específicas, em cada uma dessas quatro atividades básicas. Tais sistema são criados e mantidos por diferentes atores.

Levando-se em consideração que as informações em todos esses sistemas se sobrepõem e devem estar mutuamente acessíveis, um dos principais desafios da gestão da informação do patrimônio cultural é a interoperabilidade desses sistemas e a integração de informações (Boeuf *et al.*, 2018; Doerr, 2009).

### **1.5 Questão de partida**

Como integrar, de forma semântica, aberta e conectada, os acervos digitais de patrimônio cultural dos espaços de memória da Secretaria de Estado da Cultura do Estado de Goiás?

Considerando uma quantidade crescente de acervos culturais digitais nos mais diversos formatos e com estratégias distintas de organização da informação. Também considerando que cada fragmento de informação é relevante e atende a determinados contextos, considerando, por fim, a crescente necessidade de integração dos conteúdos de forma a gerar novos conhecimentos, essa é a questão de partida que conduziu essa investigação.

De forma acessória, outras questões complementares são elencadas para serem respondidas: Quais ferramentas e estratégias podem ser utilizadas? Quais modelos podem

ser adotados ou reutilizados? A complexidade do modelo semântico pode ser um empecilho para os curadores digitais?

Sobre os espaços de memória e seus acervos digitais, tanto em uma visão geral quanto no recorte específico dessa tese, há múltiplos olhares possíveis no que diz respeito a utilização e integração de suas fontes heterogêneas de informação. É sabido, conforme explicitado na primeira parte desse capítulo, que há muitas iniciativas e ferramentas que dão suporte a ações de digitalização do patrimônio cultural.

Não se pretende, de forma simples e direta responder à pergunta dessa tese com um sistema de informação, reduzindo toda a complexidade social envolvida a uma resposta funcional e pragmática. Pelo contrário, parte-se da visão que, com tantas ferramentas e estratégias disponíveis (porém espalhadas), há de se ter um caminho para juntá-las e, respeitadas as devidas proporções e diferenças, através de um aporte científico e tecnológico como resposta prática, porém utilizando não desprezando o conhecimento dos curadores digitais envolvidos e suas formas de trabalho e organização dos acervos.

## **1.6 Hipótese**

Apesar de não ser comum em pesquisas qualitativas a utilização de hipóteses, pretende-se, aqui, com intuito de responder à questão de partida, definir uma hipótese *pós-factum* que serviu de guia de investigação, para ser discutida e respondida direta e indiretamente ao decorrer desta tese.

Conforme Marconi e Lakatos (Marconi & Lakatos, 2003), uma das funções destas hipóteses em uma pesquisa científica é coordenar fatos já conhecidos a partir da ordenação de materiais acumulados pela observação e, sem hipóteses, corre-se o risco de gerar um “amontoado de informações estéreis” (Marconi & Lakatos, 2003, p. 131).

Thiolent (2011) afirma que a concepção das hipóteses não deve ser confundida com elaboração de testes de hipóteses, que é uma técnica estatística de aplicação restritiva (Thiolent, 2011).

O autor defende que, ao contrário da concepção tradicional de pesquisa, na pesquisa-ação (uma descrição mais aprofundada pode ser encontrada no CAPÍTULO V) a natureza argumentativa - e deliberativa - dos procedimentos está explicitamente reconhecida pelos pesquisadores, não sendo necessário tanto valor aos critérios lógicos, formais e estatísticos da pesquisa.

Para isso abre-se uma hipótese, “um modelo semântico permite o intercâmbio de informações de patrimônio cultural entre os acervos digitais”.

### **1.7 Objetivos**

Dadas a questão de partida e a hipótese declarada anteriormente, o objetivo geral desta pesquisa é desenvolver um modelo de sistema de informação semântico, aberto e conectado para acervos digitais de patrimônio cultural, cujo foco é o patrimônio cultural do Estado de Goiás.

Tal objetivo se desdobra nos seguintes objetivos específicos,

1. Realizar um estudo sobre as características dos espaços de memória e patrimônio ligados à Secretaria de Estado da Cultura do Estado de Goiás, sua utilização quanto aos sistemas de informação.
2. Realizar um estudo sobre a aplicabilidade dos modelos conceituais de representação da informação de patrimônio cultural e tecnologias da Web-Semântica.
3. Construir um modelo semântico para organização da informação de patrimônio cultural, a partir de padrões técnicos, que permita a interoperabilidade entre repositórios de patrimônio cultural, aqui denominados espaços de memória.
4. Verificar se a complexidade do modelo semântico é um empecilho para os curadores de acervos digitais de patrimônio cultural.

Pretende-se, como contributo principal desta pesquisa, fornecer um modelo semântico, à luz da iniciativa dos dados abertos conectados, que possa ser utilizado para a criação de centros de memória (arquivos, museus, bibliotecas, galerias e outros espaços de memória), os chamados GLAMs.

É pretendido, ainda, como contributos complementares, um estudo qualitativo sobre as tecnologias semânticas de armazenamento, organização e recuperação da informação de patrimônio cultural e a experimentação de protocolos e modelos conceituais desenvolvidos pela comunidade científica.

### **1.8 Objeto de estudo e limites da pesquisa**

O objeto de estudo dessa investigação é composto pelos espaços de memória da Secretaria de Estado da Cultura do Estado de Goiás (SECULT-GO) e suas relações com

os sistemas de informação de patrimônio cultural numa perspectiva de interoperabilidade semântica.

Como limites dessa investigação, temos o recorte temporal, cujos levantamentos de dados ocorreram entre 2017 e o início de 2021.

Ao se observar o problema da interoperabilidade, também sob a ótica das múltiplas perspectivas e da colaboração de grupos e atores no processo de organização e exposição da informação, é possível perceber uma possibilidade de reprodução de um modelo excludente.

De forma muito particular, vive-se em um tempo no qual a inclusão e exclusão da informação andam juntas, a dicotomia entre esses dois fenômenos é um ponto de atenção, pois a desigualdade digital pode influenciar os acervos. A exclusão digital, conforme os autores (Carvalho Junior *et al.*, 2019) diz respeito às extensas camadas das sociedades que ficaram à margem da Sociedade da Informação e da expansão das redes digitais. Portanto, apesar da oportunidade de contribuir para a preservação de uma memória para além do institucional e mesmo reconhecendo a fundamental importância desta participação, esta tese não alcança esse modelo colaborativo.

Reconhecendo as limitações e o imenso trabalho desta investigação, há de se ressaltar a importância de sistemas de informação capazes de atuar nessas redes constituídas, ou que ainda estão por se constituir, e permitir, nos artefatos e nos seus contextos, possibilidades diversas de acesso e também de preservação da memória.

## **1.9 Estrutura da Tese**

A presente tese está estruturada em 9 capítulos, incluindo este capítulo introdutório.

No CAPÍTULO II é realizado o enquadramento teórico no qual são abordados conceitos fundamentais, linguagens e os modelos de referência CIDOC-CRM e EDM.

No CAPÍTULO III apresenta-se, de forma complementar, uma revisão sistemática da literatura, a qual tem como intuito abordar outros trabalhos no estado da arte na temática desta investigação.

No CAPÍTULO IV é fornecida uma proposta de modelo genérico para o problema abordado nesta tese.

No CAPÍTULO V apresenta-se a metodologia científica adotada na presente

tese, na qual se descreve o aparato metodológico utilizado para a condução da pesquisa geradora deste trabalho, bem como outras informações acerca das estratégias e autores de em que se obteve a reflexão dos métodos científicos utilizados.

No CAPÍTULO VI são apresentados o resultado do inquérito realizado com Curadores de espaços de memória e a modelagem do protótipo de sistema de informação.

No CAPÍTULO VII são apresentados arquitetura e desenvolvimento do protótipo Guará, um Sistema de Informação semântico para o patrimônio cultural.

No CAPÍTULO VIII são encontradas discussões em torno dos resultados apresentados.

E, por fim, no CAPÍTULO IX são apresentadas as conclusões, trabalhos futuros e considerações finais.

## **CAPÍTULO II. PRESSUPOSTOS, LINGUAGENS E ESTRATÉGIAS DA CIÊNCIA DA INFORMAÇÃO**

### **2.1. Introdução**

Neste capítulo serão apresentados o enquadramento teórico, os pressupostos e a descrição das características dos elementos fundamentais constantes nesta tese. Serão abordadas as tecnologias digitais e questões relacionadas ao desafio da interoperabilidade semântica, bem como as estratégias e ferramentas utilizadas como possibilidade de construção de respostas para o problema de pesquisa.

De acordo com o que foi previamente apresentado, pretende-se apoiar esta tese nas tecnologias da Web-Semântica e outras estratégias de organização e recuperação da informação com base na Ciência da Informação, bem como nos Dados Abertos Conectados.

Capurro & Hjrland (2003), afirmam que ao invés de definir termos há maior importância em saber o quão útil o termo irá servir para seu propósito. Isso por que quase todas ciências utilizam o termo informação em um contexto específico. Assim, não se pretende aqui definir ou restaurar conceitos, mas apresentar a base de premissas para situar essa tese na Ciência da Informação, em especial no que parte dos conceitos fundamentais da CI até as ferramentas da Web Semântica.

A Web Semântica é modelada primariamente em três linguagens: RDF (*Resource Description Framework*), RDF-S (*Resource Description Framework Schema*) e OWL (*Web Ontology Language*). Essas linguagens fornecem uma base para os dados e também para recursos mais sofisticados nas camadas mais altas da Web Semântica que permitem, em especial, abordar o problema da interoperabilidade semântica em vários níveis.

Já a Web dos Dados, Dados Abertos Conectados, ou simplesmente Dados Conectados (*Linked Data*), é um conjunto de iniciativas resultantes da evolução da aplicação de ferramentas e linguagens da Web Semântica. “A origem da Web Semântica vem da busca pela externalização do conhecimento e apesar de uma longa história de avanços, a busca permanece” (Hebeler *et al.*, 2009, p. 23).

As tecnologias da Web Semântica, aliadas às iniciativas de repositórios digitais, fornecem um conjunto de possibilidades para os centros de memória e outras ações de

preservação do patrimônio cultural.

## **2.2 Pressupostos da Ciência da Informação: A informação digital e o patrimônio cultural**

Conforme Capurro & Hjarland (2003), é importante considerar a informação como uma força constitutiva da sociedade e, dessa forma, reconhecer a natureza tecnológica dos sistemas e serviços de informação (Capurro & Hjarland, 2003). Assim, parte-se do pressuposto da Ciência da Informação como ciência interdisciplinar.

Parte-se da interdisciplinaridade, que conforme Borko (1986), é por onde se pretende investigar as propriedades, o comportamento e as forças que governam os fluxos e os usos da informação, bem como suas técnicas, desde aquelas mecanizadas, para processar armazenar, recuperar e disseminar as informações (Borko, 1986).

O processo interdisciplinar (e também, multi e transdisciplinar), segundo Silva (2005) inclui a origem, disseminação, coleta, organização, armazenamento, recuperação, interpretação e uso da informação. Para o autor, “O campo deriva ou relaciona-se com a matemática, a lógica, a linguística, a psicologia, a tecnologia computacional, as artes gráficas, as comunicações, a biblioteconomia, a gestão e alguns outros campos” (A. M. da Silva, 2005, p. 22).

Em um contexto mais objetivo, parte-se também da premissa que a informação pode ser produto (Floridi, 2002) e coisa (Buckland, 1991) e que aos sistemas de informação cabem vários serviços, como organização, tratamento, armazenamento, recuperação. Para Buckland (1991), qualquer que seja o sistema de armazenamento ou recuperação da informação utiliza a informação enquanto coisa para que possa haver processamento e assim criar uma base para diversas atividades informacionais.

Desde os primeiros sistemas de organização analógicos, os fenômenos informacionais puderam ser observados como pontos fundamentais para geração e qualificação da informação. A possibilidade de construção e a visão de uma biblioteca universal de Paul Otlet a Vannevar Bush, ou mesmo de uma rede de ligações de documentos (de Berners-Lee), experimentadas nesses tempos de Sociedade em Rede, passa por vários fenômenos informacionais abordados pela Ciência da Informação.

Os fenômenos informacionais são estudados por diversos teóricos em áreas que Araújo (2014) classifica como: a) os fluxos da informação científica; b) a representação e a recuperação da informação; c) os estudos de usuário; d) a gestão do conhecimento; e)

a economia política da informação e f) os estudos métricos da informação (Araújo, 2014).

Em especial, os fenômenos relacionados à representação e a recuperação da informação, os estudos de usuário e a gestão do conhecimento têm recebido profunda atenção devido à explosão digital experimentada.

Para Saracevic (1995), a explosão da informação é um problema social que começou na ciência e se espalhou por todos os empreendimentos humanos, de forma que exige esforços e recursos maciços, uma vez que é notável a importância estratégica da informação, primeiro para o trabalho e progresso da ciência, e agora para tudo o mais na sociedade humana moderna em níveis locais e mundiais (Saracevic, 1995).

Tal explosão dos dispositivos digitais e da disponibilização massiva de informação em rede passou a exigir soluções mais sofisticadas, ágeis e específicas no trato de documentos digitais. Muitas iniciativas foram tomadas, mas ainda se verifica um desequilíbrio nas articulações teóricas e experimentos (Moura, 2009).

Moura (2009) afirma que:

A Ciência da Informação é uma área de conhecimento que privilegia o estudo dos fenômenos informacionais em seus diversos aspectos, buscando compreender e acompanhar os seus desdobramentos para uma futura disponibilização em sistemas de informação (Moura, 2009, p. 59).

Isto é, Moura (2009) reconhece a importância de compreender os vários fenômenos para o ciclo de vida dos sistemas de informação. As técnicas, estratégias, produtos, ferramentas e outros artefatos informacionais são criticamente importantes para as relações humanas com a informação, principalmente em uma sociedade em rede que tem nas tecnologias digitais suas principais linguagens e outros canais de comunicação.

Essa evolução tem exigido que teóricos da organização da informação desprendam um grande esforço intelectual em direção às teorias relacionadas aos significados, no sentido de ampliar a base para compreensão e fortalecimento da representação temática em sistemas de informação (Moura, 2011).

Capurro & Hjarland (2003) afirmam que os seres humanos são conhecedores e observadores da realidade, e portanto, processadores da informação que são usadas para propósitos racionais (Capurro & Hjarland, 2003).

A informação, portanto, é compreendida como um signo que se atualiza, de

maneira recursiva, na interface com o sujeito (Moura, 2009). O conhecimento, que também é produto da informação, passa a ser visto também como linguagem.

Já para Alvarenga (2003), a percepção, classificação e criação de conhecimentos constituem-se em prerrogativas essenciais da racionalidade humana, identificando-se nesse trajeto processos cognitivos básicos, tais como o levantamento das características e a comparação entre estas e as identificadas no que já se conhece, caracterizando-se essas ocorrências como processos de classificação/cognição (Alvarenga, 2003).

Nesse sentido, segundo Alvarenga (2003 p . 22):

No processo de tratamento ou processamento dos registros de conhecimento para fins de armazenagem nos sistemas de informação, é requerido um novo estágio de representação, desta vez partindo-se não do ser ontológico em si, mas do conhecimento sobre o ser, expresso em documentos

Carvalho Junior *et al.* (2019), analisam as iniciativas que parecem tender a um formato global e colaborativo e modificar os rumos da Ciência da Informação no qual afirmam que “percebe-se que a ênfase do profissional da Ciência da Informação se desloca das preocupações de gestão dos acervos e usuários locais para a responsabilidade compartilhada por coleções mais expansivas” (Carvalho Junior *et al.*, 2019, p. 15).

Na última década, mudanças tecnológicas decorrentes da digitalização em massa e a visível substituição de dados analógicos por objetos nato-digitais modificaram sensivelmente a produção de informação.

Isso leva à necessidade dos novos modelos teóricos, que Moura (Moura, 2009, p. 60) chama de “base para o aperfeiçoamento da fluidez informacional”. Tudo isso por meio de técnicas e tecnologias que visam construir uma interação entre usuário e informação, considerando a estrutura digital e seus impactos. Houve ainda uma modificação profunda no formato das informações, nos métodos de aquisição e construção dos modelos informacionais. Iniciativas colaborativas como a *Wikipedia* se mostraram potencialmente globais e com alta capacidade de geração de informação qualificada.

A *Wikipedia*, que é desenvolvida em um modelo *crowdsourcing*, está ativa em 282 idiomas, conta com dezenas de milhares de editores e dezenas de milhões de artigos, conforme explicita Carvalho Junior *et al.* (2019):

[...] está em direta sintonia com a de bibliotecas, arquivos e museus. Oportunidades de colaboração do movimento com as instituições de memória oferecem uma gama de resultados que podem beneficiar bibliotecas acadêmicas e de pesquisa, incluindo *crowdsourcing*, envolvimento da comunidade, modelos rápidos para publicação on-line, designações de escrita de alunos definidas na Wikipedia e envolvimento com comunidades de aprendizagem (Carvalho Junior *et al.*, 2019, p. 14).

O pressuposto inicial é que a Web é composta de repositórios digitais, sendo esta, por si só, um grande repositório. Qualquer fragmento de informação que está na Web disponível e de alguma forma acessível pode ser visto como recurso.

Disponibilizar a informação de forma que esta seja recursivamente aproveitada para geração de novas informações e, conseqüentemente, para geração de conhecimento para o proveito humano, é um dos desafios atuais da Ciência da Informação.

No domínio do patrimônio cultural é possível notar na Web um aumento importante na quantidade de repositórios digitais. São inúmeras experiências de museus, bibliotecas, arquivos, galerias e outros espaços detentores de informações de patrimônio material e imaterial que utilizam, cada vez mais, as tecnologias para se integrarem à Sociedade em Rede.

Em uma geração frenética por conhecer tudo sobre tudo, a busca pelo conhecimento passa por processos de criação e aquisição críticos que resultam em grandes fluxos de informação, e estes, de forma contraditória, aumentam o nível de dificuldade de compreensão da informação à medida que essa informação é adquirida e recursivamente ampliada.

Pensar a informação na essência desta e confrontá-la com uma rede de possibilidades de relacionamentos, abre espaço para o surgimento de sistemas inovadores, sobretudo àqueles que exploram as capacidades semânticas.

É o desejo humano é recuperar mais e mais informação onde quer esta esteja, mas para recuperá-la, precisa-se gerar mais e mais informação, sendo que essas informações exigem uma rede caótica de associações que, para nossas necessidades, precisam ser cada vez mais úteis, rápidas e assertivas.

Peter Burke (Burke, 2012, p. 141) afirma que a “forma mais óbvia e generalizada de recuperar informação é a memória humana”, entretanto, esta é limitada e os sistemas podem, de forma artificial, imitá-la para auxiliar nesse processo.

Le Coadic (1994), afirma que informação é um conhecimento inscrito sob a forma escrita, oral ou audiovisual. Esse conceito é um pressuposto nesta tese, no que diz respeito à informação que, claramente, reside em uma linha muito tênue e ao mesmo tempo circular na tríplice dado-informação-conhecimento.

Observa-se ainda, que os fragmentos de informação, com o advento das redes de sistemas colaborativos digitais, possuem capacidade de gerar novas informações quando conectados uns aos outros através de sistemas organizados para este fim.

A utilização das tecnologias digitais, seja na educação, no entretenimento, nas relações com o trabalho ou com a família, passa a um nível de necessidade básica e deixa o excluído digital à margem daquilo que Castells (2003) nomeia de o “tecido de novas vidas” (Castells, 2016, p. 7).

Essa busca incessante pela informação é reforçada por Le Coadic (1994) ao destacar que

Seja pelo simples prazer de conhecer (Freud), de estar informado sobre os acontecimentos políticos, os progressos da ciência e tecnologia, ou pelo prazer menos simples de estar a par dos últimos temas e resultados das pesquisas, de acompanhar a vanguarda do científico, o objetivo da informação permanece sendo a apreensão de sentidos ou seres em sua significação, ou seja, continua sendo o conhecimento; e o meio é a transmissão do suporte, da estrutura. (LE COADIC, 1994, p. 5).

Uma sensível mudança de paradigma de consumo da informação, no entanto, é experimentada. Tecnologias como as que permitem a informação impressa são, a cada dia, substituídas por um aparato visual totalmente associado à informação digital que muda por completo a experiência da interação com a informação.

Os hipertextos foram e continuam a ser revolucionários no sentido de permitir saltos que atendem os desejos e a organização mental de quem navega pelas páginas da Web. No entanto, em um universo crescente de informação, os serviços de ligações por *hiperlinks* encontram cada vez mais limitações.

É comum que páginas estáticas na Web percam valor informacional, também é possível que sejam esquecidas com o passar do tempo e, assim apresentarem dificuldades de serem recuperadas se não estiverem estritamente atualizadas com as linguagens específicas. Esse conjunto de informação legada aqui é chamado de dados esparsos.

Os dados esparsos, sejam em forma de metadados, planilhas ou qualquer outro artefato descritivo de objetos, nato-digitais ou não, tendem a se dissolver com o passar do tempo, uma vez que na Web dos Dados não há espaço para recursos desconectados.

### **2.3 Uma nova rede? A Web, o patrimônio cultural e o problema da interoperabilidade**

A possibilidade de que agentes automatizados possam compreender e auxiliar nos processos relacionados à informação do armazenamento até a recuperação pressupõe uma nova rede? Castells (2016) afirma que a tecnologia não determina a sociedade. “A tecnologia é a sociedade e a sociedade não pode ser entendida ou representada sem suas ferramentas tecnológicas” (Castells, 2016, p. 64) .

Baseando-se nesta e em outras afirmações de Castells, de forma antecipada e contrariando algumas tendências, é difícil apontar a Web Semântica ou mesmo a dita Web do Dados como nova rede. Como o próprio Berners-Lee afirma, estas são extensões da Web a partir de linguagens e padrões nem tão novos assim. A Web Semântica data de 1999, Os Dados Abertos Conectados datam de 2006.

A despeito da *Internet*, enquanto fenômeno tecnológico e histórico ter modificado sensivelmente as nossas relações econômico-sociais (inclusive pessoais), esta trouxe consigo desafios anteriores ao surgimento da Web. Integração e interoperação entre sistemas são exemplos desses desafios. Em outras palavras, antes de existir a Internet a interoperabilidade já era um problema a ser enfrentado.

Zeng (2019) apresenta questões cruciais relacionadas à integração entre sistemas para o mundo corporativo conforme a afirmação a abaixo.

Em um cenário muito mais amplo, os sistemas que fornecem ou dão suporte ao gerenciamento de dados e informações foram criados em todos os lugares. Eles são construídos com base nas necessidades predominantes de um domínio, organização ou aplicativo, incorporando diferentes contextos, propósitos e decisões de escopo por diferentes patrocinadores institucionais. A integração se tornou um estilo de vida para muitas organizações, e a interoperação de sistemas entre departamentos e organizações tornou-se essencial (Zeng, 2019, p. 121).

Na Sociedade em Rede é preciso se conectar; integrar sistemas passa a ser um desafio para essa conectividade. Há uma tendência na comunidade científica a apontar que a integração entre sistemas de informação passa por um processo de

interoperabilidade que exige uma *língua comum* entre os diversos sistemas e possui níveis distintos sintáticos, semânticos e lógicos para se comunicarem colaborativa e efetivamente.

Backhouse, Hsu e McDonnell afirmam que a interoperabilidade significa aplicações poderem trocar dados e serviços com consistência e meio efetivo, cobrindo diferentes plataformas de *hardware* e *software*. Esse processo deve levar em consideração vários fatores que tendem a envolver garantias em autenticação, integridade, confidencialidade e disponibilidade (Backhouse *et al.*, 2003).

Percebe-se, portanto, que o termo interoperabilidade é aplicado de diversas maneiras tanto no campo das ciências da computação e ciências da informação quanto em outras áreas (Filgueiras *et al.*, 2013). E que, de maneira genérica, o termo se refere à capacidade de troca de elementos em sistemas autônomos ou não.

Dito isso, é importante ressaltar um tipo específico de interoperabilidade que se refere aos problemas relacionados à semântica, portanto, em um nível mais alto de complexidade do que simplesmente fazer com que dois ou mais sistemas consigam compreender as mensagens *transitadas* pelo meio.

Esse tipo de interoperabilidade, conforme Boeuf *et al.* (Boeuf *et al.*, 2018) é, objetivamente, a capacidade de diferentes sistemas de informação se comunicarem de forma consistente e com um significado acordado que abrange 1) os elementos estruturados dos dados envolvidos; 2) a terminologia que acompanha os dados; e 3) os identificadores usados nos dados para os elementos factuais (locais, pessoas, objetos, entre outros).

Em Riley (2017) é possível ler que a interoperabilidade é a habilidade de múltiplos sistemas de informação com diferentes plataformas de *hardware* e *software*, estrutura de dados e interfaces para trocar dados com o mínimo de perda de conteúdo e funcionalidade (Riley, 2017).

Há ainda a se reiterar que a interoperabilidade possui em si níveis e formas variadas, e que, para tanto, precisa de múltiplas estratégias que envolvem o armazenamento, a busca e a representação da informação. Está o problema, portanto, ancorado no campo de estudos da representação e recuperação da informação (Araújo, 2009).

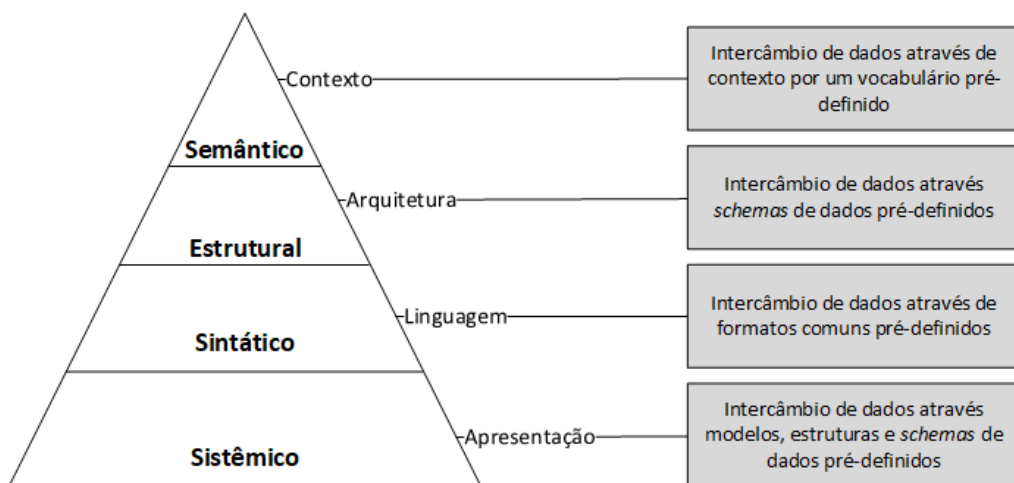
No contexto desta tese, em uma perspectiva semântica, um dos desafios é a busca

por possíveis soluções para o problema da interoperabilidade entre sistemas de informação relacionados ao patrimônio cultural. Dessa forma, pressupõe-se um problema da Web Semântica, uma vez que o interesse dos sistemas de patrimônio cultural é utilizar a *Rede* como base fundamental e plataforma colaborativa.

Zeng (2019) aponta quatro níveis de interoperabilidade, sendo que o primeiro, na base da pirâmide, exige uma habilidade específica para apresentação dos dados e está relacionada aos padrões já estabelecidos na Web, passando no segundo nível, por um desafio sintático através das linguagens da Web Semântica. Padrões e modelos de referência, como o CIDOC-CRM, estão em uma terceira camada estrutural relacionada à arquitetura de dados. E por fim, a quarta camada está relacionada ao contexto com vocabulários predefinidos em Sistemas de Organização do Conhecimento.

Na Figura 1, é possível visualizar os níveis de interoperabilidade semântica cuja complexidade aumenta de acordo com a subida da pirâmide.

**Figura 1 – Níveis de Interoperabilidade**



Fonte: Elaborada pelo autor

Geralmente, a interoperabilidade sintática é diferenciada da interoperabilidade semântica, conforme Doerr *et al.* (2019):

Interoperabilidade sintática significa que a codificação da informação dos sistemas envolvidos e dos protocolos de acesso são compatíveis, de modo que a informação pode ser processada conforme descrito acima sem erros. No entanto, isso não significa que cada sistema processa os dados de maneira consistente com o significado pretendido. Por exemplo, um sistema pode usar uma tabela

chamada “Ator” e outra chamada “Agente”. Com a interoperabilidade sintática, os dados de ambas as tabelas só podem ser recuperados como distintos, embora possam ter exatamente o mesmo significado. Para superar essa situação, a interoperabilidade semântica deve ser adicionada (Doerr *et al.*, 2020, p. 14) .

Especificamente para o patrimônio cultural, Carvalho Junior *et al.* (2019) afirmam que está surgindo um novo ambiente informacional.

Com a digitalização de seus acervos, as vantagens destas instituições em promover a interoperabilidade entre suas bases e a necessidade logística de compartilhamento de infraestrutura de *software* e *hardware* vêm promovendo a emergência de infraestruturas que permitem a operação em ambientes digitais contíguos e até permeáveis — claramente um novo ambiente informacional (Carvalho Junior *et al.*, 2019, p. 13).

Isto é, esse fenômeno, como uma via de mão dupla, se apresenta como um desafio e também como uma grande oportunidade conexão de sistemas de informação de patrimônio cultural, dado o crescente processo de digitalização de acervos e o surgimento frenético de objetos nato-digitais relacionados à uma nova cultura em expansão, a Cibercultura (Lèvy, 2014).

A interoperabilidade relacionada aos repositórios digitais de patrimônio cultural é um desses fenômenos, nos quais há múltiplas conexões e necessidade de integração entre repositórios digitais de várias tecnologias, muitas vezes distintas e que atendem a linguagens diversificadas.

Um objeto relativo à cultura de um povo possui potencial possibilidade de se conectar com muitos outros objetos. Por exemplo, uma receita de um prato tradicional, uma poesia ou uma música podem estar relacionadas a eventos históricos, a pessoas e a locais distintos que influenciou outros movimentos culturais, mas também podem ser vistos como artefatos em algum repositório digital na Web.

Não é difícil encontrar documentos em arquivos ou museus que estão relacionados a outros documentos externos e fatos históricos, conhecidos ou não do grande público.

A grande quantidade de documentos e dados disponíveis na *Internet* requer um esforço tecnológico para que, para além das pessoas, as informações sejam “inteligíveis” para as máquinas.

Abordar a interoperabilidade é levar em consideração que as informações transitadas entre os diversos repositórios possam ser consumidas e agreguem um maior poder informacional através de processamentos entre sistemas de acordo com algum contexto de utilização.

Tomem-se, como exemplo, museus da imagem e do som que possuam algum objeto relacionado à música popular brasileira. Questões relativas ao período de composição, aos instrumentos utilizados ou mesmo ao autor são fundamentais para a geração de conhecimento. Essas informações, entretanto, não precisam estar fisicamente no mesmo repositório especializado em imagem e som. Essa necessidade de conectividade informacional e colaboração entre os repositórios, sendo possível, tende a potencializar e enriquecer a informação para diversas formas de utilização (Filgueiras *et al.*, 2020).

Padilha & Café (2016), ao afirmarem que a interoperabilidade semântica pode ser vista como a capacidade que os sistemas de informação possuem de se integrarem por meio de padrões e tecnologias semânticas, reforçam que no contexto dos repositórios digitais de patrimônio cultural, pressupõem-se o emprego de padrões e protocolos específicos de museus, arquivos, galerias de arte e outras instituições de preservação do patrimônio cultural material e imaterial (Padilha & Café, 2016). E este cenário específico dos padrões e protocolos de museus, bibliotecas, arquivos e galerias, segundo Doerr (2009), apresenta uma proliferação de terminologias especializadas que nem sempre são utilizadas como meio de concordância entre especialistas.

Os padrões e protocolos, segundo Doerr (2009), são usados como ferramenta intelectual para construção de hipóteses ou descrever um fenômeno. Os assuntos, as terminologias e as estratégias de tantos sistemas especializados envolvem um grande e complexo problema de integração de esquemas para associar várias informações complementares (Doerr, 2009).

Reforça, portanto, que para que haja colaboração entre os repositórios, de forma que os dados possam ser conectáveis, abertos e, conseqüentemente, contribuam para a produção do conhecimento, para além da interligação é necessário haver um processo de interoperação semântica que leve em consideração as características fundamentais de cada objeto digital.

Assim, de forma objetiva, se faz necessário um conjunto de tecnologias possíveis

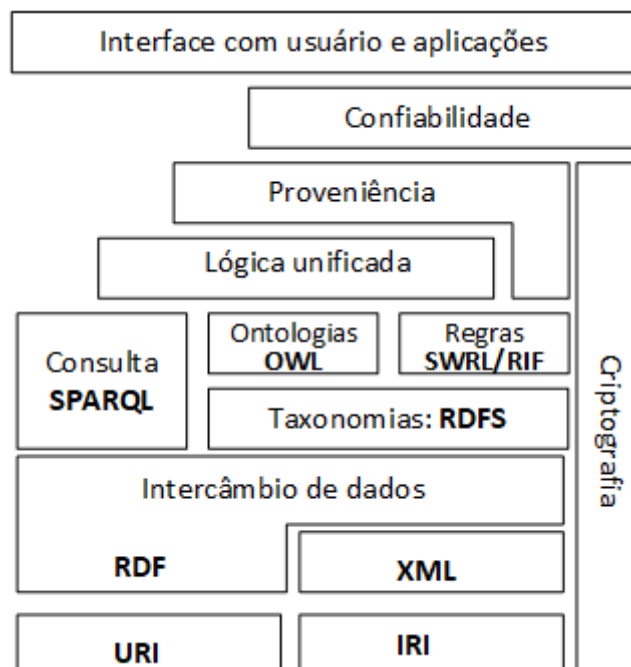
de tratar a semântica dos objetos envolvidos de forma computacional, extrapolando modelos de dados, inclusive já consolidados, que não permitem que agentes não humanos consigam compreender significados e tomar decisões.

Hebeler *et al.* (2009) reforça a necessidade dos significados para uso mais eficaz dos dados.

Uma definição para a Web Semântica começa com a definição da semântica. Semântica significa simplesmente significado. O significado permite um uso mais eficaz dos dados subjacentes. O significado geralmente está ausente da maioria das fontes de informação, exigindo que os usuários ou instruções de programação complexas o forneçam. Por exemplo, as páginas da web são preenchidas com informações e tags associadas (Hebeler *et al.*, 2009, p. 26).

Na Figura 2, é possível visualizar a complexidade das tecnologias envolvidas na Web Semântica de coisas que já existem e são utilizadas como URL e XML que estão na base da Web, e também como coisas que ainda estão por vir, como uma camada específica para “confiabilidade” e “proveniência”. O esquema é também conhecido como Bolo de noiva da Web (*Semantic Web Layer Cake*).

**Figura 2 – Esquema de Camadas da Web Semântica**



Fonte: Adaptado de (Hebeler *et al.*, 2009)

## 2.4 Linguagens e Ferramentas da Web

Pesquisadores, especialmente em Humanidades Digitais, estão começando a

usar novos conjuntos de dados para agregar informações sobre a história: coleções de dados, informações e conhecimentos que são dedicados à preservação do legado de cultura tangível e intangível herdada de gerações anteriores (Calvanese *et al.*, 2016).

Ao reforçar o texto introdutório deste capítulo, e como observado na Figura 2, por apresentar um esquema com um conjunto de camadas previstas na Web Semântica, serão apresentadas as principais tecnologias que a compõem.

#### **i) RDF - Resource Description Framework**

O *Resource Description Framework*, ou *framework* de descrição de recursos, é a base para as tecnologias semânticas da Web. Ao mesmo tempo é um modelo e uma linguagem para representar informações sobre recursos na Internet, e, conforme Curé e Blin (2015), evoluiu de um modelo de metadados para o principal componente da Web Semântica e Web dos Dados.

O primeiro objetivo do *Resource Description Framework* (RDF) era fornecer um modelo de dados de metadados para a web. Hoje em dia, surge como um modelo de dados para a Web de Dados e a Web Semântica proporcionando uma organização lógica definida em termos de algumas estruturas de dados para suportar a representação, acesso, restrições e relacionamentos de objetos de interesse em um determinado domínio de aplicação (Curé & Blin, 2015, p. 43).

Nas definições da W3C, trata-se de um protocolo a nível de metadados que tem como intenção “descrever” os recursos (ou objetos) na Web, que são, na prática, as URI’s (W3C, 2004).

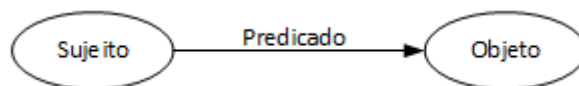
Um URI (*Uniform Resource Identifier*), que é uma das tecnologias base da Web, funciona como um identificador (endereço) único e inequívoco de um recurso que também pode ser encontrado sobre forma de IRI que, na prática, é a internacionalização dos URIs. Isto é, permite em sua estrutura caracteres de outras línguas, conforme a W3C:

IRI (*Internationalized Resource Identifier*) é uma cadeia de caracteres em formato *unicode* conforme definições em um padrão específico cuja sintaxe é definida na RFC3987. Em RDF deve ser absoluta e pode ter um identificador de fragmento (Berners-lee, 2006 s.p.).

Ainda, segundo a W3C, o RDF pode ser visto também como um arcabouço que permite fazer determinadas afirmações sobre os recursos na Web e, por apresentar-se ao mesmo tempo como linguagem e modelo de dados, proporciona uma representação

simples das *coisas*, de forma que, para além das características estruturais, as relações possam ser facilmente representadas. Uma representação básica pode ser vista na Figura 3.

**Figura 3 – Representação básica de uma tripla RDF**



Fonte: elaborada pelo Autor

Allemang, Hendler e Gandon (2020) afirmam que o RDF fornece um mecanismo que permite qualquer pessoa fazer uma declaração básica sobre qualquer coisa e coloque essas declarações em um único modelo (Allemang *et al.*, 2020).

Tais afirmações ou representações de cada recurso são modeladas em triplas, obviamente constantes dos três elementos básicos, os quais sejam: a) sujeito, b) predicado; e c) objeto. Cada tripla representa um fato sobre uma coisa que está sendo descrita – ou seja, o assunto –, em uma propriedade específica – o predicado – e com um determinado valor – o objeto (Curé & Blin, 2015).

Usando este modelo simples, o RDF permite que dados estruturados e semiestruturados sejam misturados, expostos e compartilhados entre diferentes aplicativos (W3C, 2014a).

Os URIs (ou IRIs) são a parte mais importante do RDF uma vez que estes podem representar qualquer coisa, sendo um recurso acessível através da web ou não. Essas triplas RDF (sujeito, predicado e objeto) definem um fato sobre um recurso, é justamente o conjunto dessas triplas que tem a capacidade de carregar consigo possibilidades de representação mais profunda sobre algum domínio.

Em um outro olhar, o sujeito pode ser visto como o recurso, o predicado como a propriedade e o objeto, um outro recurso ou uma informação textual simples, como um número, uma data, um nome, ou qualquer outra informação textual.

Pelo formato de representação não hierárquica torna-se fácil (e comum) modelar as informações como um grafo. Por ser uma afirmação/declaração (ou sentença) RDF, uma relação entre dois recursos por meio de uma propriedade, pode-se inferir uma estrutura em grafo cujos recursos são vértices e as propriedades são arestas que ligam tais

recursos.

Um grafo RDF, que é um conjunto de triplas RDF, cujos nós podem ser de três tipos: IRIs, literais ou nós em branco, em que um nó constante de IRI ou literal represente alguma coisa em um universo do discurso. Essas coisas são chamadas de recursos e podem ser vistas, de forma geral, como entidades ou valores relacionadas a entidades.

Qualquer coisa pode ser um recurso – coisas físicas, documentos, conceitos abstratos, informações textuais, e outros. Os recursos com IRI são referenciáveis, enquanto os literais possuem tipos de dados que os definem. Nós em branco, diferentemente daqueles que possuem literais ou IRIs referenciadas, não identificam qualquer recurso específico, mas identificam relações sem nomeá-las especificamente (W3C, 2014a).

O modelo RDF é atemporal. Isto é, um grafo RDF é uma captura instantânea e estática das informações. Este pode, no entanto, expressar informações sobre eventos ou sobre aspectos temporais, dado um vocabulário apropriado. Adicionar, modificar ou remover triplas em um RDF resulta em um novo grafo (W3C, 2014a).

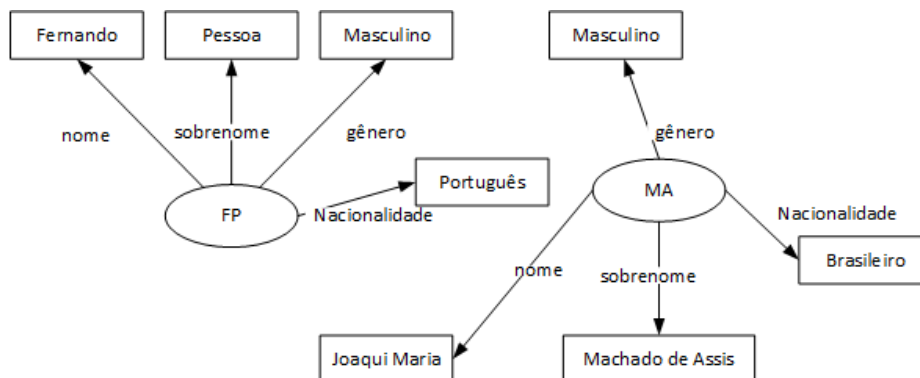
Os predicados também possuem IRIs, e são propriedades, e, de maneira simplificada, também podem ser pensados como um recurso com uma relação binária. Isto é, uma relação entre dois outros recursos. Essa é uma das características que dão aos grafos RDFs possibilidades poderosas de produzirem semântica.

Os recursos com IRIs são também chamados de referentes (*referent*). Um vocabulário RDF é um conjunto de URIs que se pretende usar em um grafo RDF. Em W3C (2014) é possível encontrar uma tripla RDF a codificar uma declaração (sentença), que é uma expressão lógica simples, em forma de conjunção (*And*) de uma afirmação sobre um universo do discurso. Há três tipos de relacionamentos entre os grafos RDF: implicação, equivalência e inconsistência.

Quanto à implicação, A implica B quando todos os arranjos que fazem A verdadeiro também fazem B verdadeiro. Quando A implica em B, se a verdade de A é presumida ou demonstrada, a verdade de B é estabelecida. Já a equivalência ocorre conforme o seguinte: todos grafos RDF são equivalentes quando fazem a mesma afirmação sobre o mundo. A e B são equivalentes somente se A implica em B e B implica em A. A inconsistência, por outro lado, diz respeito a uma contradição interna. Quando não existe um arranjo possível do mundo que torne a expressão verdadeira (W3C, 2014a).

A representação textual dos RDFs é também uma tarefa importante para o modelo, a recomendação do W3C para representação textual é usar uma serialização XML de RDF, denominada RDF / XML, mas várias outras serializações são encontradas, com o Turtle, N3 e JSON-LD.

**Figura 4 – Exemplo de grafo RDF**



Fonte: Elaborada pelo autor

A Figura 4 apresenta um exemplo de grafo RDF em que dois recursos (FP e MA) possuem declarações constantes de propriedades e valores. O quadro abaixo apresenta um algoritmo com o mesmo grafo serializado no formato RDF / XML.

**Serialização 1 – Exemplo de serialização RDF**

```

<?xml version="1.0"?>
<rdf:RDF xmlns="http://exemplo.org/exemplo/poetas#"
  xml:base="http://exemplo.org/exemplo/poetas"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:xml="http://www.w3.org/XML/1998/namespace"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
  <owl:Ontology rdf:about="http://exemplo.org/exemplo/poetas"/>
  <owl:DatatypeProperty
    rdf:about="http://exemplo.org/exemplo/poetas#gênero"/>
  <owl:DatatypeProperty
    rdf:about="http://exemplo.org/exemplo/poetas#nacionalidade"/>
  <owl:DatatypeProperty
    rdf:about="http://exemplo.org/exemplo/poetas#nome"/>
  <owl:DatatypeProperty
    rdf:about="http://exemplo.org/exemplo/poetas#sobrenome"/>
  <owl:Class rdf:about="http://exemplo.org/exemplo/poetas#Pessoa"/>
  <owl:NamedIndividual rdf:about="http://exemplo.org/exemplo/poetas#FP">
    <rdf:type rdf:resource="http://exemplo.org/exemplo/poetas#Pessoa"/>
    <gênero>Masculino</gênero>
    <nacionalidade>Português</nacionalidade>
    <nome>Fernando Antonio</nome>
    <sobrenome>Nogueira Pessoa</sobrenome>
  </owl:NamedIndividual>
  <owl:NamedIndividual rdf:about="http://exemplo.org/exemplo/poetas#MD">
    <rdf:type rdf:resource="http://exemplo.org/exemplo/poetas#Pessoa"/>
    <gênero>Masculino</gênero>
    <nacionalidade>Brasileiro</nacionalidade>
    <nome>Joaquim Maria</nome>
    <sobrenome>Machado de Assis</sobrenome>
  </owl:NamedIndividual>
</rdf:RDF>
  
```

```
</owl:NamedIndividual>  
</rdf:RDF>
```

Fonte: Elaborado pelo autor

## ii) RDF-S – Resource Description Framework Schema

A RDF-S é uma extensão semântica do RDF e apresenta-se como um vocabulário capaz de classificar os elementos do RDF, fornecendo uma capacidade de modelar os dados RDF (Berners-Lee, 2006; Samreen *et al.*, 2013; Xie & Matusiak, 2016).

Trata-se de uma linguagem com expressividade para descrever as noções básicas de *comunalidade* e *variabilidade* familiares de linguagens de objeto e outros sistemas de classes, ou seja, classes, subclasses e propriedades (Allemang *et al.*, 2020).

Através da aplicação da RDF-S é possível identificar se os elementos são recursos ou propriedade de recursos e também uma classificação hierarquizada dos elementos, restrição de domínio e alcance, permitindo que sejam possíveis determinadas inferências.

Curé e Blin (2015) afirmam que a RDF-S fornece um mecanismo para descrever melhor os tipos correspondentes em RDF (Curé & Blin, 2015, p. 61):

Mais precisamente, permite descrever grupos de recursos relacionados e suas relações. Os primeiros são chamados de conceitos ou classes, enquanto os últimos são chamados de propriedades. Todas essas descrições são fornecidas em formato RDF usando um vocabulário definido no *namespace* <http://www.w3.org/2000/01/rdf-schema#>, comumente associado ao prefixo *rdfs*.

Em linhas gerais, a RDF-S apresenta classes e propriedades. Uma classe corresponde a um grupo de recursos, que por si só já é um recurso, e, portanto, deve ser identificado por um IRI. Uma propriedade é uma relação entre um recurso – portanto, uma classe – e um objeto, que pode ser um outro recurso, um valor literal.

Um exemplo disso é que a definição de classe recebe a diretiva *rdfs:Class*, em que cada instância dessa classe deve ser definida com a propriedade *rdf:type* apontada para o elemento *rdfs:Class* correspondente.

Duas noções interessantes apresentadas na RDF-S são domínio (*rdfs:Domain*) e intervalo (*rdfs:Range*). A primeira é utilizada para afirmar que qualquer recurso a ter

uma determinada propriedade é uma instância de uma ou mais classes. Já a última, para afirmar que quaisquer valores de uma propriedade, são instâncias de uma ou mais classe.

Entretanto, em (W3C, 2014b), é possível encontrar que a RDF-S não busca enumerar todas as formas possíveis de representar o significado das classes e propriedades RDF. Em vez disso, sua estratégia é reconhecer que existem muitas técnicas por meio das quais a semântica das classes e propriedades pode ser descrita.

No quadro abaixo é possível visualizar as principais classes do RDF/RDF-S.

**Quadro 1 – Classes RDF/RDF-S**

Classe	Descrição
<i>rdfs:Resource</i>	A classe recurso. Simboliza uma coisa.
<i>rdfs:Literal</i>	A classe dos valores literais ( <i>string</i> , inteiro, etc)
<i>rdf:langString</i>	A classe de valores de literais de <i>string</i> com <i>tag</i> idioma.
<i>rdf:HTML</i>	A classe dos valores literais de HTML
<i>rdf:XMLLiteral</i>	A classe dos valores literais de XML
<i>rdfs:Class</i>	A classe das classes
<i>rdf:Property</i>	A classe das propriedades do RDF.
<i>rdfs:Datatype</i>	A classe dos tipos de dados ( <i>datatypes</i> ) RDF
<i>rdf:Statement</i>	A classe das sentenças/declarações RDF
<i>rdf:Bag</i>	A classe dos contêineres não ordenados de RDF
<i>rdf:Seq</i>	A classe dos contêineres ordenados de RDF
<i>rdf:Alt</i>	A classe de contêineres de alternativas de RDF
<i>rdfs:Container</i>	A classe de Contêineres RDF
<i>rdfs:ContainerMembershipProperty</i>	A classe de propriedades de associação do contêiner, <i>rdf: _1</i> , <i>rdf: _2</i> , ..., todos os quais são subpropriedades de 'membro'.
<i>rdf&gt;List</i>	A classe de listas RDF.

Fonte: (W3C, 2014b)

As propriedades definidas no RDF-S, em combinação com as classes, permitem que qualquer objeto possa ser classificado e obtenha significado através de relações. A tripla *sujeito – predicado – objeto* só é possível utilizando ao menos uma das propriedades a qual se define domínio e alcance. O quadro abaixo apresenta as principais propriedades do RDF/RDF-S.

**Quadro 2 – Propriedades RDF/RDF-S**

Propriedade	Descrição	Domínio	Alcance
<i>rdf:type</i>	O sujeito é uma instância de uma classe	<i>rdfs:Resource</i>	<i>rdfs:Class</i>

<i>rdfs:subClassOf</i>	O sujeito é uma subclasse de uma classe	<i>rdfs:Class</i>	<i>rdfs:Class</i>
<i>rdfs:subPropertyOf</i>	O sujeito é uma subpropriedade de uma propriedade.	<i>rdf:Property</i>	<i>rdf:Property</i>
<i>rdfs:domain</i>	Um domínio de uma propriedade do sujeito.	<i>rdf:Property</i>	<i>rdfs:Class</i>
<i>rdfs:range</i>	Um alcance de uma propriedade de um sujeito.	<i>rdf:Property</i>	<i>rdfs:Class</i>
<i>rdfs:label</i>	Um nome legível (humanamente) para um recurso.	<i>rdfs:Resource</i>	<i>rdfs:Literal</i>
<i>rdfs:comment</i>	Uma descrição para um recurso.	<i>rdfs:Resource</i>	<i>rdfs:Literal</i>
<i>rdfs:member</i>	Um membro do recurso apresentado.	<i>rdfs:Resource</i>	<i>rdfs:Resource</i>
<i>rdf:first</i>	O primeiro item em uma lista de recursos RDF	<i>rdf:List</i>	<i>rdfs:Resource</i>
<i>rdf:rest</i>	Os demais itens da lista após o primeiro item.	<i>rdf:List</i>	<i>rdf:List</i>
<i>rdfs:seeAlso</i>	Mais informações sobre o recurso	<i>rdfs:Resource</i>	<i>rdfs:Resource</i>
<i>rdfs:isDefinedBy</i>	Uma definição par ao recurso	<i>rdfs:Resource</i>	<i>rdfs:Resource</i>
<i>rdf:value</i>	Propriedades idiomáticas para valores estruturados	<i>rdfs:Resource</i>	<i>rdfs:Resource</i>
<i>rdf:subject</i>	O sujeito de uma declaração RDF	<i>rdf:Statement</i>	<i>rdfs:Resource</i>
<i>rdf:predicate</i>	O predicado de um recurso numa declaração RDF	<i>rdf:Statement</i>	<i>rdfs:Resource</i>
<i>rdf:object</i>	O objeto de uma declaração RDF.	<i>rdf:Statement</i>	<i>rdfs:Resource</i>

Fonte: (W3C, 2014b)

### iii) Ontologias

O termo ontologia é utilizado com diferentes significados e em diferentes áreas do conhecimento (Guarino *et al.*, 2009) . Tem origem na filosofia aristotélica como um ramo que visa estudar o ser e sua composição. Passou a ser amplamente utilizado nas comunidades científicas de computação e gestão da informação, em especial, para modelagens de sistemas de organização do conhecimento.

Este trabalho aborda as ontologias em um espectro computacional, portanto, mais recente e menos estabelecido tradicionalmente como o filosófico, que parte da visão da organização de conceitos através de classificações e relacionamentos de forma explícita e pragmática. Em um universo do discurso, aquilo que pode ser representado em um artefato computacional é um “objeto especial de informação”, segundo (Guarino *et al.*, 2009, p. 2)

Ontologias computacionais são um meio de modelar formalmente a estrutura de um sistema, isto é, as entidades e relações relevantes que emergem de sua

observação e que são úteis para nossos propósitos. [...] A espinha dorsal de uma ontologia consiste em uma hierarquia de conceitos de generalização / especialização, isto é, uma taxonomia. Supondo que estamos interessados em aspectos relacionados a recursos humanos, Pessoa, Gerente e Pesquisador podem ser conceitos relevantes, em que o primeiro é um superconceito dos dois últimos. Cooperar com e pode ser considerada uma relação relevante entre as pessoas. Uma pessoa concreta que trabalha em uma empresa seria então uma instância do seu conceito correspondente (Guarino *et al.*, 2009, p. 2).

As ontologias são essenciais para a Web-Semântica. Elas visam compartilhar entendimentos comuns entre pessoas e sistemas, permitir o reuso de conhecimento sobre determinado domínio, explicitar determinadas suposições sobre o domínio, separar o conhecimento do domínio do conhecimento operacional, e realizar análises a respeito deste conhecimento (Noy & McGuinness, 2001).

As ontologias são geralmente especificadas em lógica descritiva que pertence a uma família de linguagens de representação de conhecimento por serem equipadas com uma semântica formal baseada na lógica de primeira ordem; e, por ser formal, portanto, é computável. Isso permite humanos e sistemas de computador utilizá-las de forma não ambígua, e também torna possível usar a dedução lógica para inferir informações adicionais a partir dos fatos declarados de maneira explícita (Calvanese *et al.*, 2016).

Kiryakov (2006) afirma que sistemas de representação formal do conhecimento, tratam da construção de modelos do mundo, de um problema ou domínio em particular, a permitir raciocínio e interpretação automáticos, isto é, por meio de computação (Kiryakov, 2006).

Em termos de organização da informação, uma ontologia pode ser vista como uma especificação formal de conceitos em torno de algo. Define um vocabulário comum para compartilhamento de informações em um determinado domínio. Isso inclui definições de conceitos básicos para interpretação de máquinas neste domínio e as relações entre esses conceitos (Noy & McGuinness, 2001).

Diferente do sentido filosófico da ontologia de Aristóteles, em que, independente da linguagem, haverá um sistema único de categorizar determinada visão do mundo no sentido computacional, trata-se de um artefato de *software* para descrever uma determinada realidade a partir de um vocabulário, e, que possui um conjunto de suposições explícitas no sentido de aumentar os significados, isto é, a semântica dos

objetos envolvidos (Guarino, 1998). Em outras palavras, vocabulários comuns e formalizados que buscam cobrir domínios específicos e compartilhados, que torna possível especificar os termos e relações computáveis.

Gruber (1993) definiu uma ontologia como uma “especificação de uma conceitualização que é representada como uma estrutura  $\langle D, R \rangle$  na qual  $D$  é um domínio e  $R$  é um conjunto de relações relevantes em  $D$ ” (Gruber, 1993 s.p.).

Guarino (1998) traz à reflexão essa definição, afirmando que, para que esta tenha sentido, é necessário introduzir uma abordagem diferente na noção de conceitualização. Isto é, levando em conta as relações entre a ontologia, seus modelos e sua conceitualização (Guarino, 1998).

Este autor afirma ainda que uma maneira de representar intenções – e portanto relações conceituais – e enxergá-las em forma de “mundos possíveis” nos conjuntos. Enquanto relações comuns são definidas em um certo domínio, relações conceituais são definidas em um espaço de domínio.

Assim, um espaço de domínio pode ser representado com uma estrutura  $\langle D, W \rangle$ , na qual  $D$  representa o domínio e  $W$  o conjunto máximo de conceitos existentes em  $D$ . Uma conceitualização é representada pela expressão  $C = \langle D, W, R \rangle$ , na qual  $R$  é um conjunto de relações conceituais no espaço de domínio  $\langle D, W \rangle$ .

Pode-se dizer, portanto, que uma conceitualização é um conjunto de relações conceituais definidas num espaço de domínio (Guarino, 1998). Outros autores apresentam definições distintas, ora complementares ou até mesmo antagônicas em algum sentido estrito.

Martini *et al.* (2016) definem uma ontologia como:

[...] um conjunto 4-tupla  $O = (C, I, PdtO, A)$ , em que  $C$  é um conjunto de conceitos (classes ou entidades) do domínio;  $I$  é um conjunto de instâncias (indivíduos) de uma classe;  $PdtO$  é um conjunto de propriedades que indivíduos e classes podem ter:  $Pdt$  é uma propriedade de tipo de dados, ou seja, é a relação entre classes ou instâncias de classes e valores atômicos, como sequências de caracteres, números inteiros, e outros;  $PO$  é uma propriedade de objeto, ou seja, é uma relação entre classes ou instâncias de classes e outras classes ou instâncias; e  $A$  é um conjunto de axiomas (Martini *et al.*, 2016, p. 237).

Kiryakov (2006) também define uma ontologia como um conjunto de 4 tuplas,

$O = \{C,R,I,A\}$ , em que a ontologia é representada pelo conjunto de classes **C** que representam os conceitos. **R** o conjunto de propriedades que representam os relacionamentos. **I** o conjunto de indivíduos ou instâncias, e **A**, o conjunto de axiomas (Kiryakov, 2006).

Em relação ao conjunto de instâncias, a autora sustenta de forma complementar que, dependendo do nível da ontologia, este conjunto tende a ser vazio, como é possível apreciar no seguinte trecho:

Normalmente, ontologias projetadas para servir como esquemas não contêm definições de instância, mas não há restrição formal nessa direção. Traçar a fronteira entre a ontologia (isto é, a parte conceitual e consensual do conhecimento) e o restante dos dados, representados na mesma linguagem formal, nem sempre é uma tarefa trivial. Por exemplo, poderia haver uma ontologia sobre o turismo, que define as classes Localização e Hotel, bem como a relação *localizadaEm* entre elas e a categoria de atributo do hotel. As definições das classes, relações e atributos devem claramente fazer parte da ontologia. As informações sobre um hotel em particular provavelmente não fazem parte da ontologia, na medida em que não são uma questão de conceitualização e consenso, mas são apenas uma descrição, criada para algum propósito (potencialmente específico). Então, suponha que exista uma definição de Nova York como uma instância da classe *City* - pode-se argumentar que faz parte da ontologia ou apenas uma descrição de uma cidade. O fato de ser uma instância não determina necessariamente que ela não é parte da conceitualização (Kiryakov, 2006 p. 119).

Para Noy & McGuinness (2001) uma ontologia, juntamente com um conjunto de instâncias individuais de classes, constitui uma base de conhecimento. Reforçando que há uma linha tênue na qual a ontologia termina e a base de conhecimento começa (Noy & McGuinness, 2001).

As definições formais de ontologias geralmente compartilham características, independentemente da abordagem. No geral o uso de conceitos (classes ou entidades) é agrupar indivíduos que compartilham propriedades, e o uso de instâncias é agrupá-las em classes; o uso de propriedades do objeto ou propriedades do tipo de dados é para definir relações e características individuais (Martini *et al.*, 2016) .

Kiryakov (2006) reforça que é comum classificar as ontologias de acordo com o grau de generalidade destas no que se refere à conceitualização, cobertura e finalidade,

classificando as ontologias em três níveis:

1. As ontologias de topo (nível superior) representam um modelo geral do mundo, aquele adequado para uma grande variedade de tarefas, domínios e áreas de aplicação;
2. Ontologias de domínio representam uma conceituação de um domínio específico, por exemplo, na construção de estradas ou para a medicina;
3. As ontologias de aplicativos e tarefas são adequadas para intervalos específicos de aplicativos e tarefas (Kiryakov, 2006, p. 116) .

Dadas as formalizações das ontologias, de forma pragmática, é possível sintetizar que uma ontologia é composta por um conjunto de conceitos que são, de alguma maneira hierarquizados, ou seja, uma taxinomia das classes envolvidas. Tais conceitos – ou classes – podem ser abstratos, compostos, concretos, reais ou fictícios. Essa taxonomia pretende classificar esses conceitos de forma análoga às ciências biológicas.

Há ainda um conjunto de relacionamentos entre esses conceitos. Por exemplo, um objeto pertencer à uma pessoa, sugere o relacionamento *pertenceA* que liga a classe Pessoa à classe Objeto. Tais relacionamentos dependem de um conjunto de axiomas, regras que são sempre verdade. Estas são responsáveis pelas afirmações sobre as classes e suas relações com outras classes entre si, por exemplo, toda pessoa tem uma mãe.

Um conjunto de instâncias é o composto dos indivíduos que fazem parte da ontologia e que atendem estritamente as regras impostas pelos axiomas e as relações estabelecidas para as classes. Por exemplo, um indivíduo da classe Pessoa que tem por nome “João” e que tem uma mãe, que também é um indivíduo da classe Pessoa e que tem por nome “Maria”.

Kiryakov (2006) apresenta três reflexões sobre a representação ontológica e de integração de dados:

1. Semântica: se a semântica (o significado) dos dados é formalmente representada, para que uma máquina possa interpretá-lo formalmente, raciocinar e obter novos dados? Essa qualificação é diretamente relevante para o gerenciamento de raciocínio e ontologia - o raciocínio só pode ser executado com base em dados "semânticos". Dados não-semânticos podem ser adaptados para raciocínio por meio do mapeamento para uma ontologia, que é um esquema semântico que define o significado dos dados externamente. Existem casos marginais em que a especificação de uma estrutura contém elementos semânticos, por exemplo, o caso de esquemas XML. Permanecemos com uma definição relativamente estreita do que são semânticas e consideramos dados

semânticos somente quando existe alguma teoria lógica que defina o significado associado à linguagem de representação usada para representar ou interpretar os dados; 2. Estrutura: se os dados são formalmente estruturados, para que uma máquina possa interpretar e gerenciar formalmente sua estrutura? Essa distinção é importante porque as abordagens para acesso e gerenciamento automatizados (e seu desempenho típico) diferem consideravelmente entre dados estruturados e não estruturados; 3. Esquema: aqui consideramos dados esquemáticos, que determinam a estrutura e / ou a semântica de outros dados. Obviamente, existem dados esquemáticos e não esquemáticos. A qualidade do esquema é determinada pela função (pretendida) dos dados em relação a outros dados. (Kiryakov, 2006 p. 120)

Os esquemas são importantes para mediação e evolução das ontologias uma vez que determinam a consistência e a interpretação dos dados. Por exemplo, uma alteração em uma ontologia pode tornar um conjunto de dados anteriormente compatível com a versão antiga incompleto com a nova (ou vice-versa). Em muitos casos, o problema da integração de dados pode ser resolvido no nível da integração do esquema (Kiryakov, 2006).

#### **iv) OWL - *Web Ontology Language***

Como visto na Seção anterior, as ontologias podem ser vistas como vocabulários formalizados de termos que cobre um domínio específico. Estes vocabulários são compartilhados por uma comunidade de usuários que permitem especificar os termos e também os relacionamentos entre estes termos. Uma vez que os vocabulários capturam o conhecimento sobre um determinado domínio e permitem um recorte de mundo, precisam também de linguagens específicas que permitam sua expressividade.

A OWL (*Ontology Web Language*), dentro dessa perspectiva de domínio, coisas e relacionamentos, é uma linguagem para definir e instanciar ontologias na Web. Permite declarações de classes e seus relacionamentos, além de suas propriedades, regras e demais restrições.

Por ter sido projetada para o uso de aplicações que precisam expressar mais do que a simples informação para o usuário, a OWL leva como carga o conteúdo e fornece um vocabulário adicional permitindo uma semântica formal. Trata-se de uma linguagem baseada em lógica descritiva que utiliza um mecanismo de inferência para verificação de consistência e para calcular automaticamente a hierarquia das classes envolvidas, podendo ser vista como uma extensão das linguagens RDF e RDF-S.

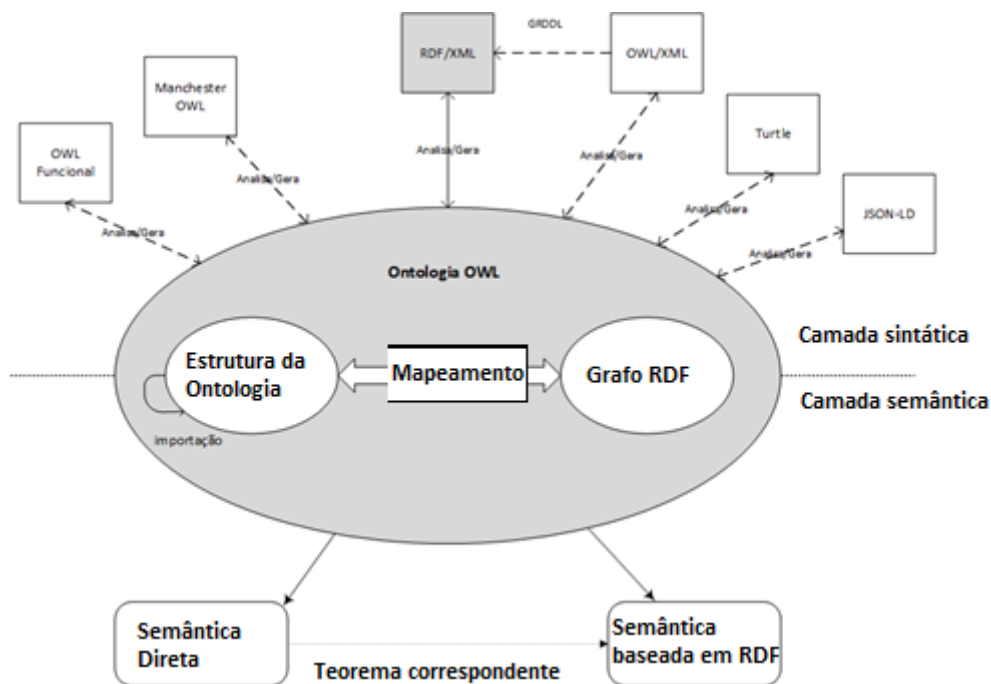
Horrigde *et al.* (2007) afirma que as ontologias OWL podem ser categorizadas em três espécies ou sublinguagens: OWL-Lite, OWL-DL e OWL-Full, de acordo com sua expressividade:

Uma característica definidora de cada sublinguagem é sua expressividade. OWL-Lite é a sublinguagem menos expressiva. OWL-Full é a sublinguagem mais expressiva. A expressividade de OWL-DL fica entre a de OWL-Lite e OWL-Full. O OWL-DL pode ser considerado uma extensão do OWL-Lite e OWL-Full uma extensão do OWL-DL (Horrigde *et al.*, 2007, p. 10).

OWL-Lite é uma sublinguagem simples para definir hierarquias entre classes e restrições simples. OWL-DL é mais expressiva por permitir adicionar restrições com base em lógica descritiva (DL), permitindo um raciocínio automatizado. Já OWL-Full possui uma altíssima expressividade e deve ser usada em casos nos quais a expressividade é mais importante que completez computacional e decidibilidade (Horrigde *et al.*, 2007).

Uma ontologia OWL pode ser vista como um grafo RDF. A estrutura conceitual da linguagem OWL está definida em um documento de especificação da W3C que, de forma geral, apresenta os papéis e funcionalidades dos elementos estruturais além de definições de sintaxe.

**Figura 5 – Estrutura básica de uma ontologia OWL**



Fonte: adaptado de (W3C, 2014a)

Como pode ser vista na Figura 5, uma ontologia OWL é composta por uma estrutura básica que pode utilizar outras ontologias a partir de um processo de importação. Essa estrutura é mapeada em um grafo RDF e vice-versa. A camada semântica da ontologia OWL permite realizar inferências a partir de lógica descritiva que são definidas através de axiomas e outras regras restritivas. A camada sintática simboliza o processo de análise (*parsing*) e geração (*produce*) da ontologia OWL em linguagens específicas de serialização como XML, Turtle, N3 e JSON-LD.

#### v) A linguagem SPARQL

Para que os grafos RDF possam ser recuperados através de consultas, se faz necessário que haja uma linguagem específica de recuperação. A SPARQL acompanha o modelo RDF para garantir que haja consultas (*queries*) em conjuntos dados RDF, e que apesar de ter objetivo similar à SQL (*Structured Query Language*), apresenta diferenças importantes em relação a esta última, conforme Curé e Blin (2015):

A peculiaridade do SPARQL é lidar com grafos RDF e suportar uma abordagem baseada em navegação para recuperação de dados. A ideia é começar a partir de um determinado nó e passar de um nó a outro seguindo certas arestas até que o nó objetivo seja alcançado. Essa abordagem é bastante diferente da execução SQL, que se baseia em junções entre tabelas. Para realizar tais operações, as variáveis devem ser suportadas nos padrões de triplas RDF permitidos na linguagem de consulta, em particular na cláusula WHERE (Curé & Blin, 2015, p. 53).

Em SPARQL, uma variável é identificada por um ponto de interrogação seguido pelo nome que deseja dar à variável. Uma variável irá, por um lado, nos permitir armazenar qualquer informação resultante do problema de combinação de grafos e, por outro lado, poderá ser combinada posteriormente em outras triplas. O Algoritmo 1 apresenta um exemplo simples de uma consulta SPARQL.

#### Algoritmo 1 – Exemplo de Consulta SPARQL

```
1 prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
2 prefix owl: <http://www.w3.org/2002/07/owl#>
3 prefix ex: <http://exemplo.org/exemplo/poetas#>
4 SELECT ?sujeito ?predicado ?objeto
5 WHERE {
6   ?sujeito ex:nome ?objeto.
7 }
8 LIMIT 10
```

Fonte: Elaborado pelo autor

Em linhas gerais o comando apresentado diz: “retorne nas variáveis *sujeito*,

*predicado e objeto* todos os nós do RDF que tenham a propriedade *nome*. O resultado será um grafo com um conjunto de triplas com o endereço do recurso, a propriedade nome, e valor de nome presente no RDF que atendam essa condição, limitado a 10 registros.

#### **vi) A Web e os Dados Abertos Conectados (*Linked Open Data*)**

Outro pressuposto importante desta tese, é que vivencia-se um processo de transformação da informação que requer que as fontes de informação sejam, cada vez mais, abertas do ponto de vista mais estrito. Para além de estarem abertos e acessíveis, esse pressuposto é que os dados sejam conectáveis.

Dados conectados dizem respeito a um método de publicação de dados estruturados que podem ser interligados e serem úteis através de consultas semânticas, baseadas em HTTP, RDF e URIs. (Mountantonakis & Tzitzikas, 2019)

Nos anos 90, Tim Berners-Lee, o inventor da *World Wide Web*, discutiu a visão de uma Web Semântica, “O primeiro passo é colocar dados na Web de uma forma que as máquinas possam naturalmente entender ou convertê-las” (Berners-Lee, 1998 s.p.).

Em 2001, Tim Berners-Lee e sua equipe descreveram as principais ideias de uma nova Web representando dados no formato RDF e usando ontologias que permitiam a criação de regras de inferência, entre outras. Posteriormente Berners-Lee (2006) destacou a necessidade emergente de integração de dados semânticos e apontou princípios para os dados conectados (*Linked Data*).

Segundo Berners-Lee (1998) a Web foi projetada para ser um espaço universal de informações e, portanto, perde o poder se houver certos tipos de coisas onde não sejam criados vínculos (Berners-Lee, 1998). Acontece que em meados dos anos 2000 a comunidade Web, sobretudo os pesquisadores envolvidos no seu desenvolvimento, percebeu que as ligações entre documentos (*links* e hipertexto) podem se apresentar muito subutilizadas enquanto possibilidade de produção do conhecimento.

Esse conhecimento, conforme destaca Le Coadic (1994), é representado por uma estrutura de conceitos e suas relações de determinados assuntos em determinados momentos (Le Coadic, 1994).

A proposta de Berners-Lee (1998) pretende fornecer uma solução para essa fragilidade semântica e desconexão dos dados na Web. A gênese da proposta trazia como pano de fundo, dentre outros objetivos, permitir que as pessoas pudessem trabalhar de

forma colaborativa e pudessem compartilhar conhecimento universalmente (Berners-Lee, 1998).

A dicotomia Web dos documentos (*hiperlinks*) versus Web dos dados (semântica) é abordada em seu clássico artigo *Linked Data*, que dentre outras afirmações afirma que

Como a Web de hipertexto, a Web de dados é construída com documentos na web. No entanto, ao contrário da web de hipertexto, onde os *links* são âncoras de relacionamentos em documentos de escritos em HTML, para os dados eles ligam coisas descritas pelo RDF (Berners-Lee, 2006 s.p.) .

Berners-Lee (2006) ainda fornece quatro expectativas para que a Web dos dados possa crescer. A primeira, uso de URIs para nomear coisas. A segunda, usar o protocolo HTTP para que pessoas possam procurar por esses nomes. A terceira, fornecer informações úteis usando padrões estabelecidos (RDF, SPARQL), quando alguém buscar um URI. E a quarta, incluir ligações para outros URIs para que eles possam descobrir mais coisas.

O crescimento dos dados conectados mostra como os princípios da Web Semântica auxiliam na expressão de dados e na aquisição de dados novos e relacionados, permitindo a criação de muito mais valor. Por exemplo, um aplicativo sobre futebol poderia explorar dados abertos da Web sobre estádios, localizações de jogos, jogadores de futebol famosos, jogos anteriores na competição ou sobre o clima. Adicioná-los ao aplicativo o tornará uma fonte de informações muito mais ricas (Berners-lee & Hara, 2013, p. 2).

Acervos culturais digitais podem se apresentar como grandes facilitadores de acesso sob várias perspectivas. Os repositórios dos ditos GLAMs (galerias, bibliotecas, arquivos e museus) oferecem, além das possibilidades de disponibilização de públicos distintos, uma enorme potencialidade de colaboração entre si para geração de conhecimento.

Nota-se aqui oportunidades de ferramentas para integração entre várias camadas dessas fontes de informação heterogênea que vão desde definições de vocabulário a serviços de trocas de informação qualificada para geração de sistemas de conhecimento.

Carvalho Junior *et al.* (2019) enfatizam que estamos diante de um novo ambiente informacional, uma vez que a digitalização de acervos das instituições culturais, aliada às

vantagens em se promover interoperabilidade entre suas bases, e a necessidade logística de compartilhar de infraestrutura de software e hardware, promovem uma emergência de infraestruturas que permitem a operação em ambientes digitais contíguos e até permeáveis (Carvalho Junior *et al.*, 2019).

Berners-Lee & O'Hara (2013) trazem uma reflexão sobre a conectividade e ubiquidade e apresentam como ponto forte o processo de transformação de identificadores de sistemas em URIs, defendendo que nem todos os sistemas sobreviveriam a essa mudança, mas aqueles que sobrevivem tendem a se tornar muito mais poderosos.

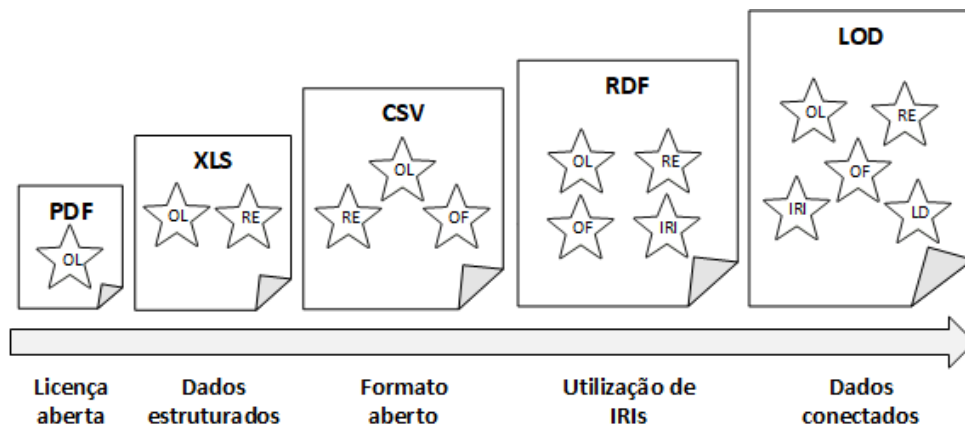
Alguns pensam que o futuro da mídia está nos aplicativos de telefone, mas um grande problema é que o conteúdo não está na Web. Há uma enorme diferença com um aplicativo da Web, que permite que os usuários compartilhem URIs para acessar diretamente as fontes de dados subjacentes, explorando o valor agregado de colocar conteúdo na Web. Não há nada errado com os aplicativos de telefone em si, mas nenhuma das informações em um aplicativo de telefone pode ser conectada. Eles não estão participando do novo mundo de pessoas que vinculam as coisas. Por exemplo, os mecanismos de pesquisa não encontrarão seu conteúdo. Existem diretrizes para a criação de aplicativos da Web, para habilitar sua função em todos os tipos de dispositivos, a partir da *MobileWeb Initiative* no W3C (Berners-lee & Hara, 2013, p. 2)

Os Dados Abertos Conectados (*LOD*), são portanto, uma mistura de Dados Conectados (*Linked Data*) e Dados Abertos (*Open Data*). Defende-se aqui, que para uma colaboração ampla entre os repositórios, que além de conectáveis, os dados sejam abertos. Iniciativas como *DBPedia*, *GeoNames*, e outros serviços, constituem o projeto *Linked Open Data (LOD)* de dados abertos conectados.

Berners-Lee (2006) propõe um espectro de desenvolvimento dos dados conectados em 5 estrelas: 1) disponibilizar com licença aberta na Web; 2) disponibilizar os dados em formato estruturado; 3) disponibilizar os dados em um formato aberto não proprietário; 4) utilizar URIs para identificar recursos; e 5) conectar com outros dados para prover contexto.

A esse nível de abertura, é possível encontrar em Berners-Lee (2012) são chamados de um esquema de representação de 5 estrelas. A Figura 6 apresenta uma visão gráfica dessa representação.

**Figura 6 – Modelo 5 estrelas dos Arquivos Abertos Conectados**



Fonte: adaptada de (Berners-Lee, 2012)

No modelo em cinco estrelas, uma estrela é dada a cada nível de abertura e conectividade do recurso. A primeira estrela é reservada a aplicações que permitam que seu recurso esteja disponível na Web com uma licença aberta (OL). A segunda estrela é reservada alcançada caso os dados estejam estruturados ao invés de imagens, por exemplo. Se estes dados estiverem em um formato com padrão não proprietário, como CSV (RE), então terá 3 estrelas. A quarta estrela é indicada àqueles recursos que forem identificados por um IRI (URI), e, portanto, possuem uma referência inequívoca que pode ser apontada ou encontrada por outro recurso. A quinta estrela é relacionada à conectividade deste recurso. Ou seja, para alcançá-la é preciso que o recurso esteja conectado a outro (LD).

## 2.5 Modelos de Referência

Os modelos de referência são fundamentais para a construção de Sistemas de Informação para o patrimônio cultural, uma vez que oferecem um arcabouço com classes e relacionamentos referentes aos diversos domínios da cultura. Serão explanados aqui dois modelos de referência estudados para este trabalho, o CIDOC-CRM e o EDM.

Inúmeros modelos e estratégias estão disponíveis à comunidade. Qualquer modelo semântico, provavelmente, produzirá resultados semelhantes no que diz respeito ao problema de pesquisa abordado. A escolha desses dois modelos está relacionada a 1) CIDOC-CRM e EDM são os modelos mais encontrados em pesquisas e publicações recentes; e 2) pretendem cobrir o patrimônio cultural como um todo, portanto, uma diversidade de espaços de memória;

### **i) Modelo de referência CIDOC-CRM**

O CIDOC-CRM é um modelo de dados da ICOM (*International Council of Museums*) que se apresenta como (ou pretende ser um) padrão para de integração de fontes de dados heterogêneas dentro do domínio do patrimônio cultural. Por ser um modelo conceitual composto de classes, propriedades, relações e regras, pode ser visto como uma ontologia de domínio.

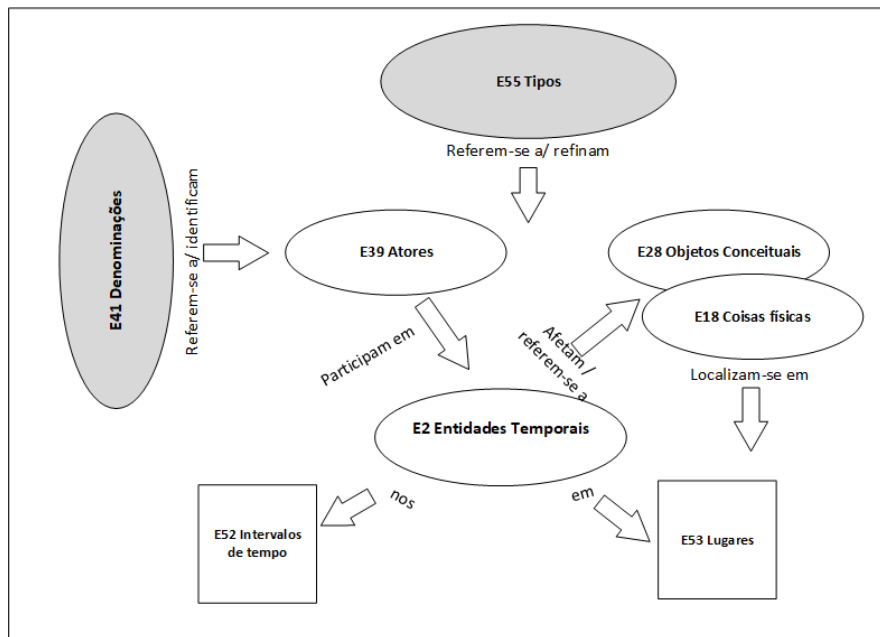
A ontologia CIDOC-CRM (ISO 21127:2014) é formal para integração e intercâmbio de informações heterogêneas de patrimônio cultural para redes de conhecimento (Doerr *et al.*, 2007). Um dos seus objetivos fundamentais é fornecer definições semânticas que visem clarificar informações locais para que estas sejam melhor aproveitadas de forma global, conforme Doerr *et al.* (2020):

A CIDOC-CRM se define e restringe-se à semântica subjacente dos esquemas de banco de dados e estruturas de documentos usados no patrimônio cultural e na documentação de museus em termos de uma ontologia formal. Esta não define nenhuma terminologia que aparece tipicamente como dados nas respectivas estruturas de dados; no entanto, prevê as relações características para seu uso. Não tem como objetivo propor o que as instituições culturais devem documentar. Em vez disso, explica a lógica do que eles atualmente documentam e, portanto, permite a interoperabilidade semântica (Doerr *et al.*, 2020, p. 2).

Doerr *et al.* (2020) afirmam que a função principal da CIDOC-CRM é apoiar a integração significativa de informações em um mundo aberto, o que significa que a própria CIDOC-CRM deve permanecer fundamentalmente aberta e as bases de conhecimento implementadas devem ser flexíveis o suficiente para receber o novo e acompanhar a evolução natural dos fenômenos sociais.

No domínio do patrimônio cultural a CIDOC-CRM se apresenta como o principal modelo utilizado, possuindo várias extensões para domínios complementares. Segundo Doerr (2009), na CIDOC-CRM são seguidos princípios estreitos para a integração global e restrições filosóficas sobre o tipo de discurso, sendo a ontologia descrita na linguagem TELOS, que é uma linguagem declarativa.

**Figura 7 – Conceitos fundamentais do CIDOC-CRM**



Fonte: Adaptado de (Doerr, 2009, p. 8)

Conforme apresentado na Figura 7, esta distingue classes individuais de propriedades, que são bidirecionais com rótulos distintos para cada direção permitindo herança múltipla tanto pra classes quanto para propriedades (Doerr, 2009).

Doerr *et al.* (2020) explica ainda que há uma centralidade nos eventos apoiando o discurso histórico. Assim, o CIDOC CRM permite, em primeiro lugar, a descrição de entidades que são, elas mesmas, processos ou evoluções com limite de tempo no decorrer do tempo usando Entidade *Temporal E2* e suas subclasses.

A função básica desta entidade é capturar o fato de algo ter acontecido ao longo do tempo. Além de permitir a descrição de uma duração temporal, as subclasses de Entidade *Temporal E2* são utilizadas para documentar as relações históricas entre objetos, semelhante ao papel dos verbos de ação em uma frase de linguagem natural.

Tudo isso relacionado ao que Doerr *et al.* (2020) afirma ser a realidade material:

Para os fins do CIDOC-CRM, a realidade material é considerada como tudo o que tem substância que pode ser percebida com os sentidos ou instrumentos. Exemplos são pessoas, uma floresta ou um ambiente de povoamento, mar, atmosfera, microestruturas celestiais ou celulares distantes, incluindo o que supomos que poderia ser potencial ou teoricamente percebido se pudéssemos estar lá, como o centro da Terra ou o Sol, e todos isso é passado. É restrito ao espaço e ao tempo. O que se passa em nossas mentes ou é produzido por nossas

mentos também é considerado parte da realidade material, pois se torna materialmente evidente para outras pessoas, pelo menos por meio de nossas declarações, comportamento e produtos (Doerr *et al.*, 2020, p. 12).

Para Santos (2016), o modelo CIDOC-CRM foi pensado como um guia para o processo, análise de requisitos e fase de modelagem conceitual para os sistemas de informação cultural, assim, fornece uma codificação das principais conceituações de domínio interdisciplinar da cultura, de forma que permita uma funcionalidade extensível.

As informações sobre o patrimônio cultural consistem em desafios específicos para o tratamento formal. Como constatamos, o modelo CIDOC-CRM é uma ontologia formal de referência que busca atender a uma crescente demanda por pesquisa orientada, estudos comparativos, transferência de dados e migração de dados entre fontes heterogêneas de conteúdos culturais (Santos, 2016, p. 57).

Para Carrasco *et al.* (2017), o modelo CIDOC-CRM é uma metodologia eficiente baseada para integrar a riqueza de informação cultural, tornando possível para os usuários a recuperação dos recursos de que necessitam a partir de diferentes coleções (Carrasco *et al.*, 2017) .

O modelo CIDOC-CRM pode ser visualizado, de uma maneira geral, através das funções suportadas pelos seus relacionamento que são: 1) identificação de itens do mundo real através de nomes do mundo real; 2) classificação de itens do mundo real; 3) decomposição parcial de coisas materiais e imateriais, entidades temporais, grupo de pessoas, lugares e tempos; 4) participação de itens persistentes em unidades temporais; 5) localização de entidades temporais no espaço-tempo e coisas materiais no espaço; 6) influência de objetos em atividades e produtos e vice-versa; e 7) referências de objetos de informação a qualquer item do mundo real (Boeuf *et al.*, 2018).

Como declarado no documento de referência da CIDOC-CRM, o escopo pretendido está no campo dos estudos dos museus mas também objetiva buscar uma harmonização do CRM com bibliotecas e arquivos, no entanto, exclui-se deste escopo os processos referentes a atividades administrativas.

[...] O CIDOC CRM destina-se especificamente a cobrir informações contextuais: os antecedentes históricos, geográficos e teóricos que dão às coleções de museus grande parte de seu significado e valor cultural. [...] O intercâmbio de informações relevantes com bibliotecas e arquivos e a harmonização do CRM do CIDOC com seus modelos se enquadram no escopo

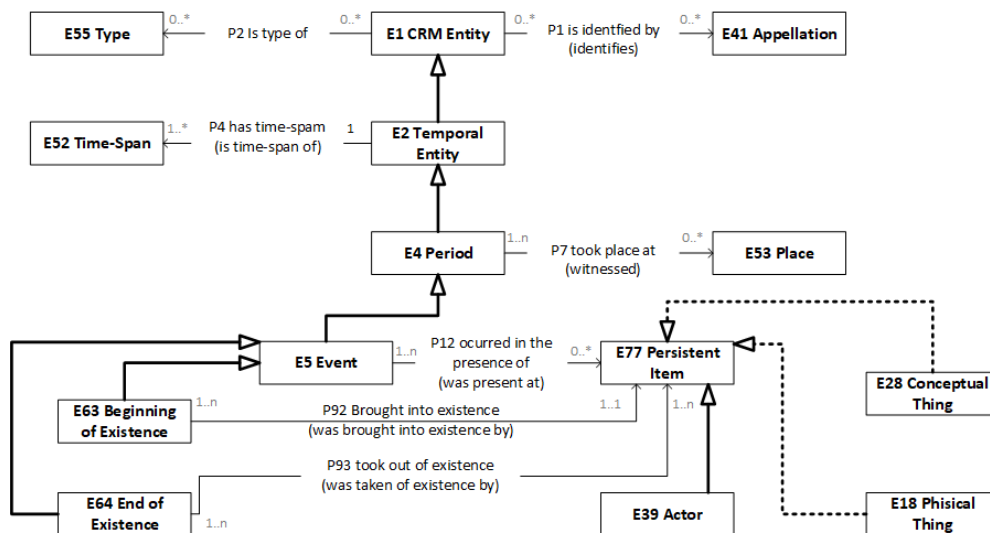
pretendido do CRM do CIDOC. [...] As informações necessárias apenas para a administração e gerenciamento de instituições culturais, como informações relacionadas a pessoal, contabilidade e estatísticas de visitantes, estão fora do escopo pretendido do CRM CIDOC (Doerr *et al.*, 2020, p. 3).

A construção dos termos do CIDOC-CRM segue uma terminologia relativamente conhecida para quem transita nas áreas das ciências da informação e computação, em especial quem conhece o modelo RDF e outros padrões de metadados.

Como uma ontologia é formal, é composta por elementos comuns de uma ontologia, tais quais classe, subclasse, propriedade, relações, regras e axiomas. Classe é definida em (Doerr *et al.*, 2020) como uma categoria de itens que compartilham uma ou mais características comuns, e que servem como critérios para identificar itens que pertencem à uma classe.

Uma *propriedade* define uma relação de um tipo específico entre duas classes, sendo caracterizada por uma intenção. Desempenha um papel análogo a um verbo gramatical, em que se devem definir seu domínio e alcance de forma análoga ao sujeito e objeto na gramática. Ao contrário de classes, as propriedades possuem dependência (Doerr *et al.*, 2020).

**Figura 8 – Classes e propriedades básicas do CIDOC-CRM**

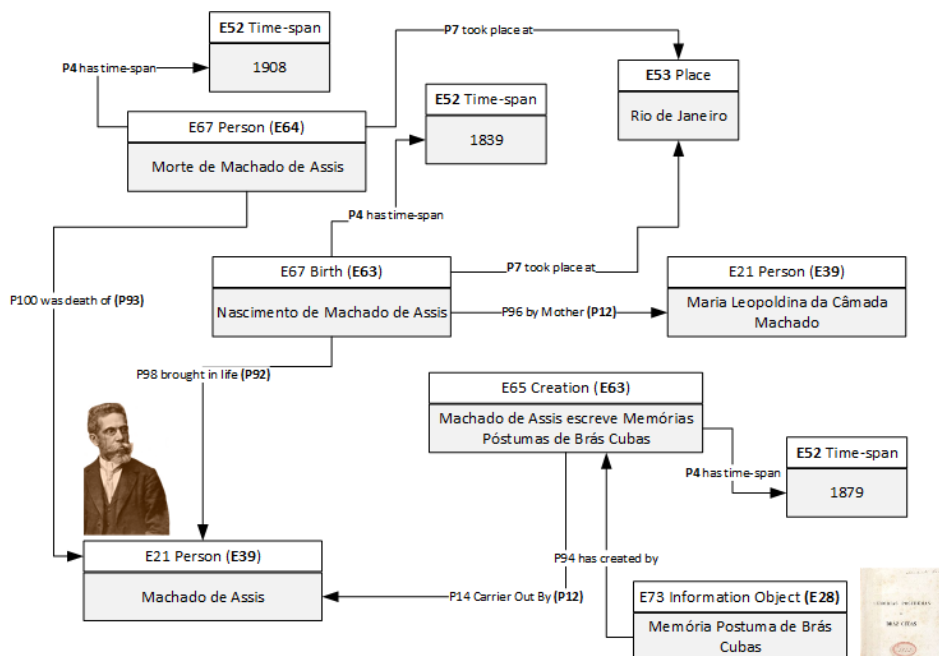


Fonte: adaptado de (Doerr *et al.*, 2020, p. 19)

A CIDOC-CRM é expressa em termos das primitivas de modelagem semântica de dados. Portanto, é formada por *Classes*, que representam noções gerais no domínio do discurso, – como a classe CRM E21 Pessoa que representa a noção de pessoa – e

*Propriedades*, que representam as relações binárias que ligam os indivíduos no domínio do discurso – como a propriedade CRM P152 *tem pai* ligando uma pessoa a um dos pais da pessoa.

**Figura 9 – Exemplo de fato em CIDOC-CRM**



Fonte: Elaborada pelo Autor.

A Figura 8 apresenta um esquema dos principais elementos do CIDOC-CRM e a Figura 9 um exemplo de fatos envolvendo o escritor brasileiro Machado de Assis, de onde é possível visualizar seu nascimento, morte e a criação da sua obra Dom Casmurro.

## ii) O modelo de referência EDM

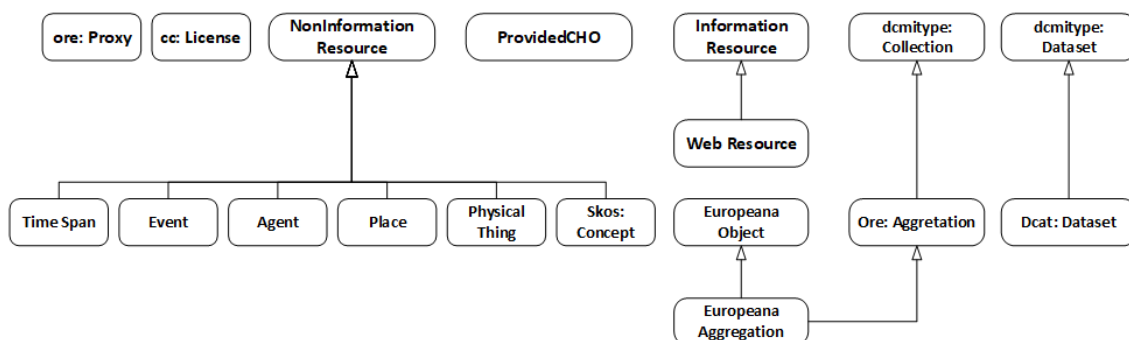
O projeto Europeana surgiu com a intenção de unir os acervos digitais das bibliotecas, museus, arquivos e outros centros culturais espalhados por toda Europa, na tentativa de disponibilizar em um único ambiente digital o acesso a essa grande massa informacional (Coneglian & Santarem Segundo, 2016).

Trata-se de uma biblioteca virtual que integra mais de 50 milhões de objetos culturais de bibliotecas, museus e arquivos. Na perspectiva de Arquivos Abertos Conectados, a Europeana permite recuperar grafos RDF a partir de consultas em um *endpoint* SPARQL. A estruturação dos dados da Europeana é orientada pelo modelo conceitual EDM – *Europeana Data Model* (Coneglian & Santarem Segundo, 2016).

O modelo EDM (*Europeana Data Model*) foi desenvolvido para suceder a ESE

(Europeana Semantic Elements). O modelo ESE é simples e permite registrar metadados sobre os objetos culturais de uma forma não-hierárquica – ou seja, não é possível registrar informação sobre partes ou componentes de um objeto –, recorrendo em parte aos metadados *Dublin Core*. A sua simplicidade, face à necessidade de registrar informação mais completa sobre os objetos, levou à proposta do EDM, que baseou-se em várias propostas da Web Semântica e em normas existentes como LIDO (museus), EAD (arquivos) e METS (bibliotecas digitais).

**Figura 10 – Classes do modelo EDM**



Fonte: (Europeana, 2017)

Na Figura 10 é apresentada uma visão hierárquica das classes do modelo EDM, onde as classes representadas por retângulos brancos são reutilizadas a partir de esquemas externos, e as em preenchidas definidas propriamente pelo EDM.

A classe *ProvidedCHO* inclui os objetos de Patrimônio Cultural que são descritos na Europeana. Funciona como uma espécie de nó central no EDM.

A classe *NonInformation Resource* se refere a todos os recursos que não são de natureza da informação, por exemplo, pessoas, lugares e objetos tangíveis. Todos esses recursos estão associados a eventos (*Event*), agentes (*Agent*), lugares (*Places*), intervalo de tempo (*Time Span*), que são classes próprias da europeana, e referem-se a conceitos, definidos pela classe externa *skos:Concept*. Estes conceitos são utilizados para contextualização dos recursos a partir de ideias ou significados específicos de acordo a descrição de sua estrutura conceitual dentro um sistema de organização do conhecimento.

A classe *EuropeanaAggregation* reúne em uma única unidade conceitual todas as informações sobre os objetos de patrimônio cultural. É partir dela que ocorre a agregação de fato da descrição dos objetos que estão em diferentes provedores de conteúdo.

## 2.6 O Centro de Memória de Goiás

No capítulo introdutório, especificamente no item 1.2 (p. 26), o Projeto Centro de Memória de Goiás foi apresentado, de forma inicial, como um dos elementos motivadores desta tese.

Tendo como pressuposto que os centros de memória, conforme afirmam de Camargo & Goulart (2015), são espaços de formação e difusão do saber que preveem a coleta, preservação e transmissão do patrimônio cultural, sobretudo com a missão de tornar a informação acessível em múltiplos contextos (Camargo & Goulart, 2015), o Projeto Centro de Memória de Goiás visa contribuir na construção e reconstrução da identidade através da digitalização e exposição de objetos do patrimônio cultural presente em espaços e acervos diversificados.

Por ser a memória é uma forma de representação da sociedade através de um conjunto de bens materiais e/ou imateriais, intelectuais ou culturais esquecidos, armazenados e selecionados por determinado grupo social, suas representações sociais são discursos memorialísticos produzidos conforme a memória, que faz lembrar e esquecer fragmentos da história (Feldman & Steindel, 2019).

Para Feldman & Steindel (2019 p. 151) “a memória é o alicerce para a constituição dos centros de memória, visto que contribui para a sensação de pertencimento a um grupo de passado comum”. Ao contribuir para o sentimento de identidade através da possibilidade da criação de memória compartilhada não só no campo real e histórico, mas no campo simbólico. Isto é, “a memória traduz informações e abarca períodos sobre determinado contexto social mediante uma necessidade do tempo-presente” (Feldman & Steindel, 2019, p. 151).

Os centros de memória são compostos de acervos que, por sua vez, são compostos de objetos (digitais ou não). Os objetos, nestes espaços, têm função de permitir a construção de uma identidade através das manifestações patrimoniais, que, segundo Dodebei (2016), é parte da cultura material. E essa cultural, segundo a autora, encontra-se em ampla expansão, sobretudo pelos museus.

As instituições que se propõem a contar a história por meio da cultura material, não param de crescer. No curso dos últimos trinta anos, o número de museus dobrou, assim como a afluência de seus visitantes. Ancoradas essencialmente nos objetos ou nos lugares, as manifestações patrimoniais se revelam um

fenômeno em expansão. Assiste-se à elaboração de políticas patrimoniais tanto em nível municipal, nacional, como internacional. Essas políticas são destinadas a proteger, a valorizar e a fazer reconhecer os objetos e os lugares patrimonializados. As crianças aprendem a história por meio da visita a exposições museológicas ou participação nas manifestações comemorativas (Dodebei, 2016, p. 230).

O trabalho de Chagas *et al.* (2018) evidenciou esse pressuposto em uma discussão sobre a criação de espaços integrados de educação patrimonial e difusão do conhecimento apresentando o projeto como um centro articular de espaços de memória, cujas ações vinculadas ao projeto pudessem agregadas em 4 núcleos, conforme esquema apresentado na Figura 11.

Estes núcleos são associados aos objetos provenientes dos espaços e aos espaços em si. Uma descrição operacional desses núcleos é encontrada em (Chagas *et al.*, 2018, p. 5):

I - Núcleo de Documentação: responsável pela higienização, catalogação e/ou digitalização de arquivos dos poderes Executivo, Legislativo e Judiciário; e arquivos cartoriais, eclesiásticos e privados. Esse núcleo não realiza a salvaguarda dos documentos e funciona como apoio a outros centros, arquivos e instituições que contenham material de interesse histórico e cultural para o Centro de Memória de Goiás.

II – Núcleo de História Oral e Imagem: se encarrega de ações que tem como objetivo o registro e catalogação de depoimentos orais e imagens das mais variadas formas de manifestação artística e cultural do povo goiano. O núcleo também é responsável pela produção de material audiovisual das atividades deste e dos outros eixos do Centro de Memória para realização de documentários, exposições e divulgação.

III – Núcleo de Arqueologia: Responsável pelas ações de arqueologia, que envolvem a coleta, catalogação e/ou organização de acervo arqueológico, bem como a sua disponibilização e apresentação em exposições ao público em geral.

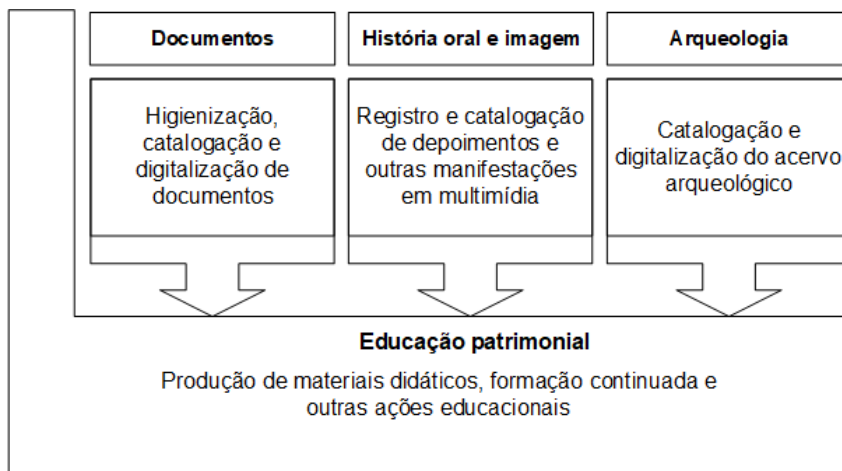
IV – Educação Patrimonial e Temas específicos da docência: refere-se às atividades relacionadas especificamente ao campo educacional, desde a produção de material didático até a formação continuada dos docentes estaduais e municipais a partir dos acervos e materiais dos outros três eixos.

Ainda na perspectiva de espaço integrado de educação e difusão do

conhecimento, o projeto tem uma característica importante se desenvolver em forma de extensão universitária. Isto é, conforme a própria essência extensionista, se constrói de forma colaborativa e participativa através da relação dialógica entre universidade e comunidade.

A Figura 11 apresenta a estruturação acadêmica do CMG, que, de acordo com os objetivos de cada núcleo, tem como eixo articulador a educação patrimonial. Reforça, portanto, o caráter e a preocupação com professores, estudantes e outros atores da educação a partir da preservação da cultura e patrimônio como identidade do povo.

**Figura 11 – Esquema dos núcleos do CMG**



Fonte: (Filgueiras *et al.*, 2020, p. 22)

## 2.7 Sumário do Capítulo

Este capítulo teve como pretensão apresentar um enquadramento teórico através de pressupostos da ciência da informação e do problema estudado, abordando as linguagens e outros elementos fundamentais das tecnologias digitais que fazem parte desta tese. Os pontos explorados reforçam o desafio da interoperabilidade semântica e aponta estratégias e ferramentas utilizadas como possibilidade de construção de respostas para o problema desta investigação. Respostas essas que esta tese pretende dar o singelo contributo apoiando-se nas tecnologias da Web-Semântica e nos Dados Abertos Conectados.

No próximo capítulo essas estratégias e tecnologias são apresentadas, em grande parte, através de um estudo do estado da arte que foi realizado através de uma revisão sistemática da literatura.

### **CAPÍTULO III. INTEROPERABILIDADE SEMÂNTICA DE PATRIMÔNIO CULTURAL: UMA REVISÃO SISTEMÁTICA**

#### **3.1 Introdução**

O objetivo deste capítulo é apresentar, em complemento ao capítulo anterior, o estado da arte da temática pesquisada através de uma revisão bibliográfica sistematizada da literatura, apontando quais principais caminhos estão sendo trilhados pelos pesquisadores de patrimônio cultural, mais precisamente no que se refere às iniciativas de interoperabilidade semântica entre Sistemas de Informação de Patrimônio Cultural.

Busca encontrar a produção das investigações publicadas em quatro grandes bases de dados científicas a partir do questionamento “quais os trabalhos de 2014 a 2021 investigaram o problema da interoperabilidade semântica em patrimônio cultural e quais estratégias são utilizadas?”. Para tanto, a revisão sistemática passou por construir uma *string* de busca, submetê-la em consulta às bases de dados selecionadas, selecionar e descartar trabalhos conforme critérios, analisar os trabalhos e construir um relatório de análise.

Um conjunto de palavras-chave genérico para ser aplicado nas bases de dados de periódicos de Ciência da Informação e Computação foi construído, partindo da questão de investigação e do problema abordado pela tese.

O conjunto genérico, composto pelos termos “interoperabilidade”, “semântica” e “patrimônio cultural”, precisou ser adaptado para cada motor de busca presente nas bases selecionadas, além de receber sinônimos possíveis para seu enriquecimento.

As bases selecionadas para pesquisa foram 1) *ACM Digital Library* e 2); *IEEE Xplorer Digital Library*; 3) *Web Of Science(Web of Knowledge)*; e 4) *Scopus (Science Direct)*, cuja escolha se deu pelo fato de serem robustas, amplamente utilizadas pela comunidade científica e possuírem trabalhos relacionados à área de Ciência da Computação e Ciência da Informação.

Após um processo de aquisição e enriquecimento através de sinônimos, as palavras-chave foram transformadas em inglês e adequadas de acordo com os motores de busca. O Quadro 3 apresenta o conjunto enriquecido de palavras-chave para cada motor de busca correspondente.

**Quadro 3 – Palavras-Chave por Base Pesquisada**

Base de Dados/Mecanismo	String de Busca Aplicada	Registros
ACM Digital Library	<i>ALL=(("semantic interoperability" OR "semantic integration" OR "semantic collaboration") AND "cultural heritage")</i>	18
IEEE	<i>((("All Metadata":semantic interoperability) OR "All Metadata":semantic integration) OR "All Metadata":semantic collaboration) AND "All Metadata":cultural heritage)</i>	18
Web of Science	<i>ALL=(("semantic interoperability" OR "semantic integration" OR "semantic collaboration") AND "cultural heritage")</i>	27
Scopus (Science Direct)	<i>TITLE-ABS-KEY ( ( "semantic interoperability" OR "semantic integration" OR "semantic collaboration" ) AND "cultural heritage" )</i>	9

Fonte: elaborado pelo autor

### 3.2 Seleção de trabalhos

Para seleção dos trabalhos levou-se em conta os registros que atendem ao conjunto de palavras-chave especificados, que tenham data de publicação igual ou posterior ao ano de 2014 e conforme critérios de inclusão e exclusão definidos dos registros encontrados em uma visão topo-base (*top-down*), partindo do título como elemento mais alto, para o texto completo, como elemento mais baixo (conforme detalhamento no CAPÍTULO IV).

#### a) Critérios de inclusão/exclusão

Para serem selecionados os artigos precisam atender 4 critérios em duas etapas:

C1 - pertencer a área da ciência da informação/computação;

C2 - ser artigo científico de anais de eventos, revistas, *journals* ou outro periódico;

C3 - ter foco no patrimônio cultural, de acordo com o que foi definido por Doerr (2009), que diz respeito a i) Gestão de acervos; ii) Conservação; iii) Pesquisa; e iv)

Apresentação (Doerr, 2009).

C4 - possuir técnicas ou estratégias de integração e/ou interoperação entre repositórios ou tecnologias.

Além destes critérios, registros duplicados também precisaram ser desconsiderados. Para os critérios C1, C2 e C3, a inclusão/exclusão foi realizada avaliação com base nos títulos e resumos dos trabalhos. Para o critério C4, foi realizada a leitura completa dos trabalhos selecionados a partir da etapa anterior.

#### **b) Critérios de avaliação**

Para avaliação dos trabalhos provenientes das etapas iniciais de seleção, foi estabelecida uma análise qualitativa a partir de oito critérios/indicadores de avaliação que estão relacionados diretamente à estratégia de cada investigação para o problema da interoperabilidade semântica de patrimônio cultural.

A1 – Se explora o modelo de referência CIDOC-CRM;

A2 – Se explora o modelo de referência EDM;

A3 – Se suas estratégias abarcam bancos de dados relacionais;

A4 – Se propõe ontologia;

A5 – Se contribui para Dados Abertos Conectados ;

A6 – Se propõe mapeamento ontológico;

A7 – Se propõe arquitetura, modelo ou sistema inovador;

A8 – Se propõe sistema ou API integrador/agregador;

### **3.3 Resultados da Busca**

A submissão das respectivas *strings* de busca nos quatro repositórios selecionados retornou o total de 72 trabalhos com duas duplicações. Todas as bases selecionadas possuem opção para busca avançada – por palavra-chave. Cada uma, no entanto, possui um conjunto específico de regras que precisam ser seguidas para a condução da pesquisa. Assim, de acordo com cada base, a *string* de busca geral foi adaptada antes de ser submetida.

Os trabalhos “*Supporting semantic interoperability in conservation-restoration domain: The PARCOURS project*” e “*ARIADNE: A Research Infrastructure for*

*Archaeology*” que foram retornadas nas bases ACM e Scopus, foram considerados apenas uma vez, o que resultou portanto em um retorno de 70 registros conforme descrito abaixo.

**a) Consulta na base da ACM Digital Library**

A consulta submetida à base de dados ACM Digital Library, que teve como string de busca *ALL=*(("semantic interoperability" OR "semantic integration" OR "semantic collaboration") AND "cultural heritage"), resultou em 18 registros. No entanto, 5 registros foram excluídos com base nos critérios propostos, resultando em 13 trabalhos que foram selecionados para a próxima etapa.

**b) Consulta à base IEEE**

A consulta submetida à base de dados IEEE Search, que teve como string de busca (((("All Metadata":semantic interoperability) OR "All Metadata":semantic integration) OR "All Metadata":semantic collaboration) AND "All Metadata":cultural heritage), obteve 18 registros com resultado. Todos os registros passaram nos critérios de exclusão, portanto, os 18 registros ficaram para a próxima etapa.

**c) Consulta à base Science Direct**

A consulta submetida à base de dados da Scopus – Science Direct, com a string de busca: (*"semantic interoperability" OR "semantic integration" OR "semantic collaboration"*) AND *"cultural heritage"* resultou em 27 registros.

Após aplicação dos critérios de inclusão/exclusão, 16 registros foram excluídos pelos critérios apresentados, resultando em 11 para o próxima etapa.

**d) Consulta à base Web of Science**

A consulta submetida à base de dados Web of Knowledge, da Web of Science, teve a string de busca *ALL=*(("semantic interoperability" OR "semantic integration" OR "semantic collaboration") AND "cultural heritage") e resultou em 9 registros.

Após aplicação dos critérios de 1 a 3, nenhum registro foi excluído, sendo os 9 registros selecionados para a próxima etapa. Há duas repetições, no entanto, com a base ACM e, assim, restam 7 registros efetivamente.

### **3.4 Tratamento dos registros retornados**

Como foram retornados 72 registros que atenderam às palavras-chave nas buscas realizadas em cada base de dados, conforme a Tabela 1, esses registros passaram por uma

primeira análise com o base no título e resumo.

**Tabela 1: Resumo dos registros retornados**

Base	Nº registros	Critério para exclusão				Excluídos	Selecionados
		C1	C2	C3	RPT		
<i>ACM Digital</i>	18	0	3	2		5	13
<i>IEEE Search</i>	18	0	0	0		0	18
<i>Science Direct</i>	27	1	9	6		16	11
<i>Web Of Science</i>	9	0	0	0	2	2	7
Total	72	1	12	8	2	23	49

Fonte: Elaborado pelo Autor

Dois registros foram encontrados em duplicidade, dessa forma, o universo de registros reduz-se a 72 e registros selecionados pelo título e resumo totalizam 49.

### 3.4.1 Aplicação dos critérios de inclusão/exclusão

Os critérios de inclusão/exclusão de C1, C2 e C3 foram aplicados com base no título, resumo e leitura parcial de cada registro encontrado.

Quanto ao primeiro critério (C1) de inclusão/exclusão, apenas um registro foi detectado e descartado por não pertencer à área de ciência da informação e computação, *StrainInfo introduces electronic passports for microorganisms* (Verslyppe *et al.*, 2014, da base Science Direct.

O artigo, da área da microbiologia, apresenta um sistema semântico denominado *StrainInfo* foi concebido para ser uma plataforma aberta que integra informações microbianas. Esta plataforma pretende formar uma base para novos métodos de pesquisa microbiológica, alavancando a vasta quantidade de informações eletrônicas disponíveis *on-line* (Verslyppe *et al.*, 2014). Apesar de extremamente interessante, por não se relacionar às áreas da Ciência da Informação e Computação conforme critério C1, o artigo não foi selecionado.

Quanto ao Critério C2, registros referentes a artigos de periódicos e anais de eventos científicos, resultou em 12 registros excluídos. As bases ACM e Science Direct trazem na sua busca avançada livros e registros de conferência, inclusive dos índices de livros, trazendo duplicação dos registros em relação aos autores de capítulos de livros.

Da consulta submetida à ACM foram excluídos, portanto, os seguintes registros:

1. CSCW '17: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing
2. JOCCH Special Issue on Digital Infrastructures for Cultural Heritage (2nd Part) Guest Editors: Massimo De Santo, Franco Niccolucci, and Julian Richards
3. WWW '19: Companion Proceedings of The 2019 World Wide Web Conference

Nos registros da Science Direct foram:

4. *Digital Curation in the Digital Humanities (index)*
5. *Investigation of Decision Process for Purchasing Foodstuff in the “Bento” Lunch Box*
6. *Introduction to digital libraries*
7. *Archives in the Digital Age (index)*
8. *Metadata*
9. *Metadata \**
10. *Metadata for transmedia resource (introduction)*
11. *Archival integration and dissemination: The trends*
12. *Information architecture and hypertextuality: concerns for digital curation*

Quanto ao Critério (C3), ter o artigo o foco em patrimônio cultural, oito registros foram excluídos. Com base nos resumos, foram notadas várias propostas interessantes quase sempre na área de ciência da computação, mas sem ligação com o patrimônio cultural, ou com estratégias que diretamente pudessem ser observadas para o contexto deste trabalho.

Nos registros da ACM, foram excluídos:

1. *Visual Recognition of Ancient Inscriptions Using Convolutional Neural Network and Fisher Vector (G. Amato et al., 2016)*
2. *Semantic Framework for Interactive Animation Generation Hui (Liang et al., 2016).*

Nos registros da Science Direct foram:

3. *Investigation of Decision Process for Purchasing Foodstuff in the “Bento” Lunch Box* (Nishimoto et al., 2015) .
4. *A Formal Model for Autonomous Planning in High Performance Systems* (F. Amato et al., 2016).
5. *Interoperable cross-domain semantic and geospatial framework for automatic change detection* (Kuo & Hong, 2016) .
6. *Information from imagery: ISPRS scientific vision and research agenda* (J. Chen et al., 2016).
7. *Live Monitoring of Earth Surface (LiMES): A framework for monitoring environmental changes from Earth Observations* (Giuliani et al., 2017).
8. *Analysis and assessment of a knowledge based smart city architecture providing service APIs* (Badii et al., 2017) .

#### **3.4.2 Trabalhos selecionados para leitura completa**

Após a aplicação dos critérios de exclusão da primeira etapa, foram selecionados 49 trabalhos que são listados abaixo conforme base de pesquisa.

Nos registros da ACM foram:

1. *A Framework Design for Information Management in Heritage Science Laboratories* (Bordalo et al., 2021).
2. *Towards Big Religious Data: RESILIENCE Research Infrastructure for Data on Religion in the Digital Age* (Büchler et al., 2020).
3. *A Semantically Annotated JSON Metadata Structure for Open Linked Cultural Data in Neo4j* (Drakopoulos et al., 2019).
4. *Context-Based Infomobility System for Cultural Heritage Recommendation: Tourist Assistant—TAIS* (Smirnov et al., 2017).
5. *Large-Scale Semantic Integration of Linked Data: A Survey* (Mountantonakis & Tzitzikas, 2019).
6. *Ontologies for the Real Property Domain* (Shi & Roman, 2018).
7. *ARIADNE: A Research Infrastructure for Archaeology* (Meghini et al., 2017).

8. *Enhancing Semantic Expressivity in the Cultural Heritage Domain: Exposing the Zeri Photo Archive as Linked Open Data* (Daquino *et al.*, 2017).
9. *Semantic Infrastructure of a Smart Museum: Toward Making Cultural Heritage Knowledge Usable and Creatable by Visitors and Professionals* (Korzun *et al.*, 2017).
10. *Supporting semantic interoperability in conservation-restoration domain: The PARCOURS project* (Niang *et al.*, 2017).
11. *Ontology-Based Data Integration System for Conservation-Restoration Data (OBDIS-CR)* (Niang *et al.*, 2016).
12. *Great War Stories Told by the People – Crowdsourced Cultural Heritage in Digital Museums Ingo* (Frommholz *et al.*, 2014).
13. *Ontological Approaches to Information Description and Extraction in the Cultural Heritage Domain* (Daquino & Tomasi, 2014).

Nos registros da IEEE foram.

14. *OntoM: An Ontological Approach for Automatic Classification* (Abu Musa *et al.*, 2020).
15. *Towards a semantic recommender system for cultural objects: Case study Draa-Tafilalet region* (Nafis *et al.*, 2020).
16. *A Matching Framework for Multimedia Data Integration Using Semantics and Ontologies* (Rinaldi & Russo, 2018).
17. *A step towards the linking of the Culture Heritage in the Egyptian museums* (Owais *et al.*, 2017).
18. *Aggregation of Linked Data : A case study in the cultural heritage domain* (N Freire *et al.*, 2018).
19. *An Edge Intelligence Empowered Recommender System Enabling Cultural Heritage Applications* (Su *et al.*, 2019).
20. *Architectural Heritage semantic data managing and sharing in GIS* (Cerutti *et al.*, 2015).
21. *CRMCR - a CIDOC-CRM extension for supporting semantic interoperability*

*in the conservation and restoration domain* (Bannour et al., 2018).

22. *Information extraction and integration for enriching cultural heritage collections* (Buranasing et al., 2016).

23. *Object-mediated innovation: Case study of adventures in the National Palace Museum in Taiwan* (H. Chen et al., 2014).

24. *Semantic enterprise service bus for cultural heritage* (Capodieci et al., 2016)

25. *Towards a semantic data infrastructure for heterogeneous Cultural Heritage data - challenges of Korean Cultural Heritage Data Model (KCHDM)* (Kim et al., 2015).

26. *Visual metaphors for semantic cultural heritage* (Damiano et al., 2015).

27. *The MONDIS project Semantic Web and the protection of historic buildings* (Cacciotti & Valach, 2015).

28. *A 'historical case' of Ontology-Based Data Access* (Calvanese et al., 2015).

29. *Personalizing the Museum Experience through Context-Aware Recommendations* (Benouaret & Lenne, 2015).

30. *An Associative Engines Based Approach supporting Collaborative Analytics in the Internet of Cultural Things* (Benedusi et al., 2015).

31. *Management and retrieval of cultural heritage multimedia collection using ontology* (Rubhasy et al., 2015).

Nos registros da Science Direct foram:

32. *Generating knowledge graphs by employing Natural Language Processing and Machine Learning techniques within the scholarly domain* (Dessì et al., 2020)

33. *Generating knowledge graphs by employing Natural Language Processing and Machine Learning techniques within the scholarly domain* (Nishanbaev, 2020)

34. *An on-line platform to unify and synchronise heritage architecture information* (Palomar et al., 2020).

35. *Formal Description and Automatic Generation of Learning Spaces Based on*

*Ontologies* (Martini *et al.*, 2016).

36. *Ontology-based data integration in EPNet: Production and distribution of food during the Roman Empire* (Calvanese *et al.*, 2016).
37. *Digital library interoperability at high level of abstraction* (Agosti *et al.*, 2015).
38. *Incorporate Cultural Artifacts Conservation Documentation to Information Exchange Standards – The DOC-CULTURE Case* (Kouis & Giannakopoulos, 2014).
39. *Knowledge-based data enrichment for HBIM: Exploring high-quality models using the semantic-web* (Quattrini *et al.*, 2017).
40. *Ontology-based affective models to organize artworks in the social semantic web* (Bertola & Patti, 2015).
41. *Providing an Application-specific Interface over a CERIF Back-end: Challenges and Solutions* (Ivanović & Houssos, 2014).
42. *Smart Space based Recommendation Service for Historical Tourism* (Varfolomeyev *et al.*, 2015).

Nos registros da Web of science foram:

43. *Bulgarian Icons in Wikidata and EDM* (Alexiev *et al.*, 2020).
44. *Inside the Meanings. The Usefulness of a Register of Ontologies in the cultural Heritage Sector* (Veninata, 2020).
45. *A knowledge-based approach to Information Extraction for semantic interoperability in the archaeology domain* (Vlachidis & Tudhope, 2015).
46. *Archaeological stratigraphy as a formal language for virtual reconstruction. Theory and practice* (Demetrescu, 2015).
47. *A pattern-based framework for best practice implementation of CRM/FRBROO* (Aalberg *et al.*, 2015).
48. *Agent-oriented and ontology-driven digital libraries: the IndianaMAS experience* (Briola *et al.*, 2017).
49. *The Iccu's Contribution To The Development Of A Digital Infrastructure For*

*Cultural Heritage* (Butto, 2017)

### 3.5 Síntese dos trabalhos

Após a leitura dos trabalhos, foi possível identificar que há inúmeras estratégias para enfrentar o problema da interoperabilidade semântica em patrimônio cultural. Na literatura, a abordagem mais comum para processos semânticos integradores é a utilização de ontologias, em que três técnicas se destacam. A primeira, a utilização de uma ontologia única que exige que os diversos repositórios atendam a um só domínio. A segunda, mais flexível, permite ontologias locais ligadas a uma ou mais ontologias superiores ou agregadores, sendo necessário construir mapeamento. E há, por fim a terceira, que é uma abordagem híbrida que mistura elementos das duas primeiras, criando-se um vocabulário compartilhado (Niang *et al.*, 2017)

Alguns questionamentos realizados na leitura dos artigos foram se as estratégias exploram CIDOC-CRM e EDM; se propõem ontologia; se realizam mapeamento ontológico; se propõem arquitetura ou sistema; se têm atuação nos Arquivos Abertos Conectados; se tratam bancos de dados relacionais; se propõem sistema ou API integrado ou agregador. Ressalte-se que os trabalhos encontrados não necessariamente atuam diretamente nesse problema.

Dessa forma, os trabalhos foram aqui divididos de acordo com o domínio do assunto principal em que atuam entre, 1) integração de metadados, esquemas e ontologias; 2) integração de multimídia e geoprocessamento; 3) repositórios e experiências semânticas; 4) proteção, conservação e restauração de patrimônio cultural; 5) reflexões sobre conceitos, linguagens e ontologias; e 6) sistemas de recomendação e espaços inteligentes .

#### i) **Integração de metadados, esquemas e ontologias**

Dos 14 trabalhos classificados nessa seção, em respeito às questões de estratégia utilizadas, 10 propõem sistema ou arquitetura, sendo que dentre eles 4 propõem API ou sistema agregador; 11 realizam mapeamento ontológico, sendo que 8 propõem ontologias. Desses trabalhos, 5 abordam o CIDOC-CRM e 3 o modelo EDM. De alguma maneira, 4 envolvem estratégias com bases de dados relacionais.

Abu Musa *et al.* (2020) no artigo, *Towards a semantic recommender system for cultural objects: Case study Draa-Tafilalet region* apresentam um modelo de ontologia que incorpora conceitos relevantes do domínio do Patrimônio Cultural no Catar. Na sua

proposta a ontologia criada é utilizada para realizar inferências sobre classificações de objetos culturais por meio de duas abordagens. A primeira, por correspondência textual (palavras-chave), que permite associação direta entre o objeto e os conceitos da ontologia. A segunda, por correspondência semântica, onde através de um banco de dados léxico (*WordNet*) busca-se todos os possíveis sinônimos para propriedades de um determinado objeto (Abu Musa *et al.*, 2020).

Drakopoulos *et al.* (2019) em *A Semantically Annotated JSON Metadata Structure For Open Linked Cultural Data In Neo4j*, propõem uma descrição JSON para itens culturais que é adaptada às necessidades do domínio da cultura digital, que pretende ser genérica para resumir a grande maioria dos itens culturais e flexível para que haja extensões mais especializadas (Drakopoulos *et al.*, 2019).

A proposta prevê agrupamentos multiníveis de construção, anotações semânticas e links para itens culturais relevantes. Como um exemplo concreto, a descrição proposta foi importada para uma instância de um banco de dados de gráficos Neo4j. Este último aproveita ao máximo os recursos oferecidos pela estrutura de metadados JSON, especialmente na ligação dinâmica direcionada entre itens culturais semelhantes.

Este trabalho apresenta três resultados: o primeiro, uma proposta de uma descrição de metadados do JSON para representar um amplo espectro de itens culturais. O segundo, um conjunto de métricas e análises de similaridade baseadas na teoria dos conjuntos e na teoria das probabilidades, respectivamente, que se beneficiam da estrutura de metadados acima mencionada. E, por fim, uma demonstração do potencial dos metadados e da análise, implementando-os em uma instância de um banco de dados Neo4j preenchido com itens culturais da região das ilhas Jônicas (Grécia).

Nuno Freire *et al.* (2018), em *Aggregation of Linked Data: A case study in the cultural heritage domain*, contribuem com um trabalho com o foco na agregação. Técnica que consiste em centralizar os metadados de várias em fontes de dados em um espaço comum (Nuno Freire *et al.*, 2018, p. 522).

A existência de muitas DLs traz descoberta desafios para seus usuários em potencial. Em todas as instituições, esses desafios são geralmente abordados por uma organização arquitetura baseada em uma organização central, que permite a descoberta de recursos de CH, coletando seus metadados associados e construindo um portal ou serviço da Web sobre eles. Essa organização tem a possibilidade de promover ainda mais o uso dos recursos por meios que não

podem ser realizados por cada DL isoladamente. Nos últimos anos, o domínio CH tem sido bem-sucedido na criação sustentável de iniciativas de agregação, como Europeana, DPLA, DigitalNZ, Trove e a National Digital Library of India, que coletam e fornecem acesso aos ativos culturais digitalizados públicos da Europa, Estados Unidos da América, Nova Zelândia, Austrália e Índia, respectivamente

Neste artigo os autores apresentam resultados de um estudo de caso piloto na Europeana (Europeana, 2017) que utilizou a Biblioteca Nacional da Holanda como provedor inicial de dados e a Dutch Digital Heritage Network como serviço intermediário que fornece conjuntos de dados para a Europeana, conforme relatam os autores:

Analisamos os requisitos para uma solução de agregação de LD e definimos um fluxo de trabalho que atende aos mesmos requisitos funcionais da solução atual da Europeana. O fluxo de trabalho foi colocado em prática no piloto e levou ao desenvolvimento de vários componentes de software para gerenciamento de conjuntos de dados, coleta de LD, análise e integração de dados (Nuno Freire *et al.*, 2018, p. 1).

Os autores relatam a experiência discutem um conjunto de esforços de se adotar essa abordagem de arquivos conectados para provedores e agregadores de dados, e reforçam experiência exigida para analistas de dados no domínio do patrimônio cultural e as ferramentas de suporte necessárias para dados semânticos.

Relatam ainda dois aspectos essenciais da experiência. A utilização do EDM como modelo comum e o compartilhamento das informações entre os provedores e o agregador. Este segundo, criticamente importante, pois “os metadados são transferidos para o agregador, mas continuam a evoluir no provedor” (Nuno Freire *et al.*, 2018, p. 523). Isto é, claramente pode suscitar em um problema de sincronização.

Buranasing *et al.* (2016) em *Information Extraction and Integration for Enriching Cultural Heritage Collections* (Buranasing *et al.*, 2016), propõem uma nova abordagem para descobrir relações semânticas entre entidades de artigos que usam a *Wikipedia* e vários objetos digitais de patrimônio cultural.

A estratégia defendida por estes autores consiste em um projeto para integrar vários assuntos e padrões de metadados de várias fontes de dados de patrimônio cultural através de padrões de relações semânticas para extrair termos e significado, a criação de um dicionário de patrimônio cultural e extração de relações entre pares de entidades em

sentenças.

Os autores propõem ainda um método para enriquecer informações sobre o patrimônio cultural com o resultado da extração de relações semânticas usando a correspondência de similaridade de sequências semânticas para alcançar conteúdos estruturados e semiestruturados, além de um enriquecimento de semântico de informações advindas de vários protocolos e modelos de metadados como *Dublin Core* e OAI-PMH.

Esta proposta consiste basicamente da criação de um dicionário de dados através de termos utilizados e seus significados, bem como da extração de relação semântica entre termos após a avaliação de determinadas perguntas. Utiliza ainda da análise de similaridade textual para enriquecer a busca semântica.

Capodieci *et al.* (2015), em *Semantic enterprise service bus for cultural heritage*, propõem uma arquitetura para um para a integração de fontes de dados estruturadas e semiestruturadas heterogêneas, permitindo que o conteúdo da informação seja mais acessível e pesquisável para os usuários (Capodieci *et al.*, 2016).

O modelo de integração de dados proposto neste artigo e visa explorar o método de integração de esquemas (*schema matching*), utilizando técnicas de integração de dados baseadas em ontologias e automatizando o processo de integração através de ferramentas para este fim (Capodieci *et al.*, 2016).

O sistema de integração de dados proposto neste artigo visa explorar o método de “*schema matching*”, possibilitando, se necessário, o aproveitamento também de dados provenientes de serviços ou aplicações web. O método apresentado usa técnicas de integração de dados baseadas em ontologia, automatizando o processo de integração por meio de ferramentas de mapeamento como Karma e D2RQ.

Kim *et al.* (2015), no artigo *Towards a semantic data infrastructure for heterogeneous Cultural Heritage data - challenges of Korean Cultural Heritage Data Model (KCHDM)*, discutem os desafios da integração de bancos dados heterogêneos relativos ao patrimônio cultural da Coreia.

Neste trabalho, os autores apresentam resultados do projeto de pesquisa ‘*K-Culture Time Machine*’ o qual foi lançado com o objetivo de ligar um banco de dados heterogêneo com um esquema de metadados abrangente e desenvolver um serviço de

informação ontológico para o patrimônio da Coreia, que tem como base os modelos de referência EDM e CIDOC-CRM

[...] propomos uma estrutura de esquema de metadados para incorporar recursos de bancos de dados heterogêneos que já possuem seus próprios esquemas de metadados. Referimo-nos ao caso do Europeana Data Model (EDM); criamos um modelo de dados baseado em ontologia, denominado Modelo de Dados do Patrimônio Cultural Coreano (KCHDM). O esquema de metadados que estamos propondo tem como objetivo funcionar como um modelo de dados superior para agregar os modelos de diferentes instituições. Além disso, o modelo de dados visa fornecer a possibilidade de uma ligação semântica entre conjuntos de dados heterogêneos. Para tanto, propomos um algoritmo que extrai padrões semânticos de textos descritivos do banco de dados patrimonial de cada instituição (Kim *et al.*, 2015, p. 4).

A ontologia KCHDM fornece informações de conteúdo chave pertencentes ao Patrimônio Cultural da Coreia do Sul, como informações sobre histórico figuras, eventos históricos, artefatos, tempo e lugar.

O ponto de partida da ontologia aponta para cinco classes conceituais, em similares àquilo que é comum a outras ontologias de patrimônio cultural: objeto, evento, ator, lugar e tempo.

Damiano *et al.* (2015), no trabalho *Visual metaphors for semantic cultural heritage*, apresentam um estudo de caso em ontologias e visualização para patrimônio cultural em um sistema denominado *Labyrinth*. Nesse sistema o usuário está imerso em um labirinto 3D em que pontos de inflexão e caminhos representam um conjunto de artefatos culturais e as relações semânticas que se mantêm entre eles (Damiano *et al.*, 2015).

A estratégia apresentada neste trabalho trata da utilização de metáforas visuais em ambientes virtuais como ferramenta para transmitir conceituação de itens do patrimônio cultural e as relações que se mantêm entre eles através de ontologias, conforme descrição realizada pelos autores:

O uso de metáforas visuais para transmitir as informações armazenadas em uma ontologia permite comunicar o modelo conceitual que esta codifica de forma imediata e envolvente, conforme demonstrado por uma linha de pesquisa estabelecida em visualização de informações. Como um estudo de caso, descrevemos a aplicação da estrutura ao projeto, desenvolvimento e edição do

sistema Labirinto. O sistema Labirinto é um aplicativo 3D que conta com uma ontologia de “arquétipos culturais” para criar e visualizar um “labirinto” de relações semânticas entre as obras contidas em um repositório (Damiano *et al.*, 2015, p. 2).

A arquitetura proposta contém um servidor de ontologias que mantém as ontologias do patrimônio cultural e fornece os serviços de raciocínio que permitem ao sistema inferir outras propriedades dos objetos a partir dos declarados. Contém também um repositório de mídia que são indexados por meio de um banco de dados relacional. Um *Web Service* que implementa a API da qual o cliente de visualização depende para consultar servidor de ontologias; e por fim, um módulo de visualização para a interação com o usuário por meio de navegação 3D.

Calvanese *et al.* (2016) no artigo intitulado *Ontology-based data integration in EPNet: Production and distribution of food during the Roman Empire*, fornecem uma abordagem baseada no paradigma Ontology-Based Data Access (OBDA), na qual os diferentes conjuntos de dados são virtualmente integrados por uma camada conceitual ontológica que fornece ao usuário um ponto de acesso claro e uma visão conceitual unificada e inequívoca para que estudiosos possam ter acesso a dados históricos do comércio de alimentos durante o Império Romano (Calvanese *et al.*, 2016).

No modelo proposto por estes autores, há uma ontologia que funciona como um modelo de referência para o sistema, e é baseado em CIDOC-CRM, no modelo de metadados EAGLE17 e outros. Foi também desenvolvido um mapeamento para conjuntos heterogêneos de dados relacionados ao patrimônio cultural, que foi gerado de forma semi-automática e, também, é fornecido um *endpoint* SPARQL para consultas externas.

Em *A ‘historical case’ of Ontology-Based Data Access*, (Calvanese *et al.*, 2015), é apresentada uma plataforma baseada na Web para facilitar o acesso de pesquisadores a dados históricos e culturais distribuídos em diferentes fontes de dados.

A abordagem faz referência ao paradigma de acesso a dados baseados em ontologia (OBDA), no qual diferentes conjuntos de dados são virtualmente integrados por uma camada conceitual que é modelada em uma ontologia (Calvanese *et al.*, 2015).

Este trabalho investigou os mecanismos e características do sistema de produção de alimentos e comércio do Império Romano. Os autores, com base no modelo CIDOC-

CRM, exploram uma lacuna dos modelos existentes em vários projetos que é a dificuldade dos termos serem entendidos diretamente por não especialistas, pois os nomes dos conceitos geralmente não são auto-explicativos (por exemplo, o nome do conceito para "*Information Carrier*" é "E84" no CIDOC-CRM) e os conceitos são definidos intencionalmente em um nível muito abstrato, a fim de serem úteis para qualquer domínio no campo das humanidades digitais (por exemplo, E75: 'Denominação de Objeto Conceitual')

Agosti *et al.* (2015), no artigo *Digital library interoperability at high level of abstraction* abordam dois modelos de referência de bibliotecas digitais, o DELOS e o 5S - Streams, Structures, Spaces, Scenarios, Societies, fornecendo uma solução ontológica que pretendem ir para além do mapeamento dos dois modelos.

Os autores propõem uma nova abordagem diferente daquilo que é normalmente utilizado para interoperação entre sistemas semânticos e que eles entendem como uma estratégia externa e *bottom-up* (Agosti *et al.*, 2015, p. 2).

A assumir que tudo em uma DL (Digital Library) deve ser exposto na Web, em vez de buscar interoperabilidade direta entre sistemas que podem não ser necessariamente apenas baseados na Web. É "*bottom-up*" porque as ontologias foram usadas apenas para descrever os recursos gerenciados por um DLS (*Digital Library System*) e não são usadas para representar os próprios conceitos que constituem o modelo DL no qual o DLS é baseado. Portanto, eles permitem a interoperabilidade semântica e integração apenas no nível de dados, ou seja, o nível mais baixo possível na arquitetura de um DLS (Agosti *et al.*, 2015, p. 2).

E portanto, "como consequência, cada DLS operacional pode não *saber* na verdade a maneira como outro DLS operacional chama e se refere às mesmas operações" (Agosti *et al.*, 2015, p. 2).

A abordagem proposta por estes autores é composta de uma ontologia denominada "*DL Foundational Models Ontology*", que para além de simplesmente mapear duas ontologias (DELOS e 5S), atua na base dos sistemas de bibliotecas digitais (DLS). Ao invés de mapear as ontologias diretamente, essa abordagem propõe a utilização de uma ontologia - *DL Foundational Model Ontology* –que representa cada conceito central no domínio das bibliotecas digitais (DL).

Ivanović & Houssos(2014), no artigo *Providing an application-specific interface over a CERIF back-end: challenges and solutions*, apresentam um caso de uso

de exposição de informação baseada no modelo conceitual CERIF - *Common European Research Information Format* – através de uma interface que permite a utilização por parte de desenvolvedores que não tenham conhecimento prévio no modelo CERIF.

O padrão CERIF descreve um modelo de dados formal que permite a interoperabilidade entre os sistemas de gestão de pesquisa e este modelo contém informações sobre pessoas, projetos, organizações, publicações, patentes, eventos, prêmios, equipamentos. As entidades CERIF são divididas em grupos. O grupo principal contém entidades Projeto, Pessoa e Organização. (Ivanović & Houssos, 2014, p. 13).

A estratégia utilizada consta da implementação de um arcabouço que permite a conversão bidirecional dos dados inseridos usando uma interface de aplicativo de modelo simples para um *back-end* e o uso do arcabouço no projeto como parte de uma infraestrutura aberta para Informações do Setor Público (Ivanović & Houssos, 2014).

O arcabouço proposto foi implementado usando a plataforma Java, com o *framework Spring*, bibliotecas de software livre escritas em Java, e uma biblioteca específica do projeto, denominada CERIF JPA.

O componente Modelo representa o modelo simples usado que não é baseado no modelo de referência CERIF. *Converters* são componentes que permitem um único conversor para cada entidade de modelo simples e fornecem mapeamentos de entidades de um modelo simples para entidades CERIF JPA e vice-versa. O componente *Repositories* inclui classes que fornecem operações básicas sobre o banco de dados CERIF e estendem as classes da biblioteca CERIF JPA implementando localizadores específicos sobre entidades CERIF necessárias para melhorar o desempenho da conversão entre entidades de modelo simples e entidades de modelo CERIF.

A camada *Services* inclui classes que fornecem operações de leitura e escrita para entidades de modelo simples, e a *API REST* é uma camada de classes que expõem operações leitura e escrita para entidades de modelo simples na *API REST*, através de acesso de algum sistema remoto que pode ser implementado por qualquer linguagem e plataforma de programação (Ivanović & Houssos, 2014).

Vlachidis & Tudhope (2015), em *A Knowledge-Based Approach to Information Extraction for Semantic Interoperability in the Archaeology Domain*, apresentam um método para indexação semântica automática de relatórios arqueológicos de “literatura

cinzenta” usando técnicas empíricas (baseadas em regras) de Extração de Informação em combinação com sistemas de organização de conhecimento específicos de domínio. Literatura cinzenta está relacionada ao tipo de literatura que não foi publicada ou foi publicada de forma não convencional, por exemplo, relatórios governamentais.

Segundo os autores, o sistema de anotação semântica (OPTIMA), realiza as seguintes tarefas (Vlachidis & Tudhope, 2015, p. 1):

Reconhecimento de Entidade Nomeada, Extração de Relação, Detecção de Negação e Desambiguação de Sentido de Palavras usando regras artesanais e recursos terminológicos para associar abstrações contextuais com classes do Modelo de Referência Conceitual CIDOC da ontologia padrão (CRM ) para o patrimônio cultural e sua extensão arqueológica, CRM-EH .

Como conclusão, os autores demonstram a efetividade do sistema proposto nos processos de indexação semântica através do reconhecimento de entidades CIDOC-CRM, usando técnicas de processamento de linguagem natural conduzidas por um uso complementar de recursos de domínio ontológico e terminológico e emprego de regras de extração baseadas em contexto para a criação de redes semânticas compostas de frases de texto não estruturado.

Demetrescu (2015), em *Archaeological stratigraphy as a formal language for virtual reconstruction. Theory and practice*, propõe o uso de uma linguagem formal para acompanhar todo o processo de reconstrução virtual de objetos de patrimônio cultural em 3D. A proposta é baseada em uma abordagem de leitura estratigráfica e visa criar um quadro comum conectando a documentação arqueológica e a reconstrução virtual nas fases iniciais da pesquisa.

A estratégia adotada para essa construção é a utilização de ferramentas e padrões usados na pesquisa arqueológica que foram "estendidos" para anotar taxonomicamente tanto a validação da hipótese quanto as fontes envolvidas (Demetrescu, 2015).

No trabalho *A pattern-based framework for best practice implementation of CRM/FRBROO* (Aalberg *et al.*, 2015), os autores apresentam um *framework* para compartilhar o conhecimento das melhores práticas relacionadas ao uso dos modelos de referência CIDOC-CRM e FRBR para desenvolvedores de software de patrimônio cultural cuja abordagem é inspirada na Engenharia de Software, conforme descrevem os

autores:

Essa abordagem é inspirada no conhecido conceito de padrões de projeto da Engenharia de Software. Os padrões de interesse para os usuários de CRM e FRBRoo são fragmentos de modelo para casos de uso comuns acompanhados de uma documentação explicando o padrão, o contexto para este padrão, bem como outras informações relevantes relacionadas ao padrão (Aalberg *et al.*, 2015, p. 4).

Estes autores afirmam que um modelo adequado deve consistir em um nome para o padrão, cenários de caso de uso e uma descrição da intenção do padrão, um fragmento de modelo gráfico, exemplos de implementação e mapeamento autorizados.

A estratégia para criação do catálogo de padrões foi utilizar um *wiki* semântico, usando a extensão *MediaWiki Semantic* (SMW) e, para o aspecto gráfico, uma extensão de uma API denominada *PlantUML* para SMW que transforma descrições textuais em diagramas UML.

Briola *et al.* (2017), com o artigo intitulado *Agent-oriented and ontology-driven digital libraries: the IndianaMAS experience*, apresentam o *IndianaMAS*, um sistema multiagente capaz de classificar e gerenciar automaticamente imagens, esboços e documentos multilíngues em um domínio de patrimônio cultural (Briola *et al.*, 2017).

Esse sistema se utiliza de agentes, ontologias e bibliotecas digitais com comunicação através de serviços *RESTful*. Na abordagem utilizada neste sistema, o patrimônio cultural é descrito formalmente em uma ontologia, que, segundo os autores, permite a integração semântica de diferentes fontes, direciona a comunicação do agente com o ambiente interno e externo e fornece uma interface abstrata e legível entre o sistema e o usuário.

## ii) **Integração de Multimídia e Geoprocessamento**

Foram encontrados cinco trabalhos que investigam questões relacionadas ao patrimônio cultural, mas dentro do domínio de multimídia e geoprocessamento. Dois desses trabalhos propõem sistemas ou API integradores. Três abordam de alguma maneira o CIDOC-CRM e dois o EDM. Três se enquadram efetivamente como propostas de Arquivos Abertos Conectados e um propõe uma nova ontologia.

Nishanbaev (2020), no artigo *A web repository for geo-located 3D digital cultural heritage models*, propõe uma nova metodologia e um repositório Web para

integrar mapas, modelos 3D e dados geoespaciais como geolocalização, que pode ser utilizado para arquivamento a longo prazo e visualização de modelos de patrimônio cultural digital em 3D geolocalizados na web. O autor defende que:

Ao contrário de muitos projetos anteriores relacionados que desenvolveram os repositórios da web do zero, o repositório da web construído usando esta metodologia é baseado em um sistema de gerenciamento de conteúdo de código aberto e fácil de implementar, nomeadamente *KeystoneJS* e outros *frameworks* associados, que são reutilizáveis e completamente extensível. Ele pode, portanto, ajudar as organizações de patrimônio cultural e os profissionais do patrimônio cultural a facilitar o rápido desenvolvimento de repositórios da web para modelos de patrimônio cultural digital 3D geo-localizados, que também podem ser estendidos de acordo com os requisitos do projeto (Nishanbaev, 2020, p. 1).

A estratégia proposta é montada por três partes. A primeira parte é a ingestão de modelos 3D para o sistema de gerenciamento de conteúdo, o que é feito por meio de uma interface gráfica de usuário (GUI). A segunda é o *back-end*, que é realizada por um CMS denominado “KeystoneJS” baseado em um ambiente de tempo de execução “Node” que permite aos desenvolvedores implementar comandos do lado do servidor na linguagem de programação *JavaScript*. A terceira parte é uma interface de *front-end* para renderização digital 3D de modelos de patrimônio cultural, que foi implementada utilizando a biblioteca *3D Java-Script* denominada *Three.js*.

Alexiev *et al.* (2020), no trabalho *Bulgarian Icons in Wikidata and EDM*, apresentam uma proposta de modelo de conhecimento em grafo para ícones com base na WikiMedia e no modelo EDM. Este modelo é composto de classes como iconógrafo, escola iconográfica, hermínia, ícone, medidas, material, técnica, sujeito iconográfico, artigo científico, periódico acadêmico, edição e ligações para classes externas pelos dados abertos conectados (Alexiev *et al.*, 2020).

O artigo descreve um esforço para contribuir com dados de ícones búlgaros para a rede através de dados abertos conectados pela Enciclopédia Virtual de Ícones da Bulgária (BIDL), cuja estratégia principal é converter e exportar os recursos para o *Wikidata* e EDM e interligar a outros dados abertos conectados, como o *Getty*.

Em *Enhancing Semantic Expressivity in the Cultural Heritage Domain: Exposing the Zeri Photo Archive as Linked Open Data*, Daquino *et al.* (2017) propõem um sistema que pretende aumentar a expressividade do valor semântico dos dados de

patrimônio cultural em um conjunto de passos. Primeiro, são propostas duas ontologias *ad hoc* para descrever todos os problemas não totalmente cobertos pelos modelos existentes – a entrada F e a ontologia de entrada do OA (Daquino *et al.*, 2017). Em seguida, realizam um mapeamento RDF com os elementos descritivos usados no catálogo atual do *Zeri Photo Archive* para convertê-lo em CIDOC-CRM e nos dois novos modelos mencionados acima os dados de origem com base nos padrões de conteúdo italiano *Scheda F* e *Scheda OA*. Por fim, criam um conjunto de dados RDF da saída do mapeamento que poderia mostrar um resultado capaz de demonstrar a complexidade do cenário apresentado.

Rinaldi & Russo (2018), no artigo intitulado *A Matching Framework for Multimedia Data Integration Using Semantics and Ontologies*, propõem uma técnica de integração de dados em uma abordagem com foco em grandes bases de conhecimento de multimídia usando formalismos conhecidos para representação do conhecimento como ontologias (Rinaldi & Russo, 2018).

A abordagem proposta é suportada por uma estrutura de software capaz de criar uma ontologia global pela integração das existentes, automatizando todo o processo de construção da ontologia.

Como resultado concreto estes autores apresentam ainda um estudo de caso sobre o domínio do patrimônio cultural para mostrar uma aplicação real da abordagem, que permite integrar dados de várias ontologias como *WordNet*, CIDOC-CRM e EDM.

No trabalho *Architectural Heritage semantic data managing and sharing in GIS*, (Cerutti *et al.*, 2015), as autoras utilizam as tecnologias da Web Semântica para gerenciar diferentes tipos de dados (espaciais, não espaciais, imagens) em diferentes visualizações, em um banco de dados exclusivo, em especial no domínio arqueológico.

Dessa maneira, segundo as autoras, “alguns esquemas foram definidos e podem ser exportados para alcançar uma interoperabilidade efetiva de dados”, com a preocupação dos dados geográficos (Cerutti *et al.*, 2015, p. 2).

Como resultado prático, as autoras apresentam uma arquitetura de um Sistema de Informações Geográficas (SIG) para representar diferentes aspectos e pontos de vista, o uso de vários níveis de detalhe sobre estruturas construídas e a adição de relações com o contexto da paisagem.

A arquitetura baseia no uso de diferentes mapas espaciais, para que mapas

regionais digitais, ou desenhos gráficos vetoriais representando características de fachadas, possam ser visualizados separadamente enquanto permanecem armazenados em um banco de dados geoespacial. A abordagem semântica levou em consideração à construção de um modelo de referência baseado em outros modelos, como o CIDOC-CRM e *CityGML*, que são integrados à ontologia criada.

### **iii) Propostas de repositórios e outras experiências semânticas**

Seis trabalhos encontrados propuseram arquiteturas, sistemas de repositórios ou experiências relacionadas a museus e outros espaços de memória. Duas com propostas de ontologias com base no CIDOC-CRM e com estratégia de mapeamento ontológico. Uma delas abordando banco de dados relacional. Um trabalho propõe uma abordagem utilizando linguagem natural.

No trabalho *Generating knowledge graphs by employing Natural Language Processing and Machine Learning techniques within the scholarly domain*, (Dessì *et al.*, 2020), os autores propõem uma nova arquitetura com base em métodos de Processamento de Linguagem Natural e Aprendizado de Máquina para extração automática de entidades e relacionamentos de publicações de pesquisa para geração em grafos de conhecimento em larga escala.

A abordagem é suficientemente genérica para abarcar várias áreas do conhecimento e a estratégia apresentada é híbrida, combinando a extração automática de conhecimento com classificação supervisionada.

Essa abordagem é composta por 4 passos, 1) Extração de entidades e triplos, que explora um conjunto de várias ferramentas de PNL e aprendizado de máquina para extrair triplos de texto, 2) Refinamento de entidades, em que as entidades resultantes são fundidas e depuradas; 3) Refinamento de triplas, onde as triplas extraídas são mesclados e as relações são mapeadas para um vocabulário comum; e 4) Seleção de triplas, em que o conjunto de triplas confiáveis são incluídas na saída criando um grafo de conhecimento composto por triplas associadas a um bom número referências consistentes semanticamente. Como resultado da proposta, os autores apresentam um grafo de conhecimento científico incluindo cem mil triplas extraídos de mais de vinte se seis resumos de artigos dentro do domínio da Web Semântica.

Bordalo *et al.* (2021), no artigo *A Framework Design for Information Management in Heritage Science Laboratories*, propõem uma arquitetura de Sistema de

Informação para Laboratórios (LIMS) específico para laboratórios de Ciências do Patrimônio, apontando um estudo de caso do laboratório Hercules, da Universidade de Évora, Portugal.

O modelo não aborda uma ontologia em específico, prevê, porém, que a gestão dos bancos de dados será realizada a partir da utilização de estratégias de vocabulários para os padrões internacionais. O sistema proposto é composto de seis partes, gestão de investigação, gestão de equipamentos, agendamentos, inventários químicos, documentações de amostragem gráficas, gerenciamento de amostras e relatórios (Bordalo *et al.*, 2021).

No trabalho *Towards Big Religious Data: RESILIENCE Research Infrastructure for Data on Religion in the Digital Age* (Büchler *et al.*, 2020), os autores apresentam um *framework* denominado *Resilience*, que serve como infraestrutura para pesquisa na área da ciência da religião com uma série de estratégias voltadas para *Big Data*, o que os autores chamam de “transição digital”, o que pode gerar perda de dados no campo dos estudos religiosos.

Owais *et al.* (2017), no trabalho intitulado *A step towards the linking of the Culture Heritage in the Egyptian museums*, os autores contribuem para melhorar o mapeamento semântico de diferentes conjuntos de dados pertencentes aos museus egípcios, propondo e implementando uma arquitetura que enriquece esses conjuntos de dados através da aplicação de técnicas de dados abertos conectados, dependendo de duas metodologias diferentes de mapeamento. A proposta é baseada na integração de dois principais modelos, o CIDOC-CRM e o R2RML, sendo esse segundo, aplicado a bancos de dados relacionais (Owais *et al.*, 2017).

Como resultado concreto dos experimentos, os autores propuseram um mapeamento dos dados para RDF e conectaram o conjunto de dados enriquecido com a DBpedia, no qual perceberam lacunas importantes no domínio dos museus (Owais *et al.*, 2017, p. 6):

Descobrimos que há uma falta de semântica nos conjuntos de dados, pois nem todos os museus estão disponíveis no DBpedia. Os conjuntos de dados dos museus precisam ser limpos e convertidos em RDF para serem semanticamente enriquecidos. Além disso, os conjuntos de dados precisam estar vinculados à DBpedia e outros conjuntos de dados relacionados em todo o mundo para enriquecer o Patrimônio Cultural. Nossa avaliação mostra que a porcentagem de

ligação foi bastante boa.

Frommholz *et al.* (2014), no artigo “*Great War Stories Told by the People: Crowdsourced Cultural Heritage in Digital Museums*”, apresentam um sistema de informação para contar a história da Primeira Guerra Mundial com base nos relatos de pessoas. Utiliza classes e propriedades do CIDOC-CRM para integração e colaboração com outros sistemas. O sistema permite que as pessoas possam colocar objetos digitais relacionados à Primeira Guerra Mundial e contar suas próprias histórias gerando assim um grande museu digital colaborativo. Como resultado concreto, apresentam o portal do projeto que pode ser acessível pela Web em <https://www.worldwar1luton.com/> (Frommholz *et al.*, 2014).

#### **iv) Proteção, conservação e restauração de Patrimônio Cultural**

Seis trabalhos relacionados a proteção, conservação e restauração do patrimônio cultural foram encontrados. Quanto às questões relacionadas às estratégias utilizadas, notou-se três propostas de arquitetura sendo duas relacionadas a sistemas integradores, são cinco propostas de ontologias. Três abordagens que utilizam CIDOC-CRM em algum momento. Três trabalhos atuam de alguma maneira com bancos de dados relacionais.

Palomar *et al.* (2020), no artigo *An on-line platform to unify and synchronise heritage architecture information*, propõem uma plataforma interdisciplinar e colaborativa, denominada *BIMlegacy*, para sincronizar informações patrimoniais especificamente relacionadas a construções, monumentos e outros elementos do Patrimônio Cultural material.

Neste trabalho, a abordagem aplicada é a modelagem de sistemas de informação de construções patrimoniais (*Heritage Building Information Modelling – HBIM*). Estas construções referem-se a prédios, sítios, paisagens, monumentos, áreas e outros bens patrimoniais materiais que possuem valor arquitetônico e cultural e precisam ser levados em consideração em projetos interdisciplinares com documentação histórica, arqueológica e artística, bem como um estudo do cenário sociocultural (Palomar *et al.*, 2020).

Todo esse campo interdisciplinar gera uma série de desafios que são abordados pelos autores, e que a plataforma proposta visa auxiliá-los (Palomar *et al.*, 2020, p. 1):

Arqueólogos, arquivistas, estruturas engenheiros ou restauradores geralmente trabalham separadamente, o que significa que dados dispersos são produzidos, informações duplicadas são geradas e as contribuições de outras partes interessadas às vezes não são levadas em consideração. Por exemplo, o arqueólogo pode pesquisar patologias de pedra sem considerar o relatório anterior do arquiteto. Essas práticas de trabalho improdutivas causam desconfiança na gestão de projetos históricos e incertezas nos custos e cronogramas para os desenvolvedores de propriedades. Ineficiências nas intervenções de arquitetura de patrimônio - conservação, reabilitação, restauração e reconstrução - fazem com que a conservação de edifícios patrimoniais seja cara e tende a comprometer a preservação de sua cultura (Palomar *et al.*, 2020, p. 1).

A plataforma proposta é composta por um *plug-in* para o software *Revit*, este *plug-in* recupera as informações necessárias do modelo *Revit* e consome uma API Web para sincronizar os dados de um servidor de banco de dados relacionais (SQL) com os dados do arquivo *Revit*.

Esta plataforma também apresenta um portal Web para inserção, edição e consulta de dados em qualquer local gráfico. Seria orientado para partes interessadas não técnicas que geralmente não trabalham com BIM (por exemplo, historiador, historiador de arte, gerente de monumento) e para visitantes externos (Palomar *et al.*, 2020).

Em Niang *et al.* (2016), no trabalho *Ontology-Based Data Integration System for Conservation-Restoration Data (OBDIS-CR)*, os autores propõem uma abordagem semântica baseada na integração de fontes de dados heterogêneas no domínio da Conservação e Restauração do Patrimônio Cultural.

Primeiro, avaliamos o contexto e as necessidades atuais; as observações confirmaram a necessidade crítica de um sistema de integração de dados aproveitando a semântica. Segundo, descrevemos o processo que seguimos para construir uma ontologia dedicada a fornecer um entendimento unificado dos dados. Terceiro, descrevemos o sistema mediador baseado em ontologia projetado para consultar efetivamente as fontes distribuídas heterogêneas (Niang *et al.*, 2016, p. 1).

Os autores também trazem contribuições quanto ao problema da descrição de objetos digitais no campo da conservação e restauração em duas vertentes: a heterogeneidade das fontes de dados e a heterogeneidade das terminologias.

Como contribuição prática, os autores apresentam o Sistema de Integração de Dados Baseado em Ontologia (OBDIS) capaz de usar a ontologia desenvolvida como suporte semântico para integrar dados de conservação-restauração de diferentes instituições de patrimônio cultural

Niang *et al.* (Niang *et al.*, 2017), em *Supporting Semantic Interoperability in Conservation-Restoration Domain: The PARCOURS Project*. Apresentam um novo trabalho, que parece ser complementar ao anterior, onde propõem uma abordagem de integração semântica no domínio da conservação/restauração que utiliza ontologias, em especial a CIDOC-CRM.

Como na maioria dos problemas de integração encontrados na literatura, esse trabalho lida com três questões clássicas: a heterogeneidade de múltiplas bases de dados, falta de poder semântico dos bancos de dados, e heterogeneidade de terminologias.

Tal abordagem, direcionada ao projeto *Parcours* a construção de ontologia de domínio, mapeamento de ontologia com fontes de dados participantes e integração dos dados e fontes de dados para uma consulta semântica integrada.

O projeto *Parcours* propõe um modelo sistema de informação para fornecer acesso unificado a vários repositórios de patrimônio cultural com enriquecimento semântico e colaboração a nível de dados conectados mas sem precisar refatorar os bancos de dados já existentes como, por exemplo, os bancos de dados relacionais.

Em *Supporting semantic interoperability in conservation-restoration domain: The PARCOURS project* (Niang *et al.*, 2017), forma complementar a (Niang *et al.*, 2016), apresentam o mesmo projeto com resultados mais concretos.

No caminho metodológico apresentado pelos autores, destacam-se três estratégias básicas:

- (1) construir uma ontologia específica de domínio, basear-se em um entendimento unificado dos dados de conservação-restauração;
- (2) mapear a ontologia compartilhada para cada fonte de dados, permitindo que cada fonte participante gerencie seu próprio banco de dados semântico, consistindo em seus dados originais agora associados ao nível semântico; e
- (3) integrar os dados de todas as fontes, para consultá-los da mesma maneira homogênea (Niang *et al.*, 2017, p. 2).

As realizações apresentadas nos dois trabalhos foram conduzidas como parte do

projeto PARCOURS (<http://projet-parcours.eu/en/>), que conforme os autores buscou “desenvolver um sistema de informação capaz de fornecer acesso unificado a fontes de informação distintas, relacionadas ao campo do patrimônio cultural em geral e aos processos de conservação-restauração em particular” (Niang *et al.*, 2017).

A estratégia constou de uma única arquitetura de ontologia com o contexto de conservação-restauração específica do projeto. As fontes de dados envolvidas no processo de integração partilham o mesmo domínio de aplicação. Estas podem, portanto, ser referidas à mesma ontologia global.

A ontologia requer uma conceituação apropriada no domínio da aplicação, de forma que uma ampla gama de fontes de restauração e conservação possam ser mapeadas, sem a necessidade de que elas tenham sua ontologia local adequada. A ontologia resultante é uma extensão do CIDOC-CRM com concentração no processo restauração-conservação.

Na abordagem em questão, tratam-se de bancos de dados já existentes e que possuem atividade e dados de patrimônio cultural estabelecidos principalmente em Sistemas Gerenciados de Bancos de Dados Relacionais. Estes autores optaram de forma justificada por adotar estratégias que agreguem semântica aos dados sem que haja refatoração no conjunto existente.

A abordagem de ontologias locais em consonância com um vocabulário genérico e ligações com outras ontologias em um mundo aberto, que parece ser interessante em vários problemas – como no problema abordado nessa tese – é negadamente justificado no trabalho em questão.

Bannour *et al.* (2018), no trabalho *CRMCR - a CIDOC-CRM extension for supporting CR semantic interoperability in the conservation and restoration domain*, apresentam uma ontologia no domínio da conservação e restauração do patrimônio cultural. Trata-se de uma ontologia de domínio dedicada à conservação e restauração de objetos culturais. O CRMCR visa fornecer um entendimento unificado dos dados de conservação-restauração e é uma extensão da ontologia de alto nível CIDOC-CRM (Bannour *et al.*, 2018).

Como estudo de caso, estes autores avaliaram o problema da interoperabilidade envolvendo dois espaços da França especializados em conservação e restauração C2RMF e LRMH, que possuem técnicas e vocabulários diferentes para os objetos culturais, e teve

como estratégia utilizar o modelo CIDOC-CRM como referência, como pode ser visto abaixo:

Como o CIDOC-CRM é uma ontologia genérica, decidimos desenvolver um modelo ontológico de conservação-restauração como uma extensão do CIDOC-CRM. Isso significa que estendemos o escopo do CIDOC-CRM ao processo de conservação-restauração, definindo novos conceitos apropriados que se especializam nos do CIDOC-CRM e estabelecendo relações apropriadas entre eles (Bannour *et al.*, 2018, p. 3).

Estes autores também utilizaram a extensão CRM-sci, que comporta dados científicos, e da integração gerou-se uma nova ontologia constante de 93 conceitos 82 propriedades do CIDOC-CRM, 22 conceitos e 24 propriedades do CRM-sci, e 63 novos conceitos e 27 novas propriedades no domínio específico da pesquisa em questão.

Quattrini *et al.* (2017), *Knowledge-based data enrichment for HBIM: Exploring high-quality models using the semantic-web*, propõem uma abordagem para resolver o problema de interoperabilidade na modelagem de sistemas de informação de construções históricas (*Heritage Building Information Modelling – HBIM*).

O artigo em questão explora a representação de grandes quantidades de modelos de informação especializados dos sistemas de informação de construções (BIM) em arquivos abertos conectados (LOD) cuja interoperabilidade é garantida por meio de ontologias e dados em RDF, sendo essa abordagem aplicada ao patrimônio cultural.

Estes autores apresentam um fluxo de trabalho que explora as vantagens das plataformas BIM e tecnologias da Web Semântica, permitindo ao usuário consultar um repositório composto por dados HBIM semanticamente estruturados e ricos (Quattrini *et al.*, 2017).

A abordagem é composta de quatro etapas principais

(i) a primeira etapa consiste na modelagem de uma ontologia com as principais necessidades de informação para o domínio de interesse, fornecendo uma estrutura de dados que pode ser aproveitada para informar a fase de enriquecimento de dados e, posteriormente, para consultar os dados de forma significativa. (ii) Posteriormente, foi realizado o enriquecimento dos dados, criando um conjunto de parâmetros compartilhados refletindo as propriedades em nossa ontologia de domínio. (iii) Para estruturar os dados em um formato legível por máquina, foi necessária uma conversão de dados para representar o

domínio (ontologia) e analisar dados de edifícios específicos, respectivamente; esta etapa é obrigatória para reaproveitar os dados de análise junto com o modelo 3D, fornecendo ao usuário final uma ferramenta de consulta. (iv) Como uma etapa final em nosso fluxo de trabalho, desenvolvemos uma aplicação web de exploração de dados demonstrativa baseada no paradigma de navegação facetada e permitindo a exploração de metadados estruturados e visualização 3D. (Quattrini *et al.*, 2017, p. 129):

Cacciotti & Valach (2015), com o artigo *The MONDIS project: Semantic Web and the protection of historic buildings*, apresentam resultados da aplicação de ferramentas ontológicas no projeto *MONDIS*, que visa a preservação de construções históricas, buscando apoiar a investigação profissional de danos e o projeto adequado de possíveis soluções, através de um sistema de conhecimento para a documentação e análise de falhas e danos a monumentos.

Segundo estes autores, o projeto *MONDIS* visa apoiar a colaboração e a experiência intercâmbio e, ao mesmo tempo, intensificar a integração das fontes de conhecimento existentes no campo da proteção do patrimônio cultural. Mais especificamente, o *MONDIS* foca em: a) documentação de danos e falhas em objetos de patrimônio cultural imóvel; b) análise das dependências lógicas que ocorrem entre as falhas e suas causas; c) medidas de mitigação de danos e planos de intervenção; e d) fatores de risco e sua prevenção (Cacciotti & Valach, 2015).

A ontologia *Monument Damage Ontology*, proposta no projeto, representa um modelo conceitual no qual os elementos mais relevantes no domínio da proteção do patrimônio cultural são organizados em uma formalmente através de entidades denominadas grupos temáticos de conceitos.

Tal seção da ontologia modela a relação semântica que permite atribuir ao objeto de patrimônio cultural características estruturais e funcionais, como os tipos estruturais e funcionais que ajudam a individualizar o esquema de resistência da construção e sua funcionalidade. Este grupo também permite modelar ontologicamente um modelo físico da construção, determinando uma organização hierárquica das partes do objeto e seus materiais relacionados. A seção de eventos inclui todos os conceitos que representam ocorrências no tempo que podem influenciar as condições do objeto. Mais especificamente, esta parte do modelo ontológico replica semanticamente como eventos como desastres naturais (Cacciotti & Valach, 2015, p. 4) .

Kouis & Giannakopoulos (2014), no artigo *Incorporate Cultural Artifacts Conservation Documentation to Information Exchange Standards – The DOC-CULTURE Case*, apresenta uma abordagem baseada no estudo de caso do projeto *DOC-CULTURE* para padronizar a documentação dos métodos de Teste de Avaliação Não Destrutivos (NDT) e seus dados de saída através da exploração do CIDOC-CRM no contexto da conservação e restauração (Kouis & Giannakopoulos, 2014).

v) **Reflexões sobre conceitos, linguagens e ontologias**

Nessa Seção foram classificados os trabalhos que trazem discussões mais genéricas ou apresentam ontologias e linguagens gerais, uma vez que nenhum deles apresenta especificamente uma arquitetura ou sistema integrador. Há no entanto duas propostas de ontologia, sendo uma delas com mapeamento ontológico.

Veninata (2020), no trabalho *Inside the Meanings. The Usefulness of a Register of Ontologies in the cultural Heritage Sector*, discute sobre as diversas iniciativas realizadas na última década no campo do patrimônio cultural e afirma a falta de uma estrutura epistemológica consolidada na modelagem conceitual dos recursos de informação, apesar das inúmeras ontologias criadas a partir dos múltiplos projetos de publicação de dados abertos conectados (Veninata, 2020).

A autora reitera a dificuldade de compreensão completa de todas as ontologias disponíveis em relação à sua área de interesse e também de não haver um processo que permita uma avaliação de uma forma suave e sistemática e confiável da sua capacidade representativa, bem como seu grau de interoperabilidade semântica.

Outra reflexão importante da autora é sobre o reaproveitamento de ontologias. Isto é, antes de criar novas soluções ontológicas, ter como boa prática – inclusive é recomendada por grande parte da literatura sobre padrões – avaliar e realizar o reaproveitamento de ontologias disponíveis no domínio do conhecimento, preferencialmente fazendo referência a ontologias padrão ou, de forma complementar, a ontologias de domínio reconhecidas, documentadas e mantidas por grupos e instituições conceituados.

Mountantonakis & Tzitzikas (2019), em *Large-Scale Semantic Integration of Linked Data: A Survey*, apresentam um artigo que analisa o trabalho realizado na área de integração de dados conectados, identificando os principais atores e casos de uso, analisam e fatoram o processo de integração de acordo com as várias dimensões e,

também, discutem os métodos usados em cada etapa com ênfase naqueles que podem ser usados para integrar vários conjuntos de dados.

Esse trabalho reforça as dificuldades em dados abertos conectados, em especial no que se refere à integração e à interoperabilidade entre as diversas fontes abertas.

A integração é um problema inerentemente difícil por vários motivos. Inicialmente, conjuntos de dados são fornecidos produzidos e gerenciados por diferentes organizações, de acordo com conceituações, modelos, esquemas e formatos com base em suas necessidades e escolhas, e esses dados são armazenados em diferentes locais e sistemas. Outra dificuldade é que as mesmas entidades ou relacionamentos do mundo real são referidos com nomes ou URIs diferentes e em diferentes idiomas naturais; os sinônimos e homônimos das línguas naturais perplexam ainda mais na identificação e na correspondência. Além disso, muitas vezes não há dados suficientes para a correspondência automatizada, porque os conjuntos de dados contêm informações complementares sobre as mesmas entidades (Mountantonakis & Tzitzikas, 2019, p. 103:3).

Os autores defendem ainda que a integração não é uma tarefa única “no sentido de que tudo pode mudar com o passar do tempo, por exemplo, esquemas, ontologias, convenções de nomenclatura, *etc.*” (Mountantonakis & Tzitzikas, 2019, p. 103:3). Os autores sugerem 6 problemas principais de integração

a) Autoridades diferentes. Os conjuntos de dados são produzidos, mantidos ou gerenciados por diferentes organizações em diferentes formatos, esquemas, modelos, locais, sistemas e licenças; b) Nomenclatura. As mesmas entidades ou relacionamentos do mundo real são referidos com diferentes URIs e nomes e em diferentes idiomas naturais, enquanto os idiomas naturais têm sinônimos e homônimos que dificultam a conexão automática; c) Complementaridade. Os conjuntos de dados geralmente contêm informações complementares; dois conjuntos de dados sobre o mesmo domínio, cada um modelando um aspecto diferente do domínio; d) Dois conjuntos de dados sobre o mesmo domínio, cada um modelando um aspecto diferente do domínio; e) Conceitualizações diferentes. Os conjuntos de dados sobre o mesmo domínio podem seguir diferentes conceituações (ou decisões de modelagem) do domínio; f) Evolução. Tudo muda: o mundo e as ontologias integração (Mountantonakis & Tzitzikas, 2019, p. 103:6).

Rubhasy *et al.* (2014), em *Management and Retrieval of Cultural Heritage*

*Multimedia Collection Using Ontology*, propõem uma abordagem baseada em ontologia na gestão e recuperação de coleções multimídia de patrimônio cultural em contraponto aos mecanismos de busca por palavra-chave como *Google* e *Yahoo*.

O objetivo é permitir a integração de diferentes tipos de coleções patrimoniais e recuperar coleções patrimoniais relevantes para uma determinada consulta em qualquer tipo de acervo. Para isso, a ontologia do patrimônio cultural é desenvolvida com base na Classificação do Patrimônio Mundial fornecida pela UNESCO / WHC. A ontologia também incluiu classificação de mídia para acomodar a natureza heterogênea da coleção multimídia. A tarefa de recuperação é feita usando a linguagem de consulta SPARQL. O resultado mostra que todo tipo de mídia de patrimônio relevante pode ser recuperado e o número de mídia não relevante recuperada é relativamente baixo (Rubhasy *et al.*, 2014, p. 255).

A ontologia proposta pelos autores apresenta uma organização do patrimônio cultural em 101 classes, 7 propriedades de objetos e 11 propriedades de dados e notadamente dá ênfase aos aspectos multimídia dos materiais envolvidos, tendo uma classe específica para mídia (Media) composta das subclasses *Animation*, *Audio*, *Image*, *Text* e *Video*.

O foco nos objetos de multimídia, segundo os autores, possui organização conforme abaixo:

A ontologia CH é projetada para descrever os conceitos (*classes*) e relações entre conceitos. Existem quatro componentes principais na ontologia: classe, propriedade dos dados, propriedade do objeto e indivíduos. A classe de ontologia é usada para representar objetos de patrimônio cultural, enquanto a propriedade de dados é usada para fornecer atributos de uma classe. A fim de fornecer relacionamento entre as classes, a propriedade do objeto é usada. O último componente, os indivíduos representam os objetos reais do domínio (Rubhasy *et al.*, 2014, p. 257)

Como resultado deste trabalho, os autores apresentam um conjunto de objetos de patrimônio cultural recuperados a partir de consultas SPARQL, mostrando que a recuperação de multimídia relacionada a esses objetos foi efetiva e com taxa baixa de perda de relevância.

Daquino & Tomasi (2014), no artigo *Ontological Approaches to Information Description and Extraction in the Cultural Heritage Domain*, apresentam uma abordagem para descrição e extração de informações de documentos de texto completo

com marcação através da utilização de ontologias (*PROles* e *HIco*) para descrever informações de forma semântica.

Neste trabalho, o conceito de proveniência é explorado como forma de criar declarações autorizadas. Os autores argumentam que, em geral, este problema é abordado por técnicas centradas no documento, e isso vai em caminho oposto ao da proposta da Web Semântica que prevê uma rede de dados. A riqueza dos documentos de texto completo seriam preservadas através da marcação utilizando as ontologias propostas pelos autores (Daquino & Tomasi, 2014).

Em Meghini *et al.* (2017), no trabalho intitulado *ARIADNE: A Research Infrastructure for Archaeology*, os autores discutem que arquivos digitais e serviços de dados tornaram-se pilares importantes para a ciência moderna, que se torna à cada dia mais colaborativa, e que a comunidade de pesquisa arqueológica foi pioneira na adoção de ferramentas digitais para aquisição, organização, análise e apresentação de resultados de pesquisas de projetos individuais (Meghini *et al.*, 2017).

O fornecimento de estratégias e serviços para compartilhamento, descoberta, acesso e reuso de dados, entretanto, não acompanhou o mesmo crescimento, assim propõem um projeto denominado *ARIADNE*, que fornece uma Infraestrutura de pesquisa avançada para redes de conjuntos de dados arqueológicos na Europa.

Esta rede financiada pela UE desenvolveu uma infraestrutura eletrônica que permite que os provedores de dados registrem e forneçam acesso aos seus recursos (conjuntos de dados, coleções) por meio do portal de dados da *ARIADNE*, facilitando a descoberta, o acesso e outros serviços nos recursos integrados. Este artigo descreve o cenário atual de repositórios de dados e serviços para arqueólogos na Europa e os problemas que dificultam a interoperabilidade entre eles. Os resultados das pesquisas da *ARIADNE* sobre as expectativas e os requisitos dos usuários também são apresentados. A seção principal do documento descreve a arquitetura da infraestrutura eletrônica, serviços principais (registro, descoberta e acesso a dados) e vários outros serviços existentes ou experimentais (Meghini *et al.*, 2017, p. 2).

Neste artigo, os autores discutem também os processos de avaliação contínua da integração de dados, e apresentam um conjunto de lições aprendidas no processo, descrevendo as perspectivas de um maior envolvimento da comunidade de pesquisa arqueológica no compartilhamento de dados por meio da plataforma *ARIADNE*.

#### vi) **Sistemas de Recomendação e espaços inteligentes**

Nessa secção foram classificados nove trabalhos com propostas de sistema ou arquitetura inteligente. Todos propõem ontologia, sendo que dois abordam especificamente o modelo de referência CIDOC-CRM e um o modelo EDM. Três trabalhos voltam seu foco para arquivos abertos conectados, e um propõe um sistema integrador, e um realiza um comparativo entre sistemas de recomendação.

Nafis *et al.* (2020) no artigo *Towards a semantic recommender system for cultural objects: Case study Draa-Tafilalet region*, apresentam um estudo comparativo de Sistemas de Recomendação (RS) semânticos existentes no campo do patrimônio cultural, a fim de extrair uma visão completa de um Sistema de Recomendação para o patrimônio cultural científico da região de Drâa-Tafilalet no Marrocos (Nafis *et al.*, 2020).

Nesse estudo, os autores apresentam ainda a estrutura de classificação do patrimônio científico cultural de Marrocos, além de uma breve descrição de estratégias e Sistemas de Recomendação. Por fim, apresentam as funcionalidades básicas de um Sistema de Recomendação em construção que tem como base uma ontologia estendida do CIDOC-CRM especificamente para o caso em estudo.

Smirnov *et al.* (2017), com o trabalho *Context-Based Infomobility System for Cultural Heritage Recommendation: Tourist Assistant – TAIS*, apresentam um sistema que intitulam infomobility – Tourist Assistant – para apoiar turistas em uma região.

O sistema resolve as duas principais tarefas relacionadas ao conceito de *infomobilidade*: análise das ações do usuário, revelação de preferências, recomendação do patrimônio cultural com base nas preferências e na situação atual da região e fornece ao usuário os meios de transporte possíveis para alcançar o patrimônio cultural correspondente. As informações multimídia sobre o patrimônio cultural incluem: descrição e imagens de texto e vídeos extraídos de fontes acessíveis da Internet – como *Wikipedia*, *Wikivoyage*, *Panoramio* e *YouTube* – (Smirnov *et al.*, 2017, p. 2).

O sistema proposto reúne um conjunto de serviços que oferecem possibilidades para organizar a troca de informações semânticas entre esses serviços usando a abordagem baseada em ontologia.

Estes serviços compartilham informações para a solução conjunta da tarefa do turista com o espaço inteligente. Esta informação forma um espaço cultural que é um

modelo de espaço físico onde o turista está. A lista do patrimônio cultural fornecido ao turista é solicitada pelo serviço de recomendação especial que implementa a classificação com base na técnica de filtragem colaborativa.

As recomendações do sistema apresentado são baseadas nas classificações definidas pelos turistas que fazem uso deste. Para fornecer ao turista diferentes meios de transporte, foi desenvolvido o serviço de transporte especial. Este serviço aciona um táxi, encontra possibilidades de compartilhamento de passeios ou calcula rotas multimodais de transporte público com base nas preferências do turista.

O artigo descreve a arquitetura do sistema baseado em serviços, desenvolvimento de serviços, suas ontologias, implementação e avaliação. O protótipo do sistema desenvolvido está acessível para download no mercado do *Google Play* para usuários de dispositivos *Android*.

Korzun *et al.* (2017), com o artigo *Semantic Infrastructure of a Smart Museum: Toward Making Cultural Heritage Knowledge Usable and Creatable by Visitors and Professionals*, apresentam o conceito de museu inteligente, apoiando-se na Internet das Coisas (IoT) e na proposta de Smart Spaces.

Nesta proposta, os serviços não se limitam à provisão direta de descrição de exposições com base em registros, como acontece nos sistemas tradicionais de informação de museus (Korzun *et al.*, 2017).

O conceito é baseado em serviços com alto nível de inteligência, quando fontes históricas adicionais podem ser usadas para enriquecer semanticamente a coleção do museu, incluindo o conhecimento adquirido dos visitantes e profissionais do museu. Um museu se torna um espaço cultural em que sua camada semântica torna o conhecimento utilizável e passível de criação por visitantes e profissionais. Nosso foco de pesquisa está na aplicação desse conceito ao estudo de caso do *Museu de História da Universidade Estadual de Petrozavodsk* (Korzun *et al.*, 2017, p. 1).

Neste trabalho foi criado um protótipo de conceito como uma fase de desenvolvimento obrigatória da engenharia de sistemas complexos para analisar a necessidade, viabilidade e abordagem técnica. O protótipo de conceito segue a abordagem de espaços inteligentes para ambientes de IoT e define soluções de *design* para criar uma infraestrutura semântica que transforma um determinado museu em sua variante inteligente (Korzun *et al.*, 2017).

Su *et al.* (2019), em *An Edge Intelligence Empowered Recommender System Enabling Cultural Heritage Applications*, apresentam uma arquitetura de Big Data para aplicações no domínio patrimônio cultural.

Além de consultar, navegar e analisar conteúdos culturais provenientes de repositórios distribuídos e heterogêneos, propomos uma nova estratégia de recomendação centrada no usuário para sugestões de itens culturais. Apesar de centralizar as operações de processamento na nuvem, a visão da inteligência de ponta foi explorada por meio de um aplicativo móvel (*Smart Search Museum*) para realizar buscas semânticas e inferência baseada em aprendizado de máquina, a fim de sugerir museus, além de outros itens de interesse para os usuários quando estes estão visitando uma cidade, explorando técnicas de recomendação em conjunto e instalações de inteligência artificial de ponta (Su *et al.*, 2019, p. 2).

Essa proposta atua no campo dos sistemas de recomendação utilizando inteligência artificial para realizar descobertas e melhorar sugestões para museus e outros espaços de patrimônio, fornecendo uma estratégia de *ranking* em um arranjo que leva em conta pontos de interesse – museus, lugares, *etc* –, dados abertos conectados de instituições e enciclopédias e comentários de usuários de mídias sociais.

Como contribuição prática, os autores apresentam um aplicativo móvel (*Smart Search Museum*) utilizado para sugerir museus e outros itens de interesse dos usuários quando visitam uma cidade. Esse aplicativo utiliza a abordagem proposta para a recomendação de conteúdos culturais, no qual o perfil do usuário é rastreado para pré-filtrar com precisão os dados usando o aprendizado de máquina, onde posteriormente, na fase de classificação, calcula três pontuações – popularidade, humor e interesse. Cada uma dessas pontuações é calculada acessando os dados coletados pelas redes sociais e pelo próprio sistema através do histórico de usuários anteriores.

R. Martini *et al.* (2016) no artigo *Formal Description and Automatic Generation of Learning Spaces based on Ontologies*, propõem uma arquitetura genérica, denominada *CaVa*, para a criação de espaços de aprendizagem virtuais, com foco no patrimônio cultural, a partir de documentos de instituições culturais (Martini *et al.*, 2016).

Conforme os autores, a proposta consiste em criar automaticamente ambientes virtuais de aprendizagem baseados em ontologias para transmitir materiais de aprendizagem relevantes para o ambiente virtual de aprendizagem, onde as ontologias são

usadas para representar os conceitos-chave e as relações semânticas em um formato compreensível pelo usuário e pela máquina.

A arquitetura proposta é construída em três módulos (Martini *et al.*, 2016, p. 240) :

Módulo A é o repositório da instituição (Documentos da Instituição); compreende os documentos reais, o repositório de banco de dados, a função ingest implementada como *Sistema de Ingestão de Documentos* (DIS), para fazer o upload dos documentos e auxiliar no gerenciamento dos dados, e a ontologia principal. O Módulo B é composto de dois componentes que são o mecanismo de especificação e as Especificações LS escritas em uma DSL projetada especificamente para este projeto. Observe que o mecanismo de especificação é gerado automaticamente a partir da gramática DSL. O Módulo C inclui os arquivos de script (scripts LS) e o navegador, onde o componente final, *Learning Space*, está localizado.

Bertola & Patti (2015), no trabalho intitulado: *Ontology-based affective models to organize artworks in the social semantic web*, propõem um método com base na análise de sentimento para a exploração e sugestão de conteúdos através de comentários textuais deixados por visitantes em acervos digitais de patrimônio cultural.

A estratégia proposta utiliza de ontologias de domínio para a construção de uma estrutura conceitual do espaço emocional dos usuários com base em um modelo psicológico das emoções humanas para reconhecer o impacto emocional das obras.

A ontologia é inspirada no modelo circunplexo de Plutchik, um modelo psicológico bem fundamentado das emoções humanas. Os usuários podem se envolver na criação do espaço emocional, por meio de uma interface gráfica interativa. O desenvolvimento de tal espaço semântico permite novas formas de acessar e explorar coleções de arte (Bertola & Patti, 2015, p. 140) .

Por ser esse modelo de categorização de sentimentos e detecção de emoção codificadas em uma ontologia OWL, a abordagem, segundo os autores, possui a dupla vantagem de permitir raciocínio tratável sobre emoções detectadas e obras de arte relacionadas e também de promover a interoperabilidade e integração de ferramentas e tecnologias da Web Semântica e Dados Abertos Conectados, através do *framework* proposto denominado *ArsEmotica* (Bertola & Patti, 2015).

Benouaret & Lenne (2015), em *Personalizing the Museum Experience through*

*Context-aware Recommendations* (Benouaret & Lenne, 2015), abordam os sistemas de recomendação como possibilidade de apresentar melhores soluções acesso a coleções em museus.

Os visitantes dos museus costumam ficar maravilhados com as informações disponíveis no espaço que estão explorando. Portanto, encontrar obras de arte relevantes para ver em um período limitado de tempo é uma tarefa crucial. (Benouaret & Lenne, 2015, p. 743)

A proposta prevê a criação de uma estrutura de recomendação que consiste em uma combinação de uma abordagem semântica para a representação do domínio do museu, utilizando ontologias e tesouros com um método de filtragem colaborativa que recebe um enriquecimento semântico.

Em uma visão geral do sistema, após o processamento inicial, há uma pós-filtragem contextual que permite a geração de um roteiro personalizado com base no ambiente físico, na localização dos visitantes e no tempo que desejam permanecer no museu (Benouaret & Lenne, 2015).

Varfolomeyev *et al.* (2015), com o artigo *Smart space based recommendation service for historical tourism*, propõem um serviço inteligente de recomendação com base na abordagem *smart spaces*.

Essa abordagem emprega métodos de projeto de serviço orientados a eventos e multiagentes com um *hub* central de informações. Cada um dos agentes troca informações por meio de ações de assinatura e publicação. Uma característica importante da tecnologia de *smart spaces* é o uso de formatos ontológicos para manter informações semanticamente descritas e inferir novos conhecimentos a partir delas. Esta tecnologia é examinada com base em dados da história de Latgale (Varfolomeyev *et al.*, 2015, p. 85).

A estratégia empregada pelos autores consiste de três partes: 1) Serviços móveis: os serviços são construídos e fornecidos para usuários móveis e usando dispositivos móveis de pessoas que transitam pelos espaços; 2) Múltiplas fontes de informação: usuários móveis equipados com dispositivos móveis pessoais, dispositivos incorporados, que representam o mundo físico com muitas "coisas" circundantes e um conjunto rico de vários serviços da web disponíveis; e 3) Troca de informações interoperáveis: os participantes do espaço inteligente são heterogêneos, desde servidores poderosos a sensores minúsculos. Tudo isso resulta em uma rede semântica conforme os autores

descrevem abaixo

O principal componente da arquitetura é o *Semantic Information Broker* (SIB), que é uma base de conhecimento central em um servidor que usa formatos semânticos RDF / OWL para os agentes trocarem e compartilharem informações. Esses agentes de software também são chamados de processadores de conhecimento (KPs), que são executados em dispositivos do ambiente. Alguns deles atuam em nome de fontes de dados, recursos e serviços externos. A troca de informações entre KPs e SIB pode ser orientada a eventos por meio de operações de publicação / assinatura (Varfolomeyev *et al.*, 2015, p. 89).

Benedusi *et al* (2015) apresentam o trabalho *An Associative Engines Based Approach supporting Collaborative Analytics in the Internet of Cultural Things*, que consiste em um sistema integrado que pretende dar suporte em fornecimento, promoção, publicação, gerenciamento e uso de bens do patrimônio cultural.

O sistema em questão (CHIS-BI) com dados combina técnicas de BI, internet das coisas e computação em nuvem é aplicado para a integração, processamento analítico e descoberta de dados de diferentes tipos de populações humanas e visitas *ad-hoc* no em exposições, museus, e eventos culturais, e também em espaços territoriais, como centros históricos, áreas turísticas, e explorações virtuais baseadas em dados dispostos na nuvem (Benedusi *et al.*, 2015).

### 3.6 Artigos excluídos após a leitura

Após a leitura completa dos artigos conforme critério de inclusão/exclusão definido (p. 80), três artigos foram excluídos, conforme o quadro abaixo, por não possuírem técnicas ou estratégias de integração ou interoperação entre repositórios e tecnologias.

**Quadro 4 – Critério C4 - possuir técnicas ou estratégias de integração e/ou interoperação entre repositórios ou tecnologias**

1	Web of science	<i>The Iccu's Contribution To The Development Of A Digital Infrastructure For Cultural Heritage</i>	Butto, Simonetta	2014
2	IEEE	<i>Object-mediated innovation: Case study of adventures in the National Palace Museum in Taiwan</i>	Chen <i>et al.</i>	2014
3	ACM	<i>Ontologies for the Real Property Domain</i>	Shi & Roman	2018

Fonte: elaborado pelo autor

No trabalho *Ontologies for the Real Property Domain*, (Shi & Roman, 2018), os autores propõem um trabalho voltado para o domínio das propriedades imobiliárias, sem conexão direta com o patrimônio cultural.

No artigo *Object-mediated innovation: Case study of adventures in the National Palace Museum in Taiwan*, (H. Chen *et al.*, 2014), os autores apresentam a experiência da produção do filme *Aventuras no Museu do Palácio Nacional em Taiwan*, com foco exclusivo na produção criativa e patrimônio cultural, e como isso pode contribuir na gestão do conhecimento. Não há ligação nem contribuição direta com o tema pesquisado nesta tese.

Buttò (2014), no artigo *The ICCU'S Contribution to the Development of a Digital Infrastructure for Cultural Heritage*, apresenta os principais projetos e atividades desenvolvidas pelo ICCU a nível nacional e internacional os quais investigam questões políticas e técnicas da relação entre o setor de Patrimônio Cultural e as infraestruturas de investigação na Itália.

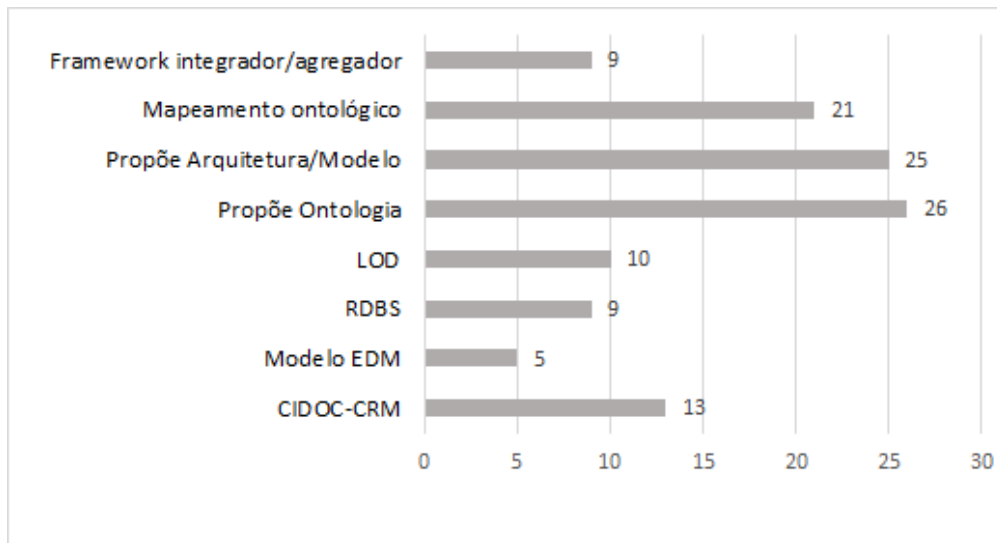
### **3.7 Sumário do capítulo**

O presente capítulo, em forma de revisão sistemática da literatura, teve o intuito de apresentar o estado da arte na temática desta tese, discutindo as práticas metodológicas e resultados de artigos publicados nos últimos anos em quatro grandes bases. Faz-se necessário destacar que muitos trabalhos não recuperados pela busca desta revisão sistematizada podem apresentar contribuições importantes nas áreas interdisciplinares relacionadas à investigação, o enfoque, no entanto, se deu em torno das bases selecionadas.

49 trabalhos foram selecionados para leitura final, mas 3 foram excluídos após a leitura. Portanto, 46 trabalhos entre 2014 e março de 2021, quando foi realizada pela última vez, a consulta nas bases de dados selecionadas para essa revisão sistematizada.

Quanto aos critérios de avaliação, como é possível visualizar na Figura 12, a maior parte propõe ontologias e soluções arquiteturais. Praticamente um terço dos trabalhos faz referência ao modelo CIDOC-CRM. Apenas cinco ao modelo EDM. Dez trabalhos analisados possuem foco voltado para os arquivos abertos conectados.

**Figura 12 – Critério de avaliação dos artigos**



Fonte: Elaborada pelo autor

O total de trabalhos recuperados (72) que atendem ao conjunto de palavras-chave estabelecido evidencia que há uma quantidade significativa de pesquisas a investigarem o problema da interoperabilidade semântica, envolvendo o patrimônio cultural. As estratégias apresentadas nos trabalhos que foram analisados nesta revisão, reforçam o uso de ontologias como elemento central e os Arquivos Abertos Conectados como grande oportunidade.

No próximo capítulo será apresentado um modelo semântico, aberto e conectado voltado para centros de memória para Estado de Goiás, que é a proposta desta tese para o problema da interoperabilidade.

## **CAPÍTULO IV. UM MODELO SEMÂNTICO, ABERTO E CONECTADO PARA O PATRIMÔNIO CULTURAL**

### **4.1 Introdução**

Este capítulo tem por finalidade apresentar a proposta desta tese, qual seja um modelo semântico, aberto e conectado voltado para centros de memória, com base na amostragem estabelecida (conforme Seção 5.3.1 p. 154) que leva em consideração as características específicas dos espaços de memória do Estado de Goiás. Tem ainda como pressuposto o potencial de interoperabilidade entre os acervos disponíveis nos diversos espaços de memória do Estado de Goiás.

Retome-se o texto introdutório desta tese para reforçar que o problema central abordado é o da interoperabilidade semântica de patrimônio cultural, especificamente no que se refere ao patrimônio material e imaterial do Estado de Goiás, tendo como objeto do estudo os espaços de memória da Secretaria de Estado da Cultura, mais precisamente os museus, bibliotecas, arquivos e galerias (GLAMs) da Superintendência de Patrimônio desta secretaria.

É pretendido, sobretudo, apresentar como proposta um modelo de sistema de informação que possa auxiliar os curadores digitais nas ações de acesso e preservação do patrimônio para atender a Sociedade em Rede e a exigência cada vez mais urgente de digitalização de pessoas, coisas e fenômenos, permitindo haver um intercâmbio entre os acervos.

Com um olhar mais pontual ao que concerne à interligação entre repositórios em um nível de contexto, esse modelo de Sistema de Informação pretende possibilitar aos curadores um modo de ampliar as informações digitais dos objetos sob a sua guarda. A proposta de interoperabilidade é constante de um conjunto de práticas metodológicas, tecnológicas e conceituais para a interoperabilidade semântica em 4 dimensões: Pessoas, Tempo, Local e Fenômeno.

Para tanto, neste capítulo são apresentados a estratégia escolhida e as técnicas utilizadas para a construção desta proposta, os esquemas de atividades e um conjunto arquitetural que prevê sua modelagem conceitual e tecnológica.

### **4.2 Estratégia utilizada**

O enquadramento teórico e o estado da arte (CAPÍTULO II e CAPÍTULO III)

permitem evidenciar que vários trabalhos abordam o problema da interoperabilidade semântica de patrimônio cultural, e que a estratégia mais comum encontrada é a utilização de ontologias nos mais variados níveis e com aplicações distintas. É recorrente encontrar nestes trabalhos muitos métodos e técnicas voltados para um repositório em específico ou para mapeamentos entre os domínios do patrimônio cultural.

As técnicas apresentadas por Freire *et al.* (2018) utilizam estratégias de agregação, que consiste em centralizar os metadados de várias fontes de dados em um nível ontológico mais alto. Já em Capodiece *et al.* (2016), a estratégia observada passa por criar uma estrutura ontológica para interoperação de fontes de dados estruturadas e semiestruturadas heterogêneas através da técnica de integração de esquemas (*schema-matching*), isto é, também utiliza mapeamento porém entre os domínios presentes em cada esquema.

Uma variação dessa técnica também é encontrada em Owais *et al.* (2017) que amplia no mapeamento semântico diferentes conjuntos de dados pertencentes aos museus egípcios, propõe e implementa uma arquitetura que enriquece esses conjuntos de objetos através técnicas de dados abertos conectados.

Na mesma direção Niang *et al.* (2016) e Niang *et al.* (2017) propõem uma abordagem de interoperabilidade com integração por uma ontologia de domínio para a restauração e conservação do Patrimônio Cultural e, de forma complementar, apresentam uma estratégia de mapeamento entre a ontologia compartilhada e cada repositório existente, criando entre eles uma associação no nível semântico.

Outros trabalhos, Palomar *et al.* (2020), evidenciaram ainda a busca pela construção de repositórios voltados para os dados abertos conectados, utilizando-se do RDF como modelo central para os processos de armazenamento e recuperação da informação, mais especificamente com a utilização da linguagem SPARQL. Palomar *et al.* (2020) utilizam técnicas de arquivos abertos conectados em uma plataforma interdisciplinar onde são sincronizadas informações patrimoniais de construções históricas.

Daquino *et al.* (2017) utilizam também as técnicas de mapeamento para criar um banco de dados RDF que permite a conversão de elementos do CIDOC-CRM para os padrões de conteúdos em duas ontologias de fotografias da Itália (Scheda OF e Scheda OA).

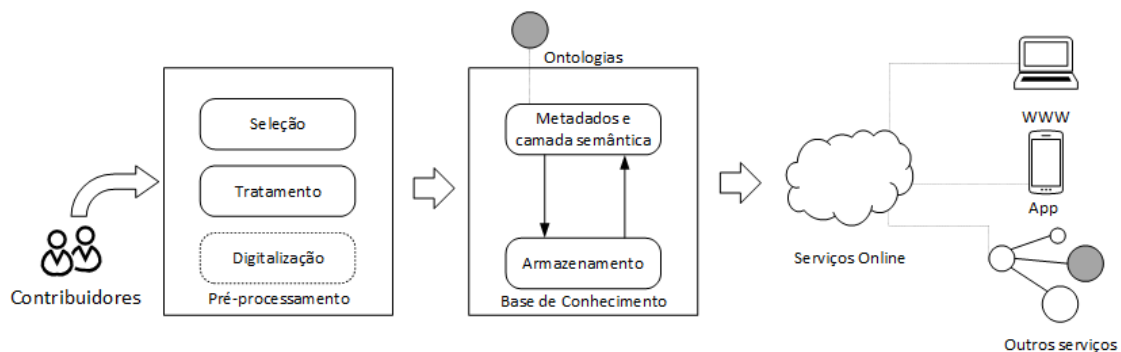
Abu Musa *et al.* (2020), por sua vez, utiliza técnicas que consistem de uma ontologia criada para realizar inferências sobre classificações de objetos culturais por palavra-chave (correspondência textual) e posteriormente por correspondência semântica através de sinônimos. Na mesma direção, Drakopoulos *et al.* (2019) apresentam uma estratégia que aplica a descrição de objetos em JSON, com agrupamentos em níveis distintos, anotações semânticas e *links* que podem ser compartilhados de forma interoperável com base em um conjunto de métricas advindas da probabilidade e teoria dos conjuntos.

A estratégia construída nesta tese passa por todos esses trabalhos encontrados no estado da arte. O ponto central, no entanto, é a concepção de um sistema de informação a partir de um conjunto de ontologias de domínio que permite uma diversidade de organização entre os repositórios, uma descrição mínima dos objetos digitais pertencentes a estes, e um intercâmbio entre eles através de uma estrutura interoperável composta por uma ontologia comum em quatro dimensões: Pessoas, Tempo, Local e Fenômeno usando como base os modelos mais simplificados como RDF/S e *Dublin Core*.

### 4.3 Ponto de partida: arquitetura genérica

Tendo em vista a estratégia escolhida para esta tese, é preciso definir como ponto de partida do modelo proposto um esquema genérico de atividades para funcionamento de um centro de memória que pode ser visualizado na Figura 13. As etapas iniciais do processo de transformação digital dos objetos são levadas em consideração, bem como uma estrutura ontológica em alto nível é necessária para fornecer serviços por meio de uma nuvem interoperável.

Figura 13 – Arquitetura geral da proposta



Fonte: Adaptada da primeira proposta arquitetural em A. Filgueiras, *et al.* (2020).

Em uma abordagem *top-down* (do geral para o detalhe), essa arquitetura

apresenta um alto nível de abstração, entendendo como pré-processamento todas as etapas anteriores às atividades de curadoria digital relacionadas à manutenção das informações dos acervos presentes nos espaços de memória. Parte-se então de um nível mais alto e abstrato em direção a um nível mais baixo e menos abstrato, que é no qual se pretende constituir os elementos específicos de interoperabilidade, objetivo maior deste modelo.

Os *Contribuidores* representam os diversos atores do sistema: estudantes, professores, técnicos e membros da comunidade que colaboram para essa produção. Os processos de seleção e tratamento – realizados pelos contribuidores – também são fenômenos que precisam ser levados em consideração para a produção do conhecimento, mas, nessa visão, estão na fronteira externa do sistema, ou seja, servem como entrada que serão processadas internamente e, portanto, não pressupõem regras ou restrições diretas sobre estes processos, requerendo apenas os objetos digitais resultantes destes.

São atividades que atendem às premissas estabelecidas na Seção 2.6, que prevê as ações com quatro eixos, conforme apresentadas na Figura 14, como um dos caminhos possíveis para a interoperabilidade entre repositórios do Patrimônio Cultural. Entretanto, ao confrontar a proposta com as definições iniciais estabelecidas no Projeto Centro de Memória de Goiás (Seção 2.6), percebeu-se que para abarcar todas as realidades dos espaços de memória observados no recorte desta tese, faz-se necessário a ampliação para acomodar os acervos em uma divisão mais orientada às suas características. Assim, ao invés de 4 eixos propõe-se uma abordagem com 5 tipos de artefatos: *arquivístico/documental*, *imagético/sonoro*, *arqueológico*, *museológico* e *bibliotecário*, o que atende aos eixos e ainda amplia as possibilidades para a classificação dos artefatos em cada espaço.

Dessa forma, independentemente das especificidades dos espaços, acervos podem dialogar dentro de uma perspectiva em que um objeto esteja associado a um ou mais eixos informacionais de um GLAM típico (Galeria, Biblioteca, Arquivo ou Museu).

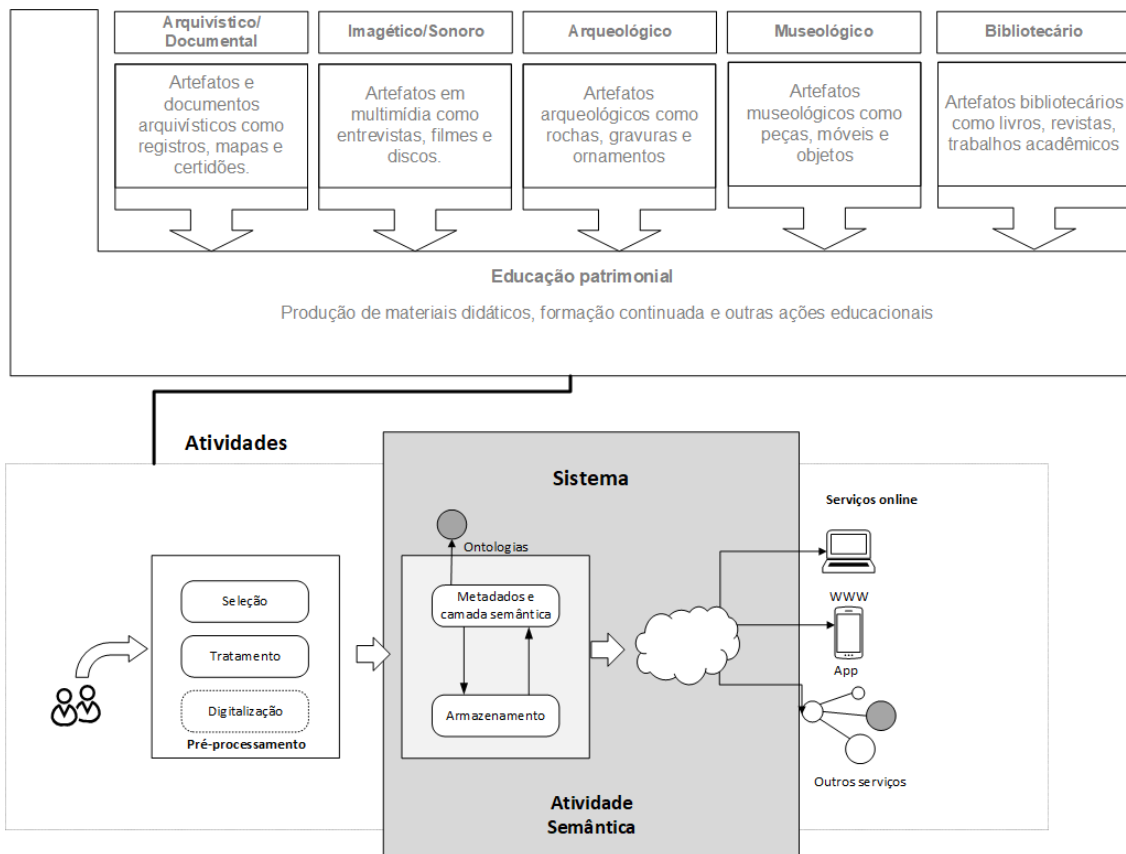
Nesta visão, as atividades comuns dos GLAMs (conforme os processos apresentados na Seção 4.3 p. 123) geram fluxos de informação que precisam ser processados, levando em consideração as características de cada acervo, mas respeitando as definições de contexto que são comuns a estes. A visualidade no nível dos repositórios e acervos são detalhadas nas seções posteriores deste capítulo

Tendo como base os eixos informacionais dos GLAMs, os objetos digitais,

provenientes da digitalização de documentos e outros elementos de mídia, permite que os recursos possíveis em áreas interdisciplinares transitem entre os espaços e sejam utilizados para a produção de conteúdos que serão explorados em vários contextos, como, por exemplo na educação patrimonial.

A Figura 14 apresenta esse processo e sua proposta de associação com a arquitetura geral definida em três atividades: operacional, semântica e de serviços.

**Figura 14 – Esquema de atividades associadas à arquitetura geral**



Fonte: Elaborada pelo Autor

O ponto de partida dessa proposta que apresenta, portanto, é a concepção de um sistema para preencher especificamente a lacuna do que foi chamado de *atividade semântica*. Isto é, pretende-se, a partir dela, garantir um sistema com uma base semântica para suportar a organização conceitual dos acervos dos espaços de memória, bem como a gestão de seus objetos físicos.

Toda essa fase, chamada de pré-processamento, deve alimentar/gerar um fluxo de armazenamento dos objetos digitais e disparar o início da composição semântica nas quatro dimensões (Tempo, Espaço, Pessoas e Fenômeno).

A centralidade nos objetos digitais vai ao encontro do que (Dodebei, 2014, p. 150) chama de “ciclo de vida dos artefatos” que, sob o ponto de vista da cultura material, pode ser utilizado, supondo-se a validade da transferência de modelos no ambiente digital, desde que seja verificada a possível materialidade destes objetos.

Para isso, a nuvem de serviços *on-line* apresentada na Figura 13, representa as estruturas tecnológicas que permitirão essa transferência por meio *Serviços* como *endpoints* SPARQL, *middlewares* e outras possibilidades de colaboração externa que o sistema possa oferecer.

De maneira complementar, porém imperativa, o próprio sistema deve oferecer, de forma mútua, uma capacidade de encontrar e conectar informações de objetos de um conjunto de repositórios, por meio de uma interface comum para interoperabilidade, disparando para cada um destes, um conjunto de informações em termos das quatro dimensões principais: Pessoa, Fenômeno, Espaço e Tempo.

Para a concepção dessa arquitetura geral o modelo proposto requer três grandes ações estruturais com níveis e complexidades distintos no que tange ao conceitual, ao tecnológico e à utilização.

#### **4.4 Concepção do modelo**

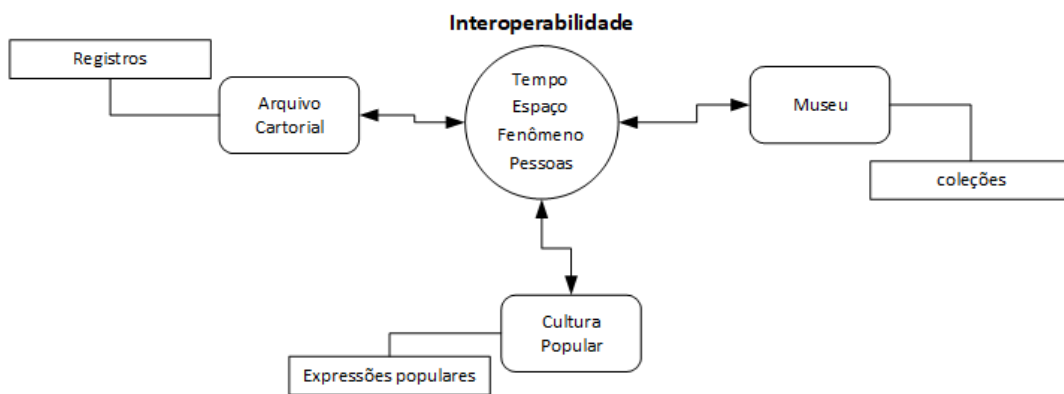
Com base nas premissas apresentadas nos capítulos iniciais e nas principais técnicas e estratégias encontradas e exploradas nos trabalhos avaliados no estado da arte, e aliando-se à experiência vivenciada, sobretudo no que se refere à arquitetura genérica utilizada como ponto de partida (Seção 4.3), propõe-se um modelo genérico de Sistema de Informação semântico, que se apresenta como um grande arcabouço composto de 4 partes. A primeira, uma proposta de sistema Web para interação com os curadores; a segunda, uma camada de armazenamento de dados RDF em tripla (TDB); a terceira uma camada semântica em alto nível para suportar as funcionalidades de um repositório de patrimônio cultural de forma interoperável, especificamente na Gestão de Acervos com ênfase nas organizações internas de cada espaço e, mais estritamente, no que concerne à organização da informação. E, por fim, a quarta, uma camada de interoperabilidade que opera em quatro dimensões para intercâmbio entre os repositórios, fazendo-se valer das classes e propriedades definidas na camada semântica.

Defende-se que este modelo genérico é capaz de suportar as características diversificadas dos centros de memória, uma vez que cada espaço pode possuir seu próprio

repositório que terá suas atividades operacionais efetivadas por um Curador digital responsável pela gestão do acervo. Assim, cada um desses espaços, independentemente de sua origem ou das características dos objetos presentes, é capaz de utilizar sua própria organização para indexação dos objetos.

Um repositório de documentos cartoriais, por exemplo, teria o foco na tipificação dos documentos, como registros, firmas, entre outros. Ou um repositório de expressões populares de uma comunidade que poderia se organizar em torno das danças, falas, mitos e costumes. Um museu, com seu repositório de objetos museológicos, se organizaria de acordo com suas coleções, e assim por diante. A Figura 15 apresenta um exemplo de repositórios possíveis.

**Figura 15 – Exemplo de Repositórios**

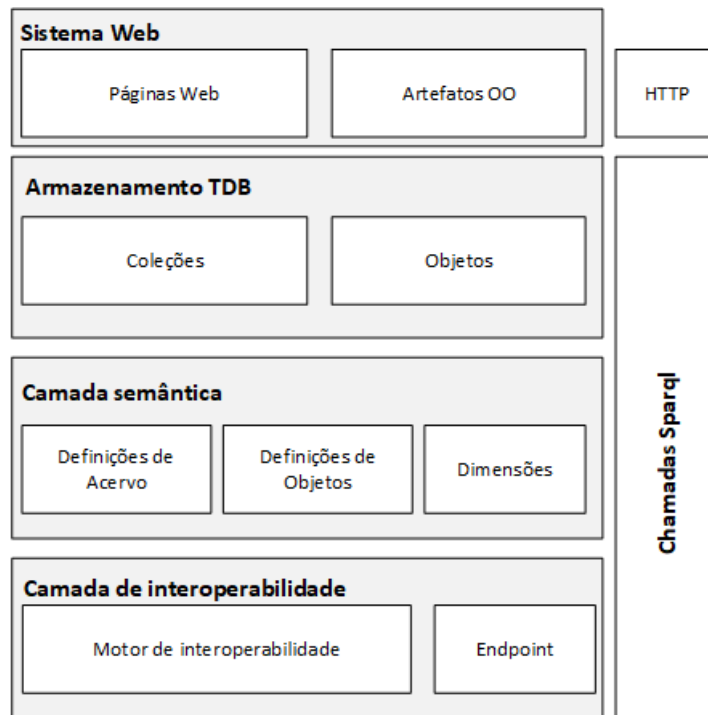


Fonte: Elaborada pelo Autor.

Cada repositório, conforme exemplo apresentado na figura acima, trata-se de um Sistema de Informação composto de uma estrutura na Web para cada espaço de memória, sendo que esse sistema utiliza de uma ontologia com a definição conceitual de cada espaço, uma ontologia para definição dos objetos desse espaço constante de um conjunto mínimo de classes e propriedades e uma última ontologia para a interoperabilidade que apresenta as quatro dimensões propostas.

Como é possível visualizar na Figura 16, este modelo de repositório pode ser organizado como um conjunto de camadas. A partir da visão arquitetural, foi possível perceber que as atividades semânticas estão presentes em todos os fluxos do sistema, e que um modelo genérico para essa arquitetura pode ser assim organizado.

**Figura 16 – Camadas do Modelo Genérico**



Fonte: Elaborada pelo Autor.

A camada mais alta representa o Sistema Web, um composto de páginas e elementos da interação humano-computador, bem como de artefatos necessários para transformação de objetos triplos a partir do paradigma de orientação a objetos ou outro paradigma que precisa ser necessariamente modelado para a tripla *sujeito-predicado-objeto*, constituindo uma organização em grafos. Um sistema Web deve responder requisições HTTP.

A segunda camada mais alta representa o armazenamento RDF através de bancos de dados em tripla (TDB). Uma vez que, havendo a modelagem do paradigma de orientação a objetos, os artefatos devem ser modelados em triplos e persistidos em um banco de dados triplo, ao que é comumente chamado de armazenamento RDF (ou *RDF Store*).

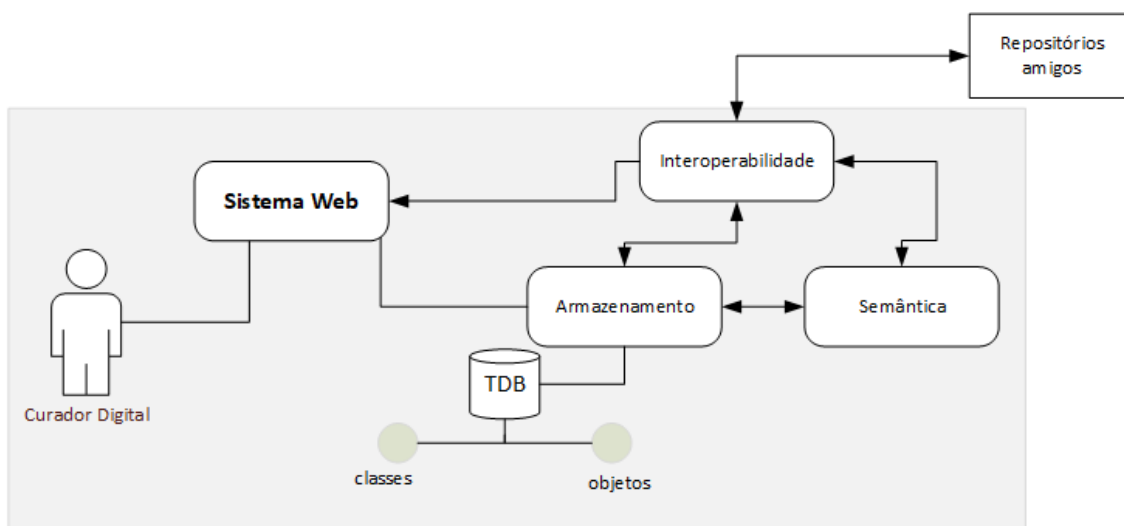
Em seguida, a terceira, denominada camada semântica, representa as ontologias de domínio para a definição dos acervos, definições dos objetos e das quatro dimensões de interoperabilidade. A camada mais baixa, a quarta, é a camada de interoperabilidade, e, por sua vez, é composta por um motor de interoperabilidade e um *end-point* SPARQL para responder a consultas externas e internas. Um motor de interoperabilidade é responsável pelo intercâmbio entre os repositórios, que se dá utilizando os *end-points* e

um conjunto de regras de repositórios “amigos”, que são aqueles repositórios conectados em uma rede colaborativa.

Este modelo em camadas, em uma visão mais ampliada – que pode ser visualizada na Figura 17 – gera uma composição de classes e seus relacionamentos em cada acervo pelos objetos constantes nas coleções com sua descrição contextual simplificada, e uma interface interoperável para permitir que outras classes e propriedades sejam livremente adicionadas pelo Curador.

O fluxo informacional inicia-se na interação do Curador Digital com o Sistema Web, em determinados eventos acionam serviços para a camada de armazenamento, que, por sua vez, aciona utiliza-se da camada semântica para realizar as ações de domínio e contexto, e, posteriormente, por meio da camada de interoperabilidade, o fluxo informacional é acrescido de uma interação com repositórios externos.

**Figura 17 – Visão dos fluxos de informação das camadas**



Fonte: Elaborada pelo Autor

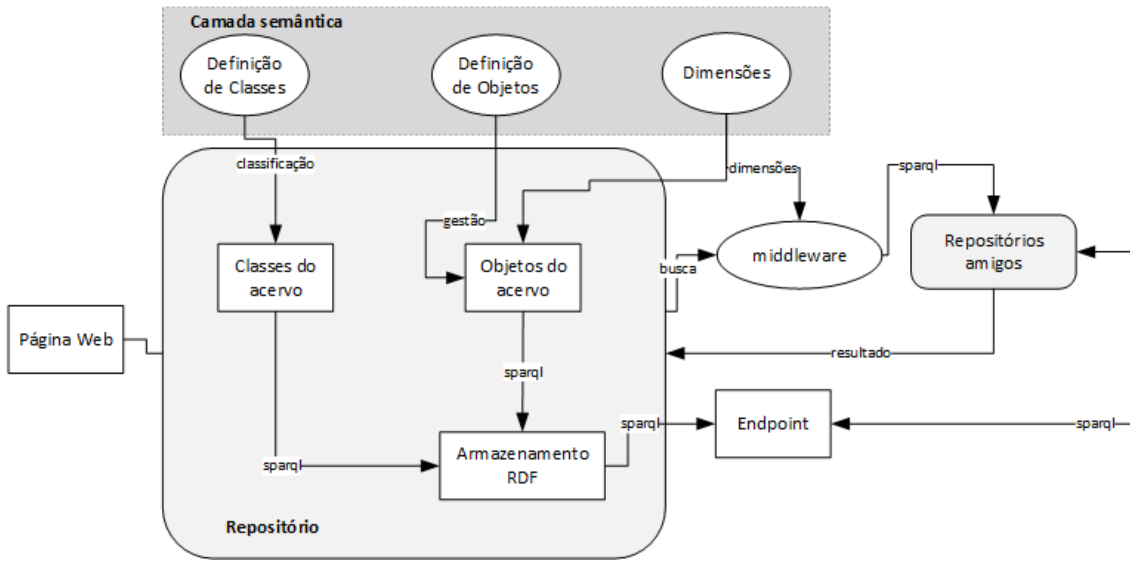
A interoperabilidade, que ocorre a partir de consultas *SPARQL* aos repositórios que fazem parte do conjunto de repositórios conhecidos (*repositórios amigos*), ocorre respeitando as quatro dimensões, Tempo, Espaço, Pessoa e Fenômeno, dentro daquilo que foi definido no ponto de partida.

A camada de interoperabilidade e camada semântica – composta pelas três ontologias para definição das classes dos acervos, dos objetos dos acervos e das dimensões de interoperabilidade – permitem que o intercâmbio ocorra entre os repositórios, partindo de cada objeto ou propriedade adicionada nos acervos através do

Curador Digital.

Dessa maneira, os repositórios funcionam como o par de repositórios demonstrados na Figura 18. Cada centro de memória possuindo seu próprio repositório com uma camada semântica comum e tendo o intercâmbio realizado através dos serviços da camada de interoperabilidade.

**Figura 18 – Esquema de Funcionamento dos Repositórios**



Fonte: Elaborada pelo autor

O repositório é composto do acervo, das classes e objetos do acervo, de um conjunto de RDFs com esses dados e um *endpoint* SPARQL para atender a camada de interoperabilidade, bem como um conjunto de páginas Web como interface com o usuário. Um *middleware* para realizar a busca de informações nos *repositórios amigos* que são pré-definidos e atendem, minimamente, às propriedades e classes básicas definidas na camada semântica.

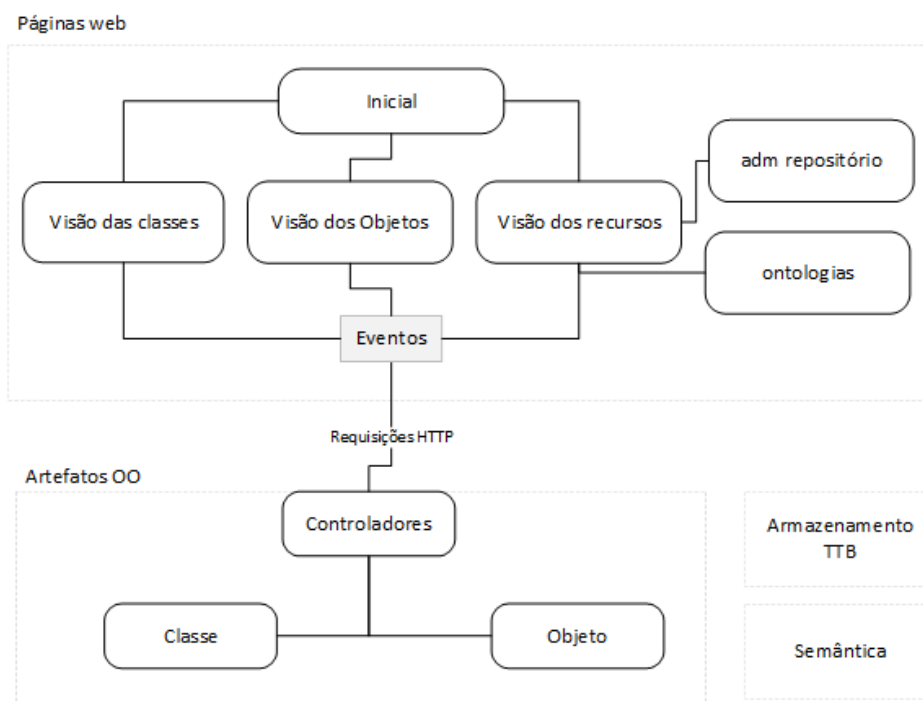
A seguir serão apresentadas as camadas do modelo com mais detalhamento de suas partes internas.

#### 4.5 Primeira camada: Sistema Web

O Sistema Web é composto por um conjunto de páginas Web e um conjunto de objetos instanciados a partir de classes definidas como artefatos de orientação a objetos cujos eventos são manipulados por controladores responsáveis pelo tratamento dos eventos e a transformação destes tanto para a criação das páginas HTML quanto para as camadas mais inferiores do modelo.

Por se tratar de um ambiente Web, a comunicação cliente/servidor obedece ao protocolo HTTP no que concerne às requisições e respostas. O Curador realiza as atividades de gestão do acervo através de ações acionadas pelos eventos (cliente) que disparam requisições para o servidor. Esse servidor, por sua vez, devolve ao Curador através de uma nova instância de página HTML. Os eventos podem ser relacionados a criação de um novo objeto, ao adição de uma nova propriedade, à criação de um novo classificador para o acervo, entre outros. A Figura 19 apresenta como devem estar disponibilizadas as partes mais internas dessa primeira camada.

**Figura 19 – Estrutura Interna do Sistema Web**



Fonte: Elaborada pelo Autor

Um artefato Web, denominado página inicial, deve permitir que o Curador tenha as opções disponíveis para a gestão do acervo a partir da manipulação dos registros referentes às classes presentes no espaço (por exemplo, uma divisão em coleções ou por tempo, ou mesmo pelo tipo dos objetos presentes). Os eventos possíveis, para tanto, referem-se às ações de pesquisa, inserção, modificação ou exclusão segura desses registros. As páginas para esses eventos são abstraídas nessa imagem em um só artefato: *Visão das Classes*. Em um espaço que houvesse uma coleção de mapas da região centro-oeste, por exemplo, poderia haver uma coleção específica para os mapas, essa classe poderia ser denominada “Mapas”.

A *Visão dos Objetos*, por sua vez, deve permitir ao Curador a manipulação dos objetos constantes no acervo. Portanto seus eventos são relacionados à descrição desses objetos e sua classificação dentro do acervo conforme organização já definida. Um mapa digitalizado, por exemplo, receberá as informações iniciais e seu apontamento para a classe “Mapas”.

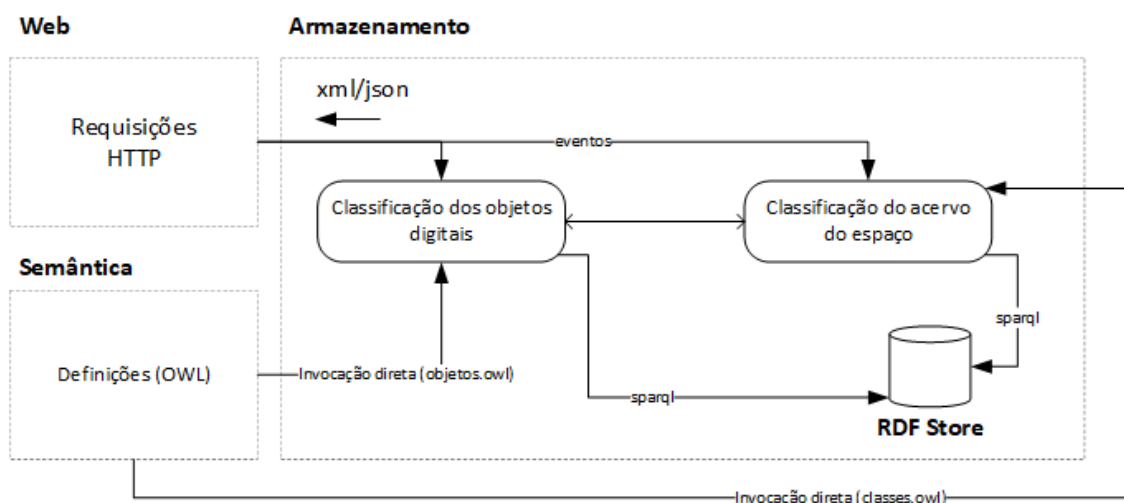
A *Visão dos Recursos* apresenta as páginas referentes à administração do repositório e dos recursos. A lista de *repositórios amigos* e das ontologias iniciais deve ser possível através dessa visão.

#### 4.6 Segunda Camada: Armazenamento (RDF Store)

A camada de armazenamento requer uma estrutura de banco de dados em tripla (TDB) para receber, persistir e recuperar os grafos RDF resultantes dessas relações. As requisições HTTP, advindas dos eventos disparados pela camada Web, são enviadas para o RDF correspondente por meio de chamadas SPARQL, que são transformadas de acordo com cada evento, objeto ou classe do acervo que está sendo manipulado.

O resultado das manipulações geradas no RDF é devolvido para a camada Web em forma de XML ou JSON. A Figura 20 apresenta um esquema do funcionamento interno da Camada de Armazenamento.

Figura 20 – Funcionamento interno da Camada de Armazenamento



Fonte: Elaborada pelo autor.

O armazenamento RDF, isto é, bancos de dados em tripla, garante que cada espaço de memória possua sua própria organização interna e possa classificar seu acervo respeitando essa organização a partir de um conjunto inicial de classes e propriedades que

são comuns a todos os espaços permitindo definir a estrutura de suas coleções, como é possível visualizar no esquema do funcionamento interno da Camada de Armazenamento presente da Figura 20.

A comunicação da Camada de Armazenamento com a Camada Semântica deve acontecer de forma a garantir os eventos relacionados à classificação do acervo do espaço através de um conjunto de classes partindo da hierarquia de maior nível (coisa) e com propriedades que garantam sua hierarquização (subclasse e superclasse). Uma ontologia OWL deve receber essas classes e propriedades.

Da mesma maneira a classificação dos objetos digitais deve seguir uma ontologia OWL que permita a identificação básica de cada objeto, como URI, nome, descrição e uma propriedade para sua tipificação em relação ao acervo classificado conforme as atividades anteriores.

Além dessa classificação básica de cada objeto, deve ser possível para o Curador adicionar propriedades presentes em ontologias estabelecidas pela comunidade. Tais propriedades são adicionadas por meio de invocação direta ou por meio de *middlewares* e outras ferramentas disponíveis na Web. O enriquecimento dos objetos digitais se dá pela ligação com outros objetos, por isso, é importante que essa ação seja possível para os curadores.

O Armazenamento RDF (*RDF Store*) possui dois bancos de dados RDFs. Um relacionado aos objetos do acervo, outro relacionados às suas classes. Se, por exemplo, uma classe “Mapa” precisa ser adicionada ao RDF referente às classes do acervo, uma requisição HTTP com o evento de adição será transformada em uma chamada SPARQL com as características conforme o quadro com o Algoritmo 2 abaixo:

#### Algoritmo 2 – Exemplo de SPARQL em um evento de inserção de classes

```
1 INSERT DATA{
2     <http://200.137.241.247/Acervo#Mapa> a owl:Class
3     rdfs:comment      "Um classe para os mapas"
4     rdfs:label        "Mapas" "
5     rdfs:subClassOf  <http://200.137.241.247/Acervo#> .
6 }
```

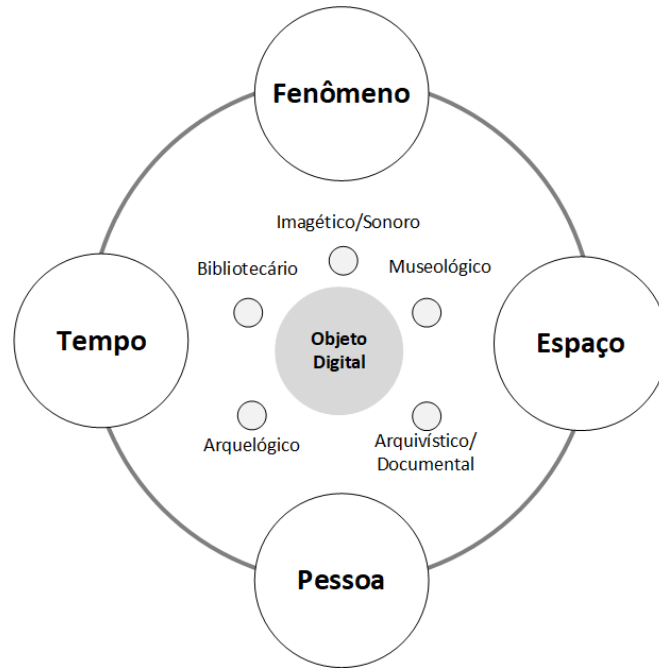
Fonte: Elaborado pelo Autor.

### 4.7 Terceira Camada: Semântica

O modelo conceitual proposto para os centros de memória apresenta, em uma

visão de partida, cinco eixos e quatro núcleos para que o conhecimento produzido transite nas quatro dimensões declaradas: Pessoa, Fenômeno, Espaço e Tempo, assumindo-se que cada objeto pertença, ao menos, a uma das formas previstas no eixos: Arqueológico, Arquivístico/Documental, Bibliotecário, Imagético/Sonoro ou Museológico.

**Figura 21 – Dimensões do Centro de Memória**



Fonte: Elaborada pelo autor.

Este modelo proposto é centrado no objeto digital que se encontra cercado das quatro dimensões, e essas por sua vez, permitindo os estados dos objetos nos cinco eixos – Arqueológico, Arquivístico/Documental, Bibliotecário, Imagético/Sonoro ou Museológico. Enquanto Fenômeno, o objeto se refere a qualquer ação, evento, acontecimento ou manifestação, que ocorreu (ou ocorre) em um Tempo, que pode ser referente a um período ou data em algum Espaço (qualquer local geográfico) e envolve, de alguma maneira, alguma Pessoa (pessoas, povo, grupo populacional).

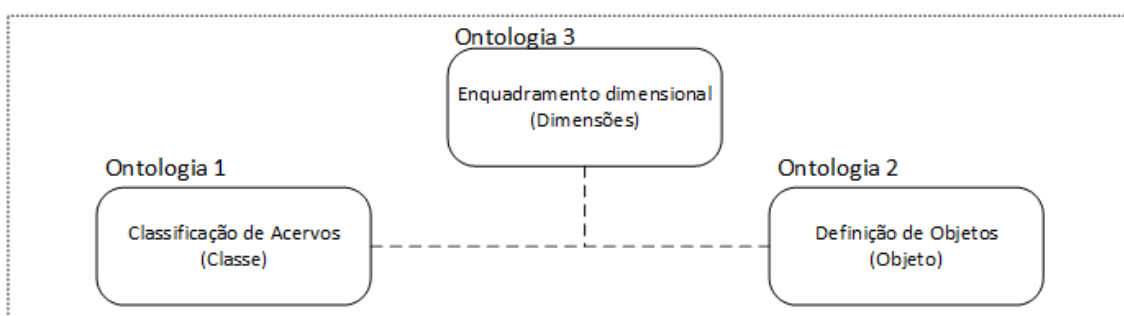
Trata-se, portanto, de um modelo de domínio – uma versão mais detalhada pode ser visualizada na Figura 26 – que pode ser utilizado como base para organização da informação atendendo os cinco eixos, que os tipos informacionais básicos de um GLAM típico.

Além do modelo genérico apresentado, que une elementos das principais técnicas encontradas, define-se aqui uma proposta de camada semântica constando de três ontologias. Uma ontologia de domínio mínimo para organização de cada espaço, uma

ontologia para descrição inicial dos objetos digitais e uma ontologia para interoperabilidade entre os centros de memória a qual leva em consideração 4 dimensões principais (Pessoa, Fenômeno, Espaço e Tempo) que, de acordo com as propriedades associadas às classes definidas em cada repositório, permitirão descrever os estados múltiplos de cada objeto constante nos acervos digitais.

É acrescentado a esta proposta um modelo de motor de interoperabilidade constante de um algoritmo para realizar o intercâmbio entre os repositórios a partir dos eventos de manutenção dos objetos disparados pelos curadores.

**Figura 22 – Esquema das Ontologias da camada semântica**



Fonte: Elaborada pelo Autor.

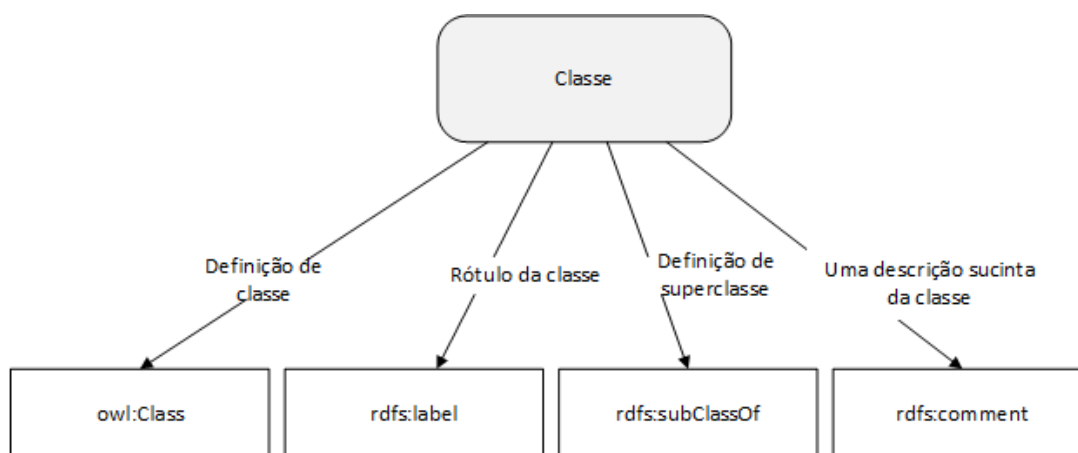
Na Figura 22 é apresentada uma visão geral do esquema com as três ontologias. A primeira ontologia é responsável pela definição dos objetos. Nela as propriedades básicas dos objetos físicos devem ser definidas para a descrição mínima de cada objeto digital do acervo. A ontologia de classificação do acervo deverá permitir, de maneira análoga, a descrição de cada Seção ou divisão do acervo de forma a garantir uma organização hierárquica dos objetos. A segunda ontologia é responsável pelas dimensões e traz a definição dos objetos dimensionais não físicos (Pessoa, Fenômeno, Espaço e Tempo) bem como as propriedades e restrições para esses fragmentos de informação que serão associados a objetos digitais físicos – isto é, presente de fato nos espaços.

#### **4.7.1 Ontologia 1: Classificação do acervo**

Nesta proposta, cada espaço de memória deve possuir sua própria organização, o que normalmente ocorre de acordo com realidades e características do próprio espaço.

Assim, a ontologia para classificação do acervo apresenta um modelo de domínio simplificado, constando de 1 classe para o acervo e 4 propriedades orientadas ao modelo RDF, cujo esquema pode ser visualizado na Figura 23.

**Figura 23 – Esquema da Ontologia de Definições de Classes**



Fonte: Elaborada pelo Autor.

A primeira propriedade, *owl:Class*, é um marcador para informar que se trata de uma ontologia OWL. A segunda, *rdfs:label*, serve para definir um rótulo para cada elemento da classificação do acervo, como uma coleção ou um tipo. A terceira, *rdfs:subClassOf*, é um marcador para hierarquizar as divisões do acervo. Uma descrição textual simplificada de cada um destes elementos é realizada através da propriedade *rdfs:comment*. O Quadro 5 apresenta uma breve descrição de cada propriedade e o IRI em que se encontra cada propriedade.

**Quadro 5 – Descrição das Propriedades da Ontologia 1**

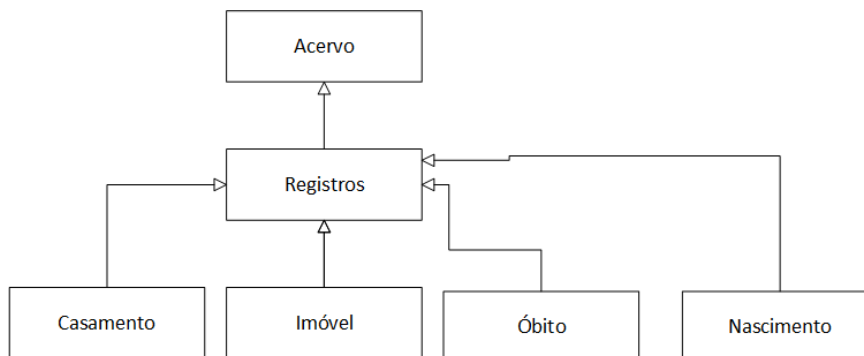
Propriedades utilizadas para classificação do acervo		
Propriedade	Descrição	IRI
<i>owl:Class</i>	Essa marcação indica que se trata de uma classe OWL.	<a href="https://www.w3.org/2002/07/owl#Class">https://www.w3.org/2002/07/owl#Class</a>
<i>rdfs:label</i>	Onde se define um nome para a classe que está sendo definida.	<a href="https://www.w3.org/TR/rdf-schema/#ch_label">https://www.w3.org/TR/rdf-schema/#ch_label</a>
<i>rdfs:subClassOf</i>	Através dessa propriedade é possível criar uma relação de especialização ou generalização entre as classes.	<a href="https://www.w3.org/TR/rdf-schema/#ch_subclassof">https://www.w3.org/TR/rdf-schema/#ch_subclassof</a>
<i>rdfs:comment</i>	Essa propriedade é usada para detalhar comentários, características gerais e outras informações sobre a classe que está sendo criada.	<a href="https://www.w3.org/TR/rdf-schema/#ch_comment">https://www.w3.org/TR/rdf-schema/#ch_comment</a>

Fonte: Elaborada pelo Autor.

Como aplicação desta ontologia, poderia se pensar em um repositório cartorial – como apresentado na parte introdutória deste capítulo. Um cartório possui registros que

podem ser divididos quanto ao tipo, normalmente, registros de casamento, registros de imóveis, de óbito e de nascimento. Para que cada objeto digital nesse cartório seja classificado em uma dessas classes do acervo, é necessário que haja uma definição terminológica, conforme Figura 24.

**Figura 24 – Exemplo de organização do acervo de um cartório**



Fonte: elaborada pelo Autor.

**Serialização 2 - Exemplo de Serialização RDF/XML da classificação de um acervo**

```

1 @prefix : <http://200.137.241.247:8080/fuseki/cartorio#> .
2 @prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
3 @prefix owl: <http://www.w3.org/2002/07/owl#> .
4 @prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
5 @prefix cmgc: <http://200.137.241.247:8080/fuseki/classes> .
6 @prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
7 @prefix dc: <https://www.dublincore.org/specifications/dublin-
8 core/dcmi-terms/> .
9 :
10   a owl:Class ;
11   rdfs:comment "Uma classe genérica para todo o acervo" ;
12   rdfs:label "Registros" ;
13   rdfs:subClassOf cmgc:Acervo .
14 :nascimento a owl:Class ;
15   rdfs:comment "Registros de nascimento"
16   rdfs:label "Nascimento" ;
17   rdfs:subClassOf : .
18 :casamento a owl:Class ;
19   rdfs:comment "Registros de casamento ;
20   rdfs:label "Casamento" ;
21   rdfs:subClassOf : .
22 :obito a owl:Class ;
23   rdfs:comment "Registros de óbito." ;
24   rdfs:label "Óbito" ;
25   rdfs:subClassOf : .
26 :imoveis a owl:Class ;
27   rdfs:comment "Contratos, manuscritos e demais registros de
28 imóveis do cartório." ;
29   rdfs:label "Imóveis" ;
30   rdfs:subClassOf : .
    
```

Fonte: Elaborada pelo autor.

Uma definição possível em RDF para o acervo de um espaço de memória cartorial hipotético, como no exemplo da figura acima, poderia possuir a materialização do RDF/XML descrito na

### Serialização 2 acima.

```

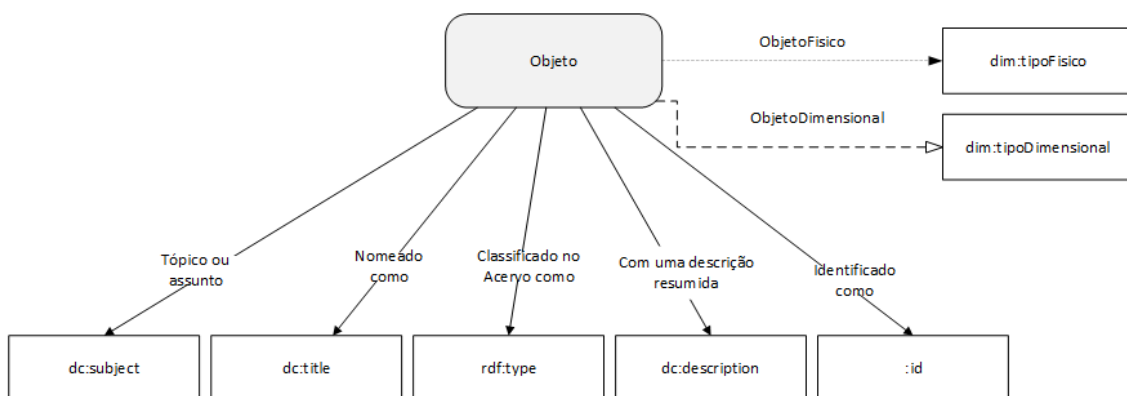
1 @prefix : <http://200.137.241.247:8080/fuseki/cartorio#> .
2 @prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
3 @prefix owl: <http://www.w3.org/2002/07/owl#> .
4 @prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
5 @prefix cmgc: <http://200.137.241.247:8080/fuseki/classes> .
6 @prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
7 @prefix dc: <https://www.dublincore.org/specifications/dublin-
8 core/dcmi-terms/> .
9 :
10 : a owl:Class ;
11 rdfs:comment "Uma classe genérica para todo o acervo" ;
12 rdfs:label "Registros" ;
13 rdfs:subClassOf cmgc:Acervo .
14 :nascimento a owl:Class ;
15 rdfs:comment "Registros de nascimento"
16 rdfs:label "Nascimento" ;
17 rdfs:subClassOf : .
18 :casamento a owl:Class ;
19 rdfs:comment "Registros de casamento ;
20 rdfs:label "Casamento" ;
21 rdfs:subClassOf : .
22 :obito a owl:Class ;
23 rdfs:comment "Registros de óbito." ;
24 rdfs:label "Óbito" ;
25 rdfs:subClassOf : .
26 :imoveis a owl:Class ;
27 rdfs:comment "Contratos, manuscritos e demais registros de
28 imóveis do cartório." ;
29 rdfs:label "Imóveis" ;
30 rdfs:subClassOf : .

```

### 4.7.2 Ontologia 2: Definição dos objetos

Nesta ontologia, os objetos digitais físicos constantes nos acervos – como uma imagem, um quadro ou uma mobília – devem ser descritos e classificados de acordo com cada classe estabelecida na estrutura inicial de classes através da propriedade *rdfs:type*.

Figura 25 – Esquema de organização inicial dos objetos



Fonte: elaborada pelo autor.

Esses objetos são denominados Objetos de Físicos, sempre fazem referência a um artefato físico presente nos espaços de memória e classificado dentro de uma das classes do acervo e normalmente dizem respeito a alguma mídia. Na ontologia número 3,

há um detalhamento sobre os tipos de objetos físicos e dimensionais. A proposta inicial de organização dos objetos segue o esquema apresentado na Figura 25.

Uma regra básica para a criação de um novo objeto é que todos os dados da definição inicial de objetos estejam completos. As quatro propriedades básicas (mínimas) a serem informadas são *rdf:type*, *dc:title*, *dc:subject* e *dc:description*. A propriedade adicional *:id* é relacionada à referência inequívoca indica que seja criado automaticamente um id, que é o identificar único do recurso que o objeto representa. Há ainda duas propriedades importantes.

A propriedade *dim:TipoFisico* é utilizada para tipificação do objeto quanto à mídia ou artefato a que este objeto representa dentro de uma visão de objetos de um GLAM. E por fim, para o caso de objetos dimensionais, há uma propriedade definida na ontologia específica (ontologia 3) que determina para qual dimensão um possível objeto dimensional se direciona.

O quadro abaixo apresenta uma descrição das propriedades iniciais do objeto.

**Quadro 6 – Descrição das propriedades da Ontologia 2**

Propriedade	Descrição	IRI
<i>:id</i>	Identificador do objeto que é gerado automaticamente pelo protótipo criando um identificador único para o recurso.	<a href="http://200.137.241.247:8080/fuseki/objetos#id">http://200.137.241.247:8080/fuseki/objetos#id</a>
<i>rdf:type</i>	Propriedade para definir que o objeto é de uma classe específica. Denota a relação “é um”, ou também o “tipo”	<a href="https://www.w3.org/TR/rdf-schema/#ch_type">https://www.w3.org/TR/rdf-schema/#ch_type</a>
<i>dc:title</i>	Conforme o vocabulário do padrão Dublin Core (DCMI), esta propriedade descreve o título (ou nome) de cada objeto.	<a href="http://purl.org/dc/elements/1.1/title">http://purl.org/dc/elements/1.1/title</a>
<i>dc:subject</i>	Esta propriedade, também do padrão Dublin Core (DCMI), apresenta o tópico ou assunto, sendo uma boa prática a utilização de URI, mas para todo o caso, texto livre quando não for possível.	<a href="http://purl.org/dc/elements/1.1/subject">http://purl.org/dc/elements/1.1/subject</a>
<i>dc:description</i>	Essa propriedade do padrão Dublin Core (DCMI) permite uma descrição mais detalhada do objeto digital em questão,	<a href="http://purl.org/dc/elements/1.1/description">http://purl.org/dc/elements/1.1/description</a>

	podendo envolver, mas não se limita a, uma tabela de conteúdos, resumo, um texto livre uma representação gráfica	
<i>dim:tipoFisico</i>	Essa propriedade é definida na ontologia de definição das dimensões para classificar os objetos quanto ao tipo de mídia que relaciona-se especificamente ao GLAM.	<a href="http://200.137.241.247:8080/fuseki/dimensoes#tipoFisico">http://200.137.241.247:8080/fuseki/dimensoes#tipoFisico</a>
<i>dim:tipoDimensional</i>	Para o caso de objetos dimensionais, essa propriedade permite que estes sejam classificados em uma das quatro dimensões (tempo, espaços, fenômeno e pessoa)	<a href="http://200.137.241.247:8080/fuseki/dimensoes#tipoDimensional">http://200.137.241.247:8080/fuseki/dimensoes#tipoDimensional</a>

Fonte: Elaborado pelo autor.

Um exemplo de materialização de um objeto é apresentado na serialização abaixo, que representa um objeto digital proveniente de um registro de batismo. Como é possível visualizar na linha 12, o objeto em questão está associado ao acervo através da declaração *rdfs:type* `<http://200.137.241.247:8080/fuseki/cartorio#nascimento>`, indicando que o objeto faz parte de uma classe de acervo “nascimento” declarada no TDB “cartório”, que é onde estão definidas as classes deste acervo de exemplo.

### Serialização 3 - Exemplo de Serialização de um objeto

```

1 @prefix : <http://200.137.241.247:8080/fuseki/objetos#> .
2 @prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
3 @prefix owl: <http://www.w3.org/2002/07/owl#> .
4 @prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
5 @prefix dim: <http://200.137.241.247:8080/fuseki/dimensoes#> .
6 @prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
7
8 <http://200.137.241.247:8080/fuseki/cartorioobj/6be7bccf-868e-4ef9-aa04-
9 971bd34c3d93>
10     a :ObjetoDigital ;
11     rdfs:type
12 <http://200.137.241.247:8080/fuseki/cartorio#nascimento> ;
13     dim:tipoFisico dim:Documental ;
14     <http://purl.org/dc/elements/1.1/description>
15         "Registro de Nascimento de João Batista da Veiga, nascido
16 aos 03/10/1881, filho de José Maria da Veiga e Ana Carmo Veiga" ;
17     <http://purl.org/dc/elements/1.1/subject>
18         "João Batita da Veiga" ;
19     <http://purl.org/dc/elements/1.1/title>
20         "Nascimento de João Batista da Veiga" .
21

```

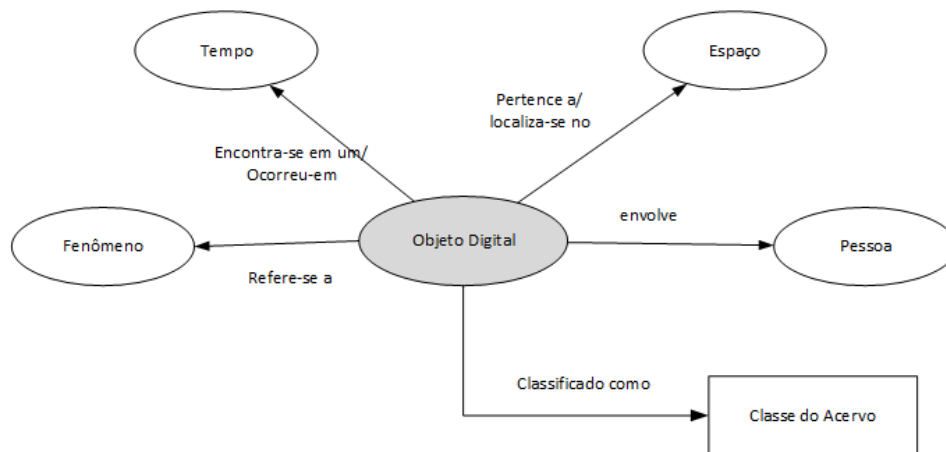
Fonte: Elaborado pelo Autor.

### 4.7.3 Ontologia 3: Dimensões

A Ontologia 3 – Dimensões – é responsável pela conexão entre as duas ontologias iniciais (definições de classes do acervo e de objetos do acervo). Esta ontologia fornece um conjunto de propriedades para permitir que os objetos digitais do acervo e os objetos dimensionais possam se interconectar. Deve permitir ainda que seja possível aplicar outras regras sobre os objetos digitais e sua tipificação quanto às características físicas ou dimensionais. Para isso, apresenta também um conjunto de propriedades que de fato trarão sentido às relações entre os objetos.

A Figura 26 abaixo apresenta uma visão dos conceitos desta ontologia, com as respectivas classes e propriedades.

**Figura 26 – Organização inicial dos conceitos**

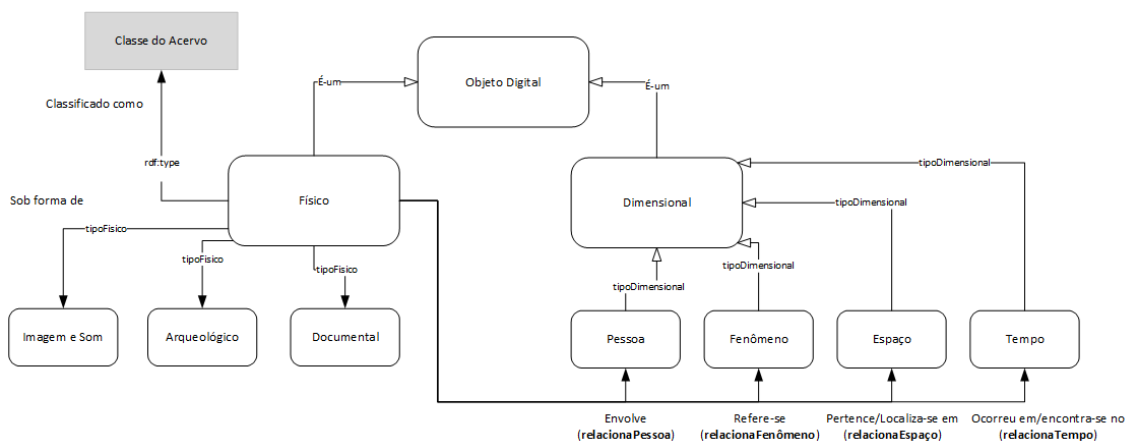


Fonte: Elaborada pelo Autor.

Um objeto digital pode ser encontrado sob várias formas, refere-se a fenômenos, e pode envolver pessoas em um determinado tempo, de acordo com as quatro dimensões apresentadas. Uma ontologia deve possuir classes e propriedades que permitam que os objetos possam ser encontrados em duas formas principais – Objetos Físicos e Objetos Dimensionais – para fornecer um domínio nas quatro dimensões e nos cinco eixos fundamentais, respeitando a classificação dos acervos que são determinadas e mantidas por cada Curador de espaço,.

A Figura 27 apresenta uma visão ampliada dos objetos, que são divididos em Objetos Físicos e Objetos Dimensionais.

**Figura 27 – Visão ampliada dos objetos**



Fonte: Elaborada pelo Autor.

Objeto Digital, neste modelo, representa uma generalização dos objetos dimensionais e objetos de físicos.

Os objetos dimensionais dizem respeito as quatro dimensões e não são, necessariamente, ligados a um artefato físico, por exemplo, a descrição de uma pessoa ou de um espaço. Já os objetos de físicos são de fato os objetos digitais materializados presentes nos espaços, como uma fotografia, uma carta ou uma vestimenta. O Quadro 7 apresenta uma descrição simplificada da ontologia em termos de suas propriedades.

**Quadro 7 – Descrição das Propriedades da Ontologia 3**

Propriedade	Descrição	IRI
<i>:relacionaPessoa</i>	Permite associar qualquer coisa a um Objeto Dimensional Pessoa.	<a href="http://200.137.241.247:8080/fuseki/dimensoes#relacionaPessoa">http://200.137.241.247:8080/fuseki/dimensoes#relacionaPessoa</a>
<i>:relacionaFenômeno</i>	Utilizada para explicitar uma relação de alguma coisa a um Objeto Dimensional Fenômeno.	<a href="http://200.137.241.247:8080/fuseki/dimensoes#relacionaFenômeno">http://200.137.241.247:8080/fuseki/dimensoes#relacionaFenômeno</a>
<i>:relacionaEspaço</i>	Propriedade que associa uma coisa a um espaço.	<a href="http://200.137.241.247:8080/fuseki/dimensoes#relacionaEspaço">http://200.137.241.247:8080/fuseki/dimensoes#relacionaEspaço</a>
<i>:relacionaTempo</i>	Para que se possa associar uma coisa a um Objeto Dimensional Tempo.	<a href="http://200.137.241.247:8080/fuseki/dimensoes#relacionaTempo">http://200.137.241.247:8080/fuseki/dimensoes#relacionaTempo</a>
<i>:tipoFísico</i>	Propriedade para classificar um Objeto Digital em um tipo físico, que envolve os artefatos físicos dos acervos	<a href="http://200.137.241.247:8080/fuseki/dimensoes#tipoFísico">http://200.137.241.247:8080/fuseki/dimensoes#tipoFísico</a>
<i>:tipoDimensional</i>	Propriedade para classificar um Objeto Digital em um tipo dimensional, que não envolve,	<a href="http://200.137.241.247:8080/fuseki/dimensoes#tipoDimensional">http://200.137.241.247:8080/fuseki/dimensoes#tipoDimensional</a>

	necessariamente, os artefatos físicos dos acervos	
<i>dim:tipoDimensional</i>	Para o caso de objetos dimensionais, essa propriedade permite que estes sejam classificados em uma das quatro dimensões (Tempo, Espaços, Fenômeno e Pessoa)	<a href="http://200.137.241.247:8080/fuseki/dimensoes#tipoDimensional">http://200.137.241.247:8080/fuseki/dimensoes#tipoDimensional</a>

Fonte: Elaborado pelo Autor.

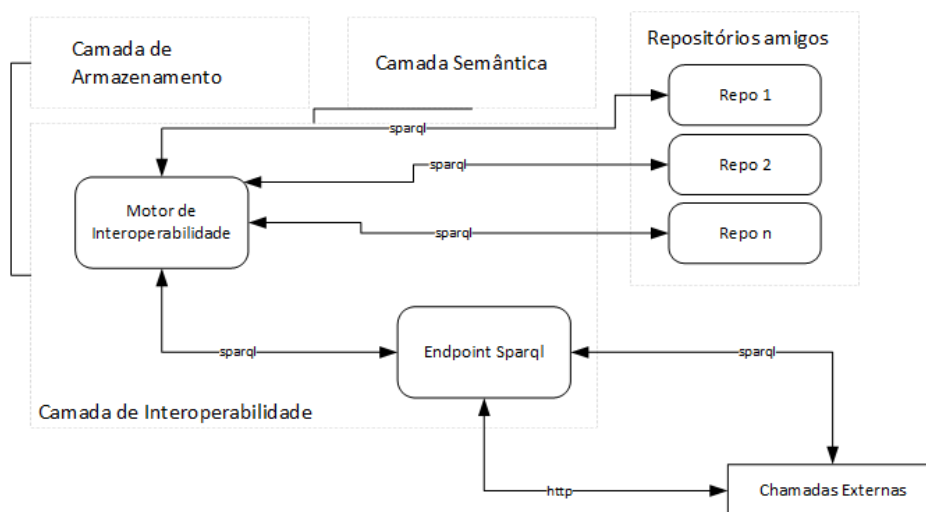
#### 4.8 Quarta Camada: Interoperabilidade

A interoperação ocorre entre os repositórios através da ligação de propriedades adicionadas pelos respectivos curadores, uma vez que os repositórios oferecem uma interface pública via SPARQL e/ou chamadas diretas por meio de *middlewares* de serviços Web.

A camada de interoperabilidade é composta por dois elementos básicos. Um motor de interoperabilidade e um *end-point* SPARQL. O primeiro é responsável pelo intercâmbio entre os repositórios digitais e o segundo permite que requisições internas e externas possam ser respondidas por meio de chamadas SPARQL.

O motor de interoperabilidade é composto por um algoritmo principal que realiza consultas SPARQL quando os *eventos* são acionados em cada repositório, e algoritmos auxiliares que são acionados pelo algoritmo principal ou por chamadas específicas. Por exemplo, se um Curador digital insere um novo objeto ou uma nova propriedade em um objeto já existente, o motor realiza uma busca nos *repositórios amigos* para encontrar informações nas quatro dimensões de interoperabilidade definidas no modelo.

Figura 28 – Esquema da Camada de Interoperabilidade



Fonte: Elaborada pelo Autor.

Como parte do motor de interoperabilidade, a lista de *repositórios amigos* é mantida de um banco de dados triplo assim como os demais objetos e classes do acervo. Essa lista de repositório possui a URI do *endpoint* SPARQL de cada repositório, bem como outras informações sobre cada um dos repositórios, como uma breve descrição e contato.

O algoritmo principal do motor de interoperabilidade aciona então a lista de *repositórios amigos* e dispara para um dos repositórios as consultas SPARQL relacionadas às dimensões e ao objeto ou propriedade que está sendo manipulado pelo Curador, de forma que sejam recuperados os grafos com relação com o objeto ou propriedade em cada uma das dimensões.

O ponto de partida do algoritmo é a busca por palavra-chave e depois, com os identificadores de cada objeto ou propriedade relacionados. Apesar de problemas significativos existentes nas consultas por palavra-chave, as questões relacionadas ao enriquecimento semântico por sinônimos, buscas por proximidade e outras técnicas não são levadas em consideração nesse momento, mas é compreendida a necessidade de utilizar tais técnicas para este tipo de busca.

### Algoritmo 3 – Algoritmo para Interoperabilidade

1. Leia (Objeto\_1)
2. Seja Dimensão={Pessoa, Fenômeno, Espaço, tempo}
3. Seja Relação={relacionaPessoa, relacionaFenômeno, relacionaEspaço, relacionaTempo}
4. Carregue a lista de *repositórios amigos*
5. Para cada repositório faça:
  - 5.1 Busque os Objetos cujo título, assunto ou descrição contenham as palavra-chave do Objeto\_1.
  - 5.2 Busque os Objetos ?x onde exista <Objeto\_1> <Relação> ?x.
6. Devolva todos os grafos ordenados por Dimensão e Repositório.
7. Apresente os dados em triplas.

Fonte: Elaborado pelo autor.

Com o objeto em memória (que pode ser um Objeto Físico ou Dimensional), o primeiro passo é carregar a lista de *repositórios amigos* para se conectar com o endereço de cada repositório.

Levando em consideração que as dimensões já são definidas conforme a propriedade *dim:tipoDimensional* da terceira ontologia, são buscados, então, conforme passo 5 do algoritmo, todos os objetos dimensionais e físicos em cada um dos *repositórios amigos*, primeiro pelo conjunto de palavras chave do objeto. Ainda em cada iteração da lista de repositórios, o algoritmo busca então todos os objetos dimensionais que tenham

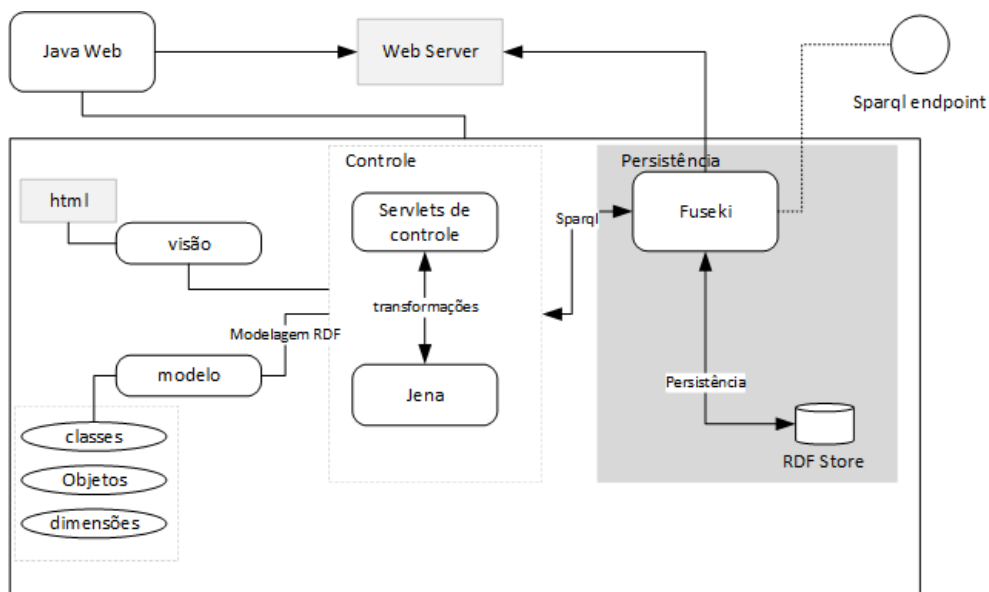
alguma relação direta com o <id> do Objeto alvo de acordo com cada uma das relações dimensionais [*relacionaPessoa/ relacionaFenômeno /relacionaEspaço/ relacionaTempo*].

#### 4.9 Um arcabouço tecnológico para operacionalização da proposta

A proposta de arcabouço é composta de tecnologias livres e já amplamente conhecidas. Uma vez que uma proposta de modelo conceitual foi estabelecida, e diante da possibilidade de utilização de diversas ferramentas para a transformação de modelo para sistema de informação. A linguagem Java, que comporta todo o projeto tecnológico, com os *frameworks*, pacotes, referências e bibliotecas necessárias, com a utilização do *Apache Maven*, para desenvolvimento de aplicações Web.

Para implementar uma arquitetura que possibilite a separação por camadas de forma mais fácil, é proposta a utilização de um framework *modelo-visão-control* (MVC). No caso, o *Java Server Faces* permite uma separação por responsabilidades tanto nos processos relacionados às transformações para visão (*html*) quanto para a manipulação dos *servlets* gerados pelas camadas mais internas do modelo, que fornecem determinados serviços como a transformação dos objetos em triplas, o armazenamento e a busca de informações via SPARQL. A Figura 29 apresenta os elementos chave dessa proposta tecnológica.

**Figura 29 – Esquema do arcabouço tecnológico**



Fonte: Elaborada pelo Autor.

Para servir as aplicações Web, a proposta é a utilização de um *Web Container*,

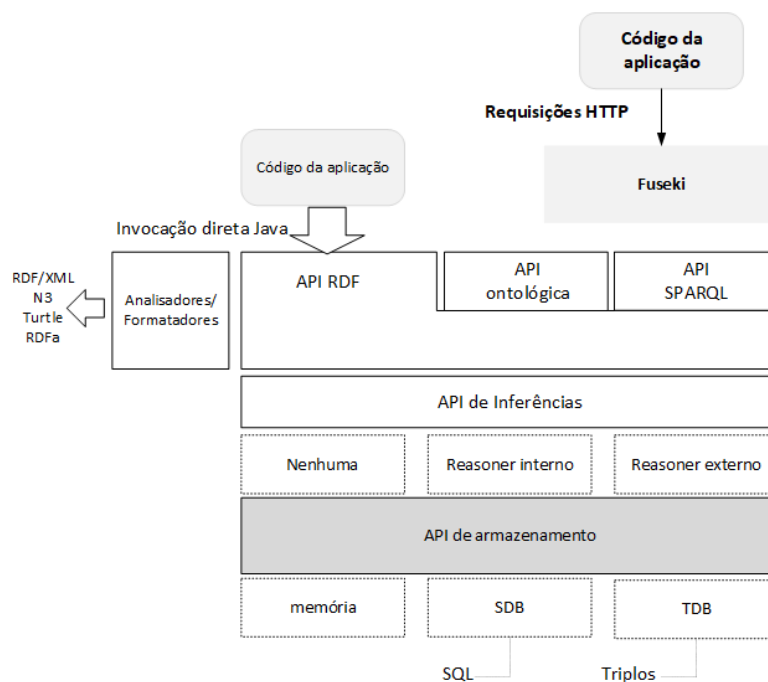
no caso, o *Apache Tomcat*. Este é um software livre que implementa as especificações para plataforma Java Web, permitindo um serviço de requisições e respostas através de protocolos como o HTTP.

Para o desenvolvimento de grafos RDF, o *framework* base *Apache Jena* com toda camada de persistência de bancos de dados em tripla realizada utilizando o *Apache Jena Fuseki* TDB.

#### 4.9.1 Apache Jena

O Apache Jena é um *framework open source* na linguagem Java voltado para o desenvolvimento aplicativos semânticos, composto de vários componentes para modelagem, manipulação, transformação e armazenamento de grafos. A Figura 30, a seguir, apresenta a visão geral do Apache Jena com seus módulos e componentes. Seus componentes são o *Fuseki*, que suporta requisições HTTP, um conjunto de classes Java em uma API RDF para manipulação de ontologias e grafos e um componente para a linguagem SPARQL.

Figura 30 – Esquema de funcionamento geral do Apache Jena



fonte: Adaptado de [https://jena.apache.org/getting\\_started/index.html](https://jena.apache.org/getting_started/index.html)

A camada de Analísadores/Formatadores possui classes para realização de transformações para linguagens de representação sintáticas de RDF com Turtle, RDF/XML e N3.

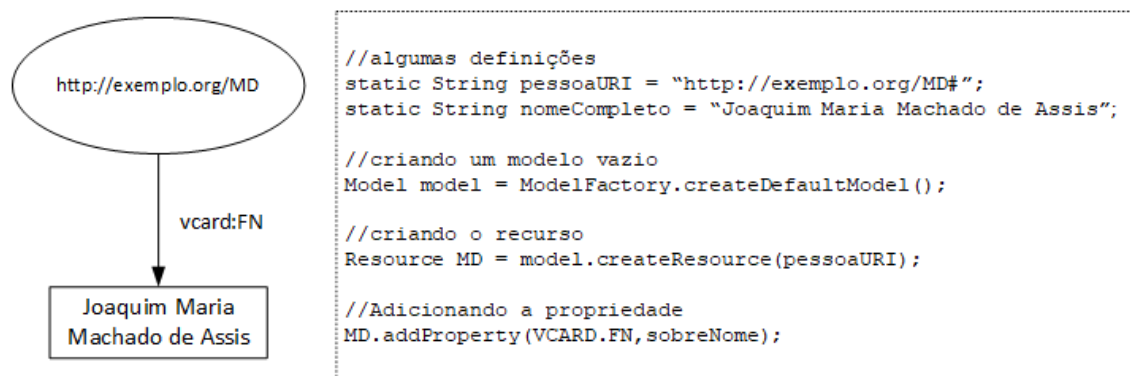
Há uma camada de inferência é composta por classes para inferência e regras para ontologias, que possui um módulo raciocinador (*reasoner*) próprio, mas também permite que sejam utilizados *reasoners* externos.

A camada de armazenamento (*Store API*) permite que os dados sejam armazenados em SQL e TDB (banco de dados triplo), que é aplicável para o caso dessa proposta. No Apache Jena, todas as informações RDF são fornecidas como uma coleção de triplas RDF em uma estrutura de dados chamada modelo. Um modelo é uma abstração Java orientada a objetos de um grafo RDF.

A API RDF permite manipular recursos, literais e nós em branco nos grafos e fornece uma classe (*Statement.class*) para representar um único triplo. A partir dela é possível recuperar o sujeito, o predicado e o objeto de qualquer triplo (Curé & Blin, 2015).

Uma instância de *RDF Model* é representada como um conjunto de sentenças RDF, ou seja, triplas. A Figura 31 traz uma representação simples de um grafo RDF na parte esquerda da imagem, e na direita, é possível visualizar a codificação desse modelo em java com a API Apache Jena.

**Figura 31 – Exemplo de modelo RDF em Apache Jena**



Fonte: Adaptada de (Apache Software Foundation, 2021)

Curé e Blin(2015) afirmam que a API de Ontologia Jena visa fornecer uma programação neutra de linguagem de ontologia através de uma interface para aplicações ontológicas. “As especificidades de cada linguagem são propagadas por meio da noção de um perfil que lista as construções, classes e propriedades permitidas” (Curé & Blin, 2015, p. 79).

Na construção do protótipo (CAPÍTULO VII p. 209), várias partes dos componentes do *Apache Jena* foram utilizados, no entanto, uma estrutura mais simplificada de triplas RDF foi construída para que objetos pudessem ser instanciados e

transformados de forma mais facilitada.

#### 4.9.2 Fuseki TDB

Para que os dados sejam armazenados e recuperados em RDF, se faz necessário utilizar alguma estratégia ou mecanismo de manipulação de triplas. O caminho mais comum na Web Semântica é a utilização de um sistema gerenciador de bancos de dados em triplas. O *Apache Jena Fuseki TDB*, que é um dos componentes do *Apache Jena*, é um servidor de triplas SPARQL que pode ser utilizado como um servidor de aplicação próprio ou rodar em um servidor de aplicação autônomo ou container web como o *Apache Tomcat*. O *Fuseki* é totalmente integrado ao mecanismo de TDB do Jena para fornecer uma camada de armazenamento persistente transacional robusta (Apache Software Foundation, 2021).

Para o protótipo que é apresentado no CAPÍTULO VII, utilizou-se uma instalação do *Apache Jena Fuseki* em um computador remoto com sistema operacional *CentOs 7*, que configurado um Web App em uma instância do *Apache Tomcat versão 9.0.3*. A estratégia foi utilizar o *Fuseki TDB* de forma similar às aplicações Java com bancos de dados relacionais suportadas pela *Java Persistence API (JPA)*, com o diferencial de realizar as consultas SPARQL de forma direta e uma transformação do modelo orientado a objetos para o modelo orientado triplas conforme as camadas já explicitadas na definição da arquitetura genérica. Para tanto, foi criada uma camada de acesso aos dados (DAO) que faz uso do driver *Jena JDBC (Java Database Connectivity)*.

Este conector implementa as principais funções de SQL em chamadas SPARQL. Basicamente a persistência e as consultas realizadas na camada de acesso aos dados do protótipo são construídas utilizando instâncias de classes desse *driver*. Os pacotes com as classes necessárias para realizar essa atividade podem ser declarados no arquivo *pom.xml* conforme quadro abaixo.

#### Serialização 4 – Parte do arquivo pom.xml de um projeto Jena/Fuseki

```
1      [...]
2      <dependency>
3          <groupId>org.apache.jena</groupId>
4          <artifactId>jena-core</artifactId>
5          <version>3.16.0</version>
6      </dependency>
7
8      <dependency>
9          <groupId>org.apache.jena</groupId>
10         <artifactId>jena-fuseki</artifactId>
11         <version>3.16.0</version>
12         <scope>test</scope>
```

```

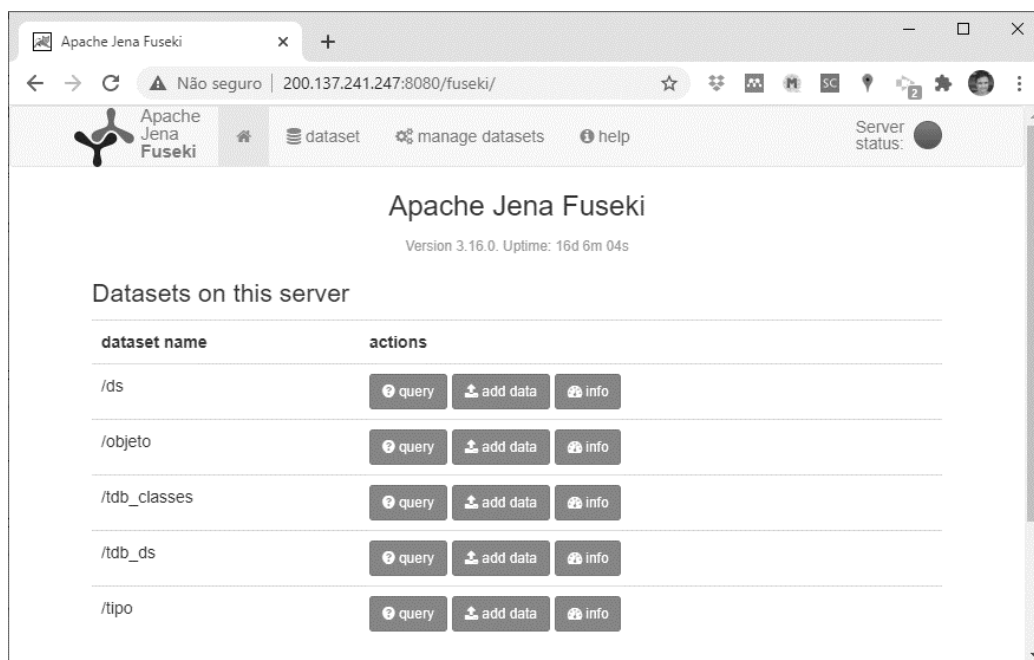
13         <type>pom</type>
14     </dependency>
15
16     <dependency>
17         <groupId>org.apache.jena</groupId>
18         <artifactId>jena-jdbc-driver-remote</artifactId>
19         <version>3.16.0</version>
20         <type>pom</type>
21     </dependency>
22     <dependency>
23         <groupId>org.apache.jena</groupId>
24         <artifactId>jena-jdbc-driver-bundle</artifactId>
25         <version>3.16.0</version>
26         <type>pom</type>
27     </dependency>
    
```

Fonte: Elaborado pelo Autor.

O cuidado para a instalação do *Apache Jena Fuseki* deve levar em consideração as permissões de usuário e as configurações do *endpoint* que devem ser realizados no sistema operacional e no arquivo de configuração *shiro.ini*. Este, cuja versão utilizada nesta tese foi a 3.16.0, fornece, dentre outras possibilidades, uma interface Web para manipulação dos bancos de dados (*Datasets*) e um *endpoint* SPARQL que é utilizado no protótipo dessa tese.

A Figura 32 apresenta a instância do servidor *Apache Jena Fuseki* rodando como Web App no servidor *tomcat* na porta 8080, e com endpoints SPARQL já definidos para cada TDB presente. As consultas SPARQL podem ser realizadas diretamente no Fuseki ou podem ser enviadas ao endpoint via requisições HTTP.

**Figura 32 – Interface Web do Apache Jena Fuseki**

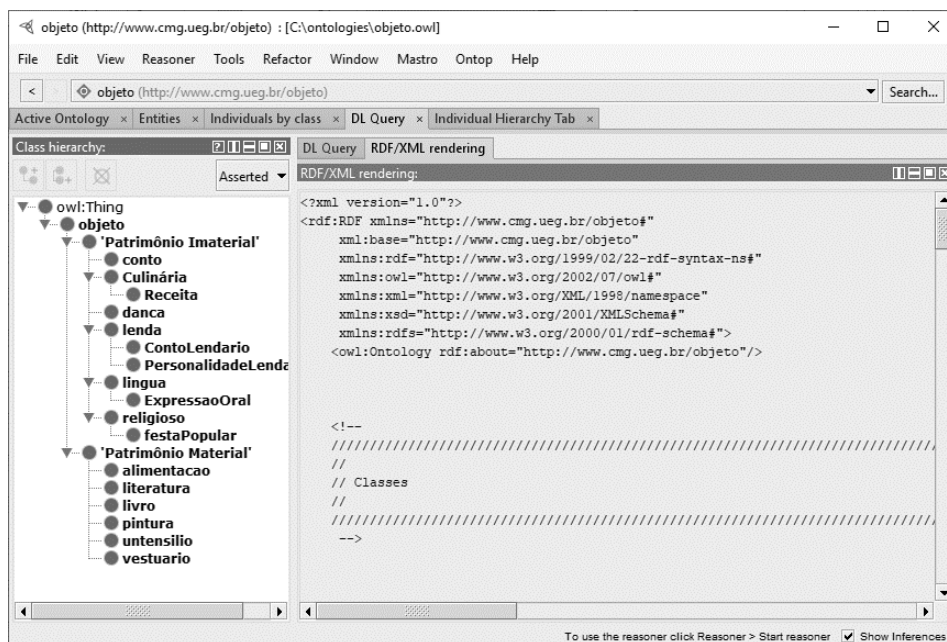


Fonte: Elaborada pelo autor.

### 4.9.3 Protégé

Apesar de não fazer parte diretamente na construção da codificação na linguagem Java, o Protégé permite a criação e manipulação de ontologias de maneira simplificada. Trata-se uma plataforma de *open source* que oferece um conjunto de ferramentas para construir modelos de domínio e aplicativos baseados em conhecimento com ontologias. Ele suporta OWL e oferece uma ferramenta de visualização amigável para navegação interativa de relacionamentos de ontologia (Briola *et al.*, 2017).

Figura 33 – Uma visão do Protégé 5.



Fonte: Elaborada pelo autor.

### 4.10 Operacionalização da proposta

Para operacionalização da proposta deste capítulo se fez necessário um diagnóstico, análise de requisitos e construção de protótipo de sistema de informação cujo detalhamento de passos está definido no CAPÍTULO V e sua materialização descrita no CAPÍTULO VII.

Um diagnóstico com os atores (curadores) de espaços de memória se faz necessário para que se possa captar as informações relacionadas espaços, acervos e objetos dos acervos, uma vez que o que a proposta requer o conhecimento sobre as estruturas tecnológicas e físicas destes com potencial para obtenção de informações de objetos nato-digitais ou digitalizados das instituições.

A intenção deste diagnóstico é compreender melhor as características tecnológicas para construir modelos e protótipos orientados às expectativas dos membros dos espaços, em especial os curadores digitais, e do potencial para interoperabilidade semântica dos seus repositórios.

Os modelos, quaisquer que sejam, precisam se preocupar com o reuso para que possam alcançar de forma plena o que Doerr (2009) define como de sistemas de informação de patrimônio. O reuso, de uma maneira ou de outra, exige que se leve em consideração as possibilidades diversas de usabilidade e conectividade com outras plataformas possíveis, para que pesquisadores, professores, estudantes e comunidade em geral possam utilizá-lo de forma efetiva.

Nesse ponto em específico, ressalte-se que a presente proposta se limita em fornecer um *endpoint* SPARQL na camada de serviços para outros sistemas e aplicativos possam fazer consumir as informações presentes nos RDFs armazenados, e outras definições de utilização requerem, no futuro, um amplo estudo das necessidades da interface com o usuário, com a finalidade de tornar as informações o mais amigável e facilmente navegável possível para um grupo heterogêneo de usuários e sistemas. Nessa proposta todas as atividades relacionadas ao reuso estão limitadas ao Curador digital nas funcionalidades de Gestão do acervo em um recorte específico do registro dos acervos e objetos digitais.

#### **4.11 Sumário do Capítulo**

O objetivo desse capítulo foi apresentar uma proposta que se compõe de uma arquitetura genérica, uma estruturação conceitual e uma estruturação tecnológica afim de encontrar um modelo de sistemas de informação para centros de memória a ser aplicado especificamente ao recorte de problema investigado. Este modelo que explora definições de domínio, tecnológicas e de utilização, traz consigo um conjunto de ferramentas possíveis para aplicação do modelo.

No próximo capítulo será apresentado o percurso metodológico desta tese detalhando os instrumentos e estratégias que foram utilizados para a aplicação da proposta apresentada neste capítulo.

## **CAPÍTULO V. PERCURSO METODOLÓGICO DA INVESTIGAÇÃO**

### **5.1 Introdução**

Neste capítulo são apresentados os métodos, as estratégias e as ferramentas metodológicas utilizadas nesta tese. Pretende-se explicitar o caminho metodológico e suas nuances desde a qualificação do problema de pesquisa até a discussão em torno de possíveis soluções.

Esta pesquisa, enquanto a sua natureza, se apresenta como teórica a partir das revisões bibliográfica e sistemática da literatura. Caracteriza-se também como uma pesquisa aplicada, de caráter empírico e exploratório, balizando-se na pesquisa-ação e na engenharia de software para produção de resultados concretos.

As técnicas de pesquisa-ação, segundo Vieira (2009), são de cunho qualitativo, uma vez que buscam informações sobre determinada questão através da interação com pessoas e seguem um caráter exploratório e não generalizável (Vieira, 2009).

Em outras palavras, os entretons envolvidos no problema e nas especificidades destes, não permitem a busca por uma resposta geral que seja aplicável para qualquer outro problema, mesmo respeitadas as estratégias metodológicas, uma vez que as condições são sensíveis aos fenômenos, especialmente em investigações que se referem ao digital e à Rede.

Assim, esta metodologia busca criar uma relação entre elementos da pesquisa-ação e da engenharia de software, ambas apoiadas na pesquisa bibliográfica, com o intuito de contribuir de forma efetiva na solução de problemas práticos.

A opção por não se limitar a um estudo de caso está na essência do que se pretende produzir enquanto resposta nessa investigação. Para além dos estudos de caso e estudos de casos múltiplos, onde seria possível um profundo conhecimento do fenômeno investigado, o método não visa propor uma solução dos problemas identificados. Já no método Pesquisa-Ação, o conhecimento inicial com análise situacional do fenômeno produz ampla visão do contexto e das práticas atuais dos participantes envolvidos (BUNDER & BARROS, 2019).

A pesquisa-ação é uma modalidade de investigação-ação fortemente desenvolvida nas ciências sociais na busca de resultados concretos (Thiolent, 2011; Tripp, 2005). Já a Engenharia de Software se apresenta sob forma de métodos técnicos e

científicos voltados para extração, compreensão e construção de respostas (soluções) através de requisitos de sistemas, que também visam produzir resultados concretos (Pressman, 2011; Sommerville, 2011).

A investigação que originou esta tese foi conduzida com base na hipótese e nos problemas norteadores declarados no capítulo introdutório, justificando-se na afirmação de Thioent (Thioent, 2011, p. 43)

Pensamos que é perfeitamente viável a flexibilização do raciocínio hipotético, de acordo com o qual a hipótese é uma suposição criativa que é capaz de nortear a pesquisa inclusive nos seus aspectos qualitativos. As hipóteses (ou diretrizes) qualitativas orientam, em particular, a busca de informação pertinente e as argumentações necessárias para aumentar (ou diminuir) o grau de certeza que podemos atribuir a elas (Thioent, 2011, p. 43).

## 5.2 Instrumentos e estratégias

Esta tese apresenta um objetivo geral, em pretende-se ser alcançado pela investigação como um todo, e quatro objetivos específicos, que pretendem-se alcançar conforme quadro abaixo.

**Quadro 8 – Detalhamento dos métodos utilizados quanto aos objetivos**

Objetivo	Método/Estratégia
1. Realizar um estudo sobre as características dos espaços de memória e patrimônio ligados à Secretaria de Estado da Cultura do Estado de Goiás (SECULT-GO) e sua utilização quanto aos sistemas de informação	Através de um diagnóstico e uma discussão acerca dos espaços de memória do objeto da investigação conforme descrição neste capítulo metodológico e análise dos resultados nos Capítulos VI e VII.
2. Realizar um estudo da aplicabilidade dos modelos conceituais de representação da informação de patrimônio cultural e tecnologias da Web-Semântica.	Através de revisão bibliográfica da literatura apresentada no CAPÍTULO II, complementando-se por uma revisão sistemática da literatura apresentada no CAPÍTULO III.
3. Construir um modelo semântico para organização da informação de patrimônio cultural, a partir de padrões técnicos, que permita a interoperabilidade entre repositórios de patrimônio cultural, aqui denominados espaços de memória	Através da apresentação de um modelo com base nas tecnologias da Web Semântica e nas características dos espaços investigados no CAPÍTULO IV, desenvolvimento e arquitetura do protótipo no CAPÍTULO VI e CAPÍTULO VII e discussão dos resultados no CAPÍTULO VIII.
4. Verificar se a complexidade do modelo semântico é um empecilho	Uma discussão sobre a complexidade do modelo semântico é apresentada no CAPÍTULO VIII.

para os curadores de acervos digitais de patrimônio cultural	
--	--

Fonte: Elaborado pelo autor

Dessa forma, essa pesquisa foi realizada de maneira incremental e iterativa constando das seguintes etapas: 1) diagnóstico inicial; 2) revisão bibliográfica e enquadramento teórico; 3) engenharia de requisitos e 4) construção de protótipos.

### **5.3 Diagnóstico inicial**

Conforme evidenciado na proposta apresentada no capítulo anterior, a construção de um diagnóstico dos repositórios digitais e espaços de memória do estado de Goiás se mostrou fundamental para a pesquisa. Sem o conhecimento das estruturas tecnológicas (e até mesmo físicas) dos espaços, torna-se difícil pensar, delimitar e propor modelos de integração de dados e aplicação de tecnologias da Web Semântica.

Assim, parte-se de um pressuposto que os espaços de memória são aqueles com potencial para obtenção de informações de objetos nato-digitais ou digitalizados que estão em instituições de alguma maneira vinculadas à Secretaria de Cultura do Estado de Goiás, e que tem como pano de fundo a necessidade de acessibilidade e potencial interoperação por meio da Rede.

De forma geral, é conhecida a realidade desses espaços, que, nos últimos anos têm encontrado dificuldades para a manutenção de seus acervos e, em especial, para torná-los acessíveis ao público. Diante disso, a intenção dessa etapa é compreender melhor as características tecnológicas para além do senso comum e, partir do contato direto com os atores envolvidos nos espaços, poder construir modelos e protótipos orientados às expectativas desses atores, coletando também dados documentais disponíveis nos sítios institucionais desses GLAMs na Web e nos repositórios de base legal.

#### **5.3.1 Amostragem da investigação**

A Secretaria de Cultura do Estado de Goiás, que foi criada em 11 de novembro de 1964, no governo de Mauro Borges Teixeira, unida à pasta de Educação (Lei 5.623), já teve, desde sua criação, diversas modificações nos governos sucessores.

Quanto aos objetivos da SECULT-GO, encontra-se a seguinte declaração:

A Secretaria de Estado de Cultura tem como objetivo, segundo a Lei nº

20.417/19, formular e executar a política estadual de desenvolvimento da cultura; bem como zelar pela conservação do patrimônio histórico e artístico do Estado, a criação e manutenção de bibliotecas, centros culturais, museus, teatros, arquivos históricos e demais instalações ou instituições de caráter cultural (SECULT, 2020).

É possível encontrar, sob às diretrizes desta instituição, a promoção, incentivo e apoio às artes cênicas, visuais, audiovisuais, a música, a literatura, bem com a cultura goiana de forma geral e o seu patrimônio, estabelecendo parcerias para a produção da cultura com escolas, universidades, organizações sociais, fundações e outras instituições de desenvolvimento da cultura (SECULT, 2020).

Segundo as diretrizes gerais do Plano Nacional de Cultura, (Plano Nacional de Cultura, 2007), “A igualdade e a plena oferta de condições para a expressão e fruição culturais são cada vez mais reconhecidas como parte de uma nova geração dos direitos humanos”.

O Plano Nacional de Cultura reforça que tais direitos, para serem efetivos, precisam ser incorporados ao cenário político e social brasileiro em uma construção colaborativa entre diferentes setores de interesse.

Assim, a presente pesquisa se apoia também em uma iniciativa interinstitucional entre a Universidade Estadual de Goiás (UEG) e a Secretaria de Estado da Cultura do Estado de Goiás (SECULT-GO), que se mostra oportuna conforme o instrumento celebrado entre as instituições, e que tem como foco maior o desenvolvimento de ações votadas para a preservação do patrimônio do estado de Goiás.

A amostragem coberta pela pesquisa é composta pelos seguintes espaços de memória: 1) Arquivo Histórico Estadual; 2) Biblioteca Estadual Pio Vargas 3) Gibiteca Jorge Braga; 4) Museu da Imagem e do Som de Goiás; 5) Museu Ferroviário de Pires do Rio; 6) Museu Pedro Ludovico; 7) Museu Zoroastro Artiaga e 8) Palácio Conde dos Arcos.

A coleta dos dados desta etapa da pesquisa se deu de três formas fundamentais 1) pesquisa nos documentos públicos e sites institucionais na Web; 2) inquérito com os curadores dos espaços culturais; e 3) entrevista com servidores, gestores e outros atores.

### **5.3.2 Pesquisa nos documentos públicos e sites institucionais**

Para a etapa da pesquisa documental foram avaliados O Plano Nacional de

Cultura do Brasil, em suas diretrizes gerais (Plano Nacional de Cultura, 2007), A lei 20.491/2019 que estabelece as competências da Secretaria de Estado da Cultura e a Resolução UEG/CsA 1045/2017 (Resolução CsA 1045/2017 - Política de Extensão Da Universidade Estadual de Goiás, 2017).

Os documentos referentes aos espaços de memória específicos foram consultados conforme o quadro abaixo.

**Quadro 9 – Sítios Eletrônicos dos Espaços de Memória**

Espaço	Sítio Eletrônico
Arquivo Histórico Estadual (AHE)	<a href="https://www.cultura.go.gov.br/index.php/centros-culturais/todas-as-unidades/2-institucional/2241-arquivo-historico-estadual">https://www.cultura.go.gov.br/index.php/centros-culturais/todas-as-unidades/2-institucional/2241-arquivo-historico-estadual</a>
Biblioteca Estadual Pio Vargas	<a href="https://www.cultura.go.gov.br/index.php/centros-culturais/todas-as-unidades/2-institucional/2242-biblioteca-estadual-pio-vargas">https://www.cultura.go.gov.br/index.php/centros-culturais/todas-as-unidades/2-institucional/2242-biblioteca-estadual-pio-vargas</a>
Gibiteca Jorge Braga	<a href="https://www.cultura.go.gov.br/index.php/centros-culturais/todas-as-unidades/2-institucional/2245-gibiteca-jorge-braga">https://www.cultura.go.gov.br/index.php/centros-culturais/todas-as-unidades/2-institucional/2245-gibiteca-jorge-braga</a>
Museu da Imagem e do Som de Goiás (MIS)	<a href="https://www.cultura.go.gov.br/index.php/centros-culturais/todas-as-unidades/2-institucional/1000-museu-imagem-e-do-som-de-goias">https://www.cultura.go.gov.br/index.php/centros-culturais/todas-as-unidades/2-institucional/1000-museu-imagem-e-do-som-de-goias</a>
Museu Ferroviário de Pires do Rio	<a href="https://www.cultura.go.gov.br/index.php/centros-culturais/todas-as-unidades/2-institucional/2246-museu-ferroviario-de-pires-do-rio">https://www.cultura.go.gov.br/index.php/centros-culturais/todas-as-unidades/2-institucional/2246-museu-ferroviario-de-pires-do-rio</a>
Museu Pedro Ludovico	<a href="https://www.cultura.go.gov.br/index.php/centros-culturais/todas-as-unidades/2-institucional/2253-museu-pedro-ludovico">https://www.cultura.go.gov.br/index.php/centros-culturais/todas-as-unidades/2-institucional/2253-museu-pedro-ludovico</a>
Museu Zoroastro Artiaga (MUZA)	<a href="https://www.cultura.go.gov.br/index.php/centros-culturais/todas-as-unidades/2-institucional/2261-museu-zoroastro-artiaga">https://www.cultura.go.gov.br/index.php/centros-culturais/todas-as-unidades/2-institucional/2261-museu-zoroastro-artiaga</a>
Palácio Conde dos Arcos	<a href="https://www.cultura.go.gov.br/index.php/centros-culturais/todas-as-unidades/2-institucional/2262-palacio-conde-dos-arcos">https://www.cultura.go.gov.br/index.php/centros-culturais/todas-as-unidades/2-institucional/2262-palacio-conde-dos-arcos</a>

Fonte: elaborado pelo autor

### 5.3.3 Inquérito com os curadores de espaços culturais

O inquérito com os curadores dos espaços de memória foi realizado por meio do aplicativo de formulários do Google (*Google Forms*) e teve o intuito principal de levar as informações dos acervos e de cada espaço de memória para que se pudesse, depois do diagnóstico inicial, passar para as próximas etapas metodológicas.

Todas as questões foram colocadas de forma aberta a partir de um roteiro estruturado. Dessa forma, os curadores responderam de acordo com a sua experiência e

conhecimento dos acervos e espaços. Junto à cada questão havia uma breve explicação sobre seu objetivo.

Após autorização por parte da Secretaria de Estado da Cultura do Estado de Goiás (SECULT-GO), o pedido de resposta do formulário, cujo endereço é <https://forms.gle/kGk7fMneVQvqFZA6>, foi encaminhado por e-mail pelo responsável pela gerência da própria secretaria. As questões são apresentadas abaixo.

- i)* Questão 1: “Por favor informe seu nome”.
- ii)* Questão 2: “Nome do Espaço (Museu, Galeria, biblioteca, outro)”
- iii)* Questão 3: “Se existe um Site institucional do espaço, informe o endereço”.
- iv)* Questão 4: “Informe o endereço de e-mail do Espaço”.
- v)* Questão 5: “Qual a composição do acervo?”
- vi)* Questão 6: “Como o acervo é organizado/inventariado?”
- vii)* Questão 7: “Quais as características dos objetos do acervo? Quais as principais classes, assuntos, termos e relações existentes?”
- viii)* Questão 8: “Quais sistemas de informação são utilizados?”
- ix)* Questão 9: “Quais as estratégias utilizadas para socialização dos conhecimentos e acesso por parte de pesquisadores e público geral?”
- x)* Questão 10: “Quais projetos/ações de digitalização atuais e passados?”
- xi)* Questão 11: “O que espera de um sistema de informação para o acervo? Quais funcionalidades precisam ser atendidas?”
- xii)* Questão 12: “Qual o perfil dos usuários?”
- xiii)* Questão 13: “Se existem, quais dificuldades mais comuns são reportadas pelos usuários?”
- xiv)* Questão 14: “Quanto às informações digitais, o que é mais importante no acervo?”
- xv)* Questão 15: “Como considera informações disponíveis em outros

acervos? Há integração ou propostas de integração?”

xvi) Questão 16: “Outras informações que julgar pertinente”

### **5.3.4 Entrevista para refinamento e levantamento das tecnologias envolvidas nos acervos**

O objetivo desta etapa da coleta foi refinar a etapa anterior, observar quais são as tecnologias utilizadas pelos espaços e como é a conectividade e interoperação com outras tecnologias e repositórios, em especial para coletar indícios sobre a utilização e do domínio das equipes em determinadas tecnologias da Web-Semântica. Se há repositório integrado, e se há acessibilidade para o público externo pela rede, se o espaço utiliza alguma estratégia ou serviço semântico, se o Curador ou alguém da equipe tem conhecimento sobre as tecnologias da Web-Semântica.

Antes de iniciar a entrevista houve uma breve apresentação sobre o projeto de pesquisa desta tese, sobre o contexto dos espaços de memória nessa investigação e os objetivos da entrevista. Houve apresentação do *Termo de Consentimento Livre e Esclarecido* onde constava que o entrevistado é livre pra responder ou deixar de responder qualquer questão e garantia que as informações pessoais e ou qualquer outra informação sensível não será divulgada ou utilizada sem autorização prévia dos entrevistados, e que as transcrições serão realizadas, preservando identidade, informações sensíveis e qualquer outra questão que possa causar qualquer constrangimento aos entrevistados.

O roteiro da entrevista previu três blocos, o primeiro apresentou as questões gerais, o segundo as questões complexas e técnicas e o terceiro foi reservado para as questões mais sensíveis.

- i) Questão 1: “Consideramos espaços de memória todo e qualquer lugar onde há esforços para preservar, manter, permitir o acesso a informações de memória. Os museus, bibliotecas, galerias de arte e arquivos são os exemplos mais comuns desses espaços. Fale um pouco sobre o espaço de memória em que atua”.
- ii) Questão 2: “Estamos vivendo em um tempo bastante peculiar. A Internet e outras novas tecnologias da informação e comunicação parecem modificar complementemente o cenário, o público o gosto, alguns autores sugerem inclusive uma nova cultura em que chama de cultura digital ou cibercultura, gostaria que você falasse um pouco sobre a

relação sobre a relação dos espaços, em especial o que atua, com a tecnologia, com as novidades, com essa nova cultura”

- iii) Questão 3: “Em relação ao acervo digital, o que se tem hoje? Como ele é armazenado?”
- iv) Questão 4: “Quanto aos sistemas existentes, quais são os sistemas de gestão do acervo? Eles atendem atualmente às demandas do espaço?”
- v) Questão 5: “Possui um sistema externo para acesso ao público? Eles atualmente atendem à demanda do público externo, na sua opinião?”
- vi) Questão 6: “Possui integração com algum outro sistema?”
- vii) Questão 7: “Tem conhecimento ou domínio de algum padrão internacional como CIDOC-CRM e EDM?”
- viii) Questão 8: “Tem conhecimento de alguma iniciativa nova de sistemas que possa ser utilizada?”
- ix) Questão 9: “Qual formação acadêmica profissional e experiência na área?”
- x) Questão 10: “Há quanto tempo atua no espaço?”
- xi) Questão 11: “Há algo que queria acrescentar?”

#### **5.4 Análise dos dados**

Para a análise dos dados levantados levou-se em consideração os aspectos dos espaços de memória, as tecnologias presentes nesses espaços e a visão dos curadores quanto as tecnologias utilizadas e com potencial de serem utilizadas no futuro. Utilizou-se da análise de conteúdo para esse fim (A. H. Silva & Fossá, 2015, p. 2):

A análise de conteúdo é uma técnica de análise das comunicações, que irá analisar o que foi dito nas entrevistas ou observado pelo pesquisador. Na análise do material, busca-se classificá-los em temas ou categorias que auxiliam na compreensão do que está por trás dos discursos

Como se trata de uma proposta balizada na pesquisa-ação, após a análise de conteúdo, utilizou-se dos elementos da Engenharia de Software para a transformação desse conteúdo em requisitos de *software*, o que resultou em modelos e protótipos cujas estratégias são descritas na próxima Seção. Os resultados do diagnóstico inicial são

discutidos detalhadamente na Seção 6.2.

### 5.5 Revisão bibliográfica e enquadramento teórico

O enquadramento teórico e, conseqüentemente, a revisão bibliográfica foram realizados conforme o CAPÍTULO II, com base nos pressupostos da ciência da informação. O CAPÍTULO III apresenta ainda de forma complementar, uma revisão sistemática da literatura.

Uma revisão sistemática, conforme Galvão e Pereira (2014), trata-se de um estudo secundário que avalia fontes primárias de pesquisa a fim de identificar, selecionar, avaliar e sintetizar evidências relevantes disponíveis. Suas etapas compreendem em elaborar pergunta de pesquisa; efetuar busca em bases científicas; selecionar trabalhos; extrair dados; avaliar os trabalhos quanto à qualidade metodológica; realizar síntese dos trabalhos (meta-análise); avaliação da qualidade das evidências; e redação e publicação dos resultados (Galvão & Pereira, 2014).

Para este trabalho, em específico, o método foi simplificado em 5 etapas. 1) construção da *string* de busca; 2) pesquisa em bases de dados; 3) seleção e exclusão de trabalhos conforme critérios; 4) análise dos trabalhos; e 5) construção do relatório de análise.

Partindo de pergunta de pesquisa (Seção 3.1 p. 79) e do problema abordado pela tese (Seção 1.5 p. 35), buscou-se construir um conjunto de palavras-chave genérico para ser aplicado em bases de dados de periódicos de Ciência da Informação e Computação, com um recorte temporal de um pouco mais de 7 anos, entre 2014 e fevereiro de 2021.

O conjunto genérico, composto pelos termos “interoperabilidade”, “semântica” e “patrimônio cultural”, precisou ser adaptado para cada motor de busca presente nas bases selecionadas, além de receber sinônimos possíveis para enriquecimento do conjunto.

As bases selecionadas para pesquisa, conforme Quadro 10, foram 1) *ACM Digital Library* e 2); *IEEE Xplorer Digital Library*; 3) *Web Of Science*(*Web of Knowledge*); e 4) *Scopus (Sciente Direct)*.

**Quadro 10 – Base de Dados Utilizadas na Pesquisa**

Base de dados	Endereço/URL
---------------	--------------

<i>ACM Digital Library</i>	<a href="https://dl.acm.org/search/advanced">https://dl.acm.org/search/advanced</a>
<i>IEEE Xplore Digital Library</i>	<a href="https://ieeexplore.ieee.org/search/advanced">https://ieeexplore.ieee.org/search/advanced</a>
<i>Web Of Science</i>	<a href="http://www.webofknowledge.com">http://www.webofknowledge.com</a>
<i>Scopus (ScienceDirect)</i>	<a href="https://www.sciencedirect.com/">https://www.sciencedirect.com/</a>

Fonte: Elaborado pelo autor.

As palavras-chave da consulta, são, em português: 1) interoperabilidade; 2) semântica; 3) patrimônio; 4) cultural. O composto inicial dessas palavras gera o conjunto de termos {*Interoperabilidade AND semântica AND patrimônio AND cultural*}. Naturalmente, se faz necessário adicionar sinônimos ao conjunto, uma vez que os termos podem ser encontrados de diversas formas nos trabalhos científicos.

O termo interoperabilidade pode ser encontrado escrito em outras formas, assim, necessário adicionar ao conjunto de palavras-chave um subconjunto de sinônimos composto pelos termos “integração”, “interoperação” e “colaboração”. Os termos patrimônio e cultural representam uma área específica, e portanto, devem ser transformados em um só termo “Patrimônio Cultural”.

Desta forma o conjunto de palavras-chave geral ficou com a seguinte estrutura na língua inglesa: ("semantic interoperability" OR "semantic integration" OR "semantic collaboration" ) AND "cultural heritage".

O detalhamento da revisão sistemática está disposto CAPÍTULO III, bem como uma discussão acerca dos trabalhos que foi apresentada na Seção 3.5 (p. 89).

## 5.6 Engenharia de requisitos

Nos ciclos de vida dos *softwares*, em especial nos anteriores à construção de um sistema de informação, é boa prática realizar um processo de modelagem, e com base nisso foram utilizadas as ferramentas de modelagem da engenharia de software (Pressman, 2011; Sommerville, 2011).

Estes autores defendem que a construção de qualquer software passa por um processo de engenharia de requisitos. Isto é, tarefas técnicas que levam a um entendimento do que é preciso no software. A engenharia de requisitos fornece um mecanismo apropriada para compreender o que o *cliente* deseja através de uma análise de suas necessidades, avaliação de viabilidade, busca de solução razoável para as ambiguidades, validação e gestão de etapas do uso (Pressman, 2011).

Por se tratar de uma atividade relacionada à pesquisa-ação, obviamente, o termo *cliente* pode figurar como algo estranho em uma construção colaborativa. Entende-se, no entanto, o cliente como o conjunto de curadores e gestores dos espaços de memória. A Engenharia de requisitos dividiu-se em duas frentes, análise de requisitos e engenharia de ontologias.

### **5.6.1 Análise de requisitos**

A análise de requisitos, segundo Pressman (2011), deve ser conduzida para estabelecer uma base sólida, sendo necessária a utilização de modelos os quais são um conjunto de elementos comportamentais, orientados a fluxo, baseados em cenários e classes.

Seguindo uma estruturação do geral para o detalhe (*top-down*), a modelagem utilizada abordou modelos de casos de uso, de classe e de comportamento, partindo das entrevistas com os curadores e outros profissionais dos espaços de memória como artefatos para geração dos requisitos.

### **5.6.2 Engenharia de ontologias**

Para a construção do protótipo do sistema, levando em consideração que a fundamental questão a ser abordada é semântica, e que a estratégia a ser adotada envolve a utilização de ontologias como estrutura de organização da informação, foi necessária a criação de ontologias, e, portanto, a metodologia utilizada para esse processo foi o método 101 conforme Noy e Mcguinness (2001).

As tarefas básicas na construção de uma ontologia são 1) definir as classes; 2) organizar as classes em uma hierarquia taxonômica; 3) definir slots e descrever os valores permitidos para estes; e 4) preencher os valores de *slots* para as instâncias (Noy & Mcguinness, 2001).

As autoras propõem sete passos interativos para essa construção:

1) Determinar o domínio e o escopo da ontologia; 2) Considerar o reuso de ontologias existentes; 3) Enumerar os termos importantes da ontologia; 4) Definir as classes e sua hierarquia; 5) Definir as propriedades das classes – *slots*; 6) Definir as facetas dos *slots* (tipos possíveis); e 7) Criar instâncias.

Este método 101, além de ser simples, é visivelmente alinhado àqueles comumente aplicados na engenharia de software e estão em harmonia com uma proposta

de protótipo com tecnologia orientada a objetos, inclusive no que se refere à característica incremental da proposta desta tese e ao universo aberto de possibilidades.

Regras fundamentais: 1) Não existe uma maneira correta de modelar um domínio - sempre existem alternativas viáveis. A melhor solução quase sempre depende do aplicativo que você tem em mente e das extensões que você prevê. 2) O desenvolvimento da ontologia é necessariamente um processo iterativo. 3) Os conceitos na ontologia devem ser próximos aos objetos (físicos ou lógicos) e relacionamentos em seu domínio de interesse. É mais provável que sejam substantivos (objetos) ou verbos (relacionamentos) em frases que descrevem seu domínio (Noy & McGuinness, 2001, p. 4).

### **5.7 Construção de protótipos**

Para a construção do protótipo apresentado nesta tese, foram consideradas questões de mais alto nível das ontologias relacionadas à sua organização e conceitos. Regras mais elaboradas e raciocínio a partir destas não foram considerados uma vez que o ponto chave é verificar requisitos pontuais, que foram simplificados para realização de simulações e provas de conceito.

A partir dos requisitos na etapa anterior foi desenvolvido um software orientado a objetos na linguagem Java utilizando de ferramentas da web-semântica. Os resultados dessa etapa estão detalhados nos capítulos VI e VII e seções correspondentes.

### **5.8 Sumário do Capítulo**

Este capítulo teve como objetivo apresentar o conjunto metodológico e os materiais utilizados nesta tese. Para tanto, foi realizado um relato dos métodos e estratégias utilizados para a pesquisa, bem como um breve relato das ferramentas utilizadas especificamente na construção do protótipo.

No próximo capítulo serão apresentados o resultado do inquérito com os curadores do espaço de memória e os artefatos da modelagem do protótipo de Sistema de Informação semântico.

## **CAPÍTULO VI. DA PESQUISA-AÇÃO À ENGENHARIA DE SOFTWARE: MODELAGEM DO PROTÓTIPO**

### **6.1 Introdução**

O objetivo deste capítulo é apresentar o resultado da pesquisa-ação desta tese, que, conforme procedimentos metodológicos explicitados no CAPÍTULO V, parte do contato com os curadores dos espaços culturais da SECULT-GO visando produzir um Sistema de Informação para os espaços de memória.

Na primeira parte, serão apresentados os resultados do inquérito e entrevistas com curadores de espaços cultural e, posteriormente, os artefatos da Engenharia de Software, resultantes do processo de modelagem, identificando e descrevendo os requisitos encontrados em casos de uso.

### **6.2 Resultados do inquérito sobre os espaços de memória da Secretaria de Estado da Cultura**

A Secretaria de Estado da Cultura do Estado de Goiás (SECULT-GO) é provedora de diversos espaços estatais relacionados ao patrimônio cultural, sendo que o inquérito, conforme Seção 5.3 (p. 154) foi realizado com os curadores de acervos de 10 espaços especificamente sob gestão da superintendência de patrimônio.

Dos 10 espaços sob gestão da Superintendência de Patrimônio obtiveram-se 08 respostas dos curadores responsáveis pelos referidos espaços. A primeira questão do inquérito refere-se ao nome do respondente. A segunda e terceira questões referem-se ao nome do espaço e e-mail institucional. Todas essas respostas foram suprimidas.

Os resultados das demais questões são apresentados abaixo.

6.2.1 Questão 4: Pergunta: “*Se existe um site institucional do espaço, informe o endereço?*”

Quanto aos sites institucionais, 1 Curador não respondeu à pergunta. 3 responderam não existir site institucional. Das 4 respostas 1 submete ao endereço geral da própria secretaria ([www.cultura.go.gov.br](http://www.cultura.go.gov.br)), 2 ao conteúdo específico e descritivo do espaço dentro deste domínio, e 1 refere-se ao e-mail institucional e página do Instagram.

6.2.2 Questão 5: Pergunta: “*Qual a composição do acervo?*”

Alusivo à composição do acervo – para conhecer a composição física e os tipos

de objetivos presentes no acervo através de uma descrição sucinta dos acervos –, esta questão buscou identificar como o próprio Curador do espaço apresenta as definições sobre a concepção do acervo. As respostas mostraram um conjunto diverso de acervos, uma vez que compostos por arquivos, bibliotecas, museus e galerias. Os principais objetos do patrimônio cultural encontrados nesses acervos são

a) arquivísticos: documentos históricos governamentais e provinciais, cartas e manuscritos, jornais e comunicações oficiais; documentos e ferramentas; Imprensa e Cinema

b) museológicos: mobiliários, porcelanatos, pratarias e pinturas; arte sacra, ferramentas da história ferroviária; numismática, indumentária, automobilístico, patrimônio botânico, e arquitetônico.

c) bibliotecários: livros literários regionais, nacionais e internacionais, e gibis clássicos.

d) arqueológicos: objetos arqueológicos (fósseis, líticos e outros); etnográficos (plumária, cestaria, cerâmicas, armas, utilitários e outros); geológico e mineral; fauna do cerrado (taxidermia); arte e cultura popular (teares, rodas de fiar, cerâmicas, têxtil e outros).

e) imagéticos e sonoros: Discos, fitas, filmes, documentários e gravações com história oral regional.

### 6.2.3 Questão 6: Pergunta: “*Como o acervo é organizado/inventariado?*”

Com a intenção de verificar existência de inventários e buscar aspectos da organização dos acervos, essa questão permitiu que o Curador discorresse de forma livre como ele consegue visualizar de maneira geral o fluxo de organização do acervo.

Nos arquivos e documentos históricos, prevaleceu a ordem cronológica, alfabética e numérica.

Nas bibliotecas e gibitecas, há predominância de classificações bibliotecárias como a CDU (Classificação Decimal Universal).

Nos museus, prevalece a catalogação por macro coleções em livros ou planilhas de arrolamento e fichas catalográficas e atas.

Quanto a realização de inventários, dos 8 espaços 6 afirmaram realizar inventário

periódico.

6.2.4 Questão 7: Pergunta: “*Quais as características dos objetos do acervo? Quais as principais classes, assuntos, termos e relações existentes?*”

Com a intenção de conhecer a forma dos objetos do acervo para realizar um esquema inicial com classes, termos e relações destes espaços, as respostas foram: acervos históricos documentais, mobiliários históricos pertencentes aos Governadores do Estado de Goiás, acervo livre e aberto de livros, cerrado goiano e sua simbologia, coleções integradas do patrimônio de Goiás, peças históricas ferroviárias, ampliações fotográficas em papel, , negativos flexíveis e de vidro; discos de vinil e cera; fitas magnéticas em diferentes formatos: DVD, DVCan, VHS, SuperVHS, BETA, BetaCan; livros impressos; documentos pessoais de fotógrafos: cartas, bilhetes de viagem, salvo conduto, tíquetes, revistas com publicações/matérias/imagens desses fotógrafos.

Há também documentos do processo histórico, social e urbanístico do desenvolvimento da cidade de Goiânia. Um acervo fonográfico eclético e diversificado que possui a maioria dos títulos oriundos de uma antiga rádio.

No acervo videográfico, há muitos títulos ligados ao Festival Internacional de Cinema e vídeo e de material de uma TV local.

As características das peças do acervo da Gibiteca, com diferentes suportes de registros, são, na sua grande maioria, materiais Impressos. Como em toda biblioteca é essencial que se disponha de um método de organização que permita a localização rápida e eficiente de uma obra, sendo que o acervo é aberto (livre acesso às estantes)

Cada livro que chega à gibiteca é registrado (tombado) e etiquetado. Os livros da Gibiteca são dispostos/separados em duas classes principais Infantil e Infante juvenil, e agrupados nas estantes por ordem alfabética das últimas letras do sobrenome do autor. Os assuntos separados são as obras de Referência divididos em Obras Gerais (dicionários, enciclopédias) artes, religião, biografia, folclore, almanaque, guias turísticos e outros.

Outras classes encontradas foram Equipamentos domésticos, objetos decorativos, objetos comemorativos, objetos religiosos, objetos pessoais, objetos de comunicação.

6.2.5 Questão 8: “*Quais sistemas de informação são utilizados?*”

Pretendeu-se, com esta questão, conhecer quais sistemas de informação foram e estão sendo utilizados, se estão na internet e se há conectividade com outros sistemas. As respostas foram: Sistema de planilhas Microsoft Excel, documentos e planilhas *do Google Docs*, Sistema de Biblioteca *Arches Library* para o caso de biblioteca e gibiteca, redes sociais (*Instagram e Facebook*) para divulgação e Banco de Dados do *Microsoft Access*.

Um dos curadores respondeu não haver qualquer sistema de informação digital. O sistema de recuperação de informações utilizado, no caso, é manual, através das fichas. Outros três espaços, segundo os curadores, possuem registros exclusivamente manuais com fichas catalográficas e documentos esparsos de apoio funcionando como índice para organização. Todos responderam não haver qualquer conectividade.

Quatro curadores informaram utilizar redes sociais para divulgação de alguns aspectos das coleções. Um espaço reportou utilizar o *YouTube* como canal de distribuição audiovisual.

#### 6.2.6 Questão 9: “*Quais as estratégias utilizadas para socialização dos conhecimentos e acesso por parte de pesquisadores e público geral?*”

O intuito dessa questão foi buscar compreender como as informações são compartilhadas de forma digital ou não.

Para o caso dos espaços que possuem arquivos e bibliotecas, não há um compartilhamento digital e a pesquisa é feita no próprio local onde é permitido ao pesquisador tirar foto sem *flash* das informações encontradas no documento. Para o público geral (não pesquisadores), esses espaços também preveem a visita direta.

Em geral, o acesso aos objetos, por parte dos pesquisadores, se dá por meio de consulta local, sendo a estratégia por contato direto solicitada diretamente a um Curador ou servidor do espaço, que faz as buscas necessárias e os direciona os interessados aos locais específicos.

Dois espaços informaram realizar exposições e trabalhos didáticos com escolas e atendimento a público geral por e-mail ou ligações telefônicas.

Na pandemia não há atendimento presencial e nem visitas programadas para pesquisadores, porém, há atendimento remoto e as demandas possíveis são atendidas através do e-mail, telefone, ou mesmo pelas redes sociais. O uso das redes sociais é estratégia de todos os espaços para compartilhar amostras para o público geral.

Os curadores realização exibição de vídeos para estudantes de nível fundamental, médio e superior, pesquisadores e usuários interessados, e neste momento de pandemia como estamos fechados procuramos atender as demandas sem que possível através do e-mail ou mesmo pelas redes sociais.

Em tempos normais (sem pandemia), há intercâmbios com escolas das redes estadual e municipal de ensino, por meio de empréstimos de vídeo que complementam o conteúdo didático; Audição de registros do Acervo Fonográfico; viabilização de cópias digitais do Acervo Fotográfico;

Os curadores realizam também promoção de oficinas nas áreas de conservação de documentos fotográficos, textuais e bibliográficos, organização documental, ação educativa, apoio técnico a instituições nas áreas de conservação de documentos fotográficos, textuais e bibliográficos.

#### 6.2.7 Questão 10: “*Quais projetos/ações de digitalização?*”

Pretendeu-se verificar se há projetos legados ou em andamento de digitalização, e ainda buscar mais informações de possíveis processos.

Quatro dos oito espaços informaram não haver qualquer projeto de digitalização legada ou atual. Há um projeto iniciado projeto de digitalização de acervos de museus e arquivos históricos é cujo o andamento é desconhecido, segundo um dos curadores.

Em um dos espaços há um projeto em andamento na fase de conclusão a digitalização de todo o acervo fotográfico. E tem se o projeto de digitalização do acervo documental. Este trabalho tem como justificativa uma maior proteção aos acervos, pois com a digitalização quando a procura de pesquisadores não à necessidade de seu manuseio, além de que estarão seguros a danos gerados com o tempo, segundo o Curador responsável.

Nos acervos audiovisuais há a produção de um catálogo com parte dos títulos disponíveis no acervo. Quando o pesquisador faz a busca é possível solicitar a cópia digital de parte de uma obra musical e há um aparelho para converter a informação do vinil para *mp3*.

Há ainda um projeto em andamento para transcrição de fitas *Umatic* para DVD e DVCan. Várias fitas em formato VHS também foram transcritas para o formato DVD com aparelho próprio para que o pesquisador ouvisitante possa utilizar. Há também um

projeto de escaneamentos e reprodução fotográfica para os itens com fragilidade no suporte e/ou imagem através de um *scanner* de mesa profissional.

6.2.8 Questão 11: Pergunta: “*O que espera de um sistema de informação para o acervo? Quais funcionalidades precisam ser atendidas?*”

Pretendeu-se levantar as necessidades gerais de sistemas de informação para o espaço no que diz respeito à digitalização, organização e recuperação da informação dos acervos. As respostas dos curadores foram, que no geral os sistemas de informação devem proporcionar as melhores condições de recuperação e segurança da informação, acesso imediato à informação desejada, facilidade de manuseio, qualidade de trabalho e produtividade, gerando maior satisfação e agilidade na busca da informação.

A agilidade nas buscas é lembrada por 7 dos 8 curadores pesquisados. Um dos curadores reitera que o processo deve acontecer on-line, outros que há a necessidade que os acervos sejam interligados a outros em rede.

Quanto à digitalização, os curadores informaram que esperam um sistema que englobe todas as partes de digitalização, organização recuperação da informação dos acervos para facilitar o uso para p pesquisador e para o trabalho da equipe da instituição e que tenha capacidade para abrigar grande quantidade de informação de dados, que possibilite filtragem e cópias de segurança.

Um dos curadores reitera que é necessário que conter informações desde a aquisição e o registro dos objetos até a digitalização, organização e disponibilização das informações do acervo a pesquisadores e público em geral, concentrando todas as informações possíveis no mesmo repositório.

Outro Curador afirma que antes de proceder qualquer digitalização, deve-se procurar saber o que se ganha com essa operação, se vale a pena levá-lo a efeito.

6.2.9 Questão 12: Pergunta: “*Qual o perfil dos usuários?*”

Essa questão teve o intuito de auxiliar na compreensão do perfil do público, suas demandas, e como o público chega ao espaço. A questão obteve as seguintes respostas.

No caso de arquivos, o mais comum é o perfil de pesquisadores das mais diversas áreas de atuação e de todos os níveis educacionais. Raramente há visitantes que não sejam pesquisadores.

No caso de museus são maioria estudantes do ensino básico e de ensino superior

bem como turistas. Há também o perfil de pesquisadores em alguns museus.

Nas bibliotecas e gibitecas há a presença de estudantes em geral e usuários diversos.

6.2.10 Questão 13: Pergunta: “*Se existem, quais dificuldades mais comuns são reportadas pelos usuários?*”

Essa questão teve o intuito de buscar compreender como se dá o fluxo da informação, em especial, nos eventos tecnológicos. As respostas, no geral, foram:

Necessidade de deslocamento para ir ao acervo físico por parte dos pesquisadores.

Ausência de informações físicas, como falta de identificação do local, por exemplo, uma placa indicando que é um espaço patrimonial. Problemas de comunicação do acervo, principalmente sobre etnologia, principalmente quando não há mediador disponível. Há também uma demanda por saber mais sobre a cultura goiana.

Ausência de um *website* em que seja possível coletar informações sem necessidade de ir ao museu físico e outras unidades de informações disponibilizadas para pesquisa na internet.

Várias informações diferentes são encontradas sobre o mesmo assunto (tipos de acervo do museu, por exemplo), quando na verdade apenas a informação correta deveria ser encontrada. Provavelmente isso acontece porque falta uma revisão da própria Comunicação Setorial da Secretaria, que acaba divulgando informações incorretas.

Falta de terminal de consulta e a consulta ao acervo de forma *on-line*. A possibilidade do uso da internet potencializa a atuação das Bibliotecas/Gibitecas para efetuar e ajudar a identificação dos endereços eletrônicos de sites que tratam a pesquisa ou mesmo os assuntos de grande procura.

6.2.11 Questão 14: Pergunta “*Quanto às informações digitais, o que é mais importante no acervo?*”

Pretendeu-se entender a prioridade do que é preciso preservar. Se é a informação descritiva, se são as informações de uso, se é a organização.

No geral as respostas tenderam a reafirmar que todas as informações a respeito do objeto são importantes, que o que pode classificá-la como mais importante ou não é o

olhar do visitante ou pesquisador, pois cada um possui interesses diferentes. Um Curador afirmou que a informação descritiva é a mais importante. Como os sistemas de registro e catalogação iniciais foram superficiais é preciso preservar a pouca informação existente

Outro Curador respondeu que o processo de digitalização de imagens também se mostra importante pois essas não precisariam passar pelo *scanner* mais de uma vez. A indexação de quais imagens já estão digitalizadas, pelo mesmo motivo anterior.

Para o caso das Bibliotecas/Gibitecas a informatização do acervo e do atendimento é imprescindível, porém a do atendimento também é importante para permitir ao usuário uma maior autonomia para realizar sua pesquisa, reservar ou mesmo renovar a obra emprestada.

6.2.12 Questão 15: Pergunta: “*Como considera informações disponíveis em outros acervos? Há integração ou propostas de integração?*”

Com essa questão pretendeu-se entender se é utilizada a interoperabilidade de alguma forma. Dois dos espaços informaram haver interoperação dentro da própria secretaria. Seis informaram não haver nenhuma iniciativa nesse sentido, pois os sistemas são desconectados.

6.2.13 Questão 16: Pergunta: “*Outras informações que julgar pertinente*”.

Nessa questão aberta, descritiva e livre, a intenção foi buscar mais informações não captadas pelas questões anteriores. Das três respostas, uma trouxe mais informações sobre o funcionamento do espaço e suas instalações, bem como sua importância para o patrimônio estadual.

Outro Curador reitera que por atender as necessidades informacionais da população goianiense é preciso que haja um maior empenho em valorizar aquele espaço desde o seu *layout*, instalação, modernização e divulgação.

E um Curador afirma que não existem muitas políticas públicas específicas para os museus e nem respaldo para que as instituições consigam cumprir suas funções científicas, educativas e sociais. Para além disso, empecilhos como a falta de recursos próprios para as instituições museológicas, a falta de valorização dos profissionais dos museus e a interferência político-partidária presente nos museus públicos impossibilitam ações que aumentariam o acesso aos museus e ao direito à memória, que deveria ser

garantido pelo Estado.

### 6.3 Resultado das entrevistas para refinamento e levantamento das tecnologias envolvidas no acervo

O objetivo desta etapa da coleta é refinar as informações obtidas através do questionário, em especial sobre tecnologias utilizadas pelos espaços e possíveis ações de conectividade e interoperação com outras tecnologias e repositórios. Seguindo o roteiro já apresentado na metodologia, obtiveram-se as respostas que permitiram gerar o Quadro 11 abaixo.

**Quadro 11 – Sistemas de Informação presentes nos Acervos**

	Site Institucional	Tour Virtual	Redes Sociais	Acervo Digitalizado	SI para Gestão do Acervo	Armazenamento em SBGD	Catálogo em Planilhas	SI de Acesso ao público	Integração com Outros SI
Arquivo Histórico Estadual	Sim	Não	Sim	Não	Não	Não	Sim	Não	Não
Biblioteca Estadual Pio Vargas	Sim	Não	Sim	Não	Sim	Sim	Não	Não	Não
Gibiteca Jorge Braga	Sim	Não	Sim	Não	Sim	Sim	Não	Não	Não
Museu da Imagem e do Som de Goiás	Sim	Sim	Sim	Parc.	Parc.	Sim	Sim	Não	Não
Museu Ferroviário de Pires do Rio	Sim	Sim	Sim	Não	Não	Não	Sim	Não	Não
Museu Pedro Ludovico	Sim	Sim	Sim	Não	Não	Não	Sim	Não	Não
Museu Zoroastro Artiaga	Sim	Sim	Sim	Parc.	Não	Não	Sim	Não	Não
Palácio Conde dos Arcos	Sim	Sim	Sim	Não	Não	Não	Sim	Não	Não

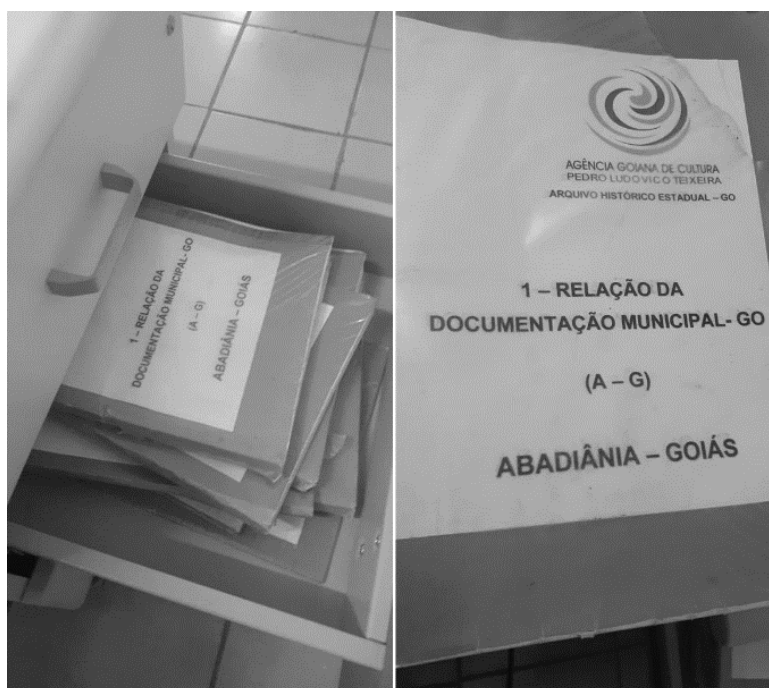
Fonte: Elaborado pelo autor.

Todos os espaços de memória pesquisados possuem sites institucionais estáticos com alguma informação sobre o espaço e acervos. Os museus possuem uma ferramenta

de *tour* virtual que permite uma visão em 360° dos espaços físicos do museu simulando uma visita.

Quanto às redes sociais, todos os espaços atuam nas redes, em especial no contexto da pandemia, realizando divulgações e conteúdos gerados a partir das coleções presentes nos acervos. Apenas dois espaços (Museu da Imagem e do Som de Goiás e Museu Zoroastro Artiaga) possuem acervo digitalizado, ambos, parcialmente.

**Figura 34 – Exemplo de Catálogo Físico**



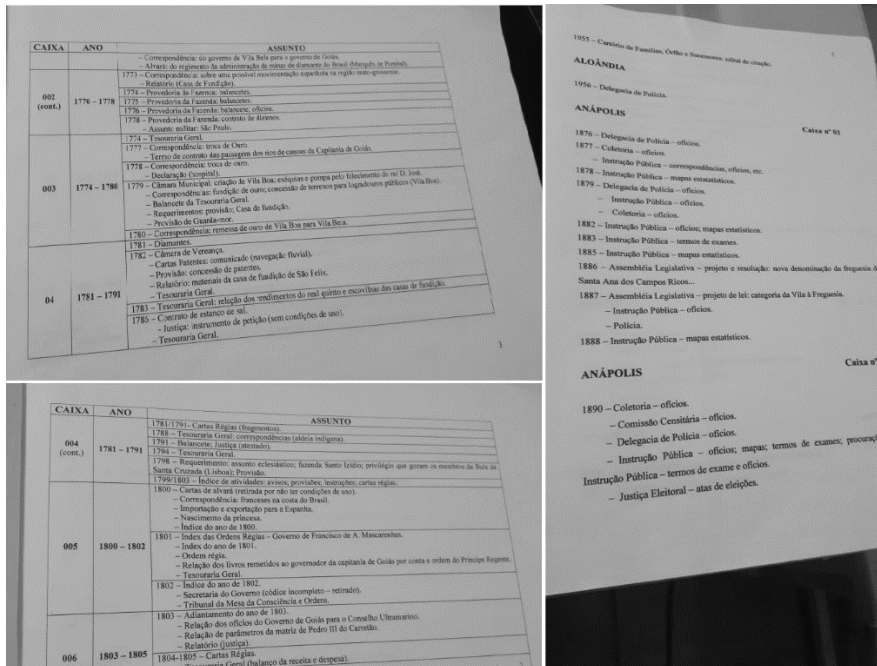
Fonte: Imagem cedida pelo Curador do Acervo.

Quanto a SI para gestão dos acervos, apenas a biblioteca e gibiteca possuem sistema específico que não atende completamente a demanda de gestão, mas permitem a organização e catalogação dos itens de forma interna. Um espaço (Museu da Imagem e do Som de Goiás) possui um SI para gestão de parte do acervo que atende também parcialmente a catalogação com imagens. Os demais não possuem sistema específico para essa finalidade.

Quanto ao catálogo em planilhas, apenas a biblioteca e gibiteca não utilizam. Os demais têm nas planilhas a principal fonte de catalogação. No geral as planilhas são impressas e disponibilizadas ao público, ou a partir dela são gerados catálogos impressos específicos para divulgação dos acervos. A Figura 34 (p. 173) apresenta um exemplo de catálogo físico impresso e disponibilizado para atendimento ao público. Já a Figura 35, abaixo, apresenta planilhas impressas que funcionam como fonte de indexação para os

objetos físicos.

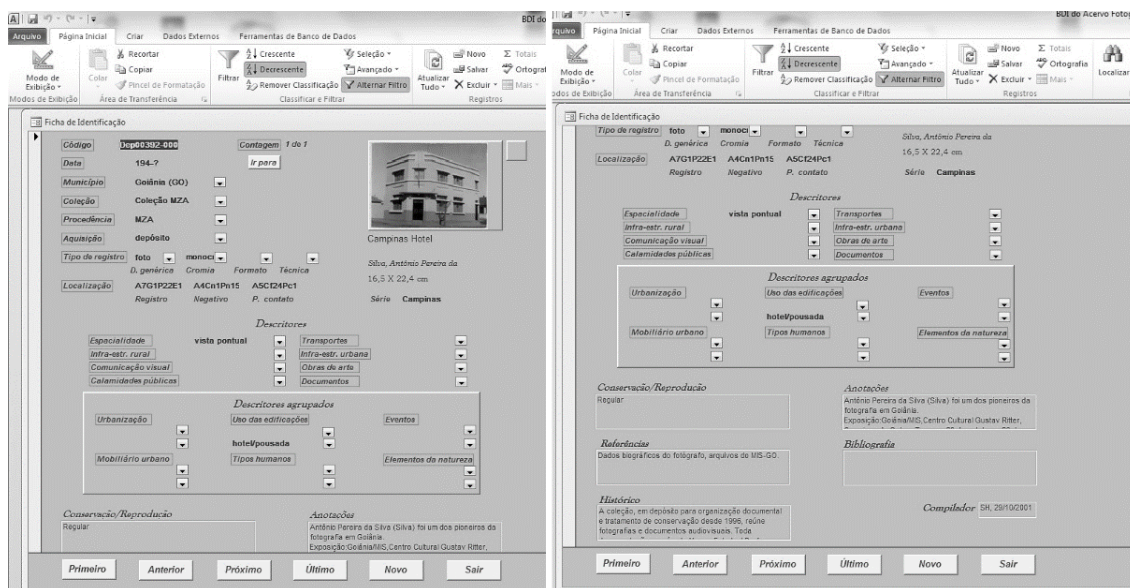
**Figura 35 – Exemplos de planilhas de catálogos**



Fonte: imagens cedidas pela Curadora do Acervo.

Quanto à utilização de SGBD, os três espaços com SI para gestão utilizam algum tipo, os demais não possuem SI, portanto, sem sistemas gerenciadores de bancos de dados. A Figura 36 apresenta uma imagem de um banco de dados relacional Access utilizado para organizar parcialmente o acervo do Museu da Imagem e do Som de Goiás.

**Figura 36 – Banco de dados em Access do Museu da Imagem e do Som de Goiás**



Fonte: Imagem cedida pela Curadora do Acervo.

Já quanto a SI para acesso ao público, tanto interna quanto externamente, nenhum dos espaços o tem. O que também não permite que haja qualquer integração com outros sistemas

#### **6.4 Modelagem do Protótipo de Sistema de Informação para os Espaços de Memória**

Como uma forma de responder ao problema abordado nessa tese e reforçado através do inquérito detalhado na Seção anterior, no CAPÍTULO IV foi apresentada uma proposta de modelo de sistema de informação semântico para centros de memória que leva em consideração o problema, as ferramentas e o estado da arte explanados nas subseções e capítulos anteriores.

Nessa Seção é apresentada de forma complementar, uma modelagem para operacionalização do protótipo em vistas de atender a proposta em um recorte específico no que tange à organização dos acervos e, especificamente, sua capacidade de interoperabilidade.

Na mesma direção da abordagem do modelo explicitada no CAPÍTULO IV, são descritos os artefatos em uma estratégia do nível mais alto para o nível mais baixo de detalhe e apresenta-se, para cada requisito ou funcionalidade extraída, uma visão geral, uma comportamental, uma visão dos fluxos, uma visão estrutural e uma visão de interação humano-computador.

Para a engenharia de requisitos desse protótipo, foram consideradas as tarefas relacionadas à digitalização e registro de acervo que normalmente são realizadas por curadores ou servidores dos espaços de memória.

Isso apoiando-se em Doerr (Doerr, 2009, p. 4) ao afirmar que “os curadores fornecem descrições básicas dos objetos que servem para sua identificação e manuseio, mas também pesquisam e justificam sua relevância”. Isto é, é responsável pelos “o quê” dos objetos e “por quê” cada um dos objetos é mantido.

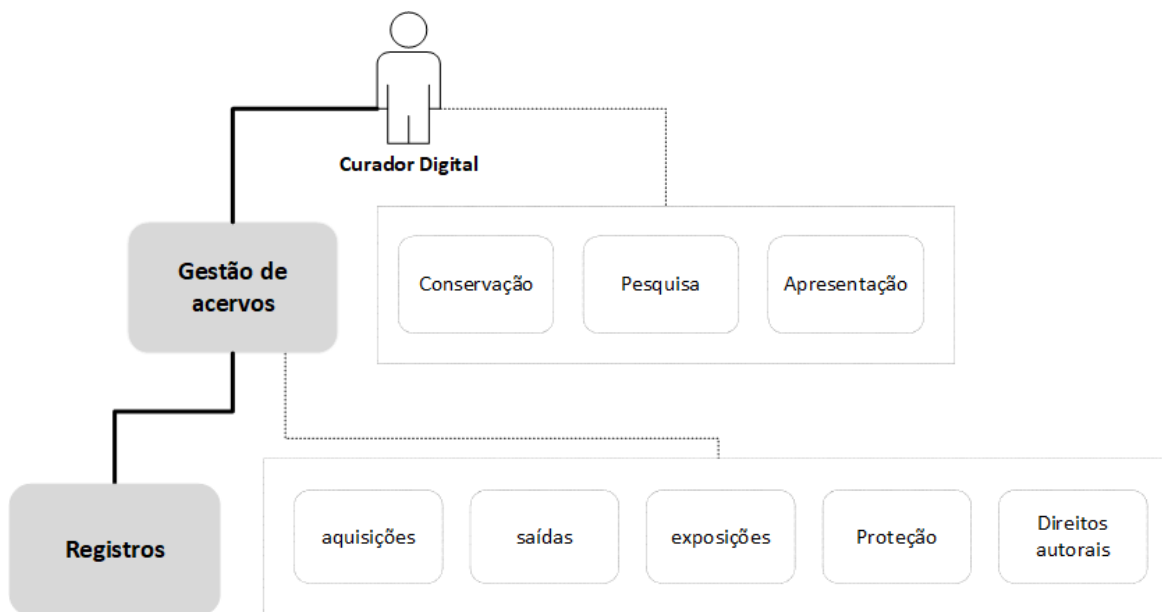
Considerando esse pressuposto e para simplificação da modelagem, os atores externos presentes nas atividades dessa funcionalidade são representados pela figura de um Curador a qual é tratada de forma genérica.

As tarefas de conservação, pesquisa, apresentação e outras relacionadas às atividades comuns dos Sistemas de Informação do Patrimônio Cultural não são cobertas

por esse protótipo.

A Figura 37 apresenta uma visão geral das funcionalidades comuns de sistemas de informação para o patrimônio cultural, com destaque à Gestão do Acervo e o processo de digitalização e registro que são cobertos por esse recorte, sem deixar de reconhecer a importância e urgência dos demais processos e funcionalidades.

**Figura 37 – Gerenciar Acervos Digitais**



Fonte: elaborada pelo Autor.

Conforme a figura acima, para a construção desse protótipo, considerou-se o requisito principal “Gestão do acervo” que foi nominado Gerenciar Acervos Digitais e se desdobrou em 6 outros requisitos funcionais conforme tabela de requisitos abaixo apresentada no Quadro 12.

**Quadro 12 – Tabela de Requisitos da funcionalidade Gerenciar Acervos Digitais**

Id.	Requisito	Descrição
R1	Definir Acervos	Funcionalidade para que o Curador possa gerenciar a estrutura do espaço em termos de criar, modificar ou descartar classes de seus acervos.
R2	Gerenciar Objetos Físicos	Nessa funcionalidade, o Curador pode adicionar, modificar ou destruir registros dos objetos digitais físicos
R3	Gerenciar Objetos Dimensionais	Nessa funcionalidade, o Curador pode adicionar, modificar ou destruir registros dos objetos digitais dimensionais, que dizem respeito a 4 dimensões (Pessoa, Fenômeno, Espaço e Tempo)

R4	Adicionar/remover Propriedades	Nessa funcionalidade, o Curador deve poder adicionar ou remover propriedades dos Objetos Físicos e Dimensionais já persistidos no seu repositório.
R5	Coletar ligações	O algoritmo de interoperabilidade deve varrer a lista de <i>repositórios amigos</i> afim de sugerir possíveis ligações entre os objetos e propriedades manipulados pelo Curador.
R6	Administrar repositório	Nessa funcionalidade, o Curador mantém os dados do repositório e a lista de <i>repositórios amigos</i> .

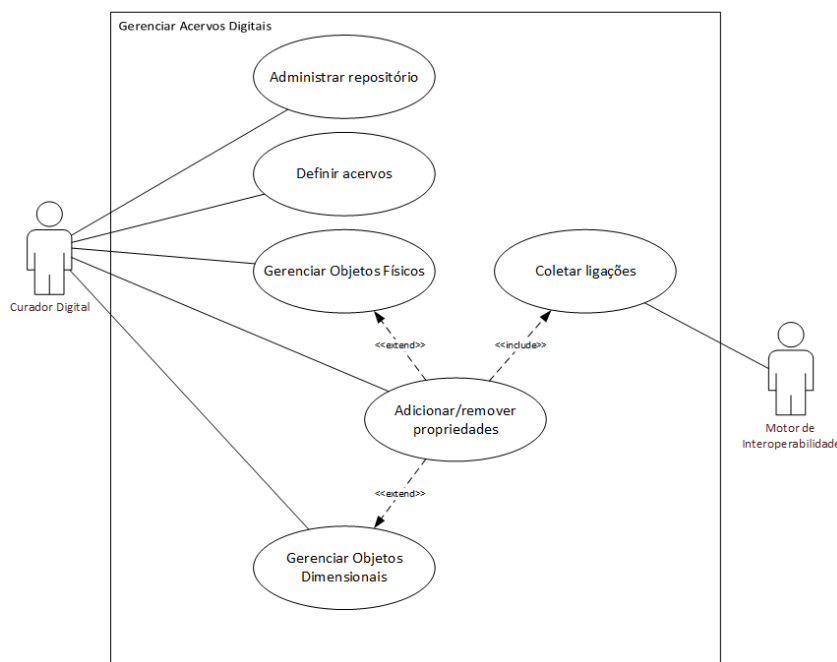
Fonte: Elaborado pelo autor.

Seguindo na estratégia de extração dos requisitos, os requisitos de R1 a R6 foram modelados em casos de uso para melhor detalhamento e visualidade gráfica.

O Curador administra o repositório, mantendo os dados deste e a lista de *repositórios amigos* (que será utilizada pelo Motor de Interoperabilidade - MI). O Curador ainda define os acervos, classificando-os de acordo com sua organização interna. Este mesmo Curador gerencia os objetos digitais (Físicos e Dimensionais), nos quais podem adicionar ou remover propriedades que disparam a coleta de ligações por meios dos algoritmos presentes no MI.

A Figura 38 apresenta o diagrama de Casos de Uso, extraído da tabela de requisitos construída no passo anterior.

**Figura 38 – Diagrama de Caso de Uso de Gerenciar Acervos**



Fonte: elaborada pelo Autor.

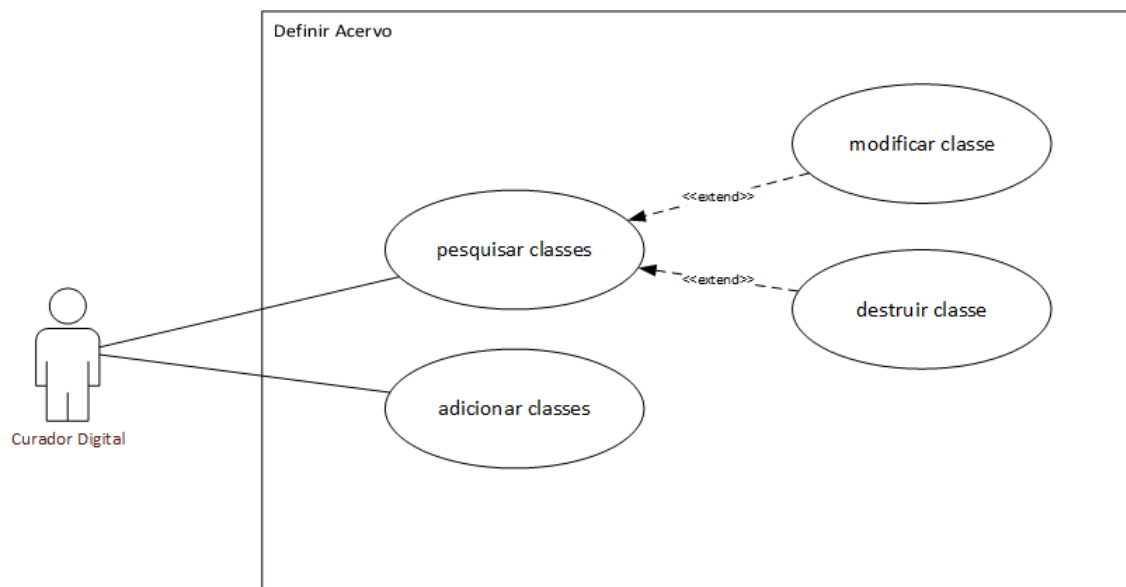
A seguir, são apresentadas as funcionalidades especificadas no Digrama de Casos de Uso da funcionalidade Gerenciar Acervos Digitais, em forma de detalhamento de casos de uso.

#### 6.4.1 Definir Acervos

Nessa funcionalidade, o Curador cria e mantém um conjunto de classes com base na ontologia OWL inicial (Seção 4.7.1 p. 135) que contém a estrutura do acervo em torno das quatro referências básicas: *owl:class*, *rdfs:label*, *rdfs:subClassOf* e *rdfs:comment*.

Atendendo ao que fora especificado na Seção 4.7.1, o elemento *owl:class* é a marcação inicial de cada classe definida na ontologia; *rdfs:label* permite definir um nome para classe do acervo; *rdfs:comment* é usada para adicionar uma descrição dessa classe e em *rdfs:subClassOf* se permite que as relações de herança nas duas direções (subclasses e superclasses) possam ocorrer.

**Figura 39 – Diagrama de Caso de uso de Definir Acervo**



Fonte: Elaborada pelo Autor.

A Figura 39 acima, apresenta os subprocessos resultantes do caso de uso principal Definir Acervo, pelo qual foi gerado um novo diagrama de caso de uso em menor nível, apontando quatro funcionalidades que são detalhadas no Quadro 13 abaixo.

O ator Curador, inicialmente, pesquisa por classes, informando ou não algum conjunto de termos e, a partir daí pode, acionar as funcionalidades de alteração e exclusão

de uma determinada classe. Também é possível adicionar uma nova classe ao seu acervo.

**Quadro 13 – Especificação do Caso de Uso Definir Acervo**

R1	Definir Acervos
Ator principal:	Curador
Fluxo de Eventos	<ol style="list-style-type: none"> <li>1. O Curador aciona o menu de classes</li> <li>2. O Curador aciona a opção pesquisar classes             <ol style="list-style-type: none"> <li>2.1 O Curador informa algum dado para pesquisa e clica em pesquisar</li> <li>2.2 O sistema carrega os dados das classes conforme ontologia inicial em uma tabela</li> </ol> </li> <li>3. O Curador seleciona uma classe e pressiona a opção alterar             <ol style="list-style-type: none"> <li>3.1 O sistema carrega uma página com a classe selecionada</li> <li>3.2 O Curador informa os dados a serem alterados</li> <li>3.3 O Curador pressiona a opção gravar</li> <li>3.4 O sistema valida consistência das informações</li> <li>3.5 O sistema persiste as informações, caso a validação seja positiva</li> </ol> </li> <li>4. O Curador seleciona a opção destruir classe             <ol style="list-style-type: none"> <li>4.1 O sistema verifica o uso da classe</li> <li>4.2 O sistema destrói a classe, caso a verificação seja negativa.</li> <li>4.3 O sistema avisa “classe em uso” caso a verificação seja positiva.</li> </ol> </li> <li>5. O Curador seleciona a opção “Criar nova classe”             <ol style="list-style-type: none"> <li>5.1 O Curador seleciona a classe de origem do acervo (superclasse)</li> <li>5.2 O Curador informa o rótulo e a descrição da nova classe</li> <li>5.3 O Curador clica em “enviar”</li> <li>5.4 O sistema verifica a consistência dos dados</li> <li>5.5 O sistema persiste os dados, em caso positivo</li> </ol> </li> </ol>
Exceções	<ol style="list-style-type: none"> <li>1. Ao tentar adicionar uma classe já existente o sistema emite um aviso</li> <li>2. A requisição não ser aceita pelo TDB por alguma má formação gramatical da SPARQL</li> <li>3. A ontologia inicial não ser localizada no TDB.</li> <li>4. Destruir uma classe em uso: emitir aviso de proibição</li> </ol>
Regras de Negócio	<ol style="list-style-type: none"> <li>1. O nome de cada classe ou subclasse criada é definido pela propriedade <i>rdfs:label</i></li> <li>2. A descrição de cada classe ou subclasse criada é definido pela propriedade <i>rdfs:comment</i></li> <li>3. A relação entre classes e subclasses, partindo da classe mais alta (coisa) é definida pela propriedade <i>rdfs:SubClassOf</i></li> <li>4. Todos os elementos precisam ser informados</li> <li>5. Os nomes de classe definidos pela propriedade <i>rdfs:label</i> devem ser únicos dentro de uma hierarquia <i>subClassOf</i></li> </ol>
Pré-Condições	A ontologia inicial precisa estar presente no TDB.
Pós-Condições	Os dados válidos informados devem ser persistidos no TDB

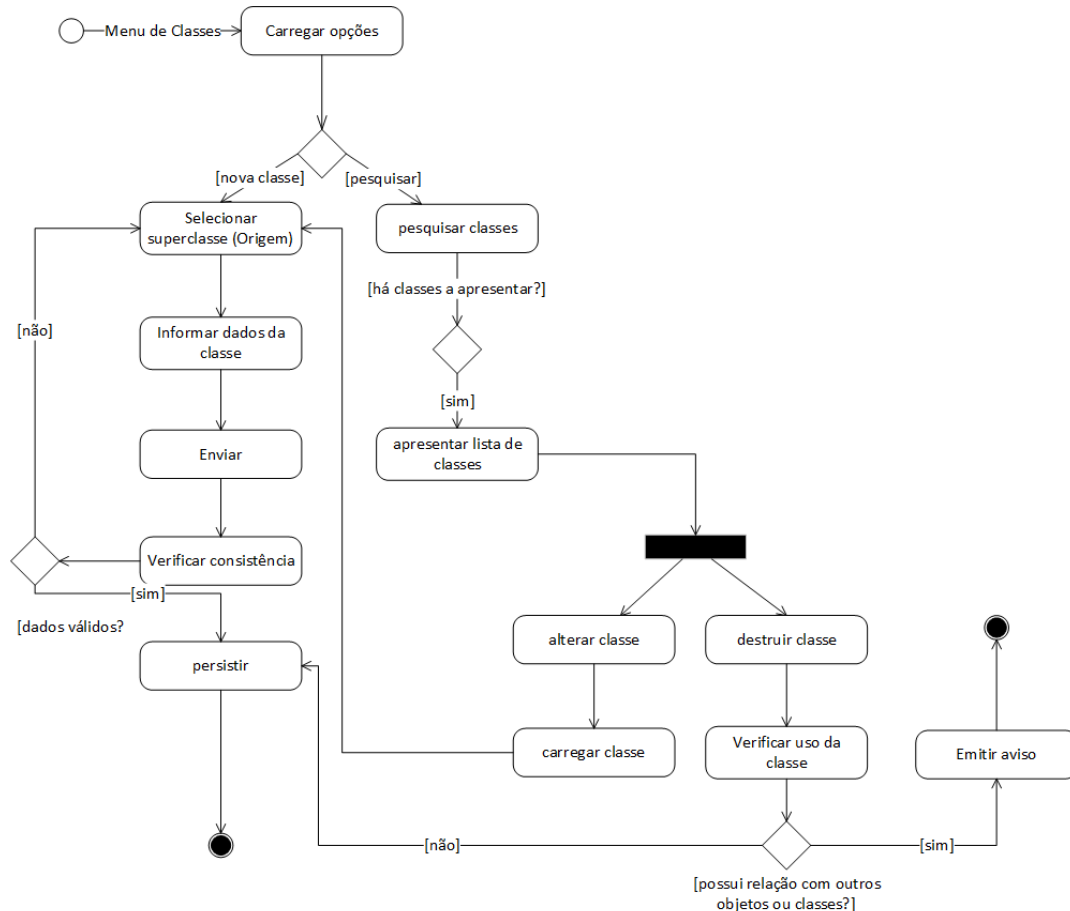
Fonte: elaborado pelo autor

Um fluxo geral, que pode ser visualizado na Figura 40, apresenta um diagrama de atividades relacionado à funcionalidade em questão. O ponto inicial do fluxo condiciona duas possibilidades. A primeira, para o caso de o Curador desejar adicionar uma nova classe, que após confirmar os dados, esses são enviados para uma atividade de verificação. Se os dados forem válidos, é acionada uma atividade para persistir os dados no TDB.

A segunda possibilidade inicial do fluxo é iniciar acionando a atividade de pesquisa, que, por sua vez, aciona o motor de acesso ao TDB e caso a consulta resulte em conjunto não vazio, carrega as classes e espera uma das atividades próximas que acionarão a funcionalidade para alterar ou destruir a classe.

Para o caso de alteração de dados, a atividade *carregar dados da classe* permite que uma página com os dados da classe seja apresentada para o usuário.

**Figura 40 – Diagrama de Atividades de Definir Acervo**



Fonte: Elaborada pelo Autor.

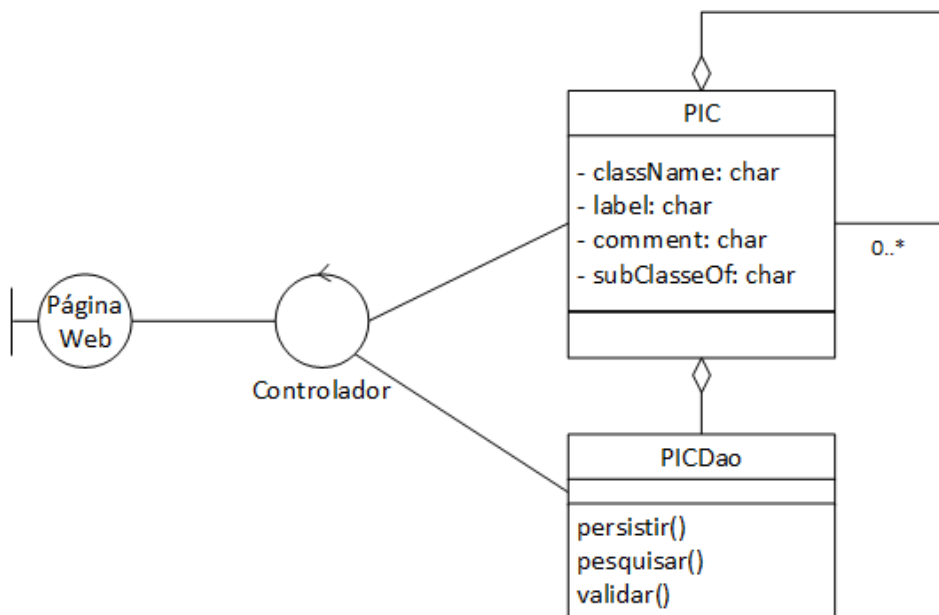
#### 6.4.1.1 Pesquisar classes

A atividade inicial dessa funcionalidade, Pesquisar Classes, deve verificar se existem dados a serem apresentados através de uma busca na ontologia inicial. Caso haja, carrega os dados em tela, e apresenta as funcionalidades de alterar e excluir as classes listadas.

A classe PIC (*Plain Initial Class*) carrega a estrutura e o comportamento da instância de cada Classe que será transformada em tripla após a persistência no referido TDB. As páginas Web são representadas por uma classe de fronteira (interação humano-computador) e os serviços acionados a partir dos eventos disparados pela página são representados pela classe Controlador. A classe PICDao possui os componentes utilizados para a conexão com o TDB.

A Figura 41 abaixo apresenta um diagrama com o esquema estrutural das classes recortado exclusivamente essa funcionalidade.

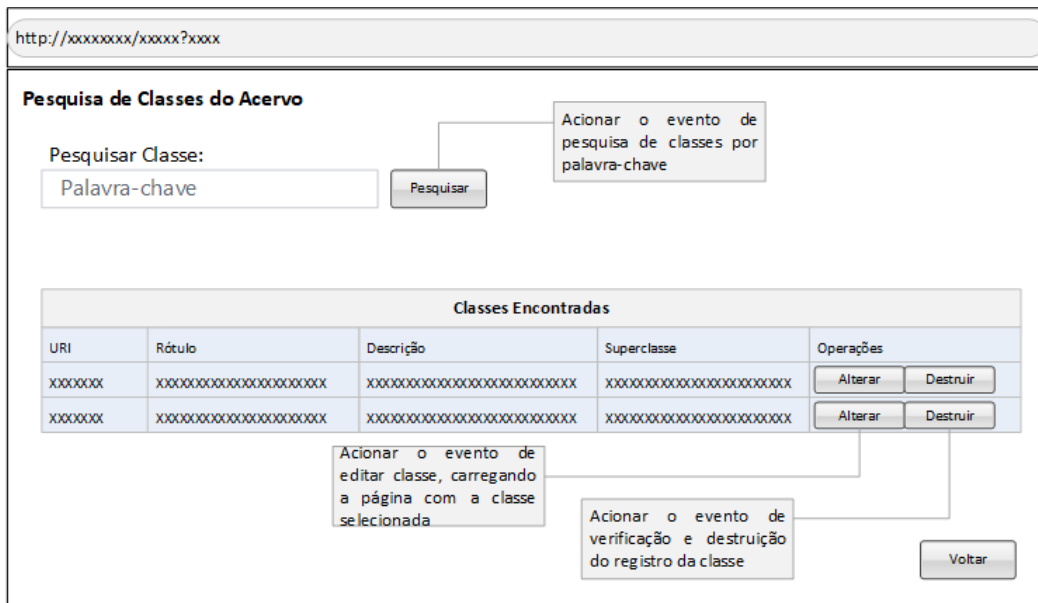
**Figura 41 – Esquema Estrutural de Definir Acervo**



Fonte: Elaborada pelo Autor.

Na Figura 42, é apresentada uma visualização da interação humano-computador de acordo com os diagramas acima. O usuário (aqui denominado Curador) aciona a funcionalidade através de uma navegação simples e pode realizar as buscas de termos constantes em qualquer uma das propriedades de cada classe do acervo de acordo com a ontologia inicial.

**Figura 42 – IHC – Pesquisar Classes**



Fonte: Elaborada pelo Autor.

As classes existentes no TDB devem ser retornadas de acordo com os critérios de busca. Um grafo RDF constante das triplas presentes no banco de dados permite que cada tripla correspondente seja transformada em um objeto que deve ser modelado em *sujeito-propriedade-objeto* de acordo com o módulo das entidades do protótipo. Caso o usuário acione a funcionalidade de editar a classe selecionada, uma página com os dados do grafo referente deverá ser aberta de acordo com a funcionalidade para modificar Classe.

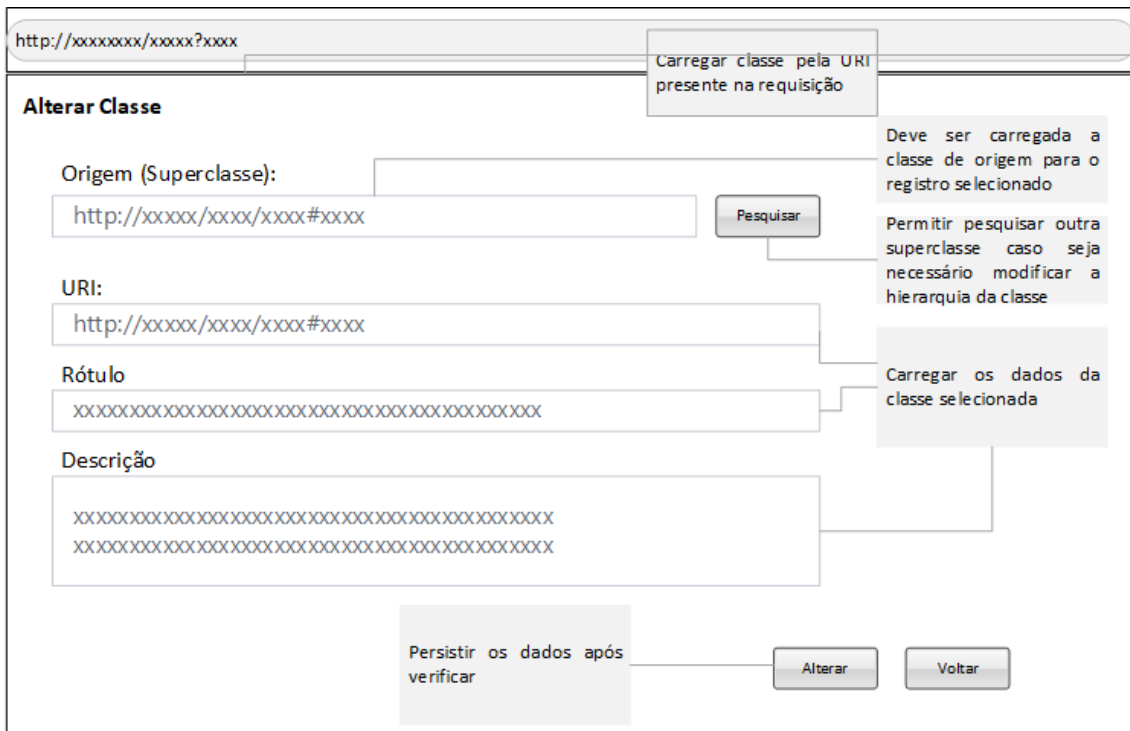
#### 6.4.1.2 Modificar Classe

A tarefa de modificação carrega os dados da classe selecionada com os dados detalhados para serem modificados. Havendo confirmação, há uma verificação se os dados são válidos antes de enviar para a tarefa de persistência destes. Caso contrário, descarta a operação e finaliza o processo.

A atividade de destruição (exclusão) verifica se há uso da classe em algum objeto nos TDBs referentes aos objetos do acervo. Caso haja, não permite a operação.

A Figura 43 apresenta uma construção de uma interação-humano computador para a funcionalidade de Modificar Classe.

**Figura 43 – IHC – Modificar Classe**

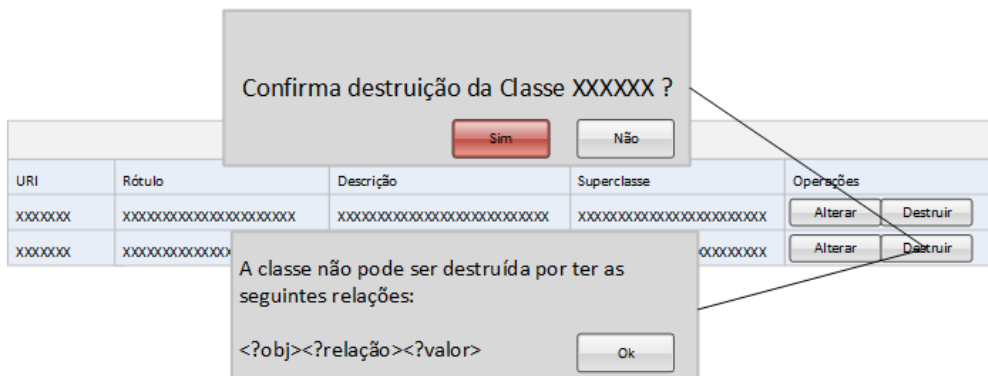


Fonte: Elaborada pelo Autor

#### 6.4.1.3 Destruir Classe

Para o caso de o Curador precisar destruir (excluir) uma classe, este deve acionar o botão de destruição (Destruir) para a classe selecionada na lista de classes carregada na pesquisa. Assim, após confirmação, um processo de verificação de utilização da classe é acionado para garantir a integridade das relações já existentes. Se Classe escolhida para ser destruída é utilizada para descrever algum objeto ou tenha alguma subclasse relacionada, a exclusão é descartada. Caso contrário, há uma confirmação de exclusão com persistência no TDB.

**Figura 44 – IHC – Destruir Classe**



Fonte: Elaborada pelo Autor.

#### 6.4.1.4 Adicionar Classes

A funcionalidade para adicionar uma nova classe apenas busca do usuário os dados referentes àquela classe conforme a determinação inicial das propriedades básicas e verifica se há alguma classe com as mesmas definições. Caso não haja, e os dados estão completos, persiste para uma nova classe no TDB referente às classes.

Essa ação está presente em um *botão* cuja *action* deve acionar um objeto instanciado na classe de controle (Controlador), que representa todas as classes controladoras no módulo de serviços que são responsáveis por gerenciar os eventos. Nesse controlador, há referências para uma instância da classe PICDao, responsável pelas ações de acesso ao banco de dados, também com instância da classe PIC, em que estão as instâncias mais baixas relacionadas ao modelo do negócio que transforma instâncias de classes em triplas RDF e vice-versa.

Em qualquer tempo, o Curador pode adicionar novas classes ao acervo. E pode fazer isso através da opção “Adicionar Nova Classe”. Conforme a Figura 45 abaixo, o Curador deve preencher todos os campos da página apresentada. A pesquisa da classe hierarquicamente superior àquela que está sendo adicionada precisa ser informada.

**Figura 45 – IHC – Adicionar Classes**

The screenshot shows a web browser window with the URL `http://xxxxxxxx/xxxxx?xxxx`. The page title is "Adicionar uma nova Classe". The form contains the following elements:

- Origem (Superclasse):** A text input field containing `http://xxxxx/xxxxx/xxxx#xxxx` and a "Pesquisar" button.
- URI:** An empty text input field.
- Rótulo:** An empty text input field.
- Descrição:** A larger empty text input field.
- Buttons:** "Pesquisar", "Inserir", and "Voltar".
- Annotations:** A callout box says "O usuário deve selecionar uma superclasse na lista" pointing to the search button. Another callout box says "Preenchimento obrigatório" pointing to the URI, Rótulo, and Descrição fields.
- Footer:** A box at the bottom contains the text "Acionar serviço de verificação e persistência".

Fonte: Elaborada pelo Autor.

#### 6.4.2 Gerenciar Objetos Físicos

Nessa funcionalidade, o Curador pesquisa, adiciona um novo, ou modifica

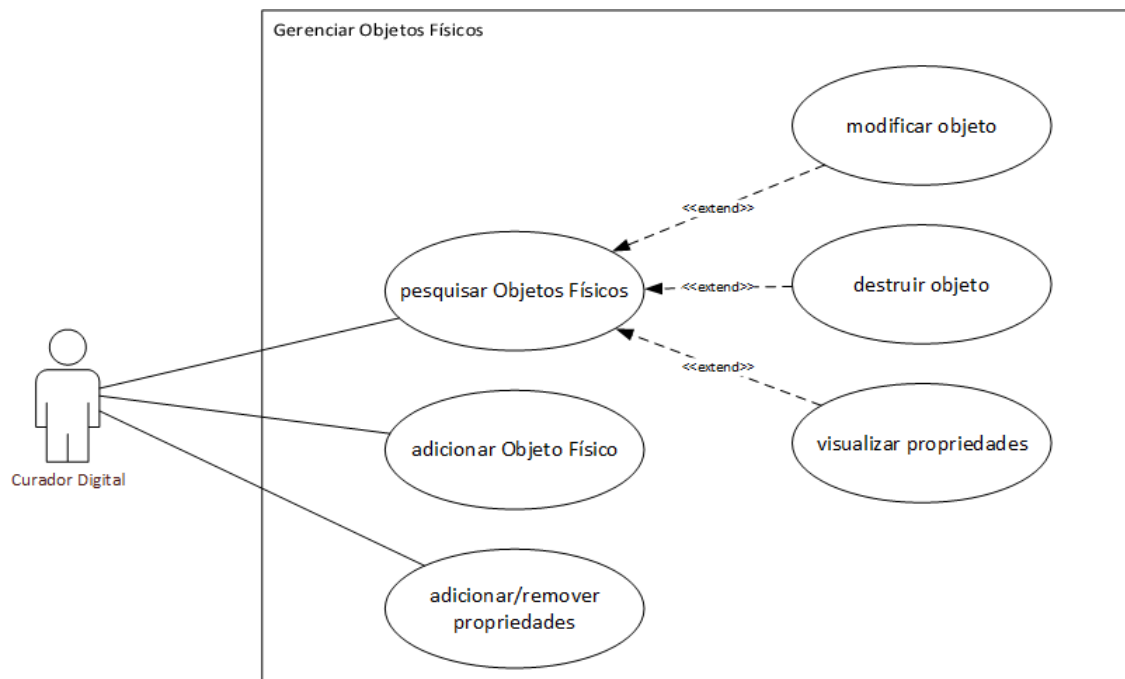
objetos digitais físicos pertencentes ao seu acervo, inclusive suas propriedades. Trata-se da etapa de criação dos objetos do espaço de memória propriamente ditos.

Os objetos representam os itens do acervo e são classificados de acordo com a estrutura de classes definida anteriormente. Importante reforçar que nesse protótipo as classes e propriedades são simplificadas conforme o modelo já apresentado.

É possível ainda visualizar todas as relações do objeto e também adicionar ou remover relações através das propriedades pré-existentes na ontologia inicial ou qualquer uma outra adicionada pelo Curador e que faça referência a algum objeto na rede.

A Figura 46 apresenta a subfuncionalidade derivada da funcionalidade principal Gerenciar Objetos Digitais, que resulta em seis processos em nível mais detalhado.

**Figura 46 – Caso de Uso Gerenciar Objetos Físicos**



Fonte: Elaborada pelo Autor.

Ao informar os dados do objeto, o ator Curador escolhe o tipo de acervo a que está relacionado o objeto que será inserido, além de informar o nome e uma descrição básica (assunto). Um identificador será gerado automaticamente. Após confirmação, os dados são persistidos caso sejam confirmados por um validador. O Quadro 14 apresenta um detalhamento da funcionalidade Gerenciar Objetos Físicos.

**Quadro 14 – Especificação do Caso de Uso Gerenciar Objetos Físicos**

R2	Gerenciar Objetos Físicos
----	---------------------------

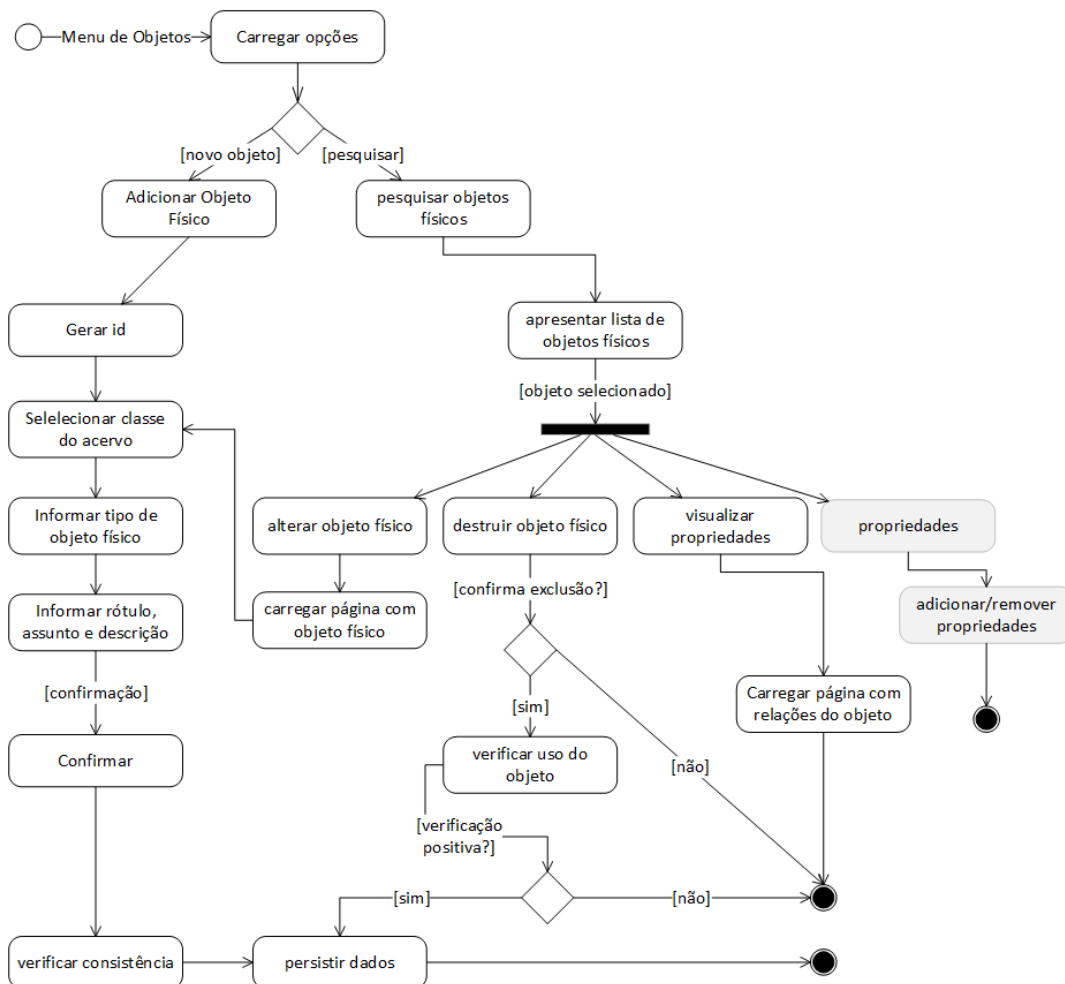
Ator principal:	Curador
Fluxo de Eventos	<ol style="list-style-type: none"> <li>1. <i>O Curador acessa a página de pesquisa de objetos</i> <ol style="list-style-type: none"> <li>1.1 <i>O Curador informa critérios de busca e aciona o botão pesquisar</i></li> <li>1.2 <i>O Sistema carrega uma lista de objetos de acordo com os critérios</i></li> </ol> </li> <li>2. <i>O Curador seleciona um objeto na lista e aciona o botão editar</i> <ol style="list-style-type: none"> <li>2.1 <i>O objeto selecionado é carregado em uma página</i></li> <li>2.2 <i>O Curador informa os dados do objeto que deseja alterar</i></li> <li>2.3 <i>O Curador seleciona a classe do acervo a qual objeto faz referência</i></li> <li>2.4 <i>O Curador clica em gravar</i></li> <li>2.5 <i>O sistema verifica a consistência dos dados, e em caso positivo, persiste os dados</i></li> </ol> </li> <li>3. <i>O Curador seleciona um objeto na lista e aciona o botão destruir</i> <ol style="list-style-type: none"> <li>3.1 <i>É emitida uma mensagem crítica confirmando a exclusão</i></li> <li>3.2 <i>Caso haja confirmação, o sistema verifica se o objeto possui alguma relação com outro objeto</i></li> <li>3.3 <i>Caso não haja nenhuma relação, o sistema persiste os dados.</i></li> </ol> </li> <li>4. <i>O Curador seleciona a opção visualizar propriedades objeto</i> <ol style="list-style-type: none"> <li>4.1 <i>O sistema carrega a página com todas as relações do objeto</i></li> </ol> </li> <li>5. <i>O Curador seleciona a opção de adicionar novo objeto físico</i> <ol style="list-style-type: none"> <li>5.1 <i>Sistema Gera identificador único</i></li> <li>5.2 <i>Seleciona um tipo de classe do acervo (rdf:type)</i></li> <li>5.3 <i>Informa o rótulo, assunto e descrição do objeto</i></li> <li>5.4 <i>Curador confirma envio.</i></li> <li>5.5 <i>Sistema verificar se dados são válidos</i></li> <li>5.6 <i>Persistir dados</i></li> </ol> </li> <li>6. <i>O Curador seleciona a opção adicionar remover propriedades</i> <ol style="list-style-type: none"> <li>6.1 <i>O sistema carrega a funcionalidade R5</i></li> </ol> </li> </ol>
Exceções	<ol style="list-style-type: none"> <li>1. <i>Usuário não informar o tipo: deve emitir aviso</i></li> <li>2. <i>Sistema não gerar o identificador único</i></li> <li>3. <i>Falha na persistência</i></li> <li>4. <i>Dados inconsistentes</i></li> </ol>
Regras de Negócio	<ol style="list-style-type: none"> <li>1. <i>Um identificador único e inequívoco deve ser gerado para cada objeto digital</i></li> <li>2. <i>O nome de cada objeto é definido pela propriedade dc:title</i></li> <li>3. <i>Uma assunto de cada objeto digital deve ser informada</i></li> <li>4. <i>O campo tipo, rdf:type, deve ser informado</i></li> <li>5. <i>O campo dc:description deve ser preenchido com uma descrição do objeto</i></li> <li>6. <i>Os nomes (dc:title) dos objetos devem ser únicos em uma hierarquia direta</i></li> </ol>
Pré-Condições	<ol style="list-style-type: none"> <li>1. <i>A ontologia inicial precisa estar presente no TDB</i></li> <li>2. <i>A classe correspondente ao tipo do acervo do objeto precisa estar presente</i></li> </ol>
Pós-Condições	<ol style="list-style-type: none"> <li>1. <i>Os dados válidos informados devem ser persistidos no TDB</i></li> </ol>

Fonte: Elaborado pelo autor.

Na Figura 47, o fluxo principal pode ser visualizado. O Curador acessa o menu de objetos físicos e tem à sua frente as opções de pesquisar objetos ou adicionar um novo objeto. Caso decida por adicionar um novo, uma página de cadastro é carregada para o Curador informar os dados iniciais do Objeto Físico.

O sistema gera automaticamente um identificador único para o objeto que está sendo criado. O Curador então deve selecionar em qual classe do acervo aquele objeto se enquadra. Posteriormente, informa qual o Tipo Físico do objeto e então informa os dados básicos, a iniciar por um nome (*label*), um assunto (*subject*) e uma descrição resumida do objeto (*description*). A classe do acervo (*rdfs:type*) também precisa ser informada.

**Figura 47 – Diagrama de Atividades de Gerenciar Objetos Físicos**



Fonte: Elaborada pelo Autor.

Somente após essas informações, é possível confirmar os dados que serão enviados através da atividade Confirmar Envio. Uma verificação é realizada e, caso seja

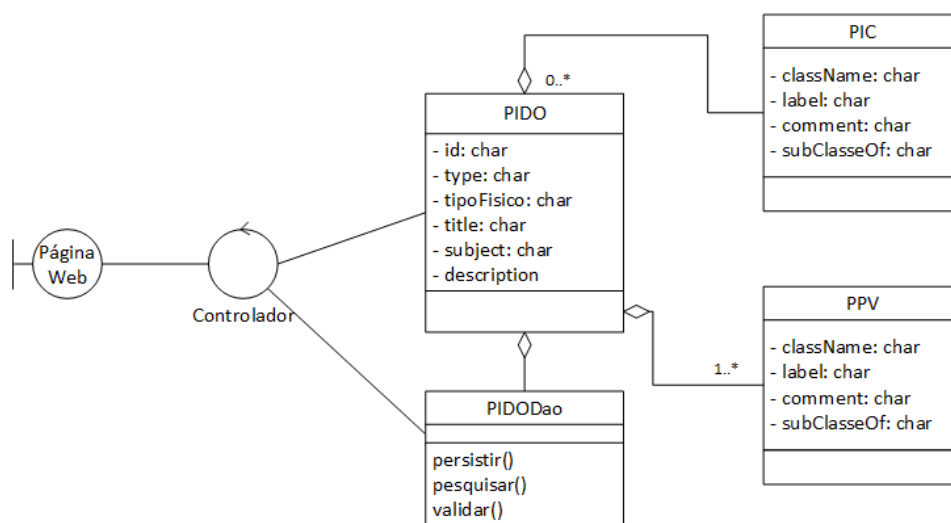
positiva, os dados são enviados para serem persistidos. Caso não seja válida, uma mensagem de erro será enviada para chegar até o Curador responsável pela ação.

Caso a opção do Curador seja por pesquisar objetos físicos, ele deve informar algum critério inicial de busca e o sistema apresenta uma lista dos Objetos Físicos que atendem ao critério informado. Caso deseje alterar o objeto, o Curador aciona o comando de alterar objeto físico para informar os dados a serem alterados e que serão persistidos caso a verificação seja positiva. Para o caso de desejar “destruir” (excluir) um Objeto Físico, o Curador deve pressionar o botão correspondente para o objeto selecionado. Em se confirmando o desejo de excluir o objeto, o sistema verifica se este é usado de alguma maneira (pesquisa de relações com outros objetos). Caso a verificação seja negativa, os dados são persistidos. Em caso positivo um alerta é apresentado ao Curador e a exclusão descartada.

Se o Curador desejar visualizar todas as relações do objeto, basta acionar a opção Visualizar Propriedades para apresentar uma página com todas as propriedades do objeto com as ligações para as URIs correspondentes de cada propriedade ou a informação textual do valor correspondente da relação na tripla (como um rótulo ou uma descrição). Se a opção do Curador for por adicionar ou remover uma propriedade, a funcionalidade específica é acionada e aquele Curador é enviado para a página correspondente.

Um diagrama de classes para a funcionalidade Gerenciar Objetos Físico pode ser visualizado na Figura 48.

**Figura 48 – Diagrama de Classes Gerenciar Objetos Físicos**



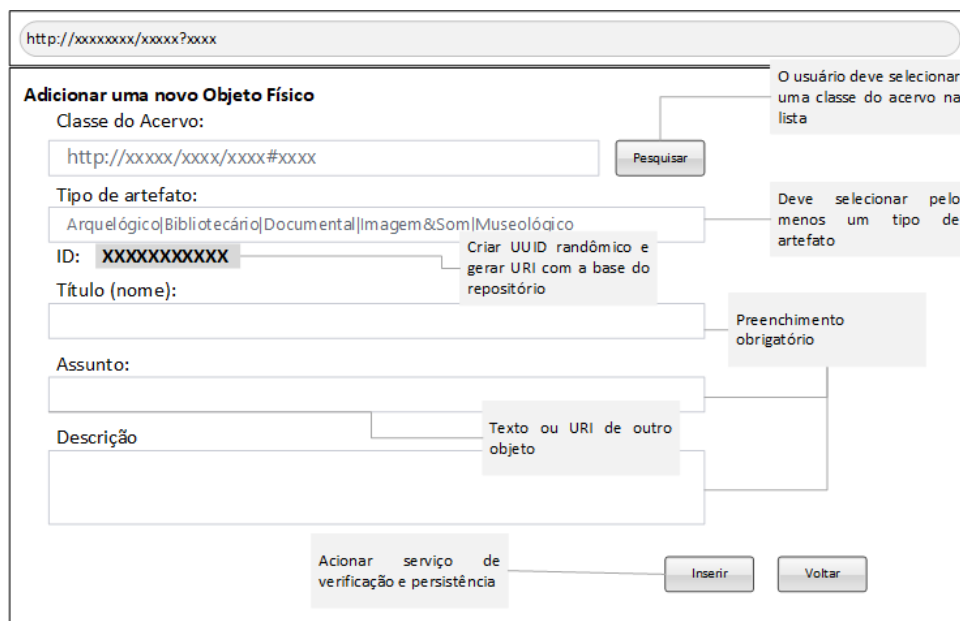
Fonte: Elaborada pelo Autor.

A classe PIDO (*Plain Initial Data Object*) possui os atributos definidos para os objetos e cada instância está relacionada com uma agregação à classe inicial (PIC) com a multiplicidade *no mínimo um e no máximo um (1..1)*. Uma classe PIC, por contrário, relaciona-se com no mínimo zero e no máximo muitas instâncias do objeto (PIDO). A classe PPV (*Plain Property Value*) apresenta uma simplificação de uma tripla ligando cada objeto a um propriedade-valor, permitindo assim as transformações e modelagens necessárias. Cada instância de PIDO carrega consigo uma lista de instâncias de PPV. Uma instância da classe PIDODao é utilizada para realizar as conexões e transações necessárias.

#### 6.4.2.1 Adicionar Objeto Físico

Na Figura 49, é apresentada uma estrutura de interação humano-computador da funcionalidade responsável por adicionar um novo objeto físico. O Curador deve poder pesquisar a classe do acervo numa lista de classes já definidas. Escolher, também em uma lista os tipos físicos relacionados àquele artefato e informar os dados básicos do objeto físico. A geração de identificador único deve ser realizada pela classe PIDO, que de posse desse *UUID* randômico gera também a URI do objeto através do endereço base do repositório (ex: *http://meurepositorio/xxxxxxxxxx*).

Figura 49 – IHC – Adicionar Novo Objeto Físico

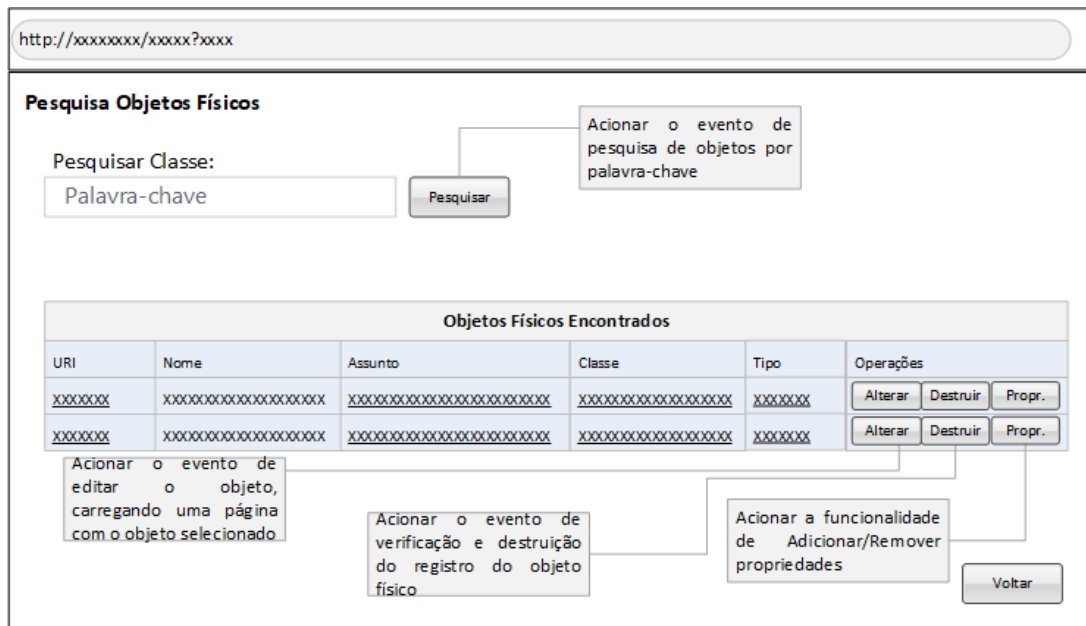


Fonte: elaborada pelo Autor.

#### 6.4.2.2 Pesquisar Objetos Físicos

A seguir, uma interação humano-computador referente a pesquisar Objetos Físicos pode ser visualizada na Figura 50. O Curador, inicia o processo sempre com possibilidade de realizar a primeira busca por palavra-chave a verificar os objetos presentes inicialmente no próprio acervo que atendem os critérios de busca em qualquer uma das relações já estabelecidas para aquele objeto.

**Figura 50 – IHC – Pesquisar Objetos Físicos**



Fonte: Elaborada pelo Autor.

### 6.4.2.3 Modificar Objeto Físico

Os objetos, uma vez criados, podem ser modificados por diversas razões. Uma delas é alguma correção de algum dado; outra, mais crítica e importante nesse protótipo, representa o adição de propriedades. As propriedades podem ser adicionadas para relacionar a objetos digitais e classes tanto internas quanto externas.

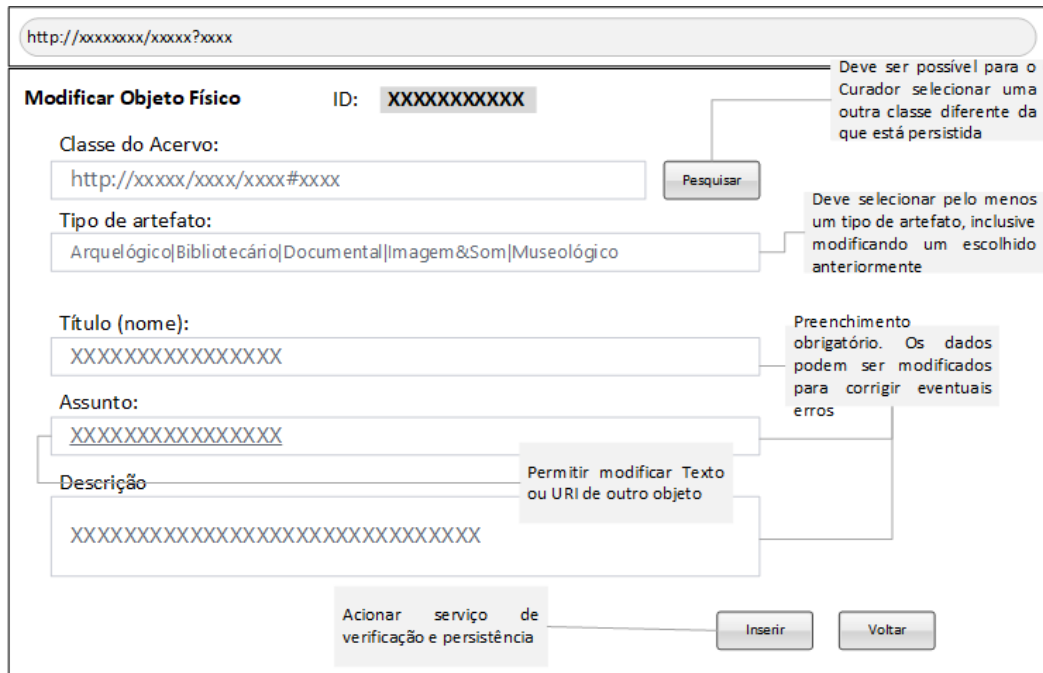
Deve ser possível ao Curador, inclusive, escolher uma classe diferente daquela que foi persistida para o objeto anteriormente, fazendo assim uma atualização da classificação do acervo.

Outra possibilidade é que o assunto do objeto seja uma referenciado por uma URI e não um texto simples. Caso o Curador, entretanto, deseje adicionar mais de um valor para a mesma propriedade (como no caso *dc:subject*), deve-se utilizar a funcionalidade de Adicionar/remover propriedades.

A Figura 51 apresenta uma interação humano-computador para a funcionalidade

Modificar Objeto Físico.

**Figura 51 – IHC – Modificar Objeto Físico**



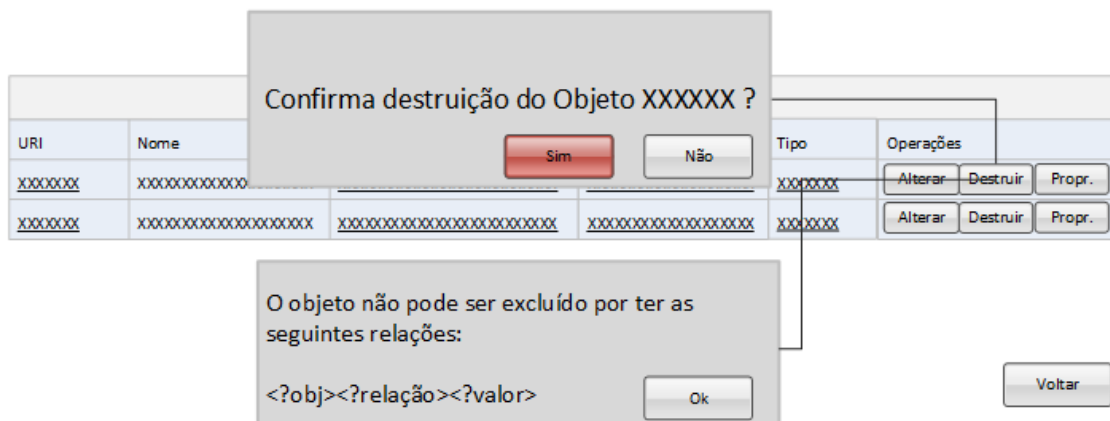
Fonte: elaborada pelo Autor.

6.4.2.4 Destruir Objeto Físico

A destruição de objeto físico está condicionada ao objeto não ter sido utilizado, devido a essa condição, a primeira interação da funcionalidade é perguntar se o Curador deseja de fato realizar a exclusão.

A Figura 52 é uma interação humano-computador da funcionalidade Destruir Objeto Físico.

**Figura 52 – IHC – Destruir Objeto Físico**



Fonte: Elaborada pelo Autor.

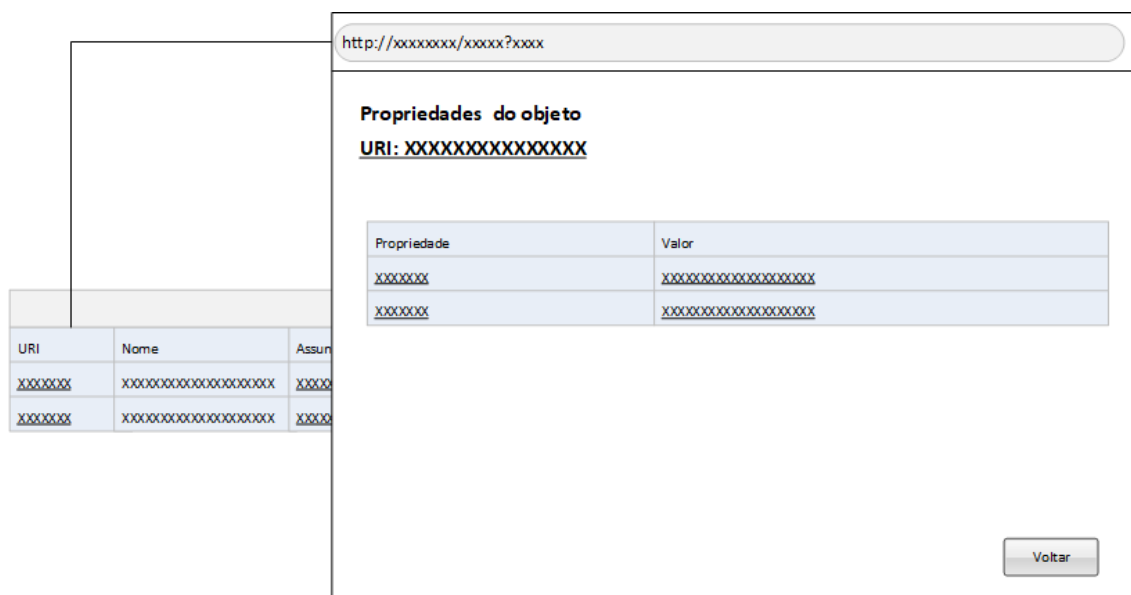
Caso a resposta seja sim, o algoritmo de pesquisa por utilização do objeto é disparado pelas camadas mais internas do sistema, verificando se aquele objeto possui alguma relação com outro objeto interno ou de *repositórios amigos*. Caso haja alguma relação, a destruição é descartada e deve-se informar o Curador. Deve ser informado, nesse caso, qual objeto e qual associação está sendo realizada para que o Curador possa navegar até o que faz a referência proibitiva da exclusão.

#### 6.4.2.5 Visualizar Propriedades

A funcionalidade de visualizar as propriedades do objeto devem ser acionadas quando o Curador clicar sob a URI do objeto, fazendo com o que o sistema dispare uma chamada (GET) para o objeto e carregue em uma página a lista com todos pares de propriedade e valor para aquela URI.

A Figura 53 é uma interação humano-computador para o Curador visualizar propriedades dos objetos.

**Figura 53 – IHC – Visualizar Propriedades**



Fonte: Elaborada pelo Autor.

#### 6.4.2.6 Adicionar/Remover propriedades

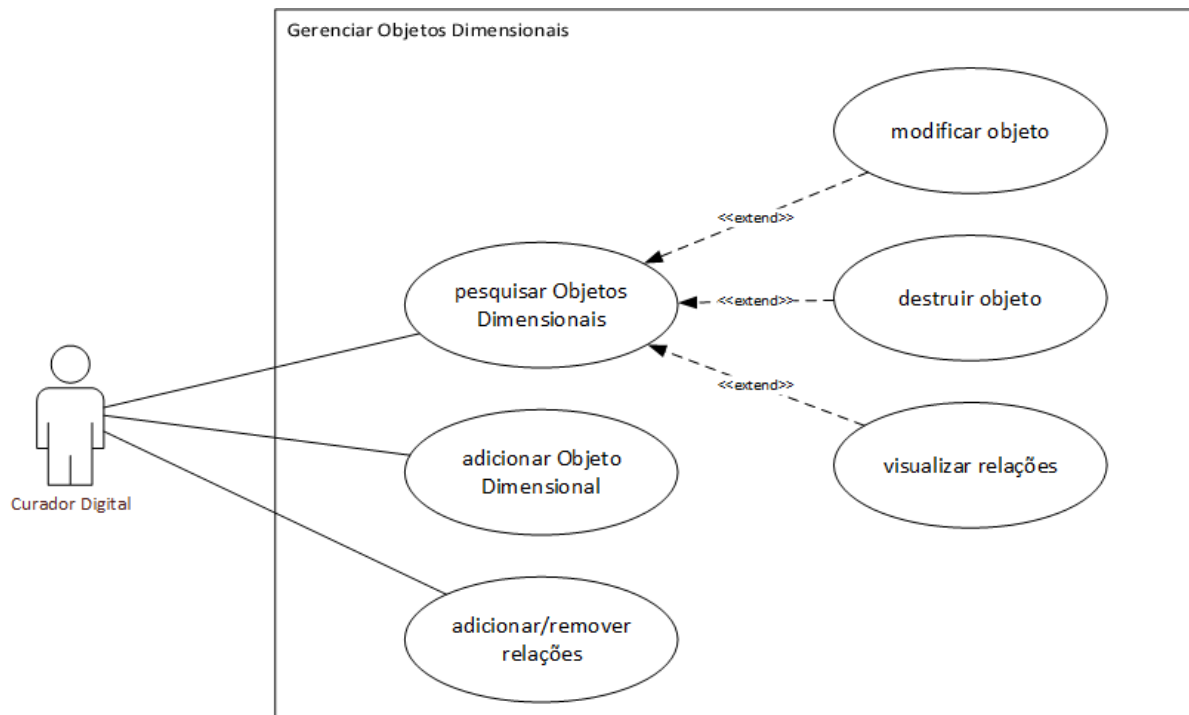
Adicionar ou remover propriedades (relações) dos objetos é uma funcionalidade para que o Curador possa corrigir eventuais erros de registro dos objetos, bem como enriquecer o objeto que está sendo manipulado com mais relações. Essa funcionalidade é uma extensão da funcionalidade descrita na Seção específica na Seção 6.4.4.

### 6.4.3 Gerenciar Objetos Dimensionais

A funcionalidade Gerenciar Objetos dimensionais deve permitir que o Curador, da mesma forma como gere os objetos físicos, adicionar ou modificar objetos digitais dimensionais pertencentes ao seu acervo. É necessário reiterar que os objetos dimensionais são os objetos não físicos, mas que possuem relação com os objetos físicos de forma que permitem, através das propriedades e classes da ontologia definida para as dimensões, adicionar relações *propriedade-valor* para enriquecimento do acervo.

A Figura 54 apresenta um diagrama de uso para a funcionalidade Gerenciar Objetos Dimensionais.

**Figura 54 – Diagrama de Caso de Uso Gerenciar Objetos Dimensionais**



Fonte: Elaborada pelo Autor.

Como é possível visualizar na Figura 54, Gerenciar Objetos Dimensionais se desdobra nos seguintes casos de uso “Pesquisar Objetos Dimensionais” - com extensões para modificar, destruir e visualizar as relações do objeto dimensional, “Adicionar Objeto Físico” e “Adicionar/remover relações”.

Um detalhamento do Diagrama de Caso de Uso acima pode ser encontrado no abaixo no Quadro 15.

**Quadro 15 – Especificação do Caso de Uso Gerenciar Objetos Dimensionais**

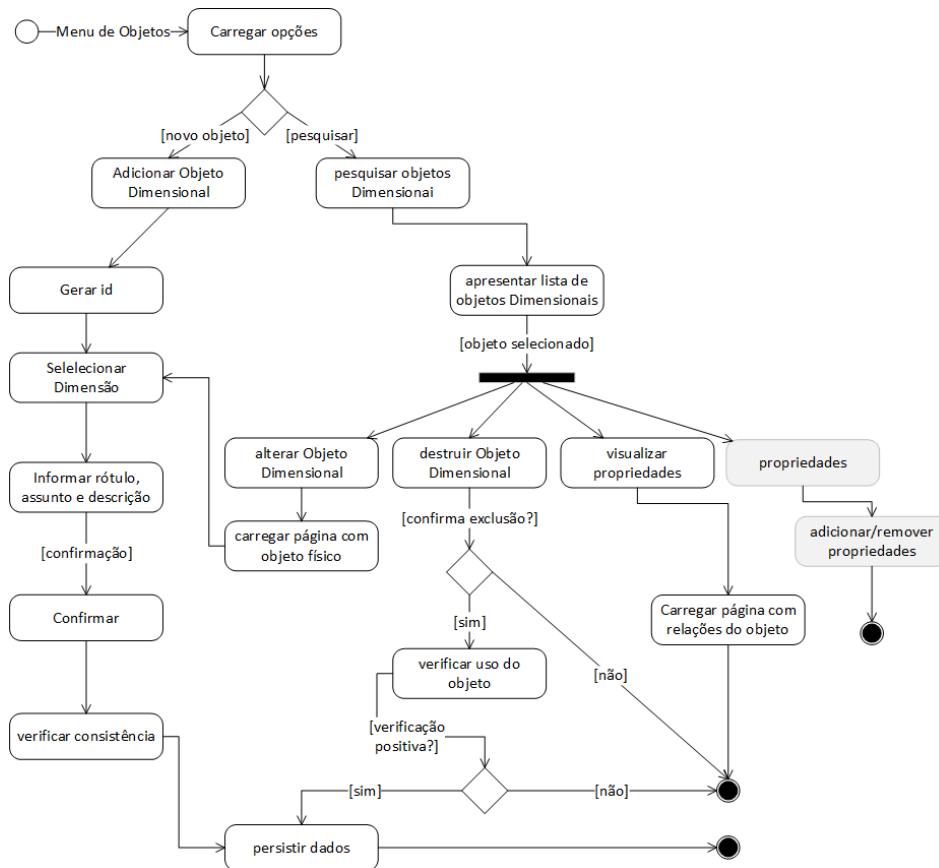
R3	Gerenciar Objetos Dimensionais
Ator principal:	Curador
Fluxo de Eventos	<ol style="list-style-type: none"> <li>1. <i>O Curador acessa a página de pesquisa de objetos dimensionais</i> <ol style="list-style-type: none"> <li>1.1 <i>O Curador informa critérios de busca e aciona o botão pesquisar</i></li> <li>1.2 <i>O Sistema carrega uma lista de objetos dimensionais de acordo com os critérios</i></li> </ol> </li> <li>2. <i>O Curador seleciona um objeto na lista e aciona o botão editar</i> <ol style="list-style-type: none"> <li>2.1 <i>O objeto selecionado é carregado em uma página</i></li> <li>2.2 <i>O Curador informa os dados do objeto que deseja alterar</i></li> <li>2.3 <i>O Curador seleciona a dimensão na qual objeto faz referência</i></li> <li>2.4 <i>O Curador clica em gravar</i></li> <li>2.5 <i>O sistema verifica a consistência dos dados, e em caso positivo, persiste os dados.</i></li> </ol> </li> <li>3. <i>O Curador seleciona um objeto na lista e aciona o botão destruir</i> <ol style="list-style-type: none"> <li>3.1 <i>É emitida uma mensagem crítica confirmando a exclusão</i></li> <li>3.2 <i>Caso haja confirmação, o sistema verifica se o Objeto Dimensional possui alguma relação com outro Objeto Físico ou Dimensional</i></li> <li>3.3 <i>Caso não haja nenhuma relação, o sistema persiste os dados.</i></li> </ol> </li> <li>4. <i>O Curador seleciona a opção visualizar propriedades Objeto Dimensional</i> <ol style="list-style-type: none"> <li>4.1 <i>O sistema carrega a página com todas as relações do Objeto Dimensional com qualquer outro Objeto Dimensional ou Físico</i></li> </ol> </li> <li>5. <i>O Curador seleciona a opção de adicionar novo Objeto Dimensional</i> <ol style="list-style-type: none"> <li>5.1 <i>Sistema Gera identificador único</i></li> <li>5.2 <i>Curador seleciona uma dimensão da ontologia Dimensoes.owl através da propriedade dim:tipoDimensional</i></li> <li>5.3 <i>Informa rótulo, assunto e descrição do Objeto Dimensional</i></li> <li>5.4 <i>Curador confirma envio</i></li> <li>5.5 <i>Sistema verificar se dados são válidos e persiste dados</i></li> </ol> </li> <li>6. <i>O Curador seleciona a opção adicionar remover Relações</i> <ol style="list-style-type: none"> <li>6.1 <i>O sistema carrega a funcionalidade R5</i></li> </ol> </li> </ol>
Exceções	<ol style="list-style-type: none"> <li>1. <i>Usuário não informar uma dimensão: deve emitir aviso</i></li> <li>2. <i>Sistema não gerar o identificador único</i></li> <li>3. <i>Falha na persistência</i></li> <li>4. <i>Dados inconsistentes</i></li> </ol>
Regras de Negócio	<ol style="list-style-type: none"> <li>1. <i>Um identificador único e inequívoco deve ser gerado para cada Objeto Dimensional</i></li> <li>2. <i>O campo Dimensão (dim:tipoDimensional) deve ser informado</i></li> <li>3. <i>O campo dc:description deve ser preenchido com uma descrição do Objeto Dimensional</i></li> </ol>

	4. Os nomes (dc:title) dos objetos devem ser únicos em uma hierarquia direta
Pré-Condições	1. A ontologia Dimensões precisa estar presente no TDB.
Pós-Condições	1. Os dados válidos informados devem ser persistidos no TDB

Fonte: Elaborado pelo Autor.

Na Figura 55, é apresentado um Diagrama de Atividades contendo o esquema com os fluxos da funcionalidade Gerenciar Objetos Dimensionais.

**Figura 55 – Diagrama de Atividades Gerenciar Objetos Digitais**



Fonte: Elaborada pelo Autor.

É possível ainda, Figura 55, visualizar todas as relações do objeto e também adicionar ou remover relações através das propriedades pré-existent na ontologia inicial ou qualquer uma outra adicionada pelo Curador e que faça referência a algum objeto na rede.

#### 6.4.3.1 Adicionar Objetos Dimensionais

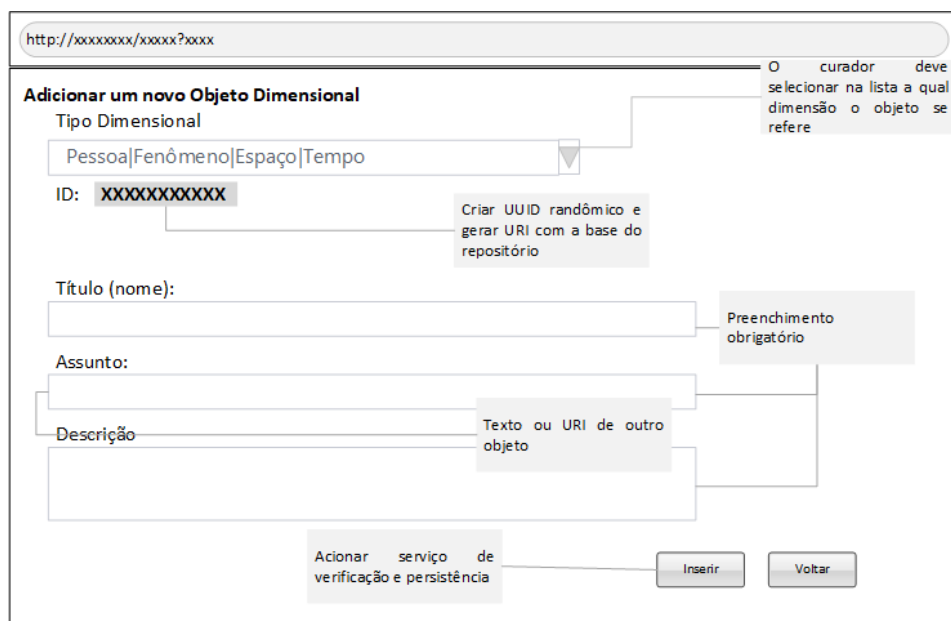
A funcionalidade Adicionar Objetos Dimensionais é similar a adicionar objetos físicos, esta, no entanto, deve permitir especificamente que o usuário selecione em uma lista o tipo dimensional daquele objeto o qual é apontado através da propriedade

*dim:tipoDimensional* para criar a relação necessária de caracterização da dimensão do objeto em Pessoa, Fenômeno, Espaço e Tempo conforme a ontologia das dimensões (Seção 4.7.3 p. 141).

O processo de verificação para persistência deve, além de testar se a dimensão e os campos foram informados corretamente, verificar se existe Objeto Dimensional com mesmo nome dentro daquele acervo, e ainda se há relações persistidas entre o objeto em questão e outros objetos do acervo interno e dos *repositórios amigos*. Antes da persistência, deve-se evitar que grafos fiquem órfãos ou com apontamentos para URIs inexistentes.

A Figura abaixo apresenta uma visão de interação humano computador para Adicionar Objetos Dimensionais.

**Figura 56 – IHC – Adicionar Objetos Dimensionais**



Fonte: Elaborada pelo Autor.

#### 6.4.3.2 Pesquisar Objetos Dimensionais

Nesta funcionalidade, assim como nos objetos físicos, o Curador inicia o processo sempre com possibilidade de realizar a primeira busca por palavra-chave, que verifica os objetos presentes inicialmente no próprio acervo dentro daqueles critérios de busca. Pode ainda informar qual o tipo Dimensional está associado ao objeto que deseja buscar, que no caso pode ser Pessoa, Fenômeno, Espaço ou Tempo, conforme definições da ontologia das dimensões (Seção 4.7.3 p. 141) As opções de alteração, destruição e



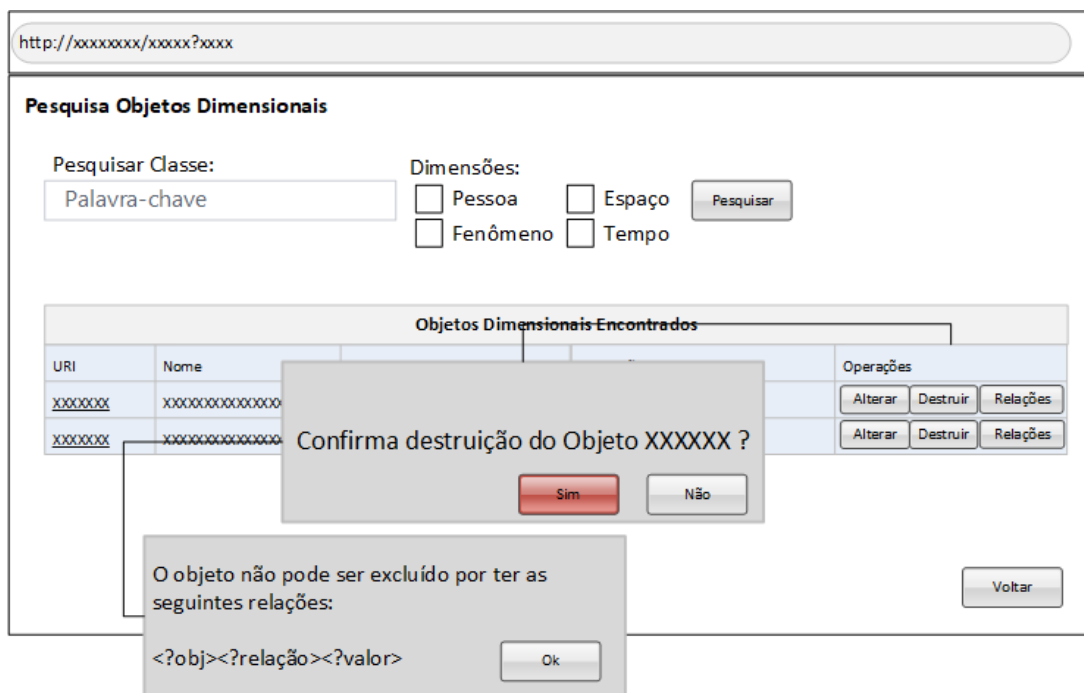
assim como no caso dos objetos físicos, permite que Curador possa modificar título, assunto e descrição, sendo possível ainda, modificar o Tipo Dimensional.

#### 6.4.3.4 Destruir Objeto Dimensional

A destruição de um Objeto Dimensional está condicionada a não utilização do objeto. Assim, é preciso também uma interação com o Curador para confirmar se este deseja, de fato, realizar a exclusão. Caso a resposta seja sim, é preciso disparar, internamente e nos *repositórios amigos*, uma pesquisa verificando se há relações com o Objeto Dimensional que está sendo destruído. Caso haja alguma relação, a destruição é descartada e o sistema deve emitir um aviso ao Curador. Deve ser informado, nesse caso, qual objeto e qual associação está sendo realizada para que o Curador possa ter conhecimento da referência que causa o impedimento da exclusão.

A Figura 59 apresenta a interação humano-computador para destruição de objetos dimensionais.

**Figura 59 – IHC – Destruir Objeto Dimensional**



Fonte: elaborada pelo Autor.

#### 6.4.3.5 Visualizar Relações

Essa funcionalidade é exatamente a mesma de Visualizar Propriedades descrita na Seção 6.4.2.5, neste o identificador único do objeto é passado para o sistema que realizar uma requisição GET e gera uma página com uma lista contendo todas as relações

daquele objeto.

#### 6.4.3.6 Adicionar/remover relações

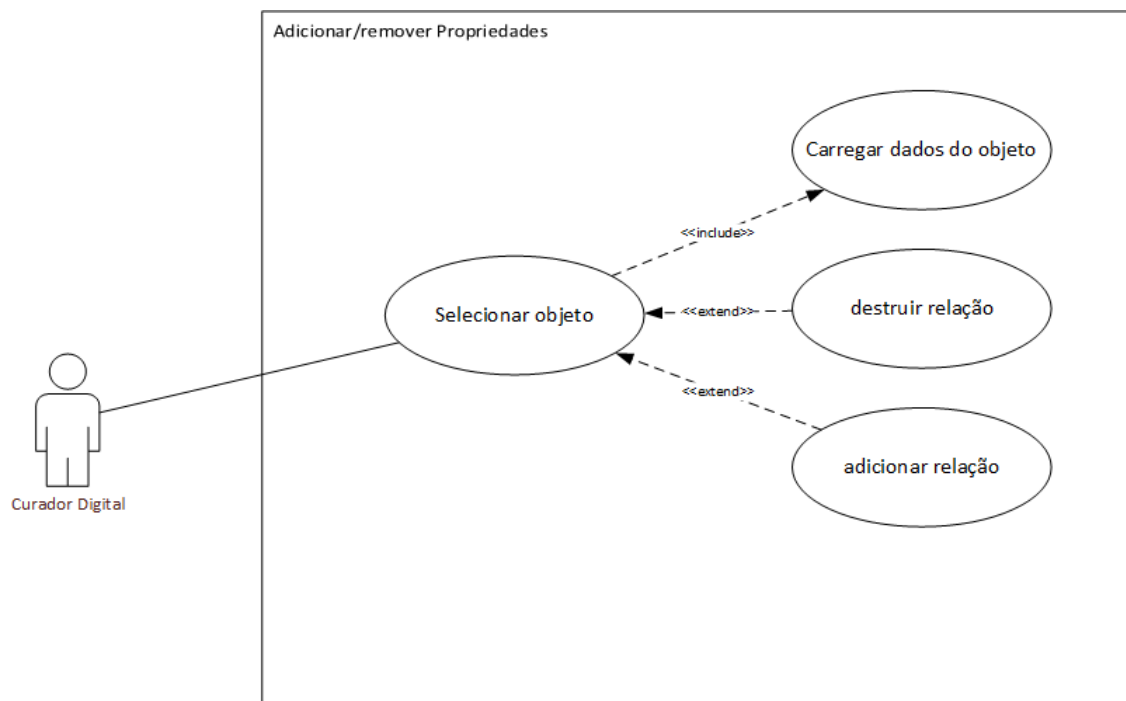
Esta funcionalidade é uma extensão da funcionalidade 6.4.4 que descrita logo abaixo.

### 6.4.4 Adicionar/Remover Propriedades

Esta funcionalidade deve permitir que o Curador adicione relações através de propriedades e valores informados para um Objeto Físico e também para um Objeto Dimensional, bem como deve permitir que uma ou mais relações sejam removidas, em caso de eventual necessidade de correção dos dados do objeto.

A Figura 60 apresenta o Diagrama de Caso de uso para a funcionalidade Adicionar/remover Propriedades

**Figura 60 – Diagrama de Caso de Uso de Adicionar/remover Propriedades**



Fonte: Elaborada pelo Autor.

Direta ou indiretamente, o Curador seleciona o objeto e o sistema carrega, obrigatoriamente, uma página todos dos dados (propriedades e valores) do objeto, estando, em seguida, disponíveis as opções para adicionar novas relações ou destruir o objeto selecionado.

O Quadro 16 abaixo apresenta um detalhamento do Diagrama de Caso de Uso

visualizado na Figura 60.

**Quadro 16 – Especificação do Caso de Uso Adicionar/Remover Propriedades**

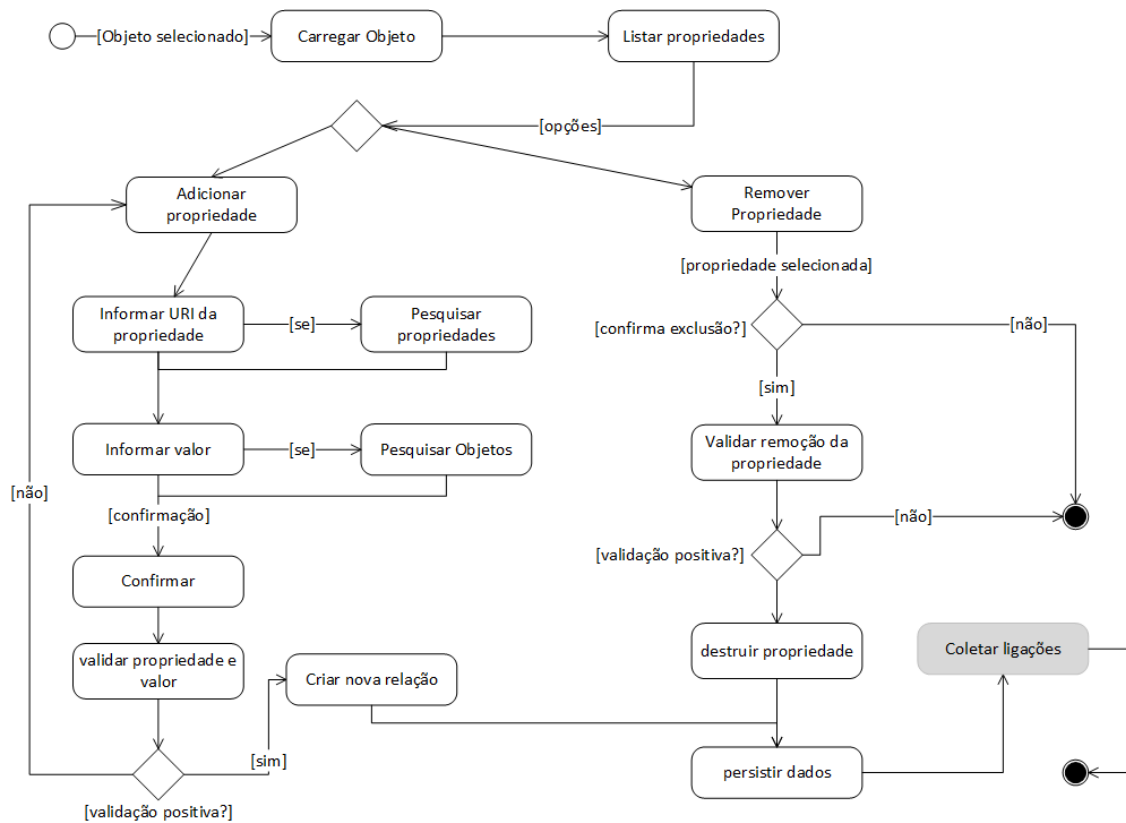
R4	Adicionar/remover Propriedades
Ator principal:	Curador
Fluxo de Eventos	<ol style="list-style-type: none"> <li>1. O evento é acionado diretamente pelo Curador ou pelas funcionalidades que o estende.</li> <li>2. O sistema carrega uma página com todas as relações listadas em tripla</li> <li>3. Caso a opção seja por adicionar propriedade             <ol style="list-style-type: none"> <li>3.1 O Curador informa o URI e Valor da propriedade</li> <li>3.2 O Curador busca, se desejar, propriedades na lista de propriedades das ontologias</li> <li>3.3 O Curador busca, se desejar, Objetos físicos e dimensionais presentes no acervo interno ou externo</li> <li>3.4 O Curador clica em adicionar</li> <li>3.5 O sistema valida os dados informados</li> <li>3.6 Em caso de validação positiva, o sistema persiste os dados. Caso contrário emite aviso com crítica.</li> </ol> </li> <li>4. Caso a opção seja por remover propriedade             <ol style="list-style-type: none"> <li>4.1 O Curador seleciona a propriedade na lista</li> <li>4.2 O Curador pressiona o botão “destruir”</li> <li>4.3 O sistema solicita confirmação</li> <li>4.4 Em caso positivo, o sistema dispara o processo de verificação</li> <li>4.5 Em caso de verificação positiva, o sistema persiste os dados, destruindo o objeto.</li> <li>4.6 Em caso negativo emite um aviso com a crítica</li> </ol> </li> </ol>
Exceções	<ol style="list-style-type: none"> <li>1. O Curador destruir uma relação definida na ontologia inicial para objetos físicos</li> <li>2. O Curador destruir uma relação definida na ontologia Dimensões para objetos dimensionais</li> <li>3. O Curador adicionar uma propriedade com URI não resolvível.</li> <li>4. O Curador adicionar um valor vazio</li> <li>5. O Curador adicionar um valor com URI não resolvível (ex. um endereço que não existe)</li> </ol>
Regras de Negócio	<ol style="list-style-type: none"> <li>1. Deve ser possível adicionar mais de uma propriedade para o mesmo objeto (ex. Nome, ou assunto).</li> <li>2. Uma URI informada pelo Curador deve ser resolvível.</li> <li>3. Relações da ontologia inicial dos objetos físicos não podem ser removidas, salvo em casos de mais de uma relação (como título e assunto) para o mesmo objeto.</li> <li>4. Os objetos dimensionais não podem ter a relação <i>dim:tipoDimensional</i> removida</li> <li>5. A relação com ID de qualquer objeto não pode ser destruída em nenhuma hipótese</li> </ol>

Pré-Condições	1. As ontologias devem estar presentes no TDB
Pós-Condições	1. O objeto modificado deve estar acessível após as operações

Fonte: Elaborado pelo Autor.

As propriedades informadas pelo Curador, nesta funcionalidade, devem possuir URI/IRI válida e resolvível, e o valor informado, de acordo com a definição da propriedade escolhida, poderá ser uma URI de um recurso ou uma informação textual, como um ano ou uma descrição, por exemplo.

**Figura 61 – Diagrama de Atividades Adicionar/remover Propriedades**



Fonte: Elaborada pelo Autor.

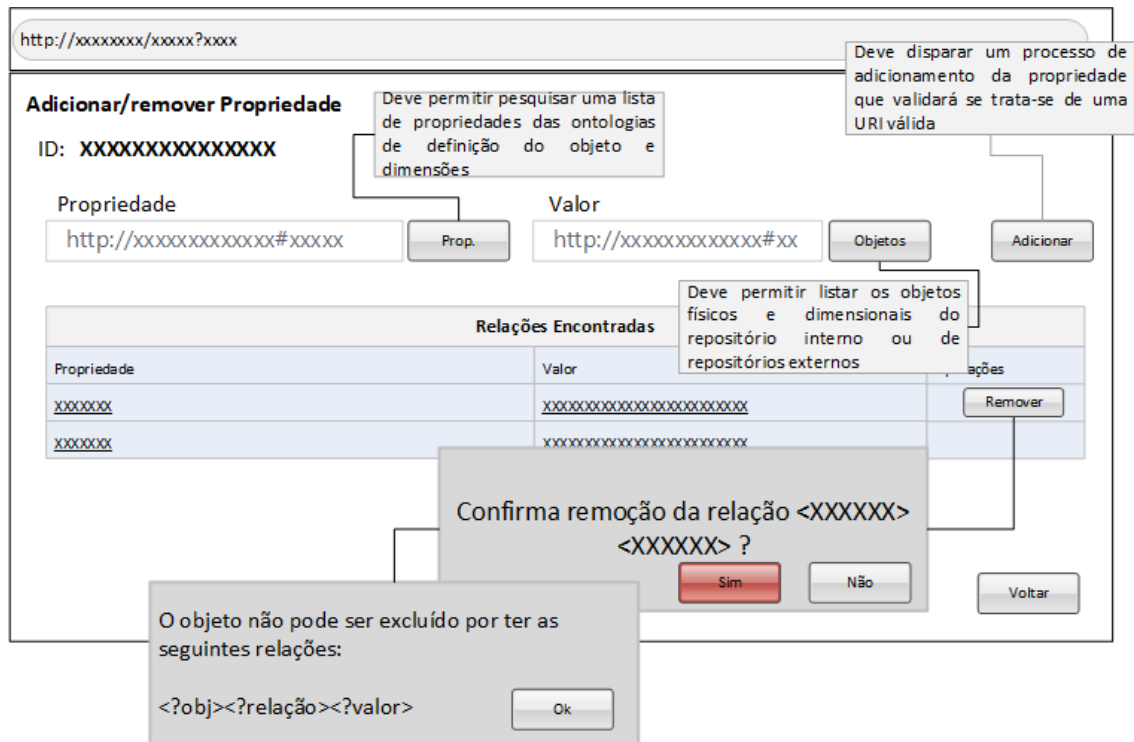
O fluxo das atividades de Adicionar/remover Propriedades se inicia, direta ou indiretamente, pelo Curador que, pode estar explorando um Objeto Dimensional ou Físico, ou poder ter acionado a funcionalidade em questão deliberadamente.

As opções, após a lista de propriedades e valores ser carrega, leva o Curador a poder escolher entre adicionar uma nova ou remover uma relação (propriedade-valor) existente. As atividades da primeira opção vão exigir que este informe URI, valor da

relação e acione a atividade de confirmação que, após validação, em caso positivo, aciona a atividade para criar uma nova relação (propriedade-valor) para o objeto e posterior persistência.

Na Figura 62 é apresentado um esquema de interação humano-computador para a funcionalidade especificada.

**Figura 62 – IHC – Adicionar/Remover Propriedades**



Fonte: Elaborada pelo Autor.

Ao listar propriedades ou objetos, o sistema deve carregar para o primeiro caso uma lista de propriedades advindas das ontologias de definição dos objetos e de dimensões (conforme Seção 4.7 p. 133). No segundo caso, o sistema deve carregar uma lista de objetos físicos e dimensionais. Isso permite que o Curador possa relacionar os objetos físicos e dimensionais com mais facilidade. Esta pesquisa por objetos físicos e dimensionais é realizada acionando o Motor de Interoperabilidade, que é quem realiza das buscas nos repositórios externos.

Quando se adiciona uma nova propriedade para um objeto ou um novo objeto, o sistema deve acionar automaticamente a busca por possíveis relações também é realizada através do Motor de Interoperabilidade. Essa funcionalidade está descrita na Seção abaixo.

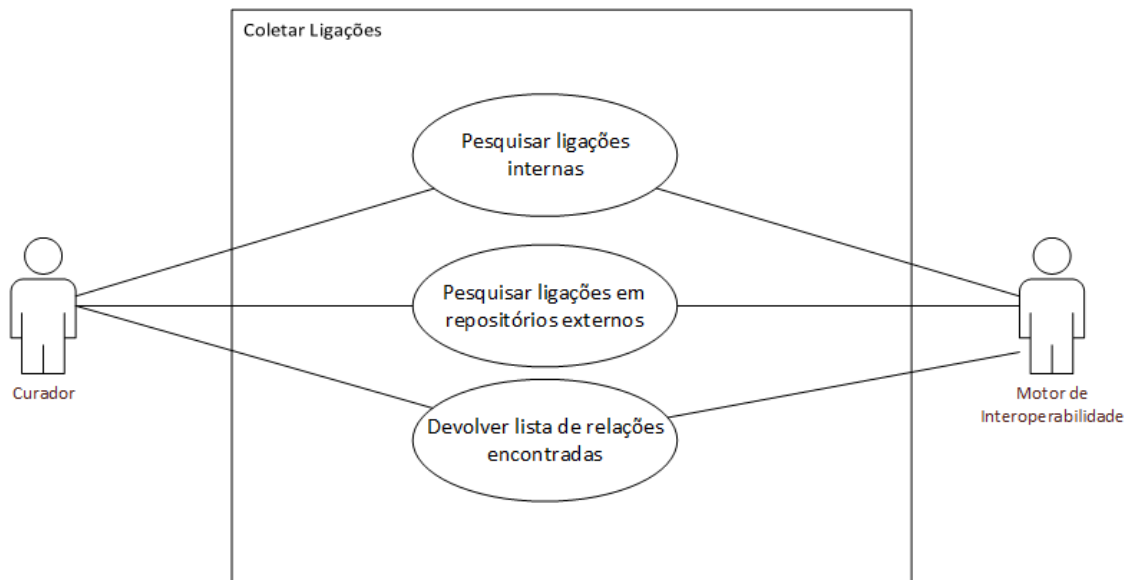
### 6.4.5 Coletar Ligações

Esta é a funcionalidade que permite que a interoperabilidade possa ocorrer conforme modelo definido na Seção 4.8 (p. 143). Isso porque a interoperação ocorre entre os repositórios através da busca de relações internas e externas após os objetos terem sido apontados com alguma propriedade pelos respectivos curadores. Ao se criar um novo objeto, por exemplo, o Curador pode querer adicionar relações com outros objetos presentes o próprio acervo ou em acervo de algum repositório externo constante na lista de *repositórios amigos*.

Ressalte-se que os repositórios oferecem uma interface pública via SPARQL e/ou chamadas diretas por meio de *middlewares* de serviços Web e estão presentes da lista de *repositórios amigos* da funcionalidade especificada na Seção 6.4.6.

Conforme o Diagrama de Caso de Uso apresentado na Figura 63, o Motor de Interoperabilidade aciona os processos responsáveis pesquisar relações internas e externas e apresenta-las ao Curador como relações em potencial para o objeto selecionado.

**Figura 63 – Diagrama de Casos de uso Coletar Ligações**



Fonte: elaborada pelo Autor.

No Quadro 17, é apresentada uma especificação para o caso de uso Coletar Ligações. O processo pode ser acionado pelo Curador de forma direta ou como uma extensão quando este estiver a gerenciar os objetos físicos, criando um novo objeto ou adicionando/removendo propriedades.

**Quadro 17 – Especificação do Caso de Uso Coletar Ligações**

R5	Coletar Ligações
Ator principal:	Curador
Ator secundário	Motor de interoperabilidade
Fluxo de Eventos	<ol style="list-style-type: none"> <li>1. O motor de interoperabilidade é acionado com um objeto selecionado</li> <li>2. O motor realiza pesquisa interna:             <ol style="list-style-type: none"> <li>2.1 Pesquisar objetos físicos por palavra-chave e utilizar a propriedade &lt;dc:relation&gt; para sugestão</li> <li>2.2 pesquisar todos os objetos dimensionais por palavra-chave e utilizar as propriedades de cada Dimensão encontrada como sugestão [relacionaPessoa relacionaFenômeno relacionaEspaço relacionaTempo]</li> <li>2.3 Pesquisa todos os objetos dimensionais que façam relação com o objeto selecionado                 <ol style="list-style-type: none"> <li>2.3.1 para cada objeto dimensional encontrado</li> <li>2.3.2 pesquisar todos os objetos físicos que façam relação com o Objeto Selecionado pela dimensão do Objeto Dimensional encontrado</li> </ol> </li> <li>2.4 O sistema prepara a lista de objetos encontrados internamente.</li> </ol> </li> <li>3. O motor realiza pesquisa nos repositórios amigos             <ol style="list-style-type: none"> <li>3.1 Pesquisar objetos físicos por palavra-chave e utilizar a propriedade &lt;dc:relation&gt; para sugestão.</li> <li>3.2 pesquisar todos os objetos dimensionais por palavra-chave e utilizar as propriedades de cada Dimensão encontrada como sugestão [relacionaPessoa relacionaFenômeno relacionaEspaço relacionaTempo]</li> <li>3.3 Pesquisa todos os objetos dimensionais que façam relação com o objeto selecionado                 <ol style="list-style-type: none"> <li>3.3.1 Para cada objeto encontrado, pesquisar todos os objetos físicos que façam relação com o Objeto Selecionado pela dimensão do Objeto Dimensional encontrado</li> <li>3.4 O sistema prepara a lista de objetos encontrados externamente</li> </ol> </li> </ol> </li> <li>4. O sistema devolve a lista separada por repositório, dimensão e tipo de objeto.</li> <li>5. O usuário clica em adicionar para adicionar uma relação da lista de relações sugeridas             <ol style="list-style-type: none"> <li>5.1 O sistema persiste a relação</li> </ol> </li> </ol>
Exceções	<ol style="list-style-type: none"> <li>1. O repositório não estar acessível</li> </ol>

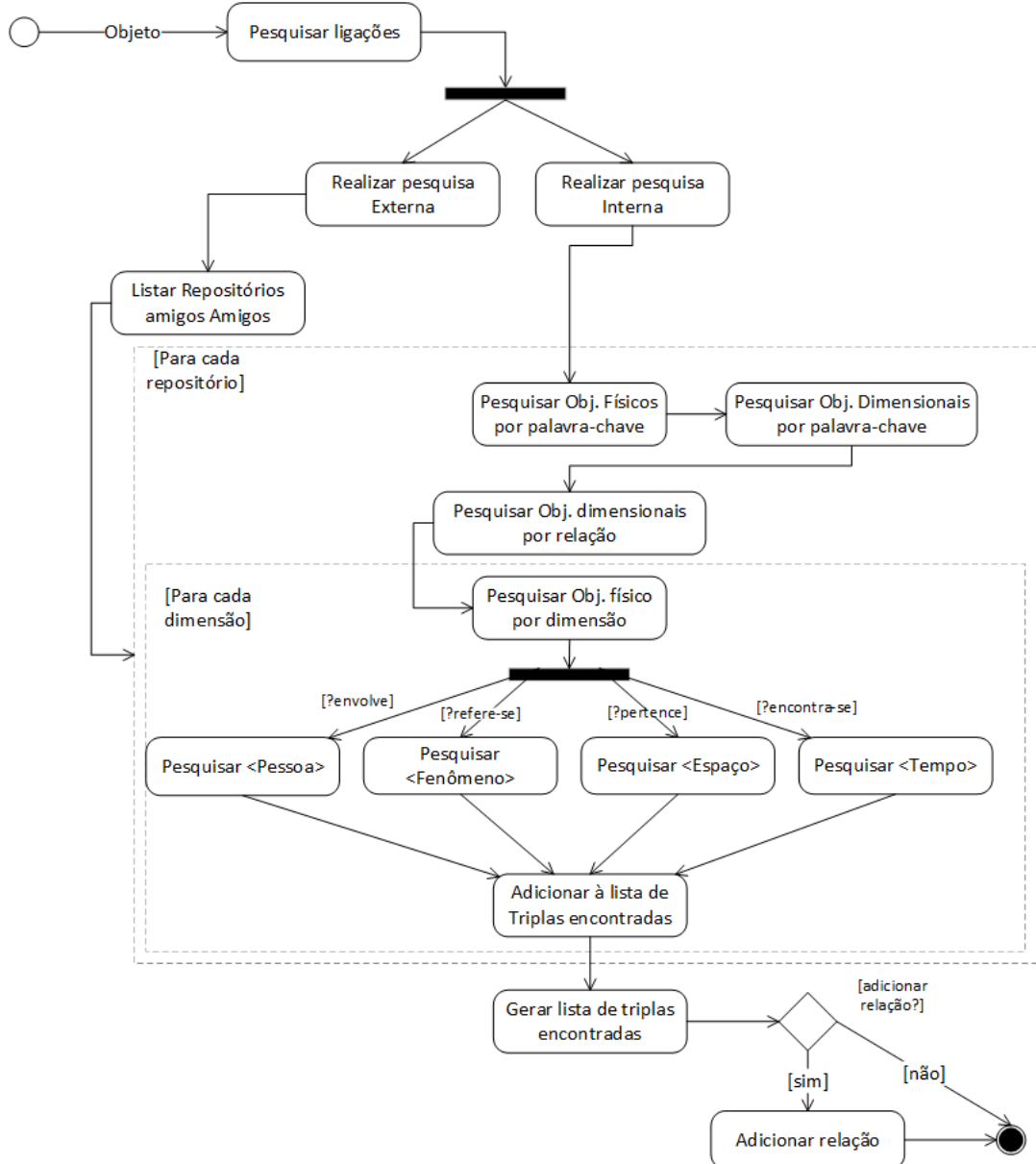
Fonte: Elaborado pelo Autor.

O módulo de interoperabilidade, composto pelo motor de interoperabilidade e o *end-point* para chamadas SPARQL, é acionado e, primeiro prepara e envia as requisições

para os repositórios digitais a buscar Objetos Dimensionais nas quatro dimensões (Pessoa, Fenômeno Espaço e Tempo) para que os *endpoints* ou *middlewares* correspondentes enviem as respostas das chamadas SPARQL em forma de triplas RDF/XML.

Na Figura 64 é apresentada uma visão do fluxo através de um Diagrama de Atividades para Coletar Ligações.

**Figura 64 – Diagrama de Atividades Coletar Ligações**



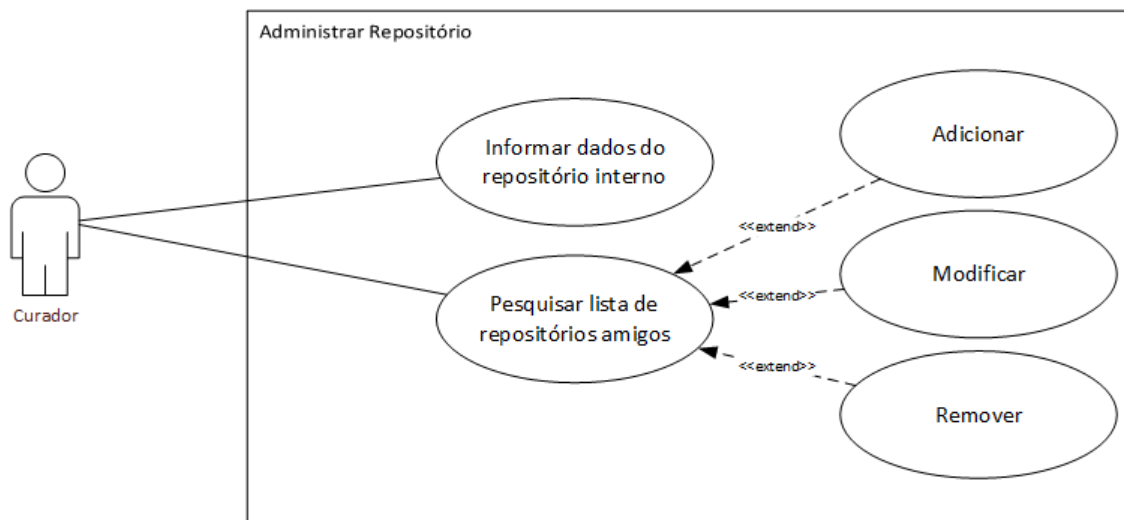
Fonte: elaborada pelo Autor.

#### 6.4.6 Administrar Repositório

Esta funcionalidade permite que o Curador possa administrar o repositório em

termos de identificação e também manter a lista de *repositórios amigos*. Na Figura 65 é apresentado um diagrama de caso de uso para a funcionalidade Administrar Repositório.

**Figura 65 – Caso de Uso Administrar Repositório**



Fonte: Elaborada pelo Autor.

O Curador pode informar os dados do repositório do espaço, bem como pesquisar e, se necessário, adicionar, modificar ou excluir alguma instância da lista de *repositórios amigos*.

O Quadro 18 apresenta um detalhamento do referido caso de uso.

**Quadro 18 – Especificação do Caso de Uso Administrar Repositório**

R6	Administrar Repositório
Ator principal:	Curador
Fluxo de Eventos	<ol style="list-style-type: none"> <li>1. O Curador abre a página do repositório               <ol style="list-style-type: none"> <li>1.1 O sistema carrega os dados do repositório</li> <li>1.2 O Curador, caso queira modifica os dados do repositório</li> <li>1.3 O Curador pressiona a opção “Gravar”</li> <li>1.4 O Sistema realiza a verificação e persiste os dados.</li> </ol> </li> <li>2. O Curador pesquisa a lista de repositórios amigos               <ol style="list-style-type: none"> <li>2.1 O Curador adiciona repositório à lista de repositórios amigos</li> <li>2.2 O Curador modifica o repositório selecionado</li> <li>2.3 O Curador remove repositório da lista de repositórios amigos</li> </ol> </li> </ol>
Exceções	<ol style="list-style-type: none"> <li>1. O Curador informa um valor inválido para URI</li> <li>2. O Curador não informa um valor para o nome do repositório.</li> </ol>
Regras de Negócio	<ol style="list-style-type: none"> <li>1. Deve ser informada uma URI válida e resolvível para os repositórios amigos.</li> </ol>

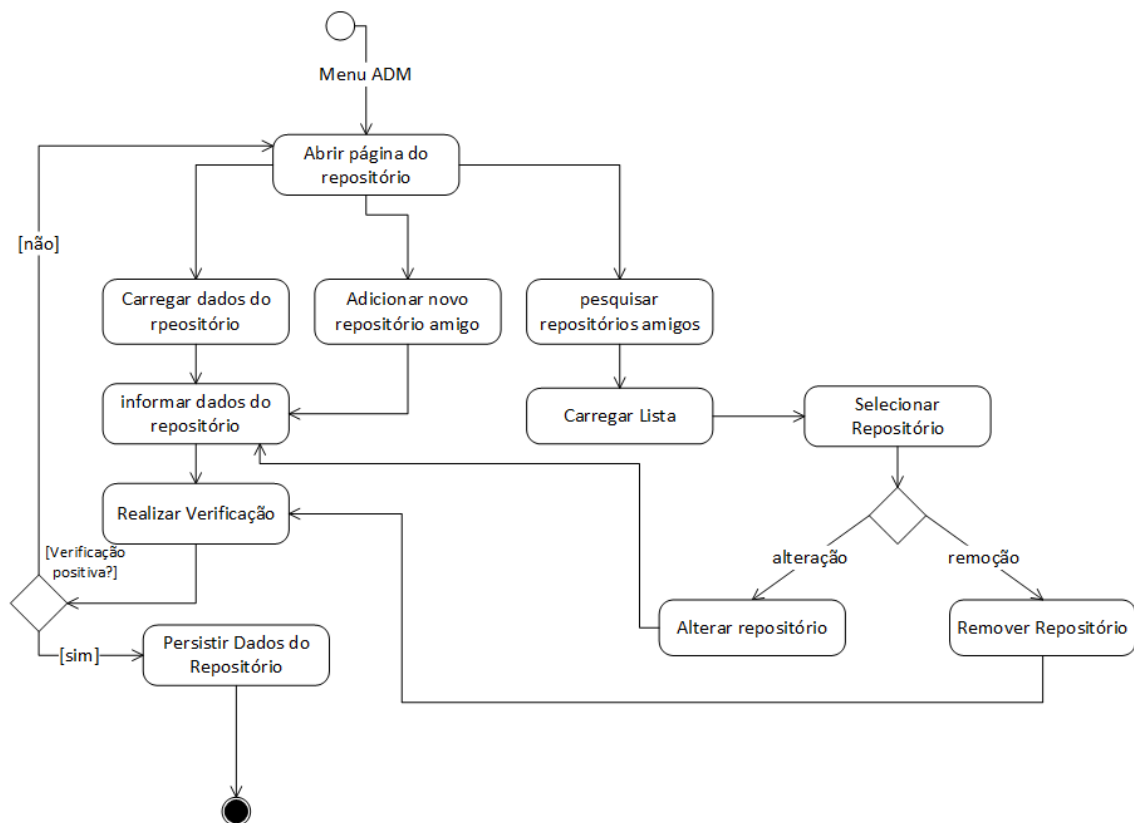
	<ol style="list-style-type: none"> <li>2. <i>Necessário informar o endereço do endpoint para pesquisa SPARQL de cada repositório.</i></li> <li>3. <i>Necessário informar um nome, descrição e um contato do repositório.</i></li> <li>4. <i>O contato deve ser um endereço de email válido.</i></li> </ol>
Pré-Condições	1. <i>O TDB dos repositórios amigos precisa estar presente e acessível</i>
Pós-Condições	1. <i>Os dados válidos informados devem ser persistidos no TDB</i>

Fonte: Elaborado pelo Autor.

O fluxo de atividades dessa funcionalidade prevê que o Curador inicie atividade visualizando os dados do repositório. De forma alternativa é possível adicionar um novo repositório à lista de *repositórios amigos*, ou, após pesquisar os repositórios, modificar ou remover um repositório da lista que é utilizada pela funcionalidade Coletar Ligações, crítica para todos os processos de interoperabilidade.

Na Figura 66 é apresentado o Diagrama de Atividades para funcionalidade Administrar Repositório.

**Figura 66 – Diagrama de Atividades Administrar Repositório**



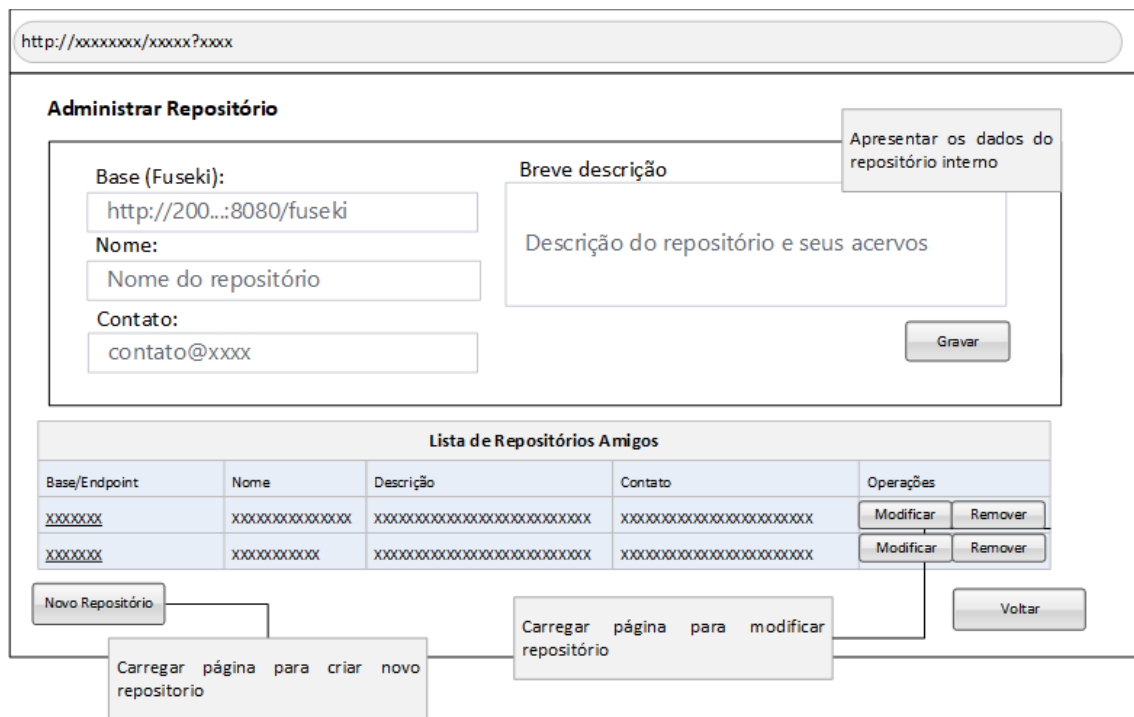
Fonte: Elaborada pelo Autor.

Os dados do repositório são apresentados em quadro e podem ser modificados, necessitando que o Curador, antes de persistir os dados, pressione o botão “gravar”. Uma

lista é apresentada para o Curador caso este pressione o botão “pesquisar”. Escolhendo a opção de modificação na operação “Modificar” para o item selecionado, é carregada uma página com os dados do repositório que podem ser modificados e persistidos. Pressionar “Remover” aciona a operação de exclusão do repositório da lista.

Uma interação humano-computador para a funcionalidade Administrar Repositórios pode ser visualizar na Figura 67.

**Figura 67 – IHC – Administrar Repositório**



Fonte: Elaborada pelo Autor.

## 6.5 Sumário do capítulo

O objetivo desse Capítulo foi apresentar os resultados da pesquisa-ação que foi conduzida para a modelagem de um Sistema de Informação semântico através de questionário e entrevista com os curadores dos espaços culturais. Detalhou-se os artefatos resultantes do processo de engenharia de requisitos que foi utilizada para o desenvolvimento do protótipo que será descrito no próximo capítulo.

## **CAPÍTULO VII. GUARÁ: UM PROTÓTIPO DE SISTEMA DE INFORMAÇÃO SEMÂNTICO PARA O PATRIMÔNIO CULTURAL**

### **7.1 Introdução**

Este capítulo tem por finalidade apresentar os resultados do processo de desenvolvimento do protótipo denominado Guará, um Sistema de Informação semântico para o Patrimônio Cultural com base na modelagem que foi apresentada no capítulo anterior.

Inicialmente serão detalhados os aspectos arquiteturais do protótipo, explorando sua estrutura modular com base no que foi estabelecido no CAPÍTULO IV. Posteriormente serão explorados modelos de interação com base em cenários de uso com repositórios e curadores hipotéticos.

### **7.2 Arquitetura e desenvolvimento do protótipo**

A construção do ambiente para o protótipo foi realizada conforme explicitado na proposta de arcabouço tecnológico da Seção 4.9. Foi utilizado o *Apache Jena* como framework de modelagem RDF e o TDB Fuseki como sistema gerenciador de banco de dados triplos e *endpoint* para atender a chamadas *SPARQL*.

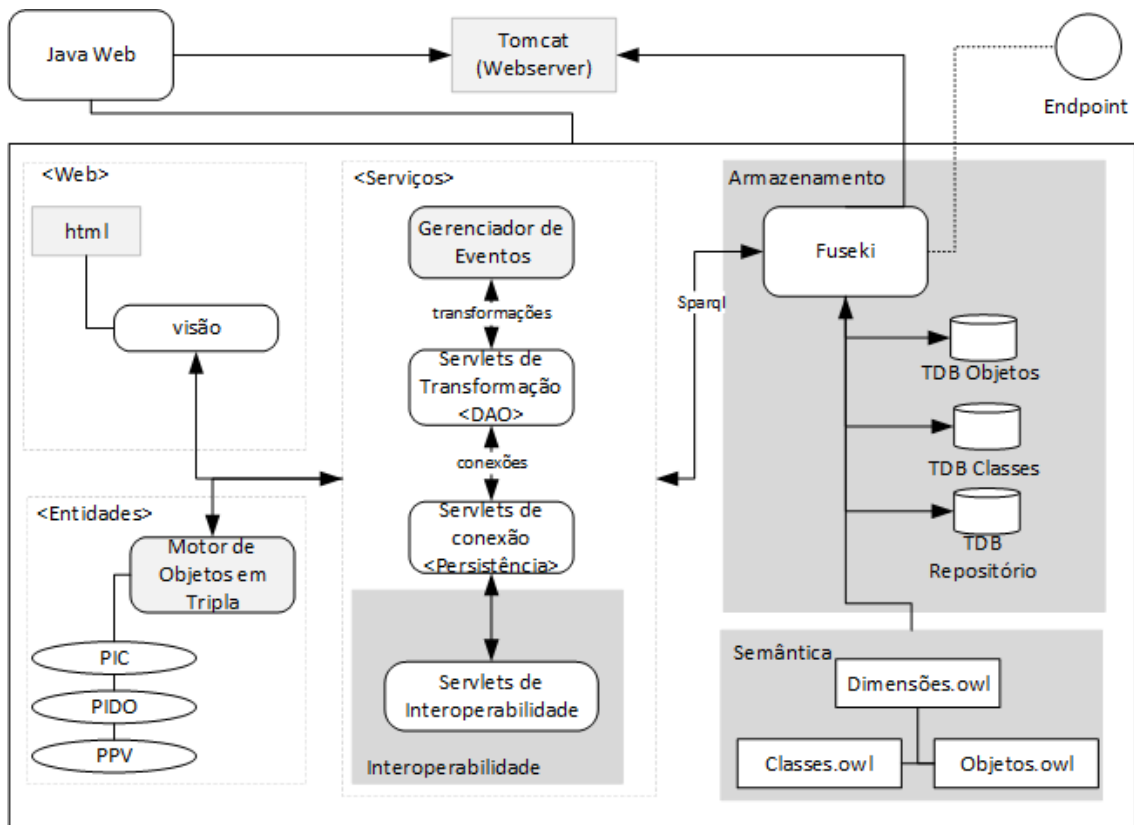
O ambiente orientado a objetos em Java também prevê a utilização do *apache Maven* para manutenção das bibliotecas dos frameworks e do próprio Java, que foi utilizado com a IDE *NetBeans*.

Para atender o modelo proposto, especificamente no que diz respeito a Seção 4.4 o protótipo se apresenta em uma arquitetura modular com cinco módulos, sendo dois externos e três internos. O Módulo Web (páginas web), o Módulo de Entidades, que é um motor de objetos em triplas; o Módulo de Serviços contendo três submódulos, um Gerenciador de Eventos, os Serviços de Transformação (DAO) e Persistência (com um Motor de Objetos em Tripla (MOT), e ainda um Motor de Interoperabilidade (MI) com os Serviços de Interoperabilidade que acionam os algoritmos de busca nos repositórios.

Os dois módulos externos dizem respeito ao armazenamento TDB e as ontologias da camada semântica que são mantidas pelo *Apache Jena Fuseki*.

A Figura 68 apresenta uma visão geral desta arquitetura modular.

**Figura 68 – Arquitetura modular do protótipo**



Fonte: elaborada pelo autor.

A tecnologia de *servlets* é utilizada para realizar as requisições e respostas *http* e transitar entre os módulos e camadas do protótipo, bem como para responder e realizar consultas externas. *Servlets* são classes java que estendem a classe *HttpServlet*. No Módulo de Serviços também são utilizadas classes especiais gerenciáveis para que o *framework* JSF possa realizar a inversão de dependências das classes do modelo e integrações com as páginas da camada de páginas Web. Essas classes especiais carregam duas anotações exigidas para as duas definições *@named* e *@requestScoped* ou outra anotação de escopo.

O funcionamento geral dessa arquitetura, de forma alinhavada com o que o modelo proposto preconiza, prevê que as Páginas Web sejam responsáveis pela interação com os usuários humanos que, por meio dos eventos disponíveis nos controles das páginas HTML, disparam ações que são transformadas em requisições para o Gerenciador de Eventos. Esse, por sua vez, aciona o Motor de Objetos em Triplas (MOT) realizando a transformação de objetos para triplas RDF e vice-versa.

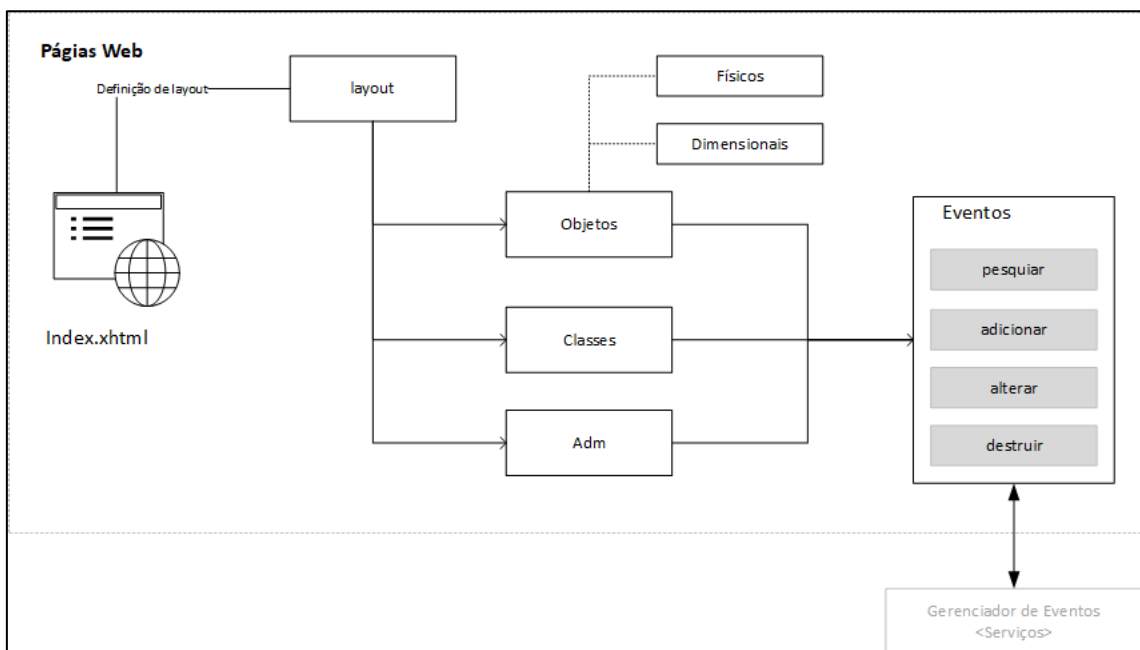
Os objetos transformados são enviados aos *servlets* de transformação (DAO) que

acionam os objetos de acesso ao banco de dados (TDB) após realizar as transformações necessárias para a linguagem SPARQL. Isso para que os grafos sejam pesquisados ou persistidos no TDB Fuseki, que mantém o serviço de Armazenamento (*RDF Store*) dos bancos de dados em tripla e das ontologias da Camada Semântica. No Fuseki há também um *endpoint* que é disponibilizado para consultas externas em SPARQL. A seguir são detalhados os módulos dessa arquitetura.

### 7.2.1 Páginas Web

Um conjunto de páginas XHTML é mantido para o funcionamento da estratégia *model-view-controller* (MVC) da especificação *Java Server Faces* (JSF). Para cada interação dos Curadores com o sistema, são geradas páginas Web que recebem estímulos e apresentam informações para os usuários. Os estímulos geram eventos que são tratados por controladores e transitados da apresentação à persistência e vice-versa.

**Figura 69 – Esquema das Páginas Web do Protótipo**



Fonte: elaborada pelo Autor.

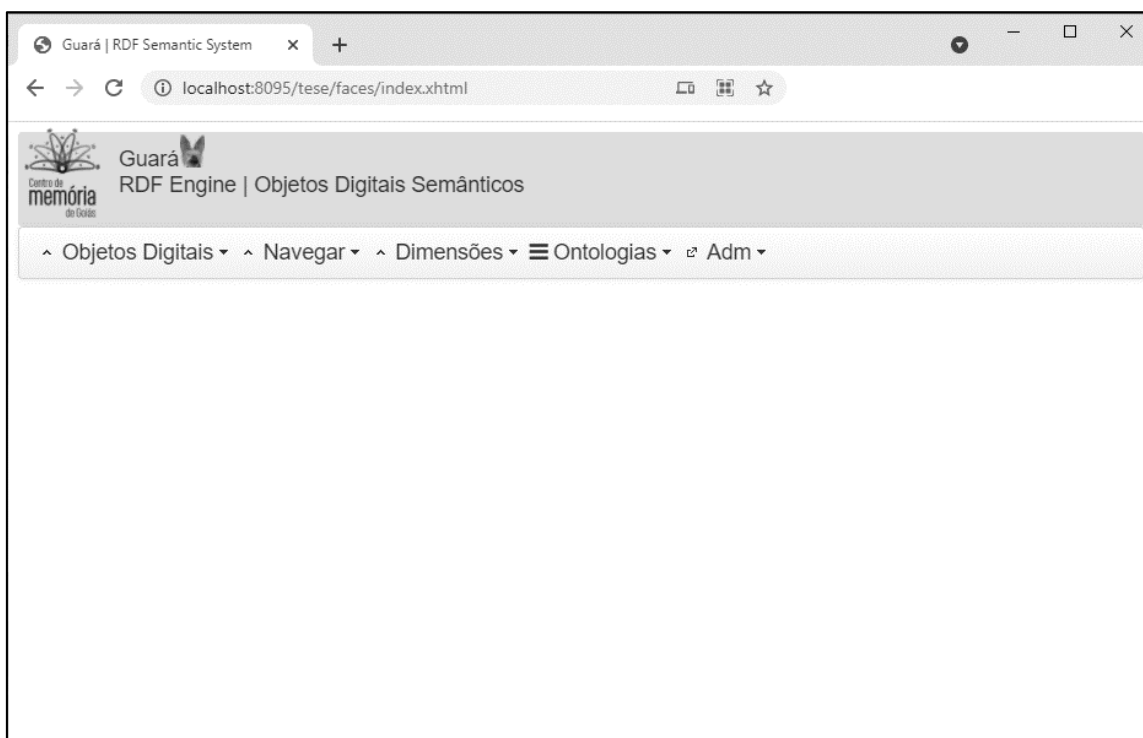
Como é possível visualizar na Figura 69 acima toda a rota é iniciada em uma página principal (*index.xhtml*) que utiliza um *layout* para criar páginas JSF em dois pacotes (Objetos, Classes e Adm). Para cada evento (pesquisar, adicionar, alterar, destruir, listar, coletar e outros) são geradas páginas XHTML específicas que acionam os eventos correspondentes no Gerenciador de Eventos, e alguma instância de uma classe controladora é também acionada para fazer a transição necessária e as requisições e

respostas de cada serviço requisitado.

A partir da página inicial, uma vez que cada espaço de memória apresenta seu próprio repositório, o Curador pode ter acesso as funcionalidades básicas do protótipo conforme Figura 70. A partir dessa página inicial o usuário tem a possibilidade navegar pelos objetos digitais, criar e manter a ontologia inicial (organização das classes do acervo) e criar e manter objetos digitais relacionados ao acervo.

Na Figura 70 é apresentada a Página Inicial do Protótipo Guará.

**Figura 70 – Página Inicial do Protótipo**



Fonte: elaborada pelo Autor.

A opção Objetos Digitais aciona as funcionalidades de gestão de Objetos Físicos e Dimensionais do acervo, permitindo que o Curador possa pesquisar e manipular os objetos digitais bem como adicionar e remover relações entre eles. Navegar permite que o Curador explore o acervo em torno das coleções nele presentes. Assim é possível ter uma visualização do acervo de acordo com cada classe e tipo presente no repositório do espaço de memória do Curador.

A opção Dimensões permite explorar o acervo em torno das quatro dimensões de interoperabilidade definidas na ontologia das Dimensões. Assim, é possível visualizar em cada uma das dimensões quais objetos físicos e dimensionais possuem relação.

A opção Ontologias permite visualizar as ontologias iniciais de classes (*classes.owl*), objetos (*objetos.owl*) e a ontologia de Dimensão. (*dimensoes.owl*), além de oferecer ao Curador a opção de explorar, modificar e adicionar classes do acervo.

Por fim, a opção Adm permite ao Curador manter os dados do repositório bem como a lista de *repositórios amigos*.

## 7.2.2 Módulo de Entidades

O Módulo de Entidades, que funciona como um Motor de Objetos em Tripla (MOT), possui as classes do paradigma de orientação a objetos para geração das instâncias que são transformadas em triplas em triplas RDF e também o caminho inverso. É composto de quatro classes simplificadas (PIC, PIDO, PPV e TriplaSimples) que permitem a instanciação das triplas dos classificadores dos recursos e dos próprios recursos, que são os objetos dos acervos. É composta também de um analisador (*parser*) para transformar os objetos em linguagem N3, RDF/XML e JSON.

A classe PIC (*Plain Initial Class*) é uma classe para simplificar a instanciação de objetos classificadores das coleções. Esta classe possui quatro atributos principais para atender elementos iniciais definidos no modelo, que são: *className*, *label*, *comment*, e *subClassOf*.

### Algoritmo 4 – Parte do código fonte da classe PIC

```
1  [...]
2  public class PIC implements Serializable, MOT {
3      private static final long serialVersionUID = -1L;
4      private String className = "";
5      private String label = "";
6      private String comment = "";
7      private String subClassOf = "";
8
9      //construtores
10
11     //getters e setter
12
13     [...]
14
15     @Override
16     public String getSPARQLInsert() {
17         String s = util.prefixos.ListaPrefixos();
18         s = s + "INSERT DATA{<" + this.className + "> a owl:Class ;\n "
19             + "          rdfs:comment      \"\" + this.comment + "\";\n "
20         "
21             + "          rdfs:label          \"\" + this.label + "\" ;\n "
22             + "          rdfs:subClassOf <" + this.subClassOf + "> ."
23             + "};";
24         return s;
25     }
26
27     @Override
28     public String getSPARQLDestroy() {
```

```

29     String s = util.prefixos.ListaPrefixos();
30     s += "DELETE WHERE { <" + this.className + ">"
31         + "          rdfs:comment ?x;\n "
32         + "          rdfs:label ?y;\n "
33         + "          rdfs:subClassOf ?z ."
34         + "};";
35
36     return s;
37 }
38
39 [...]
40 }
41

```

Fonte: elaborado pelo autor

### 7.2.3 Módulo de Serviços

O módulo de serviços possui três submódulos, o Gerenciador de Eventos, que controla os eventos advindos do mundo exterior por meio de classes de controle, o motor com objetos de acesso ao TDB (DAO), que possui classes de conexão com o banco de dados em tripla e os servlets de conexão (Persistência) que garantem as conexões com o TDB.

#### 7.2.3.1 Gerenciador de Eventos

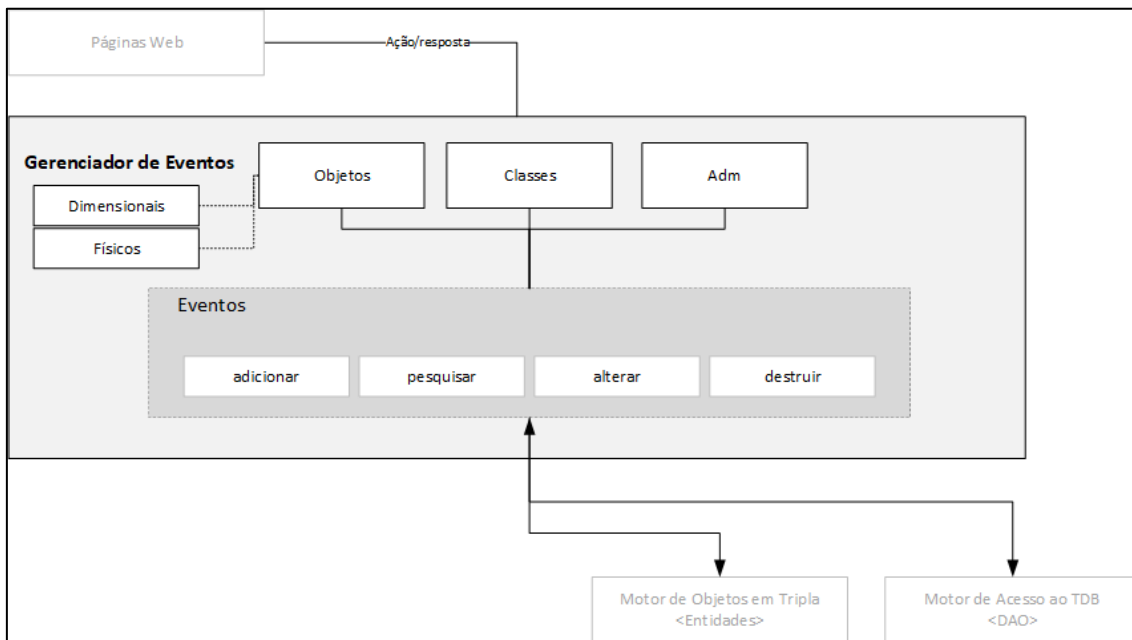
O Gerenciador de eventos tem a responsabilidade de acionar o Motor de Objetos em Triplas (MOT) para realizar as transformações necessárias. Trata-se de uma camada de serviços que permite que as interações com os diversos usuários sejam acionadas por meio de métodos específicos dos objetos presentes nessa camada. Todas as classes de controle são definidas nesta e apontadas para *servlets* gerenciáveis.

Objetos e Classes possuem um controlador específico com os eventos principais e com um serviço para cada um destes eventos. Dessa forma, as operações listar ou pesquisar classes estão no servlet *ListarClasses*. O servlet *SelecionarClasses* permite a operação responsável por listar todas as classes para uma eventual seleção. *EditarClasse* é um *servlet* que permite carregar uma classe para ser editada. O *PICController* é um *servlet* controlador dos objetos para as operações de inicializar, bem como, adicionar um novo objeto. As ações de editar e listar os objetos, de forma similar às operações nas classes, possuem, cada uma, um controlador específico para as instâncias de cada classe.

Listar objetos permite carregar lista de objetos que, caso sejam alterados ou recebam uma propriedade extra, serão acionados os eventos específicos do *servlet* *EditarObjeto*, assim por diante.

A Figura 71 apresenta um esquema funcional do gerenciador de eventos.

**Figura 71 – Esquema Funcional do Gerenciador de Eventos**



Fonte: Elaborada pelo Autor.

### 7.2.3.2 Serviços de Transformação (DAO)

O módulo Serviços de Transformação (DAO) contém todo o mecanismo de consulta e escrita no banco de dados em tripla. Seguindo a lógica de mapeamento objeto relacional com classes que permitem o acesso a dados (DAO), ele permite que os gerados pelo Motor de Objetos em Triplas sejam transformados em consultas SPARQL e persistidos, ou, no caminho inverso, possa acionar o mesmo motor para transformar os grafos retornados pelas consultas SPARQL para os objetos em tripla.

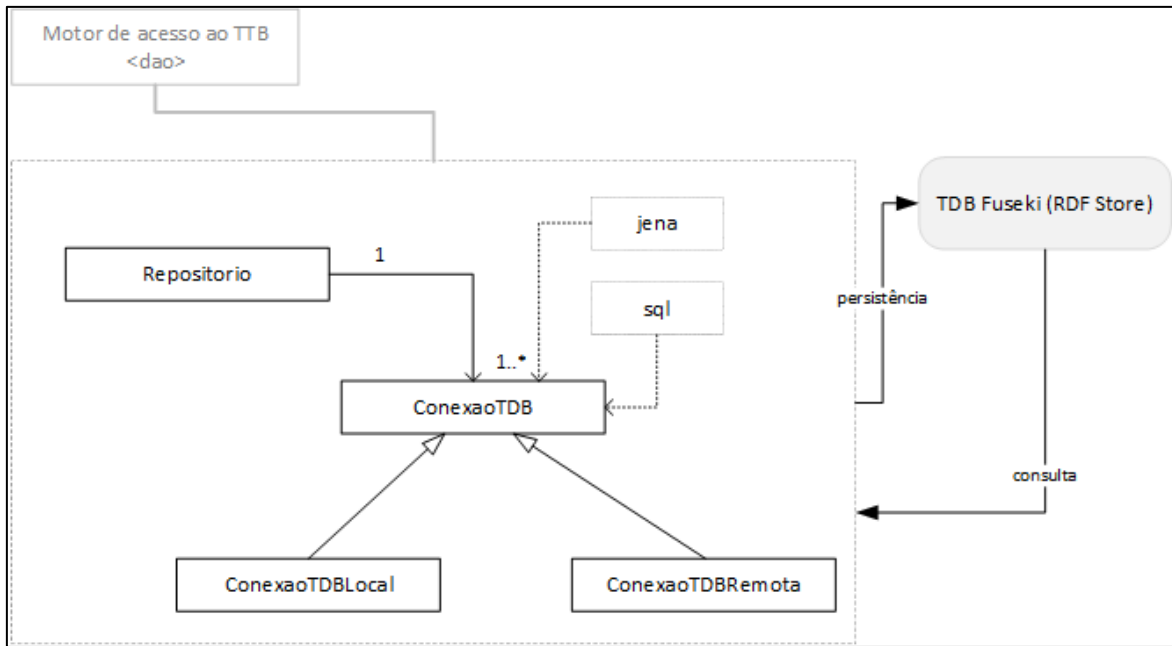
Este módulo é composto por 4 classes, *PICDao* responsável pelas transações das classes, *ObjetoDigitalDao* e *PIDODao*, responsáveis pelas transações dos objetos dimensionais e físicos e, *RepositorioDao*, responsável pelas transações com as instâncias de repositórios.

### 7.2.3.3 Serviços de Conexão (Persistência)

O motor de acesso ao TDB é composto por uma classe *ConexaoTDB*, que superclasse de duas classes específicas para conexão remota ou local. A depender da utilização do repositório, este pode ser remoto ou local.

Na Figura 72 é apresentado um esquema da Camada de Persistência do protótipo.

**Figura 72 – Esquema da Camada de Persistência**



Fonte: elaborada pelo Autor

As classes provenientes no Motor de Acesso ao TDB são utilizadas para persistência com o banco de dados em triplas RDF. Na prática, esta é uma camada DAO (*Data Access Object*) que utiliza as classes específicas do Apache Jena e do pacote SQL do Java.

Os objetos da camada DAO interagem com os objetos da camada Entidades. A camada Entidades, por sua vez, é capaz de transformar os objetos instanciados em SPARQL para que sejam persistidos no banco de dados. No caminho de volta, é a partir desses objetos que é possível transformar as triplas retornadas de acordo com cada objeto em uma instância *PIC* ou *PIDO*.

O Algoritmo 5 apresenta o código simplificado da classe responsável pela conexão com os bancos de dados em tripla RDF. Os Métodos recebem como parâmetros consultas SPARQL.

**Algoritmo 5 – Código simplificado de Classe para Conexão com o TDB**

```

1 //imports
2 [...]
3 public abstract class ConexaoTDB {
4     protected Statement st;
5     public String executa(String sql) {
6         try {
7             st.execute(sql);
8             return "Sucesso";
9         } catch (SQLException ex) {

```

```

10         return ex.getMessage();
11     }
12 }
13 public ResultSet getResultset(String sql) throws SQLException {
14     ResultSet r;
15     try {
16         r = st.executeQuery(sql);
17         return r;
18     } catch (SQLException ex) {
19         throw new Exception(ex.getMessage());
20     }
21 }
22
23 public List<Object> getResultList(String sql) {
24     ResultSet r;
25     List<Object> l = new ArrayList();
26     try {
27         r = st.executeQuery(sql);
28         l = (List<Object>) r;
29         return (List<Object>) r;
30     } catch (SQLException ex) {
31         return null;
32     }
33 }
34
35 public void fecha() {
36     try {
37         st.close();
38     } catch (SQLException ex) {
39
40     }
41 }
42 [...]
43 }

```

Fonte: Elaborado pelo autor

Uma extensão da classe de conexão é criada para o caso de conexão remota e outra para o caso de conexão local. Há uma sobrecarga de métodos, conforme o código simplificado no Algoritmo 6 abaixo.

#### Algoritmo 6 – Código simplificado de Conexão Remota com o TDB

```

1 //[...] imports
2 public class ConexaoTDBRemota extends ConexaoTDB {
3     private String FusekiPath = TDBUtil.getUrlRemota();
4     private String ds;
5     public ConexaoTDBRemota(String ds, String fusekiPath) {
6         this.FusekiPath = fusekiPath;
7         this.ds = ds;
8         Conecta();
9     }
10    public ConexaoTDBRemota(String ds) {
11        this.ds = ds;
12        Conecta();
13    }
14    private void Conecta() {
15
16        String URL =
17            "jdbc:jena:remote:query=" + this.FusekiPath + this.ds +
18            "/query"
19            + "&update=" + this.FusekiPath + this.ds + "/update";
20
21        try {
22            RemoteEndpointDriver.register();
23            Connection con = DriverManager.getConnection(URL);
24            st = con.createStatement();
25

```

```

26         } catch (SQLException ex) {
27
28         }
29
30     }
31
32     @Override
33     public String executa(String sql) {
34         try {
35             return super.executa(sql);
36         } catch (Exception ex) {
37
38         }
39
40     }
41
42     @Override
43     public void fecha() {
44         super.fecha();
45     }
46
47     @Override
48     public ResultSet getResultset(String sql) throws Exception {
49         return super.getResultset(sql);
50
51     }
52
53     @Override
54     public List<Object> getResultList(String sql) {
55         return super.getResultList(sql);
56
57     }
58
59 }

```

Fonte: Elaborado pelo Autor.

#### 7.2.3.4 Serviços de Interoperabilidade

As classes responsáveis pela interoperabilidade, que representam a camada de interoperabilidade do modelo estabelecido, estão acomodadas no Módulo de Serviços e possuem os algoritmos necessários para a interoperabilidade, que é um ponto central do modelo definido no CAPÍTULO IV (Seção 4.8 p. 143).

A primeira classe de interoperabilidade é o *servlet getObject* que, uma vez acionado, devolve em formato JSON ou HTML todas informações do objeto requisitado via HTTP. Conforme Seção 4.8 (Camada de Interoperabilidade – p. 143), o Algoritmo de interoperabilidade aciona os repositórios em busca de relações possíveis com o objeto alvo. A Figura 64 – Diagrama de Atividades Coletar Ligações na página 205 apresenta um conjunto de atividades do algoritmo dividindo em quatro funções de coleta a) pesquisar Objetos Físicos por palavra-chave; b) pesquisar Objeto Digitais por palavra-chave; c) pesquisar Objetos Dimensionais por relação e d) pesquisar Objetos Físicos por dimensão.

Abaixo são apresentados os algoritmos para cada uma das atividades

mencionadas.

a) Pesquisar Objetos Físicos por palavra-chave:

Abaixo, o Algoritmo 7, é responsável por buscar os Objetos Dimensionais por palavra-chave, verificando se nome, assunto e descrição dos objetos presentes no acervo pesquisado possuem o nome do objeto alvo.

**Algoritmo 7 – Pesquisar Objetos Físicos por Palavra-chave**

```
1 PREFIX obj: <http://200.137.241.247:8080/fuseki/objetos#>
2 PREFIX cls: <http://200.137.241.247:8080/fuseki/classes#>
3 PREFIX dim: <http://200.137.241.247:8080/fuseki/dimensoes#>
4 PREFIX owl: <http://www.w3.org/2002/07/owl#>
5 prefix dc: <http://purl.org/dc/elements/1.1/>
6 prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
7 prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
8 SELECT
9 distinct ?id ?nome ?assunto ?descricao ?classe
10 WHERE {
11     ?id dim:tipoFisico ?tipo.
12     ?id dc:title ?nome. ?id dc:subject ?assunto.
13     ?id dc:description ?descricao.
14     ?id rdfs:type ?classe
15
16     FILTER(REGEX(?nome, 'nome_do_objeto', 'i')
17             || REGEX(?assunto, ' nome_do_objeto ', 'i')
18             || REGEX(?descricao, ' nome_do_objeto ', 'i')).
19 }
```

Fonte: Elaborado pelo Autor.

b) Pesquisar Objetos Digitais por palavra-chave:

O Algoritmo 8 abaixo, de forma similar ao anterior, busca Objetos Digitais Dimensionais por palavra-chave.

**Algoritmo 8 – Pesquisar Objetos Dimensionais por Palavra-chave**

```
1 PREFIX obj: <http://200.137.241.247:8080/fuseki/objetos#>
2 PREFIX cls: <http://200.137.241.247:8080/fuseki/classes#>
3 PREFIX dim: <http://200.137.241.247:8080/fuseki/dimensoes#>
4 PREFIX owl: <http://www.w3.org/2002/07/owl#>
5 PREFIX dc: <http://purl.org/dc/elements/1.1/>
6 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
7 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
8 SELECT
9 DISTINCT ?id ?nome ?assunto ?descricao ?dimensao
10 WHERE {
11     ?id dc:title ?nome. ?id dc:subject ?assunto.
12     ?id dc:description ?descricao. ?id dim:tipoDimensional ?dimensao.
13     FILTER(REGEX(?nome, 'nome_do_objeto', 'i')
14             || REGEX(?assunto, 'nome_do_objeto', 'i')
15             || REGEX(?descricao, 'nome_do_objeto', 'i')).
16 }
```

Fonte: Elaborado pelo Autor.

c) Pesquisar Objetos Relacionais por relação

O método para pesquisar Objetos Dimensionais por relação aciona o Algoritmo 9 abaixo.

A propriedade relação pode ser `dim:relacionaPessoa`, `dim:relacionaFenômeno`, `dim:relacionaEspaço` ou `dim:relacionaTempo`. O nó “val” recebe uma lista dos valores dos grafos <sujeito - predicado – valor> encontrados.

### Algoritmo 9 – Pesquisar Objetos Dimensionais por Relação

```
1 PREFIX obj: <http://200.137.241.247:8080/fuseki/objetos#>
2 PREFIX cls: <http://200.137.241.247:8080/fuseki/classes#>
3 PREFIX dim: <http://200.137.241.247:8080/fuseki/dimensoes#>
4 PREFIX owl: <http://www.w3.org/2002/07/owl#>
5 PREFIX dc: <http://purl.org/dc/elements/1.1/>
6 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
7 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
8 SELECT
9     ?val
10 WHERE {
11     <id_objeto> <relação> ?val. }
12 }
13 --relação =
14 {relacionaPessoa|relacionaFenômeno|relacionaEspaço|relacionaTempo}
```

Fonte: Elaborado pelo Autor.

O algoritmo para pesquisar objetos dimensionais por relação (Algoritmo 9) é dividido em quatro partes, uma para cada dimensão (Pessoa, Fenômeno, Espaço e Tempo) e busca as relações com o objeto alvo para construir uma lista de ligações prováveis com base nas relações já existente. Digamos que um Objeto Físico *OF1* possua uma relação *P1* com o Objeto Dimensional *OD1*; e um Objeto Físico *OF2* também possui relação *P1* com *OD1*. O Algoritmo apresenta o objeto *OF2* como uma oportunidade possível para ligação com *OF1*, que por serem físicos, podem ser associados pela propriedade *dim:relation* (*dc:relation*). Caso o objeto encontrado seja um Objeto Dimensional, a propriedade que associa os objetos deve ser preservada, caso seja uma Pessoa, deve ser sugerida a propriedade *dim:relacionaPessoa*.

#### d) Pesquisar Objetos Físicos por Dimensão

Nessa etapa são recuperados grafos correspondentes aos Objetos Físicos que estejam associados à mesma dimensão do objeto alvo. Isto é, se um objeto físico *OF1* possui uma ligação *P1* com *OD1*, o sistema busca todos os objetos físicos *OFn* cuja relação <*OFn P1 OD1*> seja existente.

O Algoritmo 10 abaixo apresenta a consulta SPARQL do método de pesquisar Objetos Físicos por Dimensão.

### Algoritmo 10 – pesquisar Objetos Físicos por Dimensão

```

1 PREFIX obj: <http://200.137.241.247:8080/fuseki/objetos#>
2 PREFIX cls: <http://200.137.241.247:8080/fuseki/classes#>
3 PREFIX dim: <http://200.137.241.247:8080/fuseki/dimensoes#>
4 PREFIX owl: <http://www.w3.org/2002/07/owl#>
5 PREFIX dc: <http://purl.org/dc/elements/1.1/>
6 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
7 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
8 SELECT
9     ?sub
10 WHERE {
11     ?sub <relação> ?id_do_objeto. }
12 }
13 --relação =
14 {relacionaPessoa|relacionaFenômeno|relacionaEspaço|relacionaTempo}
    
```

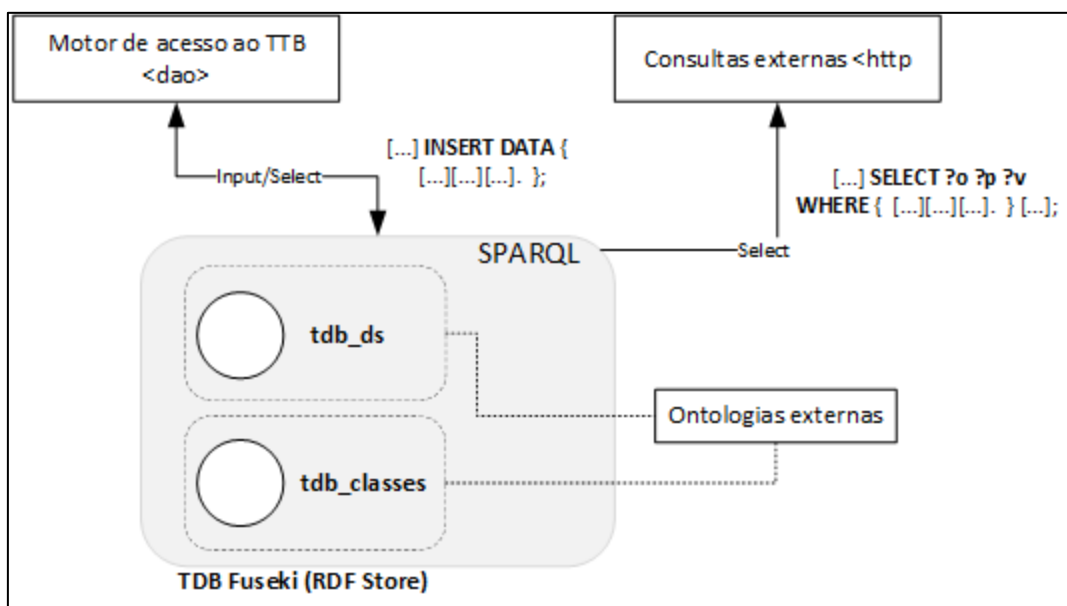
Fonte: Elaborado pelo Autor.

A propriedade (relação) pode ser *dim:relacionaPessoa*, *dim:relacionaFenômeno*, *dim:relacionaEspaço* ou *dim:relacionaTempo*. O nó “sub” recebe uma lista dos sujeitos dos grafos <sujeito - predicado – valor> encontrados.

#### 7.2.4 Armazenamento (RDF Store)

Uma instância do TDB *Fuseki* com um *endpoint* público permite as consultas por meio de chamadas HTTP e linguagem SPARQL. Dessa forma, qualquer outro sistema ou mecanismo de consulta pode fazer referência aos objetos do repositório. A Figura 73 apresenta uma visão do Esquema de Armazenamento (RDF Store).

Figura 73 – Esquema de Armazenamento RDF (RDF Store)



Fonte: elaborada pelo Autor

A Ontologia inicial com a definição dos objetos é armazenada em um TDB específico o qual pode ser acionado pelo endereço do repositório *fuseki* (<http://200.137.241.247:8080/fuseki/objetos>). Da mesma maneira, a Ontologia inicial para definição das classes do acervo, (definidas em *classes.owl*) pode ser acessada pela URI <http://200.137.241.247:8080/fuseki/classe>. Por fim, a Ontologia das dimensões – (*dimensoes.owl*) pode ser acessada em <http://200.137.241.247:8080/fuseki/dimensoes>.

Os TDBs para armazenamento dos objetos (físicos e dimensionais), e das classes do acervo de cada repositório são instanciados separadamente em cada instalação precedida pela base do repositório, neste caso <http://200.137.241.247:8080/fuseki/>.

### 7.2.5 Semântica

A camada semântica do protótipo é o segundo módulo externo e composto pelas três ontologias definidas no modelo conforme Seção 4.7 .

Como resultado da primeira ontologia (Ontologia 1) para a classificação dos acervos (conforme Seção 4.7.1), uma definição em OWL a partir da serialização RDF/XML é apresentada na Serialização 5 abaixo.

#### Serialização 5 - RDF/XML da ontologia para Classificação dos acervos

```
1 @prefix :      <http://200.137.241.247:8080/fuseki/classes#> .
2 @prefix rdf:   <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
3 @prefix owl: <http://www.w3.org/2002/07/owl#> .
4 @prefix xsd:   <http://www.w3.org/2001/XMLSchema#> .
5 @prefix rdfs:  <http://www.w3.org/2000/01/rdf-schema#> .
6
7 rdfs:comment a owl:DatatypeProperty .
8
9 rdfs:label a owl:DatatypeProperty .
10
11 rdfs:subClassOf a owl:DatatypeProperty .
12
13 :Classe a owl:Class .
14
15 <http://200.137.241.247/ontologies/classes>
16 a owl:Ontology .
17
```

Fonte: Elaborada pelo autor.

A segunda ontologia (Ontologia 2), por sua vez, prevê a definição inicial dos objetos do acervo, conforme proposta apresentada na Seção 4.7.2 (p. 138). Dessa forma, uma possível definição em OWL a partir da serialização RDF/XML é apresentada abaixo:

#### Serialização 6 - Serialização em RDF/XML da ontologia de Definições de Objetos

```
1 @prefix :      <http://200.137.241.247:8080/fuseki/objetos#> .
2 @prefix rdf:   <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
3 @prefix owl: <http://www.w3.org/2002/07/owl#> .
4 @prefix xsd:   <http://www.w3.org/2001/XMLSchema#> .
```

```

5 @prefix dim: <200.137.241.247:8080/fuseki/dimensoes#> .
6 @prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
7 @prefix dc: <http://purl.org/dc/elements/1.1> .
8
9 <http://200.137.241.247:8080/fuseki/objetos>
10     a          owl:Ontology .
11
12 rdf:type a          owl:DatatypeProperty .
13
14 :id      a          owl:DatatypeProperty .
15
16 <http://purl.org/dc/elements/1.1/description>
17     a          owl:DatatypeProperty .
18
19 <http://purl.org/dc/elements/1.1/subject>
20     a          owl:DatatypeProperty , owl:ObjectProperty .
21
22 <http://purl.org/dc/elements/1.1/title>
23     a          owl:DatatypeProperty .
24
25 :ObjetoDigital a owl:Class .
26
27 dim:tipoDimensional a owl:DatatypeProperty .
28
29 dim:tipoFisico a owl:DatatypeProperty .

```

Fonte: Elaborada pelo Autor.

A Serialização 7 apresenta a estruturação da Ontologia 3, conforme Seção 4.7.3 (p. 141), serializada no formato RDF/XML, onde é possível visualizar suas classes e propriedades, bem como as restrições iniciais.

#### Serialização 7 - Serialização da ontologia de definição das dimensões

```

1 @prefix : <http://200.137.241.247:8080/fuseki/dimensoes#> .
2 @prefix owl: <http://www.w3.org/2002/07/owl#> .
3 @prefix rdfs: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
4 @prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
5 @prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
6
7 <http://200.137.241.247:8080/fuseki/dimensoes>
8     a          owl:Ontology .
9
10 rdfs:subClassOf a owl:ObjectProperty .
11
12 :Espaço a          owl:Class ;
13     rdfs:subClassOf :ObjetoDimensional .
14
15 :Fenômeno a          owl:Class ;
16     rdfs:subClassOf :ObjetoDimensional .
17
18 :Pessoa a          owl:Class ;
19     rdfs:subClassOf :ObjetoDimensional .
20
21 :Tempo a          owl:Class ;
22     rdfs:subClassOf :ObjetoDimensional .
23
24
25 <http://purl.org/dc/elements/1.1/relation>
26     a          owl:ObjectProperty .
27
28 :Arqueológico a          owl:Class ;
29     rdfs:subClassOf :ObjetoFisico .
30
31 :Documental a          owl:Class ;
32     rdfs:subClassOf :ObjetoFisico .
33

```

```

34 :ImagemSom a owl:Class ;
35     rdfs:subClassOf :ObjetoFisico .
36
37 :ObjetoDimensional a owl:Class .
38
39 :Museológico a owl:Class ;
40     rdfs:subClassOf :ObjetoFisico .
41
42 :Bibliotecário a owl:Class ;
43     rdfs:subClassOf :ObjetoFisico .
44
45 :tipoDimensional a owl:ObjectProperty ;
46     rdfs:domain owl:Thing ;
47     rdfs:range :Espaço , :Fenômeno , :Pessoa , :Tempo ;
48     rdfs:subPropertyOf rdfs:subClassOf .
49
50 :tipoFisico a owl:ObjectProperty ;
51     rdfs:domain owl:Thing ;
52     rdfs:range :Arqueológico , :Documental , :ImagemSom ;
53     rdfs:subPropertyOf rdfs:subClassOf .
54
55 :ObjetoFisico a owl:Class .
56
57 :relacionaTempo a owl:ObjectProperty , owl:DatatypeProperty ;
58     rdfs:domain owl:Thing ;
59     rdfs:range :Tempo , xsd:dateTime , xsd:dateTimeStamp ,
60     xsd:string ;
61     rdfs:subPropertyOf <http://purl.org/dc/elements/1.1/relation> .
62
63 :relacionaPessoa a owl:ObjectProperty ;
64     rdfs:domain owl:Thing ;
65     rdfs:range :Pessoa ;
66     rdfs:subPropertyOf <http://purl.org/dc/elements/1.1/relation> .
67
68 :relacionaEspaço a owl:ObjectProperty ;
69     rdfs:domain owl:Thing ;
70     rdfs:range :Espaço ;
71     rdfs:subPropertyOf <http://purl.org/dc/elements/1.1/relation> .
72
72 :relacionaFenômeno a owl:ObjectProperty ;
74     rdfs:domain owl:Thing ;
75     rdfs:range :Fenômeno ;
76     rdfs:subPropertyOf <http://purl.org/dc/elements/1.1/relation> .

```

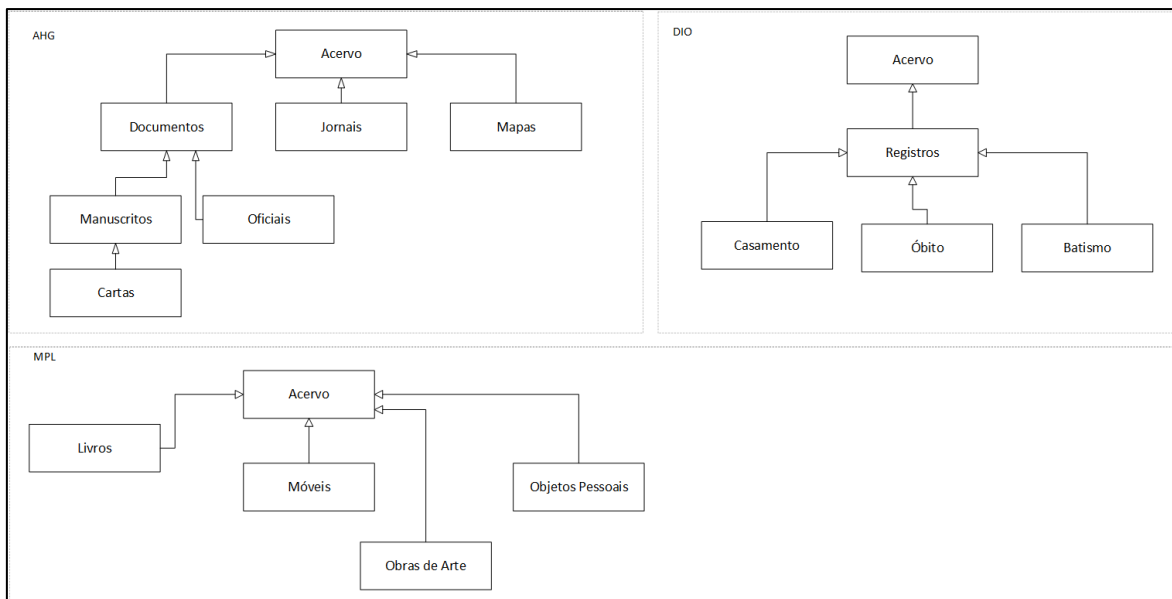
Fonte: Elaborada pelo Autor.

### 7.3 Exemplos de interação

A seguir são apresentados exemplos de interação de partir de cenários-exemplo. Para visualizar o aspecto funcional do protótipo e demonstrar as atividades que os curadores estão aptos a realizar, o comportamento esperado pelo sistema em algumas situações e, em especial, o que diz respeito à interoperabilidade entre os repositórios e as tarefas junto aos acervos. Para esses exemplos, consideramos uma situação hipotética com três repositórios e três curadores, respectivamente.

A Figura 74 apresenta uma visão da estrutura de classes dos repositórios de exemplo.

**Figura 74 – Estrutura dos Acervos dos Repositórios de Exemplo**



Fonte: Elaborada pelo Autor.

### Repositório 1 - AHG

O primeiro repositório, gerido pelo Curador 1, simboliza um espaço arquivístico (inspirado no Arquivo Histórico de Goiás) constante de um acervo de documentos, jornais e outros artefatos. Este repositório será chamado de AHG. Este acervo é dividido em Documentos manuscritos e oficiais, mapas e jornais.

O endereço do repositório está disponível em <http://200.137.241.247:8080/ahg>. O TDB com suas classes do acervo está acessível pelo *endpoint* <http://200.137.241.247:8080/fuseki/ahgclass>, e o TDB com seus objetos do acervo em <http://200.137.241.247:8080/fuseki/ahgobj>.

### Repositório 2 - DIO

O Segundo repositório, denominado DIO, é um espaço de memória de um arquivo diocesano que contém os registros de batismo, casamento e óbito, inspirado na antiga Diocese de Vila Boa. Esse repositório é gerido pelo Curador 2. O endereço do repositório está disponível em <http://200.137.241.247:8080/diocese>. O TDB com suas classes do acervo está acessível pelo *endpoint* <http://200.137.241.247:8080/fuseki/dioceseclass>, e o TDB com seus objetos do acervo em <http://200.137.241.247:8080/fuseki/dioceseobj>.

### Repositório 3 - MPL

O terceiro, um espaço museológico inspirado no Museu Pedro Ludovico, o qual chamaremos de MPL. Este último é gerido pelo Curador 3. O endereço do repositório está disponível em <http://200.137.241.247:8080/mpl>. O TDB com suas classes do acervo está acessível pelo *endpoint* <http://200.137.241.247:8080/fuseki/mplclass>, e o TDB com seus objetos do acervo em <http://200.137.241.247:8080/fuseki/mplobj>.

Os três acervos possuem documentos digitais, que podem ou não serem complementares entre si. Os curadores podem adicionar classes de acervo para classificar os objetos, registrar os objetos que podem ser físicos ou dimensionais e gerir suas relações com outros objetos.

#### Cenários:

1. A primeira interação do exemplo está relacionada a adicionar uma classe. O Curador 1, que gere o espaço (AHG), utiliza as atividades de gestão do acervo que compete Definir Acervo (pesquisar, modificar e destruir classes) e adiciona a primeira classe.
2. Na segunda interação, o Curador 2, que gere o espaço 2 (aqui chamado DIO), adiciona um Objeto Físico utilizando as funcionalidades de gestão do acervo que diz respeito a pesquisar, modificar e destruir objetos físicos conforme Seção 6.4.2 (p. 184).
3. Na terceira interação, o Curador 3, também na funcionalidade de Gerir Acervos Digitais, adiciona um Objeto Dimensional conforme Seção 6.4.3 (p. 193).
4. Na quarta interação, o Curador 2 Adiciona uma propriedade conforme Seção 6.4.4 (p. 199).
5. Na quinta o sistema aciona o motor de interoperabilidade com a funcionalidade de Coletar Ligações conforme 6.4.5 (p. 203).

#### **7.3.1 Cenário 1: Adicionando uma Classe no acervo**

##### *7.3.1.1 Descrição*

A partir das atividades de gestão do acervo, é possível disparar a funcionalidade para adicionar uma nova classe através da ontologia inicial. A Ontologia OWL inicial de

definições e classificação do acervo é recuperada do endereço (<http://200.137.241.247:8080/fuseki/classes>) conforme Seção 4.7.1 (p. 135), na qual cada nova classe é criada respeitando a relação “:Classe a rdfs:Class” e associada a estrutura superior através da relação “:Classe rdfs:subClassOf <URI\_da\_Classe> .” O TDB do acervo está presente no tdb “*aghclass*”, na mesma base do repositório, sendo acessível pelo endereço <http://200.137.241.247:8080/fuseki/aghclass>.

Neste estado inicial o TDB das classes do repositório conta apenas com a classe primitiva “Acervo”, o Curador 1 irá adicionar uma classe denominada “Documentos” para agrupar os documentos existentes no espaço.

A situação do TDB “*aghclass*”, no início da operação, pode ser observada na serialização Serialização 8 abaixo:

#### Serialização 8 - Situação Inicial do TDB de classes do Acervo no Cenário 1.

```
1 @prefix : <http://200.137.241.247:8080/fuseki/ahgclass#> .
2 @prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
3 @prefix owl: <http://www.w3.org/2002/07/owl#> .
4 @prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
5 @prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
6 @prefix cls: <http://200.137.241.247:8080/fuseki/classes#> .
7
8 :Acervo a owl:Class ;
9         rdfs:comment "Classe que agrupa todos os registros" ;
10        rdfs:label "Acervo" .
11
12 <http://200.137.241.247:8080/fuseki/ahgclass>
13     a owl:Ontology .
```

Fonte: Elaborada pelo Autor.

#### 7.3.1.2 Ações do Curador 1

O Curador aciona a opção Ontologias/Adicionar classe e é confrontado com a ontologia de definição de classes já presentes no acervo e com a única classe “Acervo” presente no TDB das classes do acervo.

O Curador 1, após esse processo inicial, seleciona a superclasse (#Acervo), informa a URI da classe (#Documento), informa o rótulo para as classes (que se chamará “Documentos”) e uma descrição, que será “Classe para abrigar todos os documentos do acervo do arquivo”.

Este curador, então, aciona o botão de “Gravar”. Após a gravação, pesquisa as classes do acervo para verificar se a classe “Documento” foi criada.

A Figura 75 apresenta a primeira interação do Curador 1 ao inserir nova classe

de acervo.

**Figura 75 – Tela da Ação no Cenário 1.**

The screenshot shows a web form titled "Incluir Nova Classe" (Add New Class) in the Guará RDF Engine interface. The form includes the following fields and controls:

- Origem: (Superclasse):** A text input field containing the URI `http://200.137.241.247:8080/fuseki/ahgclass#Acervo` and a "Pesquisar" (Search) button.
- URI: (owl:Class):** A text input field containing `http://200.137.241.247:8080/fuseki/ahgclass#Documento`.
- Rótulo: (rdfs:label):** A text input field containing the label "Documentos".
- Comentário/Descrição: (rdfs:comment):** A text area containing the description "Classe para abrir os documentos do acervo do Arquivo." and a "Gravar" (Save) button.
- At the bottom of the form, there are "Gravar" (Save) and "Voltar" (Back) buttons.

Fonte: Elaborada pelo Autor.

### 7.3.1.3 Resultados Esperados

Ao final do processo, espera-se que a classe informada pelo Curador tenha sido gravada e ao se acionar a opção pesquisar, a classe “Documento” deve ter sido adicionada às classes do acervo.

### 7.3.1.4 Comportamento do sistema

A Figura 76 mostra a tela com o resultado da ação disparada pelo Curador 2.

**Figura 76 – Tela do resultado da ação no Cenário 1**

The screenshot shows the "Classes" management page in the Guará RDF Engine interface. It features a search bar with the text "pesquisar" and a table listing the classes. The table has the following columns: "Classe: (owl:Class)", "Rótulo: (rdfs:label)", "Comentário/Descrição: (rdfs:comment)", "Filha de: (rdfs:subClassOf)", and "Ação".

Classe: (owl:Class)	Rótulo: (rdfs:label)	Comentário/Descrição: (rdfs:comment)	Filha de: (rdfs:subClassOf)	Ação
<a href="http://200.137.241.247:8080/fuseki/ahgclass#Acervo">http://200.137.241.247:8080/fuseki/ahgclass#Acervo</a>	Acervo	Classe que agrupa todos os registros		<a href="#">Editar</a> <a href="#">Destruir</a>
<a href="http://200.137.241.247:8080/fuseki/ahgclass#Documento">http://200.137.241.247:8080/fuseki/ahgclass#Documento</a>	Documentos	Classe para abrir os documentos do acervo do Arquivo.	<a href="http://200.137.241.247:8080/fuseki/ahgclass#Acervo">http://200.137.241.247:8080/fuseki/ahgclass#Acervo</a>	<a href="#">Editar</a> <a href="#">Destruir</a>

Fonte: Elaborada pelo Autor.

Utilizando uma conexão como TDB “ahgclass” via API do Jena com a URL “jdbc:jena:remote:query=http://200.137.241.247:8080/fuseki/ahgclass/query&update=http://200.137.241.247:8080/fuseki/ahgclass/”, o sistema realizou as verificações pertinentes nos dados informados pelo Curador e persistiu o grafo do TDB correspondente executando a seguinte consulta SPARQL.

#### Algoritmo 11 – SPARQL do Cenário 1 – Inserindo uma Nova Classe.

```
1 [...]
2 PREFIX obj: <http://200.137.241.247:8080/fuseki/objetos#>
3 PREFIX cls: <http://200.137.241.247:8080/fuseki/classes#>
4 PREFIX dim: <http://200.137.241.247:8080/fuseki/dimensoes#>
5 PREFIX owl: <http://www.w3.org/2002/07/owl#>
6
7 INSERT DATA{
8 <http://200.137.241.247:8080/fuseki/ahgclass#Documento> a owl:Class ;
9     rdfs:comment      "Classe para abrigar todos os documentos do
10 Arquivo";
11     rdfs:label        "Documentos" ;
12     rdfs:subClassOf
13 <http://200.137.241.247:8080/fuseki/ahgclass#Acervo> .
14 }
```

Fonte: Elaborado pelo Autor.

Ao final da operação, ao se executar a chamada SPARQL "SELECT ?class ?label ?description ?subclassof WHERE{?class a owl:Class. OPTIONAL { ?class rdfs:label ?label} OPTIONAL { ?class rdfs:comment ?description} OPTIONAL { ?class rdfs:subClassOf ?subclassof} } um grafo é retornado com as informações persistidas conforme as triplas resultantes do arquivo JSON abaixo, gerado pelo *endpoint*.

#### Serialização 9 – Retorno em JSON da operação de Inserir Nova Classe

```
1 { "head": {
2   "vars": [ "class" , "label" , "description" , "subclassof" ]
3 } ,
4 "results": {
5   "bindings": [
6     {
7     "class": { "type": "uri" , "value":
8 "http://200.137.241.247:8080/fuseki/ahgclass#Acervo" } ,
9     "label": { "type": "literal" , "value": "Acervo" } ,
10    "description": { "type": "literal" , "value": "Classe que agrupa
11 todos os registros" }
12    } ,
13    {
14    "class": { "type": "uri" , "value":
15 "http://200.137.241.247:8080/fuseki/ahgclass#Documento" } ,
16    "label": { "type": "literal" , "value": "Documentos" } ,
17    "description": { "type": "literal" , "value": "Classe para
18 abrigar os documentos do acervo do Arquivo." } ,
19    "subclassof": { "type": "uri" , "value":
20 "http://200.137.241.247:8080/fuseki/ahgclass#Acervo" }
21    }
22  ]
23 }
```

Fonte: Elaborado pelo Autor.

Ao final da interação e, ao se observar o resultado da consulta, nota-se que o

resultado foi o esperado e a classe “Documentos” foi persistida no TDB “*ahgclass*” com um novo grafo referente às triplas “:Documento a owl:Class ; rdfs:comment “Classe para abrigar os documentos do acervo do Arquivo.” ; rdfs:label “Documentos” ; rdfs:subClassOf :Acervo .”.

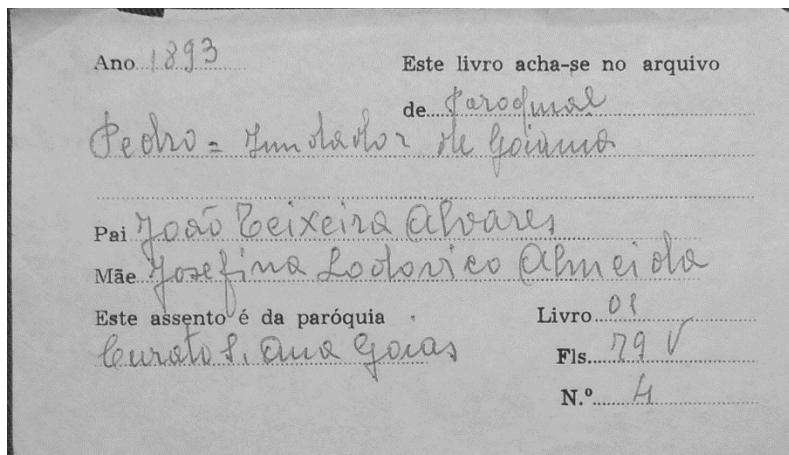
### 7.3.2 Cenário 2: Adicionando um Objeto Físico

#### 7.3.2.1 Descrição

Neste cenário, o Curador 2 Adiciona um novo Objeto Físico. Leva-se em consideração que o TDB das classes do acervo “*dioceseclass*” já se encontra com as classes definidas, no entanto no TDB de objetos “*dioceseobj*” ainda não há nenhum objeto e, portanto, o Curador 2 irá adicionar o primeiro objeto que se trata de um artefato do tipo Registro de Batismo.

Este curador adicionará um Objeto Físico, que é um registro de batismo de Pedro Ludovico Teixeira, o qual consta naquela paróquia no livro 01, Fls 79 V No. 4 conforme ficha de arquivo da imagem abaixo.

**Figura 77 – Ficha de batismo de Pedro Ludovico Teixeira**



A imagem mostra uma ficha de batismo manuscrita em português. O formulário contém os seguintes dados:

Ano	1893	Este livro acha-se no arquivo
		de
Pedro = filho de		
Pai	João Beixeira Alvares	
Mãe	Josefina Ludovico Almeida	
Este assento é da paróquia		Livro
curato S. Ant. Goias		Fls.
		N.º

Fonte: Imagem cedida pela Arquidiocese de Goiás, exemplo retirado de um processo de digitalização de uma atividade piloto do Projeto Centro de Memória de Goiás.

Esta funcionalidade permite ao Curador, Conforme Seção 6.4.2 (p. 184), pesquisar, modificar e destruir objetos físicos. Para que este possa disparar a funcionalidade de adicionar um novo Objeto Físico a estrutura inicial deve estar presente, com a classe referente à coleção a qual o objeto faz parte, já adicionada e acessível no endereço <http://200.137.241.247:8080/fuseki/aghclass>.

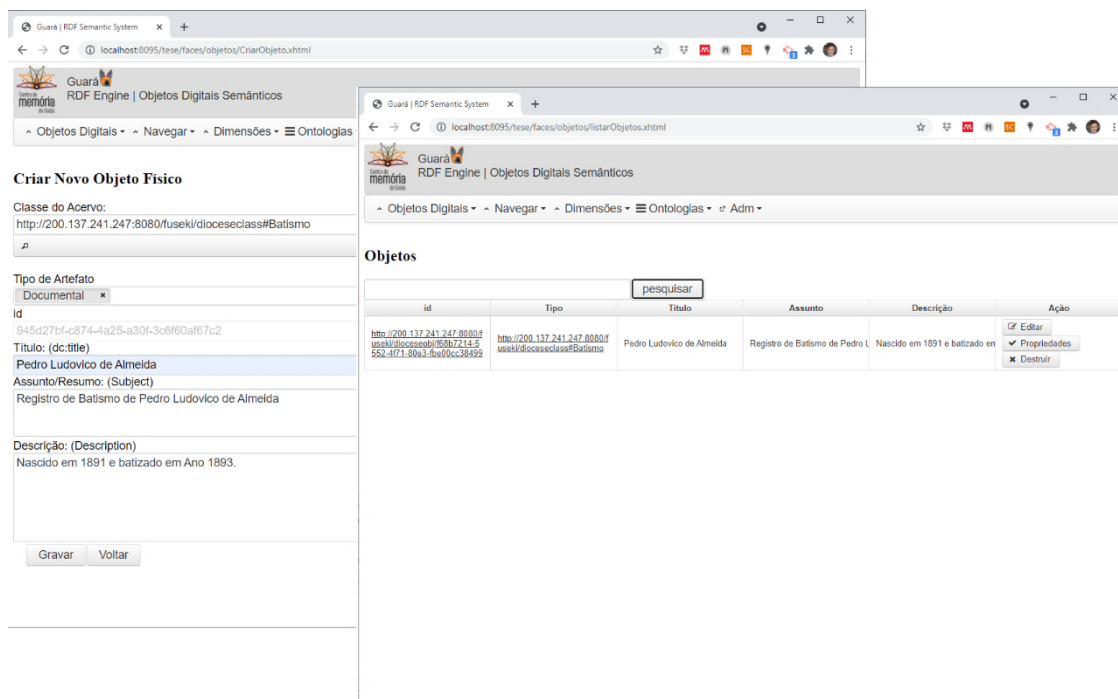
### 7.3.2.2 Ações do Curador 2

O Curador 2 aciona a opção Objetos Digitais/Novo Objeto Físico e informa os dados do Objeto Físico, selecionando primeiro qual classe aquele objeto pertence no acervo.

Este curador seleciona a classe do acervo, seleciona na lista os tipos de artefato alcançados pelo objeto, informa o título do objeto, informa um assunto e uma descrição sucinta e aciona em seguida a opção de gravar.

O Curador 2 aciona a opção Objetos/Pesquisar objetos para verificar se o objeto criado se encontra na lista de objetos persistidos no TDB. A Figura 78 mostra as ações do Curador 2 nos processos de gravação e pesquisa do Objeto Físico desse exemplo de interação.

**Figura 78 – Ações do Curador no Exemplo de Interação 2**



Fonte: elaborada pelo Autor

### 7.3.2.3 Resultados Esperados

Ao final do processo, espera-se que o Objeto Físico descrito pelo Curador tenha sido gravado após terem sido realizadas as devidas validações. Também que, ao se acionar a opção pesquisar, o objeto tenha sido criado e recebido um id único persistido no Acervo. Ao clicar no *link* da *URI* do objeto, que o Sistema tenha presente todas as relações do objeto selecionado em uma nova janela, bem como essas relações também devem ser

carregadas na opção “Propriedades”. Caso o usuário tenha clicado em Editar o objeto, o sistema deverá estar com o objeto criado para ser alterado.

#### 7.3.2.4 Comportamento do Sistema

O Sistema conectou-se com os dois bancos de dados (uma conexão para as classes do acervo e outra para os objetos). Para manipular as classes, utilizou uma conexão com o TDB “*dioceseclass*” via API do *Jena* através da URL “*jdbc:jena:remote:query=http://200.137.241.247:8080/fuseki/dioceseclass/query*” e, para manipular os objetos, com o TDB de objetos através da URL “*jdbc:jena:remote:query=http://200.137.241.247:8080/fuseki/dioceseobj/query*” &update=*http://200.137.241.247:8080/fuseki/dioceseobj*”. O sistema realizou as verificações pertinentes nos dados informados pelo Curador e persistiu o grafo do TDB correspondente executando a seguinte consulta SPARQL.

#### Algoritmo 12 – SPARQL do Cenário 2 – Inserindo uma Novo Objeto Físico.

```
1 [...]
2
3 PREFIX obj: <http://200.137.241.247:8080/fuseki/objetos#>
4 PREFIX cls: <http://200.137.241.247:8080/fuseki/classes#>
5 PREFIX dim: <http://200.137.241.247:8080/fuseki/dimensoes#>
6 PREFIX owl: <http://www.w3.org/2002/07/owl#>
7
8 INSERT DATA
9 {<http://200.137.241.247:8080/fuseki/dioceseobj/f68b7214-5552-4f71-80e3-
10 fbe00cc38499> dc:title "Pedro Ludovico Teixeira";
11   a dim:ObjetoFisico; dim:tipoFisico dim:Documental;
12
13   rdfs:type <http://200.137.241.247:8080/fuseki/dioceseclass#Batismo>;
14
15   dc:subject "Registro de Batismo de Pedro Ludovico Teixeira";
16
17   dc:description "Nascido em 1891. Batizado em 1893. Fls 01. Livro 04.
18   Pedro Ludovico Teixeira, filho de João Teixeira Alvares e Josefina
19   Ludovico Almeida."
20 }
21
22
```

Fonte: Elaborado pelo Autor.

Ao final da operação, ao se executar a chamada SPARQL, “SELECT DISTINCT ?s ?v ?o ?q WHERE{ ?s a dim:ObjetoFisico.?s dc:title ?v. ?s dc:subject ?o. ?s rdfs:type ?q. filter (regex(?s,',' , 'i') || regex(?v,',' , 'i') )}” o sistema retornou um grafo com as informações persistidas conforme as triplas resultantes do arquivo JSON que foi gerado pelo *endpoint* segundo a Serialização 10 descrita abaixo.

**Serialização 10 – Resultado em JSON da operação de Inserir Novo Objeto Físico**

```

1 { "head": {
2   "vars": [ "s" , "v" , "o" , "q" ]
3   } ,
4   "results": {
5     "bindings": [
6       {
7         "s": { "type": "uri" , "value":
8 "http://200.137.241.247:8080/fuseki/dioceseobj/f68b7214-5552-4f71-80e3-
9 fbe00cc38499" } ,
10        "v": { "type": "literal" , "value": "Pedro Ludovico Teixeira" }
11      } ,
12        "o": { "type": "literal" , "value": "Registro de Batismo de
13 Pedro Ludovico Teixeira" } ,
14        "q": { "type": "uri" , "value":
15 "http://200.137.241.247:8080/fuseki/dioceseobj/Batismo" }
16      }
17    ]
18  }
19 }

```

Fonte: Elaborado pelo Autor.

Na consulta SPARQL processada, *s* equivale ao identificador do objeto; *v* ao nome; *o* ao assunto e *q* à classe do acervo. Ao final da interação e ao se observar o resultado da consulta, nota-se ter alcançado o resultado conforme esperado. O objeto *f68b7214-5552-4f71-80e3-fbe00cc38499* foi persistido no TDB “*dioceseobj*” com um novo grafo referente às triplas adicionadas.

Ao consultar o objeto mencionado pela sua URI, o sistema enviou uma solicitação GET através de um serviço do Gerenciador de Serviços (*services.getObject<id>*), que por sua vez, conforme ( Seção 6.4.2.2 e 6.4.2.5 ), enviou para o TDB a seguinte consulta SPARQL: `SELECT DISTINCT ?prop ?val WHERE{<http://200.137.241.247:8080/fuseki/dioceseobj/f68b7214-5552-4f71-80e3-fbe00cc38499> ?prop ?val.}`; e obteve como resposta a serialização em JSON abaixo que apresenta uma lista de objetos em pares “propriedade – valor” para o Objeto Físico pesquisado.

**Serialização 11 – Resultado em JSON da Operação de Visualizar Propriedades**

```

1 { "head": {
2   "vars": [ "prop" , "val" ]
3   } ,
4   "results": {
5     "bindings": [
6       {
7         "prop": { "type": "uri" , "value":
8 "http://purl.org/dc/elements/1.1/title" } ,
9         "val": { "type": "literal" , "value": "Pedro Ludovico Teixeira"
10      }
11    } ,
12    {
13      "prop": { "type": "uri" , "value":
14 "http://www.w3.org/1999/02/22-rdf-syntax-ns#type" } ,
15    }

```

```

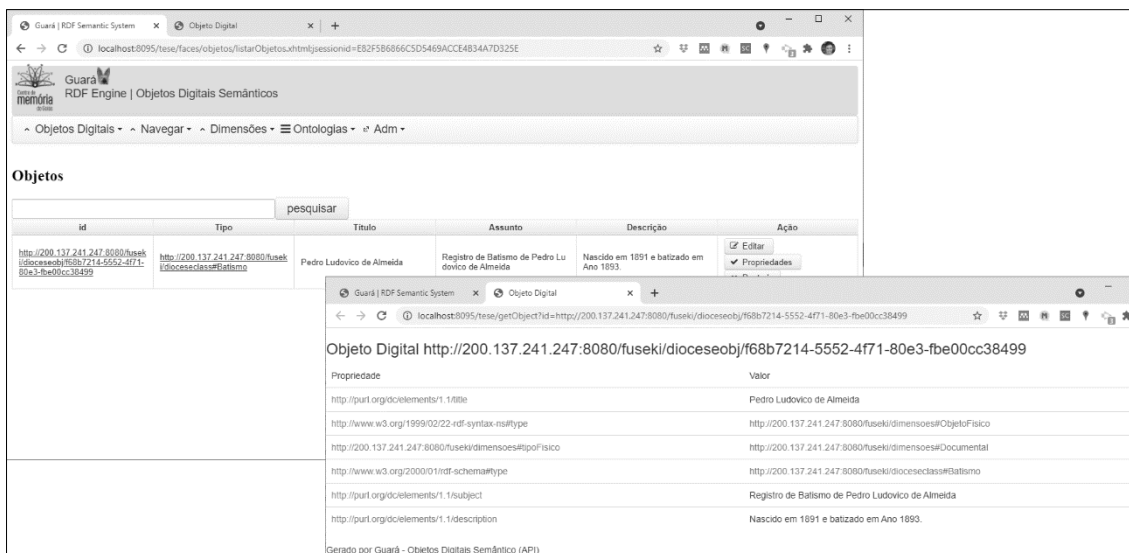
16     "val": { "type": "uri" , "value":
17 "http://200.137.241.247:8080/fuseki/dimensoes#ObjetoFisico" }
18   } ,
19   {
20     "prop": { "type": "uri" , "value":
21 "http://200.137.241.247:8080/fuseki/dimensoes#tipoFisico" } ,
22     "val": { "type": "uri" , "value":
23 "http://200.137.241.247:8080/fuseki/dimensoes#Documental" }
24   } ,
25   {
26     "prop": { "type": "uri" , "value":
27 "http://www.w3.org/2000/01/rdf-schema#type" } ,
28     "val": { "type": "uri" , "value":
29 "http://200.137.241.247:8080/fuseki/dioceseclass#Batismo" }
30   } ,
31   {
32     "prop": { "type": "uri" , "value":
33 "http://purl.org/dc/elements/1.1/subject" } ,
34     "val": { "type": "literal" , "value": "Registro de Batismo de
35 Pedro Ludovico Teixeira" }
36   } ,
37   {
38     "prop": { "type": "uri" , "value":
39 "http://purl.org/dc/elements/1.1/description" } ,
40     "val": { "type": "literal" , "value": "Nascido em 1891 e
41 batizado em Ano 1893." }
42   }
43 ]
44 }
}

```

Fonte: Elaborada pelo Autor.

A Figura 79 apresenta a operação de pesquisar o objeto e, após ter sido persistido, a visualização de propriedades na segunda parte da figura, que foi gerada a partir da ação do Curador 2 no link gerado pela URI do objeto.

**Figura 79 – Resultados da Ação do Curador 2 na Interação 2**



Fonte: Elaborada pelo Autor.

Como pôde ser observado nas serializações apresentadas na Figura 78 e na

Figura 79, os resultados foram alcançados. Na primeira parte da imagem é possível visualizar o registro do objeto criado com as opções da funcionalidade de pesquisar acervos, e na segunda parte, é possível verificar todas as relações do objeto que foram capturados pelo serviço de visualizar propriedades.

### **7.3.3 Cenário 3: Adicionando um Objeto Dimensional**

#### *7.3.3.1 Descrição*

Nesse cenário o Curador 3, responsável pelo espaço de memória MPL deve criar um Objeto Dimensional. Da mesma forma que Cenário 3, consideremos o estado do TDB das classes do acervo “*mplclass*” as classes do acervo já definidas, e o TDB de objetos “*mpleobj*” sem nenhum objeto. Dessa forma, o Curador 3 irá adicionar o primeiro objeto que trata-se do Objeto Dimensional da dimensão Pessoa.

Adicionará um Objeto Dimensional que é uma representação da pessoa “Pedro Ludovico Teixeira”, que foi Governador do Estado de Goiás após as revolução de 1930 e idealizador e fundador da nova capital do Estado, Goiânia, em 1933.

Conforme Seção 6.4.3 (p. 193), essa funcionalidade permite que o Curador, da mesma forma que com os objetos físicos, pesquise, modifique ou destrua objetos dimensionais. Objetos físicos não possuem ligação direta com as classes do acervo, a condição para essa funcionalidade, é a utilização da ontologia das Dimensões (*dimensoes.owl*).

#### *7.3.3.2 Ações do Curador 3*

O Curador aciona a opção Objetos Digitais/Novo Objeto Dimensional e informa os dados do objeto e, selecionando o tipo de objeto dimensional, que no caso é o tipo Pessoa definido na ontologia Dimensões, informa o título do objeto, um assunto e uma descrição sucinta, acionando em seguida a opção “gravar”. Após esse processo, o Curador aciona a opção Objetos/Pesquisar objetos para verificar se o objeto criado se encontra na lista de objetos persistidos no TDB.

**Na**

Figura 80 é apresentada a captura de tela da ação do Curador 3 ao adicionar a Pessoa “Pedro Ludovico Teixeira”.

**Figura 80 – Ação do Curador 3 ao Adicionar Novo Objeto Dimensional**

Guará | RDF Semantic System

localhost:8095/tese/faces/objetos/CriarObjetoDim.xhtml

Guará  
RDF Engine | Objetos Digitais Semânticos

Objetos Digitais | Navegar | Dimensões | Ontologias | Adm

### Criar Novo Objeto Dimensional

Tipo  
Pessoa

id  
a6966469-1260-415d-a4c2-1452d896a4ff

Título: (dc:title)  
Pedro Ludovico de Almeida

Assunto/Resumo: (Subject)  
Pedro Ludovico de Almeida

Descrição: (Description)  
Pedro Ludovico de Almeida. Interventor e Governador de Goiás após a revolução de 1930. Idealizar e fundador da nova capital do estado, Goiânia, em 1933.

Gravar Voltar

Fonte: Elaborada pelo Autor.

#### 7.3.3.3 Resultados Esperados

Ao final do processo, espera-se que o Objeto Dimensional adicionado pelo Curador, após terem sido realizadas as devidas validações, tenha sido persistido. Após essa ação, ao se acionar a opção pesquisar, espera-se que o objeto criado tenha recebido um id único e tenha sido persistido no Acervo. Ao clicar no *link* da *URI* do objeto, espera-se que o Sistema tenha apresentado todas as relações do objeto selecionado em uma nova janela, bem como essas relações também devem ser carregadas na opção “Propriedades”. Caso o usuário tenha clicado em Editar o objeto, espera-se que o sistema tenha carregado o objeto criado para ser alterado.

#### 7.3.3.4 Comportamento do Sistema

O Sistema conectou-se com os dois bancos de dados (uma conexão para as classes do acervo e outra para os objetos), utilizando para manipular as classes uma conexão como TDB “*mplclass*” via API do Jena com a URL “*jdbc:jena:remote:query=http://200.137.241.247:8080/fuseki/mplclass/query*”, e para manipular os objetos, com o TDB de objetos através da URL

“*jdbc:jena:remote:query=http://200.137.241.247:8080/fuseki/mplobj/query*”&update=*http://200.137.241.247:8080/fuseki/dioceseobj*”. O sistema realizou as verificações pertinentes nos dados informados pelo Curador e persistiu o grafo no TDB correspondente, executando a consulta SPARQL conforme Algoritmo 13 abaixo.

### Algoritmo 13 – SPARQL do Cenário 3 – Inserindo uma Novo Objeto Dimensional.

```

1 [...]
2 PREFIX obj: <http://200.137.241.247:8080/fuseki/objetos#>
3 PREFIX cls: <http://200.137.241.247:8080/fuseki/classes#>
4 PREFIX dim: <http://200.137.241.247:8080/fuseki/dimensoes#>
5 PREFIX owl: <http://www.w3.org/2002/07/owl#>
6 INSERT DATA{<http://200.137.241.247:8080/fuseki/mplobj/5cf7b91d-dcc0-
7 48ec-9428-84fff2b1f321> dc:title "Pedro Ludovico Teixeira";
8   a dim:ObjetoDimensional; dim:tipoDimensional dim:Pessoa;dc:subject
9   "Pedro Ludovico Teixeira";
10  dc:description "Pedro Ludovico Teixeira. Interventor e Governador de
11  Goiás após a revolução de 1930. Idealizar e fundador da nova capital do
12  estado, Goiânia, em 1933. ".
13 }

```

Fonte: Elaborado pelo Autor.

Ao final da operação, ao se executar a chamada SPARQL “SELECT DISTINCT ?s ?v ?o ?q ?x WHERE{ ?s dc:title ?v.optional{?s dc:subject ?o.} ?s dim:tipoDimensional ?q. optional{?s dc:description ?x.}filter (regex(?s,',' , 'i') || regex(?v, ',' , 'i') )}”, o sistema retornou um grafo com as informações persistidas conforme as triplas resultantes do arquivo JSON que foi gerado pelo *endpoint* de acordo com a Serialização 12 abaixo.

### Serialização 12 – Resultado em JSON da operação de Inserir Novo Objeto Físico

```

1 { "head": {
2   "vars": [ "s" , "v" , "o" , "q" , "x" ]
3 } ,
4 "results": {
5   "bindings": [
6     {
7       "s": { "type": "uri" , "value":
8 "http://200.137.241.247:8080/fuseki/mplobj/5cf7b91d-dcc0-48ec-9428-
9 84fff2b1f321" } ,
10      "v": { "type": "literal" , "value": "Pedro Ludovico Teixeira" }
11    ,
12      "o": { "type": "literal" , "value": "Pedro Ludovico Teixeira" }
13    ,
14      "q": { "type": "uri" , "value":
15 "http://200.137.241.247:8080/fuseki/dimensoes#Pessoa" } ,
16      "x": { "type": "literal" , "value": "Pedro Ludovico Teixeira.
17 Interventor e Governador de Goiás após a revolução de 1930. Idealizador
18 e fundador da nova capital do estado, Goiânia, em 1933. " }
19    }
20  ]
21 }
22 }

```

Fonte: Elaborada pelo Autor.

Na consulta SPARQL processada, *s* equivale ao identificador do objeto; *v* ao nome; *o* ao assunto e *q* à Dimensão ao que o Objeto Dimensional se refere. Ao final da

interação e ao se observar o resultado da consulta, nota-se que o resultado esperado foi alcançado, uma vez que o objeto *5cf7b91d-dcc0-48ec-9428-84fff2b1f321* foi persistido no TDB “*dioceseobj*” no endereço <http://200.137.241.247:8080/fuseki/mplobj> com um novo grafo referente às triplas adicionadas.

Ao consultar as propriedades do objeto que está sendo manipulado, o sistema enviou uma solicitação via Gerenciador de Serviços que, por sua vez (conforme Seção 6.4.2.5 p. 192), enviou para o TDB a seguinte consulta SPARQL: `SELECT DISTINCT ?prop ?val WHERE{<http://200.137.241.247:8080/fuseki/mplobj/5cf7b91d-dcc0-48ec-9428-84fff2b1f321> ?prop ?val.}`; e obteve como resposta a presente na serialização JSON abaixo que apresenta uma lista de objetos em pares “propriedade – valor” para o Objeto Físico pesquisado.

### Serialização 13 – Resultado em JSON da Operação de Visualizar Propriedades

```

1  { "head": {
2    "vars": [ "prop" , "val" ]
3  } ,
4  "results": {
5    "bindings": [
6      {
7        "prop": { "type": "uri" , "value":
8  "http://purl.org/dc/elements/1.1/title" } ,
9        "val": { "type": "literal" , "value": "Pedro Ludovico Teixeira"
10     }
11      } ,
12      {
13        "prop": { "type": "uri" , "value":
14  "http://purl.org/dc/elements/1.1/subject" } ,
15        "val": { "type": "literal" , "value": "Pedro Ludovico Teixeira"
16     }
17      } ,
18      {
19        "prop": { "type": "uri" , "value":
20  "http://purl.org/dc/elements/1.1/description" } ,
21        "val": { "type": "literal" , "value": "Pedro Ludovico Teixeira.
22  Interventor e Governador de Goiás após a revolução de 1930. Idealizador
23  e fundador da nova capital do estado, Goiânia, em 1933. " }
24      } ,
25      {
26        "prop": { "type": "uri" , "value":
27  "http://www.w3.org/1999/02/22-rdf-syntax-ns#type" } ,
28        "val": { "type": "uri" , "value":
29  "http://200.137.241.247:8080/fuseki/dimensoes#ObjetoDimensional" }
30      } ,
31      {
32        "prop": { "type": "uri" , "value":
33  "http://200.137.241.247:8080/fuseki/dimensoes#tipoDimensional" } ,
34        "val": { "type": "uri" , "value":
35  "http://200.137.241.247:8080/fuseki/dimensoes#Pessoa" }
36      }
37    ]
38  }
39 }

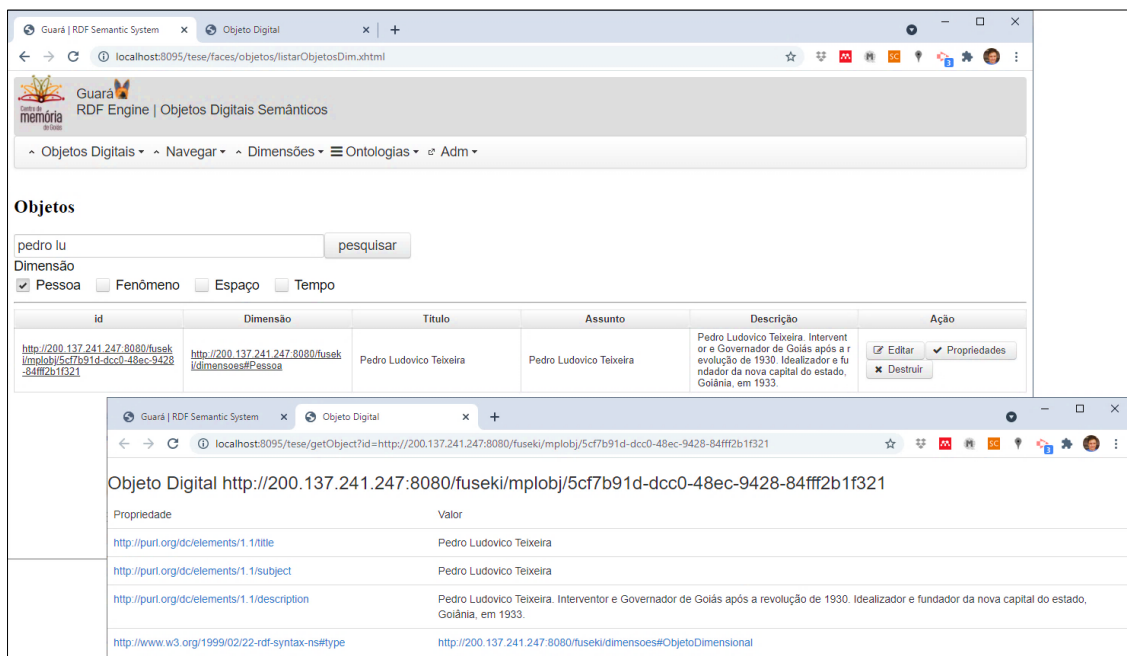
```

Fonte: Elaborada pelo Autor.

Abaixo, a Figura 81 apresenta os resultados das operações executadas pelo

Curador 3 na interação deste cenário, mostrando, na primeira parte da imagem, o objeto após ter sido persistido e, a visualização de suas propriedades na segunda parte da figura, que foi gerada a partir da ação do Curador 2 no *link* gerado pela URI do objeto.

**Figura 81 – Resultado da Ação do Curador 3 na Interação 3**



Fonte: Elaborada pelo Autor.

### 7.3.4 Cenário 4: Adicionando uma Nova Relação

#### 7.3.4.1 Descrição

Nesse cenário, o Curador 2 adiciona uma nova relação em um objeto, conforme funcionalidade descrita na Seção 6.4.4 (p. 199), que mostra o como Curador pode adicionar ou remover relações (propriedade – valor) dos objetos físicos e dimensionais.

O Curador 2 irá adicionar uma relação do registro de batismo, criado no Cenário 2, com a Pessoa “Pedro Ludovico Teixeira”, criada no Cenário 3. Isto é, no repositório do espaço “DIO” será criada uma relação do objeto “Registro de Batismo de Pedro Ludovico Teixeira” com o objeto “Pedro Ludovico Teixeira”, que está no repositório MPL.

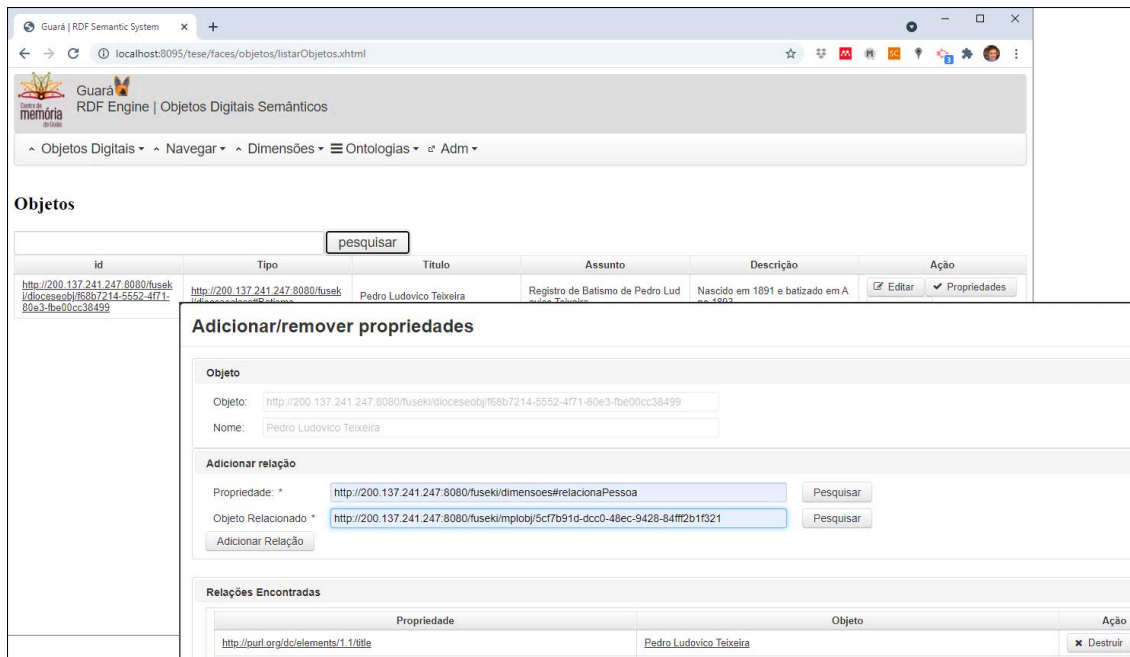
#### 7.3.4.2 Ações do Curador 2

O Curador 2 aciona a opção Objetos Digitais/Pesquisar Objetos Físicos, e na lista, seleciona o objeto *f68b7214-5552-4f71-80e3-fbe00cc38499*, que é o objeto criado no Cenário 1. Clica na opção Adicionar Propriedades. Com as relações carregadas,

adiciona a propriedade *relacionaPessoa* com o Objeto Dimensional Pessoa que foi criado no Cenário 2. Preenche propriedade e valor e clica em Adicionar Propriedade. Após o processo ser concluído, este Curador verifica a lista de propriedades.

A Figura 82 abaixo apresenta a captura de tela das ações do Curador 2 ao adicionar uma nova propriedade.

**Figura 82 – Ações do Curador no Cenário 3 – Adicionar Nova Propriedade**



Fonte: Elaborada pelo Autor.

#### 7.3.4.3 Resultados Esperados

Ao final do processo Adicionar Nova Propriedade, espera-se que a relação adicionada pelo Curador tenha sido persistida com a propriedade e valor correspondentes, após terem sido realizadas as devidas validações. Após essa ação, espera-se que a relação criada apareça da lista de relações do objeto através da função “Propriedades”.

#### 7.3.4.4 Comportamento do Sistema

O Sistema conectou-se com o TDB de objetos do acervo através da URL “*jdbc:jena:remote:query=http://200.137.241.247:8080/fuseki/diocesej/query*” &update =*http://200.137.241.247:8080/fuseki/dioceseobj*”. O sistema carregou as propriedades e valores de todas as relações do objeto. O Curador informou a nova propriedade e o novo valor da relação e pressionou a opção “Adicionar Relação”.

O Algoritmo 14 abaixo mostra a consulta SPARQL que foi utilizada para a ação.

#### Algoritmo 14 – SPARQL do Cenário 2 – Inserindo uma Nova Relação ao Objeto

```
1 [...]
2 PREFIX obj: <http://200.137.241.247:8080/fuseki/objetos#>
3 PREFIX cls: <http://200.137.241.247:8080/fuseki/classes#>
4 PREFIX dim: <http://200.137.241.247:8080/fuseki/dimensoes#>
5 PREFIX owl: <http://www.w3.org/2002/07/owl#>
6 INSERT DATA{
7   <http://200.137.241.247:8080/fuseki/dioceseobj/f68b7214-5552-4f71-
8   80e3-fbe00cc38499>
9   <http://200.137.241.247:8080/fuseki/dimensoes#relacionaPessoa>
10  <http://200.137.241.247:8080/fuseki/mplobj/5cf7b91d-dcc0-48ec-
11  9428-84fff2b1f321>}
```

Fonte: Elaborado pelo Autor.

Ao final da operação, ao se executar a chamada SPARQL “SELECT DISTINCT ?prop ?val WHERE{ <http://200.137.241.247:8080/fuseki/dioceseobj/f68b7214-5552-4f71-80e3-fbe00cc38499> ?prop ?val.}” o sistema retornou um grafo com as relações do objeto persistidas conforme as triplas resultantes do arquivo JSON, que foi gerado pelo *endpoint* conforme a Serialização 14 abaixo.

#### Serialização 14 – Resultado em JSON das relações manipuladas no Cenário 4.

```
1 { "head": {
2   "vars": [ "prop" , "val" ]
3   } ,
4   "results": {
5     "bindings": [
6       {
7         "prop": { "type": "uri" , "value":
8         "http://purl.org/dc/elements/1.1/title" } ,
9         "val": { "type": "literal" , "value": "Pedro Ludovico Teixeira"
10        }
11      } ,
12      {
13        "prop": { "type": "uri" , "value":
14        "http://www.w3.org/1999/02/22-rdf-syntax-ns#type" } ,
15        "val": { "type": "uri" , "value":
16        "http://200.137.241.247:8080/fuseki/dimensoes#ObjetoFisico" }
17      } ,
18      {
19        "prop": { "type": "uri" , "value":
20        "http://200.137.241.247:8080/fuseki/dimensoes#tipoFisico" } ,
21        "val": { "type": "uri" , "value":
22        "http://200.137.241.247:8080/fuseki/dimensoes#Documental" }
23      } ,
24      {
25        "prop": { "type": "uri" , "value":
26        "http://www.w3.org/2000/01/rdf-schema#type" } ,
27        "val": { "type": "uri" , "value":
28        "http://200.137.241.247:8080/fuseki/dioceseobj#Batismo" }
29      } ,
30      {
31        "prop": { "type": "uri" , "value":
32        "http://purl.org/dc/elements/1.1/subject" } ,
33        "val": { "type": "literal" , "value": "Registro de Batismo de
34        Pedro Ludovico Teixeira" }
35      } ,
36      {
37
```

```

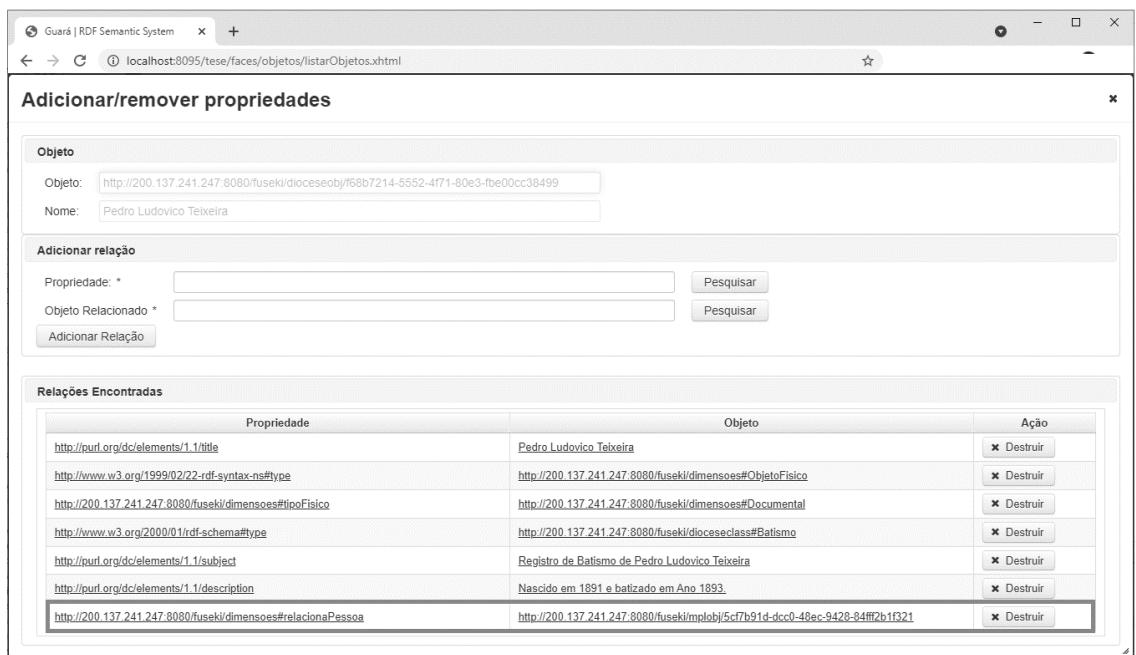
38     "prop": { "type": "uri" , "value":
39 "http://purl.org/dc/elements/1.1/description" } ,
40     "val": { "type": "literal" , "value": "Nascido em 1891 e
41 batizado em Ano 1893." }
42   } ,
43   {
44     "prop": { "type": "uri" , "value":
45 "http://200.137.241.247:8080/fuseki/dimensoes#relacionaPessoa" } ,
46     "val": { "type": "uri" , "value":
47 "http://200.137.241.247:8080/fuseki/mplobj/5cf7b91d-dcc0-48ec-9428-
48 84fff2b1f321" }
49   }
50 ]
51 }}
    
```

Fonte: Elaborado pelo Autor.

Na consulta SPARQL processada, *prop* equivale à propriedade que está associada ao objeto; *val* equivale ao valor da propriedade que pode ser uma URI ou um valor textual. O objeto *b3872c51-e75a-40b6-8f34-99dba1d5cf7e* que está persistido no TDB “*dioceseobj*”, no endereço <http://200.137.241.247:8080/fuseki/dioceseobj>, possui todas as propriedades apresentadas na serialização que foi mostrada acima.

Como é possível visualizar na Figura 83 abaixo, os resultados das operações executadas pelo Curador 2 na interação deste, cenário mostram o objeto recebeu a nova propriedade <<http://200.137.241.247:8080/fuseki/dimensoes#relacionaPessoa>> tendo como valor o objeto <<http://200.137.241.247:8080/fuseki/mplobj/5cf7b91d-dcc0-48ec-9428-84fff2b1f321>> que se encontra no repositório do espaço 2 (MPL).

**Figura 83 – Resultado da Ação do Curador 3 no Cenário 4**



Fonte: Elaborada pelo Autor.

### 7.3.5 Cenário 5: Coletar ligações entre repositórios

#### 7.3.5.1 Descrição

Nessa ação, que é acionada pelo Curador indiretamente quando esta pesquisa por objetos ou inclui um novo Objeto Digital, o motor de interoperabilidade é acionado com o intuito de coletar ligações possíveis entre objetos do repositório interno e dos *repositórios amigos*. Faz isso primeiro por palavra-chave e em seguida pelas relações conforme Seção 6.4.5 (p. 203) e Seção 7.2.3.4 (p. 218).

O estado dos repositórios para esse cenário se encontra conforme as seguintes definições: Repositório 1 - (AHG) possui um Objeto Físico do tipo “Mapa” referente a um mapa de Goiânia, da fundação da capital em 1933 com *id ad3816a0-8e2b-4a8a-befd-96a69beb0511*. Repositório 2 - (DIO) possui um Objeto Físico do tipo “Registro de Batismo” referente ao batismo de “Pedro Ludovico Teixeira”, fundador de Goiânia, com *id f68b7214-5552-4f71-80e3-fbe00cc38499*. Repositório 3 - (MPL) possui um Objeto Dimensional do tipo “Pessoa”, referente a “Pedro Ludovico Teixeira”, com *id 5cf7b91d-dcc0-48ec-9428-84fff2b1f321*.

**Quadro 19 – Estado dos Repositórios no Cenário 5**

URI	Descrição	Repositório	Tipo
ad3816a0-8e2b-4a8a-befd-96a69beb0511	Mapa de Goiânia	AHG	Físico
f68b7214-5552-4f71-80e3-fbe00cc38499	Registro de Batismo de Pedro Ludovico	DIO	Físico
5cf7b91d-dcc0-48ec-9428-84fff2b1f321	Pedro Ludovico de Almeida	MPL	Dimensional
b3872c51-e75a-40b6-8f34-99dba1d5cf7e	Livro de Memórias de Pedro Ludovico de Almeida	MPL	Físico

Fonte: Elaborado pelo Autor.

#### 7.3.5.2 Ações do Curador 3

O Curador adiciona um novo Objeto Físico do tipo “Livro” para que Motor de Interoperabilidade possa pesquisar por relações potenciais do objeto. Trata-se do Livro de memórias de Pedro Ludovico Teixeira.

#### 7.3.5.3 Resultados Esperados

Espera-se que o Sistema colete, nos *repositórios amigos* (DIO e AGH), a lista de relações possíveis para o objeto com base no que foi definido na Seção e que o sistema persista as relações que o Curador adicionar após a lista de relações para o objeto ser exibida.

#### 7.3.5.4 *Comportamento do Sistema*

O adição do Objeto Físico “Livro”, conforme Seção 7.3.2 (p. 230), permitiu que o objeto fosse persistido no TDB de objetos (“*mplobj*”) com o id ***b3872c51-e75a-40b6-8f34-99dba1d5cf7e***.

O Sistema acionou o Motor de Interoperabilidade através do método *getRelation(objeto)* que executou as consultas SPARQL invocando o algoritmo de interoperabilidade primeiro pelas palavras-chave depois pelas relações do objeto. Este processo seguiu os seguintes passos:

a) Buscas no Repositório AHG

Na primeira interação, o motor de interoperabilidade enviou para o repositório AHG a SPARQL do Algoritmo para Pesquisar Objetos Físicos por palavra-chave e não retornou nenhum registro, já que o único registro no TDB é um mapa que não tem a palavra-chave procurada no nome, no assunto ou na descrição. Na segunda interação, o motor de interoperabilidade também não retornou nenhum registro, uma vez que ainda não há nenhum Objeto Dimensional no repositório. Na terceira interação, o algoritmo de consultar Objetos Dimensionais foi disparado e também não encontrou nenhum resultado. O mesmo ocorreu para a quarta interação.

b) Buscas no Repositório DIO

Na primeira interação o Motor de Interoperabilidade encontrou o objeto ***f68b7214-5552-4f71-80e3-fbe00cc38499*** que possui no título a palavra-chave do título do objeto alvo “Pedro Ludovico Teixeira”. Trata-se do Registro de Batismo de Pedro Ludovico Teixeira. Na segunda interação nada foi encontrado, já que o TDB não possui ainda nenhum Objeto Dimensional. Na quarta interação também é retornado o objeto ***f68b7214-5552-4f71-80e3-fbe00cc38499***, referente ao registro de batismo de Pedro Ludovico.

c) Buscas no Repositório MPL

Na primeira interação o Motor de Interoperabilidade encontrou o objeto de id ***b3872c51-e75a-40b6-8f34-99dba1d5cf7e*** referente ao Livro “Memórias de Pedro Ludovico Teixeira”, que se trata do mesmo objeto alvo; portanto, não aproveitada. A segunda interação, retornou o objeto com id ***5cf7b91d-dcc0-48ec-9428-84fff2b1f321***, referente à Pessoa “Pedro Ludovico Teixeira”, que suscita a possível relação

*dim:relacionaPessoa* entre o objeto alvo e Objeto Físico retornado. A terceira interação não retorna nenhuma informação. O Curador 3 não adicionou nenhuma relação extra. Na quarta interação também é retornado o objeto *b3872c51-e75a-40b6-8f34-99dba1d5cf7e*.

O Quadro 20 apresenta um resumo dos objetos retornados em cada interação do Motor de Interoperabilidade deste Cenário 5.

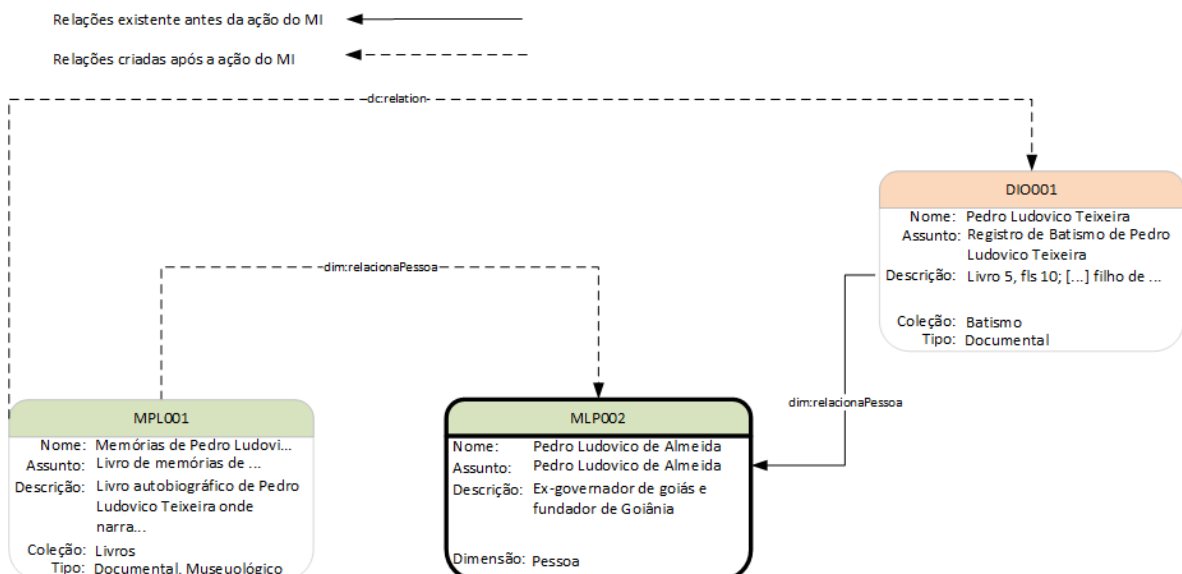
**Quadro 20 – Objetos Retornados pelo Motor de Interoperabilidade no Cenário 5**

Repo.	Por palavra-chave		3ª. Interação - Objetos Dimensionais por relação	4ª. Interação - Objetos Físicos por Dimensão
	1ª Interação - Objetos Físicos	2ª Interação - Objetos Dimensionais		
AHG	-	-	-	-
DIO	f68b7214-5552-4f71-80e3-fbe00cc38499	-	-	f68b7214-5552-4f71-80e3-fbe00cc38499
MPL	<u>b3872c51-e75a-40b6-8f34-99dba1d5cf7e</u>	5cf7b91d-dcc0-48ec-9428-84fff2b1f321	-	b3872c51-e75a-40b6-8f34-99dba1d5cf7e

Fonte: Elaborado pelo Autor.

Ao final das interações a lista de possíveis relações é formada sugerindo a relação do objeto alvo *id b3872c51-e75a-40b6-8f34-99dba1d5cf7e* com o objeto *f68b7214-5552-4f71-80e3-fbe00cc38499*, e *5cf7b91d-dcc0-48ec-9428-84fff2b1f321*. Como é possível verificar na Figura abaixo, após a ação do Motor de Interoperabilidade, duas novas relações foram adicionadas ao objeto alvo.

**Figura 84 – Ilustração do Resultado da Ação de Interoperabilidade no Cenário 5**



Fonte: Elaborada pelo Autor.

#### **7.4 Sumário do capítulo**

O objetivo desse Capítulo foi apresentar os resultados da operacionalização do protótipo do Guará, um Sistema de Informação semântico para o Patrimônio Cultural. Foram apresentados, neste capítulo, detalhes arquiteturais e exemplos de interação a partir de cenários de uso.

No próximo capítulo será apresentada uma análise dos resultados, bem como uma discussão das questões levantadas nesta tese.

## **CAPÍTULO VIII. DOS ESPAÇOS DE MEMÓRIA AO SISTEMA DE INFORMAÇÃO: UMA ANÁLISE DOS RESULTADOS**

### **8.1 Introdução**

Neste capítulo é realizada uma análise acerca dos resultados da tese proposta, que foram apresentados no capítulo anterior, em torno das evidências levantadas e dos objetivos declarados na Seção 1.7, os quais, de forma resumida, são 1) Realizar um estudo sobre as características dos espaços de memória e patrimônio ligados à Secretaria de Estado da Cultura do Estado de Goiás (SECULT-GO) e sua utilização quanto aos sistemas de informação; 2) Realizar um estudo da aplicabilidade dos modelos conceituais de representação da informação de patrimônio cultural e tecnologias da Web-Semântica; 3) Construir um modelo semântico para organização da informação de patrimônio cultural, a partir de padrões técnicos, que permita a interoperabilidade entre repositórios de patrimônio cultural, aqui denominados espaços de memória; e 4) Verificar se a complexidade do modelo semântico é um empecilho para os curadores de acervos digitais de patrimônio cultural.

### **8.2 Os espaços de memória da Secretaria de Estado de Cultura-GO**

Para alcançar o primeiro objetivo proposto nesta tese, buscou-se especificar as características dos espaços de memória da Secretaria de Estado de Cultura-GO e sua utilização quanto aos sistemas de informação. Faz-se necessário, para tanto, iniciar a discussão em torno da memória, que aqui, tem em si a intenção de ir além do ato simples de armazenar objetos digitais.

O Estado de Goiás possui uma riqueza multicultural que mistura saberes, práticas, costumes e outros aspectos que foram e são influenciados por diferentes povos, desde os nativos que aqui habitavam no momento de sua colonização. O patrimônio cultural, em especial o material, relacionado ao povo goiano pode ser encontrado em diversos lugares tais como Galerias, Arquivos, Bibliotecas e Museus que conservam, de alguma forma, importantes acervos do patrimônio cultural.

Dito isso, percebe-se que um dos desafios impostos pela própria essência dos centros de memória, reside em garantir acesso à população de forma interativa e com um certo grau de interconectividade com outras fontes de conhecimento disponíveis. Para que os curadores desses espaços possam fazê-lo digitalmente, de acordo com o estado da arte dos Sistemas de Informação para o Patrimônio Cultural, é preciso atender a uma série

de premissas: que haja uma estrutura mínima tecnológica; que os acervos estejam armazenados de forma organizada; que a estratégia seja clara; e que os dados sejam acessíveis para outros sistemas, como os mecanismos de busca ou mesmo outras tecnologias, em especial aquelas relacionadas aos dados abertos.

Como apresentado no CAPÍTULO IV, para que isso seja realizado, se faz necessário levar em consideração os padrões estabelecidos pelas comunidades e as tecnologias experimentadas para esse fim. Tecnologias estas que tem se desenvolvido a partir de um grande conjunto de estratégias, desde o foco mais básico na gestão e organização do patrimônio, até os modelos mais avançados relacionados à engenharia do conhecimento que passam, inevitavelmente, pela Web Semântica.

Doerr (Doerr, 2009) afirma que há uma pressão política imposta às organizações para estas disponibilizem seus acervos em formato digital. Tais acervos são advindos da representação digital de objetos físicos produzidos em grande escala e em disciplinas altamente especializadas.

Os resultados do inquérito e entrevista aplicados aos curadores dos espaços permitiram uma visualização mais efetiva de uma evidente precariedade digital, e até mesmo, informacional, dos acervos e dos espaços em si. No geral, os curadores trouxeram informações genéricas e também detalhes físicos, organizacionais e tecnológicos sobre os espaços, reforçando, em essência, a importância das informações para a preservação da memória, mas com a sensação que, em relação aos Sistemas de Informação, pouco foi realizado.

Quanto aos espaços deste estudo, tem-se o Museu da Imagem e do Som de Goiás (MIS), que é um museu localizado na praça cívica de Goiânia no centro de Goiânia e que abriga documentos sonoros, de cera, rolos de áudio, fotográficos, negativos, bases plásticas, positivos ampliação em papel, videográficos com filmes de 16mm, 32mm, que Abriga também materiais como cartas e diários e possui uma biblioteca especializada em imagem, som e música com seu acervo parcialmente digitalizado com algo em torno 45 mil itens entre CDs, fitas cassetes e discos.

Tem-se a Biblioteca Pio Vargas, um grande espaço com dois andares e um térreo onde fica o acervo geral com literatura estrangeira, literatura nacional, livros didáticos, enciclopédias e outros. No andar superior, é abrigado o acervo goiano onde registra, resgata e guarda o patrimônio da literatura Goiânia com mais de 30 mil livros.

Segundo a curadora do espaço, trata-se de um orgulho um órgão público abrigar este acervo, mas ao mesmo tempo causa tristeza o fato de o povo Goiano não ter conhecimento dele - *Um problema de divulgação, inclusive de sinalizações físicas visuais*, completa ela.

A despeito das sensações e frustrações da curadora, a impressão que se tem é que não há mesmo como a população ter conhecimento do espaço e muito menos acesso ao acervo. O *Digital* tem muito a contribuir nesses espaços para além da divulgação, no acesso.

Tem-se o mesmo sentimento a respeito da Gibiteca Jorge Braga, que possui revistas, livros e gibis históricos em quadrinhos do início do século passado, constante inclusive de obras raras. O acervo é bem dividido nas classes CDU de forma física, mas não existe um sistema de informação que acompanha essa classificação.

**Figura 85 – Gibiteca Jorge Braga**



Fonte: Imagem feita pelo Autor com autorização da Curadora do Espaço.

No Arquivo Histórico Estadual, o espaço é para organizar, abrigar e preservar a informação histórica produzida pela administração pública do estado desde a província. O acervo é composto por documentos históricos permanentes manuscritos, datilografados e cópias do período colonial, imperial e republicano.

Existe uma documentação avulsa de 1731 a 1975. Documentação do ouro, de

escravos, sesmarias e outros documentos diversificados deste período. Há ainda uma documentação manuscrita no período de 1724 a 1932 e uma hemeroteca de jornais, diários e correios oficiais de 1835 a 2011.

Outro acervo riquíssimo do Arquivo Histórico, é a documentação de praticamente todos os municípios goianos e do estado do Tocantins. Além de uma biblioteca com mais de 2000 livros históricos e de arquivologia geral do Brasil. Há ainda teses, dissertações e outros trabalho de pesquisadores que utilizam o arquivo, quase sempre sobre Goiás e Tocantins.

A Figura 86 apresenta uma ilustração criada com imagens do acervo do Arquivo Histórico de Goiás durante as atividades de pesquisa.

**Figura 86 – Imagens do Acervo do Arquivo Histórico do Estado de Goiás**



Fonte: imagem criada pelo Autor com autorização da Curadora do Acervo.

Já o Museu Pedro Ludovico, é um museu também localizado em Goiânia onde viveu Pedro Ludovico Teixeira, que foi o fundador da cidade. Ele quem finalmente executou o plano de transferência da Capital dos anos 30 para os anos 40. Foi interventor e depois governador do estado. Logo após seu falecimento, em 1976, um decreto estadual determinou que fosse criado um centro de memória homenageando Pedro Ludovico e a capital, pela grande importância política deste ator político e do grande feito da criação

de uma nova capital. Os atos de criação deste espaço remetem textualmente ao termo “centro de memória”.

Trata-se de um museu para preservar a memória de Goiânia e do fundador. “Mas não é um memorial exclusivo de Pedro Ludovico”, conforme palavras do Curador. Este defende que apesar de haver um recorte político e, de se tratar da antiga residência da família de Pedro Ludovico – que, portanto, possui, dentre outros, os seus objetos pessoais – é um museu da Cidade de Goiânia, que segundo o profissional, trata-se de um título que precisa ser fixado como “identidade”.

**Figura 87 – Objetos do Acervo do Museu Pedro Ludovico**



Fonte: Ilustração feita com fotografias retiradas pelo Autor nas atividades de pesquisa com Autorização do Curador do Espaço.

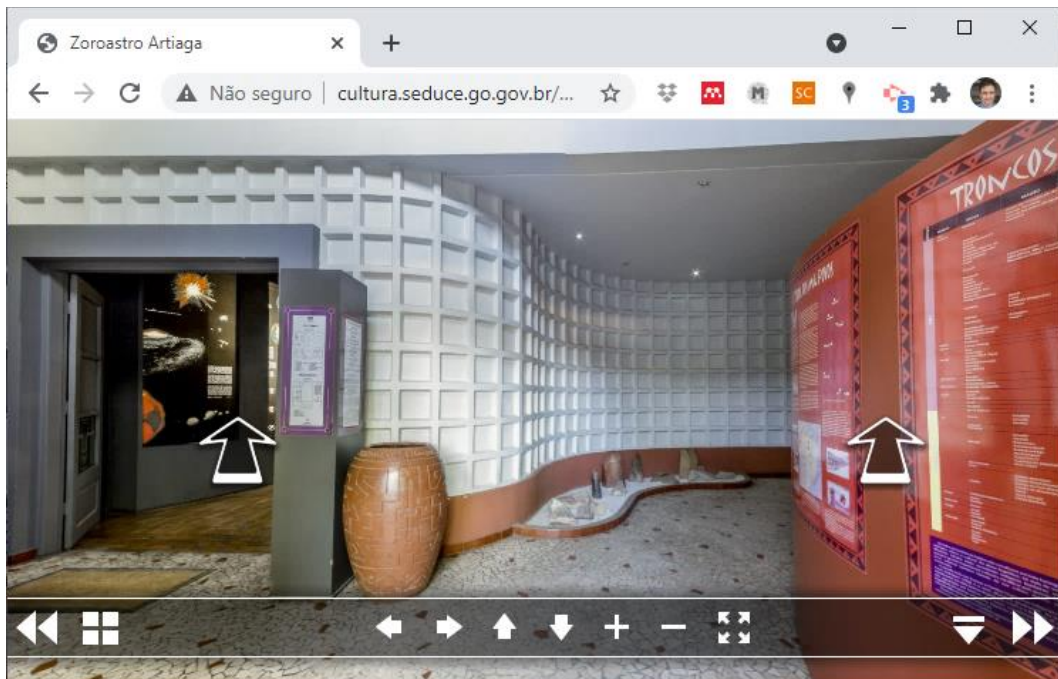
O Museu Goiano Professor Zoroastro Artiaga, ou MUZA, é um museu público estadual fundado em 1946, leva o nome de seu primeiro diretor. Um geólogo e grande estudioso multidisciplinar da época. O museu nasce a partir de um acervo cujas primeiras coleções foram coletadas por uma comissão para divulgar as riquezas e potencialidades do estado de Goiás. Em um tempo onde havia algumas exposições de produtos regionais, nacionais e internacionais, uma edição, no Rio de Janeiro, recebeu esses objetos da riqueza natural e cultural do estado. Terminada essa exposição, na ocasião do batismo cultural de Goiânia, o acervo foi organizado e criado o espaço para abrigar esse material com o intuito de ser um museu da história e cultura de Goiás.

O Museu Ferroviário de Pires do Rio é um museu destinado à estrada de ferro e materiais referentes à história da ferrovia e da cidade de Pires do Rio-GO. Já o Palácio Conde dos Arcos, é um museu onde era casa do Governador do estado na antiga capital. Até hoje os governadores se instalam no espaço durante o ato simbólico de transferência da capital, que ocorre uma vez por ano na Cidade de Goiás, antiga Vila-Boa. Esse espaço foi aberto à visitação pública na década 70 e é composto por mobiliários e outras peças das famílias dos governadores que lá residiram.

Sobre os sites institucionais na Web, os dados mostraram que há sites institucionais de forma estática em todos os espaços, porém, com poucas informações efetivas sobre os acervos, mas sendo possível, no caso específico dos museus, visitas virtuais (tour virtual).

As técnicas de museus inteligentes (*smart museum*) são exploradas em diversos trabalhos como Korzun *et al.* (2017), Smirnov *et al.* (2017) e Varfolomeyev *et al.* (2015). Estas técnicas, normalmente relacionadas a sistemas de recomendação com base na experiência do usuário e coleções dos museus, não são encontrados na visita virtual dos museus, conforme os curadores. O *tour virtual* destes espaços trata-se de uma ferramenta estática que permite uma visualização em 360° de alguns locais do espaço.

**Figura 88 – Imagem do Tour Virtual de Museus – SECULT-GO**



Fonte: <http://www.seduca.go.gov.br/museuvirtual/zoroastro.html>

A Figura 88 acima apresenta uma visão do *tour* virtual presente no site institucional da Secretaria de Estado da Cultura (SECULT-GO).

As informações dos objetos, bem como sua classificação e contexto, não são levadas em consideração na ferramenta, permitindo que o usuário possa visualizar as galerias e os objetos no espaço virtual, mas com uma experiência frágil de interação.

Em Martini *et al.* (2017) há uma proposta de sistema de arquivos históricos com base no CIDOC-CRM e a criação automática de visualizações ao que os autores nomeiam “Salas virtuais”. A classificação utilizando as classes e propriedades do CIDOC-CRM pode facilitar a integração com museus e outros espaços.

Quanto aos questionamentos em relação à composição dos acervos, pretendeu-se captar informações sobre aspectos físicos os tipos de objetos presentes, e permitiu-se conhecer características peculiares dos objetos.

Nos acervos arquivísticos com documentos históricos diversos, em especial relacionados ao governo do estado e da província, que são anteriores à república, as classes principais encontradas foram Documentos Históricos, Cartas, Manuscritos, Jornais, Mapas e Notícias.

Do acervo Documental, que compõe o arquivo geral, observa-se a existência de três áreas físicas de Arquivos setoriais de acordo com a natureza histórica dos documentos: Documentação Avulsa (1731-1973), alusiva a documentos históricos da Tesouraria Provincial, tais como: folhas de pagamento, ofícios, balancetes, atos, decretos, relatórios, regulamentos e leis e documentação dos Municípios. Manuscrita (1724 a 1973), dentre os quais, registros gerais de correspondências, requerimento de sesmarias, provisões, patentes, leis e decretos. Assim como registros e minutas da correspondência dos governantes da capitania e da província com o reino e autoridades provinciais e registros das Secretaria de Polícia e obras públicas.

O acervo arquivístico possui uma hemeroteca com um vasto conjunto de objetos da imprensa goiana com jornais como o Correio Oficial, Diário oficial do Estado, o Tocantins, Província de Goiás, a Imprensa, Voz do Povo, o Democrata, o Social, Folha de Goiás, Diário da manhã e o Popular.

No contexto da proposta apresentada nesta tese, os objetos arquivísticos podem ser classificados de acordo com sua própria organização interna, podendo fazer referência a qualquer tipo de organização externa através da adição de propriedades em cada objeto.

Uma questão peculiar observada no Arquivo Histórico é uma forma de catalogação e recuperação da informação (quase totalmente manual) onde a dimensão mais visualizada é o Tempo (período de tempo, ano, década).

Já o acervo bibliotecário, composto de livros literários, gibis e obras raras, está presente em uma biblioteca, uma gibiteca; e em acervos específicos de livros em três museus e no arquivo histórico.

As bibliotecas costumam atender a uma perspectiva de organização internacional, já os acervos bibliotecários presentes em outros espaços não seguem nenhuma padronização.

A proposta apresentada nesta tese não visa a operacionalização de empréstimos ou outras atividades presentes em bibliotecas, mas permite que, uma vez criado um acervo, os objetos possam ser relacionados e possuir descrições com grandes modelos de dados já estabelecidos.

Em Agosti *et al.* (2015) é possível encontrar uma proposta de integração de bibliotecas em uma abordagem *bottom-up*, dos modelos DELOS e o 5S (*Streams, Structures, Spaces, Scenarios, Societies*), fornecendo uma solução ontológica para integração entre os modelos.

Já o acervo museológico é bastante diversificado. No domínio arqueológico foram identificados objetos fósseis e líticos de pedra lascada ou polida. Outras classes identificadas foram etnografia (plumária, cestaria, cerâmicas, armas e utilitários); geológico e mineral; arte sacra; fauna do cerrado (taxidermia); arte e cultura popular (teares, rodas de fiar, cerâmicas e têxtil), mobiliário, cerâmicas, pinturas, prataria, numismática, indumentária, automobilístico, botânico, arquitetônico e ferroviário. Há ainda um acervo de imagem e som compostos objetos fonográficos, videográficos e fotográfico.

Os questionamentos sobre inventários e ações de organização buscaram verificar existência da prática de inventários, que conforme Doerr (2009), é uma das funcionalidades importantes na gestão do acervo. Além disso, pretendeu-se buscar aspectos da organização dos acervos em relação a esses inventários.

Percebeu-se, que, no geral, 6 dos 8 espaços realizam inventários e mantém os catálogos resultantes em planilhas, não havendo nenhum processo informatizado mais elaborado para este fim. A organização informacional dos espaços facilita os inventários

e vice-versa. Padilha & Café (2016) mostram que o cruzamento de inventários é especialmente útil para identificar inconsistências e diferenças terminológicas dos objetos.

Seguindo nos questionamentos sobre classes e objetos do acervo foi possível conhecer a forma dos objetos e obter uma visão inicial das classes, termos e relações.

No domínio dos arquivos, as principais classes identificadas foram: documentos oficiais, manuscritos, jornais, certidões, mapas, cartas, leis e decretos, com ênfase a períodos e pessoas. Martini *et al.* (2017) fornecem uma estratégia ontológica para arquivos e criação de exposições com base no CIDOC-CRM para metadados de baixo e alto nível.

No domínio museológico percebeu-se a presença de documentos histórico e urbanísticos, mobiliários, equipamentos domésticos, pinturas, objetos decorativos, objetos comemorativos, objetos religiosos, objetos pessoais, objetos de comunicação.

Owais *et al.* (2017) implementam uma arquitetura de mapeamento e enriquecimento semântico no domínio dos museus voltados para os dados abertos conectados com base no CIDOC-CRM e utilizando a DBPedia para enriquecimento. Frommholz *et al.* (2014), apresentam um sistema de informação para contar a história a partir do relato de pessoas utilizando classes e propriedades do CIDOC-CRM.

A proposta apresentada nesta tese visa colaborar com os modelos e iniciativas já existentes, em especial, as ontologias e estratégias já consolidadas, como CIDOC-CRM, EDM, *Schema.org*, *Dublin Core*, e outros.

No domínio arqueológico, várias classes e termos foram identificados, como fósseis, pinturas, líticos, plumária, cestaria, cerâmicas, armas, utilitários; geológico e mineral; fauna do cerrado (taxidermia). No domínio da Imagem e Som foram encontrados músicas, vídeos, documentários, discos, filmes, festivais, cinema, tv, intérpretes, fotos, negativos, fitas magnéticas, DVD, VHS e outros artefatos. No domínio bibliotecário foram encontradas classes como impressos, livros, revistas, religião, folclore, almanaque, artes, enciclopédia, guias, autores. Já em um domínio multidisciplinar foram encontradas classes como arte e cultura popular, história, religião e personalidades.

### 8.3 Modelos conceituais de representação da informação de patrimônio cultural e tecnologias da Web-Semântica

Para atender o segundo objetivo desta tese, que foi realizar um estudo da aplicabilidade dos modelos conceituais de representação da informação de patrimônio cultural e tecnologias da Web-Semântica, se faz necessário retomar rapidamente o que já foi apresentado no enquadramento teórico e estado da arte. Os principais modelos de representação utilizados para patrimônio cultural são o CIDOC-CRM e EDM. Ambos os modelos podem ser aplicados em parte, tanto no que diz respeito as classes como nas propriedades presentes em cada um desses modelos.

Para o funcionamento geral do protótipo apresentado, entretanto, é preciso levar em consideração algumas questões fundamentais. A primeira delas, simplificar para os curadores as tarefas relacionadas à manutenção da classificação do acervo e dos dados dos objetos digitais de cada um destes acervos. A segunda, cada espaço de memória possui uma organização própria constante de objetos e domínio próprios.

Modelos mais simplificados do próprio RDF-S e *Dublin Core* foram utilizados para definir duas estruturas básicas iniciais e uma estrutura dimensional. Uma ontologia inicial para os acervos, aqui denominada Classes (*classes.owl*, presente em <http://200.137.241.247:8080/fuseki/classes#>), que permite a classificação dos acervos com quatro propriedades básicas: *rdfs:Class*, *rdfs:label*, *rdfs:comment* e *rdfs:subClassOf*. Dessa forma, foram utilizadas como classes básicas aquelas advindas do *schema* RDF (<http://www.w3.org/2000/01/rdf-schema#>).

A outra estrutura inicial, denominada Objetos (*objetos.owl*, presente em <http://200.137.241.247:8080/fuseki/objetos#>) apresenta uma classificação mínima para os objetos do acervo constando das propriedades *owl:Class*, *rdf:type*, *dc:title*, *dc:subject*, *dc:description*; e mais duas propriedades dimensionais: *dim:TipoFisico*, *dim:tipoDimensional*. O padrão *Dublin Core* foi utilizado como base para a descrição dos objetos, bem como a tipificação por meio da propriedade *type* do RDF/S.

Uma terceira ontologia foi desenvolvida (*dimensoes.owl*, presente em <http://200.137.241.247:8080/fuseki/objetos#dimensoes>) para que a interoperabilidade semântica possa ocorrer nas quatro dimensões propostas pelo modelo e classificar os objetos em tipos físicos, aqueles presentes nos espaços fisicamente, e os dimensionais, objetos não físicos que se referem a uma das dimensões.

As classes definidas são armazenadas no TDB Fuseki (RDF Store) específico, contendo a descrição de cada classe de acordo com essas quatro propriedades. A funcionalidade descrita na Seção 6.4.1 (p. 178) **Definir Acervo** apresentou com detalhes as operações envolvidas nesse processo.

Essa estrutura criada para os acervos, permite classificar cada objeto digital de acordo com cada classe estabelecida ontologia inicial através da propriedade *rdfs:type*. Já a organização dos objetos respeita a uma ontologia de definição objetos conforme o esquema apresentado na Seção 4.7.2 (p. 138).

Uma regra básica para a criação de um novo Objeto Digital é que todos os dados da definição inicial de objetos estejam completos. As quatro propriedades básicas (mínimas) a serem informadas são *rdf:type*, *dc:title*, *dc:subject* e *dc:description*. As propriedades *dim:tipoFisico* e *dim:TipoDimensional* são propriedades de objeto (*owl:objectProperty*) e é quem vão determinar a natureza deste. O sistema cria automaticamente um *id*, que é o identificar único do recurso que o objeto representa.

A ontologia das dimensões (Seção 4.7.3 p. 141) apresenta as quatro classes dimensionais: Pessoa, Fenômeno, Espaço e Tempo. Apresenta também cinco classes físicas relacionadas ao tipo do acervo: Arqueológico, Bibliotecário, Documental, Imagem e Som, e Museológico; além das classes ObjetoFisico e ObjetoDimensional para combinar com as propriedades dos tipos dimensionais e físicos, e as propriedades de objeto (*owl:objectProperty*): *relacionaPessoa*, *relacionaFenômeno*, *relacionaEspaço* e *relacionaTempo*, respectivamente associadas às dimensões Pessoa, Fenômeno, Espaço e Tempo.

No que se refere às tecnologias da Web-Semântica e dos dados abertos conectados, na Seção 2.4 do CAPÍTULO II (p. 51) foram apresentadas características gerais e detalhes dos principais componentes dessas tecnologias.

Um passo para que os dados abertos conectados sejam efetivos é permitir que estes sejam acessados por meio de um componente público em linguagem acessível e computável. Conforme a proposta aqui apresentada, os elementos essenciais para que um Sistema de Informação atenda às especificações da Web Semântica, envolvem, construir ontologias com as classes e propriedades referentes ao domínio; utilizar RDF/RDF-S e suas extensões; e, utilizar um armazenamento em tripla para os RDFs gerados. Na abordagem deste modelo, isso é possível através de consultas SPARQL em um *endpoint*

que responde a requisições HTTP em um servidor TDB *Apache Fuseki*.

Na Seção 4.9 (p. 145) foi fornecida uma proposta de arcabouço tecnológico com uma descrição sobre o *Jena* e o *Fuseki* como *framework* e sistema gerenciador de bancos de dados em tripla. O *Jena* é estruturado em componentes para modelagem, manipulação, transformação e armazenamento de grafos nas linguagens da Web Semântica e o *Fuseki* é um TDB de fácil manipulação. As classes de conexão fornecidas pelo *framework* tornam a atividade transformação e persistência simplificadas para os desenvolvedores. Não foram realizados, entretanto, testes à exaustão no que diz respeito à latência ou mesmo armazenamento em grande escala.

#### **8.4 Um Modelo de Sistema de Informação Semântico**

O terceiro objetivo a ser alcançado por esta tese, foi construir um modelo semântico para organização da informação de patrimônio cultural, a partir de padrões técnicos, que permita a interoperabilidade entre repositórios de patrimônio cultural, aqui denominados espaços de memória. No CAPÍTULO IV, esse modelo foi estabelecido por meio de uma proposta de acordo como estado da arte e a realidade dos espaços de memória pesquisados.

Há de acompanhar este modelo os pressupostos de Doerr (2009) que afirma que os Sistema de Informação do Patrimônio Cultural, normalmente, apresentam ao menos uma dessas quatro funcionalidades: Gestão de acervos; Conservação; Pesquisa; e Apresentação. O recorte para a construção do protótipo foi realizado em torno da primeira. A gestão de acervo, segundo (Doerr, 2009, p. 3), relaciona-se às atividades de “aquisição (e digitalização), registro, abandono, inventário, concessões (saídas), exposições, seguros, direito autorais e zonas de proteção”. Em um recorte ainda mais específico, apenas as atividades de digitalização e registro foram cobertas pela proposta.

O modelo foi proposto na Seção 4.4 (p. 126) e se apresenta como um arcabouço composto de 4 partes: 1) Sistema Web; 2) Camada de Armazenamento TDB; 3) Camada Semântica; e 4) Camada de Interoperabilidade. Esta última com o objetivo de operar nas quatro dimensões e permitir intercâmbio entre os repositórios, fazendo-se valer das classes e propriedades definidas na Camada Semântica.

A operacionalização do protótipo, dentre outros objetivos, pretendeu mostrar que este modelo genérico é capaz de suportar as características diversificadas dos centros de memória, com a premissa base que cada espaço tenha seu próprio repositório e suas

atividades operacionais efetivadas por um Curador digital responsável pela gestão do acervo.

Pôde-se observar, nos exemplos do CAPÍTULO VII (Seção 7.3 p. 224), que repositórios de GLAMs com uma diversidade de classes de acervo e objetos foram manipulados sem grandes dificuldades.

Na Seção 4.5 (p. 130) foi apresentada a proposta de um Sistema Web composto por um conjunto de páginas Web, um conjunto de objetos controladores e de objetos instanciados a partir de classes definidas na transformação RDF/Objetos. As atividades de gestão do acervo através de ações acionadas pelos eventos que dispararão requisições para o servidor

A Seção 4.6 (p. 132) apresentou uma proposta de camada de armazenamento com uma estrutura de banco de dados em tripla (TDB) para receber, persistir e recuperar os grafos RDF resultantes as relações entre objetos e classes. Essa camada atende requisições HTTP e chamadas SPARQL que são transformadas de acordo com cada evento, objeto ou classe do acervo que está sendo manipulado. O resultado das manipulações geradas no RDF é serializado em XML ou JSON.

Na página 133, especificamente na Seção 4.7, foi apresentada uma proposta de camada semântica com um modelo conceitual em uma visão com cinco eixos e quatro núcleos, para que o conhecimento produzido transite nas quatro dimensões (Pessoa, Fenômeno, Espaço e Tempo), assumindo em forma de objetos do Arqueológico, do Arquivístico/Documental, do Bibliotecário, do Imagético/Sonoro ou do Museológico. Este modelo com três ontologias de domínio simplificadas foi apresentado de detalhadamente nas Seções 4.7.1. (p. 135), 4.7.2 (p. 138) e 4.7.3 (p. 141).

A despeito das limitações do modelo, que é genérico e com poucas restrições, a informação é centrada no objeto que pode transitar nas quatro dimensões apresentadas, bem como assumir múltiplos eixos dentro dos GLAMs (podendo ser documental e museológico, por exemplo). Estes objetos foram divididos em Objetos Dimensionais e Objetos Físicos. Os primeiros dizem respeito às quatro dimensões e podem descrever um conjunto informacional que não acompanha nenhum objeto físico. Por exemplo, a descrição de uma pessoa, um recorte de tempo ou um espaço. Já os Objetos de Físicos são de fato os objetos digitais materializados presentes nos acervos, como uma arma antiga ou mapa histórico.

Dentro das características dos objetos digitais presentes nos repositórios, se faz necessário reafirmar que há um caráter de estado momentâneo individual. Ou seja, um objeto não pode ser completamente fechado em si como um bloco monolítico de informações. As propriedades e relacionamentos precisam ser vivas e em cada momento pode surgir um novo estado a partir de um novo conjunto de relações, que são manipuladas ou decididas pelos curadores dos acervos, após uma interação com o mecanismo de interoperabilidade, ou mesmo por sua ação individual na Web.

A Camada de Interoperabilidade do modelo proposto foi apresentada na Seção 4.8 (p. 143). Desde os primeiros passos deste trabalho, essa pesquisa busca criar estruturas tecnológicas para permitir transferência de informação semântica por meio Serviços Web como *endpoints* SPARQL, *middlewares* e outras possibilidades de colaboração externa onde os espaços possam, de forma mútua, adquirir uma capacidade de encontrar e conectar informações de objetos de um conjunto de repositórios conhecidos, aqui chamados “*repositórios amigos*”. Defende-se que este modelo, nas quatro dimensões apresentadas, permite um aumento nas ligações entre os objetos de cada acervo.

Tanto a funcionalidade de adicionar novas propriedades, quanto a coleta de ligações por meio do Motor do Interoperabilidade, podem contribuir nesse sentido. A simplificação do modelo em alto nível permite que as características diversificadas dos espaços de memória sejam respeitadas, independentemente de sua origem ou das características dos objetos dos acervos.

Assim, cada espaço pode utilizar sua própria organização para indexação dos itens do acervo e possuir seu repositório de forma independente, mas com condições de se conectar a outros repositórios e realizar intercâmbio de informações.

A camada semântica do modelo é composta de uma arquitetura genérica com três classes ontológicas de domínio para o patrimônio cultural. A primeira fornece uma descrição e classificação inicial para o acervo, a segunda para os objetos e terceira para a classificação em quatro dimensões (Pessoas, Fenômeno, Espaço e Tempo).

A modelagem conceitual é vista como ponto crucial dos trabalhos encontrados no estado da arte, uma vez que partir dos conceitos que se dá o processo organização, fundamental para recuperar a informação. Uma boa modelagem conceitual, portanto, pode ajudar em melhor indexação e usabilidade dos recursos. Assim, defende-se que uma integração de repositórios de patrimônio cultural requer uma combinação de múltiplas

fontes de dados (Buranasing *et al.*, 2016), fazendo-se, portanto, necessário uma concepção em alto nível que possa abrigar as classes e propriedades possíveis nos repositórios.

### 8.5 Protótipo do Sistema de Informação Guará

De forma combinada, o terceiro e o quarto objetivo declarados nessa tese dizem respeito ao que foi apresentado no CAPÍTULO VI. A construção de um protótipo de Sistema de Informação Semântico através do modelo apresentado no CAPÍTULO IV. Este protótipo de Sistema de Informação semântico – denominado Guará – pode ser acessado no endereço <http://200.137.241.247:8080/guara>.

Niang *et al.* (2017), na mesma direção, propôs um Sistema de Informação semântico para o Patrimônio Cultural, constando de uma única arquitetura de ontologia com o contexto da conservação-restauração a qual os atores denominaram “ontologia global” (Niang *et al.*, 2017). A proposta apresentada na Seção 4.4 do CAPÍTULO IV (p. 126) e desenvolvida em forma de protótipo no CAPÍTULO VII, segue estratégia parecida. Entretanto, ao invés de restringir ao acesso único ou a uma só entrada, aponta para os múltiplos acessos e múltiplos repositórios, o que torna a tarefa de interoperabilidade potencialmente mais complexa. Nesse ponto em específico, o modelo aqui proposto avança em permitir o intercâmbio entre os múltiplos repositórios através de uma camada de interoperabilidade.

De forma complementar à proposta de Aalberg *et al.* (2015), o modelo usa artefatos da Engenharia de Software em uma visão do geral para o detalhe. Nesta abordagem, os autores propõem uso de modelos e artefatos da Engenharia de Software e defendem que um modelo adequado deve consistir em uma estrutura de domínio (nomes), cenários de caso de uso com uma descrição da intenção, um fragmento de modelo gráfico, exemplos de implementação e mapeamentos. A proposta descrita nos Capítulos IV e VI estes artefatos da Engenharia de Software com base nas premissas da pesquisa-ação e avançou na construção de algoritmos e operacionalização do Sistema de Informação por meio do protótipo.

Na Seção 6.4 (p. 175) foi apresentada uma modelagem do protótipo, em termos de artefatos da engenharia de software, a qual identificou no recorte seis requisitos que foram descritos em uma abordagem do geral para o detalhe. *R1 – Definir Acervos; R2 – Gerenciar Objetos Físicos; R3 – Gerenciar Objetos Digitais; R4 – Adicionar/remover*

*propriedades; R5 – Coletar ligações; e R6 – Administrar repositório.*

Cada uma das camadas do modelo proposto no CAPÍTULO IV foi detalhada dentro desses requisitos apresentados, em uma arquitetura utilizada para desenvolvimento deste conforme Seção 7.2 do CAPÍTULO VII (p. 209). A divisão em camadas de modelo, visão e controle permitiu um tratamento dos serviços apresentados no modelo de forma a transitar as informações a partir dos eventos disparados entre as camadas, em especial no que se refere às ações do Curador.

A construção de arcabouço em camadas também foi proposta por Ivanović & Houssos (2014) em um modelo em composto por *Converters, Repositories, e Services* contendo uma API para fornecer serviços web para conversão e mapeamento de entidades CERIF (*common European Research Information Format*). Ao invés de oferecer uma API para transformação, a proposta aqui apresentada fornece um *endpoint* para consultas SPARQL e métodos de serviços por meio de requisições HTTP de forma aberta.

A camada de Páginas Web (Seção 7.2.1 p. 211) foi ilustrada dentro da estratégia *modelo-visão-controle* (MVC) e da especificação do Java Server Faces (JSF) que foi utilizada para a maior parte das interações. A partir dele foi possível gerar páginas Web de acordo com os eventos que são tratados por controladores e transitados da apresentação até o TDB.

Como pôde ser visto, o Módulo de Entidades é acionado como um Motor de Objetos em Tripla (MOT) com classes e métodos para geração das instâncias do paradigma de Orientação a Objetos para o modelo triplas RDF e também o caminho inverso.

Foram apresentadas quatro classes para esse módulo: PIC (*Plain Initial Class*), ou classe inicial plana, que apresenta as definições básicas das classes do acervo conforme a Ontologia inicial de classes; PIDO (*Plain Initial Data Object*) ou Objeto de Dados Inicial Plano da mesma forma das PIC trouxe uma definição para objetos; PPV (*Plain Property Value*) ou Propriedade-Valor Planos; e TriplaSimples que permitem a instanciação das triplas dos classificadores dos recursos e dos próprios recursos, que são os objetos dos acervos.

O Módulo de Serviços foi apresentado na Página Seção 7.2.3 (p. 214) com ênfase ao Gerenciador de Eventos. Como foi possível visualizar na Seção 7.2.3.1, é o Gerenciador de eventos quem tem a responsabilidade de manter o diálogo entre os eventos

disparados pelos curadores e os serviços executados pelas camadas mais profundas.

Foi possível ver ainda, que o acesso aos TDBs também é realizado por meio de serviços em uma subcamada específica DAO. Os objetos DAO interagem com os objetos das entidades (modelo) e transformam objetos instanciados em SPARQL para que sejam persistidos nos bancos de dados.

Na Seção 7.2.3.4 (p. 218), pôde ser visto que a interoperabilidade é possível através de uma estratégia de manter um conjunto de repositórios com *endpoints* conectáveis denominados “*repositórios amigos*” e realizar intercâmbio através das propriedades dos Objetos Físicos e Dimensionais presentes nos acervos. É possível também que o Curador possa gerir as propriedades dos objetos de forma manual.

O algoritmo apresentado na Seção 4.8 (p. 143) foi codificado, dentro do protótipo, em métodos das classes do Motor de Interoperabilidade, que envolvem o envio de chamadas SPARQL. O algoritmo lista todos os repositórios da lista de “*repositórios amigos*” e pesquisa relações potenciais para o objeto, primeiro com base nas palavras-chave e num segundo momento, com base nas relações do objeto alvo.

O Motor de Interoperabilidade (MI) dispara os algoritmos para as consultas SPARQL relacionadas às dimensões e ao objeto ou propriedade, e recupera os grafos correspondentes.

A despeito dos problemas significativos existentes nas consultados por palavra-chave, as questões relacionadas ao enriquecimento semântico por sinônimos, buscas por proximidade e outras técnicas, não são levadas em consideração nesse momento, mas compreende-se a necessidade de utilizar tais técnicas para este tipo de busca.

Levando em consideração as dimensões definidas na relação  $\langle x \rangle$  *dim:tipoDimensional*  $\langle y \rangle$ , são enviadas consultas SPARQL para os Objetos Dimensionais e Físicos  $x$  e  $y$  em cada um dos repositórios, primeiro pelo conjunto de palavras chave do objeto. Em seguida são retornados os Objetos Dimensionais que tenham alguma relação direta com o  $\langle id \rangle$  do Objeto alvo de acordo com cada uma das relações dimensionais pelas relações  $\langle x \rangle$  *dim:[relacionaPessoa| relacionaFenômeno|relacionaEspaço| relacionaTempo]*  $\langle y \rangle$ .

Em estratégia similar à encontrada em Abu Musa *et al.* (2020), o algoritmo busca realizar inferências sobre classificações de objetos culturais por meio de duas abordagens. A primeira, por correspondência textual (palavras-chave), que permite associação direta

entre o objeto e os conceitos da ontologia (Abu Musa *et al.*, 2020). A segunda abordagem proposta por estes autores busca correspondência semântica através de um banco de dados léxico para ampliar os sinônimos de propriedades de um determinado objeto. A abordagem proposta nesta tese, no entanto, é limitada nesse quesito, uma vez que não utiliza dicionários sinônimos (como o *WordNet*) e faz inferência direta de acordo com as relações do objeto alvo.

Foi possível visualizar na Seção 7.2.4 (p. 221), que o protótipo possui dois módulos externos. O primeiro de armazenamento TDB composto pelo *Fuseki* que está instalado no endereço <http://200.137.241.247:8080/fuseki>. A Seção 4.9 (p. 145) do CAPÍTULO IV apresentaram os detalhes da instalação do *Fuseki* como *Web-App* do *Apache Tomcat*.

A Seção 7.3 (p. 224) apresentou experimentos do protótipo com a simulação de ações de Curadores e repositórios em 5 cenários. Para os cenários foram apresentados 3 repositórios hipotéticos (AHG, DIO e MPL) associados a 3 curadores dos espaços, respectivamente.

No Cenário 1 o Curador 1 adicionou uma primeira classe no repositório AHG conforme Seção 7.3.1 (p. 226).

Foi possível verificar, ao final da interação que o resultado foi alcançado e a classe “Documentos” foi persistida no TDB “*ahgclass*” com um novo grafo referente às triplas “:Documento a owl:Class ; rdfs:comment “Classe para abrigar os documentos do acervo do Arquivo.” ; rdfs:label “Documentos” ; rdfs:subClassOf :Acervo .”.

No Cenário 2 Curador 2 (aqui chamado DIO) adicionou um Objeto Físico conforme Seção 7.3.2 (p. 230). Ao final da operação percebeu que o objeto foi criado um novo grafo no TDB de objetos (“*dioceseobj*”). O Objeto Físico *f68b7214-5552-4f71-80e3-fbe00cc38499* que se refere a um registro de batismo de “Pedro Ludovico Teixeira”.

No Cenário 3 o Curador 3, ainda na funcionalidade Gerir Acervos Digitais, adicionou um Objeto Dimensional conforme Seção 7.3.3 (p. 235). O objeto *5cf7b91d-dcc0-48ec-9428-84fff2b1f321* persistido nesse cenário refere-se à Pessoa “Pedro Ludovico Teixeira”.

No Cenário 4, conforme Seção 7.3.4 (p. 239), Curador 2 adicionou uma propriedade ao objeto *f68b7214-5552-4f71-80e3-fbe00cc38499* que se encontrava no

TDB “*Dioceseclass*”. Foi possível observar que se deu a criação de uma nova relação através da propriedade *dim:relacionaPessoa* entre este objeto (“Registro de Batismo de Pedro Ludovico”) e o objeto `<http://200.137.241.247:8080/fuseki/mplobj/5cf7b91d-dcc0-48ec-9428-84fff2b1f321>`, Pessoa “Pedro Ludovico de Almeida” que se encontra no repositório do espaço 2 (MPL). Essa associação se deu de forma direta nesse experimento, e foi possível verificar que é possível o Curador adicionar manualmente ligações com objetos de forma manual.

No Cenário 5, o sistema acionou o Motor de Interoperabilidade (MI) com a funcionalidade Coletar Ligações, conforme Seção 7.3.5 (p. 243). Foi possível verificar a persistência de um novo Objeto Físico “Livro” no TDB de objetos (“*mplobj*”) com o *id* `b3872c51-e75a-40b6-8f34-99dba1d5cf7e` e, posteriormente, a ação do Motor de Interoperabilidade (MI) que disparou as consultas SPARQL invocando o algoritmo de interoperabilidade, primeiro pelas palavras-chave depois pelas relações do objeto nos repositórios listados.

Ao final das interações, foi gerada uma lista de possíveis relações do objeto alvo que foi persistido (`b3872c51-e75a-40b6-8f34-99dba1d5cf7e`) com o objeto `f68b7214-5552-4f71-80e3-fbe00cc38499`, e `5cf7b91d-dcc0-48ec-9428-84fff2b1f321`, e as duas novas relações foram adicionadas ao objeto alvo. Foi possível verificar que Motor de Interoperabilidade (MI) permite o intercâmbio entre os “*repositórios amigos*” através do algoritmo de interoperabilidade, mas sem adicionar automaticamente as relações, deixando essa atividade para a confirmação do Curador.

## 8.6 O Ambiente semântico e a complexidade para os Curadores

A criação de um ambiente semântico é o que visa possibilitar a interoperação entre os repositórios e, conseqüentemente, uma integração entre os diversos acervos que compõem o patrimônio cultural, cujas informações são encontradas em formato digital nos espaços de memória.

Nas palavras de Carvalho Junior *et al.* (2019) “a demanda pela interoperabilidade dos acervos dos diferentes domínios arquivísticos é pautada pela possibilidade de integração dos conteúdos diversos via web” (Carvalho Junior *et al.*, 2019, p. 1), e isso é capaz de promover, ao mesmo tempo que exige, novas possibilidades de acesso e de processamento das informações de patrimônio cultural pela sociedade.

O grande ator nessa demanda, aqui denominado Curador Digital, ou

simplesmente Curador, assume papéis fundamentais na construção dessa interoperabilidade, complementando atividades daquele especialista, que conforme (Floridi, 2002) *apud* Carvalho Junior *et al.* (2019) trata-se de um “cuidador do ambiente semântico”.

O quarto objetivo, colocado em forma de questão, busca saber a complexidade existente no modelo é um empecilho para que os curadores mantenham um ambiente semântico. Não foi possível realizar testes específicos com os curadores no que se refere aos modelos de representação com o protótipo criado.

Aqui é preciso retomar, a despeito das questões semânticas envolvidas, que os repositórios de patrimônio cultural são sistemas de informação, que segundo, Doerr (2009) podem ser distinguidos, principalmente, pelas seguintes funções de Gestão, Conservação, Pesquisa, e Apresentação, e que para cobrir essas várias funções há um grande número de desafios que precisam ser transformados em processos sistematizados uma vez que os centros de memória possuem características próprias.

Especialmente o desafio de *suportar os vários modelos de organização da informação*, uma vez que há uma alta diversidade de dados de memória aliadas às facetas exigidas pelas funções, e as expectativas de utilização geram um cenário de alta complexidade.

Aalberg *et al.* (2015) afirmam que o compartilhamento e a disseminação de dados semânticos são uma tendência importante para instituições de memória em todo o mundo, mas a reutilização, integração e exploração eficientes desses dados requerem um alto nível de interoperabilidade semântica (Aalberg *et al.*, 2015).

A alternativa ao ambiente semântico seria criação ou utilização de sistemas de informação convencionais, mas o entendimento é que isso pode levar a um grande retrabalho no futuro, sem falar na possível perda de informações pela ausência de um modelo que leve em consideração o contexto dos objetos digitais, da utilização dos profissionais e da expectativa dos usuários.

Apesar de não ter sido possível realizar testes aprofundados com os curadores, pelo que foi exposto nas respostas do questionário e na entrevista com os curadores dos espaços de memória, foi possível obter que estes espaços possuem uma precariedade quanto à utilização dos sistemas de informação mais simples, e isso pode levar à uma certa dificuldade para utilização de sistemas mais complexos, com uma grande

quantidade de classes e propriedades. No entanto, o que se defende aqui é que a simplificação do modelo apresentado, partindo de uma estrutura *top-down*, e permitindo que os próprios curadores possam organizar suas classes, é um fator facilitador.

### **8.7 Sumário do capítulo**

Este capítulo teve como objetivo maior discutir os resultados desta tese com base no enquadramento teórico e estado da arte, presentes nos CAPÍTULOS II e III, na proposta apresentada no CAPÍTULO IV, e nos resultados apresentados no CAPÍTULO VI e no CAPÍTULO VII. Toda discussão foi conduzida seguindo a ordem dos objetivos apresentados no capítulo introdutório.

No próximo capítulo serão apresentadas as alegações finais, conclusões e contribuições deste trabalho bem como os apontamentos para trabalhos futuros.

## **CAPÍTULO IX. CONCLUSÃO E TRABALHOS FUTUROS**

### **9.1 Introdução**

O objetivo deste capítulo é realizar o fechamento da presente tese. Aqui, pretende-se discutir se a hipótese apresentada é válida ou não, e, com base nos objetivos do trabalho, discorrer em torno dos resultados, em especial do sistema de informação semântico já apresentado e discutido. Visa ainda apresentar as considerações finais, limitações e perspectivas de trabalho futuro.

### **9.2 Sumário da investigação**

Como explicitado CAPÍTULO I, o ponto de partida dessa pesquisa foi um problema prático, que é o da interoperabilidade semântica de patrimônio cultural. Dessa forma, pretendeu-se produzir resultados práticos. Tanto para a qualificação do problema quanto construção de propostas, partiu-se da cooperação com grupos de pessoas envolvidas na gênese do Projeto Centro de Memória de Goiás, mas, principalmente com curadores e servidores dos espaços de memória presentes, nas instituições que fizeram parte investigação e em todo o espectro de memória e preservação.

Num primeiro momento de interação envolvendo profissionais da educação, servidores de governos, estudantes e outros autores que afetam e são afetados pelo projeto, quer seja pela iminente ameaça da perda memória de um coletivo específico, quer seja pela dificuldade de ter acesso à memória. Em um segundo momento, delimitando o espectro da investigação, foi realizado um recorte em uma parte específica de atuação do projeto que envolve a parceria com a Secretaria de Estado de Cultura do Estado de Goiás (SECULT-GO).

Nos capítulos II e III foram apresentados o enquadramento teórico e estado da arte. No CAPÍTULO II descrevem-se os pressupostos da ciência da informação e os principais elementos da Web-Semântica, linguagens e tecnologias utilizadas para a construção da proposta presente nesta tese.

De forma complementar, o CAPÍTULO III apresentou uma revisão sistemática da literatura com os artigos científicos nas bases *ACM Digital Library*, *IEEE Xplorer*, *Science Direct* e *Web of Science*, correspondendo à temática da interoperabilidade semântica em Sistemas de Informação de patrimônio cultural.

No CAPÍTULO IV foi apresentada a proposta da presente tese, composta de um

modelo de Sistema de Informação semântico, aberto e conectado para centros de memória, a fim de contribuir na integração dos acervos digitais de patrimônio cultural através da interoperabilidade semântica entre os repositórios.

No CAPÍTULO V foi apresentado o percurso metodológico presente nesta tese, que se apoia na Pesquisa-ação e na Engenharia de Software, ambas de base empírica, para construção de resultados concretos.

Apenas a investigação bibliográfica poderia se apresentar empobrecida, uma vez que vários problemas não estão descritos ou disponíveis de forma fácil em referências bibliográficas, ou mesmo acessíveis em canais convencionais como a *Internet*. Como é possível encontrar em (Thiolent, 2011, p. 21) a abordagem empírica da investigação-ação “encontra um contexto favorável quando os pesquisadores não querem limitar sua investigação a aspectos burocráticos das pesquisas convencionais e buscam ser ativos no equacionamento dos problemas encontrados”.

No CAPÍTULO VI foram apresentados os resultados constantes do inquérito e entrevista realizados com os curadores de acervos culturais da Secretaria de Estado de Cultura de Goiás (SECULT-GO), e a modelagem do protótipo de Sistema de Informação semântico resultante da proposta desta tese, dando ênfase aos artefatos da Engenharia de Software, com base nos requisitos do recorte realizado.

No CAPÍTULO VII foi apresentado o Guará, um protótipo de Sistema de Informação semântico para o patrimônio cultural que permite o intercâmbio entre múltiplos repositórios a partir de uma visão da informação em quatro dimensões (Pessoa, Fenômeno, Espaço e Tempo).

No CAPÍTULO VIII foi realizada uma análise dos resultados e uma discussão detalhada a partir dos objetivos declarados nessa tese, dando ênfase ao que foi proposto, em especial aos mecanismos internos do protótipo.

### **9.3 Questão de pesquisa, questões acessórias e hipótese**

Quanto à questão inicial de pesquisa, “*como integrar de forma semântica, aberta e conectada os acervos digitais de patrimônio cultural dos espaços de memória da Secretaria de Estado da Cultura do Estado de Goiás?*”, apresentaram-se nessa tese uma proposta de modelo genérico (CAPÍTULO IV) e um conjunto de tecnologias e linguagens, bem como, uma estratégia possível para integrar os repositórios de forma semântica com base em fundamentos teóricos e estado da arte (CAPÍTULO II e III).

Em um caminho oposto ao da tentativa de apresentar um modelo pronto e completo, aqui buscou-se apresentar uma trajetória da pesquisa realizada e, de forma alinhada aos objetivos desta tese, apresentar respostas, por mais simplificadas que sejam, que possam contribuir para o problema da interoperabilidade semântica, que foi evidenciado e qualificado nos capítulos iniciais, mas que levaram em conta um problema concreto referente ao patrimônio cultural do estado de Goiás.

Ressalte-se que, a partir do que foi apresentado e discutido, se faz necessário enfatizar que não se trata de um conjunto trivial de atividades. Desde as etapas iniciais desta pesquisa percebeu-se a complexidade que envolve aspectos humanos, sociais, políticos e tecnológicos, sendo estes últimos especialmente críticos para o problema abordado nessa tese.

Entretanto, defende-se aqui que, em se utilizando dos mesmos materiais e métodos, mesmo em situação peculiar, é possível que os resultados apresentados no CAPÍTULO VI e CAPÍTULO VII e discutidos no CAPÍTULO VIII, não possam ser replicados de forma similar, levando em consideração que os diversos fenômenos aqui explorados são altamente subjetivos.

A questão de partida, que de forma indireta foi respondida pelos diversos capítulos e seções dessa tese, pode ter a seguinte resposta resumida. “Através de um conjunto ferramental que envolve uma modelagem conceitual com utilização de ontologias, um sistema de informação com armazenamento em grafos RDF, atendendo as especificações da Web Semântica e dos Dados Abertos Conectados”

A hipótese deste trabalho remeteu-se a testar se “*Um modelo semântico permite o intercâmbio de informações de patrimônio cultural entre os acervos digitais*”, que teve na proposta apresentada no CAPÍTULO IV e os resultados presentes no CAPÍTULO VI e CAPÍTULO VII resposta positiva.

Pode-se dizer, com base no que foi exposto, que o protótipo de Sistema de Informação apresentado permite que múltiplos repositórios possam, através de uma camada de serviços e relações criadas através das técnicas de arquivos abertos conectados, intercambiar informações de forma colaborativa que permite uma relação de interoperabilidade entre os acervos digitais.

Uma questão complementar à essa hipótese é verificar se “*A complexidade do modelo semântico não é um empecilho para os curadores*”. Como ficou evidenciado no

estado da arte, são inúmeras as iniciativas que promovem sistematizações em torno do patrimônio cultural. O projeto *Europeana*, como destaque, atua como ponto de agregação de milhares de arquivos, bibliotecas e museus europeus com um acervo de mais de 50 milhões de itens digitalizados.

Tanto o modelo EDM quanto o CIDOC-CRM, como pode ser visto no CAPÍTULO II e Seção 8.3 (p. 256), permitem uma organização detalhada de objetos do patrimônio cultural. Mas a ausência de alternativas voltadas especificamente para o patrimônio cultural do estado de Goiás torna difícil explorar e delimitar sua utilização.

Foi apresentado nessa tese, entretanto, um protótipo de sistema de informação semântico que leva em consideração as características internas dos espaços de memória, em uma alternativa diferente, na qual ao invés de construir agregadores, criam-se repositórios capazes de conectarem entre si.

A despeito do desafio semântico, o que se percebeu através do questionário e entrevista com os curadores de memória, é que os espaços possuem em geral uma estrutura física e tecnológica precária. Os curadores possuem bom conhecimento sobre os espaços. Todos os espaços possuem alguma forma de organização usual. Nenhum dos curadores, entretanto, possui conhecimento avançado sobre metodologias, padrões internacionais, ou tecnologias semânticas, como mostra a tabela de resultados sobre o inquérito e entrevista com estes curadores na Seção 6.3 (p. 172).

Com base nessas questões, isso leva a crer que, a complexidade dos modelos semânticos pode ser um empecilho ou pode criar empecilhos para os repositórios digitais nestes espaços. Defende-se aqui, no entanto, que o modelo simplificado é capaz de diminuir tais impactos, uma vez que a complexidade cresce de acordo com o tamanho os acervos e possibilidades de interligações.

#### **9.4 Objetivos do trabalho**

O objetivo geral desse foi desenvolver um modelo de sistema de informação semântico, aberto e conectado para acervos digitais de patrimônio cultural, especificamente voltado para patrimônio cultural do Estado de Goiás.

Esta tese, como um todo, pretendeu atingir este objetivo levando em consideração as nuances dos espaços de memória da Secretaria de Estado da Cultura e tendo como partida o projeto Centro de Memória de Goiás. Especificamente no CAPÍTULO IV foi apresentada a proposta dessa tese, composta de um modelo de sistema

de informação com base em duas modelagens principais, uma conceitual e outra tecnológica.

Quanto aos objetivos específicos, o primeiro foi “Realizar um estudo sobre as características dos espaços de memória e patrimônio ligados à Secretaria de Estado da Cultura do Estado de Goiás (Bibliotecas, Museus, Arquivos e Galerias), e sua utilização quanto aos sistemas de informação”, conforme discussões do capítulo anterior, teve seu cumprimento através de um diagnóstico que também foi utilizado para a construção do sistema proposto, uma vez que se trata de uma investigação-ação.

O segundo objetivo, que foi “Realizar um estudo sobre a aplicabilidade dos modelos conceituais de representação da informação de patrimônio cultural e tecnologias da Web-Semântica” que, aqui defende-se, foi alcançado através da fundamentação teórica e estado da arte.

O Terceiro objetivo declarado, que foi “Construir um modelo semântico para organização da informação de patrimônio cultural, a partir de padrões técnicos, que permita a interoperabilidade entre repositórios de patrimônio cultural, aqui denominados espaços de memória”, está irrefutavelmente ligado ao primeiro e segundo objetivo. O modelo apresentado no CAPÍTULO IV, o protótipo apresentado CAPÍTULO VI, e a discussão sobre o protótipo no CAPÍTULO VIII mostram que o objetivo foi alcançado.

Já o quarto objetivo específico, *Verificar se a complexidade do modelo semântico é um empecilho para os curadores de acervos digitais de patrimônio cultural*, alcançou-se de forma indireta, qual seja, analisar a partir das respostas dos questionários e entrevistas, pela experiência e pelo estado da arte apresentado. Como discutido na Seção 7.3.

## **9.5 Contributos do trabalho**

De forma não pretensiosa e sob a premissa que a Ciência se desenvolve muitas vezes de pequenas contribuições que, juntadas podem oferecer caminhos, defende-se a partir desta tese os seguintes contributos.

1) uma metodologia prática para criação de sistemas de informação semânticos para o patrimônio cultural em quatro camadas;

2) uma ontologia simplificada que permite a organização de acervos de patrimônio cultural sob uma perspectiva interna;

3) uma ontologia de descrição dos objetos que permite que eles sejam classificados de forma mínima e simplificada, constante de propriedades de classes já estabelecidas na comunidade;

4) uma ontologia para interoperabilidade semântica em 4 dimensões (Pessoa, Fenômeno, Espaço e Tempo);

5) uma estratégia de transformação de objetos em triplas que podem ser persistidas através de SPARQL;

6) um diagnóstico sobre os espaços de memória da secretaria de Estado da Cultura do Estado de Goiás (SECULT-GO) e seus sistemas de informação;

7) um Algoritmo de interoperabilidade que busca potenciais ligações por palavra-chave e relações entre os objetos dos acervos.

## **9.6 Limitações do trabalho**

O trabalho apresenta algumas limitações. Uma se refere ao reuso que não foi abordado à medida da sua importância. Outra limitação está relacionada aos espaços. O recorte apresentado levou em consideração os espaços de memória da Secretaria de Estado de Cultura que possui parceria com Universidade Estadual de Goiás no Projeto Centro de Memória de Goiás. Portanto, instituições como IBRAM, IPHAN e outras relacionadas ao patrimônio não foram envolvidas nesse processo.

Há uma limitação semântica em adicionar mais propriedades e classes às ontologias, e conseqüentemente, regras e axiomas para aumentar o poder semântico e permitir o raciocínio e recuperação da informação baseada em regras, como a utilização do SWRL.

Há uma limitação de alcance conceitual e de outros elementos das ontologias do protótipo. Foi criada uma versão simplificada devido às inúmeras dificuldades de realizar esse trabalho de forma completa. Isso poderia durar anos, senão décadas.

## **9.7 Trabalhos futuros**

- Complementar as demais funcionalidades de um Sistema de Informação de Patrimônio Cultural.
- Construir uma ontologia com o mapeamento das classes do modelo para as principais ontologias (CIDOC-CRM, EDM, *Schema.org*).

- Criar mais possibilidades para exploração e reuso.
- Criar portal dos professores.
- Adicionar classes, propriedades, axiomas e regras às ontologias permitindo raciocínio e recuperação baseada em regras.

## **9.8 Recomendações**

Certa vez, em uma ação anterior à presente investigação que originou esta tese, uma colega, jovem pesquisadora, apresentou um HD externo contendo centenas, talvez milhares de imagens de um acervo documental riquíssimo, referente ao século XVIII e XIX. Havia ali algum tipo de organização nas pastas em que os arquivos estavam armazenados e também havia uma planilha com algumas informações que a pesquisadora achava pertinente. A mesma pesquisadora, revelou sobre trabalhos de outros historiadores que também estavam a coletar e digitalizar outros acervos históricos, e que cada acervo, a depender de sua forma, local, idade, e o olhar do pesquisador, possuía uma estratégia de escolha, descrição e organização dos objetos pesquisados. Ao final daquele encontro, a pesquisadora em questão e seu interlocutor (o autor desta tese) dividiram a frustração que reside na dificuldade de conceber sistemas capazes de absorver toda a complexidade existente nas coleções de patrimônio cultural e nas suas relações com a sociedade. Como avançar para além da simples disponibilização das imagens e informações cruas? O compartilhamento dessa riqueza em suas mais diversas perspectivas continua sendo um desafio e tanto.

Dito isso, cabe nessas alegações finais reforçar que os sistemas de informação não existem de forma isoladas. Eles possuem um ciclo de vida que envolvem diversos contextos e levam em conta aspectos organizacionais, tecnológicos e sociais. Diante dessa perspectiva, as boas práticas da construção de softwares devem ser levadas em consideração respeitando, além das técnicas, os ambientes com fatores internos e externos envolvidos.

Recomenda-se, portanto, como ponto de partida (ou de convergência) crucial para os Sistemas de Patrimônio Cultural, que os curadores, pesquisadores e membros dos grupos culturais sejam cada vez mais ouvidos e que suas angústias e experiências sejam levadas em consideração durante todo o percurso da construção dos artefatos dos sistemas.

## 9.9 Publicações resultantes da investigação

- Filgueiras Alison Carlos, Gouveia, Feliz Alberto Ribeiro and Paula Roberta Chagas. 2020. “An open, semantic and connected model for digital collections of cultural heritage of Goiás State”. *In proceedings. 2nd International Conference on Museum Big Data*. Nicosia, Cyprus.
- Chagas, Paula Roberta, Feliz Alberto Ribeiro Gouveia, Alison Carlos Filgueiras, and Marcos Antônio da Torres. 2018. “Goiás State University as Locus for Heritage Preservation: Genesis of a Memory Center.” *In 6th International Conference on Heritage and Sustainable Development*, 741–48. Granada-ES.
- Filgueiras, Alison Carlos, Feliz Ribeiro Gouveia, and Paula Roberta; Chagas. 2020. “Interoperabilidade Semântica Em Patrimônio Cultural: Estratégias e Ferramentas Para Concepção de Um Centro de Memória.” *Revista Sapiência: Sociedade, Saberes e Práticas Educacionais* 155: 136–55.
- Filgueiras, Alison Carlos and Feliz Ribeiro Gouveia. “Plataformas Semânticas de Patrimônio Cultural: uma abordagem em Linked Open Data para o Centro de Memória de Goiás”. *In proceedings SITEGI*. 2019. Porto.
- Chagas, Paula Roberta, Gouveia Feliz Ribeiro, Filgueiras, Alison Carlos, França, Thyago Madeira. 2019. “A Extensão Universitária como estratégia para o desenvolvimento sustentável: o caso do Centro de Memória de Goiás. *In 1<sup>st</sup> International Conference on Humanitarian Action and Cooperation for Development*. Porto.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Aalberg, T., Vennesland, A., & Farrokhnia, M. (2015). A pattern-based framework for best practice implementation of CRM/FRBROO. *Communications in Computer and Information Science*, 539, 438–447. [https://doi.org/10.1007/978-3-319-23201-0\\_44](https://doi.org/10.1007/978-3-319-23201-0_44)
- Abu Musa, T. H., Bouras, A., Belhi, A., & Gasmi, H. (2020). OntoM: An Ontological Approach for Automatic Classification. *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*, 329–334. <https://doi.org/10.1109/ICIoT48696.2020.9089449>
- Agosti, M., Ferro, N., & Silvello, G. (2015). Digital library interoperability at high level of abstraction. *Future Generation Computer Systems*, 55, 129–146. <https://doi.org/https://doi.org/10.1016/j.future.2015.09.020>
- Alexiev, V., Tarkalanov, P., Georgiev, N., & Pavlova, L. (2020). Bulgarian Icons in Wikidata and EDM. *DIGITAL PRESENTATION AND PRESERVATION OF CULTURAL AND SCIENTIFIC HERITAGE*, 10, 45–63.
- Allemang, D., Hendler, J., & Gandon, F. (2020). *Semantic Web for the Working Ontologist: Modeling in RDF, RDFS and OWL* (3rd ed.). Association for Computing Machinery. <https://doi.org/10.1145/3382097>
- Alvarenga, L. (2003). REPRESENTAÇÃO DO CONHECIMENTO NA PERSPECTIVA DA CIÊNCIA DA INFORMAÇÃO EM TEMPO E ESPAÇO DIGITAIS REPRESENTATION OF THE KNOWLEDGE IN THE PERSPECTIVE OF THE. *Enc. Bibli: R. Eletr. Bibliotecon. Ci. Inf*, 8, 18–40.
- Amato, F., Mazzeo, A., Mazzocca, N., & Moscato, F. (2016). A Formal Model for Autonomous Planning in High Performance Systems. *Procedia Computer Science*, 58, 503–508. <https://doi.org/10.1016/j.procs.2016.09.085>
- Amato, G., Falchi, F., & Vadicamo, L. (2016). Visual Recognition of Ancient Inscriptions Using Convolutional Neural Network and Fisher Vector. *J. Comput. Cult. Herit.*, 9(4). <https://doi.org/10.1145/2964911>
- Apache Software Foundation. (2021). *An Introduction to RDF and the Jena RDF API*. [https://jena.apache.org/tutorials/rdf\\_api.html](https://jena.apache.org/tutorials/rdf_api.html)
- Araújo, C. A. Á. (2009). Correntes teóricas da ciência da informação. *Ciência Da Informação*, 38(3), 192–204. <https://doi.org/10.1590/S0100-19652009000300013>
- Araújo, C. A. Á. (2014). Fundamentos Da Ciência Da Informação: Correntes Teóricas E O

- Conceito De Informação. *Perspectivas Em Gestão & Conhecimento*, 4, 57–79.  
<http://periodicos.ufpb.br/ojs2/index.php/pgc.ISSN:2236-417X.PublicaçãosobLicença>.
- Araújo, C. A. Á. (2017). TEORIAS E TENDÊNCIAS CONTEMPORÂNEAS DA CIÊNCIA DA INFORMAÇÃO. *Inf. Pauta*, 2, 11–34.
- Backhouse, J., Hsu, C., & McDonnell, A. (2003). Toward public-key infrastructure interoperability. *Communications of the ACM*, 46(6), 98–100.  
<https://doi.org/10.1145/777313.777345>
- Badii, C., Bellini, P., Cenni, D., Difino, A., Nesi, P., & Paolucci, M. (2017). Analysis and assessment of a knowledge based smart city architecture providing service APIs. *Future Generation Computer Systems*, 75, 14–29.  
<https://doi.org/https://doi.org/10.1016/j.future.2017.05.001>
- Bannour, I., Marinica, C., Bouiller, L., Pillay, R., Darrieumerlou, C., Malavergne, O., Kotzinos, D., & Niang, C. (2018). CRMCR - a CIDOC-CRM extension for supporting semantic interoperability in the conservation and restoration domain. *2018 3rd Digital Heritage International Congress (DigitalHERITAGE) Held Jointly with 2018 24th International Conference on Virtual Systems & Multimedia (VSMM 2018)*, 1–8.  
<https://doi.org/10.1109/DigitalHeritage.2018.8810098>
- Benedusi, P., Chianese, A., Marulli, F., & Piccialli, F. (2015). An Associative Engines Based Approach Supporting Collaborative Analytics in the Internet of Cultural Things. *2015 10th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC)*, 533–538. <https://doi.org/10.1109/3PGCIC.2015.56>
- Benouaret, I., & Lenne, D. (2015). Personalizing the Museum Experience through Context-Aware Recommendations. *2015 IEEE International Conference on Systems, Man, and Cybernetics*, 743–748. <https://doi.org/10.1109/SMC.2015.139>
- Berners-lee, T. (2006). *Linked Data*. <https://www.w3.org/DesignIssues/LinkedData.html>
- Berners-Lee, T. (1998). *Realising the Full Potential of the Web*.  
<https://www.w3.org/1998/02/Potential.html>
- Berners-Lee, T. (2006). *Linked Data*. <https://Www.W3.Org/DesignIssues/LinkedData.Html>.  
<https://www.w3.org/DesignIssues/LinkedData.html>
- Berners-Lee, T. (2012). *5 Stars Open Data*. <https://5stardata.info/pt-BR/>
- Berners-lee, T., & Hara, K. O. (2013). The read – write Linked Data Web Subject Areas : *Philosophical Transactions of the Royal Society*, 'Web science: a new frontier'.  
<https://royalsocietypublishing.org/doi/pdf/10.1098/rsta.2012.0513>

- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific American*, 284(5), 34–43. <https://doi.org/10.1038/scientificamerican0501-34>
- Bertola, F., & Patti, V. (2015). Ontology-based affective models to organize artworks in the social semantic web. *Information Processing & Management*, 52(1), 139–162. <https://doi.org/https://doi.org/10.1016/j.ipm.2015.10.003>
- Boeuf, P. Le, Doerr, M., Ore, C. E., & Stead, S. (2018). *Definition of the CIDOC Conceptual Reference Model*.
- Bordalo, R., Bottaini, C., & Candeias, A. (2021). A Framework Design for Information Management in Heritage Science Laboratories. *J. Comput. Cult. Herit.*, 14(1). <https://doi.org/10.1145/3417304>
- Borko, H. (1986). What Is Information Science? *Soviet Education*, 28(10–11), 51–54.
- Plano Nacional de Cultura, 2007 (2007).
- Briola, D., Deufemia, V., Mascardi, V., & Paolino, L. (2017). Agent-oriented and ontology-driven digital libraries: the IndianaMAS experience. *Software - Practice and Experience*, 47(11), 1773–1799. <https://doi.org/10.1002/spe.2494>
- Büchler, M., Riegert, S., Alpi, F., & Cadeddu, F. (2020). Towards Big Religious Data: RESILIENCE Research Infrastructure for Data on Religion in the Digital Age. *Proceedings of the 2nd International Conference on Digital Tools & Uses Congress*. <https://doi.org/10.1145/3423603.3424007>
- Buckland, M. (1991). Information as Thing. *Journal of the American Society for Information Science*, 42(5), 351.
- BUNDER, J., & BARROS, G. G. (2019). *O Estudo De Caso E a Pesquisa-Ação: Compreensão Teórica E Evidências Empíricas*. *October*, 1561–1565. <https://doi.org/10.14393/sbqp19140>
- Buranasing, W., Phoomvuthisarn, S., & Buranarach, M. (2016). Information extraction and integration for enriching cultural heritage collections. *2016 11th International Conference on Knowledge, Information and Creativity Support Systems (KICSS)*, 1–6. <https://doi.org/10.1109/KICSS.2016.7951425>
- Burke, P. (2003). *Uma História Social do Conhecimento: de Gutemberg a Diderot*.
- Burke, P. (2012). *Uma História Social do Conhecimento II: da Enciclopédia à Wikipedia* (Zahar (ed.)).
- Bush, V. (1945). As we may think. *THE ATLANTIC MONTHLY*.

- Butto, S. (2017). The ICCU'S Contribution to the Development of a Digital Infrastructure for Cultural Heritage. *ARCHEOLOGIA E CALCOLATORI*, 9, 53–59.
- Cacciotti, R., & Valach, J. (2015). The MONDIS project Semantic Web and the protection of historic buildings. *2015 Digital Heritage*, 2, 307–313.  
<https://doi.org/10.1109/DigitalHeritage.2015.7419512>
- Calvanese, D., Liuzzo, P., Mosca, A., Remesal, J., Rezk, M., & Rull, G. (2016). Ontology-based data integration in EPNet: Production and distribution of food during the Roman Empire. *Engineering Applications of Artificial Intelligence*, 51, 212–229.  
<https://doi.org/https://doi.org/10.1016/j.engappai.2016.01.005>
- Calvanese, D., Mosca, A., Remesal, J., Rezk, M., & Rull, G. (2015). A 'historical case' of Ontology-Based Data Access. *2015 Digital Heritage*, 2, 291–298.  
<https://doi.org/10.1109/DigitalHeritage.2015.7419510>
- Camargo, A. M., & Goulart, S. (2015). *Centros de Memória: Uma proposta de definição* (C. S. Culturas (ed.); 1a ed.). Sesc-SP.
- Capodieci, A., Mainetti, L., & Carrozzo, S. (2016). Semantic enterprise service bus for cultural heritage. *2016 12th International Conference on Innovations in Information Technology (IIT)*, 1–8. <https://doi.org/10.1109/INNOVATIONS.2016.7880044>
- Capurro, R., & Hjarland, B. (2003). Theorizing Information and Information Use The Concept of Information. *Annual Review of Information Science and Technology*, 343–411.
- Carrasco, L., Thaller, M., & Vidotti, S. A. B. G. (2017). Ontologia Cidoc CRM no contexto dos ambientes digitais de patrimônios culturais. *Liinc Em Revista*, 11(1), 208–222.  
<https://doi.org/10.18617/liinc.v11i1.789>
- Carvalho Junior, J. M. C., Martins, D. L., & Germani, L. B. (2019). GLAM e Instituições de Memória em Rede: uma 'Infosfera' de Floridi? *PragMatizes*.
- Castells, M. (2016). *A Sociedade em Rede* (17a Edição).
- Cerutti, E., Noardo, F., & Spanò, A. (2015). Architectural Heritage semantic data managing and sharing in GIS. *2015 1st International Conference on Geographical Information Systems Theory, Applications and Management (GISTAM)*, 1–8.
- Chagas, P. R., Gouveia, F. A. R., Filgueiras, A. C., & Torres, M. A. da. (2018). Goiás State University as locus for heritage preservation: genesis of a memory center. *6th International Conference on Heritage and Sustainable Development*, 741–748.
- Chen, H., Chen, S., & Wang, C. Y. (2014). Object-mediated innovation: Case study of adventures in the National Palace Museum in Taiwan. *Proceedings of PICMET '14*

*Conference: Portland International Center for Management of Engineering and Technology; Infrastructure and Service Integration*, 2214–2225.

- Chen, J., Dowman, I., Li, S., Li, Z., Madden, M., Mills, J., Paparoditis, N., Rottensteiner, F., Sester, M., Toth, C., Trinder, J., & Heipke, C. (2016). Information from imagery: ISPRS scientific vision and research agenda. *ISPRS Journal of Photogrammetry and Remote Sensing*, 115, 3–21. <https://doi.org/https://doi.org/10.1016/j.isprsjprs.2015.09.008>
- Coneglian, C. S., & Santarem Segundo, J. E. (2016). Europeana no Linked Open Data: conceitos de Web Semântica na dimensão aplicada das Humanidades Digitais. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência Da Informação*, 22(48), 88. <https://doi.org/10.5007/1518-2924.2017v22n48p88>
- Curé, O., & Blin, G. (2015). *RDF database systems : triples storage and SPARQL query processing*.
- Damiano, R., Lombardo, V., & Lieto, A. (2015). Visual metaphors for semantic cultural heritage. *2015 7th International Conference on Intelligent Technologies for Interactive Entertainment (INTETAIN)*, 100–109.
- Daquino, M., Mambelli, F., Peroni, S., Tomasi, F., & Vitali, F. (2017). Enhancing Semantic Expressivity in the Cultural Heritage Domain: Exposing the Zeri Photo Archive as Linked Open Data. *J. Comput. Cult. Herit.*, 10(4). <https://doi.org/10.1145/3051487>
- Daquino, M., & Tomasi, F. (2014). Ontological Approaches to Information Description and Extraction in the Cultural Heritage Domain. *Proceedings of the Third AIUCD Annual Conference on Humanities and Their Methods in the Digital Ecosystem*. <https://doi.org/10.1145/2802612.2802633>
- Davies, J., Studer, R., & Warren, P. (2006). Semantic Web Technologies: Trends and Research in Ontology-based Systems. In *Semantic Web Technologies: Trends and Research in Ontology-based Systems*. <https://doi.org/10.1002/047003033X>
- Demetrescu, E. (2015). Archaeological stratigraphy as a formal language for virtual reconstruction. Theory and practice. *JOURNAL OF ARCHAEOLOGICAL SCIENCE*, 57, 42–55. <https://doi.org/10.1016/j.jas.2015.02.004>
- Dessì, D., Osborne, F., Reforgiato Recupero, D., Buscaldi, D., & Motta, E. (2020). Generating knowledge graphs by employing Natural Language Processing and Machine Learning techniques within the scholarly domain. *Future Generation Computer Systems*, 116, 253–264. <https://doi.org/https://doi.org/10.1016/j.future.2020.10.026>
- Dodebei, V. (2014). Memória do conhecimento: em busca de sustentabilidade para os objetos digitais. *Ciência Da Informação*, 43(1), 145–153.

<http://search.ebscohost.com/login.aspx?direct=true&db=llf&AN=112718974&site=ehost-live>

- Dodebei, V. (2016). *Objetos & memória*. 227–243.
- Doerr, M. (2009). Ontologies for Cultural Heritage. In S. Staaf & R. Studer (Eds.), *Handbook on Ontologies* (Second Edi, pp. 463–486). Springer. <https://doi.org/10.1007/978-3-540-92673-3>
- Doerr, M., Bruseker, G., Bekiari, C., Ore, C. E., Velios, T., & Stead, S. (2020). Definition of the CIDOC Conceptual Reference Model Produced by the ICOM / CIDOC Documentation Standards Group , continued by the CIDOC CRM Special Interest Group. In *Definition of the CIDOC Conceptual Reference Model Produced* (Issue CIDOC-CRM).
- Doerr, M., Ore, C.-E., & Stead, S. (2007). The CIDOC Conceptual Reference Model - A New Standard for Knowledge Sharing. *Tutorials, Posters, Panels and Industrial Contributions at the 26th International Conference on Conceptual Modeling - Volume 83*, 83(Er), 51–56. <http://dl.acm.org/citation.cfm?id=1386957.1386963>
- Drakopoulos, G., Spyrou, E., Voutos, Y., & Mylonas, P. (2019). A Semantically Annotated JSON Metadata Structure for Open Linked Cultural Data in Neo4j. *Proceedings of the 23rd Pan-Hellenic Conference on Informatics*, 81–88. <https://doi.org/10.1145/3368640.3368659>
- Europeana. (2017). *Definition of the Europeana Data Model*. <https://pro.europeana.eu/resources/standardization-tools/edm-documentation>
- Feldman, D., & Steindel, G. E. (2019). As relações entre Centros de Memória e Ciência da Informação. *INCID: Revista de Ciência Da Informação e Documentação*, 10(1), 147–166. <https://doi.org/10.11606/issn.2178-2075.v10i1p147-166>
- Filgueiras, A., Da Silva, J., & Vincenzi, A. (2013). A prototype for querying heterogeneous data sources on the Web. *Proceedings of the IADIS International Conference WWW/Internet 2013*.
- Filgueiras, A., Gouveia, F., & Chagas, P. (2020). Interoperabilidade semântica em patrimônio cultural: estratégias e ferramentas para concepção de um centro de memória. *Revista Sapiência: Sociedade, Saberes e Práticas Educacionais*, 155, 136–155.
- Floridi, L. (2002). On defining library and information science as applied philosophy of information. *Social Epistemology*, 16(1), 37–49. <https://doi.org/10.1080/02691720210132789>
- Freire, N, Meijers, E., Valk, S. de, Voorburg, R., Isaac, A., & Cornelissen, R. (2018).

- Aggregation of Linked Data : A case study in the cultural heritage domain. *2018 IEEE International Conference on Big Data (Big Data)*, 522–527.  
<https://doi.org/10.1109/BigData.2018.8622348>
- Freire, Nuno, Meijers, E., Voorburg, R., & Isaac, A. (2018). Aggregation of cultural heritage datasets through the Web of Data. *Procedia Computer Science*, *137*, 120–126.  
<https://doi.org/10.1016/j.procs.2018.09.012>
- Frommholz, I., Graves, D., Liu, H., Kumar, A., & Brady, G. (2014). Great War Stories Told by the People: Crowdsourced Cultural Heritage in Digital Museums. *Proceedings of the 14th ACM/IEEE-CS Joint Conference on Digital Libraries*, 419–420.
- Galvão, T. F., & Pereira, M. G. (2014). Revisões sistemáticas da literatura: passos para sua elaboração. *Epidemiologia e Serviços de Saúde*, *23*(1), 183–184.  
<https://doi.org/10.5123/s1679-49742014000100018>
- Giuliani, G., Dao, H., De Bono, A., Chatenoux, B., Allenbach, K., De Laborie, P., Rodila, D., Alexandris, N., & Peduzzi, P. (2017). Live Monitoring of Earth Surface (LiMES): A framework for monitoring environmental changes from Earth Observations. *Remote Sensing of Environment*, *202*, 222–233.  
<https://doi.org/https://doi.org/10.1016/j.rse.2017.05.040>
- Gruber, T. R. (1993). A Translation Approach to Portable Ontology Specifications. *Knowl. Acquis.*, *5*(2), 199–220. <https://doi.org/10.1006/knac.1993.1008>
- Guarino, N. (1998). *Formal Ontology in Information Systems: Proceedings of the 1st International Conference June 6-8, 1998, Trento, Italy* (1st ed.). IOS Press.
- Guarino, N., Oberle, D., & Staab, S. (2009). What Is an Ontology? In S. Staab & R. Studer (Eds.), *International Handbooks on Information Systems* (2nd ed., pp. 1–17).  
<https://doi.org/10.1007/978-3-540-92673-3>
- Hebeler, J., Fisher, M., Blace, R., & Perez-lopez, A. (2009). *Semantic*. Willey.
- Horridge, M., Jupp, S., Moulton, G., Rector, A., Stevens, R., & Wroe, C. (2007). A Practical Guide To Building OWL Ontologies Using Protégé 4 and CO-ODE Tools. In *October*.
- Ivanović, D., & Houssos, N. (2014). Providing an Application-specific Interface over a CERIF Back-end: Challenges and Solutions. *Procedia Computer Science*, *33*, 11–17.  
<https://doi.org/https://doi.org/10.1016/j.procs.2014.06.003>
- Kim, S., Ahn, J., Suh, J., Kim, H., & Kim, J. (2015). Towards a semantic data infrastructure for heterogeneous Cultural Heritage data - challenges of Korean Cultural Heritage Data Model (KCHDM). *2015 Digital Heritage*, *2*, 275–282.

<https://doi.org/10.1109/DigitalHeritage.2015.7419508>

- Kiryakov, A. (2006). Ontologies for Knowledge Management. In J. Davies, R. Studer, & P. Warren (Eds.), *Handbook on Ontologies*.
- Korzun, D., Varfolomeyev, A., Yalovitsyna, S., & Volokhova, V. (2017). Semantic Infrastructure of a Smart Museum: Toward Making Cultural Heritage Knowledge Usable and Creatable by Visitors and Professionals. *Personal Ubiquitous Comput.*, 21(2), 345–354. <https://doi.org/10.1007/s00779-016-0996-7>
- Kouis, D., & Giannakopoulos, G. (2014). Incorporate Cultural Artifacts Conservation Documentation to Information Exchange Standards – The DOC-CULTURE Case. *Procedia - Social and Behavioral Sciences*, 147, 495–504. <https://doi.org/https://doi.org/10.1016/j.sbspro.2014.07.144>
- Kuo, C.-L., & Hong, J.-H. (2016). Interoperable cross-domain semantic and geospatial framework for automatic change detection. *Computers & Geosciences*, 86, 109–119. <https://doi.org/10.1016/j.cageo.2015.10.011>
- Le Coadic, Y.-F. (1994). *A Ciência da Informação*.
- Lèvy, P. (2014). *Cibercultura* (Editora 34 (ed.); 3a Ed. 2a). Editora 34.
- Liang, H., Chang, J., Wang, M., Chen, C., & Zhang, J. J. (2016). Semantic Framework for Interactive Animation Generation. *Proceedings of the 15th ACM SIGGRAPH Conference on Virtual-Reality Continuum and Its Applications in Industry - Volume 1*, 137–145. <https://doi.org/10.1145/3013971.3013998>
- Lima, F. R. B., Santos, P. L. V. A. C., & Santarém Segundo, J. E. (2016). Padrão de metadados no domínio museológico. *Perspectivas Em Ciência Da Informação*, 21(3), 50–69. <https://doi.org/10.1590/1981-5344/2639>
- Marconi, M., & Lakatos, E. (2003). Fundamentos de metodologia científica. In *Editora Atlas S. A.* <https://doi.org/10.1590/S1517-97022003000100005>
- Martini, R., Librelotto, G., & Henriques, P. R. (2016). Formal Description and Automatic Generation of Learning Spaces Based on Ontologies. *Procedia Computer Science*, 96, 235–244. <https://doi.org/https://doi.org/10.1016/j.procs.2016.08.136>
- Meghini, C., Scopigno, R., Richards, J., Wright, H., Geser, G., Cuy, S., Fihn, J., Fanini, B., Hollander, H., Niccolucci, F., Felicetti, A., Ronzino, P., Nurra, F., Papatheodorou, C., Gavrilis, D., Theodoridou, M., Doerr, M., Tudhope, D., Binding, C., & Vlachidis, A. (2017). ARIADNE: A Research Infrastructure for Archaeology. *ACM JOURNAL ON COMPUTING AND CULTURAL HERITAGE*, 10(3, SI). <https://doi.org/10.1145/3064527>

- Menezes, S. M. P. de. (2006). Conservação e Memória. *Cadernos de Sociomuseologia*, 2003, 9–11.
- Mountantonakis, M., & Tzitzikas, Y. (2019). Large-Scale Semantic Integration of Linked Data: A Survey. *ACM Comput. Surv.*, 52(5). <https://doi.org/10.1145/3345551>
- Moura, M. A. (2009). *Informação, ferramentas ontológicas e redes sociais: a interoperabilidade na construção de tesouros e ontologias*. 589–604.
- Moura, M. A. (2011). Interoperabilidade Semântica e Ontologia Semiótica: a construção e o compartilhamento de conceitos científicos em ambientes colaborativos online. *Informação & Informação*, 16(supl), 165–179. <https://doi.org/10.5433/1981-8920.2011v16nesp.p165>
- Nafis, F., Fararnni, K. A., Yahyaouy, A., & Aghoutane, B. (2020). Towards a semantic recommender system for cultural objects: Case study Draa-Tafilalet region. *2020 International Conference on Intelligent Systems and Computer Vision (ISCV)*, 1–4. <https://doi.org/10.1109/ISCV49265.2020.9204187>
- Niang, C., Marinica, C., Leboucher, É., Bouiller, L., Capderou, C., & Bouchou, B. (2016). Ontology-Based Data Integration System for Conservation-Restoration Data (OBDIS-CR). *Proceedings of the 20th International Database Engineering & Applications Symposium*, 218–223. <https://doi.org/10.1145/2938503.2938545>
- Niang, C., Marinica, C., Markhoff, B., Leboucher, E., Malavergne, O., Bouiller, L., Darrieumerlou, C., & Laissus, F. (2017). Supporting semantic interoperability in conservation-restoration domain: The PARCOURS project. *Journal on Computing and Cultural Heritage*, 10(3). <https://doi.org/10.1145/3097571>
- Nishanbaev, I. (2020). A web repository for geo-located 3D digital cultural heritage models. *Digital Applications in Archaeology and Cultural Heritage*, 16, e00139. <https://doi.org/https://doi.org/10.1016/j.daach.2020.e00139>
- Nishimoto, H., Hamada, A., Takai, Y., & Goto, A. (2015). Investigation of Decision Process for Purchasing Foodstuff in the “Bento” Lunch Box. *Procedia Manufacturing*, 3, 472–479. <https://doi.org/10.1016/j.promfg.2015.07.210>
- Noy, N. F., & McGuinness, D. L. (2001). *Ontology Development 101: A Guide to Creating Your First Ontology*.
- OKF. (2020). *Open Knowledge Foundation: Open Data Handbook*. <https://okfn.org/opendata/>
- Owais, N. N., Al-Feel, H. T., & Hassanein, E. E. (2017). A step towards the linking of the Culture Heritage in the Egyptian museums. *2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC)*, 1–7.

<https://doi.org/10.1109/CCWC.2017.7868440>

- Padilha, R. C., & Café, L. M. A. (2016). A interoperabilidade semântica entre acervos de museus: discutindo o caso dos Museus da Imagem e do Som. *Em Questão*, 23(1), 113. <https://doi.org/10.19132/1808-5245231.113-128>
- Palomar, I. J., García Valdecabres, J. L., Tzortzopoulos, P., & Pellicer, E. (2020). An online platform to unify and synchronise heritage architecture information. *Automation in Construction*, 110, 103008. <https://doi.org/https://doi.org/10.1016/j.autcon.2019.103008>
- Pressman, R. S. (2011). *Engenharia de Software: Uma Abordagem Profissional* (7th ed.). McGraw Hill Brasil.
- Quattrini, R., Pierdicca, R., & Morbidoni, C. (2017). Knowledge-based data enrichment for HBIM: Exploring high-quality models using the semantic-web. *Journal of Cultural Heritage*, 28, 129–139. <https://doi.org/https://doi.org/10.1016/j.culher.2017.05.004>
- Riley, J. (2017). *Understanding Metadata: What Is Metadata and What Is It For?* National Information Standards Organization (NISO). [http://www.niso.org/apps/group\\_public/download.php/17446/Understanding Metadata.pdf](http://www.niso.org/apps/group_public/download.php/17446/Understanding%20Metadata.pdf)
- Rinaldi, A. M., & Russo, C. (2018). A Matching Framework for Multimedia Data Integration Using Semantics and Ontologies. *2018 IEEE 12th International Conference on Semantic Computing (ICSC)*, 363–368. <https://doi.org/10.1109/ICSC.2018.00074>
- Rubhasy, A., Paramartha, A. A. G. Y., Budi, I., & Hasibuan, Z. A. (2014). Management and retrieval of cultural heritage multimedia collection using ontology. *2014 The 1st International Conference on Information Technology, Computer, and Electrical Engineering*, 255–259. <https://doi.org/10.1109/ICITACEE.2014.7065752>
- Rubhasy, A., Yudhi Paramartha, A. A. G., Budi, I., & Hasibuan, Z. A. (2015). Management and retrieval of cultural heritage multimedia collection using ontology. *2014 1st International Conference on Information Technology, Computer, and Electrical Engineering: Green Technology and Its Applications for a Better Future, ICITACEE 2014 - Proceedings*, 255–259. <https://doi.org/10.1109/ICITACEE.2014.7065752>
- Samreen, S., Mirza, J. S., & Rasheed, A. (2013). RDF and OWL Ontology Building of Web Applications. *Research Journal of Information Technology*, 5(4), 109–117. <https://doi.org/10.19026/rjit.5.5795>
- Santos, H. P. (2016). Modelo CIDOC CRM: interoperabilidade semântica de informações culturais. *Brazilian Journal of Information Science: Research Trends*, 10(1), 56–62. <http://revistas.marilia.unesp.br/index.php/bjis/article/view/5060/3975>

- Saracevic, T. (1995). Interdisciplinary nature of information science. *Ciência Da Informação*, 24(1), 36–41.
- SECULT. (2020). *Secretaria de Cultura de Goiás*. <https://www.cultura.go.gov.br/>
- Shi, L., & Roman, D. (2018). Ontologies for the Real Property Domain. *Proceedings of the 8th International Conference on Web Intelligence, Mining and Semantics*. <https://doi.org/10.1145/3227609.3227661>
- Silva, A. H., & Fossá, M. I. T. (2015). Análise De Conteúdo: Exemplo De Aplicação Da Técnica Para Análise De Dados Qualitativos. *Análise De Conteúdo: Exemplo De Aplicação Da Técnica Para Análise De Dados Qualitativos*, 16(1), 1–14. <https://doi.org/10.18391/qualitas.v16i1.2113>
- Silva, A. M. da. (2005). Ciência da Informação e Sistemas de Informação: (re)exame de uma relação disciplinar. *Prisma.Com*, 2–46. <http://revistas.ua.pt/index.php/prisma.com/article/view/657/pdf>
- Smirnov, A. V, Kashevnik, A. M., & Ponomarev, A. (2017). Context-Based Infomobility System for Cultural Heritage Recommendation: Tourist Assistant--TAIS. *Personal Ubiquitous Comput.*, 21(2), 297–311. <https://doi.org/10.1007/s00779-016-0990-0>
- Sommerville, I. (2011). *Engenharia de software*. PEARSON BRASIL. <https://books.google.com.br/books?id=H4u5ygAACAAJ>
- Souza, L. de, Câmara, B. de C., Rautenberg, S., & Dall'agnol, J. M. H. (2018). Enriquecendo Dados Abertos Conectados Do Qualisbrasil Com Melhores Práticas W3c Para Publicação E Consumo De Dados Na Web. *XIX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB*.
- Su, X., Sperli, G., Moscato, V., Picariello, A., Esposito, C., & Choi, C. (2019). An Edge Intelligence Empowered Recommender System Enabling Cultural Heritage Applications. *IEEE Transactions on Industrial Informatics*, 15(7), 4266–4275. <https://doi.org/10.1109/TII.2019.2908056>
- Thiolent, M. (2011). *METODOLOGIA DA PESQUISA-AÇÃO* (18th ed.). Cortez.
- Tripp, D. (2005). *Pesquisa-ação: uma introdução m etodológica*. 443–466.
- Resolução CsA 1045/2017 - Política de Extensão da Universidade Estadual de Goiás, (2017).
- Unesco. (2020). *UNESCO*. <https://whc.unesco.org/en/conventiontext/>
- Varfolomeyev, A., Korzun, D., Ivanovs, A., Soms, H., & Petrina, O. (2015). Smart Space based Recommendation Service for Historical Tourism. *Procedia Computer Science*, 77, 85–91. <https://doi.org/https://doi.org/10.1016/j.procs.2015.12.363>

- Veninata, C. (2020). Inside the Meanings. The Usefulness of a Register of Ontologies in the cultural Heritage Sector. *DIGITAL PRESENTATION AND PRESERVATION OF CULTURAL AND SCIENTIFIC HERITAGE*, 11(2), 45–58. <https://doi.org/10.4403/jlis.it-12624>
- Verslyppe, B., De Smet, W., De Baets, B., De Vos, P., & Dawyndt, P. (2014). StrainInfo introduces electronic passports for microorganisms. *Systematic and Applied Microbiology*, 37(1), 42–50. <https://doi.org/https://doi.org/10.1016/j.syapm.2013.11.002>
- Vieira, S. (2009). *Como Elaborar Questionários* (Atlas (ed.)).
- Vlachidis, A., & Tudhope, D. (2015). A knowledge-based approach to Information Extraction for semantic interoperability in the archaeology domain. *Journal of the Association for Information Science and Technology*, 67(5), 1138–1152. <https://doi.org/10.1002/asi.23485>
- W3C. (2014a). *RDF 1.1 Concepts and Abstract Syntax: W3C Recommendation 25 February 2014*. <https://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/>
- W3C. (2014b). *RDF Schema 1.1: W3C Recommendation 25 February 2014*. [https://www.w3.org/TR/rdf-schema/#ch\\_introduction](https://www.w3.org/TR/rdf-schema/#ch_introduction)
- W3C. (2021). *Linked Open Data*. <https://www.w3.org/standards/semanticweb/data>
- Xie, I., & Matusiak, K. K. (2016). *Chapter 5 - Metadata* (I. Xie & K. K. B. T.-D. D. L. Matusiak (eds.); pp. 129–170). Elsevier. <https://doi.org/https://doi.org/10.1016/B978-0-12-417112-1.00005-3>
- Zeng, M. L. (2019). Interoperability. *Knowledge Organization*, 46(2), 122–146. <https://doi.org/10.5771/0943-7444-2019-2-122.Abstract>