

Uso de modelos de dados multidimensionais para ampliação da transparência ativa

Ricardo Ceasar Gonçalves Sant'Ana*

Fernando de Assis Rodrigues**

Resumo Ambientes para acesso a dados governamentais, via Tecnologias de Informação e Comunicação, podem ampliar possibilidades de acompanhamento pelo cidadão, retroalimentando futuras demandas. O objetivo deste estudo é identificar nos dados disponíveis via transparência ativa, a existência de elementos que permitam a elaboração de propostas de modelos dimensionais, propiciando a antecipação de demandas de acesso a dados. Como referencial teórico-metodológico, o texto utiliza os conceitos Business Intelligence e Citizen Intelligence. Como resultado, foi elaborada a proposta de um modelo dimensional a partir da consulta de despesas diárias, disponível no Portal de Transparência do Governo Federal.

Palavras-chave Transparência Pública, Tecnologias de Informação e Comunicação, Coleta de Dados, Citizen Intelligence, Data Warehouse.

Use of multidimensional data models to increase active public transparency

Abstract Environments for access to government data, via Information and Communications Technologies, may expand possibilities for citizen monitoring, providing feedback for future demands. The aim of this study is to identify, in the available data via active transparency, the existence of elements that allow the construction of new proposals of dimensional models, enabling an anticipation of demands on data access. The theoretical-methodological framework, the text uses the concepts Citizen Intelligence and Business Intelligence. As a result, a dimensional model was proposed, building on a dimensional model from a daily expenses query, available in the Transparency home-page of the Brazilian Federal Government.

Keywords Public Transparency, Information and Communication Technologies, Collecting Data, Citizen Intelligence, Data Warehouse.

* Doutor em Ciência da Informação. Universidade Estadual Paulista - UNESP - Campus Experimental de Tupã. Endereço: Rua Domingos da Costa Lopes, 780 - Jd. Itaipu - 17602-496 – Tupã - SP. Telefone: (14) 3404-4200. E-mail: ricardosantana@tupa.unesp.br

** Mestre em Ciência da Informação. Universidade Estadual Paulista - UNESP - Faculdade de Filosofia e Ciências - Campus de Marília. Endereço: Avenida Hygino Muzzi Filho, 737 - Mirante - 17.525-000 – Marília – SP. Telefone: (14) 3402-1300. E-mail: fernando@elleth.org

Introdução

A transparência das ações governamentais perante a sociedade, mediante acesso a dados e informações do Estado, é parte integrante das discussões sobre tendências de modernização dos modelos de administração pública, principalmente nas democracias representativas – onde cidadãos elegem representantes diretamente ou indiretamente para a composição dos poderes executivo e legislativo.

Esse novo modelo de administração pública

[...] busca redistribuir competências e recursos de coordenação entre diferentes níveis institucionais e organizacionais, governamentais e não-governamentais, permitindo o pluralismo institucional nas funções públicas, ao contrário do antigo modelo de monopólio estatal. (MALIN, 2006, p. 1)

Conforme o modelo citado, a transparência das atividades do Estado perante a sociedade tem como uma de suas premissas fortalecer a participação dos cidadãos na administração pública. Esse fortalecimento pode ser garantido com a construção de ambientes democráticos que, entre outras características, criem possibilidades de novos fluxos informacionais entre Estado e sociedade, garantindo assim uma maior visibilidade sobre a gestão da coisa pública. Nas democracias representativas, ampliar esses mecanismos de controle da sociedade civil sobre a administração pública significa ir além do voto, ou seja, trata-se do comprometimento em criar condições para o acompanhamento social na administração pública além dos processos eleitorais (BOHMAN, 1996).

Ambientes para acesso a dados governamentais, por meio das Tecnologias de Informação e Comunicação (TIC), podem ampliar possibilidades de ação do cidadão, retroalimentando a demanda do usuário por informações governamentais e também criando uma motivação extra por transparência nos atores envolvidos com os setores estatais (SANT'ANA, 2008; SANT'ANA, 2009).

Os recursos informacionais mantidos em ambientes informacionais digitais pelo poder público podem ser analisados sob dois focos de utilização: a) de uso interno, que dão respaldo aos processos internos de administração e gestão da coisa pública e que são utilizados por funcionários vinculados ao poder público, e b) de uso externo, resultado da aplicação de TIC em ambientes para interação entre a administração pública e entidades externas, como empresas, cidadãos (ou grupos de cidadãos) ou mesmo outras organizações públicas (SANT'ANA, 2009).

As vantagens na adoção de recursos tecnológicos como ferramenta de respaldo para a gestão pública são evidentes e embasadas em vários estudos de áreas como a Administração, a Ciência da Informação e a Ciência da Computação. Todavia, órgãos governamentais ainda não exploram inúmeras possibilidades oferecidas no uso de Tecnologias de Informação e Comunicação de uso externo. Uma parcela dos sítios oficiais do Governo Federal utiliza recursos tecnológicos apenas

para deixar disponíveis informações tais como quadros de avisos, horários de funcionamento, páginas institucionais, legislação, portarias, diretrizes, listas com contatos telefônicos e eletrônicos (RODRIGUES, 2012).

Assim, explorar o potencial do uso externo das TIC para intensificar a relação entre Estado e sociedade torna-se elemento-chave para uma interação mais direta e eficiente – proporcionando o aumento da transparência do Estado perante a sociedade, bem como a repercussão de suas ações, o que fortalece a democracia e fomenta a mobilização de grupos na defesa de direitos e dos interesses coletivos e, também, para que o cidadão tenha condições de estabelecer uma reflexão crítica, minimizando as assimetrias informacionais nas relações sociedade-governo (SANT'ANA, 2008; SANT'ANA, 2009).

A Constituição Federal de 1988, artigo cinco, inciso XXXIII, destaca que:

[...] todos têm direito a receber dos órgãos públicos informações de seu interesse particular, ou de interesse coletivo ou geral, que serão prestadas no prazo da lei, sob pena de responsabilidade, ressalvadas aquelas cujo sigilo seja imprescindível à segurança da sociedade e do Estado. (BRASIL, 1988, p. 1)

Consta também da Constituição Federal, artigo 216, segundo parágrafo, que a administração pública deve criar mecanismos para consulta a quem necessitar de seus documentos.

Cabe à administração pública, na forma da lei, a gestão da documentação governamental e as providências para franquear sua consulta a quantos dela necessitem. (BRASIL, 1988, p. 1)

Em 18 de novembro de 2011 foi sancionada a Lei Federal 12.527/2011, que revogou a Lei 11.111/2005 a partir do dia 16 de maio de 2012. Esta Lei, também conhecida como Lei de Acesso à Informação (LAI), adiciona novas obrigações às instituições públicas (BRASIL, 2005; BRASIL, 2011).

Uma das questões importantes que a Lei de Acesso à Informação contextualiza é a do sigilo das informações governamentais: o acesso a qualquer informação governamental é garantido; e o sigilo torna-se uma exceção – em antítese às abordagens anteriores. (BRASIL, 2011)

Os órgãos governamentais, independentemente de a qual esfera sejam vinculados, devem adotar dois modelos para a realização da transparência de seus dados. O primeiro é o da forma ativa – os órgãos governamentais devem tornar disponíveis conjuntos de dados para livre acesso à população; e o segundo é a forma passiva – em que órgãos governamentais devem oferecer mecanismos para receber requisições dos cidadãos para acesso a conjuntos de dados ainda não disponíveis. Em ambos os modelos, a disponibilidade não é realizada de forma concentrada: ou seja, cada órgão deve elaborar, em seus sites oficiais, mecanismos para o atendimento dos

aspectos legais ligados à transparência ativa e passiva, seguindo diretrizes encontradas na legislação brasileira.

Em adição, são importantes tanto os aspectos legais ligados à promoção do acesso aos dados governamentais, como também a criação de elementos que promovam o uso intensivo das TIC no desenvolvimento de aplicações que auxiliem o processo de visualização destes dados – pois a visualização de dados é um dos componentes que impulsionam a transparência pública (FIERRO; GIL-GARCIA, 2012).

Os estudos vinculados à visualização de dados são multidisciplinares. São originários de diversas áreas do conhecimento, tais como o *Design*, a Ciência da Computação e a Ciência da Informação, entre outras.

A visualização de dados é definida como uma “ferramenta ou método para interpretar os dados imagéticos alimentados por um computador e para a geração de imagens a partir de complexos conjuntos de dados multidimensionais” (MCCORMICK et. al. apud OWEN, 1993, p. 340, tradução nossa).

Senay e Ignatius (1999) colocam que uma das questões mais importantes na elaboração de visualizações de dados é o mapeamento de atributos a partir dos dados selecionados para elaboração *a posteriori* de um gráfico o mais simplificado possível, porém garantindo efetivamente a transmissão do conteúdo informativo dos dados. Entretanto, um mesmo conjunto de dados apresenta diversas possibilidades de mapeamento de seus atributos, levando a diferentes modelos e técnicas para a construção da visualização. Ou seja, um mesmo conjunto de dados pode apresentar múltiplas visualizações, dependendo de quais atributos sejam selecionados para as elaborações.

Esta multiplicidade dos estudos de visualização de dados é devida a sua própria essência. Por exemplo: ante a concepção e geração de um gráfico, um resumo, um quadro ou uma tabela a partir de conjuntos de dados governamentais, são necessários inúmeros passos, como a coleta dos dados no recurso informacional, o armazenamento dos dados (em uma base de dados), o tratamento e processamento dos dados, a escolha dos elementos gráficos, cores, outros fatores. Além de fatores estéticos e computacionais, também é importante o público alvo a que se destina a construção visual dos dados governamentais.

Na legislação brasileira já é prevista uma forma de simplificação do processo de visualização dos dados:

É dever do Estado garantir o direito de acesso à informação, que será franqueada, mediante procedimentos objetivos e ágeis, de forma transparente, clara e em linguagem de fácil compreensão. (BRASIL, 2011, p. 1, grifo nosso)

A “Lei de Acesso à Informação” também torna obrigatório o uso da infraestrutura da internet como instrumento de disseminação e acesso aos dados governamentais.

[...] os órgãos e entidades públicas deverão utilizar todos os meios e instrumentos legítimos de que dispuserem, sendo obrigatória a divulgação em sítios oficiais da rede mundial de computadores (internet). (BRASIL, 2011, p. 1, grifo nosso)

Em novembro de 2004, a Controladoria-Geral da União (2004) iniciou as atividades do sítio denominado Portal da Transparência do Governo Federal, no intuito de assegurar o acesso a dados sobre aplicações dos recursos financeiros públicos, com o objetivo de aumentar a transparência da gestão pública perante o cidadão, para que este possa acompanhar o destino dos recursos financeiros públicos e auxiliar no processo de fiscalização.

Contudo, sítios do Estado que disponibilizam, pela transparência ativa, conjuntos de dados governamentais são elaborados na tentativa de atender à necessidade informacional de qualquer interessado, em um antítese aos moldes de elaboração de aplicativos ou *web sites* desenvolvidos para atender públicos específicos, como um aplicativo para auxílio a gestão de uma indústria ou um aplicativo para gestão de pontos de vendas em varejo. Este é um fator que torna complexa a construção de visualizações de dados a partir da coleta de conjuntos de dados governamentais, pois nem sempre as tabelas ou os gráficos presentes nos sítios do Estado irão atender todas as demandas informacionais da sociedade.

Em suma, é importante salientar a necessidade da existência de gráficos, sínteses, quadros e outros artefatos de visualização nos sítios do Estado, mas estes também devem contemplar a possibilidade de coleta dos dados primários – o mais granular, com o maior nível de detalhamento – para criarem oportunidades de construções de novas percepções e visualizações sobre os dados governamentais, a partir da elaboração de uma estrutura prévia que permita que visualizações também sejam criadas fora dos sítios do Estado - principalmente devido à pluralidade do público interessado nesse tipo de informação.

Considera-se o acesso a dados primários como um dos fatores-chave de sucesso na construção de alternativas para acesso à dados evitando-se eventuais falhas de interpretação e ainda permitindo variados caminhos de análise em função do alto grau de liberdade de combinações e possíveis processamentos que um dado primário oferece. (SANT'ANA; RODRIGUES, 2013, p. 50)

Portanto, o objetivo deste estudo é identificar nos dados governamentais disponíveis via transparência ativa, a existência de elementos que permitam a elaboração de propostas de modelos dimensionais a partir do conceito de *Data Warehouse*, propiciando a antecipação de demanda de acesso a dados.

Como referencial teórico-metodológico, o texto utiliza os conceitos *Business Intelligence*, *Data Warehouse*, *Data Mart* e *Citizen Intelligence*. Como resultado, foi elaborada a proposta de um

modelo dimensional a partir da consulta ‘Detalhamento Diário das Despesas’, disponível no Portal de Transparência do Governo Federal.

Business intelligence

Nos ambientes empresariais, gestores podem contar com o apoio de processos de transformação de conjuntos de dados em informações para dar suporte a suas decisões. O conjunto de teorias, metodologias e tecnologias para auxiliar o processo de tomada de decisão é o conceito denominado *Business Intelligence* (BI), que está diretamente relacionado à própria tomada de decisão; ao uso de Sistemas Gerenciadores de Banco de Dados (SGBD) que permitam a construção de banco de dados compatíveis com um *Data Warehouse* (DW) para geração de relatórios e análises desse dados; e a processos de descobrimento de padrões nos conjuntos de dados armazenados por programas de computadores, como o *data mining* (BARBIERI, 2001).

O *Data Warehouse* é definido como uma base de dados orientada a assuntos, integrada, variável no tempo, não volátil no auxílio ao processo de tomada de decisão gerencial (INMON, 1992). Diferente da base de dados transacionais, utilizada e estruturada para uso e apoio às atividades diárias de uma empresa ou órgão público, o *Data Warehouse* é estruturado com foco específico para a realização de consultas e análises orientadas à uma demanda específica (KIMBALL, 1996).

Não apenas instituições privadas podem fazer uso de *Data Warehouse* para a melhoria do processo de tomada de decisão: órgãos, setores, entidades governamentais e agências de fomento, entre outras áreas, podem adotar o *Data Warehouse* para apoio à decisão. Por exemplo, Mussi (2005) conclui que, ao adotar um *Data Warehouse*, a Agência Nacional de Vigilância Sanitária (ANVISA) cria a possibilidade de buscar e interpretar

[...] informações armazenadas, garantindo maior acuricidade nas tomadas de decisão; permite estabelecer uma aproximação integrada e colaborativa para capturar, criar, organizar e usar todos os ativos de informação de uma empresa; consegue antecipar as mudanças bruscas no mercado; conhecer o negócio; ter uma visão clara sobre novos negócios; auxilia na implementação de novas ferramentas gerenciais; ter conhecimento sobre novas tecnologias, produtos ou serviços que possam afetar direta e indiretamente o seu negócio. (MUSSI, 2005, p. 6)

Para a elaboração de uma base de dados em um modelo dimensional, proposto pelas teorias do *Data Warehouse*, é importante determinar e identificar dois termos: fato e dimensão. O *Data Warehouse* é formado por diversas tabelas representando fatos e dimensões. Cada conjunto interligado de fatos e dimensões é denominado *Data Mart*.

O *Data Mart* é um conjunto de dados flexível, baseado em dados em seu estado mais granular extraídos de uma base de dados transacional, apresentados em um modelo dimensional mais adaptável a consultas de usuários do que em uma base de dados transacional. O Data Mart representa dados de um único processo de negócios de uma instituição, como, por exemplo, uma venda (KIMBALL; ROSS, 2002).

Por exemplo, uma base de dados transacional de uma empresa no ramo varejista possui uma tabela para armazenar dados sobre vendas (Figura 1). Essa tabela contém diversos atributos, e em cada um desses atributos é alocado um tipo de dado, como o identificador numérico único para cada venda (conhecido como chave primária), a data em que foi realizada, o valor total, o nome do funcionário que realizou a venda e o nome do departamento.

Figura 1: Exemplo de tabela "Venda" em uma base de dados transacional

	Venda
	Identificador único (chave) Data Valor Total Funcionário Departamento

Fonte: Autores

Ao iniciar a elaboração de um modelo dimensional a partir da base de dados transacional desta empresa hipotética, é necessário identificar os atributos que sejam fatos. Uma tabela de fatos (ou tabela fato) é a tabela primária, ou seja, a tabela central, ponto de partida e pré-requisito a ser elaborada em um modelo dimensional, onde devem ser armazenados os atributos quantificáveis. Por exemplo, a elaboração da tabela de fato "Venda" contém os atributos: valor total de venda; e os identificadores únicos numéricos (chaves estrangeiras) que representarão cada funcionário, cada departamento e cada data de venda (Figura 2). Estes identificadores únicos são necessários para interligar cada registro de dado contido na tabela de fato "Venda" com as tabelas contendo informações dos departamentos, das datas de vendas e dos funcionários (KIMBALL; ROSS, 2002).

Figura 2: Exemplo de tabela de fato Venda em um modelo dimensional

	Venda (fato)
	Identificador Único da Data Identificador Único do Funcionário Identificador Único do Departamento Valor Total

Fonte: Autores

Ou seja, ao invés de armazenar na tabela "Fato" atributos descritivos com o nome do funcionário, o nome do departamento e a descrição da data da realização da venda, cria-se um vínculo através de identificadores únicos representando estes atributos, que estarão presentes nas tabelas de dimensões.

Assim como a tabela de "fato", as tabelas de dimensões também são componentes essenciais para o funcionamento de um *Data Mart*. As tabelas de dimensões contêm as descrições textuais para cada identificador único (chaves estrangeiras) encontrado na tabela fato. Cada uma dessas chaves estrangeiras deve fazer relação com uma tabela de dimensão.

Por exemplo, o identificador único do funcionário deve estar diretamente ligado a uma tabela de dimensão chamada "Funcionário", e essa tabela de dimensão deve conter não só o elemento descritivo que represente seu nome, mas também um contexto muito mais amplo: descrições como seu gênero, sua idade, o turno em que exerce sua função, o Registro de Identificação (RG), o Cadastro de Pessoa Física (CPF), entre outros (Figura 3).

Figura 3: Exemplo de tabela de dimensão "Funcionário" em um modelo dimensional

	Funcionário (dimensão)
	Identificador Único do Funcionário Nome Gênero Idade Turno RG CPF ...

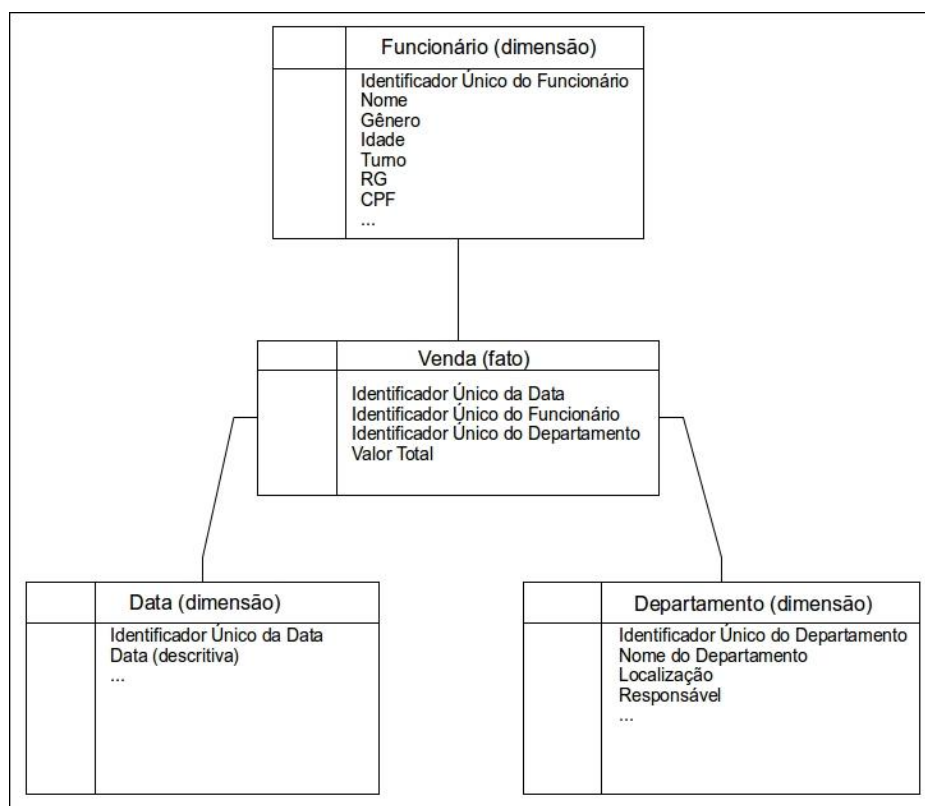
Fonte: Autores

As tabelas de dimensões são as interfaces de entrada para as consultas aos dados em um *Data Mart* e os atributos dessas tabelas servem como recurso primário nas consultas a serem realizadas (KIMBALL; ROSS, 2002).

Por fim, a junção do relacionamento de todas as tabelas que representam as dimensões com a tabela que representa os fatos gera o modelo dimensional (Figura 4). Esse modelo de

representação, com a tabela de fato ao centro e as tabelas de dimensões no entorno é conhecido como esquema estrela (*star schema*), adotado nesse estudo.

Figura 4: Exemplo de um modelo dimensional em esquema estrela



Fonte: Autores

Com o modelo dimensional desenvolvido, a realização de consultas para atender objetivos específicos à tomada de decisão é simplificada, pois o processo inicia-se a partir dos atributos encontrados em cada dimensão. Por exemplo, para recuperar dados de qual o valor vendido apenas pelos funcionários que trabalham em um determinado turno, é necessário iniciar a consulta da tabela de fato "venda" a partir da dimensão funcionário - evitando o uso de dados desnecessários das dimensões data e departamento para atender a resposta. Em outra percepção, para atender a consulta de qual o total de vendas em um determinado dia, só há a necessidade

utilizar a tabela de fato venda a partir da dimensão data. É importante ressaltar que o sucesso dos resultados das consultas em um modelo dimensional está ligado diretamente à formulação das perguntas corretas.

Um dos usos do modelo dimensional em instituições ocorre na tentativa de antecipar a demanda a partir de informações armazenadas nas bases de dados das empresas. Com o modelo dimensional é possível aplicar técnicas para reconhecer padrões ou necessidades de uma determinada ação em curto ou longo prazo. Com isso, os gestores podem prever certos acontecimentos e preparar ações para aproveitar, da melhor maneira, futuras oportunidades ou evitar crises.

Este processo de desenvolvimento de um modelo dimensional pode ser aplicado em contextos como nas consultas a dados governamentais elaboradas para atender à demanda por transparência ativa.

Citizen intelligence

Com novas possibilidades oferecidas em ambientes colaborativos voltados ao fortalecimento da relação Sociedade-Estado, surge o conceito *Citizen Intelligence* (CI). O *Citizen Intelligence* é o conjunto de teorias, fontes e métodos que determinam elementos necessários na construção de mecanismos para suporte a tomada de decisão dos cidadãos a partir de dados governamentais. Entretanto, diferente do *Business Intelligence* que é utilizado no apoio a decisões internas de uma instituição, o *Citizen Intelligence* é completamente transparente não só aos atores envolvidos, mas a toda a sociedade; e os resultados são focados no interesse coletivo e público e não nos interesses de um órgão governamental ou em um grupo gestor de uma indústria ou empresa. Por ser transparente, o processo e o resultado do *Citizen Intelligence* pode ser compartilhado sem restrições, pois não há sigilo nos métodos para a elaboração do processo e nos resultados obtidos (P2P FOUNDATION, 2011).

Quando um determinado grupo ou cidadãos têm interesse em um conjunto de dados governamentais, esse interesse tem origem em uma reflexão – em uma questão. Por exemplo, um grupo pode se interessar em quantificar recursos financeiros gastos em um município em saúde pública ou qual o valor diário repassado a projetos diariamente pelo Governo Federal. Este tipo de problemática é concomitante às questões que um modelo dimensional de *Data Warehouse* tenta responder, porém em outro cenário.

Portanto, a coleta de dados governamentais não é realizada por si só (apenas coletar por coletar): a coleta de dados governamentais pressupõe responder uma questão originária de cidadãos ou grupos.

A elaboração de um modelo dimensional, em esquema estrela, para realizar consultas nos conjuntos de dados governamentais pode auxiliar nesse processo de coleta de dados, com o intuito de responder futuras questões de cidadãos ou grupos, pois a própria ideia da transparência ativa é uma antecipação de demanda: deixar disponível nos sítios do Estado conjuntos de dados governamentais antes que cidadãos precisem solicitar o acesso via transparência passiva.

Nesse cenário, o grande diferencial do uso de *Business Intelligence* e *Data Warehouse* é a necessidade de se antecipar futuras demandas dos gestores, e, de forma correlata, o *Citizen Intelligence* também requer que se tenha um mapeamento das dimensões para que seja possível uma estrutura capaz de atender a estas futuras demandas de acesso a dados.

Modelo dimensional para transparência

O Portal da Transparência do Governo Federal possui um recurso informacional denominado Despesas, onde é possível o acompanhamento diário de informações relativas ao repasse financeiro de verbas federais para as esferas estaduais e municipais, e para o pagamento de serviços e materiais.

Dentro deste recurso informacional, a consulta ‘Detalhamento Diário das Despesas’ apresenta “dados detalhados e diariamente atualizados sobre os atos praticados pelas unidades gestoras do Poder Executivo Federal no decorrer da execução das suas despesas”.

Na forma atual de consulta (Figura 5), é possível refinar os resultados a partir de sete (7) filtros, sendo eles: Período, Fase da Despesa, Órgão Superior, Órgão / Entidade Vinculada, Unidade Gestora, Elemento de Despesa e Favorecido.

Figura 5: Formas de refinamento na consulta Detalhamento Diário das Despesas

Detalhamento Diário das Despesas

Consulta Avançada [Consulta Rápida](#) | [Consulta por Documento](#)

Período: 01/07/2013 a 14/07/2013 Formato: dd/mm/aaaa

Fase da Despesa: Empenho Liquidação Pagamento

Órgão Superior: Todos (período de 1 dia ou favorecido específico) ▼

Órgão / Entidade Vinculada: (Escolha o Órgão Superior) ▼

Unidade Gestora: (Escolha o Órgão / Entidade Vinculada) ▼

Elemento de Despesa: Todos ▼

Favorecido: Fornecer CNPJ, CPF, UG-Gestão ou outros (sem pontuações)

Fonte: Controladoria-Geral da União

Uma característica importante no refinamento são as restrições de períodos de datas a serem consultados: os resultados da consulta só retornam dados por períodos de trinta e um (31) dias se selecionado um órgão superior, ou retornam dados de apenas um (1) dia se não selecionado um órgão superior. Essa restrição é necessária no modelo atual de consulta pois, como é um ponto de

acesso aos dados contidos em uma base de dados transacional, a quantidade de registros de despesas do Governo Federal gerados diariamente é grande.

Em adição, só é possível consultar dados das despesas a partir de "Favorecidos" se o cidadão souber de antemão o número do Cadastro Nacional de Pessoa Jurídica (CNPJ) ou do Cadastro de Pessoa Física (CPF), ou do Cadastro de Unidade Gestora (UG-Gestão). Este último é necessário quando o destino da despesa for um órgão governamental, como uma prefeitura ou uma câmara de vereadores.

Portanto, se um cidadão deseja identificar despesas de um determinado órgão em um período anual, terá de realizar doze (12) consultas para o atendimento do seu objetivo. Caso o objetivo desse cidadão for o de encontrar o valor total gasto no período anual por este órgão, ainda terá que calcular o total fora do sítio oficial, pois o resultado da consulta é uma lista ordenada por data de forma crescente e dividida em páginas com quinze (15) registros cada, contendo os atributos: Data, Fase, Documento, Espécie, Órgão Superior, Órgão/Entidade Vinculada, Unidade Gestora, Elemento de Despesa, Favorecido e Valor (R\$) (Figura 6).

Figura 6: Resultados da consulta Detalhamento Diário das Despesas em forma de lista

Resultado da Consulta									
Página 1 de 5876									
1 2 3 4 5 > >> Página: nº página Ir									
Data	Fase	Documento	Espécie	Órgão Superior	Órgão / Entidade Vinculada	Unidade Gestora	Elemento de Despesa	Favorecido	Valor (R\$)
11/07/2013	Pagamento	2013NL000001	Original	MINISTERIO DA DEFESA	COMANDO DA MARINHA	COORD-GERAL PROG. DESENV. SUBMAR. NUCLEAR	MATERIAL DE CONSUMO	DCN CHERBOURG	345,662,590.15(*)
11/07/2013	Pagamento	2013OB800237	Original	MINISTERIO DA DEFESA	COMANDO DA AERONAUTICA	SECRETARIA DE ECON. E FINANÇAS DA AERONAUTICA	PRINCIPAL CORRIGIDO DIVIDA CONTRATUAL RESG.	BBVA-BANCO BILBAO VIZCAYA ARGENTARIA S.	7,405,426.25
11/07/2013	Pagamento	2013NL000003	Original	MINISTERIO DA DEFESA	COMANDO DA MARINHA	COORD-GERAL PROG. DESENV. SUBMAR. NUCLEAR	MATERIAL DE CONSUMO	DCN CHERBOURG	5,487,058.29(*)
11/07/2013	Pagamento	2013OB800063	Original	MINISTERIO DA DEFESA	COMANDO DO EXERCITO	D CONT - SETORIAL FINANCEIRA		BANCO SANTANDER (BRASIL) S.A.	4,096,512.00
11/07/2013	Pagamento	2013OB800238	Original	MINISTERIO DA DEFESA	COMANDO DA AERONAUTICA	SECRETARIA DE ECON. E FINANÇAS DA AERONAUTICA	JUROS SOBRE A DIVIDA POR CONTRATO - LC141/12	BBVA-BANCO BILBAO VIZCAYA ARGENTARIA S.	1,963,941.92
11/07/2013	Pagamento	2013OB800312	Original	MINISTERIO DA DEFESA	COMANDO DA MARINHA	ARM OMP S - I APOIO DE BASE	OUTROS SERVICOS DE TERCEIROS-PESSOA JURIDICA	BANCO DO BRASIL SA (*)	1,445,138.57
11/07/2013	Pagamento	2013OB802333	Original	MINISTERIO DA DEFESA	COMANDO DA AERONAUTICA	GRUPAMENTO DE APOIO LOGISTICO	OUTROS SERVICOS DE TERCEIROS-PESSOA JURIDICA	DIGEX AIRCRAFT MAINTENANCE S.A.	798,805.04
11/07/2013	Pagamento	2013OB800405	Original	MINISTERIO DA DEFESA	COMANDO DA AERONAUTICA	COMISSAO DE IMPLANT.DO TIST.DE CONTR.DO E AER	Múltiplo	OMNISYS ENGENHARIA LTDA	692,563.60
11/07/2013	Pagamento	2013OB801510	Original	MINISTERIO DA DEFESA	COMANDO DA MARINHA	CENTRO TECNOLOGICO DA MARINHA EM SAO PAULO	MATERIAL DE CONSUMO	VILLARES METALS SA	690,590.25
11/07/2013	Pagamento	2013OB801496	Original	MINISTERIO DA DEFESA	COMANDO DA MARINHA	CENTRO TECNOLOGICO DA MARINHA EM SAO PAULO	OUTROS SERVICOS DE TERCEIROS-PESSOA JURIDICA	MOODY INTERNATIONAL BRASIL LTDA.	639,867.05
11/07/2013	Pagamento	2013OB803559	Original	MINISTERIO DA DEFESA	COMANDO DA AERONAUTICA	GRUPAMENTO DE INFRA-EST E APOIO DE S J CAMPOS	OUTROS SERVICOS DE TERCEIROS-PESSOA JURIDICA	BANCO DO BRASIL SA (*)	588,114.70
11/07/2013	Pagamento	2013OB801497	Original	MINISTERIO DA DEFESA	COMANDO DA MARINHA	CENTRO TECNOLOGICO DA MARINHA EM SAO PAULO	OUTROS SERVICOS DE TERCEIROS-PESSOA JURIDICA	MOODY INTERNATIONAL BRASIL LTDA.	565,022.95
11/07/2013	Pagamento	2013NS000635	Original	MINISTERIO DA DEFESA	COMANDO DA AERONAUTICA	COMISSAO AERONAUTICA BRASILEIRA NA EUROPA	MATERIAL DE CONSUMO	THALES RAYTHEON SYSTEMS	537,727.57(*)
11/07/2013	Pagamento	2013OB800403	Original	MINISTERIO DA DEFESA	COMANDO DA AERONAUTICA	COMISSAO DE IMPLANT.DO SIST.DE CONTR.DO E AER	OUTROS SERVICOS DE TERCEIROS-PESSOA JURIDICA	ATECH - NEGOCIOS EM TECNOLOGIAS S.A.	529,595.26
11/07/2013	Pagamento	2013OB800917	Original	MINISTERIO DA DEFESA	COMANDO DA AERONAUTICA	PARQUE DE MATERIAL AERONAUTICO DE SAO PAULO	OUTROS SERVICOS DE TERCEIROS-PESSOA JURIDICA	IAS INCREASE AVIATION SERVICE LTDA	420,950.61

(*) Este favorecido é um intermediário que recebe o recurso e repassa-o para os favorecidos finais.

(*) Valor em moeda estrangeira convertido para real (R\$).

Página 1 de 5876 1 2 3 4 5 > >> Página: nº página Ir

Liinc e

Fonte: Controladoria-Geral da União

Caso exista a necessidade de consultar outras informações referentes a cada despesa, os registros do atributo Documento apresentam um vínculo (*hyperlink*) com informações do contrato. A Figura 7 apresenta o modo em que os dados são encontrados no resultado da consulta.

Figura 7: Detalhamento de documentos vinculados em cada despesa

DADOS BÁSICOS									
Fase:	Empenho								
Documento:	2013NE800749								
Data:	11/07/2013								
Tipo de Empenho:	ESTIMATIVO								
Órgão Superior:	36000 - MINISTERIO DA SAUDE								
Órgão / Entidade Vinculada:	36901 - FUNDO NACIONAL DE SAUDE								
Unidade Gestora Emitente:	250061 - HOSPITAL FEDERAL DOS SERVIDORES DO ESTADO								
Gestão:	00001 - TESOURO NACIONAL								
Favorecido:	00.210.051/0001-48 - BERKELEY EQUIPAMENTOS MEDICOS LTDA								
Valor:	R\$ 1,362,000.00								
DADOS DETALHADOS									
Observação do Documento:	REFORÇO PARA PAGAMENTO REF.3ºTRIM./2013.								
Esfera:	2 - ORÇAMENTO DE SEGURIDADE SOCIAL								
Fonte de Recursos:	51 - CONTR.SOCIAL S/O LUCRO DAS PESSOAS JURIDICAS								
Unidade Orçamentária:	36901 - FUNDO NACIONAL DE SAUDE								
Funcional Programática									
Função:	10 - SAUDE								
Subfunção:	302 - ASSISTENCIA HOSPITALAR E AMBULATORIAL								
Programa:	2015 - APERFEICOAMENTO DO SISTEMA UNICO DE SAUDE (SUS)								
Ação:	6217 - ATENCAO A SAUDE NOS SERVICOS AMBULATORIAIS E HOSPITALARES DO MINISTERIO DA SAUDE								
Subtítulo (localizador):	6506 - ATENCAO A SAUDE NOS SERVICOS AMBULATORIAIS E HOSPITALARES DO MINISTERIO DA SAUDE - NO MUNICIPIO DO RIO DE JANEIRO - RJ (HOSPITAL GERAL DOS SERVIDORES)								
Linguagem Cidadã:									
Categoria de Despesa:	3 - Despesas Correntes								
Modalidade de Aplicação:	90 - Aplic. Diretas (Gastos Diretos do Governo Federal)								
Elemento de Despesa:	39 - OUTROS SERVICOS DE TERCEIROS-PESSOA JURIDICA								
Processo Nº:	33433005712200921								
Modalidade de Licitação:	PREGAO								
Referência da Dispensa ou Inexigibilidade:									
Nº Convênio / Contrato de Repasse / Termo de Parceria / Outros:									
Detalhamento do Gasto									
Subitem da Despesa	Quantidade	Valor Unitário (R\$)	Valor Total (R\$)	Descrição					
12 - LOCAÇÃO DE MAQUINAS E EQUIPAMENTOS	3	454,000.00	1,362,000.00	LOCAÇÃO / MANUTENÇÃO DE EQUIPAMENTOS PERMANENTES 000004405 Contratação de empresa especializada para Locação, com Manutenção Preventiva e Corretiva, de equipamentos para o Centro Cirúrgico Geral, CTI Geral e Unidade Coronariana do Hospital Federal dos Servidores do Estado, em consonância com o disposto no item 6 e subitens 6.1 a 6.10 do Termo de Referência.					
DOCUMENTOS RELACIONADOS									
Data	Fase	Documento	Espécie	Órgão Superior	Órgão / Entidade Vinculada	Unidade Gestora	Elemento de Despesa	Favorecido	Valor (R\$)
25/02/2013	Empenho	2013NE800090	Original	MINISTERIO DA SAUDE	FUNDO NACIONAL DE SAUDE	HOSPITAL FEDERAL DOS SERVIDORES DO ESTADO	Outros Serviços de Terceiros - Pessoa Jurídica	BERKELEY EQUIPAMENTOS MEDICOS LTDA	454,000.00
Recomenda-se visualizar em resolução 1024x768									
Fonte: SIAFI									

Fonte: Controladoria-Geral da União

Na forma atual da consulta aos dados de despesas diárias, o resultado apresenta os atributos organizados e distribuídos em três grupos: Dados Básicos, Dados Detalhados e Documentos Relacionados.

Somente neste momento são apresentados ao cidadão atributos descritivos que identifiquem a esfera vinculada à despesa, a fonte dos recursos financeiros que irá pagar a despesa e número identificador do favorecido (como o CPF, CNPJ ou UG-Gestão).

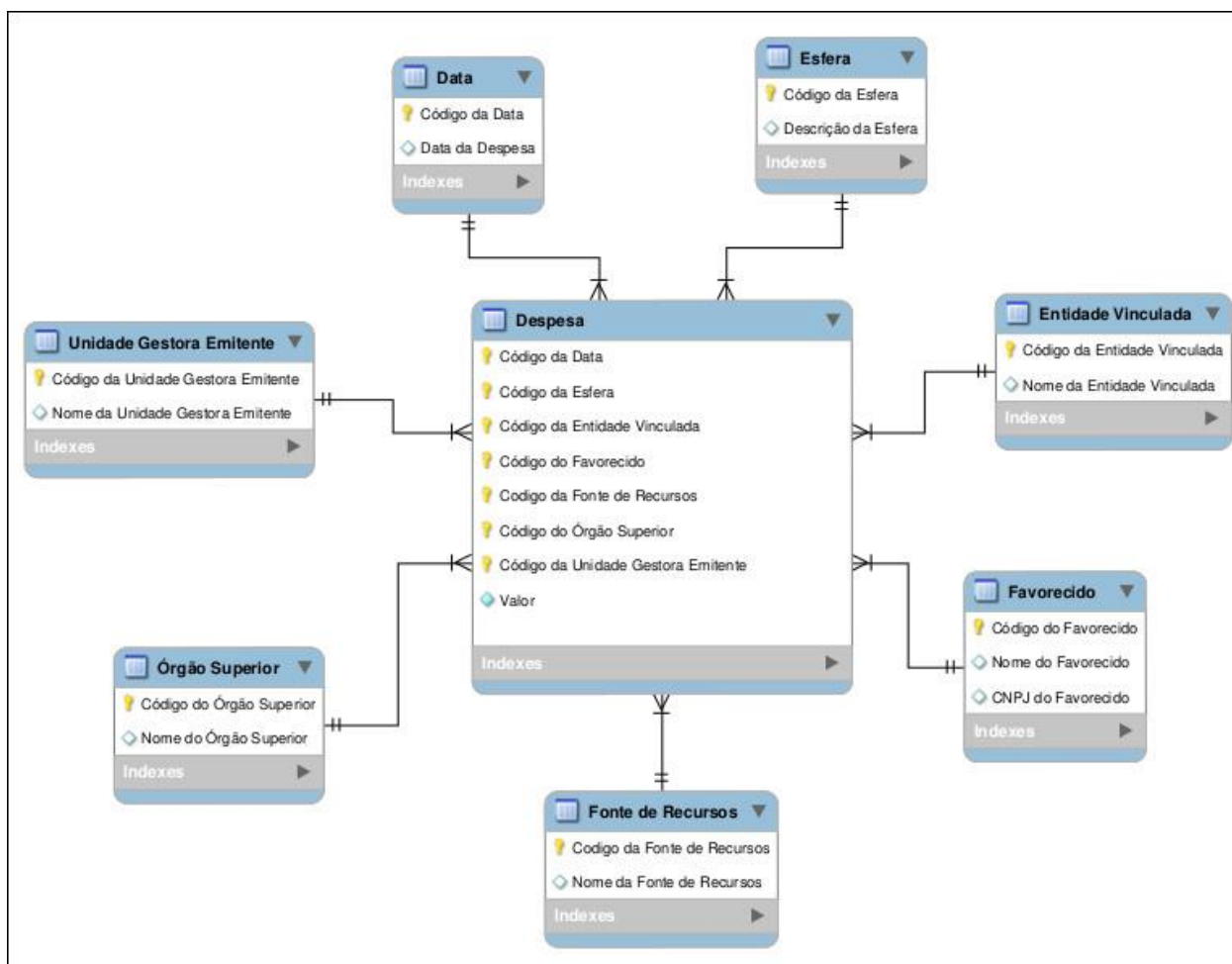
Portanto, caso o cidadão tenha interesse em elaborar um gráfico exibindo valores (em reais) do total de despesas repassadas anualmente para uma empresa privada, porém apenas as que sejam oriundas de uma determinada esfera, o processo de coleta se torna inviável: por exemplo, seria necessário realizar doze coletas para cada ano, uma para cada mês do ano escolhido, com todos os resultados vinculados à instituição privada selecionada; entrar em todos os documentos vinculados aos contratos para identificar a esfera; transformar os resultados coletados em páginas da consulta em registros em uma planilha eletrônica; filtrar apenas a esfera de interesse; e gerar uma somatória com o total anual.

A elaboração de um *Data Mart* pode auxiliar neste tipo de limitação nas consultas diretas aos dados armazenados em um banco de dados transacional, pois o fato "Despesa" pode ser vinculado a dimensões que sejam pontos de entrada.

Os atributos encontrados na listagem resultante da consulta atual, juntamente com os atributos encontrados nos agrupamentos de Dados Básicos e Dados Detalhados podem ser reorganizados em um modelo dimensional como proposto por Kimball e Ross (2002), pois é possível identificar atributos fatos (como valor da despesa) e atributos dimensionais, como a esfera, a fonte dos recursos, a data, o órgão superior, o órgão ou entidade vinculada, a unidade gestora emitente e o favorecido.

Partindo da identificação dos próprios atributos no detalhamento de despesas e na reorganização dos atributos em tabelas de fatos e de dimensões é possível a elaboração de uma proposta de modelo dimensional em esquema estrela (Figura 8) para a consulta 'Detalhamento Diário das Despesas' voltado para a recuperação de dados. A tabela "Fato" está ao centro (despesa) com cada uma das dimensões em seu entorno. Os atributos identificadores únicos estão identificados com uma chave na cor amarela; os atributos descritivos (contendo textos) estão identificados com um losango preenchido com a cor branca; e os atributos numéricos estão identificados com um losango preenchido com a cor azul claro.

Figura 8: Proposta de modelo dimensional no esquema estrela (*star schema*)



Fonte: Autores

Com isso, os atributos de cada uma das tabelas de dimensões podem representar um ponto de partida para diferentes tipos de consultas de um mesmo fato (as despesas), criando uma percepção temática nos elementos de origem e destino das despesas, para um público específico (no caso, cidadãos interessados neste tipo de informação).

Conclusões

A construção de um modelo dimensional por um *Data Warehouse* pode contribuir à construção de novas consultas aos recursos informacionais encontrados no Portal da Transparência do Governo Federal, que está ligado diretamente ao processo de transparência ativa. A elaboração de conjuntos de *Data Mart* são elementos-chave para que agentes públicos antecipem demandas

por dados governamentais, pois podem auxiliar no processo de coleta de dados para satisfazer questionamentos da sociedade.

Por exemplo, através do *Data Mart* proposto, o processo de coleta de dados governamentais que tratam diretamente com os valores das despesas e o caminho percorrido entre a origem e o destino de uma despesa poderia ser simplificado. Em adição, as dimensões podem conter mais atributos descritivos, que não estão presentes na recuperação das informações no modelo transacional proposto na consulta atual. O modelo dimensional poderia trazer mais informações sobre as esferas, os órgãos superiores, favorecidos, entre outros.

Além disso, não existem impedimentos para que um *Data Warehouse* de dados governamentais seja elaborado com conjuntos de *Data Mart* atendendo a outros cenários, tais como *Data Mart* centrados em outros fatos (como convênios, transferências de recursos, entre outros) ou com dimensões elaboradas ao atendimento de outros interesses (como dimensões diferentes para fato despesa). Ambos os cenários são importantes na formação de um *Data Warehouse* e, também, no processo de acesso a dados da gestão pública com mais qualidade e eficiência - e em prol da concepção de *Citizen Intelligence*.

Também é importante salientar que, apesar da proposta de elaboração de um *Data Warehouse* de dados governamentais transparente à sociedade, os sítios governamentais devem continuar disponibilizando acesso aos dados primários e às consultas já existentes. O *Data Warehouse* tem papel complementar a essas consultas, principalmente auxiliando a interpretação das próprias consultas existentes.

Artigo recebido em 15/07/2013 e aprovado em 09/09/2013

Referências

BARBIERI, C. **BI - Business intelligence modelagem & tecnologia**. Rio de Janeiro: Excel Books, 2001. p. 424.

BOHMAN, J. **Public deliberation, pluralism, complexity and democracy**. London: MIT Press, 1996.

BRASIL. **Constituição da República Federativa do Brasil de 1988**. Portal do Planalto. Brasília, DF. Disponível em: <www.planalto.gov.br/ccivil_03/Constituicao/Constituicao.htm>. Acesso em: 20 jan. 2012.

_____. Lei número 11.111, de 05 de maio de 2005. Regulamenta a parte final do disposto no inciso XXXIII do caput do art. 5º da Constituição Federal e dá outras providências. Portal do Planalto. Brasília, DF, 5 maio 2005. Disponível em: <http://www.planalto.gov.br/ccivil_03/ Ato2004-2006/2005/Lei/L11111.htm>. Acesso em: 31 jan. 2012.

_____. Lei número 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. Portal do Planalto. Brasília, DF, 18 nov. 2011. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2011/Lei/L12527.htm>. Acesso em: 31 jan. 2012.

BRASIL. Controladoria Geral da União. **Portal da transparência do Governo Federal**. Brasília, 2006. Disponível em: <<http://www.cgu.gov.br/>>. Acesso em: 20 jun. 2013.

_____. Detalhamento das despesas diárias. In: _____. **Portal da transparência do Governo Federal**. Brasília, 2004. Disponível em: <<http://www.portaltransparencia.gov.br/despesasdiarias/>>. Acesso em: 20 jun. 2013.

FIERRO, A. E.; GIL-GARCIA, J. R. Transparency websites as tools for decision making in a democratic government. In: TRANSATLANTIC CONFERENCE ON TRANSPARENCY RESEARCH, Utrecht, Holanda: Utrecht University, 2012. Disponível em <<http://www.transparencyconference.nl/wp-content/uploads/2012/05/Fierro.docx>>. Acesso em: 6 jul. 2013.

INMON, W.H. **Building the data warehouse**. /s.n./ John Wiley & Sons, 1992. 543 p.

KIMBALL, R. **The data warehouse toolkit**. /s.n./ John Wiley & Sons, 1996. 388 p. ISBN 978-0-471-15337-5.

KIMBALL, R.; ROSS, M. **The data warehouse toolkit: the complete guide to Dimensional Modeling**. /s.n./ John Wiley & Sons, 2002. 393 p. ISBN 0-471-20024-7.

MALIN, A. M. B. Gestão da informação governamental: em direção a uma metodologia de avaliação. **DataGramZero**, v. 7, n. 5, out., 2006. Disponível em: <http://www.dgz.org.br/out06/Art_02.htm>. Acesso em: 29 fev. 2012.

MUSSI, C. Data warehouse - a experiência da Agência Nacional de Vigilância Sanitária do Brasil. In: CONGRESO INTERNACIONAL ONLINE SOBRE TECNOLOGÍAS DE LA INFORMACIÓN EN CIENCIAS DE LA VIDA, 1. (TICvida'2005), 2005, Madrid.

P2P FOUNDATION. **Introduction to Citizen Intelligence Sources and Methods**. P2P Foundation, 11 fev. 2011. Disponível em: <http://p2pfoundation.net/Introduction_to_Citizen_Intelligence_Sources_and_Methods>. Acesso em: 1 jul. 2013.

RODRIGUES F. A. Restrições tecnológicas e de acesso a dados disponíveis sobre destinos de repasses financeiros federais para a saúde pública em ambientes informacionais digitais. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 13. Rio de Janeiro, 2012. ISSN: 978-85-62454-02-8. Disponível em: <<http://www.eventosecongressos.com.br/metodo/enancib2012/arearestrita/pdfs/19435.pdf>>. Acesso em: 1 jul. 2013.

SANT'ANA, R. C. G. **Mensuração da disponibilização de informações e do nível de interação dos ambientes informacionais digitais da administração municipal com a sociedade.** Tese de Doutorado em Ciência da Informação – Universidade Estadual Paulista, Marília, São Paulo, 2008. 153 p.

_____. **Tecnologia e gestão pública municipal:** mensuração da interação com a sociedade. São Paulo: Cultura Acadêmica, 2009. 181 p.

SANT'ANA, R. C. G.; RODRIGUES, F. A. Visualização de afinidades entre parlamentares mediante dados de votações no Senado Brasileiro. **Informação & Sociedade:** Estudos, v. 23, n. 1, jan./ abr. 2013. Disponível em: <<http://www.ies.ufpb.br/ojs2/index.php/ies/article/view/12903>>. Acesso em: 1 jul. 2013.

SENAY, H.; IGNATIUS, E. **Rules and principles of scientific data visualization.** /s.n./ ACM SIGGRAPH Education Committee. Association for Computing Machinery, 1999. Disponível em: <<http://www.siggraph.org/education/materials/HyperVis/percept/visrules.htm>>. Acesso em: 20 jun. 2013.