

Ordinance ontology: ontologia para representação de portarias emitidas por instituições federais de ensino

Ordinance ontology: ontology for representing ordinances issued by education federal institutions

Anderson Veiga da SILVA¹  0000-0002-9788-8354

Mark Douglas de Azevedo JACYNTHO²  0000-0003-3910-1442

Resumo

As instituições federais de ensino do Brasil geram e publicam uma grande quantidade de documentos eletrônicos para as mais diversas atividades realizadas. Existem vários tipos de documentos, cada um com uma finalidade e, dentre eles, as portarias são utilizadas para, entre outras coisas, oficializar atos administrativos que definem competências e estabelecem normas e procedimentos. Para exemplificar, apenas no ano de 2019 o Instituto Federal Fluminense publicou cerca de 2500 portarias e essas, assim como os outros tipos de documentos oficiais, são publicadas em formato de documento de texto simples, sem qualquer tipo de estruturação do conteúdo, o que impossibilita a realização de consultas refinadas baseadas no conteúdo descrito. Para permitir a estruturação do conteúdo, a *Web Semântica* fornece um conjunto de tecnologias e padrões para representação formal de informações relacionadas a determinado domínio de conhecimento. Essa representação formal é baseada em um modelo chamado ontologia, que permite descrever objetos e propriedades do domínio de conhecimento, além de possibilitar a inferência de novos conhecimentos com base em axiomas pré-estabelecidos. Este artigo apresenta o desenvolvimento e a validação de uma ontologia para descrever as portarias publicadas por instituições federais de ensino, a fim de possibilitar a estruturação do conhecimento contido dentro desses documentos e, conseqüentemente, permitir o uso dessas informações estruturadas para a realização de pesquisas precisas mais rebuscadas. É esperado que este trabalho forneça uma base para o desenvolvimento de novas ontologias que possibilitem a descrição dos demais tipos de documentos publicados por instituições públicas de ensino.

Palavras-chave: Instituições Federais de Ensino. Ontologia. Portarias. *Web Semântica*.

Abstract

Federal educational institutions in Brazil generate and publish a large number of electronic documents for most of the activities performed. There are several kinds of documents, each one with a purpose, and, among these kinds, ordinances are used to formalize administrative

¹ Instituto Federal Fluminense, Coordenação do Curso Técnico em Informática. Av. Dário Vieira Borges, 235, Parque do Trevo, Bom Jesus do Itabapoana, Rio de Janeiro, Brasil. Correspondência para/Correspondence to: A. V. SILVA. E-mail: anderson.silva@ifff.edu.br

² Instituto Federal Fluminense, Programa de Pós-Graduação em Sistemas Aplicados à Engenharia e Gestão. Campos dos Goytacazes, Rio de Janeiro, Brasil.

Artigo elaborado a partir de dissertação de A. V. SILVA, intitulada "Proposição de ontologia para representação de portarias emitidas por instituições federais de ensino". Instituto Federal Fluminense, 2021.

Recebido em 1 de novembro de 2021 e aprovado em 17 de novembro de 2021.

Como citar este artigo/How to cite this article

Silva, A. V.; Jacyntho, M. D. A. *Ordinance ontology: ontologia para representação de portarias emitidas por instituições federais de ensino. Transinformação*, v. 34, e210072, 2022. <http://doi.org/10.1590/2318-0889202234e210072>



acts that define competencies and establish rules and procedures, among other finalities. To exemplify, in 2019, the Fluminense Federal Institute published around 2500 ordinances and these, as well as the other kinds of official documents, are published in plain text format, without any type of content structuring, which makes it impossible to perform refined queries based on the described content. To allow content structuring, the Semantic Web provides a set of technologies and standards necessary for representing information related to a given domain of knowledge. This formal representation is defined as ontology and it allows to describe objects and properties of a domain of knowledge, in addition to enabling inference of new knowledge based on pre-established axioms. This paper presents the development and validation of an ontology to describe ordinances published by federal educational institutions, in order to enable the structuring of the knowledge contained within these documents and, thus, to allow the use of this structured information to perform refined searches. It is also expected that this work provides a basis for the development of new ontologies, enabling descriptions of the other kinds of documents published by public educational institutions.

Keywords: Federal Educational Institutions. Ontology. Ordinances. Semantic Web.

Introdução

No Brasil, existem vários órgãos e instituições públicas que disponibilizam informações na *web* através de documentos de texto com grande valor agregado, porém essas informações não possuem uma estrutura formal que facilite seu processamento por um agente computacional, havendo sempre a necessidade de interpretação por um agente humano.

As instituições federais de ensino do Brasil são autarquias federais que atuam em diversos níveis e modalidades de ensino, desde o ensino básico até a pós-graduação. Por serem órgãos ligados à administração pública, toda comunicação oficial deve ser feita por meio de publicação de documentos que seguem diretrizes e regras previamente estabelecidas em manuais de redação oficial.

Existem vários tipos de documentos que podem ser elaborados pelas instituições federais de ensino com a finalidade de publicizar atos normativos ou disponibilizar informações de interesse público, como por exemplo: resoluções, pareceres, ordens de serviço, editais, portarias, dentre outros. As portarias, em especial, são documentos emitidos por órgãos públicos e utilizados como instrumento pelo qual uma autoridade pública expede instruções sobre a organização e o funcionamento de serviço, sobre questões de pessoal e outros atos de sua competência (Brasil, 2018).

No contexto educacional, as portarias são utilizadas como documento oficial para publicizar atos administrativos com diversas finalidades, como, por exemplo, conceder promoções e progressões, autorizar afastamentos de servidores, designar chefia para um determinado setor, nomear novos servidores, dentre outros. Porém, as portarias, assim como os diversos tipos de documentos gerados por essas instituições, normalmente são publicadas como arquivos de texto no formato *Portable Document Format* (PDF), sem qualquer tipo de estruturação formal do conteúdo.

Nesse cenário, existem documentos de texto que são publicados com informações não estruturadas e que dependem única e exclusivamente da interpretação humana para que seja extraído valor dessas informações. Seria ideal que agentes computacionais tivessem também a capacidade de compreender o significado (semântica) das informações contidas nesses documentos, a fim de extrair informações úteis deles, pois a capacidade de processamento de agentes computacionais é muito maior do que a de um ser humano, o que permite que eles sejam capazes de analisar uma quantidade bem maior de informações em menos tempo. Foi para atingir esse objetivo que a *Web* semântica foi proposta por Tim Berners-Lee (Berners-Lee; Hendler; Lassila, 2001).

A *Web* Semântica pode ser vista como uma coleção de padrões e tecnologias que permitem às máquinas entenderem o significado das informações publicadas na *web*, fazendo com que essas possam ser processadas automaticamente em larga escala (Yu, 2014).

Para permitir que dados disponibilizados na *web* fossem estruturados, o *World Wide Web Consortium* (W3C) criou um modelo de representação de dados chamado *Resource Description Framework* (RDF). O RDF é um modelo de dados em grafo usado para representação de informações sobre recursos na *web* e destina-se particularmente a representar metadados sobre esses recursos (Manola; Miller, 2004), que são descritos em RDF por meio de triplas compostas por sujeito (recurso), predicado (propriedade) e objeto (valor).

A descrição de recursos utilizando o padrão RDF deve utilizar termos e propriedades que possam ser compartilhados com outros recursos. Esses termos e propriedades são pré-determinados em um vocabulário comum, permitindo que recursos de diferentes fontes possam ser analisados por um agente computacional de forma equivalente. Esses vocabulários são também chamados de ontologias. A rigor, ontologia é um modelo de representação formal do conhecimento de um domínio, onde um conjunto de conceitos e seus relacionamentos são descritos e, em seguida, publicados como termos de um vocabulário (Breitman; Casanova; Truskowski, 2007).

Assim, com o auxílio das tecnologias relacionadas à *Web Semântica*, é possível desenvolver sistemas computacionais que consigam analisar e interpretar informações contidas em qualquer tipo de documento devidamente estruturado (como as portarias, por exemplo), e, assim, permitir a realização de buscas mais elaboradas e semânticas.

Como exemplo, o Instituto Federal Fluminense (IFF) é uma das diversas instituições federais de ensino existentes no Brasil que publica as portarias emitidas no formato de documento de texto. Nesse caso, as portarias são divulgadas no endereço eletrônico do Centro de Documentação Digital (CDD) e, através de um levantamento próprio realizado nesse endereço, pode-se constatar que, apenas no ano de 2019, cerca de 2500 portarias foram publicadas. Essa grande quantidade de portarias possui conhecimentos importantes subjacentes que estão obscuros em documentos de texto, sem qualquer tipo de organização formal. As portarias possuem itens que determinam qual a sua finalidade ou seu tipo, quais servidores são citados, a data de sua publicação, dentre outras informações específicas de cada tipo de portaria. Para realizar uma consulta que retorne todas as portarias de progressão e promoção relacionadas a um determinado servidor, por exemplo, é preciso realizar o trabalho manual de verificação do conteúdo de cada documento até encontrar todos os que satisfaçam os critérios dessa busca.

Esse cenário expõe claramente a necessidade de se estruturar formalmente as informações relevantes contidas nas portarias publicadas, pois existem diversos critérios de busca que podem ser combinados para a realização de consultas ainda mais específicas.

Este artigo propõe, tendo como base as portarias publicadas pelo Instituto Federal Fluminense, uma ontologia acerca das portarias publicadas por instituições federais de ensino do Brasil com o objetivo de possibilitar a estruturação formal do conhecimento contido dentro desses documentos e, conseqüentemente, permitir que agentes computacionais realizem buscas mais precisas e processem esse conhecimento, utilizando as tecnologias e padrões relacionados à *Web Semântica*.

Fundamentos

Web Semântica

A *Web Semântica* é uma extensão da *web* convencional. Ela é construída vinculando as páginas da *web* original a um conjunto de dados (semi)estruturados que indica a semântica da página vinculada. Um agente inteligente, que seja capaz de entender esse conjunto de dados estruturados, pode, em escala global, realizar ações inteligentes e tomar decisões fundamentadas (Yu, 2014).

A ideia da *Web Semântica* é definir um modelo padrão para a representação de dados, combinado com um conjunto de vocabulários (ontologias) de uso comum, permitindo que a semântica dos dados possa ser anexada a eles de forma a construir um ambiente mais homogêneo (Laufer, 2015).

A *Web Semântica* pode ser vista como um conjunto de padrões e tecnologias que torna possível que os dados publicados na *web* sejam inteligíveis por máquinas, dando a elas a capacidade de interpretar e relacionar as informações com base em ontologias, regras e axiomas pré-estabelecidos. Além disso, com base na descrição e modelagem feita de um determinado domínio de conhecimento, a *Web semântica* também fornece às máquinas o poder de inferir novas informações com base nas informações já descritas.

Resource Description Framework

O *Resource Description Framework* (RDF) é um modelo abstrato de descrição de metadados. Apesar do RDF possuir o termo *framework* (arcabouço) em seu nome, ele é um modelo de dados e, portanto, é diferente dos frameworks normalmente utilizados na computação. O RDF pode ser considerado um framework de conceitos.

Segundo Jacyntho e Schwabe (2016, p. 1, tradução nossa):

Um modelo de dados é apenas uma forma de visualização dos dados. O já estabelecido modelo relacional vê os dados através de relações e tuplas. O modelo em grafos RDF, baseado em triplas, é uma representação natural de vários tipos de aplicações, onde entidades são fortemente conectadas entre si. Em contraste com sistemas de gerenciamento de bancos de dados legados, essas aplicações consideram as propriedades multivaloradas tão desejáveis na modelagem de dados reais que trabalham com propriedades multivaloradas por padrão. Consultas de propriedades multivaloradas ou de valor único são feitas exatamente do mesmo modo, sem preocupações com a necessidade de se associar a uma terceira tabela para modelar um relacionamento n-para-n.

Segundo Breitman, Casanova e Truskowski (2007), o RDF é uma linguagem de propósito geral para representar informações sobre recursos na *web*. Destina-se principalmente a representar metadados sobre recursos *web*, mas também pode ser utilizado para representar informações sobre objetos que podem ser identificados na *web*, mesmo quando eles não podem ser recuperados diretamente dela.

Por ser um modelo abstrato, o RDF independe de qualquer tipo de sintaxe. Os conceitos fundamentais relacionados ao modelo RDF são: *statements*, recursos e propriedades.

Statement (declaração, afirmação), no contexto da *Web Semântica*, representa uma unidade fundamental de conhecimento descrita no modelo RDF. Essa unidade é representada por uma tripla, composta da seguinte estrutura: “recurso – propriedade – valor” ou “sujeito – predicado – objeto”. Recursos (sujeitos) representam qualquer coisa do mundo real, identificada por um endereço *web* (URI), cujas propriedades possam ser descritas; as propriedades (predicados) representam os relacionamentos entre um recurso e um determinado valor e os valores (objetos) podem ser qualquer valor literal ou até mesmo outros recursos. As propriedades também são identificadas por URIs e são descritas em ontologias.

SPARQL

O W3C definiu um conjunto de especificações chamado *SPARQL Protocol and RDF Query Language* (SPARQL), que provê linguagens e protocolos para consultar e manipular dados de grafos RDF na *web* ou em bases de dados RDF (The W3C SPARQL Working Group, 2013).

A linguagem de consulta SQL está para o modelo de dados relacional assim como a linguagem de consulta SPARQL está para o modelo de dados RDF. Enquanto a linguagem SQL realiza consultas baseadas na seleção de tuplas do produto cartesiano entre relações, a linguagem SPARQL realiza consultas fazendo seleções por meio de modelos denominados *triple pattern* e *graph pattern*.

Triple patterns são semelhantes a triplas RDF, exceto que cada recurso, propriedade ou valor pode ser substituído por variáveis. Um *graph pattern* corresponde a um conjunto de *triple patterns*. Utilizando *graph pattern* é possível empregar uma mesma variável em diferentes *triple patterns*. O resultado de um *graph pattern* corresponde a um subgrafo que combina com os critérios definidos na consulta (Harris; Seaborne; Prud'hommeaux, 2013).

Ontologias

RDF é um modelo de dados semântico. Em outras palavras, para reduzir a ambiguidade de recursos e propriedades, especialmente quando se realiza a integração de triplas oriundas de diferentes fontes de dados, é preciso (re)utilizar um vocabulário (ontologia) que defina formalmente conceitos e relacionamentos comuns entre eles, permitindo que a máquina interprete, inequivocamente, os dados (Alsharif, 2013).

Uma ontologia representa formalmente um determinado domínio de conhecimento, definindo seus principais conceitos (classes) e relacionamentos entre eles. Outros relacionamentos também são encontrados nas ontologias, definidos por meio de propriedades que descrevem características ou atributos das classes e que relacionam instâncias das classes. Assim, uma ontologia é um conjunto de axiomas que codifica o conhecimento de um determinado domínio em um formato estruturado formal que possibilita que a máquina o entenda (Jacyntho, 2012).

Segundo Noy e McGuinness (2001), uma ontologia representa um determinado domínio de conhecimento e consiste em: classes, que representam os principais conceitos do domínio; propriedades, que representam relacionamentos e atributos das classes, e axiomas, que são regras de restrição sobre o domínio e instâncias das classes.

Para criação de ontologias padrão, o W3C propõe duas linguagens de criação de ontologias (meta-ontologias), a saber: *RDF Schema* (RDFS) e *Ontology Web Language* (OWL). A *RDF Schema* é uma extensão semântica do RDF. Ela provê mecanismos para descrever recursos e suas relações (Brickley; Guha, 2014) e é usada para construir ontologias leves em RDF. De forma simples, vocabulários descritos através da RDFS consistem em definições de classes e propriedades (Heath; Bizer, 2011).

A RDFS fornece mecanismos para descrever classes e propriedades, porém possui algumas limitações. A *Ontology Web Language* (OWL) foi concebida para ser uma linguagem que aumenta o poder de expressividade de *RDF Schema* e, conseqüentemente, fornece meios para que relações mais complexas possam ser descritas, aumentando o poder de inferência de uma ontologia.

A OWL é uma linguagem projetada para representar conhecimento rico e complexo sobre coisas, bem como seus relacionamentos. É uma linguagem baseada em lógica computacional, de forma que todo conhecimento expresso em OWL pode ser raciocinado por programas de computadores, seja para verificar a consistência desse conhecimento ou para tornar explícito conhecimentos implícitos (Hitzler *et al.*, 2012). O propósito da OWL é o mesmo da *RDF Schema*: definir ontologias que incluem classes, propriedades e relacionamentos para um determinado domínio de conhecimento. A OWL, no entanto, fornece maior expressividade, permitindo a construção de aplicações que possuem a capacidade de raciocinar e inferir novas informações. A OWL pode ser vista como a junção de *RDF Schema* com novos termos que fornecem melhor expressividade. Dessa forma, todos os termos presentes na *RDF Schema* também podem ser utilizados em documentos descritos em OWL (Yu, 2014).

Segundo Jacyntho e Moraes (2021), as ontologias trazem dois diferentes benefícios:

- Fornecem semântica explícita aos dados e, por conseguinte, poder de inferência para as máquinas, possibilitando, através de raciocinadores, a descoberta de novos conhecimentos;
- Promovem o compartilhamento e reuso de um vocabulário, garantindo um entendimento comum de um domínio de conhecimento.

Trajatória metodológica ao longo do desenvolvimento da ontologia de portarias proposta

Nesta seção é apresentado o processo de desenvolvimento da ontologia referente ao domínio de conhecimento das portarias publicadas por instituições federais de ensino. Para atingir esse objetivo, foi adotado o método de desenvolvimento *Ontology Development 101*.

O método *Ontology Development 101* foi concebido por Noy e McGuinness (2001) e se propõe a ser um guia prático e direto para construção de ontologias. Esse método é composto de sete etapas simples que visam nortear e cobrir todo o processo de desenvolvimento de uma ontologia, a saber:

- 1) Determinar o escopo da ontologia;
- 2) Considerar o reuso de ontologias existentes;
- 3) Enumerar termos importantes do domínio de conhecimento;
- 4) Definir a hierarquia de classes;
- 5) Definir as propriedades da ontologia;
- 6) Definir as restrições sobre as propriedades;
- 7) Criar instâncias.

Definição do escopo

A ontologia desenvolvida neste trabalho utiliza como referência as portarias emitidas pelo Instituto Federal Fluminense, as quais são publicadas no portal do Centro de Documentação Digital, que pode ser acessado no seguinte endereço eletrônico: <<http://cdd.iff.edu.br/documentos/portarias>>.

Com base na análise de portarias publicadas, foram identificadas as seguintes informações básicas presentes em portarias e que são relevantes para a ontologia proposta:

- Número: identificador numérico da portaria;
- Data: data de publicação da portaria;
- Setor: setor que publicou a portaria;
- Assinante/Emissor: servidor responsável por assinar/emitir a portaria;
- Considerações: fundamentações que validam os atos;
- Atos: atos normativos presentes na portaria;
- Tipo: classificação de cada ato normativo presente em uma portaria;
- Endereço: endereço eletrônico do arquivo de texto publicado referente à portaria.

Outra forma de definir o escopo é por meio da elaboração de um conjunto de Questões de Competência (QC) para as quais a ontologia deve ser capaz de oferecer respostas. A seguir são apresentadas as questões de competências elencadas para a ontologia de portarias.

- QC1: Quais portarias foram publicadas em um determinado intervalo de tempo?
- QC2: Quais portarias foram publicadas por um determinado setor do instituto?
- QC3: Quais portarias foram publicadas com uma determinada finalidade (tipo)?
- QC4: Quais portarias relacionam-se a um determinado servidor?
- QC5: Quais portarias foram assinadas por determinada autoridade competente?
- QC6: Quais portarias referenciam outras portarias?
- QC7: Quais portarias possuem uma determinada palavra-chave em seu conteúdo?
- QC8: Quais portarias possuem uma determinada condição ou consideração?
- QC9: Quais portarias foram demandadas por um determinado setor?

Com base nas questões de competência levantadas, foram definidas novas informações, presentes nas portarias, que precisavam ser modeladas na ontologia a fim de complementar aquelas levantadas inicialmente. A seguir são descritas essas novas informações:

- Documento: documento oficial que pode ser referenciado em uma portaria;
- Mandato: mandato exercido por um servidor responsável por emitir portarias;
- Referência: referência a qualquer recurso, como pessoas, documentos, entre outros;
- Servidor: servidor público referenciado em algum ato normativo;
- Vínculo: vínculo empregatício associado a um servidor público.

Foi adotado o nome *Ordinance Ontology* (Ontologia de Portarias) para a ontologia proposta neste trabalho e o prefixo ord: foi utilizado para representação do seu respectivo *namespace*: <http://purl.org/ordinance-ontology/>.

De acordo com o processo de desenvolvimento de ontologias adotado, após a definição do escopo é preciso considerar a reutilização de classes e propriedades presentes em ontologias já existentes. A reutilização de ontologias evita a duplicidade de termos para representar entidades do mesmo domínio e permite uma maior interoperabilidade entre sistemas semânticos, visto que utilizam termos de um vocabulário comum.

Foram feitas pesquisas no catálogo *Linked Open Vocabularies* (LOV) a fim de identificar termos em ontologias já existentes que pudessem ser reaproveitados no domínio das portarias. Como resultados dessas pesquisas, algumas classes e propriedades de outras ontologias foram utilizadas no processo de construção da *Ordinance Ontology*.

O Quadro 1 apresenta cada uma das ontologias reutilizadas no desenvolvimento deste trabalho, associadas aos seus respectivos prefixos e *namespaces*.

Após o levantamento das ontologias a serem reusadas, foram definidos os termos pertinentes que deveriam estar presentes na ontologia.

O levantamento desses termos é importante para identificar possíveis classes e propriedades que serão definidas na ontologia e, no processo de definição do escopo da ontologia, foi possível enumerar termos relevantes, como, por exemplo: portarias, atos, condições, setores, emissores, data de publicação, entre outros.

Quadro 1 – Lista de ontologias reutilizadas no processo de desenvolvimento.

Prefixo	Namespace	Ontologia	Propósito
foaf	http://xmlns.com/foaf/0.1/	<i>Friend of a Friend</i>	É um vocabulário utilizado para descrever pessoas, suas atividades e suas relações com outras pessoas ou objetos (Brickley; Miller, 2014).
org	http://www.w3.org/ns/org#	<i>The Organization Ontology</i>	É um vocabulário projetado para permitir a publicação de informações em organizações e em estruturas organizacionais, incluindo organizações governamentais. Seu objetivo é prover uma ontologia genérica e reutilizável que possa ser estendida ou especializada em situações particulares (Reynolds, 2014).
schema	http://schema.org/	<i>Schema.org</i>	Vocabulário utilizado para descrever os mais variados tipos de informações em sites com o intuito de auxiliar os mecanismos de busca.
skos	http://www.w3.org/2004/02/skos/core#	<i>Simple Knowledge Organization System</i>	É uma recomendação do W3C utilizada para representação de diferentes tipos de sistemas de organização do conhecimento, como tesouros, esquemas de classificação e taxonomias (Miles; Bechhofer, 2009).

Fonte: Elaborado pelos autores (2021).

Com base nas informações levantadas, foi criado um modelo conceitual da ontologia proposta, representado por meio de um diagrama de classes da *Unified Modeling Language* (UML) adaptado, onde as classes são representadas por retângulos, as propriedades de dados (*data properties*) são representadas por atributos descritos dentro das classes e as propriedades de objeto (*object properties*) são representadas por associações entre classes.

Com o objetivo de permitir maior alcance e visibilidade da ontologia proposta, todas as classes, propriedades e anotações foram descritas no idioma inglês. A Figura 1 apresenta o diagrama com o modelo conceitual da ontologia proposta para representação de portarias, mostrando suas classes, propriedades e restrições.

Desenvolvimento da ontologia

Os próximos passos do processo de desenvolvimento estão bem entrelaçados. Esses passos consistem em definir a hierarquia de classes, propriedades e restrições. Nesta seção, são apresentadas de forma detalhada todas as propriedades e classes apresentadas na Figura 1, bem como as restrições existentes.

Classes

A classe *schema:Thing* (coisa) pertence à ontologia *Schema.org* e é reusada como a entidade de domínio mais genérica da ontologia; ou seja, o tipo que está no topo da hierarquia de classes, abaixo apenas da *owl:Thing* (conjunto universo). Ela representa o conceito mais genérico do qual todas as outras classes descendem direta ou indiretamente. Por ser uma classe de topo de hierarquia, todas as suas propriedades são herdadas pelas que estão abaixo dela.

A classe *foaf:Agent* (agente) pertence à ontologia FOAF e representa qualquer agente; ou seja, qualquer indivíduo que faça algo. Dentre esses agentes, destacam-se pessoas e organizações e os papéis que elas desempenham.

A classe *OfficialDocumentIssuer* (emissor de documento oficial) representa, de forma genérica, qualquer entidade responsável por emitir um documento oficial. Essa classe possui uma relação de equivalência com um axioma que determina que qualquer indivíduo que emita pelo menos um documento oficial pertence a ela.

A classe *Party* (parte) representa um conceito genérico relacionado tanto a pessoas quanto a organizações. É subclasse de *foaf:Agent* e *schema:Thing*. Possui uma relação de equivalência com a união das classes disjuntas *org:Organization* e *foaf:Person*, determinando seu particionamento por essas duas.

A classe *org:Organization* (organização) pertence à ontologia *The Organization Ontology* e representa um conjunto de pessoas organizadas em uma comunidade ou em alguma estrutura social, comercial ou política. No contexto da ontologia, representa as instituições de ensino e suas subdivisões. É subclasse de *Party* e possui uma relação de equivalência com duas classes homônimas, que pertencem às ontologias FOAF e *Schema.org*, indicando que todo indivíduo que pertença à classe *org:Organization* também pertence às classes *foaf:Organization* e *schema:Organization*. Essa classe possui uma restrição de cardinalidade sobre a propriedade *org:subOrganizationOf*, indicando que uma organização pode ser suborganização de, no máximo, uma outra organização.

A classe *foaf:Person* (pessoa) pertence à ontologia FOAF e representa pessoas. No contexto da ontologia, representa, de forma genérica, qualquer pessoa que esteja relacionada a um documento oficial. É subclasse de *Party* e possui uma relação de equivalência com a classe homônima presente na ontologia *Schema.org*.

A classe *PublicServant* (servidor público) representa um grupo de pessoas que mantém um vínculo profissional com uma instituição pública. Essa classe é utilizada para representar os servidores públicos relacionados às portarias publicadas e possui uma relação de equivalência com uma classe anônima gerada a partir de uma restrição existencial sobre a propriedade *entailedTo*, indicando que qualquer indivíduo que esteja relacionado a um *Entailment* (vínculo empregatício) por meio dessa propriedade é um servidor público.

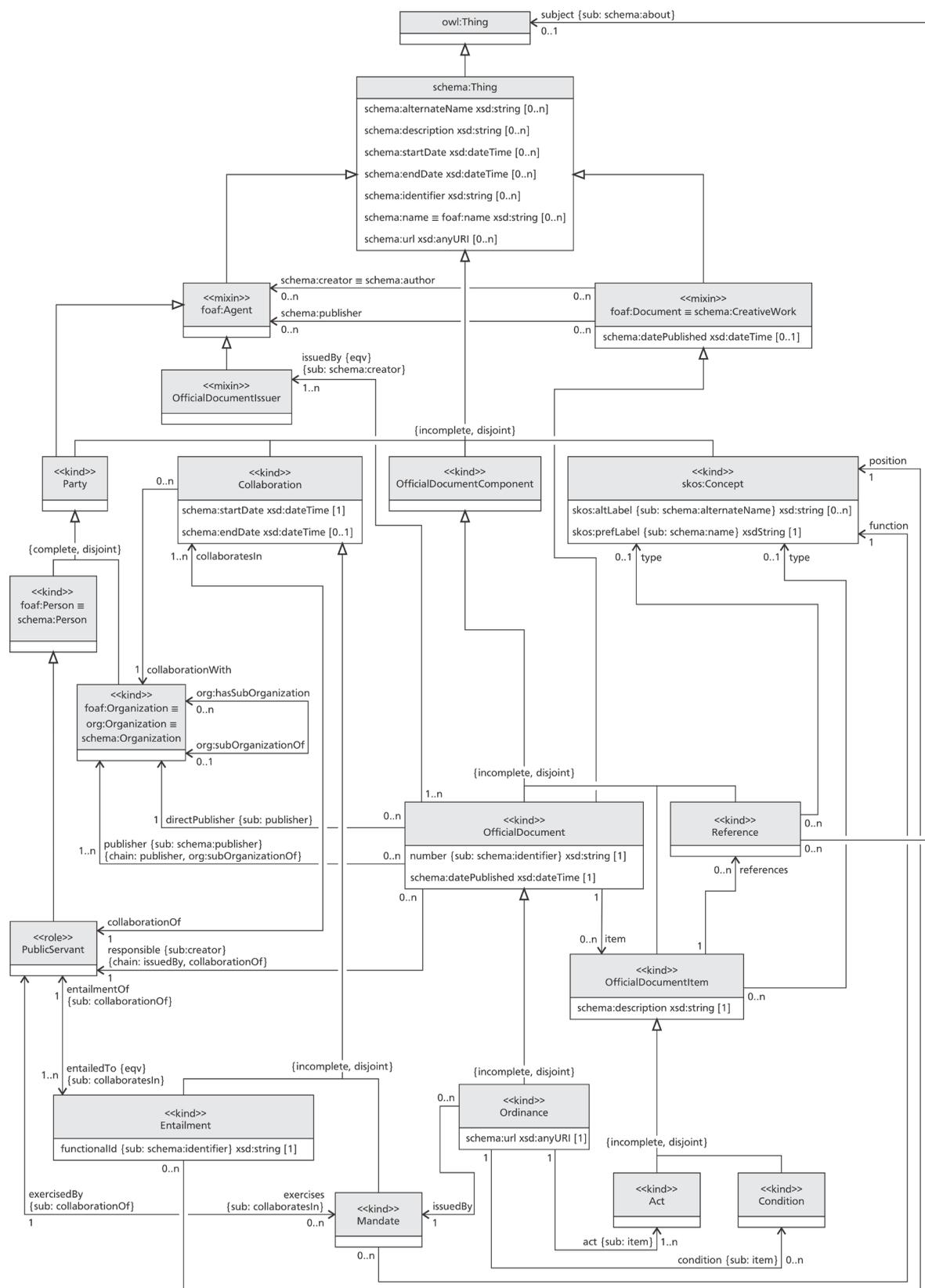


Figura 1 – Modelagem conceitual da Ordinance Ontology.

Fonte: Elaborada pelos autores (2021).

A classe *Collaboration* (colaboração) é um conceito genérico que representa a relação de colaboração entre um servidor público e uma organização. Essa classe possui restrições sobre as propriedades *schema:startDate* e *schema:endDate*, indicando que uma relação de colaboração deve possuir, exatamente, uma data de início e, no máximo, uma data de fim.

A classe *Entailment* (vínculo) representa um vínculo empregatício de um servidor público; ou seja, qual o cargo exercido por ele em uma determinada instituição. Servidores podem estar relacionados a mais de um vínculo, os quais são identificados pela respectiva matrícula. Essa classe possui uma restrição, indicando que indivíduos pertencentes a ela estão relacionados, através da propriedade *position*, a um determinado cargo.

A classe *Mandate* (mandato) representa um mandato eletivo que pode ser exercido por um servidor público durante um determinado período de tempo. Essa classe possui uma restrição, indicando que indivíduos pertencentes a ela estão relacionados, através da propriedade *function*, a uma determinada função.

A classe *foaf:Document* pertence à ontologia FOAF e representa um documento. Documento é um conceito mais abrangente que indica qualquer tipo de registro de informações. Essa classe tem uma relação de equivalência com a classe *schema:CreativeWork*, pertencente à ontologia *Schema.org*. Também possui uma restrição de cardinalidade sobre a propriedade *schema:datePublished*, indicando que um documento pode ter uma data de publicação no máximo.

A classe *OfficialDocumentComponent* (componente de documento oficial) representa, de forma genérica, todos os tipos de componentes relacionados aos documentos oficiais.

A classe *OfficialDocument* (documento oficial) representa, de forma genérica, um documento oficial publicado por uma instituição de ensino. Apesar do foco deste trabalho ser um tipo específico de documento – a portaria –, a definição dessa classe surge para permitir que a ontologia seja extensível para novos tipos de documentos oficiais no futuro. Essa classe possui uma relação de equivalência com a classe anônima originada a partir da restrição existencial sobre a propriedade *issuedBy*, indicando que um documento oficial é emitido por, pelo menos, um emissor. Além disso, possui uma restrição existencial sobre a propriedade *publisher*, indicando que um documento oficial deve ser publicado por, pelo menos, uma organização. Ela ainda possui restrições de cardinalidade sobre as propriedades *schema:datePublished* e *number*, indicando que um documento oficial deve ter, exatamente, uma data de publicação e um número associado.

A classe *OfficialDocumentItem* (item de documento oficial) representa, de forma genérica, itens presentes em documentos oficiais. Essa classe possui uma restrição de cardinalidade sobre a propriedade *type*, indicando que cada item de documento deve ter, no máximo, um tipo associado.

A classe *Ordinance* (portaria) representa uma portaria emitida por uma instituição. As portarias são compostas por dois itens de documento específicos, a saber: atos e condições. Essa classe possui uma relação existencial sobre a propriedade *act*, a qual indica que, para um indivíduo ser uma portaria, ele deve estar relacionado a, pelo menos, um ato normativo. Além disso, possui uma restrição existencial sobre a propriedade *issuedBy*, indicando que uma portaria só pode ser emitida por um mandato ou por um servidor que exerça um mandato. Também é definida uma restrição de cardinalidade sobre essa propriedade, indicando que uma portaria só pode estar vinculada a um mandato apenas. Ela ainda possui restrições de cardinalidade sobre a propriedade *schema:url*, indicando que cada portaria deve ter exatamente um URL.

A classe *Act* (ato) representa os atos normativos promulgados pela autoridade competente que está presente em uma portaria.

A classe *Condition* (condição) representa as condições ou pré-requisitos que validam os atos normativos presentes em portarias.

A classe *Reference* (referência) representa qualquer recurso que possa ser referenciado por um item de documento – no caso de portarias, um ato normativo ou uma condição. Normalmente, portarias podem referenciar

peças, outras portarias, outros documentos (atas, editais, leis etc.), setores, dentre outras coisas. Essa classe possui restrições de cardinalidade sobre as propriedades *type* e *subject*, indicando que uma referência pode ter, no máximo, um tipo associado e, no máximo, um assunto.

A classe *skos:Concept* (conceito) pertence à ontologia SKOS e representa um conceito (categoria). Essa classe descreve a estrutura conceitual de um sistema de organização de conhecimento (*tesauro*), que pode ser definida de maneira hierárquica. No contexto da ontologia proposta, é utilizada para classificar tipos de portarias, funções exercidas por mandatos, cargos vinculados a servidores públicos e tipos de referências presentes em itens de documentos. Ela possui uma relação de equivalência com uma restrição existencial sobre a propriedade *skos:prefLabel*, indicando que conceitos devem possuir, pelo menos, um rótulo associado.

Propriedades de objetos

A propriedade *collaborationOf* (colaboração de) relaciona uma colaboração a um servidor público. É inversa de *collaboratesIn* e também uma propriedade funcional que indica que dois ou mais URIs relacionados à mesma colaboração representam o mesmo indivíduo.

A propriedade *entailmentOf* (vínculo de) relaciona um vínculo empregatício a um servidor público. É subpropriedade de *collaborationOf* e inversa de *entailedTo*, sendo também uma propriedade funcional que indica que um vínculo empregatício não pode estar relacionado a mais de um servidor público distinto; ou seja, caso essa propriedade associe um determinado vínculo a dois ou mais URIs, todos eles serão considerados o mesmo indivíduo.

A propriedade *exercisedBy* (exercido por) relaciona um mandato a um servidor público. É subpropriedade de *collaborationOf* e inversa de *exercises*, sendo também uma propriedade funcional, indicando que um mandato não pode estar relacionado a mais de um servidor público.

A propriedade *collaboratesIn* (colabora em) relaciona um servidor público a uma colaboração. É inversa de *collaborationOf*. Essa propriedade é utilizada para representar, de forma genérica, a relação entre um servidor público e uma organização através de um vínculo de colaboração.

A propriedade *entailedTo* (vinculado a) relaciona um servidor público a um vínculo empregatício. É subpropriedade de *collaboratesIn* e inversa de *entailmentOf*. Por ser inversa de *entailmentOf*, é inferido que essa propriedade é funcional inversa.

A propriedade *exercises* (exerce) relaciona um servidor público a um mandato eletivo. É subpropriedade de *collaboratesIn* e inversa de *exercisedBy*. Por ser inversa de *exercisedBy*, é inferido que essa propriedade é funcional inversa.

A propriedade *collaborationWith* (colaboração com) relaciona uma colaboração a uma organização. Essa propriedade indica a organização à qual um servidor público está relacionado através de um vínculo de colaboração. É uma propriedade funcional que indica que se dois ou mais URIs estiverem vinculados à mesma colaboração, eles serão considerados o mesmo indivíduo.

A propriedade *function* (função) relaciona um mandato a um conceito. Essa propriedade indica a função exercida por um servidor público através de um mandato (reitor, diretor, coordenador etc.). É uma propriedade funcional que indica que um mandato pode estar relacionado a apenas uma função.

A propriedade *position* (posição) relaciona um vínculo empregatício a um conceito. Ela indica o cargo que um determinado servidor exerce em um determinado vínculo que ele mantém com o serviço público. É uma propriedade funcional que indica que um vínculo empregatício pode estar relacionado a apenas um cargo.

A propriedade *type* (tipo) relaciona qualquer recurso a um conceito. No contexto da ontologia, essa propriedade pode ser utilizada para classificar (ou categorizar) itens de documentos (atos ou condições) e referências.

A propriedade *org:hasSubOrganization* (tem sub-organização) pertence à ontologia ORG e relaciona uma organização a outra. É inversa de *org:subOrganizationOf*. No contexto da ontologia, essa propriedade representa o relacionamento de uma instituição ou setor com seus respectivos subsetores, de forma hierárquica.

A propriedade *org:subOrganizationOf* (suborganização de) pertence à ontologia ORG e relaciona uma organização a outra. É inversa de *org:hasSubOrganization*. Essa propriedade indica que uma organização está subordinada a outra.

A propriedade *item* (item) relaciona um documento oficial a um item de documento. Essa propriedade indica um item descrito em um documento oficial de forma genérica.

A propriedade *act* (ato) relaciona uma portaria a um ato normativo. É subpropriedade de *item*.

A propriedade *condition* (condição) relaciona uma portaria a uma condição. É subpropriedade de *item*.

A propriedade *references* (referência) relaciona um item de documento (ato ou condição) a uma referência genérica. No contexto da ontologia, cada item presente em uma portaria pode fazer referência a qualquer coisa; logo, essa propriedade permite indicar um recurso que representa uma referência de um determinado item. Cada referência pode ser classificada (ou categorizada) através de um conceito (via propriedade *type*) e está relacionada a qualquer coisa que se deseja relacionar ao item de documento.

A propriedade *schema:about* (sobre) pertence à ontologia Schema.org e relaciona dois recursos de forma genérica. Essa propriedade indica o assunto ou tópico associado a um recurso.

A propriedade *subject* (assunto) relaciona uma referência a qualquer coisa. É subpropriedade de *schema:about*. Essa propriedade indica um indivíduo que está sendo referenciado por uma referência de um item de documento.

A propriedade *schema:creator* (criador) pertence à ontologia Schema.org e relaciona um documento ou trabalho criativo a um agente. É equivalente à propriedade *schema:author*.

A propriedade *issuedBy* (emitido por) relaciona um documento oficial a um emissor de documentos. É subpropriedade de *schema:creator* e é definida como funcional. Essa propriedade define que todo indivíduo que emite um documento oficial deve ser classificado como um emissor de documento oficial. No contexto da ontologia, ela normalmente relaciona um mandato a um documento, indicando que o mandato exercido por um servidor é considerado o emissor do documento oficial.

A propriedade *responsible* (responsável) relaciona um documento oficial a um servidor público. É subpropriedade de *schema:creator*. Essa propriedade indica qual é o servidor público responsável por emitir o documento oficial, através de um mandato. Ela é definida por uma cadeia de propriedades, indicando que, se um documento foi emitido por um mandato e esse mandato é exercido por um servidor, logo, o servidor é o responsável pelo documento emitido.

A propriedade *schema:publisher* (publicador) pertence à ontologia Schema.org e relaciona um trabalho criativo a um agente de forma genérica.

A propriedade *publisher* (publicador) definida na ontologia relaciona um documento oficial a uma organização. É subpropriedade de *schema:publisher*. Essa propriedade indica qual a organização ou setor que publicou um documento oficial, seja de forma direta ou indireta. Para isso, foi definida uma cadeia de propriedades indicando que um documento também é publicado pelos setores presentes nos níveis hierárquicos superiores ao setor que o publicou.

A propriedade *directPublisher* (publicador direto) relaciona um documento oficial a uma organização. É subpropriedade de *publisher* e também uma propriedade funcional que indica apenas o setor que publicou diretamente um documento oficial.

Propriedades de Dados

A propriedade *schema:name* (nome) pertence à ontologia *Schema.org* e representa o nome de um recurso, sendo herdada por todas as classes da ontologia. É subpropriedade de *rdfs:label* e possui uma relação de equivalência com a propriedade homônima da ontologia FOAF.

A propriedade *schema:alternateName* (nome alternativo) pertence à ontologia *Schema.org* e indica um nome alternativo atribuído a um recurso.

A propriedade *skos:prefLabel* (rótulo preferido) pertence à ontologia SKOS e é utilizada para indicar o nome preferencial de um determinado conceito. É subpropriedade de *schema:name* e é definida como uma propriedade funcional.

A propriedade *skos:altLabel* (rótulo alternativo) pertence à ontologia SKOS e é utilizada para indicar o nome alternativo de um determinado conceito. É subpropriedade de *schema:alternateName*.

A propriedade *schema:description* (descrição) pertence à ontologia *Schema.org* e é utilizada para indicar a descrição de um recurso, sendo herdada por todas as classes da ontologia. É subpropriedade de *rdfs:comment*.

A propriedade *schema:identifier* (identificador) pertence à ontologia *Schema.org* e é utilizada para indicar um valor que representa um identificador para o recurso.

A propriedade *functionalId* (identificador funcional) relaciona um vínculo empregatício de um servidor público. É subpropriedade de *schema:identifier* e é definida como uma propriedade funcional. Por ser uma propriedade funcional, ela não permite que dois valores distintos sejam relacionados a um determinado vínculo.

A propriedade *number* (número) é utilizada para indicar o número que identifica um documento oficial. É subpropriedade de *schema:identifier*. Embora seja uma propriedade utilizada para identificação de recursos, ela não representa um identificador exclusivo para cada documento, sendo possível que dois documentos distintos possuam o mesmo número. Essa propriedade identifica exclusivamente um documento publicado por um determinado setor em um ano específico.

A propriedade *schema:startDate* (data de início) pertence à ontologia *Schema.org* e é utilizada para representar a data inicial vinculada a um recurso. No contexto da ontologia, essa propriedade pode indicar a data inicial da relação de colaboração entre uma pessoa e uma organização.

A propriedade *schema:endDate* (data de fim) pertence à ontologia *Schema.org* e é utilizada para representar a data final vinculada a um recurso. No contexto da ontologia, essa propriedade pode indicar a data de encerramento da relação de colaboração entre uma pessoa e uma organização.

A propriedade *schema:datePublished* (data de publicação) pertence à ontologia *Schema.org* e relaciona um documento ou trabalho criativo à sua data de publicação. Essa propriedade é utilizada para indicar a data de publicação de um documento.

A propriedade *schema:url* (endereço eletrônico) pertence à ontologia *Schema.org* e relaciona um documento oficial ao seu endereço *web*.

O arquivo OWL da ontologia pode ser obtido no endereço do endereço: <https://purl.org/ordinance-ontology/>.

Validação da ontologia

A avaliação da ontologia corresponde a uma importante etapa no seu processo de desenvolvimento e, de acordo com o processo adotado no desenvolvimento da ontologia utilizada neste estudo, para validá-la é preciso verificar se o modelo proposto fornece respostas para todas as questões de competência levantadas na definição do escopo.

A fim de validar a ontologia desenvolvida para representação de portarias, foi realizado um estudo de caso realista, utilizando como base as informações contidas em uma portaria emitida pelo Instituto Federal Fluminense (IFF). Além do estudo de caso apresentado nesta seção, na dissertação de mestrado (Silva, 2021) da qual este trabalho faz parte foi realizado um estudo de caso mais abrangente, no qual uma vasta quantidade de portarias do IFF foi automaticamente mapeada em RDF usando a ontologia proposta, viabilizando, com sucesso, consultas precisas e elaboradas. Por razões de espaço, esse estudo maior foi omitido, mas pode ser visto na referida dissertação.

Para dar suporte ao contexto em que os estudos de caso foram realizados, foram criadas, na ontologia, instâncias para representar todos os setores do IFF que têm competência para emitir portarias, bem como instâncias para representar as funções exercidas por servidores públicos através de mandatos eletivos.

Este estudo de caso utilizou a ontologia desenvolvida para representar a maior quantidade possível de informações presentes em uma portaria. Para isso, foi preciso mapear as informações contidas nos documentos e representá-las em triplas RDF manualmente.

Este estudo de caso pode ser dividido nas seguintes etapas:

- 1) Seleção de uma portaria para ser utilizada no estudo de caso;
- 2) Identificação das informações relevantes a serem representadas;
- 3) Descrição das informações identificadas no formato de triplas RDF utilizando os termos definidos na *Ordinance Ontology*;
- 4) Inserção do documento de texto contendo as triplas RDF em uma base de dados RDF;
- 5) Realização de consultas SPARQL a fim de responder todas as questões de competência.

Primeiramente, foi selecionada para a realização deste estudo de caso a portaria de número 363 emitida em 11 de novembro de 2020 no portal do CDD do Instituto Federal Fluminense, a qual pode ser acessada através do seguinte endereço eletrônico: <<http://cdd.iff.edu.br/documentos/portarias/reitoria/progep/2020/novembro/portaria-13/view/++widget++form.widgets.arquivo/@@download/Portaria+N.%C2%BA+363%2C+de+11+de+no+vembro+de+2020.pdf>>.

A Figura 2 apresenta a transcrição do conteúdo contido na portaria.

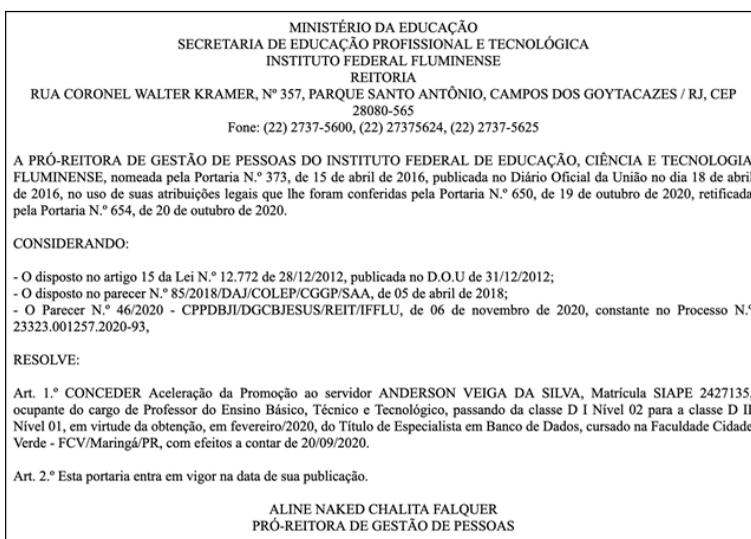


Figura 2 – Transcrição da portaria utilizada no estudo de caso realista.

Fonte: Elaborada pelos autores (2021) baseada na portaria de número 363 emitida em 11 de novembro de 2020 pelo Instituto Federal Fluminense.

Após a análise do conteúdo da portaria, foi possível distinguir informações relevantes que puderam ser descritas utilizando a ontologia desenvolvida. O Quadro 2 apresenta as informações identificadas.

Quadro 2 – Informações identificadas na portaria de exemplo.

Propriedade	Valor
Setor que publicou	Pró-Reitoria de Gestão de Pessoas
Número	363
Descrição	Portaria nº 363 de 11 de novembro de 2020
Data de publicação	11/11/2020
Emissor	Aline Naked Chalita Falquer
Condições	- O disposto no artigo 15 da lei nº 12.772 de 28/12/2012, publicada no ... - O disposto no parecer N.º 85/2018/DAJ/COLEP/CGGP/SAA, de ... - O parecer nº 46/2020 - CPPDCBJI/DGCBJESUS/REIT/IFFLU, de ...
Ato normativo	- Conceder Aceleração da Promoção ao servidor Anderson Veiga da Silva...
URL	http://cdd.iff.edu.br/documentos/portarias/reitoria/progep/2020/novembro/portaria-13/view/++widget++form.widgets.arquivo/@@download/Portaria+N.%C2%BA+363%2C+de+11+de+novembro+de+2020.pdf

Fonte: Elaborado pelos autores (2021).

Analisando cada uma das condições presentes na portaria, foi possível identificar que outros documentos são referenciados em cada uma das condições. A primeira condição cita a lei nº 12.772, a segunda condição cita o parecer nº 85 emitido por um setor de um determinado órgão e a terceira condição cita o parecer nº 46 de 2020 emitido pela Comissão Permanente de Pessoal Docente do *campus* Bom Jesus do Itabapoana. Além disso, foi possível observar que o ato normativo presente nessa portaria descreve a aceleração da promoção do servidor Anderson Veiga da Silva.

Portanto, além das informações descritas no Quadro 2, também foi possível identificar documentos e servidores públicos que foram referenciados nos itens da portaria, os quais podem ser representados através da ontologia.

Como este é um estudo de caso realista e tem por objetivo descrever em um grafo RDF todas as informações relacionadas à portaria de forma detalhada, foram feitas consultas no Sistema Unificado de Administração Pública (SUAP) a fim de identificar informações funcionais dos servidores referenciados na portaria. Também foram definidos, para este exemplo, indivíduos na própria ontologia para representar unidades organizacionais da instituição.

A seguir são apresentadas todas as informações identificadas na portaria analisada, descritas em triplas RDF utilizando a sintaxe *Turtle*.

```
@base <http://purl.org/ordinance-ontology/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix schema: <http://schema.org/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
<progep_363_2020-11-11> a <Ordinance> ;
<number> "363"^^xsd:string ;
schema:description "Portaria nº 363 de 11 de novembro de 2020"^^xsd:string ;
schema:datePublished "2020-11-11"^^xsd:dateTime ;
schema:url "http://cdd.iff.edu.br/documentos/portarias/reitoria/progep/2020/novembro/portaria-13/view/++widget++form.widgets.arquivo/@@download/Portaria+N.%C2%BA+363%2C+de+11+de+novembro+de+2020.pdf"^^xsd:anyURI ;
<directPublisher> <progep> ;
```

<issuedBy> <mandate_1> ;
 <act> <act_1> ;
 <condition> <condition_1>, <condition_2>, <condition_3> .
 <aline_naked> a <PublicServant> ;
 schema:name "Aline Naked Chalita Falquer"^^xsd:string ;
 <exercises> <mandate_1> ;
 <entailedTo> <entailment_1> .
 <anderson_veiga> a <PublicServant> ;
 schema:name "Anderson Veiga da Silva"^^xsd:string ;
 <entailedTo> <entailment_2> .
 <mandate_1> a <Mandate> ;
 <function> <pro_rector> ;
 schema:startDate "2020-03-02"^^xsd:dateTime .
 <entailment_1> a <Entailment> ;
 <functionalId> "269405"^^xsd:string ;
 <position> <administration_assistant> ;
 schema:startDate "1993-05-31"^^xsd:dateTime .
 <entailment_2> a <Entailment> ;
 <functionalId> "2427135"^^xsd:string ;
 <position> <teacher> ;
 schema:startDate "2017-10-11"^^xsd:dateTime .
 <teacher> a skos:Concept ;
 skos:prefLabel "Professor"^^xsd:string .
 <administration_assistant> a skos:Concept ;
 skos:prefLabel "Assistente em Administração"^^xsd:string .
 <condition_1> a <Condition> ;
 schema:description "O disposto no artigo 15 da Lei N.º 12.772 de 28/12/2012, publicada no D.O.U. de 31/12/2012"^^xsd:string ;
 <references> <reference_1> .
 <condition_2> a <Condition> ;
 schema:description "O disposto no Parecer N.º 85/2018/DAJ/COLEP/CGGP/SAA, de 05 de abril de 2018"^^xsd:string ;
 <references> <reference_2> .
 <condition_3> a <Condition> ;
 schema:description "O Parecer N.º 46/2020 - CPPDCBJI/DGCBJESUS/REIT/IFFLU, de 06 de novembro de 2020, constante no Processo N.º 23323.001257.2020-93"^^xsd:string ;
 <references> <reference_3> .
 <act_1> a <Act> ;
 schema:description "Art. 1.º CONCEDER Aceleração da Promoção ao servidor ANDERSON VEIGA DA SILVA, Matrícula SIAPE 2427135, ocupante do cargo de Professor do Ensino Básico, Técnico e Tecnológico, passando da Classe D I Nível 02 para a Classe D II Nível 01, em virtude da obtenção em fevereiro/2020, do Título de Especialista em Banco de Dados, cursado na Faculdade Cidade Verde - FCV/Maringá/PR, com efeitos a contar de 20/09/2020."^^xsd:string ;
 <type> <promotion_acceleration> ;
 <references> <reference_4> .
 <reference_1> a <Reference> ;
 <subject> <law_12772> .
 <reference_2> a <Reference> ;
 <subject> <opinion_85_2018> .
 <reference_3> a <Reference> ;
 <subject> <opinion_46_2020> .
 <reference_4> a <Reference> ;

```

<subject> <anderson_veiga> .
<promotion_acceleration> a skos:Concept ;
skos:prefLabel "Aceleração de Promoção"^^xsd:string .
<law_12772> <OfficialDocument> ;
schema:description "Lei Nº 12.772/2012"^^xsd:string .
<opinion_85_2018> a <OfficialDocument> ;
schema:description "Parecer Nº 85/2018/DAJ/COLEP/CGGP/SAA"^^xsd:string .
<opinion_46_2020> a <OfficialDocument> ;
schema:description "Parecer Nº 46/2020 - CPPDCBJI/DGCBJESUS/REIT/IFFLU"^^xsd:string ;
<directPublisher> <cppd_bom_jesus> .
<cppd_bom_jesus> a foaf:Organization ;
schema:description "Comissão Permanente de Pessoal Docente - Bom Jesus"^^xsd:string .

```

O arquivo RDF gerado foi então importado para uma base de dados RDF com o objetivo de permitir que consultas SPARQL sobre os dados fossem realizadas. Neste estudo de caso, foi utilizado o banco de dados *GraphDB*, desenvolvido pela *Ontotext*.

O *GraphDB* é uma base de dados baseada em grafos altamente eficiente e robusta, que fornece suporte ao RDF e à SPARQL (Ontotext, 2021). O *GraphDB* foi escolhido para realização deste estudo de caso por ser um sistema de fácil instalação e utilização, por fornecer uma interface para realização de consultas SPARQL com suporte nativo à inferência ontológica e, também, por trabalhar com quase todos os formatos de serialização RDF.

O próximo passo deste estudo de caso consistiu em realizar consultas SPARQL sobre o grafo RDF armazenado no *GraphDB*, visando responder as questões de competência que foram levantadas no processo de desenvolvimento da ontologia.

Muitas consultas SPARQL realizadas neste trabalho puderam ser simplificadas por conta das inferências obtidas através de axiomas definidos na ontologia

A seguir são apresentadas todas as consultas SPARQL para cada questão de competência levantada.

- QC1: Quais portarias foram publicadas em um determinado intervalo de tempo?

Para responder a essa questão de competência, é preciso consultar os recursos que são portarias (*ord:Ordinance*) e que foram publicados (*schema:datePublished*) em um determinado intervalo de tempo. A consulta a seguir retorna todas as portarias cuja data de publicação esteja entre 01/11/2020 e 31/03/2021.

```

SELECT ?ordinance WHERE {
  ?ordinance a ord:Ordinance .
  ?ordinance schema:datePublished ?date .
  FILTER (?date > "2013-01-01"^^xsd:dateTime && ?date < "2021-03-31"^^xsd:dateTime) .
}

```

- QC2: Quais portarias foram publicadas por um determinado setor do instituto?

Para responder a essa questão de competência, é preciso consultar os recursos que são portarias (*ord:Ordinance*) e que foram publicados (*schema:publisher*) por um determinado setor da instituição. A consulta a seguir retorna todas as portarias que foram publicadas pela Pró-Reitoria de Gestão de Pessoas do Instituto Federal Fluminense.

```

SELECT ?ordinance WHERE {
  ?ordinance a ord:Ordinance .
  ?ordinance schema:publisher ord:progep.
}

```

- QC3: Quais portarias foram publicadas com uma determinada finalidade (tipo)?

Para responder a essa questão de competência, é preciso consultar as portarias que possuem atos normativos (*ord:act*) explicitamente classificados com um determinado tipo (*ord:type*). Neste exemplo, não foi definido que o indivíduo consultado tem que ser uma portaria (*ord:Ordinance*), pois essa informação é inferida pelo domínio da propriedade *ord:act*. A consulta a seguir retorna todas as portarias que tenham atos normativos que descrevem a aceleração da promoção de um servidor.

```
SELECT ?ordinance ?type WHERE {
  ?ordinance ord:act ?act .
  ?act ord:type ord:promotion_acceleration .
}
```

- QC4: Quais portarias relacionam-se a um determinado servidor?

Para responder a essa questão de competência, é preciso consultar as portarias cujos atos normativos possuem referências (*ord:references*) onde o assunto (*ord:subject*) seja um servidor público. A consulta seguinte retorna todas as portarias que tenham atos normativos que façam referência ao servidor Anderson Veiga da Silva.

```
SELECT ?ordinance WHERE {
  ?ordinance ord:act ?act .
  ?act ord:references ?reference .
  ?reference ord:subject ?public_servant .
  ?public_servant a ord:PublicServant .
  ?public_servant schema:name "Anderson Veiga da Silva"^^xsd:string .
}
```

- QC5: Quais portarias foram assinadas por determinada autoridade competente?

Para responder a essa questão de competência, é preciso consultar as portarias que foram emitidas (*ord:issuedBy*) durante um determinado mandato que seja exercido (*ord:exercisedBy*) por um determinado servidor público. A consulta a seguir retorna todas as portarias que foram emitidas pela servidora Aline Naked Chalita Falquer.

```
SELECT ?ordinance WHERE {
  ?ordinance ord:issuedBy ?mandate .
  ?mandate ord:exercisedBy ?public_servant .
  ?public_servant schema:name "Aline Naked Chalita Falquer"^^xsd:string .
}
```

- QC6: Quais portarias referenciam outras portarias?

Para responder a essa questão de competência, é preciso consultar as portarias cujos itens (*ord:item*) possuem uma referência (*ord:references*) onde o assunto (*ord:subject*) seja outra portaria. Neste exemplo, a portaria escolhida não faz referência a outras portarias. A consulta a seguir retorna todas as portarias que fazem referência a outra portaria qualquer.

```
SELECT DISTINCT ?ordinance WHERE {
  ?ordinance ord:item ?item .
  ?item ord:references ?reference .
  ?reference ord:subject ?document .
  ?document a ord:Ordinance .
}
```

- QC7: Quais portarias possuem uma determinada palavra-chave em seu conteúdo?

Para responder a essa questão de competência, é preciso consultar um determinado termo na descrição da portaria ou em algum de seus itens (atos ou condições). A consulta a seguir retorna todas as portarias cujo termo "conceder" esteja presente.

```
SELECT DISTINCT ?ordinance WHERE {
  ?ordinance a ord:Ordinance .
  {
    ?ordinance ord:item ?item .
    ?item schema:description ?search .
  } UNION {
    ?ordinance schema:description ?search .
  }
  FILTER REGEX(?search, "conceder", "i") .
}
```

- QC8: Quais portarias possuem uma determinada condição (documento)?

Para responder a essa questão de competência, é preciso consultar as portarias cujas condições (*ord:condition*) fazem referência a um determinado documento. A consulta seguinte retorna todas as portarias cujas condições fazem referência à lei nº 12.772, de 2012.

```
SELECT ?ordinance WHERE {
  ?ordinance ord:condition ?condition .
  ?condition ord:references ?reference .
  ?reference ord:subject ord:law_12772 .
}
```

- QC9: Quais portarias foram demandadas por um determinado setor?

Para responder a essa questão de competência, é preciso consultar as portarias que, em suas condições, fazem referências (*ord:references*) a documentos publicados (*ord:directPublisher*) por determinado setor (*org:Organization*) da instituição. São considerados demandantes os setores que publicam documentos que estão presentes nas condições das portarias. A consulta seguinte retorna todas as portarias que foram demandadas pela CPPD do campus Bom Jesus do Itabapoana, pertencente ao IFF.

```
SELECT ?ordinance WHERE {
  ?ordinance ord:condition ?condition .
  ?condition ord:references ?reference .
  ?reference ord:subject ?official_document .
  ?official_document ord:directPublisher ord:cppd_bom_jesus .
}
```

Resultados e Discussão

Como pode ser observado no estudo de caso realizado, a ontologia fornece classes e propriedades para descrever, de maneira formal, informações presentes em uma portaria publicada por uma instituição federal de ensino.

As consultas SPARQL realizadas para responder as questões de competência permitem obter informações utilizando critérios de pesquisa que não seriam aplicáveis ao conteúdo disponibilizado originalmente nos documentos de texto referentes às portarias.

É possível, inclusive, aplicar filtros de pesquisa em portarias que são referenciadas por outras portarias. Por exemplo, é possível retornar quais são as portarias que referenciam outras portarias que foram publicadas por um determinado setor, ou em um determinado intervalo de tempo específico. São diversas as possibilidades de combinações de critérios de pesquisa que podem ser adotadas, através da utilização da ontologia desenvolvida.

Segundo Noy e McGuinness (2001), uma ontologia deve definir um vocabulário comum para pessoas que precisam compartilhar informações sobre um domínio. Isso inclui definições de conceitos básicos, sobre o domínio e a relação entre eles, que possam ser inteligíveis por máquinas. A ontologia desenvolvida atende a esses critérios, visto que conceitos relacionados ao contexto das portarias foram definidos formalmente e regras foram estabelecidas para que um agente computacional tenha a capacidade de interpretar o conhecimento descrito, permitindo inclusive a inferência de novas informações.

Segundo Breitman, Casanova e Truszkowski (2007), ontologias devem representar modelos que tornam explícitos o vocabulário utilizado em aplicações semânticas, garantindo que a comunicação seja feita livre de ambiguidades. O reuso de termos presentes em ontologias já existentes na construção da ontologia de portarias garante que os dados possam ser interpretados por outros agentes computacionais que já fazem uso dessas ontologias, sem que haja qualquer tipo de ambiguidade entre os termos e seus respectivos significados.

Devido ao escopo deste trabalho de pesquisa visar um problema muito específico, poucos trabalhos relacionados foram encontrados a fim de que fosse realizada uma comparação da ontologia desenvolvida com alguma já existente. A seguir são apresentados trabalhos relacionados ao tema deste artigo e que utilizam ontologias para descrever dados presentes em documentos de texto não estruturados ou dados presentes em documentos gerados por instituições de ensino.

Em seu trabalho denominado "Uso de ontologias para a extração de informações em atos jurídicos em uma instituição pública", Batres *et al.* (2005) apresentam uma ferramenta automatizada que utiliza técnicas de mineração de texto para extração de informações contidas em documentos jurídicos publicados por uma instituição federal de ensino. O trabalho é realizado utilizando-se atos administrativos publicados. Ao final, os dados extraídos são descritos através de uma ontologia criada para representar esses atos. As portarias são mencionadas no trabalho, porém os autores decidiram utilizar apenas atos administrativos. Esse trabalho é o que possui maior similaridade com o trabalho de pesquisa proposto, visto que busca representar documentos publicados por instituições públicas de ensino através de modelos ontológicos, além de fornecer uma ferramenta de extração de informações.

No trabalho denominado "*Use of Linked Data principles for semantic management of scanned documents*", Monteiro e Jacyntho (2016) propõem uma ferramenta para gerenciamento semântico de documentos digitalizados. Como resultado, a ferramenta fornece meios para realizar anotações semânticas sobre informações relacionadas aos documentos digitalizados, utilizando, para isso, uma ontologia que fornece termos que representam o domínio de conhecimento das informações relacionadas a documentos. A ferramenta também realiza o *mashup* das informações anotadas com recursos disponíveis na DBpedia a fim de enriquecê-las com dados presentes em outra fonte.

Em "*Ontology-based higher educational information systems*", Nyitrai e Varga (2012) mostram que, em instituições de ensino superior, muitos documentos e informações são gerados para cada atividade realizada e que a criação de uma ontologia serviria para descrever esses documentos. É então proposta nesse trabalho a criação de uma ontologia para descrever a estrutura organizacional interna de universidades, os programas acadêmicos, os cursos ofertados, pessoas relacionadas de alguma forma ao ambiente da universidade e os documentos gerados por ela. Para a criação dessa ontologia são reutilizadas as ontologias AISO, DCTerms, FOAF e NACE. Os documentos são representados de forma genérica pela ontologia FOAF. Por fim, é realizado um estudo de caso com a ontologia proposta, utilizando-a como base para executar consultas SPARQL.

Em "*Development of University Ontology for aSPOCMS*", Dwivedi e Kumar (2013) apresentam a construção de uma ontologia para representar o domínio de conhecimento de uma universidade, aplicando essa ontologia a um sistema de gerenciamento de conteúdo semântico já existente denominado aSPOCMS. Esse sistema visa possibilitar o gerenciamento e processamento de arquivos e documentos de departamentos de instituições

educacionais em um ambiente denominado *paperless* (sem papel). A ontologia descreve de forma genérica a estrutura e fluxos de processos de universidades na Índia e permite a realização de consultas inteligentes para recuperação de informações no ambiente institucional das universidades.

Conclusão

Extrair informações de valor de dados publicados na Internet ainda é uma tarefa muito trabalhosa, pois demanda esforço humano para buscar, interpretar, classificar e relacionar informações oriundas das mais variadas fontes de dados – um processo tedioso e sujeito a erros. A utilização das tecnologias relacionadas à *Web* semântica reduz drasticamente esse esforço, permitindo que máquinas se utilizem de sua capacidade de processamento para realizar essa tarefa com base em informações formalmente estruturadas.

Nesse sentido, devido ao grande número de portarias que diariamente são publicadas por instituições públicas de ensino, o desenvolvimento de uma ontologia para representação de portarias permite que dados importantes presentes nesses documentos possam ser estruturados e relacionados de forma que os computadores possam interpretá-los corretamente e realizar tarefas que seriam muito trabalhosas para os seres humanos.

Este trabalho demonstrou que o uso de tecnologias relacionadas à *Web Semântica* permite modelar informações relativas ao domínio de conhecimento das portarias e representá-las de forma concisa e precisa por meio de grafos RDF.

De certa forma, este trabalho contribui para a otimização do serviço público deste país, pois possibilita que informações de grande valor, que antes estavam escondidas dentro de arquivos de texto sem qualquer tipo de estruturação, agora estejam explicitamente estruturadas e disponíveis para serem utilizadas de forma adequada. Além disso, também representa mais um passo em direção à construção de uma *web* semântica, na qual os dados estejam devidamente organizados e disponibilizados para serem diretamente interpretados por máquinas, em prol do benefício humano.

Este artigo fornece algumas possibilidades para continuação desta pesquisa em trabalhos vindouros, as quais estão descritas a seguir:

Extensão da ontologia para contemplar explicitamente detalhes de diferentes tipos de portaria

Cada portaria emitida pode conter informações inerentes ao seu tipo. Por exemplo, portarias de convocação normalmente fazem referência a um determinado edital relacionado a um concurso público, aos servidores convocados e seus respectivos cargos, dentre outras informações específicas do contexto.

É possível especializar a ontologia proposta de forma que, para cada tipo de portaria existente, seja criada uma classe correspondente, que contenha propriedades e relacionamentos específicos, sendo possível fornecer, assim, uma estruturação das informações com mais detalhamento e possibilitar que consultas e inferências ainda mais rebuscadas sejam realizadas.

Extensão da ontologia para contemplar outros tipos de documentos oficiais

A ontologia proposta também pode ser especializada com o objetivo de representar outros tipos de documentos oficiais emitidos por instituições públicas de ensino, como ordens de serviço, editais, pareceres, notas técnicas, dentre outros tipos. Os documentos oficiais já estão representados de forma genérica na ontologia desenvolvida; logo, a especialização da ontologia se daria através da criação de novos tipos para representar novos documentos, bem como suas propriedades e relacionamentos com outros recursos.

Contribuição

A. V. Silva colaborou com a concepção e desenvolvimento. M. D. A. Jacyntho colaborou na concepção, orientação, revisão e aprovação.

Referências

- Alsharif, M. *Semantic Web Core Technologies*. [S.l.:s.n.], 2013. Disponível em: <https://www.cse.wustl.edu/~jain/cse570-13/ftp/semantic/index.html>. Acesso em: 21 dez. 2020.
- Batres, E. J. Q. et al. Uso de ontologias para a extração de informações em atos jurídicos em uma instituição pública. *Revista Eletrônica de Biblioteconomia e Ciência da Informação*, v. 10, n. 19, p. 73-88, 2005.
- Berners-Lee, T.; Hendler, J.; Lassila, O. The Semantic Web. *Scientific American*, v. 284, n. 5, p. 34-43, 2001.
- Brasil. Presidência da República. *Manual de redação da Presidência da República 2018*. Brasília: Presidência da República, 2018. Disponível em: <http://www4.planalto.gov.br/centrodeestudos/assuntos/manual-de-redacao-da-presidencia-da-republica/manual-de-redacao.pdf>. Acesso em: 21 dez. 2020.
- Breitman, K. K.; Casanova, M. A.; Truskowski, W. *Semantic web: concepts, technologies and applications*. New York: Springer, 2007.
- Brickley, D.; Guha, R. V. *Rdf Schema 1.1*. [S.l.]: W3C, 2014. Disponível em: <https://www.w3.org/TR/rdf-schema/>. Acesso em: 22 dez. 2020.
- Brickley, D.; Miller, L. *Foaf Vocabulary Specification*. [S.l.:s.n.], 2014. Disponível em: <http://xmlns.com/foaf/spec/>. Acesso em: 21 dez. 2020.
- Dwivedi, S. K.; Kumar, A. Development of University Ontology for ASPOCMS. *Journal of Emerging Technologies in Web Intelligence*, v. 5, n. 3, p. 213-221, 2013.
- Harris, S.; Seaborne, A.; Prud'hommeaux, E. *Sparql 1.1 Query Language*. [S.l.]: W3C, 2013. Disponível em: <https://www.w3.org/TR/sparql11-query/>. Acesso Em: 21 Dez. 2020.
- Heath, T.; Bizer, C. *Linked data: evolving the web into a global data space*. San Rafael: Morgan & Claypool, 2011.
- Hitzler, P. et al. *Owl 2 Web Ontology Language Primer*. 2nd ed. [S.l.]: W3C Recommendation, 2012. Disponível em: <https://www.w3.org/TR/owl2-primer/>. Acesso em: 19 abr. 2021.
- Jacyntho, M. D. A. *Um modelo de bloqueio multigranular para RDF*. 2012. Tese (Doutorado em Informática) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro. 2012.
- Jacyntho, M. D. A.; Morais, M. D. Ontology-based decision-making. In: Sakira, J. et al. *Web Semantics: cutting edge and future directions in healthcare*. [S.l.]: Elsevier, 2021. p. 195-209.
- Jacyntho, M. D.; Schwabe, D. A multigranularity locking model for RDF. *Journal of Web Semantics*, v. 39, p. 25-46, 2016.
- Laufer, C. *Guia de Web Semântica*, 2015. Disponível em: https://ceweb.br/media/docs/publicacoes/13/Guia_Web_Semantica.pdf. Acesso em: 26 nov. 2020.
- Manola, F.; Miller, E. *RDF Primer*. 2004. Disponível em: <https://www.w3.org/TR/2004/REC-rdf-primer-20040210/>. Acesso em: 29 nov. 2020.
- Miles, A.; Bechhofer, S. *Skos simple knowledge organization system reference*. [S.l.]: W3C, 2009. Disponível em: <https://www.w3.org/TR/2009/REC-skos-reference-20090818>. Acesso em: 21 dez. 2020.
- Monteiro, L. L. P.; Jacyntho, M. D. A. Use of linked data principles for semantic management of scanned documents. *Transinformação*, v. 28, n. 2, p. 241-251, 2016.
- Noy, N. F.; McGuinness, D. L. *Ontology development 101: a guide to creating your first ontology*. São Francisco: Stanford Knowledge Systems Laboratory, 2001.
- Nyitrai, E.; Varga, B. Ontology-based higher educational information systems. *Pollack Periodica*, v. 7, n. 3, p. 139-148, 2012.
- Ontotext. General: GraphDB SE 9.7.0 *Ontotext*, 2021. Documentation. Disponível em: <https://graphdb.ontotext.com/documentation/standard/index.html>. Acesso em: 28 abr. 2021.
- Reynolds, D. The organization ontology. W3C, 2014. Disponível em: <https://www.w3.org/TR/vocab-org/>. Acesso Em: 21 Dez. 2020.
- Silva, A. V. *Proposição de ontologia para representação de portarias emitidas por instituições federais de ensino*. 2021. Dissertação (Mestrado em Sistemas Aplicados à Engenharia e Gestão) - Instituto Federal Fluminense, Campos dos Goytacazes, 2021.
- The W3C SPARQL Working Group. SPARQL 1.1 Overview. W3C, 2013. Disponível em: <https://www.w3.org/TR/sparql11-overview/>. Acesso em: 21 dez. 2020.
- Yu, L. *A developer's guide to the semantic web*. Berlin: Springer, 2014.