

# DA NECESSIDADE E VIABILIDADE DA ADOÇÃO DO PADRÃO LOD PELO CONGRESSO NACIONAL: um estudo no contexto do orçamento público

Marcus Vinícius Chevitarese Alves\*  
Marcello Peixoto Bax\*\*

## RESUMO

O objetivo da pesquisa foi analisar a necessidade e a viabilidade da adoção do padrão LOD pelo Congresso Nacional no que concerne à informação sobre o Orçamento Federal. A pesquisa inicia-se com uma breve revisão dos fundamentos teóricos da Web Semântica (WS). Em seguida foi empreendida pesquisa acerca das principais tecnologias da WS, especialmente aquelas associadas ao padrão Dados Abertos e Ligados. Posteriormente, foi elaborado e aplicado um estudo de casos múltiplos, no qual se avaliou os sistemas de informação orçamentários do Congresso Nacional que possuem interface Web: o Siga Brasil (Senado) e o Fiscalize (Câmara dos Deputados). Dando prosseguimento, avaliou-se a viabilidade da adoção do padrão Dados Abertos e Ligados pelo Congresso Nacional, no que diz respeito aos dados e informações orçamentárias. Para isso, ampliou-se a revisão bibliográfica inicial, de modo a se ter mais clareza sobre os princípios do padrão em estudo, bem como das tecnologias e metodologias necessárias para a sua implantação. Além disso, fez-se outro estudo de casos múltiplos, no qual se avaliou duas aplicações computacionais baseadas nesse padrão. Os resultados da pesquisa indicam a necessidade e a viabilidade da adoção do padrão Dados Abertos e Ligados pelo Congresso Nacional no contexto das informações associadas ao Orçamento Público.

**Palavras-chave:** Direito à informação. Transparência administrativa. Orçamento Público. Tecnologias da informação e da comunicação. Web Semântica.

\* Mestre em Administração Pública pela Fundação João Pinheiro, Brasil. Analista Legislativo da Câmara dos Deputados, Brasil.  
E-mail: mchevita@gmail.com

\*\* Doutor em Informática pela Université Montpellier 2 - Sciences et Techniques, França. Professor Associado da Universidade Federal de Minas Gerais, Brasil.  
E-mail: bax@eci.ufmg.br

## I INTRODUÇÃO

Conforme Platt Neto (2009, p. 76), “[...] as entidades que compõem a administração pública brasileira são obrigadas pela Constituição Federal a prestar contas do uso de recursos públicos e a respeitar o princípio da publicidade [...]”. Tal determinação ensejou a criação de órgãos de controle, e a edição de normas infraconstitucionais, tais como a Lei de Responsabilidade Fiscal (BRASIL, 2000) e a Lei da Transparência (BRASIL, 2009). Entretanto, há

indicativos de que as normas e as instituições governamentais de controle, isoladamente, são incapazes de coibir a malversação de recursos públicos (cf. CARLOS et al., 2008; MELLO et al., 2002). A Administração Pública está atenta à questão. Diversos produtos de informação foram criados, como os portais de governo. Contudo, alguns problemas ficaram evidenciados, dos quais destacamos dois: (1) dados da Administração Pública fragmentados, pouco integrados, em formatos distintos e, muitas vezes, proprietários e, portanto, de difícil

reutilização (ALANI et al., 2007); (2) aplicações de exploração e visualização dos dados possibilitam baixa flexibilidade na sua manipulação. A literatura denomina essa forma de disseminação como *broadcasting* (W3C Escritório Brasil, 2010), ou seja, um emissor (no caso, o governo) e vários receptores (os cidadãos e organizações civis). Portanto, nesse caso, se o usuário da informação quiser explorar a informação de forma diferente daquelas que foram previstas e implementadas pelos sistemas governamentais, enfrentará grandes dificuldades.

Esses problemas estimularam a criação de um conjunto de recomendações denominado “Dados Abertos” (*Open Data*), definido como “[...] a disponibilização de informações governamentais representadas em formato aberto e acessível de tal modo que possam ser reutilizadas, misturadas com informações de outras fontes, gerando novos significados [...]” (W3C Escritório Brasil, 2010, p. 8). Com a Web Semântica (WS) emergiu um novo padrão, denominado Dados Ligados (*Linked Data*), que vem absorvendo o padrão anterior, dando origem aos *LOD*<sup>1</sup> (*Linked and Open Data – LOD*). Conforme Bizer et al (2009), os dados ligados referem-se a dados publicados na Web de modo que sejam legíveis por máquina, seus significados sejam explicitamente definidos, estejam ligados a outros conjuntos de dados e, por sua vez, possam ser ligados a partir de conjuntos de dados externos. A estratégia de *LOD* tem o potencial de agregar valor aos dados, tornando-os informações úteis aos usuários. No caso dos dados em posse do governo, esses usuários são todos os cidadãos, que são, por assim dizer, seus legítimos donos. Nesse contexto, é importante destacar também a Lei de Acesso à Informação – LAI (BRASIL, 2011), a qual é clara no que tange à necessidade de se tornar disponíveis as informações no formato “dados abertos”. Já existe uma gama de tecnologias e ferramentas das quais governo e sociedade podem lançar mão para auxiliar a implementação dessa lei. No entanto, como ressalta Klischewski *et al.* (2006), o governo eletrônico enfrenta um desafio para migração de dados para um formato ‘semântico’: afinal, trata-se de “páginas HTML feitas à

mão e páginas web geradas por sistemas gerenciadores de conteúdo projetados para exibir conteúdo para humanos [não para máquinas]” (p. 1). Os autores explicam que, para a Web Semântica (WS), é necessário enriquecer o conteúdo de modo a torná-lo legível por máquinas. Isso pode ser feito: (1) reorganizando o código dos programas e ou conteúdo; (2) criando uma “camada semântica” extra. Esta pesquisa irá explorar projetos a segunda abordagem, por ser a menos onerosa, em tese. O foco será a análise dos sistemas de informações relativas ao Orçamento Público, na esfera federal.

Mello et al. (2002) afirmam: “[...] o conhecimento do processo orçamentário [...] e a prestação de contas é de fundamental importância para o controle social e o combate à corrupção [...]” (p. 172). Nesse contexto, o problema de pesquisa é: como o Congresso Nacional pode cumprir seu papel institucional e legal, especialmente os estabelecidos pela Constituição Federal e pela LAI, no que concerne à abertura dos dados sobre o Orçamento Público? O objetivo da pesquisa é analisar a necessidade e a viabilidade da adoção do padrão *LOD* pelo Congresso Nacional no que concerne ao Orçamento Público. A pesquisa possui dois pressupostos: (P1) Os sistemas Web de informações orçamentárias do Congresso Nacional hoje disponíveis não atendem plenamente os requisitos estabelecidos pela LAI; (P2) A adoção do padrão *LOD* pelo Congresso Nacional, no tocante à informação sobre o Orçamento Federal, é viável e possibilita o atendimento aos requisitos da LAI.

A fim de cumprir o objetivo proposto, foram brevemente revisados os fundamentos teóricos da WS, base do padrão *LOD*. Pesquisou-se acerca das principais tecnologias da WS e do *LOD*. Para investigar o pressuposto (P1), foi realizado um estudo de casos múltiplos, no qual se avaliou os sistemas de informação orçamentários do Congresso Nacional de interface Web: o Siga Brasil (Senado) e o Fiscalize (Câmara dos Deputados). No intuito de confirmar ou refutar o pressuposto (P2), avaliou-se a viabilidade da adoção do padrão *LOD* pelo Congresso Nacional, no que diz respeito às informações orçamentárias. Para isso, ampliou-se a revisão

<sup>1</sup> É comum, na literatura da área, o uso da expressão *Linked Data* como sinônimo de *Linked and Open Data*. Optou-se pela segunda forma, para enfatizar o caráter aberto dos dados.

bibliográfica inicial, de modo a se ter mais clareza sobre os princípios do padrão LOD e das tecnologias e metodologias necessárias para a sua implantação. Fez-se ainda outro estudo de casos múltiplos, no qual se avaliou duas aplicações baseadas nesse padrão.

## 2 FUNDAMENTOS TEÓRICOS

Como organizar, tratar e recuperar informação de forma eficiente em um repositório tão grande e dinâmico como a Web? Para Freitas (2003, p. 4) [...] a falta de mecanismos capazes de captar a semântica do conteúdo das páginas da Web criou uma forte demanda de serviços que se ajustam à classe de serviços estudada em Inteligência Artificial, que passou a ser vista como alternativa para melhor tratamento dos problemas de informação na Internet [...]. Segundo ele, “[...] dois tipos de solução foram propostos, que não são mutuamente exclusivas [...]”. A primeira consiste em “[...] dotar os sistemas de inteligência e autonomia para percorrer e selecionar informação relevante na rede, deduzindo as informações úteis [...]”. Essa abordagem é conhecida como *agentes inteligentes*. Já a segunda abordagem intenta “[...] fazer com que as páginas possuam semântica clara e definida e que agentes possam raciocinar sobre esta semântica [...]” (FREITAS, 2003, p. 4). Segundo o autor, foi essa a abordagem que deu origem à WS. No entanto, todo o potencial da WS só poderá ser atingido com o uso dos agentes inteligentes. Houve uma grande evolução nas tecnologias e estudos envolvendo a WS. Atualmente, seus padrões são definidos pelo W3C<sup>2</sup>, para quem a WS possibilita a transição da Web de *documentos* para a Web de *dados*. Os aspectos tecnológicos serão descritos na Seção 3. Antes é importante descrever algumas das teorias subjacentes à WS, ainda que sucintamente.

O termo *semântica* possui diversos significados, sendo objeto de várias disciplinas. É necessário discutir como a semântica pode ser expressa, o que leva à reflexão sobre a Representação do Conhecimento, i.e.,

construir um *modelo*, expresso em alguma *linguagem*. O grau de formalismo do modelo depende do grau de formalismo da linguagem. Allemang e Hendler (2011, p. 15) explicam que quando um modelo “[...] recai sobre as particularidades do contexto do seu leitor para que seu conteúdo seja interpretado, [...] este modelo é *informal*”; isto é, “[...] o modelo carece de um formalismo no qual o significado dos termos [...] podem ser definidos de forma unívoca [...]”. Além da questão do grau de *formalismo* de um modelo, existe outro aspecto importante, que é sua *expressividade*. A WS busca representar o conhecimento de forma que a semântica seja acessível também às máquinas. Mas que tipo de semântica pode ser acessível às máquinas? Existem as mais ‘pobres’, ou seja, menos expressivas, mas facilmente computáveis; por outro lado, existem aquelas ‘ricas’, mas intratáveis do ponto de vista computacional. Almeida e Souza (2011) estudaram a expressividade e o formalismo de alguns instrumentos para organização da informação, com ênfase na WS. Para “representar conhecimento” é necessário algum tipo de *linguagem*. Linguagem, nesse sentido, transcende o conceito de linguagem verbal. Feita a ressalva, é importante registrar que o tipo de Representação do Conhecimento tratado pela pesquisa é aquele *intrínseco* à WS, ou seja, assentado sobre linguagens formais.

No contexto de um sistema *aberto* como a Web, dois conjuntos de dados “podem utilizar terminologias diferentes para referir-se à mesma informação”; por outro lado, existe a possibilidade “[...] de uma mesma terminologia estar sendo utilizada com significados diferentes, por aplicações distintas [...]” (MORAIS ; SOARES, 2004, p. 3). Logo é necessário um mecanismo de possibilite a *interoperabilidade semântica* entre os recursos. Na WS, esse é o papel das ontologias. Uma ontologia estabelece um vocabulário controlado (terminologia) e uma rede de relacionamentos para esse conjunto específico de conceitos ou classes. Para que essa rede seja processável por dispositivos computacionais, é necessário haver uma representação formal. Vieira et al (2008) afirmam ter havido muitas pesquisas no campo da representação do conhecimento aliadas ao raciocínio automático. O foco é a definição de linguagens formais para representar o conhecimento, combinadas a métodos que permitam fazer inferên-

<sup>2</sup> World Wide Web Consortium: entidade responsável pela definição de padrões para a Web. Disponível em: <<http://www.w3c.org/sobre>>.

cias a partir do conhecimento expresso nessas linguagens. Os *métodos de inferências* têm origem na lógica formal. Sowa (2011a) afirma que a lógica de primeira ordem (LPO) é um subconjunto ou superconjunto da maior parte das linguagens usadas para representação do conhecimento. Ademais, a LPO é também um subconjunto de todas as linguagens naturais. As principais linguagens artificiais utilizadas na WS – RDF, RDFS e a maior parte da OWL – são compatíveis com a LPO.

As Lógicas de Descrições, ou lógicas descritivas, constituem uma “[...] evolução dos formalismos de representação do conhecimento baseados em objetos (como as redes semânticas), ao qual corresponde um *subconjunto estruturado da lógica de primeira ordem* [...]” (VIEIRA et al., 2008, p. 22, grifo nosso). Os autores esclarecem que as Lógicas de Descrição (DL – *Description Logics*) tem menor expressividade que a LPO. Mas possuem duas características fundamentais: “grande capacidade de representação de conhecimentos” e “métodos de dedução [computacionalmente] eficientes para os serviços de raciocínio associados” (VIEIRA et al., 2008, p. 22). A OWL (*Ontology Web Language*), adotada pelo W3C, é baseada em DL.

A WS não é meramente um conjunto de tecnologias: possui forte embasamento teórico, que remonta à Filosofia Clássica, passando pela Lógica, Linguística, Ciência da Computação, entre outras. Mas, como em outras áreas, não se pode prescindir da tecnologia para a materialização de ideias. A próxima seção explora as tecnologias de suporte à WS.

### 3 TECNOLOGIAS DA WEB SEMÂNTICA: RDF, RDFS, OWL E SPARQL

A WS consiste em diversas camadas interdependentes. A camada de infraestrutura, fornece os componentes básicos, denominados recursos, e os meios pelos quais esses serão referenciados e/ou descritos pelos componentes das demais camadas. Um *recurso* é qualquer conteúdo que tenha *identidade*. Assim, um recurso pode ser uma simples sequência de caracteres, uma imagem, um documento eletrônico completo ou mesmo uma coleção de outros recursos. A *referência* a um recurso é feita por um

identificador único, denominado URI (*Universal Resource Identifier*). Enquanto o URI *identifica* um recurso, a intenção do RDF (*Resource Description Framework*) é *descrevê-lo*. Por meio de sentenças, o RDF descreve e estrutura este recurso em termos de suas propriedades e das relações com outros recursos.

O RDF é uma linguagem formal definida como uma “[...] lógica baseada em assertivas, na qual cada tripla expressa uma proposição simples [...]” (W3C, 2004b, Não paginado). A técnica usada para construir a semântica do RDF é a *teoria dos modelos*. A teoria dos modelos “assume que a linguagem refere-se a um ‘mundo’, e descreve as condições que um mundo precisa satisfazer para atribuir significado para cada expressão na linguagem”. Um “mundo em particular é chamado *interpretação*”, logo a teoria dos modelos “poderia ser mais bem denominada ‘teoria da interpretação’” (W3C, 2004b, Não paginado).

O RDFS estende o RDF, conferindo-lhe maior expressividade. O RDFS especifica como usar o RDF para descrever vocabulários, a partir de um conjunto predefinido de vocabulários (W3C, 2004c). Esse conjunto é formado por classes e propriedades.

A linguagem recomendada pelo W3C para a descrição de ontologias é a OWL (*Web Ontology Language*): “[...] OWL é uma *linguagem* [...] projetada para *representar conhecimento* rico e complexo sobre coisas, grupo de coisas e relações entre coisas [...]” (W3C, 2012, Não paginado, tradução nossa). Ou seja, é a linguagem para descrever ontologias, na acepção computacional do termo. Importante ressaltar que a combinação dos mecanismos do RDFS e da OWL confere às ontologias da WS um grau de expressividade maior que alguns recursos tradicionais de representação da informação e do conhecimento, como os tesouros e as taxonomias. Conforme Bräscher (2008, p. 44), a ontologia “[...] agrega valor aos demais esquemas de representação, por meio de uma semântica mais profunda e também sob os aspectos conceitual, relacional e tecnológico [...]”.

O RDF e OWL provêm uma forma simples de representar dados distribuídos: no entanto, uma representação de dados é inútil sem os meios adequados para o acesso aos dados (ALLEMANG ; HENDLER, 2011). A linguagem SPARQL foi desenvolvida com o intuito de

preencher essa lacuna. SPARQL (*SPARQL Protocol And RDF Query Language* – Protocolo e Linguagem de Consulta RDF) é feita para dados em RDF, sendo baseada em padrões de grafos e correspondência (*matching*) entre subgrafos. Conforme discutido, uma das características distintivas da WS em relação à Web tradicional é a representação do conhecimento apoiada em modelos e linguagens formais. Quando se utiliza apenas linguagens naturais, as relações e inferências não estão totalmente explícitas e padronizadas, o que as impede de serem processados computacionalmente de forma eficiente. Ou seja, boa parte do conteúdo hoje existente na Web ‘tradicional’ depende da interpretação humana, mesmo quando seria interessante que fossem processáveis automaticamente. E esse é justamente o caso dos dados governamentais que, salvo raras exceções previstas legalmente, são públicos. Na próxima seção será visto como a utilização dos princípios e tecnologias da WS tornam os dados efetivamente ‘abertos’, ou seja, amplamente disponíveis e inteligíveis, inclusive por sistemas computacionais.

#### 4 O PADRÃO LOD NA REPRESENTAÇÃO E DISSEMINAÇÃO DA INFORMAÇÃO ORÇAMENTÁRIA

O processo orçamentário compreende as fases de elaboração e execução das leis orçamentárias – PPA (Plano Plurianual), LDO (Lei de Diretrizes Orçamentárias) e LOA (Lei Orçamentária Anual) (BRASIL, 2013). O PPA “[...] estabelece as diretrizes, os objetivos e as metas da Administração Pública, para as despesas de capital e outras delas decorrentes, e para as relativas aos programas de duração continuada [...]” (BEZERRA FILHO, 2006, p. 22). Já a LDO “[...] estabelece metas e prioridades da administração pública para o exercício seguinte [...]”. Finalmente a LOA estabelece o orçamento propriamente dito, composto de Orçamento Fiscal, Orçamento de Investimentos das Estatais e Orçamento da Seguridade Social. Como o foco do estudo de caso será o acompanhamento da execução orçamentária da despesa, os principais sistemas de informação são apresentados a seguir.

#### 4.1 Análise dos Sistemas de Informação Orçamentária do Congresso Nacional frente à LAI

O Congresso Nacional possui dois sistemas que divulgam informações orçamentárias via Web: o Sistema Fiscalize<sup>3</sup> e o Sistema Siga Brasil<sup>4</sup>. Os sistemas serão avaliados em termos da conformidade com o art. 8º da LAI, a fim de verificar o pressuposto (P1) desta pesquisa. As críticas serão classificadas como limitações ou não-conformidades em relação as dispositivos do art. 8º, §3º, da LAI. No último caso, haverá uma referência ao dispositivo. É importante mencionar que tais sistemas foram concebidos antes da promulgação da LAI. Portanto, o objetivo principal da avaliação foi mostrar a necessidade de se desenvolver estratégias complementares para o cumprimento pleno da LAI, uma vez que os sistemas avaliados, isoladamente, são insuficientes, conforme será visto. E, dentre as estratégias complementares possíveis, defende-se aqui a utilização de uma plataforma baseada no padrão LOD.

##### 4.1.1 Sistema Fiscalize

O Sistema Fiscalize<sup>5</sup> extrai informação do armazém de dados orçamentários da Câmara dos Deputados (DW Orçamento), e, por meio da tecnologia *Oracle Reports*, gera relatórios em formato estático (PDF). Sua atualização é diária, o que implica que ele atende ao art. 8º, §3º, inc. VI, da LAI; suas fontes são o Sistema Integrado de Administração Financeira (SIAFI operacional) e sua variante SIAFI Gerencial, o que, em tese, garante a autenticidade e a integridade das informações disponíveis para acesso – LAI, art. 8º, §3º, inc. V.

O sistema possui quatro módulos. Para a avaliação será descrito o módulo Transferências a Municípios, com foco nas informações sobre execução de convênios, conforme mencionado. Como os demais módulos têm estrutura e funcionamento análogos, a avaliação de um

3 <<http://www2.camara.leg.br/atividade-legislativa/orcamentobrasil/fiscalize/transferencia>>.

4 <[http://www9.senado.gov.br/portal/page/portal/orcamento\\_senado/SigaBrasil](http://www9.senado.gov.br/portal/page/portal/orcamento_senado/SigaBrasil)>. Neste trabalho, foi utilizada a opção ‘Acesso Livre’.

5 Informações obtidas verbalmente junto à Seção de Soluções de Inteligência de Negócios (SESIN) do Centro de Informática da Câmara dos Deputados.

aplica-se aos demais. O funcionamento do módulo se dá pelos seguintes passos: (1) o usuário seleciona o tipo de relatório: valores pagos ou valores empenhados/conveniados. Não é possível visualizar ambos os valores em um mesmo relatório; (2) o usuário seleciona período de tempo coberto pelo relatório: o mês e o ano. Não é possível selecionar um intervalo de tempo envolvendo dois ou mais meses; (3) o usuário seleciona o município destinatário do recurso. Não é possível selecionar mais de um município simultaneamente; (4) o usuário solicita a geração do relatório. O relatório é gerado apenas no formato PDF.

Como o relatório é gerado apenas em PDF, trata-se de uma não-conformidade e, ao mesmo tempo, de uma limitação. O fato de o sistema não possibilitar a geração dos dados em formatos abertos simples, como o CSV, o qual pode ser lido pelos principais editores de planilha eletrônica disponíveis, é uma não-conformidade relativa a dois dispositivos da LAI: art. 8º, §3º, inc. II: “[...] possibilitar a gravação de relatórios em diversos formatos eletrônicos, inclusive abertos e não proprietários, tais como planilhas e texto, de modo a facilitar a análise das informações [...]”; art. 8º, §3º, inc. III: “[...] possibilitar o acesso automatizado por sistemas externos em formatos abertos, estruturados e legíveis por máquina [...]” (BRASIL, 2012). Trata-se ainda de uma limitação, já que não está disponível em formatos abertos e ligados, como o RDF. Em tese, caso os metadados necessários estivessem disponíveis, seria possível a transformação dos dados para o RDF. Ocorre que, embora o PDF seja um formato aberto, o processo de extração dos dados e eventuais metadados existentes exigiria um esforço considerável de ‘raspagem de dados’ – conforme será visto no caso “Ligado nos Políticos” (Seção 3.3.2).

As demais observações são no sentido das limitações em termos de flexibilidade na combinação e visualização dos dados. Para ilustrar, não seria possível fazer análises comparativas do tipo ‘comparar os valores empenhados e conveniados transferidos da União para municípios com a mesma população, dentro de um mesmo estado’. Naturalmente,

seria possível obter esta informação a partir da junção das informações obtidas de vários relatórios separados e de consultas a bases oficiais, como o IBGE; no entanto, além de ser trabalhoso, a princípio, exigiria um agente humano. Caso a informação estivesse disponível em RDF, no formato de dados abertos e ligados, ela poderia ser recuperada por um agente computacional. Nesse caso, poderia haver um vocabulário compartilhado sobre municípios, por meio de uma ontologia, entre dois conjuntos de dados: por exemplo, um oriundo do armazém de dados da Câmara dos Deputados e outro, do IBGE.

Ademais, a linguagem utilizada no Fiscalize não pode ser considerada de fácil compreensão para um leigo, pois contém siglas sem tradução (como FNDE e LOA, por exemplo) e muitos termos técnicos, como ‘Unidade Orçamentária’, ‘Projeto-Atividade’ e ‘Empenho’. É fato que no Portal Orçamento Brasil existe uma página denominada ‘Entenda o Orçamento’<sup>6</sup>. No entanto, a página em si não apresenta informações explicativas sobre o orçamento, mas contém *hyperlinks* para documentos, como uma cartilha. No entanto, mesmo essa cartilha está em formato PDF, ou seja, não está em um formato aberto e legível por máquina. O problema da falta de clareza da linguagem é uma não-conformidade com o disposto na LAI, art. 8º, §3º, inc. I: “[...] Conter ferramenta de pesquisa de conteúdo que permita o acesso à informação de forma objetiva, transparente, clara e em linguagem de fácil compreensão [...]”.

Por último, em relação à LAI, o Fiscalize não divulga como a informação é estruturada; ou seja, não são explicitados os metadados que mostram como as categorias e os itens de informação se relacionam. Portanto, a falta de divulgação dos metadados é uma não-conformidade com o disposto na LAI, art. 8º, §3º, inc. IV: “[...] Divulgar em detalhes os formatos utilizados para estruturação da informação [...]”. A síntese do resultado da avaliação do Sistema Fiscalize em relação à LAI pode ser visto no quadro 6.

<sup>6</sup> Vide <<http://www2.camara.leg.br/atividade-legislativa/orcamentobrasil/entenda>> .

Quadro 6 - Síntese da Avaliação do Sistema Fiscalize frente à Lei de Acesso à Informação

Requisito	Dispositivo Legal	Conformidade
Conter ferramenta de pesquisa de conteúdo que permita o acesso à informação de forma objetiva, transparente, clara e em linguagem de fácil compreensão.	LAI, art. 8º, §3º, inc. I	Atende parcialmente
Possibilitar a gravação de relatórios em diversos formatos eletrônicos, inclusive abertos e não proprietários, tais como planilhas e texto, de modo a facilitar a análise das informações.	LAI, art. 8º, §3º, inc. II	Não atende
Possibilitar o acesso automatizado por sistemas externos em formatos abertos, estruturados e legíveis por máquina.	LAI, art. 8º, §3º, inc. III	Não atende
Divulgar em detalhes os formatos utilizados para estruturação da informação.	LAI, art. 8º, §3º, inc. IV	Não atende
Garantir a autenticidade e a integridade das informações disponíveis para acesso.	LAI, art. 8º, §3º, inc. V	Atende
Manter atualizadas as informações disponíveis para acesso	LAI, art. 8º, §3º, inc. VI	Atende

Fonte: Dados da pesquisa

#### 4.1.2 Sistema Siga Brasil

O Siga Brasil extrai dados do armazém de dados orçamentários do Senado e, por meio de técnicas de modelagem dimensional de dados e da tecnologia *Business Objects*, gera informações orçamentárias por meio de *relatórios dinâmicos*. Sua atualização é diária, o que implica que ele atende ao art. 8º, §3º, da LAI. Suas fontes são as bases de dados de três sistemas de informação do Governo e do Legislativo Federal (BRASIL, 2013): Sistema Integrado de Administração Financeira (SIAFI), Sistema de Elaboração Orçamentária (SE-LOR) e Sistema Integrado de Orçamento Público (SIOP). Aliás, tal fato, em tese, também garante a autenticidade e a integridade das informações disponíveis para acesso - LAI, art. 8º, §3º, inc. V.

É importante, nesse ponto, mostrar algumas diferenças relativas ao Sistema Fiscalize, no que tange à conformidade com a LAI. No Siga Brasil, por exemplo, é possível, a partir de então, exportá-lo para PDF, XLS (Excel) e CSV, em conformidade com o que dispõe o art. 8º, §3º, inc. II da LAI, da referida lei. Outra diferença é a possibilidade de fazer modificações no relatório, tais como: reordená-lo com base em algum outro critério; acrescentar cálculos estatísticos simples e acrescentar filtros. É possível ainda observar alguns relacionamentos entre as categorias;

no entanto, esses metadados só podem ser visualizados. Portanto, o Siga Brasil atende parcialmente o disposto no art. 8º, §3º, inc. IV da LAI. O sistema apresenta ainda metadados descritivos, que explicam alguns termos utilizados. Mesmo assim a linguagem não é de fácil compreensão para o público leigo. Dessa forma, o Siga Brasil atende apenas parcialmente o disposto no art. 8º, §3º, inc. I da LAI.

A despeito dos avanços do Siga Brasil em relação ao Fiscalize, em termos de dados abertos persistem alguns problemas. Apesar de ser bastante flexível e possibilitar a geração em formatos não proprietários, como o CSV, o acesso aos dados é indireto, via uma aplicação, e não diretamente, via um URI. Com isso, não seria possível automatizar o processo de extração dos dados. O formato CSV, por sua vez, é mais fácil de ser transformado em RDF, se comparado ao PDF. No entanto, o Siga Brasil não fornece os metadados semânticos em formatos recomendáveis pelo W3C, como o RDFS e a OWL. Na verdade, os metadados podem ser visualizados, mas não estão disponíveis em um local físico acessível por um URI baseado em HTTP. Logo, o Siga Brasil não atende o disposto no art. 8º, §3º, inc. III da LAI. A síntese do resultado da avaliação do Siga Brasil em relação à LAI é apresentada no quadro 7.

Quadro 7 - Síntese da Avaliação do Sistema Siga Brasil frente à Lei de Acesso à Informação

Requisito	Dispositivo Legal	Nível de Atendimento
Conter ferramenta de pesquisa de conteúdo que permita o acesso à informação de forma objetiva, transparente, clara e em linguagem de fácil compreensão.	LAI, art. 8º, §3º, inc. I	A t e n d e parcialmente
Possibilitar a gravação de relatórios em diversos formatos eletrônicos, inclusive abertos e não proprietários, tais como planilhas e texto, de modo a facilitar a análise das informações.	LAI, art. 8º, §3º, inc. II	Atende
Possibilitar o acesso automatizado por sistemas externos em formatos abertos, estruturados e legíveis por máquina.	LAI, art. 8º, §3º, inc. III	Não atende
Divulgar em detalhes os formatos utilizados para estruturação da informação.	LAI, art. 8º, §3º, inc. IV	A t e n d e parcialmente
Garantir a autenticidade e a integridade das informações disponíveis para acesso.	LAI, art. 8º, §3º, inc. V	Atende
Manter atualizadas as informações disponíveis para acesso	LAI, art. 8º, §3º, inc. VI	Atende

Fonte: Dados da pesquisa

A análise efetuada indica que os Sistemas Fiscalize e Siga Brasil, isoladamente, não atendem os requisitos estabelecidos pelo art. 8º da Lei de Acesso à Informação, no tocante às informações orçamentárias. Ou seja, o pressuposto P1 foi confirmado.

Serão discutidos e ilustrados a seguir aspectos referentes à aplicabilidade do padrão LOD para a construção de aplicações que atendam plenamente o disposto na legislação e permitam, no futuro, a interligação com outras informações de interesse governamental e público, as quais hoje, embora disponíveis, encontram-se dispersas e em formatos distintos, prejudicando sua interoperabilidade.

#### 4.2 O Padrão LOD

LOD (do inglês *Linked and Open Data*) deriva de dois outros conceitos: dados abertos (*open data*) e dados ligados (*linked data*). Conforme Bizer et al. (2009), os dados ligados “[...] referem-se a dados publicados na Web de modo que sejam legíveis por máquina, seus significados sejam explicitamente definidos, estejam ligados a outros conjuntos de dados e, por sua vez, possam ser ligados a partir de conjuntos de dados externos [...]”. Outro conceito, extraído da organização Linkeddata.org é: um termo usado para descrever uma boa prática recomendada para expor, compartilhar e conectar pedaços de dados, in-

formações e conhecimentos na WS usando URIs e RDF” (LINKEDDATA, 2013). Portanto, a despeito do nome ‘dados ligados’, o conceito transcende a exposição somente dos dados: por meio da formalização das relações entre as categorias associadas aos dados e suas instâncias, ou seja, os dados propriamente ditos, também informações e conhecimento a respeito dos dados são compartilhados. E o uso de URI e RDF, ao invés de URL e HTML, é fundamental. Segundo Heath and Bizer (2011), os *links* HTML indicam que dois documentos são relacionados, mas deixa ao usuário a tarefa de inferir a *natureza* do relacionamento. O RDF possibilita a quem publica os dados definir, *de forma explícita*, a natureza da conexão. A ideia de dados ligados de Berners-Lee (2006) foi sendo refinada por outros, como Heat e Bizer, que sintetizam os princípios dos dados ligados no artigo (HEAT ; BIZER, 2011). Esse princípios têm sido também incorporados na esfera governamental, no contexto do chamado Governo Aberto (*Open Government*). O fenômeno é denominado Dados Abertos de Governo (OGD – *Open Government Data*) e é descrito por Bauer e Kaltenböck (2012, p. 10). Em particular, no âmbito do governo, essa estratégia é denominada ‘LOD de governo’ – LOGD (*Linked Open Government Data*). Ela vem emergindo da Web de dados ligados como uma forma de facilitar a abertura, a interligação e o reuso dos dados abertos de governo (DING et al., 2012).

É relevante notar que os princípios dos dados ligados, aplicados ao governo, são compatíveis com os requisitos estabelecidos pela LAI. O atendimento ao Princípio 1 (“Usar URIs baseadas em HTTP para identificar objetos e conceitos abstratos”, de modo que “os clientes HTTP podem procurar a URI usando o protocolo HTTP e recuperar a descrição do recurso referenciado pelo URI”) e ao Princípio 3 (“Usar um único modelo de dados para publicar dados estruturados na Web: o RDF”) implica no atendimento a três dispositivos da LAI: (1) art. 8º, §3º, inc II: “possibilitar a gravação de relatórios em diversos formatos eletrônicos, inclusive abertos e não proprietários, tais como planilhas e texto, de modo a facilitar a análise das informações”; (2) art. 8º, §3º, inc. III: “possibilitar o acesso automatizado por sistemas externos em formatos abertos, estruturados e legíveis por máquina”; (3) art. 8º, §3º, inc. IV: “divulgar em detalhes os formatos utilizados para estruturação da informação”.

Os dados de governo ligados podem ser produzidos em duas etapas: conversão e aprimoramento (*enhancement*). A etapa de conversão é constituída, por sua vez, de duas fases principais. Na primeira, os dados de governo “brutos” (*raw data*) são “limpos” e gravados mediante representação baseada na estrutura RDF. Na segunda, estes dados convertidos podem ser referenciados por meio de URIs, de forma que tanto os dados quanto suas ontologias podem ser estendidas por terceiros: cidadãos, ONGs, outros órgãos de governo (BERNERS-LEE, 2009). Uma vez convertidos para a estrutura RDF, pode-se iniciar a etapa de aprimoramento (*enhancement*), a qual pode ser feita em quatro fases principais: adicionar dados aos dados de governo preexistentes; derivar dados de governo; ligar dados de governo; integrar dados de governo. Portanto, a implementação de *linked data* depende das tecnologias URI, RDF e SPARQL, brevemente descritas na Seção 2. No escopo deste texto, o importante é notar que, nesse processo, significados vão sendo adicionados aos dados, por meio da sofisticação da sintaxe das estruturas que os descrevem. Assim, as aplicações conseguem transformar esses dados em informações mais ricas e contextualizadas, as quais, por sua vez, podem ser associadas a outras, possibilitando uma recuperação mais precisa e completa. A seguir serão descritas duas aplicações construídas a partir dos pressupostos dos LOD e que utilizam as tecnologias mencionadas.

#### 4.2.1 Data-Gov Wiki

Segundo Ding et al. (2010) o Data-gov Wiki é um projeto que está sendo conduzido pelo Tetherless World Constellation, no Instituto Politécnico Rensselaer. Eles estão pesquisando o uso de tecnologias da WS na exploração e visualização de conjuntos de dados abertos de governo. A maior parte dos dados é oriunda do portal de governo norte-americano [data.gov](http://data.gov)<sup>7</sup>; o restante é oriundo de bases de dados de outros países, bem como fontes não-governamentais. Há uma transformação paulatina de dados brutos em informações úteis e acessíveis a diversos públicos, de forma flexível e variada. Uma das etapas na estratégia da Data-Gov Wiki consiste em mostrar o valor dos dados (previamente enriquecidos), usando aplicações de exploração e visualização. Para isso, foram construídos vários aplicativos demonstrativos. Para cada um deles, há uma página na Data-Gov Wiki, descrevendo seus objetivos e citando as tecnologias utilizadas na sua construção.

As tecnologias utilizadas nesta aplicação são: RDF, SPARQL, SparqlProxy e Google Visualization API<sup>8</sup>. Todas são gratuitas; ademais, as duas primeiras são ainda abertas e padronizadas pelo W3C. Além disso, são *aplicações genéricas*, ou seja, basta que os dados sejam passados em um formato predeterminado, a saída será produzida e a interação, possibilitada, não importando o conteúdo desses dados. Com isso, não há necessidade de se produzir novos aplicativos para cada contexto de uso, economizando recursos preciosos dos governos e das organizações civis. Além disso, são apresentadas também informações da origem dos dados, como a identificação do conjunto de dados (*Dataset 10040*) e suas características mais relevantes.

Este exemplo indica que é possível, a partir do uso das tecnologias disponíveis e utilizando o padrão LOD, atender ao que dispõe o inc. I do art. 3 da LAI : “[...] Conter ferramenta de pesquisa de conteúdo que permita o acesso à informação de forma objetiva, transparente, clara e em linguagem de fácil compreensão [...]”.

<sup>7</sup> Disponível em: <<http://www.data.gov>>.

<sup>8</sup> A lista completa das tecnologias utilizadas em cada aplicativo encontra-se em <[http://data-gov.tw.rpi.edu/wiki/Property:Technology\\_used](http://data-gov.tw.rpi.edu/wiki/Property:Technology_used)>.

#### 4.2.2 Ligado nos Políticos

Este projeto<sup>9</sup> foi proposto e desenvolvido por Araújo e Souza (2011). Foram utilizados como fontes de dados: os portais da Administração Pública do Tribunal Superior Eleitoral (TSE), do Senado e da Câmara dos Deputados; os portais das ONGs (Organização Não Governamentais) Políticos Brasileiros, Ficha Limpa e do projeto Excelências da Transparência Brasil (ARAÚJO; SOUZA, 2011). Os seguintes dados foram coletados: “dados pessoais, dados da eleição, divulgação de bens, dados parlamentares, lideranças, missões, mandatos, afastamentos, pronunciamentos, comissões, proposições e ocorrências” (p. 8). Os autores relatam que poucos dados estão disponíveis em formatos mais abertos e de mais fácil tratamento, como CSV; a maior parte estava disponível somente em formato HTML. Por isso, “foi necessária a criação e utilização de *Web Crawlers* para extrair os dados de uma forma metódica e automatizada” (p. 8). Esta técnica para extrair dados disponíveis na Internet e que não estão em formatos abertos é denominada *screen scraping* ou “raspagem de dados” (ARAÚJO ; SOUZA, 2011).

Também esta aplicação, assim como a anterior (Data.gov Wiki), demonstra o potencial das soluções baseadas em LOD na representação e disseminação da informação governamental, atendendo aos requisitos da Lei de Acesso à Informação. As aplicações indicam ainda a viabilidade técnica desse tipo de solução. Na próxima seção a aplicação dos LOD será discutida no âmbito da representação e disseminação da informação orçamentária.

#### 4.3 Viabilidade do LOD no Contexto da Informação Orçamentária

Pelo exposto, a conjunção dos dados abertos de governo e as tecnologias e ferramentas da WS já é uma realidade. Em particular, os casos apresentados buscam, justamente, a construção dos dados ligados (*linked data*) como etapa fundamental para agregar valor ao conjunto de dados original, enriquecendo-o com a associação a outros conjuntos de dados conexos. Ou seja, a utilização de conjuntos de dados (*datasets*) expressos em formatos abertos e referenciáveis,

no caso, o RDF, e associados a vocabulários comuns, também abertos e referenciáveis, expressos em RDFS e OWL.

No caso dos sistemas Siga Brasil e Fiscalize, ao contrário, os conjuntos de dados estão em bases de dados proprietárias e a semântica das relações entre as categorias (classes) e indivíduos (instâncias) está implícita nos sistemas e nas próprias bases. Isso acarreta algumas desvantagens: (1) os dados não podem ser acessados diretamente por uma aplicação via URI; (2) qualquer interligação entre os conjuntos de dados demanda grande esforço, uma vez que a semântica deverá ser inferida (por um processo de ‘engenharia reversa’), uma vez que não está explícita e formalizada, o que leva a (3) os usuários finais da informação, mesmo que tenham acesso aos dados, não conhecem as relações entre eles: e isso, certamente, reduz o valor da informação.

O caso das informações sobre transferências a municípios, explorado na seção anterior, ilustra bem essa situação, uma vez que estão dispersas em diversos sistemas: Selor, Siafi, Siconv e Sistema de Acompanhamento de Obras da Caixa Econômica Federal. No Selor encontram-se informações sobre a elaboração orçamentária, nos quais os programas de trabalho são definidos. O Siafi contém informações sobre a execução orçamentária desses programas de trabalho. O Siconv possui o cadastro das informações das transferências voluntárias (incluindo convênios), os quais, como se viu, são associados a programas de trabalho. Finalmente, o Sistema de Acompanhamento de Obras possui informações do andamento de obras, as quais podem ser objetos de contratos de repasse firmados entre a União e os municípios. Portanto, os dados são complementares, mas não estão disponíveis a partir de um só ponto de acesso, nem são totalmente processáveis por máquina. Um dos pressupostos desta pesquisa (P2) é que a construção de aplicações e conjuntos de dados (*datasets*) baseadas no padrão LOD é uma solução viável para a representação e disseminação da informação orçamentária, não só atendendo ao que determina a Lei de Acesso à Informação, mas permitindo a interligação entre dados e metadados, aumentando, assim, o valor dessa informação.

Baseado no que foi visto, seriam quatro as etapas para a construção de uma aplicação de dados ligados e informações orçamentárias

<sup>9</sup> Disponível em: <<http://ligadonospoliticos.com.br>>.

(adaptado de DING et al., 2010): (1) Converter os dados orçamentários em formato legível por máquina - no caso, RDF; (2) Enriquecer e ligar os dados orçamentários por meio de vários tipos de processamento (p. ex: extração, normalização e mapeamento); (3) Mostrar o valor dos dados orçamentários enriquecidos usando aplicações interessantes de exploração e visualização; (4) Mostrar a proveniência do conhecimento sobre os dados orçamentários por meio de metadados e processos de rastreabilidade.

Uma forma de avaliar a viabilidade da adoção do LOD é analisar a viabilidade das etapas citadas. Com esse intuito, será explicado no que consiste cada etapa. Serão destacados exemplos de tecnologias usadas em cada uma delas, bem como aplicações que as utilizam. Além disso, a título de ilustração das etapas 1 e 2, será apresentado um exercício de modelagem de uma ontologia.

#### 4.3.1 A Viabilidade da Etapa de Conversão

Segundo DING et al. (2010, p. 39), o processo de conversão é “direto e extensível”. Inicialmente os dados ‘crus’ de governo (*raw data*) “[...] passam por um processo de ‘limpeza’ e preservação por meio de representações baseadas em RDF [...]”. Posteriormente, “[...] esses conjuntos de dados convertidos passam a ser navegáveis via URIs, de forma que tanto os conjuntos de dados quanto suas ontologias possam ser estendidos por terceiros [...]”. Para se iniciar a conversão define-se uma ou mais fontes de dados orçamentários. Conforme a Seção 3.2, são fontes as bases de dados dos sistemas a seguir: Sistema de Elaboração de Leis Orçamentárias (Selor), Sistema Integrado de Administração Financeira (Siafi), Siafi Gerencial, Sistema de Gestão de Convênios e Contratos de Repasse (Siconv) e Sistema de Acompanhamento de Obras da Caixa Econômica Federal (Siurb)<sup>10</sup>. Naturalmente, não há necessidade de se converter todas as bases para se passar à próxima etapa. O trabalho pode ser feito de forma iterativa e incremental. Importante ressaltar que várias dessas bases de dados estão disponíveis, sendo que algumas delas em formato de dados abertos, ainda que não de dados ligados. A base de dados do Selor é

hospedada e gerida pelo Senado e contém dados sobre elaboração orçamentária: PPA, LDO e LOA (BRASIL, 2007); ela ainda não está disponível diretamente em formatos abertos. A base de dados do Siconv está disponível nos formatos CSV e XML e com acesso via HTTP e já é aderente ao padrão dados abertos<sup>11</sup>, mas não ao padrão dados ligados. Os dados do Siafi Gerencial são transmitidos diariamente tanto à Câmara dos Deputados, quanto ao Senado - são esses dados que alimentam os armazéns de dados, que, por sua vez, são utilizados pelos Sistemas Fiscalize e Siga Brasil, conforme Seção 3.2.

Portanto, entre as principais fontes de dados, restaria apenas o acesso às bases de dados do Siafi e do Siurb. Acredita-se, portanto, que os dados do Siafi Gerencial disponíveis hoje no Congresso Nacional sejam suficientes para se iniciar o trabalho de conversão. Quanto aos dados do Siurb, há a possibilidade de se firmar um convênio com seu gestor, a Caixa Econômica Federal, para a geração dos dados em algum formato aberto, que facilitasse a conversão para RDF. De qualquer forma, as bases já disponíveis fornecem dados suficientes para, após a realização das três etapas de construção dos LOD, prover o acompanhamento integrado de grande parte da elaboração e da execução orçamentária.

A etapa de conversão exige também o uso de tecnologias, as quais poderiam ser desenvolvidas *interna corporis*; no entanto, felizmente, várias delas já estão disponíveis, em licenças *free and open source* (FOS). Algumas delas são destacadas no Quadro 8. Para conhecer as tecnologias citadas e como utilizá-las, recomenda-se a leitura de Hebler et al. (2009) e Allemang e Hendler (2009), além dos *websites* citados no rodapé da própria tabela. Portanto, a etapa de conversão é viável e deve ser feita de forma iterativa e incremental, uma vez que já existem bases de dados e tecnologias disponíveis. A conversão poderia se iniciar com as bases já existentes em formatos abertos (SICONV, LOA 2012). Posteriormente, poderiam ser convertidas aquelas as quais o Congresso Nacional tem acesso (Selor, Siafi Gerencial). Futuramente, na medida em que outras bases sejam abertas (Siurb, Siafi), elas também poderiam ser incorporadas. Para isso, seria necessário que as ontologias fossem estendidas, o que corresponde à etapa de enriquecimento.

<sup>10</sup> Vide <[http://www1.caixa.gov.br/gov/gov\\_social/municipal/assistencia\\_tecnica/siurb\\_index.asp](http://www1.caixa.gov.br/gov/gov_social/municipal/assistencia_tecnica/siurb_index.asp)>.

<sup>11</sup> Vide <<http://api.convenios.gov.br/siconv/doc/>>.

Quadro 8 - Tecnologias para conversão de dados para RDF, manipulação e gravação

Tecnologia	Descrição
csv2rdf4lod <sup>(1)</sup>	Permite a conversão de dados em formato 'valores separados por vírgulas' ( <i>Comma-Separated Values</i> – CSV) para RDF.
Virtuoso Triple Store <sup>(2)</sup>	Permite armazenar triplas RDF.
Apache Jena™ <sup>(3)</sup>	Provê uma coleção de ferramentas escritas em Java que auxiliam o desenvolvimento de aplicações baseadas em WS e dados ligados.
RDFLib <sup>(4)</sup>	Biblioteca baseada na linguagem de programação Python, permite gerar, manipular e recuperar informação a partir de triplas RDF.
pyrdfa3 <sup>(5)</sup>	Consegue extrair dados expressos em RDFa 1.1 a partir de (X)HTML, SVG e XML. Baseada em Python.
R2RML <sup>(6)</sup>	Uma linguagem para expressar mapeamentos customizados de bases de dados relacionais para conjuntos de dados RDF.

**Fontes:** (1) <[http://logd.tw.rpi.edu/tutorial/getting\\_started\\_csv2rdf4lod\\_create\\_verbatim\\_rdf\\_conversions\\_tabular\\_data](http://logd.tw.rpi.edu/tutorial/getting_started_csv2rdf4lod_create_verbatim_rdf_conversions_tabular_data)>; (2) <<http://www.openlinksw.com/wiki/main/>>; (3) <<http://jena.apache.org/>>; (4) <<https://github.com/RDFLib>>; (5) <<https://github.com/RDFLib>>; (6) <<http://www.w3.org/TR/r2rml/>>.

### 4.3.2 A Viabilidade da Etapa de Enriquecimento

O enriquecimento visa transformar os dados em RDF, para que possam ser ligados com outros conjuntos de dados, por meio do uso de ontologias comuns. Uma das formas, segundo Ding et al. (2010) é fazer com que as propriedades usadas nos dados RDF sejam associadas a termos oriundos de ontologias bem conhecidas, tais como FOAF e Dublin Core. Essa abordagem permite que os conjuntos de dados possam ser usados por ferramentas da WS, como o Tabulator (navegador para dados ligados). Quando se navega em um conjunto de dados, ele procura por valores associados a marcadores 'rdfs:label' (ou subpropriedades correspondentes), os quais exibirá no lugar das URIs, tornando os dados mais legíveis para humanos. Outra estratégia é ligar os conjuntos de dados à nuvem LOD. Tanto para a etapa de conversão, quanto para a etapa de *enriquecimento e ligação dos dados orçamentários* (2), há a necessidade da definição de vocabulários comuns, por meio de ontologias.

Apresenta-se um exercício de modelagem, o qual ilustrará alguns passos da metodologia

proposta por Uchold. Essa metodologia foi escolhida por ser baseada em *cenários de motivação*, o qual é justamente o caso da informação orçamentária: há um cenário que motiva a adoção de ontologias que permitam a descrição dos dados orçamentários em RDF e a ligação destes com outras bases de interesse público que hoje estão dispersas e, às vezes, pouco acessíveis. A figura 34 dá uma visão geral do método.

Figura 34 - Método de Uchold para a construção de ontologias (skeletal methodology)



**Fonte:** BREITMAN (2006).

Prosseguindo a ilustração baseada na metodologia, na primeira etapa, o *propósito* seria o descrito no quadro 9:

Quadro 9 - Exercício de modelagem – Etapa ‘Identificar o Propósito’

“Descrever dados referentes à execução orçamentária da União em RDF, de modo a facilitar a construção de aplicações que permitam uma visão integrada entre a elaboração e a execução orçamentária pela sociedade e pelo próprio governo”.

Fonte: Dados da pesquisa.

A segunda etapa, *capturar*, consiste em “[...] definir textualmente conceitos e relacionamentos [...]” (BREITMAN, 2006, p. 70). No caso do Orçamento Público, há um ponto de partida bastante útil: as categorias e os conceitos são comuns às diversas aplicações, pois são definidos por normas federais, que devem ser seguidas por todas elas. Assim, por exemplo, o termo ‘empenho’ é utilizado nos diversos sistemas para designar o mesmo conceito, o qual é

associado à categoria na qual se enquadra todos os empenhos específicos (indivíduos ou instâncias dessa categoria). O mesmo se dá em relação aos termos ‘convênio’, ‘programa’, ‘conveniente’, ‘concedente’, ‘unidade orçamentária’ etc. Isso significa que já existe um vocabulário implicitamente compartilhado entre as aplicações.

Continuando o exercício de modelagem, a partir dos conceitos descritos em 3.1, são definidas as seguintes relações (Quadro 10).

Quadro 10 - Exercício de modelagem – Etapa ‘Capturar’

Tipo de Relação	Descrição em Linguagem Natural (LN)
Relação entre ‘despesa’ e ‘empenho’	‘Todo item de despesa é empenhado por uma nota de empenho’.
Relação entre ‘despesa’, ‘empenho’ e ‘liquidação’	‘Todo item de despesa é liquidado por uma nota de liquidação, desde que esse item já tenha sido empenhado’ (ou seja, desde que haja uma reserva orçamentária associada).
Relação entre ‘despesa’, ‘pagamento’ e ‘liquidação’	‘Todo item de despesa é realizado (pago) por uma ordem de pagamento de uma despesa, desde que esse item já tenha sido liquidado’ (ou seja, o pagamento da despesa já tenha sido autorizado).

Fonte: Dados da pesquisa.

A próxima etapa (terceira), *codificar*, consiste na representação dos conceitos e relacionamentos definidos na etapa de captura, por meio de uma linguagem formal (BREITMAN, 2006, p. 71). É necessário “[...] decidir se cada termo será modelado como classe, entidade ou relacionamento [...]”.

Continuando o exercício proposto e utilizando os exemplos construídos anteriormente, ter-se-iam as classes, relacionamentos expressos no Quadro 11 e os axiomas – em lógica de primeira ordem e em lógica descritiva – expressos no Quadro 12.

Quadro 11 - Exercício de modelagem – Etapa ‘Codificar’ – Classes e Relacionamentos

Expressão	Nome OWL	Símbolo	Tipo
‘item de despesa’	ItemDeDespesa	ID(x)	Classe
‘nota de empenho’	NotaDeEmpenho	NE(x)	Classe
‘nota de liquidação’	NotaDeLiquidacao	NL(x)	Classe
‘ordem de pagamento’	OrdemDePagamento	OP(x)	Classe
‘é liquidado por’	ehLiquidadoPor	lp(x,y)	Relacionamento / Propriedade
‘é empenhado por’	ehEmpenhadoPor	ep(x,y)	Relacionamento / Propriedade
‘é realizado por’	ehRealizadoPor	rp(x,y)	Relacionamento / Propriedade

Fonte: dados da pesquisa.

Quadro 12 - Exercício de modelagem – Etapa ‘Codificar’ – Axiomas em LPO e DL

Axioma em Linguagem Natural	Axioma em Lógica de Primeira Ordem	Axioma em Lógica Descritiva
Todo item de despesa é empenhado por uma nota de empenho.	$\exists v(\forall x(ep(x, v) \rightarrow NE(v)))$ $\exists y(\forall x(ep(x, y) \rightarrow ID(x)))$	(1) $T \sqsubseteq \forall ev. NE$ (2) $T \sqsubseteq \forall ep. ID$
Todo item de despesa é liquidado por uma nota de liquidação, desde que esse item já tenha sido empenhado.	$\exists y(\forall x(lp(x, y) \rightarrow NL(y)))$ $\exists y(\forall x(lp(x, y) \rightarrow ID(x) \wedge \exists z(ep(x, z) \wedge NE(z))))$	(3) $T \sqsubseteq \forall lv. NL$ (4) $T \sqsubseteq \forall lp. ID \sqcap \exists ep. NE$
Todo item de despesa é realizado (pago) por uma ordem de pagamento de uma despesa, desde que esse item já tenha sido liquidado.	$\exists y(\forall x(rp(x, y) \rightarrow OP(y)))$ $\exists y(\forall x(rp(x, y) \rightarrow ID(x) \wedge \exists z(lp(x, z) \wedge NL(z))))$	(5) $T \sqsubseteq \forall rp. OP$ (6) $T \sqsubseteq \forall rp. ID \sqcap \exists lp. NL$

Fonte: Dados da pesquisa.

Os axiomas da última coluna do Quadro 12 indicam a imagem (*range*) e o domínio (*domain*) das propriedades – exemplos: axiomas (1) e (2), respectivamente. O símbolo **T** representa uma tautologia, e equivale à classe Thing da OWL. Para inferir os axiomas, foi necessário o uso de equivalências lógicas, omitidas por simplicidade. Para referências sobre lógica de primeira ordem,

recomenda-se Mortari (2001). Para equivalências entre Lógica de Primeira Ordem (FOL) e Lógica Descritiva (DL), recomenda-se Grosz et al. (2003) e Baader et al. (2003). Os axiomas da última coluna podem ser convertidos para OWL, uma vez que esta linguagem é um tipo de lógica descritiva, conforme foi visto. Existem ferramentas que podem auxiliar as etapas de captura, codificação, integração, validação e

codificação. A principal delas é a plataforma Protégè<sup>12</sup> (STANFORD, 2013).

Finda a etapa de codificação, pode-se iniciar a etapa de integração. No caso dos dados orçamentários, já existe uma ontologia, denominada Ontologia da Classificação da Despesa do Orçamento Federal, desenvolvida pela Secretaria de Orçamento e Finanças (SOF), e expressa em OWL. Com base nessa ontologia, foi gerado um conjunto de dados (*dataset*), referentes à base de dados da Lei Orçamentária Anual (LOA) de 2012. O *dataset*, descrito em RDF, contém a descrição de todos os itens de despesa que estavam programados para serem executados em 2012, com seus valores de dotação inicial. Ele já está disponível no portal <<http://orcamento.dados.gov.br>>. O arquivo OWL contém a descrição dos metadados associados aos itens de despesa e como eles se relacionam. Existem ainda manuais de utilização disponíveis na Web (BRASIL, 2012).

O Quadro 13 mostra as classes definidas nessa ontologia e as relaciona ao Manual Técnico do Orçamento de 2012. A Figura 35 exibe um mapeamento entre a ontologia e alguns recursos expressos em RDF. Hipoteticamente, poderia ser criada, a partir do exercício proposto, uma ontologia OWL de Execução da Despesa compatível com a Ontologia de Classificação da Despesa do Orçamento Federal. A classe ItemDeDespesa (Item de Despesa) seria o elo entre ambas. A Figura 36 ilustra essa ideia. Trata-se de uma ontologia simplificada. Mas ela indica que a etapa de enriquecimento no contexto dos dados do Orçamento Federal é viável, utilizando os métodos citados na literatura, como o Uschold, bem como as ontologias e tecnologias disponíveis. Uma ontologia mais sofisticada demandaria a execução da etapa de avaliação, além de novas iterações das etapas já mencionadas, as quais ampliariam o vocabulário, por meio de novas classes, propriedades e axiomas. Após sucessivas iterações, caso a ontologia fosse julgada suficientemente madura, ela poderia ser, finalmente, documentada.

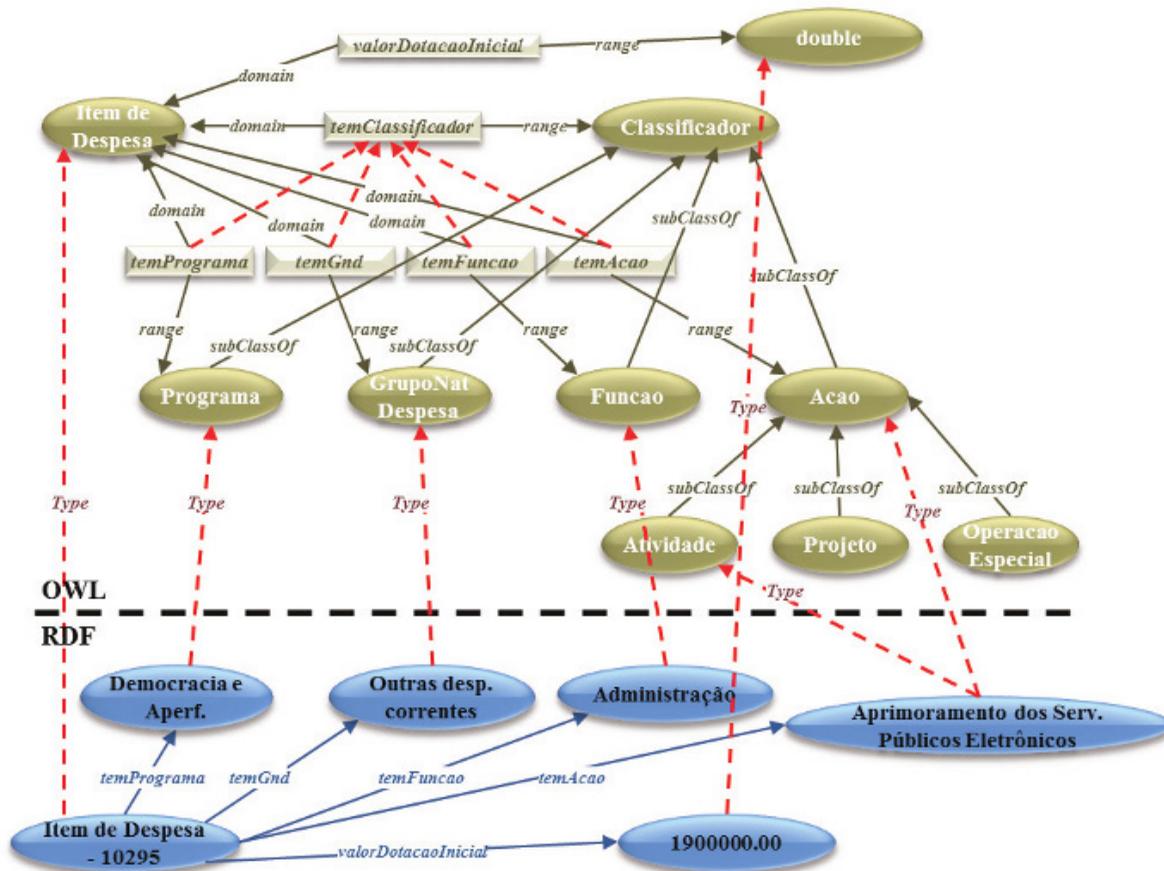
Quadro 13 - Ontologia da Classificação da Despesa do Orçamento Federal

Classes	MTO 2012
<b>Categoria Economica</b>	Categoria Econômica
<b>Classificador</b>	Sem correspondência. Classe de suporte ao conceito de classificação, das quais, todas as outras classes que se referem ao conceito de classificação da despesa se especializam. Tem como propriedades a descrição ( <i>label</i> ) e o código ( <i>codigo</i> ).
<b>ElementoDespesa</b>	Elemento de Despesa
<b>Esfera</b>	Esfera
<b>FonteRecursos</b>	Fonte de Recursos
<b>GrupoNatDespesa</b>	Grupo de Natureza da Despesa (GND)
<b>CategoriIndicadorUsoaEconomica</b>	Indicador de Uso
<b>ItemDeDespesa</b>	Sem correspondência. Cada elemento da classe ItemDeDespesa é caracterizado por ter propriedades ( <i>object properties</i> ) específicas para cada uma das subclasses da classe Classificadores, além de possuir um <i>data type property</i> denominado ValorDotaçãoInicial que é do tipo double.
<b>Localizador</b>	Subtítulo (Localizador do gasto)
<b>ModalidadeAplicacao</b>	Modalidade de Aplicação
<b>Programa</b>	Programa
<b>Acao</b>	Ação
<b>Projeto</b>	Projeto
<b>Atividade</b>	Atividade
<b>OperacaoEspecial</b>	Operação Especial
<b>Funcao</b>	Função
<b>Subfuncao</b>	Subfunção
<b>UnidadeOrcamentaria</b>	Unidade Orçamentária (UO)
<b>Orgao</b>	Órgão
<b>ResultadoPrimario</b>	Identificador de Resultado Primário

Fonte: <<http://vocab.e.gov.br/2012/08/loa2012#modelo>>.

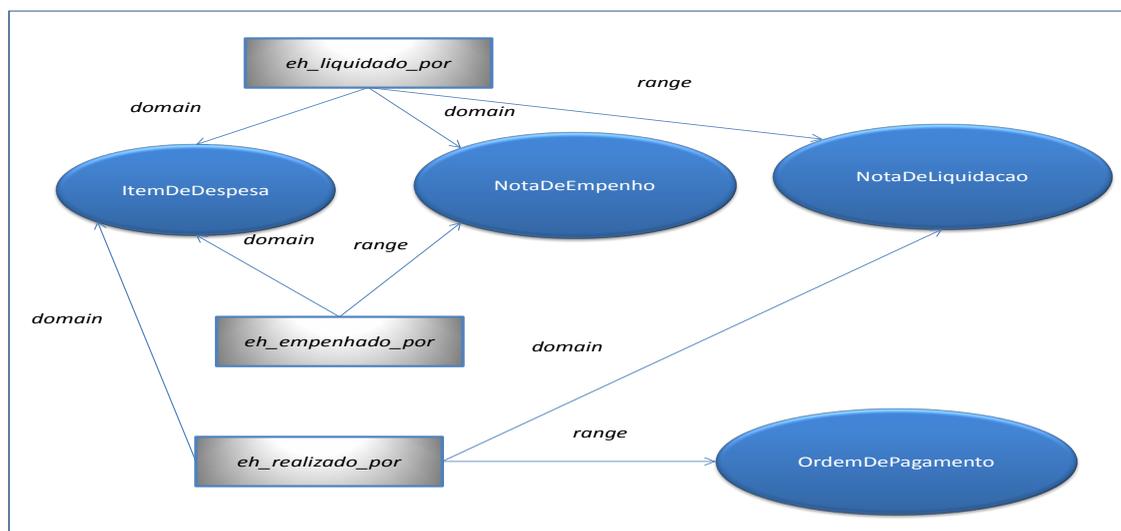
<sup>12</sup> Disponível em: <<http://protege.stanford.edu/>>.

Figura 35 - Mapeamento entre a Ontologia de Classificação da Despesa do Orçamento Federal e um recurso



Fonte: <<http://vocab.e.gov.br/2012/08/loa2012#modelo>>.

Figura 36 - Exercício de modelagem - Etapa Integrar



Fonte: Elaborada pelo autor.

### 4.3.3 A viabilidade da Etapa de Visualização

Estando os dados abertos, ligados e enriquecidos, é necessário mostrar o seu valor. Conforme visto nas aplicações Data.gov Wiki e Ligado nos Políticos, uma vez que os dados estão no padrão *LOD*, é possível usar várias tecnologias de busca, exploração e visualização de dados. Existem ainda ferramentas de visualização baseadas em *mashups* (BIZER et al.,

2006). Um exemplo já mencionado nesta pesquisa são aplicações que possuem uma interface para navegação por coordenadas geográficas. O Quadro 14 apresenta algumas que podem ser usadas para visualização de dados. A figura 38 exibe uma aplicação que ilustra a utilização dessas tecnologias. Em relação à busca, a principal tecnologia é o SPARQL, descrita no capítulo 2. O portal citado também possui uma interface de busca SPARQL, que funciona sobre uma plataforma Open Link Virtuoso. Ela pode ser acessada pelo endereço: <<http://orcamento.dados.gov.br/sparql/>>.

Quadro 14 - Ferramentas para visualização de dados do tipo 'mashup'

Tecnologia	Descrição	Endereço
<u>Google Visualization API</u>	Permite acessar múltiplas fontes de dados estruturados, e exibi-los por diversas formas de visualizações.	<a href="http://code.google.com/apis/visualization/">http://code.google.com/apis/visualization/</a>
<u>Google Visualization GeoMap</u>	Extensão da Google Visualization API que permite exibir dados referenciados geograficamente	<a href="http://code.google.com/apis/visualization/documentation/gallery/geomap.html">http://code.google.com/apis/visualization/documentation/gallery/geomap.html</a>
<u>Google Visualization Pie Chart</u>	Extensão da Google Visualization API que permite construir gráficos estatísticos.	<a href="http://code.google.com/apis/visualization/documentation/gallery/piechart.html">http://code.google.com/apis/visualization/documentation/gallery/piechart.html</a>
<u>Simile Exhibit</u>	Permite criar páginas web com busca textual, filtragem, mapas interativos e linhas de tempo.	<a href="http://simile-widgets.org/exhibit/">http://simile-widgets.org/exhibit/</a>

Fonte: Elaborado pelo autor a partir do portal Data.gov Wiki: <[http://data-gov.tw.rpi.edu/wiki/The\\_Data-gov\\_Wiki](http://data-gov.tw.rpi.edu/wiki/The_Data-gov_Wiki)>.

Há, portanto, uma gama de tecnologias, que possibilita a visualização. Isso indica a viabilidade da implementação da etapa de visualização de *LOD* pelo Congresso Nacional, provendo múltiplas formas de acesso aos dados orçamentários pelos cidadãos, empresas, organizações não-governamentais e pelo próprio governo. Com o uso dessas tecnologias associadas às demais descritas ao longo desta seção, os dados orçamentários poderiam ser interligados com dados socioeconômicos, urbanísticos, demográficos e outros, possibilitando não apenas transparência, mas uma visão integrada do país aos gestores públicos, valiosa para o planejamento e acompanhamento das ações do Estado. Naturalmente chegar nesse ponto requer um esforço considerável, mas, a estratégia

baseada em dados ligados é, até o presente momento, a que provê maior flexibilidade e interoperabilidade.

### 4.3.4 A Viabilidade da Etapa de Proveniência do Conhecimento sobre os Dados

Assegurar a proveniência dos dados aumenta a credibilidade da informação gerada, p.ex. qual é a base de origem e a data em que foi feita a extração dos dados. Ding et al. (2010) explicam que uma estratégia é a utilização de metadados de proveniência. No projeto Data.gov Wiki (acima) eles usaram vários metadados: *título*, para o título do conjunto de dados; *agência* para o órgão governamental responsável pela publicação; *cobertura geográfica*, para localidades

geográficas mencionadas no conjunto de dados, categoria, para a classificação do conjunto de dados, e data de alteração, para quando o conjunto de dados foi modificado; etc.

A implementação desses metadados pode ser trabalhosa, mas é plenamente viável. Para os conjuntos de dados que ainda não estão disponíveis diretamente em CSV, como o Siafi Gerencial, duas estratégias podem ser adotadas: solicitar ao órgão gestor a geração e a disponibilização do conjunto de dados convertido, ou solicitar apenas a licença para a publicação, uma vez que, como foi mencionado, tanto a Câmara quanto o Senado já possuem cópias desses conjuntos de dados, porém em formato relacional e proprietário. O registro da proveniência dos dados e o acesso aos dados originais seriam essenciais, uma vez que correspondem a um requisito da Lei de Acesso à Informação: mais precisamente, ao inc. IV do art. 3º: “garantir a autenticidade e a integridade das informações disponíveis para acesso”. Trata-se, portanto, de uma etapa necessária e viável.

Como conclusão verificou-se o não-atendimento pleno dos requisitos da LAI pelos sistemas Siga Brasil e Fiscalize, no que tange às informações orçamentárias, conforme Seção 3.2. Tal fato confirma o pressuposto (P1) e leva a uma tentativa de solução, que passa, necessariamente, pela adoção do padrão *dados abertos* de forma mais ampla pelo Congresso Nacional. A segunda questão é o fato que a LAI não determina, explicitamente, a adoção do padrão LOD. Os argumentos deste trabalho, especialmente nas seções 4.3 e 4.4, indicam o potencial e a viabilidade da sua adoção, por todos os motivos que foram expostos, os quais podem ser sintetizados em: já existem conjuntos de dados em formatos abertos, que podem ser convertidos com relativa facilidade; o Congresso Nacional possui acesso ou gerencia importantes bases de dados orçamentários em formato relacional, as quais também podem ser convertidas; além disso, já existem conjuntos de dados em formatos abertos e ligados, que podem se somar às demais. Ademais, o Congresso Nacional, como ator imprescindível em várias fases do processo orçamentário, pode e deve contribuir, tanto na questão da modelagem quanto no fornecimento de dados. Finalmente, existem tecnologias disponíveis em qualidade e quantidade, já testadas com sucesso em outros

projetos. Dessa forma, o pressuposto (P2), relativo à adoção do padrão LOD pelo Congresso Nacional no que tange ao Orçamento também é confirmado, com a ressalva de que a adoção pode e deve ser paulatina, trabalhando cada base de dados de acordo com a sua disponibilidade e com a capacidade dos setores envolvidos.

## 5 CONSIDERAÇÕES FINAIS

Nas democracias contemporâneas, a transparência é palavra de ordem. No entanto, a transparência não pode ser meramente formal: ela deve ser efetiva; ou seja, deve ser materializada por meio de ações concretas, alcançando o maior público possível. As Novas Tecnologias da Informação e Comunicação não podem ser ignoradas nesse processo. A Web, em particular, é a principal plataforma tecnológica para a efetivação da transparência, uma vez que, além do seu longo alcance, possibilita a interatividade governo-sociedade. A Lei de Acesso à Informação – LAI, ao mesmo tempo em que vem ao encontro dos anseios por transparência, está em sintonia com as novas tecnologias, pois indica, claramente, a necessidade da exposição dos dados de forma aberta e acessível por máquinas.

A Web vem evoluindo drasticamente: a WS, que poderia ser considerada apenas uma visão há alguns anos, possui, atualmente, um real potencial de implementação. Conforme visto ao longo deste texto, ela é um resultado de anos de estudos, com forte embasamento teórico e possui, como aliada, uma vasta gama de tecnologias que lhe dão suporte. Os órgãos públicos não deveriam ignorar tal fato. Na verdade alguns deles, ainda que de forma incipiente se comparados a outros países, vem tomando iniciativas no sentido de adotar padrões e tecnologias da WS para a concretização da transparência, por meio dos chamados LOD. Uma dessas iniciativas é a criação do portal de Dados Abertos do Orçamento Federal: <<http://orcamento.dados.gov.br>>. Trata-se de iniciativa recente, conduzida pelo Ministério do Planejamento. Por isso mesmo, contém apenas os dados da elaboração orçamentária de 2012.

O Congresso Nacional, como um dos atores institucionais responsáveis pela elaboração e fiscalização da execução orçamentária, não poderia ficar alijado desse processo. Contudo,

apesar do avanço na abertura dos dados, tanto a Câmara quanto o Senado não possuem ainda iniciativa de peso na adoção dos LOD no que tange à área orçamentária. Até o presente momento, nenhuma das Casas Legislativas tornou disponível nenhum conjunto de dados no padrão LOD. Ademais, conforme demonstrado, os dois principais sistemas de informações orçamentárias baseados na Web no âmbito do Congresso Nacional – Fiscalize e Siga Brasil – não satisfazem plenamente os requisitos da Lei de Acesso à Informação.

Os estudos de caso descritos indicam a viabilidade da adoção ampla do padrão LOD no contexto do Orçamento Público. Naturalmente existe uma questão operacional, que envolve aspectos gerenciais e tecnológicos. Os aspectos gerenciais poderiam contar com o suporte da Comunidade TI Controle<sup>13</sup>, que já apoia importantes projetos, como o LexML<sup>14</sup>. Essa Comunidade coordenaria as etapas necessárias para a ampliação e aprimoramento dos LOD de orçamento, desde a identificação das fontes de dados até a construção e integração de ontologias e esquemas. A execução das etapas poderia ser

dividida entre os diversos órgãos envolvidos, de acordo com sua capacidade técnica e operacional. Reconhece-se que algumas etapas, como a construção de ontologias necessitam de conhecimento especializado. Porém, tanto a Câmara dos Deputados quanto o Senado possuem um corpo técnico qualificado na área orçamentária, por meio das suas consultorias, bem como nas áreas de Ciência da Computação/ Informação e Biblioteconomia. Além disso, poderiam ser firmados convênios com universidades e institutos de pesquisa, numa interação benéfica a todos os envolvidos.

O estudo apresentado, embora preliminar, indica que a aplicação do padrão LOD nas informações relativas ao Orçamento Público seria proveitosa tanto para o Congresso Nacional e demais atores envolvidos – por utilizar uma gama de técnicas, tecnologias e ferramentas maduras e flexíveis, quanto para a sociedade em geral, que ganharia poder de (re)elaborar a informação pública de acordo com suas necessidades. E isso seria um importante passo no cumprimento dos requisitos legais e institucionais de transparência e controle social.

## **NECESSITY AND FEASIBILITY OF ADOPTING LOD STANDARD BY NATIONAL CONGRESS: a study in the context of the public budget**

### **ABSTRACT**

*The objective of the research was to analyze the necessity and feasibility of the adoption of the Linked Open Data (LOD) standard by the Brazilian National Congress regarding information about the Federal Budget. The survey begins with a brief review of the theoretical foundations of the Semantic Web (SW). Then a research on the key technologies of SW was undertaken, concerning especially those associated with the LOD standard. Subsequently, we designed and implemented a study of multiple cases in which we evaluated the information systems used for the management of Congress' budget. We analyzed specially those which have web interface: Siga Brasil (Senate) and Fiscalize (House of Representatives). Continuing, we evaluated the feasibility of adopting the LOD standard by the Congress with respect to budget information. For this, we expanded the initial literature review, in order to have more clarity on the principles of the LOD and the technologies and methodologies necessary for its implementation. Finally, we made another case study, which evaluated two computer applications based on this standard. The results indicate the necessity and feasibility of the adoption of the LOD by the Congress in the context of information associated with the Public Budget.*

### **Keywords:**

*Right to information. Administrative transparency. Public Budget. Information technology and communication. Semantic Web.*

---

Artigo recebido em 10/07/2013 e aceito para publicação em 07/03/2014

---

<sup>13</sup> Comunidade de Tecnologia da Informação Aplicada ao Controle (TIControle), reúne representantes dos Poderes Legislativo, Executivo, Judiciário, do Ministério Público e da Advocacia-Geral da União. Foi criada em 2006 e tem por finalidade contribuir para o incremento da eficiência, eficácia e efetividade na gestão pública. Disponível em: <<http://www.ticontrole.gov.br>>.

<sup>14</sup> O LexML é uma rede de informação legislativa e jurídica que pretende organizar, integrar e dar acesso às informações disponibilizadas nos diversos portais de órgãos do governo na Internet. Disponível em: <<http://lexml.gov.br>>.

## REFERÊNCIAS

ALANI, H. et al. Unlocking the Potential of Public Sector Information with Semantic Web Technology. In: **ISWC – International Semantic Web Conference, 2007**.

ALLEMANG, D.; HENDLER, J. Semantic Web for the Working Ontologist: effective modeling in RDFS and OWL. 2nd. ed. Waltham: Morgan Kaufmann, 2011.

ALMEIDA, M. B.; SOUZA, R. R. Avaliação do espectro semântico de instrumentos para organização da informação. **Enc. Bibli. R. Eletr. Bibliotecon. Ci. Inf.**, Florianópolis, v. 16, n. 31, p.25-50, 2011.

ARAÚJO, L. R.; SOUZA, J. F. Aumentando a transparência do governo por meio da transformação de dados governamentais abertos em dados ligados. **Revista Eletrônica de Sistemas de Informação**, v. 10, n. 1, 2011. Disponível em: <<http://revistas.facecla.com.br/index.php/reinfo/article/download/880/pdf>>. Acesso em: maio de 2012.

BAADER, F.; HORROCKS, I.; SATTler, U. Description Logics as Ontology Languages for the Semantic Web. In: Festschrift in honor of Jörg Siekmann, Lecture Notes in Artificial Intelligence. Berlin: Springer Verlag, 2003, p. 228-248.

BERNERS-LEE, T. **Linked data**. [S.l.]: W3C, 2009. Disponível em: <<http://www.w3.org/DesignIssues/LinkedData.html>>. Acesso em: maio de 2012.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The semantic web: a new form of web content that is meaningful to computers will unleash a revolution of new possibilities. **Scientific American**, New York, maio 2001. Disponível em: <<http://trac.assembla.com/soray/export/230/user/Marcell/readings/Semantics/The%20Semantic%20Web.pdf>>. Acesso em: maio de 2012.

BEZERRA FILHO, J. E. **Contabilidade Pública: teoria, técnica de elaboração de balanços e 500 questões**. 2. ed. Rio de Janeiro: Elsevier, 2006.

BIZER, C.; CYGANIAK, R.; HEATH, T. **How to publish linked data on the web**. [S.l. : s.n.], 2007. Disponível em: <<http://www4.wiwiss.fu-berlin.de/bizer/pub/LinkedDataTutorial>>. Acesso em: maio de 2012.

BRÄSCHER, M. **Tesauros, Taxonomias e Ontologias**. IV Encontro de Arquivos e Bibliotecas da Administração Pública Federal. Brasília: 2008.

BRASIL. **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: <[http://www.planalto.gov.br/ccivil\\_03/Leis/LCP/Lcp101.htm](http://www.planalto.gov.br/ccivil_03/Leis/LCP/Lcp101.htm)>. Acesso em: 8 maio. 2013.

BRASIL. **Lei Complementar nº 131, de 27 de Maio de 2009**. Acrescenta dispositivos à Lei Complementar nº 101, de 4 de maio de 2000, que estabelece normas de finanças públicas voltadas para a responsabilidade na gestão fiscal e dá outras providências, a fim de determinar a disponibilização, em tempo real, de informações pormenorizadas sobre a execução orçamentária e financeira da União, dos Estados, do Distrito Federal e dos Municípios. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/Leis/LCP/Lcp131.htm](http://www.planalto.gov.br/ccivil_03/Leis/LCP/Lcp131.htm)>. Acesso em: 8 maio. 2013.

BRASIL. **Lei nº 12.527, de 18 de Novembro de 2011**. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências.

BRASIL. MPOG. **Orçamento Federal em Formato Aberto: manual de referência**. 2012. Disponível em: <[https://www.siof.planejamento.gov.br/rdf/manual\\_referencia\\_2012.pdf](https://www.siof.planejamento.gov.br/rdf/manual_referencia_2012.pdf)>. Acesso em: jan. 2013.

BRASIL. MPOG. O Sistema De Gestão Das Transferências Voluntárias Da União – Siconv Como Ferramenta Para Captação De Recursos Federais Pelos Municípios. 2013. Disponível em: <[https://www.convenios.gov.br/portal/avisos/CARTILHA\\_SICONV\\_PARA\\_MUNICIPIOS\\_-\\_Jan\\_13-1.pdf](https://www.convenios.gov.br/portal/avisos/CARTILHA_SICONV_PARA_MUNICIPIOS_-_Jan_13-1.pdf)>. Acesso em: 2 abr. 2013.

BRASIL. SENADO. Portal de Trabalho Siga Brasil. Brasília: Senado, 2007. Disponível em: <[http://www9.senado.gov.br/dav\\_portal/portal/orcamento\\_senado/arquivo/Outros/SigaBrasil/390-Introducao\\_ao\\_SIGA\\_BRASIL.pdf](http://www9.senado.gov.br/dav_portal/portal/orcamento_senado/arquivo/Outros/SigaBrasil/390-Introducao_ao_SIGA_BRASIL.pdf)>. Acesso em: 27 mar. 2013.

BRASIL. SENADO. Sistema Siga Brasil. Disponível em: <[http://www9.senado.gov.br/portal/page/portal/orcamento\\_senado](http://www9.senado.gov.br/portal/page/portal/orcamento_senado)>. Acesso em: 27 mar. 2013.

CARLOS, F. A. et al. Uma discussão sobre a Criação de Indicadores de Transparência na Gestão Pública como Suporte ao Ciclo da Política Pública. **Revista de Contabilidade do Mestrado em Ciências Contábeis da UERJ**, Rio de Janeiro, v.13, n.2, maio./ago., 2008.

CHEVITARESE ALVES, M. V. **Transparência orçamentária efetiva e Internet**: um estudo a partir dos portais eletrônicos de orçamento da Câmara dos Deputados e do Senado Federal. 2009. Monografia (especialização) - Curso em Instituições e Processos Políticos do Legislativo, Câmara dos Deputados, Centro de Formação, Treinamento e Aperfeiçoamento (Cefor), 2009. Disponível em: <<http://bd.camara.gov.br/bd/handle/bdcamara/5644>>. Acesso em: 27 mar. 2013.

DING, L. et al. **Data-gov Wiki**: Towards Linking Government Data. [S.l.]: Association for the Advancement of Artificial Intelligence, 2010. Disponível em: <<http://data-gov.tw.rpi.edu/2010/linkedai-2010-datagov.pdf>>. Acesso em: maio 2012.

FREITAS, F. L. G. Ontologias e a WS. **Jornada de Minicursos em Inteligência Artificial**, SBC, v. 8, 2003.

GROSOFF, I. H.; VOLZ, R.; DECKER, S. **Description Logic Programs: Combining Logic Programs with Description Logic**. Working Paper 4437-03, February 2003. [S.l.]: MIT Sloan School of Management, 2003. Disponível em: <[http://papers.ssrn.com/sol3/Delivery.cfm/SSRN\\_ID460986\\_code285952.pdf?abstractid=460986&mirid=1](http://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID460986_code285952.pdf?abstractid=460986&mirid=1)>. Acesso em: 12 mar. 2013.

HEATH, T.; BIZER, C. Linked Data: Evolving the Web into a Global Data Space. In:

**Synthesis Lectures on the Semantic Web: Theory and Technology**, 1:1, 1-136. [S.l.]: Morgan & Claypool: 2011. [Versão online]. Disponível em: <<http://linkeddatatool.com/editions/1.0/#htoc6>>. Acesso em: 27 mar. 2013.

HEBELER, J. et al. **Semantic Web Programming**. Indianapolis: Wiley, 2009.

KLISCHEWSKI, R.; TAGAMOA, A.; KHAMES, A. Migrating small governments' websites to the semantic web. In: The Semantic Web meets eGovernment, AAAI 2006, Stanford University, California. **Proceedings...** Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.106.8669&rep=rep1&type=pdf>>. Acesso em: maio 2012.

MORAIS, E.F. ; SOARES, M. B. **WS para Máquinas de Busca**. [S.l.]: UFMG, 2004. Disponível em: <<http://www.dcc.ufmg.br/~nivio/cursos/pa03/seminarios/seminario7/seminario7.pdf>>. Acesso em: fev. 2013.

MORTARI, C. A. **Introdução à lógica**. São Paulo: Unesp, 2001.

SOWA, J. F. (2011a). Controlled Natural Languages For Representing Ontology. In: 4º Seminário de Pesquisa em Ontologias no Brasil (Ontobras), Porto Alegre, UFRGS, 2011. **Anais...** Porto Alegre: UFRGS, 2011. Disponível em: <<http://www.inf.ufrgs.br/ontobras-most2011/arquivos/curso1/OntobrasMostSlidesMinicurso1Sowa.pdf>>. Acesso em: 10 fev. 2013.

VIEIRA, R. et al. WS: ontologias, lógica de descrição e inferências. In: TEIXEIRA, C.; BARRERE, E.; ABRAÃO, I. (Org.). **Web e Multimídia: Desafios e Soluções (WebMedia 2005 - Minicursos)**. Porto Alegre: SBC, 2005, v. 1, p. 127-167. Disponível em: <<http://www.inf.pucrs.br/~rvieira/cursos/webmidia.pdf>>. Acesso em: 10 fev. 2013.

W3C. **RDF Primer**. (2004a). W3C Recommendation 10 February 2004. Disponível em: <<http://www.w3.org/TR/rdf-primer/>>. Acesso 9 fev. 2013.

\_\_\_\_\_. **RDF Semantics**. (2004b). W3C Recommendation 10 February 2004. Disponível em: <<http://www.w3.org/TR/rdf-mt/>>. Acesso em: 9 fev. 2013.

\_\_\_\_\_. **RDF Vocabulary Description Language 1.0: RDF Schema**. (2004c). W3C Recommendation 10 February 2004. Disponível em: <<http://www.w3.org/TR/rdf-schema/>>. Acesso em: 10 fev. 2013.

\_\_\_\_\_. **OWL 2 Web Ontology Language Document Overview (Second Edition)**. (2012a). Disponível em: <<http://www.w3.org/TR/2012/REC-owl2-overview-20121211/>>. Acesso em: 10 fev. 2013.

\_\_\_\_\_. **OWL 2 Web Ontology Language Direct Semantics (Second Edition)**. (2012b). Disponível em: <<http://www.w3.org/TR/2012/REC-owl2->

[direct-semantics-20121211/#Class\\_Expression\\_Axioms](http://www.w3.org/TR/2012/REC-owl2-direct-semantics-20121211/#Class_Expression_Axioms)>. Acesso em: 10 fev. 2013.

\_\_\_\_\_. **OWL 2 Web Ontology Language Mapping to RDF Graphs (Second Edition)**. (2012c). Disponível em: <[http://www.w3.org/TR/2012/REC-owl2-direct-semantics-20121211/#Class\\_Expression\\_Axioms](http://www.w3.org/TR/2012/REC-owl2-direct-semantics-20121211/#Class_Expression_Axioms)>. Acesso em: 10 fev. 2013.

W3C ESCRITÓRIO BRASIL. O governo de inovação na Copa 2014: uso de redes sociais e dados abertos. In: Seminário de Inovação em Governo Eletrônico. 2., 2010, Porto Alegre. **Anais...** Porto Alegre: [s.n.], 2010. Disponível em: <[http://www.procergs.rs.gov.br/uploads/1285856001W3C\\_Seminario\\_Inovacao\\_eGov\\_POA\\_17092010.pdf](http://www.procergs.rs.gov.br/uploads/1285856001W3C_Seminario_Inovacao_eGov_POA_17092010.pdf)>. Acesso em: maio de 2012.