



**APLICAÇÃO DE INFERÊNCIA SEMÂNTICA PARA EXPLORAÇÃO
CONCEITUAL DE BASES DE DADOS RELACIONAIS**

DANIEL ALVES DA SILVA

**DISSERTAÇÃO DE MESTRADO EM ENGENHARIA ELÉTRICA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

**FACULDADE DE TECNOLOGIA
UNIVERSIDADE DE BRASÍLIA**

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

**APLICAÇÃO DE INFERÊNCIA SEMÂNTICA PARA EXPLORAÇÃO
CONCEITUAL DE BASES DE DADOS RELACIONAIS**

DANIEL ALVES DA SILVA

ORIENTADOR: RAFAEL TIMÓTEO DE SOUSA JÚNIOR

DISSERTAÇÃO DE MESTRADO EM ENGENHARIA ELÉTRICA

PUBLICAÇÃO: PPGEE.DM-561/2014.

BRASÍLIA / DF, 23 de Maio de 2014.

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

**APLICAÇÃO DE INFERÊNCIA SEMÂNTICA PARA EXPLORAÇÃO
CONCEITUAL DE BASES DE DADOS RELACIONAIS**

DANIEL ALVES DA SILVA

**DISSERTAÇÃO DE MESTRADO ACADÊMICO SUBMETIDA AO DEPARTAMENTO DE
ENGENHARIA ELÉTRICA DA FACULDADE DE TECNOLOGIA DA UNIVERSIDADE DE BRASÍLIA,
COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA OBTENÇÃO DO GRAU DE MESTRE.**

APROVADA POR:

**RAFAEL TIMÓTEO DE SOUSA JÚNIOR, Dr., ENE/UNB
(ORIENTADOR)**

**FLÁVIO ELIAS GOMES DE DEUS, Dr., ENE/UNB
(EXAMINADOR INTERNO)**

**ROBSON DE OLIVEIRA ALBUQUERQUE, Dr., ABIN
(EXAMINADOR EXTERNO)**

Brasília, 23 de Maio de 2014.

FICHA CATALOGRÁFICA

SILVA, DANIEL A.

Aplicação de Inferência Semântica para Exploração Conceitual de Bases de Dados Relacionais. [Distrito Federal] 2014.

xiii, 115p, 210 x 297 mm (ENE/FT/UnB, MESTRE, Engenharia Elétrica, 2014).

Dissertação de Mestrado - Universidade de Brasília, Faculdade de Tecnologia. Departamento de Engenharia Elétrica.

1. Sistemas Distribuídos

2. *Web Semântica*

3. Ontologia

4. Serviços de Dados

I. ENE/FT/UnB.

II. Título (Série)

REFERÊNCIA BIBLIOGRÁFICA

Silva, A, Daniel. (2014). Aplicação de Inferência Semântica para Exploração Conceitual de Bases de Dados Relacionais. Dissertação de Mestrado em Engenharia Elétrica, Publicação PPGEE.DM-561/2014. Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 115p.

CESSÃO DE DIREITOS

AUTOR: DANIEL ALVES DA SILVA

TÍTULO: APLICAÇÃO DE INFERÊNCIA SEMÂNTICA PARA EXPLORAÇÃO
CONCEITUAL DE BASES DE DADOS RELACIONAIS.

GRAU: Mestre ANO: 2014

É concedida à Universidade de Brasília permissão para reproduzir cópias desta dissertação de DANIEL ALVES DA SILVA e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. Do mesmo modo, a Universidade de Brasília tem permissão para divulgar este documento em biblioteca virtual, em formato que permita o acesso via redes de comunicação e a reprodução de cópias, desde que protegida a integridade do conteúdo dessas cópias e proibido o acesso a partes isoladas desse conteúdo. O autor reserva outros direitos de publicação, e nenhuma parte desta dissertação de DANIEL ALVES DA SILVA pode ser reproduzida sem a autorização por escrito do autor.

DANIEL ALVES DA SILVA
QNA 54, Lote 21 – Taguatinga Norte/DF
CEP: 71215-200

AGRADECIMENTOS

Ao meu orientador e amigo, Prof. Dr. Rafael Timóteo de Sousa Júnior, que me orientou e apoiou no momento mais difícil da minha vida.

Aos professores e amigos da Faculdade de Tecnologia, em especial ao Fabio L. L. Mendonça e ao Valério A. Martins.

Ao CDT/UnB, órgão de apoio ao desenvolvimento tecnológico da UnB, que me deu suporte na qualidade de bolsista de projetos em que pude inserir minha pesquisa.

Aos colegas do Laboratório Latitude e ao corpo administrativo da Faculdade de Tecnologia que me incentivaram para o desenvolvimento deste trabalho.

Agradeço também ao Ministério do Planejamento e toda a equipe da Secretaria do Orçamento Federal, que sempre ofereceu apoio às equipes de pesquisadores da Universidade de Brasília, à Ministra do Planejamento Miriam Belchior, ao Secretário de Orçamento Federal José Roberto Fernandes Junior, ao Secretário-Adjunto de Orçamento Federal, Antônio Carlos Paiva Futuro, e em especial a Ex-Secretária de Orçamento Federal, Célia Correa, e ao Coordenador Geral de Tecnologia e Informação, Carlos Eduardo Lacerda Veiga, pelo apoio, companheirismo e incentivo na busca dos melhores resultados, características de grandes líderes.

À minha amiga Lícia Albanese, exemplo de filha, profissional e mãe. Nunca esquecerei sua grande e qualificada ajuda em minha publicação e no meu desenvolvimento pessoal e profissional.

À minha esposa e família pelo apoio e incentivo que foi dado durante todo o tempo em que estive envolvido neste trabalho.

DEDICATÓRIA

Dedico este trabalho para meus avós, Terezinha de Lourdes Silva e José Alves da Silva Neto (*in memorian*), e a toda minha família.

Também dedico a minha querida sogra (*in memorian*) e a minha esposa.

RESUMO

APLICAÇÃO DE INFERÊNCIA SEMÂNTICA PARA EXPLORAÇÃO CONCEITUAL DE BASES DE DADOS RELACIONAIS.

A presente dissertação propõe uma extensão ao modelo ontológico do Orçamento Federal Brasileiro, de modo a apresentar valores de execução orçamentária correlacionados com as restrições originais do orçamento planejado. A hipótese é a de que tal extensão permite aumentar as capacidades dos sistemas existentes de análise e execução orçamentária e suporta a análise estratégica comparativa entre o orçamento planejado em relação aos resultados da sua execução. Para verificar essa hipótese, indo além do modelo anterior que oferece apenas informações anuais do planejamento orçamentário, o novo modelo proposto acrescenta uma nova granularidade para séries temporais e critérios adicionais de classificação. A avaliação dessa solução é feita pela implementação do modelo de modo a produzir *datasets* do orçamento em formato RDF (*resource description framework*), conforme recomendado pelo *World Wide Web Consortium* (W3C) como linguagem de escolha para representação do conhecimento. Desse modo, os resultados finais são publicados e acessados em um terminal de consultas denominado *endpoint*, visando dar suporte a um monitoramento público e *just-in-time* para o planejamento e para a execução do orçamento. Estatísticas de acesso apontam que a utilização desse *endpoint* comprova que a extensão proposta permite acesso ao conhecimento contido na execução orçamentária, mesmo por pessoas e sistemas externos ao órgão de planejamento do orçamento, o que confirma a hipótese da dissertação.

ABSTRACT

APPLICATION OF INFERENCE FOR SEMANTIC EXPLORATION CONCEPT OF RELATIONAL DATABASES

This master thesis proposes an extension to the Ontological model of the Brazilian Federal Budget, in order to present the budget execution values correlated with the original restrictions defined in the planned budget. The hypothesis is that this extension allows increasing the capabilities of existing budget analysis and execution systems since it supports the strategic comparative analysis between the planned budget and the results of its execution. To verify this hypothesis, going beyond the previous model that only provides annual information of budget planning, the proposed new model adds a new granularity for time series and additional classification. The evaluation of this solution is done by implementing the model in order to produce the budget datasets in RDF (Resource Description Framework), the format recommended by World Wide Web Consortium (W3C) as the chosen language for knowledge representation. Thus, the final results are published and accessed on an endpoint, aiming to support public and just-in-time monitoring for the planning and execution of the budget. Utilization data of this endpoint shows that the proposed extension allows access to the knowledge contained in budget execution, even by external people and systems of the Government Federal Budget, which confirms the hypothesis of this master thesis.

ÍNDICE

1 - INTRODUÇÃO	1
1.1 MOTIVAÇÃO	3
1.2 OBJETIVOS	4
1.2.1 Objetivo Geral	4
1.2.2 Objetivos Específicos.....	5
1.3 ORGANIZAÇÃO DO TRABALHO	6
2 - FUNDAMENTAÇÃO TEÓRICA	7
2.1 CONCEITOS DE ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO.....	7
2.2 Sistemas de Organização do Conhecimento	8
2.2.1 Tesouro.....	12
2.2.2 Ontologias	13
2.3 Linguagens de representação do conhecimento	20
2.3.1 XML.....	20
2.3.2 Serialização RDF	21
2.3.2.1 <i>TURTLE</i>	22
2.3.2.2 <i>N-TRIPLES</i>	23
2.3.3 Vocabulários	24
2.3.4 Linguagens Ontológicas.....	25
2.4 CONCEITOS DE INTEROPERABILIDADE.....	27
2.4.1 Interoperabilidade Ontológica	28
2.4.2 <i>Web Semântica e Linked Data</i>	28
2.4.3 <i>Linked Open Data - LOD</i>	33
2.4.4 URI.....	35
2.4.5 SPARQL.....	36
2.4.6 Dados Governamentais Abertos	37
2.4.7 Experiências de Dados Governamentais Abertos	38
3 - FERRAMENTAS UTILIZADAS	41
3.1 Ferramentas para construção de ontologias	41
3.2 Ferramentas de Triplificação.....	43
3.2.1 Plataforma D2RQ.....	44
3.2.2 Ferramenta <i>Dump-RDF</i>	45

3.3	Sistemas de Armazenamento de Ontologias	46
3.3.1	<i>Framework Jena</i>	46
3.3.2	Triplestore TDB.....	48
3.3.3	Servidor <i>Fuseki</i>	48
4	SOLUÇÃO PROPOSTA	50
4.1	MODELO ORÇAMENTÁRIO	50
4.2	MODELO ONTOLÓGICO	54
4.3	MODELO SISTÊMICO	61
4.3.1	Base relacional.....	61
4.3.2	Arquivo de mapeamento D2RQ.....	62
4.3.3	<i>Dump-RDF</i>	62
4.3.4	Arquivo RDF gerado.....	63
4.3.5	TDB.....	63
4.3.6	Terminal <i>Fuseki</i>	63
4.3.7	Questões de Interface Visual.....	63
5	PROVA DE CONCEITO E RESULTADOS OBTIDOS	65
5.1	PROVA DE CONCEITO	65
5.1.1	Extensão ontológica.....	65
5.1.2	Geração do RDF	67
5.1.3	Carga do <i>Triple store</i>	70
5.1.4	SPARQL <i>Endpoint</i>	72
5.1.5	Interface da aplicação	73
5.1.5.1	Interface de Consulta Livre	76
5.1.5.2	Interface de Consulta Personalizada	78
5.2	RESULTADOS OBTIDOS	84
6	CONCLUSÕES	90
6.1	Trabalhos futuros.....	92
	ARTIGOS PUBLICADOS	93
	REFERÊNCIAS BIBLIOGRÁFICAS	94
	APÊNDICES.....	100

ÍNDICE DE TABELAS

Tabela 2.1 Níveis de classificação do conhecimento.....	7
Tabela 2.2 Linha do tempo dos sistemas de organização do conhecimento.....	9
Tabela 3.1 Ferramentas para construção, uso e edição de ontologias.....	41
Tabela 3.2 Características das ferramentas de extração sobre dados estruturados.	44
Tabela 4.1 Relação de Classes e seus significados.	57
Tabela 5.1 Quantidade de triplas por classe e quantidade total.	84
Tabela 5.2 Item de despesa da LOA 2013 em RDF e serializada em <i>N-Triples</i>	85
Tabela 5.3 Item de despesa da LOA 2013 em RDF e serializada em <i>N-Triples</i> (Continuação)	86

ÍNDICE DE FIGURAS

Figura 2.1 Tipos de Sistemas de Organização do Conhecimento.	12
Figura 2.2 Tipos de Ontologia de acordo com o seu nível de dependência.	19
Figura 2.3 Exemplo de uma tripla RDF.	21
Figura 2.4 Exemplo de um grafo RDF.	22
Figura 2.5 Exemplo de tripla do orçamento federal em RDF serializado no formato <i>Turtle</i>	23
Figura 2.6 Pilha de Normalizações da <i>Web Semântica</i>	29
Figura 3.1 Manual de execução da ferramenta <i>Dump-rdf</i>	45
Figura 3.2 Arquitetura do <i>Framework Jena</i>	47
Figura 4.1 Representação simplificada da relação das ontologias do planejamento e de execução orçamentária.	52
Figura 4.2 Item de despesa da LOA de 2013.	53
Figura 4.3 Ontologia da Classificação das Despesas do Orçamento Federal.	54
Figura 4.4 Modelo arquitetural da solução proposta.	61
Figura 4.5 Mapeamento do classificador Programa	62
Figura 4.6 Ontologia formato N-Triples	63
Figura 4.7 Resultado de consulta SPARQL do orçamento.	64
Figura 5.1 Modelagem de Classes no <i>protegé</i>	66
Figura 5.2 Parâmetros de conexão utilizados pelo arquivo de mapeamento.	68
Figura 5.3 Mapeamento da classe <i>loa:UnidadeOrçamentaria</i>	69
Figura 5.4 Exemplo de consulta SPARQL.	73
Figura 5.5 Tela inicial de acesso.	74
Figura 5.6 Tela de acesso ao Dados abertos.	75
Figura 5.7 Formulário de consulta livre.	76
Figura 5.8 Formulário de seleção de consulta.	77
Figura 5.9 Relatório gerado.	77
Figura 5.10 Formulário de seleção de tipo de consulta.	78
Figura 5.11 Formulário de identificação do usuário.	78
Figura 5.12 Formulário de solicitação de senha de acesso.	79
Figura 5.13 Formulário de identificação do usuário.	79
Figura 5.14 Adicionando uma consulta personalizada.	80
Figura 5.15 Formulário de gerenciamento de consultas.	80
Figura 5.16 Formulário de consultas gravadas.	81

Figura 5.17 Formulário de consultas gravadas.	81
Figura 5.18 Tela de resultado da consulta.	82
Figura 5.19 Formulário de seleção de consultas.....	83
Figura 5.20 Relatório em formato PDF.....	83
Figura 5.21 Mapeamento de informações entre os arquivos de sentenças RDF e OWL.....	88

LISTA DE ACRÔNIMOS

API	<i>Application Programming Interface</i> – Interface de Programação de Aplicações
BI	<i>Business intelligence</i> – Inteligência de Negócios
IEEE	<i>Institute of Electrical and Electronics Engineers</i>
HTML	<i>Hypertext Markup Language</i> – Linguagem de Marcação de Hipertextos
LOD	<i>Linked open data</i>
OWL	<i>OWL Web Ontology Language</i>
RDF	<i>Resource Description Framework</i>
RDF Schema	<i>RDF Vocabulary Description Language</i>
URI	<i>Uniform Resource Identifier</i>
URL	<i>Uniform Resource Locator</i>
XML	<i>Extensible Markup Language</i>
W3C	<i>World Wide Web Consortium</i>
<i>Web</i>	<i>World Wide Web</i>
<i>WEB Services</i>	Modelo de integração de sistemas, utilizado na comunicação entre aplicações diferentes por meio de serviços no ambiente <i>Web</i>
W3C	<i>World Wide Web Consortium</i> - Comunidade internacional na qual as organizações-membros, funcionários de tempo integral e agentes de setores públicos trabalham juntos para desenvolver padrões para <i>Web</i> (<i>Web standards</i>).
XML	<i>Extensible Markup Language</i> - Linguagem de Marcação Extensível.

1 - INTRODUÇÃO

A grande quantidade de dados disponibilizada mundialmente impõe um grande desafio não somente para o seu armazenamento, mas também para o seu melhor aproveitamento e acesso. O movimento mundial pela evolução da inteligência na busca e processamento desses dados nos remete à terceira geração da *Internet*, denominada por Berners-Lee (2001) como *Web Semântica*. Tal tecnologia vem ao encontro da necessidade cada vez maior dos governos modernos e democráticos em prestar contas à sua população de forma ágil e transparente.

Castells (1999) denomina esse novo modo de desenvolvimento, com base na tecnologia da informação, como sociedade informacional. Além disso, afirma que nessa nova sociedade a fonte de produtividade acha-se na tecnologia de geração de conhecimentos, de processamento de informação e de comunicação de símbolos. Na verdade, conhecimento e informação são elementos cruciais em todos os modos de desenvolvimento.

Nesse contexto, De Masi (2000) define gestão social do conhecimento como um método capaz de propiciar à sociedade um mecanismo flexível, bilateral e colaborativo de comunicação entre empresas, cidadãos e Estado em busca da democratização e socialização do conhecimento, fomentando a criação de redes de inteligência coletiva para a construção e a validação do conhecimento em prol de toda uma coletividade.

Atento a esse movimento, o governo brasileiro tem buscado melhorias no tratamento da informação de forma a suportar uma maior transparência e o controle social de suas ações, principalmente por intermédio da implementação de soluções voltadas ao suporte da Lei de Responsabilidade Fiscal (Brasil, 2000) e a Lei de Acesso à Informação (Brasil, 2011). Uma das ações de destaque é o acesso transparente e público aos controles do planejamento e uso dos recursos orçamentários.

No entanto, para que os dados governamentais publicados sejam realmente acessíveis para a população em geral, é necessário um tratamento prévio das informações devido ao seu volume e complexidade. Tal acesso pela população pode ser feito de forma automatizada, desde que a informação esteja em um formato estruturado e interoperável (Lathrop e Ruma, 2010).

A força da Lei de Acesso à Informação e a tecnologia de dados abertos não são suficientes para reverter décadas de sistemas legados, sendo muitos desses ainda fora de plataformas *Web*. Nesse novo paradigma semântico, é clara a necessidade de definir padrões estruturados para os sistemas de informações de governo.

Atualmente, a informação gerada é muitas vezes encapsulada em gráficos, imagens e formatos proprietários, tornando-se assim um grande obstáculo ao processamento automático, bem como para a interoperabilidade entre aplicações do próprio governo, que dispense um enorme esforço em aplicações de *Web Services*, *data warehouse* e *business intelligence* para interoperarem e fornecerem as informações requisitadas e necessárias à tomada de decisões. Verifica-se que isso também se aplica a entidades privadas que necessitam cada vez mais se integrar aos seus fornecedores e às necessidades do mercado globalizado, que impõem grandes desafios sistêmicos com aquisições e fusões. No entanto, nem todas as esferas ou unidades da Federação contam com estrutura tecnológica ou recursos para promover essas grandes e contínuas empreitadas tecnológicas.

A *World Wide Web Consortium* (W3C)¹ vem definindo padrões, tais como o *Resource Description Framework* - RDF², que estão criando alternativas para as instituições públicas e privadas ingressarem no mundo da *Web Semântica*. Dessa forma, por meio do uso de dados estruturados e com relacionamentos definidos junto a outros repositórios, o reaproveitamento e a difusão de tais dados se torna possível e natural.

A exemplo do projeto de dados abertos do Reino Unido³, o governo brasileiro lançou um sítio de Dados Abertos⁴, com o objetivo de promover a publicação, além de divulgar aplicações que façam a manipulação destes.

Em 2012, resultou-se o projeto do acordo de cooperação técnica entre a Secretaria de Orçamento Federal do Ministério do Planejamento, Orçamento e Gestão – SOF/MP e a Universidade de Brasília. Araújo *et al.* (2012) propõe um modelo de transparência interoperável que utiliza a estruturação de dados no formato RDF para organizar o acesso

¹ <http://www.w3.org/>

² <http://www.w3.org/RDF/>

³ <http://www.data.uk/>

⁴ <http://dados.gov.br/>

público à LOA – Lei Orçamentária Anual. Trata-se nesse caso de dados estáticos, pois após a aprovação pelo congresso e a sanção presidencial, estes dados não podem ser alterados. Tal trabalho teve como foco a formação de uma base de conhecimento visando aumentar a transparência e o acesso ao orçamento em atenção a outros esforços que vêm sendo dispendidos pelo Governo Federal, para tornar fácil e acessível o conhecimento da distribuição do orçamento no Brasil. Porém, a simples divulgação de dados do orçamento não permite a sua análise completa. Para tanto, faz-se necessário também um modelo conceitual, alinhado aos dados publicados e que apresente *just-in-time* os pagamentos associados a tal orçamento, no mesmo ano em análise.

Motivada pela Lei de Acesso à Informação (Brasil, 2011) e com o objetivo de ampliar a informação disponibilizada, a Secretaria de Orçamento Federal e o Ministério do Planejamento, Orçamento e Gestão – SOF/MP, mais uma vez, em parceria técnica com a Universidade de Brasília, realizaram um trabalho adicional voltado à formação de um modelo compatível que pudesse não só disponibilizar os dados estáticos, como suportar a comparação do planejamento orçamentário em relação à correspondente execução orçamentária na sua dinâmica. Entende-se que esses dados favorecem outros usos, tal como o provimento de dados para acompanhamento estratégico setorizado e classificado dessa execução orçamentária.

Assim, propondo uma contribuição que vai além dos trabalhos precedentes, a presente dissertação busca realizar uma prova de conceito, objetivando ampliar o nível de informações com um novo provimento de dados governamentais de forma estruturada e legível por máquina, com foco nos dados da execução orçamentária, haja vista que estes dados representam efetivamente os indicadores da qualidade da gestão pública por colocar em relevo a utilização dos recursos públicos.

1.1 MOTIVAÇÃO

Tendo em vista a enorme quantidade de dados governamentais publicados no Brasil, bem como a forma como são publicados, evidencia-se a dificuldade na utilização de sistemas computacionais para realizar esse trabalho, uma vez que tais dados estão em formato sintático muitas vezes proprietário, ou até mesmo publicados apenas na forma de gráficos e

bases de dados legadas sem integração com os demais sistemas governamentais. Por outro lado, o processamento manual desses dados é geralmente dispendioso ou mesmo inviável.

Para resolução dessa situação, é motivação central dessa proposta a disponibilização de dados governamentais em formato aberto, seguindo as recomendações da *Web Semântica*⁵ e possibilitando aos desenvolvedores um reaproveitamento mais eficiente dos dados. De forma complementar, o desenvolvimento da aplicação da prova de conceito pretende demonstrar que a publicação de dados governamentais, em formato aberto, simplificará etapas de desenvolvimento de aplicações, visto que os dados já estarão tratados e prontos para uso. Em particular, os dados de planejamento e de execução do orçamento federal são de grande interesse para tal forma de publicação, haja vista que caracterizam diretamente as condições de realização dos gastos públicos.

Dessa forma, por intermédio da evolução do modelo ontológico do orçamento, será projetada uma aplicação semântica para publicação dos dados da execução orçamentária como prova de conceito, dentro dos padrões estabelecidos pela W3C.

1.2 OBJETIVOS

O presente trabalho objetiva especificar e implementar uma solução para disponibilização de conceitos e dados oriundos do orçamento federal brasileiro por meio da formação de um ambiente de exploração de dados triplificados, ou seja, convertidos em formato RDF, com basea em uma ontologia para uso comum por comunidades públicas, organizações governamentais e outros.

1.2.1 Objetivo Geral

O objetivo geral deste trabalho é implementar um modelo de automação para publicação de dados governamentais em formato aberto, mais especificamente os dados da execução orçamentária brasileira a partir da evolução da ontologia do orçamento público federal e do desenvolvimento de uma *Framework* para apresentar os dados estruturados

⁵ <http://www.w3c.br/Padroes/WebSemantica>

semanticamente. Espera-se, também, mostrar a facilidade no desenvolvimento de aplicações quando os dados estão em formato adequado.

Para comprovar a hipótese de que é possível democratizar o acesso à informação estimulando a participação efetiva da população na fiscalização da utilização dos recursos públicos, por meio de um modelo ontológico do orçamento federal que abranja a execução diária do orçamento, a presente pesquisa descreve a implementação de um modelo arquitetural capaz de extrair dados de repositórios relacionais, estruturar tais dados em formato RDF, usando a sintaxe *N-Triples*, disponibilizando-os por intermédio de *triple stores*, além de implementar *endpoints* para a exploração conceitual. Desse modo, o modelo proposto é submetido a uma prova de conceito para efeito de avaliação dos resultados.

1.2.2 Objetivos Específicos

Para o cumprimento do objetivo geral proposto, foram estabelecidos os seguintes objetivos específicos:

- Atualizar a ontologia de domínio do orçamento federal brasileiro para exploração dos dados de execução orçamentária.
- Desenvolver e disponibilizar fontes integradas e automatizadas para extração de dados do Orçamento Federal Brasileiro, transformando-os e armazenando-os em formato triplicado, ou seja, uma estrutura de texto em sintaxe RDF composta de sujeito, predicado e objeto, compatível com os padrões brasileiros de dados abertos.
- Propor um acordo conceitual em nível federal, replicável dentro de todas as esferas de governo para obter dados do orçamento federal.
- Desenvolver e disponibilizar aplicações clientes para acesso e exploração desses dados triplicados.
- Atender aos padrões de interoperabilidade preconizados pelo Governo Federal (e-PWG, 2010).

1.3 ORGANIZAÇÃO DO TRABALHO

Para um melhor entendimento e organização, este trabalho é dividido em Capítulos conforme relacionado a seguir.

O Capítulo 2 apresenta os principais conceitos acerca de representação de dados e interoperabilidade, bem como iniciativas governamentais no escopo da *Web Semântica*.

O Capítulo 3 reúne todas as ferramentas utilizadas na prova de conceito desta dissertação.

O Capítulo 4 apresenta os processos e o modelo arquitetural proposto neste trabalho, tratando a modelagem ontológica, com a definição da hierarquia de classes e seus atributos e propriedades. Ainda, neste capítulo, é apresentado o padrão estruturado RDF para o povoamento do banco com essas instâncias, assim como as inferências sobre essas mesmas instâncias.

O Capítulo 5 aborda a implementação da prova de conceito que apresentará a base de dados utilizada pela aplicação, algumas de suas consultas e seus resultados.

O Capítulo 6 apresenta a conclusão deste trabalho com a discussão dos resultados obtidos a partir da exploração semântica dos dados, ilustrando os consequentes avanços no processo governamental, assim como as vantagens do modelo proposto.

O apêndice A apresenta a parte mais relevante dos códigos da ontologia.

2 - FUNDAMENTAÇÃO TEÓRICA

Nesse capítulo são apresentados os conceitos básicos necessários para o suporte metodológico, científico e tecnológico essenciais para o êxito do trabalho.

Inicialmente, são apresentados os conceitos básicos de representação do conhecimento, sistemas de organização do conhecimento. Dando prosseguimento são discutidos conceitos básicos de interoperabilidade, dados governamentais abertos e alguns casos de uso pelo mundo.

2.1 CONCEITOS DE ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO

Para Dahlberg (2006), a organização do conhecimento é a ciência que ordena a estruturação e sistematização dos conceitos, de acordo com suas características, que podem ser definidas como elementos de herança do objeto, assim como a aplicação dos conceitos e classes dos conceitos ordenados pela indicação de valores dos referentes conteúdos dos objetos ou assuntos. A partir dessa organização do conhecimento, criam-se ferramentas que apresentam a interpretação organizada e estruturada do objeto, chamados de sistemas de organização do conhecimento.

Segundo Brachman 1979 (*apud* Campos, 2001), a representação do conhecimento pode ser classificada em níveis conforme a Tabela 2.1:

Tabela 2.1 Níveis de classificação do conhecimento

Nível	Descrição
Lógico	Nível de formalização, não existe preocupação com a semântica em termos dos conceitos e de suas relações, o foco está direcionado para uma dada “sintaxe”.
Epistemológico	A noção genérica de um conceito é dada nos fundamentos de estruturação de conhecimento, ou seja, especifica-se a estrutura dos conceitos e seus relacionamentos.
Ontológico	Tem o objetivo de limitar o número de possibilidades de interpretação do conceito em um determinado contexto, usando-se de um formalismo para representar o conteúdo do

	conceito. Portanto, trata-se do processo de organização e classificação de um dado domínio, trabalhando com definições de conceitos que nele estão inseridos.
Conceitual	Os conceitos possuem a priori uma interpretação definida.

Bräscher e Café (2008) definem a organização do conhecimento como o processo de modelagem do conhecimento para construções de representações do conhecimento. A partir da análise do conceito, assim como de suas características, é possível determinar o espaço que um conceito ocupa dentro de um domínio e suas relações com os outros conceitos que formam um determinado sistema notacional.

Nesse trabalho, aborda-se a representação do conhecimento num escopo restrito às representações construídas pelo homem. Considera-se, portanto, que esses tipos de representações são modelos de abstração do mundo real construídos para determinada finalidade, caracterizando, assim, os sistemas de organização do conhecimento como tipos de representação do conhecimento.

2.2 SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO

De acordo com Bräscher e Café (2008) os sistemas de organização do conhecimento são objetos de interesse para a Ciência da Informação, pois têm papel importante na padronização da terminologia adotada para organização e recuperação de informações ao delimitar o uso de termos e definir conceitos e relações de alguma área do conhecimento de forma compartilhada e consensual.

Ainda de acordo com Bräscher e Café (2008), a organização do conhecimento é o processo de modelagem que visa construir representações do conhecimento e busca a análise do conceito e de suas características para estabelecer a posição que cada conceito ocupa em um domínio específico, assim como suas relações com os demais conceitos.

Dahlberg (2006) afirma que os sistemas de organização do conhecimento são sistemas conceituais, ou tipos de representações do conhecimento que nascem do processo de organização do conhecimento. Esse autor (*ibidem*) aponta que os sistemas de organização

do conhecimento também podem ser aplicados para mapear objetos informacionais, representando assim os assuntos dos documentos como, por exemplo, em um sistema de informação, os quais podem ser utilizados nos processos de classificação e indexação. A qualidade obtida na recuperação da informação dependerá substancialmente desses instrumentos. Portanto, os padrões de organização devem ser definidos desde a concepção do sistema para permitir que a informação seja encontrada posteriormente.

A representação do conhecimento é uma questão que preocupa a documentação desde sua origem. O problema agora é relevante em muitas outras situações além dos documentos e índices: a estrutura de registros e arquivos de bases de dados; a estrutura de dados nos programas de computador; a estrutura sintática e semântica da linguagem natural; a representação do conhecimento em inteligência artificial; e os modelos de memória humana. Em todos esses campos, é necessário decidir como o conhecimento pode ser representado de forma que tais representações possam ser manipuladas (Vickery, 1986).

De acordo com Vickery 2008 (*apud* Bräscher e Carlan, 2010) os sistemas de organização do conhecimento podem ser divididos em quatro grupos considerando uma linha do tempo, conforme apresentado na Tabela 2.2:

Tabela 2.2 Linha do tempo dos sistemas de organização do conhecimento

Período	Descrição
Era da pré-coordenação	Os sistemas de organização do conhecimento eram estruturas estáticas e atendiam às necessidades dos sistemas manuais de organização e recuperação da informação, como índices e catálogos. Incluem-se aqui as listas de cabeçalhos de assunto e as classificações.
Era da pós-coordenação	Os sistemas de organização do conhecimento tornam-se mais dinâmicos e possibilitam que cada um de seus elementos (termos) sejam manipulados de forma independente para representar os assuntos de cada documento. Exemplos de sistemas de organização do conhecimento dessa era são vocabulários controlados (listas de termos autorizados para uso na indexação e recuperação da informação) e tesouros.

Era da <i>Internet</i>	Os sistemas de organização do conhecimento que se destacam são as classificações hierárquicas que orientam o usuário na escolha do termo que melhor expressa sua questão de busca, como por exemplo, os elos estabelecidos por meio de URL entre itens da <i>Web</i> e os índices das ferramentas de busca, compostos de palavras extraídas dos conteúdos dos objetos informacionais.
Era da <i>Web Semântica</i>	Os sistemas de organização do conhecimento dessa era diferenciam-se dos demais por serem projetados para uso por agentes inteligentes. O principal exemplo são as ontologias.

De acordo com Hodge (2000)

Sistemas de organização do conhecimento englobam todos os tipos de instrumentos usados para organizar a informação e promover o gerenciamento do conhecimento. Incluem os esquemas de classificação que organizam materiais em nível geral (como livros em estantes), cabeçalhos de assunto que provêm acesso mais detalhado e listas de autoridades que controlam versões variantes de chaves de acesso à informação (nomes geográficos e nomes de pessoas). Incluem, ainda, esquemas menos tradicionais, tais como redes semânticas e ontologias (Hodge, 2000).

Em sua classificação, Hodge 2000 (*apud* Bräscher e Carlan, 2010) agrupa os sistemas de organização do conhecimento em três categorias:

1) Listas de termos: englobam listas de termos geralmente acompanhados de suas definições. Compreendem:

- Lista de autoridades: listas de termos que controlam as variações de nomes para entidades, como nomes de países, indivíduos ou instituições.
- Glossários: lista de termos com definições, geralmente de um assunto ou domínio específico.
- Dicionários: lista alfabética de palavras e suas definições. Seu escopo é mais geral que o dos glossários.
- *Gazetteers*: dicionário de nomes de lugares e acidentes geográficos, tais como cidades, rios, vulcões. Quando georeferenciados, os *gazetteers* apresentam as coordenadas para a localização de lugares na superfície da Terra.

2) Classificações e categorias: reúnem sistemas de organização do conhecimento que enfatizam a criação de classes de assuntos. São eles:

- Cabeçalhos de assunto: conjunto de termos controlados que representam os assuntos de uma coleção.
- Esquemas de classificação, taxonomias e esquemas de categorização: esquemas utilizados para agrupar entidades em classes mais gerais. Os esquemas de categorização são geralmente usados para reunir os termos de tesouros em tópicos. As taxonomias são usadas em modelos orientados a objeto e em sistemas de gestão do conhecimento para indicar grupos de objetos com base em características particulares.

3) Listas de relacionamentos: agrupam sistemas de organização do conhecimento que privilegiam a conexão entre termos e conceitos. Englobam:

- Tesouros: conjunto de termos que representam conceitos e as relações de equivalência, hierárquicas e associativas que se estabelecem entre eles.
- Redes semânticas: estrutura de conceitos e termos em forma de rede ou teia, os conceitos são nós e os relacionamentos expandem-se a partir dos nós.
- Ontologia: estrutura de conceitos e representação dos relacionamentos complexos entre eles, incluindo regras de inferência e axiomas.

Zeng e Salaba (2005) sintetizam a classificação proposta por Hodge (2000) em um gráfico conforme a Figura 2.1.

Tipos de Sistemas de Organização do Conhecimento

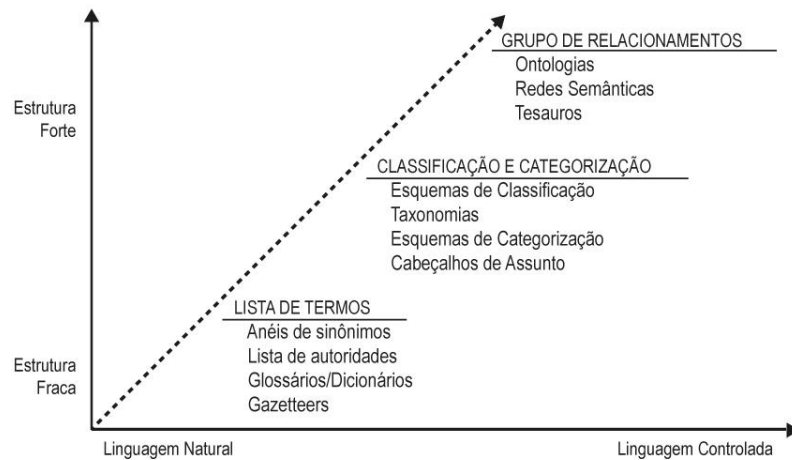


Figura 2.1 Tipos de Sistemas de Organização do Conhecimento.
(Adaptado de Zeng & Salaba, 2005.)

2.2.1 Tesouro

O termo tesouro vem do latim – *thesaurus* e do grego – *thesaurós*, que significa tesouro, armazenamento, repositório. Cavalcanti (1978) define tesouro como uma lista estruturada de termos associados, empregada por indexadores para descrever um documento com a desejada especificidade e para permitir aos usuários a recuperação da informação que procuram.

Gomes (1990) esclarece que os tesouros são classificados quanto à tipologia, de acordo com a(s) língua(s) que englobam, em monolíngues e multilíngues; quanto ao tipo de conceitos, são macrotesouros (representam conceitos mais amplos) e microtesouros (representam conceitos específicos); e, de acordo com a abrangência temática, são multidisciplinares e de disciplina específica.

Os tesouros são instrumentos de controle terminológico, utilizados em sistemas de informação para traduzir a linguagem dos documentos, dos indexadores e dos pesquisadores numa linguagem controlada, usada na indexação e na recuperação de informações.

De acordo com o IBICT (1984), as principais finalidades de um tesouro são:

a) controlar os termos usados na indexação mediante um instrumento que traduza a linguagem natural dos autores, usuários e indexadores para uma linguagem mais controlada;

b) uniformizar, mediante essa linguagem documentária, os procedimentos de indexação de profissionais em uma instituição ou numa rede cooperativa;

c) limitar o número de termos necessários à explicitação dos conceitos expostos pelos autores de uma área;

d) auxiliar na tarefa de recuperação da informação, fornecendo termos adequados para a estratégia de busca.

Guinchat e Menou (1994) apontam que as vantagens do tesauro são sua especificidade, maleabilidade e capacidade de descrever as informações de forma completa. Os tesouros auxiliam o indexador a localizar o conceito num campo semântico e o usuário a identificar assuntos relacionados que podem ser de interesse na busca. O emprego das novas tecnologias da informação possibilita diferentes formas de apresentação, mais amigáveis e compreensíveis, para orientar os usuários na navegação, ampliando ou restringindo sua busca.

Os tesouros são padronizados por meio das normas, como a ISO 25964-1:2011⁶ e a ISO 25964-1:2011⁷.

2.2.2 Ontologias

Historicamente o termo ontologia, da mesma forma que tesauro e taxonomia, tem origem no grego *ontos*, ser, e *logos*, palavra. O termo original é a palavra aristotélica “categoria”, que pode ser usada para classificar alguma coisa. Essas categorias serviam como base para classificar qualquer entidade objetiva e subjetiva existente no mundo. Nesse contexto, visualiza-se a ligação entre a origem da palavra ontologia e sua utilização na representação do conhecimento.

⁶ http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=53657

⁷ http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=53657

As ontologias estão presentes na Filosofia, Ciências Cognitivas, Ciência da Computação, Linguística, Ciência da Informação, Medicina etc. Dada a grande variedade de áreas de conhecimento, o tema apresenta características próprias em cada domínio.

O termo Ontologia é tradicionalmente utilizado na Filosofia para designar o estudo do ser, do que existe, das condições para a existência, entre outras. O termo foi adotado pela ciência da computação e a definição mais aceita nesse contexto é a de Grubber (1993): “Ontologia é uma especificação explícita e formal de uma conceitualização compartilhada de um domínio do conhecimento”. O termo “formal” nessa definição conota a capacidade de ser processada computacionalmente.

Neches e Fikes (1991) destacam que as ontologias, no sentido computacional, tornaram-se relevantes a partir dos anos 90 com o desenvolvimento na área de inteligência artificial. Uma ontologia possibilita a especificação de noções genéricas como tempo e espaço ou ligadas a um domínio de conhecimento específico, como engenharia ou ciências biológicas. O termo “compartilhado” se refere ao aspecto de que uma ontologia representa um conhecimento consensual acordado por um grupo como resultado de uma construção social.

Em ciência da computação, as ontologias se referem a um tipo especial de objeto ou artefato que tem por objetivo representar um modelo de um determinado domínio do conhecimento. Ontologias são meios de modelar formalmente a estrutura de um sistema, isto é, as suas entidades relevantes e as relações entre elas e que são úteis para um propósito determinado. Ontologias podem ser codificadas em arquivos por meio da linguagem OWL - *Ontology Web Language* (Guarino, 1998).

Silva *et al.* (2009) afirma que Vickery foi um dos primeiros do campo a dar atenção ao termo ontologia na Ciência da Informação, e ressalta questões presentes na pesquisa de ontologias que também são abordadas pela Biblioteconomia como a categorização de conceitos, princípio básico da Teoria da Classificação. Expõe que, apesar da similaridade evidente, os autores da Inteligência Artificial não referenciam trabalhos importantes da Biblioteconomia, como por exemplo, os métodos e técnicas empregadas por Lancaster

(2004) na construção de vocabulários direcionados a sistemas de recuperação da informação.

Na *Web* semântica a ideia principal é tornar a informação lida e entendida automaticamente. Nesse contexto, a ontologia passa a ser o principal instrumento de representação do conhecimento armazenado. A estrutura da *Web* semântica permite que as informações sejam armazenadas de maneira organizada e padronizada e, por conseguinte, recuperada (Berners-Lee, 2001).

Na *Web* semântica, as ontologias especificam descrições para conceitos, com semântica expressiva, explícita e bem definida, possivelmente interpretável por máquina, que processam automaticamente as informações encontradas na *Web*. Para tanto, usam linguagens compatíveis (OWL, *Extensible Markup Language* – XML – e XML *Scheme*, RDF e RDF *Scheme*) com o ambiente digital.

Para a organização e recuperação da informação, o conceito empregado é proveniente da proposta de Berners-Lee (2001) e está associado ao contexto da *Web* Semântica. Para o autor, ontologia é um documento ou arquivo que define formalmente as relações entre termos.

Segundo Almeida e Bax (2003), uma ontologia é criada por especialistas que definem as regras e as combinações entre os termos e suas relações em um domínio do conhecimento. Na interpretação desses autores, a ontologia visa melhorias nos processos de representação e recuperação da informação. Missikoff *et al* (2003) afirmam que por meio de um relacionamento semântico mais rico, a ontologia é capaz de representar um conhecimento conceitual. De forma sumária, os autores afirmam que uma ontologia pode ser vista como um tesouro enriquecido que permite ir além do relacionamento entre termos. Contrapondo esse conceito, Feitosa (2006) afirma que,

[...] do ponto de vista da representação do conhecimento, uma ontologia não deve ser concebida apenas como um vocabulário informal, ou mesmo como uma linguagem de termos estruturados – como um tesouro, por exemplo –, mas requer uma possibilidade de interpretação algorítmica dos seus significados e, por conseguinte, uma representação em uma linguagem formal, cujo processamento dos significados pode ser realizado por máquinas.

O Escritório Brasileiro do Consórcio *World Wide Web* (W3C)⁸ aborda a ontologia como “a definição de termos utilizados na descrição e na representação de uma área de conhecimento”. Essa definição parece nortear as principais definições sobre ontologia, posto que as definições mais aceitas possuem características comuns com a definição da W3C.

Durante os últimos anos, as atenções têm se voltado para a metodologia de construção de ontologias, as quais podem ser construídas a partir do reuso de ontologias disponíveis. Diferentes processos de construção de ontologias tendem a estar associados com diferentes tipos de ontologias.

Para cada domínio, as abordagens para construção de ontologias, geralmente, são específicas e limitadas e dependem dos interesses para os quais são construídas. Mas, independentemente do domínio, a elaboração de uma ontologia requer um trabalho complexo e dispendioso.

Gruber (1993) afirma que uma ontologia é formada por:

a) classes: representações dos conceitos de um domínio arranjadas em uma taxonomia, herdando as características ou propriedades de suas classes-pai, de forma similar ao modelo orientado a objetos;

b) relações: representação do tipo de interação entre os conceitos de um domínio;

c) axiomas: estruturas de restrição e interpretação dos conceitos envolvidos;

d) instâncias: representação dos objetos específicos (os próprios dados).

Sobre o uso das ontologias, são várias as abordagens apresentadas na *Literatura da Ciência da Informação*, não ficando restrito ao controle da terminologia, mas sendo esse o principal uso e o mais abordado pelos autores.

Almeida e Bax (2003) afirmam que as ontologias podem melhorar os processos de recuperação de informação ao organizar o conteúdo de fontes de dados em um

⁸ <http://www.w3c.br/Home/WebHome>

determinado domínio, uma vez que a complexidade na construção das relações entre os termos vão muito além das relações presentes em um tesouro, por exemplo.

Para Brandão e Lucena (2002), a utilização das ontologias possibilita, além de definir conceitos provenientes de domínios do conhecimento (o que evita ambiguidades), definir relacionamentos entre esses conceitos e, principalmente, realizar inferências, fato que a torna bastante importante na recuperação da informação. Os autores afirmam ainda que as ontologias passaram a ser consideradas aliadas da Ciência da Informação, no que diz respeito à organização e à representação do conhecimento, consideradas como modelos de linguagem documentária.

Assim como Breitman (2005), as autoras Campos e Gomes (2008) acreditam que a utilização da ontologia está diretamente ligada ao contexto da *Web* semântica, sendo necessária para que essa venha a funcionar de forma efetiva, permitindo aos computadores o acesso a coleções estruturadas de informações e a conjuntos de regras de inferência definidas por intermédio das ontologias e sua forma de organização.

Apesar das ontologias possuírem esses elementos típicos, eles não são obrigatórios. Com isso,

uma ontologia pode assumir vários formatos, mas necessariamente deve incluir um vocabulário de termos e alguma especificação de seu significado. Esse deve abranger definições e uma indicação de como os conceitos estão inter-relacionados, o que resulta na estruturação do domínio e nas restrições de possíveis interpretações de seus termos (Breitman, 2005).

Como exemplos básicos dos aspectos taxonômicos de uma ontologia, tem-se:

- o conceito “Tempo”, que tem em suas divisões humanas naturais as classes de anos, meses de um ano, dias de um mês, entre outras divisões;
- o conceito de “Geografia”, que no Brasil pode ser descrita pela hierarquia de regiões, unidade federativas das regiões, municípios das unidades federativas etc;

Outro ponto importante é que o escopo de aplicação da ontologia se diferencia na visão de alguns autores:

Segundo Guarino (1998), as ontologias são classificadas, considerando o seu conteúdo nas seguintes categorias:

- a) Ontologias genéricas: referem-se a conceitos de uma forma geral, os quais independem de um domínio particular. Por exemplo, espaço, tempo, evento.
- b) Ontologias de domínio: apresentam conceitos sobre um determinado domínio, utilizando um vocabulário a partir de um domínio mais genérico. Por exemplo, ontologias de veículos, documentos, computação, medicina.
- c) Ontologias de tarefas: tratam de conceituações para resolução de problemas, sem levar em consideração o domínio no qual ocorrem, ou seja, delineiam o vocabulário de acordo com uma atividade ou tarefa genérica. Por exemplo, a tarefa de realizar um diagnóstico, ou vender uma roupa.
- d) Ontologias de aplicação: relacionam-se com conceitos inerentes a um domínio e tarefa específica. Esses conceitos geralmente descrevem funções realizadas por entidades do domínio, que exercem uma determinada atividade. Por exemplo, uma ontologia para uma aplicação que trabalhe com carros de luxo.
- e) Ontologias de representação: explicam as conceituações que fundamentam os formalismos de representação de conhecimento. Por exemplo, ontologia de frames, utilizada em Ontolíngua.

Guarino (1998) sugere que ontologias sejam criadas de acordo com seu nível de generalidade, como é mostrado na Figura 2.4. Uma ontologia de tarefa ou de domínio são especializações dos termos introduzidos por uma ontologia genérica. Já conceitos de uma ontologia de aplicação devem ser especializações dos termos das ontologias de domínio e de tarefa correspondentes.

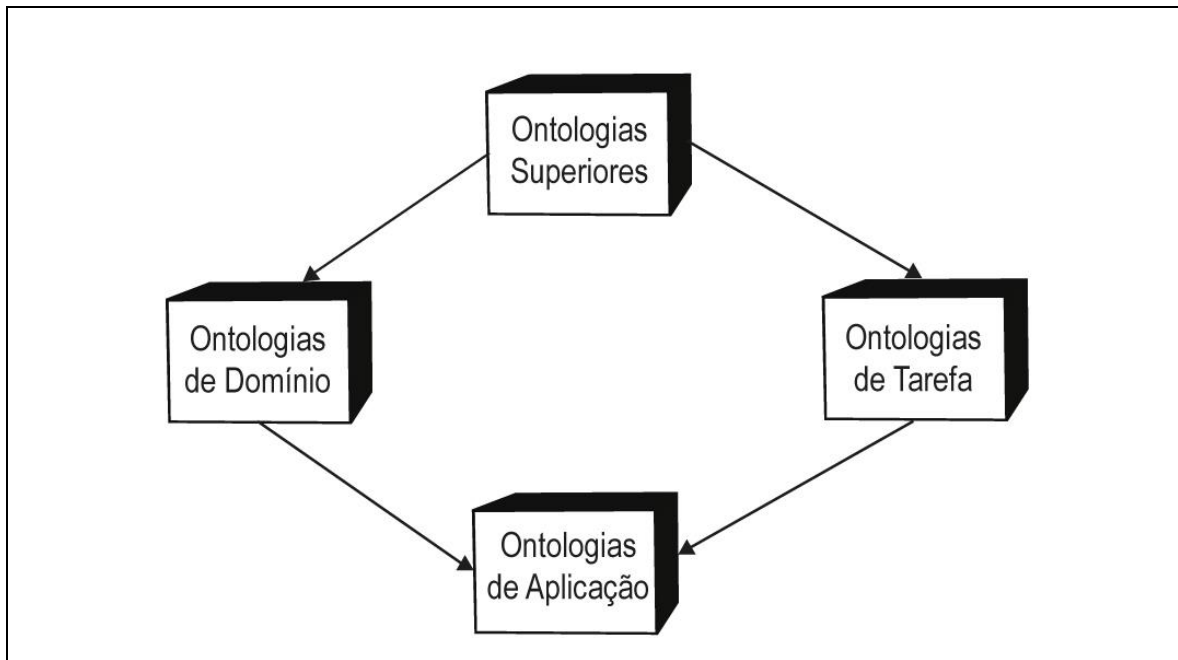


Figura 2.2 Tipos de Ontologia de acordo com o seu nível de dependência.
(Adpatado de Guarino, 1998.)

Segundo Guarino (1993) alguns princípios, se usados com precisão, garantem a qualidade da ontologia, quais sejam:

- a) Clareza: os programas usam diferentes modelos e abstrações na resolução de seus problemas. Na definição do conhecimento, deve-se ter a objetividade de definir apenas o que se presume ser útil na resolução da classe de problemas a ser atingida. Definições completas, com condições necessárias e suficientes, devem ter precedência sobre definições parciais.
- b) Legibilidade: as definições devem guardar correspondência com as definições correntes e informais. A ontologia deve usar um vocabulário normalmente compartilhável, com o jargão e terminologia usados por especialistas do domínio.
- c) Coerência: as inferências derivadas da ontologia definida devem ser corretas e consistentes do ponto de vista formal e informal.
- d) Extensibilidade: a ontologia deve permitir extensões e especializações monotonicamente e com coerência, sem a necessidade de revisão de teoria, que

consiste na revisão lógica automática de uma base de conhecimento em busca de contradições.

- e) Mínima codificação: devem ser especificados conceitos genéricos independentemente de padrões estabelecidos para mensuração, notação e codificação, garantindo a extensibilidade. Essa genericidade é limitada pela clareza. Tal característica se contrapõe à programação imperativa orientada a objetos, porque, por exemplo, atributos de objetos trazem tipos básicos e informações sobre a implementação.
- f) Mínimo compromisso ontológico: para maximizar o reuso, apenas o conhecimento essencial deve ser incluído, gerando a menor teoria possível acerca de cada conceito e permitindo a criação de conceitos novos, mais especializados ou estendidos.

2.3 LINGUAGENS DE REPRESENTAÇÃO DO CONHECIMENTO

2.3.1 XML

A Extensible Markup Language (XML) é um formato com base em texto simples para representar informações estruturadas: documentos, dados, configuração, livros, transações, faturas e muito mais. Foi derivada a partir de um formato padrão mais antigo denominado SGML (ISO 8879)⁹, a fim de ser mais apropriado para o uso da *Web*.

De acordo com Souza e Alvarenga (2004):

O XML é uma recomendação formal do W3C e, em determinados aspectos, assemelha-se ao HTML. Ambas são derivadas do SGML e contêm *tags* para descrever o conteúdo de um documento. Mas, enquanto o HTML tem como objetivo controlar a forma com que os dados serão exibidos, o XML se concentra na descrição dos dados que o documento contém. Além disso, o XML é flexível no sentido de que podem ser acrescentadas novas *tags* à medida que forem necessárias, bastando para isso que estejam descritas em um DTD específico; ou seja, qualquer comunidade de desenvolvedores pode criar suas marcações (*tags*) específicas que sirvam aos propósitos de descrição de seus dados. Isto possibilita que os dados sejam descritos com mais significado, abrindo caminho para

⁹ <http://www.w3.org/XML/9712-reports.html>

embutirmos semântica em documentos da *World Wide Web* e nas *intranets* (Souza e Alvarenga, 2004)

A W3C define os parâmetros de validade e formação adequada de sintaxe de um documento XML, esta é uma orientação clara de como obter êxito na composição do documento XML que pode ser encontrada na Extensible Markup Linguagem XML 1.0 (XML, 2004).

Na presente dissertação o XML deve ser entendido com linguagem base para o desenvolvimento da sintaxe RDF em seus diversos formatos.

2.3.2 Serialização RDF

Arquivos em RDF são adequados para representar dados em que é utilizada uma forma simples de expressar afirmações sobre esses recursos. A unidade de informação utilizada por esse modelo é chamada tripla, que é composta pelos seguintes recursos: sujeito, predicado (propriedade) e objeto, formando assim uma afirmação (*statement*). Para que um recurso seja identificado dentro do contexto apresentado pela *Web Semântica* é atribuído a ele um *Uniform Resource Identifier* (URI), ou seja, tanto sujeito, predicado e objeto estão associados a um identificador de recursos. A Figura 2.3 apresenta um exemplo de uma tripla utilizando o vocabulário da LOA.

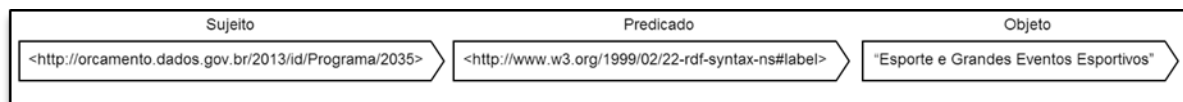


Figura 2.3 Exemplo de uma tripla RDF.

Um conjunto de triplas forma um grafo de modelo de dados, no qual uma tripla é representada por dois “nós” e a ligação entre eles. Portanto, um arquivo RDF é constituído por uma série de triplas interligadas, que resultam em afirmações a respeito dos relacionamentos entre recursos. Cada nó pode ser um URI que identifica um recurso, um dado do tipo Literal, ou até mesmo um nó em branco. A Figura 2.4 exemplifica tal estrutura.

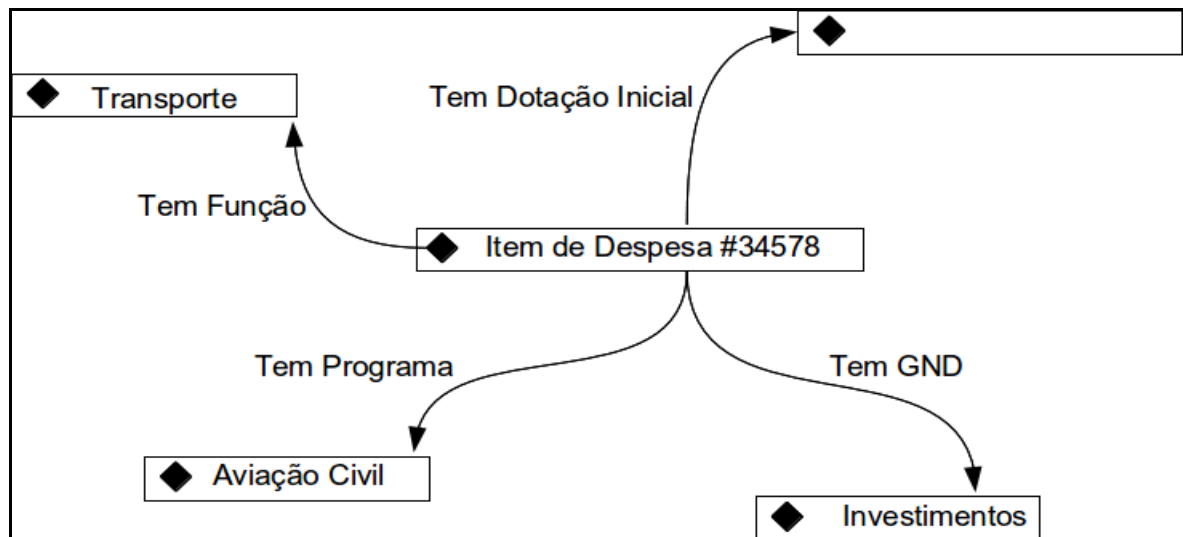


Figura 2.4 Exemplo de um grafo RDF.

As descrições RDF apresentadas nas Figuras 2.3 e 2.4 foram ilustradas por intermédio de gráficos. Para construir e disponibilizar um arquivo RDF, é necessário que ele esteja na forma textual. O RDF não é uma linguagem própria e, portanto, faz-se o uso de outras sintaxes para se expressar textualmente.

Existem outras sintaxes que facilitam a leitura e escrita de arquivos RDF serializados, dentre elas o *Turtle*, que é um formato mais completo, é ao mesmo tempo mais simples, no que tange a leitura e escrita por humanos, além de englobar as características da forma de serialização do formato *N-Triples*. Este *N-Triples* é uma versão simplificada do *Turtle*, pois não apresenta abreviações e uso de prefixos para URIs, aumentando a performance de seu processamento por máquinas e complicando a leitura e escrita humana.

2.3.2.1 *TURTLE*

O RDF não é uma linguagem própria, portanto faz-se o uso de outras sintaxes para se expressar textualmente. A linguagem mais utilizada para esse fim e recomendada pela W3C é o RDF/XML, na qual é realizada uma serialização de um grafo RDF em um documento no formato XML.

```

@prefix siop: <http://localhost/rdf-loa2012/ontologia/> .
@prefix siop-i: <http://localhost/rdf-loa2012/instancias/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

siop-i:item_de_despesa-10295    rdf:type                siop:ItemDeDespesa .
siop-i:item_de_despesa-10295    siop:temEsfera         siop-i:esfera-10 .
siop-i:item_de_despesa-10295    siop:temUnidadeOrcamentaria siop-i:unidade_orcamentaria-47101 .
siop-i:item_de_despesa-10295    siop:temFuncao        siop-i:funcao-04 .
siop-i:item_de_despesa-10295    siop:temSubfuncao     siop-i:subfuncao-126 .
siop-i:item_de_despesa-10295    siop:temPrograma      siop-i:programa-2038 .
siop-i:item_de_despesa-10295    siop:temAcao          siop-i:acao-4838 .
siop-i:item_de_despesa-10295    siop:temLocalizador   siop-i:localizador-48380001 .
siop-i:item_de_despesa-10295    siop:temElementoDespesa siop-i:elemento_despesa-00 .
siop-i:item_de_despesa-10295    siop:temFonteRecursos siop-i:fonte_recursos-100 .
siop-i:item_de_despesa-10295    siop:temGnd           siop-i:gnd-3 .
siop-i:item_de_despesa-10295    siop:temIndicadorUso  siop-i:indicador_uso-0 .
siop-i:item_de_despesa-10295    siop:temModalidadeAplicacao siop-i:modalidade_aplicacao-90 .
siop-i:item_de_despesa-10295    siop:temResultadoPrimario siop-i:resultado_primario-2 .
siop-i:item_de_despesa-10295    siop:temValorDotacaoInicial "1900000"^^xsd:double .

```

Figura 2.5 Exemplo de tripla do orçamento federal em RDF serializado no formato *Turtle*.

O RDF/XML¹⁰ apresenta restrições no uso de grafos e XML *Namespaces*, como alternativa a W3C recomenda o uso do *Turtle Terse RDF Triple Language – Turtle*¹¹, que apresenta a possibilidade de incluir *namespaces* em sua sintaxe, ou seja, é possível utilizar abreviações atribuindo nomes para URIs. Também são usados outros recursos para simplificar sua escrita, criando assim uma maneira compacta de representar triplas. Pelos motivos citados, a conversão RDF da LOA 2012 foi feita considerando o formato *Turtle*, que é uma das notações RDF mais simples. Nesse caso, os recursos são escritos diretamente com seu URI (entre < >) e também na ordem de sujeito, predicado e objeto, no qual um ponto final (.) encerra a tripla, conforme a Figura 2.5. No entanto a sintaxe não se mostrou eficiente para o processamento de grandes volumes de dados, principalmente os da execução orçamentária, que são atualizados diariamente.

2.3.2.2 N-TRIPLES

O *N-Triples*¹² é um formato de armazenamento e transmissão de dados, padronizado pelo W3C, que foi desenvolvido pelo Grupo de Trabalho RDFCore W3C¹³. É um formato com base em linha, texto simples, serialização de RDF (*Resource Description*

¹⁰ <http://www.w3.org/TR/REC-rdf-syntax/>

¹¹ <http://www.w3.org/TR/2014/REC-turtle-20140225/>

¹² <http://www.w3.org/TR/2014/REC-n-triples-20140225/>

¹³ <http://www.w3.org/2001/sw/RDFCore/Overview.html>

Framework) gráficos, e um subconjunto da *Turtle* (*Terse RDF Triplo Language*) formato. *N-Triples* não deve ser confundida com a notação 3 (três) que é um super conjunto de *Turtle*.

O *N-Triples* foi projetado para ser um formato mais simples do que a notação 3 e o *Turtle* e, portanto, um *software* mais fácil para analisar e gerar. No entanto, o *N-Triples* carece de alguns dos atalhos fornecidos por outras serializações RDF, podendo ser oneroso para digitar grandes quantidades de dados à mão, além disso, é difícil de ler.

Para o presente trabalho, foi adotado o padrão RDF na sintaxe *N-Triples*, uma vez que esse formato é um modelo recomendado pela W3C para descrever recursos encontrados na *Web*, e o processo de conversão dos dados orçamentários em um arquivo RDF apresentam um desempenho superior no formato *N-Triples*.

2.3.3 Vocabulários

Para Sayão e Marcondes (2008) a interoperabilidade tem sido um item crítico no desenvolvimento e operação de sistemas de repositórios e de bibliotecas digitais distribuídos funcionando em rede.

De acordo com Sayão e Marcondes (2008):

Para diferentes conteúdos só poderão ser integrados e reusados, no sentido de terem aproveitadas as sinergias uns dos outros, se estiverem ancorados por sistemas que permitam um alto grau de interoperabilidade (Sayão e Marcondes, 2008)

Existem algumas iniciativas que têm como objetivo munir publicadores de dados com vocabulários e ontologias, com a finalidade de aumentar a qualidade dos dados publicados. Dentre essas iniciativas, destacam-se os projetos: *Friend of a Friend* (FOAF) que tem o objetivo de descrever pessoas e os relacionamentos que a cercam; o projeto *Description of a Project* (DOAP)¹⁴ que visa descrever projetos, especialmente os de código aberto; o *Simple Knowledge Organization System* (SKOS)¹⁵; e a própria

¹⁴ *Description of a Project* (DOAP) é um vocabulário RDF Schema e XML para descrever projetos de *software*, em especial, *software* livre e de código aberto.

¹⁵ www.w3.org/2004/02/skos/intro

DBpedia indiretamente.

No tema de Dados Abertos, a Infraestrutura Nacional de Dados Abertos (INDA)¹⁶ recomenda a utilização do Vocabulário Controlado do Governo Eletrônico (VCGE)¹⁷ para atribuição de semântica aos dados governamentais publicados no Portal dados.gov.br. No entanto, o VCGE não possui relações semânticas entre as classes, além da relação gênero/espécie.

2.3.4 Linguagens Ontológicas

Ontologias são construídas de formas diferentes, vários tipos de linguagens podem ser usadas, incluindo linguagens de programação lógica, como o Prolog¹⁸. O mais comum, no entanto, são linguagens que têm sido especificamente construídas para dar suporte à ontologia. O Modelo *Open Knowledge Base Connectivity* (OKBC)¹⁹ e a linguagem *Knowledge Interchange Format* (KIF)²⁰ (e o sucessor *Common Logic* –(CL)²¹ são exemplos que se tornaram base para outras linguagens ontológicas.

Várias linguagens têm como base o pensamento lógico, na computação são conhecidas como linguagens de lógica descritiva (Franconi, 2000).

Dentre essas linguagens incluem-se:

- RDF: é uma linguagem gráfica utilizada para representar informações sobre fontes na *Internet*, é uma linguagem ontológica básica. As fontes são descritas em termos do uso das propriedades e dos valores das propriedades, utilizando expressões

¹⁶ <http://www.governoeletronico.gov.br/acoes-e-projetos/Dados-Abertos/inda-infraestrutura-nacional-de-dados-abertos>

¹⁷ <http://www.governoeletronico.gov.br/acoes-e-projetos/e-ping-padrees-de-interoperabilidade/vcge>

¹⁸ Prolog é uma linguagem de programação que se enquadra no paradigma de Programação em Lógica Matemática. É uma linguagem de uso geral que é especialmente associada com a inteligência artificial e linguística computacional.

¹⁹ *Open Knowledge Base Connectivity* (OKBC) é um protocolo é uma API para acesso ao conhecimento em sistemas de representação do conhecimento, tais como repositórios de ontologias e bancos de dados objeto-relacional.

²⁰ *Knowledge Interchange Format* (KIF) é uma linguagem orientada para o intercâmbio de conhecimento entre os diferentes programas de computador.

²¹ *Common Logic* (CL) é uma estrutura para uma família de linguagens lógicas, com base na lógica de primeira ordem, que visa facilitar a troca e transmissão de conhecimento em sistemas fundamentados em computador.

RDF. As expressões são representadas como triplas, que consistem de um sujeito, predicado e objeto (S, P, O). O RDF é escrito em XML e utiliza URIs (*Unique Resource Identifiers*) para identificar fontes. Em relação aos seus pontos mais fracos, o RDF usa apenas propriedades binárias. Essa restrição parece séria, pois muitas vezes usamos predicados com mais de dois argumentos. Outro problema com o RDF é relativo à manipulação de propriedades. O RDF *Schema* estende o RDF semanticamente para permitir usar classes de fontes e as propriedades que serão usadas com elas. O RDF e o RDF *Schema* possuem capacidades básicas para descrever vocabulários que descrevem fontes. No entanto, algumas outras características são desejáveis, tais como:

- restrição de cardinalidade;
 - especificação de que propriedades são transitivas;
 - especificação de propriedades inversas;
 - especificação de cardinalidades e/ou propriedades, quando utilizado com uma determinada classe;
 - capacidade de descrever novas classes, combinando classes existentes (utilizando interseções e sindicatos) e com uso da negação.
- OWL (Dean e Schieber, 2009): é construída e tem como base em RDF (OWL estende RDF(S) semanticamente), e seu antecessor à linguagem DAML + OIL. OWL é uma linguagem de ontologia que foi feita primariamente para descrever e definir classes. As classes são, portanto, os blocos de construção básicos de uma ontologia OWL, a qual possui um rico conjunto de construtores para modelagem, que permite usabilidade por vários usuários.

A OWL foi proposta como padrão pelo W3C agregando diversos pontos positivos das linguagens anteriores e está sendo utilizada pela *WEB Semântica*. Ela está subdividida em três sublinguagens: OWL *Lite*²², OWL DL²³ e OWL *Full*²⁴.

Criadas para comunidades específicas de usuários, a OWL *Lite* dá suporte principalmente aos usuários que precisam de uma classificação hierárquica e restrições simples. Por exemplo, ele suporta restrições de cardinalidade podendo só permitir valores de 0 ou 1. Ainda, proporciona um rápido caminho de migração para léxicos e outras classificações, além disso, tem também uma menor complexidade formal do que a OWL DL.

A OWL DL inclui todos os construtores da linguagem OWL, mas eles só podem ser utilizados sob certas restrições. Por exemplo, enquanto uma classe pode ser uma subclasse de muitas classes, uma classe não pode ser uma instância de outra classe.

Já a OWL *Full*, possibilita alta expressividade e a liberdade sintática de RDF, sem garantias computacionais. Uma classe OWL *Full* pode tratar simultaneamente um conjunto de indivíduos e um único indivíduo.

Devido ao fato de o projeto necessitar basicamente de uma hierarquia de classificação e funcionalidades com restrições simples, a linguagem OWL *Lite* foi escolhida por não apresentar grande complexidade para a implementação de suas funcionalidades.

2.4 CONCEITOS DE INTEROPERABILIDADE

Existem diversas estratégias diferentes para promover a interoperabilidade entre sistemas. De acordo com o foco deste trabalho serão considerados os conceitos básicos da Interoperabilidade Semântica.

²² OWL *Lite* é um subconjunto da linguagem OWL (*Web Ontology Language*) que é usada para definir e instanciar ontologias na *Web*; dá suporte principalmente aos usuários que precisam de uma classificação hierárquica e restrições simples.

²³ OWL DL é um subconjunto da linguagem OWL (*Web Ontology Language*) que é usada para definir e instanciar ontologias na *Web*; dá suporte aqueles usuários que querem a máxima expressividade.

²⁴ OWL *Full* é um subconjunto da linguagem OWL (*Web Ontology Language*) que é usada para definir e instanciar ontologias na *Web*; dá suporte àqueles usuários que querem a máxima expressividade e a liberdade sintática do RDF sem nenhuma garantia computacional.

2.4.1 Interoperabilidade Ontológica

Interoperabilidade, tal como definido pelo IEEE (2000), é “a capacidade de dois ou mais sistemas ou componentes trocar informações e usar as informações que foram trocadas”. Um sistema (ou componente) pode interagir com qualquer outro sistema (ou componente), desde que eles possam entender as informações que foram trocadas entre si. Brodie, 1992 (*apud* Buranarach, 2004) dá uma definição funcional de interoperabilidade em sistemas de informação, qual seja:

Interoperabilidade: Dois componentes (ou objetos) X e Y podem interoperar (são interoperáveis) se X pode enviar pedidos de serviços (ou mensagens) de R para Y com base em um mútuo entendimento de R por X e Y, e Y pode retornar respostas S a X fundamentado numa mútua compreensão de como S (respectivamente) respostas a R por X e Y. Em suma, o sistema X pode interoperar com o sistema Y se os pedidos de X podem ser respondidos apropriadamente por Y.

Para que as informações disponíveis na *Web* sejam utilizadas pelos diferentes sistemas, é necessário, em um primeiro momento, que essas sejam localizadas, acessadas, compreendidas e processadas pelos mesmos. Em um segundo momento, devido à heterogeneidade do conteúdo dessas informações, que podem ser classificadas segundo os aspectos estrutural, sintático e semântico (Sheth, 1999), é preciso compatibilizar esses conteúdos entre si.

2.4.2 Web Semântica e *Linked Data*

Apesar da grande evolução das atuais interfaces de consulta, a taxa de recuperação da informação é afetada pela grande quantidade de dados não estruturados cuja indexação é dificultada, o que atrapalha dos métodos de classificação de conteúdo ou *ranking* pelos quais os conteúdos são recuperados de acordo com sua relevância.

Atenta a esta problemática, a W3C – *World Wide Web Consortium* – (*SemanticWeb*, 2004), através de seus pesquisadores e parceiros, propôs à *Web Semântica* o objetivo de estruturar o conteúdo das páginas *Web*, de forma que a informação possa ser interpretada pelas máquinas.

A proposta não foi a de uma *Web* separada da atual, mas sim uma extensão da mesma, com base em documentos – as ontologias – descrevendo relacionamentos entre objetos e

contendo informação semântica dos mesmos para automatizar o processamento pelas máquinas.

Segundo Antoniou e Harmelen (2008), a *Web Semântica* e *Linked Data* nascem do conceito de compartilhamento de informações, lembrando um modelo informacional anárquico.

As ontologias são utilizadas quando se necessita de um vocabulário compartilhado para troca de informações e permite a interoperabilidade semântica entre conceitos, habilitando o computador a processar o dado de acordo com sua semântica, e não apenas de acordo com o conteúdo de seu texto (Allemang e Vender, 2011).

Na Figura 2.6, temos a representação da “pilha” da *Web Semântica*. Observa-se que na base da pilha encontra-se o Unicode, que permite a representação de alfabetos modernos da escala mundial, e o *Uniform Resource Identifier* (URI), que permite a identificação de recursos na *Web*.

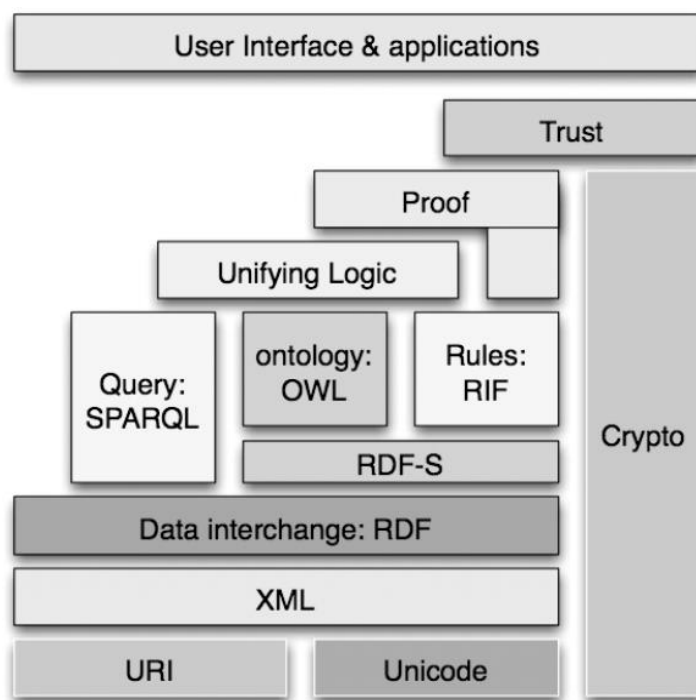


Figura 2.6 Pilha de Normalizações da *Web Semântica*.
(Bratt, 2006)

A *Extensible Markup Language* (XML) proposta na Figura 2.6, permite a interoperabilidade entre diferentes aplicações que partilhem dados com formatos distintos, além do mecanismo de *namespaces* que traz a desambiguação de termos iguais. O RDF compreende as especificações do modelo e sua sintaxe, possibilitando a descrição dos recursos por meio de suas propriedades e valores. As informações em RDF são representadas na forma de triplas sujeito-predicado-objeto. O sujeito e o objeto são os recursos que se relacionam por meio da propriedade (ou predicado).

De acordo com Antoniou e Harmelen (2008) O RDF Sema (RDF-S) é uma extensão semântica do RDF, utilizado para a descrição do vocabulário RDF, possibilitando a definição de taxonomias de recursos em termos de uma hierarquia de classes, criando uma ontologia mais rica. Ainda diretamente acima do RDF, temos a *Simple Protocol and. RDF Query Language* (SPARQL)²⁵. Assim como a *Structured Query Language* (SQL)²⁶ permite a consulta e modificação das bases de dados relacionais, a SPARQL permite a realização de consultas em bases de dados no formato RDF (Antoniou e Harmelen, 2008). Essa linguagem tornou-se recomendação da W3C e muito se assemelha com o SQL.

A linguagem de programação de ontologias indicada pela W3C é a *Web Ontology Language* (OWL). Essa linguagem estende o RDF-S e, conseqüentemente, o RDF. Segundo Antoniou e Harmelen, (2008), a linguagem OWL permite descrever formalmente, de modo mais eficiente, os aspectos semânticos dos termos utilizados e seus respectivos relacionamentos, possibilitando representações mais abrangentes das linguagens RDF e RDF *Schema* e favorecendo uma maior interoperabilidade. Na seção 3.1, ela será melhor detalhada.

A camada ontológica é formada pela OWL que é uma linguagem com base em regras. As regras (*rules*) permitem uma espécie de “introdução lógica” que será incorporada a *Unifying Logic*, possibilitando o uso de lógicas avançadas, em que agentes computacionais possam realizar inferências automáticas a partir das relações existentes entre os recursos informacionais, podendo, inclusive, inferir novas informações (Antoniou e Harmelen, 2008).

²⁵ <http://www.w3.org/TR/rdf-sparql-query/>

²⁶ <http://en.wikipedia.org/wiki/SQL>

A camada *Proof* possibilita a verificação/comprovação da coerência lógica dos recursos e a camada *Trust* garante que as informações estejam representadas de modo correto, possibilitando a confiabilidade das informações. Segundo Berners-Lee (2001), na camada *Trust* as assinaturas digitais funcionam como um mecanismo de prevenção de inconsistências nessa *Web Semântica*

Dessa maneira, não só humanos, mas principalmente as máquinas, poderão processar a informação de forma muito mais eficaz e eficiente. Essa evolução possibilitará tanto uma maior comunicação, i.e., troca de informações entre as aplicações dessa nova *Web* quanto, conseqüentemente, a interoperabilidade semântica entre elas.

De acordo com Bizer *et al.* (2007), a proposta de dados abertos interligados oferece grande potencial ao conectar recursos informacionais por meio de *links* semânticos, os quais são significativos também para programas. Ao contrário, *links* convencionais nada mais são (além de uma eventual etiqueta textual significativa para usuários humanos) do que meios para que programas navegadores, a partir de um recurso, acessem outro, sem explicitar qual o significado da ligação entre os recursos. Sendo significativos para programas, *links* semânticos podem ser processados de forma mais rica por eles, explorando e enriquecendo cognitivamente o significado (legível por máquina) da ligação entre ambos os recursos.

Segundo Isele *et al.* (2010), as tecnologias de dados interligados não são propriamente tecnologias de recuperação de informações, pois o principal objetivo, após a recuperação de um recurso informacional, é interligá-lo com outros que lhe agreguem valor semântico, cultural, cognitivo, econômico ou científico. A dimensão atual da proposta de dados interligados mostrando as instituições e seus recursos de informação já envolvidos é comentada por Bizer *et al.* (2007).

A proposta da *Web Semântica* descreve uma *Web* qualitativamente diferente da atual, na qual, ao invés dos conteúdos serem estruturados para serem exibidos e lidos por pessoas, terão uma semântica “inteligível” por programas, os chamados “agentes de *software*”; ou seja, programas que não sejam criados especificamente para processarem somente um tipo determinado de dado. As tecnologias da *Web Semântica* são a evolução da *Web* atual em direção às mais amplas possibilidades de interação e de recuperação semântica da

informação. A *Web* atual é chamada de *Web* Sintática pelos formuladores do projeto da *Web* Semântica por trabalhar apenas com padrões sintáticos de forma. Uma busca por conteúdo é processada por programas como uma busca de padrões de textos idênticos aos fornecidos como parâmetro, que são, em última instância, meros padrões de *bits* acesos ou apagados.

As diferentes linguagens-padrões da *Web* Semântica para estruturação de conteúdos são apresentadas por Owl (2004) por ordem crescente de “expressividade semântica”, ou seja, dispõe de um vocabulário capaz de expressar um leque cada vez mais amplo de significados:

- XML - sintaxe para estruturar documentos. Não impõe restrições semânticas.
- XML *Schema* - restringe a estrutura de documentos XML.
- RDF - modelo de dados para objetos e relacionamentos. Semântica simplificada.
- RDF *Schema* - vocabulário para descrever propriedades e classes expressas em RDF. Semântica para generalização.
- OWL - vocabulário para descrever classes e propriedades: relacionamentos (ex. disjunção), cardinalidade (ex. exatamente um), igualdade, tipos mais expressivos de propriedades, características de propriedades (ex. simetria) e enumeração dos elementos de uma classe.

Segundo Berners-Lee *et al.* (2001), o padrão RDF seria fundamental e marcaria um ponto de inflexão na sequência de linguagens apresentada, pois ali estaria o primeiro “lampejo” do que seria uma semântica inteligível por programas. O RDF é justamente o padrão básico da proposta de dados abertos interligados, por intermédio do qual recursos são descritos por meio de uma ou mais assertivas sobre eles; conteúdos são publicados de forma estruturada enquanto uns e outros são interconectados por meio de *links* diferentes dos convencionais, que expressam a natureza ou semântica da ligação entre os recursos sendo ligados (Bizer *et al.*, 2007). A semântica inteligível por programas se daria porque o RDF seria, no ordenamento das linguagens da *Web* mencionado, a primeira linguagem (a

segunda seria a OWL) que pressupõe uma semântica; essa é composta por sujeito, predicado e objeto de uma assertiva, os elementos de seu metamodelo (Guizzardi, 2005).

2.4.3 *Linked Open Data - LOD*

Dado o embasamento sobre os principais conceitos e padrões envolvidos, pode-se, finalmente, falar sobre Dados Abertos Interligados (*Linked Open Data - LOD*)²⁷, o qual pode ser encarado como um direcionamento, uma abordagem para representação e interligação de dados. É um conjunto de diretrizes sobre como publicar dados na *internet* e, por consequência, como relacioná-los entre si, criando algo semelhante a um enorme banco de dados sobre tudo que há na *Web*: pessoas, objetos, organizações etc. É comum a referência, nesse contexto, à “nuvem” de LOD.

Berners-Lee (2006) é o principal mentor da iniciativa LOD. Ele sustenta que para LOD funcionar são necessárias quatro regras (ou “expectativas de comportamento”, como o mesmo prefere chamar). São elas:

- 1^a) uso de URIs para identificar coisas, conceitos, objetos, relacionamentos e tudo mais;
- 2^a) essas URIs devem ser HTTP, de modo que sejam referenciáveis para que possam ser buscadas por máquinas e pessoas;
- 3^a) informação útil deve ser provida a URI do objeto buscado, usando formatos padrões (RDF, OWL etc);
- 4^a) relacionamentos devem ser colocados para outras URIs que tenham alguma relação com a URI buscada.

A primeira regra apenas diz que tudo deve ter um identificador único global, referenciando-se, assim, um único objeto. Dada uma URI, uma e apenas uma entidade ou relacionamento (por exemplo) deve ser retornado.

²⁷ <http://www.linkeddata.org>.

A segunda, complementar à primeira, diz que as URIs devem ser do tipo HTTP, pois elas devem ser buscadas via navegadores de documentos e de dados por humanos ou máquinas (agentes).

A terceira fala sobre padronização desses dados. Imagine se cada empresa, pessoas ou entidade que gerasse dados na *Web* o fizesse num formato diferente. Seria quase impossível agregar e relacionar os dados, visto que cada formato teria sua sintaxe e sua semântica.

A quarta é um apelo para que relacionamentos sejam criados, já que o principal motivador de LOD é exatamente a conectividade entre os dados, para que os mesmos possam ser acessados de diversas origens diferentes. Esses relacionamentos podem usar vocabulários específicos (FOAF, DBPROP etc), podem indicar que existe outro endereço para o mesmo recurso, utilizando a expressão “rdf:sameAs”, entre outros diversos tipos.

Por mais simples que as regras sejam, inicialmente, em 2006, vários dados, apesar de publicados, não tinham conexões relevantes com outros dados. Um tanto quanto indignado com isso e preocupado com o futuro da iniciativa, Berners-Lee (2009) simplificou ainda mais as regras em sua apresentação na conferência global Technology, Entertainment, Design - TED²⁸ em 2009, reformulando as quatro regras em três:

- 1^a) todos os tipos de coisas conceituais na *Web* têm “nomes” que começam com HTTP;
- 2^a) quando buscar por uma URI, deve-se receber um conjunto de informações relevantes, ou seja, informações de pessoas que procuram por aquilo que queiram saber;
- 3^a) quando receber informações, não serão apenas altura e largura de alguma coisa: tem que haver relacionamentos! E os mesmos devem, também, começar com HTTP (URI).

Essas “novas” regras não são tão novas assim: elas só enfatizam as ideias por trás da iniciativa LOD. Dados compartilhados na *Web* devem ter relacionamentos e informações relevantes.

²⁸ <http://www.ted.com/>

2.4.4 URI

Usada como identificador de recursos em redes, mais especificamente a *Internet*, uma *Uniform Resource Identifier* (URI)²⁹ pode ser classificada como *Uniform Resource Name* (URN)³⁰, *Uniform Resource Locator* (URL)³¹ ou uma interseção das duas últimas. As URNs funcionam como o nome de uma pessoa, enquanto a URL faz o papel de endereço dessa pessoa.

Um exemplo típico de URN é o sistema do Número Padrão Internacional de Livro (*International Standard Book Number* - ISBN)³². Nesse padrão que, em 1972, se tornou na norma ISO 2108 para identificação de livros, é utilizado o seguinte formato: (urn:isbn:0-486-27557-4 que cita, sem ambiguidade, uma edição específica de *Romeo e Julieta* de Shakespeare). Para acessar esse livro, deve-se prover o endereço URL do recurso, que pode ser, por exemplo, “file:///home/username/RomeoAndJuliet.pdf”.

De forma geral, uma URI é formada por um identificador de protocolo (“http”, “ftp”, “file”, entre outros), o caractere “:” e o caminho para achar o recurso.

Por fim, uma URI pode ser absoluta e relativa. Absoluta é quando um recurso é referenciado não considerando o local no qual o mesmo está sendo referenciado (http://www.google.com). Relativa, ao contrário, leva em consideração o contexto no qual aquele recurso é referenciado (“.././include.js”, por exemplo).

Para o escopo deste trabalho, URI são identificadores globalmente únicos que são usados para referenciar um determinado objeto da *Internet*, seja conceito, documento, propriedade etc. Essas URIs são do tipo absoluta e HTTP (começam com HTTP). Conforme exemplo a seguir retirado do e-PING (2013):

<http://estruturaorganizacional.dados.gov.br/doc/idade-organizacional/completa.xml?codigounidade=1232> (consulta coletiva, usando a {funcao} completa, especificando o formato xml, e com o parâmetro “codigo-unidade com o valor 1232). Obs.: é uma consulta coletiva, mesmo que retorne apenas um resultado, pois foi utilizado um *slug* do tipo {funcao} e não {identificador}.)

²⁹ <http://www.w3.org/Addressing/URL/uri-spec.html>

³⁰ <http://www.ietf.org/rfc/rfc2141.txt>

³¹ <http://www.w3.org/Addressing/>

³² <http://www.isbn.bn.br/>

2.4.5 SPARQL

De acordo com Bizer *et al.* (2007) os endpoints SPARQL podem ser classificados como genéricos ou específicos. Um endpoint genérico funciona sobre qualquer base de dados RDF, que pode ser armazenada localmente ou acessada pela web. Um *endpoint* específico está ligado à apenas uma base de dados, que não pode estar conectada com outros endpoints.

Um *endpoint* SPARQL³³ pode ser configurado para retornar resultados em diferentes formatos, por exemplo, uma tabela no formato HTML, construída através de uma transformação XSL de um arquivo XHTML resultante da consulta realizada. Quando acessado por alguma aplicação, os resultados são serializados em formatos RDF processáveis por máquinas, como por exemplo, RDF/XML, *Turtle* (Terse RDF Triple Language) e *N-Triples* (W3C RDF Core WG Internal Working Draft).

Para formular uma consulta SPARQL, é necessário informar dentro da cláusula WHERE quais triplas devem ser analisadas e selecionar na cláusula SELECT dos recursos a serem recuperados, possibilitando inclusão de filtros para a pesquisa. Para representar os recursos nas triplas que estão sendo consultadas, são utilizadas variáveis de consulta, as quais são identificadas por serem termos quaisquer, antecedidos por um símbolo de interrogação (?).

Nesse sentido, visando a disponibilização aberta para consultas, são utilizados pontos de acesso SPARQL (*SPARQL endpoints*), que são interfaces que podem ser utilizadas por usuários ou aplicações em contratos de serviços para realizar consultas em uma base de dados RDF utilizando a linguagem SPARQL. Sua função é servir como ponto de acesso aos dados, recebendo uma requisição de uma consulta e ativando um motor de consulta que acessa o *triple store*, retornando o resultado. O *endpoint* pode ter como *front-end* uma aplicação *Web* ou *stand-alone*.

O *endpoint* pode ser acessado por intermédio de uma URL, na qual é disponibilizada uma interface para que a consulta seja informada e posteriormente os resultados exibidos. Se o usuário for uma aplicação, o acesso ao *endpoint* é realizado por alguma API especializada,

³³ <http://www.w3.org/TR/rdf-sparql-query/>

como, por exemplo, no caso desta dissertação, a biblioteca *Jena* de desenvolvimento de aplicações para a *Web Semântica* em Java.

2.4.6 Dados Governamentais Abertos

Segundo a W3C Brasil (2009), Dados Abertos Governamentais é uma metodologia para a publicação de dados do governo em formatos reutilizáveis, visando ao aumento da transparência e maior participação política por parte do cidadão, além de gerar diversas aplicações desenvolvidas colaborativamente pela sociedade.

Nascido em 2009, o *Open Data*, como é conhecido em inglês, movimentou comunidades em todo o mundo para cobrar informações públicas dos governos. Dessa forma os cidadãos podem colaborar com os processos de governo e com o controle social das políticas públicas.

No caso do Brasil, vários órgãos da Administração Pública já publicam seus dados na *web*, na forma de relatórios e balanços para que os cidadãos possam acompanhar os resultados das ações de governo.

O conceito denominado “governo aberto” é o da disponibilização, por intermédio da *Internet*, de informações e dados governamentais de domínio público para a livre utilização pela sociedade. É parte integrante do conceito que, à sociedade, seja garantido acesso aos dados primários, de forma que se possa combiná-los, cruzá-los e, enfim, produzir novas informações e aplicações, colaborando com o governo na geração de conhecimento social a partir das bases governamentais (Lathrop e Ruma, 2010).

Esse sistema segue o princípio que assegura o direito de acesso irrestrito aos documentos, dados e informações no âmbito da administração pública, salvo nas hipóteses de sigilo previstas na Constituição e legislação em vigor, e ainda ressalvadas as exceções disciplinadas por regulamentação específica do órgão ou entidade da administração pública.

Os órgãos e entidades da administração possuem inúmeras bases que contêm dados e informações de caráter público (não sigilosos) e cujo acesso, quando disponível, normalmente é feito de forma parcial e por caminhos extremamente burocratizados e

pouco transparentes. Certamente, muitos dados e informações estão disponíveis nos portais e páginas que os órgãos e entidades mantêm na *Internet*, mas esse acesso normalmente é parcial e fragmentado, o que impede e dificulta que o interessado possa trabalhar, analisar, cruzar e integrar os dados e informações segundo o foco e os interesses próprios. Praticamente inexistem, salvo raríssimas exceções, a disponibilização de bases de dados e informações de forma eletrônica e por meio de padrões abertos.

Disponibilizar essas bases de dados e informações, não sigilosas, para livre acesso do cidadão interessado de forma eletrônica e por meio de padrões abertos, vai ao encontro dos objetivos expressos na Lei de Acesso à Informação (Brasil, 2011), que estabelece as diretrizes e políticas para gestão do conhecimento e inovação no âmbito da administração pública no que diz respeito.

2.4.7 Experiências de Dados Governamentais Abertos

O Brasil a exemplo de outros países já disponibiliza os seus dados a partir das premissas do governo aberto. Destacam-se entre os países que já possuem políticas de dados abertos:: Estados Unidos, Reino Unido, Austrália e Nova Zelândia (Agune *et al.*, 2010). Esses países criaram políticas nacionais de acesso aos dados públicos na *Web* e criaram portais que disponibilizam as informações públicas em diferentes formatos. Além disso, muitas vezes eles incentivam os cidadãos a criarem novos aplicativos e modos de utilizarem a informação publicada em seus portais.

O governo da Austrália criou o portal Data Austrália³⁴ em 2010, com objetivo de ser o banco de dados de informações públicas governamentais do país. A proposta principal foi encorajar a população a utilizar as informações disponibilizadas pelo portal para desenvolvimento de *Ganges* e outros *softwares* visando o controle social de todas as informações armazenadas pelo portal.

³⁴ <https://data.govt.nz/>

O portal do governo da Nova Zelândia³⁵ foi criado para ofertar dados e informações das agências e órgãos públicos do governo. Ele possui espaço para discussão por intermédio de fóruns que podem ser propostos pelos cidadãos.

O exemplo do Reino Unido pretende ofertar dados e informações da gestão pública aos cidadãos para ampliação do controle social do governo. Possui espaço para *wiki*, comunidade virtual, envio de novas aplicações *Web* e RSS.

Nos Estados Unidos o presidente Barack Obama foi fundamental na criação de políticas públicas para a promoção da transparência que incentiva a disponibilização dos dados públicos em formato aberto.

Através do memorando *Transparency and Open Government* (Obama, 2009) enviado a todos os chefes de governo, o governo norte americano, firmava o seguinte compromisso:

criar níveis sem precedentes de abertura no governo. O presidente convoca os representantes do Poder Executivo a trabalhar para assegurar a confiança e estabelecerem um sistema de transparência, participação pública e colaboração. A abertura visa fortalecer a democracia, promover a eficiência e dar efetividade ao governo (Silva, 2010).

A política *Open Government Initiative*, visou garantir a efetivação desse memorando. Entre as suas ações está o portal *Data.gov*³⁶, criado em maio de 2009. O objetivo do *Data.gov*, portal da transparência do governo federal, é ofertar informações e expandir o uso criativo dos dados governamentais por meio de aplicações *Web*, encorajando ideias inovadoras para oferecer maior transparência das ações governamentais. O portal possui informações sobre centenas de agências e instâncias governamentais separadas por órgão. Informa estatísticas de visitantes mensais, diários, anuais. Também possui espaço para participação dos cidadãos que desejam expressar suas vontades.

Além das experiências governamentais, já existem organismos não governamentais que visam promover e apoiar a adoção da política de dados abertos. Por exemplo, a Fundação *Sunlight* e o Projeto *Open Government Data*³⁷, da *Access Info* e Fundação *Open*

³⁵ <https://data.govt.nz/>

³⁶ <http://Data.gov>

³⁷ <http://opengovernmentdata.org/>

*Knowledge*³⁸. Essas experiências advogam em favor da adoção da política de dados abertos por governos e realizam diversas atividades desde pesquisa até incentivo à criação de soluções e padrões para formatos abertos. Em suma, apesar da discussão e da existência recentes de disponibilização de dados abertos pelo governo, já existem exemplos de governos que adotam essa teoria como política pública de promoção da transparência.

³⁸ <https://okfn.org/>

3 - FERRAMENTAS UTILIZADAS

3.1 FERRAMENTAS PARA CONTRUÇÃO DE ONTOLOGIAS

Como afirma Campos (2001), “um modelo conceitual deve ser visto, também, como um espaço comunicacional em que transpomos o mundo fenomenal para um espaço de representação”. Para tanto, algumas ferramentas têm sido utilizadas como auxílio no desenvolvimento de ontologias, garantindo o projeto lógico da modelagem, ao passo que possibilitam a visualização dos conceitos e das relações e, muitas vezes, permitem a interação do usuário com o mundo representacional por meio de uma interface gráfica. Segue, na Tabela 3.1, uma breve descrição de algumas ferramentas:

Tabela 3.1 Ferramentas para construção, uso e edição de ontologias.

Fonte: Almeida e Bax (2003).

Ferramentas	Breve descrição
GKB-Editor (<i>Generic Knowledge Base Editor</i>)	Ferramenta para navegação e edição de ontologias por meio de sistemas de representação com basea em <i>frames</i> . Oferece interface gráfica, em que os usuários podem editar diretamente a base de conhecimento e selecionar a parte que é de seu interesse.
JOE (<i>Java Ontology Editor</i>)	Ferramenta para construção e visualização de ontologias. Proporciona gerenciamento do conhecimento em ambientes abertos, heterogêneos e com diversos usuários. As ontologias são visualizadas como um diagrama entidade-relacionamento, como o gerenciador de arquivos do <i>MS Windows</i> ou como uma estrutura em árvore.
<i>OntoEdit</i>	É um ambiente gráfico para edição de ontologias que permite inspeção, navegação, codificação e alteração de ontologias. O modelo conceitual é armazenado usando um modelo de ontologia que pode ser mapeado em diferentes linguagens de representação. As ontologias são armazenadas em bancos relacionais e podem ser implementadas em XML, <i>FLogic</i> , RDF(S) e DAML+OIL.
<i>Protégé</i>	Ambiente interativo para projeto de ontologias, de código aberto, que oferece uma interface gráfica para edição de ontologias e uma arquitetura para a criação de ferramentas com base em conhecimento. A Arquitetura é modulada e permite a inserção de novos recursos.
<i>WebODE</i>	Ambiente para engenharia ontológica que dá suporte à maioria das atividades de desenvolvimento de ontologias. A integração com outros sistemas é possível, importando e exportando ontologias de linguagens de marcação.

WebOnto	Ferramenta que possibilita a navegação, criação e edição de ontologias, representadas na linguagem de modelagem OCML. Permite o gerenciamento de ontologias por interface gráfica, inspeção de elementos, verificação da consistência da herança e trabalho cooperativo. Possui uma biblioteca com mais de cem ontologias.
---------	--

Destaca-se que a interface gráfica da ferramenta *Protégé*³⁹ permite o acesso a todas as suas funções por meio de abas para edição. Um projeto desenvolvido no *Protégé* integra a modelagem de classes que descrevem um assunto particular, a criação de uma ferramenta de aquisição de conhecimento, a inserção de exemplos específicos dos dados que compõem a base de conhecimento e a execução de diversas aplicações. A base de conhecimento resultante é utilizada para resolver problemas e responder perguntas a respeito do domínio. Uma aplicação é o produto final criado quando a base de conhecimento é usada para resolver um problema específico. Por fim, essa ferramenta permite o reuso das ontologias e das aplicações uma vez desenvolvidas (User Guide, 2005).

A escolha do editor de ontologias e bases de conhecimento *Protégé*, como ferramenta a ser utilizada nesta pesquisa, foi favorecida pelos seguintes aspectos:

- ferramenta gratuita e de código aberto, não apresentando custos financeiros para a sua utilização;
- arquitetura modulada que permite a inserção de novos recursos por intermédio de *plug-ins* ou extensões desenvolvidas para sua customização;
- desenvolvida em Java e, portanto, multiplataforma: funciona em ambientes *Windows*, *Mac OS X*, *Linux*, e outros;
- possui interface gráfica interativa e amigável, ou seja, de fácil utilização;

³⁹ <http://protege.stanford.edu>

- é multiusuário: permite que vários usuários editem simultaneamente uma mesma ontologia, promovendo maior interatividade durante a representação, o uso e a visualização de conhecimento;
- é extensível, facilitando a inclusão de gráficos, tabelas e mídias como, imagem, som e vídeo;
- suporta diferentes tipos de formatos de armazenamento, tais como: OWL, RDF, XML e HTML, para serem utilizados de acordo com as aplicações, inclusive externas à ferramenta;
- é amplamente difundida e utilizada: conta com uma comunidade ativa de usuários por todo mundo que realiza pesquisas e projetos que otimizam o uso da ferramenta.

3.2 FERRAMENTAS DE TRIPLIFICAÇÃO

Existem diversas abordagens, sem consenso ou padrão estabelecido, sobre qual a melhor forma de executar essa extração. Dentre as ferramentas usadas nesse tipo de extração, destacam-se o *Triplify*⁴⁰, o D2RQ⁴¹ e o Virtuoso⁴². Vale destacar que essas ferramentas possuem outras funcionalidades, além daquelas relacionadas com a extração de dados. Por exemplo, o Virtuoso é, principalmente, um repositório de triplas. A Tabela 3.3 apresenta um comparativo entre essas três últimas ferramentas, expondo características importantes na hora de um desenvolvedor escolher uma ou outra.

Pela Tabela 3.2, é possível observar que as 3 ferramentas foram desenvolvidas em plataformas diferentes. Apenas D2RQ e Virtuoso disponibilizam SPARQL *endpoints* para consultas e ambos têm como linguagem de mapeamento entre modelos um derivativo de RDF, enquanto o *Triplify* usa o próprio SQL. Ainda sobre mapeamentos, o D2RQ é o único das três a prover uma ferramenta para gerar o mapeamento limitado automaticamente. Em termos de escalabilidade, o Virtuoso apresenta um desempenho melhor.

⁴⁰ <http://triplify.org/Overview>

⁴¹ <http://d2rq.org/>

⁴² <http://virtuoso.openlinksw.com/dataspace/doc/dav/wiki/Main/>

Tabela 3.2 Características das ferramentas de extração sobre dados estruturados.

Ferramenta	<i>Triplify</i>	Virtuoso	D2RQ
Tecnologia	Linguagem de <i>scripts</i> (PHP)	Solução de <i>middleware</i>	Java
Endpoint SPARQL	Não possui	Possui	Possui
Linguagem de mapeamento	SQL	Base em RDF	Base em RDF
Geração de mapeamento	Manual	Manual	Semiautomática
Escalabilidade	Média a alta, mas não usa SPARQL	Alta	Média

Melhorar a automação do processo de extração é o grande desafio nessa modalidade. Uma sintaxe declarativa que sirva de diretiva para essa extração é o próximo passo para sobrepujar os problemas dessa modalidade. Nesse sentido, as soluções da plataforma D2RQ se mostraram mais eficientes, conforme detalhamento de suas funcionalidades nas próximas seções:

3.2.1 Plataforma D2RQ

A grande maioria das instituições armazena seus dados em SGBDs relacionais, entretanto existem ferramentas que, a partir desse modelo convencional, oferecem serviços relacionados ao paradigma da *Web Semântica*.

A ferramenta D2RQ é um *software Open Source* que permite realizar operações em um grafo RDF a partir dos dados armazenados em um banco de dados relacional. As principais funcionalidades desta ferramenta são:

- realizar consultas SPARQL em um banco de dados relacional;
- utilizar o conteúdo do banco de dados relacional como padrão *Linked Data*;
- converter os dados armazenados no banco de dados relacional para arquivo no formato RDF;
- acessar o banco de dados relacional utilizando o *Jena API*⁴³.

⁴³ Será descrito na seção 3.3.1.

O *software* D2RQ utiliza um arquivo de mapeamento para realizar a tradução do banco de dados relacional para o grafo RDF. Esse arquivo de configuração é escrito em formato RDF, no qual o esquema do banco relacional é transcrito para um vocabulário ou ontologia de domínio. Dessa maneira, cada classe é representada por uma tabela, na qual as colunas simbolizam os atributos da mesma.

Com a definição do arquivo de mapeamento, é possível acessar o grafo RDF a partir do servidor D2R *Server* (também disponibiliza um *Endpoint* SPARQL), ou realizando consultas SPARQL com o uso do “d2r-query”, ou utilizar o módulo “*Dump-rdf*” para converter todo o conteúdo do banco de dados relacional para um único arquivo RDF.

3.2.2 Ferramenta *Dump-RDF*

O *Dump-rdf* é uma das ferramentas que fazem parte do *software* D2RQ, seu objetivo é converter em um único arquivo RDF todas as informações armazenadas no banco de dados relacional. Para isso, é necessário informar apenas o arquivo de mapeamento e as informações de saída do arquivo a ser gerado. A forma de execução em linha de comando pode ser vista na Figura 3.1.

```
dump-rdf [-f format] [-b baseURI] [-o outfile.ttl]
          [--verbose] [--debug]
          mapping-file.ttl
```

Figura 3.1 Manual de execução da ferramenta *Dump-rdf*.

Pode-se observar na Figura 3.1 que a *flag-b* indica qual será o padrão das URIs que serão geradas no arquivo RDF, como por exemplo, uma instância da classe “siop:ItemDeDespesa” definida na ontologia do orçamento, caso possua a URI base <http://orcamento.dados.gov.br/2000/id> terá a URI completa <http://orcamento.dados.gov.br/2000/id/Exercicio/2000>;

A ferramenta *Dump-rdf*, realiza uma conversão materializada dos dados, ou seja, a triplificação completa. Já o D2R *Server* provê uma virtualização do grafo RDF a partir do arquivo de mapeamento, utilizando um conceito semelhante ao de uma SQL *View*.

3.3 SISTEMAS DE ARMAZENAMENTO DE ONTOLOGIAS

Por se apresentar em formato de arquivo, o texto se torna um grande desafio de como lidar com grandes ontologias. Dada a OWL como linguagem padrão para a descrição de ontologias segundo a W3C, faz-se necessário a análise de ferramentas que auxiliem no armazenamento, manipulação e recuperação das informações a partir de arquivos OWL.

A Literatura apresenta duas estratégias básicas para o armazenamento de ontologias (Harrison *et al.*, 2005):

- Armazenar a ontologia no sistema de arquivos da máquina na forma de arquivos de texto: o problema desta abordagem deriva do fato de que o sistema de arquivos de uma máquina não provê escalabilidade, dificultando o compartilhamento das informações e não oferecendo suporte às consultas.
- Armazenar a ontologia em banco de dados utilizando um SGBD: a principal dificuldade desta estratégia é que o SGBD requer que a ontologia tenha uma estrutura padronizada, o que é muito difícil de acontecer em casos reais.

No caso de uso a ser apresentado nesta dissertação, foi utilizado o *triple store* da solução D2QR, uma vez que o volume e o perfil de acesso aos RDFs não são grandes.

3.3.1 *Framework Jena*

Jena é um projeto mantido pela *Apache Software Foundation*⁴⁴, que consiste em um *Framework Java open-source* para construção de aplicações que seguem o paradigma da *Web* semântica. A seguir, na Figura 3.2, a arquitetura geral do *Framework*:

⁴⁴ <http://www.apache.org/>

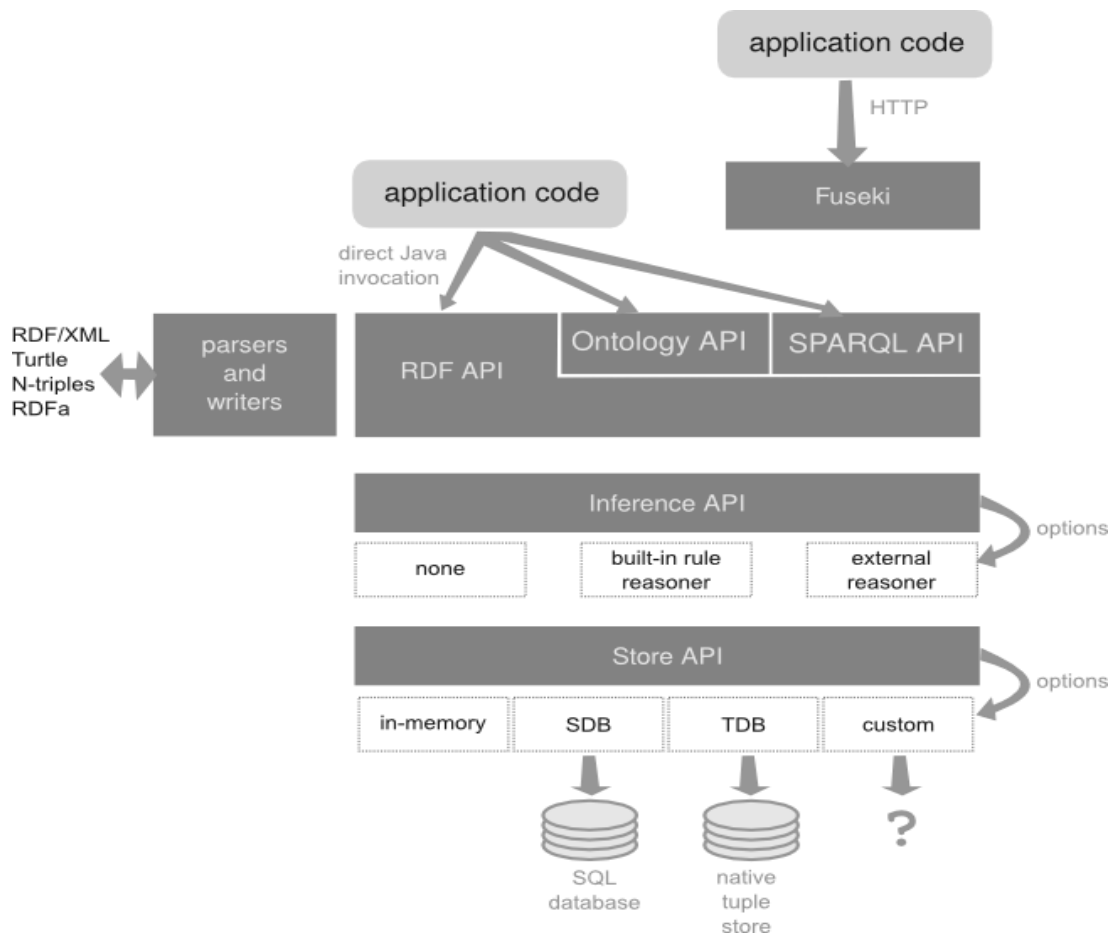


Figura 3.2 Arquitetura do *Framework Jena*.
(Fonte: Jena, 2007)

Os componentes do *Framework* abordam diversas tecnologias relacionadas à *Web* semântica, como: RDF, SPARQL, *endpoints*, ontologias, *Triplestores*, inferência etc. Na Figura 3.2, todos os componentes que têm o nome terminado com “API” representam um conjunto de bibliotecas Java com um determinado propósito, já os outros componentes são ferramentas nativas ou de terceiros.

Dentro do contexto do projeto de abertura dos dados da LOA e a execução orçamentária pela SOF, existiam três principais necessidades para se trabalhar com os dados em RDF gerados após um processo de triplificação, a saber:

- uma forma de persistir esses dados em algum tipo de banco de dados, permitindo um meio de consulta mais otimizado, fazendo uso de alguma estrutura de índices se

for o caso (originalmente o RDF nada mais é que um arquivo de texto serializado seguindo alguma notação);

- uma forma de permitir o acesso aos dados, via alguma tecnologia que ofereça um serviço de acesso via rede;
- uma forma de interagir com os dados, via a linguagem de programação Java.

Após análise do *Framework Jena*, foram identificados os componentes que atendem as três necessidades, os quais serão descritos nas subseções a seguir.

3.3.2 Triplestore TDB

O *Framework Jena* oferece três formas padrões de armazenamento de dados RDF:

- *in-memory*: oferece a carga e a leitura direta dos dados na memória principal do computador (forma transiente);
- *Simple Database System (SDB)*: um sistema de acesso a bancos relacionais como se fossem grafos RDF virtuais (forma persistente);
- *Triplestore Database System (TDB)*: uma *Triplestore* nativa para armazenamento de grafos RDF (forma persistente).

O componente *in-memory* não foi escolhido por não persistir os dados, ou seja, não armazená-los em memória secundária. O SDB também não foi escolhido por utilizar um banco de terceiros (MySQL, *Postgres*, Oracle etc) para conseguir emular um grafo RDF virtual, mas para o projeto havia a necessidade de ser gerado um arquivo RDF real que seria publicado para a população. O TDB então foi escolhido por fornecer uma forma nativa de banco de dados para RDF (comumente chamado de *Triplestore*).

3.3.3 Servidor Fuseki

O TDB oferece um sistema de banco de dados, mas ele sozinho não consegue possibilitar que uma aplicação ou pessoa tenha acesso a esses dados para consulta SPARQL em uma rede. Para oferecer esse tipo de serviço é necessário um *endpoint* SPARQL, uma aplicação

que atende consultas SPARQL de aplicações clientes em uma determinada porta, a maioria dos *endpoints* ainda oferecem alguma página *Web* com um terminal interativo para receber consultas SPARQL feitas por humanos.

O *endpoint* do *Framework Jena*⁴⁵ é o *Fuseki*⁴⁶, o qual consegue oferecer esse serviço para qualquer uma das tecnologias de armazenamento do *Framework*, no projeto ele é utilizado juntamente como o TDB.

⁴⁵ <http://jena.apache.org/>

⁴⁶ http://jena.apache.org/documentation/serving_data/

4 - SOLUÇÃO PROPOSTA

Tendo em vista a crescente movimentação para publicação de dados governamentais em formato aberto no Brasil, este trabalho propõe uma aplicação que converte dados relacionais para formato ontológico segundo o padrão de dados abertos, disponibilizando-os para consulta.

A estratégia adotada no presente trabalho foi estender uma ontologia de classificação da despesa do orçamento federal, contemplando as categorias e conceitos definidos no Manual Técnico do Orçamento - MTO 2013 (Brasil, 2012).

O grande objetivo dessa ontologia é a geração do conhecimento por intermédio da mensuração das ações do governo de modo a possibilitar a realização de uma avaliação dos resultados alcançados dentro do tempo e, por consequência, revisar o uso para uma melhor destinação dos recursos públicos.

Os dados tratados para formato estruturado serão armazenados em aplicações conhecidas como *triple stores* (W3C, 2004), que permitem armazenar dados no formato RDF, possibilitando a construção de consultas na linguagem SPARQL e em diferentes repositórios publicados na *Web*.

4.1 MODELO ORÇAMENTÁRIO

Sanches (2004) descreve que as despesas públicas no orçamento brasileiro são executadas em quatro estágios ou etapas distintas, segundo a estrutura técnico-jurídica vigente: fixação ou planejamento, empenho, liquidação e pagamento. O autor ainda afirma que a etapa de fixação ou planejamento da despesa pode ser definida como sendo a materialização pela publicação da Lei Orçamentária Anual ou do ato de abertura do crédito adicional.

Empenho “é o ato emanado de autoridade competente que cria para o Estado obrigação de pagamento pendente ou não de implemento de condição” (Brasil, 1964).

A liquidação consiste na verificação do direito adquirido pelo credor, tendo por base os títulos e documentos comprobatórios do respectivo crédito, e tem por objetivo apurar a regularidade do objeto, a importância e a quem se deve o pagamento (Brasil, 1964).

Conforme Sanches (2004), o pagamento como etapa final é constituído pela formalização do despacho do ordenador de despesa, autorizando-o, e pela entrega de numerário ao credor por meio de cheque nominativo, ordens de pagamentos ou crédito em conta. Essa etapa só pode ser efetuada após a regular liquidação da despesa.

A elaboração da Lei Orçamentária Anual e as alterações efetivadas ao longo do exercício são realizadas no Sistema Integrado de Planejamento e Orçamento – SIOP (Brasil, 2012).

A contabilização dos valores orçados, bem como as operações e os registros relacionados a cada etapa da execução da despesa são contabilizados no Sistema Integrado de Administração Financeira – SIAFI (Brasil, 2012). O SIOP atualiza o SIAFI com as informações das dotações vigentes na Lei Orçamentária Anual.

Ao longo do exercício financeiro, cada rubrica é anotada para permitir a verificação de quanto foi orçado (dotação inicial; etapa de fixação ou planejamento), quanto foi empenhado (empenho), quanto foi liquidado (liquidação) e quanto foi pago efetivamente (pagamento).

Na Figura 4.1, é possível visualizar uma representação simplificada da relação das ontologias do planejamento e de execução orçamentária.

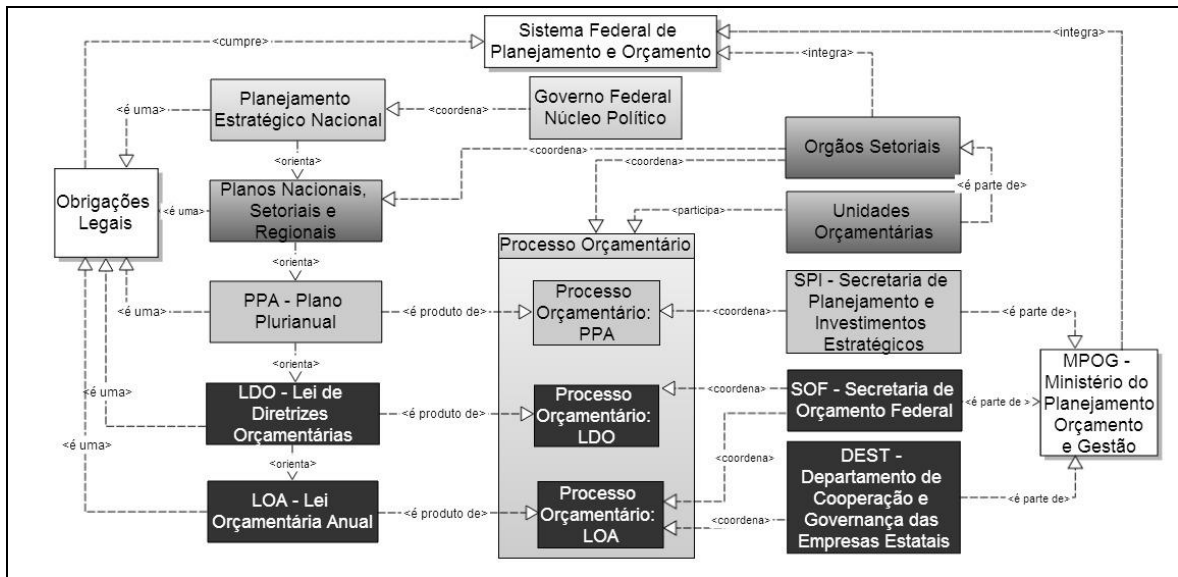


Figura 4.1 Representação simplificada da relação das ontologias do planejamento e de execução orçamentária.

De acordo com Araújo *et al.* (2012),

[...] a base para a compreensão do orçamento público é o sistema de classificação. É por meio desse sistema, que o orçamento é organizado, ou seja, segmentado com base em critérios. Essa estrutura permite que os técnicos e gestores públicos consigam estratificar os dados e estabelecer as relações entre os valores financeiros do orçamento e os fenômenos da administração pública associados (ex. gasto em quê, para quê, sob a responsabilidade de quem).

A estrutura de execução orçamentária é composta fundamentalmente dos mesmos elementos existentes na estrutura de classificações utilizadas na elaboração do orçamento, isto é, de uma lista de itens de despesa com seus respectivos valores. Assim, cada item é representado por um código de 37 dígitos, estruturado em 13 campos, que os vinculam aos critérios de classificação da despesa. O significado de cada critério de classificação e os códigos correspondentes podem ser encontrados no Manual Técnico do Orçamento 2013 – MTO 2013 (Brasil, 2012).

A título de exemplo, é possível considerar um item de despesa da LOA de 2013 com os seus respectivos atributos qualitativos e quantitativos, como está ilustrado na Figura 4.2.

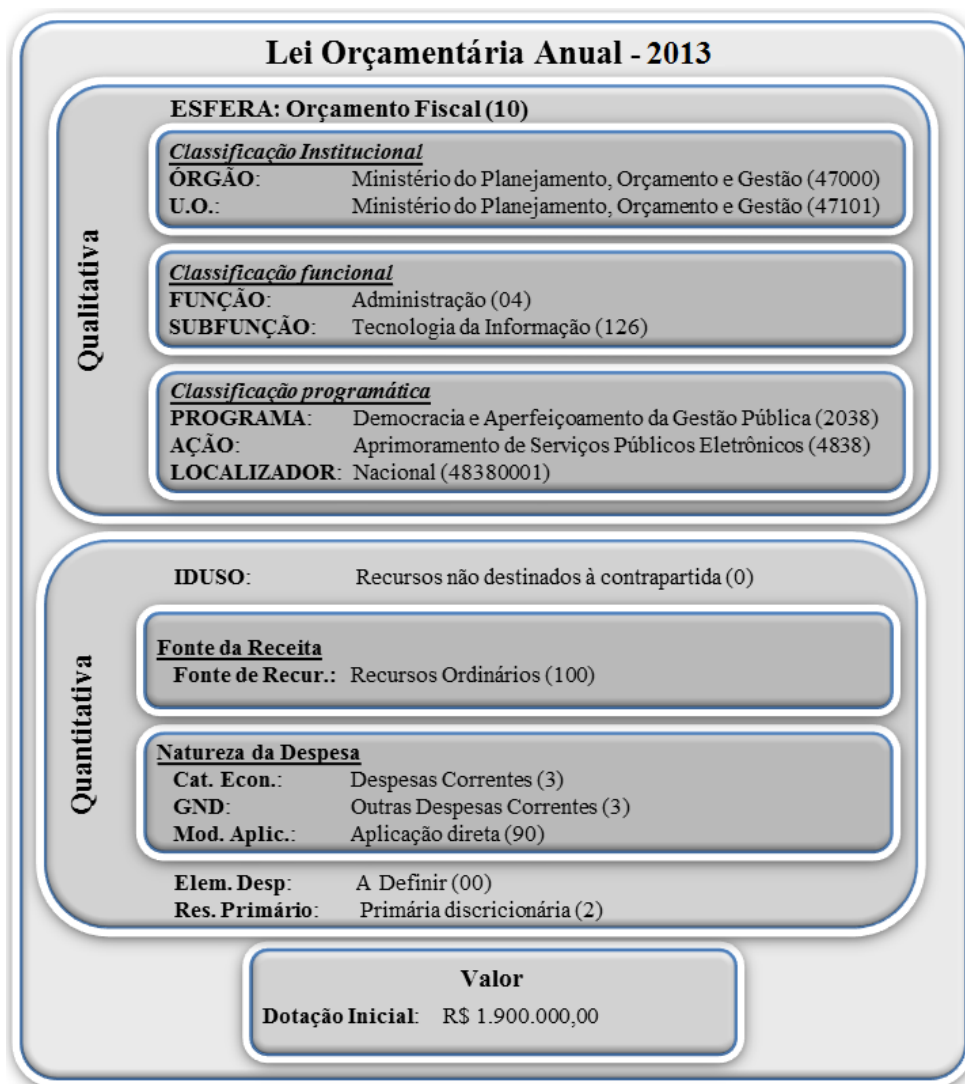


Figura 4.2 Item de despesa da LOA de 2013.
(Brasil, 2012.)

Embora o novo modelo tenha aumentado com as novas exigências, o presente trabalho fornece uma maneira fácil de integrar essas estruturas por meio de relações ontológicas. Assim, novas questões podem ser respondidas por essa expansão da estrutura existente, como a proporção da quantidade gasta entre órgãos ou na avaliação da aplicação do orçamento associado a diversos programas e ações governamentais.

4.2 MODELO ONTOLÓGICO

Para o acompanhamento da execução, utilizou-se como ponto de partida a mesma ontologia das classificações da despesa do orçamento federal descrita em Araújo *et al.* (2012), que é composta pelos itens de despesa e os seus classificadores orçamentários. Neste trabalho, a ontologia foi ampliada, com a adição de novas classes e relações, sinalizadas pelo retângulo cinza, conforme apresentado na Figura 4.3.

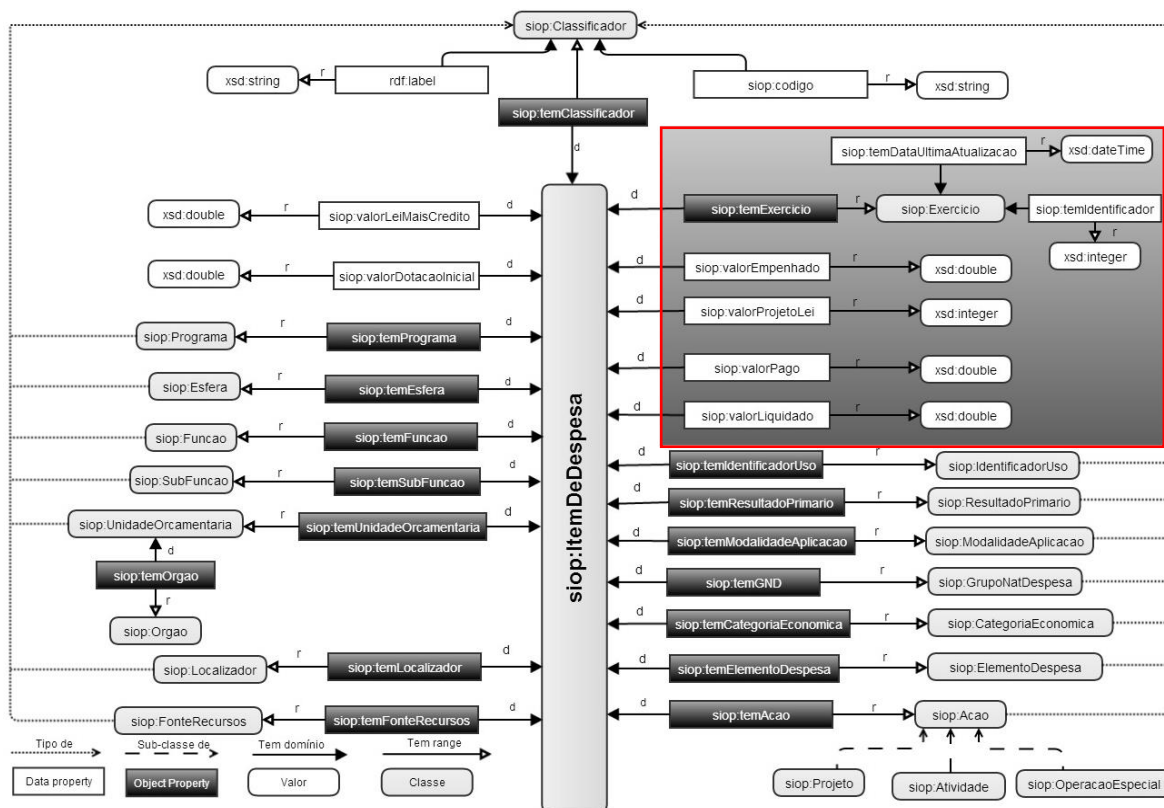


Figura 4.3 Ontologia da Classificação das Despesas do Orçamento Federal.
(Adaptado de Araújo *et al.*, 2012)

Conforme descrito por Araújo *et al.* (2012) na Figura 4.3, os conceitos identificados são descritos como classes OWL, identificadas por retângulos na cor cinza, e as relações entre elementos de classes por retângulos pretos. O *range* (interseções ou tipos de dados) das relações é representado por uma letra ‘r’, enquanto o *domain* é representado pela letra ‘d’. Linhas tracejadas representam relações de subclasse. O prefixo “siop” foi assumido para as propriedades de objeto (*object property*). As propriedades de tipos de dados (*data type property*) são representadas por retângulos brancos contendo seu tipo. Subentende-se aqui

que os elementos da classe Item de Despesa são definidos por terem uma relação unívoca (1:1) com um elemento de cada uma das classes referentes aos diferentes tipos de classificadores vigentes no orçamento. Apesar de não constar, estão implícitas: (i) a classe Classificadores, que reúne todos os classificadores orçamentários apresentados na Figura 4.3, e (ii) a classe Orçamento Anual, que contempla os itens de despesa e outros elementos significativos. Ainda na Figura 4.3, é possível visualizar na área de destaque retangular cinza no canto superior direito a extensão proposta para a ontologia, a inserção da classe “Exercício” e dos atributos: “Exercício”, “DataUltimaAtualizacao”, “valorEmpenhado”, “valorLiquidade” e “valorPago”.

Para que a execução orçamentária pudesse ser representada, a classe “Exercício” foi adicionada, levando em consideração que o orçamento é elaborado e executado por ano, desta forma tornou-se possível a publicação da execução orçamentária dos exercícios anteriores.

Considerando a atualização diária dos dados de execução, o atributo, “temDataUltimaAtualizacao” foi adicionado a partir da classe “Exercício”, o que nada mais é que um controle da posição dos dados para que seja sempre igual a menos um dia da data corrente.

Conforme descrito no MTO 2013 (Brasil, 2012), um identificador único pode ser utilizado para representar uma determinada dotação orçamentária, representado pela relação “temIdentificador”. Alguns valores são fundamentais para acompanhar a execução do orçamento federal, pois demonstram o que efetivamente está sendo realizado. Esses valores são representados nas propriedades “valorEmpenhado”, “valorProjetoLei”, “valorLiquidade” e “valorPago”. A propriedade “valorEmpenhado” representa o valor com destinação certa, previsto para ser realizado, mas que ainda não o foi. A propriedade “valorDotação” representa o valor que consta na aprovação do orçamento, o que está aprovado e previsto para ser gasto, e que é fixado após aprovação do orçamento, porém de suma importância para análises de execução orçamentária. A propriedade “valorLiquidade” representa o que estava empenhado e efetivamente foi autorizado. Autorizar o valor não significa necessariamente transferir o recurso, mas conceder, de fato, a autorização legal para realização, levando-se em conta que o valor foi empenhado

anteriormente. Por último, a propriedade “valorPago” representa o que efetivamente saiu dos cofres públicos para realizar a despesa.

O orçamento, para ser descrito, deve conter todas essas características qualitativas e quantitativas associadas. Na ontologia proposta, a unidade básica de informação do orçamento é chamada de “Item de Despesa”, que é a principal classe e se refere a um detalhamento do orçamento, em cuja parte fica descrita a relação dos classificadores do orçamento com o valor de dotação inicial. Portanto, o orçamento pode ser entendido como um conjunto de itens de despesa que são caracterizados por valor de dotação inicial. Vinculado a esse valor, há uma série de classificadores que permitem dividir o orçamento em diferentes formas de acordo com o interesse. Nesse caso, Classificador é a classe maior que se relaciona à classe “ItemDeDespesa” pela relação “itemClassificador” e deve conter dois tipos de dados (*data type property*) associados: (i) um *label*, que descreve o classificador, e (ii) um código que individualiza o classificador.

A classe “ItemDeDespesa” contém relações explícitas (*object properties*) com diversas classes relativas aos classificadores definidos para o orçamento de 2013. Mais especificamente, é definida por um conjunto de restrições que declara que um item de despesa deve ter: (i) um valor de dotação (*data type property*), e (ii) um conjunto de relações (*object property*) funcionais, ou seja, relações com único indivíduo das classes Esfera, Unidade Orçamentária, Função, Subfunção, Programa, Ação, Localizador, Grupo de Natureza da Despesa, Modalidade de Aplicação, Elemento de Despesa, Fonte de Recursos, Indicador de Uso e Resultado Primário. No caso da relação com o Órgão, ela é descrita de forma implícita, pois quem realmente possui a ligação com o Órgão é a Unidade Orçamentária. O mesmo caso ocorre com a Categoria Econômica, pois quem tem a relação explícita é o Grupo de Natureza da Despesa.

Na Figura 4.3 existem também subclasses que são particularizações de uma classe maior. Por exemplo, a classe Ação possui as subclasses Atividade, Projeto e Operação Especial, que são particularizações do conceito de Ação. Na prática, um Item de Despesa tem obrigatoriamente uma relação com Ação, mais especificamente com uma dessas três subclasses. Na Tabela 4.1, estão descritas as classes principais da ontologia básica de orçamento.

Tabela 4.1 Relação de Classes e seus significados.

Classes	Descrição
CategoriaEconomic	Classe nomeada que contém dois elementos, em princípio, cujos <i>labels</i> e códigos são: (i) despesas correntes - código 3, e (ii) despesas de capital - código 4.
Classificador	O orçamento anual pode ser visto sob diferentes tipos de classificações. A classe “Classificadores” possui as subclasses relativas aos tipos de classificação que o orçamento pode ter e é definida pela restrição de <i>data type property</i> . Ou seja, os elementos dessa classe precisam ter um <i>label</i> e um código (<i>data type properties</i>) para pertencerem a essa classe.
ElementoDespesa	Classe nomeada cujos elementos referem-se aos indivíduos que têm por finalidade identificar os objetos de gasto, tais como vencimentos e vantagens fixas, juros, diárias, material de consumo, serviços de terceiros prestados sob qualquer forma, subvenções sociais, obras e instalações, equipamentos e material permanente, auxílios, amortização e outros que a Administração Pública utiliza para consecução de seus fins.
Esfera	Classe nomeada cujos elementos representam os tipos de orçamento definidos para o Governo: orçamento fiscal (10), orçamento da seguridade social (20) e orçamento de investimento (30).
FonteRecursos	Classe nomeada cujos elementos referem-se aos indivíduos que representam a origem dos recursos para realizar a despesa declarada para o indivíduo relacionado da classe ItemDeDespesa. Em outras palavras, essa classe define a origem do dinheiro a ser gasto no item de despesa relacionado.
GrupoNatDespesa	Classe cujos elementos são agregadores de elemento de despesa com as mesmas características quanto ao objeto de gasto, conforme a seguinte discriminação: 1. Pessoal e encargos sociais; 2. Juros e encargos da dívida; 3. Outras despesas correntes; 4. Investimentos; 5. Inversões financeiras; 6. Amortização da dívida.
IndicadorUso	Classe nomeada cujos elementos completam a informação concernente à aplicação dos recursos e destina-se a indicar se os recursos compõem contrapartida nacional de empréstimos ou de doações ou destinam-se a outras aplicações, constando da LOA e de seus créditos adicionais.
ItemDeDespesa	Classe definida por restrições cujos indivíduos representam os itens de despesa definidos para cada orçamento anual, com ênfase especial para a LOA 2012. Cada elemento da classe “ItemDeDespesa” é caracterizado por ter <i>object properties</i> específicas para cada uma das subclasses da classe “Classificadores”, além de possuir um <i>data type property</i> denominado ValorDotacaoInicial do tipo <i>double</i> .
Localizador	Elementos dessa classe são objetos que identificam o local definido para realização de uma determinada ação.
ModalidadeAplicacao	Indica se os recursos serão aplicados mediante transferência financeira, inclusive a decorrente de descentralização orçamentária para outros níveis de Governo, seus órgãos ou entidades, ou diretamente para entidades privadas sem fins lucrativos e outras instituições; ou então diretamente pela unidade detentora do crédito orçamentário, ou por outro órgão ou entidade no âmbito do mesmo nível de Governo. Seu principal objetivo é eliminar a dupla contagem dos recursos transferidos ou descentralizados.

Classes	Descrição
Programa	Os elementos dessa classe são programas orientados para a realização dos objetivos estratégicos definidos para o período do PPA (Plano Plurianual), ou seja, quatro anos. Programas podem ser temáticos (que retratam no PPA a agenda de governo e orienta a ação governamental), ou de gestão, manutenção e serviços (instrumentos do PPA que classificam um conjunto de ações destinadas ao apoio, gestão e manutenção da atuação governamental).
Acao	Os elementos dessa classe são operações da qual resultam produtos (bens ou serviços) que contribuem para atender ao objetivo de um programa. Incluem-se aqui as transferências obrigatórias ou voluntárias a outros entes da Federação e a pessoas físicas ou jurídicas, na forma de subsídios, subvenções, auxílios, contribuições, entre outros e financiamentos.
Projeto	Os elementos dessa classe são ações limitadas no tempo e das quais resulta um produto que concorre para a expansão ou aperfeiçoamento da ação do governo.
Atividade	Tipo de ação que se realiza de modo contínuo e permanente, das quais resulta um produto ou serviço necessário à manutenção da ação do Governo.
OperacaoEspecial	Tipo de ação cujas despesas não contribuem para a manutenção, expansão ou aperfeiçoamento das ações de governo. Ações dessa subclasse não resultam em produto e não geram contraprestação direta sob a forma de bens ou serviços.
Funcao	Classe de elementos que representam o maior nível de agregação das diversas áreas de atuação do setor público. Reflete a competência institucional do órgão, como, por exemplo, cultura, educação, saúde, defesa, que guarda relação com os respectivos ministérios.
Subfuncao	Classe de elementos que representam um nível de agregação imediatamente inferior à classe “Funcao” e deve evidenciar cada área de atuação governamental por intermédio da identificação da natureza das ações.
UnidadeOrcamentaria	Classe nomeada que representa os órgãos, denominados aqui de Unidades Orçamentárias, que são os destinatários das dotações orçamentárias declaradas ou autorizações de gasto declaradas na classe ElementoDeDespesa.
Orgao	Os elementos dessa classe referem-se aos órgãos orçamentários que possuem vinculados a si os elementos da classe “UnidadeOrcamentaria” por meio da propriedade “temOrgao”. Órgãos não correspondem necessariamente a uma estrutura administrativa, como ocorre, por exemplo, com alguns fundos especiais e com alguns órgãos tais como (i) Transferências a Estados, Distrito Federal e Municípios, (ii) Encargos financeiros da União, (iii) Operações oficiais de crédito, (iv) Refinanciamento da dívida pública mobiliária federal, e (v) Reserva de contingência.
ResultadoPrimario	Classe nomeada cujos elementos referem-se à apuração de resultado primário previsto na LDO (Lei de Diretrizes Orçamentárias), devendo constar no PLOA e na respectiva lei em todos os GNDS, identificando, de acordo com a metodologia de cálculo de necessidades de financiamento, cujo demonstrativo constará no anexo à LOA.

Com relação às propriedades de objeto (*object property*), a maior parte delas tem como domínio a classe “ItemDeDespesa” e como *range*, as subclasses de classificadores descritas na Figura 4.3. Além disso, existem propriedades de objeto que relacionam indivíduos das classes “CategoriaEconomicas” com os da classe “GrupoNatDespesa” e entre os indivíduos da classe “UnidadeOrcamentaria” e os da classe “Orgao”. A descrição de cada uma dessas relações está descrita na Tabela 4.2.

Tabela 4.2 Relação de Propriedades e seus significados.

Object property	Descrição
temClassificador	Essa <i>object property</i> relaciona a classe <i>ItemDeDespesa</i> (<i>domain</i>) com a classe <i>Classificadores</i> (<i>range</i>).
temAcao	Propriedade funcional que tem como domínio a classe <i>ItemDeDespesa</i> e <i>range</i> a classe Ação.
temCategoriaEconomica	Propriedade (<i>object property</i>) que relaciona elementos da classe <i>GrupoNatDespesa</i> (<i>domain</i>) aos elementos da classe <i>CategoriaEconomica</i> . (<i>range</i>).
temElementoDespesa	Propriedade funcional que tem como domínio a classe <i>ItemDeDespesa</i> e <i>range</i> a classe <i>ElementoDespesa</i> .
temEsfera	Propriedade funcional que tem como domínio a classe <i>ItemDeDespesa</i> e <i>range</i> a classe Esfera.
temFonteRecursos	Propriedade funcional que tem como domínio a classe <i>ItemDeDespesa</i> e <i>range</i> a classe <i>FonteRecursos</i> .
temFuncao	Propriedade funcional que tem como domínio a classe <i>ItemDeDespesa</i> e <i>range</i> a classe <i>Funcao</i> .
temGnd	Propriedade funcional que tem como domínio a classe <i>ItemDeDespesa</i> e <i>range</i> a classe <i>GrupoNatDespesa</i> .(GND)
temIndicadorUso	Propriedade funcional que tem como domínio a classe <i>ItemDeDespesa</i> e <i>range</i> a classe <i>IndicadorUso</i> .
temLocalizador	Propriedade funcional que tem como domínio a classe <i>ItemDeDespesa</i> e <i>range</i> a classe <i>Localizador</i> .
temModalidadeAplicacao	Propriedade funcional que tem como domínio a classe <i>ItemDeDespesa</i> e <i>range</i> a classe <i>ModalidadeAplicacao</i> .
temOrgao	<i>Object property</i> que relaciona elementos da classe (<i>domain</i>) <i>UnidadeOrcamentaria</i> com os elementos da classe (<i>range</i>) <i>Orgao</i> .
temPrograma	Propriedade funcional que tem como domínio a classe <i>ItemDeDespesa</i> e <i>range</i> a classe <i>Programa</i> .
temResultadoPrimario	Propriedade funcional que tem como domínio a classe <i>ItemDeDespesa</i> e <i>range</i> a classe <i>ResultadoPrimario</i> .
temSubfuncao	Propriedade funcional que tem como domínio a classe <i>ItemDeDespesa</i> e <i>range</i> a classe <i>SubFuncao</i> .
temUnidadeOrcamentaria	Propriedade funcional que tem como domínio a classe <i>ItemDeDespesa</i> e <i>range</i> a classe <i>UnidadeOrcamentaria</i> .

Nessa ontologia, as propriedades de tipo de dado (*data type property*) são também importantes para restringir e definir classes. Por exemplo, “ItemdeDespesa” possui uma propriedade “codigo” que representa o código do classificador. Portanto, qualquer descrição que tome essa ontologia como referência será considerada inválida se não contiver esses elementos descritos. Além desse tipo de dado, existem os tipos associados aos classificadores que são suficientes para definir os elementos para o contexto da LOA 2013. A Tabela 4.3 contém a descrição dos tipos de dados utilizados nessa ontologia.

Tabela 4.3 Relação de Tipos de Dados e seus significados.

<i>Data Type Propertytype property</i>	Descrição
Código	Código do classificador. Esse <i>data type</i> é herdado para as subclasses da classe Classificador e é utilizado como uma restrição para definir essas subclasses.
<i>Label</i>	Refere-se ao nome do classificador. Esse <i>data type</i> é herdado para as subclasses da classe Classificador e é utilizado como uma restrição para definir subclasses.

Além das tabelas e da descrição apresentadas aqui, esse documento contém a descrição OWL completa das classes e relações apresentadas no Apêndice A.

Para a construção da ontologia do domínio LOA, foi escolhida, dentre as opções mencionadas, a ferramenta *Protégé*, em sua versão 4.3.0, que foi desenvolvido pelo *Stanford Medical Informatics* na escola de medicina da universidade de Stanford (Califórnia, EUA), com o apoio de diversos colaboradores. Essa ferramenta dispõe de uma interface gráfica para edição de ontologias e uma arquitetura para a criação de ferramentas com base em conhecimento, podendo ser usada tanto por desenvolvedores de sistema como por especialistas em domínio para criar bases de conhecimento, permitindo representar facilmente o conhecimento de uma área. Esse editor é capaz de tratar classes, com sua definição e exemplos, simultaneamente, assim, um exemplo singular pode ser usado no nível de uma definição de classe, e uma classe pode ser armazenada como um exemplo. Similarmente, os atributos empregados dentro da classe podem ser elevados ao mesmo nível que uma classe. (*Protégé*, 2005).

4.3 MODELO SISTÊMICO

O modelo sistêmico proposto para solucionar o problema de publicação de visões de RDF de dados relacionais, é fundamentado na plataforma D2RQ, pelo uso de ferramenta *Dump-RDF*, que possibilitou o correto mapeamento dos dados relacionais para os seus respectivos valores em RDF. A Figura 4.4 apresenta o diagrama arquitetural e representa os componentes da solução.

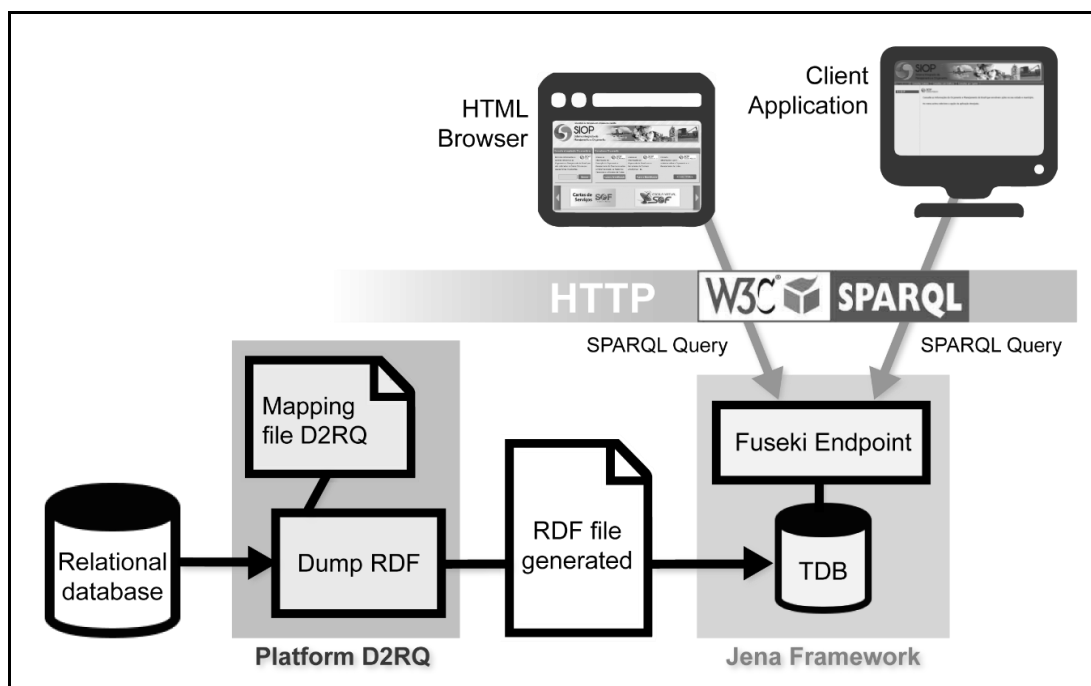


Figura 4.4 Modelo arquitetural da solução proposta.

No diagrama da Figura 4.4, pode-se observar os seguintes itens da arquitetura proposta:

4.3.1 Base relacional

Os dados do Orçamento Federal estavam originalmente armazenados em um banco de dados relacional, no SGBD PostgreSQL, versão 9.0. Os registros da LOA se encontravam inicialmente em uma única tabela, na qual possuía todas as informações referentes a todos os exercícios. A partir disso foi executada uma transformação dos dados para um modelo relacional customizado, no qual foi realizada uma filtragem para que cada Classificador Orçamentário originasse uma nova tabela. Outra tabela foi utilizada para armazenar os

valores orçamentários e o relacionamento com os seus respectivos classificadores. Esse modelo novo foi replicado para cada exercício.

4.3.2 Arquivo de mapeamento D2RQ

As propriedades mapeadas definem quais campos do modelo relacional se tornarão propriedades da classe no arquivo RDF final. Um exemplo de mapeamento do classificador Programa é apresentado a seguir na Figura 4.5, com um comentário da estrutura da tabela relacional utilizada, desconsiderando o objeto para configuração do banco de dados:

```
@prefix map: <#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix d2rq: <http://www.wiwiw.fu-berlin.de/suhl/bizer/D2RQ/0.1#> .
@prefix loa: <http://vocab.e.gov.br/2012/08/loa2012#> .
# Namespace instance passed in command line: <http://orcamento.dados.gov.br/2013/id/>

#CREATE TABLE Programa (cod_programa character(4), des_programa varchar(300));

map:Programa a d2rq:ClassMap;
d2rq:dataStorage map:database;
d2rq:uriPattern "Programa/@@rdf:programa.cod_programa@@";
d2rq:class loa:Programa; .

map:Programa_label a d2rq:PropertyBridge;
d2rq:belongsToClassMap map:Programa;
d2rq:property rdf:label;
d2rq:column "rdf:programa.des_programa"; .

map:Programa_codigo a d2rq:PropertyBridge;
d2rq:belongsToClassMap map:Programa;
d2rq:property siop:codigo;
d2rq:column "rdf:programa.cod_programa"; .
```

Figura 4.5 Mapeamento do classificador Programa

4.3.3 Dump-RDF

A ferramenta componente da plataforma D2RQ, será responsável por triplicar os dados relacionais, gerando um arquivo RDF como saída. Essa ferramenta é capaz de gerar mapeamentos automáticos da base relacional, porém, não são suficientes para descrever as relações entre as tabelas relacionais e a ontologia proposta. Para efetuar esta conversão, na arquitetura proposta, foi necessário descrever a relação entre a ontologia e o modelo de dados relacional em um arquivo de mapeamento, conforme descrito no item 4.3.2.

4.3.4 Arquivo RDF gerado

Após o processo de triplificação, um documento RDF é criado no formato de serialização *N-Triples*. Considerando o mapeamento realizado para a tabela relacional programada, um registro convertido para o formato *N-Triples*, levando em consideração a ontologia, é exemplificado na figura 4.6:

```
@prefix loa: <http://vocab.e.gov.br/2012/08/loa2012>.
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

loa:program-2035 rdf:type loa: Programa.
loa:program-2035 loa:code "2035" .
loa:program-2035 rdf:label "Apoio à Realização da Copa do Mundo FIFA 2014" .
```

Figura 4.6 Ontologia formato N-Triples

4.3.5 TDB

TDB é o componente do *Jena Framework* que atua como um *triple store*, oferecendo armazenamento e consulta de dados no formato RDF. Tendo como insumo um arquivo RDF serializado, esse componente persiste nos dados que ele representa, aumentando o desempenho de consultas em relação a uma consulta direta ao arquivo.

4.3.6 Terminal Fuseki

O fluxo das atividades para se chegar ao momento ao qual há a possibilidade de uma consulta dos dados orçamentários é relativamente simples: a ferramenta *RDF-Dump*, a partir do arquivo de mapeamento e tendo como fonte uma base de dados relacional, triplifica os dados e gera um arquivo RDF, o qual é carregado no TDB e, a partir desses dados, o servidor *Fuseki* consegue receber consultas SPARQL de um usuário, entregá-las ao TDB e, ao obter uma resposta, devolver o resultado ao usuário requisitante.

4.3.7 Questões de Interface Visual

A Interface do sistema SIOP apresentado na prova de conceito tem como principal finalidade propiciar uma melhor experiência por parte do usuário na manipulação de dados

e na navegabilidade no sistema. Para atender a essas finalidades, são utilizados conceitos de usabilidade centrada no usuário.

O Sistema SIOP Dados Abertos foi concebido dentro das quatro diretrizes propostas na Cartilha de Codificação para Padrões *Web* do Governo Eletrônico (e-PWG)⁴⁷ que são:

- a) Páginas leves: as páginas devem ter preferencialmente até 50 kb, somados código, conteúdo e imagens. Recomenda-se que as páginas não ultrapassem 70kb de tamanho.
- b) Separação da forma do conteúdo: a estrutura e o conteúdo da página devem ser separados de sua codificação e visual. O posicionamento, a forma e a cor devem ser criados a partir de folhas de estilo *Cascading Style Sheets (CSS)*⁴⁸.
- c) Páginas em conformidade com os padrões *Web*: padrões *Web* são especificações abertas que preveem a acessibilidade desses documentos ao maior grupo de indivíduos possível.
- d) Páginas independentes de navegador e plataforma: a utilização de *tags* ou comandos proprietários podem comprometer, prejudicar e até impedir a visualização e o perfeito funcionamento da página em outros navegadores ou sistemas operacionais.

Na Figura 4.7, é exemplificado uma saída do módulo de Dados Abertos do SIOP.

	Projeto Lei	Dotação Inicial	Dotação Atual	Empenhado	Liquidado	Pago
Orgão: 47000 - Ministério do Planejamento, Orçamento e Gestão	870.000,00	870.000,00	870.000,00	50.000,00	15.934,95	15.934,95
UO: 47101 - Ministério do Planejamento, Orçamento e Gestão	870.000,00	870.000,00	870.000,00	50.000,00	15.934,95	15.934,95
Programa: 2038 - Democracia e Aperfeiçoamento da Gestão Pública	870.000,00	870.000,00	870.000,00	50.000,00	15.934,95	15.934,95
04.121.2038.20U0 - Gestão e Aprimoramento do Planejamento	870.000,00	870.000,00	870.000,00	50.000,00	15.934,95	15.934,95
Localizador: 20U0.0001 - Nacional	870.000,00	870.000,00	870.000,00	50.000,00	15.934,95	15.934,95
PO: 0001 - Gestão da Tecnologia e da Informação de Planejamento e Investimentos Estratégicos	620.000,00	620.000,00	620.000,00	0,00	0,00	0,00
PO: 0003 - Participação do Brasil no Conselho Sul-Americano de Infraestrutura e Planejamento da UNASUL (COSIPLAN)	250.000,00	250.000,00	250.000,00	50.000,00	15.934,95	15.934,95

Figura 4.7 Resultado de consulta SPARQL do orçamento.

⁴⁷ <http://epwg.governoeletronico.gov.br/cartilha-codificacao>

⁴⁸ <http://www.w3.org/Style/CSS/>

5 - PROVA DE CONCEITO E RESULTADOS OBTIDOS

Neste tópico serão apresentados os passos executados na prova de conceito que vão desde a extensão ontológica até a recuperação da informação, já em formato RDF através da consulta SPARQL.

Vale ressaltar que em todo momento houve a preocupação de que o terminal de consulta se apresentasse de forma simples, cognitiva e ergonômica de forma a diminuir o nível de frustração do usuário e proporcionar o acesso simplificado à informação.

Ainda aqui serão disponibilizados os resultados desta prova de conceito quanto ao formato dos dados triplicados, volume de classes triplicadas e quantidade de acessos à aplicação de consulta durante o período da prova de conceito.

Em suma, o grande objetivo dessa prova de conceito é apresentar de forma simples os processos necessários para a manipulação de dados abertos, com isso, estimular outros desenvolvedores a criar outras aplicações capazes de interpretar dados estruturados abertos.

5.1 PROVA DE CONCEITO

Para execução da prova de conceito, foi utilizado um Item de Despesa da LOA de 2013. Conforme descrito na proposta do trabalho, foi executada a transformação dos dados estruturados em sentenças RDF.

5.1.1 Extensão ontológica

A partir da ontologia disponível em <https://www1.siop.planejamento.gov.br/acessopublico>, foi utilizado o *Protégé* para realização da extensão ontológica, inserindo novas classes e propriedades, conforme o exposto no tópico 4.2 do presente trabalho.

Primeiramente foi necessário transcrever a classe “Exercício” criada no diagrama para a ontologia. A figura 5.1, ilustra a ferramenta *Protégé* na aba de definição de classes, no momento da criação da classe, com diversas classes e subclasses definidas no diagrama de hierarquia de classes.

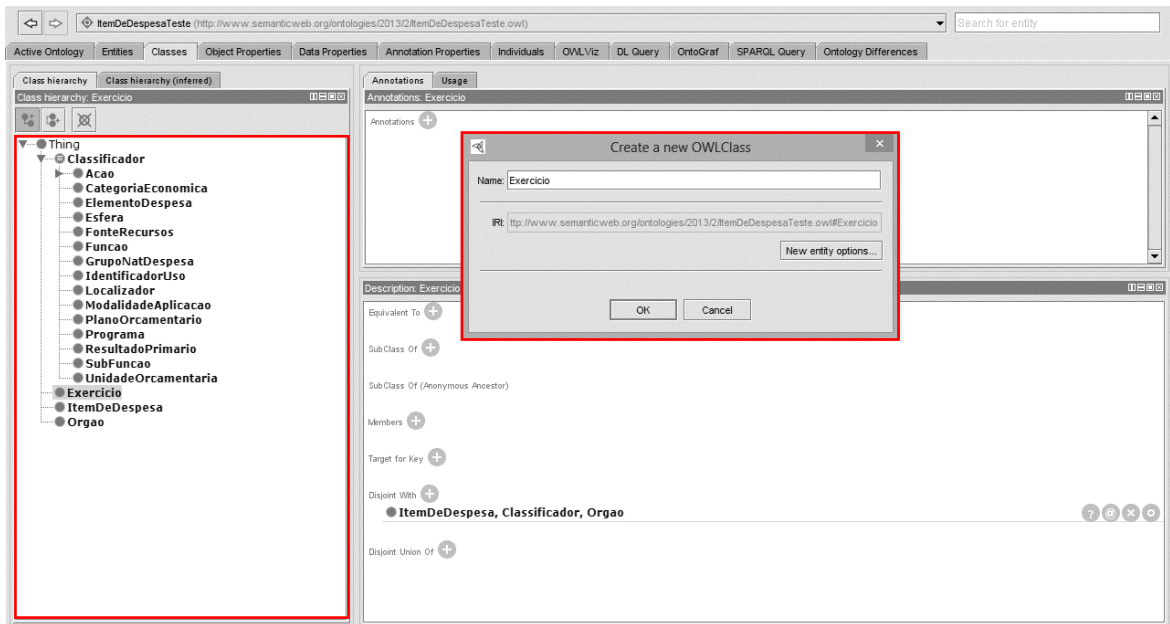


Figura 5.1 Modelagem de Classes no *protegé*.

A partir do diagrama de classes, pode-se identificar especificações da modelagem de dados, que são tratadas de modo distinto no momento da criação da ontologia.

Foram criados os novos atributos das classes, que na ontologia são tratados como *Datatype Properties* (propriedades/atributos): “temDataUltimaAtualizacao”, “valorEmpenhado”, “valorLiquidado”, “valorPago”. *Datatype Properties* são propriedades ou atributos das classes que possuem como domínio uma ou mais classes da ontologia e como escopo, tipos de dados determinados, como *string*, *float*, *int* e *boolean*. Por padrão o *namespace* <http://www.w3.org/2001/XMLSchema>, representado pelo prefixo *xsd*, é utilizado por conter as definições de tipos de dados.

Todas essas propriedades exercem a função apenas de atributos das classes, contudo outras podem ser utilizadas para criar regras, que restringem as características para que instâncias pertençam a determinadas classes.

Enquanto os atributos de classes são transformados em *Datatype Properties*, os relacionamentos entre classes, apresentados no diagrama, são transformados em *Object properties* no desenvolvimento da ontologia. De acordo com a definição, cada *Object property* tem como opção domínio de uma ou mais classes, e como contra-domínio também, uma ou mais classes. Ao se instanciar uma classe, um *Object property* terá o valor

uma instância da classe especificada como contradomínio. Nesse caso, é possível observar a relação sujeito, predicado e objeto, princípio fundamental da modelagem na *Web Semântica*.

Além das características de domínio e escopo, respectivamente *Domain* e *Range*, existem especificações adicionais e opcionais para as propriedades. Essas características são utilizadas para impor ainda mais restrições necessárias para as instâncias dessas classes. As especificações que marcam a propriedade como funcional (*Functional*), funcional inversa (*InverseFunctional*) e inversa (*Inverse*) merecem destaque pela contribuição para a modelagem.

A propriedade definida como funcional indica que, para um sujeito que pertence à instância de uma classe, existe apenas um objeto para determinada propriedade. Inversamente, a definição de uma propriedade como funcional inversa indica que existe apenas um sujeito para um dado predicado (propriedade) e objeto. Ainda mais interessante é a característica inversa das propriedades, na qual é definida a reciprocidade de uma determinada propriedade, ou seja, a direção do relacionamento, que vai do domínio do sujeito até o contradomínio do objeto, também tem correta aplicação inversa.

O resultado final da extensão da ontologia do orçamento pode ser observado no Apêndice A em representação OWL/XML que foi gerada utilizando o *Protégé*.

5.1.2 Geração do RDF

A ferramenta *Dump-rdf*, presente na plataforma D2RQ, foi escolhida para realizar a conversão dos dados persistidos no esquema relacional para o formato RDF. Para que essa tarefa seja concluída, é necessário configurar um arquivo de mapeamento, no qual as tabelas e colunas do modelo relacional serão interpretadas como classes e relações da ontologia de domínio. Esse arquivo de configuração é escrito no formato RDF, por intermédio da serialização *Turtle*.

Para formalizar o mapeamento do modelo relacional para o modelo ontológico, deve ser utilizada uma linguagem específica definida pela própria ferramenta, chamada *D2RQ Mapping Language*. Utilizando os recursos desta sintaxe, é possível determinar as relações

entre tabelas e classes, juntamente com seus atributos. Também são definidos parâmetros de conexão com o banco de dados relacional, conforme a Figura 5.2.

```
@prefix map: <#> .
@prefix db: <> .
@prefix loa: <http://vocab.e.gov.br/2012/08/loa#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix d2rq: <http://www.wiwiss.fu-berlin.de/suhl/bizer/D2RQ/0.1#> .
@prefix jdbc: <http://d2rq.org/terms/jdbc/> .

##### Database Connection #####
map:database a d2rq:Database;
    d2rq:jdbcDriver "org.postgresql.Driver";
    d2rq:jdbcDSN "jdbc:postgresql://<%=ip_bd_dump_rdf%>/<%=base_bd_dump_rdf%>";
    d2rq:username "<%=user_bd_dump_rdf%>";
    d2rq:password "<%=pass_bd_dump_rdf%>";
    jdbc:autoReconnect "true";
    jdbc:zeroDateTimeBehavior "convertToNull";
    .
```

Figura 5.2 Parâmetros de conexão utilizados pelo arquivo de mapeamento.

A primeira parte desse mapeamento é composta por um conjunto de *namespaces* que evidenciam os diferentes contextos utilizados. O prefixo “loa” representa a ontologia de domínio criada pela SOF. Em seguida, são informados os parâmetros de conexão com o banco de dados relacional, assim como URL, *driver* de conexão, usuário e sua respectiva senha.

As tabelas que compõem o esquema relacional construído para esse processo correspondem a classes do modelo ontológico. Portanto, cada uma dessas tabelas foi relacionada com a cláusula “d2rq:Classmap” indicando a representação de uma classe, na qual a propriedade “d2rq:uriPattern” especifica a coluna a ser usada para identificar as instâncias desta classe.

As relações entre as classes são mapeadas por meio das propriedades “d2rq:PropertyBridge” e “d2rq:belongsToClassMap”. A cláusula “d2rq:column” classifica a propriedade como um *Datatype Property*, fazendo a ligação de uma classe a *Literais* de determinada coluna. Por outro lado, as relações entre instâncias, os chamados *Object properties*, são identificados pelas expressões “d2rq:refersToClassMap” e “d2rq:join”.

Logo, uma chave estrangeira que relaciona duas tabelas se transforma em uma relação entre indivíduos de uma ou mais classes.

Na Figura 5.3, é possível observar o mapeamento da classe `loa:UnidadeOrcamentaria`:

```
##### Table unidade_orcamentaria #####
#### Class ####
map:UnidadeOrcamentaria a d2rq:ClassMap;
  d2rq:dataStorage map:database;
  d2rq:uriPattern "UnidadeOrcamentaria/@@rdf.unidade_orcamentaria.cod_unidade_orcamentaria@";
  d2rq:class loa:UnidadeOrcamentaria;
.

#### Datatype Properties ####
map:UnidadeOrcamentaria_label a d2rq:PropertyBridge;
  d2rq:belongsToClassMap map:UnidadeOrcamentaria;
  d2rq:property rdf:label;
  d2rq:column "rdf.unidade_orcamentaria.des_unidade_orcamentaria";
.

map:UnidadeOrcamentaria_codigo a d2rq:PropertyBridge;
  d2rq:belongsToClassMap map:UnidadeOrcamentaria;
  d2rq:property loa:codigo;
  d2rq:column "rdf.unidade_orcamentaria.cod_unidade_orcamentaria";
.

#### Object Properties ####
map:UnidadeOrcamentaria_orgao_cod a d2rq:PropertyBridge;
  d2rq:belongsToClassMap map:UnidadeOrcamentaria;
  d2rq:property loa:temOrgao;
  d2rq:refersToClassMap map:Orgao;
  d2rq:join "rdf.unidade_orcamentaria.cod_orgao => rdf.orgao.cod_orgao";
.
```

Figura 5.3 Mapeamento da classe `loa:UnidadeOrcamentaria`.

Após o arquivo de mapeamento ser escrito, é possível realizar a transformação concreta dos dados persistidos no banco de dados relacional para um único arquivo RDF. Essa conversão é realizada pela ferramenta *Dump-rdf*, sendo necessário apenas informar o formato de serialização do arquivo de saída, o padrão de URL a ser utilizado e o próprio arquivo de mapeamento. A linha de execução desse comando para realizar a triplificação dos dados do Orçamento Federal pode ser vista a seguir:

- **`./Dump-rdf -f N-TRIPLE -b http://orcamento.dados.gov.br/2013/id/mapping_loa.ttl > loa2013.nt`**

Ao fim da execução do comando, o arquivo `loa2013.nt` deve conter todos os dados do Orçamento Federal em RDF. O formato de serialização *N-TRIPLE* representa a informação de maneira simples, não fazendo uso de *namespaces* e abreviações. Entretanto, reduz drasticamente o tempo da triplificação se comparado a outros tipos mais utilizados, como por exemplo, o *TURTLE*, justificando assim sua escolha.

No início do projeto RDF, a ferramenta *Triplify* era utilizada para realizar a triplificação dos dados, porém apresentava um conjunto de limitações que tornava necessário efetuar alterações manuais no arquivo RDF gerado. As principais desvantagens e dificuldades encontradas na utilização do *Triplify* foram:

- gerar instâncias com um padrão definido de URI;
- criar relações entre classes presentes em mais de uma tabela no banco de dados relacional;
- triplificar colunas que apresentam valor igual a zero.

A ferramenta D2RQ se mostrou muito superior, pois sua linguagem de mapeamento permite especificar com detalhes as relações existentes entre o modelo relacional e o modelo ontológico. Após a definição desse arquivo de configuração, a triplificação pode ser realizada com apenas um comando, como mostrado nas seções anteriores.

5.1.3 Carga do *Triple store*

Nessa etapa do processo, o arquivo RDF gerado pela triplificação é carregado em um *triple store* TDB, ou seja, todas as informações deverão ser persistidas em um banco de dados específico para grafos. O *upload* dos arquivos para o *triple store* pode ser realizado de duas maneiras diferentes, por intermédio do *Jena Framework*.

O servidor *Fuseki* disponibiliza um *Endpoint* para que usuários e aplicações realizem consultas SPARQL em um *triple store* TDB. Uma das ferramentas presentes nesse serviço é o *s-put*, que permite carregar um arquivo RDF em um grafo. Nesse caso, é necessário informar o endereço do *Endpoint*, a URL referente ao grafo que será criado e o arquivo RDF, a saber:

- **`./s-put http://localhost:3030/ds/data http://orcamento.dados.gov.br/2013/loa2013.nt`**

O comando descrito ilustra a execução do *s-put*, no qual o arquivo *loa2013.nt* é carregado em um *Endpoint* local. Caso o grafo informado já exista, os seus dados são excluídos e as

novas informações são inseridas em sequência. Para a utilização dessa ferramenta, é necessário que o servidor *Fuseki* esteja disponível, portanto a transferência do arquivo RDF ocorre por meio do protocolo HTTP.

Outra forma de realizar o *upload* de arquivos RDF para um *triple store* TDB é utilizando a ferramenta *tdbloader*. Ao contrário do *s-put*, esse comando permite carregar os registros do arquivo RDF sem a necessidade de utilizar a rede, representando assim uma alternativa *standalone*. Seu desempenho também é consideravelmente superior, pois é capaz de carregar cerca de 50 mil triplas por segundo.

O *tdbloader* compõe um conjunto de ferramentas que estão presentes no projeto *Apache Jena*. Por apresentar um excelente desempenho, foi escolhido para realizar a carga de arquivos RDF no *triple store* TDB da SOF. Para seu funcionamento, é necessário informar a localização do diretório TDB e o arquivo RDF. Caso a cláusula “*-graph*” não seja informada, os registros serão salvos no grafo *default*. A execução do comando pode ser vista a seguir:

- **`./tdbloader --loc=$FUSEKI_HOME/tdb_2000-2013/ $D2RQ_HOME/ loa2013.nt`**

Para que o TDB possa armazenar os dados, ele constrói uma tabela de nós, que relaciona cada nó (recurso RDF) com uma chave de 8 bytes que representa um número inteiro, e vice-versa. Uma tabela de índices é criada a partir dessas chaves, assim é possível melhorar as consultas SPARQL que são realizadas nesse *triple store*. Quando ocorre alguma exclusão de triplas, os registros são removidos apenas da tabela de índices, ou seja, ainda ocupam espaço na tabela de nós.

O arquivo RDF contendo os dados da execução financeira é gerado diariamente, os registros do dia anterior sempre devem ser excluídos e atualizados com a posição atual. Devido ao fato relatado no parágrafo anterior, os dados do TDB não podem ser removidos apenas com algum *update*. Os registros seriam excluídos, mas não na tabela de nós, o que resulta em um aumento significativo de espaço em disco.

Dessa forma, os arquivos do diretório TDB devem ser excluídos a cada atualização da execução financeira em formato RDF. Todos os arquivos em RDF devem ser carregados novamente por intermédio do *tdbloader*.

5.1.4 SPARQL Endpoint

O servidor *Fuseki* oferece um *endpoint* de acesso aos dados armazenados no TDB, fornecendo a URL <http://host:3030/ds/query> acessível em uma rede para consulta dos dados utilizando SPARQL por qualquer aplicação. Além disso, o *Fuseki* também oferece certos serviços via uma interface *Web* acessível pela URL <http://host:3030/>, dentre elas existe uma página com *textareas* para enviar consultas SPARQL ou SPARUL, possibilitando que qualquer usuário possa enviar consultas ao *endpoint* via seu *browser*.

Estando com o *Fuseki* devidamente instalado e os dados RDF da LOA inseridos no TDB, os seguintes passos foram executados:

- a) Iniciação do serviço do *Fuseki*, indicando em que diretório está o TDB:

```
$ cd /path/to/Fuseki
```

```
$ ./Fuseki-Server --loc=/path/to/tdb /ds
```

- b) Acesso à URL <http://localhost:3030/>
- c) Acesso ao link “Control Panel”
- d) Acesso à função “Select”

Nesse momento, serão exibidos dois terminais: um para efetuar SPARQL *queries*, e outro para SPARQL *updates*, uma opção de inserção de arquivos RDF (*File upload*). A figura 5.4 ilustra um exemplo de consulta SPARQL, a qual faz um somatório de todos os valores orçamentários dos itens de despesa do ano de 2013.

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX siop: <http://vocab.e.gov.br/2012/08/loa#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>

SELECT (SUM(?loa) as ?totalLoa) (SUM(?ploa) as ?totalPloa) (SUM(?loacre) as ?totalCreditos)
      (SUM(?emp) as ?totalEmpenhado) (SUM(?liq) as ?totalLiquidado) (SUM(?pago) as ?totalPago)
WHERE
{
  ?item a siop:ItemDeDespesa;
  siop:temExercicio [
    siop:identificador 2013 ;
  ];
  siop:valorDotacaoInicial ?loa;
  siop:valorProjetoLei ?ploa;
  siop:valorLeiMaisCredito ?loacre;
  siop:valorEmpenhado ?emp;
  siop:valorLiquidado ?liq;
  siop:valorPago ?pago;
}

```

Figura 5.4 Exemplo de consulta SPARQL.

Dessa forma, qualquer usuário é capaz de efetuar consultas SPARQL sem a necessidade de criação de uma aplicação para tal.

Para realizar consultas nas informações presentes no *endpoint* citado e em conformidade com recomendação da W3C (W3C, 2008), a implementação final utiliza-se da linguagem de consulta SPARQL (*SPARQL Protocol and RDF Query Language*), criada de acordo com os requisitos estabelecidos pelo *RDF Data Access Use Cases and Requirements Working Group*, que estabelece a sintaxe que possibilita a criação de comandos para manipular grafos, sendo considerada uma tecnologia chave no contexto da *Web Semântica*.

5.1.5 Interface da aplicação

A estrutura padrão de arquitetura da informação e interface gráfica da tela inicial de acesso (*login*) do sistema SIOP foi dividido em:

- a) área de acesso público;
- b) área de formulário de acesso;
- c) área de informação ao usuário sobre o Sistema SIOP;
- d) área de serviços;
- e) *banners* de divulgação;
- f) rodapé de assinatura do sistema.

Para acessar o “Acesso ao Cidadão”, no SIOP a partir do acesso público, deve-se clicar em um dos dois botões destacados na Figura 5.5.



Figura 5.5 Tela inicial de acesso.

Para acessar os “Dados Abertos RDF”, deve-se clicar no botão destacado na Figura 5.6.

Ministério do Planejamento, Orçamento e Gestão

SIOP Sistema Integrado de Planejamento e Orçamento

SIOPLEGIS Pesquisar Legislação...

Acesso Público | Solicitação de Acesso | Fique por Dentro | Serviços | Suporte ao Usuário

Você está aqui: Início » Acesso Público

CONSULTA PÚBLICA AO ORÇAMENTO FEDERAL

A Secretaria de Orçamento Federal – SOF, alinhada a compromissos definidos pela Lei de Acesso à Informação – LAI (Lei nº 12.527/2011), que estabeleceu o uso de tecnologia em prol da transparência, publica os dados da LOA 2012 em formato aberto (formato RDF).

Consulta Livre ao Orçamento

Consulte informações sobre o Orçamento da União, com várias opções de filtro (palavra chave, órgão, unidade, função, subfunção) e gere relatórios em PDF e Excel.

[Consulta Livre](#)

Minha Consulta ao Orçamento

Esta funcionalidade do Acesso Público possibilita a criação e reutilização de consultas personalizadas, bem como a publicação de consultas para serem acessadas por outros usuários da sociedade.

[Consulta Personalizada](#)

Dados Abertos do Orçamento Federal

Baixe aqui o Orçamento Federal dos anos anteriores.

Escolha um Exercício

[Manual de Referência](#) [Carta de Serviços](#)

[Ontologia do Orçamento Federal Brasileiro em OWL](#)

Figura 5.6 Tela de acesso ao Dados abertos.

Foram implementados dois tipos de consultas aos Dados Abertos – RDF, conforme a Figura 5.6:

- Consulta Livre: para pesquisa por palavra-chave, órgão, unidade, função e subfunção.
- Consulta Personalizada: criação de consultas personalizadas que podem ser armazenadas e compartilhadas com outros usuários.

Ainda na Figura 5.6, foi implementada uma área de conteúdo mais específica para usuários técnicos, com acesso aos itens:

- Orçamento Federal dos anos anteriores: acesso ao orçamento em formato RDF dos anos 2000 a 2014.
- Ontologia do Orçamento em formato OWL.
- Manual de Referência da Aplicação.
- Carta de Serviço ao Usuários do Sistema.

5.1.5.1 Interface de Consulta Livre

A Consulta Livre permite que diversos parâmetros sejam configurados para a busca da informação desejada, conforme a Figura 5.7.

Selecione os filtros

Exercício: 2013 Palavra chave:

Órgão: 47000 - M. do Plan. Orçamento e Gestão Unidade: 47101 - M. do Plan. Orçamento e Gestão

[Pesquisa Avançada ▲](#)

Pesquisa Avançada

Programa: Todos Ação: Todas

Plano orçamentário: Selecione uma ação Função: Todas

Subfunção: Todas

Resultados para a pesquisa solicitada

Exercício	Unidade	Programa	Ação	Plano Orçamentário
2013	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	20U0 - Gestão e Aprimoramento do Planejamento	@ - Valor não detalhado (Reabertura de crédito)
2013	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	20U0 - Gestão e Aprimoramento do Planejamento	0001 - Gestão da Tecnologia e da Informação de Planejamento e Investimentos Estratégicos
2013	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	20U0 - Gestão e Aprimoramento do Planejamento	0003 - Participação do Brasil no Conselho Sul-Americano de Infraestrutura e Planejamento da UNASUL (COSIPLAN)
2013	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	20U0 - Gestão e Aprimoramento do Planejamento	0004 - Elaboração, Monitoramento e Avaliação de Políticas, Planos e Programas de Governo, Gestão e Revisão do Plano Plurianual (PPA)
2013	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	8861 - Gestão e Aprimoramento do Processo Orçamentário	@ - Valor não detalhado (Reabertura de crédito)

Página 1 de 22 | Registros: 108 Itens por página 5

Figura 5.7 Formulário de consulta livre.

As opções desejadas são marcadas na caixa de seleção, indicadas na Figura 5.8, e automaticamente essas opções aparecem no quadro “Minha Seleção”.

Resultados para a pesquisa solicitada						
Exercício	Unidade	Programa	Ação	Plano Orçamentário		
<input type="checkbox"/>	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	20U0 - Gestão e Aprimoramento do Planejamento	@ - Valor não detalhado (Reabertura de crédito)		
<input checked="" type="checkbox"/>	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	20U0 - Gestão e Aprimoramento do Planejamento	0001 - Gestão da Tecnologia e da Informação de Planejamento e Investimentos Estratégicos		
<input checked="" type="checkbox"/>	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	20U0 - Gestão e Aprimoramento do Planejamento	0003 - Participação do Brasil no Conselho Sul-Americano de Infraestrutura e Planejamento da UNASUL (COSIPLAN)		
<input checked="" type="checkbox"/>	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	20U0 - Gestão e Aprimoramento do Planejamento	0004 - Elaboração, Monitoramento e Avaliação de Políticas, Planos e Programas de Governo, Gestão e Revisão do Plano Plurianual (PPA)		
<input type="checkbox"/>	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	8861 - Gestão e Aprimoramento do Processo Orçamentário	@ - Valor não detalhado (Reabertura de crédito)		

Página 1 de 22 | Registros: 108

Itens por página 5

Gerar Relatório

Minha Seleção

Exercício	Unidade	Programa	Ação	Plano Orçamentário			
2013	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	20U0 - Gestão e Aprimoramento do Planejamento	0003 - Participação do Brasil no Conselho Sul-Americano de Infraestrutura e Planejamento da UNASUL (COSIPLAN)			<input type="button" value="Excluir"/>
2013	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	20U0 - Gestão e Aprimoramento do Planejamento	0004 - Elaboração, Monitoramento e Avaliação de Políticas, Planos e Programas de Governo, Gestão e Revisão do Plano Plurianual (PPA)			<input type="button" value="Excluir"/>
2013	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	20U0 - Gestão e Aprimoramento do Planejamento	0001 - Gestão da Tecnologia e da Informação de Planejamento e Investimentos Estratégicos			<input type="button" value="Excluir"/>

Página 1 de 1 | Registros: 3

Itens por página 10

Figura 5.8 Formulário de seleção de consulta.

Conforme a Figura 5.8, clicando em “Gerar Relatório”, o sistema gerará um arquivo em formato PDF que poderá ser visualizado ou salvo pelo usuário. O relatório gerado pode ser observado na Figura 5.9 a seguir.

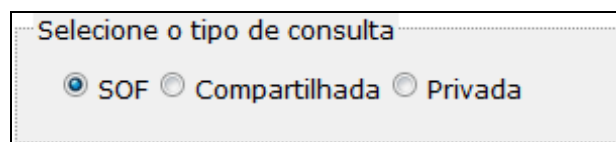
Consulta:	Autor:					
	Projeto Lei	Dotação Inicial	Dotação Atual	Empenhado	Liquidado	Pago
Orgão: 47000 - Ministério do Planejamento, Orçamento e Gestão	3.900.000,00	3.900.000,00	3.900.000,00	777.766,35	301.981,22	250.189,51
UO: 47101 - Ministério do Planejamento, Orçamento e Gestão	3.900.000,00	3.900.000,00	3.900.000,00	777.766,35	301.981,22	250.189,51
Programa: 2038 - Democracia e Aperfeiçoamento da Gestão Pública	3.900.000,00	3.900.000,00	3.900.000,00	777.766,35	301.981,22	250.189,51
04.121.2038.20U0 - Gestão e Aprimoramento do Planejamento	3.900.000,00	3.900.000,00	3.900.000,00	777.766,35	301.981,22	250.189,51
Localizador: 20U0.0001 - Nacional	3.900.000,00	3.900.000,00	3.900.000,00	777.766,35	301.981,22	250.189,51
PO: 0001 - Gestão da Tecnologia e da Informação de Planejamento e Investimentos Estratégicos	620.000,00	620.000,00	620.000,00	0,00	0,00	0,00
PO: 0003 - Participação do Brasil no Conselho Sul-Americano de Infraestrutura e Planejamento da UNASUL (COSIPLAN)	250.000,00	250.000,00	250.000,00	50.000,00	15.934,95	15.934,95
PO: 0004 - Elaboração, Monitoramento e Avaliação de Políticas, Planos e Programas de Governo, Gestão e Revisão do Plano Plurianual (PPA)	3.030.000,00	3.030.000,00	3.030.000,00	727.766,35	286.046,27	234.254,56

Figura 5.9 Relatório gerado.

5.1.5.2 Interface de Consulta Personalizada

A Consulta Personalizada possibilita a criação e reutilização de consultas específicas, bem como a publicação de consultas para serem acessadas por outros usuários da sociedade.

Existem três tipos de consultas disponíveis de acordo com a Figura 5.10:



Selecione o tipo de consulta

SOF Compartilhada Privada

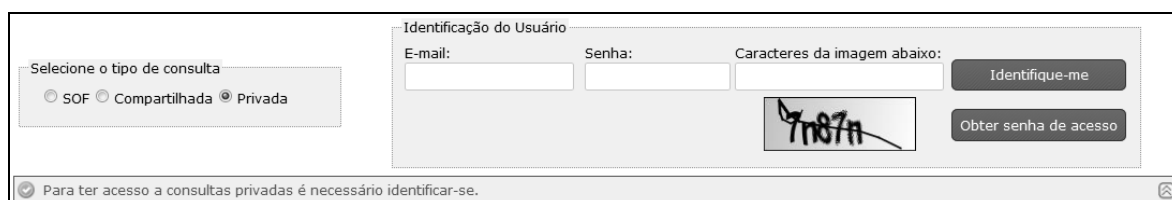
Figura 5.10 Formulário de seleção de tipo de consulta.

- SOF: Consultas criadas pela SOF. Todos poderão utilizar essa consulta.
- Compartilhada: consultas criadas por pessoas previamente cadastradas. A solicitação de cadastro deve ser feita à SOF. Todos poderão utilizar essa consulta.
- Privada: consultas criadas pelo usuário e que serão utilizadas exclusivamente por ele. É preciso cadastrar um *e-mail* e senha no SIOP.

A forma de busca e execução da consulta será igual para os três tipos de consultas personalizadas.

O exemplo na Figura 5.11 será o da consulta privada, pois além da busca e execução, esse tipo inclui também o armazenamento de uma nova consulta.

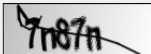
O usuário da consulta privada deve primeiramente identificar-se através de seu *e-mail* e senha.



Identificação do Usuário

Selecione o tipo de consulta

SOF Compartilhada Privada

E-mail: Senha: Caracteres da imagem abaixo: 

Identifique-me

Obter senha de acesso

Para ter acesso a consultas privadas é necessário identificar-se.

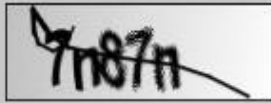
Figura 5.11 Formulário de identificação do usuário.

Caso ainda não haja cadastro, basta clicar em “Obter senha de acesso” e informar um *e-mail* e a digitação de caracteres de validação. Clicando em “Obter senha”, o sistema gerará uma senha temporária e a enviará ao *e-mail* informado, conforme a Figura 5.12.

Obter senha de acesso

E-mail:

Caracteres da imagem abaixo:



Obter Senha

Cancelar

Figura 5.12 Formulário de solicitação de senha de acesso.

A senha temporária deve ser trocada por uma definitiva. Para isso, basta clicar em “Trocar Senha”, conforme a Figura 5.13.

Identificação do Usuário

E-mail: Sair Trocar Senha

Senha Atual:

Nova Senha:

Confirmação Nova Senha:

Trocar Senha

Cancelar

Figura 5.13 Formulário de identificação do usuário.

Sendo identificado, a criação de consultas será feita por intermédio do botão “Adicionar Consulta Personalizada Privada”, conforme a Figura 5.14.

Figura 5.14 Adicionando uma consulta personalizada.

Os filtros da consulta deverão ser determinados e em seguida deve-se clicar em “Salvar”, conforme a Figura 5.15.

Figura 5.15 Formulário de gerenciamento de consultas.

A consulta salva ficará disponível para o usuário sempre que ele se identificar no sistema.

A Figura 5.16 mostra uma consulta privada, porém as consultas do tipo SOF e compartilhada aparecerão da mesma maneira.

Consulta Personalizada Privada		
Adicionar Consulta Personalizada Privada Excluir		
Código	Título	Criador
<input type="checkbox"/>	0A1B Ações do Ministério do Planejamento	@planejamento.gov.br

Página 1 de 1 | Registros: 1 Itens por página 10

Figura 5.16 Formulário de consultas gravadas.

Para executar a consulta, basta clicar em sua linha correspondente. Os filtros serão automaticamente carregados e o usuário deverá clicar em “Pesquisar”, como pode ser observado na Figura 5.17.

Cadastro

Voltar Salvar Copiar Consulta

Identificação da Pesquisa:

Código*: 0A1B Título*: Ações do Ministério do Planejamento

Descrição: Exemplo de consulta. Ações do Ministério do Planejamento.

Ocultar a área de pesquisa

Selecione os filtros

Exercício: 2013 Palavra chave:

Órgão: 47000 - M. do Plan. Orçamento e Gestão Unidade: 47101 - M. do Plan. Orçamento e Gestão

Pesquisa Avancada

Pesquisar

Figura 5.17 Formulário de consultas gravadas.

O resultado da pesquisa será exibido em seguida, conforme a Figura 5.18

Resultados para a pesquisa solicitada					
Exercício	Unidade	Programa	Ação	Plano Orçamentário	
<input type="checkbox"/>	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	20U0 - Gestão e Aprimoramento do Planejamento	@ - Valor não detalhado (Reabertura de crédito)	
<input type="checkbox"/>	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	20U0 - Gestão e Aprimoramento do Planejamento	0001 - Gestão da Tecnologia da Informação de Planejamento e Investimentos Estratégicos	
<input type="checkbox"/>	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	20U0 - Gestão e Aprimoramento do Planejamento	0003 - Participação do Brasil no Conselho Sul-Americano de Infraestrutura e Planejamento da UNASUL (COSIPLAN)	
<input type="checkbox"/>	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	20U0 - Gestão e Aprimoramento do Planejamento	0004 - Elaboração, Monitoramento e Avaliação de Políticas, Planos e Programas de Governo, Gestão e Revisão do Plano Plurianual (PPA)	
<input type="checkbox"/>	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	8861 - Gestão e Aprimoramento do Processo Orçamentário	@ - Valor não detalhado (Reabertura de crédito)	

Página 1 de 22 | Registros: 108

Itens por página: 5

Minha Seleção					
Exercício	Unidade	Programa	Ação	Plano Orçamentário	
Atenção! Não foi encontrado nenhum item para o filtro informado. Tente novamente					

Figura 5.18 Tela de resultado da consulta.

A seleção do que estará efetivamente nos relatórios é feita por meio da marcação das caixas de seleção ao lado dos itens desejados.

A Figura 5.19 exibe dois itens selecionados.

Resultados para a pesquisa solicitada

Exercício	Unidade	Programa	Ação	Plano Orçamentário
<input type="checkbox"/>	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	20U0 - Gestão e Aprimoramento do Planejamento	@ - Valor não detalhado (Reabertura de crédito)
<input checked="" type="checkbox"/>	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	20U0 - Gestão e Aprimoramento do Planejamento	0001 - Gestão da Tecnologia e da Informação de Planejamento e Investimentos Estratégicos
<input checked="" type="checkbox"/>	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	20U0 - Gestão e Aprimoramento do Planejamento	0003 - Participação do Brasil no Conselho Sul-Americano de Infraestrutura e Planejamento da UNASUL (COSIPLAN)
<input type="checkbox"/>	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	20U0 - Gestão e Aprimoramento do Planejamento	0004 - Elaboração, Monitoramento e Avaliação de Políticas, Planos e Programas de Governo, Gestão e Revisão do Plano Plurianual (PPA)
<input type="checkbox"/>	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	8861 - Gestão e Aprimoramento do Processo Orçamentário	@ - Valor não detalhado (Reabertura de crédito)

Página 1 de 22 | Registros: 108

Itens por página 5

Gerar Relatório

Minha Seleção

Exercício	Unidade	Programa	Ação	Plano Orçamentário	
2013	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	20U0 - Gestão e Aprimoramento do Planejamento	0001 - Gestão da Tecnologia e da Informação de Planejamento e Investimentos Estratégicos	<input type="button" value="Excluir"/>
2013	47101 - Ministério do Planejamento, Orçamento e Gestão	2038 - Democracia e Aperfeiçoamento da Gestão Pública	20U0 - Gestão e Aprimoramento do Planejamento	0003 - Participação do Brasil no Conselho Sul-Americano de Infraestrutura e Planejamento da UNASUL (COSIPLAN)	<input type="button" value="Excluir"/>

Página 1 de 1 | Registros: 2

Itens por página 10

Figura 5.19 Formulário de seleção de consultas.

Ainda na Figura 5.19, clicando em “Gerar Relatório”, o sistema gerará um arquivo em formato PDF que poderá ser visualizado ou salvo pelo usuário, conforme exibido na Figura 5.20.

	Projeto Lei	Dotação Inicial	Dotação Atual	Empenhado	Liquidado	Pago
Orgão: 47000 - Ministério do Planejamento, Orçamento e Gestão	870.000,00	870.000,00	870.000,00	50.000,00	15.934,95	15.934,95
UO: 47101 - Ministério do Planejamento, Orçamento e Gestão	870.000,00	870.000,00	870.000,00	50.000,00	15.934,95	15.934,95
Programa: 2038 - Democracia e Aperfeiçoamento da Gestão Pública	870.000,00	870.000,00	870.000,00	50.000,00	15.934,95	15.934,95
04.121.2038.20U0 - Gestão e Aprimoramento do Planejamento	870.000,00	870.000,00	870.000,00	50.000,00	15.934,95	15.934,95
Localizador: 20U0.0001 - Nacional	870.000,00	870.000,00	870.000,00	50.000,00	15.934,95	15.934,95
PO: 0001 - Gestão da Tecnologia e da Informação de Planejamento e Investimentos Estratégicos	620.000,00	620.000,00	620.000,00	0,00	0,00	0,00
PO: 0003 - Participação do Brasil no Conselho Sul-Americano de Infraestrutura e Planejamento da UNASUL (COSIPLAN)	250.000,00	250.000,00	250.000,00	50.000,00	15.934,95	15.934,95

Figura 5.20 Relatório em formato PDF.

5.2 RESULTADOS OBTIDOS

A evolução da ontologia do orçamento com a adição da classe “Exercício” possibilitou consultas de exercícios anteriores, ampliando o universo de dados disponibilizados e consequentemente analisados.

Os atributos: “valorEmpenhado”, “valorLiquidado” e “valorPago” agora são populados diariamente por intermédio de extração de dados do SIOP e do SIAFI que fornecem os dados necessários para a geração diária do RDF de execução orçamentária com 4.484.487 triplas disponíveis no portal <https://www1.siop.planejamento.gov.br/acessopublico> e distribuídas entre as classes, conforme a Tabela 5.1:

Tabela 5.1 Quantidade de triplas por classe e quantidade total.

Classe	Quantidade de triplas
loa:ItemDeDespesa	4.393.782
loa:Exercicio	3
loa:Esfera	6
loa:Orgao	159
loa:UnidadeOrcamentaria	1.640
loa:Funcao	87
loa:Subfuncao	300
loa:Programa	354
loa:Acao	7.044
loa:Localizador	44.568
loa:PlanoOrcamentario	35.860
loa:IndicadorUso	21
loa:FonteRecursos	315
loa:CategoriaEconomica	9
loa:GrupoNatDespesa	21
loa:ModalidadeAplicacao	45
loa:ElementoDespesa	216
loa:ResultadoPrimario	12
rdfs:Class e rdf:Property	45
Total	4.484.487

As sentenças RDF foram criadas tomando como base a definição de conceitos e relações da ontologia OWL descrita na proposta (Capítulo 4) e deram origem a uma descrição RDF serializada em formato *N-Triples* que está apresentada na Tabela 5.2.

Tabela 5.2 Item de despesa da LOA 2013 em RDF e serializada em *N-Triples*.

	Sujeito	Predicado	Objeto
1	<http://orcamento.dados.gov.br/2013/id/ItemDeDespesa/66165>	<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>	<http://vocab.e.gov.br/2013/09/loa#ItemDeDespesa> .
2	<http://orcamento.dados.gov.br/2013/id/ItemDeDespesa/66165>	<http://vocab.e.gov.br/2013/09/loa#temAcao>	<http://orcamento.dados.gov.br/2013/id/Acao/20DB> .
3	<http://orcamento.dados.gov.br/2013/id/ItemDeDespesa/66165>	<http://vocab.e.gov.br/2013/09/loa#temCategoriaEconomica>	<http://orcamento.dados.gov.br/2013/id/CategoriaEconomica/4> .
4	<http://orcamento.dados.gov.br/2013/id/ItemDeDespesa/66165>	<http://vocab.e.gov.br/2013/09/loa#temIndicadorUso>	<http://orcamento.dados.gov.br/2013/id/IndicadorUso/0> .
5	<http://orcamento.dados.gov.br/2013/id/ItemDeDespesa/66165>	<http://vocab.e.gov.br/2013/09/loa#temElementoDespesa>	<http://orcamento.dados.gov.br/2013/id/ElementoDespesa/20> .
6	<http://orcamento.dados.gov.br/2013/id/ItemDeDespesa/66165>	<http://vocab.e.gov.br/2013/09/loa#temGND>	<http://orcamento.dados.gov.br/2013/id/GrupoNatDespesa/4> .
7	<http://orcamento.dados.gov.br/2013/id/ItemDeDespesa/66165>	<http://vocab.e.gov.br/2013/09/loa#temExercicio>	<http://orcamento.dados.gov.br/2013/id/Exercicio/2013> .
8	<http://orcamento.dados.gov.br/2013/id/ItemDeDespesa/66165>	<http://vocab.e.gov.br/2013/09/loa#temFonteRecursos>	<http://orcamento.dados.gov.br/2013/id/FonteRecursos/100> .
9	<http://orcamento.dados.gov.br/2013/id/ItemDeDespesa/66165>	<http://vocab.e.gov.br/2013/09/loa#temFuncao>	<http://orcamento.dados.gov.br/2013/id/Funcao/27> .
10	<http://orcamento.dados.gov.br/2013/id/ItemDeDespesa/66165>	<http://vocab.e.gov.br/2013/09/loa#temEsfera>	<http://orcamento.dados.gov.br/2013/id/Esfera/10> .
11	<http://orcamento.dados.gov.br/2013/id/ItemDeDespesa/66165>	<http://vocab.e.gov.br/2013/09/loa#temLocalizador>	<http://orcamento.dados.gov.br/2013/id/Localizador/5110120DB0001> .
13	<http://orcamento.dados.gov.br/2013/id/ItemDeDespesa/66165>	<http://vocab.e.gov.br/2013/09/loa#temUnidadeOrcamentaria>	<http://orcamento.dados.gov.br/2013/id/UnidadeOrcamentaria/51101> .
14	<http://orcamento.dados.gov.br/2013/id/ItemDeDespesa/66165>	<http://vocab.e.gov.br/2013/09/loa#temModalidadeAplicacao>	<http://orcamento.dados.gov.br/2013/id/ModalidadeAplicacao/90> .
15	<http://orcamento.dados.gov.br/2013/id/ItemDeDespesa/66165>	<http://vocab.e.gov.br/2013/09/loa#temSubfuncao>	<http://orcamento.dados.gov.br/2013/id/Subfuncao/811> .
16	<http://orcamento.dados.gov.br/2013/id/ItemDeDespesa/66165>	<http://vocab.e.gov.br/2013/09/loa#temResultadoPrimario>	<http://orcamento.dados.gov.br/2013/id/ResultadoPrimario/2> .
17	<http://orcamento.dados.gov.br/2013/id/ItemDeDespesa/66165>	<http://vocab.e.gov.br/2013/09/loa#temPrograma>	<http://orcamento.dados.gov.br/2013/id/Programa/2035> .
18	<http://orcamento.dados.gov.br/2013/id/ItemDeDespesa/66165>	<http://vocab.e.gov.br/2013/09/loa#temPlanoOrcamentario>	<http://orcamento.dados.gov.br/2013/id/PlanoOrcamentario/565> .
19	<http://orcamento.dados.gov.br/2013/id/ItemDeDespesa/66165>	<http://vocab.e.gov.br/2013/09/loa#valorEmpenhado>	“2201211.28”^^<http://www.w3.org/2001/XMLSchema#decimal> .
20	<http://orcamento.dados.gov.br/2013/id/ItemDeDespesa/66165>	<http://vocab.e.gov.br/2013/09/loa#valorLiquidadado>	“1892881.28”^^<http://www.w3.org/2001/XMLSchema#decimal> .
21	<http://orcamento.dados.gov.br/2013/id/ItemDeDespesa/66165>	<http://vocab.e.gov.br/2013/09/loa#valorPago>	“1892881.28”^^<http://www.w3.org/2001/XMLSchema#decimal> .

Na tabela 5.2, cada linha representa uma única tripla, em que o sujeito de todas as triplas descritas são o mesmo recurso <http://orcamento.dados.gov.br/2013/id/ItemDeDespesa/66165>. A linha 1 define que o recurso é do tipo “ItemDeDespesa”, da linha 2 até a linha 18 são definidos quais os classificadores que esse item de despesa tem relação; da linha 19 a 21 são definidos os valores financeiros que o item possui.

Conforme destacado no estado da arte, o formato de serialização *N-Triples* não possui prefixos para os *namespaces*, então eles ficam descritos completamente nos recursos, classes e propriedades. Nessa parte explicativa do RDF de 2013 estão contidos 4 diferentes *namespaces*, (1) <http://orcamento.dados.gov.br/2013/CLASSE/CODIGO> - Representam os recursos (instâncias) dos dados orçamentários, (2) http://vocab.e.gov.br/2013/09/loa#PROPRIEDADE_OU_CLASSE - Representam as propriedades e classes da ontologia da Classificação das Despesas do Orçamento Federal Brasileiro definidos no arquivo OWL disponibilizado, (3) <http://www.w3.org/1999/02/22-rdf-syntax-ns#PROPRIEDADE> - Representa o vocabulário padrão do RDF, neste caso foi utilizada a sua propriedade *type* para indicar a que classe pertence a um recurso, (4) <http://www.w3.org/2001/XMLSchema#PROPRIEDADE> - Representa o vocabulário padrão de definição de tipos de dados.

Tabela 5.3 Item de despesa da LOA 2013 em RDF e serializada em *N-Triples* (Continuação)

	Sujeito	Predicado	Objeto
1	http://orcamento.dados.gov.br/2013/id/Acao/20DB	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://vocab.e.gov.br/2013/09/loa#Acao .
2	http://orcamento.dados.gov.br/2013/id/Acao/20DB	http://vocab.e.gov.br/2013/09/loa#codigo	“20DB” .
3	http://orcamento.dados.gov.br/2013/id/Acao/20DB	http://www.w3.org/1999/02/22-rdf-syntax-ns#label	“Apoio à Realização da Copa do Mundo FIFA 2014” .
4	http://orcamento.dados.gov.br/2013/id/Programa/2035	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://vocab.e.gov.br/2013/09/loa#Programa .
5	http://orcamento.dados.gov.br/2013/id/Programa/2035	http://vocab.e.gov.br/2013/09/loa#codigo	“2035” .
6	http://orcamento.dados.gov.br/2013/id/Programa/2035	http://www.w3.org/1999/02/22-rdf-syntax-ns#label	“Esporte e Grandes Eventos Esportivos” .

7	<http://orcamento.dados.gov.br/2013/id/GrupoNatDespesa/4>	<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>	<http://vocab.e.gov.br/2013/09/loa#GrupoNatDespesa> .
8	<http://orcamento.dados.gov.br/2013/id/GrupoNatDespesa/4>	<http://vocab.e.gov.br/2013/09/loa#codigo>	“4” .
9	<http://orcamento.dados.gov.br/2013/id/GrupoNatDespesa/4>	<http://www.w3.org/1999/02/22-rdf-syntax-ns#label>	“Investimentos” .
10	<http://orcamento.dados.gov.br/2013/id/Funcao/27>	<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>	<http://vocab.e.gov.br/2013/09/loa#Funcao> .
11	<http://orcamento.dados.gov.br/2013/id/Funcao/27>	<http://vocab.e.gov.br/2013/09/loa#codigo>	”27” .
12	<http://orcamento.dados.gov.br/2013/id/Funcao/27>	<http://www.w3.org/1999/02/22-rdf-syntax-ns#label>	“Desporto e Lazer” .

A tabela 5.3 traz a parte do arquivo onde estão listadas as propriedades e atributos de alguns dos recursos que se relacionaram com o item de despesa explicado anteriormente, mais especificamente foram definidos os seus tipos, códigos e descrições. Na Figura 5.21 estão apresentadas, de forma gráfica, as sentenças RDF relativas ao exemplo das tabelas 5.2 e 5.3 e seus respectivos tipos e significados diante da especificação OWL descrita na proposta de modelo ontológico no item 4.2.

Dessa forma, o arquivo RDF resultante contém linhas com uma sentença para registrar que um determinado Item de Despesa tem:

- Valor Empenhado igual a R\$ 2.201.211,28.
- Valor Liquidado igual a R\$ 1.892.881,28.
- Valor Pago igual a R\$ 1.892.881,28.

Também se pode observar várias outras sentenças para afirmar características desses valores, tais como o programa, a ação, a função e os demais classificadores vinculados a esse item de despesa.

A Figura 5.21 ilustra um diagrama com as sentenças RDF citadas como exemplo nas tabelas 5.2/5.3, bem como tipos e significados diante da especificação OWL, apresentada no item 4.2 Figura 4.3.

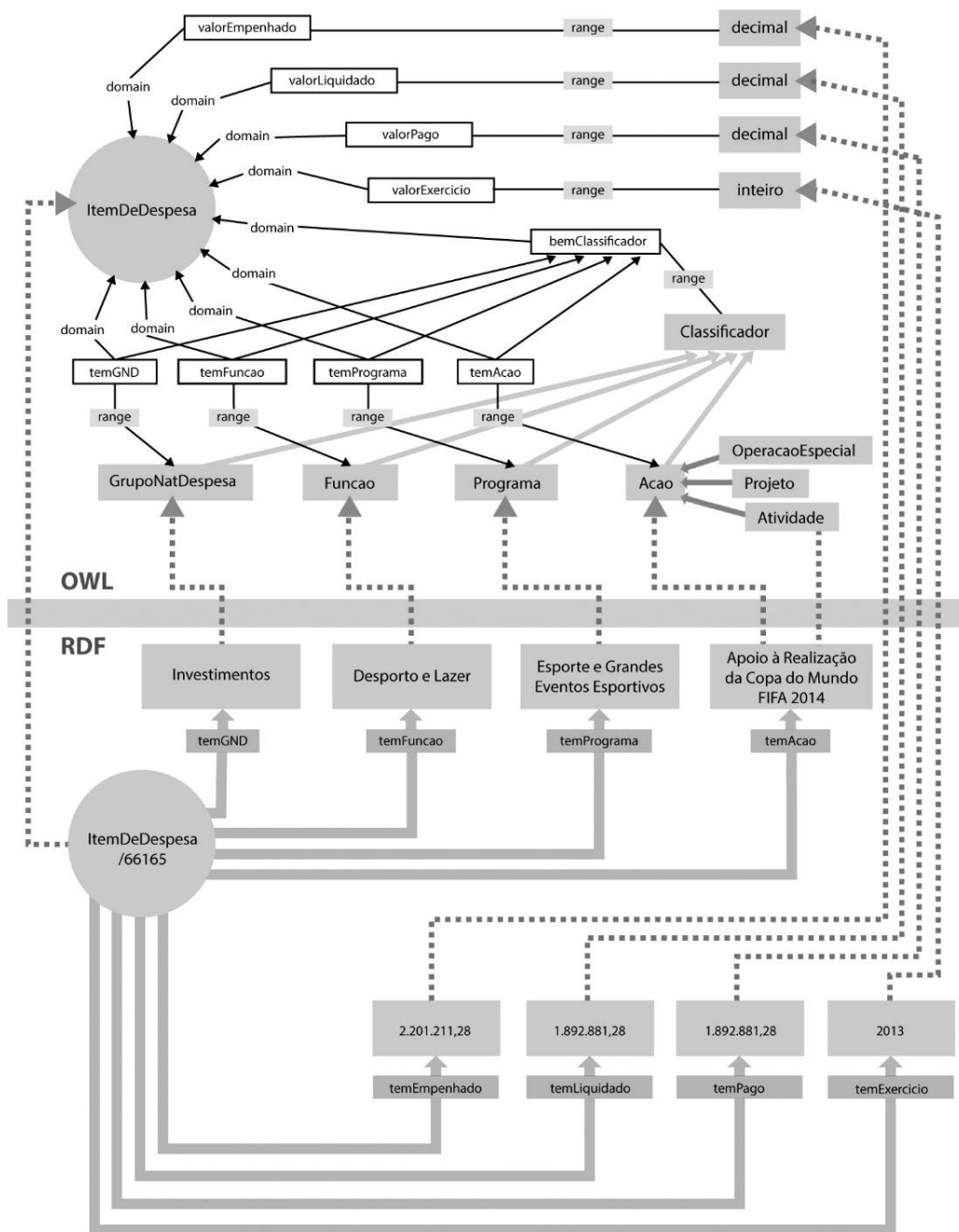


Figura 5.21 Mapeamento de informações entre os arquivos de sentenças RDF e OWL.

Como parte da iniciativa brasileira de dados abertos, o resultado desse projeto foi disponibilizado pelo grupo responsável pelo portal em um *endpoint* acessível na *internet* (orcamento.dados.gov.br/sparql) e uma aplicação que oferece uma interface de navegação pelos dados da LOA (orcamento.dados.gov.br). O primeiro propiciou o uso dos dados por

usuários técnicos e o segundo para usuários comuns. As duas iniciativas tiveram como insumo o arquivo RDF gerado da LOA.

Nesse sentido, o arquivo RDF gerado não somente alimenta o *triple store* que oferece os dados em uma consulta via *endpoint*, mas também promove a interoperabilidade com outras arquiteturas compatíveis com RDF.

6 - CONCLUSÕES

Os resultados alcançados neste trabalho são mais um passo para facilitar a acessibilidade e ampliar a democratização da informação governamental no Brasil, uma vez que possibilitam o acesso simplificado ao conteúdo das descrições das ações orçamentárias, permitindo a visualização do orçamento público e sua execução diária, ajudando a promover cada vez mais o debate e a participação da sociedade organizada na aplicação dos recursos federais de forma transparente e de fácil acesso.

Pela aplicação do modelo proposto, fica disponível, até mesmo ao cidadão não técnico, o uso de uma ferramenta de pesquisa e interpretação das informações do orçamento federal de forma transparente. Já em mãos especializadas, reconhecido o potencial de disponibilização da informação estruturada em RDF, esses dados podem ser explorados de diversas formas adicionais por meio de ferramentas de *On-line Analytical Processing* (OLAP), acessando como serviços os modelos RDF nos *endpoints* definidos.

A utilização de ontologias visa evoluir o processo de busca, recuperação e compartilhamento da informação, ao caracterizar e criar taxonomias sobre conceitos de um domínio. Quando se define classes, relações, axiomas etc, possibilita-se a redução de diferenças terminológicas ou semânticas, possibilitando a criação de um vocabulário comum que promove a eficiência do processamento automático da informação.

A arquitetura proposta possibilita não só a geração de dados estruturados, mas a capacidade de criação de indicadores de eficiência do gasto público por intermédio de uma ontologia modelada para o acompanhamento da execução orçamentária.

No cumprimento dos objetivos específicos, temos:

- A atualização da ontologia de domínio do Orçamento Federal Brasileiro para exploração dos dados de execução orçamentária, pela inclusão da classe “Exercicio” e dos atributos “valorEmpenhado”, “valorLiquidadado” e “valorPago”.
- A estruturação de um processo automatizado de extração, transformação e armazenamento dos dados orçamentários em formato compatível com os padrões

brasileiros de dados abertos, utilizando-se de solução de triplificação com base no *DumpRDF*, de armazenamento no TDB e disponibilização por meio de terminais *Fuseki*.

- A disponibilização de aplicações com base em clientes para acesso e exploração desses dados triplificados, que facilitam e tornam transparente ao usuário a construção de consultas SPARQL junto aos terminais *Fuseki* disponibilizados.

Com a proposta concretizada por intermédio da operacionalização da prova de conceito, fica cumprida a motivação deste trabalho que era a disponibilização de dados governamentais em formato aberto seguindo as recomendações da W3C.

Outra iniciativa que se concretiza com este trabalho de pesquisa é a disponibilização do RDF da LOA para publicação no portal Dados Abertos (www.dados.gov.br), promovendo a interoperabilidade entre os portais que utilizam plataformas diferentes para manipulação de arquivos RDF através de uma aplicação que oferece uma interface de navegação pelos dados da LOA (orcamento.dados.gov.br). O primeiro propiciando o uso dos dados por usuários técnicos e o segundo para usuários comuns.

O Brasil, comparado aos projetos do Reino Unido e EUA, ainda tem um longo caminho a percorrer em relação a dados abertos governamentais, visto que a quantidade de *data sets* disponíveis ainda é medida em dezenas. Outro ponto que tem que ser aprimorado é a definição de políticas de incentivo de desenvolvimento de aplicações *open source* para o consumo e processamento destes dados, pois para a população em geral, os dados estruturados em si não representam muita coisa, e o governo não conseguiria desenvolver *endpoints* em quantidade necessária que atendesse a toda a demanda informacional. Portanto, isso tem que feito de forma colaborativa.

6.1 TRABALHOS FUTUROS

Durante o estudo foram levantadas algumas questões com o intuito de explorar as oportunidades que não puderam ser tratadas neste trabalho. São elas:

- Desenvolvimento de uma camada de interface fundamentada em *meta-modelos*. Desta maneira o crescimento do modelo ontológico seria suportado de forma automática, possibilitando que a aplicação tenha uma evolução simplificada e baixo custo de manutenção.
- Mudança de ferramenta para disponibilização do SPARQL *Endpoint* - uma desvantagem de utilizar o SPARQL *Endpoint* oferecido pelo D2R-Server é que um de seus prerequisites é a existência de uma base relacional, para que os mapeamentos sejam feitos em tempo real. Em algumas situações, é interessante que o *Endpoint* não possua esta dependência, necessitando apenas de um conjunto de dados RDF para executar consultas. Desta forma, uma possível melhoria para este trabalho seria adotar uma ferramenta independente de banco de dados relacional para a disponibilização do SPARQL *Endpoint*.
- Integração com diversos *mashups* e APIs, como por exemplo, o Google MAPS, para exploração georeferenciada das triplas RDF com coordenadas geográficas.
- Implementação de axiomas de domínio. Os axiomas devem ser necessários e suficientes para expressar as questões de competência e para caracterizar suas soluções. Além disso, qualquer solução para uma questão de competência deve ser descrita pelos axiomas da ontologia e deve ser consistente com ele, com, por exemplo, uma restrição do seguinte tipo: uma ação de projeto não pode existir em duas esferas diferentes.
- Buscar soluções que auxiliem o desenvolvimento ontológico em todo o seu ciclo de vida: integração, fusão, reengenharia, avaliação, tradução etc.

ARTIGOS PUBLICADOS

1. Silva, D. A., Sousa Jr, R. T., Canedo, E. D., Martins, V. A., Medonça, F. L. L. and Veiga, C. E. L. Ontology Improvement Within The Brazilian Federal Government Budget Process. 9th International Conference on Knowledge Management, Montreal, Canada, 11/2013.
2. Silva, D. A., Sousa Jr, R. T., Veiga, C. E. L., Martins, V. A., and Exposto, E. N. A Extensão da Ontologia do Orçamento Federal com uma Classe de Geolocalização permite a Consideração de Indicadores de Desenvolvimento Humano nas Decisões da Política Orçamentária Brasileira, 9ª Conferência Ibérica de Sistemas e Tecnologias de Informação, Barcelona, Espanha, 06/2014.

REFERÊNCIAS BIBLIOGRÁFICAS

- Agune, R. M., Gregorio F., A. S., Bolliger, S. P. (2010). “Governo aberto SP: disponibilização de bases de dados e informações em formato aberto”. In: *III Congresso CONSAD de Gestão Pública*, Brasília, Brasil.
- Allemang, D. and Hendler, J. (2011). *Semantic Web for the Working Ontologist: Effective Modeling in RDFS*. ELSEVIER, second edition, *Massachusetts, United States*.
- Almeida, M. B., Bax, M. P. (2003). “Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção”. In: *Ciência da Informação*, Brasília, 32 (3), 7-20. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/View/17/11>>. Acessado em outubro de 2012.
- Antoniou, G. and Harmelen, F. (2008). “A Semantic Web Primer”. In: *The MIT Press*, Second Edition, London, England.
- Araújo, L., Cruz, F. W. (2012). “Uma ontologia das classificações da despesa do Orçamento Federal”. In: *V Seminário de Pesquisa em Ontologias do Brasil / VII International Workshop on Metamodels, Ontologies, Semantic Technologies*, Pernambuco, Brasil.
- Bax, M. P. (2001). “Introdução às linguagens de marcas”. In: *Ciência da Informação*, vol. 30, nº. 1, Brasília. Disponível em: <http://www.ibict.br/cienciainformacao/viewarticle.php?id=256>. Acessado em junho de 2012.
- Berners-Lee, T., Hendler, J., Lassila, O. (2001) “The Semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities”. In: *Scientific American*. New York. Disponível em: <<http://www.cs.nyu.edu/rgrimm/teaching/reading/semantic-Web.pdf>>. Acessado em fevereiro de 2013.
- Berners-Lee T. (2006). “Linked Data Design Issues”. Disponível em: <<http://www.w3.org/DesignIssues/LinkedData.html> W3C>. Acessado em março de 2014.
- Berners-Lee T. (2009), “Tim Berners-Lee on the Web”, TED (Technology, Entertainment,

Design) Talk, February 2009, Long Beach, California, USA, http://www.ted.com/talks/lang/eng/tim_berners_lee_on_the_next_web.htm.

Acessado em março de 2013.

Bizer, C., Cyganiak, R., Heath, T. (2007) “How to publish Linked Data on the Web”.

Disponível em: <<http://www4.wiwiss.fu-berlin.de/bizer/pub/LinkedDataTutorial/>>.

Acessado em Novembro de 2012.

Brandão, A. A. F., Lucena, C. J. P. (2002). “Uma Introdução à engenharia de ontologias no contexto da Web semântica”. Disponível em: <ftp://139.82.16.194/pub/docs/techreports/02_29_brandao.pdf>. Acessado em setembro de 2012.

Bräscher, M., Café, L. (2008). “Organização da Informação ou Organização do Conhecimento?” In: *Encontro Nacional de Pesquisa em Ciência da Informação. Anais...* São Paulo: ECA/USP. Programa de Pós-Graduação em Ciência da Informação. ENANCIB. 1835 pdf. Set./out. 2008. CD-ROM.

Brasil (1964) Lei nº 4.320, de 17 de março de 1964. Estatui Normas Gerais de Direito Financeiro para elaboração e controle dos orçamentos e balanços da União, dos Estados, dos Municípios e do Distrito Federal. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/l4320.htm>. Acessado em Maio de 2013

Brasil (2000). Lei Complementar nº 101, de 4 de maio de 2000. Estabelece normas de finanças públicas voltadas para a responsabilidade na gestão fiscal e dá outras providências. Disponível em: <www.planalto.gov.br/ccivil_03/leis/lcp/lcp101.htm>. Acessado em abril de 2013.

Brasil (2011) Lei nº 12.527, de 18 de novembro de 2011. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2011/Lei/L12527.htm>. Acessado em fevereiro de 2013.

Brasil (2012) “Manual Técnico Orçamentário”, Versão 2013, Brasília: Ministério do Planejamento, Orçamento e Gestão – Secretaria de Orçamento Federal. Disponível em: <http://www.orcamentofederal.gov.br/informacoes-orcamentarias/manual-tecnico/MTO_2013_4.pdf>. Acessado em Outubro de 2012.

Brasil. (2012a) Lei nº 12.708, de 17 de agosto de 2012. Dispõe sobre as diretrizes para a elaboração e execução da Lei Orçamentária de 2013 e dá outras providências. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2012/Lei/L12708.htm>. Acessado em maio de 2013.

- Bratt, S. (2006). "Semantic web and other w3c technologies to watch". Disponível em <[http://www.w3.org/2006/Talks/1023-sb-W3CTechSemWeb/#\(1\)](http://www.w3.org/2006/Talks/1023-sb-W3CTechSemWeb/#(1))>. Acesso em fevereiro de 2013.
- Breitman, K. (2005). *Web Semântica: a Internet do futuro*. LTC, Rio de Janeiro.
- Brickley, D., Guha, R. V. and McBride, B. (2004). "RDF Vocabulary Description Language 1.0: RDF Schema". W3C Recommendation. Disponível em <<http://www.w3.org/TR/rdf-Schema/>>. Acessado em Setembro de 2012.
- Buranarach, M. A. (2004). *Framework for the Organization and Discovery of Information Resources in a WWW Environment Using Association, Classification e Deduction*. Tese de Doutorado, University of Pittsburgh, 220p. Disponível em: <http://www.sis.pitt.edu/~dist/programs/PhD/dissertations/Buranarach_M_Dissertation_13Dec2004_revised.pdf>. Acessado em maio de 2013.
- Campos, M. L. A. A. (2001). *Organização de Unidades do Conhecimento em Hipertextos: o modelo conceitual como um espaço comunicacional para a realização da autoria*. Tese de Doutorado em Ciência da Informação. Programa de Pós-Graduação e Ciência da Informação da Escola de Comunicação da Universidade Federal do Rio de Janeiro, 198p.
- Campos, M. L. A., Gomes, H. E. (2008). "Taxonomia e Classificação: o princípio de categorização". In: *DataGramaZero*, Rio de Janeiro, **vol.9**, 213-222. Disponível em: <http://www.dgz.org.br/ago08/Art_01.htm>. Acessado em abril de 2013.
- Castells, M. (1999). *A sociedade em rede*. PAZ E TERRA. São Paulo.
- Cavalcanti, C.R. (1978). *Indexação e tesouro: metodologia e técnicas*. In: ABDF. Brasília.
- Cheng, H. G. (1997). *Representing and Reasoning about Semantic Conflicts in Heterogeneous Information Sources*. Phd, MIT. Massachusetts, USA.
- Dahlberg, I. (2006). "Knowledge organization: a new science?". In: *Knowledge Organization*, **vol. 33**, n° 4, 11-19.
- Dean, M. and Scheiber, G. "OWL Web Ontology Language Reference, W3C Recommendation 12 November 2009". Disponível em <<http://www.w3.org/TR/owl-ref/>>. Acessado em Setembro de 2013.
- De Masi, D. (2000). *A Sociedade Pós-Industrial*. 3ª. Edição. SENAC, São Paulo.
- Extensible markup language XML 1.0. Recommendation. (2004). Disponível em: <<http://w3.org/TR/REC-xml/>>. Acessado em setembro de 2012.

- Farquar, A., Fikes, R., Rice, R. (1996). "The Ontolingua Server: a Tool for Collaborative Ontology Construction". In: *Int. Journal of Human-Computer Studies*, 46, 707-727.
- Feitosa, A. (2006). "Organização da informação na Web: das tags à Web semântica". THESAURUS. Brasília.
- Felicíssimo, C. H. (2004). *Interoperabilidade Semântica na Web: Uma Estratégia para o Alinhamento Taxonômico de Ontologia*. Dissertação de Mestrado, PUC-Rio, 180p.
- Franconi, E. (2000). "First Order Logic and ER Schema". In: *University of Manchester Manchester, England*.
- Gomes, H. E. (1990). "Manual de elaboração de tesouros monolíngues". In: *Brasília: o Programa*. Estado, 78p.
- Gruber, T. R. (1996). "What is an ontology?". Disponível em: <<http://ksl.stanford.edu/kst/what-is-an-ontology.html>>. Acessado em setembro de 2012.
- Gruber, T. R. (1993). "A translation approach to portable ontology specifications". In: *Appeared in Knowledge Acquisition*, vol. 5, n. 2, 1-26. Disponível em: <<http://tomgruber.org/writing/ontolingua-kaj-1993.pdf>>. Acessado em julho de 2012.
- Guarino, N. (1998). "Formal Ontology and Information Systems". In: *Proceedings of FOIS'98*. N. GUARINO, IOS Press, Trento, Itália.
- Guinchat, C., Menou, M. (1994). "Introdução geral às ciências e técnicas da informação e documentação". FBB e IBICT. Brasília 540 p.
- Guizzardi, G. (2005). "Ontological foundations for structural conceptual models". In: *Enschede, The Netherlands*. CTIT, Telematica Institut, Enschede, Holanda.
- Hodge, G. (2000). "Systems of Knowledge Organization for Digital Libraries: beyond traditional authorities files". Washington, DC, CLIR. Disponível em: <<http://www.clir.org/pubs/reports/pub91/contents.htm>>. Acessado em maio 2013.
- IEEE Standards Information Network (2000). "IEEE 100. The authoritative dictionary of IEEE standards terms" IEEE, Seventh Edition. New York, USA.
- Ibict. (1984). "Diretrizes para elaboração de tesouros monolíngues". EDITORA. Brasília.
- Isele, R., Anja J., Bizer, C., Volz, J. (2010). "Silk - A Link Discovery Framework for the Web of Data". Disponível em: <<http://www4.wiwiss.fu-berlin.de/bizer/silk/>>. Acessado em abril de 2012.

- Jena. (2007). “A Semantic Web Framework”. Disponível em <<http://jena.sourceforge.net>>. Acessado em agosto de 2012
- Lancaster, F. W. (2004). “Indexação e resumos: teoria e prática”. 2^a ed. BRIQUET DE LEMOS LIVROS. Brasília 452p.
- Lathrop, D., Ruma, L. (2010). “Open Government: Transparency, Collaboration and Participation in Practice”. In: *O’Reilly*, California, USA.
- Martins, V. A., Costa, J. P. C. L. and Sousa Jr, R. T., (2012). “Architecture of a Collaborative Business intelligence Environment based on an Ontology Repository and Distributed Data Services”. In: *Proc. 4th International Conference on Knowledge Management and Information Sharing*, Barcelona, Spain.
- Missikoff, M., Velardi, P., Fabriani, P. (2003). “Text Mining Techniques to Automatically Enrich a Domain Ontology”. Disponível em: <<http://www.springerlink.com/content/m717014453174720/>>. Acessado em agosto de 2013.
- Necib, C. C. B., Freytag, C. J. (2004). “Using Ontologies for Database Query Reformulation”. In: *Proceeding on the 18 th conference on Advances in Databases and Information Systems (ADBIS’2004)*, Budapest, Hungary.
- Noy, N. F., Mcguinness, D. L. (2001). “Ontology Development 101: A Guide to Creating Your First Ontology”. In: *Stanford University*, Stanford. Disponível em: <<http://www.ksl.stanford.edu/people/dlm/papers/ontology101/ontology101-noy-mcguinness.html>>. Acessado em setembro de 2012.
- Noy, N. F., Mcguinness, D. L. (2001). “Ontology Development 101: A Guide to Creating Your First Ontology”. In: *Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, March 2001*. Disponível em: <http://protege.stanford.edu/publications/ontology_development/ontology101.pdf>. Acessado em setembro de 2012.
- Obama, B. (2009). “Memorandum on Transparency and Open Government”. Disponível em <www.whitehouse.gov/the_press_office/Transparency_and_Open_Government> Acessado em maio de 2012.
- OWL – Ontology Web Language OverView (2004). Disponível em: <<http://www.w3.org/TR/owl-features/>>. Acessado em Setembro de 2012.
- R. Harrison, D. Obst, and C.W. Chan. (2005) “Design of an ontology management framework”. *Cognitive Informatics. ICCI. Fourth IEEE Conference* 260- 266.

- Disponível em <<http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=10264>>. Acessado em dezembro de 2012.
- Manola, F.; Miller, E. (2004) “W3C Rdf Primer”. Disponível em: <<http://www.w3.org/TR/2004/REC-rdf-primer-20040>>. Acessado em maio de 2012.
- Sanches, O. (2004). “Dicionário de Orçamento, Planejamento e Áreas Afins”. PRISMA/OMS2ª ed. Brasília, Brasil.
- Sayão, L. F., Marcondes, C. H. (2008). “O desafio da interoperabilidade e as novas perspectivas para as bibliotecas digitais”. In: *TransInformação*, Campinas, **vol. 20**, nº 2, 133-148.
- Sheth, P. (1999). “Changing Focus on Interoperability in Information Systems: From System, Syntax, Structure to Semantics. Interoperating Geographic Information Systems”. Edited by Good-child M. & Egenhofer, M. & Fegeas, R. & Kottman, C. Kluwer. ACADEMIC PUBLISHERS, Massachusetts, USA. 5-29.
- Silva, D. B. (2010). *Transparência na esfera pública interconectada* Dissertação de Mestrado, Faculdade Cásper Líbero, São Paulo, 114p.
- Souza, R. R., Alvarenga, L. (2004). “A Web Semântica e suas contribuições para a ciência da informação”. In: *Ciência da Informação*, Brasília, **vol. 33**, nº 1, 132-141. Disponível em: <<http://www.ibict.br/cionline/viewarticle.php?id=71>> Acessado em março de 2013.
- User guide: contents (2005). Disponível em: <http://protege.stanford.edu/doc/users_guide/index.html> Acessado em outubro de 2012.
- Vickery, B. C. (1986). “Knowledge representation: a brief review”. In: *Journal of Documentation*. London, **vol. 24**, nº 3, 145-159.
- Vickery, B. (2008). “On knowledge organisation”. Última edição. Disponível em: <<http://www.lucis.me.uk/knowlorg.htm#start>>. Acessado em Março de. 2013.
- W3C Escritório Brasil (2009). “Melhorando o acesso ao governo com o melhor uso da web”. Comitê Gestor da Internet no Brasil. São Paulo, 1ª ed.. Disponível em: <<http://www.w3c.br/divulgacao/pdf/gov-web.pdf>>. Acessado em setembro de 2013.
- Zeng, M. L., Salaba, A. (2005). “Toward an International Sharing and Use of Subject Authority Data”. In: *FRBR Workshop, OCLC*. Kent State University. Ohio, USA. Disponível em: www.oclc.org/research/events/frbr.../zeng/Zeng_Salaba.ppt. Acessado em agosto de 2013.

APÊNDICES

APÊNDICE A – ONTOLOGIA DO ORÇAMENTO EXTENDIDA EM FORMATO OWL/RDF.

```
<rdf:RDF                                xmlns="http://www.semanticweb.org/owl/owlapi/turtle#"
xmlns:loa="http://vocab.e.gov.br/2013/09/loa#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:owl="http://www.w3.org/2002/07/owl#"
xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xml:base="http://www.semanticweb.org/owl/owlapi/turtle">

<owl:Ontology rdf:about="http://vocab.e.gov.br/2013/09/loa#">

<rdfs:comment>Ontologia do Orçamento Federal</rdfs:comment>

</owl:Ontology>

<!--

////////////////////////////////////
//
// Annotation properties
////////////////////////////////////

-->
<!-- http://www.w3.org/1999/02/22-rdf-syntax-ns#label -->
<owl:AnnotationPropertyrdf:about="http://www.w3.org/1999/02/22-rdf-syntax-
ns#label"/>
<!--
```

```

////////////////////////////////////
//
// Object Properties
////////////////////////////////////

-->
<!-- http://vocab.e.gov.br/2013/09/loa#temAcao -->
<owl:ObjectProperty rdf:about="http://vocab.e.gov.br/2013/09/loa#temAcao">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AsymmetricProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#IrreflexiveProperty"/>
<rdfs:range rdf:resource="http://vocab.e.gov.br/2013/09/loa#Acao"/>
<rdfs:subPropertyOf
rdf:resource="http://vocab.e.gov.br/2013/09/loa#temClassificador"/>
</owl:ObjectProperty>
<!--
http://vocab.e.gov.br/2013/09/loa#temCategoriaEconomica
-->
<owl:ObjectProperty
rdf:about="http://vocab.e.gov.br/2013/09/loa#temCategoriaEconomica">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AsymmetricProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#IrreflexiveProperty"/>
<rdfs:range rdf:resource="http://vocab.e.gov.br/2013/09/loa#CategoriaEconomica"/>
<rdfs:subPropertyOf
rdf:resource="http://vocab.e.gov.br/2013/09/loa#temClassificador"/>
</owl:ObjectProperty>
<!-- http://vocab.e.gov.br/2013/09/loa#temClassificador -->
<owl:ObjectProperty
rdf:about="http://vocab.e.gov.br/2013/09/loa#temClassificador">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AsymmetricProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#IrreflexiveProperty"/>
<rdfs:range rdf:resource="http://vocab.e.gov.br/2013/09/loa#Classificador"/>
<rdfs:domain rdf:resource="http://vocab.e.gov.br/2013/09/loa#ItemDeDespesa"/>
</owl:ObjectProperty>
<!--
http://vocab.e.gov.br/2013/09/loa#temElementoDespesa
-->
<owl:ObjectProperty
rdf:about="http://vocab.e.gov.br/2013/09/loa#temElementoDespesa">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AsymmetricProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>

```

```

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#IrreflexiveProperty"/>
<rdfs:range rdf:resource="http://vocab.e.gov.br/2013/09/loa#ElementoDespesa"/>
<rdfs:subPropertyOf
rdf:resource="http://vocab.e.gov.br/2013/09/loa#temClassificador"/>
</owl:ObjectProperty>
<!-- http://vocab.e.gov.br/2013/09/loa#temEsfera -->
<owl:ObjectProperty rdf:about="http://vocab.e.gov.br/2013/09/loa#temEsfera">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AsymmetricProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#IrreflexiveProperty"/>
<rdfs:range rdf:resource="http://vocab.e.gov.br/2013/09/loa#Esfera"/>
<rdfs:subPropertyOf
rdf:resource="http://vocab.e.gov.br/2013/09/loa#temClassificador"/>
</owl:ObjectProperty>
<!-- http://vocab.e.gov.br/2013/09/loa#temExercicio -->
<owl:ObjectProperty rdf:about="http://vocab.e.gov.br/2013/09/loa#temExercicio">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AsymmetricProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#IrreflexiveProperty"/>
<rdfs:range rdf:resource="http://vocab.e.gov.br/2013/09/loa#Exercicio"/>
<rdfs:domain rdf:resource="http://vocab.e.gov.br/2013/09/loa#ItemDeDespesa"/>
</owl:ObjectProperty>
<!-- http://vocab.e.gov.br/2013/09/loa#temFonteRecursos -->
<owl:ObjectProperty
rdf:about="http://vocab.e.gov.br/2013/09/loa#temFonteRecursos">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AsymmetricProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#IrreflexiveProperty"/>
<rdfs:range rdf:resource="http://vocab.e.gov.br/2013/09/loa#FonteRecursos"/>
<rdfs:subPropertyOf
rdf:resource="http://vocab.e.gov.br/2013/09/loa#temClassificador"/>
</owl:ObjectProperty>
<!-- http://vocab.e.gov.br/2013/09/loa#temFuncao -->
<owl:ObjectProperty rdf:about="http://vocab.e.gov.br/2013/09/loa#temFuncao">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AsymmetricProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#IrreflexiveProperty"/>
<rdfs:range rdf:resource="http://vocab.e.gov.br/2013/09/loa#Funcao"/>
<rdfs:subPropertyOf
rdf:resource="http://vocab.e.gov.br/2013/09/loa#temClassificador"/>
</owl:ObjectProperty>
<!-- http://vocab.e.gov.br/2013/09/loa#temGND -->
<owl:ObjectProperty rdf:about="http://vocab.e.gov.br/2013/09/loa#temGND">

```

```

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AsymmetricProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#IrreflexiveProperty"/>
<rdfs:range rdf:resource="http://vocab.e.gov.br/2013/09/loa#GrupoNatDespesa"/>
<rdfs:subPropertyOf
rdf:resource="http://vocab.e.gov.br/2013/09/loa#temClassificador"/>
</owl:ObjectProperty>
<!--
  http://vocab.e.gov.br/2013/09/loa#temIdentificadorUso
-->
<owl:ObjectProperty
rdf:about="http://vocab.e.gov.br/2013/09/loa#temIdentificadorUso">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AsymmetricProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#IrreflexiveProperty"/>
<rdfs:range rdf:resource="http://vocab.e.gov.br/2013/09/loa#IdentificadorUso"/>
<rdfs:subPropertyOf
rdf:resource="http://vocab.e.gov.br/2013/09/loa#temClassificador"/>
</owl:ObjectProperty>
<!--  http://vocab.e.gov.br/2013/09/loa#temLocalizador  -->
<owl:ObjectProperty rdf:about="http://vocab.e.gov.br/2013/09/loa#temLocalizador">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AsymmetricProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#IrreflexiveProperty"/>
<rdfs:range rdf:resource="http://vocab.e.gov.br/2013/09/loa#Localizador"/>
<rdfs:subPropertyOf
rdf:resource="http://vocab.e.gov.br/2013/09/loa#temClassificador"/>
</owl:ObjectProperty>
<!--
  http://vocab.e.gov.br/2013/09/loa#temModalidadeAplicacao
-->
<owl:ObjectProperty
rdf:about="http://vocab.e.gov.br/2013/09/loa#temModalidadeAplicacao">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AsymmetricProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#IrreflexiveProperty"/>
<rdfs:range
rdf:resource="http://vocab.e.gov.br/2013/09/loa#ModalidadeAplicacao"/>
<rdfs:subPropertyOf
rdf:resource="http://vocab.e.gov.br/2013/09/loa#temClassificador"/>
</owl:ObjectProperty>
<!--  http://vocab.e.gov.br/2013/09/loa#temOrgao  -->
<owl:ObjectProperty rdf:about="http://vocab.e.gov.br/2013/09/loa#temOrgao">

```

```

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AsymmetricProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#IrreflexiveProperty"/>
<rdfs:range rdf:resource="http://vocab.e.gov.br/2013/09/loa#Orgao"/>
<rdfs:domain
rdf:resource="http://vocab.e.gov.br/2013/09/loa#UnidadeOrcamentaria"/>
</owl:ObjectProperty>
<!--
  http://vocab.e.gov.br/2013/09/loa#temPlanoOrcamentario
-->
<owl:ObjectProperty
rdf:about="http://vocab.e.gov.br/2013/09/loa#temPlanoOrcamentario">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AsymmetricProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#IrreflexiveProperty"/>
<rdfs:range rdf:resource="http://vocab.e.gov.br/2013/09/loa#PlanoOrcamentario"/>
<rdfs:subPropertyOf
rdf:resource="http://vocab.e.gov.br/2013/09/loa#temClassificador"/>
</owl:ObjectProperty>
<!-- http://vocab.e.gov.br/2013/09/loa#temPrograma -->
<owl:ObjectProperty rdf:about="http://vocab.e.gov.br/2013/09/loa#temPrograma">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AsymmetricProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#IrreflexiveProperty"/>
<rdfs:range rdf:resource="http://vocab.e.gov.br/2013/09/loa#Programa"/>
<rdfs:subPropertyOf
rdf:resource="http://vocab.e.gov.br/2013/09/loa#temClassificador"/>
</owl:ObjectProperty>
<!--
  http://vocab.e.gov.br/2013/09/loa#temResultadoPrimario
-->
<owl:ObjectProperty
rdf:about="http://vocab.e.gov.br/2013/09/loa#temResultadoPrimario">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AsymmetricProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#IrreflexiveProperty"/>
<rdfs:range rdf:resource="http://vocab.e.gov.br/2013/09/loa#ResultadoPrimario"/>
<rdfs:subPropertyOf
rdf:resource="http://vocab.e.gov.br/2013/09/loa#temClassificador"/>
</owl:ObjectProperty>
<!-- http://vocab.e.gov.br/2013/09/loa#temSubfuncao -->
<owl:ObjectProperty rdf:about="http://vocab.e.gov.br/2013/09/loa#temSubfuncao">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AsymmetricProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>

```

```

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#IrreflexiveProperty"/>
<rdfs:range rdf:resource="http://vocab.e.gov.br/2013/09/loa#Subfuncao"/>
<rdfs:subPropertyOf
rdf:resource="http://vocab.e.gov.br/2013/09/loa#temClassificador"/>
</owl:ObjectProperty>
<!--
  http://vocab.e.gov.br/2013/09/loa#temUnidadeOrcamentaria
-->
<owl:ObjectProperty
rdf:about="http://vocab.e.gov.br/2013/09/loa#temUnidadeOrcamentaria">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AsymmetricProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#IrreflexiveProperty"/>
<rdfs:range
rdf:resource="http://vocab.e.gov.br/2013/09/loa#UnidadeOrcamentaria"/>
<rdfs:subPropertyOf
rdf:resource="http://vocab.e.gov.br/2013/09/loa#temClassificador"/>
</owl:ObjectProperty>
<!--

////////////////////////////////////
/////
//
// Data properties
//

////////////////////////////////////
/////

-->
<!-- http://vocab.e.gov.br/2013/09/loa#codigo -->
<owl:DatatypeProperty rdf:about="http://vocab.e.gov.br/2013/09/loa#codigo">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
<rdfs:domain rdf:resource="http://vocab.e.gov.br/2013/09/loa#Classificador"/>
<rdfs:domain rdf:resource="http://vocab.e.gov.br/2013/09/loa#Orgao"/>
<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:DatatypeProperty>
<!--
  http://vocab.e.gov.br/2013/09/loa#dataUltimaAtualizacao
-->
<owl:DatatypeProperty
rdf:about="http://vocab.e.gov.br/2013/09/loa#dataUltimaAtualizacao">

```

```

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
<rdfs:domain rdf:resource="http://vocab.e.gov.br/2013/09/loa#Exercicio"/>
<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#dateTime"/>
</owl:DatatypeProperty>
<!-- http://vocab.e.gov.br/2013/09/loa#identificador -->
<owl:DatatypeProperty
rdf:about="http://vocab.e.gov.br/2013/09/loa#identificador">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
<rdfs:domain rdf:resource="http://vocab.e.gov.br/2013/09/loa#Exercicio"/>
<rdfs:domain rdf:resource="http://vocab.e.gov.br/2013/09/loa#Localizador"/>
<rdfs:domain rdf:resource="http://vocab.e.gov.br/2013/09/loa#PlanoOrcamentario"/>
<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#integer"/>
<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:DatatypeProperty>
<!--
http://vocab.e.gov.br/2013/09/loa#valorDotacaoInicial
-->
<owl:DatatypeProperty
rdf:about="http://vocab.e.gov.br/2013/09/loa#valorDotacaoInicial">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
<rdfs:domain rdf:resource="http://vocab.e.gov.br/2013/09/loa#ItemDeDespesa"/>
<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#decimal"/>
</owl:DatatypeProperty>
<!-- http://vocab.e.gov.br/2013/09/loa#valorEmpenhado -->
<owl:DatatypeProperty
rdf:about="http://vocab.e.gov.br/2013/09/loa#valorEmpenhado">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
<rdfs:domain rdf:resource="http://vocab.e.gov.br/2013/09/loa#ItemDeDespesa"/>
<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#decimal"/>
</owl:DatatypeProperty>
<!--
http://vocab.e.gov.br/2013/09/loa#valorLeiMaisCredito
-->
<owl:DatatypeProperty
rdf:about="http://vocab.e.gov.br/2013/09/loa#valorLeiMaisCredito">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
<rdfs:domain rdf:resource="http://vocab.e.gov.br/2013/09/loa#ItemDeDespesa"/>
<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#decimal"/>
</owl:DatatypeProperty>
<!-- http://vocab.e.gov.br/2013/09/loa#valorLiquidado -->
<owl:DatatypeProperty
rdf:about="http://vocab.e.gov.br/2013/09/loa#valorLiquidado">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>

```



```

<rdfs:domain rdf:resource="http://vocab.e.gov.br/2013/09/loa#ItemDeDespesa"/>
<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#decimal"/>
</owl:DatatypeProperty>
<!-- http://vocab.e.gov.br/2013/09/loa#valorPago -->
<owl:DatatypeProperty rdf:about="http://vocab.e.gov.br/2013/09/loa#valorPago">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
<rdfs:domain rdf:resource="http://vocab.e.gov.br/2013/09/loa#ItemDeDespesa"/>
<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#decimal"/>
</owl:DatatypeProperty>
<!-- http://vocab.e.gov.br/2013/09/loa#valorProjetoLei -->
<owl:DatatypeProperty
rdf:about="http://vocab.e.gov.br/2013/09/loa#valorProjetoLei">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
<rdfs:domain rdf:resource="http://vocab.e.gov.br/2013/09/loa#ItemDeDespesa"/>
<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#decimal"/>
</owl:DatatypeProperty>
<!-- http://www.w3.org/1999/02/22-rdf-syntax-ns#label -->
<owl:DatatypeProperty rdf:about="http://www.w3.org/1999/02/22-rdf-syntax-
ns#label">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:DatatypeProperty>
<!--

////////////////////////////////////
/////
//
// Classes
//

////////////////////////////////////
/////

-->
<!-- http://vocab.e.gov.br/2013/09/loa#Acao -->
<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#Acao">
<rdfs:subClassOf rdf:resource="http://vocab.e.gov.br/2013/09/loa#Classificador"/>
</owl:Class>
<!-- http://vocab.e.gov.br/2013/09/loa#Atividade -->
<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#Atividade">
<rdfs:subClassOf rdf:resource="http://vocab.e.gov.br/2013/09/loa#Acao"/>
</owl:Class>

```

```

<!--
  http://vocab.e.gov.br/2013/09/loa#CategoriaEconomica
-->
<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#CategoriaEconomica">
<rdfs:subClassOf rdf:resource="http://vocab.e.gov.br/2013/09/loa#Classificador"/>
</owl:Class>
<!-- http://vocab.e.gov.br/2013/09/loa#Classificador -->
<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#Classificador">
<owl:equivalentClass>
<owl:Class>
<owl:intersectionOf rdf:parseType="Collection">
<owl:Restriction>
<owl:onProperty rdf:resource="http://vocab.e.gov.br/2013/09/loa#codigo"/>
<owl:someValuesFrom rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:Restriction>
<owl:Restriction>
<owl:onProperty rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#label"/>
<owl:someValuesFrom rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:Restriction>
</owl:intersectionOf>
</owl:Class>
</owl:equivalentClass>
</owl:Class>
<!-- http://vocab.e.gov.br/2013/09/loa#ElementoDespesa -->
<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#ElementoDespesa">
<rdfs:subClassOf rdf:resource="http://vocab.e.gov.br/2013/09/loa#Classificador"/>
</owl:Class>
<!-- http://vocab.e.gov.br/2013/09/loa#Esfera -->
<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#Esfera">
<rdfs:subClassOf rdf:resource="http://vocab.e.gov.br/2013/09/loa#Classificador"/>
</owl:Class>
<!-- http://vocab.e.gov.br/2013/09/loa#Exercicio -->
<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#Exercicio">
<owl:equivalentClass>
<owl:Class>
<owl:intersectionOf rdf:parseType="Collection">
<owl:Restriction>
<owl:onProperty
rdf:resource="http://vocab.e.gov.br/2013/09/loa#dataUltimaAtualizacao"/>
<owl:someValuesFrom rdf:resource="http://www.w3.org/2001/XMLSchema#dateTime"/>
</owl:Restriction>
<owl:Restriction>
<owl:onProperty rdf:resource="http://vocab.e.gov.br/2013/09/loa#identificador"/>

```

```

<owl:allValuesFrom rdf:resource="http://www.w3.org/2001/XMLSchema#integer"/>
</owl:Restriction>
</owl:intersectionOf>
</owl:Class>
</owl:equivalentClass>
</owl:Class>
<!-- http://vocab.e.gov.br/2013/09/loa#FonteRecursos -->
<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#FonteRecursos">
<rdfs:subClassOf rdf:resource="http://vocab.e.gov.br/2013/09/loa#Classificador"/>
</owl:Class>
<!-- http://vocab.e.gov.br/2013/09/loa#Funcao -->
<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#Funcao">
<rdfs:subClassOf rdf:resource="http://vocab.e.gov.br/2013/09/loa#Classificador"/>
</owl:Class>
<!-- http://vocab.e.gov.br/2013/09/loa#GrupoNatDespesa -->
<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#GrupoNatDespesa">
<rdfs:subClassOf rdf:resource="http://vocab.e.gov.br/2013/09/loa#Classificador"/>
</owl:Class>
<!-- http://vocab.e.gov.br/2013/09/loa#IdentificadorUso -->
<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#IdentificadorUso">
<rdfs:subClassOf rdf:resource="http://vocab.e.gov.br/2013/09/loa#Classificador"/>
</owl:Class>
<!-- http://vocab.e.gov.br/2013/09/loa#ItemDeDespesa -->
<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#ItemDeDespesa">
<owl:equivalentClass>
<owl:Class>
<owl:intersectionOf rdf:parseType="Collection">
<owl:Restriction>
<owl:onProperty rdf:resource="http://vocab.e.gov.br/2013/09/loa#temAcao"/>
<owl:someValuesFrom rdf:resource="http://vocab.e.gov.br/2013/09/loa#Acao"/>
</owl:Restriction>
<owl:Restriction>
<owl:onProperty
rdf:resource="http://vocab.e.gov.br/2013/09/loa#temCategoriaEconomica"/>
<owl:someValuesFrom
rdf:resource="http://vocab.e.gov.br/2013/09/loa#CategoriaEconomica"/>
</owl:Restriction>
<owl:Restriction>
<owl:onProperty
rdf:resource="http://vocab.e.gov.br/2013/09/loa#temElementoDespesa"/>
<owl:someValuesFrom
rdf:resource="http://vocab.e.gov.br/2013/09/loa#ElementoDespesa"/>
</owl:Restriction>

```

```

<owl:Restriction>
<owl:onProperty rdf:resource="http://vocab.e.gov.br/2013/09/loa#temEsfera"/>
<owl:someValuesFrom rdf:resource="http://vocab.e.gov.br/2013/09/loa#Esfera"/>
</owl:Restriction>
<owl:Restriction>
<owl:onProperty rdf:resource="http://vocab.e.gov.br/2013/09/loa#temExercicio"/>
<owl:someValuesFrom rdf:resource="http://vocab.e.gov.br/2013/09/loa#Exercicio"/>
</owl:Restriction>
<owl:Restriction>
<owl:onProperty
rdf:resource="http://vocab.e.gov.br/2013/09/loa#temFonteRecursos"/>
<owl:someValuesFrom
rdf:resource="http://vocab.e.gov.br/2013/09/loa#FonteRecursos"/>
</owl:Restriction>
<owl:Restriction>
<owl:onProperty rdf:resource="http://vocab.e.gov.br/2013/09/loa#temFuncao"/>
<owl:someValuesFrom rdf:resource="http://vocab.e.gov.br/2013/09/loa#Funcao"/>
</owl:Restriction>
<owl:Restriction>
<owl:onProperty rdf:resource="http://vocab.e.gov.br/2013/09/loa#temGND"/>
<owl:someValuesFrom
rdf:resource="http://vocab.e.gov.br/2013/09/loa#GrupoNatDespesa"/>
</owl:Restriction>
<owl:Restriction>
<owl:onProperty
rdf:resource="http://vocab.e.gov.br/2013/09/loa#temIdentificadorUso"/>
<owl:someValuesFrom
rdf:resource="http://vocab.e.gov.br/2013/09/loa#IdentificadorUso"/>
</owl:Restriction>
<owl:Restriction>
<owl:onProperty rdf:resource="http://vocab.e.gov.br/2013/09/loa#temLocalizador"/>
<owl:someValuesFrom
rdf:resource="http://vocab.e.gov.br/2013/09/loa#Localizador"/>
</owl:Restriction>
<owl:Restriction>
<owl:onProperty
rdf:resource="http://vocab.e.gov.br/2013/09/loa#temModalidadeAplicacao"/>
<owl:someValuesFrom
rdf:resource="http://vocab.e.gov.br/2013/09/loa#ModalidadeAplicacao"/>
</owl:Restriction>
<owl:Restriction>
<owl:onProperty rdf:resource="http://vocab.e.gov.br/2013/09/loa#temPrograma"/>
<owl:someValuesFrom rdf:resource="http://vocab.e.gov.br/2013/09/loa#Programa"/>

```

```

</owl:Restriction>
<owl:Restriction>
<owl:onProperty
rdf:resource="http://vocab.e.gov.br/2013/09/loa#temResultadoPrimario"/>
<owl:someValuesFrom
rdf:resource="http://vocab.e.gov.br/2013/09/loa#ResultadoPrimario"/>
</owl:Restriction>
<owl:Restriction>
<owl:onProperty rdf:resource="http://vocab.e.gov.br/2013/09/loa#temSubfuncao"/>
<owl:someValuesFrom rdf:resource="http://vocab.e.gov.br/2013/09/loa#Subfuncao"/>
</owl:Restriction>
<owl:Restriction>
<owl:onProperty
rdf:resource="http://vocab.e.gov.br/2013/09/loa#temUnidadeOrcamentaria"/>
<owl:someValuesFrom
rdf:resource="http://vocab.e.gov.br/2013/09/loa#UnidadeOrcamentaria"/>
</owl:Restriction>
<owl:Restriction>
<owl:onProperty
rdf:resource="http://vocab.e.gov.br/2013/09/loa#valorDotacaoInicial"/>
<owl:someValuesFrom rdf:resource="http://www.w3.org/2001/XMLSchema#decimal"/>
</owl:Restriction>
<owl:Restriction>
<owl:onProperty rdf:resource="http://vocab.e.gov.br/2013/09/loa#valorEmpenhado"/>
<owl:someValuesFrom rdf:resource="http://www.w3.org/2001/XMLSchema#decimal"/>
</owl:Restriction>
<owl:Restriction>
<owl:onProperty
rdf:resource="http://vocab.e.gov.br/2013/09/loa#valorLeiMaisCredito"/>
<owl:someValuesFrom rdf:resource="http://www.w3.org/2001/XMLSchema#decimal"/>
</owl:Restriction>
<owl:Restriction>
<owl:onProperty rdf:resource="http://vocab.e.gov.br/2013/09/loa#valorLiquidado"/>
<owl:someValuesFrom rdf:resource="http://www.w3.org/2001/XMLSchema#decimal"/>
</owl:Restriction>
<owl:Restriction>
<owl:onProperty rdf:resource="http://vocab.e.gov.br/2013/09/loa#valorPago"/>
<owl:someValuesFrom rdf:resource="http://www.w3.org/2001/XMLSchema#decimal"/>
</owl:Restriction>
<owl:Restriction>
<owl:onProperty
rdf:resource="http://vocab.e.gov.br/2013/09/loa#valorProjetoLei"/>
<owl:someValuesFrom rdf:resource="http://www.w3.org/2001/XMLSchema#decimal"/>

```

```

</owl:Restriction>
</owl:intersectionOf>
</owl:Class>
</owl:equivalentClass>
</owl:Class>
<!-- http://vocab.e.gov.br/2013/09/loa#Localizador -->
<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#Localizador">
<owl:equivalentClass>
<owl:Restriction>
<owl:onProperty rdf:resource="http://vocab.e.gov.br/2013/09/loa#identificador"/>
<owl:allValuesFrom rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:Restriction>
</owl:equivalentClass>
<rdfs:subClassOf rdf:resource="http://vocab.e.gov.br/2013/09/loa#Classificador"/>
</owl:Class>
<!--
http://vocab.e.gov.br/2013/09/loa#ModalidadeAplicacao
-->
<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#ModalidadeAplicacao">
<rdfs:subClassOf rdf:resource="http://vocab.e.gov.br/2013/09/loa#Classificador"/>
</owl:Class>
<!-- http://vocab.e.gov.br/2013/09/loa#OperacaoEspecial -->
<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#OperacaoEspecial">
<rdfs:subClassOf rdf:resource="http://vocab.e.gov.br/2013/09/loa#Acao"/>
</owl:Class>
<!-- http://vocab.e.gov.br/2013/09/loa#Orgao -->
<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#Orgao">
<owl:equivalentClass>
<owl:Class>
<owl:intersectionOf rdf:parseType="Collection">
<owl:Restriction>
<owl:onProperty rdf:resource="http://vocab.e.gov.br/2013/09/loa#codigo"/>
<owl:someValuesFrom rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:Restriction>
<owl:Restriction>
<owl:onProperty rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#label"/>
<owl:someValuesFrom rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:Restriction>
</owl:intersectionOf>
</owl:Class>
</owl:equivalentClass>
</owl:Class>
<!--

```

```

http://vocab.e.gov.br/2013/09/loa#PlanoOrcamentario
-->
<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#PlanoOrcamentario">
<owl:equivalentClass>
<owl:Restriction>
<owl:onProperty rdf:resource="http://vocab.e.gov.br/2013/09/loa#identificador"/>
<owl:allValuesFrom rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:Restriction>
</owl:equivalentClass>
<rdfs:subClassOf rdf:resource="http://vocab.e.gov.br/2013/09/loa#Classificador"/>
</owl:Class>
<!-- http://vocab.e.gov.br/2013/09/loa#Programa -->
<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#Programa">
<rdfs:subClassOf rdf:resource="http://vocab.e.gov.br/2013/09/loa#Classificador"/>
</owl:Class>
<!-- http://vocab.e.gov.br/2013/09/loa#Projeto -->
<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#Projeto">
<rdfs:subClassOf rdf:resource="http://vocab.e.gov.br/2013/09/loa#Acao"/>
</owl:Class>
<!--
http://vocab.e.gov.br/2013/09/loa#ResultadoPrimario
-->
<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#ResultadoPrimario">
<rdfs:subClassOf rdf:resource="http://vocab.e.gov.br/2013/09/loa#Classificador"/>
</owl:Class>
<!-- http://vocab.e.gov.br/2013/09/loa#Subfuncao -->
<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#Subfuncao">
<rdfs:subClassOf rdf:resource="http://vocab.e.gov.br/2013/09/loa#Classificador"/>
</owl:Class>
<!--
http://vocab.e.gov.br/2013/09/loa#UnidadeOrcamentaria
-->
<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#UnidadeOrcamentaria">
<owl:equivalentClass>
<owl:Restriction>
<owl:onProperty rdf:resource="http://vocab.e.gov.br/2013/09/loa#temOrgao"/>
<owl:someValuesFrom rdf:resource="http://vocab.e.gov.br/2013/09/loa#Orgao"/>
</owl:Restriction>
</owl:equivalentClass>
<rdfs:subClassOf rdf:resource="http://vocab.e.gov.br/2013/09/loa#Classificador"/>
</owl:Class>
<!--

```

```

////////////////////////////////////
/////
//
// General axioms
//

////////////////////////////////////
/////

-->
<rdf:Description>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AllDisjointClasses"/>
<owl:members rdf:parseType="Collection">
<rdf:Description rdf:about="http://vocab.e.gov.br/2013/09/loa#Classificador"/>
<rdf:Description rdf:about="http://vocab.e.gov.br/2013/09/loa#Exercicio"/>
<rdf:Description rdf:about="http://vocab.e.gov.br/2013/09/loa#ItemDeDespesa"/>
<rdf:Description rdf:about="http://vocab.e.gov.br/2013/09/loa#Orgao"/>
</owl:members>
</rdf:Description>
<rdf:Description>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AllDisjointClasses"/>
<owl:members rdf:parseType="Collection">
<rdf:Description rdf:about="http://vocab.e.gov.br/2013/09/loa#Acao"/>
<rdf:Description
rdf:about="http://vocab.e.gov.br/2013/09/loa#CategoriaEconomica"/>
<rdf:Description rdf:about="http://vocab.e.gov.br/2013/09/loa#ElementoDespesa"/>
<rdf:Description rdf:about="http://vocab.e.gov.br/2013/09/loa#Esfera"/>
<rdf:Description rdf:about="http://vocab.e.gov.br/2013/09/loa#FonteRecursos"/>
<rdf:Description rdf:about="http://vocab.e.gov.br/2013/09/loa#Funcao"/>
<rdf:Description rdf:about="http://vocab.e.gov.br/2013/09/loa#GrupoNatDespesa"/>
<rdf:Description rdf:about="http://vocab.e.gov.br/2013/09/loa#IdentificadorUso"/>
<rdf:Description rdf:about="http://vocab.e.gov.br/2013/09/loa#Localizador"/>
<rdf:Description
rdf:about="http://vocab.e.gov.br/2013/09/loa#ModalidadeAplicacao"/>
<rdf:Description
rdf:about="http://vocab.e.gov.br/2013/09/loa#PlanoOrcamentario"/>
<rdf:Description rdf:about="http://vocab.e.gov.br/2013/09/loa#Programa"/>
<rdf:Description
rdf:about="http://vocab.e.gov.br/2013/09/loa#ResultadoPrimario"/>
<rdf:Description rdf:about="http://vocab.e.gov.br/2013/09/loa#Subfuncao"/>
<rdf:Description
rdf:about="http://vocab.e.gov.br/2013/09/loa#UnidadeOrcamentaria"/>

```



```
</owl:members>
</rdf:Description>
<rdf:Description>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AllDisjointClasses"/>
<owl:members rdf:parseType="Collection">
<rdf:Description rdf:about="http://vocab.e.gov.br/2013/09/loa#Atividade"/>
<rdf:Description rdf:about="http://vocab.e.gov.br/2013/09/loa#OperacaoEspecial"/>
<rdf:Description rdf:about="http://vocab.e.gov.br/2013/09/loa#Projeto"/>
</owl:members>
</rdf:Description>
</rdf:RDF>
```