



Centro de Convenções Ulysses Guimarães Brasília/DF – 16, 17 e 18 de abril de 2013

INFRAESTRUTURA NACIONAL DE DADOS ABERTOS

Augusto Herrmann Batista
Nitai Bezerra da Silva
Christian Moryah Contiero Miranda



INFRAESTRUTURA NACIONAL DE DADOS ABERTOS

Augusto Herrmann Batista
Nitai Bezerra da Silva
Christian Moryah Contiero Miranda

RESUMO

A publicação de dados abertos pelos governos permite que a sociedade acesse e reutilize dados brutos para fins diversos. Assim, surgem oportunidades de serviços inovadores, novos negócios, e maior participação do cidadão na gestão pública e nas atividades do governo. Para potencializar essas iniciativas, o Ministério do Planejamento criou a Infraestrutura Nacional de Dados Abertos (INDA) – que é um conjunto de políticas e de meios para catalisar a publicação e o acesso a dados abertos do governo. A INDA é a política do governo brasileiro para dados abertos. Neste trabalho, o Ministério apresentará como funciona a INDA, seus objetivos, como está sendo elaborado o plano de ação e a relação dela com a Parceria para Governo Aberto. Também apresenta o Portal Brasileiro de Dados Abertos: o dados.gov.br, um de seus principais produtos



INTRODUÇÃO

Governos de todo o mundo, e em que o Brasil não é exceção, enfrentam o problema de ter a maioria de seus dados armazenados em silos isolados, alguns em bases de dados proprietárias e em formatos fechados, atrás de sistemas quase sempre não interoperáveis. Com o passar dos anos, o volume de dados tem crescido muito e estamos cada vez mais dependentes de sistemas de informação. Nesse cenário, a falta de interoperabilidade entre diferentes departamentos governamentais torna-se ainda mais crítica. Na tentativa de atenuar o problema, muitos governos começaram a investir em iniciativas de interoperabilidade baseada em padrões abertos. O conjunto de padrões brasileiros da e-PING¹ é uma dessas iniciativas.

A publicação de dados abertos governamentais é amplamente conhecida como uma forma de empoderar o cidadão e promover a transparência. Mas uma grande utilidade tem sido pouco explorada: o uso de dados abertos por setores de governo, para suprir as suas próprias necessidades de informação. O consumo de dados já publicados por outra agência ou órgão, na forma de Dados Abertos, referenciando-os em seus próprios sistemas de informação, apresenta várias vantagens.

No presente texto, serão explorados alguns desses problemas de interoperabilidade e as vantagens que os dados abertos podem trazer como uma solução alternativa a esses problemas.

OBJETIVOS

O objetivo geral deste trabalho é considerar a possibilidade de utilizar a Infraestrutura Nacional de Dados Abertos, principal projeto de publicação de dados abertos do executivo federal, como um caminho para se atingir a interoperabilidade entre sistemas de informação de governo, no que tange aos pontos de integração que utilizem dados públicos.

¹ Padrões de Interoperabilidade de Governo Eletrônico - disponível em <http://www.governoeletronico.gov.br/acoes-e-projetos/e-ping-padroes-de-interoperabilidade>, acesso em 18 jul 2012.



A ampliação do volume de dados abertos publicados é uma das consequências indiretas do uso dos dados abertos para integração, devido ao aumento na demanda interna por dados abertos e à priorização dos projetos de abertura de dados. Como consequência, esse maior volume de dados disponíveis tende a estimular a economia relacionada ao uso de dados abertos governamentais pela sociedade civil, na medida em que agentes econômicos privados os utilizam na criação de novos processos de negócio e na otimização dos processos existentes. A primeira ocorre quando empresas utilizam os dados públicos para gerar novos serviços e aplicações úteis à sociedade. A última pode ocorrer, por exemplo, quando um comerciante decide onde instalar uma nova filial com base em informações demográficas ou na existência de determinados serviços públicos no local. Assim, visa-se também alcançar esse resultado indireto ao propor que os dados abertos sejam um instrumento de interoperabilidade.

O uso de dados ligados objetiva implementar uma plataforma onde os cidadãos possam ter acesso a dados estruturados semanticamente, o que promove um maior grau de transparência, na medida em que a informação está melhor organizada, pode ser mais facilmente encontrada e melhor compreendida. Esse novo conjunto de boas práticas, aliado ao melhor uso da plataforma web, aumentam as possibilidades de busca dessas informações pelos cidadãos.

METODOLOGIA

O presente documento constitui em uma reflexão baseada nas experiências empíricas vivenciadas pelos autores no contexto da implementação da Infraestrutura Nacional de Dados Abertos e dos compromissos do Ministério do Planejamento no Plano de Ação da Parceria para Governo Aberto.

DESAFIOS DA INTEROPERABILIDADE DOS SISTEMAS DE GOVERNO

Antes de ponderar sobre a interoperabilidade entre sistemas de governo, é necessário relembrar a trajetória do desenvolvimento do governo eletrônico do Brasil, pois aí encontra-se a raiz de muitas barreiras à integração dos sistemas de informação. Algumas das dificuldades do Estado começam com um crescimento



desenfreado e despadronizado de alguma coisa, esse é mais um exemplo: a informatização não padronizada da Administração Pública acabou trazendo problemas de integração de sistemas para diversos países.

Considerando a natural divisão de papéis e competências dentro de uma estrutura de Estado, cada organização ou área ficou responsável pela gestão de informações do seu próprio negócio. Esse modo de administrar influenciou e engessou o processo da informatização do governo brasileiro: a Tecnologia da Informação (TI) não foi tratada de forma sistêmica, sua natureza de "área meio" motivou cada organização a gerir sua própria TI. De fato, não havia uma metodologia consolidada para fazer isso. O processo de informatização de cada área de negócio seguiu uma arquitetura da informação que refletia a visão e o modelo da sua respectiva organização, de forma isolada das demais.

Iniciadas as necessidades de se integrar as informações das diversas áreas (que são, no caso do Brasil, descentralizadas) do governo e o crescimento do volume das informações geridas por ele, foram surgindo problemas de interoperabilidade.

A princípio acreditava-se que as limitações eram simplesmente tecnológicas, problemas clássicos de interoperabilidade sintática. O uso do mesmo protocolo ou da mesma sintaxe de comunicação foram as primeiras barreiras para fazer os sistemas conversarem, sendo estes resolvidos com relativa facilidade.

Depois de alguns anos de trabalho, percebeu-se que a questão não era exclusivamente tecnológica. A causa raiz das complicações era a forma como as informações eram tratadas e organizadas. Cada departamento tinha uma forma diferente de interpretar e processar as informações de seus negócios, discordâncias semânticas sobre quais são as entidades, sua denominação, de que forma elas devem ser descritas e os relacionamentos entre elas.

Um exemplo clássico desse problema de interoperabilidade semântica é o caso órgão-departamento: o sistema A modela órgão como qualquer unidade organizacional, desde uma autarquia até uma coordenação-geral dentro de um ministério. Já o sistema B modela órgão somente como aquela entidade superior que agrupa todas as unidades organizacionais em uma autoridade "máxima" (uma organização como um todo), designando outra denominação para as unidades administrativas internas. Esse exemplo, da modelagem conceitual de órgão, é um problema recorrente na administração pública brasileira.



POSSÍVEIS SOLUÇÕES

A resolução de problemas de interoperabilidade sintática, como colocado anteriormente, é feita através da uniformização do canal de comunicação e da mensagem trocada pelos sistemas. Na prática, são acordados quais protocolos de troca serão utilizados, as linguagens, os formatos de arquivo e a codificação de caracteres. Esse problema pode ser considerado trivial, pois no mercado já existem empresas, metodologias e ferramentas especializadas em sua resolução.

Os novos desafios da integração de Sistemas de Informação do governo brasileiro, giram em torno da integração de conceitos entre as áreas de negócio das organizações e desenvolvimento dessas visões em seus respectivos sistemas, para que não existam mais divergências conceituais no momento das integrações.

Analizando um cenário ideal, um sistema distribuído com dois ou mais componentes autônomos e interdependentes em alguma funcionalidade (a que demanda a interoperação), podemos desenhar duas alternativas:

1. Geração de uma camada de integração, na qual todos os sistemas entendam a mesma estrutura (por exemplo, um arquivo XML, com um esquema bem especificado), sem precisar, necessariamente, alterar o modelo de dados dos sistemas. Apenas a estrutura da mensagem ou transação é considerada, e o consenso se dá apenas no contexto destas. Não há maiores atenções quanto às estruturas dos modelos de dados internos de cada sistema;
2. Normalizar os conceitos das entidades de negócio com os responsáveis pelo negócio (e também os responsáveis pelo sistema) e modelar os respectivos domínios da informação dos sistemas, de maneira que essa uniformização seja feita numa conceituação comum e mapeada para os seus respectivos modelos de dados. Dessa forma, garante-se que cada aspecto do negócio seja tratado sob um sentido que é de comum acordo a ambos.

Consideremos a primeira solução. Dentre as duas, ela é a resposta mais barata e mais rápida quando há necessidade de integração.



Acordam as partes (responsáveis pelos sistemas) sobre protocolos, linguagens, formatos de arquivos e codificação de caracteres e sobre o modelo de dados utilizado na integração. Nesse momento é normal que algum dos sistemas precise fazer uma adaptação entre seu modelo de dados para o modelo do arquivo de troca, ou seja, um “de-para” de uma ou mais entidades que não são representadas exatamente da mesma maneira no arquivo. É importante esclarecer que essa mudança não é feita no modelo de dados, ela é feita no momento da serialização das entidades para o arquivo, para garantir que o outro sistema consiga interpretá-lo.

Esse método permite integrações rápidas, mas a interoperabilidade ocorre apenas como uma troca de mensagens fortemente acoplada e de utilidade pontual para aquela integração específica. Havendo uma nova necessidade de integração entre os mesmos sistemas, ou a necessidade de compartilhar as mesmas informações com um terceiro, todo o processo e os seus custos têm necessariamente que se repetir, sem que praticamente nada possa ser reaproveitado. Além disso, permanece a falta de acordo entre os respectivos modelos de dados dos sistemas. Até agora analisamos uma integração entre dois sistemas, mas o que acontece quando aumentamos as variáveis?

Havendo sucesso na primeira integração, o número de sistemas que dela participam pode aumentar. Isso é comum no governo pois os negócios têm demandas e ofertas muito específicas e por vezes precisam trocar informações. Se cada área é responsável por um negócio, seus sistemas vão refletir essa realidade.

Na medida que novos componentes vão se integrando, estes também terão que fazer adaptações, para que o modelo de integração permaneça o mesmo. É aqui que os problemas começam a aparecer.

Fazer a integração de dois sistemas que não possuem o mesmo modelo de dados ou cujas entidades trocadas não são estruturadas da mesma forma, possui uma complexidade. Fazer isso com 3 sistemas eleva a complexidade consideravelmente. Conforme o número aumenta, a proporção de aumento do custo também. Tendo-se n sistemas a integrar, seria necessário projetar $n \times (n-1)$ interfaces entre os mesmos. Além disso, após a sua implantação, manter essa infraestrutura e garantir sua manutenção ao longo do tempo pode tornar-se algo inviável.



Uma forma de atenuar esse problema, reduzindo o esforço necessário e os custos a médio e longo prazo, é produzir um novo modelo global, que represente as entidades e relacionamentos que são de comum acordo entre os domínios de informação de todos os sistemas que compartilham alguma informação. Os modelos de dados existentes são, então, mapeados para esse modelo comum, e a troca de informações entre os mesmos se dá por interfaces descritas conforme esse modelo. Essa é a segunda hipótese de solução.

Esse modelo comum traz consigo ainda outras vantagens. Ele é a visão que um agente externo ao sistema tem de seus dados. Essa visão serve não somente para projetar as interfaces de trocas de mensagens entre sistema, mas também para expor a parte pública dos dados ao cidadão, na internet, como forma de promover a transparência governamental. Isso, como veremos adiante, é uma condição necessária, mas não suficiente, para promover uma política de dados abertos.

Quando há uma preocupação em tornar essa modelagem o mais reaproveitável possível, de modo a maximizar o seu reúso, o modelo comum utilizado deve ser uma ontologia. Ao se projetar uma ontologia, preocupa-se não apenas com o registro das entidades, sua nomenclatura e os relacionamentos entre elas que são pertinentes ao negócio, mas também em documentar o seu significado. Registram-se as descrições e estabelecem-se as relações semânticas com entidades internas e externas que não necessariamente servem diretamente ao propósito do negócio, mas que evitam ambiguidades oriundas da terminologia utilizada. Por exemplo, ao modelar uma entidade chamada “Servidor”, estabelecer o relacionamento como subclasse de “Pessoa” deixa claro que não se trata de uma máquina servidora de software.

O uso de ontologias para descrever o domínio da informação dos sistemas da administração pública pode ter ainda outras vantagens. O desenvolvimento de novos sistemas, ao utilizar e expandir as ontologias existentes, já podem nascer com conceituação consensuada com os demais sistemas que utilizam a ontologia, diminuindo a impedância e os custos de futuras integrações.



Além disso, os dados públicos divulgados na internet para cumprir os requisitos de transparência governamental podem ser descritos, também, conforme a ontologia desenvolvida. Isso facilita à sociedade o cruzamento dos dados com outras fontes de diversas origens, sejam elas governamentais ou não. Os governos de alguns países no mundo têm feito justamente isso, usando os padrões da web semântica e os princípios dos dados ligados na publicação de dados abertos.

WEB SEMÂNTICA E DADOS LIGADOS

A Web Semântica pode ser descrita como a web dos dados. A World Wide Web consiste num conjunto de documentos hipertexto ligados entre si, e que ao seguir essas ligações é possível chegar a qualquer ponto da “teia” a partir de qualquer outro ponto. De forma análoga, a Web Semântica tem como ideia central a ligação entre dados relacionados entre si, formando um grafo conectado globalmente, no qual é possível navegar pelos dados e chegar a qualquer outro ponto no mundo.

A ideia da Web Semântica foi lançada por Tim Berners-Lee, Jim Hendler e Ora Lassila, num artigo para a revista *Scientific American* (2001). Ao longo dos anos, vem sido promovida pelo *World Wide Web Consortium* (W3C), que tem se empenhado em desenvolver e publicar um conjunto de padrões técnicos relacionados ao uso da Web Semântica.

O mais básico dos padrões, o RDF (*Resource Description Framework*) fornece uma estrutura de grafos para suportar os dados, os quais são registrados por meio de triplas que o descrevem, que consistem em tuplas contendo sujeito, predicado e objeto. O sujeito e objeto são nós no grafo, e o predicado é uma aresta, ou relacionamento.

A mesma infraestrutura também serve para suportar os metadados, tais como tipo (classe), sua descrição e seus relacionamentos (e.g. especialização, equivalência). Esses metadados podem ser representados, por exemplo, por meio de vocabulários ou de ontologias. Vários dos padrões técnicos do W3C para a Web Semântica, tais como o RDF Schema e OWL (*Web Ontology Language*), estabelecem como se dá o seu uso em RDF, e ainda, o uso regras de inferência.



Estas, por sua vez, permitem a uma máquina processar as assertivas existentes em um grafo e inferir novas informações factuais. Por exemplo, se Sócrates é um filósofo e todo filósofo é um homem, Sócrates é um homem.

Se, por um lado, inicialmente enfatizou-se na Web Semântica, quase que exclusivamente, o aspecto da descrição dos dados por ontologias e a inferência de novos fatos usando lógica descritiva, por outro foi deixado de lado o aspecto de rede global, inicialmente envisionado por Berners-Lee. Até que, em 2006, publicou o documento *Linked Data: Design Issues* (Dados Ligados: Questões de Projeto) (T Berners-Lee - W3C, 2006 - ACM Press, 2006), no qual voltou a enfatizar aspectos necessários à Web Semântica para que se concretizasse em uma rede global de dados interconectados, o que ele chama de *Giant Global Graph* (Grafo Gigante Global).

Segundo ele, para atingir esse ideal, é necessário seguir os seguintes princípios:

- 1 Usar URIs como nomes para coisas;
- 2 Usar URIs do tipo HTTP, para que as pessoas possam consultar por esses nomes;
- 3 Quando alguém consultar uma URI, forneça informações úteis, usando os padrões (tais como RDF e SPARQL);
- 4 Inclua ligações para outras URIs, para que as pessoas possam descobrir mais coisas.

Esses conceitos têm sido cada vez mais utilizado por empresas e governos em todo o mundo para publicar dados na Web Semântica. Exemplos disso são os dados sobre vários assuntos (e.g. estrutura organizacional governamental, despesas públicas) publicados pelo governo do Reino Unido no portal data.gov.uk e os dados geoespaciais publicados pelo Grupo de Engenharia de Ontologias da Escola Politécnica de Madri, na Espanha, no portal geo.linkeddata.es.

DADOS ABERTOS GOVERNAMENTAIS, CONCEITO E HISTÓRIA

Segundo a Fundação do Conhecimento Aberto (Open Knowledge Foundation), Dados Abertos são:



“...dados que podem ser livremente usados, reutilizados e redistribuídos por qualquer pessoa - sujeitos no máximo, à exigência de atribuição da fonte e compartilhamento pelas mesmas regras.

...

Se você está se perguntando porque é tão importante ser claro no que significa aberto e porque essa definição é usada, a resposta é simples: **interoperabilidade**.

É importante esclarecer que o objetivo dos dados serem abertos é a interoperabilidade. Mais do que a interoperabilidade tradicional, feita por dois sistemas conhecidos, os princípios de Dados Abertos procuram fazer com que os dados tenham tamanha “abertura” que possam ser interoperados inclusive com sistemas desconhecidos, a serem utilizados por pessoas de fora das organizações governamentais (ou de outros departamentos, ou de outras organizações), em momentos oportunos, independente de conhecimento dos seus produtores.

Considerando que os sistemas consumidores de dados abertos não precisam desenvolver as suas interfaces de leitura no mesmo momento que os fornecedores de dados desenvolvem as suas APIs, essa forma de interoperabilidade poderia ser chamada de *interoperabilidade assíncrona*. Isto é, a disponibilização dos dados abertos se dá em um determinado momento e, em momentos posteriores, os vários consumidores de dados desenvolvem os seus respectivos códigos de integração.

A conceituação de dados abertos foi escrita e evoluída à partir das experiências de Dados Abertos ao redor do mundo. Nessas experiências, todas ocorridas nos últimos anos, os cidadãos têm exigido de seus governantes cada vez mais possibilidade de participação e transparência de suas ações. Algumas vezes através de ativismo presencial, outras através de ações na web.

Como resposta, alguns governos têm instituído algo chamado de Governo Responsivo (Braga et al, 2008), Governo 2.0 (O'Reilly, 2009) ou, em sua forma mais conhecida, Governo Aberto (Lathrop e Ruma, 2010). Tratam-se de mecanismos de fortalecimento da participação social nas atividades e decisões do Estado, uso de novas tecnologias e da internet para o aumento da transparência das informações públicas e reaproveitamento dos dados públicos para a criação de serviços inovadores para o governo e para o cidadão.



No Brasil, a movimentação em direção aos Dados Abertos foi distribuída e assíncrona, ocorreu em vários estados e esferas diferentes, de forma separada, antes de receber uma orientação legal. O primeiro caso de publicação de dados abertos foi do Tribunal de Contas dos Municípios do Estado do Ceará².

Em seguida, motivado por pressão social, o Ministério do Planejamento, Orçamento e Gestão iniciou uma política de disseminação de dados. A política, além de iniciar um processo de publicação que resultou na abertura de dois sistemas do Ministério (SICONV³ e SICAF⁴), desdobrou-se em um projeto que culminou na instituição da Infraestrutura Nacional de Dados Abertos - INDА.

Ao longo do tempo outras iniciativas surgiram, como o projeto de Governo Aberto do estado de São Paulo, projetos no Rio de Janeiro, Rio Grande do Sul, Alagoas, Pernambuco e Minas Gerais.

A Infraestrutura Nacional de Dados Abertos - INDА, no que tange aos seus objetivos de interoperabilidade, pretende reunir organizações comprometidas com a publicação de seus dados para padronizar e consensuar as melhores práticas de publicação de dados abertos. Isso em um primeiro momento, na segunda fase do projeto, ele pretende consensuar não só os padrões e tecnologias, mas os conceitos e a forma como as diferentes áreas da administração pública visualizam e modelam seus negócios e seus sistemas.

Essa crescente tendência em direção ao aumento da transparência e da participação social em governos de todo o mundo culminou, no ano de 2011, com a formação da Parceria para Governo Aberto, firmada entre o Brasil e outros 7 países. A Parceria visa a troca de experiências entre países sobre Governo Aberto e a adoção de compromissos públicos com o objetivo de ampliar a participação social, a transparência pública e o uso responsável dos recursos públicos. A iniciativa partiu inicialmente de dois países, o Brasil e os Estados Unidos. O Brasil, devido ao seu histórico positivo em transparência ativa (isto é, aquela em que o Estado divulga proativamente as informações, sem que tenham sido solicitadas), foi convidado a co-liderar a Parceria para Governo Aberto durante o primeiro ano de sua existência.

² Disponível em <http://api.tcm.ce.gov.br/>.

³ Sistema de Gestão de Convênios e de Contratos de Repasses, que controla os processos de transferências voluntárias da União a Estados, Municípios e OSCIPs.

⁴ Sistema de Cadastramento Unificado de Fornecedores, parte do Comprasnet que controla o cadastro de fornecedores, no qual empresas precisam se registrar para participar de licitações com a Administração Pública Federal.



Em 2012 foi líder, tendo a liderança sido passada, ainda no mesmo ano, para o Reino Unido. Outros países são incentivados a aderir à Parceria, mas para isso precisam satisfazer alguns critérios objetivos mínimos de qualificação, os quais consideram o grau de transparência pública efetivamente implantado no país, assim como propor compromissos para ampliá-la.

Também no ano de 2011, em 18 de novembro, a presidente da república sancionou a Lei 12.527, ou lei de acesso à informação, que trata de pedidos de acesso à informação e transparência ativa em todas as esferas e poderes do estado brasileiro. A lei passou a valer em 2012, mas desde o momento de sua sanção, trouxe preocupações a alguns órgãos, por vários motivos. Um motivo importante é a situação precária quanto à gestão documental e gestão da informação em que se encontravam os órgãos e entidades da administração federal. Outro, que foi uma inovação importante na lei brasileira, foi a previsão da publicação de forma que os dados de forma estruturada, para que possam ser reaproveitados por máquina, e sem restrições de reuso. Apesar da Lei não citar a expressão “dados abertos” de maneira explícita, essas provisões são essencialmente similares aos princípios de Dados Abertos. Essas mudanças provocaram as diferentes esferas e poderes de governo a repensar os seus métodos de gestão e publicação de informação para a sociedade.

Com a publicação de Dados Abertos, o Estado atende não apenas ao anseio do cidadão por mais transparência e participação na gestão pública, pelo desenvolvimento de visualizações de dados e aplicações integradas de interesse público e privado. Atende também o próprio Estado ao satisfazer mais prontamente as necessidades de integração de dados entre sistemas de informações, economizando recursos públicos de diversas formas. Por exemplo, recursos são reaproveitados quando os mesmos dados abertos são servidos a vários órgãos que têm a necessidade da informação. Economiza-se, também, eliminando os trâmites burocráticos que tradicionalmente ocorrem para viabilizar uma integração de sistemas, após os dados públicos já estarem abertos e disponíveis na internet.



ECONOMIA DOS DADOS ABERTOS

Outra motivação para a utilizar do método de publicação de Dados Abertos são os impactos econômicos que ele pode gerar. Esse aspecto pode despertar interesse nos governos, no intuito de melhorar sua economia, e nas organizações que visem lucro ou publicidade. Incluir o mercado nessas iniciativas potencializa suas chances de sucesso.

Os ganhos para as empresas são certos. Hoje no Brasil, centenas, talvez milhares de empresas comprem bases de dados do governo ou de outras empresas, para desenvolver suas soluções. São bases de endereços, bases geográficas, bases de saúde que poderiam ser obtidas diretamente do governo, mediante acesso aos dados já públicos ou pedidos de acesso à informação (oferecidos pela lei de acesso à informação).

Embora a princípio os visionários dos dados abertos visualizem desenvolvimento de aplicações para o público, nada impede que as empresas tomem vantagem dos dados públicos para fazer sistemas de apoio à tomada de decisão de seus clientes, por exemplo. Na verdade, isso já é feito, mas normalmente os dados são comprados.

Além disso, novas oportunidades de negócio são criadas diariamente. Publicação de dados sobre trânsito em formatos amigáveis já são realidade em vários locais do país, dados sobre linhas e paradas de ônibus, aplicativos que identificam serviços próximos ao cidadão, programas para acompanhamento sobre ações do poder legislativo, informações precisas sobre imóveis à venda, como antigos donos, ano de construção, etc., já existem no mercado internacional.

Os ganhos para o governo também são consideráveis. Vamos enumerar alguns, que são os mais evidentes, mas vale considerar que essa lista não está completa em nenhum lugar do mundo, pois até agora ninguém foi capaz de estimar o verdadeiro potencial dos Dados Abertos.

1. Geração de emprego: com novas oportunidade de negócio, novas empresas estão sendo criadas ou antigas empresas estão sendo ampliadas para geração de valor agregado com base nos dados públicos.



2. Economicidade com estudos sobre os dados do governo: algumas organizações estão deixando de pagar consultorias especializadas em análise de seus dados, porque essas análises já estão disponíveis na web, através de aplicativos gerados gratuitamente por cidadãos engajados.
3. Redução dos gastos com transparência passiva: diariamente, o governo brasileiro recebe inúmeros pedidos de informação, alguns formalmente, através da LAI, outros informalmente, feitos pelo próprio governo, de outros órgãos, de outras esferas ou de outros poderes. A resposta a esses pedidos demanda tempo, especialistas, servidores públicos e infraestrutura, para que os dados sejam extraídos e preparados para entrega, o custo disso nunca foi estimado no Brasil, mas certamente há ônus considerável para suas organizações. A publicação de dados abertos agilizaria o processamento desses pedidos de diversas formas. Primeiramente, teria efeito na redução da própria demanda, dado que muitas vezes o potencial demandante encontraria por si próprio a informação desejada, não chegando nem mesmo a fazer a solicitação. Os dados abertos também podem ajudar o receptor dos pedidos de acesso à informação a encontrá-la, seja diretamente em pesquisa nos conjuntos de dados já publicados, seja indiretamente, pelas melhorias na gestão da informação provocadas pela publicação dos dados abertos. Facilitam ainda a resposta, uma vez que o solicitante recebe apenas uma indicação (endereço) de onde a informação pode ser encontrada. Há ainda outro benefício ao solicitante, que é a possibilidade de obter atualizações da informação desejada sem a necessidade de um novo pedido, caso o ofertante mantenha seus dados abertos atualizados.
4. Aumento da produção: a geração de emprego, lucro de empresas, redução de custos dos cofres públicos, oferecem resultados consideráveis para os governos, segundo David Eaves [Case Study: how Open Data Saved Canada \$3.2 Billion], a publicação de dados abertos permitiu que ao governo canadense identificar 3.2 bilhões que haviam sido desviados dos cofres públicos por falsas organizações de caridade.



Uma análise de 2011, realizada por Grahan Vickery [Review of recent studies on PSI re-use and related market developments] estima que os ganhos com o reuso de dados públicos na Europa tenha sido de 32 bilhões de Euros em 2010.

Considerando que essas motivações são suficientes para a publicação de Dados Abertos Governamentais, é importante observar que, para fazer os dados serem usados e representarem vantagens para as empresas e os próprios publicadores (governo), eles precisam ter qualidade e sua disponibilidade precisa ser garantida. Se um sistema distribuído precisa de disponibilidade dos diversos componentes para poder funcionar, também precisará qualquer serviço que utilize dados do governo.

Segundo Jeni Tennison [Open Data Business Models], do Instituto de Dados Abertos do Reino Unido, *"Para que as empresas passem a utilizar os dados abertos governamentais em seus processos de negócio, e a eles agregar valor econômico e gerar riqueza, é necessário primeiro que o governo garanta a sustentação das iniciativas no longo prazo."*

Uma forma de garantir essa estabilidade é colocando o consumo de dados abertos nos processos internos de trabalho do próprio governo. Ou seja, promover a interoperabilidade de dados públicos pela publicação e consumo de dados abertos pelo governo. Isso é benéfico para a administração pública sob pelo menos dois aspectos: a agilidade na integração de sistemas de informação e a redução de custos.

CONCLUSÃO

Assim, considera-se a possibilidade do uso dos dados abertos para promover a interoperabilidade de sistemas de informação de governos como uma medida complementar que traz uma série de vantagens.

Uma delas é agilização do processo, evitando-se o tempo e esforço usualmente gastos com a burocracia da tramitação formal e análise de um pedido, bem como com as avaliações de pertinência, segurança da informação e privacidade e implementação de um canal de dados que normalmente ocorrem para



cada pedido de acesso. Com dados abertos, publicam-se os dados públicos uma única vez, e estes podem ser utilizados por um número ilimitado de outras organizações públicas, economizando recursos. O procedimento usual, lento, fica relegado apenas às informações que tenham restrições de sigilo e privacidade, o que representa apenas uma fração das necessidades de informação de outras organizações públicas.

Outra vantagem é que, a partir do momento que o próprio governo passa a depender dos dados abertos já publicados para o funcionamento de seus próprios sistemas de informação, há uma pressão interna, que se soma à pressão externa que normalmente vem da sociedade, para que os serviços sejam mantidos em funcionamento, com alta disponibilidade. Isso garante a sustentabilidade aos dados abertos publicados, beneficiando mais uma vez o cidadão.

Por fim, o conhecimento e experiência adquiridos pelos servidores públicos e pelos prestadores de serviços, durante o desenvolvimento de soluções para a publicação de dados abertos, referente aos padrões e tecnologias utilizadas, é útil também para a implementação de interoperabilidade para a troca de informações que não são públicas.

Portanto, muito embora a abertura de dados não seja a solução completa para problemas de interoperabilidade de sistemas de informação, traz vantagens suficientes para que ela seja colocada em prática sempre que se considere as formas de integração de sistemas e para que faça parte, de forma indissociável, de uma política de interoperabilidade.



REFERÊNCIAS

Berners-Lee, Tim; Hendler, James; Lassila, Ora. The Semantic Web. Sicientific Americam Magazine, 2001.

Berners-Lee, Tim. Linked Data: Design Issues. World Wide Web Consortium, 2006. Disponível em: <http://www.w3.org/DesignIssues/LinkedData.html> Acesso em 8 jul 2012.

Braga, Lamartine; et al. O papel do Governo Eletrônico no fortalecimento da governança do setor público. Revista do Serviço Público, v59 (1), p5-21. Brasília: Escola Nacional de Administração Pública, 2008.

Eaves, David. Case Study: how Open Data Saved Canada \$3.2 Billion. [eaves.ca](http://eaves.ca/2010/04/14/case-study-open-data-and-the-public-purse/). Disponível em <http://eaves.ca/2010/04/14/case-study-open-data-and-the-public-purse/> . Acesso em 21 jul 2012.

Heath, Tom; Bizer, Christian. Linked Data: Evolving the Web into a Global Data Space. Synthesis Lectures on the Semantic Web: Theory and Technology, 1:1, 1-136. Morgan & Claypool, 2011. Disponível em <http://linkeddatabook.com/editions/1.0/>. Acesso em 22 jul 2012.

Lathrop, David; Ruma, Laurel. Open Government: Transparency, Collaboration and Participation in Practice. O'Reilly, 2010.

Magalhães, Darcilene; Knight, Peter; da Costa, Eduardo Moreira. Will the Soccer World Cup of 2014 Help Bridge the Social Gap through the Promotion of ICT and E-government in Brazil? In: The Global Information Technology Report 2008-2009. World Economic Forum, 2009.

Reino Unido. Open Data White Paper. A document that seeks to lay out what citizens, businesses and the public sector can expect from government to help unlock the benefits of Open Data. Disponível em <http://data.gov.uk/library/open-data-white-paper> Acesso em 18 jul 2012.

Van Berlo, Davied. Civil Servant 2.0, New ideas and practical tips for working in government 2.0. Disponível em <http://boek.ambtenaar20.nl/> Acesso em 15 jul 2012.

Vickery, Grahan. Review of recent studies on PSI re-use and related market developments. Disponível em http://ec.europa.eu/information_society/policy/psi/docs/pdfs/report/psi_final_version_formatted.docx . Acesso em 19 jul 2012.

Tennison, Jeni. Open Data Business Models. Disponível em <http://www.jenitennison.com/blog/node/172> . Acesso em 20/08/2012.



AUTORIA

Augusto Herrmann Batista – Ministério do Planejamento, Orçamento e Gestão.

Endereço eletrônico: augusto.herrmann@planejamento.gov.br

Nitai Bezerra da Silva – Ministério do Planejamento, Orçamento e Gestão

Endereço eletrônico: nitai.silva@planejamento.gov.br

Christian Moryah Contiero Miranda – Ministério do Planejamento, Orçamento e Gestão,

Endereço eletrônico: christian.miranda@planejamento.gov.br

