



Universidade de Brasília - UnB
Faculdade UnB Gama - FGA
Engenharia de Software

Desenvolvimento de Ontologia para Estruturas Organizacionais do Governo Brasileiro.

Autor: Pedro Henrique de Brito Chaves
Orientador: Prof. Dr. Edgard Costa Oliveira

Brasília, DF
2015



Pedro Henrique de Brito Chaves

Desenvolvimento de Ontologia para Estruturas Organizacionais do Governo Brasileiro.

Monografia submetida ao curso de graduação em (Engenharia de Software) da Universidade de Brasília, como requisito parcial para obtenção do Título de Bacharel em (Engenharia de Software).

Universidade de Brasília - UnB

Faculdade UnB Gama - FGA

Orientador: Prof. Dr. Edgard Costa Oliveira

Brasília, DF

2015

Pedro Henrique de Brito Chaves

Desenvolvimento de Ontologia para Estruturas Organizacionais do Governo Brasileiro. / Pedro Henrique de Brito Chaves. – Brasília, DF, 2015-
87 p. : il. (algumas color.) ; 30 cm.

Orientador: Prof. Dr. Edgard Costa Oliveira

Trabalho de Conclusão de Curso – Universidade de Brasília - UnB
Faculdade UnB Gama - FGA , 2015.

1. Dados Abertos. 2. Web Semântica. I. Prof. Dr. Edgard Costa Oliveira.
II. Universidade de Brasília. III. Faculdade UnB Gama. IV. Desenvolvimento de
Ontologia para Estruturas Organizacionais do Governo Brasileiro.

CDU 02:141:005.6

Pedro Henrique de Brito Chaves

Desenvolvimento de Ontologia para Estruturas Organizacionais do Governo Brasileiro.

Monografia submetida ao curso de graduação em (Engenharia de Software) da Universidade de Brasília, como requisito parcial para obtenção do Título de Bacharel em (Engenharia de Software).

Trabalho aprovado. Brasília, DF, :

Prof. Dr. Edgard Costa Oliveira
Orientador

Prof. Dra. Fernanda Lima
Convidado 1

Prof. Dr. Guilherme Novaes Ramos
Convidado 2

Brasília, DF
2015

Resumo

Com a criação lei de acesso a informação em 2011, vários modelos da administração pública do governo federal estão sofrendo transformações a fim de se adequarem as novas mudanças. No contexto deste trabalho, temos o sistema que controla a estruturara organizacional do poder executivo federal, o SIORG – Sistema de Informações Organizacionais do Governo Federal. A objetivo principal desse sistema é funcionar como um cadastro único de unidades organizacionais, a fim de terminar com a miscelânea de conceitos que existem no governo atualmente, decorrentes das várias ambiguidades geradas por outros sistemas que utilizam estes conceitos. Partindo dessa problemática e observando o crescimento da participação dos cidadãos na vigília eletrônica das ações do governo federal, a criação de uma ontologia para representar esse domínio passa a ser um requisito. Neste trabalho de conclusão de curso serão apresentados os principais motivos e razões para construção dessa ontologia, assim como a criação de um roteiro próprio para a criação de um modelo conceitual. Apresenta-se também uma modelagem conceitual do domínio em questão, fruto da aplicação do roteiro definido neste trabalho.

Palavras-chaves: Dados Abertos. Web Semântica. Estrutura Organizacional.

Abstract

With creation of access to information law in 2011, several models of public administration of the federal government are experiencing transformations to fit the new changes. In the context of this research, we have the system that controls the organizational structure of the federal executive power, the SIORG - Organizational Information System of the Federal Government. The major goal of this system is to function as a single register of organizational units in order to end the hodgepodge of concepts that exist in current government, arising from Ambiguities generated by other systems using these concepts. Based on this problem and observing the growth of citizen participation in the electronic waking of the actions of the federal government, creating an ontology to represent this domain becomes a requirement. In this graduation work will be presented the main motives and reasons for construction this ontology, as well as creating their own roadmap for creating a conceptual model. We also present a conceptual model of the domain in question, the fruit of implementation of the roadmap defined in this work.

Key-words: Open Data. Semantic Web. Organizational Structure.

Lista de ilustrações

Figura 1 – Diagrama de implantação da INDA.	25
Figura 2 – Portal de dados abertos.	26
Figura 3 – Países participantes da OGP.	27
Figura 4 – Número total de Websites por ano.	33
Figura 5 – Arquitetura da Web Semântica.	35
Figura 6 – Trecho de código XML destacando dados de um professor.	37
Figura 7 – Documento XML alternativo ao da figura 6.	37
Figura 8 – Forma abstrata de visualizar triplas.	39
Figura 9 – Visualização de triplas através de um grafo.	39
Figura 10 – Trecho de um documento RDF que representa o grafo da figura 9.	40
Figura 11 – Interoperabilidade considerando a diversidade do ecossistema.	46
Figura 12 – Categorias de ontologias.	49
Figura 13 – Método 101 para desenvolvimento de ontologias	50
Figura 14 – Fases da metodologia On-to-Knowledge	53
Figura 15 – Ciclo de vida da <i>methontology</i>	54
Figura 16 – Técnicas de elicitação de requisitos	56
Figura 17 – Representação do Roteiro proposto	61
Figura 18 – Tela principal do SIORG.	64
Figura 19 – Hierarquia de uma unidade organizacional.	73
Figura 20 – Hierarquia de uma Organização Pública.	74
Figura 21 – Hierarquia de uma Entidade.	74
Figura 22 – Hierarquia de uma Unidade Administrativa.	75
Figura 23 – Características principais das Relações entre classes.	76
Figura 24 – Relação dos atributos básicos que compõem uma unidade organizacional.	77
Figura 25 – Modelagem Conceitual.	79
Figura 26 – Árvore taxonômica da ontologia de estrutura organizacional.	80
Figura 27 – Interface do editor de ontologia protégè.	81
Figura 28 – Estrutura Organizacional reduzida da presidência da república.	82

Lista de tabelas

Tabela 1 – Oito principios dos dados aberto governamentais.	30
Tabela 2 – As três leis dos dados abertos governamentais.	30
Tabela 3 – Conceitos Sobre Interoperabilidade.	43
Tabela 4 – Conceitos sobre ontologia.	47
Tabela 5 – Conceitos sobre ontologia.	72
Tabela 6 – Classificação dos Termos da ontologia.	73

Lista de abreviaturas e siglas

CSV	Comma-separated values
DOU	Diário Oficial da União
DTD	Document Type Definition
e-PING	Padrões de Interoperabilidade de Governo Eletrônico
INDA	Infraestrutura nacional de Dados Abertos
IRI	Internationalized Resource Identifier
IP	Internet Protocol
JSON	JavaScript Object Notation
ODT	OpenDocument format
OGD	Open Government
OGP	Open Government Partnership
OWL	Web Ontology Language
RDF	Resource Description Framework
SIORG	Sistema de Informações Organizacionais do governo federal
TCP	Transmission Control Protocol
URI	Uniform resource identifier
URL	Uniform Resource Locator
W3C	World Wide Web Consortium
WEB	World Wide Web
XML	Extensible Markup Language

Sumário

1	INTRODUÇÃO	17
1.1	Objetivos	19
1.1.1	Objetivos Gerais	19
1.1.2	Objetivos Específicos	19
1.2	Metodologia	19
1.3	Organização do Trabalho	21
2	GOVERNO ABERTO	23
2.1	INDA	24
2.2	OGP	26
3	DADOS ABERTOS	29
4	WEB SEMÂNTICA	33
4.1	O que é Web Semântica	33
4.2	Arquitetura da Web Semântica	35
4.2.1	URI/IRI e UNICODE	36
4.2.2	XML, NAMESPACE e XML Schema	36
4.2.3	RDF e RDF Schema	38
4.2.4	Ontologia	40
4.2.5	Lógica	41
4.2.6	Prova	41
4.2.7	Validação	42
5	INTEROPERABILIDADE SEMÂNTICA	43
6	ONTOLOGIAS	47
6.1	Definição de ontologias	47
6.2	Metodologias de desenvolvimento	49
6.2.1	Ontology Development 101	49
6.2.2	On-to-knowledge	51
6.2.3	Methontology	53
7	ROTEIRO DIDÁTICO PARA CONSTRUÇÃO DE ONTOLOGIAS	55
7.1	Iniciação	55
7.1.1	Elaborar cronograma	55
7.1.2	Definir propósito da ontologia	56

7.1.3	Determinar domínio e escopo	56
7.1.4	Levantar recursos de conhecimento	56
7.1.5	Mapear ontologias existentes	57
7.2	Aquisição de conhecimento	57
7.2.1	Definir questões de competência	57
7.2.2	Enumerar termos importantes	57
7.2.3	Selecionar termos reutilizáveis	58
7.2.4	Classificar os termos	58
7.2.5	Definir hierarquia de Classes	58
7.2.6	Relacionar propriedades às classes	59
7.2.7	Relacionar atributos às classes	59
7.2.8	Definir modelo conceitual	59
7.3	Implementação	60
7.3.1	Mapear modelo conceitual para ontologia	60
7.4	Validação	60
7.4.1	Validar ontologia perante suas fontes de conhecimento	60
7.4.2	Validar ontologia perante o usuário	60
7.4.3	Validar ontologia perante as questões de competência	61
8	ESTUDO DE CASO - SIORG	63
9	CONSTRUÇÃO DA ONTOLOGIA	65
9.1	Definir propósito da ontologia	65
9.2	Determinar domínio e escopo	65
9.3	Levantar recursos de conhecimento	66
9.4	Mapear ontologias existentes	67
9.5	Definir Questões de competência	67
9.6	Enumerar termos importantes	68
9.7	Selecionar termos reutilizáveis	71
9.8	Classificar Termos	72
9.9	Definir Hierarquia de Classes	73
9.10	Relacionar propriedades às classes	76
9.11	Relacionar atributos às classes	77
9.12	Definir o modelo conceitual	78
9.13	Mapear modelo conceitual para ontologia	79
10	CONCLUSÃO E CONSIDERAÇÕES FINAIS	83
	REFERÊNCIAS	85

1 Introdução

Atualmente, com avanço tecnológico que estamos sofrendo, surgem diversas ferramentas que aumentam a capacidade da sociedade em assumir seus direitos e obrigações cívicas. A inclusão digital, a informatização dos procedimentos governamentais e a integração entre diversos repositórios de dados públicos gera crescentes demandas da população por mais transparência e participação através de meios tecnológicos.

Durante muito tempo o Brasil sofreu com a falta de leis que pudessem amparar essa disseminação dos dados, leis específicas que determinassem prazos e formas para que o poder público atenda a pedidos de informação pública por parte da comunidade.

O pontapé inicial aconteceu com a lei de nº 12.527, publicamente conhecida como LAI – lei de acesso a informação. Ela foi a primeira lei criada que trata de acesso a informações públicas.

Depois dessa lei, surgiram várias iniciativas da parte do governo na área de dados abertos governamentais. Uma grande conquista foi a criação da INDA – Infraestrutura Nacional de Dados Abertos. A INDA é um conjunto de padrões, tecnologias, procedimentos e mecanismos de controle necessários para atender às condições de disseminação e compartilhamento de dados e informações públicas no modelo de Dados Abertos, em conformidade com o disposto na e-PING. A INDA é a política do governo brasileiro para dados abertos.

No âmbito internacional, o Brasil foi um dos fundadores da OGP (*Open Government Partnership*). A OGP é uma iniciativa internacional que pretende difundir e incentivar globalmente práticas governamentais relacionadas a transparência dos governos, ao acesso à informação pública e à participação social.

Quando estamos falando de governo aberto, não podemos deixar de falar sobre o conceito de Dados Abertos. De acordo com a *Open Definition*¹: “dado aberto é um dado que pode ser livremente utilizado, reutilizado e redistribuído por qualquer um”.

Este conceito resume bem a ideia sobre o que é um Dado Aberto, mas existem diversos princípios e leis que devem ser respeitados para que o dado publicado seja considerado aberto. Essas leis e princípios serão explicitados nos capítulos seguintes.

Para que haja um padrão de publicação desses dados, o governo vem criando uma série de documentos, cartilhas, manuais, guias que ajudem tanto as organizações como seus próprios órgãos a publicarem seus dados.

Dentro do governo existem diversos sistemas que publicam Dados Abertos, mas

¹ Open Definition. Disponível em: <<http://opendefinition.org/od/>>

para esse trabalho focaremos em um em específico, o novo SIORG. O novo SIORG (Sistema de Informações Organizacionais do Governo Federal) é um sistema informatizado que controla as estruturas organizacionais do poder executivo federal. Ele fornecerá uma identificação única para as unidades organizacionais que serão utilizadas em todos sistemas estruturantes possibilitando uma visão global e integrada de suas diversas dimensões: pessoal, orçamento e finanças, custos entre outras.

O problema que estava ocorrendo na atual gestão era que existiam diversos sistemas que precisavam da estrutura organizacional para realizar suas tarefas, só que, em vez dessa informação está centralizada, cada sistema tinha sua tabela de estruturas, causando uma miscelânea de códigos organizacionais e dificultando a interoperabilidade entre eles.

Hoje o SIORG conta com 38 órgãos, 211 entidades, 71.392 unidades administrativas e mais de 100.000 cargos comissionados e de confiança. Devido a essa complexidade, viu-se a necessidade de um planejamento melhor para o armazenamento eficaz desses dados.

O foco desse trabalho realizar alguns passos para construção de uma ontologia que represente a estrutura organizacional do governo federal brasileiro. Serão aplicados alguns conceitos de Web Semântica, bem como de interoperabilidade semântica. Os objetivos principais e como isso será feito será descrito nos capítulos seguintes.

1.1 Objetivos

1.1.1 Objetivos Gerais

Neste trabalho, temos como objetivo o desenvolvimento de uma ontologia para representar estruturas organizacionais do Governo Federal Brasileiro.

1.1.2 Objetivos Específicos

Os objetivos de específicos desse trabalho são:

1. Propor um roteiro didático para construção de modelos conceituais de domínios no âmbito do governo federal brasileiro;
2. Definir uma estrutura semântica do domínio de estruturas organizacionais no contexto do sistema de informações organizacionais do governo federal - SIORG;
3. Iniciar mapeamento da modelagem conceitual para ontologia utilizando a ferramenta *protegè*.

1.2 Metodologia

Neste trabalho estamos realizando uma pesquisa exploratória. Esse tipo de método é muito utilizado quando queremos realizar um estudo preliminar do principal objetivo da pesquisa que será realizada, ou seja, familiarizar-se com o fenômeno que está sendo investigado, de modo que a pesquisa subsequente possa ser concebida com uma maior compreensão e precisão.

Basicamente, essa pesquisa é um trabalho que envolve um levantamento bibliográfico de todos assuntos que circundam o contexto de um objetivo geral.

Para este trabalho, foram levantados alguns temas relacionados ao objetivo principal: governo aberto, Dados Abertos, Web Semântica, interoperabilidade semântica e Ontologia. Para cada um desses assuntos, foi realizada uma pesquisa em diversos tipos de bibliografias com o intuito de produzir um arcabouço teórico que pudesse servir de auxílio para o entendimento da pesquisa posterior.

Quanto aos procedimentos foi escolhido como estudo de caso o sistema novo SIORG, onde se pode aplicar todos temas que foram identificados mais acima.

A pesquisa foi realizada através de uma parceria da UnB (Universidade de Brasília) com o Ministério do Planejamento (MPOG), onde foi estabelecido um termo de cooperação que definia uma série de objetivos e metas a serem cumpridos, um deles se tornando o objetivo deste trabalho.

O primeiro passo para a realização desse trabalho foi o ingresso no projeto por meio do professor Dr. Edgar Costa, um dos pesquisadores contratados para realização do projeto. Depois desse primeiro contato, foram realizadas inúmeras reuniões entre UnB e MPOG com o intuito de esclarecer melhor a problemática atual e definir uma espécie de cronograma final contendo as datas da maioria dos entregáveis do projeto.

Depois dessas várias reuniões realizadas e da coleta de todo o material existente sobre o estudo de caso, foram definidas metodologias individuais para a concepção de cada objetivo específico deste trabalho. Abaixo serão listados esses objetivos e suas respectivas metodologias.

1. Propor um roteiro didático para construção de modelos conceituais de domínios no âmbito do governo federal brasileiro.

Para a realização desse objetivo, serão realizados os seguintes passos metodológicos:

- a) levantamento das metodologias de construção de ontologias mais comumente usadas. O foco dessa pesquisa será metodologias que descrevem etapas de modelagem de dados que podem ser adaptadas para a finalidade específica deste objetivo.
- b) Realizar um estudo mais aprofundado dessas metodologias escolhidas a fim de coletar informações sobre as melhores estratégias para elaboração de modelos conceituais.
- c) Definição do roteiro didático próprio com base em todas as informações coletadas sobre as metodologias estudadas.

Esse roteiro conterá todas as etapas necessárias para a produção de um modelo conceitual.

2. Definir uma estrutura semântica do domínio de estruturas organizacionais no contexto de sistema de informações organizacionais do governo federal SIORG;

Para a realização desse objetivo, serão realizados os seguintes passos metodológicos:

- a) Identificar toda a documentação existente do SIORG: manuais, vídeos, *wiki*, entre outros;
- b) Elaborar um glossário com os termos identificados no domínio;
- c) Analisar e selecionar ferramenta para construção de modelos conceituais;
- d) Aplicar o roteiro definido no objetivo específico 1 e gerar um modelo conceitual do SIORG.

Vale ressaltar que o modelo conceitual a ser produzido não represente só os conceitos do SIORG, ou seja, o modelo será desenvolvido de uma forma geral o suficiente

para permitir o seu reaproveitamento em futuras modelagens conceituais de estruturas organizacionais em outras esferas ou poderes da administração pública do Brasil. Outro ponto que está intrínseco no roteiro para produção do modelo, é que, sempre que possível, as classes e propriedades serão relacionadas com ontologias internacionalmente utilizadas pelas comunidades, com intuito de uma possível integração com essas ontologias no futuro.

1.3 Organização do Trabalho

Este trabalho está organizado em capítulos. O Capítulo 2 apresenta os principais conceitos sobre governo aberto, INDA e OGP. Já o capítulo 3 trás não só principais conceitos sobre Dados abertos, como também leis e princípios que os circundam. O capítulo 4 fala sobre Web Semântica. São definidos diversos conceitos sobre o assunto, bem como a descrição de todas as camadas de sua arquitetura. O capítulo 5 trás definições sobre o conceito de interoperabilidade semântica e como ela está relacionada com o contexto do governo federal. O capítulo 6 é voltado unicamente para o assunto de ontologias, como sua definição e metodologias de desenvolvimento. O capítulo 7 define um roteiro didático, uma espécie de guia, para ajudar os desenvolvedores de ontologias. O capítulo 8 fala sobre o estudo de caso abordado nesse trabalho – SIORG. O capítulo 9 simula a aplicação do roteiro definido nesse trabalho no estudo de caso supracitado. E por fim o capítulo 10 que se refere a conclusão. Ela relata a importância do projeto, bem como os resultados alcançados na sua conclusão.

2 Governo Aberto

De acordo com Daniela Siva (SILVA, 2010), A publicidade dos atos de governo é um princípio democrático, que no Brasil aparece expressa no artigo 5º da Constituição: “todos são iguais perante a lei, sem distinção de qualquer natureza, garantindo-se aos brasileiros e aos estrangeiros residentes no País a inviolabilidade do direito à vida, à liberdade, à igualdade, à segurança e à propriedade”. No inciso XXXIII deste artigo, determina-se que “todos têm direito a receber dos órgãos público informações de seu interesse particular, ou de interesse coletivo ou geral, que serão prestadas no prazo da lei, sob pena de responsabilidade, ressalvadas aquelas cujo sigilo seja imprescindível à segurança da sociedade e do Estado”. No artigo 37 da Constituição, fica explícito que “a administração pública direta e indireta de qualquer dos Poderes da União, dos estados, do distrito federal e dos municípios obedecerá aos princípios de legalidade, impessoalidade, moralidade, publicidade e eficiência”.

Atualmente, com avanço tecnológico que estamos sofrendo, surgem diversas ferramentas que aumentam a capacidade da sociedade em assumir seus direitos e obrigações cívicas. A inclusão digital, a informatização dos procedimentos governamentais e a integração entre diversos repositórios de dados públicos gera crescentes demandas da população por mais transparência e participação através de meios tecnológicos.

Mesmo com o crescente interesse da sociedade atual nos dados públicos e com o amparo da constituição, até então no Brasil ainda não existiam leis específicas que determinassem prazos e formas para que o poder público atenda a pedidos de informação pública por parte da comunidade.

Em 2011, foi sancionada a lei N° 12.527, mais conhecida como lei de acesso a informação (LEI..., 2011). Foi a primeira lei criada que trata de acesso as informações públicas. Antes de ser aprovada, ela passou por algumas reformulações sugeridas pela comunidade Hacker, afim de garantir o acesso a dados abertos. A lei engloba todos os oito princípios de dados abertos.

Ainda de acordo com Daniela Silva (SILVA, 2010), os “Governos e departamentos interessados em fazer a abertura de seus dados, portanto, podem seguir como princípio as determinações do projeto de Lei de Acesso à Informação Pública, que está de acordo com as possibilidades de gerar cruzamentos, visualizações e serviços garantidos pelas novas tecnologias em rede.”

2.1 INDA

A lei de acesso a informação (LEI... , 2011) foi o pontapé inicial para várias iniciativas na área de dados abertos governamentais (OGD¹ – open government data). Hoje, o governo vem promovendo diversas ações tanto na esfera administrativa quanto na legislativa, dando apoio a abertura de dados.

Em 13 de abril de 2012, foi publicado no Diário Oficial da União (DOU), a instrução normativa de N° 4 que institui o Plano de Ação Nacional sobre Governo Aberto, o qual estabelece o compromisso do governo de implantar a Infraestrutura Nacional de Dados Abertos - INDA.

Como disposto no site de dados abertos do governo federal, a INDA é um conjunto de padrões, tecnologias, procedimentos e mecanismos de controle necessários para atender às condições de disseminação e compartilhamento de dados e informações públicas no modelo de Dados Abertos, em conformidade com o disposto na e-PING². A INDA é a política do governo brasileiro para dados abertos.

Todas essas diretrizes da INDA são melhores descritas no seu plano de ação (PLANO... , 2012). Esse plano apresenta três finalidades principais:

- Auxiliar as organizações integrantes da INDA a cumprir a Lei de Acesso à Informação, no que se refere à transparência ativa pela publicação de dados governamentais abertos;
- Nortear os órgãos e entidades integrantes da INDA, quanto à visão, estratégia e política de abertura de dados na administração pública federal, para os anos de 2013 e 2014;
- Servir como base para criação de planos de publicação de dados abertos na INDA, conforme disposto no Art. 6º, VII, alínea c da Instrução Normativa 04 de 12 de abril de 2012.

¹ OGD - Disponível em: <<http://opengovernmentdata.org/>>.

² e-PING - Padrões de Interoperabilidade de Governo Eletrônico. Disponível em:<<http://eping.governoeletronico.gov.br/>>

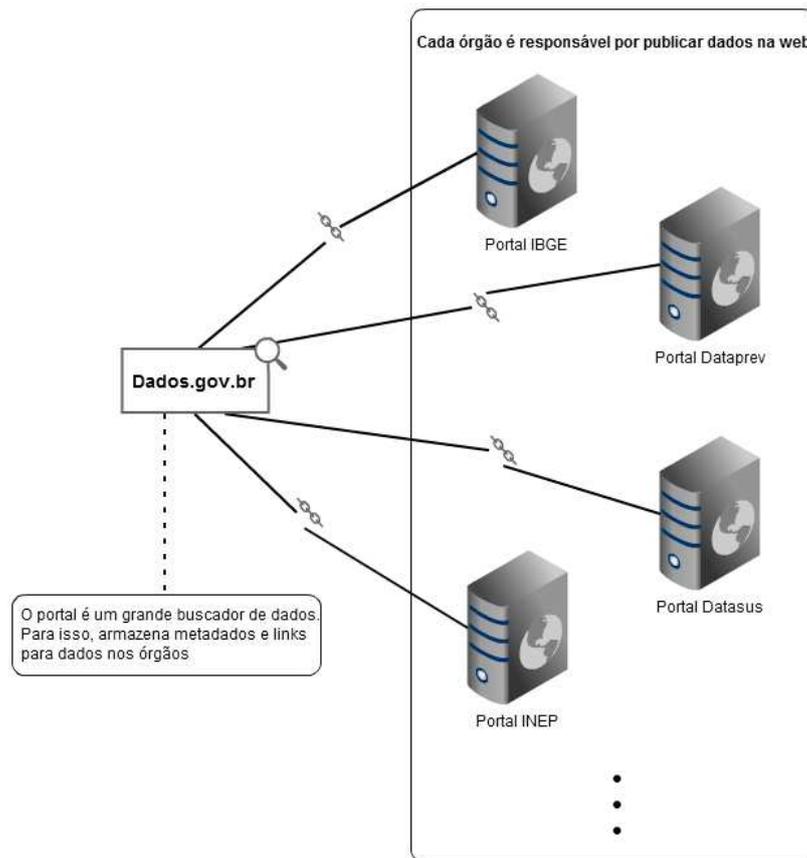


Figura 1: Diagrama de implantação da INDA. Extraído de (PLANO..., 2012)

Como podemos ver na figura 1, a INDA é responsável por manter o portal brasileiro de dados abertos. Esse portal é um catálogo central que mantém um conjunto de metadados sobre as informações disponibilizadas pelas organizações da INDA. Informações como nome do dado, URL (ou endereço web) do dado, autor do dado, responsável pela manutenção do dado, formato do dado (odt, csv, json, xml, etc) são catalogadas para garantir que o usuário encontre o que está procurando.

Desse modo, o portal funcionará como um grande buscador de dados. Cada organização participante da INDA será responsável por publicar seus dados na web através de portal próprio, cadastrar seus metadados no portal (endereço, nome, data da coleta, assunto, etc) e posteriormente garantir a disponibilidade desses dados em seu ambiente próprio.

Para que essas organizações possam participar da INDA, elas deverão seguir uma série de recomendações para que os dados disponibilizados sejam mais úteis, reutilizáveis e fáceis de encontrar. Um importante documento que contém inúmeras boas práticas para esse processo de publicação é a Cartilha Técnica para publicação de Dados Abertos (MINISTERIO DO PLANEJAMENTO, 2012), disponível no portal.

The screenshot shows the 'Dados Abertos' portal. At the top, there is a yellow navigation bar with links for 'BRASIL', 'Acesso à informação', 'Faltam 6 dias para a Copa', 'Participe', 'Serviços', 'Legislação', and 'Canais'. Below this is a green header with the text 'Dados Abertos GOVERNO FEDERAL' and social media icons. The main content area includes a search bar with the text 'Pesquisa ...' and a 'PEQUISAR' button. Below the search bar, it says 'em 238 conjuntos de dados com 2132 recursos (o que é isto?)'. There are three main sections: 'Dados em destaque' with featured datasets, a table of 'Conjunto de dados', and 'Dados por etiqueta' and 'Notícias' sections.

Conjunto de dados	Mês	Totais
Lista de Eleitores Filiados aos Partidos Políticos	609	1657
Microdados do Censo Escolar	450	863
Destinação do Lixo Domiciliar	437	1072
Obras do PAC - Programa de Aceleração do Crescimento	422	1187
População Economicamente Ativa por sexo	302	786

Figura 2: Portal de dados abertos. Extraído de (DADOS..., 2014)

2.2 OGP

A parceria para Governo Aberto ou OGP³ (do inglês open Government Partnership), lançada em 2011, é uma iniciativa internacional que pretende difundir e incentivar globalmente práticas governamentais relacionadas a transparência dos governos, ao acesso à informação pública e à participação social (PARTICIPATING..., 2014).

Inicialmente ela foi composta por oito países: África do sul, Brasil, Estados Unidos, Filipinas, Indonésia, México, Noruega e Reino Unido. Esses países são considerados os fundadores da OGP. Eles oficializaram essa parceria quando assinaram a Declaração de Governo Aberto e apresentaram seus planos de ação. Atualmente, a OGP conta com 63 países.

³ OGP - Disponível em: <<http://www.opengovpartnership.org/>>

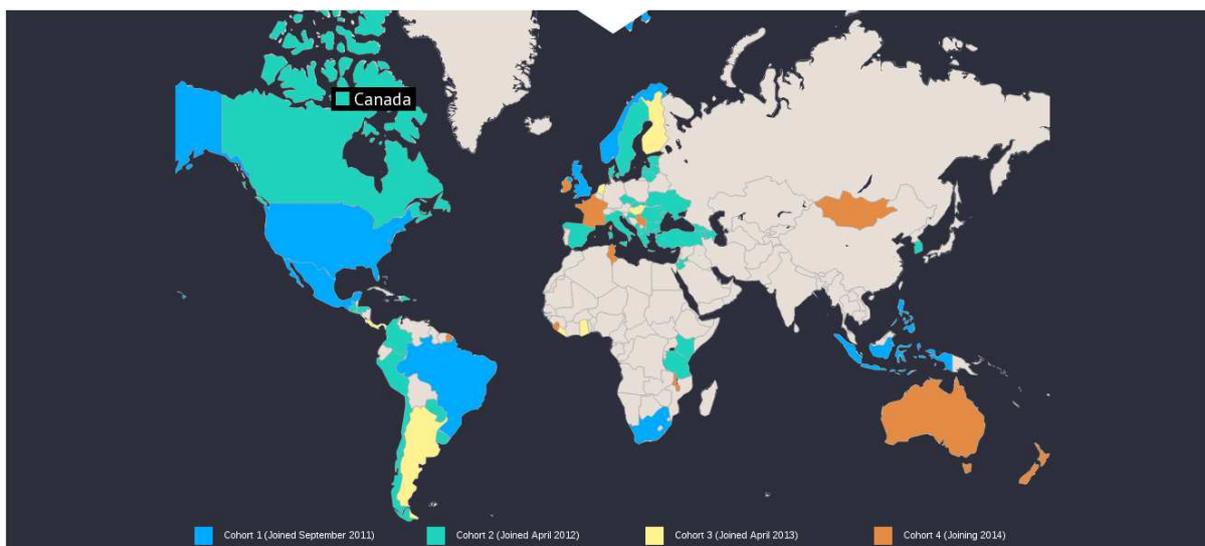


Figura 3: Países participantes da OGP. Extraído de ([PARTICIPATING...](#), 2014)

Como um dos co-fundadores da OGP, o Brasil está fortemente empenhado em reforçar a transparência das ações do governo, prevenção e combate à corrupção, a promoção dos ideais democráticos com a participação dos cidadãos na tomada de decisões e melhoria dos serviços públicos. Ao longo dos últimos 10 anos, o país desenvolveu várias iniciativas para melhorar o seu quadro legal, a participação do cidadão e fomentar o uso da tecnologia para uma maior abertura.

Todas as medidas a serem executadas pelo Brasil, ficam registradas em seu plano de ação. No primeiro plano, o Brasil se comprometeu com 32 compromissos. O país conseguiu implementar total ou parcialmente cerca de 90

Responsável com congregar nações e organizações da sociedade civil líderes em transparência e governo aberto, a OGP é um veículo para se avançar mundialmente no fortalecimento das democracias e dos direitos humanos, na luta contra a corrupção e no fomento de inovações e tecnologias para transformar a governança do século XXI mais transparente.

3 Dados Abertos

De acordo com a definição da Open Definition: “dado aberto é um dado que pode ser livremente utilizado, reutilizado e redistribuído por qualquer um”. A definição completa pode ser encontrada em (OPEN..., 1999), mas abaixo serão listados alguns pontos importantes para o contexto em questão:

- Disponibilidade e Acesso: os dados devem estar disponíveis como um todo e sob custo não maior que um custo razoável de reprodução, preferencialmente possíveis de serem baixados pela internet. Os dados devem também estar disponíveis de uma forma conveniente e modificável.
- Reutilização e Redistribuição: os dados devem ser fornecidos sob termos que permitam a reutilização e a redistribuição, inclusive a combinação com outros conjuntos de dados.
- Participação Universal: todos devem ser capazes de usar, reutilizar e redistribuir - não deve haver discriminação contra áreas de atuação ou contra pessoas ou grupos. Por exemplo, restrições de uso ‘não-comercial’ que impediriam o uso ‘comercial’, ou restrições de uso para certos fins (ex.: somente educativos) excluem determinados dados do conceito de ‘abertos’.

Os dados abertos podem vir de qualquer fonte. Existem dados abertos na ciência, dados de empresas privadas e o mais importante para esse trabalho, Dados Abertos no governo.

Dados abertos governamentais são dados produzidos pelo governo e colocados à disposição das pessoas de forma a tornar possível não apenas sua leitura e acompanhamento, mas também sua reutilização em novos projetos, sítios e aplicativos; seu cruzamento com outros dados de diferentes fontes; e sua disposição em visualizações interessantes e esclarecedoras. (MINISTERIO DO PLANEJAMENTO., 2011).

Apenas o fato de uma organização publicar seus dados na Web não os tornam abertos. Para que eles sejam considerados Dados Abertos, a organização deve respeitar uma série de princípios, onde os dados devem ser:

Completos	Todos os dados públicos são disponibilizados. Dados públicos são dados que não se submetem à limitações válidas de privacidade, segurança ou privilégio.
Primários	Os dados são coletados na sua fonte, com o maior nível possível de granularidade, não estando em formas agregadas ou modificadas.
Atualizados	Os dados são disponibilizados tão rápido quanto seja necessário para preservar seu valor.
Acessíveis	Os dados estão disponíveis para o maior escopo possível de usuários e para o maior escopo possível de finalidades.
Não-discriminatórios	Os dados estão disponíveis para todos, sem necessidade de registro.
Não-proprietários	Os dados são disponibilizados num formato do qual nenhuma entidade tem controle exclusivo.
Livres de licenças	Os dados não estão sujeitos a nenhuma forma de direito autoral, patente, propriedade intelectual ou segredo industrial. Restrições razoáveis de privacidade, segurança e privilégio podem ser permitidas.

Tabela 1: Oito princípios dos dados aberto governamentais.

Todos os oito princípios citados na tabela acima foram definidos num evento que reuniu mais de 30 ativistas em prol da abertura dos governos. O evento foi realizado em Sebastopol, na Califórnia, e tinha o objetivo de “desenvolver um entendimento mais robusto de porque dados governamentais abertos são essenciais para a democracia”.

Analisando os princípios elencados na tabela 1, podemos perceber que eles garantem que a disponibilização de dados governamentais seja orientada de acordo a possibilitar a apropriação desses dados por parte dos cidadãos, que podem reutilizá-los na rede.

Um outro ativista que contribuiu também para a definição de dados governamentais foi o David Eaves ([EAVES, 2009](#)). Em 30 de setembro de 2009, Eaves apresentou o painel *Conference for Parliamentarians: Transparency in the Digital Era*, em um evento no Canadá, onde o objeto era debater e refletir sobre o novo paradigma que o mundo digital inaugura para o direito a informação. Como parte desse painel, Eaves apresentou três leis dos dados governamentais abertos, que seguem descritas no quadro abaixo:

Se o dado não pode ser encontrado e indexado na web, ele não existe.
Se não estiver aberto e disponível em formato compreensível por a máquina, ele não pode ser reaproveitado.
Se algum dispositivo legal não permitir sua replicação, ele não é útil.

Tabela 2: As três leis dos dados abertos governamentais.

O conceito de dados governamentais abertos, portanto, se relaciona com um entendimento de que a forma como os governos disponibilizam suas informações permite que a inteligência coletiva crie melhores formas de trabalhar com elas do que os próprios governos poderiam fazer (SILVA, 2010).

Para que as organizações consigam publicar seus dados, o governo vem disponibilizando uma série de documentos que podem servir de guias para os próprios órgãos do governo, desenvolvedores ou qualquer interessado em abertura de dados. São alguns exemplos, entre outros:

- Manual dos Dados Abertos: Governo;
- Manual dos Dados Abertos: Desenvolvedores;
- Cartilha para desenvolvedores;
- Folheto sobre Dados Abertos;
- Cartilha Técnica para Publicação de Dados Abertos no Brasil;
- Guia de Abertura de Dados;
- Arquitetura Técnica Referencial para Abertura de Dados;
- Guia de Dados Abertos (*Open Knowledge Foundation*).

Com esse arcabouço teórico e com a divulgação da importância de se publicar os dados de forma aberta pelo governo federal, espera-se que a população e o governo sejam beneficiados de várias formas, como por exemplo:

- Transparência e controle democrático;
- Participação popular nas decisões das entidades que abriam dados;
- Empoderamento dos cidadãos;
- Melhores ou novos produtos e serviços privados;
- Inovação através do reuso dos dados como informação;
- Melhoria na eficiência e na efetividade de serviços governamentais;
- Medição do impacto das políticas;
- Conhecimento novo a partir da combinação de fontes de dados e padrões.

4 Web semântica

4.1 O que é Web Semântica

A ideia de Web Semântica surgiu em 2001, após a publicação de um artigo através da revista *Scientific American*, denominado: *The semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities* (Web Semântica: um novo formato de conteúdo para a Web que tem significado para computadores vai iniciar uma revolução de novas possibilidades.”. Este artigo foi elaborado por Tim Berners-Lee, James Hendler e Ora Lassila (EBERTZ, 2013).

Para que possamos entendê-la melhor, precisamos compreender como funciona a Web¹ atual e como ela chegou no que é hoje.

No início da internet, as páginas eram desenvolvidas por programadores de *software* (BREITMAN, 2005). Essas páginas eram feitas exclusivamente para apresentação da informação, ou seja, o processo de interpretação ficava todo a cargo dos seres humanos.

Com o passar dos anos, a internet foi ficando cada vez mais popular. De acordo com o site *internet live stats* (TOTAL..., 2014), no final de 2014, a web contará com aproximadamente 1 bilhão de *websites*.

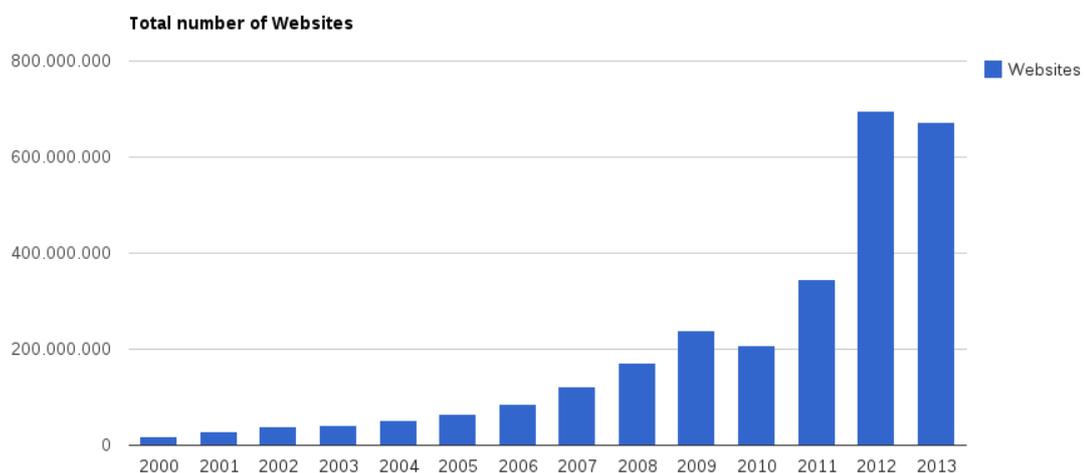


Figura 4: Número total de *Websites* por ano. Extraído de (TOTAL..., 2014)

O grande problema desse avanço é que a maioria dos *Websites* criados ainda mantém sua característica inicial, ou seja, ainda são feitos para as pessoas interpretarem e

¹ Web - *World Wide Web* (em inglês). É sistema hipertextual que opera através da Internet.

não as máquinas. De acordo com a Karin Breitman (BREITMAN, 2005), essa Web atual pode ser definida como Web Sintática.

Segue um exemplo para que possamos entender melhor esse conceito: vamos supor que você esteja pensando em tirar umas férias. Você deseja visitar um lugar quente e tropical e reservou um orçamento de 3.000 reais para a sua viagem. Você deseja ficar em um bom lugar, mas não quer que isso custe muito em seu orçamento. Você também quer fazer um bom negócio com as passagens de avião (STRICKLAND, 2014).

Com os recursos da Web (sintática) que temos atualmente, nós teríamos que pesquisar bastante para encontrarmos a melhor opção. Teríamos que visitar vários sites de passagens aéreas e de hotéis e ainda comparar os preços entre eles. Seria um processo muito trabalhoso.

Em (BREITMAN, 2005), são enumerados os maiores problemas que temos com os atuais mecanismos de busca na Internet, através de ferramentas do tipo *Google*², *Yahoo*³ e *Bing*⁴, por exemplo, como se segue:

- Grande numero de páginas encontradas, porém com pouca precisão – Por exemplo, ao realizar uma busca por TCP/IP no *Google*, temos aproximadamente 14.900.000 resultados. Mesmo encontrando páginas relevantes, esse resultado seria de pouca utilidade caso a maioria das páginas fossem de pouca relevância.
- Resultados são muitos sensíveis ao vocabulário – em determinados casos, até a ordem em que as palavras são digitadas tem impacto nos resultados. Muitas vezes os documentos relevantes acabam usando terminologias diferente das nossas.
- Resultados são páginas individuais – em muitos casos temos um grande número de páginas no resultado que pertencem a um mesmo site. Seria mais interessante ter algum tipo de organização geográfica dos resultados. Ao final, temos de extrair manualmente as porções desses documentos de interesse.

O que podemos concluir dessas situações citadas acima, é que a Internet se tornou um meio para se compartilhar documentos entre pessoas, ao invés de ser um meio em que a troca de dados e informações pudessem ser processadas automaticamente.

No meio desse caos , Tim Berners-Lee, considerado por muitos o criador da Internet, apostou no aparecimento de uma Web mais organizada, mais conectada. Ele a chamou de Web Semântica.

² Google - Disponível em: <<https://www.google.com.br/>>

³ Yahoo - Disponível em: <<https://br.yahoo.com/>>

⁴ Bing - Disponível em: <<https://www.bing.com/>>

De acordo com Bernes-Lee, Hendler e Lassila: “A Web Semântica é uma extensão da Web atual, na qual é dada à informação um significado bem definido, permitindo que computadores e pessoas trabalhem em cooperação”.

A ideia central é encontrar uma maneira de categorizar o conteúdo da Web de forma padronizada, facilitando seu acesso (BREITMAN, 2005). A Web Semântica não se trata de uma nova rede de informações, mas sim de um projeto para aplicar conceitos inteligentes na internet atual. Nela cada informação vem com um significado bem definido e não se encontra mais solta no mar de conteúdo, permitindo uma melhor interação com o usuário.(PRADA, 2008).

4.2 Arquitetura da Web Semântica

Não sabemos ainda como a web semântica será efetivamente construída, mas já existe uma arquitetura definida pela W3C (*World Wide Web Consortium*). Segue abaixo sua representação:

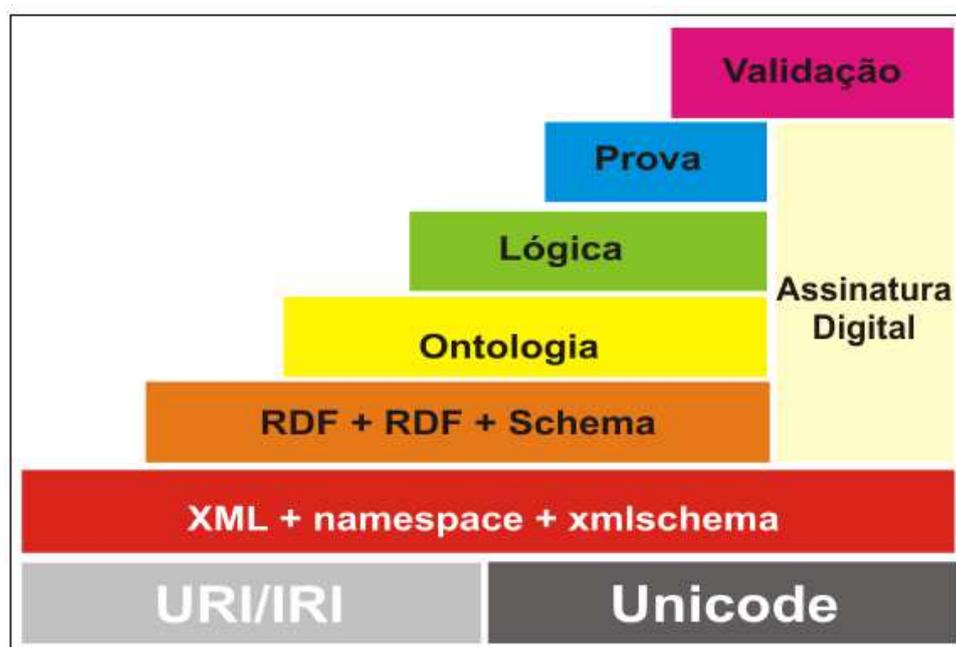


Figura 5: Arquitetura da Web Semântica. Extraído de (JORGE; REIS, 2013)

Como podemos perceber, a arquitetura da web semântica é formada por 7 camadas, cada uma com uma ferramenta e tecnologia diferente (ALVES, 2003). Tim Berners-Lee, definiu uma estrutura em camadas que reflete os passos que devem ser dados para que o projeto da Web Semântica seja realizado de uma forma incremental (FERNEDA, 2003). As próximas seções falaram um pouco sobre essas sete camadas.

4.2.1 URI/IRI e UNICODE

Vamos realizar uma abordagem *bottom-up* (nota rodapé), ou seja, vamos começar a análise pela camada mais inferior. Esta camada é composta pela URI (*Uniform Resource Identifier*) e UNICODE que são padrões para a descrição e estabelecimento de identificadores universais do recurso e códigos internacionais de dados (SEGUNDO, 2003). Esses dois elementos são responsáveis pela designação de uma identificação mínima dos recursos na rede (ALVES, 2005).

Segundo (RIBEIRO, 2013), o URI é um padrão para identificar um recurso físico ou abstrato de maneira única e global. Ele estabelece uma forma padrão para a identificação de recursos. Através da utilização de URI faz-se a referência para recursos representados na Web Semântica. No contexto da Internet, o conceito de URI já é bem utilizado. Na Web é utilizado um tipo de URI chamado URL. Através da URL é possível endereçar documentos utilizando protocolos específicos da Internet como http e ftp (ROSA, 2002).

Já o UNICODE, de acordo com (ROSA, 2002), é uma linguagem que define uma forma padrão para a representação de caracteres. Unicode proporciona uma forma única para a representação de um caractere não importando a plataforma, o programa nem a linguagem que está sendo utilizada. A utilização de Unicode na Web Semântica proporciona a capacidade de troca de símbolos de maneira universal, requisito fundamental para o sucesso desta nova proposta de representação de informação na Internet.

4.2.2 XML, NAMESPACE e XML Schema

Essa camada 2, também chamada de camada sintática, é responsável pelo estabelecimento correto da sintaxe de descrição dos dados.

O XML é uma linguagem de marcação que, diferentemente do HTML, permite a criação e o uso de tags personalizadas, fornecendo assim uma maneira simples de organizar e estruturar os dados existentes em uma determinada aplicação (EXTENSIBLE... , 2009). Atualmente, o XML é a linguagem padrão recomendada pelo W3C⁵ para troca de dados via Web.

Hoje, a Web Semântica exige uma descrição formal da semântica dos dados, de tal forma a evitar ambiguidades e permitir a interpretação de informações por parte das aplicações. Neste contexto, XML provê uma sintaxe bem definida, sendo atualmente utilizado na maioria das aplicações existentes na Web (FILHO; LOSCIO, 2009).

⁵ W3C - Disoinível em: <<http://www.w3c.br/Home/WebHome>>

```
<Professor>
  <id>12345</id>
  <Nome>Fernando Wagner</Nome>
  <Area>Banco de Dados</Area>
</Professor>
```

Figura 6: Trecho de código XML destacando dados de um professor.

```
<Professor id = '12345'>
  <Descricao>
    <Nome>Fernando Wagner</Nome>
    <Area>Banco de Dados</Area>
  </Descricao>
</Professor>
```

Figura 7: Documento XML alternativo ao da figura 6.

Um fato importante a ser levado em consideração no XML, é que um mesmo conjunto de dados podem ser representados de várias formas. Se observarmos figura 7, por exemplo, podemos perceber que ela é uma versão alternativa a figura 6. Essa característica pode causar algumas discordâncias entre aplicações.

Para contornar esses problemas, foram criadas linguagens que definem esquemas para XML. São uma espécie de contrato, onde todas as partes envolvidas por um contexto de aplicação devem escrever seus documentos XML seguindo o padrão de estruturação especificado no esquema XML correspondente. Dentre as linguagens para definição de esquemas XML, destacam-se DTD⁶ e XML Schema (FILHO; LOSCIO, 2009).

De acordo com (ROSA, 2002), XML Schema é uma ferramenta que permite a definição e a descrição de estruturas e de conteúdos de documentos XML. Através dessa linguagem, define-se o formato válido de um documento XML, incluindo quais elementos e atributos são permitidos ou não, quais são as suas localizações, o número de ocorrências de cada elemento e outras características, Ou seja, proporciona mecanismos para a definição de gramáticas para correção de documentos XML.

O último elemento dessa camada, são os chamados *namespaces*. Eles são classificados como um método para qualificar nomes de elementos e atributos usados em documentos XML, através da associação de referências URI. Através desse mecanismo de espaço de nomes, é possível a combinação de documentos com a utilização de vocabulário compartilhado. Através do mecanismo de espaço de nomes definido em XML, é possível

⁶ DTD - *Document Type Definition*. Disponível em: <<http://www.w3schools.com/DTD/>>

compartilhar e reutilizar a definição de outros esquemas XML sem que haja problemas de colisão de nomes (ROSA, 2002) .

4.2.3 RDF e RDF Schema

Essa camada também pode ser chamada de camada de dados. Ela está diretamente relacionada com a representação, o processamento e a codificação dos metadados (ALVES, 2005). Para isso estão presentes nessa camada a arquitetura de metadados RDF e o RDF Schema, que são ferramentas responsáveis por expressar significados e promover a interoperabilidade entre metadados e padrões ou formatos de metadados (SEGUNDO, 2003).

Segundo Rosa (ROSA, 2002), RDF (*Resource Description Framework*) é uma linguagem para representação de informação na Web. Trata-se de uma infra-estrutura que fornece a habilidade para codificação, troca e reutilização de metadados. RDF define um modelo de dados para descrição de semântica de dados para o entendimento do computador. É o fundamento para o processamento de metadados (informação sobre informação).

O RDF veio como uma alternativa para o XML, que mesmo sendo recomendado pelo W3C e amplamente utilizado em aplicações, tinha muitas limitações para descrever adequadamente a semântica de uma informação. Com o RDF, conseguimos expressar como os elementos devem se relacionar.

Como já foi dito, o RDF é um modelo de dados. Esse modelo possibilita a definição de afirmações, chamadas sentenças, sobre um recurso.

De acordo com (FILHO; LOSCIO, 2009), Entende-se por um recurso, “qualquer coisa” sobre a qual se quer expressar uma ideia. Um recurso pode estar relacionado com dados ou com outros recursos através das sentenças. Uma sentença é estruturada no formato sujeito + predicado + objeto onde:

- Sujeito: Tem como valor o recurso do qual se quer escrever uma sentença.
- Predicado: Especifica um relacionamento entre sujeito e objeto. O predicado é especificado através de propriedades, que são relações binárias, geralmente nomeadas por um verbo e permitem relacionar um recurso a dados ou a outros recursos.
- Objeto: Denomina o recurso ou dado que se relaciona ao sujeito.

Por causa de seu formato, uma sentença também é chamada de Tripla. Logo, um documento de RDF pode ser visto como um conjunto de triplas que descrevem informações sobre recursos de um certo domínio. Abaixo, vemos uma forma abstrata de representar essas triplas:

Subject	Predicate	Object
http://www.w3.org/RDF/Validator/run/p91002043177	http://uni.org/uni/elements/1.1/nomeDocente	"Berna Farias"
http://www.w3.org/RDF/Validator/run/CK120	http://uni.org/uni/elements/1.1/nomeDisciplina	"Banco de Dados I"
http://www.w3.org/RDF/Validator/run/CK120	http://uni.org/uni/elements/1.1/EnsinadoPor	http://www.w3.org/RDF/Validator/run/p91002043177

Figura 8: Forma abstrata de visualizar triplas. Extraído de (FILHO; LOSCIO, 2009)

Realizando uma breve análise da figura 8, percebemos que ela é composta por dois recursos: “p91002043177” e “CK120”.

Na primeira tripla, temos o recurso “p91002043177” recebendo o nome “Berna Farias”.

Na segunda, o recurso “CK120” recebe o nome “Banco de Dados”.

E na terceira tripla, temos a descrição de um relacionamento entre os dois recursos. Essa Relação foi criada através do predicado “EnsinadoPor”.

Uma outra forma de representação muito utilizada é a de grafos. Esses grafos são compostos de nós e arcos, que representam os recursos, suas propriedades e valores. Abaixo, é mostrado uma representação de um grafo da figura 8.

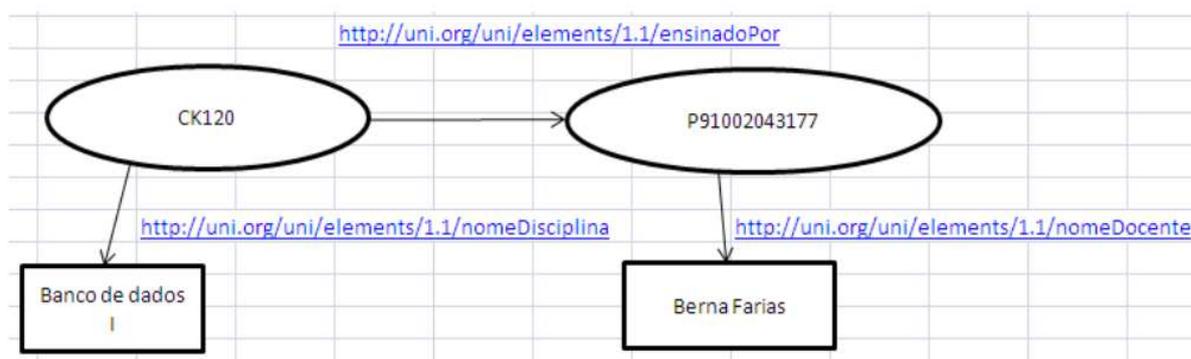


Figura 9: Visualização de triplas através de um grafo. Extraído de (FILHO; LOSCIO, 2009)

Na figura 9, Os chamados Nós literais são representados pelas caixas e os Nós referência são representados pelas URIs nas elipses.

Para Conseguimos uma certa compatibilidade entre a escrita e os grafos RDFs, foi criada uma linguagem chamada RDF/XML. Com isso, um documento escrito em RDF

pode ser facilmente transformado em um grafo caso ele siga a sintaxe da linguagem. Abaixo, segue uma representação do grafo em uma linguagem com base em XML:

```
1: <?xml version="1.0"?>
2: <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
3:   xmlns:uni="http://uni.org/uni/elements/1.1/">
4:
5: <rdf:Description rdf:ID="P91002043177"/>
6:
7: <rdf:Description rdf:about="P91002043177">
8:   <uni:nomeDocente>Berna Farias</uni:nomeDocente>
9: </rdf:Description>
10:
11: <rdf:Description rdf:about="CK120">
12:   <uni:nomeDisciplina>Banco de Dados I</uni:nomeDisciplina>
13:   <uni:ensinadoPor rdf:resource="P91002043177"/>
14: </rdf:Description>
15:
16: </rdf:RDF>
```

Figura 10: Trecho de um documento RDF que representa o grafo da figura 9 . Extraído de (FILHO; LOSCIO, 2009)

O RDF Schema tem uma função parecida com o XML Schema da seção anterior. De acordo com (ROSA, 2002): é uma linguagem que define a estrutura válida para dos documentos RDF. RDF e RDF Schema são recomendações do consórcio W3C que definem o padrão para a representação de metadados. São a base de todas as linguagens para expressar semântica da Web Semântica, devido à adoção pelo consórcio W3C.

Portanto, considerando que o RDF pode conter inconsistências, o RDFS visa solucionar tais problemas provendo construtores que permitem especificar formalmente um esquema (FILHO; LOSCIO, 2009). O estratégia é pra funcionar da seguinte forma: todas as sentenças descritas num documento RDF deverá obedecer à semântica descrita no esquema RDFS, onde esse esquema nada mais é que a modelagem do domínio de interesse.

4.2.4 Ontologia

O RDF conseguiu suprir algumas deficiências do XML em relação a semântica de dados, mas comparado a ontologias ele se torna bastante limitado. Ele basicamente serve para escrever sentenças sem que haja qualquer descrição formal de um domínio (FILHO; LOSCIO, 2009). As ontologias surgiram para suprir algumas características que o RDF não contempla. De acordo com (FILHO; LOSCIO, 2009), são algumas dessas características:

- Restrições de propriedades: Muitas vezes precisamos impor restrições nos valores que uma propriedade pode assumir. Por exemplo, não conseguimos dizer em RDF/RDFS que um time de futebol tem que ter, no mínimo, onze jogadores para poder disputar uma partida.
- Disjunção de classes: No domínio-alvo pode acontecer de classes (conceitos) serem disjuntos. Por exemplo, homem e mulher são dois conceitos disjuntos, pois uma pessoa não pode ser do sexo masculino e feminino ao mesmo tempo. Em RDF/RDFS, é possível somente expressar relações de hierarquia como mulher é subclasse de pessoa.
- Combinação entre classes: RDF/RDFS não permite que se criem novos conceitos utilizando uma combinação de conceitos já especificados usando, por exemplo, a união ou interseção destes.
- Características de propriedades: Também não é possível especificar na camada de RDF/RDFS algumas características de propriedades como, por exemplo, a transitividade de valores.

Essa camada é importante, pois além de ter a definição dos significados e semântica dos dados é nela que estão estabelecidos os esquemas classificatórios utilizados pelos agentes de softwares (SEGUNDO, 2003). Vamos descreve-la melhor num outro capítulo.

4.2.5 Lógica

De acordo com (ROSA, 2002), a camada de Lógica proporciona a definição de semântica em linguagem formal habilitando a execução de serviços inteligentes. É composta principalmente por regras de inferência, com as quais os agentes poderão se utilizar para relacionar e processar informação.

Vamos a um exemplo extraído de (FERNEDA, 2003): imaginando que uma revendedora de veículos define que quem vender mais do que 20 produtos em um ano será categorizado como Super Vendedor. Um programa pode seguir essa regra e fazer uma simples dedução: “José vendeu 25 veículos, portanto José é um Super Vendedor”.

Depois de definido um sistema que segue a lógica, ou seja, as regras de inferência, podemos construir a prova.

4.2.6 Prova

De acordo com (ROSA, 2002): De posse das regras de inferência da camada imediatamente inferior a esta (camada de prova), os agentes podem ter mais poder para raciocinar sobre conceitos e relacioná-los na camada de ontologia. Esta é a camada na

qual pode-se obter explicações (provas) sobre as respostas dadas por agentes que consomem alguma informação com o objetivo de verificar se a dedução foi correta.

Com essa camada, podemos relacionar vários conceitos de lógica processadas pelo agentes para a construção de prova.

Podemos citar um outro exemplo de (FERNEDA, 2003) para entendermos melhor: os registros da empresa mostram que Maria vendeu 15 automóveis e 8 caminhões. O sistema define que automóveis e caminhões são produtos da empresa. As regras matemáticas dizem que $15 + 8 = 23$, que é maior que 20. Existe uma regra que diz que quem vende mais de 20 produtos é classificado como Super Vendedor. O computador junta as regras para provar que Maria é uma Super Vendedora.

4.2.7 Validação

Segundo (SEGUNDO, 2003), essa última camada da Web Semântica é responsável pelo estabelecimento de verdades, ou seja, pelo estabelecimento de autenticidade, confiabilidade e validade dos dados na Web Semântica.

Por causa dessas verdades, essa camada também é chamada de camada de confiança. De acordo com (ROSA, 2002), A camada de confiança (*Trust*) conjuntamente com a camada de assinatura digital (*digital signature*) proporciona mecanismos para prevenção de inconsistências na Web Semântica. Através de aplicações criadas neste nível, é possível criar agentes que saibam dizer, identificar e validar algum tipo de informação. Trata-se de outra característica importante da Web Semântica e muito importante no ambiente da Internet, na qual blocos de dados encriptados podem ser utilizados para garantir a autenticidade das fontes e a confiabilidade da informação que os agentes consultam.

5 Interoperabilidade Semântica

Como vimos nos capítulos anteriores, um dos maiores problemas atualmente do governo brasileiro é a falta de capacidade que os sistemas têm de se comunicar. Os motivos são vários entre eles a falta de padronização na hora da disponibilização dos dados. Todos esses conceitos estão altamente relacionados com um outro, a interoperabilidade.

Abaixo, segue uma tabela contendo conceitos de interoperabilidade de várias iniciativas pelo mundo:

Interoperabilidade refere-se à capacidade de dois ou mais sistemas computacionais quaisquer de interagir e trocar dados para obter resultados conforme o esperado.	(KAMADA, 2011).
Interoperabilidade define um conjunto de procedimentos, definição de dados, padrões técnicos e aspectos de implementação que possibilitam a cooperação administrativa através do intercâmbio de dados (com semântica definida) entre organizações e indivíduos que tem seus processos implementados ou suportados por TIC – Tecnologia da informação e Comunicação.	(ESTADO..., 2009).
Interoperabilidade provê habilidade para trocar informação entre diferentes sistemas e usar de maneira eficiente a informação que foi trocada.	(SEMANTIC..., 2008).
Interoperabilidade se refere à habilidade para combinar diferentes aspectos de diferentes organizações, com o objetivo de interagir em benefício das próprias organizações, através do intercâmbio e compartilhamento de informações e conhecimentos entre seus processos (de negócio) automatizados em sistemas, legados ou não, independente de tecnologias, de linguagens e de plataformas de <i>software</i> e <i>hardware</i> .	(ESTADO..., 2009).
Interoperabilidade é a capacidade de organizações díspares e diferentes interagirem para atingir objetivos comuns, convencionados e vantajosos para todas as partes, envolvendo o compartilhamento de informações e conhecimento entre as organizações por intermédio dos processos comerciais existentes, pela troca de dados entre seus respectivos sistemas da tecnologia da informação e comunicação (TIC).	(EUROPEAN..., 2008).

Tabela 3: Conceitos Sobre Interoperabilidade.

O conceito pode parecer simples, mas a discussão que o termo traz sobre todos os aspectos envolvidos bem como a implantação de fato da interoperabilidade, são bastante complexas.

Como descrito em (ESTADO..., 2009), as organizações enfrentam grandes desafios em termos de interoperabilidade porque os sistemas que hoje precisam interoperar

foram construídos, em épocas diferentes, como entidades independentes e monolíticas. Mas como sabemos, estamos vivendo uma época em que cada vez mais o mundo se globaliza, onde mais e mais processos de negócio vão surgindo, atravessando múltiplos domínios administrativos e organizacionais e tonando os aspectos de interoperabilidade em recursos computacionais fatores competitivos cruciais.

Dentro desse contexto podemos citar alguns obstáculos que dificultam a interoperabilidade entre os sistemas atuais:

- A maioria dos sistemas atualmente, principalmente os legados, sofrem falta de uso de padrões de dados, metadados, linguagens e infraestrutura para que esses interoperem;
- Existi uma certa dificuldade em “evangelizar” o assunto de interoperabilidade entres os diversos atores (gestores, desenvolvedores, clientes ,etc) do sistema, de modo que eles precisam usar padrões desde o início, mesmo que o sistema não interopere com ninguém em sua fase inicial.
- De acordo com (KAMADA, 2011), existi um obstáculo que refere-se à necessidade de contornar rapidamente as barreiras políticas e legais entre os atores das diversas esferas e níveis de governos e empresas e as restrições de segurança e sigilo de informações.

Em meio desses impasses, o governo federal atua em iniciativas como e-PING para reduzir esses obstáculos.

A arquitetura e-PING – Padrões de Interoperabilidade de Governo Eletrônico – define um conjunto mínimo de premissas, políticas e especificações técnicas que regulamentam a utilização da Tecnologia de Informação e Comunicação (TIC) na interoperabilidade de serviços de Governo Eletrônico, estabelecendo as condições de interação com os demais Poderes e esferas de governo e com a sociedade em geral (PADROES. . . , 2014).

Os padrões e as políticas contidas no e-PING não são impostas de forma obrigatória para sociedade. No entanto, caso queiramos interoperar com o governo federal, esses padrões e essas políticas devem ser respeitados.

Para os órgãos do governo federal – Poder Executivo brasileiro – a adoção dos padrões e políticas contidos na e-PING é obrigatória (Portaria SLTI/MP n° 5, de 14 de julho de 2005) (PADROES. . . , 2014).

Uma dimensão da interoperabilidade importante para o contexto deste trabalho é a interoperabilidade semântica. De acordo com (EUROPEAN. . . , 2008), interoperabilidade semântica refere-se à habilidade que um sistema deve ter de combinar dinâmica e auto-

maticamente suas informações com outras, recebidas de outros sistemas, e processá-las para produzir determinado valor.

Para que a interoperabilidade semântica ocorra, é necessário que a interação e a troca de dados entre sistemas estejam livres de ambiguidades, de modo que os dados recebidos por um sistema receptor sejam “entendidos” exatamente como foram enviados pelo sistema emissor.

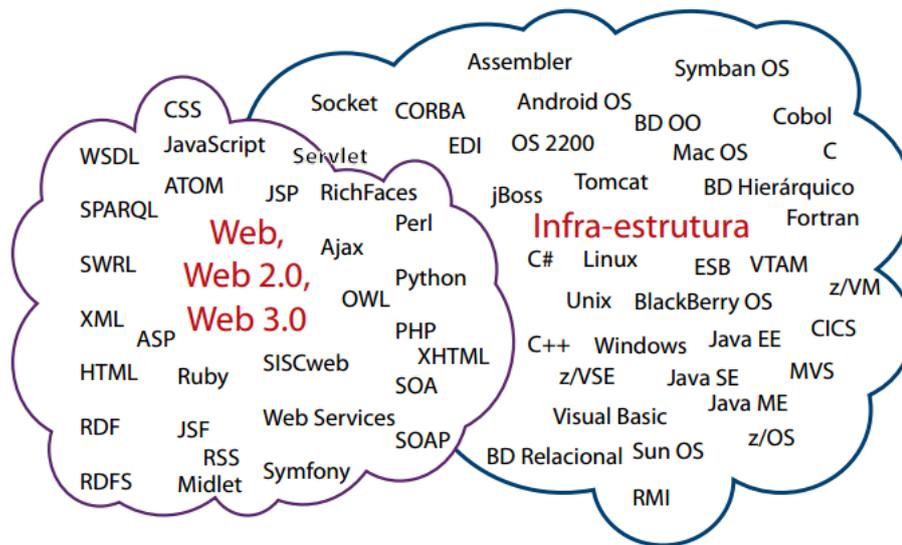
Para que isso ocorra, a troca de dados não deve se limitar apenas aos formatos e tipo de dados trocados, mas sobretudo ao conhecimento sobre os dados que devem ser compartilhados entre as partes (KAMADA, 2011).

Logo, a interoperabilidade semântica trata fundamentalmente da agregação e uso de metadados para “carregar” informações e conhecimento junto aos dados (KAMADA, 2011).

Fazendo uma correlação com os capítulos anteriores desse trabalho, percebemos que o conceito de interoperabilidade semântica tem uma forte relação com a Web Semântica e *open linked data*¹.

Para (KAMADA, 2011), toda essa conciliação semântica deve considerar a diversidade de tecnologias, linguagens, ferramentas, ambientes, plataformas operacionais e desenvolvimento de software, novos, legados e futuros, que coexistem nesse grande ecossistema, como representado na figura abaixo:

¹ Linked Data. Disponível em: <<http://linkeddata.org>>



Sistemas que precisam **interoperar** devem considerar sistemas que foram construídos em diferentes épocas, linguagens, tecnologias, plataformas, como entidades independentes e monolíticas.

Figura 11: Interoperabilidade considerando a diversidade do ecossistema. Extraído de (KAMADA, 2011)

Podemos perceber que para conseguir realizar a interoperabilidade semântica devemos resolver um problema que já foi discutido nesse trabalho, a heterogeneidade semântica, que hoje é um dos maiores impasses para integração de sistemas de informação.

Resumindo, a heterogeneidade semântica se deve ao fato de que mudanças de significado que ocorrem, seja dentro de um contexto ao longo do tempo seja por diferenças de requisitos em diferentes domínios, necessariamente resultam em diferentes modelos de informação (KAMADA, 2011).

De forma geral, temos que conciliar todas essas incompatibilidades semânticas entre os objetos e suas relações nos diferentes domínios, como por exemplo: diferenças de contexto e lógica em esquemas de banco de dados, diferenças entre nomes com mesmos conceitos, entre outros.

Logo, vimos como o conceito de interoperabilidade é importante no contexto do governo federal para promover a cooperação tanto entre os órgãos como sociedade. Podemos afirmar também a necessidade do estabelecimento de políticas de informação que viabilizem a produção e a recepção de informações. Enfim, encarando essas mudanças como uma forma de evolução, futuramente podemos ter um governo mais claro, com uma maior participação da população.

6 Ontologias

6.1 Definição de ontologias

Na seção 4.2.4 foi descrito resumidamente algumas vantagens de se utilizar ontologias para representar a semântica de dados. Ela veio para suprir algumas características que nem o RDF nem o XML contemplam, como: restrições de propriedades, disjunção de classes, combinação entre classes, características de propriedades, etc. Neste capítulo, o assunto ontologia será mais aprofundado. Abaixo, segue uma tabela com as diferentes visões sobre o significado de ontologia:

Ontologias são especificações formais e explícitas de conceitualizações compartilhadas, são modelos conceituais que capturam e explicitam o vocabulário utilizado nas aplicações semânticas, servem como base para garantir uma comunicação livre de ambiguidades. Ontologias serão a língua franca da Web Semântica.	(BREITMAN, 2005).
Uma ontologia é uma especificação explícita de uma conceitualização. [...] Em tal ontologia, definições associam nomes de entidades no universo do discurso (por exemplo, classes, relações, funções etc. Com textos que descrevem o que os nomes significam e os axiomas formais que restringem a interpretação e o uso desses termos) [...].	(GRUBER, 1992).
A ontologia é definida como uma especificação formal e explícita de uma conceitualização compartilhada, onde especificação formal quer dizer algo que é legível para os computadores, explícita são os conceitos, propriedades, relações, funções, restrições e axiomas explicitamente definidos, conceitualização representa um modelo abstrato de algum fenômeno do mundo real e compartilhada significa conhecimento consensual.	(NICO; BORST, 1997).
Ontologias são vocabulários de termos que podem ser usados em anotações semânticas e cujo o significado é formalmente especificado e pode ser acessado e processado automaticamente.	(PATEL-SCHNEIDER,).
[...] ontologia se refere a um artefato constituído por um vocabulário usado para descrever uma certa realidade, mais um conjunto de fatos explícitos e aceitos que dizem respeito ao sentido pretendido para as palavras do vocabulário. Este conjunto de fatos tem a forma da teoria da lógica de primeira ordem, onde as palavras do vocabulário aparecem como predicados unários ou binários.	(GUARINO, 1997).

Tabela 4: Conceitos sobre ontologia.

Podemos perceber que em todas as definições citadas acima a palavra vocabulário aparece exaustivamente, esse é o principal objetivo de uma ontologia. Definir um vocabu-

lário do domínio de um problema e restrições sobre as possíveis combinações de termos na sua modelagem, compartilhar conhecimento comum, reutilizar o conhecimento sobre um domínio, tornar explícitas proposições assumidas sobre um domínio, separar conhecimento sobre um domínio de conhecimento operacional e analisar o conhecimento sobre um domínio, são as razões mais comuns para o desenvolvimento de uma ontologia(DEVEDZI; VLADAN, 2002).

Na prática, uma ontologia define uma “linguagem”, conjunto de termos, que será utilizada para formular consultas (ALMEIDA, 2003). A ontologia define as regras de combinação entre os termos e seus relacionamentos, estes relacionamentos são criados por especialistas, e os usuários formulam consultas usando os conceitos especificados (MORAIS; AMBROSIO, 2007). Segundo (GUARINO, 1997), as ontologias podem se dividir nas seguintes categorias:

- ontologias genéricas: descrevem conceitos bastante gerais, tais como, espaço, tempo, matéria, objeto, evento, ação, etc., que são independentes de um problema ou domínio particular;
- ontologias de domínio: expressam conceituações de domínios particulares, descrevendo o vocabulário relacionado a um domínio genérico, tal como Medicina e 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, diagnose ou vendas;
- ontologias de aplicação: descrevem conceitos dependentes do domínio e das tarefas particulares. Estes conceitos frequentemente correspondem a papéis desempenhados por entidades do domínio quando da realização de uma certa atividade;
- ontologias de representação: explicam as conceituações que fundamentam os formalismos de representação de conhecimento.

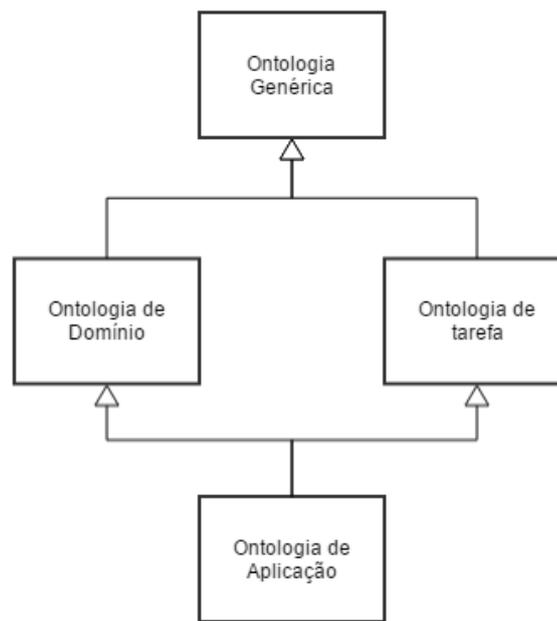


Figura 12: Categorias de ontologias. Extraído de (GUARINO, 1998)

De acordo com a figura 12, *Guarino* propõe que as ontologias devem ser construídas segundo seus níveis de generalidade, ou seja, os conceitos das ontologias de domínio ou tarefa devem ser especializações dos termos de uma ontologia genérica e os conceitos de uma ontologia de aplicação devem ser especializações dos termos de uma ontologia de domínio e/ou tarefa. O objetivo deste trabalho é desenvolver uma ontologia de domínio, pois será proposto para um contexto mais particular, dentro de um domínio mais genérico. Logo, todos os passos propostos nos próximos capítulos são para a construção de uma ontologia de domínio.

6.2 Metodologias de desenvolvimento

6.2.1 Ontology Development 101

Esse método funciona como um “guia para criação da sua primeira ontologia”, apresenta uma abordagem iterativa para desenvolvimento de ontologias, vale ressaltar que não há uma maneira “correta” ou metodologia para o desenvolvimento de ontologias (NOY, 2000), sempre vão existir várias alternativas. A melhor solução depende da aplicação.



Figura 13: Método 101 para desenvolvimento de ontologias. Extraído de (BREITMAN, 2005)

Na figura 13, temos os sete passos proposto pelo método proposto. Mais abaixo vamos descreve-los resumidamente.

1. Passo 1: Determinar o domínio e o escopo da ontologias.

Para definir o domínio e escopo, as ontologias devem responder a algumas questões básicas (BREITMAN, 2005):

- Que domínio se deseja cobrir com a ontologia?
- Com que propósito(s) será utilizada a ontologia?
- Para que informações a ontologia deve fornecer respostas?
- Quem vai utilizar e manter a ontologia?

Essas questões nada mais são do que suas *competency questions*.

2. Passo 2: Considerar o reuso de outras ontologias.

É recomendável sempre verificar antes se existe alguma ontologia que codifique os termos requeridos ou se é possível refinar algum modelo existente para o nosso domínio de aplicação.

É importante realizar esse passo porque a ontologia que estamos propondo poderá interoperar com outras ontologias existentes.

3. Passo 3: Enumerar termos importantes da ontologia.

É muito importante ter uma lista contendo os principais termos do domínio que queremos definir ou explicar para os usuários. Nesse momento ainda não há preocupação em detalhar os seus relacionamentos.

4. Passo 4: Definir classes e a hierarquia de classes.

Esse passo e o seguinte ocorrem quase sempre de forma paralela. O que acontece de fato é que ao definir uma classe, nós já definimos suas propriedades.

Posteriormente, devemos definir uma hierarquia para essas classes, utilizando algumas estratégias como: *top-Down*, *bottom up* ou combinação.

5. Passo 5: Definir as propriedades das classes. (*slots*)

Uma vez definida as classes, é necessário descrever as estruturas internas de seus conceitos. Os termos da lista definido no passo 3 que não viraram uma classe são grandes candidatos para se tornarem propriedades de alguma classe.

De modo geral, existem vários tipos de propriedades relativas a classes: intrínsecas, extrínsecas, partes, estrutura, relacionamentos com outras classes, relacionamentos com outros itens, entre outras.

6. Passo 6: Definir valores das propriedades.

As classes (“slots”) podem ter diferentes tipos de valores (“Facetas”) como strings, números, booleanos, etc. Além desses vários valores, os atributos podem conter cardinalidades, por exemplo uma classe Pessoa pode ter as propriedades nome e telefone, o nome é considerado uma cardinalidade simples, pois uma pessoa só tem um nome e já o telefone é classificado como cardinalidade múltipla, pois uma pessoa pode ter vários telefones.

7. Passo 7: Criar instâncias.

Esse último passo se resume apenas em criar instâncias individuais das classes criadas. Utilizando o exemplo do passo anterior, criar uma instância para uma classe Pessoa seria o mesmo que: nome-> Fulano de Tal; telefone-> 9999-9999.

Essa metodologia foi escolhida por se tratar de ser um guia, algo muito semelhante ao roteiro didático que será proposto por este trabalho, esse guia dá uma visão clara de todo o processo iterativo para o desenvolvimento de uma ontologia. Boa parte de seus passos vão servir como subsídio para o roteiro.

6.2.2 On-to-knowledge

On-to-Knowledge(OTK) é uma metodologia de desenvolvimento de ontologias fruto da cooperação de várias entidades europeias (??), que tinha como intuito desenvolver ontologias para serem empregadas em sistemas de gestão do conhecimento. Essa metodologia abrange aspectos desde as fases iniciais de um projeto de gestão do conhecimento (GC) até o final da implantação da aplicação de GC baseada em ontologia. A OTK

tem um ciclo de vida espiral e é dividida em cinco fases. De acordo com (STAAB et al., 2000) e (SURE; STAAB; STUDER, 2003), essas fases são descritas da seguinte forma:

- *Estudo de Viabilidade: nesta fase é verificada a viabilidade econômica e técnica do projeto, inicialmente identificando as áreas problema e de oportunidade e potenciais soluções e, posteriormente, colocando-as dentro de uma perspectiva organizacional mais ampla. Esta fase baseia-se no estudo de viabilidade descrito na metodologia CommonKADS.*
- *Início da Ontologia: o produto de saída da fase kickoff (começar) é um documento de especificação de requisitos descrevendo qual ontologia deveria suportar e esboçar a área planejada da aplicação de ontologia, que também deveria guiar o engenheiro de ontologia a decidir sobre inclusão, exclusão e a estrutura hierárquica de conceitos na ontologia. Nesta fase também deveriam ser consideradas as ontologias já desenvolvidas e potencialmente reutilizáveis. Em resumo, ela deveria descrever claramente as seguintes informações: objetivo da ontologia, domínio e escopo, aplicações suportadas pela ontologia, fontes de conhecimento (por exemplo: especialistas de domínio, gráficos da organização, planos de negócio, dicionários, listas de índice, esquemas de banco de dados, etc.), usuários potenciais e cenários de uso, questionário de competência (isto é, uma visão geral de possíveis questões a serem feitas ao sistema, indicando o escopo e conteúdo da ontologia de domínio) e ontologias potencialmente reutilizáveis.*
- *Refinamento: o objetivo desta fase é produzir uma ontologia alvo madura e orientada à aplicação, de acordo com a especificação dada pela fase kickoff. Nesta fase, a ontologia é formalizada e são adicionados relações e axiomas.*
- *Avaliação: esta fase serve como uma prova para a utilidade das ontologias desenvolvidas e seu ambiente associado. No primeiro passo, o engenheiro de ontologia verifica se a ontologia alvo está de acordo com o documento de especificação de requisitos de ontologia e se a mesma suporta ou “responde” às questões de competência analisadas na fase kickoff. No segundo passo, a ontologia é testada no ambiente de aplicação alvo.*
- *Manutenção: para refletir as mudanças que ocorrem na vida real, as ontologias têm que ser mantidas frequentemente, como ocorre com outras partes do software.*

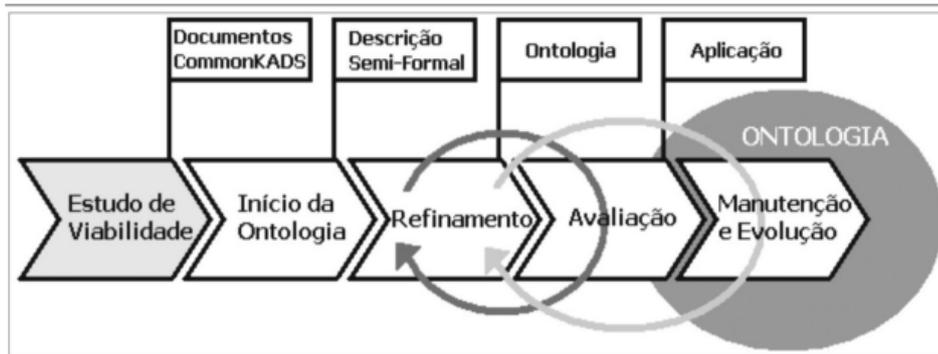


Figura 14: Fases da metodologia On-to-Knowledge. Extraído de (RAUTENBERG; TODESCO, 2008)

Um das características que fez a OTK ser uma das escolhidas como base foi sua preocupação com as fases iniciais da ontologia, principalmente com a definição do escopo (questões de competência) e o estudo de viabilidade. Como um dos objetivos deste trabalho é desenvolver um modelo conceitual, produto de uma fase antes da construção da ontologia, a OTK se demonstrou muito pertinente.

6.2.3 Methontology

O *Methontology* é um *framework* desenvolvido no laboratório de inteligência Artificial do Politécnico de Madri que fornece apoio automatizado para a construção de ontologias (BREITMAN, 2005). Essa metodologia é fortemente influenciada por metodologias de Engenharia de Software e Engenharia de Conhecimento. O *Methontology* é baseado na construção de ontologia a partir do conhecimento de um domínio (MORAIS; AMBROSIO, 2007). Segundo seus autores, Assunción Gómez-Pérez, Mariano Fernández e Natalia Juristo, essa metodologia se divide em três grupos contendo as seguintes atividades (BREITMAN, 2005):

- Atividades de gerenciamento de ontologias – elaboração de cronogramas, controle, garantia da qualidade.
- Atividades ligadas ao desenvolvimento de ontologias – estudo do ambiente, estudo de viabilidade, especificação, conceitualização, formalização, implementação, manutenção e uso.
- Atividades de Suporte – aquisição do conhecimento, avaliação, integração, documentação, gerência da configuração, alinhamento.

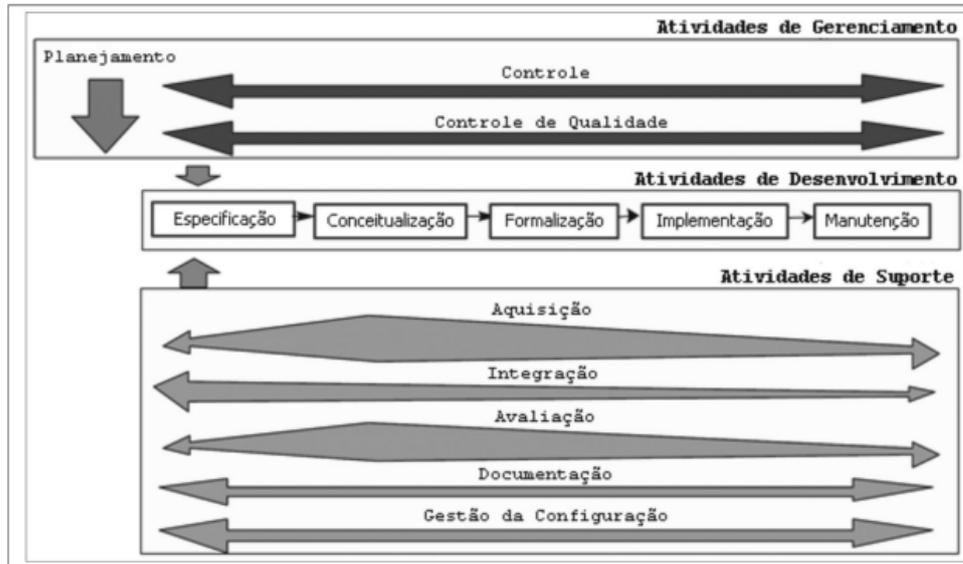


Figura 15: Processo de desenvolvimento e ciclo de vida da *methontology*. Extraído de (RAUTENBERG; TODESCO, 2008)

Esta metodologia prevê um ciclo de vida baseado na prototipagem de ontologias, de acordo com a evolução do processo de desenvolvimento (especificação, conceitualização, formalização, implementação e manutenção) (RAUTENBERG; TODESCO, 2008). Na figura 15 logo acima, temos a representação dos grupos com suas respectivas atividades aqui já citadas.

A característica principal da *methontology* é a definição de um rico conjunto de artefatos, dos quais muitos ajudaram a inspirar o roteiro didático proposto por este trabalho. Existem várias outras metodologias além dessas citadas acima, em (ALMEIDA, 2003), pode ser encontrado mais informações sobre outras metodologias.

7 Roteiro Didático para construção de ontologias

Como foi dito, atualmente não existe uma metodologia suficientemente madura que sirva para desenvolver qualquer tipo de ontologia. Compreendido isso, o roteiro proposto uniu as melhores práticas de algumas metodologias, sendo as *Methontology*, *On-to-Knowledge* e *Ontology Development 101*.

A *Ontology Development 101* contribuiu com uma visão simplificada como se dá todo o processo de construção de ontologias, influenciando bastante na fase de iniciação e aquisição de conhecimento.

A *On-to-Knowledge* também contribui para a construção das fases iniciação e aquisição de conhecimento, mas com um foco um pouco diferente. A OTK auxiliou a determinar tarefas de especificação de requisitos da ontologia, como questões de competência, definição de escopo, determinação do propósito da ontologia, entre outras.

E por fim temos a *Methontology*, as fases de implementação (modelo conceitual) e validação foram as mais influenciadas por essa metodologia. A geração de alguns artefatos também são heranças dessa metodologia.

Nas próximas seções, será dissertado sobre todas as fases do roteiro proposto bem como suas tarefas.

7.1 Iniciação

Essa fase é bem explorada nas três metodologias tidas como base para este roteiro, apresenta tarefas que ajudam a responder questionamentos como: por que esta ontologia está sendo construída? Quais serão seus usuários? Será que já existem ontologias que representam o domínio em questão? A iniciação é formada por cinco tarefas. São elas:

7.1.1 Elaborar cronograma

Essa tarefa é importante para manter o controle sobre o projeto, no cronograma deve ser identificado todas as fases com suas respectivas tarefas. O cronograma pode ser simples, precisa apresentar apenas os principais marcos. O nível de formalismo desse cronograma é escolhido pelo próprio usuário.

7.1.2 Definir propósito da ontologia

Essa de fato é a primeira tarefa relacionada a construção da ontologia, seu principal objetivo é determinar com que propósito a ontologia será utilizada, quais as suas competências, ou seja, seus usos e propósitos. O intuito é descobrir por que a ontologia deve ser construída.

7.1.3 Determinar domínio e escopo

De acordo com (NOY, 2000), existem alguns questionamentos que se respondidos podem ajudar a determinar o escopo e o domínio da ontologia:

- Que domínio se deseja cobrir com a ontologia?
- Para que informações a ontologia deve fornecer repostas?
- Quem vai utilizar e manter a ontologia?

Além de ajudar a determinar o escopo e o domínio da ontologia, estas questões vão dar origem a vários requisitos da ontologia. Na engenharia de requisitos, existem várias técnicas que podem ajudar o desenvolvedor a levantar os requisitos necessários para o seu projeto, a escolha da técnica vai depender de cada situação. Abaixo segue um quadro com as vantagens e desvantagens de algumas dessas técnicas:

Técnicas	Entrevista	Leitura Dinâmica de documentos	Questionários	Workshop de requisitos	Observação	Análise de protocolo	Enfoque antropológico	Bases de requisitos não funcionais
Vantagens	Contato direto com atores, possibilidade de validação imediata.	Facilidade de acesso às fontes de informação; volume de informação.	Padronização de perguntas; tratamento estatístico.	Múltiplas opiniões, criação coletiva.	Baixo custo, pouca complexidade da tarefa.	Fatos não observáveis; melhor compreensão dos fatos.	Visão de dentro para fora contextualizada	Reutilização de conhecimento; antecipação de aspectos implementacionais; identificação de conflitos.
Desvantagens	Conhecimento tácito, diferenças culturais.	Dispersão das informações; volume de trabalho requerido para identificação dos fatos.	Limitação das respostas; pouca interação; participação.	Dispersão, custo.	Dependência do ator (observador); superficialidade e decorrente da pouca exposição ao universo de informações.	Foco na performance; o que se diz não é o que se faz.	Tempo; pouca sistematização.	Custo de construção de base; RNF falsa impressão de completude

Figura 16: Técnicas de elicitação de requisitos. Adptado de (BREITMAN, 2005)

7.1.4 Levantar recursos de conhecimento

O propósito dessa fase é levantar todas as fontes de conhecimento que podem ajudar na identificação dos conceitos contidos no domínio em questão. Essas fontes podem ser ontológicas ou não-ontológicas.

Os recursos ontológicos podem ser qualquer ontologia ou vocabulário existente que seja relacionado com a área de negócio e que esteja em plena utilização.

Os recursos não-ontológicos podem ser de vários tipos, podem ser modelos como modelos de Entidade-Relacionamento, modelos de banco de dados, diagrama de classes, modelos conceituais, enfim, tudo que possa conter conceitos sobre o domínio. Outros documentos que podem ser utilizados são artefatos do projeto como especificações de requisitos, manuais de sistemas, documentos de visão, negócio, livros, artigos, leis, etc.

7.1.5 Mapear ontologias existentes

Sempre vale a pena verificar se alguém já codificou os termos em uma ontologia ou se é possível refinar um modelo existente para o nosso domínio de aplicação (BREITMAN, 2005). A ideia é pesquisar ontologias correlacionadas dos quais pode-se aproveitar conceitos já estabelecidos.

7.2 Aquisição de conhecimento

Essa fase contém tarefas de conceitualização e formalização da ontologia, é a fase que o engenheiro de conhecimento tem mais contato com os especialistas do domínio. É nela que vão ser abstraídos a maioria dos elementos da ontologia a ser construída, como classes, propriedades, instâncias, principais termos, questões de competência, entre outros.

7.2.1 Definir questões de competência

Questões de Competência (QC) são questões que vão ajudar a definir o escopo da ontologia, o intuito é desenvolver questionamentos que futuramente a ontologia deverá respondê-las, ou seja, além de ajudar a definir alguns requisitos da ontologia, as QCs vão funcionar como um teste de validação no final do projeto: Será que a ontologia contém informações suficientes para responder as questões de competência?

7.2.2 Enumerar termos importantes

É nesta fase que se inicia o levantamento dos principais conceitos da ontologia, a partir de uma análise das fontes de conhecimento e das questões de competência, é necessário abstrair os principais termos utilizados pelos especialistas do domínio. Neste primeiro momento, não é necessário se preocupar com redundâncias ou em detalhar muito os termos, pois isso já será feito nos próximos passos.

Ao definir os conceitos, deve-se preocupar com as ambiguidades e inconsistências, ou seja, deve existir um consenso entre o engenheiro de conhecimento e os especialistas do domínio, pois uns dos objetivos principais da ontologia é garantir que todos tenham um mesmo entendimento sobre um universo de informação.

7.2.3 Selecionar termos reutilizáveis

Após mapear os termos das fontes de conhecimento, o próximo passo é selecionar, caso exista, termos de outras ontologias ou vocabulários que podem ser aproveitados. O intuito é não “reinventar a roda”, ou seja, se já existe algo pronto e a comunidade já tem conhecimento e aceita, não é necessário criar um conceito para explicar algo que já exista.

7.2.4 Classificar os termos

Depois de listar todos os termos, é o momento de classificá-los. Essa classificação consiste em determinar se o termo é uma classe, propriedade, instância ou um atributo.

Uma classe representa um conjunto ou coleção de indivíduos(instâncias) - objetos, pessoas, coisas - que compartilham de um grupo de características que os distinguem dos demais (BREITMAN, 2005). Classe é a representação concreta de um conceito do domínio, geralmente é organizada em taxonomias.

Propriedades representam relações entre instância ou entre instâncias e atributos.

Instâncias são objetos do domínio, são utilizadas para representar elementos da própria ontologia. É a concretização de uma classe.

Os atributos são valores que caracterizam uma instância.

7.2.5 Definir hierarquia de Classes

Com a lista de termos já devidamente classificada, o próximo passo é definir a hierarquia entre as classes já definidas. Como descrito por Karim Breitman em (BREITMAN, 2005), existem várias estratégias para se definir uma hierarquia, entre elas estão a *top-down*, *bottom up* e combinação.

A estratégia *top-down*(topo-para-baixo) é a mais comum de todas, se baseia em definir primeiramente os termos mais genéricos (pai ou superclasse) e posteriormente os termos mais específicos (filhos ou subclasses) abaixo dessas superclasses.

A estratégia *bottom up*(Baixo-para-cima) é exatamente o contrário. Primeiramente é definido as classes mais específicas para depois se identificar possíveis grupamentos (superclasses).

E por fim a combinação, essa técnica utiliza um misto das duas estratégias acima citadas. Geralmente é a mais utilizada por usuários inexperientes, pois é comum que os conceitos levantados no mundo real sejam os intermediários, ou seja, nem muito geral nem muito específico.

O objetivo dessa tarefa é identificar possíveis casos de herança e generalização.

7.2.6 Relacionar propriedades às classes

As classes isoladamente não conseguem responder as questões de competência que foram definidas a uns passos atrás. A lista de termos também definida neste roteiro, contém alguns conceitos que representam as propriedades das classes e o objetivo desta tarefa é determinar a que classe pertencem estas propriedades.

Para cada classe, vão existir “relações” que vão evidenciar de forma explícita o seu relacionamento com as demais, geralmente o termos vistos como verbos são grandes candidatos para se tornarem relações. Outro ponto de deve ser definido são suas respectivas cardinalidades, que cada propriedade apresenta a sua, fica a cargo do engenheiro de conhecimento defini-las.

7.2.7 Relacionar atributos às classes

Como já foi dito, atributos são valores que caracterizam instâncias, a intenção desta tarefa é determinar a que classes esses termos pertencem. Se imaginarmos que pessoa é uma classe, e que Pedro é uma instância dessa classe, então data de nascimento, idade, sexo, estado civil, CPF e RG são possíveis exemplos de atributos.

Deve ser definido também os tipos desse atributos, ou seja, se são do tipo literal, número, booleano, entre outros.

7.2.8 Definir modelo conceitual

Esta é a principal tarefa da fase de aquisição de conhecimento. Como descrito por (BRASCHER; CAFE, 2008), a representação do conhecimento se constitui numa estrutura conceitual que representa modelos de mundo, capazes de descrever e fornecer explicações sobre os fenômenos observados; ela é fruto de um processo de análise de domínio e procura refletir uma visão consensual sobre a realidade que se pretende representar.

O modelo conceitual é o principal produto antes de partimos para a construção da ontologia propriamente dita. Este modelo é mais formal que a lista de termos e menos formal se comparado a ontologia.

A forma de representação das classes e relacionamentos fica a cargo do engenheiro de conhecimento. Muitos autores utilizam UML para tal representação, que vale ressaltar que na modelagem conceitual são representados apenas os itens mais relevantes, ou seja, atributos não são representados nessa fase, apenas classes e relacionamentos, ajudando a dar foco no que é essencial para o entendimento do mundo real.

Logo, podemos concluir que a modelagem conceitual é uma representação visual das classes e relacionamentos definidos nas tarefas anteriores.

7.3 Implementação

A fase de implementação se resume em transformar o modelo conceitual em uma ontologia passível de processamento automático, ou seja, computável. Para que isso aconteça, a mesma deve ser implementada em alguma linguagem, como por exemplo: OWL, OIL, DAML+OIL.

7.3.1 Mapear modelo conceitual para ontologia

Essa tarefa consiste em representar o modelo conceitual em uma linguagem ontológica. Dentre as linguagens citadas acima, a OWL é a mais recomendada, pois é a indicada pelo W3C, e esta linguagem foi baseada nas outras duas, OIL e DAML+OIL.

A OWL é uma linguagem para definir e instanciar ontologias na *Web*. Ela foi projetada de modo a atender as necessidades das aplicações para Web Semântica (BREITMAN, 2005).

A ideia é utilizar o *protegè* como ferramenta de construção, como já foi descrito neste trabalho, ela é uma ferramenta muito poderosa para edição de ontologias e é a mais utilizada pela comunidade atualmente.

7.4 Validação

Avaliar uma ontologia significa verificar se a mesma satisfaz os requisitos definidos em sua construção (MORAIS; AMBROSIO, 2007). Neste caso, a ontologia será avaliada em três aspectos: conformidade perante as fontes de conhecimento, aos usuários e as fontes de competência.

7.4.1 Validar ontologia perante suas fontes de conhecimento

Essa tarefa consiste em verificar a coerência do conhecimento representado pela ontologia com o que é expresso nas fontes de conhecimento. A necessidade de realizar essa averiguação, é garantir que os conceitos estejam bem definidos, concisos, onde não haja oportunidades para interpretações ambíguas.

7.4.2 Validar ontologia perante o usuário

Essa tarefa, como o próprio nome já diz, consiste reunir os especialistas do domínio e os usuários para verificar a usabilidade e utilidade da ontologia. O *feedback* dessas personalidades são muito importantes para a refinamentos sucessivos da ontologia, essa validação pode ser feito através de reuniões, onde o especialista de conhecimento apresenta a ontologia e posteriormente os usuários apresentam seus questionamentos.

7.4.3 Validar ontologia perante as questões de competência

E por fim, validar a ontologia perante as questões de competência (QCs) definidas na fase de aquisição de conhecimento e esta tarefa consiste basicamente em localizar na ontologia fatos que respondam as questões definidas. Só assim garantimos que a ontologia está pronta ou não para ser utilizada em larga escala. Caso ela não esteja num nível satisfatório de qualidade, ela deverá passar por mais refinamentos até que a mesma responda um número aceitável de QCs.

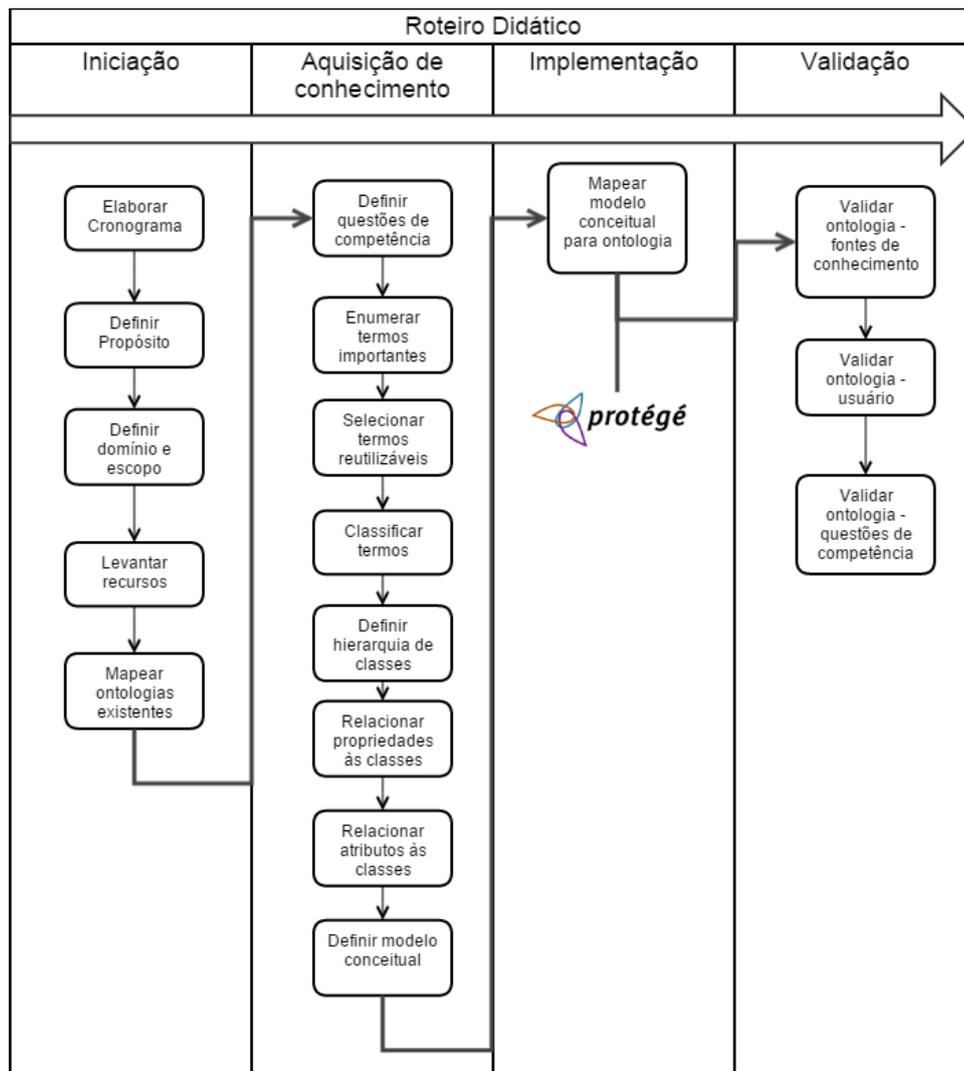


Figura 17: Representação do Roteiro proposto

8 Estudo de caso - SIORG

O SIORG – Sistema de Organização e Inovação Institucional do Governo Federal, normatizado pelo decreto 6.944, de 2009, é um sistema estruturante das atividades de desenvolvimento organizacional dos órgãos e entidades da administração direta, autárquica e fundacional do Poder Executivo Federal, apoiado na construção de espaços de articulação, intercâmbio e construção de consensos entre dirigentes e técnicos das unidades dos órgãos e entidades com atuação em temas de gestão ([MINISTERIO DO PLANEJAMENTO, ORCAMENTO E GESTAO, 2014](#)).

Com a criação da lei de acesso a informação em 2011, vários modelos da administração pública do governo federal estão em constantes mudanças para se adequarem as novas diretrizes. Este sistema foi criado com o intuito de controlar a estrutura organizacional do poder executivo federal, funcionando como um cadastro único de unidades organizacionais.

De acordo com ([PLANEJAMENTO, 2014](#)), o SIORG apresenta os seguintes objetivos:

- Modernizar os Sistemas de Informações Organizacionais da APF (Administração Pública Federal);
- Maior transparência da estrutura organizacional da APF para a sociedade;
- Agilizar e tornar mais confiáveis as informações, proporcionando ao administrar melhor acompanhamento das políticas e diretrizes definidas pelo governo federal, contribuindo para o processo de Reforma do Estado;
- Recuperar em tempo real o histórico das estruturas organizacionais do Estado.

Atualmente, o cadastro do Siorg conta com 250 órgãos/entidades e mais de 69.000 unidades administrativas cadastrados, localização de órgãos em mais de 1.400 cidades, descrição detalhada das finalidades para mais de 1.400 órgãos e competências de mais de 23.500 - unidades organizacionais ([MINISTERIO DO PLANEJAMENTO, ORCAMENTO E GESTAO, 2014](#)).

Vale ressaltar que todas as atividades de desenvolvimento organizacional dos órgãos e entidades que constituem o SIORG são fundamentadas em leis, ou seja, qualquer modificação na estrutura organizacional tem que está descrito numa lei ou decreto. As principais são: constituição federal de 1988; Decreto-Lei 200, de 1967; Lei 10.683, de 2003; Lei 12.527, de 2011; Decreto 6.944 de 2009 e Instruções Normativas 3 e 5, de 2010.

Atualmente o sistema encontra-se no ar¹ e já está sendo utilizado por alguns órgãos do governo. Dentre suas principais funcionalidades estão:

- Consultar Estruturas
- Cadastrar/Extinguir Órgão/Entidade
 - Consultar propostas
 - Cadastrar nova proposta
 - Consultar pendências
- Cadastrar informações Complementares
- Comparar Estruturas

Unidade Organizacional

Selecione a Unidade Organizacional

Hierarquia:

- República Federativa do Brasil
 - União
 - Presidência da República**
 - Advocacia-Geral da União
 - Casa Civil da Presidência da R
 - Controladoria-Geral da União
 - Gabinete de Segurança Instituc
 - Ministério da Agricultura, Pecuá
 - Ministério da Ciência, Technolog
 - Ministério da Cultura
 - Ministério da Defesa
 - Ministério da Educação
 - Ministério da Fazenda
 - Ministério da Integração Nacion
 - Ministério da Justiça

Legenda

Hierarquia: ● Órgão ◻ Unidade Administrativa
● Ente ◻ Entidade ◻ Unidade Colegiada

Normalização: ➔ Lei/Decreto ➔ Ato Interno

Unidade Organizacional | Informações Complementares

Informações Básicas

Código: 26 Esfera: Federal
 Denominação: Presidência da República Poder: Executivo
 Sigla: PR Nat. Jurídica: Administração Direta
 Categoria: Estrutura Básica Subnatureza Jurídica:
 Tipo de Alteração: Lei / Decreto

Competência/Finalidade

Competência:
 À Presidência da República,
 I - Compete privativamente ao Presidente da República: Nomear e exonerar os Ministros de Estado;
 II - Exercer, com o auxílio dos Ministros de Estado, a direção superior da Administração Federal;
 III - Sancionar, promulgar e fazer publicar as leis, bem como expedir decretos e regulamentos para sua fiel execução;
 IV - Vetar projetos de lei, total ou parcialmente;
 V - Dispor sobre a organização e o funcionamento da Administração Federal, na forma da lei;
 VI - Manter relações com Estados estrangeiros e acreditar seus representantes diplomáticos;
 VII - Celebrar tratados, convenções e atos internacionais, sujeitos a referendo do Congresso Nacional
 VIII - Decretar o Estado de defesa e o Estado de sítio; Decretar e executar

Figura 18: Tela principal do SIORG. Extraído de (PLANEJAMENTO, 2014).

¹ O SIORG está disponível em: <https://siorg.planejamento.gov.br/siorg-cidadao-webapp/apresentacao.jsf>

9 Construção da Ontologia

Considerando que vimos toda parte teórica e metodológica que envolve o desenvolvimento de uma ontologia, neste capítulo conduziremos um estudo para a modelagem conceitual do SIORG bem como a construção de uma versão preliminar da ontologia utilizando o *Protégé*¹. Os mapas conceituais produzidos neste capítulo utilizará o software *Cmaps Tools*², ele é uma ferramenta própria para elaborar esquemas conceituais e representá-los graficamente.

9.1 Definir propósito da ontologia

O propósito dessa ontologia é fornecer uma base para que o ministério do planejamento possa futuramente publicar dados na Web utilizando o que é considerado, no cenário internacional, o estado da arte dos padrões da Web Semântica, definidos pela W3C. A ontologia pretende também atender outros propósitos mais específicos como:

- Normalizar o vocabulário de estruturas organizacionais e facilitar a comunicação e o entendimento entre os gestores e usuários dos sistemas de informações
- Construir uma base semântica de conhecimento que possa, após sua formalização, atender a portais semânticos (ex: interface web navegável) e conectar-se ao *Linked Open Data*
- Publicação e reúso dos dados
- Permitir possíveis integração com outras ontologias

9.2 Determinar domínio e escopo

A ontologia deve cobrir o domínio de estruturas organizacionais do governo brasileiro, mais precisamente do poder executivo federal. Atualmente, a principal fonte detentora desse conhecimento é o SIORG, que será a principal fonte de conhecimento estudada para esta ontologia.

O SIORG além de representar estruturas organizacionais, apresenta também um módulo sobre Cargos e Funções. Foi definido que esta ontologia abordará somente o módulo de estrutura organizacional, ou seja, qualquer conceito sobre cargo ou função não fará parte desse domínio.

¹ O Protégé é um editor de ontologias. Pode ser acessado em: <<http://protege.stanford.edu/>>

² <http://cmap.ihmc.us/>

Em relação aos possíveis usuários dessa ontologia, temos como principal beneficiado o governo federal brasileiro. O público externo como empresas e cidadãos também terão acesso a esse projeto, eles poderão contribuir tanto com a manutenção da ontologia como o desenvolvimento de aplicativos para a sociedade.

9.3 Levantar recursos de conhecimento

Os recursos de conhecimento foram divididos em dois grupos: recursos ontológicos e recursos não ontológicos. Para a obtenção destes recursos, foi realizado uma reunião com os especialistas domínio onde os mesmos forneceram toda sua base de conhecimento sobre o SIORG. As outras fontes ficaram a cargo do especialista de conhecimento.

Abaixo segue uma lista com as fontes não-ontológicas utilizadas neste projeto:

- Manual do SIORG - Cadastro de Estruturas Organizacionais
- Documentação do novo Web Service SIORG
- Vídeos de treinamento do sistema SIORG.
- Manual de orientação para arranjo institucional de órgãos e entidades do poder executivo federal
- Constituição federal de 1988
- Decreto-Lei 200, de 1967
- Lei 10.683, de 2003
- Lei 12.527, de 2011
- Decreto 6.944, de 2009
- Livros sobre direito administrativo
- Modelo de Entidade-Relacionamento do SIORG

Em relação ao grupo de recursos ontológicos, não foi encontrado muitas fontes dentro desse domínio, pois existem muitas diferenças entre as estruturas organizacionais de cada país. Utilizamos apenas três ontologias como referência:

- *Federal Enterprise Architecture Reference Model Ontology*³ – FEA-RMO
- *Central Government Ontology*⁴
- *An organization ontology*

³ <<http://notes.3kbo.com/fea-rmo>>

⁴ <<http://lov.okfn.org/dataset/lov/vocabs/cgov>>

9.4 Mapear ontologias existentes

Como já foi citado acima, foram identificadas apenas três ontologias sobre o domínio de estruturas organizacionais governamentais, a FEA-RMO, a *Central Government Ontology* e a *An organization ontology*.

Como descrito em (BATISTA; MIRANDA; ARAUJO, 2011), a FEA-RMO foi desenvolvida pela empresa *Top Quadrant* para o governo dos EUA, em 2005. O desenvolvimento da FEA-RMO teve foco principal na Web Semântica e nos recursos de expressividade que a linguagem de ontologias da web (OWL) e seu modelo de inferências permite representar. O modelo também partiu de uma arquitetura corporativa já documentada, FEA, para então representá-la num modelo ontológico formal. Essa arquitetura se divide em um núcleo, cinco perspectivas (modelos de referência) sobre a estrutura organizacional e modelos de “ponte” que fazem os mapeamentos entre estas.

O *Central Government Ontology* é um modelo de estrutura organizacional do governo do Reino Unido que se deu durante o desenvolvimento do projeto de Linked Data para o portal ⁵, que também desenvolveu ontologias para outras áreas de governo. Além disso, referenciam conceitos de ontologias externas que são utilizadas pelas comunidades de prática (e.g., FOAF, Dublin Core, VOID, SKOS, ontologia de organizações do W3C, etc.) (BATISTA; MIRANDA; ARAUJO, 2011).

A *An organization ontology*⁶ é uma ontologia que foi desenvolvida pela empresa *Epimorphics*, especializada em desenvolver soluções para *Linked Data*⁷. A *An organization ontology* foi criada por uma necessidade de publicar informações relativas à estrutura organizacional de um governo.

Eles acharam que a melhor abordagem fosse desenvolver uma ontologia pequena, genérica e reutilizável, que continha apenas as informações fundamentais sobre estrutura organizacional, possibilitando que os desenvolvedores pudessem ampliá-la e especializá-la para domínios mais específicos.

9.5 Definir Questões de competência

O levantamento das questões de competência (Qcs) foi realizado através de uma reunião entre os especialistas do domínio (gestores do SIORG), a equipe da SLTI (Secretaria de Logística e Tecnologia da Informação) e os especialistas do conhecimento.

Primeiramente foi realizado uma rápida apresentação sobre o tema, pois alguns integrantes não tinham o conhecimento sobre o assunto. Após o nivelamento do conheci-

⁵ <data.gov.uk>

⁶ <<http://www.epimorphics.com/public/vocabulary/org.html>>

⁷ Linked Data é uma abordagem poderosa para compartilhamento de informações em toda Web.

mento, partimos para o real motivo da reunião. A discussão se iniciou com o especialista do conhecimento, onde o mesmo apresentou um conjunto de questões para serem validadas pelas outras equipes. No final da reunião, foi gerada uma lista com todas as questões, conforme abaixo:

- Qual é a estrutura de uma organização?
- Como uma organização é decomposta em unidades administrativas?
- Qual dispositivo legal oficializa uma unidade? e um departamento? uma divisão? Coordenação?
- Como os órgãos se relacionam entre si?
- Qual o número máximo de níveis de uma organização?
- Quais tipos de unidade posso criar no meu departamento?
- Qual a sigla de uma unidade organizacional?
- Quais organizações públicas estão representadas em um órgão colegiado?
- Qual a missão, objetivos e competências de uma unidade organizacional?
- Em que esfera/poder da Administração Pública o órgão encontra-se inserido?
- Quando foi criado(a) um órgão ou unidade administrativa?
- Quando foi extinto(a) um órgão ou unidade administrativa?
- Qual é o número de inscrição do órgão no CNPJ?
- Qual é o endereço/telefone/website/e-mail do órgão ou unidade administrativa?
- O imóvel que o órgão ou unidade administrativa ocupa está adaptado para acesso a pessoas com deficiência?
- Quais são os dias e horários de funcionamento do órgão ou unidade administrativa?

9.6 Enumerar termos importantes

Todos os termos identificados nessa fase foram retirados das questões de competência, dos recursos de conhecimento e das reuniões realizadas com as equipes do SIORG e SLTI. A ideia é detectar aqueles termos que mais se repetem dentro dos documentos analisados e aqueles que são mais mencionados nas reuniões.

É muito importante nessa fase observar as redundâncias e ambiguidades. Termos que forem parecidos, ou até mesmo diferentes, mas que se referem a mesma coisa têm que serem unidos em um só.

Abaixo segue a lista geral de termos com suas definições:

1. **Autarquia:** O serviço autônomo, criado por lei, com personalidade jurídica, patrimônio e receita próprios, para executar atividades típicas da Administração Pública, que requeiram, para seu melhor funcionamento, gestão administrativa e financeira descentralizada.
2. **Categoria:** Categorias que caracterizam uma unidade organizacional. Exemplos: Administração Geral; Forças Armadas; Repartição no Exterior, etc...
3. **Código:** Código sequencial único que identifica uma unidade organizacional.
4. **Competência:** Competência da unidade organizacional.
5. **Denominação:** Nome da unidade organizacional.
6. **Empresa Pública:** A entidade dotada de personalidade jurídica de direito privado, com patrimônio próprio e capital exclusivo da União, criado por lei para a exploração de atividade econômica que o Governo seja levado a exercer por força de contingência ou de conveniência administrativa podendo revestir-se de qualquer das formas admitidas em direito.
7. **Entidade :** É uma unidade da Administração pública Indireta do Poder Executivo Federal. Tem personalidade jurídica e patrimônio próprios, com autonomia administrativa e financeira. criado por lei de iniciativa privativa do Chefe do Poder Executivo e com regulamento aprovado por Decreto, nos casos das entidades de direito público, ou ato previsto no Código Civil, quando de direito privado; criado para exercício de competência pública executiva, descentralizada, mantendo vínculo com o órgão da administração direta responsável pela direção superior de sua área de atuação, para fins de coordenação e supervisão ministerial, porém sem subordinação. As entidades se subdividem em Autarquias, Fundações Públicas, Empresas Públicas e Sociedades de Economia Mista
8. **Esfera:** É o domínio sobre o qual uma unidade administrativa exerce poder. O Poder Executivo é dividido em três esferas: Federal, Estadual/Distrital e Municipal.
9. **Finalidade:** Finalidade da unidade organizacional.
10. **Missão:** Missão da unidade organizacional.

11. Natureza jurídica: É uma classificação da constituição jurídico-institucional das entidades públicas nos cadastros da administração pública do país. Exemplo: administração direta.
12. Organização pública: Unidade orgânica da Administração Pública, provida de estrutura hierarquizada, com a prerrogativa de propor alterações na sua própria estrutura organizacional.
13. Órgão: É uma Unidade Organizacional da Administração Direta do Poder Executivo Federal, chefiada por Ministro de Estado ou autoridade equivalente. Não tem personalidade jurídica e vontade própria. É um centro de competência governamental ou administrativa, instituído para o desempenho de funções estatais, cuja atuação é imputada à pessoa jurídica a que pertence. Criado por lei para auxiliar o Presidente da República no exercício do Poder Executivo, competindo-lhe, privativamente, exercer as funções de direção superior da administração pública. São órgãos os Ministérios e demais unidades da Presidência da República a quem a lei estender a condição de Ministério e que se enquadrarem em todos os seguintes aspectos conceituais mencionados acima. A lista completa de órgãos pode ser encontrada na Lei 10.683, de 2003.
14. Poder: Refere-se aos três poderes da União: Executivo, Legislativo e Judiciário.
15. Sigla: Sigla da unidade Organizacional.
16. Subnatureza jurídica: É um subconjunto de natureza jurídica. Exemplo: Autarquia Especial é uma subnatureza jurídica de Autarquia.
17. Sociedade de economia mista: a entidade dotada de personalidade jurídica de direito privado, criada por lei para a exploração de atividade econômica, sob a forma de sociedade anônima, cujas ações com direito a voto pertençam em sua maioria à União ou a entidade da Administração Indireta.
18. Tipo unidade organizacional: Campo que define o tipo da unidade organizacional. Pode ser: Ente, Órgão, Entidade, Unidade Administrativa, ou Unidade Colegiada.
19. Unidade Administrativa: É uma Unidade Organizacional que compõem a estrutura do Órgão ou Entidade. Possui um conjunto de competências desdobradas das competências do Órgão ou Entidade da qual fazem parte ao qual se subordinam diretamente; normalmente, a unidade administrativa não detém autonomia administrativa e financeira, mas há casos em que, por força das competências que exercem, a lei ou o regulamento lhes concedem autonomias específicas necessárias ao adequado cumprimento de suas competências. São Unidades Administrativas, as Secretarias, Diretorias e outras subdivisões dos Ministérios, órgãos da Presidência da República, das autarquias e fundações.

20. Unidade Colegiada: São unidades organizacionais que tem composição pluripessoal, constituído por representantes de órgãos ou entidades do Poder Público, do setor privado ou da sociedade civil, segundo a natureza da representação. São Órgãos Colegiados: a COFIEIX (Decreto n 3.502, de 2000), a CAMEX (Decreto 4.732, de 2003), o CADE (Decreto 7.738, de 2012), dentre outros.
21. Unidade Organizacional: Qualquer item da Estrutura Organizacional. No âmbito do Siorg, as unidades organizacionais são especializadas em: Órgãos, Entidades, Unidades Administrativas e Unidades Colegiadas.
22. Composição
23. Vinculação
24. Subordinação
25. Supervisão

9.7 Selecionar termos reutilizáveis

Já vimos que atualmente não existe nenhuma ontologia na área de estrutura organizacional que abrange todo o domínio aqui estudado, mas existem alguns termos que podem ser reaproveitados na ontologia proposta por este trabalho.

Segue abaixo os termos principais que serviram de inspiração para esta ontologia:

<i>Organization</i>	<i>Represents a collection of people organized together into a community or other social, commercial or political structure. The group has some common purpose or reason for existence which goes beyond the set of people belonging to it and can act as an Agent. Organizations are often decomposable into hierarchical structures. It is recommended that SKOS lexical labels should be used to label the Organization. In particular skos:prefLabel for the primary (e.g. legally recognized name), skos:altLabel for alternative names (trading names, colloquial names) and skos:notation to denote codes from a code list. Alternative names: Collective Body Org Group</i>	<i>An organization ontology</i>
<i>Organizational Unit</i>	<i>An Organization such University Support Unit which is part of some larger FormalOrganization and only has full recognition within the context of that FormalOrganization, it is not a Legal Entity in its own right. Units can be large and complex containing other Units and even FormalOrganizations. Alternative names: OU Unit Department</i>	<i>An organization ontology</i>
<i>Organization</i>	<i>The Organization class represents a kind of Agent corresponding to social institutions such as companies, societies etc.”</i>	FOAF

Tabela 5: Conceitos sobre ontologia.

Podemos observar que além de serem poucos termos, são os mais genéricos. Devido a estrutura organizacional dos demais países serem diferentes, aproveitamos apenas uma pequena parte dessas ontologias já construídas.

9.8 Classificar Termos

Após definimos a lista dos termos mais importantes, é hora de classificá-los, essa tarefa consiste apenas em definir quem é classe, propriedade, atributo ou instância. É importante observar que podem existir alguns termos nessa fase que não estão descritos na lista de termos, pois muitos deles estão implícitos nos recursos de conhecimento.

Abaixo segue um quadro com os termos já classificados:

Classe	Propriedade	Atributo
Autarquia, Empresa Pública, Entidade, Organização pública, Órgão, Sociedade de economia mista, Unidade administrativa, Unidade Colegiada, Unidade organizacional, Divisão, Coordenação, Secretaria, Subsecretaria, Coordenação-Geral, Serviço, Departamento, Seção.	Composição, Vinculação, Subordinação, Supervisão.	Categoria, Código, Competência, Denominação, Esfera, Finalidade, Missão, Natureza jurídica, Poder, Sigla, Subnatureza, Tipo unidade organizacional.

Tabela 6: Classificação dos Termos da ontologia.

Podemos observar que nesse momento não foram listados as instâncias. Atualmente, o SIORG conta com mais de 250 órgãos/entidades e 69.000 unidades administrativas, tornando-se inviável sua representação nesta fase. Posteriormente, serão construídos alguns exemplos com essas instâncias.

9.9 Definir Hierarquia de Classes

Como mencionado neste trabalho, existem três estratégias que podem ser utilizadas para definir hierarquia de classes. Foi optado pela combinação, que basicamente é um misto das outras duas estratégias: *top-down* e *bottom up*.

As hierarquias foram definidas da seguinte forma: primeiramente foram identificadas classes que tinham algumas características em comum. O segundo passo foi descobrir se existia alguma classe que poderia funcionar como uma classe Pai (superclasse), gerando assim uma generalização e se por acaso essa classe não exista, verificar a necessidade de criar uma superclasse ou não. E, posteriormente, repetir esses passos para as outras classes. No final, foram gerados os seguintes grupos de classes:



Figura 19: Hierarquia de uma unidade organizacional.

A hierarquia na figura 19 foi definida de forma simples, pois estava bem explícita nos recursos de conhecimento, mais precisamente no manual do SIORG. Segue abaixo um trecho do documento que valida a hierarquia acima:

[...] No âmbito do Siorg, as unidades organizacionais são especializadas em: Órgãos, Entidades, Unidades Administrativas e Unidades Colegiadas.



Figura 20: Hierarquia de uma Organização Pública.

Essa hierarquia 20 já apresenta uma particularidade que não vimos na anterior. O conceito organização pública não existe nos recursos de conhecimento, pois ela foi criada devido a uma necessidade descoberta durante a execução dessa tarefa.

Foi observado que as classes órgão e entidade apresentam grandes similaridades como suas competências, finalidades, estruturas, formas de provimento e extinção, atributos, entre outros. Em consequência disso, foi criado a superclasse organização pública, responsável por conter as características em comum dos órgãos e entidades.

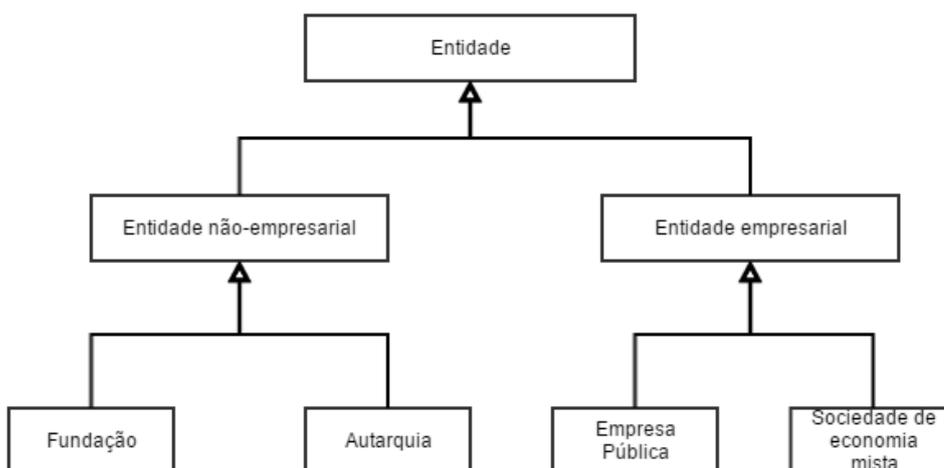


Figura 21: Hierarquia de uma Entidade.

A estrutura da figura 21 seguiu a mesma linha da figura anterior, ou seja, também foram utilizados conceitos que de certa forma não estavam explícitos nos recursos de

conhecimento. Como vimos na seção responsável por enumerar os termos mais importantes, as entidades se subdividem em Autarquias, Fundações Públicas, Empresas Públicas e Sociedades de economia mista.

Estudando mais a fundo os recursos de conhecimento e com ajuda dos especialistas do SIORG, notou-se que os pares Autarquias – Fundações e Empresas Públicas – Sociedades de economia mista tinham algumas características em comum. Foi criado então as classes Entidade não-empresarial e Entidade Empresarial. A primeira possui natureza meramente administrativa e a segunda pode prestar serviço público que permita a exploração no mundo empresarial ou pode exercer atividade econômica de interesse coletivo.

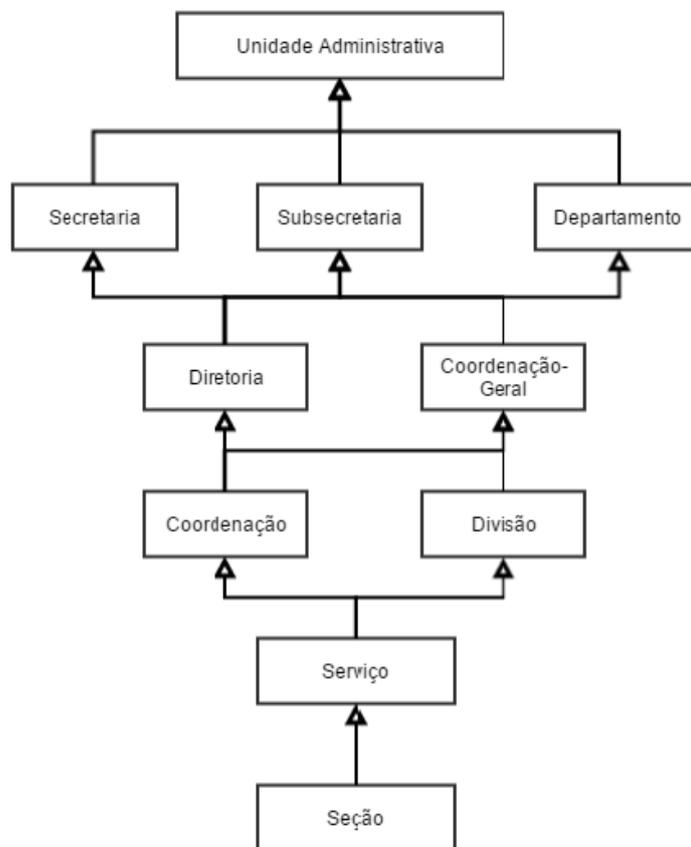


Figura 22: Hierarquia de uma Unidade Administrativa.

E por fim temos a hierarquia da unidade administrativa, como já vimos, são as unidades administrativas que vão compor os órgãos e entidades. Ao analisar os recursos de conhecimento, percebemos que essa unidade se dividia em vários níveis, como foi demonstrado pela figura 22, criando assim uma hierarquia entre eles. Vale ressaltar que não existe obrigatoriedade de um órgão ou entidade possuir todos os níveis.

9.10 Relacionar propriedades às classes

Propriedade é uma terminologia utilizada no protegê para representar as relações, no primeiro momento foram identificadas quatro relações principais entre as classes. Nessa tarefa, foi optado por não descrever as relações com os atributos. Vale ressaltar que todas essas relações estão explícitas de forma clara nos recursos de conhecimento, mais precisamente no decreto-lei 20 de 1967⁸.

Abaixo segue um quadro com algumas características dessas relações:

Relacionamento	Definição			
	Cardinalidade	Domínio (Domain)	Escopo (Range)	Relacionamento Inverso
Composto de	Como o próprio nome já diz, serve para demonstrar a composição de alguns elementos da estrutura organizacional.			
	1,N	Unidade organizacional, Organização Pública, Unidade administrativa	Órgão, Entidade, Unidade Colegiada, Unidade administrativa, fundação, autarquia, empresa pública, sociedade de economia mista	Compõe
Subordinado a	Demonstra uma relação de subordinação entre órgãos, entre entidades ou entre órgão/entidade e unidade colegiada.			
	1,N	Órgão, Entidade, Unidade colegiada	Órgão, Entidade	É subordinado de
Vinculado a	Demonstra a relação de vinculação entre os membros da administração direta (órgão, por exemplo) e indireta (entidades, por exemplo).			
	N,1	Entidade	Órgão	Tem vinculação
Supervisiona	Todo órgão/entidade da administração pública direta ou indireta está sujeito a supervisão de um ministério (órgão)			
	1,N	Órgão	Órgão, Entidade	É supervisionado por

Figura 23: Características principais das Relações entre classes.

A relação de composição é a mais simplória de todas elas, apenas serve para organizar a estrutura organizacional, ou seja, indicar como os órgãos, entidades, fundações, entre outros, se estruturam, como eles são formados de fato.

As relações de subordinação e vinculação podem ser confundidas em algumas

⁸ <http://www.planalto.gov.br/ccivil_03/decreto-lei/del0200.htm>

situações. A diferença básica entre elas é que a subordinação é um fenômeno de caráter interno, dentro da mesma administração (direta ou indireta), e a vinculação é acontece quando há uma relação entre um elemento da administração direta e outro da administração indireta.

E por fim temos a relação de supervisão. De acordo com a lei 200, de 1967, todo órgão ou entidade pode estar submetido a supervisão ministerial, ou seja, a relação é de 1 para N, um ministério pode supervisionar vários órgãos ou entidades mas uma entidade ou órgão só pode ter um ministério como supervisor.

9.11 Relacionar atributos às classes

Nessa fase, foram levantados todos os atributos básicos que compõe uma unidade organizacional, ou seja, todos os atributos que aqui serão citados compõem os órgãos, entidades, unidades colegiadas e unidades administrativas. Vale ressaltar que todos esses atributos foram extraídos do manual do sistema SIORG (??).

Abaixo será descrito a definição de cada atributo levantado:

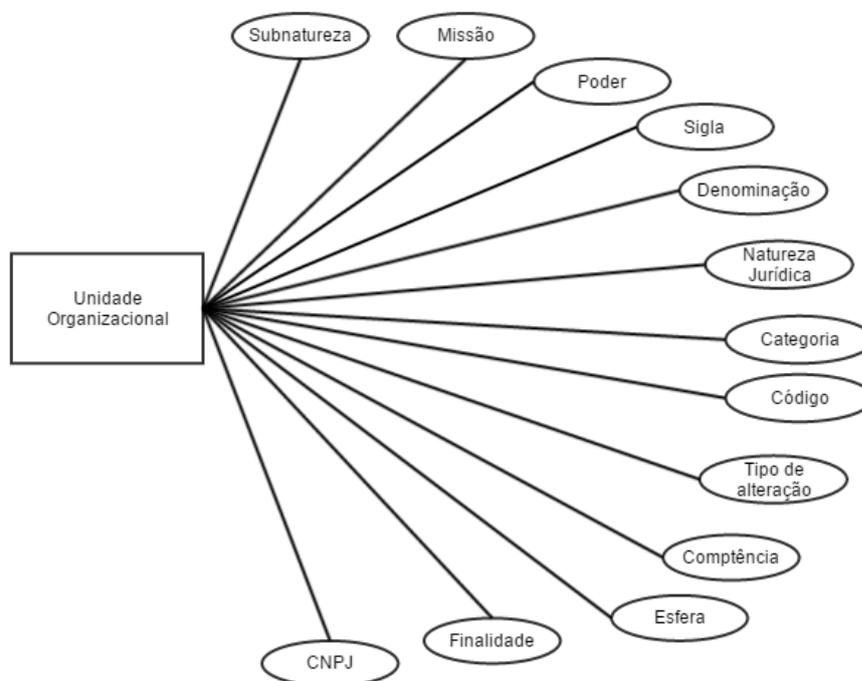


Figura 24: Relação dos atributos básicos que compõem uma unidade organizacional.

- CNPJ: CNPJ da unidade.

- Categoria: Categorias que caracterizam uma unidade organizacional. Exemplos: Administração Geral; Forças Armadas; Repartição no Exterior.
- Código: Código sequencial único que identifica uma unidade organizacional.
- Competência: Competência da unidade organizacional.
- Denominação: nome completo e oficial da unidade.
- Esfera: É o domínio sobre o qual uma unidade administrativa exerce poder. O Poder Executivo é dividido em três esferas: Federal, Estadual/Distrital e Municipal.
- Finalidade: Finalidade da unidade organizacional.
- Missão: Missão da unidade organizacional.
- Natureza Jurídica: É uma classificação da constituição jurídico-institucional das entidades públicas nos cadastros da administração pública do País. Exemplo: Administração Direta.
- Poder: Refere-se aos três poderes da União: Executivo, Legislativo e Judiciário.
- Sigla: sigla da unidade.
- Subnatureza Jurídica: É um subconjunto de natureza jurídica. Exemplo: Autarquia Especial é uma subnatureza jurídica de Autarquia.
- Tipo alteração: Informar se a unidade foi criada por meio de Ato/Decreto ou ato interno

9.12 Definir o modelo conceitual

Partes do modelo conceitual já foram desenvolvidas de forma isolada nas tarefas anteriores. Nessa fase vamos realizar a união de todas as classes com seus respectivos relacionamentos para termos uma visão geral de como funciona a estrutura organizacional do poder executivo federal. Lembramos que não será exibido os atributos das classes para não dificultar o entendimento do modelo.

Segue abaixo a representação da modelagem conceitual. Vale enfatizar que esse modelo ainda pode sofrer alterações.

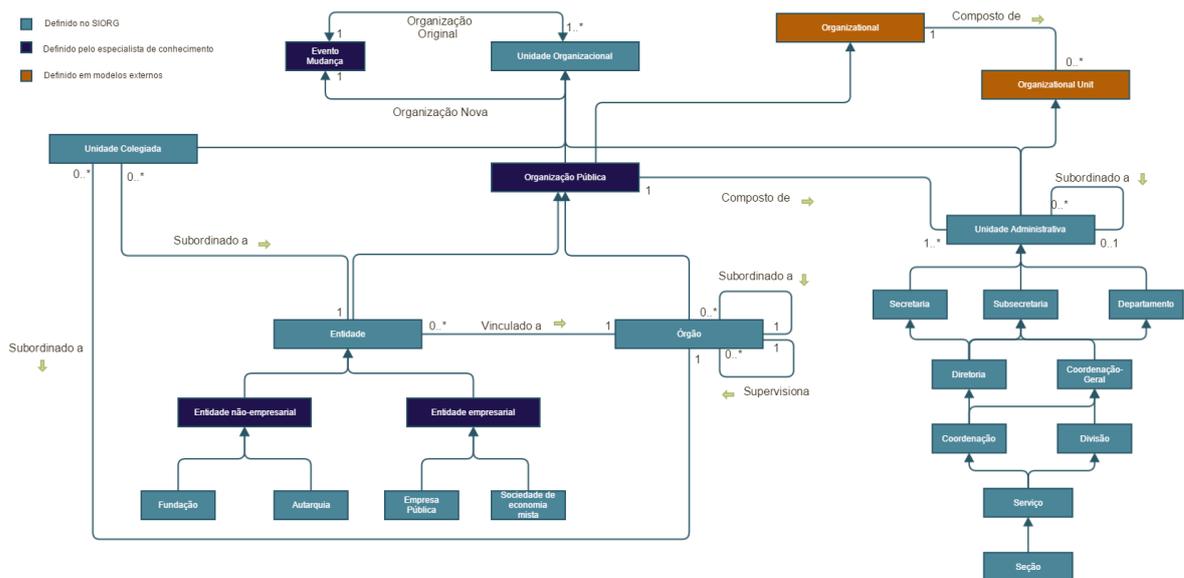


Figura 25: Modelagem Conceitual.

Podemos perceber que existem três tipos de elementos contidos no modelo: definidos pelo SIORG, definidos pelo engenheiro de conhecimento e definido em modelos externos. O primeiro, garante que todos os elementos ali representados foram retirados das fontes de conhecimento do SIORG.

O segundo tipo define elementos que foram criados pelo especialista de conhecimento em conjunto com a equipe de especialistas do SIORG. Na tarefa referente a criação de hierarquia de classes, já foi explicado o motivo para a criação dessas classes adicionais.

E por fim o terceiro tipo, os elementos ali simbolizados são heranças de modelos conceituais externos. Esses elementos que já foram citados na tarefa selecionar os termos reutilizáveis, funcionando aqui no modelo final como pontos de ligação com ontologias externas.

9.13 Mapear modelo conceitual para ontologia

Essa tarefa exige muito cuidado por parte do engenheiro de ontologia, pois não basta apenas transcrever o modelo conceitual para a ontologia, existem algumas inferências que são necessárias para que a ontologia seja funcional. Como sugerido nesse roteiro didático, o *protegè* será a ferramenta utilizada para a construção da ontologia.

Vale ressaltar que a ontologia aqui definida é uma versão preliminar. ela não conterá todos atributos necessários nem suas restrições, pois a ontologia completa é um objetivo para trabalhos futuros. Nessa ontologia, vamos focar nas principais classes e em algumas inferências necessárias para que ontologia possa representar o modelo conceitual da forma

mais fiel possível.



Figura 26: Árvore taxonômica da ontologia de estrutura organizacional.

A árvore taxonômica acima já é um produto do *protegè*, como podemos observar, as classes mais à esquerda são as classes mais gerais enquanto no lado direito temos as classes mais específicas. As setas indicam relação de subclasse com uma classe. Por exemplo, órgão e entidade são subclasses de organização pública.

Observando essa taxonomia, notamos que existem duas classes que não estão na modelagem conceitual, a Estrutura Organizacional e Ato Decreto. Essas classes são resultados de algumas inferências feitas pelo engenheiro de ontologia.

A classe Estrutura Organizacional foi criada para auxiliar na criação de várias estruturas organizacionais, pois atualmente o SIORG contempla apenas elementos do poder executivo federal (presidência da república). Futuramente, o SIORG poderá se estender para os poderes legislativo e judiciário, justificando então existência da classe, outro motivo dessa classe existir é que foi a única forma encontrada para a representação de uma estrutura organizacional, ou seja, precisávamos de uma classe que abrangesse um conjunto de unidades organizacionais.

A classe Ato Decreto deixou de ser um atributo da unidade organizacional para se tornar uma classe, percebemos que ela continha informações suficientes para ser carac-

terizada como tal, informações como uma gama de atributos e subclasses. Com isso, fica mais fácil de localizar quais órgãos e entidades foram gerados por um ato ou decreto “X”.

Abaixo segue uma foto da tela do *protegè*. Nela, podemos observar como que as classes (*Class*), relações (*Object Property*) e atributos *Data Property* foram construídos e nomeados. Ressaltamos que os nomes de algumas classes e relações sofreram pequenas modificações para se adaptarem a realidade do *protegè*.

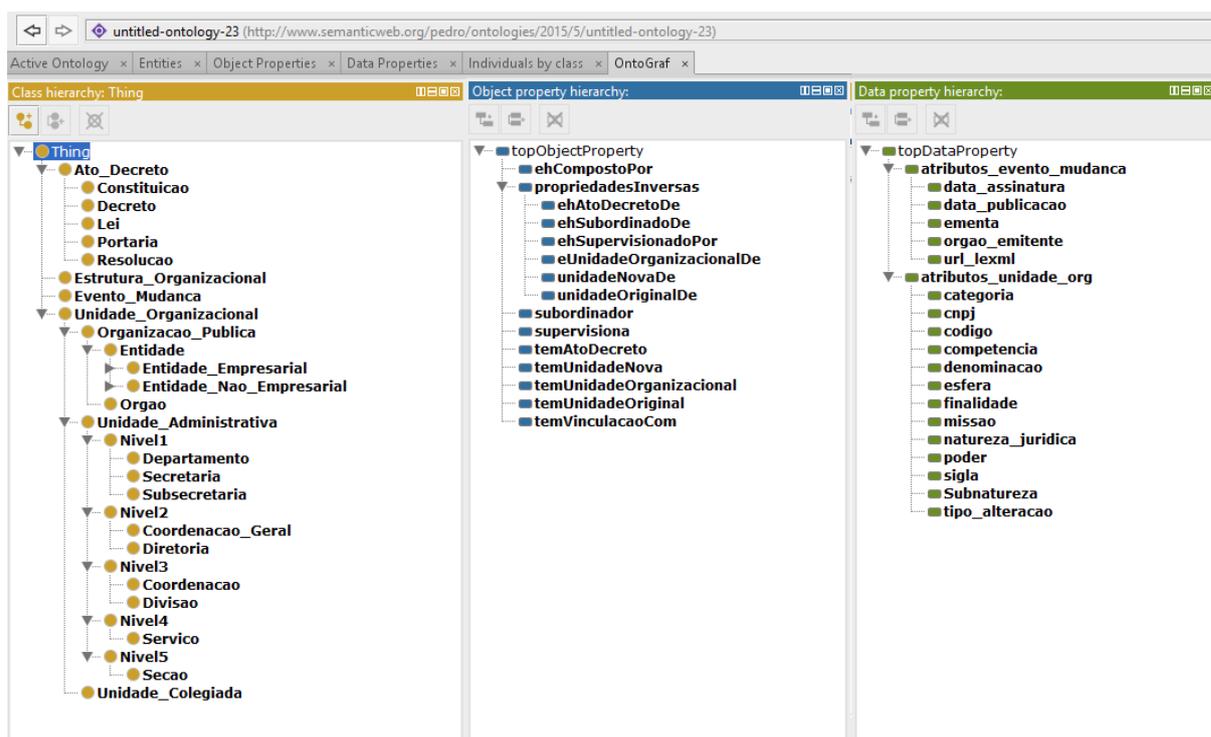


Figura 27: Interface do editor de ontologia *protegè*.

Para exemplificar como seria a criação de uma estrutura organizacional no *protegè*, foi decidido utilizar um escopo bem reduzido, pois seria inviável criar instâncias para todos os elementos que compõem a presidência da república. Segue abaixo o modelo gerado pela ferramenta.

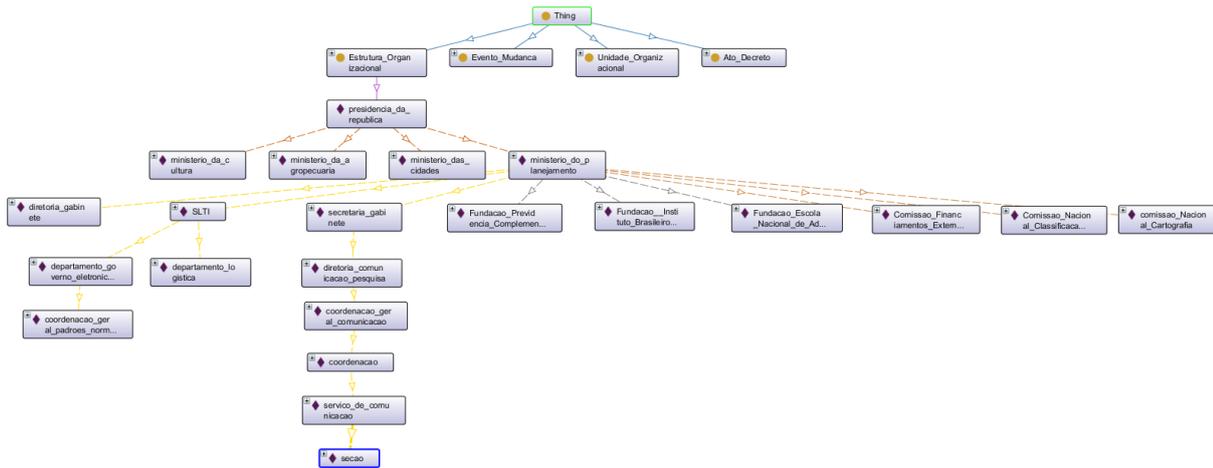


Figura 28: Estrutura Organizacional reduzida da presidência da república.

Podemos observar que apenas as classes mais gerais foram apresentadas, pois se fossemos representar todas as classes o digrama ficaria muito poluído, dificultando seu entendimento. Foi decidido então demonstrar apenas as instâncias das outras classes.

A estrutura organizacional é representado pela instância Presidência da república. Através da relação *temUnidadeOrganizacional* (seta laranja), a estrutura é formada por milhares de unidades organizacionais, no caso apenas quatro ministérios.

Para exemplificarmos as outras relações, detalhamos ainda mais a estrutura do ministério do planejamento (MP). Vale lembrar que esse ministério é composto por muito mais unidades organizacionais, mas seria inviável representá-los nesse momento, nesse exemplo, as comissões são unidades colegiadas, as fundações são entidades, e o restante são unidades administrativas.

Logo, as setas laranjas que conecta o MP às comissões representam a relação subordinador, ou seja, as comissões são subordinadas ao órgão MP. As setas cinzas que liga o MP às fundações representam a relação *temVinculacaoCom*. E por fim, as setas amarelas, que são responsáveis por representar a relação *ehCompostoPor*.

Analisando somente as unidades administrativas, podemos perceber claramente seus níveis representados, quanto mais abaixo, maior o seu nível. No caso, Seção foi unidade administrativa com nível mais alto representado na estrutura, nível 5.

10 Conclusão e Considerações finais

Podemos concluir que o objetivo final deste trabalho é algo inédito para o governo federal brasileiro. O pontapé inicial para abertura de dados do governo só foi dado no final de 2011, com a criação da lei de acesso a informação, ou seja, é um assunto muito novo no âmbito da administração pública.

Outro ponto que podemos perceber durante todo o trabalho é a complexidade do tema. Conceitos como Web Semântica, ontologias, interoperabilidade semântica, entre outros, são muito novos e pouco explorados, dificultando ainda mais a concepção do objetivo final, que é gerar uma ontologia para estrutura organizacional do poder executivo federal.

Durante todo o estudo, foram levantados recursos de conhecimento na área de estrutura organizacional afim de nos ajudar a alcançar os objetivos definidos neste trabalho. Nesse sentido, podemos considerar que os capítulos de dois a cinco nos auxiliaram a entender melhor o contexto em que a ontologia iria ser construída. O capítulo seis, sobre ontologia, foi essencial para o desenvolvimento do roteiro didático, produto do objetivo um.

O roteiro didático desenvolvido nesse trabalho funciona como um guia prático, ou seja, ele determina fases e tarefas que o desenvolvedor da ontologia deve seguir para chegar a ontologia final. Vale ressaltar que o roteiro supracitado foi resultado da união de três metodologias: a *On-to-Knowledge*, *METHONTOLOGY* e do guia *Ontology Development 101*. A sua estrutura final pode ser vista na figura 17.

Um dos principais produtos desse roteiro, é a modelagem conceitual da área de conhecimento em que ele é aplicado. Essa modelagem é exatamente o marco do objetivo dois, logo, o capítulo sete foi o meio e o fim para o cumprimento desse objetivo.

Vimos na prática, que o roteiro funciona. É claro que ele deve passar por muitos outros testes, mas para este trabalho ele funcionou perfeitamente. Aplicando o roteiro no estudo de caso descrito do capítulo oito, o SIORG, conseguimos realizar os objetivos dois e três, ou seja, conseguimos definir o modelo conceitual do SIORG e também conseguimos iniciar a construção da ontologia. O modelo conceitual final pode ser visto na figura 25 e a ontologia nas figuras 28, 26 e 27.

Um ponto a ser observado no roteiro é a fase da validação. Como a proposta foi apenas iniciar uma ontologia, ainda existem diversos itens a serem revistos como atributos, relações, restrições, cardinalidades, entre outros. Sem esses itens, ficamos impossibilitado de realizar uma validação da ontologia preliminar construída.

Como trabalho futuro, existem pontos no roteiro e na ontologia a serem evoluídos. No roteiro, desenvolver a fase de implementação. Hoje, temos apenas uma tarefa geral: mapear modelo conceitual para ontologia. Ao desenvolver a ontologia, percebemos que existem diversas tarefas que poderiam ser descritas no roteiro.

Em relação a ontologia, identificar os atributos restantes bem como suas relações. Outro fator muito importante, são as restrições. Todas elas devem ser levantadas, documentadas e implementadas. A ontologia só poderá ser publicada após aplicarmos as validações definidas no roteiro.

Por fim, espero que este trabalho possa contribuir de alguma forma com o governo no que diz respeito a publicação de dados abertos. Espero que, com a ontologia aqui proposta, o governo possa agilizar e tornar mais confiáveis as informações, proporcionando ao administrar melhor acompanhamento das políticas e diretrizes definidas pelo governo federal, sem esquecer da maior transparência para a sociedade.

Referências

- ALMEIDA, M. B. Uma visao geral sobre ontologias: pesquisa sobre definicoes, tipos, aplicacoes, metodos de avaliacao e de construcao. *Ci. Inf., Brasilia*, v. 32, n. 3, dec 2003. Citado 2 vezes nas páginas 48 e 54.
- ALVES, R. C. V. *Web Semantica: uma analise focada no uso de metadados*. Tese (Doutorado) — Universidade Estadual Paulista, Faculdade de Filosofia e Ciencias, 2005. Citado 2 vezes nas páginas 36 e 38.
- BATISTA, A. H.; MIRANDA, C. M. C.; ARAUJO, D. B. de. 2011. Citado na página 67.
- BRASCHER, M.; CAFE, L. Organizacao da informacao ou organizacao do conhecimento. *IX ENANCIB. Diversidade cultural e politicas de informacao.*, 2008. Citado na página 59.
- BREITMAN, K. K. *Web Semantica: a Internet do futuro*. [S.l.]: LTC editora, 2005. Citado 10 vezes nas páginas 33, 34, 35, 47, 50, 53, 56, 57, 58 e 60.
- DADOS Abertos governo federal. 2014. Disponível em: <<http://dados.gov.br/>>. Citado na página 26.
- DEVEDZI; VLADAN. Understanding ontological engineering. *Commun. ACM*, v. 45, n. 4, apr 2002. Citado na página 48.
- EAVES, D. The three laws of open government data. *Eaves.ca*, sep 2009. Disponível em: <<http://eaves.ca/2009/09/30/three-law-of-open-government-data/>>. Citado na página 30.
- EBERTZ, S. J. D. Introducao a web semantica. *DEV MEDIA*, 2013. Disponível em: <<http://www.devmedia.com.br/introducao-a-web-semantica/26181>>. Citado na página 33.
- ESTADO da Arte em Interoperabilidade Semantica. 2009. Disponível em: <http://www.softwarepublico.gov.br/5cqualibr/2-documentos-tecnicos/view/vetor-interoperabilidade/D2.3_Interoperabilidade_semantica.pdf>. Citado na página 43.
- EUROPEAN Interoperability Framework. 2008. Disponível em: <<http://ec.europa.eu/idabc/servlets/Doc?id=31597>>. Citado 2 vezes nas páginas 43 e 44.
- EXTENSIBLE Markup Language. 2009. Disponível em: <<http://www.w3.org/XML/>>. Citado na página 36.
- FERNEDA, E. *Recuperacao de Informacao: Analise sobre a contribuicao da Ciencia da Computacao para a Ciencia da Informacao*. Tese (Doutorado) — Universidade de Sao Paulo, Escola de Comunicacao e Artes, 2003. Citado 3 vezes nas páginas 35, 41 e 42.
- FILHO, F. W.; LOSCIO, B. F. *Web Semantica: Conceitos e Tecnologias*. 2009. Disponível em: <<http://www.ufpi.br/subsiteFiles/ercemapi/arquivos/files/minicurso/mc9.pdf>>. Citado 5 vezes nas páginas 36, 37, 38, 39 e 40.

GRUBER, T. What is an ontology? *International Journal of Human-Computer Studies*, 1992. Citado na página 47.

GUARINO, N. Understanding, building and using ontologies. *Int. J. Hum. Comput. Stud.*, v. 46, mar 1997. Citado 2 vezes nas páginas 47 e 48.

GUARINO, N. Formal ontology and information systems. *FOIS 98*, jun 1998. Citado na página 49.

JORGE, E.; REIS, U. *Web Semantica: O Futuro das Aplicacoes*. 2013. Disponível em: <<http://www.devmedia.com.br/revista-java-magazine-85/18489>>. Citado na página 35.

KAMADA, A. Interoperabilidade semantica. os desafios e obstaculos na implementacao de padroes que propiciam interoperabilidade em sistemas computacionais. *Linux Magazine Especial*, p. 26–27, jun 2011. Disponível em: <<http://repositorio.cti.gov.br/repositorio/bitstream/10691/234/1/Interoperabilidade%20semantica.pdf>>. Citado 4 vezes nas páginas 43, 44, 45 e 46.

LEI de acesso a informacao. 2011. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm>. Citado 2 vezes nas páginas 23 e 24.

MINISTERIO DO PLANEJAMENTO. *Manual dos dados abertos: governo*. [S.l.], 2011. Disponível em: <http://www.w3c.br/pub/Materiais/PublicacoesW3C/Manual_Dados_Abertos_WEB.pdf>. Citado na página 29.

MINISTERIO DO PLANEJAMENTO. *Cartilha tecnica para publicacao de Dados Abertos no Brasil*. [S.l.], 2012. Disponível em: <<http://dados.gov.br/cartilha-publicacao-dados-abertos/>>. Citado na página 25.

MINISTERIO DO PLANEJAMENTO, ORCAMENTO E GESTAO. *Manual do Siorg - Cadastro de Estruturas Organizacionais*. [S.l.], 2014. Citado na página 63.

MORAIS, E. A. M.; AMBROSIO, A. P. L. *Ontologias: conceitos, usos, tipos, metodologias, ferramentas e linguagens*. [S.l.], 2007. Citado 3 vezes nas páginas 48, 53 e 60.

NICO, W.; BORST. *Construction of engineering Ontologies for Knowledge Sharing and Reuse*. Tese (Doutorado) — Universite Twente, Enschede, 1997. Citado na página 47.

NOY, D. M. N. *Ontology Development 101: A Guide to Creating Your First Ontology*. 2000. Citado 2 vezes nas páginas 49 e 56.

OPEN Definition. 1999. Disponível em: <<http://opendefinition.org/od/>>. Citado na página 29.

PADROES de Interoperabilidade de Governo Eletronico - e-PING. 2014. Disponível em: <<http://eping.governoeletronico.gov.br/>>. Citado na página 44.

PARTICIPATING Countries. 2014. Disponível em: <<http://www.opengovpartnership.org/countries>>. Citado 2 vezes nas páginas 26 e 27.

PATEL-SCHNEIDER, P. F. *Position Paper: A Comparison of Two Modelling Paradigms in the Semantic Web*. Citado na página 47.

PLANEJAMENTO, M. do. 2014. Citado 2 vezes nas páginas 63 e 64.

PLANO de Acao para Implantacao da INDA Infraestrutura nacional de Dados Abertos. 2012. Disponível em: <<http://wiki.gtinda.ibge.gov.br/MainPage.ashx>>. Citado 2 vezes nas páginas 24 e 25.

PRADA, R. O que e web semantica? *TecMundo*, mar 2008. Disponível em: <<http://www.tecmundo.com.br/web/800-o-que-e-web-semantica-.htm>>. Citado na página 35.

RAUTENBERG, S.; TODESCO, J. L. A methodology for the development of ontologies. *Revista Ciencias Exatas e Naturais*, v. 10, n. 2, dec 2008. Citado 2 vezes nas páginas 53 e 54.

RIBEIRO, A. L. As camadas da arquitetura da web semantica. *Blog Elementos da Web Semantica*, 2013. Disponível em: <<http://adagenor.blogspot.com.br/2008/03/as-camadas-da-arquitetura-da-web.html>>. Citado na página 36.

ROSA, P. A. *Web Semantica*. Tese (Doutorado) — Universidade de Sao Paulo, Instituto de Matematica e Estatistica, dec 2002. Citado 6 vezes nas páginas 36, 37, 38, 40, 41 e 42.

SEGUNDO, J. E. S. Linguagem xml como base na busca da interoperabilidade e organizacao da informacao. *Simposio em filosofia e ciencias.*, 2003. Citado 4 vezes nas páginas 36, 38, 41 e 42.

SEMANTIC Interoperability. Report about Technological State-of-the-Art. 2008. Disponível em: <<http://archive.today/MMbkl>>. Citado na página 43.

SILVA, D. B. da. *Transparencia na esfera publica interconectada*. Tese (Doutorado) — Faculdade Casper Libero, Escola de comunicacao, 2010. Citado 2 vezes nas páginas 23 e 31.

STAAB, S. et al. *Knowledge Processes and Ontologies*. 2000. Citado na página 52.

STRICKLAND, J. *Como funcionara a Web 3.0*. 2014. Disponível em: <<http://tecnologia.hsw.uol.com.br/web-30.htm>>. Citado na página 34.

SURE, Y.; STAAB, S.; STUDER, R. *On-To-Knowledge Methodology (OTKM)*. 2003. Citado na página 52.

TOTAL number of Websites. 2014. Disponível em: <<http://www.internetlivestats.com/total-number-of-websites/>>. Citado na página 33.