



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE ARTES E COMUNICAÇÃO
DEPARTAMENTO DE CIÊNCIA DA INFORMAÇÃO
CURSO DE BIBLIOTECONOMIA

JÉSSICA OLIVEIRA DA SILVA

**ONTOLOGIA DA ARQUIVÍSTICA: construção de uma ontologia de domínio com base
no texto da Lei 8.159**

RECIFE

2024

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE ARTES E COMUNICAÇÃO
DEPARTAMENTO DE CIÊNCIA DA INFORMAÇÃO
CURSO DE BIBLIOTECONOMIA

JÉSSICA OLIVEIRA DA SILVA

**ONTOLOGIA DA ARQUIVÍSTICA: construção de uma ontologia de domínio com base
no texto da Lei 8.159**

Trabalho de Conclusão de Curso apresentado ao Departamento de Ciência da Informação, do Centro de Artes e Comunicação, da Universidade Federal de Pernambuco, como requisito para a obtenção do título de Bacharel em Biblioteconomia.

Orientador(a): Prof. Dra. Márcia Ivo Braz.

RECIFE

2024

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Silva, Jéssica Oliveira da.

Ontologia da Arquivística: construção de uma ontologia de domínio com base
no texto da Lei 8.159 / Jéssica Oliveira da Silva. - Recife, 2024.
105 p. : il.

Orientador(a): Márcia Ivo Braz

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de
Pernambuco, Centro de Artes e Comunicação, Biblioteconomia, 2024.

Inclui referências, apêndices, anexos.

1. arquivística. 2. ontologias de domínio. 3. lei de arquivos. 4. Protégé. 5.
engenharia do conhecimento. I. Braz, Márcia Ivo. (Orientação). II. Título.

020 CDD (22.ed.)



Serviço Público Federal
Universidade Federal de Pernambuco
Centro de Artes e Comunicação
Departamento de Ciência da Informação

FOLHA DE APROVAÇÃO

ONTOLOGIA DA ARQUIVÍSTICA: construção de uma ontologia de domínio com base no texto da Lei 8.159

JÉSSICA OLIVEIRA DA SILVA

Trabalho de Conclusão de Curso submetido à Banca Examinadora, apresentado no Curso de Biblioteconomia, do Departamento de Ciência da Informação, da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de Bacharel em Biblioteconomia.

TCC aprovado em 13 de março de 2024

Banca Examinadora:

MÁRCIA IVO BRAZ - Orientador(a)
Universidade Federal de Pernambuco - DCI

ANTÔNIO DE SOUZA SILVA JÚNIOR - Examinador(a) 1
Universidade Federal de Pernambuco - DCI

FELIPE MOZART DE SANTANA NASCIMENTO – Examinador(a) 2
Doutor em Ciência da Informação (PPGCI/UFPE)

Dedico este trabalho à Marlene Olivia de Oliveira, minha mãe, minha
melhor amiga, e a pessoa que sempre acreditou em mim e nos meus
sonhos.

AGRADECIMENTOS

Agradeço primeiramente à minha mãe, por ser essa pessoa guerreira, de bons valores, disciplinada e perseverante, que sempre tem me apoiado em todos os aspectos da minha vida. Você foi e continua sendo minha primeira e maior incentivadora e eu tenho plena certeza que nunca teria chegado aonde estou ou me tornado a mulher que sou, se não fosse pelos seus ensinamentos, seus conselhos e a inspiração que tenho em te ver como mulher e melhor mãe do mundo. Eu amo muito você de todo o meu coração.

Agradeço aos meus amigos, a turminha do TI Tancredo, como nos chamávamos inicialmente, pois primeiramente foi a distância entre nossas casas e a UFPE que nos uniu, mas posteriormente foi a aproximação que desenvolvemos ao nos reconhecer como amigos e nos apoiarmos. Paulinha com sua paciência e compreensão absurdas, é um ser evoluído que admiro muito. Diogo, que mesmo aparentemente não demonstrando, é temperamental mas tem um coração enorme. Bia que sempre nos incentiva a pensar de forma diferente e que junto com Bea nos fazia dar muitas risadas pelas palhaçadas que faziam com seus jeitos loucos de ser. Débora, que mesmo sendo sempre organizada e certinha (capricorniana como a minha mãe), gostava de animar com o grupo e sempre fazia ou falava algo que nos tirava risadas. E Thayná, que sempre foi a animadora oficial do grupo, com seu jeito espontâneo de ser.

Agradeço a Lukas, que mesmo tendo nos aproximado mais tardiamente na carreira acadêmica, foi essencial em vários aprendizados de vida e um ombro amigo em todos os momentos que precisei. E Marília, que apesar da distância (da zona oeste para a zona sul é quase como atravessar o nosso país Pernambuco), sempre se mostrou presente e por ser uma pessoa centrada e calma, também me apoiou, incentivou e aconselhou em vários momentos essenciais. Eu amo todos vocês, meus amigos.

Agradeço aos meus queridos docentes. Antônio, eu sempre brinquei com sua insistência em participar da minha banca de TCC, mas depois foi justamente essa insistência em me ver formada que me deu forças para superar minhas dificuldades e me refazer nos momentos que precisei, para chegar ao ponto que sempre foi nosso objetivo: a conclusão da trajetória acadêmica de graduação com o diploma em mãos. Márcia, nos poucos meses que a tive como orientadora, percebi como sua postura profissional é admirável, sempre solícita e disposta a me ajudar, sua paciência e orientações também foram essenciais para a bela conclusão desta jornada. Agradeço também a Thais, pelas oportunidades que me deu como sua monitora de disciplina e por ser a pessoa que me ajudou na definição do tema do meu

trabalho. Murilo, por ter aceitado que eu participasse de um de seus projetos de PIBIC, assim me iniciando na pesquisa científica. E também André, por ter me chamado para ser sua primeira orientanda de PIBIC e pelos momentos que precisou ter muita paciência comigo.

Também agradeço a todo o DCI, com Tereza e os outros funcionários que sempre se prestaram a auxiliar os discentes quando necessário. Todos contribuíram imensamente na minha jornada como estudante e como pessoa, e por isso, sou eternamente grata!

Faço um agradecimento especial para Felipe, que entrou de cabeça no meio da minha jornada de escrita dessa pesquisa, e se mostrou uma pessoa perspicaz e incrivelmente inteligente, sempre me guiando para os melhores caminhos a seguir a fim de que este projeto fosse um sucesso. Agradeço também ao CNPq, pelas oportunidades de iniciação científica que me assistiram no meu papel como pesquisadora e me fizeram ganhar gosto pelo ofício. E agradeço também à UFPE, por milagrosamente não ter me jubilado nos 4 anos (brincadeira não foi tudo isso) que passei somente escrevendo esse texto.

Por fim, mas não menos importante, agradeço a mim mesma por apesar de ter pensado em desistir diversas vezes, nunca ter de fato seguido por este rumo, e com o incentivo e apoio de todos os mencionados acima, como entre tantos outros que não cabem nesta página, ter superado todos os obstáculos para finalmente concluir este curso e cumprir o desejo da garotinha de 13 anos de idade que sonhava em ser bibliotecária. Eu também amo você.

“Twilight, again,” he murmured.
“Another ending. No matter how perfect
the day is, it always has to end”. (Meyer,
2007, p. 495).

RESUMO

Após as revoluções tecnológicas e criação das TICs (Tecnologias da Informação e Comunicação) é inegável a explosão informacional e seus impactos na busca e disponibilização de informação no meio digital. Pensar novas formas de organizar este aglomerado informacional disponível na Web, é de extrema importância no contexto atual e um dos compromissos assumidos pela Ciência da Informação. Após reflexões nesse âmbito, o objetivo geral da presente pesquisa é elaborar uma ontologia voltada para a área da Arquivística, com base textual na Lei nº 8.159 de 8 de janeiro de 1991, que dispõe sobre a política nacional de arquivos públicos e privados. A pesquisa foi realizada através de levantamento bibliográfico e estudo de caso, e se classifica como pesquisa exploratória de natureza aplicada e inédita. A implementação da ontologia foi feita utilizando o software Protégé versão desktop 5.6.3, com base nas metodologias Methontology, Ontology Development 101, e OntoForInfoScience em linguagem OWL. A Arquivontology foi disponibilizada na sua linguagem de programação para acesso aberto em repositório do Github. Com isto, trata-se também de uma ferramenta da engenharia do conhecimento de uso por mineradores de dados, inteligências artificiais, entre outras tecnologias, para possibilitar facilidade de busca, recuperação e conseqüente entendimento sobre o tema apresentado.

Palavras-chave: arquivística; ontologias de domínio; lei de arquivos; protégé; engenharia do conhecimento.

ABSTRACT

After the technological revolutions and the creation of ICTs (Information and Communication Technologies), the information explosion and its impacts on the search and availability of information in the digital environment are undeniable. Thinking about new ways of organizing this cluster of information available on the Web is extremely important in the current context and one of the commitments made by Information Science. After reflections in this context, the general objective of this research is to develop an ontology focused on the area of Archivistics, based on the text of the Law 8,159 of January 8, 1991. The research was carried out through a bibliographic survey and case study, and It is classified as exploratory research of an applied and unprecedented nature. The implementation of the ontology was carried out using the Protégé software, desktop version 5.6.3, based on the Methontology, Ontology Development 101, and OntoForInfoScience methodologies in OWL language. Arquivontology was made available in its programming language for open access in a Github repository. Therefore, it is also a tool for use by data miners, artificial intelligence, among other technologies, to enable easy search and understanding of the topic presented.

Keywords: archival; domain ontologies; archives law; protégé; knowledge engineering.

LISTA DE FIGURAS

Figura 1 - Representação do processo de recuperação da informação	29
Figura 2 - Componentes de uma ontologia	32
Figura 3 - Classificações das ontologias	35
Figura 4 - Descritores da pesquisa usados em meta-buscadores	42
Figura 5 - Metodologias para construção de uma ontologia	49
Figura 6 - Fluxograma de Procedimentos da Arquivontology	54
Figura 7 - Requerimentos para especificação da ontologia	55
Figura 8 - Trecho da Lei 8.159/1991	58
Figura 9 - Extração Terminológica no texto da Lei 8.159/1991	67
Figura 10 - Ontologia de gestão de documentos feito no Protégé	71
Figura 11 - Taxonomia da Arquivontology no CmapTools	72
Figura 12 - Taxonomia da classe “documentação”	74
Figura 13 - Taxonomia da classe “arquivos”	75
Figura 14 - Taxonomia da classe “instituição arquivística”	76
Figura 15 - Taxonomia da classe “administração de arquivos”	76
Figura 16 - Taxonomia da classe “gestão de documentos”	76
Figura 17 - Taxonomia da classe “política nacional de arquivos”	78
Figura 18 - Versão 5.6.3 em desktop do Protégé	80
Figura 19 - Classes gerais em desktop do Protégé desktop 5.6.3	81
Figura 20 - Classes gerais em desktop do Protégé desktop 5.6.3	82
Figura 21 - Classes gerais em desktop do Protégé desktop 5.6.3	84
Figura 22 - Anotações em desktop do Protégé desktop 5.6.3	85
Figura 23 - Object Properties no Protégé desktop 5.6.3	86
Figura 24 - Data Properties no Protégé desktop 5.6.3	87
Figura 25 - Instâncias no Protégé desktop 5.6.3	87
Figura 26 - Avaliação da Arquivontology no OOPS!	88

LISTA DE QUADROS

Quadro 1 - Sintaxes possíveis a partir da linguagem OWL	38
Quadro 2 - Técnicas de Pesquisa	41
Quadro 3 - Metodologias da pesquisa	43
Quadro 4 - Metodologias para construção de uma ontologia	45
Quadro 5 - Metodologias para construção de uma ontologia	46
Quadro 6 - Metodologias para construção de uma ontologias	48
Quadro 7 - Metodologias para construção de uma ontologias	49
Quadro 8 - Procedimentos metodológicos para construção da Arquivontology	51
Quadro 9 - Representação por Triplas	59
Quadro 10 - Especificação de Requisitos da ontologia	63
Quadro 11 - Questões de Competência da Arquivontology	65
Quadro 12 - Termos extraídos x Termos traduzidos	67
Quadro 13 - Representação RDF em triplas	78

LISTA DE ABREVIATURAS E SIGLAS

AAB	Associação de Arquivistas Brasileiros
ABNT	Associação Brasileira de Normas Técnicas
AN	Arquivo Nacional
BDA	Base de Dados em Arquivística
BRAPCI	Base de Dados em Ciência da Informação
CONARQ	Conselho Nacional de Arquivos
DIBRATE	Dicionário Brasileiro de Terminologia Arquivística
DOU	Diário Oficial da União
LAI	Lei de Acesso à Informação
NBR	Norma Técnica Brasileira
OC	Organização do Conhecimento
OI	Organização da Informação
OOPS!	Ontology Pitfall Scanner!
PNAPP	Política Nacional de Arquivos Públicos e Privados
RC	Representação do Conhecimento
RDF	Resource Description Framework
RI	Representação da Informação
SciELO	Scientific Electronic Library Online
SINAR	Sistema Nacional de Arquivos
SOC	Sistemas de Representação e Organização do Conhecimento
TIC	Tecnologias da Informação e Comunicação
UAb	Repositório institucional da Universidade Aberta
VCGE	Vocabulário Controlado de Governo Eletrônico
W3C	World Wide Web Consortium
WWW	World Wide Web

SUMÁRIO

1 INTRODUÇÃO.....	13
1.2 Justificativa.....	16
1.3 Objetivos.....	17
2 ARQUIVÍSTICA.....	18
2.1 Definição do termo “Arquivo”.....	18
2.2 Arquivística ou Arquivologia?.....	19
2.3 Cientificidade da Arquivística.....	21
2.4 Origem dos Arquivos.....	21
2.5 Arquivística Brasileira.....	22
2.6 Legislação Arquivística.....	24
2.7 Lei dos Arquivos.....	25
3 ONTOLOGIAS NA REPRESENTAÇÃO DO CONHECIMENTO.....	27
3.1 Organização e Representação do Conhecimento.....	27
3.2 Sistemas de Organização e Representação do Conhecimento.....	28
3.3 O que são Ontologias?.....	30
3.4 Porque Ontologias?.....	32
3.5 Classificações de Ontologias.....	33
3.6 Ontologias de domínio no contexto da Arquivística.....	36
3.7 Linguagens para representação em ontologias.....	37
4 METODOLOGIA.....	39
4.1 Objetivos de Pesquisa.....	39
4.2 Natureza da Pesquisa.....	40
4.3 Abordagem da Pesquisa.....	40
4.4 Procedimentos da Pesquisa.....	41
5 CRITÉRIOS E COMPONENTES NA CONSTRUÇÃO DA ARQUIVONTOLOGY..	44
5.1 Metodologias para construção da ontologia.....	44
5.2 Procedimentos para construção da Arquivontology.....	51
5.3 Especificação do Domínio e Escopo (Etapa 1).....	55
5.4 Reuso de ontologias (Etapa 2).....	56
5.5 Aquisição e extração do conhecimento (Etapa 3).....	56
5.6 Conceitualização e estruturação de dados (Etapa 4).....	58
5.7 Implementação em software (Etapa 5).....	60
5.8 Avaliação da ontologia (Etapa 6).....	61
5.9 Documentalização e Divulgação (Etapa 7).....	62
6 ARQUIVONTOLOGY.....	63
6.1 Etapa 1: Especificação do Escopo e Domínio.....	63
6.2 Etapa 2: Extração e Tradução Terminológica.....	66
6.3 Etapa 3: Reuso de ontologias.....	71

6.4 Etapa 4: Conceitualização e Estruturação Taxonômica.....	71
6.5 Etapa 5: Implementação em sistema.....	79
6.6 Etapa 6: Avaliação da ontologia.....	88
6.7 Etapa 7: Documentalização e Divulgação da ontologia.....	89
7 CONSIDERAÇÕES FINAIS.....	90
REFERÊNCIAS.....	92
APÊNDICE A – TAXONOMIA DA ARQUIVONTOLOGY.....	100
APÊNDICE B – TRECHO DA ARQUIVONTOLOGY EM OWL.....	104

1 INTRODUÇÃO

A Arquivística¹, inserida nos estudos da Ciência da Informação é, desde o século XIX, um importante domínio do conhecimento prático e teórico, que contém algumas concepções como a classificação arquivística e a gestão de documentos em suporte físico e digital, os quais são frutos dos resultados do que a sociedade produz, usa, compartilha e compreende, estando sintetizados e agrupados independente do formato. Para realizar este que é o seu propósito, a Arquivística se debruça sobre a teoria, a metodologia e a prática das funções do arquivo, além dos princípios e técnicas a serem observados na produção, organização, guarda, preservação e utilização dos arquivos públicos e privados, de caráter administrativo, histórico, jurídico, entre outros (Arquivo Nacional, 2005).

No Brasil, a Arquivística começa a ter mais destaque na década de 1970, com a criação da Associação de Arquivistas Brasileiros (AAB), cujo primeiro objetivo foi determinado por elaborar um glossário com a finalidade de estabelecer a legitimação da área como um campo científico a partir da especificação de seu objeto, por meio da adequada denominação em uma terminologia; documento este que foi publicado no I Congresso Brasileiro de Arquivística (Calderon, 2013). No entanto, muito antes disso, em 1838, a criação do Arquivo Nacional (AN) denominado na época de Arquivo Público do Império, foi o que se tornou o marco histórico primordial para o desenvolvimento da prática e da teoria arquivística no país, e hoje é considerado o órgão de maior responsabilidade e destaque da área (Macedo; Lima, 2020).

Outro fator de grande importância durante este período de trajetória inicial no aumento da relevância da Arquivística brasileira, foi a elaboração dos dois projetos de lei durante os anos 1980, que viriam posteriormente a se tornar a Lei N° 8.159/1991, a qual se consolidou como uma das normativas de maior reconhecimento e significância para a legislação arquivística no país, em um movimento de reestruturação e readequação organizacional que teve como principal ator político-institucional o AN, mas que também contou com o apoio da sociedade, em uma mobilização em prol da redemocratização do país; e o Estado, tendo esse último o papel de suscitar alguma transparência a respeito dos documentos por ele produzidos (Barros, 2021).

¹ O termo utilizado na referência é Arquivologia, mas tendo sido considerado que os dois termos se tratam de descritores com similaridade semântica, e como neste texto foi preterido um termo sobre outro, como será elucidado na seção 2.2 Arquivística ou Arquivologia?, resolveu-se por utilizar o termo Arquivística.

Desse modo, o estabelecimento da Lei 8.159/1991, também conhecida como “Lei de Arquivos”, - que dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências - (Brasil, 1991), tornou-se um dos principais avanços, podendo até ser considerado o maior, no estabelecimento de mecanismos de normatização relacionados aos acervos públicos e privados de interesse social proveniente de órgãos públicos, instituições de caráter público, entidades privadas, e pessoas físicas ou jurídicas, para dar a garantia de preservação e acesso aos documentos arquivísticos produzidos pela sociedade, devido ao seu caráter administrativo, jurídico, fiscal e histórico (Mintegui; Karpinski; Barros, 2019). Além de consolidar a responsabilidade do Conselho Nacional de Arquivos (CONARQ) como o órgão detentor e responsável pela Política Nacional de Arquivos, tal como órgão central do Sistema Nacional de Arquivos (SINAR).

Hoje, devido a democratização de acesso à informação, principalmente após a publicação da Lei 12.527, de 18 de novembro de 2011, conhecida como Lei de Acesso à Informação (LAI), em conjunto com as constantes transformações digitais pelo uso de TIC's - Tecnologias da Informação e Comunicação, que impactaram diretamente nos hábitos de representação e recuperação de dados, toda a legislação arquivística brasileira referente a gestão de arquivos, normativas, deliberação de órgãos de responsabilidade, entre outros aspectos relacionados ao tema, se encontra disponível na internet, a World Wide Web (WWW), especificamente em sites do governo, além de sua publicação na base de dados do Diário Oficial da União (DOU).

Esta facilidade de acesso às informações mediante a disponibilidade online, leva ao entendimento presumido de que os dados estão prontamente dispostos às necessidades informacionais dos usuários (sociedade). No entanto, com estas mudanças e evoluções tecnológicas também surgiram novas problemáticas na organização de informações, agora em meio digital, como o acúmulo exponencial de documentos atualmente disponíveis na internet. Segundo o Senado Federal (Brasil, 2000), o site de legislação brasileira, disponível para consultas desde 28 de fevereiro de 2000, permite a realização de pesquisas textuais na base de dados de mais de 120.000 referências de normas jurídicas do Senado Federal, cadastradas desde 1940.

A partir desta informação, já desatualizada em 24 anos, percebe-se melhor como a grande massa documental em meio digital é uma adversidade gerada em contrariedade com a viabilização de acesso informacional e que ainda acaba por se desdobrar em outros diversos

problemas, gerando um efeito dominó. As consequências acarretadas inerentemente são a baixa precisão e alta revocação durante os processos de busca e recuperação de informações, em razão da dificuldade de realizar um tratamento linguístico adequado, visto que não é possível organizar adequadamente a estruturação desses dados de uma forma rápida, devido a tamanha quantidade de informações. Outra consequência ruim da situação posta, é a falta de interoperabilidade entre sistemas, já que pela falta de uma linguagem estruturada entendível por diversas máquinas além de humanos, não existe uma interligação entre o que é produzido e postado digitalmente, gerando repetição e perda de dados.

Este equívoco de entendimento fica ainda mais compreensível quando também se percebe que “atualmente existe uma diversidade de sistemas de busca e são os métodos mais utilizados na recuperação de informação na WWW²”, como afirma Deters e Adaime (2003). E a maioria dos usuários não consegue expressar bem o que deseja inferir:

A dificuldade na formulação da necessidade de informação pelo usuário, ocorre também em grande parte por [...] o usuário estar consciente que precisa da informação, mas não consegue nem sequer a sua definição em linguagem natural. Portanto, transpô-la, para a linguagem suportada pelo sistema automático de RI é muito mais difícil, consequentemente o usuário pode vir a formular uma consulta inadequada e a probabilidade do sistema retornar documentos não relevantes para a necessidade do usuário aumentar (Deters; Adaime, 2003, p. 191).

No entanto, os baixos níveis de instrução às técnicas de pesquisa ou até o total desconhecimento sobre a forma de utilização desses mecanismos de busca, não é culpa do usuário em si, especialmente levando-se em consideração buscas pela terminologia de uma área do conhecimento. Mesmo consultas em textos jurídicos, onde houve um esforço por parte dos órgãos responsáveis pela estruturação desses dados, os termos técnicos podem gerar imprecisão de sentido e acabar por confundir o usuário, uma vez que a semântica dos descritores utilizados, podem levar a ambiguidades, polissemias, entre outros ruídos da comunicação. Esta é uma questão que deve ser pensada pelos engenheiros do conhecimento que buscam estruturar a informação para facilidade de recuperação pelos indexadores de busca, ou mesmo pelos órgãos competentes a realizar intervenções neste aspecto.

Esta situação, por exemplo, não passa despercebida pelo Estado, já que com o estabelecimento do Governo Digital (evolução do e-Gov, Governo Eletrônico), houveram mudanças nos protocolos e regulamentos sobre a disposição de textos normativos em meio digital para a promoção do acesso, transparência e prestação de serviços públicos pelo poder

² WWW ou World Wide Web é a internet como a conhecemos hoje.

público para o benefício da população. Um exemplo de um desses serviços foi a criação do Vocabulário Controlado de Governo Eletrônico (VCGE), que se configura como uma lista de termos no governo federal de âmbito geral, para auxiliar a população na busca, por indexar as informações do governo de uma forma simples e entendível para o público geral (Brasil, 2019). É notável como mesmo antes da consolidação da internet como um meio essencial para as interações sociais, a Lei Complementar nº 95 de 26/02/1998, já estabelecia algumas disposições sobre a elaboração, redação, alteração e consolidação de atos normativos, de acordo com o art. 59 da Constituição Federal (Brasil, 1998).

Como exemplificado no inciso II do Art. 4º, do Marco Civil da Internet, é de responsabilidade da União a promoção “do acesso à informação, ao conhecimento e à participação na vida cultural e na condução dos assuntos públicos” (Brasil, 2014). Assim, fica a cargo de profissionais especializados e organizações do Governo incumbidas e envolvidas nos processos de produção, organização, representação, e disseminação de tais dados normativos na Web Semântica, ter a noção da indiscutível importância que o controle terminológico específico de uma área do conhecimento impõe sobre a interoperabilidade online, bem como os impactos que causam na recuperação precisa de informações e o entendimento sem conturbações de sentido pelos usuários.

Portanto, levando-se em consideração essa problemática, apresentam-se as ontologias como ferramentas de modelagem do conhecimento, tratamento terminológico e a impor a Lei de Arquivos para sua disseminação e acesso pelos usuários digitais que buscam sobre conceitos desta área do saber, foi proposto um questionamento: uma ontologia de domínio com base textual na Lei 8.159, poderia ser uma ferramenta útil em sistemas informacionais e aplicação pelos indexadores para o aperfeiçoamento de IA ao realizar recuperação com alta precisão de busca? O intuito deste trabalho é responder este questionamento.

1.2 Justificativa

A Organização e Representação do Conhecimento dentro do domínio da arquivística, especificamente os conceitos dispostos na Lei de Arquivos justifica o presente trabalho, pois demanda aos profissionais da informação a responsabilidade de criar e utilizar as ferramentas de modelagem do conhecimento - como tesouros e ontologias - para que ao realizar pesquisas com os termos relacionados ao domínio, a informação possa ser encontrada de forma mais precisa e certa ao comparar os buscadores de navegadores e sites na web, ou que consultam

textos jurídicos, pois não dispõem da utilização de descritores estruturados em ferramentas de OR e OC. Acrescenta-se à justificativa também o fato da inexistência de ontologias específicas à Arquivística: “Se, no campo teórico, ainda há espaço para novos estudos que relacionem ontologia e Arquivologia, no que tange às ferramentas e ontologias para representar a produção de documentos arquivísticos também há bastante a ser feito” (Löw, 2021). A última justificativa para a realização deste trabalho com relação ao domínio escolhido, é pela familiaridade com a prática da arquivística que a autora possui, adquirida desde sua trajetória acadêmica como nas primeiras oportunidades de experiência profissional.

1.3 Objetivos

Diante das questões levantadas, temos como objetivo geral deste trabalho elaborar uma ontologia voltada para a área da Arquivística, com base textual na Lei nº 8.159 de 8 de janeiro de 1991, que dispõe sobre a política nacional de arquivos públicos e privados.

De modo específico, temos os seguintes objetivos:

- ❖ Realizar levantamento bibliográfico sobre a Arquivística, ontologias e as metodologias para construção de ontologias que sirvam ao propósito do trabalho;
- ❖ Definir os procedimentos utilizados com o propósito da construção de uma ontologia do domínio da Arquivística;
- ❖ Definir as ferramentas utilizados com o propósito da construção de uma ontologia de domínio voltada para a temática em questão;
- ❖ Realizar análise comparativa entre os termos listados no DIBRATE e os termos extraídos da Lei dos Arquivos;
- ❖ Divulgar a ontologia em repositório de acesso aberto.

O presente trabalho está estruturado conforme o seguinte percurso: inicialmente o referencial teórico abarca a Arquivística, suas definições as principais temáticas de discussão deste universo e finalizando a parte teórica com uma seção sobre as ontologias e a representação do conhecimento. Em seguida, a metodologia descreve as particularidades da pesquisa, passando para um capítulo com os critérios e componentes da construção da Arquivontology, uma seção dedicada à Arquivontology e encerrando com as considerações finais. Fazem parte deste volume ainda o Apêndice A, com a taxonomia da Arquivontology e Apêndice B, que corresponde à Arquivontology em linguagem OWL.

2 ARQUIVÍSTICA

2.1 Definição do termo “Arquivo”

A terminologia é a determinação de uma linguagem própria utilizada por uma área técnica ou ciência, com base na especificidade de um conceito estabelecida por teóricos e profissionais técnicos. e que não pode, sob pena de graves prejuízos ao entendimento da área do saber, conviver com equívocos, polissemias e dubiedades. Segundo Bellotto (2007), não existe necessidade de dicionários ou glossários, ou seja, a terminologia é dispensável, quando um domínio do conhecimento:

“[...] consegue compor a teoria, estabelecer a metodologia e operacionalizar as práticas usando – na transmissão dos conceitos e na denominação dos objetos e das ações, enfim, na sua expressão e comunicação – tão somente os recursos da língua, do linguajar corrente, sem sombra de desvios de interpretação [...]” (Bellotto, 2007, p. 48).

A autora afirma que a questão terminológica é um grande desafio para a área da Arquivística, principalmente no que se refere o que talvez seja o seu conceito principal, o qual é a definição do termo “arquivos”:

Os problemas da falta de um instrumento terminológico para uma área técnica são evidentes e, para o caso dos arquivos, eles têm sido apontados e analisados. A verdade é que a existência do controle terminológico, como já firmamos, além de facilitar o entendimento entre os profissionais nacionais e estrangeiros, pode aumentar, pela precisão da pesquisa e das denominações, a qualidade técnica dos trabalhos, afora se constituir em instrumental útil para a formação e treinamento de candidatos ou iniciantes na profissão (Bellotto, 2007, p.48).

Estando assim ciente desta conjuntura, objetiva-se nessa seção, colaborar com a discussão do tema e determinar quais serão os conceitos utilizados neste trabalho para o termo “arquivo”, para fins de esclarecimento entre os pesquisadores da área, além de procurar facilitar a recuperação da informação quando se busca por textos sobre esta temática. Partindo de um breve resumo histórico do termo até as principais definições encontradas hoje.

Arquivo é um termo derivado do latim *archivum*, proveniente da palavra grega *archeion*, composta de dois elementos: ARKHAIOS (antigo) e EPO (dispor, ter cuidado). E possui característica polissêmica, pois a ambiguidade de significado do termo esteve presente

desde a Antiguidade Clássica, por indicar tanto o local público onde os atos escritos eram emanados por uma autoridade, como o móvel que servia para a guarda dos referidos atos (Silva, 2016). Segundo Duranti (1993 *apud* Silva, 2016), “[...] a compreensão grega clássica considerava que o lugar público incluía todos os elementos necessários para a existência do arquivo: o lugar, os documentos e o responsável pela custódia”. O principal objetivo desses antigos arquivos era possibilitar o acesso ao público dos escritos como leis, tratados políticos, listas de vencedores das Olimpíadas, entre outros.

Em contrapartida, a apresentação mais antiga da definição da palavra, afirma Silva (2016), está presente na obra jurídica romana que foi o Código de Justiniano, o qual com a afirmação do jurista Ulpiano: “*locus publicus in quo instrumenta deponuntur*” (lugar público onde os documentos são guardados) estabeleceu a definição do termo como um lugar público de evidências íntegras da memória atestada. Essa concepção se manteve até o século XIX, quando a custódia jurídica de local de preservação, acesso e proteção dos documentos “[...] cedeu lugar à outra concepção, que passou a considerar como elemento essencial do conceito de arquivo não mais o lugar de conservação, mas sim os próprios documentos custodiados pelos arquivos” (Sandri, 1968, p. 109 *apud* Silva, 2016).

Hoje, os arquivos são definidos segundo o Dicionário Brasileiro de Terminologia Arquivística do Arquivo Nacional (Arquivo Nacional, 2005) e a NBR 9578 (ABNT, 1986), como o conjunto de documentos produzidos, acumulados e recebidos por uma entidade coletiva, sendo esta de direito público ou privado, pessoa física ou jurídica, no decorrer de suas atividades, funções e estrutura, possuindo assim natureza orgânica de acumulação, e conservado ou preservado para fins de registro administrativo, de prova ou de importância histórica, independentemente do suporte de acesso em que se encontram os documentos.

2.2 Arquivística ou Arquivologia?

Um ponto de extrema importância de discussão na literatura da área, que cabe destacar neste trabalho, é outra ambivalência de sentido, agora presente entre os termos Arquivística e Arquivologia, pois desde sempre as duas palavras vêm continuamente se manifestando no vocabulário popular, bem como na própria literatura cinzenta da área para tratar sobre o mesmo assunto. A definição apresentada na NBR 9578 (ABNT, 1986, p.3) sobre a Arquivística, corresponde aos “Princípios e técnicas a serem observados na

constituição, organização, desenvolvimento e utilização dos arquivos”. Já o termo Arquivologia é definido como “Estudo, ciência e arte dos arquivos”.

É interessante notar que o Dicionário Brasileiro de Terminologia Arquivística do Arquivo Nacional (Arquivo Nacional, 2005, p. 37), só apresenta a definição de Arquivologia como “Disciplina que estuda as funções do arquivo, os princípios e técnicas a serem observados na produção, organização, guarda, preservação e utilização dos arquivos. Também chamada arquivística”. Ou seja, o termo Arquivística se torna um não-descritor de Arquivologia no documento, não sendo autorizado na indexação, pois é entendido como um sinônimo da primeira palavra.

Pode-se compreender assim, que isto decorre do fato da Arquivística estar atrelada às práticas utilizadas desde o aparecimento dos primeiros arquivos, pelas pessoas que estavam responsáveis pela custódia dos documentos ali salvaguardados. Enquanto a Arquivologia foi nomeada como a disciplina que trata das questões teóricas de gestão documental, como afirma a definição de Santos (2016), no qual a Arquivologia se categoriza como uma disciplina contemporânea que se ocupa dos estudos referentes às funções dos arquivos, bem como as técnicas de produção, organização, acesso, guarda, preservação e utilização dos acervos documentais de caráter administrativo, jurídico e histórico.

O fato de a Arquivística estar, muito antes do surgimento da ciência, consolidada como uma prática baseada em técnicas empíricas para o arranjo e conservação de arquivos, impactou a forma como é vista pelos teóricos de outras áreas. A elaboração de princípios, metodologias e até uma terminologia específica utilizada pelos profissionais da área é muito recente, pois a disciplina é relativamente moderna, e surge com o nome de Arquivologia (Herrera, 1991). Em vista disso, esses dois termos e suas diferentes interpretações de conceituação são amplamente discutidos e definidos por diversos autores desde o seu surgimento até os dias atuais. Para este trabalho em específico, o que foi levado em consideração é a prática de gestão e organização documental em arquivos públicos e privados, e assim, seguindo o pensamento dos autores previamente apresentados, o termo Arquivística, foi escolhido para representar o domínio.

2.3 Cientificidade da Arquivística

Outro tema de destaque dentro da área, que também gera grande discussão nas produções científicas, se trata da identidade da arquivologia enquanto disciplina científica. Como aponta Jardim (2012), até o século passado, o conhecimento gerado sobre Arquivística, veio a partir das práticas de gerenciamento de instituições e serviços arquivísticos. A autora afirma: “Não por acaso, boa parte desse conhecimento plasmou-se em manuais que, até 20 anos atrás, constituíam a base quase exclusiva daquilo que se entende por Arquivologia”.

Os modos do fazer arquivísticos eram registrados, comunicados e ensinados com forte ênfase nas dimensões empíricas e pragmáticas no seu caráter de execução. Essa visão predominante não impediu o amadurecimento de vários aspectos teóricos da disciplina, entretanto acabou por prejudicar severamente o status científico com a constituição de um cenário teórico com evidente densidade em seus princípios e fundamentos (Jardim, 2012).

Levando-se em conta a qualidade de procedimentos práticos no qual o termo Arquivística detém na literatura quando se realiza a comparação com o termo Arquivologia, o qual é voltado na interpretação dos autores que definiram os conceitos anteriormente expostos como uma disciplina voltada para a teoria do domínio. Em razão disso, este se torna um outro motivo pelo qual foi preferido o termo Arquivística em deferência ao termo Arquivologia. É pertinente ressaltar que também não é o objetivo deste trabalho definir a cientificidade dos dois termos, pois não cabem na proposta desta pesquisa, mas ao levantar esse questionamento e construir mesmo que minimamente o tema dentro da literatura da comunidade, é a razão da existência dessa seção.

2.4 Origem dos Arquivos

No que se refere ao surgimento dos arquivos, Reis (2006) afirma que os mais antigos de que se tenha conhecimento, remontam ao 4º milênio a. C., junto das Civilizações do Vale do Nilo e Mesopotâmia, o que corresponde ao que alguns autores determinam como o período da Idade Antiga (Arquivos dos Palácios). A divisão de períodos para o entendimento da trajetória dos arquivos varia a depender da perspectiva do autor, há aqueles que, por exemplo, desconsiderem o surgimento da escrita e determinam os registros pré-históricos como os primórdios do que entendemos hoje como arquivos.

É notável portanto, que a estruturação linear histórica de surgimento dos arquivos, bem como a própria trajetória de definição do termo, depende de vários fatores, como a dispersão historiográfica que cada autor traz em seu texto pela adoção de critérios próprios, além da dificuldade de encontrar fontes que façam essa síntese. Assim, compreender a trajetória dos arquivos não é uma tarefa fácil, o que acarreta na produção de resultados bibliográficos desiguais e fragmentados. No entanto, para este trabalho se levou em consideração a “perspectiva diacrônica” de Silva *et al* (2009), que estudam a história dos arquivos desde o berço das civilizações pré-clássicas (arquivos pré-clássicos), passando pela prática greco-romana (arquivos clássicos), até a prática medieval (arquivos medievais), e a moderna (arquivos modernos) quando, na opinião dos autores, tem-se a configuração da Arquivística como disciplina.

Em especial, os arquivos modernos surgem logo após a Segunda Guerra Mundial, com a produção e acúmulo exacerbados de documentos, adiantam-se novas mudanças na arquivística. A produção de documentos, consequência do desenvolvimento tecnológico, cresce numa progressão geométrica e é necessário que se estabeleçam parâmetros para a sua administração, ou seja, busca da economia e eficácia na produção, no uso e na destinação final dos documentos, até mesmo por problemas práticos referentes ao espaço físico para guardá-los. Torna-se impossível lidar com a massa de documentos, produzida ou acumulada pela administração pública; é pois necessário racionalizar o seu crescimento e tratamento (Cortês, 1996).

2.5 Arquivística Brasileira

Os primeiros conjuntos documentais de que se tem conhecimento e comprovação história no Brasil, adveio das mediações realizadas entre o corpo político-administrativo instaurado por Portugal, ainda no período colonial, por suas possessões ultramarina, haja vista que os arquivos eram e são produtos das funções e atividades administrativas. Desta forma, as instituições coloniais brasileiras produziam os documentos a partir das estruturas estabelecidas pela metrópole sobre sua colônia (Henrique; Silva; Dorneles, 2017).

Esta concepção é de extrema importância para o entendimento de todo o percurso histórico dos arquivos e concomitante da própria Arquivística brasileira, uma vez que grande parte da respectiva memória do país contida em documentos de caráter jurídico, administrativo ou de cunho histórico cultural, se deve ao surgimento e manutenção dos

arquivos públicos brasileiros. Como também delimita os marcos temporais de período da Arquivística no Brasil: arquivos no Brasil Colônia, seguidos de arquivos no Brasil Império, para arquivos no Brasil República, chegando ao período dos arquivos na Ditadura Brasileira e por fim, os arquivos Brasileiros atuais, como resume Cortês (1996):

No que se refere ao tratamento da questão de documentos de arquivo no Brasil, encontramos momentos distintos nas diversas Constituições, que começam com a criação do Arquivo Imperial em 1824, passando pelo descaso com relação aos arquivos ou pelas restrições em períodos de ditadura e pelas questões técnicas e metodológicas na década de 70, até chegar às discussões sobre o acesso a documentos produzidos e acumulados pela administração pública na nova constituição de 1988 (Cortês, 1996, p.43).

É justamente neste segundo período que acontecem as primeiras movimentações para a criação do Arquivo Público do Império, hoje conhecido como Arquivo Nacional, com base no artigo 70 do capítulo VI da Constituição de 1824:

Art. 70. Assignada a Lei pelo Imperador, referendada pelo Secretario de Estado competente, e sellada com o Sello do Imperio, se guardará o original no Archivo Publico, e se remetterão os Exemplares della impressos a todas as Camaras do Imperio, Tribunaes, e mais Logares, aonde convenha fazer-se publica [sic] (Brasil, 1824 *apud* Souza, 2013).

Na época foi o então deputado pernambucano Pedro de Araújo Lima, durante a Assembleia Geral Constitutiva e Legislativa instalada em abril de 1823 que votou pela criação do órgão (Costa, 2000). No entanto, somente 14 anos depois que a resolução para criação do Arquivo Público do Império seria instituída, com “[...] a dupla missão de fortalecer as estruturas do Estado emergente e consolidar a ideia do regime monárquico em um contexto continental marcado pela hegemonia das repúblicas” segundo dos Santos (2022, p. 8).

E apesar de sua inexpressiva presença ao longo da trajetória histórica do país, interpretação essa feita por Célia Leite Costa (1997 *apud* Santos, 2022), o AN é um órgão que hoje desempenha seu papel como principal instituição arquivística de toda a organização e administração de arquivos públicos na estrutura administrativa do Estado brasileiro, vinculando-se ao Conselho Nacional de Arquivos (CONARQ), o qual é a unidade mais importante do Sistema Nacional de Arquivos (SINAR). Como ficou evidente, discutir a Arquivística pelo viés histórico, é também discutir a história do AN, pois os dois estão intrinsecamente conectados. Ademais, esta não é uma tarefa fácil:

Na história da arquivologia no Brasil, as instituições arquivísticas – em particular, os arquivos públicos – ainda carecem de pesquisas mais sistemáticas. Há ainda um longo caminho para alcançarmos uma visão mais global dessa trajetória histórica, de forma a rompermos os limites do eixo Rio-São Paulo (SANTOS, 2022, p.14).

Apesar disto, espera-se que com essa pequena contribuição para a literatura da área sobre este viés histórico, seja visto e sirva como um incentivo para a criação de novos textos com pesquisas aprofundadas e com diferentes recortes de período, para aumentar o escopo de referencial sobre a história da Arquivística brasileira e estabelecer a importância dos arquivos na organização, guarda e disseminação de informações, na história do país.

2.6 Legislação Arquivística

A primeira menção em um documento oficial sobre a gestão de arquivos foi na Constituição de 1824, como mencionado anteriormente, a qual implicava a guarda do documento original dessa mesma Constituição no Arquivo Público. Esta determinação estava ligada aos estabelecimentos implementados em Portugal para lidar com a massa documental que, até então a sua colônia na época, produzia e acumulava durante suas atividades políticas, jurídicas e financeiras. Quase duzentos anos depois, na Constituição Federal (CF), de 1988, em seu inciso III do art. 23 diz atualiza esta primeira determinação com uma nova determinação “é competência comum da União, dos Estados, do Distrito Federal e dos Municípios: [...] proteger os documentos, as obras e outros bens de valor histórico, artístico e cultural, os monumentos, as paisagens naturais notáveis e os sítios arqueológicos [...]”(Brasil, 1988).

Este é o primeiro momento em que uma normativa determina uma ação relacionada à prática da Arquivística no país. Cortês também faz um recorte preciso sobre esta temática:

Bastos e Araújo (1990), dividem o tratamento constitucional da questão documental no Brasil em três períodos distintos. O primeiro período vai da outorga da Constituição política do Império até a Constituição de 1937, quando não existia uma consciência arquivística e o Arquivo Nacional funcionava apenas como um depósito de documentos. O segundo período, a partir da Constituição de 1946, quando tem início a discussão sobre os documentos enquanto acervo arquivístico. E o terceiro período a partir do processo constituinte de 1986, quando cresce o movimento para elaboração de uma Lei Nacional de Arquivos (Cortês, 1996, p.43).

Hoje, o cenário da legislação arquivística no país é bem diferente, e conta com a disponibilização em site do governo da Coletânea da Legislação Arquivística Brasileira e Correlata, produzida pelo CONARQ (2017), cujo conteúdo é o compilado de todas as leis federais, decretos federais, medidas provisórias, resoluções, portarias federais, instruções

normativas federais e as próprias resoluções do CONARQ dispostos nos poderes legislativo e judiciário, e no âmbito municipal, estadual e federal. Diante dessa multiplicidade de normativas dispostas, uma das que mais ganha destaque é a Lei 8.159/1991.

2.7 Lei dos Arquivos

A Lei Nº 8.159³, de 8 de janeiro de 1991, também conhecida como Lei de Arquivos, dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências. Seu propósito se respalda na garantia da gestão documental, acesso, e preservação a documentos de arquivos públicos e privados, independentemente da espécie, gênero, forma, formato ou suporte em que se encontram. Como instrumento de apoio à administração, à cultura, ao desenvolvimento científico e como elementos de prova e informação de valores históricos, especialmente no caso de conjuntos documentais do poder público (Barros, 2021).

Conta com 5 capítulos: “Capítulo I: Disposições Gerais”, “Capítulo II: Dos Arquivos Públicos”, “Capítulo III: Dos Arquivos Privados”, “Capítulo IV: Da Organização e Administração de Instituições Arquivísticas Públicas”, “Capítulo V: Do Acesso e do Sigilo de Documentos Públicos”, o qual foi revogado face a Lei 12.527, de 18 de novembro de 2011 (Lei de Acesso à Informação), e por fim, com as “Disposições Finais”.

No capítulo I, Disposições gerais, a lei afirma o compromisso do Estado com os documentos, ao afirmar no Art. 1º que “É dever do Poder Público a gestão documental e a proteção especial a documentos de arquivos, como instrumento de apoio à administração, à cultura, ao desenvolvimento científico e como elementos de prova e informação” (Brasil, 1991). Neste mesmo capítulo também dá a definição de arquivos e gestão de documentos, e já introduz as noções de acesso às informações de interesse particular ou de interesse coletivo ou geral contidos nos documentos públicos. No capítulo II, a lei define o que são arquivos públicos, também esclarece sobre as fases que perpassa um arquivo dentro da Teoria das Três Idades e insere os conceitos de alienabilidade e transferência, bem como põe responsabilidade sobre a eliminação de documentos públicos.

Já no capítulo III, designa o conceito e determina a gestão dos documentos de arquivos privados cujo conteúdo seja de interesse público. No capítulo IV, categoriza os tipos de arquivos por subordinação das estruturas de poder relativas aos arquivos públicos. No capítulo 5, cujo conteúdo foi revogado, discorria sobre o as determinações de acesso e sigilo

³ Disponível em: https://www.planalto.gov.br/ccivil_03/leis/L8159.htm.

de documentos públicos, e por fim, as considerações estabelece as competências e cria os órgãos de responsabilidade sobre a Política Nacional de Arquivos.

Um dos aspectos mais importantes encontrados na lei, é a criação do CONARQ, vinculado ao AN, e como órgão central do SINAR. Bem como, a promulgação de responsabilidade sobre a criação da Política Nacional de Arquivos, pelo Conselho:

Em 2002, a Lei 8.159/91 foi regulamentada pelo Decreto 4.073, estabelecendo que cabe ao CONARQ - Conselho Nacional de Arquivos definir a política nacional de arquivos públicos e privados e ao SINAR - Sistema Nacional de Arquivos implementar essa política, visando à gestão, à preservação e ao acesso aos documentos de arquivo (Pontes, 2015).

A Política Nacional de Arquivos Públicos e Privados (PNAPP), é um conjunto de planos e programas contendo ações, diretrizes, princípios norteadores, programas de trabalho, sistemas, documentos e serviços que voltados à intervenção da produção, tramitação, classificação, organização, guarda, acesso e preservação dos conjuntos de documentos de arquivos públicos e privados no Brasil. Contudo, a Lei de Arquivos não se caracteriza como a PNAPP, pois esta vai além da limitação de um texto legislativo:

Um aspecto muito frequente é confundir-se legislação arquivística com política arquivística. A legislação arquivística fornece elementos normalizadores à política arquivística, mas não é em si mesma uma política. Muitas vezes a legislação arquivística tende a ser considerada o marco zero de uma nova era arquivística. É compreensível tal expectativa já que uma legislação adequadamente concebida pode ser um poderoso instrumento a favor da gestão, uso e preservação dos arquivos. A viabilidade dessa legislação torna-se comprometida se não for simultaneamente instrumento e objeto de uma política arquivística (Jardim, 2008, p.7).

Assim, através da interpretação da autora, fica evidente como apesar de sancionar a obrigatoriedade de criação da política, além de criar tais órgão de responsabilidade, a lei dispõe sobre alguns aspectos importantes aos arquivos, mas se limita a tal. Enquanto a PNAPP tem um objetivo muito maior a cumprir. No entanto, não diminui o papel da Lei de Arquivos em definir aspectos essenciais da Arquivística e com isto, sua divulgação é essencial para encorajar a participação no desenvolvimento e implementação de normas na prática do domínio.

3 ONTOLOGIAS NA REPRESENTAÇÃO DO CONHECIMENTO

3.1 Organização e Representação do Conhecimento

Os termos Organização da Informação (OI) e Organização do Conhecimento (OC) são bastante importantes para o aporte teórico da Ciências a Informação, sendo parte tanto do seu conceito, de acordo com o determinado por alguns autores, como um de seus princípios fundamentais, além de compor parte da divisão de áreas do conhecimento da CI. E como já deve ter ficado evidente durante este trabalho, a terminologia presta um papel essencial dentro de uma área do saber, assim, ao reiterar essas discussões terminológicas além de dar dimensionamento à literatura sobre o tema, pode facilitar a busca e recuperação de pesquisas relacionadas ao assunto, auxiliando, portanto, os pesquisadores e pessoas que atuam no campo do domínio.

Em especial a OC, cujo produto desta pesquisa se insere nos estudos desse conceito e assim, tem o foco desta seção, é definida por Café e Brascher (2011) como:

A Organização do Conhecimento (OC), como campo de estudo, está fundamentada essencialmente em análises de cunho semântico. Relações semânticas são estabelecidas por meio da análise das características ou propriedades dos conceitos, as quais permitem identificar diferenças e semelhanças que evidenciam determinados tipos de relacionamentos (Café; Brascher, 2011, p. 25).

Ou seja, os objetos naturais da realidade são conceitualizados, e a partir da existência de diferenças e semelhanças entre esses ditos objetos se criam relacionamentos entre classes e conceitos. No entanto, a forma como deve ser feita a OC desses conceitos e relacionamentos dependerá exclusivamente do domínio em que se inserem em conjunto com a visão e agrupamento de critérios determinados pelo responsável da tarefa:

A organização dos conceitos em classes é um processo arbitrário, no qual selecionamos os relacionamentos a serem apresentados, segundo as peculiaridades do domínio que pretendemos representar. Ao analisar um domínio, identificamos facetas e podemos agrupar os conceitos e relacioná-los de diferentes maneiras (Café; Brascher, 2011, p. 26).

Outro ponto de relevância dentro deste tópico, é como o tipo de Sistema de Organização do Conhecimento (SOC) em que se deseja construir tem grande impacto sobre como é feita a designação entre termos e seus relacionamentos. Em resumo, a OC se destina a organizar o conhecimento de entidades naturais da realidade através de conceitos e relações, de uma forma que represente um modelo de visão limitada, tendo em vista que esses conceitos e seus relacionamentos podem ser entendidos como “[...] abstrações que selecionam

características importantes para determinada finalidade e que ignoram detalhes ou complexidades que são importantes apenas para determinados objetivos” (Sowa, 1984, p. 344 *apud* Café; Brascher, 2011, p. 26).

Na OC, existe uma divisão entre dois agrupamentos que caracterizam esta ciência, definidos como: Processos de organização e representação do conhecimento: indexação, catalogação, análise de assunto, classificação; Sistemas de organização e representação do conhecimento: ferramentas geradas de e para esses processos. Sendo estes últimos os primordiais para a efetiva organização (Hjørland, 2008 *apud* Barros; Bastos; Santos, 2022).

3.2 Sistemas de Organização e Representação do Conhecimento

Ao considerarmos a Organização do Conhecimento (OC) uma ciência que ordena a estruturação e sistematização conceitual de acordo com suas características, que podem ser definidas como elementos de herança do objeto, e nos permite o mapeamento de conceitos e classes de relacionamentos pela indicação de valores, em determinado recorte de qualquer domínio através da representação de assuntos de documentos, os SOC visam efetivar essas aplicações (Dahlberg, 2006).

Os Sistemas de Organização do Conhecimento (SOC), segundo Carlan e Medeiros (2011, p. 56) são “[...] ferramentas semânticas com vocabulários estruturados e formalizados, usadas para o tratamento e a recuperação da informação, tanto no ambiente web como no tradicional, inclusive no desenvolvimento da Web Semântica⁴”. Ainda segundo as autoras:

O termo Sistema de Organização do Conhecimento é a tradução para o português do original inglês “Knowledge Organization System” (KOS), proposto pelo Networked Knowledge Organization Systems Working Group na 1ª Conferência da ACM Digital Libraries em 1998, Pittsburgh, Pennsylvania. Assim como a sigla KOS, adotou-se o correspondente SOC em português” (Carlan; Medeiros, 2011, p. 54).

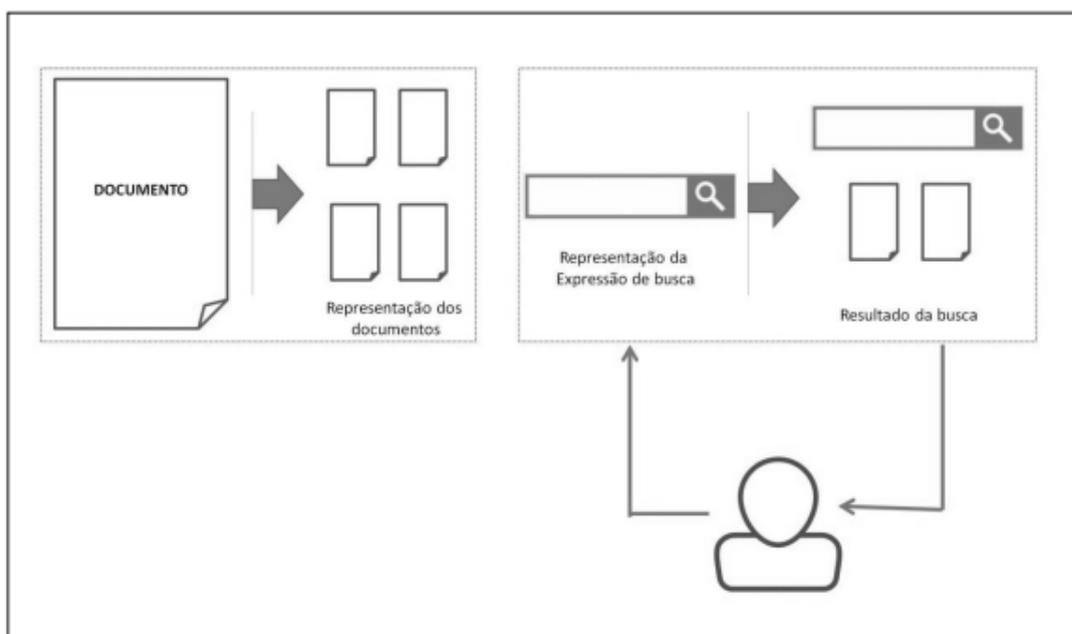
Dessa forma, através desses instrumentos que explicitam os relacionamentos semânticos de conceitos e classes e a forma como os diversos assuntos se entrelaçam através da essência dos registros produzidos por estas comunidades, é possível ter um instrumento que reflita essas perspectivas.

No âmbito da Ciência da Informação, conforme Carlan (2010), os SOC são instrumentos empregados para traduzirem os conteúdos dos documentos que se encontram

⁴ “A Web Semântica não é uma Web separada, mas sim uma extensão da Web atual, na qual a informação é dado um significado definido, favorecendo que computadores e pessoas trabalhem em cooperação” (Berners-Lee; Hendler, 2001).

estruturados em um sistema de informação, contribuindo com a sua organização. Nascimento (2018) explica que um SOC tem por objetivo a recuperação da informação, e apresenta a ilustração, Figura 1, a seguir:

Figura 1 - Representação do processo de recuperação da informação



Fonte: Nascimento (2018, p. 46).

Conforme o autor explica na Figura 1 acima, dentro de um processo de recuperação, o documento é a fonte das informações que podem ser representadas sob diferentes pontos de vista conforme o perfil de usuário que busca este material, a exemplo, nesta pesquisa, de informações sobre a Arquivística. Já a etapa de recuperação dos documentos se refere a descrição e detalhamento do conteúdo informacional de cada um desses registros de modo a facilitar a sua localização e recuperação por parte do usuário, que aparece como interagente no processo de representação da expressão de busca, conforme os termos que utilizará para sanar a sua necessidade informacional e por fim o resultado deste processo que deve apresentar os documentos com informações relevantes.

Existem diversos tipos de SOC, que dão conta de finalidades definidas, a exemplo das Taxonomias, que têm como objetivo ordenar hierarquicamente os conceitos de um determinado domínio, e os Tesouros, que busca representar as questões terminológicas de determinado campo do saber, apresentando o campo conceitual através das relações de equivalência, hierárquicas e associativas. Por conseguinte, temos as ontologias, objeto de

estudo aqui apresentado, também como um SOC, que iremos tratar mais detalhadamente adiante. Ainda sobre isto, Carlan e Medeiros (2011) apontam:

O World Wide Web Consortium (W3C⁵) desenvolveu um modelo e especificação, denominado Simple Knowledge Organization System (SKOS), para dar suporte aos tradicionais SOC (tesauro, taxonomia, vocabulário controlado e sistemas de classificação) de forma que seja entendido pela máquina. SKOS usa uma sintaxe flexível de XML/RDF que fornece estrutura para publicação dos termos usados nos SOC e seus relacionamentos para dar suporte às buscas, mapeamentos e conexões entre os diferentes SOC (Carlan; Medeiros, 2011, p. 56).

Entre os sistemas previamente mencionados, a ontologia vem cada vez mais se destacando como um Sistema de Organização e Representação do Conhecimento por capturar um quadro da representação das entidades que existem dentro de uma determinada porção do mundo em um determinado nível de generalidade, incluindo uma taxonomia dos tipos de entidades e relações que existem no mundo sob uma determinada perspectiva e por poderem ser usadas por pessoas, bancos de dados, em técnicas e aplicações de raciocínio indutivo e inferências de compartilhamento de informações sobre um dado domínio (Biagetti, 2021).

3.3 O que são Ontologias?

Esta é a tradução adaptada do título no artigo publicado por Gruber em 1993, cuja definição de ontologias se tornou talvez a mais repercutida e utilizada. O autor define ontologia como “*is a specification of a conceptualization*”, que em tradução livre: “é uma especificação de uma conceitualização”. Na Ciência da Computação, as ontologias são inferidas como ferramentas dentro da Engenharia do Conhecimento⁶ que possibilitam capturar, organizar e modelar o conhecimento e cujo conceito tem denotações diferentes quando voltados às Ciências Humanas, a exemplo da Filosofia, com a definição de “um ramo da metafísica que se preocupa com as coisas que existem” (Blackburn, 1996 *apud* Mendonça, 2015). Até mesmo a forma como é escrita foi diferenciada:

Os filósofos, habitualmente, referem-se ao termo com “O” maiúsculo e compreendem “Ontologia” como uma descrição básica das coisas que verdadeiramente existem no mundo. Por outro lado, os cientistas da informação e da computação compreendem “ontologia” como um artefato de engenharia e um componente importante de sistemas de informação, referindo-se ao termo como “o” minúsculo, a fim de diferenciá-lo da Filosofia (Mendonça, 2015, p. 51).

⁵ “O World Wide Web Consortium (W3C) é a organização que lidera o desenvolvimento de padrões para a Web. Tim Berners-Lee, fundador, prevê uma rede interligada de recursos de informação que orienta o desenvolvimento de padrões da Web e aponta o caminho para a criação de uma Web Semântica” (Brooks, 2009).

⁶ A Engenharia de Conhecimento busca capturar o conhecimento das organizações e dos indivíduos e formalizá-lo de maneira independente de domínio de modo a permitir sua apropriação e reuso em sistemas ou processos” (Abel; Fiorini, 2013, p. 2).

Mas todos esses aspectos vinculados à área do saber no qual o termo foi cunhado, acabam por congregam características junto à Ciência da Informação, de modo que Martins (2006), compreende as ontologias dentro da CI como: “[...] a conceitualização formal explícita de um conhecimento relativo a uma área ou domínio”. Ontologia como o “modelo de dados onde a informação é representada de maneira estruturada, visando agilizar a geração de conhecimento”, formada por um conjunto de elementos e os relacionamentos entre eles em termos únicos, também é um conceito com o mesmo viés apresentado por Ferreira (2020).

Já Edmond-Neal e Hlava (2005, p. 96 *apud* Pinheiro; Ferrez, 2014), define o termo ontologia dentro da CI como uma "coleção de definições de palavras e conceitos que representam uma área do conhecimento e incluem relações, propriedades e funções, formatadas para facilitar o compartilhamento da informação, sobretudo, entre computadores, constituindo um componente essencial da Web Semântica.

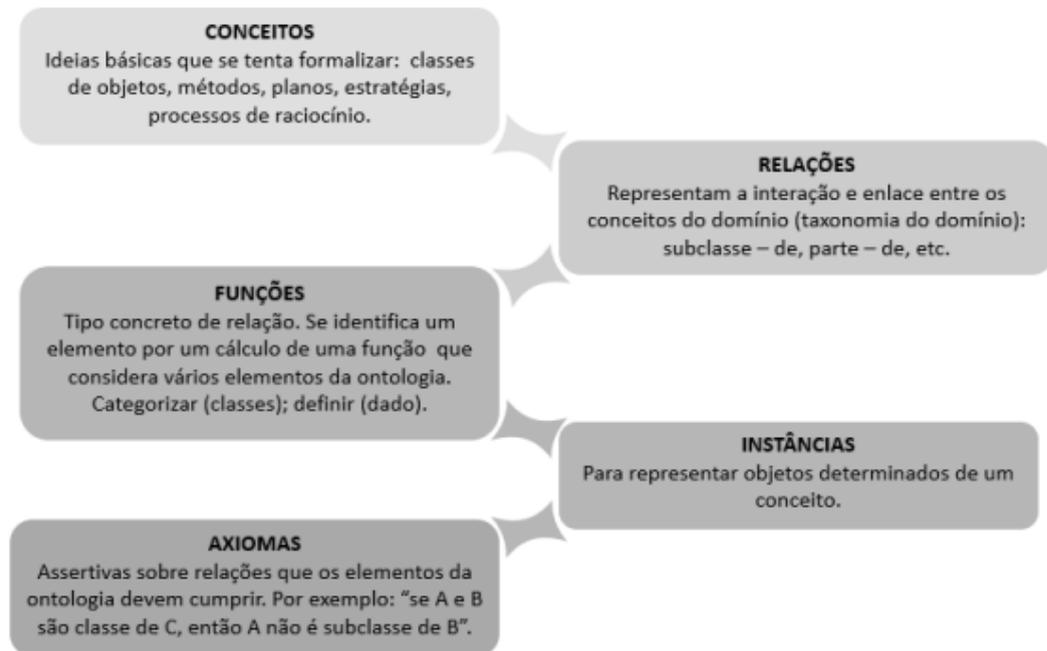
Sales e Café (2008 *apud* Nascimento 2018, p. 53) consideram que as ontologias são criadas e desenvolvidas em meio informático, com os seguintes objetivos: i) representação da realidade em um domínio, interoperabilidade de sistemas; ii) fornecimento de uma base conceitual para projetos de pesquisas e execução; iii) possibilidade de serviços baseados em operacionalizações semânticas; integração do conhecimento do conhecimento humano e os processamentos computacionais; iv) melhoria na consistência e reuso de informações e compartilhamento do conhecimento.

Moreiro González (2011, p. 77) destaca que é um sistema que:

[...] inclui definições de conceitos básicos no domínio e relacionamentos entre eles.
[...] As ontologias codificam o conhecimento sobre um domínio, chegando a codificar conhecimento que cobre mais de um domínio. Desta forma, elas fazem com que o conhecimento seja reutilizável. Então, uma ontologia é uma descrição explícita e formal de conceitos em um domínio de discurso (classes, também chamadas conceitos), propriedades de cada conceito, descrevendo várias características e atributos do conceito (slots – funções ou propriedades), e restrições sobre os slots (facetas – restrições de uma função). Uma ontologia com um conjunto de elementos de classe constitui uma base de conhecimentos.

Nesse sentido, consta que são dois os componentes necessários para a construção de uma ontologia, a saber: conceitos e relações, além de contarem com com funções, instâncias e axiomas, conforme aponta Nascimento (2022, p. 107), que também define cada componente na Figura 2 adiante:

Figura 2 - Componentes de uma ontologia



Fonte: Nascimento (2018, p. 107).

Fica explícito pela esquematização gráfica do autor, que ontologias consistem de relacionamentos complexos entre conceitos, que incluem regras e axiomas para suportar o raciocínio lógico que envolve o domínio. Ontologias formais funcionam como vocabulários controlados ou taxonomias, mas se diferenciam dos dois tipos de SOC, por possuir algo que é elementar para toda ontologia: a representação dos relacionamentos entre classes e instâncias por propriedades, necessariamente implantadas em softwares desenvolvidos para este propósito e serem entendíveis por máquinas além de humanos (Mazzocchi, 2018).

Outro aspecto importante a ser notado na ontologia é que o termo “[...] conceitualização corresponde a uma coleção de objetos, conceitos e outras entidades que se assume existirem em um domínio e os relacionamentos entre eles” (Genesereth; Nilsson, 1987 *apud* Almeida; Bax, 2003, p. 8). Como a palavra que a designa bem, ontologia deriva do grego *ontos* (ser) e *logos* (palavra), é a descrição da realidade por uma visão abstrata.

3.4 Porque Ontologias?

A Recuperação de Informações, especialmente em ambiente digital, vem crescendo exponencialmente a cada ano que passa.

Através da consulta formulada pelo usuário, o sistema de recuperação de informação é capaz de selecionar as informações (documentos) relevantes para a necessidade do

usuário. A forma utilizada pelo sistema de RI para selecionar a informação relevante é identificar a similaridade entre as informações armazenadas (índice) no sistema com a necessidade de informação descrita na expressão da consulta (Mendonça, 2015, p. 51).

Para tanto, as ontologias se mostram excelentes ferramentas de estruturação terminológica alinhado com os propósitos de busca e recuperação de informações. “Ontologias foram pensadas no contexto da Web Semântica para trabalharem associadas a sistemas de informação na web - os “web services”” (Breitman, 2005, p. 141).

Essa estruturação de dados agrega na interoperabilidade entre sistemas e também estabelece os termos que devem ser utilizados em determinadas áreas, orientando o usuário de uma SRI a utilização dos descritores definidos pelos especialistas do domínio, e quais são os relacionamentos entre os conceitos que os termos representam. O estabelecimento de mecanismos de estruturação terminológica para facilitar a organização e recuperação em meio digital com base em ferramentas tornou-se uma das principais bases do conhecimento teórico-práticas inseridos nos estudos para bibliotecários. E dentre as diversas ferramentas de SOC desenvolvidas pela Biblioteconomia, as ontologias em específico tem ganhado grande destaque nos últimos anos.

Para Mendonça (2015, p. 52), 5 motivos para os quais uma ontologia pode ter as seguintes aplicações:

- Para dar suporte à extração de informação, por exemplo, das grandes bibliotecas médicas ou da literatura científica sobre a área de medicina, que, em grande parte, encontra-se disponível em diversas fontes na web.
- Na tradução de linguagem natural, na qual as ontologias podem ser úteis para o processo de parsing e para resolver problemas de ambiguidade.
- Para a integração automática de um conjunto de vocabulários padronizados ou dicionários de dados relacionados a um domínio específico. A ontologia, neste caso, pode auxiliar na construção de um único vocabulário padrão, resolvendo um problema comum em terminologias, conhecido como fusão.
- Ontologias podem alimentar um banco de dados com informações sobre categorias (conceitos) existentes no mundo/no domínio, além das propriedades referentes a esses conceitos, bem como as relações existentes entre eles.
- Para a integração de bancos de dados, de softwares ou de modelos de negócio.

3.5 Classificações de Ontologias

As ontologias podem ser classificadas quanto a natureza de seu conteúdo, seu grau de formalismo, aplicação, entre outros aspectos. Para Guarino (1997; 1998) e Breitman (2005) existem vários tipos de classificações de ontologias quanto a natureza do que representam, os autores destacam alguns:

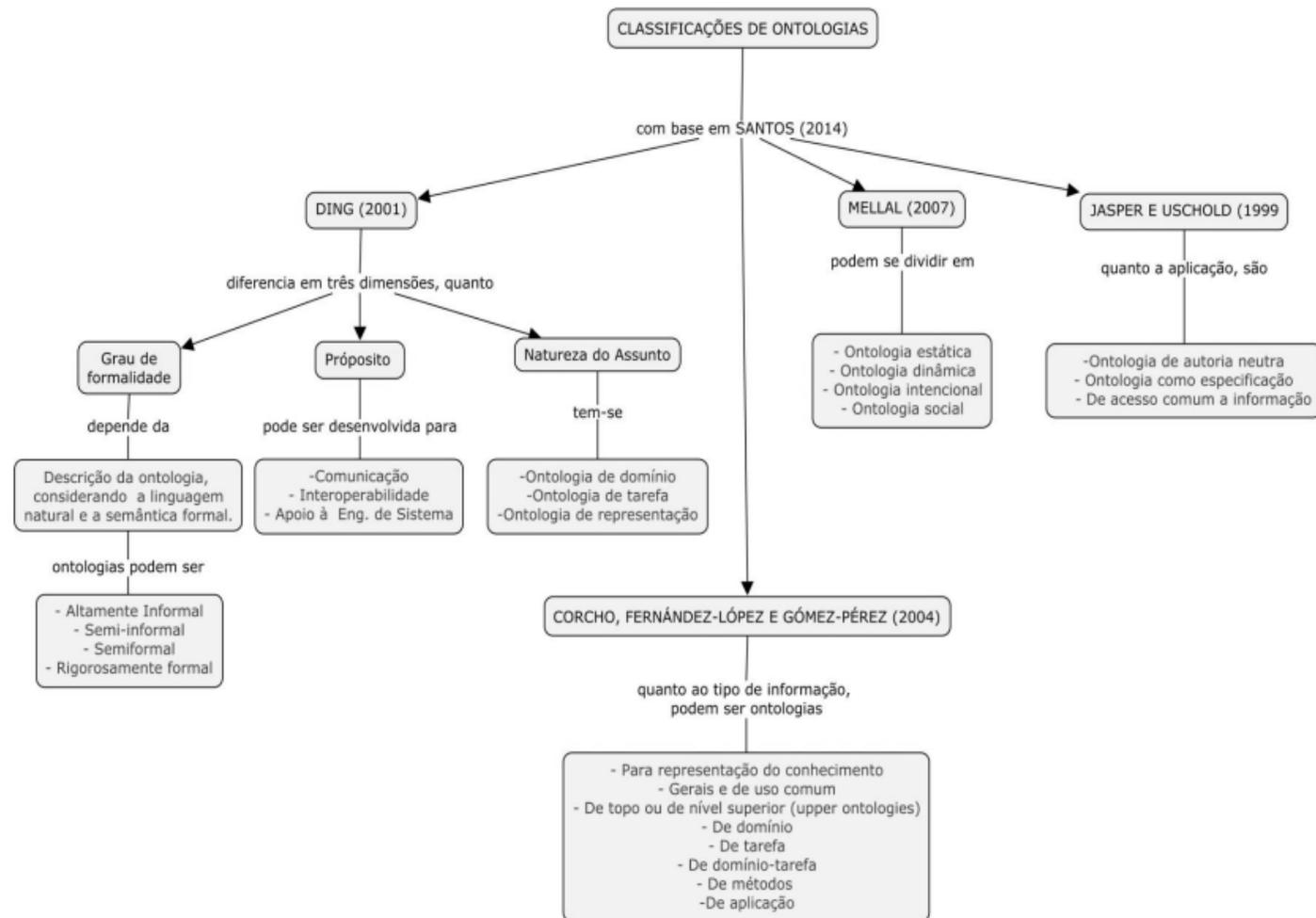
- **Ontologias Genéricas:** expressam teorias básicas do mundo, de caráter bastante abstrato, aplicáveis a qualquer domínio, conhecimento de senso comum. Tipicamente, ontologias genéricas definem conceitos tais como coisa, estado, evento, processo, ação, entre outros., com o intuito de serem especializados na definição de conceitos em uma ontologia de domínio;
- **Ontologias de Domínio:** expressam conceituações particulares, descrevendo o vocabulário relacionado a um domínio genérico tal como direito;
- **Ontologias de Tarefas:** expressam conceituações sobre a resolução de problemas, independentemente do domínio em que ocorram, isto é, descrevem o vocabulário relacionado a uma atividade ou tarefa genérica, tal como processo de vendas;
- **Ontologia de Aplicação:** expressam conceitos dependentes do domínio e da tarefa particulares. Esses conceitos frequentemente correspondem a papéis desempenhados por entidades do domínio quando da realização de uma determinada atividade;
- **Ontologia de Representação:** expressam os compromissos ontológicos embutidos em formalismo de conhecimento.

Quanto ao grau de formalidade, Uschold e Gruninger (1996), estabelecem a divisão entre: i) altamente informais, quando expressa livremente em linguagem natural; ii) semi-informais, quando expressa em linguagem natural de forma estruturada e restrita; iii) semi-formais, quando expressa em linguagem artificial definida formalmente; e iv) rigorosamente formais, onde os termos são definidos com semântica formal, teoremas e provas. Quanto às aplicações que uma ontologia pode ter:

[...] **autoria neutra**, quando, por exemplo, um aplicativo é descrito em uma única língua e depois convertido para o uso em diversos sistemas, reutilizando-se as informações; **especificação**, baseada em uma ontologia de domínio, que é utilizada para documentação e manutenção no desenvolvimento de softwares; e de **acesso comum a informação**, quando, por exemplo, um vocabulário é inacessível e a ontologia torna a informação possível de ser entendida, proporcionando conhecimento compartilhado dos termos (Morais; Ambrósio, 2007).

A diversidade de formas pelas quais as ontologias podem ser classificadas, foi mapeada e sintetizada na Figura 3 proposta por Nascimento (2018, p. 55):

Figura 3 - Classificações das ontologias



Fonte: Nascimento (2018, p. 55)

3.6 Ontologias de domínio no contexto da Arquivística

Dentro das classes de ontologias propostas pelos autores na seção anterior, as do tipo domínio possuem grande espaço de discussão e difusão no meio teórico acadêmico, assim como utilização no meio digital tendo em vista que uma das vantagens de utilizar essas “[...] ontologias integradas no ambiente de internet é a possibilidade do compartilhamento e da atualização do conhecimento em tempo real para facilitar a padronização da informação e dos procedimentos no contexto da administração pública” (Martins, 2006).

E como foi verificado durante o levantamento bibliográfico, a não existência de uma ontologia publicada sobre os conceitos que englobam o domínio da Arquivística enquanto prática e ciência. Esta prerrogativa abre espaço para o que esta pesquisa se dispõe. No entanto, uma questão que deve ficar bem esclarecida durante a produção desta pesquisa, é que apesar de ser utilizado o texto da Lei 8.159/1991, como base para a construção da ontologia, o propósito da pesquisa é criar uma ontologia de domínio voltados aos saberes da Arquivística, e não como uma ontologia de domínio sob a perspectiva jurídica⁷ do assunto tratado na lei.

Essa nuance é difícil de identificar, segundo Valente (1995 *apud* Batres; Oliveira; Viçosa; Gabrielli; Amorim; Moreira, 2005):

[...] para se construir uma ontologia sobre uma lei é necessário adotar-se uma perspectiva baseada numa visualização sistêmica que distinga o sistema do ambiente e que determine como estes dois elementos interagem. Ressalta que, ao contrário de outros domínios, como da mecânica e da medicina, no domínio das leis não é fácil distinguir qual é o sistema e qual o meio ambiente. Isto ocorre porque há dois sistemas que podem ser representados:

- O sistema legal é uma organização legal composta por vários sub-organismos, e nesta perspectiva o modelo do sistema legal conteria informações sobre a estrutura do sistema legal em termos de suas agências, estrutura interna etc.;
- O domínio da aplicação legal que é o domínio da sociedade que o sistema legal pretende regular.

Assim, através deste excerto fica evidente que esta pesquisa não se dispõe a construir uma ontologia sobre o sistema legal, pois seu viés está voltado para o domínio de aplicação da lei. Considerando que na atualidade, toda a legislação brasileira se encontra disponível para conhecimento da nação em meio digital como previamente mencionado, o aumento de ontologias do tipo se justifica, uma vez que organizar a vastidão de informação jurídica está a cada dia se tornando um desafio para os sistemas informatizados.

⁷ As ontologias legais ou ontologias jurídicas, como são reconhecidas terminologicamente no Brasil, são as chamadas ontologias cujo domínio envolve a representação do conhecimento de normativas jurídicas e legislativas, entre outros aparatos de administração e instituição de regras estabelecidas pelo Estado.

E um dos preceitos da engenharia do conhecimento é a utilização de Inteligências Artificiais (IA) para modelar o conhecimento de especialistas humanos em resolver problemas informacionais (Ramos Júnior, 2008). Porém, como mencionado brevemente antes, os objetivos deste trabalho se pautam na recuperação de informações através de metabuscadores por parte dos usuários geral sobre o domínio da Arquivística, e as ontologias se mostram a melhor ferramenta para executar tal finalidade.

3.7 Linguagens para representação em ontologias

A ontologias são esqueletos semanticamente estruturados, indexados em linguagem natural, que possui:

1) classes, organizadas em estrutura conceitual hierárquica; 2) relações definem os relacionamentos existentes entre classes e conceitos; 3) axiomas representam sentenças verdadeiras; 4) instâncias representam os elementos de um domínio associados a conceitos específicos ao domínio; 5) funções, que representam um tipo especial de relações (Brito, 2018).

E para representar estas estruturas de forma entendível para máquinas, ou seja, em linguagens de programação, existem diversas linguagens que podem ser utilizadas para expressar as ontologias, tais como: HTML, XML, RDF, RDF-S, SHOE, OIL, DAML e OWL (Breitman, 2005).

A linguagem recomendada pela W3C para construção de ontologias é a OWL, originada em linguagens de marcação, que permite a escrita, publicação e o compartilhamento de ontologias na web. Foi projetada para uso em aplicações que demandam processamento de conteúdo da informação em detrimento daquelas que apenas apresentam as informações aos humanos (W3C, 2004 *apud* Emygdio 2021).

Ainda segundo Emygdio (2021), existem duas alternativas tradicionais de representação semântica: i) OWL 2 DL e ii) OWL 2 FULL:

- OWL 2 DL (Semântica direta): provê significado para ontologias por meio da Lógica de Descrição⁸. Possui características computacionais desejáveis para sistemas de raciocínio.
- OWL 2 FULL: extensão da semântica presente em RDFS, utilizada para visualização de grafos RDF.

Em ambos os perfis, são providas sintaxes para formatos distintos nos quais as ontologias podem ser armazenadas e compartilhadas por meio de ferramentas e aplicações, são elas: RDF/XML, OWL/XML, como detalha o quadro 1 adiante:

⁸ DL (Description Logic) ou lógica de descrição “é uma área de pesquisa que estuda formalismos para representação de conhecimento [...] baseadas em uma semântica formal provendo mecanismos de inferência consistentes, completos, decidíveis, e que geram satisfazibilidade e consequência lógica” (Barker-Plummer *et al.*, 2011 *apud* Lima, 2017 p. 55).

Quadro 1 - Sintaxes possíveis a partir da linguagem OWL

Sintaxe	Especificação	Status	Objetivo
RDF/XML	Mapeamentos para grafos RDF RDF/XML	Obrigatório	Formato obrigatório, podendo ser reconhecido por qualquer <i>software</i> OWL.
OWL/XML	Serialização XML	Opcional	Processamento simples com ferramentas XML
Sintaxe Funcional	Especificação Estrutural	Opcional	Visualização simples da estrutura formal da OWL
Sintaxe Manchester	Sintaxe Manchester	Opcional	Leitura/Escrita simples de Ontologias em Lógica de Descrição
Turtle	Mapeamentos para grafos RDF Turtle	Opcional	Leitura/Escrita simples de triplas RDF

Fonte: Emygdio (2021).

Para os propósitos deste trabalho foi escolhida a linguagem OWL 2 FULL, uma vez que é a compatível com o que se propõe aqui, por ser relacionar com outras ontologias garantindo o padrão de interoperabilidade, e assim possibilitando o reuso caso sejam criadas e publicadas outras ontologias do domínio ou até de ontologias para diferentes aplicações, pois a utilização de uma linguagem padrão para expressar ontologias busca facilitar a interoperabilidade entre diferentes aplicações. Bem como por ser uma linguagem em compatível com vários softwares de criação de ontologias em OWL, bem como formatos RDF e RDFS, como é o caso do Protégé 5.6.3, que é implementado na linguagem Java e amplamente utilizado (Nascimento, 2015).

4 METODOLOGIA

Esta pesquisa busca elaborar uma ontologia do domínio da Arquivística com base textual na Lei 8.159/1991, também conhecida como Lei de Arquivos, para promover a interoperabilidade entre sistemas e facilitar a busca e recuperação de dados sobre o domínio. Desta forma, para atingir os objetivos propostos serão expostos aqui os fundamentos teóricos e procedimentos operacionais escolhidos de maneira a sustentar a discussão do problema apresentado e garantir sua validação científica, uma vez que para Gil (2002) “Como toda atividade racional e sistemática, a pesquisa exige que as ações desenvolvidas ao longo de seu processo, sejam efetivamente planejadas”.

4.1 Objetivos de Pesquisa

Para Köshe (2011), existem diversas possibilidades de pesquisa, no entanto, um trabalho deve se utilizar de apenas uma: “[...] a que leva em conta o procedimento geral que é utilizado para investigar o problema”. A partir desta afirmação é importante perceber as nuances que existem entre os métodos e procedimentos metodológicos, como é o caso do bibliográfico, e o objetivo geral da pesquisa, a qual se define como sua tipologia. Segundo esse critério, o autor distingue no mínimo três tipos de pesquisa:

A começar pela: i) bibliográfica: pesquisas que tentam explicar uma questão através do conhecimento adquirido em variadas fontes informacionais como livros, artigos, entre outros; a ii) experimental: pesquisas que buscam analisar um problema levantando hipóteses e manipulando variáveis para avaliar como se dão suas relações e quais resultados se chegam; e a iii) descritiva: pesquisas que, como o próprio nome já esclarece, visam descrever as características de uma determinada população dentro das situações e condições em que existem através de técnicas padronizadas para coleta de dados (Köshe, 2011).

Porém, existem outras tipologias relatadas em outros textos sobre metodologias, como são os casos da iv) exploratória: pesquisas que têm por finalidade apresentar e caracterizar a natureza das variáveis sobre um determinado assunto a fim de o conhecer melhor e v) explicativa: pesquisas que tem por finalidade registrar, analisar e interpretar os fenômenos estudados, buscando dessa forma identificar suas razões, fatores determinantes para sua existência e suas causas (Lakatos; Marconi, 1991).

Isto posto, é possível classificar o trabalho em questão com a metodologia de pesquisa de caráter exploratório, uma vez que envolve o levantamento e desenvolvimento de mecanismos de representação da informação através da concepção e disseminação de uma ontologia, como forma de estruturar os metadados no domínio da Arquivística e obter mais familiaridade com o domínio, levantando-se a hipótese de que a representação da informação da área pela ferramenta ontológica facilita a recuperação de dados por indexadores e a interoperabilidade entre sistemas.

4.2 Natureza da Pesquisa

Quanto à natureza de uma pesquisa, é identificada de acordo com Silveira e Córdova (2009, p. 31) como básica ou fundamental: “Objetiva gerar conhecimentos novos, úteis para o avanço da Ciência, sem aplicação prática prevista. Envolve verdades e interesses universais”, ou aplicada: “Objetiva gerar conhecimentos para aplicação prática, dirigidos à solução de problemas específicos. Envolve verdades e interesses locais”.

Assim, a natureza do presente trabalho é de caráter aplicado pois visa a geração de conhecimento para solução de um problema específico que é a estruturação de dados da Arquivística na Web Semântica, e é dirigida à busca da verdade para determinada aplicação prática em situação particular, com a criação de uma ontologia do domínio para resolução da questão. Também pode-se inferir que se trata de uma investigação original e inédita concebida pelo interesse em adquirir novos conhecimentos sobre a temática e considerando-se a inexistência de trabalhos consolidados com a perspectiva da criação de uma ontologia no domínio da Arquivística.

4.3 Abordagem da Pesquisa

A abordagem qualitativa “é baseada na interpretação dos fenômenos observados e no significado que carregam, ou no significado atribuído pelo pesquisador, dada a realidade em que os fenômenos estão inseridos. Considera a realidade e a particularidade de cada sujeito objeto da pesquisa” enquanto a abordagem quantitativa “é uma abordagem ou método que emprega medidas padronizadas e sistemáticas, reunindo respostas pré-determinadas, facilitando a comparação e a análise de medidas estatísticas de dados” (Nascimento, 2016).

Desta forma, a abordagem desta pesquisa é identificada como de aplicação qualitativa, haja visto que o trabalho se propõe a apresentar um estudo sobre a construção de uma

ontologia de um domínio específico do conhecimento. Pois a interpretação é construída com base em fontes de pesquisa da área e nas ideias de especialistas sobre o tema. Em resumo, é preciso enfatizar que os conhecimentos desenvolvidos no domínio, do ponto de vista ontológico, não podem ser mensuráveis quantitativamente.

4.4 Procedimentos da Pesquisa

Em relação aos métodos e procedimentos técnicos realizados durante uma pesquisa, são inúmeros em suas possibilidades, entretanto para uma melhor esquematização, foi criado o Quadro 2 - Técnicas de Pesquisa, para compactar e exemplificar alguns dos métodos mais conhecidos e aplicados, segundo Gil (1999), Köshe (2011) e Marconi e Lakatos (1991):

Quadro 2 - Técnicas de Pesquisa

MÉTODOS OU PROCEDIMENTOS DE PESQUISA	
TÉCNICA	DESCRIÇÃO
Estudo de Caso	“O estudo de caso é caracterizado pelo estudo profundo e exaustivo de um ou de poucos objetos, de maneira a permitir o seu conhecimento amplo e detalhado, tarefa praticamente impossível mediante os outros tipos de delineamentos considerados” (Gil, 1999).
Bibliográfica	“A pesquisa bibliográfica é desenvolvida a partir de material já elaborado, constituído principalmente de livros e artigos científicos. [...] a pesquisa bibliográfica se utiliza fundamentalmente das contribuições dos diversos autores sobre determinado assunto” (Gil, 1999, p. 50)
Experimental	“[...] o delineamento experimental consiste em determinar um objeto de estudo, selecionar as variáveis que seriam capazes de influenciá-lo, definir as formas de controle e de observação dos efeitos que a variável produz no objeto” (Gil, 1999, p. 51).
Documental	“A pesquisa documental assemelha-se muito à pesquisa bibliográfica. A única diferença entre ambas está na natureza das fontes. [...] vale-se de materiais que não receberam ainda um tratamento analítico, ou que ainda podem ser reelaborados de acordo com os objetivos da pesquisa” (Gil, 1999, p. 51)
Ex-Post-Facto	“[...] estuda as relações entre duas ou mais variáveis de um dado fenômeno sem manipulá-las. [...] não há a manipulação a priori das variáveis. É feita a constatação de sua manifestação a posteriori” (Köshe, 2011, p. 124).
Levantamento Campo/ Survey	“Consiste na observação de fatos e fenômenos tal como ocorrem espontaneamente, na coleta de dados a eles referentes e no registro de variáveis que se presume relevantes, para analisá-los” (Marconi, Lakatos, 1991, p. 186).

Fonte: Autora (2024).

Desta forma, pode-se determinar que esta pesquisa se utiliza primeiramente do método bibliográfico, visto que para cumprir o primeiro dos objetivos específicos, foi realizada uma revisão bibliográfica com os seguintes temas: ontologias e Arquivística. Nos quais são

coletadas informações sobre ambos e através dos dados coletados desenvolve-se a compreensão de maneira interpretativa destas duas áreas de estudo.

Para isto, a busca de informações se deu através de ferramentas de busca como os meta-buscadores Google Acadêmico (Scholar Google)⁹, MetaCrawler¹⁰ e Dogpile. Foram consultados textos disponíveis na internet, bases de dados e bibliotecas digitais da Ciência da Informação e da Arquivística, como a Base de Dados em Ciência da Informação (BRAPCI), Base de Dados em Arquivística (BDA), Scientific Electronic Library Online (SciELO), e revistas científicas como a Acervo - Revista do Arquivo Nacional, Springer Nature, entre outras, e repositórios como o Attena - Repositório Digital da UFPE e Repositório AbERTO - Repositório institucional da Universidade Aberta (UAb).

Livros em formato físico também foram utilizados, apesar de em menor quantidade devido aos vários contrapontos para obtenção desse tipo específico de suporte de fontes informacionais. Neste segundo caso, o local de obtenção de foi o acervo das bibliotecas central e descentralizadas do campus Recife da UFPE. Com relação aos descritores utilizados bem como as expressões de busca criadas para recuperar os documentos nos buscadores de navegadores na internet, segue figura mostrando os termos utilizados em conjunto com aspas e operadores booleanos¹¹ para formar as expressões de busca:

Figura 4 - Descritores da pesquisa usados em meta-buscadores

DESCRITORES DE PESQUISA (EQUAÇÃO DE BUSCA):
• "arquivos" AND "arquivística"
• "legislação arquivística" OR "legislação arquivística brasileira"
• "sistemas de representação do conhecimento"
• "metodologias de ontologias"
• "ontologias de domínio"
• "construção de ontologias" AND "desenvolvimento de ontologias"
• "engenharia de ontologias" OR "projeto de ontologia"
• "ontologia legal" OR "ontologia jurídica"
• "ontologies project" OR "ontology construction method"
• "ontology development" OR "ontology engineering"
• "ontologies methodology"
• "domain ontologies"
• "law ontology" OR "legal ontology development"
• "legal ontology" AND "knowledge development"

⁹ <https://scholar.google.com/>.

¹⁰ <https://www.metacrawler.com/>.

¹¹ "Utilização de operadores booleanos (delimitadores): representados pelos termos conectores AND, OR e NOT. Esses termos permitem realizar combinações dos descritores que serão utilizados na busca, sendo AND uma combinação restritiva, OR uma combinação aditiva e NOT uma combinação excludente" (Santos; Pimenta; Nobre, 2007, p.510).

Fonte: Autora (2024).

O levantamento bibliográfico sobre arquivística revelou o contexto no qual o domínio se insere, e esclareceu alguns pontos relevantes de discussão na área. Já a revisão sobre as ontologias enquanto sistema de representação do conhecimento (SOC), como também suas metodologias de construção ou aplicação colaborou para a elaboração da Arquivontology. Mesmo caso quanto às metodologias pesquisadas, já que a interpretação é desenvolvida com base na literatura da área e nas análises feitas pela autora deste trabalho.

E como segundo e último procedimento, é o estudo de caso que investiga a prática da construção de uma ontologia com o objetivo de descrever e explicar o processo praticado de construção da ontologia. Essa proposta estabelece uma criação de uma ontologia de domínio, devido ao seu objetivo que se encontra relacionado ao fornecimento de um instrumento que utilize da lista de termo especializadas da área que constituem os conceitos desse domínio, sobre seus relacionamentos, sobre as atividades e sobre princípios que os governam (Jasper; Uschold, 1999).

Segue abaixo quadro resumido para visualização de todos os detalhes metodológicos:

Quadro 3 - Metodologias da pesquisa

METODOLOGIAS DO TRABALHO			
TIPO/OBJETIVO	NATUREZA	ABORDAGEM	PROCEDIMENTOS / MÉTODOS
pesquisa exploratória	aplicado	qualitativa	bibliográfico
	inédito		estudo de caso

Fonte: Autora (2024).

5 CRITÉRIOS E COMPONENTES NA CONSTRUÇÃO DA ARQUIVONTOLOGY

5.1 Metodologias para construção da ontologia

O desenvolvimento de ontologias ocorre desde o final dos anos 1990, e estes são somente alguns exemplos das diversas metodologias encontradas atualmente para criação desta ferramenta de estruturação conceitual (Almeida, 2020). Escolher uma metodologia se tornou uma dinâmica exaustiva de seleção sobre quais as possibilidades que preenchem melhor os requisitos de escopo do que se pretende construir.

É importante ressaltar que algo chamou atenção durante a fase de revisão bibliográfica, foram os apontamentos feitos por diversos autores de construção de ontologias por bases metodológicas, desde os pioneiros nesta área até os mais recentes, mostrando como ainda é recorrente nos textos pesquisados da área, a situação em como esses métodos em sua maioria são considerados incompletos, apesar das repetidas tentativas de criação de metodologias que abrangesse todos os processos essenciais.

Portanto, ao analisar as metodologias listadas, se percebeu como as mais recentes estavam sempre tentando reutilizar seus antecessores para acrescentar etapas e tornar todo o procedimento mais completo e complexo. Ainda que esta meta, em geral, não tenha sido alcançada com êxito total. Partindo-se deste pressuposto, vias metodológicas foram analisadas para determinar quais seriam as mais adequadas para a construção da Arquivontology, uma vez que a incorporação de várias metodologias durante o processo de construção de uma ontologia, e até a modificação para customização própria, parece ser o caminho mais recomendado pelas pesquisas recuperadas.

Durante o levantamento, encontrou-se no texto de Mendonça (2015), uma ampla revisão bibliográfica sobre as principais práticas metodológicas para construção de ontologias, e esta foi compactada em um quadro, facilitando o processo de levantamento das principais metodologias de construção ontológica, e o qual foi fundamental para contemplar as principais metodologias que geralmente são utilizadas na construção de ontologias, e se estas poderiam se encaixar no escopo proposto pela Arquivontology.

Essa revisão feita pelo autor inicialmente teve como foco analisar o propósito geral de cada metodologia, para posteriormente entender as suas aplicações práticas e, principalmente, suas etapas no desenvolvimento de ontologias. Assim, segue os quadros com resumo das metodologias revisadas por Mendonça (2015):

Quadro 4 - Metodologias para construção de uma ontologia

Metodologia	Origem e Propósito Principal	Domínio de aplicação
Metodologia de Gruninger e Fox – TOVE (GRUNINGER e FOX, 1995)	Metodologia criada com base no desenvolvimento do projeto <i>Toronto Virtual Enterprise (TOVE)</i> , cujo objetivo era o de criar um modelo de senso comum ou conhecimento compartilhado sobre empresas. Esta metodologia serviu de base para o projeto e avaliação de ontologias integradas em domínios corporativos, incluindo propostas de construção de ontologias e extensões de ontologias já existentes.	Negócios (empresarial)
Metodologia de Uschold e King – ENTERPRISE (USCHOLD e KING, 1995)	Método desenvolvido com base na prática da construção da ontologia de alto nível <i>Enterprise Ontology</i> . Tem como propósito principal descrever o conhecimento sobre domínios corporativos ou de negócios.	Negócios (empresarial)
<i>Methontology</i> (GÓMEZ-PEREZ, FERNANDEZ-LOPES e VICENTE, 1996)	Metodologia desenvolvida no Laboratório de Inteligência Artificial da Universidade Politécnica de Madri entre 1996 e 1997, que possibilita a construção de uma ontologia por reengenharia sobre outra ontologia, utilizando-se do conhecimento do domínio tratado. Tal metodologia pode ser usada, segundo seus autores, em quaisquer domínios do conhecimento, embora seu uso mais conhecido e citado ocorreu na criação de uma ontologia no domínio da química.	Diversos
Método <i>Kactus</i> (BERNARAS, LARESGOTTI, CORERA, 1996)	Método recursivo derivado do projeto Kactus que permitiu a reutilização de conhecimento em sistemas de complexidade técnica, tal como o domínio de redes elétricas, e a construção de ontologias nesse domínio como suporte a tais sistemas.	Sistemas de complexidade técnica
Método <i>Sensus</i> (SWARTOUT et al., 1996).	Método derivado da ontologia <i>Sensus</i> , a qual foi desenvolvida pelo grupo <i>Information Sciences Institute (ISI)</i> com o propósito de ser usada para fins de processamento de linguagem natural. O método <i>Sensus</i> propõe alguns processos para estabelecer as ligações entre os termos específicos e os termos da ontologia de alto nível, que corresponde à ontologia <i>Sensus</i> . Na prática, o método <i>Sensus</i> foi aplicado no desenvolvimento de uma ontologia no domínio de planejamento de uma operação militar aérea.	Diversos
Método 101 (NOY e GUINNESS, 2001)	Método concebido a partir da experiência no desenvolvimento de uma ontologia de vinhos e alimentos, utilizando o editor de ontologias <i>Protégé</i> ³⁸ .	Diversos
Método CYC (REED, LENAT, 2002)	Método usado na construção da ontologia CYC, que considera o conhecimento consensual do mundo e é indicada pelos autores na criação de ontologias para fundamentar sistemas inteligentes.	Diversos

Fonte: Mendonça (2015, p. 130).

Quadro 5 - Metodologias para construção de uma ontologia

<p><i>On-to-Knowledge Methodology</i> (OTKM) (SURE, STAAB e STUBER, 2003)</p>	<p>Metodologia desenvolvida para a construção de ontologias para aplicações de gestão do conhecimento, com o foco em Processo de Conhecimento (<i>Knowledge Process</i>) e em Conhecimento do Processo Meta (<i>Knowledge Meta Process</i>). Na prática, tal metodologia foi aplicada em um estudo de caso em gestão de competências de uma empresa internacional localizada na Suíça – a <i>Swiss Life</i>.</p>	<p>Gestão do conhecimento empresarial</p>
<p>Metodologia <i>UP for ONtology</i> (UPON) (DE NICOLA, MISSIKOFF e NAVIGLI, 2009)</p>	<p>Metodologia de construção de ontologias derivada e baseada no padrão de engenharia de software conhecido como Processo Unificado – do inglês <i>Unified Software Development Process</i> ou <i>Unified Process</i> (UP) – do qual foram derivados metodologias de software como o <i>Rational Unified Process</i> (RUP) e o PProcesso para Aplicativos eXtensíveis e Interativos (PRAXIS). Apesar da <i>UP for ONtology</i> poder ser usada em diferentes domínios do conhecimento, segundo seus autores, seu uso mais comum é no domínio do <i>e-bussines</i>.</p>	<p>Negócios (<i>e-bussines</i>)</p>
<p>Metodologia <i>NeON</i> (SUAREZ-FIGUEROA, 2010)</p>	<p>Metodologia para construção de redes ontológicas baseado em um desenvolvimento colaborativo e argumentativo de ontologias. Tal metodologia foi desenvolvida em uma abordagem híbrida que combina o trabalho metodológico da área de Engenharia de Software e algumas metodologias para construção de ontologias, especificamente, a <i>Methontology</i>, a <i>On-To-Knowledge</i>, a <i>DILIGENT</i> e outros métodos ontológicos, como o de Grüniger e Fox. A <i>NeOn Methontology</i> inclui: (i) o Glossário <i>NeOn</i> de Processos e Atividades (ao todos tem-se 59 processos e atividades definidos), o qual identifica e define os processos e atividades potencialmente envolvidos quando redes ontológicas são construídas colaborativamente; e (ii) duas redes ontológicas sobre os modelos de ciclo de vida</p>	<p>Diversos</p>
<p>Metodologia <i>MFPFO</i> (LIM, LIU e LEE, 2011)</p>	<p>Metodologia de construção de ontologia multi-facetada, anotada semanticamente, para a modelagem de uma família de produtos. Tal metodologia é capaz de sugerir, automaticamente, anotações semanticamente relacionadas, baseadas no design e no repositório de construção.</p>	<p>Domínios que possuem uma família de produtos</p>
<p>Ciclo de Vida de Schiessl e Bräscher (2011)</p>	<p>Embora não seja propriamente uma metodologia de construção de ontologias, o ciclo de vida ontológico, descrito por Schiessl e Bräscher (2011), inclui todas as etapas necessárias no processo de construção de ontologias, destacando o papel de cada etapa e as tarefas contidas em cada uma delas.</p>	<p>Diversos</p>

Fonte: Mendonça (2015, p. 131).

Ao realizar um novo levantamento bibliográfico com base nas metodologias listadas no quadro do autor, ficou perceptível que existe uma preferência dos construtores de ontologias por certas metodologias mais conhecidas em meio digital para o seu desenvolvimento. Como afirma Brito (2018):

[...] é possível observar que a Methontology, desenvolvida por Gómez-Pérez, Corcho, Fernández-López (2005), a OntoForInfoScience, desenvolvida por Mendonça (2014) e a Ontology Development 101, desenvolvida por Noy e McGuinness (2008) são respectivamente as metodologias mais usadas no contexto da construção de ontologias na Ciência da Informação no Brasil.

O *Ontology Development 101: A Guide to Creating Your First Ontology* de Noy e McGuinness (2001). Este guia estabelece sete passos para a criação de uma ontologia básica, sendo estes: 1) Determinação do domínio e escopo da ontologia; 2) Consideração de reuso de ontologias existentes; 3) Enumeração de termos importantes na ontologia; 4) Definição de classes e as suas hierarquias; 5) Definição das propriedades das classes (*slots*); 6) Definição das facetas (restrições) das propriedades (*slots*) e 7) criação de instâncias.

Outra metodologia que teve grande incidência de retorno nas pesquisas feitas, foi a METHONTOLOGY de Fernández, Gómez-Pérez e Juristo (1997). Esta se divide entre quatro etapas: 1) Especificação da ontologia; 2) Aquisição do conhecimento; 3) Implementação da ontologia e 4) Avaliação da ontologia. Já a OntoForInfoScience, proposta pelo brasileiro Mendonça (2015) e criada especialmente para a CI, compreende uma pré-etapa e oito etapas para o desenvolvimento de ontologias. São elas: 1) especificação; 2) aquisição e extração de conhecimento; 3) conceitualização; 4) fundamentação ontológica; 5) formalização; 6) avaliação; 7) documentação e 8) disponibilização, e se baseia nas duas metodologias previamente mencionadas bem como na metodologia NeOn (Mendonça, 2015).

E a terceira que teve grande destaque foi a própria OntoForInfoScience, já que incorpora elementos das metodologias NeOn (Suárez-Figueroa *et al.*, 2008), METHONTOLOGY e *Ontology Development 101: A Guide to Creating Your First Ontology*. Esta metodologia adota um procedimento de 8 passos e por ser uma das mais recentes no cenário de construção ontológica, especialmente por ser voltada diretamente para os profissionais da área, como bibliotecários e arquivistas, demonstrou uma complexidade de questões levantadas em seus procedimentos, inclusive com etapas que ainda não haviam sido estabelecidas, como é o caso da etapa 0: avaliação da necessidade da ontologia.

Desse modo, a ontologia trabalhada nesta pesquisa combinou elementos metodológicos da METHONTOLOGY de (Gómez-Pérez; Corcho; Fernández-López, 2004),

OntoForInfoScience de (Mendonça, 2015), e Ontology Development 101: A Guide to Creating Your First Ontology de Noy e McGuinness (2001), haja vista a utilização das mesmas com grande ênfase nos trabalhos estudados e por seus procedimentos serem adequados ao que se propõe a pesquisa.

Como apontado por Brito (2018):

- Ontology Development 101 – contribuirá com uma metodologia clara, detalhada e simplificada de como se comporta o processo interativo para desenvolvimento de ontologias.
- OntoForInfoScience – apresenta de forma detalhada todas as etapas e cada passo utilizado para o desenvolvimento de uma ontologia.
- Methontology – possui um rico processo de documentação e verificação de ontologias, sendo eficaz no processo de construção do relatório técnico da Arquivontology.

Com relação a apresentação dos procedimentos e métodos que cada metodologia incorpora, também foi criado um quadro por Mendonça (2015), que esquematiza com riqueza de detalhes, todas as técnicas exploradas por cada uma das metodologias elencadas. E como já foram determinadas quais as metodologias que serão utilizadas na Arquivontology, seguem os seus procedimentos específicos:

Quadro 6 - Metodologias para construção de uma ontologias

<p><i>Methontology</i> (GÓMEZ-PEREZ, FERNANDEZ- LOPES e VICENTE, 1996)</p>	<ol style="list-style-type: none"> 1. Especificação: identificar o propósito da ontologia, incluindo os usuários pretendidos, cenários de uso, o grau de formalidade requerido, etc., e o escopo da ontologia, incluindo o conjunto de termos a ser representados. 2. Aquisição de Conhecimento: etapa realizada em paralelo com a etapa (1). Esta etapa é não-prescritiva e, assim, qualquer método pode ser usado, embora seja mais comum as entrevistas com especialistas e análises de textos do domínio tratado. 3. Conceitualização: identificação de termos do domínio, tais como conceitos, instâncias, relações verbais e propriedades, com cada termo tendo uma representação intermediária. 4. Integração: deve-se analisar termos de outras ontologias, tal como o padrão Ontolingua, que possam ser reutilizados na ontologia em construção. 5. Implementação: a ontologia deve ser representada formalmente através de uma linguagem formal, tais como a Ontolingua e a lógica descritiva. 6. Avaliação: são usadas técnicas baseadas nos métodos de validação e verificação dos sistemas de base de conhecimento. Há também um conjunto de diretrizes para avaliar incompletudes, inconsistências e redundâncias. 7. Documentação: recomenda-se a especificação de uma documentação em linguagem da natural ao final de cada fase do ciclo de vida do desenvolvimento da ontologia.
--	--

Fonte: Mendonça (2015, p.133).

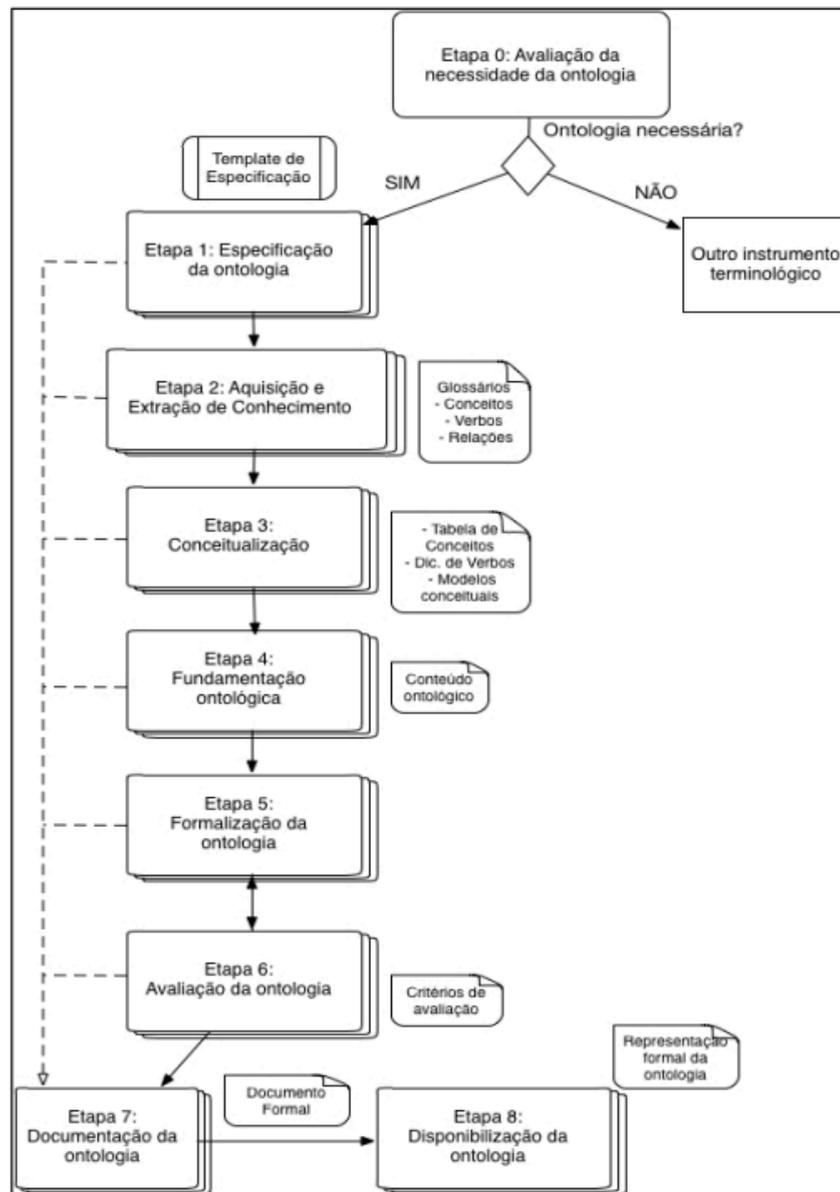
Quadro 7 - Metodologias para construção de uma ontologias

Método 101 (NOY e GUINNESS, 2001)	<ol style="list-style-type: none"> 1. Determinar escopo da ontologia; 2. Considerar o reuso de termos de outras ontologias; 3. Enumerar termos; 4. Definir classes na ontologia; 5. Organizar as classes em uma taxonomia; 6. Definir propriedades (slots) e descrever seus valores permitidos (facetadas), através de restrições; 7. Adicionar valores de slots para as instâncias, isto é, criar instâncias.
---	---

Fonte: Mendonça (2015, p.134).

A metodologia OntofoInfoScience de Mendonça (2015), segue um fluxograma definido pelo autor, como mostra a Figura 5, que contém 8 passos para a execução de seus procedimentos. Começando pela etapa 0, que define a necessidade da ontologia, a etapa 1 segue em paralelo com o que é estabelecido nas duas metodologias mencionadas anteriormente: especificação da ontologia, onde será definido o escopo e propósito da ontologia, junto com a classe de usuário, o uso pretendido, o tipo de ontologia, a delimitação do escopo, o grau de formalidade, o limite coberto pelo domínio e as questões de competência. Na etapa 2, a aquisição e extração de conhecimento. Etapa 3 se desdobra na conceitualização, com as ações de identificação, análise e negociação dos conceitos de um determinado domínio. A 4ª etapa se refere a busca por outras ontologias para a fundamentação ontológica, e a etapa 5 é a formalização da ontologia. Já a etapa 6 busca fazer uma avaliação da ontologia, a etapa 7 busca documentar todo o processo de construção ontológico e por fim, mas não menos importante, a etapa 8 disponibilizar a ontologia.

Figura 5 - Metodologias para construção de uma ontologia



Fonte: Mendonça (2015, p.134).

A descrição detalhada de todas essas etapas que incluem os métodos, técnicas e procedimentos que cada uma das metodologias propõe, de forma esquematizada e simplificada, contribuiu muito para que o entendimento sobre as construções ontológicas fosse mais rápido e prático. E assim, também auxiliou na confirmação de que as metodologias escolhidas previamente para construção da ontologia da pesquisa, poderiam de fato ser usadas na criação da Arquivontology. Para obter sucesso nos objetivos propostos por esta pesquisa, a seguir serão indicadas todas as etapas para construção da Arquivontology, com base nos procedimentos metodológicos discutidos nesta seção.

5.2 Procedimentos para construção da Arquivontology

Prosseguindo com a etapa do planejamento de criação da ontologia, adiante se encontra esquematizada o percurso de ações a serem cumpridas com base nas metodologias adotadas pelos autores abordados na seção anterior:

Quadro 8 - Procedimentos metodológicos para construção da Arquivontology

PROCEDIMENTOS METODOLÓGICOS		
ETAPA	AÇÃO	MÉTODO
1º Etapa: Especificação do domínio e escopo	Definir o domínio e escopo geral, classe de usuários, propósito geral, uso pretendido, grau de formalidade, linguagem, local de disponibilização, delimitação do escopo de cobertura, fonte de referência da extração de termos e questões de competência.	Utilizar o quadro disponível na seção 4.1 da METHONTOLOGY como modelo para responder aos questionamentos, com o acréscimo de informações pela OntoForInfoScience , que não estavam contempladas na primeira metodologia.
2º Etapa: Reuso de ontologias como requisito de interoperabilidade	Buscar fontes de referência através do levantamento bibliográfico de todas ontologias que podem servir de base para a construção desta nova ontologia, seja por similaridade de domínio ou por similaridade estrutural.	Utilizar o método de busca por repositórios específicos de ontologias para reuso como o <i>Ontolingua ontology library</i> ou o Ontobee, proposto pelo Ontology Development 101: A Guide to Creating Your First Ontology e também descrito na seção 4.4 da METHONTOLOGY e como etapa 4 da OntoForInfoScience .
3º Etapa: Aquisição de conhecimento do domínio	Realizar análise da Lei nº 8.159/1991, para executar a extração manual de todos os conceitos e inferir seus relacionamentos, com o auxílio do Dicionário Brasileiro de Terminologia Arquivística do Arquivo Nacional (Arquivo Nacional, 2005), para completar a etapa com a tradução terminológica técnica.	Utilizar o método de análise formal de textos, proposto pela etapa 2 da metodologia OntoForInfoScience e na seção 4.2 da METHONTOLOGY . O qual consiste na análise dos documentos de referência do domínio e identificação de estruturas textuais, tais como definição e afirmação, e o tipo de conhecimento que tais estruturas podem

		representar na ontologia: conceitos (classes), propriedades, instâncias, relações, etc. (Mendonça, 2015).
4º Etapa: Estruturação dos dados da ontologia	Estruturar os termos definidos para uso nesta ontologia pela etapa 2, em uma hierarquização entre classes e subclasses e instâncias. Também nesta etapa as sentenças de inferência entre os relacionamentos (propriedades), serão definidas. Uma esquematização do mapa conceitual contendo todas as classes, instâncias e propriedades deve ser criada por meio do software CmapTool na versão de ambiente web, também conhecido como CmapCloud.	Utilizar o método de listagem de termo para criação de um glossário de acordo com o passo 3, 4, 5 e 6 da Ontology Development 101: A Guide to Creating Your First Ontology , em conjunto com a seção 4.3 Conceitualização da METHONTOLOGY . E em simultaneidade com esta primeira fase, realizar a definição de classes e propriedades, presente no passo 4 e 5 da Ontology Development 101: A Guide to Creating Your First Ontology e etapa 5 da OntoForInfoScience , se utilizando do método combinado: este método utiliza estruturação top-down e bottom-up simultaneamente. A definição de propriedades através de sentenças verbais deve ser feita pelo esquema de Representação por Triplas e representado em mapa conceitual.
5º Etapa: Implementação da ontologia	Inserir os termos estruturados em ambiente digital do software Protégé, ou na versão desktop 5.6.3. Incluindo classes, sub classes, instâncias, e as propriedades definidas serão divididas entre: data properties e object properties. Também devem ser incluídos os nomes das classes, sinônimos e definições.	Utilizar o software Protégé versão de ambiente digital ou a versão desktop 5.6. para incluir todas as classes, instâncias (<i>individuals</i>) e propriedades divididas entre (<i>object properties</i> e <i>data properties</i>). Esta etapa está listada como a seção 4.3 na METHONTOLOGY , e etapa 6 do OntoForInfocScience .
6º Etapa: Avaliação da ontologia	Avaliar a ontologia pela ferramenta OOPS! (Ontology Pitfall Scanner!), de acordo com os seguintes critérios	Utilizar a ferramenta OOPS! (Ontology Pitfall Scanner!). Esta etapa está mencionada na seção 4.6 na

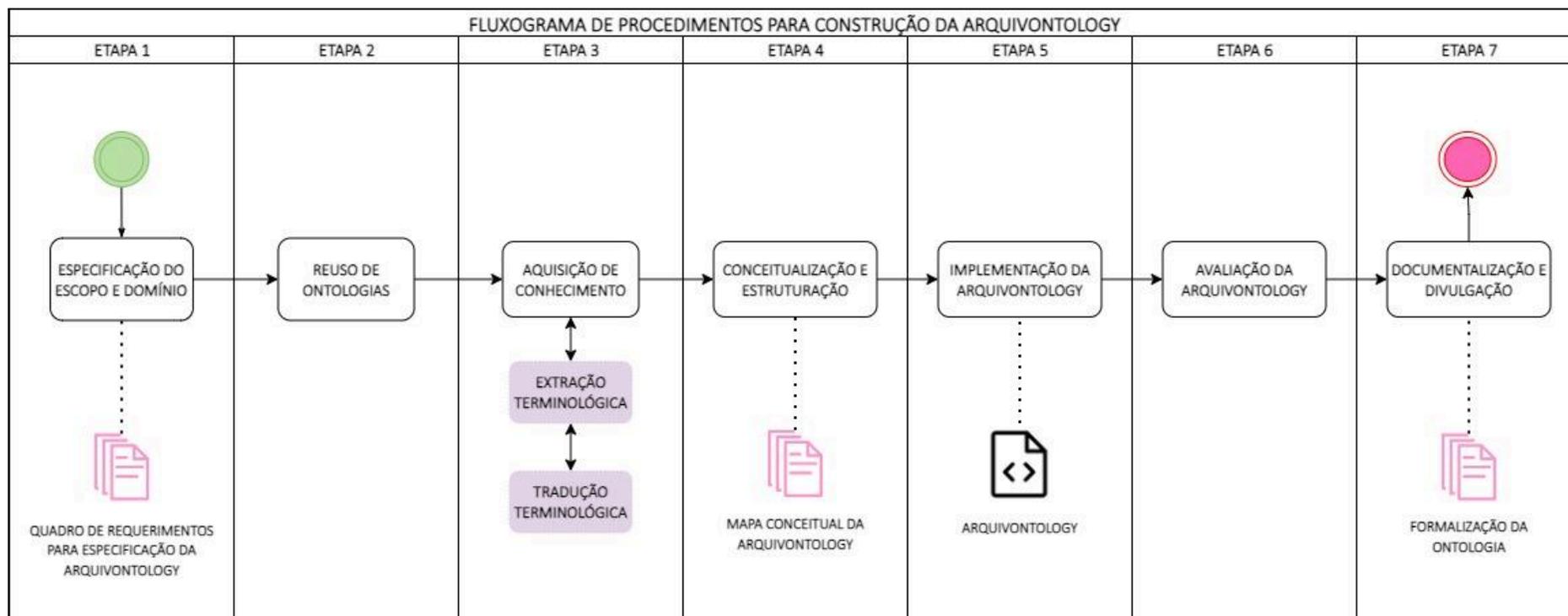
	apresentados pela ferramenta: erros menores, erros importantes e erros críticos.	METHONTOLOGY , e como a etapa 6 do OntoForInfocScience .
7º Etapa: Documentalização da ontologia e Divulgação da ontologia	Documentar todas as etapas para construção da ontologia em um documento que possa ser divulgado para acesso aberto ao público geral. Divulgar a ontologia por sua <i>mark-up language</i> de escolha, como o padrão OWL, ou outros tipos de linguagem, em um espaço de acesso aberto para recuperação de dados por qualquer tipo de indivíduo que queira reutilizar a ontologia. Como é o caso de repositórios como o Github.	Apresentar todos os resultados das ações aqui planejadas em uma seção dentro desta pesquisa, relatando todo o processo de construção da Arquivontology. Divulgar em formato aberto o código na linguagem OWL da Arquivontology, para acesso dentro de repositório criado na plataforma GitHub. Esta etapa está listada na seção 4.7 na METHONTOLOGY , e etapa 6 do OntoForInfocScience .

Fonte: Autora (2024).

Nas próximas seções deste trabalho, serão dissecadas todas as etapas vislumbradas anteriormente para melhor entendimento de todos os métodos e procedimentos que serão efetuados para construção da Arquivontology.

A seguir pode ser visualizado um fluxograma criado para resumir todas as etapas das atividades propostas com base nas metodologias mencionadas anteriormente, para a construção da Arquivontology. Bem como os documentos criados a partir destes métodos, e o produto final, o qual é a ontologia em si:

Figura 6 - Fluxograma de Procedimentos da Arquivontology



Fonte: Autora (2024).

5.3 Especificação do Domínio e Escopo (Etapa 1)

Neste primeiro estágio dentro dos processos de criação de uma ontologia, como modelo para delimitação do escopo e domínio do projeto, será utilizado o quadro da METHONTOLOGY, como mostra a Figura adiante, que determina: 1) o domínio da ontologia; 2) a data de sua produção; 3) responsável pela conceitualização; 4) responsável pela implementação; 5) o propósito da ontologia; 6) o nível de formalidade; 7) o escopo do projeto e 8) as fontes de referência para extração de dados:

Figura 7 - Requerimentos para especificação da ontologia

ONTOLOGY REQUIREMENT SPECIFICATION DOCUMENT	
Domain:	Chemicals
Date:	May, 15th 1996
Conceptualized-by:	Asunción Gómez-Pérez
Implemented-by:	Mariano Fernandez-López
Purpose:	Ontology about chemical substances to be used when information about chemical elements is required in teaching, manufacturing, analysis, etc. This ontology could be used to ascertain, e.g., the atomic weight of the element Sodium.
Level of Formality:	Semi-formal.
Scope:	List of 103 elements of substances: <i>Lithium, Sodium, Chlorine, ...</i> List of concepts: <i>Halogens, noble-gases, semi-metal, metal, ...</i> At least information about the following properties: <i>atomic-number, atomic-weight, atomic-volume-at-20-degrees-celsius, boiling-point, density-at-20-degrees-celsius, electronegativity, electron-affinity, and symbol.</i>
Sources of Knowledge:	<i>Handbook of Chemistry and Physics. 65th edition. CRC-Press, Inc. 1984-1985.</i>

Fonte: Gómez-Pérez; Corcho; Fernández-López (2004, p.37).

E além desses requisitos, serão adicionados 1) o cenário informação que a ontologia pretende ocupar; 2) o reuso de ontologias do mesmo domínio ou similares (caso existam); 3) o grupo alvo de usuários que se pretende atingir com esta ontologia; 4) o uso pretendido da ontologia; 5) o tipo de ontologia que será construído; 6) qual a linguagem de implementação escolhida; 7) qual a ferramenta (software) de implementação escolhido e por fim, 8) qual o local de disponibilização e manutenção da ontologia. Esses incrementos, que somam 16 requerimentos nesta etapa, foram somados em parte pela metodologia OntofoInfoScience, como é o caso dos itens 3, 4 e 5. Como também por adições de autoria própria.

Outra ação desta primeira etapa é a definição das questões de competência presente no passo 1 da *Ontology Development 101: A Guide to Creating Your First Ontology*, e visam perguntar sobre o propósito geral e os objetivos específicos que a ontologia é capaz de atender. A etapa 0 da metodologia *OntoForInfoScience* será incluída nesta etapa, ao invés de ser uma etapa própria, tendo em vista que acredita-se ser compatível a incorporação desse item nas questões de competência. A primeira questão de competência sobre a ontologia desenvolvida deve ser: Esta ontologia é necessária para organização de dados do domínio? E o segundo questionamento pode ser: Existe outro aparato de SOC que pode ser utilizado ao invés de uma ontologia?

5.4 Reuso de ontologias (Etapa 2)

O reuso de ontologias está determinado como o passo 2 do *Ontology Development 101: A Guide to Creating Your First Ontology*. Esta etapa também está compreendida na *METHONTOLOGY* em sua etapa 4.4 Integração, e na *OntoForInfoScience* corresponde a etapa 4, “[...] envolve a seleção e utilização de ontologias de fundamentação para a ontologia em desenvolvimento”, como afirma Mendonça (2015, p.186). Nesta etapa, devem ser levantadas todas ontologias sobre o domínio ou com características semelhantes a que se deseja criar, com o intuito de reutilizar o conhecimento que já foi organizado para integrar os sistemas de organização do conhecimento.

A busca em repositórios específicos, voltados a ontologias de reuso como é o caso do *Ontolingua Ontology Library* ou o servidor *Ontobee*¹², é recomendado como um método pelo passo 2 do *Ontology Development 101: A Guide to Creating Your First Ontology*. Em caso de não haver aparatos ontológicos criados e publicados que preencham os requisitos para reuso, a criação da nova ontologia deve ser feita desde o princípio.

5.5 Aquisição e extração do conhecimento (Etapa 3)

Esta etapa é contemplada no tópico 4.2 da *METHONTOLOGY*, apesar desta etapa ser uma atividade independente do processo de desenvolvimento ontológico. E está presente no *OntoForInfoScience* como a etapa 2. Já no *Ontology Development 101: A Guide to Creating Your First Ontology*, aparece de forma bem simplória como o passo 3: enumeração de termos

¹² Disponível em: <https://ontobee.org/>.

da ontologia, indicando somente a necessidade levantar uma lista de descritores, mas não informando como isto deve ser feito.

Já no caso das duas metodologias previamente mencionadas, os autores citam um método de análise formal e informal, onde são identificadas inicialmente as estruturas textuais do domínio, como definição, afirmação, para em seguida, distinguir o tipo de elemento a cada uma se refere, como classes, propriedades, instâncias e relacionamentos entre conceitos.

O método que deve ser adotado para construção da Arquivontology é a de análise manual e informal de textos, em que os próprios desenvolvedores visam buscar as relações entre conceitos, em que verbos podem indicar uma relação conceitual e nomes (substantivos, pronome, adjetivos). Esta técnica foi escolhida por entender que “[...] a análise humana tende a ser mais criteriosa nesse processo, devido à capacidade interpretativa inerente ao ser humano” (Mendonça, 2015).

O critério de frequência pode ser considerado para a seleção dos candidatos a termos, porém o critério de maior importância deve ser a análise semântica-lexical, em que o terminólogo responsável pela extração se utiliza do contexto de significação da especificidade um elemento linguístico, para diferenciar entre termo (unidade dos discursos especializados) e palavra (unidade da língua geral), como exemplificam as autoras:

Todos os candidatos a termo selecionados iam sendo armazenados no que denominamos estrutura conceitual. A estrutura conceitual constitui uma representação da realidade no âmbito do domínio que se toma como objeto de estudo. Essa representação procura recolher e organizar todas as ramificações que são próprias do referido domínio, de modo a refletir, em forma de esquema, a realidade da área em questão [...] .Nessa estrutura, são indicados todos os campos e subcampos nocionais contemplados no trabalho terminológico, o que permite que cada candidato a termo seja armazenado no campo nocional correspondente. Isso permite que cada termo seja determinado pela posição que ocupa no sistema nocional representado, explicitando, dessa maneira, as relações hierárquicas e não-hierárquicas entre eles. Assim, esses campos ou subcampos nocionais constituíam listas que eram submetidas aos profissionais da área para que apontassem, dentre todos os termos constantes de determinado campo, aqueles que fossem relevantes (Teline; Almeida; Aluísio, 2003).

Outra ação contemplada nesta etapa, é a tradução terminológica de todos os termos extraídos da lei. Para realizar esta tradução será usado como texto comparativo o Dicionário Brasileiro de Terminologia Arquivística (DIBRATE). Esta tradução terminológica técnica, usará o discernimento da similaridade gramatical, mas o maior critério utilizado deverá ser o semântico, observando-se as definições disponibilizadas no texto de referência. Um quadro comparativo entre os termos extraídos e os termos traduzidos pode ser criado, para facilitar a visualização de todos os componentes desta etapa.

5.6 Conceitualização e estruturação de dados (Etapa 4)

Nesta etapa deve ser utilizado o método de listagem de termo para criação de um glossário de acordo com o passo 3 da *Ontology Development 101: A Guide to Creating Your First Ontology*, em conjunto com a etapa 4.3 Conceitualização da METHONTOLOGY e etapa 3 e 5 da *OntoforInfoScience*. Em simultaneidade com a primeira fase de listagem por um glossário de termos, deve ser realizada a ação de definir as classes e propriedades, presente no passo 4 e 5 da *Ontology Development 101: A Guide to Creating Your First Ontology*, se utilizando do método combinado: este método utiliza estruturação *top-down* e *bottom-up* simultaneamente. Em que classes gerais e classes específicas vão sendo definidas ao mesmo tempo e podem ser modificadas conforme determinação dos desenvolvedores da ontologia.

Também nesta etapa as sentenças de inferência (sintaxes) entre os relacionamentos (propriedades), serão definidas através uma ferramenta de modelagem para representar o Conjunto de Relações Conceituais em um formato gráfico de Modelos Conceituais, como afirma Mendonça (2015). O método que deve ser utilizado para determinar valores possíveis para cada um dos conceitos contidos na lista de termos da *Arquivontology* é a Representação por Triplas através da análise semântica-lexical feita pelos desenvolvedores. Neste procedimento, os conceitos são divididos entre sujeito (classe), predicado (propriedade) e objeto (subclasse ou instância).

Após conclusão da etapa anterior de extração e tradução, as frases contidas na lei serão revisadas para inferência e entendimento de seu contexto semântico, para auxiliar na proposição dos verbos que irão compor o esquema de Representação por Triplas.

Figura 8 - Trecho da Lei 8.159/1991



Subchefia para Assuntos Jurídicos

LEI Nº 8.159, DE 8 DE JANEIRO DE 1991.

[Regulamento](#)

Dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências.

[Vide Decreto nº 4.553, de 27.12.02](#)

O PRESIDENTE DA REPÚBLICA, faço saber que o Congresso Nacional decreta e eu sanciono a seguinte Lei:

CAPÍTULO I

DISPOSIÇÕES GERAIS

Art. 1º - É dever do Poder Público a gestão documental e a proteção especial a documentos de arquivos, como instrumento de apoio à administração, à cultura, ao desenvolvimento científico e como elementos de prova e informação.

Art. 2º - Consideram-se arquivos, para os fins desta Lei, os conjuntos de documentos produzidos e recebidos por órgãos públicos, instituições de caráter público e entidades privadas, em decorrência do exercício de atividades específicas, bem como por pessoa física, qualquer que seja o suporte da informação ou a natureza dos documentos.

Fonte: Autora (2024)

A análise sintática que pode ser feita através do trecho de texto da lei exposto na figura 8, levando-se em consideração os termos que foram delimitados na etapa anterior, é estabelecida em quadro contendo a Representação Tripla, da seguinte forma:

- Consideram-se arquivos para os fins desta Lei, os conjuntos de **documentos**.

Quadro 9 - Representação por Triplas

REPRESENTAÇÃO POR TRIPLAS		
SUJEITO	PREDICADO	OBJETO
arquivos	são compostos por	documentos

Fonte: Autora (2024).

Este método foi escolhido uma vez que na linguagem RDF (Resource Description Framework) e RDF(S) ou RDF-Schema, expressões são formadas através de declarações conhecidas como triplas, onde o sujeito representa as “coisas” reais ou abstratas do universo, o predicado descreve as relações entre sujeitos e o objeto representam especificações do sujeito. Este esquema busca tornar a semântica da linguagem natural (LD) compreensível para máquinas, através dos marcadores que utiliza, baseados no padrão Dublin Core¹³. Este acréscimo de metadados descrito por uma sentença verbal, deve ser usado para criar as propriedades relativas aos elementos (classes) e entidades (instâncias) da ontologia.

¹³ O Dublin Core é um padrão de metadados, composto por 15 elementos de descrição de conteúdo, e foi planejado para facilitar a representação do conhecimento de recursos eletrônicos.

Para realizar a representação gráfica, na qual humanos podem discernir a informação de maneira mais fácil e coerente, é criada uma estrutura conhecida como “mapa conceitual”, os termos são geralmente dispostos em figuras ovais e os seus relacionamentos são representados por flechas nomeadas. Já para que máquinas (computadores) possam interpretar os dados contidos em ontologias, são feitas representações em formato de triplas (relação entre conceitos): sujeito (conceito inicial), predicado (distinção entre relacionamentos e objeto (conceito ao qual o conceito inicial se refere) (Ferreira, 2020). Assim, apesar de estabelecer uma esquematização por meio do software CmapTools, em ambiente web, também chamado de CmapCloud, deve ser criada, para apresentar a representação gráfica.

5.7 Implementação em software (Etapa 5)

Para implementação da ontologia, o software escolhido utilizando-se ambiente web foi a ferramenta Protégé. O Protégé é uma das aplicações mais citadas nos trabalhos de referência, para construção de ontologia, por ser um ambiente interativo de carácter intuitivo para construção multiplataforma, com interface gráfica para edição de ontologias em diversas linguagens de marcação semântica, inclusive a OWL, com o uso das ferramentas RDF(S) e XML. O valor adicionado às propriedades, classes e instâncias criadas na ontologia devem ser em XSD: String, para representar os elementos textuais descritos na ontologia.

Todos os termos devem ser escritos em português, sem acentuação, com letra minúscula e com traços para marcar o espaçamento entre termos que possuem duas palavras. Com base na estruturação definida na etapa de Conceitualização e Estruturação, os termos devem ser dispostos no sistema entre: classes, propriedades (*object properties* e *data properties*), instâncias (*individuals*). A divisão de propriedades dentro do sistema deve considerar o seguinte critério: *object properties* são as preposições verbais criadas com o auxílio da diagramação por Representação em Triplas e dentro do Modelo Conceitual. Já *data properties*, advém do glossário de termos criado na etapa 2, que representam propriedades de classes, como elementos ou propósitos em suas definições expostas no DIBRATE.

Serão criadas anotações dentro do sistema para todas as classes e subclasses, as anotações terão as seguintes instâncias de propriedades:

- ★ **rdf:label:** que define o nome do objeto representado, para a linguagem de programação esta seria a “etiqueta” que identifica o termo;

- ★ **definition:** que agrega a ontologia, a informação sobre a definição que todas as classes e subclasses possuem. Essas definições foram extraídas do Dicionário de Terminologia Arquivística, ou no caso dos termos inclusos, foram criadas novas definições com base no texto do DIBRATE. A segunda anotação foi criada pela desenvolvedora da ontologia;
- ★ **synonyms:** também foi criada na ontologia, com o propósito de definir os sinônimos de um termo.

As propriedades de classes *object properties*, devem ser vinculadas às classes e subclasses, seguindo o modelo estabelecido no mapa conceitual criado no CmapTools. Em vista disso, a possibilidade de leitura incorreta pelo sistema, é uma possibilidade aceitável dentro do presente escopo de ontologia.

5.8 Avaliação da ontologia (Etapa 6)

A avaliação está descrita como a etapa 4.6 na METHONTOLOGY, e etapa 6 do OntoForInfocScience. Em que após o desenvolvimento ontológico, esta etapa tem o propósito de avaliar a ontologia buscando verificar se os conceitos foram inseridos de forma correta pela avaliação do sistema e assim, as questões de competência levantadas foram contempladas. Para esta ação será utilizada a ferramenta OOPS!¹⁴ (*Ontology Pitfall Scanner!*), de acordo com as orientações de Brito (2018). o código em OWL da Arquivontology será gerado e inserido por entrada direta para a ferramenta realizar a análise.

Em caso de erros, a desenvolvedora deverá notificar no documento gerado pela etapa final de Documentalização e Divulgação da ontologia (etapa 7). Caso os erros não comprometam o propósito geral da ontologia, a mesma poderá ser disponibilizada, caso o erro vá contra o propósito geral da ontologia, a mesma não deverá ser disponibilizada e será sugerida reanálise (Brito, 2018). Os erros serão analisados de acordo com o grau determinado pela ferramenta, sendo estes: 1) erros menores, podem ou não ser corrigidos (são erros de fácil correção) e não comprometem a ontologia; 2) erros importantes, devem ser corrigidos mas

¹⁴ OOPS! É uma ferramenta baseada na web, independente de qualquer ambiente de desenvolvimento de ontologias, para detectar possíveis armadilhas que possam levar a erros de modelagem. Esta ferramenta tem como objetivo auxiliar os desenvolvedores de ontologias durante a atividade de validação de ontologias, que pode ser dividida em diagnóstico e reparo. Atualmente, OOPS! fornece mecanismos para detectar automaticamente uma série de armadilhas, ajudando assim os desenvolvedores na atividade de diagnóstico (Poveda-Villalón; Suárez-Figueroa, 2012). Disponível em: <https://oops.linkeddata.es/>.

não comprometem a ontologia mas devem ser corrigidos e 3) erros críticos, que devem ser investigados posteriormente de forma aprofundada, e podem comprometer o propósito da Arquivontology.

5.9 Documentalização e Divulgação (Etapa 7)

Esta seção agrega as duas ações finais da construção da Arquivontology e correspondem a etapa 4.7 da METHONTOLOGY e etapa 7 da OntoForInfocScience. A documentação da Arquivontology de acordo com todas as etapas aqui apresentadas na seção 6 desta pesquisa. Com relação à divulgação, cuja metodologia está descrita na 8 da OntoForInfocScience, e como afirma Mendonça (2015):

[...] a disponibilização da ontologia em um algum meio eletrônico de fácil acesso e visualização aos usuários desta ontologia. Essa é uma medida fundamental para disseminar o conhecimento representado na ontologia e torná-la um instrumento útil para organização e uso da informação do domínio que ela representa.

Inicialmente seria feita a disponibilização do código em OWL da Arquivontology em formato aberto, em que não há proprietário, cuja especificação esteja documentada publicamente e seja de livre conhecimento e implementação, livre de patentes ou qualquer outra restrição legal quanto a sua utilização, no e-GoV - Repositório de Vocabulários e Ontologias de Governo Eletrônico, local para acesso a todas as referências ontológicas do Governo Eletrônico Federal, já que o texto utilizado como base foi uma lei ordinária. Porém, o link para o para acesso ao E-VoG não se encontra atualmente disponível, assim, deverá ser criado um repositório no Github¹⁵, para disponibilização da ontologia.

¹⁵ O GitHub é uma plataforma de hospedagem de código-fonte e arquivos.

6 ARQUIVONTOLOGY

De acordo com a metodologia de construção propostas e procedimentos para criação da Arquivontology, nesta seção serem explanados os dados que correspondem à documentação formal da ontologia, onde todas as ações registradas durante o processo de construção serão elencadas, de forma a melhorar sua clareza e facilitar sua utilização, manutenção e reutilização. Para cumprir com tais propósitos, optou-se por separar em um único capítulo toda a documentação formal da ontologia, e subdividi-lo de acordo com cada uma das etapas executadas na criação da Arquivontology.

6.1 Etapa 1: Especificação do Escopo e Domínio

Seguindo os procedimentos operacionais para avaliação da questão exposta neste trabalho, e por consequente, a construção de uma ontologia do domínio da Arquivística com base no texto da Lei 8.159/1991 (Lei de Arquivos), seguimos para a primeira meta proposta na metodologia, a qual é a definição do escopo e domínio. De acordo com o que foi exemplificado no item 4.1 do texto de Fernández, Gómez-Pérez e Juristo (1997), a METHONTOLOGY, em conjunto com o primeiro e segundo passos da Ontology Development 101: A Guide to Creating Your First Ontology de Noy e McGuinness (2001), através de modificações de autoria própria, foi criado o Quadro 10 - Especificação dos Requisitos da Arquivontology, para elucidar as primeiras questões levantadas:

Quadro 10 - Especificação de Requisitos da ontologia

ESPECIFICAÇÃO DE REQUISITOS DA ARQUIVONTOLOGY	
Domínio:	Arquivística
Data:	2024
Conceitualizada por:	Jéssica Oliveira da Silva (UFPE)
Implantada por:	Jéssica Oliveira da Silva (UFPE)
Propósito:	Ontologia com o propósito de representar as relações entre os conceitos do domínio da Arquivística através de um controle terminológico e uma estruturação taxonômica para aplicação em meio digital. Isto auxiliará os algoritmos de busca e recuperação de informações na web; possibilitará

	reutilização da ontologia através do compartilhamento; e irá prover suporte semântico à interoperabilidade entre sistemas computacionais.
Cenário:	Não existência de um instrumento ontológico implementado e divulgado em sistema, para representação semântico-lexical dos conceitos gerais referentes à prática e teoria da Arquivística enquanto campo do conhecimento, dificultando a recuperação de informações sobre o domínio de forma lógica e estruturada.
Reuso:	Nas pesquisas feitas durante a etapa metodológica de levantamento bibliográfico não foram encontradas ontologias do domínio da Arquivística implementadas e divulgadas no Protégé nas versão 5.6.3 ou similares para reuso, cujo viés estivesse focado nos principais conceitos da área e suas aplicações e teorias.
Grupo alvo de usuários:	<ul style="list-style-type: none"> ● Desenvolvedores e mantenedores de ontologias; ● Bibliotecários, arquivistas e profissionais da engenharia do conhecimento; ● Comunidade acadêmica e público geral (usuários) que utilizam meta-buscadores na web; ● Programadores, indexadores de busca, agentes de software e funcionários do governo eletrônico;
Uso pretendido:	Representação de dados estruturados de linguagem natural para linguagem de programação OWL sobre o domínio da Arquivística com base textual na Lei 8.159/1991 dentro da Web Semântica para padronização e interoperabilidade.
Escopo:	Esta ontologia tem como objetivo estruturar dados sobre o domínio da Arquivística para facilitar a recuperação de informações e auxiliar na interoperabilidade entre sistemas.
Tipo de ontologia:	Domínio
Aplicação da ontologia:	De acesso comum à informação
Grau de Formalidade:	Formal
Linguagem de Implementação:	padrão OWL e ferramentas RDF(S)/XML
Ferramenta utilizada para	Protégé versão 5.6.3

implementação da ontologia:	
Fontes informacionais utilizadas para extração de termos:	BRASIL. Lei nº 1.859, de 8 de janeiro de 1991. Dispõe sobre a regulamentação das profissões de Arquivista e de Técnico de Arquivo, e dá outras providências. Brasília, 1978. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l8159.htm .
Local da ontologia e manutenção:	Disponível no Github através do link. https://github.com/JessicaOliveiraUFPE/Arquivontology

Fonte: Autora (2024)

Como exibido no Quadro 10 - Especificação de Requisitos da ontologia, esta estruturação teve como intuito elucidar às seguintes perguntas: a) qual o limite do domínio representado; b) qual o propósito para construção da ontologia; c) qual o cenário que a ontologia pretende ocupar; d) quais são as ontologias que podem ser reutilizadas (caso existam); e) qual o grupo de usuários visado; f) qual o uso pretendido; g) qual o seu tipo; h) qual o nível de formalidade; i) qual a linguagem utilizada; j) quais fontes textuais foram utilizadas para extração de termos; e k) qual o local de acesso e guarda, e quem será responsável pela sua manutenção.

Com relação às perguntas que a ontologia pretende responder, também chamadas de Questões de Competência, que se encontra presente na etapa inicial de desenvolvimento da ontologia, dentro do passo 1 do Ontology Development 101, etapa de especificação da METHONTOLOGY, e etapa 1 da OntoForInfocScience, seguem os questionamentos:

Quadro 11 - Questões de Competência da Arquivontology

QUESTÕES DE COMPETÊNCIA DA ARQUIVONTOLOGY
<ul style="list-style-type: none"> → A ontologia pode ser usada para estruturação de dados referentes à Arquivística? → A ontologia contempla os principais assuntos referentes à Arquivística? → Quais são os principais conceitos da Arquivística? → Quais são os tipos de arquivos? → Como é a organização sistemática de subordinação das estruturas de poder relativas aos arquivos públicos? → Quais são os tipos de documentos de arquivo? → Quais são os principais conceitos da Gestão de Documentos? → Quais são os princípios na Administração de arquivos?

→ Quais são os tipos de instituições arquivísticas?

Fonte: Autora (2024).

Essas são todas as perguntas que a ontologia pretendia responder. Se encerrando a etapa 1 e após isto, seguindo para a próxima etapa, onde houve a aquisição do conhecimento sobre o domínio escolhido como conteúdo de sua representação.

6.2 Etapa 2: Extração e Tradução Terminológica

Para realizar esta que é a segunda etapa de criação da Arquivontology, o primeiro passo a ser executado foi a leitura integral do texto da Lei 8.159/1991 por especialista da área, com a finalidade de identificar as frases e descritores técnicos do domínio utilizados no documento base, quais suas frequências de aparição, bem como entender suas relações semânticas. O primeiro método utilizado para fazer a extração foi a análise formal do documento, seguido pelo critério de frequência. Esta análise quantitativa em conjunto com a identificação de relevância por análise semântico-lexical foram utilizadas como critérios para a diferenciação entre os termos (unidade das comunicações especializadas) a serem usados na ontologia e palavras (unidade da língua geral)¹⁶ que não tinham propósito para a meta implementada.

Para logo em sequência se dar a extração terminológica feita de forma manual dos termos, com assistência de especialistas da área ou por fontes balizadoras do conhecimento representado, para fins de autenticidade durante o processo de extração dos termos considerados pertinentes para o domínio da Arquivística, como mostra a imagem a seguir:

¹⁶ A diferenciação entre “termo (unidade dos discursos especializados) e palavra (unidade da língua geral)” é discutida no texto: Linguagens especializadas em corpora : modos de dizer e interfaces de pesquisa [recurso eletrônico] / organizadoras, Cristina Lopes Perna, Heloísa Koch Delgado, Maria José Finatto. – Dados eletrônicos. – Porto Alegre : EDIPUCRS, 2010. 255 p. disponível em: <https://shorturl.at/itDFW>.

Figura 9 - Extração Terminológica no texto da Lei 8.159/1991

METODOLOGIA DA ARQUIVONTOLOGY
ETAPA II - AQUISIÇÃO DE CONHECIMENTO SOBRE O DOMÍNIO

LEI 8.159/1991

- ☑ A ~~lei 8159~~ dispõe sobre a ~~política nacional de arquivos públicos e privados (política nacional de arquivos)~~ e dá outras providências.
- ☑ É dever do Poder Público a ~~gestão documental (gestão de documentos)~~ e a ~~proteção especial a documentos de arquivos [proteção legal de arquivos]~~, como ~~instrumento de apoio à administração (valor administrativo)~~, à ~~cultura (valor informativo)~~, ao ~~desenvolvimento científico (valor intrínseco)~~ e como ~~elementos de prova (valor probatório)~~ e ~~informação~~.
- ☑ Consideram-se ~~arquivos~~, para os fins desta Lei, os ~~conjuntos de documentos (documentação) produzidos (produção) e recebidos (entrada de documentos) [acumulação]~~ por órgãos públicos, instituições de caráter público e entidades privadas, em ~~decorrência do exercício de atividades específicas (organicidade)~~, bem como por pessoa física, qualquer que seja o ~~suporte da informação (suporte)~~ ou a ~~natureza dos documentos (natureza do conteúdo)~~.
- ☑ Considera-se ~~gestão de documentos (gestão de documentos)~~ o ~~conjunto de procedimentos e operações técnicas (processamento técnico)~~ referentes à sua

Fonte: Autora (2024).

Em seguida, foi realizada a tradução terminológica de todos os 113 termos extraídos. Esta tradução terminológica teve como base para texto comparativo o Dicionário Brasileiro de Terminologia Arquivística (DIBRATE)¹⁷. Ou seja, os termos extraídos foram contrapostos com os estabelecidos pelo dicionário como um instrumento consolidado da linguagem especializada na área, e como consequência, na maior parte das vezes foi preterido o termo convencionado pela literatura especializada. No entanto, em alguns casos foi necessário realizar o acréscimo de termos não listados na DIBRATE, já que estes foram considerados integrantes, bem como importantes, para representação do domínio.

Quadro 12 - Termos extraídos x Termos traduzidos

LISTAGEM DE TERMOS EXTRAÍDOS X TERMOS TRADUZIDOS

#	TERMO EXTRAÍDO	TERMO TRADUZIDO
1	acesso	acesso

¹⁷ Disponível em:
https://www.gov.br/conarq/pt-br/centrais-de-conteudo/publicacoes/dicionrio_de_terminologia_arquivistica.pdf.

2	administração de arquivos	administração de arquivos
3	administração de documentos	gestão de documentos
4	administração de instituições arquivísticas públicas	administração de arquivos
5	alienados	alienação
6	aquisição	entrada de documentos
7	arquivamento	arquivamento
8	arquivo nacional	arquivo nacional
9	arquivos	arquivo
10	arquivos de entidades religiosas	arquivo de entidade religiosa
11	arquivos do distrito federal	arquivo do distrito federal
12	arquivos do ministério da aeronáutica	arquivo do ministério da aeronáutica
13	arquivos do ministério da marinha	arquivo do ministério da marinha
14	arquivos do ministério das relações exteriores	arquivo do ministério das relações exteriores
15	arquivos do ministério do exército	arquivo do ministério do exército
16	arquivos do poder executivo	arquivo do poder executivo
17	arquivos do poder jurídico	arquivo do poder jurídico
18	arquivos do poder legislativo	arquivo do poder legislativo
19	arquivos estaduais	arquivo estadual
20	arquivos federais	arquivo federal
21	arquivos municipais	arquivo municipal
22	arquivos privados	arquivo privado
23	arquivos públicos	arquivo público
24	arquivos públicos dos territórios	arquivo público
25	atividades administrativas	arquivo administrativo
26	avaliação	avaliação
27	categorias de sigilo	grau de sigilo
28	classificação	classificação
29	classificação de documentos	grau de sigilo
30	conjunto de documentos produzidos e recebidos	acumulação
31	conjuntos de documentos	documentação
32	conselho nacional de arquivos	conselho nacional de arquivos
33	consulta	consulta
34	critérios de organização	processamento técnico
35	depositado a título revogável	empréstimo
36	depositados	armazenamento
37	depósito	depósito
38	desfigurar documento	desfiguração
39	destruição	eliminação
40	destruir documento	eliminação
41	dispersão	desfiguração
42	divulgação	disseminação da informação
43	doados	doação
44	documentação pública	documento público
45	documentos correntes	arquivo corrente
46	documento de interesse público e social	interesse público social
47	documento em decorrência do exercício de atv.	organicidade
48	documento privado	documento privado
49	documento público	documento público
50	documento sigiloso	documento sigiloso

51	documentos de arquivos	documento classificado
52	documentos do poder executivo	documento do poder executivo
53	documentos do poder judiciário	documento do poder judiciário
54	documentos do poder legislativo	documento do poder legislativo
55	documentos intermediários	arquivo intermediário
56	documentos permanentes	arquivo permanente
57	elemento de prova	valor probatório
58	eliminação	eliminação
59	exibição reservada	disseminação da informação
60	fase corrente	fase corrente
61	fase intermediária	fase intermediária
62	fontes relevantes	fonte de referência
63	franqueado	acesso
64	gestão de documentos	gestão de documentos
65	gestão documental	gestão de documentos
66	guarda	armazenamento
67	guarda permanente	guarda permanente
68	imprescritíveis	imprescritibilidade
69	inalienáveis	inalienabilidade
70	informação	informação
71	instituição arquivística do distrito federal	instituição arquivística do distrito federal
72	instituição arquivística estadual	instituição arquivística estadual
73	instituição arquivística federal	instituição arquivística federal
74	instituição arquivística municipal	instituição arquivística municipal
75	instituição arquivística pública	instituição arquivística pública
76	instrumento de apoio à administração	valor administrativo
77	instrumento de apoio à cultura	valor informativo
78	instrumento de apoio ao desenv científico	valor intrínseco
79	lei nº 8.159	lei 8159
80	movimentação	tramitação
81	natureza dos documentos	natureza do conteúdo
82	operações técnicas	processamento técnico
83	organização	classificação
84	organização de arquivos	administração de arquivos
85	organização de instituições arquivísticas	administração de arquivos
86	órgão produtor	entidade produtora
87	perda da unidade documental	perda da unidade documental
88	política nacional de arquivos	política nacional de arquivos
89	possuidor	entidade custodiadora
90	preservação	preservação
91	procedimentos e operações técnicas	processamento técnico
92	produção	produção
93	proprietário	entidade custodiadora
94	proteção especial a documentos de arquivo	proteção legal dos arquivos
95	recebimento	entrada de documentos
96	recolhimento	recolhimento
97	registros civis	documento pessoal
98	restrito	restrição de acesso
99	sigilo	sigilo
100	sigilosos	proteção de dados
101	sistema nacional de arquivos	sistema nacional de arquivos
102	suporte da informação	suporte
103	tramitação	tramitação

104	transferência	transferência
105	transferência à instituição sucessora	herança de fundos
106	transferência para o exterior	sucessão arquivística
107	uso	uso
108	valor histórico	valor permanente
109	valor informativo	valor informativo
110	valor permanente	valor permanente
111	valor probatório	valor probatório
112	vinculação de arquivos	jurisdição arquivística
113	violação do sigilo	violação do sigilo

Fonte: Autora (2024).

Como exemplificado no Quadro 12 - Termos extraídos x Termos traduzidos, todos os 113 termos extraídos foram traduzidos com base na lista do dicionário da área, e os outros termos acrescentados estão destacados em negrito. Estas são em sua maioria situações de representação de unidades específicas do mundo real, como por exemplo: “arquivo do ministério da marinha” ou o próprio “arquivo nacional”, e por este motivo não estão listados no DIBRATE. Os acréscimos remanescentes são de termos encontrados no texto que representam conceitos não listados no dicionário, e portanto, se mostram importantes de serem acrescentados à representação do domínio, como é o caso de “sigilo” e “produção”.

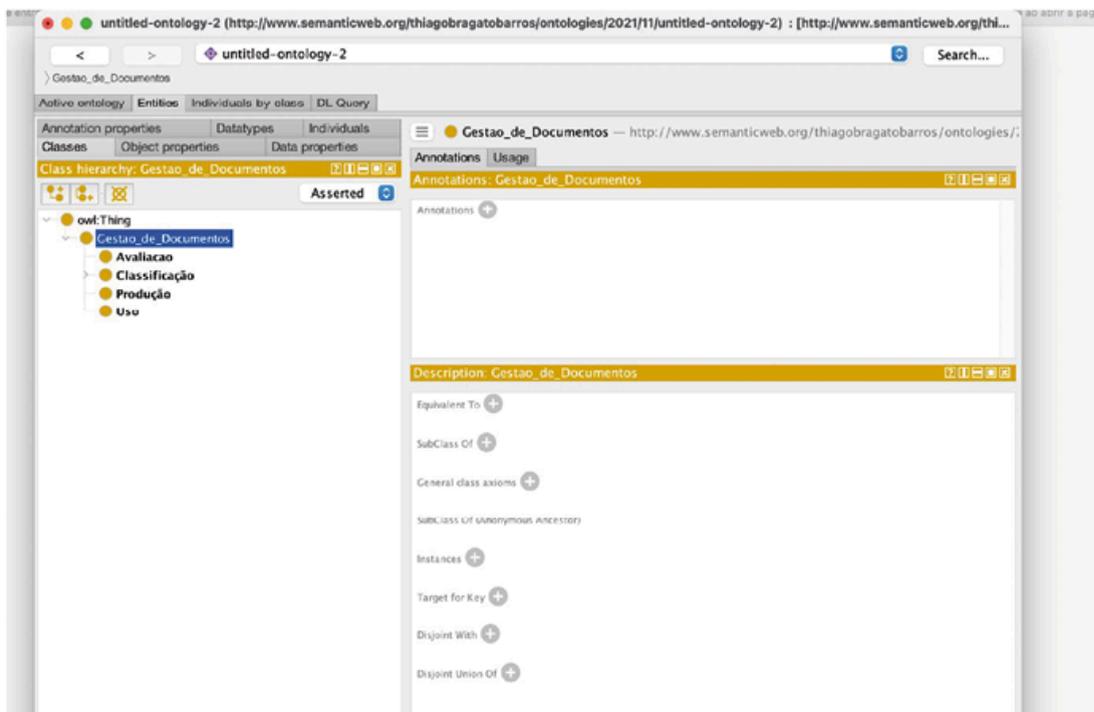
O único termo listado que posteriormente foi retirado da lista, foi o termo “lei 8159”, pois, mesmo tendo o entendimento de que a um descritor representando a lei mais importante da área e a qual serviu como base textual para criação do trabalho tenha sua relevância, esta ontologia é focada nos conceitos principais do domínio da Arquivística e não pretende fazer uma representação da estrutura da lei, como é o caso das ontologias jurídicas explanadas na seção 3.5 Tipos de ontologias. Assim, em razão disto se decidiu pela retirada do termo, e desta forma este descritor foi listado no Quadro 12 em negrito e grifado em roxo.

Este exercício possibilitou uma comparação de alguns dos termos extraídos com relação aos encontrados no DIBRATE, quanto à preferência e aplicabilidade de uso geral. Como foi o caso por exemplo do descritor “gestão de documentos”, o qual tinha como sinônimo com maior recorrência na literatura cinzenta, assim como outros textos e aparições em sites diversos, como a própria lei analisada, em contraponto ao termo “gestão documental” listado no dicionário. Com isto, levanta-se o questionamento quanto à possibilidade de revisão da lista de descritores do documento técnico, tendo em vista a preferência atual da literatura especializada por termos que não se encontram representados no DIBRATE.

6.3 Etapa 3: Reuso de ontologias

Esta etapa está presente na METHONTOLOGY na etapa 4.4 Integração, e se encontra no passo 2 do Ontology Development 101 e na etapa 4 da OntoForInfocScience, E como esclarecido na seção 4.2 da metodologia, não foram identificados durante o levantamento bibliográfico, nenhuma ontologia implementada no sistema Protégé , sobre o domínio da Arquivística. O documento mais próximo dos parâmetros estabelecidos por esta pesquisa, foi o pequeno exercício de Barros, Bastos e Santos (2022), que estruturava o conceito de gestão de documentos, como mostra a Figura 10:

Figura 10 - Ontologia de gestão de documentos feito no Protégé



Fonte: Barros, Bastos e Santos (2022, p.11).

6.4 Etapa 4: Conceitualização e Estruturação Taxonômica

A primeira parte desta etapa foi a criação de uma hierarquia da lista de termos determinados para representar o domínio, com o propósito de criar uma taxonomia dos conceitos, definindo desde já as classes e subclasses principais da ontologia. Esta taxonomia inicialmente foi feita por uma simples listagem dos termos conforme sua colocação hierárquica (Anexo A - Taxonomia da Arquivontology).

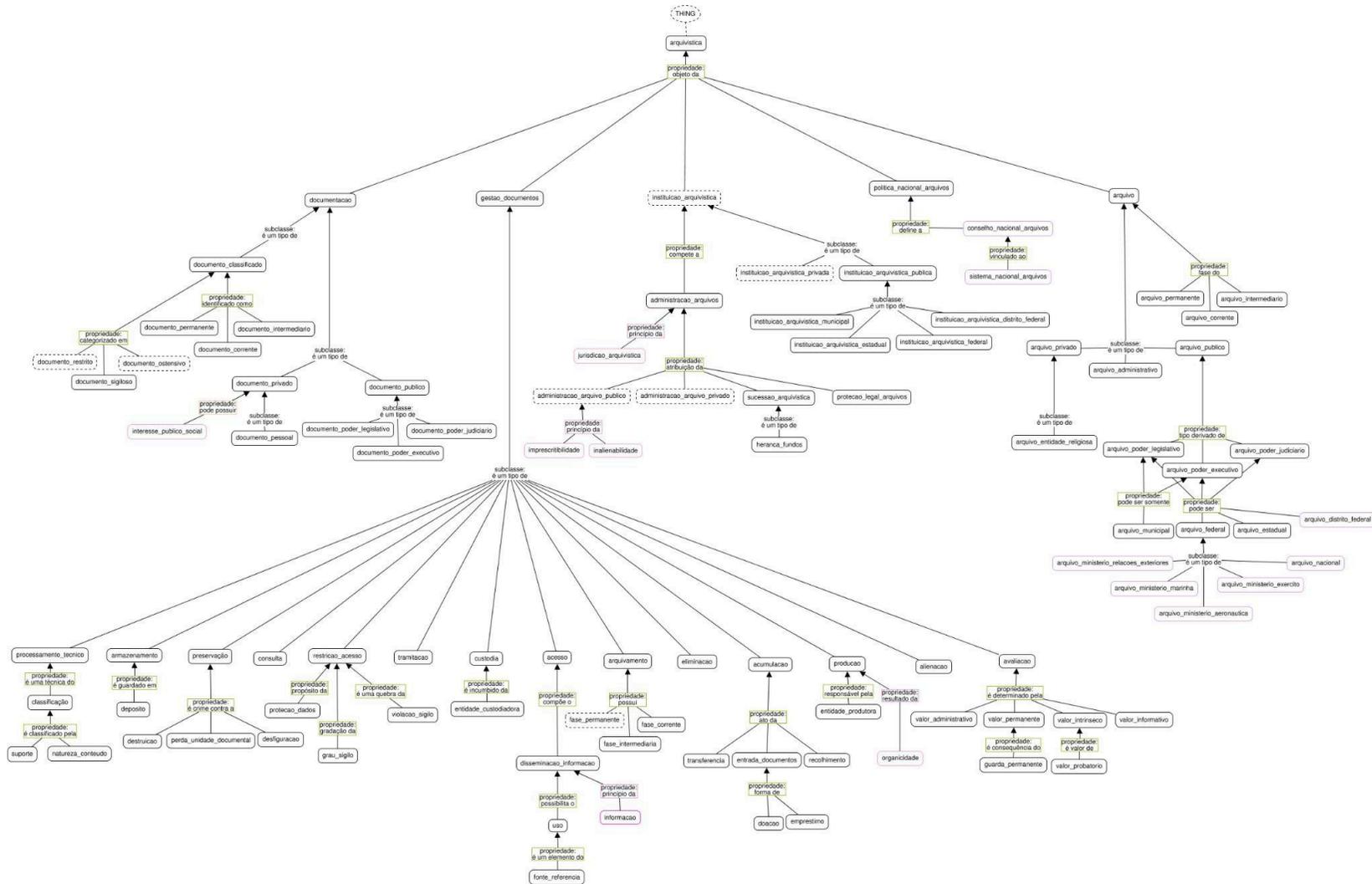
Esta estruturação considera as classes principais ou gerais como os elementos mais acima ou à esquerda, e as subclasses como os elementos abaixo ou mais à direita do ordenamento. Os termos que foram definidos como instâncias, estão representados pelos termos mais à direita, em negrito e na cor roxa, para diferenciação. Enquanto os termos que foram inferidos como propriedades das classes e subclasses, estão em negrito, itálico e grifados na cor roxa.

O critério pelo qual esses termos foram entendidos como propriedades, foi o de análise semântico-lexical, durante a extração manual e análise informal do texto da lei, pois os descritores em suas definições se vinculam a outros sujeitos como propriedades ou elementos essenciais, ao invés de subcategorias ou especificações das ações que as classes gerais representam, e assim foram aferidos pelos desenvolvedores da ontologia como propriedades.

Em seguida, com propósitos de facilitar a visualização de todos os termos definidos hierarquicamente, bem como as características de suas propriedades e suas instâncias, foi gerada uma esquematização através de um mapa conceitual pelo software CmapTools¹⁸ na versão em nuvem, chamada de Cmap Cloud, como ilustrado a seguir:

Figura 11 - Taxonomia da Arquivontology no CmapTools

¹⁸ CmapTools é um software de mapeamento de conceito desenvolvido pelo Florida Institute for Human and Machine Cognition. Disponível em: <https://cmap.ihmc.us/cmaptools/>.

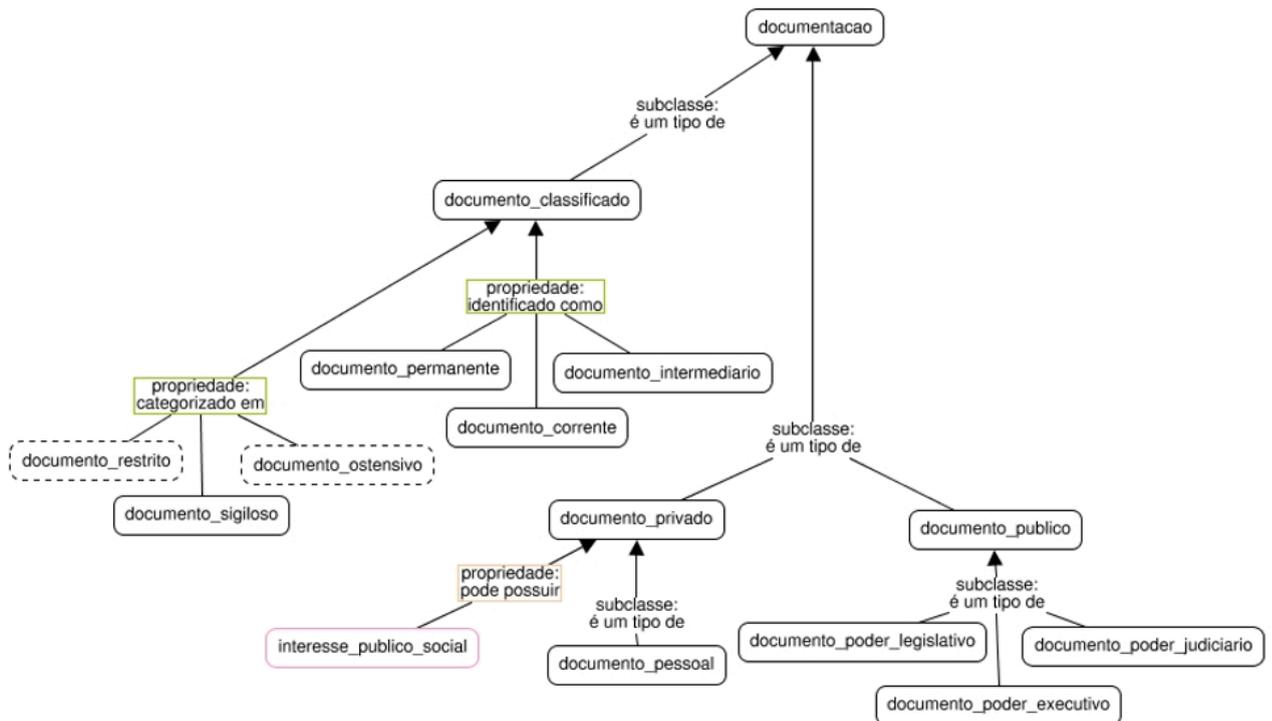


Fonte: Autora (2024)

A Figura 11 - Taxonomia da Arquivontology no CmapTools, esquematiza a totalidade de 97 termos usados da ontologia (entre classes, subclasses, instâncias e algumas das propriedades). A classe “arquivística”, termo geral da ontologia, está representada na parte superior da imagem e em suas subdivisões ou subclasses são divididas entre os 6 termos gerais: “arquivo”, “instituição arquivística”, “documentação”, “administração de arquivos”, “gestão de documentos” e “política nacional de arquivos”. Logo abaixo desses termos seguem suas subclasses e por fim, instâncias, as quais somente existem em algumas classes.

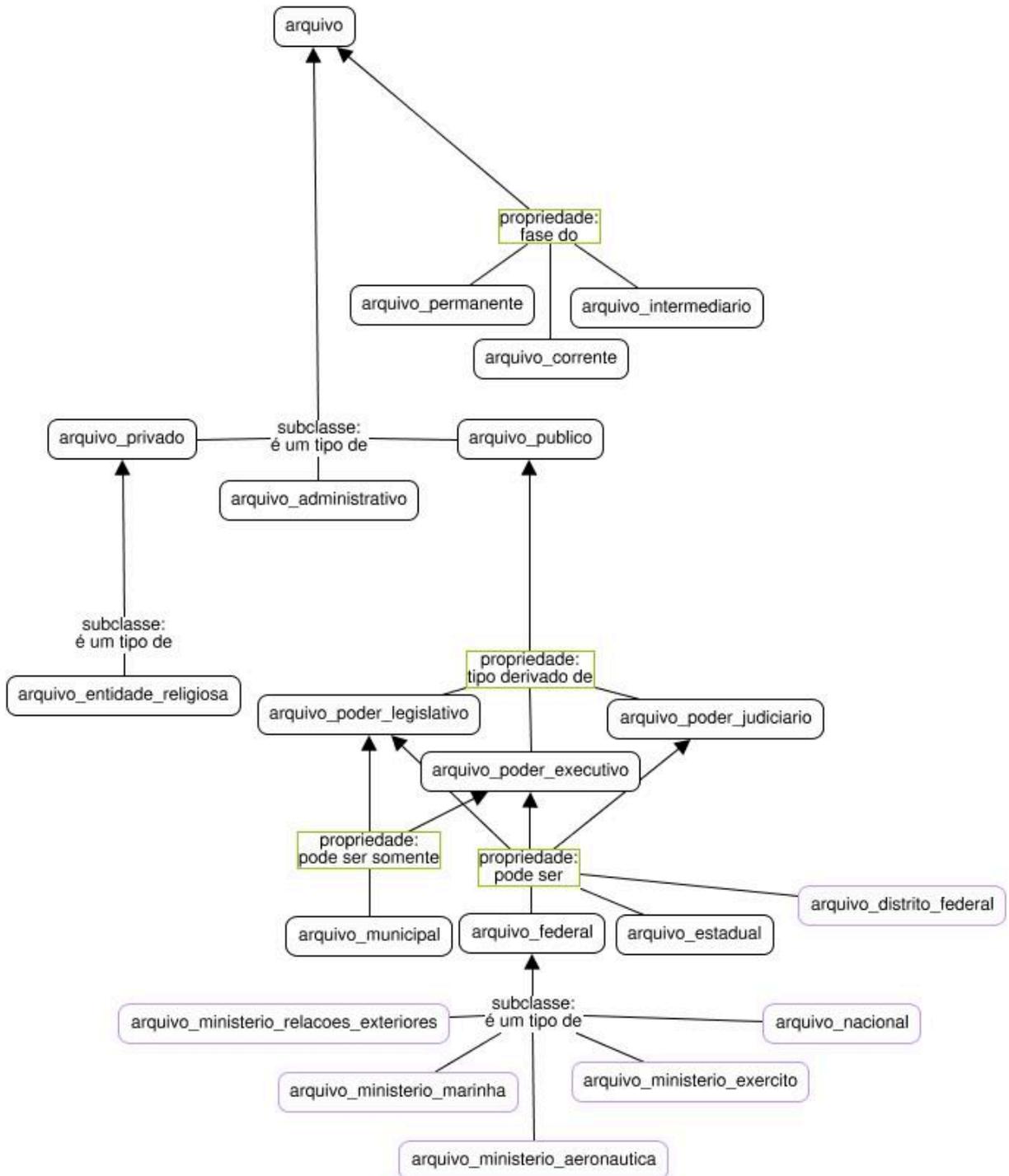
E por fim, os relacionamentos entre classes e subclasses também foram representados através de sentenças, também nomeadas de predicados ou propriedades dentro da metodologia de representação em triplas da modelagem do conhecimento. Os termos que representam as classes, subclasses e instâncias estão representados dentro de balões. Sendo os balões roxos as instâncias, e os balões tracejados, os termos incluídos na ontologia, pois não estavam contemplados pelo DIBRATE, durante a etapa 2. Os relacionamentos entre conceitos estão representados pelas frases com setas apontadas para cima, entre os balões. Quando os relacionamentos representam propriedades estão cercados por uma linha verde, enquanto os relacionamentos que representam somente as subclasses dentro do sistema, não estão grifadas. Para visualização em detalhe, seguem as imagens das classes separadamente:

Figura 12 - Taxonomia da classe “documentação”



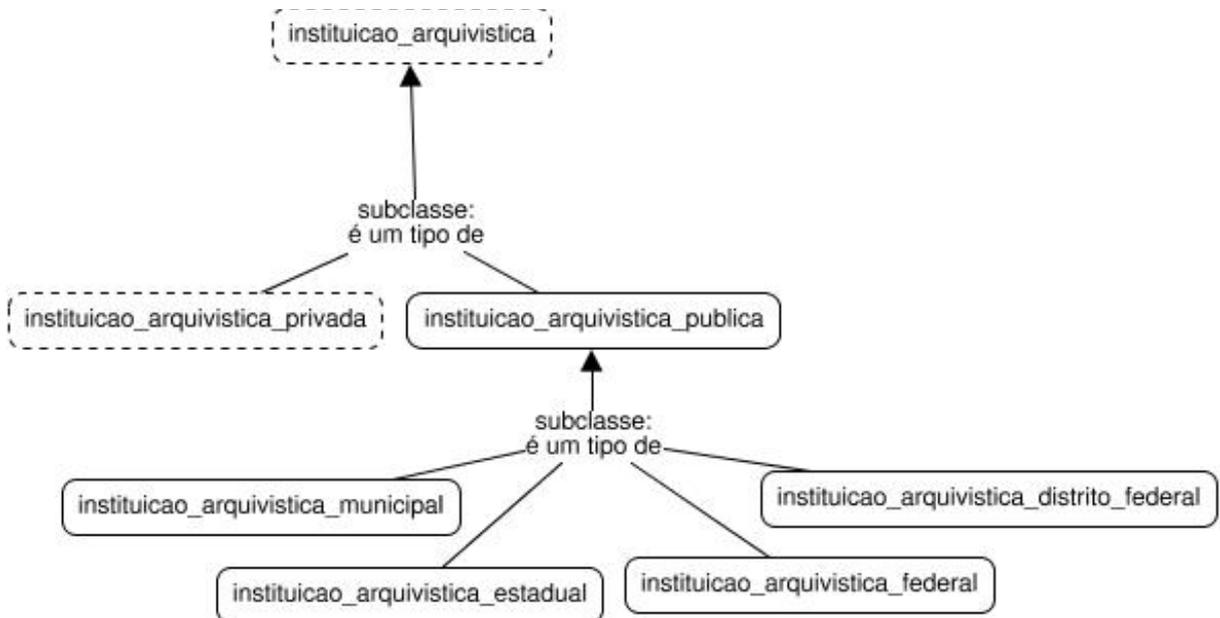
Fonte: Autora (2024).

Figura 13 - Taxonomia da classe “arquivos”



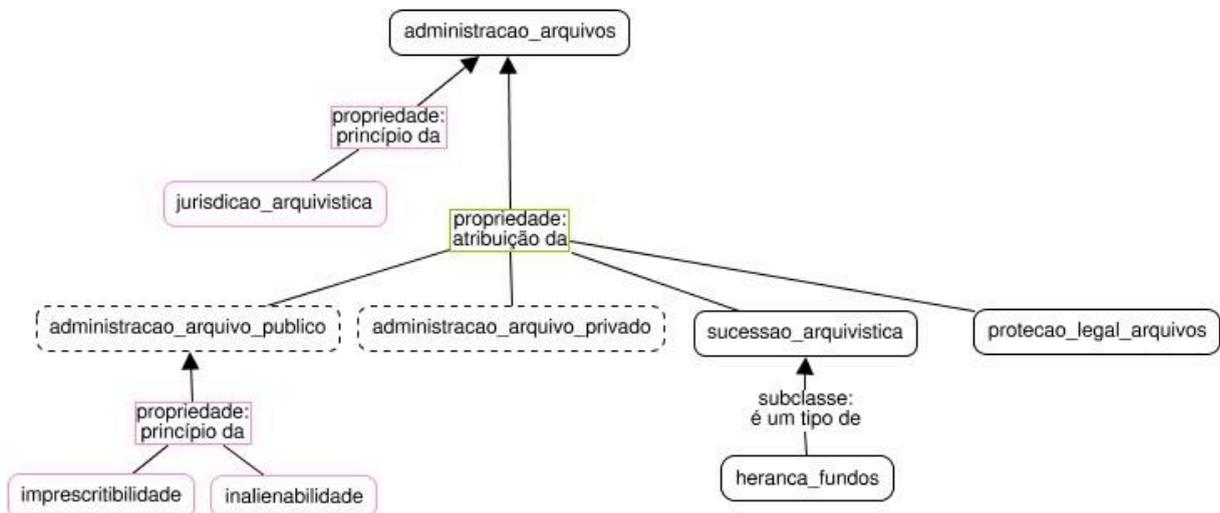
Fonte: Autora (2024).

Figura 14 - Taxonomia da classe “instituição arquivística”



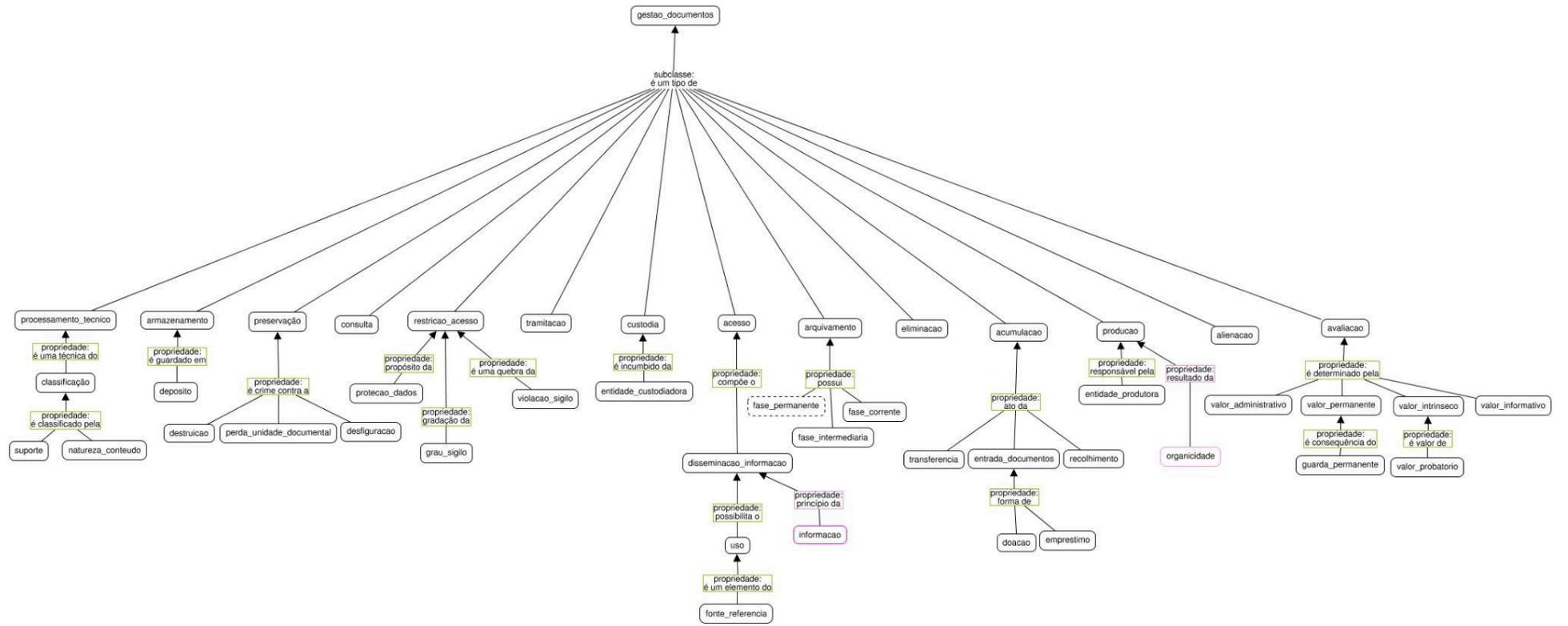
Fonte: Autora (2024).

Figura 15 - Taxonomia da classe “administração de arquivos”



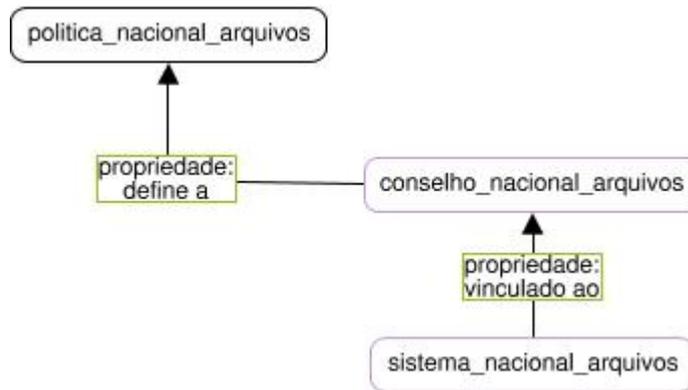
Fonte: Autora (2024).

Figura 16 - Taxonomia da classe “gestão de documentos”



Fonte: Autora (2024).

Figura 17 - Taxonomia da classe “política nacional de arquivos”



Fonte: Autora (2024).

Como exemplificado nas imagens acima, os conceitos foram hierarquizados de forma vertical por classes gerais e subclasses graficamente representados com os termos escritos dentro de balões, com diferenciação entre termos traduzidos e termos adicionados pela linha do balão: linha simples para o primeiro tipo e tachado para o segundo, e os relacionamentos por frases com setas entre os balões. Dentro deste diagrama também foi possível definir quais seriam as instâncias, observado pelos balões na cor roxa, das classes “arquivo federal” e “instituições arquivísticas públicas”, as únicas que possuem tais objetos, e os termos considerados propriedades de classes, constituídos de balões na cor rosa.

Como elucidado na metodologia, o modelo RDF utiliza a Representação em Triplas para estruturação em base de dados, assim como método selecionado para promover uma visualização simples deste sistema, e definir quais seriam as frases relacionais entre os diversos tipos de relacionamentos disponíveis: entre duas classes, entre uma instância e uma classe, entre uma propriedade e uma classe ou entre duas instâncias, foi elaborado um quadro que exemplifica alguns relacionamentos de todas as propriedades criadas:

Quadro 13 - Representação RDF em triplas

REPRESENTAÇÃO EM TRIPLAS		
SUJEITO	PREDICADO	OBJETO
arquivo	é objeto da	arquivística
administração de arquivos	compete a	instituição arquivística
proteção legal de arquivos	atribuição da	administração de arquivos
documento corrente	identificado como	documento classificado
documento restrito	caracterizado como	documento classificado

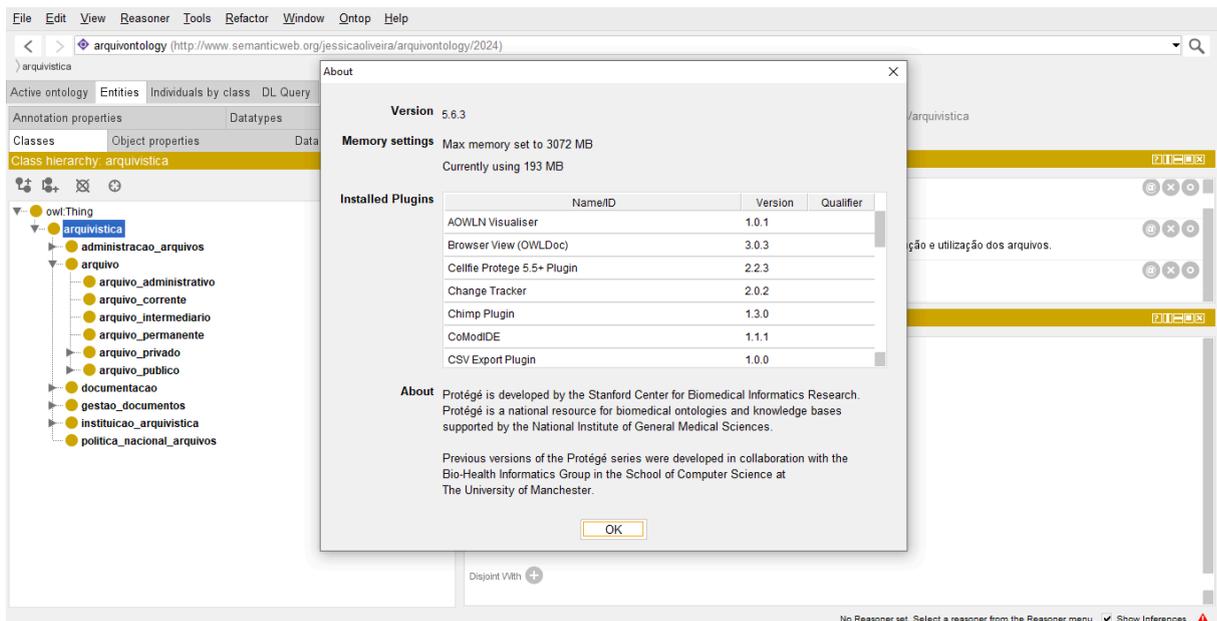
conselho nacional de arquivos	define a	política nacional de arquivos
sistema nacional de arquivos	vinculado ao	conselho nacional de arquivos
arquivo intermediário	fase do	arquivo
arquivo do poder legislativo	tipo derivado do	arquivo público
arquivo estadual	pode ser	arquivo judiciário
arquivo municipal	pode somente ser	arquivo executivo
disseminação da informação	compõe o	acesso
classificação	é uma técnica do	processamento técnico
suporte	é classificado pela	classificação
depósito	local de guarda do	armazenamento
desfiguração	é crime contra o	preservação
proteção de dados	é propósito da	restrição de acesso
violação do sigilo	é quebra	restrição de acesso
grau de sigilo	é gradação da	restrição de acesso
entidade custodiadora	é incumbido da	custódia
uso	possibilita a	disseminação de informação
fonte de referência	é elemento do	uso
arquivamento	possui fase	fase corrente
recolhimento	ato da	acumulação
doação	forma de	entrada de documentos
entidade produtora	responsável pela	produção
valor administrativo	é determinado pela	avaliação
guarda permanente	é consequência do	valor permanente
valor probatório	é valor do	valor intrínseco

Fonte: Autora (2024).

6.5 Etapa 5: Implementação em sistema

A etapa de implementação especificamente para o sistema Protégé, utilizou a inclusão de classes gerais e subclasses pelo método *top-down*. Inicialmente a implementação foi realizada em software, o escolhido para tal ação foi o Protégé, em sua versão web. Contudo, esta versão apesar de ter disponibilidade de acesso em qualquer dispositivo por estar em nuvem, facilitando o processo de construção da ontologia, se mostrou posteriormente limitante com relação a vários dos recursos disponíveis da ferramenta, e por isso, durante a etapa de implementação o software online foi substituído pela versão 5.6.3 desktop do Protégé, como mostra a imagem adiante:

Figura 18 - Versão 5.6.3 em desktop do Protégé



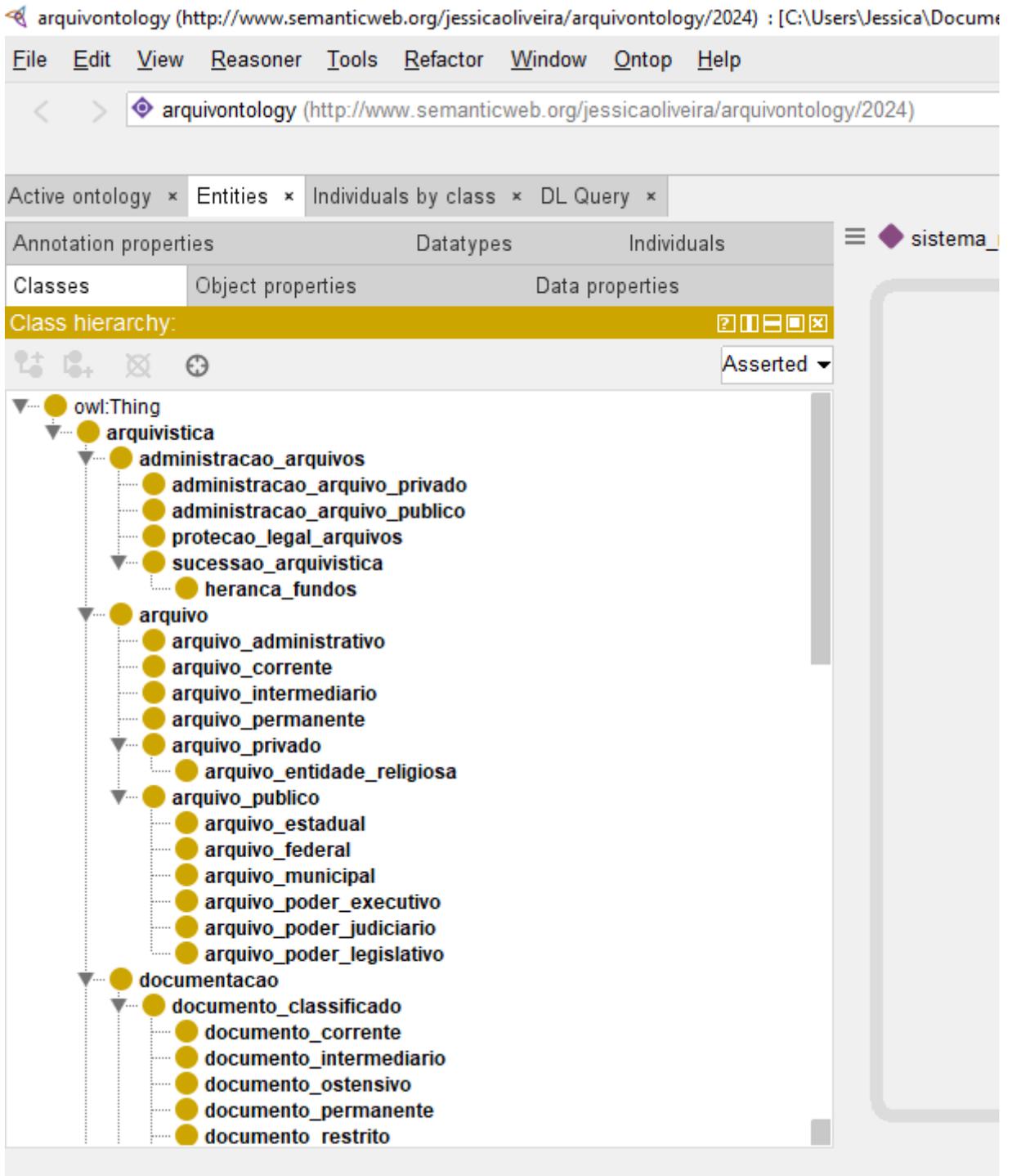
Fonte: Autora (2024).

A superclasse da ontologia foi nomeada como “arquivística” e representa o domínio em si, suas subclasses são: “administração de arquivos”, “arquivo”, “instituição arquivística”, “documentação”, “gestão de documentos” e “política nacional de arquivos”. Esses são os 6 termos gerais que descrevem as ramificações da superclasse “arquivística”. Para a classe geral “administração de arquivos”, foram criadas as seguintes subclasses: “administração de arquivos privados”, “administração de arquivos públicos”, “proteção legal de arquivos” e “sucessão arquivística”. E a subclasse da subclasse “sucessão arquivística”, é “herança de fundos”

Já a classe geral “arquivo”, possui as subclasses “arquivo administrativo”, “arquivo corrente”, “arquivo intermediário”, “arquivo permanente”, “arquivo privado” e “arquivo público”. A subclasse de arquivo privado é “arquivo de entidade religiosa”. E as subclasses de “arquivo público” são: “arquivo estadual”, “arquivo municipal”, “arquivo federal”, “arquivo do poder executivo”, “arquivo do poder legislativo” e “arquivo do poder judiciário”.

Em contrapartida, a classe geral “documentação”, possui as seguintes subclasses: “documento classificado”, “documento privado” e “documento público”. Dentro da subclasse “documento classificado” estão as subclasses “documento corrente”, “documento intermediário”, “documento permanente”, “documento ostensivo”, “documento restrito” e “documento sigiloso”. Como estão ilustradas nas figuras 19, 20 e 21.

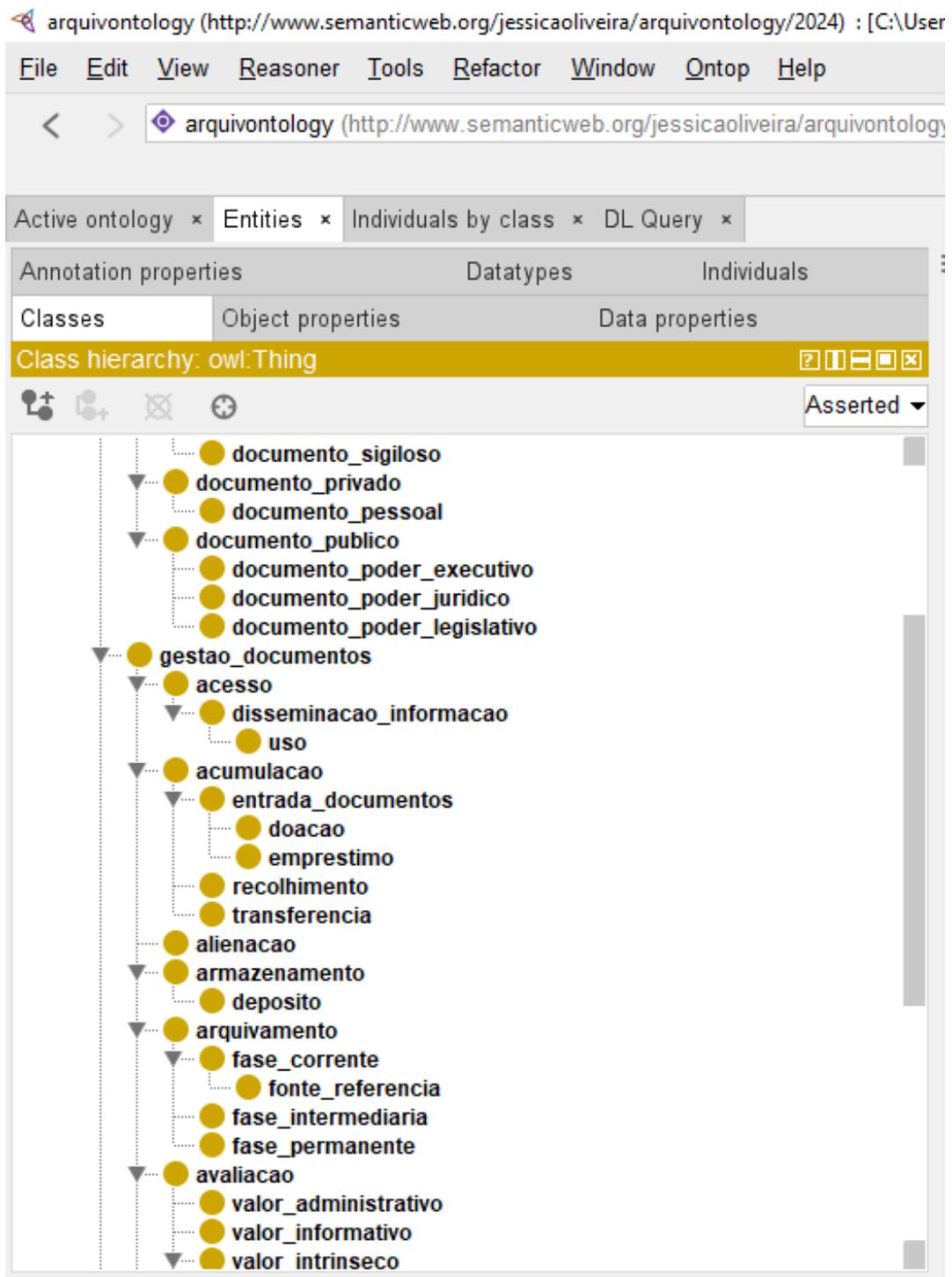
Figura 19 - Classes gerais em desktop do Protégé desktop 5.6.3



Fonte: Autora (2024)

Na subclasse “documento privado” está a subclasse “documento pessoal” e dentro da subclasse “documento público” estão as subclasses “documento do poder executivo”, “documento do poder judiciário” e “documento do poder legislativo”.

Figura 20 - Classes gerais em desktop do Protégé desktop 5.6.3



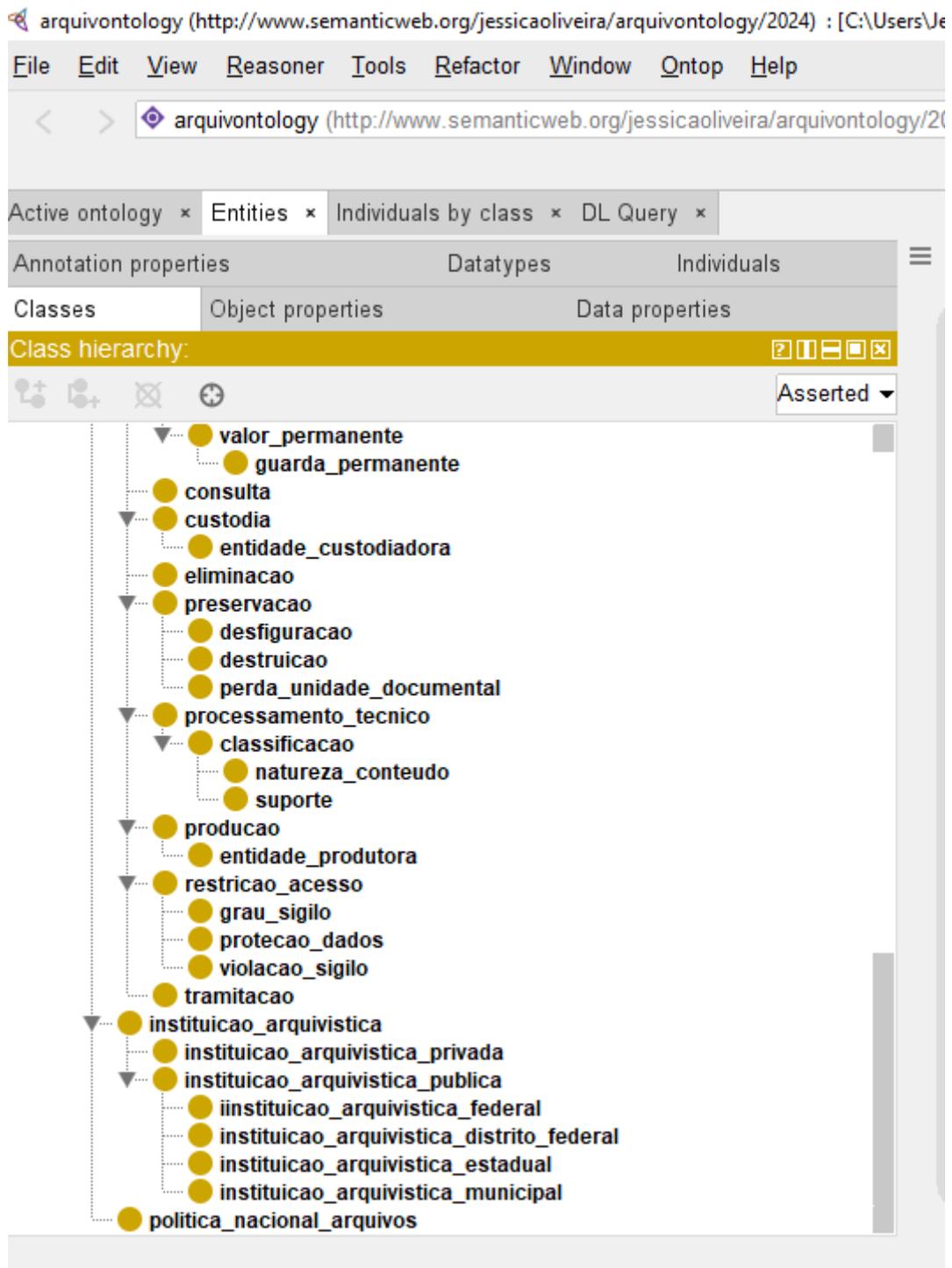
Fonte: Autora (2024).

A subclasse “gestão de documentos” a que possui mais subclasses de toda a ontologia tem: “acesso”, “acumulacao”, “alienação”, “armazenamento”, “arquivamento”, “avaliação”, “consulta”, “custódia”, “eliminação”, “preservação”, “processamento técnico”, “produção”, “restrição de acesso” e “tramitação”. A subclasse de “acesso” é a “disseminação de informação” cuja subclasse é “uso”. As subclasses de “acumulação” é “entrada de documentos”, “recolhimento” e “tramitação”. Enquanto a subclasses de “entrada de documentos” são “doação” e “empréstimo”.

A subclasse de “armazenamento” é “depósito”, já as subclasses de “arquivamento” são: “fase corrente”, que possui sua própria subclasse (“fonte de referência”), “fase intermediária” e “fase permanente”. As subclasses de “avaliação” são: “valor administrativo”, “valor informativo”, “valor intrínseco” e “valor permanente”. E a subclasse de “valor intrínseco” é “valor probatório” enquanto a subclasse de “valor permanente” é “guarda permanente”. A subclasse de “custódia” é “entidade custodiadora”, e as subclasses de “preservação” são: “desfiguração”, “destruição” e “perda da unidade documental”.

Já o termo “processamento técnico”, possui subclasse “classificação”, e este também possui subclasses que são “natureza do conteúdo” e “suporte”. A subclasse de “produção” é “entidade produtora”, enquanto que as subclasses de “restrição de acesso” são: “grau de sigilo”, “proteção de dados” e “violação de sigilo”. E para finalizar a lista de classes gerais com subclasses, temos “instituição arquivística” que possui duas subclasses: “instituição arquivística privada” e “instituição arquivística pública”, e dentro desta última, temos as subclasses: “instituição arquivística federal”, “instituição arquivística estadual”, “instituição arquivística do distrito federal” e “instituição arquivística municipal”.

Figura 21 - Classes gerais em desktop do Protégé desktop 5.6.3

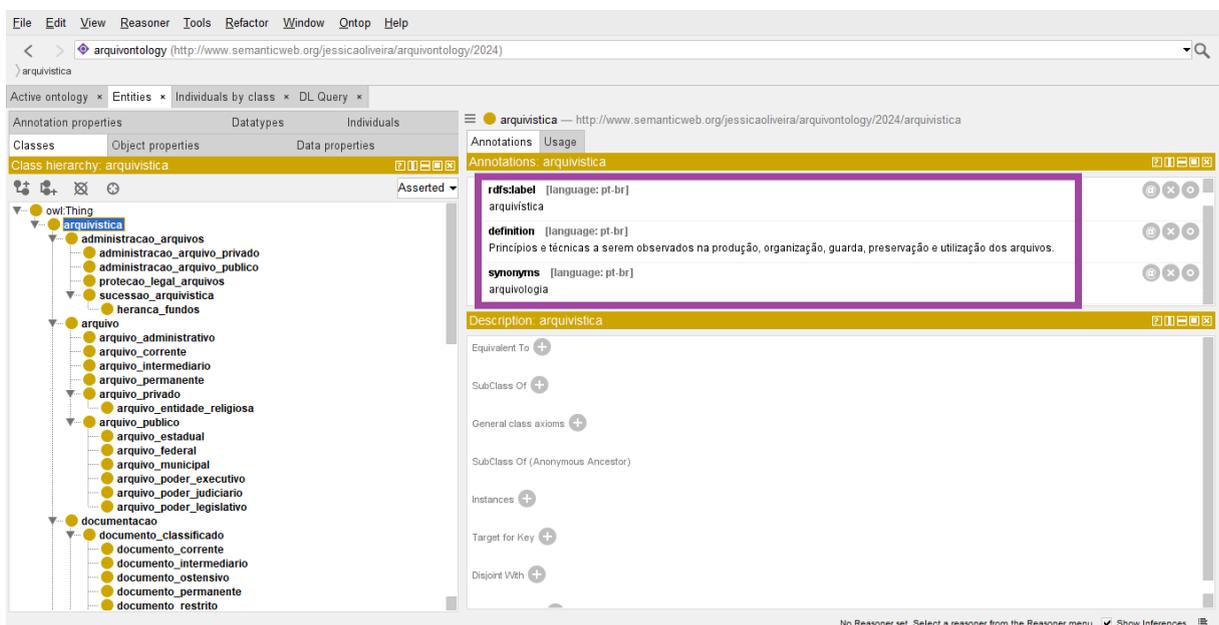


Fonte: Autora (2024)

No campo de anotação do Protégé, foram criadas 3 anotações: **rdf:label**, que define o nome do objeto representado, para a linguagem de programação esta seria a “etiqueta” que identifica o termo. A segunda anotação foi criada pela desenvolvedora da ontologia, é a

definition, que agrega a ontologia, a informação sobre a definição que todas as classes e subclasses possuem. Essas definições foram extraídas do Dicionário de Terminologia Arquivística, ou no caso dos termos inclusos, foram criadas novas definições com base no texto do DIBRATE. Já a terceira anotação, **synonyms**, e também foi criada na ontologia para definir os sinônimos de um termo. Todas as anotações foram feitas pelo em XML Schema (XSD)¹⁹, com valor *string*, como será exibido adiante:

Figura 22 - Anotações em desktop do Protégé desktop 5.6.3

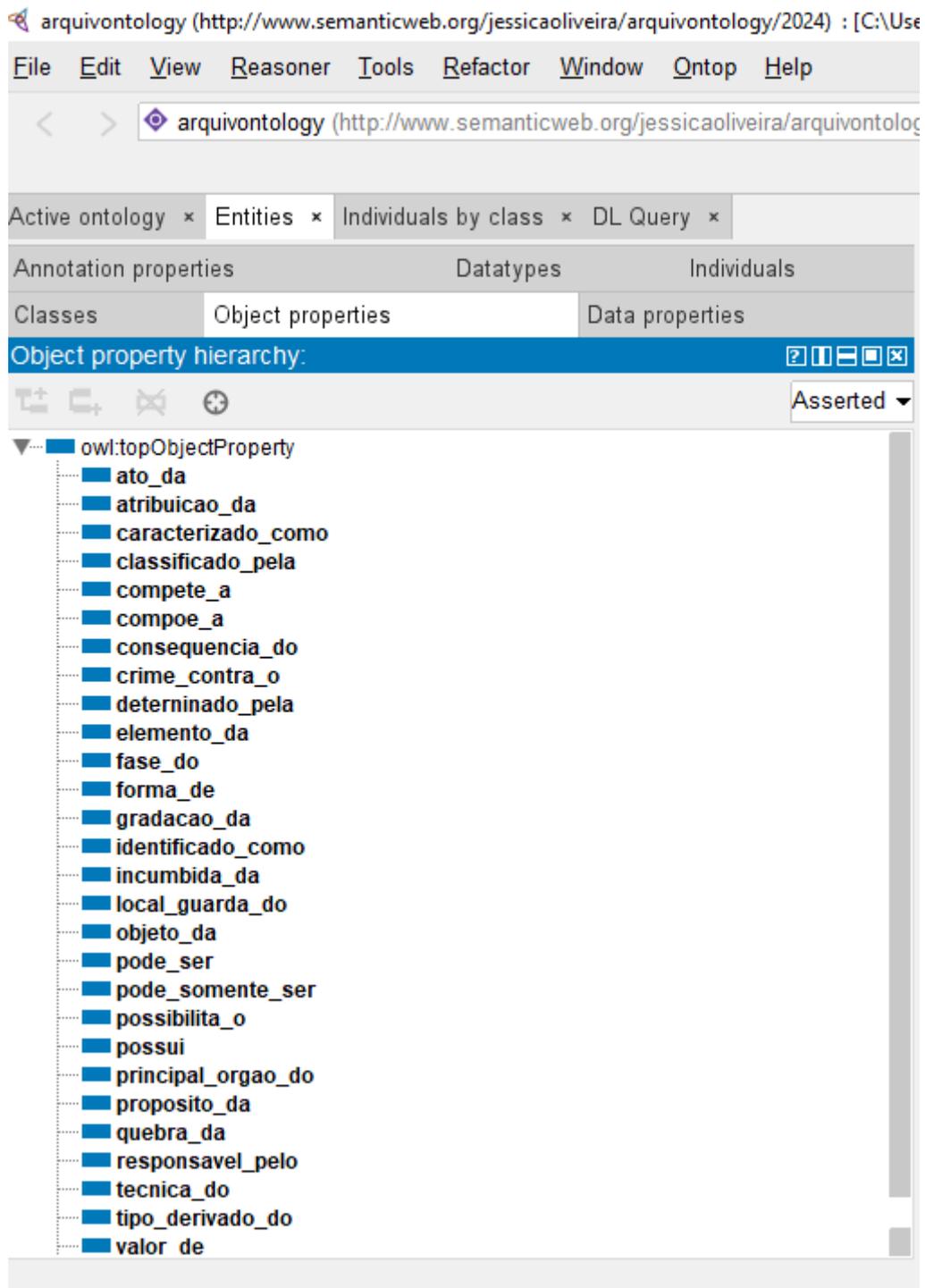


Fonte: Autora (2024).

A implementação das propriedades que traduzem os relacionamentos entre os classes e instâncias, foi feita dentro do Protégé, como *object properties*, como explicitado na seção 5.7 Conceitualização e Estruturação de dados, somando um total de 29 propriedades. Como mostra a seguinte figura:

¹⁹ A especificação do tipo de dados no esquema XML do W3C, define vários tipos de dados para validar o conteúdo do elemento e o valor do atributo. String é o valor que se dá aos elementos textuais de uma ontologia.

Figura 23 - Object Properties no Protégé desktop 5.6.3

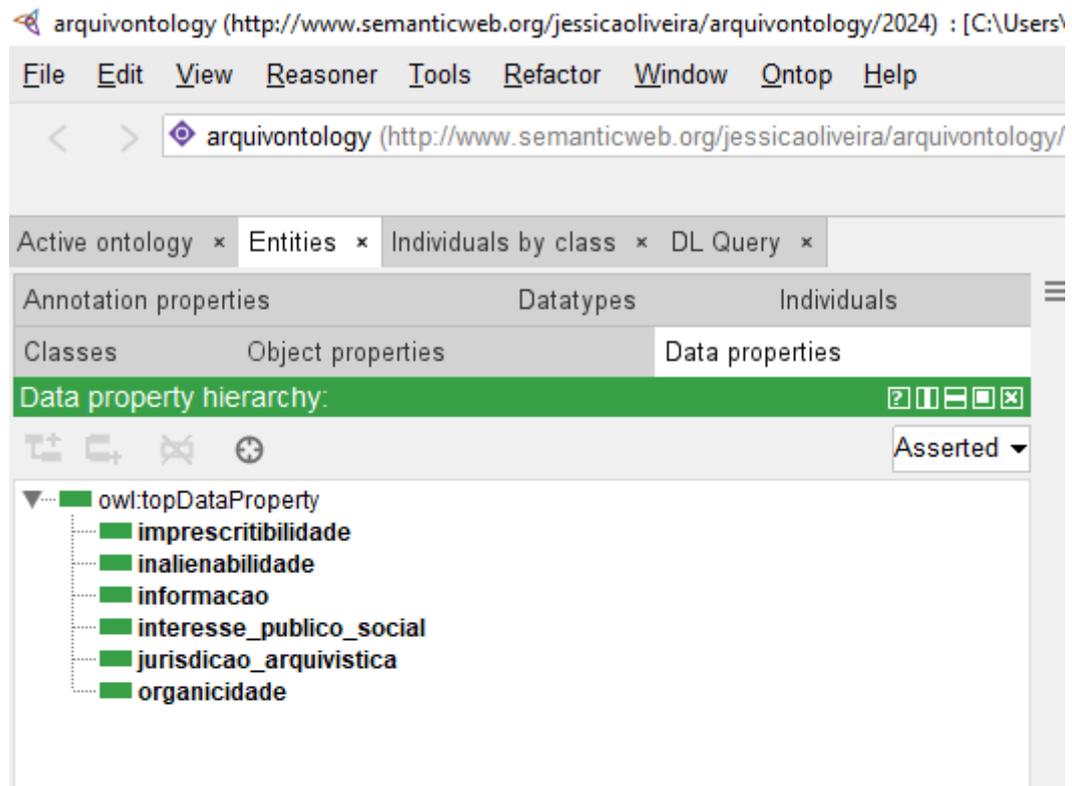


Fonte: Autora (2024)

Já a implementação das propriedades que representam elementos ou se remetem contextualmente às classes, subclasses e instâncias, foram descritas dentro do Protégé, como *data properties*. Foram 6 termos extraídos e identificados propriedades, por serem interpretados como elementos ou propósitos vinculados em vista da definição que levam no

DIBRATE. São eles: “imprescritibilidade”, “inalienabilidade”, “informação”, “interesse público social”, “jurisdição arquivística” e “organicidade”. Como mostrado figura adiante:

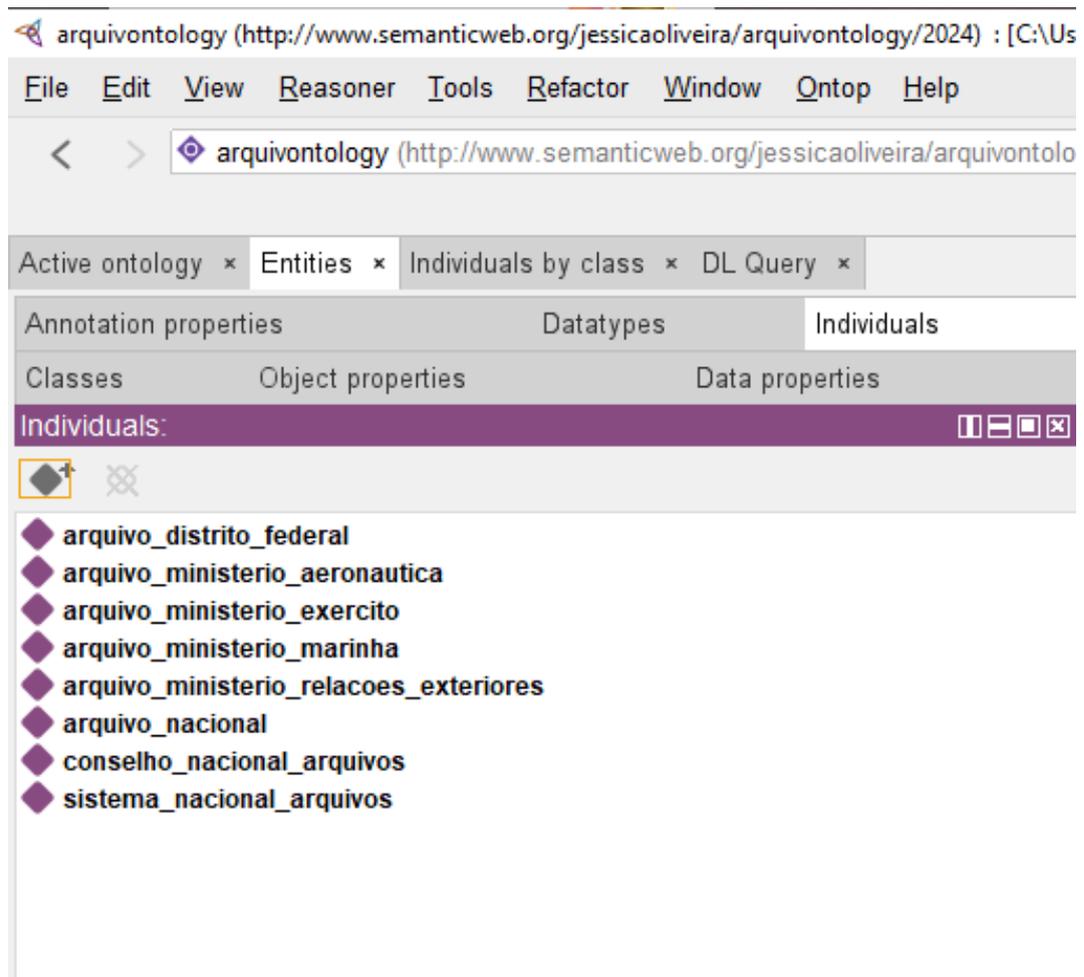
Figura 24 - Data Properties no Protégé desktop 5.6.3



Fonte: Autora (2024).

Somente foram definidos 8 termos como instâncias (individuals) de classes e subclasses dentro da ontologia, tendo em vista que estes elementos representam a um objeto particular ou específico da realidade. Os termos são “arquivo do distrito federal”, “arquivo do ministério da aeronáutica”, “arquivo do ministério do exército”, “arquivo do ministério da marinha”, “arquivo do ministério das relações exteriores”, “arquivo nacional”, “conselho nacional de arquivos” e “sistema nacional de arquivos”, ilustradas adiante:

Figura 25 - Instâncias no Protégé desktop 5.6.3



Fonte: Autora (2024).

6.6 Etapa 6: Avaliação da ontologia

Como proposto na seção 5.8 desta pesquisa, o código em linguagem OWL da Arquivontology foi analisado pela ferramenta OOPS!, e o resultado foi contemplado conforme figura abaixo:

Figura 26 - Avaliação da Arquivontology no OOPS!

The screenshot shows the OOPS! Pitfall Catalogue interface. At the top, there is a navigation bar with the OOPS! logo and links for Pitfall Catalogue, Service, Feedback, and About Us. Below the navigation bar, there is a text box explaining the three levels of importance in pitfalls: Critical, Important, and Minor. The main content area is titled "Pitfalls detected:" and contains a table with five rows of results.

Pitfall ID	Description	Count	Severity
P08	Missing annotations.	122 cases	Minor
P11	Missing domain or range in properties.	11 cases	Important
P13	Inverse relationships not explicitly declared.	29 cases	Minor
P19	Defining multiple domains or ranges in properties.	14 cases	Critical
P41	No license declared.	Ontology*	Important

Fonte: Autora (2024)

Pela análise da ferramenta, existem 3 situações de erro: 1) erros menores: com 122 casos com falta de anotações e relação de inversão não explícita; 2) erros importantes: uma licença não declarada e 11 casos de falta de *domain* e *range* em propriedades e 3) erros críticos: 14 casos de definição de múltiplo *domain* e *range*. Todos os erros menores e importantes podem ser facilmente corrigidos durante a revisão da ontologia. Já os casos de erro crítico indicam uma avaliação mais aprofundada das propriedades estabelecidas, pois talvez a criação de classes e subclasses tenha ganhado o foco durante o desenvolvimento e esta segunda parte precise de uma nova verificação, com possíveis acréscimos de propriedades. Essas são as sugestões inferidas com base na etapa de avaliação.

6.7 Etapa 7: Documentalização e Divulgação da ontologia

A formalização em documento da ontologia foi realizada durante as seções 6.1 (Especificação do Escopo e Domínio), 6.2 (Extração e Tradução Terminológica), 6.3 (Reuso de Ontologias), 6.4 (Conceptualização e Estruturação) e 6.5 (Implementação em Software) desta pesquisa. Já a divulgação em repositório no GitHub do código em OWL da Arquivontology foi realizada através do seguinte link que será disponibilizado posteriormente.

7 CONSIDERAÇÕES FINAIS

A busca pela consensualidade deve ser uma constante, uma vez que ontologias são para uso geral e suas aplicações incorrem em todo o sistema de representação, busca e recuperação de informações em contexto digital. Portanto, os termos utilizados nesta ontologia podem e devem ser discutidos quanto à sua pertinência e viabilidade linguística, bem como a incorporação de novos termos e alterações entre relacionamentos. Ontologias são ferramentas que estão em perpétua evolução pois acompanham o desenvolvimento sociocultural, político e econômico de um país.

Partindo da escassez de materiais acerca de ontologias voltadas ao domínio arquivístico, o objetivo geral deste trabalho foi elaborar uma ontologia voltada para a área da Arquivística, com enfoque na Lei nº 8.159 de 8 de janeiro de 1991, através das etapas de levantamento bibliográfico sobre o tema, a definição das metodologias, ferramentas e procedimentos e da utilização do Sistema Protégé, de modo que consideramos os tópicos específicos realizados satisfatoriamente, e assim cumprindo o objetivo inicial.

Como limitações, colocamos a falta de materiais na temática arquivística para embasar os componentes da ontologia, resultando em certa morosidade no desenvolvimento das etapas, assim como a necessidade de aprender a utilizar corretamente o Sistema Protégé.

Outro ponto importante a ser colocado é a noção de que a Arquivontology não se constitui de uma ferramenta que pretende ser a única visão e tentativa de representação da linguagem legal do domínio da Arquivística. Principalmente levando-se em consideração que devido à natureza de um trabalho de conclusão de curso de graduação, há a necessidade de aprofundamentos e aperfeiçoamentos posteriores, como por exemplo na expansão de termos que poderiam ser adicionados para abranger melhor as relações entre conceitos, ficando assim esta tarefa, para uma possibilidade futura.

Também como sugestão, colocamos a necessidade de atualizar os termos do Dicionário Brasileiro de Terminologia Arquivística, o que pode trazer novas contribuições sobre a área. Sobre trabalhos que podem ser realizados acerca das ontologias, vemos a Ciência da Informação como uma infinidade de possibilidades tanto em questões teóricas, como aplicadas, como implantação de ontologias em sistemas e bases de dados já existentes, a exemplo do domínio governamental. Por fim, avaliamos a experiência desta pesquisa como

positiva e esperamos que as contribuições para a Ciência da Informação, Biblioteconomia e Arquivística sejam proveitosas.

REFERÊNCIAS

ABEL, Mara; FIORINI, Sandro Rama. Uma revisão da engenharia do conhecimento: evolução, paradigmas e aplicações. **International Journal of Knowledge Engineering and Management**, Florianópolis, v. 2, n. 2, p. 1-35, mar./maio 2013. Disponível em: <https://periodicos.ufsc.br/index.php/ijkem/article/view/81475>. Acesso em 24 jan. 2024.

ALMEIDA, M. B.; BAX, M. P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Ciência da Informação**, Brasília, DF, v. 32, n. 3, p. 7-20, set./dez. 2003. Disponível em: <https://www.scielo.br/j/ci/a/LR68syZsPSSmwwPHrNXmC8N/?format=pdf&lang=pt>. Acesso em: 9 mar. 2024.

ARQUIVO NACIONAL (Brasil). **Dicionário brasileiro de terminologia arquivística**. Rio de Janeiro: Arquivo Nacional, 2005. 232 p. Disponível em: https://www.gov.br/conarq/pt-br/centrais-de-conteudo/publicacoes/dicionrio_de_terminologia_arquivistica.pdf. Acesso em: 13 jun. 2023.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 9578**: arquivos: terminologia. Rio de Janeiro: ABNT, 1986. 4p.

BARROS, T. H. B. A Lei de Arquivos do Brasil: apontamentos sobre seu impacto 30 anos depois. **Revista do Arquivo**, São Paulo, v. 7, n. 13, p. 152-155, dez. 2021. Disponível em: <https://lume.ufrgs.br/handle/10183/237197>. Acesso em: 01 ago. 2023.

BARROS, T. H. B. Os arquivos, a arquivística e o discurso: alguns marcos históricos e conceituais. **Informação Arquivística**, Rio de Janeiro, v. 2, n. 1, p. 135-157, jan./jun. 2013. Disponível em: <https://www.aaerj.org.br/ojs/index.php/informacaoarquivistica/article/view/23/15>. Acesso em: 18 abr. 2023

BARROS, T. H. B.; BASTOS, C. M. C.; SANTOS, A. C. R. Sistemas de organização do conhecimento no contexto da arquivologia: aportes metodológicos para seu desenvolvimento. **Acervo: Revista do Arquivo Nacional**, Rio de Janeiro, v. 35, n. 2, p. 1-20, maio/ago. 2022. Disponível em: <https://revista.an.gov.br/index.php/revistaacervo/article/view/1812>. Acesso em: 9 mar. 2024.

BATRES, Eduardo Jaime Quirós; OLIVEIRA, Alcione de Paiva; GABRIELLI, Bruno Vitorim; AMORIM, Vinci Pegoretti; MOREIRA, Alexandra. Uso de ontologias para a extração de informações em atos jurídicos em uma instituição pública. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, [S. l.], v. 10, n. 19, p. 73–88, 2005. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2005v10n19p73>. Acesso em: 24 fev. 2024.

BELLOTTO, Heloísa Liberalli. A terminologia das áreas do saber e do fazer: o caso da arquivística. **Acervo: Revista do Arquivo Nacional**, Rio de Janeiro, v. 20, n. 1-2, p. 47–56,

jan./dez. 2007. Disponível em:

<https://revista.arquivonacional.gov.br/index.php/revistaacervo/article/view/71>. Acesso em: 4 set. 2023.

BERNERS-LEE, Tim; HENDLER, James. Publishing on the semantic web. **Nature**, London, v. 410, n. 6832, p. 1023-1024, 2001. Disponível em:

<https://www.nature.com/articles/35074206>. Acesso em: 24 jan. 2024.

BIAGETTI, Maria Teresa. Ontologies as Knowledge Organization Systems. **Knowledge Organization**, Germany, v. 48, n. 2, p. 152-176, 2021. Disponível em:

https://web.archive.org/web/20220312052742id_/https://www.nomos-elibrary.de/10.5771/0943-7444-2021-2-152.pdf. Acesso em: 24 jan. 2024.

BRASCHER, Marisa.; CAFÉ, Lígia. Organização da informação ou organização do conhecimento? *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 9., 2008, São Paulo. **Anais eletrônicos [...]**. São Paulo: USP; Rio de Janeiro: ANCIB, 2008. Disponível em:

[https://cmapspublic.ihmc.us/rid=1KR7TM7S9-S3HDKP-5STP/BRASCHER%20CAF%20C3%89\(2008\)-1835.pdf](https://cmapspublic.ihmc.us/rid=1KR7TM7S9-S3HDKP-5STP/BRASCHER%20CAF%20C3%89(2008)-1835.pdf). Acesso em: 13 ago. 2023.

BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil**. São Paulo: Atlas, 2000. Brasília: Senado, 1988.

BRASIL. Lei Complementar nº 95, de 26 de fevereiro de 1998. Dispõe sobre a elaboração, a redação, a alteração e a consolidação das leis, conforme determina o parágrafo único do art. 59 da Constituição Federal, e estabelece normas para a consolidação dos atos normativos que menciona. **Diário Oficial da União**: Brasília, DF, 27 fev. 1998. Disponível em:

http://www.planalto.gov.br/ccivil_03/Leis/LCP/Lcp95.htm. Acesso em: 25 set. 2023.

BRASIL. Lei nº 6.546, de 4 de julho de 1978. Dispõe sobre a regulamentação das profissões de Arquivista e de Técnico de Arquivo, e dá outras providências. **Diário Oficial da União**: Brasília, DF, 5 jul. 1978. Disponível em:

https://www.planalto.gov.br/ccivil_03/leis/1970-1979/l6546.htm. Acesso em: 10 ago. 2023

BRASIL. Lei nº 8.159, de 8 de janeiro de 1991. Dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências. **Diário Oficial da União**: Brasília, DF, p. 455, 9 jan. 1991. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l8159.htm Acesso em 23 jan. 2024.

BRASIL. Lei nº 12.527, de 18 de novembro de 2011. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências. **Diário Oficial da União**: Brasília, DF, ano 148, n.221-A, p. 1, 18 nov. 2011. Disponível em:

https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm. Acesso em 14 jan. 2024.

BRASIL. Lei nº 12.965, de 23 de abril de 2014. Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. **Diário Oficial da União**: Brasília, DF, ano 151, n. 77, p. 1, 24 abr. 2014. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm. Acesso em: 10 ago. 2023.

BRASIL. Senado Federal. **Legislação Brasileira na Internet**. São Paulo, 2000. Trabalho apresentado no CONIP 2000 - VI Congresso Nacional de Informática Pública, realizado em São Paulo no período de 23-25 de maio de 2000. Disponível em: <https://www.interlegis.leg.br/comunicacao/noticias/2006/03/legislacao-brasileira-na-internet>. Acesso em: 10 ago. 2023.

BRASIL. **Vocabulário Controlado** (vocabulário controlado do Governo Eletrônico - VCGE, interface de comunicação, Interoperabilidade, ePING). *In*: Governo digital. Disponível em: <https://www.gov.br/governodigital/pt-br/governanca-de-dados/vocabulario-controlado-do-governo-eletronico>. Acesso em: 08. fev. 2024.

BREITMAN, K. **Web Semântica: a Internet do futuro**. Rio de Janeiro: LTC, 2005.

BRITO, Thiago Henrique da Silva. **A representação do Frevo na Web Semântica**. Trabalho de Conclusão de Curso (Bacharelado em Gestão da Informação) – Universidade Federal de Pernambuco, Recife, 2018. Disponível em: <https://repositorio.ufpe.br/handle/123456789/46486>. Acesso em: 22 set. 2023.

BROOKS, Terrence. World wide web consortium (W3C). *In*: BATES, Marcia J.; MAACK, Mary Niles (ed.). **Encyclopedia of library and information sciences**. Boca Raton: CRC Press, 2010. p. 5695-5699.

CAFÉ, Lígia; MEDEIROS, Marisa Bräscher Basílio. Organização do Conhecimento: teorias semânticas como base para estudo e representação de conceitos. **Informação & Informação**, Londrina, v. 16, n. 3, p. 25–51, jan./jun. 2011. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/10388>. Acesso em: 23 dez. 2023. DOI: 10.5433/19818920.2011v16nesp.p25.

CALDERON, Wilmara Rodrigues. **O arquivo e a informação arquivística: da literatura científica à prática pedagógica no Brasil**. São Paulo: Cultura Acadêmica, 2013. (Coleção PROPG Digital - UNESP). ISBN 9788579834868. Disponível em: <http://hdl.handle.net/11449/109305>. Acesso em: 01 jun. 2023.

CARLAN, E. **Sistemas de Organização do Conhecimento: uma reflexão no contexto da Ciência da Informação**. 2010. 100f. Dissertação (Mestrado) – Departamento de Ciência da Informação e Documentação, Universidade de Brasília, Brasília, DF, 2010. Disponível em http://www.repositorio.unb.br/bitstream/10482/7465/1/2010_ElianaCarlan.pdf. Acesso em janeiro de 2024.

CARLAN, Eliana; MEDEIROS, Marisa Brascher Basílio. Sistemas de Organização do Conhecimento na visão da Ciência da Informação. **Revista Ibero-Americana de Ciência da**

Informação, Brasília, DF, v. 4, n. 2, p. 53-73, ago./dez. 2011. Disponível em: <https://periodicos.unb.br/index.php/RICI/article/view/1675>. Acesso em: 24 jan. 2024.

CORTES, Maria Regina Persechini Armond. **Arquivo público e informação: acesso à informação nos arquivos públicos estaduais do Brasil**. Belo Horizonte: UFMG/E.B., 1996. Disponível em: https://repositorio.ufmg.br/bitstream/1843/BUOS-96UPHB/1/dissertacao_mariaregina.pdf. Acesso em: 13 fev. 2024.

DAHLBERG, Ingetraut. **Knowledge Organization: A New Science?**. 2006. Disponível em: <https://www.nomos-elibrary.de/10.5771/0943-7444-2006-1-11.pdf>. Acesso em: 24 jan. 2024.

DETERS, J. I.; ADAIME, S. F. Um estudo comparativo dos sistemas de busca na web. *In*: ENCONTRO DE ESTUDANTES DE INFORMÁTICA DO TOCANTINS, 5., 2003, Palmas, TO. **Anais** [...]. Palmas, TO: CEULP/ULBRA, 2003. p. 189-200. Disponível em: <https://ulbra-to.br/encoinfo/wp-content/uploads/2020/03/sistemasbuscaweb-encoinfo2003.pdf>. Acesso em: 09 fev. 2024.

EMYGDIO, Jeanne Louize et al. **Interoperabilidade semântica orientada por ontologia para a Ciência da Informação: a metodologia Onto4All-Interoperability como resultado de estudo de caso no domínio de energia**. 2021. Disponível em: <https://repositorio.ufmg.br/handle/1843/42050>. Acesso em: 02 fev. 2024.

SANTOS, E. P. R. Um ensaio historiográfico sobre a história dos arquivos e da arquivologia no Brasil: balanço e perspectivas. **Acervo: Revista do Arquivo Nacional**, Rio de Janeiro, v. 35, n. 3, p. 1–34, set./dez. 2022. Disponível em: <https://revistaacervo.an.gov.br/index.php/revistaacervo/article/view/1888>. Acesso em: 9 jan. 2024.

FERNÁNDEZ-LÓPEZ, Mariano; GÓMEZ-PÉREZ, Asunción; JURISTO, Natalia. **Methontology: from ontological art towards ontological engineering**. 1997.

FERREIRA, Thiago Castro. *In*: **Aula 1.2: Ontologias e suas Vantagens | Introdução a Ontologias e à Web Semântica**. 2020. Disponível em: <https://www.youtube.com/watch?v=iDlagGj-akU&list=PLLrIHSmC0Mw6IEj060psXrt3BSATBFVzV&index=4>. Acesso em: 23 set. 2023.

GIL, Antonio Carlos. **Métodos e técnicas da pesquisa social**. 5. ed. Atlas: São Paulo, 1999.

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

GUARINO, N. Formal ontology and information systems. PROCEEDINGS OF FORMAL ONTOLOGY AND INFORMATION SYSTEMS. 1998, [s.l.]. **Anais**... [s.l.]. 1998. p. 3–15. Disponível em: https://www.researchgate.net/publication/272169039_Formal_Ontologies_and_Information_Systems. Acesso em: 13 fev. 2024.

GUARINO, N. Understanding, building and using ontologies. **Journal Human-Computer Studies**. Academic Press, v.46, n.2-3, p.293-310, 1997. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S1071581996900919>. Acesso em: 13 fev. 2024.

HENRIQUE, J. M.; SILVA, R. N.; DORNELES, S. L. Olhares sobre a história dos arquivos e da arquivologia no brasil. **Pesquisa Brasileira em Ciência da Informação e Biblioteconomia**, v. 12, n. 1, 2017. DOI: 10.22478/ufpb.1981-0695.2017v12n1.33763 Acesso em: 11 ago. 2023.

HERRERA, A. H. **Archivística general: teoría y práctica – actualizada y aumentada**. Sevilla : Diputación Provincial de Sevilla, 1991. 5a. ed.

HJØRLAND, B. **What is knowledge organization (KO)?** Knowledge Organization, v. 35, n. 2, p. 86-101, 2008. Disponível em: https://www.isko.org/cyclo/knowledge_organization. Acesso em: 22 jan. 2024.

JARDIM, J. M. A Pesquisa em Arquivologia: um cenário em construção. In: VALENTIM, M. L. P., (Org.). **Estudos avançados em Arquivologia**. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2012, p. 135- 153. Disponível em: https://ebooks.marilia.unesp.br/index.php/lab_editorial/catalog/view/36/95/2723. Acesso em: 09 mar. 2024.

JASPER, Robert et al. **A framework for understanding and classifying ontology applications**. In: Proceedings 12th Int. Workshop on Knowledge Acquisition, Modelling, and Management KAW. 1999. p. 16-21. Disponível em: <https://www.cs.man.ac.uk/~horrocks/Teaching/cs646/Papers/uschold99.pdf> Acesso em: 07 de jan. 2024.

KÖCHE, José Carlos. **Fundamentos de pesquisa científica: teoria da ciência e iniciação à pesquisa**. Petrópolis, RJ: Vozes, 2011. Disponível em: https://btux.com.br/professorbruno/wp-content/uploads/sites/10/2018/07/K%C3%B6che-Jos%C3%A9-Carlos0D0AFundamentos-de-metodologia-cient%C3%ADfica-_-teoria-da0D0Acia-e-inicia%C3%A7%C3%A3o-%C3%A0-pesquisa.pdf Acesso em: 07 fev. 2024.

LAKATOS, Eva Maria; MARCONI, Marina de Andrade. **Metodologia científica**. São Paulo: Atlas, v. 3, 1991.

LIMA, G. A. Organização e representação do conhecimento e da informação na web: teorias e técnicas. **Perspectivas em Ciência da Informação**, [S. l.], p. 57–97, 2020. Disponível em: <https://periodicos.ufmg.br/index.php/pci/article/view/22283> Acesso em: 13 ago. 2023.

LIMA, Carlos Eduardo Teixeira. **Um Chatbot para criação e desenvolvimento de ontologias com lógica de descrição**. 2017. Dissertação de Mestrado. Universidade Federal de Pernambuco. Disponível em: <https://repositorio.ufpe.br/handle/123456789/25236>. Acesso em: 04 fev. 2024.

LÖW, Marieta Marks. Proveniência arquivística e semântica: uso de ontologias de BFO para representação da produção documental e contexto. 2021. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Biblioteconomia e Comunicação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2021. Disponível em:

<https://www.lume.ufrgs.br/handle/10183/238060>. Acesso em: 09 mar. 2024.

MACEDO, I. M.; LIMA, G. L. Q. Formações históricas da teoria arquivística no Brasil: uma revisão a partir do surgimento dos cursos de bacharelado em arquivologia. **Páginas A&B, Arquivos e Bibliotecas (Portugal)**, n. 13, p. 46-66, 2020. Disponível em:

<http://hdl.handle.net/20.500.11959/brapci/145269>. Acesso em: 13 ago. 2023.

MARTINS, Maria Cleci Coti. Menção honrosa. Ontologia legal: estudo sobre a modelagem do conhecimento legal no contexto do direito tributário. Capítulo do livro: **Administração Pública: Prêmio de Criatividade e Inovação Auditor-Fiscal da Receita Federal José Antônio Schöntag: 5º Prêmio Schöntag: monografias premiadas**. Brasília: SRF, 2006. Disponível em: <https://repositorio.enap.gov.br/handle/1/4563>. Acesso em 25 fev. 2024.

MAZZOCCHI, Fulvio. **Knowledge organization system (KOS): an introductory critical account. KO Knowledge Organization**, v. 45, n. 1, p. 54-78, 2018. Disponível em: <https://www.nomos-elibrary.de/10.5771/0943-7444-2018-1-54.pdf>. Acesso em: 13 de jan. 2024.

MENDONÇA, F. M. **OntoForInfoScience: metodologia para construção de ontologias pelos cientistas da informação: uma aplicação prática no desenvolvimento da ontologia sobre componentes do sangue humano (Hemonto)**. 2015. Tese (Doutorado)- Universidade Federal de Minas Gerais, Belo Horizonte, Brasil, 2015. Disponível em: <http://hdl.handle.net/1843/BUBD-A35H3K>. Acesso em: 13 ago. 2023.

MEYER, Stephenie. **Twilight**. Boston: Little, Brown and Company, 2005.)

MINTEGUI, E.; KARPINSKI, C.; BARROS, T. H. B. As formações discursivas da arquivologia na formulação e proposta de alteração da lei de arquivos brasileira. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 20., 2019, Florianópolis. **Anais [...]**. Florianópolis: ANCIB, 2019. Disponível em: <https://lume.ufrgs.br/handle/10183/202757>. Acesso em: 13 ago. 2023.

MORAIS, Edison Andrade Martins; AMBRÓSIO, Ana Paula L. Ontologias: conceitos, usos, tipos, metodologias, ferramentas e linguagens. **Relatório Técnico–RT-INF-001/07**, 2007. Disponível em: https://ww2.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF_001-07.pdf. Acesso em: 20 fev. 2024.

NASCIMENTO, Felipe Mozart de Santana. **Uso Estratégico da Ontologia para Organização e Gestão da Informação Jurídica**. 2018. 105f. Dissertação (Mestrado em Ciência da Informação) – Universidade Federal de Pernambuco, Recife, 2018. Disponível em: <https://repositorio.ufpe.br/handle/123456789/29646>. Acesso em 14 dez. 2023.

NASCIMENTO, Felipe Mozart de Santana. **Sistema de Organização do Conhecimento: um mapeamento de conceitos e processos para a Ciência da Informação no Brasil**. 2022. Tese (Doutorado em Ciência da Informação) - Universidade Federal de Pernambuco, Recife, 2022.

NASCIMENTO, F. P. **Metodologia da pesquisa científica: teoria e prática: como elaborar TCC**. Brasília: Thesaurus, 2016.

NASCIMENTO, M. B. VENÂNCIO, R. P. Instrumentos de gestão de documentos na legislação brasileira, 1978-2021. **Perspectivas em Ciência da Informação**, v. 27, 2022. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/212016>. Acesso em: 19 abr. 2023.

NOY, N. F.; MCGUINNESS, D. L. **Ontology Development 101: A Guide to Creating Your First Ontology**. Disponível em: https://protege.stanford.edu/publications/ontology_development/ontology101.pdf. Acesso em: 02 jan 2024.

PINHEIRO, Lena Vania Ribeiro; FERREZ, Helena Dodd. **Tesouro Brasileiro de Ciência da Informação**. Rio de Janeiro; Brasília, DF: IBICT, 2014. Disponível em: <http://sitehistorico.ibict.br/publicacoes-e-institucionais/tesouro-brasileiro-de-ciencia-da-informacao-1/tesouro-brasileiro-de-ciencia-da-informacao/resolveuid/1c60ede36f47aee60c48957ef6db7510>. Acesso em: 9 mar. 2024.

PINTO, J. A.; ALMEIDA, M. B. Ontologias públicas sobre governo eletrônico: Uma Revisão Sistemática da Literatura. **Brazilian Journal of Information Science: research trends**, [S. l.], v. 14, n. 3 - jul-set, p. e020003, 2020. DOI: 10.36311/1940-1640.2020.v14n3.10105. Disponível em: <https://revistas.marilia.unesp.br/index.php/bjis/article/view/10105>. Acesso em: 22 ago. 2023.

PONTES, Carlos Eduardo. **Legislação arquivística brasileira: suficiência e aplicação**. 2015.

Poveda-Villalón, María, Asunción Gómez-Pérez, and Mari Carmen Suárez-Figueroa. "OOPS!(Ontology Pitfall Scanner!): An on-line tool for ontology evaluation." **International Journal on Semantic Web and Information Systems (IJSWIS)** 10.2 (2014): 7-34.

PROTEGE. **Ontology editor and knowledge-base framework**. Disponível em: <http://protege.stanford.edu/>. Acesso em: 18 ago. 2023.

REIS, L. O arquivo e arquivística evolução histórica. **Biblios**, v. 7, n. 24, abril-junho, 2006. Disponível em: <https://www.redalyc.org/articulo.oa?id=16172402>. Acesso em: 06 ago. 2023.

SANTOS, Cristina Mamédio da Costa; PIMENTA, Cibele Andrucioli de Mattos; NOBRE, Moacyr Roberto Cuce. A estratégia PICO para a construção da pergunta de pesquisa e busca de evidências. **Revista latino-americana de enfermagem**, v. 15, p. 508-511, 2007.

SANTOS, H. P. História e Arquivologia: relação, individualização e mídias contemporâneas. *In: ENCONTRO ESTADUAL DE HISTÓRIA*, 8., 2016, Feira de Santana. **Anais [...]**. Bahia: Associação Nacional de História, 2016. Disponível em:

https://www.encontro2016.bahia.anpuh.org/resources/anais/49/1473775054_ARQUIVO_ArquivologiaeHistoria_relacionamento_COMPLETO.pdf. Acesso em: 15 ago. 2023.

SILVA, M. **A polissemia do termo "arquivo"**. In: XVII Encontro Nacional de Pesquisa em Ciência da Informação. 2016. Disponível em: <http://www.ufpb.br/evento/lti/ocs/index.php/enancib2016/enancib2016/paper/viewFile/38799/2444>. Acesso em: 24 jul. 2023.

SILVEIRA, Denise Tolfo; CÓRDOVA, Fernanda Peixoto. **A pesquisa científica. Métodos de pesquisa**. Porto Alegre: Editora da UFRGS, 2009. p. 33-44, 2009. Disponível em: <https://www.lume.ufrgs.br/bitstream/handle/10183/213838/000728731.pdf>. Acesso em: 13 fev. 2024.

SOUZA, Louise Gabler. O Arquivo Público do Império como lugar de memória: a administração de Joaquim Pires Machado Portela (1873-1898). In: XXVII Simpósio Nacional de História. Rio Grande do Norte: UFRN, 2013. Disponível em: http://mapa.an.gov.br/images/Arquivo_Nacional_Machado_Portela.pdf. Acesso em: 26 fev. 2024.

TELINÉ, Maria Fernanda; ALMEIDA, Gladis Maria de Barcellos; ALUÍSIO, Sandra Maria. Extração manual e automática de terminologia: comparando abordagens e critérios. In: **16th Brazilian Symposium on Computer Graphics and Image Processing-SIBGRAPI**. 2003. Disponível em: http://www.nilc.icmc.usp.br/til/til2003_English/oral/Teline_Almeida_Aluisio_37.pdf. Acesso em: 20 fev. 2024.

USCHOLD, Mike; GRUNINGER, Michael. Ontologies: Principles, methods and applications. **The knowledge engineering review**, v. 11, n. 2, p. 93-136, 1996.

APÊNDICE A – TAXONOMIA DA ARQUIVONTOLOGY

arquivística

politica_nacional_arquivos

arquivo

arquivo_privado

arquivo_entidade_religiosa

arquivo_publico

arquivo_poder_legislativo

arquivo_municipal

arquivo_estadual

arquivo_distrito_federal

arquivo_federal

arquivo_poder_executivo

arquivo_municipal

arquivo_estadual

arquivo_distrito_federal

arquivo_federal

arquivo_poder_judiciario

arquivo_estadual

arquivo_distrito_federal

arquivo_federal

arquivo_ministerio_relacoes_exteriores

arquivo_ministerio_marinha

arquivo_ministerio_exercito

arquivo_ministerio_aeronautica

arquivo_nacional

conselho_nacional_arquivos

sistema_nacional_arquivos

arquivo_administrativo

arquivo_corrente

arquivo_intermediario

arquivo_permanente

instituicao_arquivistica

instituicao_arquivistica_privada

instituicao_arquivistica_publica

instituicao_arquivistica_municipal

instituicao_arquivistica_estadual

instituicao_arquivistica_distrito_federal

instituicao_arquivistica_federal

arquivo_nacional

documentacao

documento_privado

documento_pessoal

interesse_publico_social

documento_publico

documento_poder_legislativo

documento_poder_executivo

documento_poder_judiciario

documento_classificado

documento_corrente

documento_intermediario

documento_permanente

documento_ostensivo

documento_restrito

documento_sigiloso

administracao_arquivos

administracao_arquivo_privado

jurisdicao_arquivistica

administracao_arquivo_publico

jurisdicao_arquivistica

imprescritibilidade***inalienabilidade***

protecao_legal_arquivos

sucessao_arquivistica

heranca_fundos

gestao_documentos

acesso

uso

fonte_referencia

disseminacao_informacao

informacao

restricao_acesso

protecao_dados

sigilo

grau_sigilo

violacao_sigilo

acumulacao

entrada_documento

doacao

emprestimo

recolhimento

transferencia

arquivamento

avaliacao

fase_corrente

fase_intermediaria

fase_permanente

valor_administrativo

valor_intrinseco

valor_informativo

- valor_permanente
 - valor_probatorio
 - guarda_permanente
- consulta
- custodia
 - entidade_custodiadora
 - alienacao
 - inalienabilidade
 - imprescritibilidade
- eliminacao
- preservacao
 - protecao_legal_arquivos
 - destruicao
 - desfiguracao
 - perda_unidade_documental
- processamento_tecnico
 - classificacao
 - natureza_conteudo
 - suporte
- producao
 - entidade_produtores
 - organicidade
- tramitacao

APÊNDICE B – TRECHO DA ARQUIVONTOLOGY EM OWL

```

<?xml version="1.0"?>
<rdf:RDF xmlns="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024/"
xml:base="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024/"
xmlns:owl="http://www.w3.org/2002/07/owl#"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
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#"
xmlns:arquivontology="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024#">
<owl:Ontology rdf:about="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024"/>
<!-- ////////////////////////////////// Annotation properties ////////////////////////////////// -->
  <!-- http://www.semanticweb.org/jessicaoliveira/arquivontology/2024/definition -->
  <owl:AnnotationProperty
rdf:about="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024/definition"/>
  <!-- http://www.semanticweb.org/jessicaoliveira/arquivontology/2024/synonyms -->
  <owl:AnnotationProperty
rdf:about="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024/synonyms"/>
  <!-- ////////////////////////////////// Object Properties ////////////////////////////////// -->
  <!-- http://www.semanticweb.org/jessicaoliveira/arquivontology/2024#ato_da -->
  <owl:ObjectProperty rdf:about="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024#ato_da">
  <rdfs:domain rdf:resource="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024#recolhimento"/>
  <rdfs:domain rdf:resource="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024#transferencia"/>
  <rdfs:domain
rdf:resource="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024/entrada_documentos"/>
  <rdfs:range rdf:resource="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024/acumulacao"/>
  <rdfs:label xml:lang="pt-br">ato da</rdfs:label>
</owl:ObjectProperty>
  <!-- http://www.semanticweb.org/jessicaoliveira/arquivontology/2024#atribuicao_da -->
  <owl:ObjectProperty
rdf:about="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024#atribuicao_da">

```

```

<rdfs:domain
rdf:resource="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024/administracao_arquivo_privado"
/>

<rdfs:domain
rdf:resource="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024/administracao_arquivo_publico"
/>

<rdfs:domain
rdf:resource="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024/protecao_legal_arquivos"/>

<rdfs:domain
rdf:resource="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024/sucessao_arquivistica"/>

<rdfs:range
rdf:resource="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024/administracao_arquivos"/>

<rdfs:label xml:lang="pt-br">atribuição da</rdfs:label>

</owl:ObjectProperty>

<!-- http://www.semanticweb.org/jessicaoliveira/arquivontology/2024#caracterizado_como -->

<owl:ObjectProperty
rdf:about="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024#caracterizado_como">

<rdfs:domain
rdf:resource="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024#documento_ostensivo"/>

<rdfs:domain
rdf:resource="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024#documento_restrito"/>

<rdfs:domain
rdf:resource="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024#documento_sigiloso"/>

<rdfs:range
rdf:resource="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024/documento_classificado"/>

<rdfs:label xml:lang="pt-br">carcterizado como</rdfs:label>

</owl:ObjectProperty>

<!-- http://www.semanticweb.org/jessicaoliveira/arquivontology/2024#classificado_pela -->

<owl:ObjectProperty
rdf:about="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024#classificado_pela">

<rdfs:domain
rdf:resource="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024#natureza_conteudo"/>

<rdfs:domain rdf:resource="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024#suporte"/>

<rdfs:range rdf:resource="http://www.semanticweb.org/jessicaoliveira/arquivontology/2024#classificacao"/>

<rdfs:label xml:lang="pt-br">é classificado pela</rdfs:label>

(...)

<!-- Generated by the OWL API (version 4.5.26.2023-07-17T20:34:13Z) https://github.com/owlcs/owlapi -->

```