DADOS ABERTOS PARA A
DEMOCRACIA NA ERA DIGITAL

Ministério das Relações Exteriores

Ministro de Estado Embaixador Antonio de Aguiar Patriota
Secretário-Geral Embaixador Ruy Nunes Pinto Nogueira

Fundação Alexandre de Gusmão

Presidente Embaixador Gilberto Vergne Saboia

Ministério da Fazenda

Ministro de Estado Guido Mantega

Serviço Federal de Processamento de Dados - SERPRO

Diretor - Presidente Marcos Vinícius Ferreira Mazoni

A *Fundação Alexandre de Gusmão*, instituída em 1971, é uma fundação pública vinculada ao Ministério das Relações Exteriores e tem a finalidade de levar à sociedade civil informações sobre a realidade internacional e sobre aspectos da pauta diplomática brasileira. Sua missão é promover a sensibilização da opinião pública nacional para os temas de relações internacionais e para a política externa brasileira.

Ministério das Relações Exteriores
Esplanada dos Ministérios, Bloco H
Anexo II, Térreo, Sala 1
70170-900 Brasília, DF
Telefones: (61) 3411-6033/6034/6847
Fax: (61) 3411-9125
Site: www.funag.gov.br

CONSEGI 2011
IV Congresso Internacional Software Livre e Comércio Eletrônico

Dados Abertos para a Democracia na Era Digital

Brasília, 2011

Direitos de publicação reservados à
Fundação Alexandre de Gusmão
Ministério das Relações Exteriores
Esplanada dos Ministérios, Bloco H
Anexo II, Térreo
70170-900 Brasília – DF
Telefones: (61) 3411-6033/6034
Fax: (61) 3411-9125
Site: www.funag.gov.br
E-mail: funag@itamaraty.gov.br

Equipe Técnica:
Henrique da Silveira Sardinha Pinto Filho
André Yuji Pinheiro Uema
Fernanda Antunes Siqueira
Fernanda Leal Wanderley
Juliana Corrêa de Freitas
Pablo de Rezende Saturnino Braga

Programação Visual e Diagramação:
Juliana Orem

Programação Visual da Capa:
Rômulo Geraldino

Impresso no Brasil 2011

Dados abertos para a Democracia na Era Digital. – Brasília: Fundação Alexandre de Gusmão, 2011.
84 p.

I. Tecnologia da Informação. II. Governança da Internet. III. Tecnologia Digital.

ISBN: 978.85.7631.293-2

CDU: 65.011.56

Ficha catalográfica elaborada pela Bibliotecária
Sonale Paiva – CRB /1810

Depósito Legal na Fundação Biblioteca Nacional conforme Lei n° 10.994, de 14/12/2004.

Sumário

Apresentação, 7
Marcos Vinícius Ferreira Mazoni

Política Externa para a Sociedade da Informação: a contribuição da diplomacia para a democracia na era digital, 11
Benedicto Fonseca Filho
Livia Oliveira Sobota

Rede de Informações Pública e Aberta - DadosGov - Um modelo aberto para a troca de informações, 21
Ednylton Maria Franzosi
Miriam Chaves
Gilberto Paganotto
Marcus Vinicius da Costa
Sérgio Assis Rodrigues
Jano Moreira de Souza

Segurança eletrônica em *Web* Semântica, 53
Jose Maria Leocadio

A Disseminação de Dados Governamentais como Serviço Público – Os Dados Abertos Governamentais e a Experiência Brasileira, 63
Christian Moryah Contiero Miranda

A Infraestrutura Nacional de Dados Espaciais, um projeto colaborativo para interoperabilidade e disseminação de dados abertos no Governo Brasileiro, 75
Hesley da Silva Py

Apresentação

O papel dos dados abertos para seguir construindo um novo Brasil

Um dos principais desafios para o avanço da democracia é o desenvolvimento contínuo das ferramentas de interação entre governo e sociedade. As novas tecnologias da informação e comunicação, com a internet no centro da revolução, formam um cenário ideal para ampliar o acesso e a utilização dos dados públicos pelos cidadãos.

É preciso evoluir o Governo Eletrônico sempre, da mesma forma que surgem as versões aperfeiçoadas dos sistemas, linguagens e tecnologias. A própria Web alcança um novo patamar, agora denominada 3.0. Seguindo a mesma lógica evolutiva, o e-Gov e a e-Democracia avançam com o ingresso do conceito dos dados abertos.

A novidade impõe-se às estratégias políticas e aos processos de governança, aumenta a exigência da qualificação das formas de interação entre governos, pessoas e países. Exige também maior eficiência nos serviços ofertados, na prestação de contas, transparência, colaboração e integração.

Nos países preocupados em fortalecer a democracia, a disponibilização de dados públicos na Web torna-se requisito básico para a consolidação de um Governo 2.0. Fornecer os dados de forma aberta significa aproximar o cidadão do Estado, representa enriquecer a própria democracia. Não há dúvida de que o acesso à informação pública é uma prerrogativa de todo cidadão, e a possibilidade de reutilizá-la, recombiná-la e dar a ela novos significados é uma ampliação necessária deste direito.

Mesmo sem um conceito unânime sobre dados abertos, uma definição começa a ganhar aceitação: a da Open Knowledge Foundation. Em resumo, ela estabelece que um dado - ou uma porção de conhecimento - é aberto se estiver livre para uso, reuso e redistribuição. Um princípio que pode ser aplicado à música, filmes, livros, imagens e outros materiais.

Trata-se de uma filosofia que demanda a disponibilidade de dados para todos, sem restrições de copyright, patentes ou outros mecanismos de controle. Um caráter similar à lógica adotada em diversas comunidades de software de código aberto - movimentos que precisam estreitar ainda mais as suas relações.

O governo brasileiro aposta na iniciativa de disponibilização de dados nos ambientes Web para que eles possam ser transformados por pessoas e máquinas. Informações públicas compartilhadas em formato bruto e aberto que poderão ser reutilizadas e transformadas para atender a cada necessidade do cidadão, com diferentes possibilidades de cruzamentos e análises.

Empresas de tecnologia podem também trabalhar com os dados abertos. E, na medida em que atuam como impulsionadores de um conceito inovador, fomentam novos empreendimentos. Todos estão aptos a ser agentes dessa mudança, desenvolvedores de software ou não, podendo sugerir e criar serviços que não foram imaginados por entidades governamentais ou até mesmo que não poderiam ser implementados por estas. Abre-se espaço para uma nova forma de parceria entre Estado e sociedade, favorecendo a emergência da inteligência coletiva.

Os dados abertos estão fundamentados em três pilares: transparência, participação e colaboração. A transparência promove a responsabilidade de informar os cidadãos sobre o que o governo está fazendo e que ações se pretende tomar. A participação permite aos cidadãos contribuir com suas ideias e competências, auxiliando o poder público a elaborar políticas eficazes e abrangentes, graças à informação amplamente dispersa em nossa sociedade. Essa colaboração aprimora a eficácia do Estado, encorajando a cooperação entre a sociedade, os diferentes níveis de governo e a iniciativa privada.

O Brasil é pioneiro nas discussões sobre dados abertos na América Latina. Para manter esta posição de referência, deve investir na implementação de quatro importantes passos: publicar informações governamentais online, qualificar estes dados, construir a cultura de

governo aberto e criar um arcabouço de políticas favoráveis para guiar e consolidar o processo.

O país só tem a ganhar com a abertura de dados dos governos para a sociedade. A livre combinação de dados geográficos, financeiros, demográficos, orçamentários, educacionais e outros, acessíveis em formato aberto na Web, é a garantia de geração de conhecimento e democratização da informação. Com isso, aumenta-se a participação civil nas decisões políticas, aperfeiçoam-se os serviços prestados e estimula-se o crescimento do Brasil.

Com engajamento acadêmico, adoção de políticas e padrões tecnológicos e com a participação constante e direta de empresas e cidadãos, é viável construir um modelo de dados abertos no governo brasileiro. O presente livro analisa, sob a perspectiva de Estado, a necessidade de abraçar de forma holística os conceitos de disponibilização de dados, transparência, colaboração e participação para uma interação mais completa e ampla entre governo e sociedade.

Boa Leitura!

Marcos Vinicius Ferreira Mazoni
Presidente do Serpro

Política Externa para a Sociedade da Informação: a contribuição da diplomacia para a democracia na era digital

Benedicto Fonseca Filho[1]
Livia Oliveira Sobota[2]

Proposta de artigo para a publicação do IV Congresso Internacional *Software* Livre e Governo Eletrônico (CONSEGI 2011) sobre "Dados abertos para a democracia na era digital".

Introdução

A Internet e as tecnologias digitais de uma forma geral são fenômenos recentes, mas já mudaram, quantitativa e qualitativamente, a forma como as pessoas e instituições comunicam-se e organizam-se. Seja acentuando a velocidade e abrangência de trocas de informações que já eram realizadas por outros meios de comunicação, seja possibilitando novos tipos de intercâmbios, a Internet trouxe consequências importantes para as relações internacionais.

Entre os novos desafios que se apresentam aos governos poderíamos mencionar a garantia do acesso da população à Internet e às novas tecnologias digitais; a utilização de tais meios para modernizar a administração pública e conferir maior transparência na relação com os cidadãos; e a indução do desenvolvimento dessas tecnologias com o fortalecimento dos respectivos setores produtivos. O governo brasileiro, em

[1] Diretor do Departamento de Temas Científicos e Tecnológicos - DCT/ MRE.
[2] Subchefe da Divisão da Sociedade da Informação - DI/ DCT/ MRE

particular, tem buscado responder a esses desafios com políticas públicas nas áreas de inclusão digital; governo eletrônico; ciência, tecnologia e inovação; desenvolvimento produtivo; e impulso à televisão digital.

A construção dos modelos de governança dessas novas tecnologias constitui um desafio da maior complexidade para os governos, pois começa na necessária distinção entre as decisões de ordem técnica, que não requerem sua participação, e as de ordem política, que a exigem. Essa distinção, embora tão difícil quanto afirmar a diferença entre o técnico e o político, faz-se necessária para assegurar, de um lado, que os governos não se envolvam no cotidiano dos processos técnicos (que necessitam de liberdade, agilidade e flexibilidade) e, de outro, que não deixem de atuar no que se refere à proteção do interesse público e dos direitos dos usuários.

Não é tarefa simples identificar o interesse público no que diz respeito às redes digitais. Diversas formas de interação são aproveitadas por usuários com finalidades igualmente diversas. Se para alguns as redes são oportunidades de negócios, para outros são ferramentas de interação em redes sociais, participação política e mobilização social, desenvolvimento colaborativo de novas tecnologias, compartilhamento de conhecimento e valorização de manifestações culturais. Há, também, os que as aproveitam para praticar crimes cibernéticos e promover discursos de ódio. A definição do interesse público deve estar, portanto, relacionada aos direitos desses usuários à comunicação, ao conhecimento, à segurança, à privacidade, à expressão e à honra – que devem coexistir de forma equilibrada –, e como são afetados pelos processos de governança.

O papel da Política Externa para a Sociedade da Informação

A discussão de questões ligadas às tecnologias digitais entrou definitivamente na agenda internacional no processo que culminou na Cúpula Mundial sobre Sociedade da Informação (CMSI ou WSIS, do inglês *World Summit on the Information Society*), que se realizou em duas fases – Genebra (2003) e Túnis (2005). A CMSI envolveu temática abrangente, cuja síntese pode ser encontrada na "Declaração de Princípios e Plano de Ação de Genebra" e na "Agenda de Túnis", documentos que, por representarem o consenso dos Estados, têm servido para orientar o debate internacional sobre a matéria.

O Itamaraty tem dedicado atenção crescente ao tema. Em 2009, foi criada, no âmbito do Ministério das Relações Exteriores, a Divisão da Sociedade da Informação (DI), vinculada ao Departamento de Temas Científicos e Tecnológicos. Compete à DI, entre outras atribuições, acompanhar, dirigir e orientar a posição oficial brasileira em relação às tecnologias da informação e comunicações (TICs), governo eletrônico, inclusão digital, governança da Internet e temas conexos; acompanhar, instruir e supervisionar a posição oficial brasileira em reuniões, conferências, organismos, negociações e foros globais, regionais e bilaterais na área de sua competência; e negociar aspectos externos das políticas públicas relativas à Sociedade da Informação e temas conexos.

A atuação do MRE nos foros onde o assunto é discutido reflete os princípios gerais da política externa brasileira, entre os quais a defesa dos direitos humanos e do multilateralismo, a promoção da integração sul-americana e a construção de uma ordem internacional orientada para a paz, a segurança e o desenvolvimento. Ao fazê-lo, procura contribuir para a consecução dos objetivos das políticas públicas nacionais.

Nesse contexto, a política externa para a Sociedade da Informação envolve três frentes principais: governança da Internet, inovação (alta tecnologia em TICs) e utilização das TICs para o desenvolvimento, as quais serão, a seguir, examinadas em maior profundidade.

Governança da Internet

Considerando-se que a Internet organiza-se em "camadas", os principais temas de sua governança são, entre outros: (i) na "camada técnica", a administração do Sistema de Nomes de Domínios (DNS), incluindo seus servidores raiz, e a neutralidade de rede; e (ii) na "camada de uso", a utilização de padrões abertos, o acesso ao conhecimento (e às condições para sua produção e circulação), a promoção da diversidade e do multilinguismo na rede, a facilitação do comércio eletrônico, a proteção dos direitos humanos, a preservação da privacidade e da segurança e o combate a crimes cibernéticos. São relevantes também os debates sobre a necessidade de reduzir os custos de interconexão que tanto encarecem o acesso à Internet nos países em desenvolvimento; implementar Pontos de Troca de Tráfego (PTTs) que deem eficiência ao

tráfego de dados; e tratar temas emergentes, como as redes sociais e a computação em nuvem.

O Brasil tem buscado participar ativamente da constituição do regime de governança global da Internet e defendido, com base na Agenda de Túnis, a participação isonômica entre governos nesse processo, que deve ser aberto, democrático e transparente. Essa posição brasileira, ao mesmo tempo crítica e construtiva, é coerente com as políticas internas para a Internet, exemplificadas pela gestão do "código-país" brasileiro (o ".br"), que se encontra a cargo do Comitê Gestor da Internet no Brasil (CGI.br). Em linha com o "espírito de Túnis", o CGI.br envolve a participação de representantes do governo, da comunidade técnica, da sociedade civil e do meio empresarial, que baseiam sua atuação nos "Princípios para a Governança e o Uso da Internet no Brasil" desenvolvidos no âmbito do próprio CGI.br, a saber: liberdade, privacidade e direitos humanos; governança democrática e colaborativa; universalidade; diversidade; inovação; neutralidade da rede; inimputabilidade da rede; funcionalidade, segurança e estabilidade; padronização e interoperabilidade; e ambiente legal e regulatório que preserve a dinâmica da Internet como espaço de colaboração.

Internacionalmente, no que se refere à governança em sentido estrito, relativa aos "recursos críticos" da Internet, o modelo tem como núcleo, desde 1998, a "Corporação da Internet para Atribuição de Nomes e Números – ICANN", entidade privada sediada na Califórnia. Autodefinida como uma "corporação internacional sem fins lucrativos", a ICANN está vinculada ao Departamento de Comércio dos EUA por contrato segundo o qual se delega a ICANN o exercício das funções de administração do sistema de domínios de primeiro nível, tanto genéricos (gTLDs) quanto códigos de país (ccTLDs), a coordenação de parâmetros para protocolos técnicos da Internet e a alocação de números IP (Protocolo de Internet).

A relação entre a ICANN e o governo norte-americano é regida, atualmente, pela "Afirmação de Compromissos – AoC", de setembro de 2009, documento que, embora represente avanços em relação ao arranjo anterior em matéria de transparência, não solucionou o problema da desigualdade entre os governos no desempenho de suas funções no regime de governança da Internet. A participação dos governos nessa entidade se dá por meio de "Comitê Consultivo

Governamental – GAC", que se reúne a cada quadrimestre à margem da reunião geral da ICANN.

No âmbito das Nações Unidas, como resultado dos debates da CMSI, o Fórum de Governança da Internet (IGF) foi instituído pela Agenda de Túnis com o objetivo de "discutir políticas relacionadas a elementos-chave da governança da Internet para estimular a sustentabilidade, o vigor, a segurança, a estabilidade e o desenvolvimento da Internet". O IGF é aberto a participação de todas as partes interessadas ("*stakeholders*"): governos, empresas, sociedade civil, comunidade acadêmica e indivíduos. O Brasil teve atuação expressiva em todas as reuniões do IGF – realizadas em Atenas, na Grécia (2006), Rio de Janeiro (2007), Hyderabad, na Índia (2008), Sharm-El-Sheik, no Egito (2009) e Vilna, na Lituânia (2010) –, bem como em seu "Grupo Consultivo Multissetorial" (MAG), criado para assessorar o Secretário-Geral da ONU na preparação do IGF.

Caberia mencionar, ainda, os processos de seguimento da CMSI, representados pelas reuniões da Comissão de Ciência e Tecnologia para o Desenvolvimento da Organização das Nações Unidas (CSTD/ONU) e do Fórum CMSI (*WSIS Forum*, que se reúne sob os auspícios da União Internacional de Telecomunicações), ambas realizadas anualmente em Genebra.

Inovação: alta tecnologia em TICs

No que se refere à Ciência, Tecnologia e Inovação (CTI), o Itamaraty tem buscado contribuir para a construção e o aprofundamento de parcerias do Brasil com outros países e organismos internacionais que ajudem a fortalecer o Sistema Nacional de CTI e superar os gargalos da indústria brasileira de TICs, em linha com o Plano de Ação em Ciência, Tecnologia e Inovação para o Desenvolvimento Nacional (PACTI 2007 – 2010). Esse plano estabelece, em sua linha de ação nº8, o objetivo de "promover e apoiar atividades de formação e capacitação de recursos humanos em tecnologias da informação e comunicação (TICs), incentivar as atividades de P,D&I e de produção, por meio da cooperação entre ICTs e empresas, da instalação e da ampliação de empresas de manufatura e de serviços no País."

Constituem, nesse contexto, setores de interesse prioritário para o Brasil para a construção de parcerias internacionais, como os de *software*,

semicondutores, supercomputadores, computação em nuvem, Internet do futuro e Internet das coisas.

TICs para o desenvolvimento

A Agenda de Túnis reconhece a "crescente importância do papel das TICs não apenas como um meio de comunicação, mas também como habilitadoras do desenvolvimento e ferramentas para o cumprimento de metas e objetivos acordados internacionalmente, dentre eles as Metas de Desenvolvimento do Milênio". Reconhece, ainda, que "enfrentar o hiato digital ( ) requererá investimentos adequados e sustentáveis na infraestrutura e nos serviços de TICs, bem como na construção de capacidades e na transferência de tecnologia".

Nesse contexto, cabe recordar que uma das prioridades do governo brasileiro refere-se à implementação do Programa Nacional de Banda Larga (PNBL), lançado em 2010, com o objetivo de massificar o acesso à Internet em banda larga no Brasil, com maior capilaridade e velocidade e menores preços. O PNBL deverá ensejar intensa cooperação com os demais países da América do Sul. Ao melhor integrar seu território nas redes digitais, o Brasil colaborará de forma decisiva para o processo regional de integração de infraestruturas e, paralelamente, para a produção de conteúdos latino-americanos.

Nesse mesmo contexto, a adoção do padrão nipo-brasileiro de televisão digital (ISDB-T) pela quase totalidade dos países da América do Sul descortina amplas possibilidades de cooperação, não apenas no nível técnico, mas ainda no que se refere ao desenvolvimento de indústrias ligadas à produção de equipamentos e de conteúdos digitais interativos, e também no que tange à discussão dos impactos sociais, culturais e políticos da implementação do padrão.

Dentre os fóruns voltados para a utilização das TICs para o desenvolvimento, destaca-se, em âmbito regional, a Estratégia para a Sociedade da Informação na América Latina e no Caribe (eLAC), cujo Plano de Ação prevê, entre outros objetivos, a promoção de ações em áreas como acesso (massificação da infraestrutura de acesso à Internet em banda larga), governo eletrônico (visto como

direito dos cidadãos, devendo ser transacional e participativo), meio ambiente (em especial o uso das TICs para combater a mudança do clima e prevenir desastres naturais, bem como políticas relativas ao lixo eletrônico), seguridade social (incluindo *Telessaúde*); e desenvolvimento produtivo e inovação.

Conclusão: perspectivas e desafios

A diversidade brasileira habilita a política externa do País a dialogar com as distintas realidades encontradas ao redor do mundo. De um lado, por exemplo, unimo-nos aos demais países em desenvolvimento na busca da ampliação do acesso à Internet, ainda concentrado e caro; de outro, somos reconhecidos internacionalmente pela capacidade inovadora e criativa do brasileiro na Internet, exemplificada pela participação ativa do País na rede, sua grande comunidade de desenvolvedores de *softwares* livres, bem como pela informatização de nosso sistema bancário e de nosso processo eleitoral.

Nesse diálogo, experiências exitosas como, entre outras, a do CGI.br e a do Fórum do Sistema Brasileiro de Televisão Digital Terrestre (SBTVD), propiciam ao País diversas possibilidades de cooperação, tanto com países em desenvolvimento quanto com países desenvolvidos. É com base em sua experiência, portanto, que o Brasil participa – por vezes liderando – dos debates fundamentais sobre temas da Sociedade da Informação.

A novidade e o dinamismo desses temas exigem da política externa para a Sociedade da Informação um contínuo aperfeiçoamento, em especial no que se refere ao processo preparatório às reuniões internacionais (nos quais deve ser buscada coordenação interna não apenas com atores governamentais, mas também com a sociedade civil e a iniciativa privada) e ao aprendizado sobre a utilização dessas novas ferramentas na prática diplomática.

"Abertura", a palavra-chave do desenvolvimento da Internet e de sua governança, inspira a prática cotidiana da política externa brasileira para a Sociedade da Informação. E inspira também essa edição do "Congresso Governo Eletrônico e *Software* Livre – CONSEGI", que tem cumprido importante papel na aproximação do governo, a diplomacia incluída, com os temas mais atuais do mundo digital.

Referências

Mais informações podem ser obtidas nos seguintes sítios:

Fórum de Governança da Internet (IGF)
http://www.intgovforum.org

Grupo Consultivo "*Multistakeholder*" do IGF (IGF MAG)
http://www.intgovforum.org/cms/magabout

Fórum WSIS
http://www.itu.int/wsis/implementation/2010/forum/geneva/

Comissão de Ciência e Tecnologia para o Desenvolvimento da Organização das Nações Unidas (CSTD / ONU)
http://www.unctad.org/Templates/Page.asp?intItemID=2696

Corporação da Internet para a Designação de Nomes e Números (ICANN)
http://www.icann.org/

Comitê Consultivo Governamental da ICANN (ICANN GAC)
http://gac.icann.org/

Registro de Endereços da Internet para a América Latina e o Caribe (LACNIC)
http://lacnic.net/sp/index.html

Comitê Gestor da Internet no Brasil (CGI.br)
www.cgi.br

Global Alliance for Information and Communication Technologies and Development (GAID)
http://www.un-gaid.org

UNESCO *Information for All Program* (IFAP)
http://portal.unesco.org/ci/en/ev.php-URL_ID=1627&URL_DO=DO_TOPIC&URL_SECTION=201.html

Estratégia para a Sociedade da Informação na América Latina e no Caribe (eLAC)
http://www.eclac.org/socinfo/elac/

Rede de Informações Pública e Aberta DadosGov - Um modelo aberto para a troca de informações

Ednylton Maria Franzosi[1,3] – ednylton.franzosi@serpro.gov.br
Miriam Chaves[2] – miriam.chaves@planejamento.gov.br
Gilberto Paganotto[1] – gilberto.paganotto@serpro.gov.br
Marcus Vinicius da Costa[1] – marcus-vinicius.costa@serpro.gov.br
Sérgio Assis Rodrigues[3] – sergio@cos.ufrj.br
Jano Moreira de Souza[3] – jano@cos.ufrj.br

Resumo: Este artigo apresenta uma estratégia de abordagem de solução que considera a integração de projetos, em andamento no âmbito de Governo Brasileiro, em uma arquitetura de informações como forma de fortalecer uma política permanente de democratização da informação e melhoria do governo eletrônico, e-Gov. A proposta combina padrões de organização de dados e metadados, apresentados de forma pública e aberta em um catálogo, temático, de informações e uma arquitetura de informações que combina a coleta de dados em sistemas, ad-hoc*, de apoio à decisão, com aquelas providas pelos sistemas estruturantes e corporativos de Governo.*

O trabalho mostra ainda um estudo de caso, DadosGov, utilizando o modelo proposto no Comitê de Organização de Informações da Presidência da República, COI-PR, além de orientações futuras acerca das peculiaridades da proposta.

[1] SERPRO – Serviço Federal de Processamento de Dados – Rio de Janeiro – RJ – Brasil
[2] Ministério do Planejamento
[3] COPPE/UFRJ – Departamento de Ciência da Computação, Universidade Federal do Rio de Janeiro, Brasil

1. Introdução

O governo brasileiro tem realizado esforços no sentido de melhorar a qualidade das informações para a tomada de decisão, no que diz respeito ao planejamento, controle e execução das ações governamentais.

Embora persistam alguns problemas e as iniciativas não sejam uniformes para toda Federação, existe uma nova mentalidade permeando aquelas ações visando tornar mais eficazes os instrumentos de operacionalização das políticas públicas e, principalmente, proporcionar ao cidadão um nível de atendimento mais aprimorado [1].

O maior problema existente e que dificulta consideravelmente a melhoria da gestão pública é a existência de uma grande massa de informações não tratadas e não disseminadas aos tomadores de decisão. O estudo de Fresneda & Goulart [2] que aplicou questionários entre servidores federais perguntando quais são os principais entraves para a melhoria da gestão de conhecimento no Governo apontou:

Quadro 1 - Principais entraves para a melhoria da gestão de conhecimento no Governo

1 - Existência de grande massa de informações não tratadas e não disseminadas aos tomadores de decisões governamentais	81%
2 - Competências individuais e coletivas não exploradas	81%
3 - Insuficiente colaboração intraorganizacional	50%
4 - Baixa utilização de trabalho em grupo, colaborativo e virtual	50%
5 - Dificuldades de promover o aprendizado coletivo, constituindo-se em uma barreira à capacidade de criação e inovação	43%

Ao analisar o Quadro 1, observa-se que as informações para tomada de decisão estão indisponíveis para o uso. Oitenta e um por cento das informações não são disseminadas aos tomadores de decisão. Outros problemas apontados no estudo são derivados do primeiro; sem informação não se exercitam competências, sem a

melhoria das competências não se potencializam os resultados e as equipes envolvidas não apreendem com o coletivo, diminuindo consideravelmente a capacidade de criar e inovar.

2. O legado de conhecimento do Governo

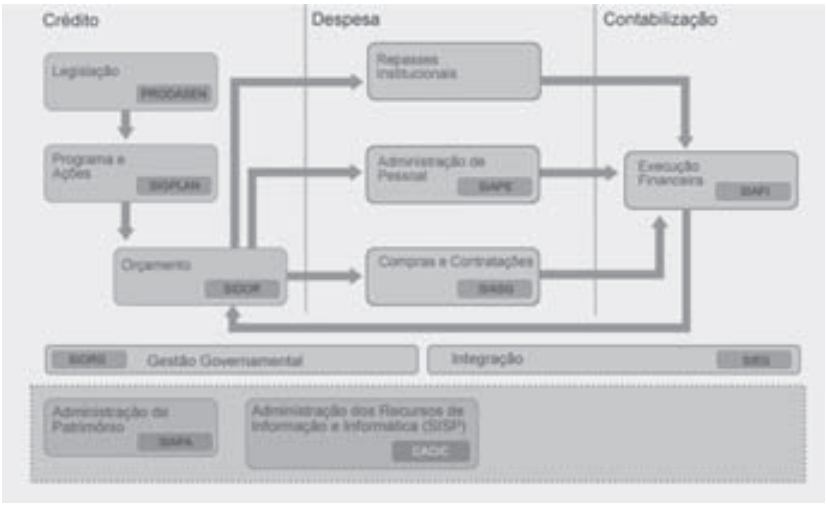
Grande parte das informações para tomada de decisão está nos chamados Sistemas de Gestão Administrativa, SGA, ou Sistemas Estruturantes como são mais conhecidos. Os sistemas de gestão administrativa, SGAs ou Estruturantes, são aqueles sistemas de informação, transversais na estrutura orgânica de Governo que atendem às necessidades finalísticas de mais de um órgão de Governo.

Cada SGA é responsável por um processo de gestão administrativa como:

- Elaboração e Acompanhamento do Orçamento, SIDOR;
- Administração de Recursos Humanos, SIAPE;
- Alocação de Cargos às Estruturas do Governo Federal, SIORG;
- Compras Governamentais, SIASG;
- Acompanhamento dos Programas e Ações do Governo, SIGPLAN;
- Administração Financeira e Contábil do Governo, SIAFI;
- Patrimônio da União, SIAPA entre outros.

O diagrama abaixo mostra a relação entre estes sistemas. A integração mostrada no diagrama é apenas conceitual. Os sistemas estruturantes ainda não estão totalmente integrados. Nos últimos anos diversas melhorias foram sendo implementadas visando a integração plena de suas funcionalidades.

Figura 1 - Ciclo orçamentário-financeiro e os sistemas estruturantes de Governo [3]



Uma lista de pontos focais de melhorias nos Sistemas Estruturantes de Governo é listada abaixo e procura resolver as fragilidades identificadas em trabalho de relatoria desenvolvido pelo SERPRO [4]:

- Melhorar a oferta de informações para apoio à decisão diminuindo a demanda *ad-hoc* de informações para um conjunto não integrado de sistemas gerenciais;
- Incentivar a utilização de padrões, preconizados pela e-PING, para troca de informações entre sistemas de informação;
- Organizar a demanda não estruturada de informação para apoio à decisão por parte de clientes e gestores;
- Melhorar a estruturação dos processos decisórios;
- Atender as necessidades de informações dos órgãos setoriais visando diminuir a proliferação dos subsistemas nestes órgãos, hoje ainda, com baixa integração com os sistemas estruturantes.
- Retornar ao local de origem o dado atualizado nas bases de dados. A coleta de dados realizada pelos sistemas estruturantes de forma centralizada ainda não retorna aos sistemas locais os mesmos dados que ali são necessários;

- Criar mecanismos automáticos de acesso aos dados instituindo aparatos legais que estabeleçam critérios para acesso a dados públicos considerando sua importância política e sigilo inibindo barreiras e facilidades pessoais de acesso;
- Estabelecer política de uso das informações e sua divulgação. Instituir uma documentação automática dos serviços de informações em um catálogo de serviços de Governo que garanta a ampla divulgação do conteúdo semântico das informações disponíveis, as regras para obtenção de informações etc.;
- Implementar arquitetura de informações que forneça um modelo e uma metodologia de aquisição e utilização da informação; e
- Utilizar o catálogo de informações como instrumento para o cruzamento de informações entre os Sistemas de Governo.

Os sistemas corporativos, aqueles desenvolvidos nos âmbitos de Ministérios, Fundações e Autarquias, com objetivo das políticas finalísticas destes órgãos, não foram objeto de estudo deste artigo, mas acredita-se que reúnam grande parte das fragilidades identificadas nos SGAs.

3. Abertura de dados, com qualidade, para a melhoria da gestão pública

A imensa quantidade de código dos Sistemas Estruturantes e Corporativos de Governo, somado à complexidade do negócio que apoiam e as inúmeras diferentes plataformas computacionais que os suportam remetem a diversas tentativas de evolução e integração desses sistemas. Esta dificuldade natural em implementar melhorias apontadas em ambientes complexos incentivou, ao longo de anos, a criação de um Cinturão de Dados, CD para apoio ao processo decisório de Governo. Este CD é constituído quase que integralmente com planilhas e pequenos bancos de dados, preparados *ad hoc* como veremos mais adiante. Os dados dos SGAs, muitas vezes extraídos e consolidados em planilhas, são acrescidos com outras informações declarativas, outros documentos, notícias etc., constituindo uma base importante de informações para tomada de decisão no nível estratégico de Governo.

O presente artigo propõe uma reflexão de como melhorar a utilidade, a disponibilidade e a qualidade das informações de Governo ao conjugar o foco das implementações de melhorias nos SGAs com a utilização dos dados, extraídos do CD, apresentados em formato padrão na *web*. A abordagem considera que atender as áreas estratégicas de Governo, com informações fornecidas e validadas pelos gestores destas áreas em uma Rede de Colaboração que também provê dados abertos à sociedade, atente a quatro expectativas que se potencializam:

- Responder a demanda reprimida de informações gerenciais aos níveis decisórios de Governo;
- Fornecer mecanismos que permitam a avaliação da qualidade, completude, proveniência e governança dos dados, com foco no que é importante em cada um dos sistemas fornecedores, estruturantes e corporativos, de Governo;
- Disponibilizar, a partir das diferentes demandas das diversas áreas do governo, dados governamentais num ambiente que facilite a mineração e análise desses dados com uso de múltiplas ferramentas; e
- Organizar a oferta de Dados Abertos de Governo para a sociedade de forma planejada e organizada.

Estas expectativas estão sendo trabalhadas em dois projetos em execução no âmbito do Governo Federal.

O primeiro, o projeto DadosGov, apesar de não ser um SIG, sistema de informações gerenciais para apoio à decisão, na sua essência, apresenta um catálogo de informações organizado em árvores temáticas com mais de mil séries históricas das ações de Governo atualizadas por gestores de mais de 40 órgãos, o que representa um esforço de integração e padronização dos dados existentes no cinturão de dados.

O segundo, o Modelo Global de Dados, MGD, mapeia os dados registrando a estrutura, a semântica e os processos a partir da visão de negócio em um repositório. A visão de processo, através das regras de transformação, permite rastrear o ciclo de vida do dado, da sua criação aos seus diversos usos e eventual descarte [5]. Assim, o conjunto de processos e metadados poderia vir a apontar a residência daquelas séries históricas apresentadas no Catálogo de Informações do DadosGov.

O MGD se apresenta como instrumento de avaliação da qualidade, completude, proveniência e governança dos dados apresentados no Catálogo de Informações do DadosGov.

A abertura dos dados, iniciativa atual de vários governos, impulsiona o Brasil a tomar medidas, ainda que iniciais, para dispor desta tecnologia. No paradigma dos Dados Abertos o cidadão é parceiro do Governo para a melhoria dos serviços de *e-gov*. A qualidade dos dados ganha dimensão de exigência, já que o exercício da cidadania, apesar de parceiro, é externo e fiscalizador dos governos, diferente da gestão do Estado em que os dados públicos, em uso por gestores, podem apresentar incorreções, toleradas pela capacidade de julgamento dos atores públicos que trabalham com estas informações. A gestão do conhecimento tem como princípio organizacional a utilização de informação, experiência, conhecimento e habilidades existentes nas organizações para gerar valor [7]. Assim, gestores e executivos de Governo, em uma rede de colaboração, podem dispor de elementos que enriquecem o entendimento da informação incentivando o reúso e planejando o seu retorno com valor agregado.

Na medida em que a qualidade dos dados é exigência e o momento é da abertura dos dados, a documentação destes dados, através de bons metadados e proveniência das informações, são imprescindíveis. A união de dados e metadados das séries históricas, acessadas através de ontologia de Governo elaborada a partir da integração das árvores temáticas do DadosGov e o MGD como o repositório destes metadados, das regras que documentem as transformações dos dados desde a origem até a apresentação das séries históricas na *Web*, pode viabilizar uma estratégia diferenciada de abertura dos dados das outras práticas de abertura de dados pelo mundo.

Acessando os sítios do *Open Data* norte-americano (http://www.data.gov), do Reino Unido (http://data.gov.uk), da Alemanha (http://offenedaten.de) e do Canadá (http://www.datadotgc.da) não se tem a possibilidade de acesso aos dados por diferentes árvores temáticas, tão pouco uma abordagem *top-down*, de grandes assuntos até os itens primários de acesso aos dados.

Um grupo de representantes de Governo vem tratando das questões de abertura dos dados:

- Presidência da República com sua experiência em catalogação de informações e o projeto DadosGov desenvolvido pelo SERPRO;
- Secretaria de Logística e Tecnologia da Informação (SLTI) do Ministério do Planejamento Orçamento e Gestão (MP), com a Infraestrutura Nacional de Dados Abertos, INDA, conforme missão institucional de regular a área de Tecnologia da Informação e Comunicação (TIC) no governo;
- SERPRO e DATAPREV como provedores de infraestrutura tecnológica e de soluções;
- outros parceiros como a W3C, Universidades e fornecedores que discutem padrões e protótipos para os dados abertos.

3.1. Arquiteturas de Informação Transitórias

A abordagem proposta, de integração de esforços de projetos que visam a uma melhoria da gestão pública, se torna importante na medida em que a problemática apresentada, de melhoria dos sistemas de Governo, dificilmente é resolvida por um único caminho, de uma única vez e em pouco tempo. É necessária uma estratégia evolutiva de solução considerando arquiteturas de informações transitórias e complementares, planejadas em horizontes de tempo que garantam a geração de ondas de informações para apoio ao processo decisório, única forma de sustentar e justificar a modernidade tecnológica e projetos de modelagem de longo prazo. Não é a tecnologia que justifica a modernidade, mas sim a melhoria constante das informações e do processo decisório sustentado pela inovação tecnológica.

Conforme Murat Erder e Pierre Pureur em *Transitional Architectures for Enterprise Evolution* [8], estas arquiteturas devem considerar platôs de infraestrutura tecnológica e de conteúdo que possam ser disponibilizados a baixos custos e atendendo a demanda reprimida de serviços de informações pelos clientes e usuários dos sistemas informatizados de Governo.

O desafio é organizar gestores para que esta colaboração e modernização dos sistemas estruturantes e corporativos se deem segundo padrão de organização do conhecimento de forma confortável para todos, de modo a provocar a participação daqueles que conhecem os diversos assuntos de Governo e assim aproximar metodologias, vivencias e culturas distintas.

A reconstrução dos sistemas estruturantes de governo em um único sistema seria tarefa praticamente impossível. A taxa média de até três

atualizações diárias dos códigos destes sistemas, somados à complexidade dos assuntos, dinâmica de mudança dos requisitos e a constante alternância dos gestores envolvidos são algumas das características que levariam a um tempo de projeto muito grande a custos exorbitantes com sérios riscos de não atender às necessidades de informação quando finalizado.

A abordagem por aproximações sucessivas com resultados de melhorias evolutivas e parciais é o indicado como abordagem. Assim, o primeiro platô de infraestrutura tecnológica e de conteúdo considera a implementação dos componentes apresentados no quadro 1. Ressalta-se neste platô 1 a utilização de ferramentas que vão permitir acesso automático, sem interferência humana, aos dados nas fontes primárias

Organizar a integração das iniciativas apresentadas e planejar a abertura dos dados, ainda "*not linked*", são ações no contexto do primeiro platô. A questão de *linked* data prevê conhecimento dos domínios de governo. As modelagens deste conhecimento, em ontologias, devem surgir a médio e longo prazo e serão consideradas em platôs adiante planejados. Estes outros platôs subsidiariam a estratégia evolutiva até o platô, objetivo final, da Arquitetura Referencial de Informações, AR, dispondo mecanismos de *software* integradores e configurando uma solução distribuída de serviços, aos moldes da SOA, *Service Oriented Architecture*, prevista na e-PING [9], com procedimentos de acesso e tratamento automáticos, máquina a máquina.

Os Platôs intermediários, identificados como marcos evolutivos do projeto de integração, contariam com o Modelo Global de Dados como fotografias de evolução de cada uma das camadas de conteúdo adicionadas ao longo do projeto.

Os itens seguintes de 4 a 7 apresentam maiores detalhes sobre os projetos DadosGov, MGD e o padrão de organização da dados e qual a relação dos seus componentes com a AR.

4. Arquitetura Referencial de Interoperabilidade, AR

A AR é uma arquitetura de interoperabilidade para a troca de dados, um repositório para o registro de metadados, serviços, gestão do conhecimento do serviço e também um sistema de Gestão de informações resultantes dos serviços [10].

A utilização dos conceitos da AR neste artigo visa organizar os componentes do projeto DadosGov e do Modelo Global de Dados de

forma a facilitar a descrição e compreensão do platô inicial de tecnologia e conteúdo propostos.

A AR define duas visões: visão estratégica e a visão técnica. A visão estratégica apresenta uma abordagem dirigida aos executivos de Governo como forma a facilitar o entendimento da visão SOA pretendida e proposta na e-PING. A visão estratégica compreende três camadas lógicas, conhecidas como i3gov: integração, inteligência e informação. As camadas lógicas da AR representam um modelo conceitual com propósito de facilitar o entendimento e organizar os componentes físicos da arquitetura [11].

A camada de integração compreende os serviços de infraestrutura e os serviços que dão suporte a interoperabilidade da Rede de Desenvolvimento Colaborativo em operação.

A camada de Inteligência é responsável por manter as informações necessárias para a execução e gerência dos serviços *Web* e outros: o padrão de catalogação de dados, os metadados do Catálogo de Informações, a ontologia do Vocabulário Controlado de Governo Eletrônico, as árvores temáticas e o repositório dos componentes do MGD.

A camada de informação oferece séries históricas organizadas em grupos de informações a gestores e sociedade e mecanismos assistidos para a apresentação dos resultados dos serviços.

Interagindo com as camadas lógicas, a AR define atores que se utilizam e gerenciam os componentes físicos da visão técnica da AR.

Figura 2. Visão Estratégica, Atores e Camadas Lógicas da AR

O quadro 2, a seguir, apresenta os componentes físicos da Visão Técnica da AR disponíveis na Rede de Desenvolvimento Colaborativo. Parte destes serviços e funcionalidades já está em operação no platô 1 e outros serviços serão acrescidos no correr do projeto até o platô objetivo final.

Quadro 2 – Visão técnica da AR e componentes físicos da Rede de Desenvolvimento Colaborativo conforme Arquitetura Referencial de Interoperabilidade, AR

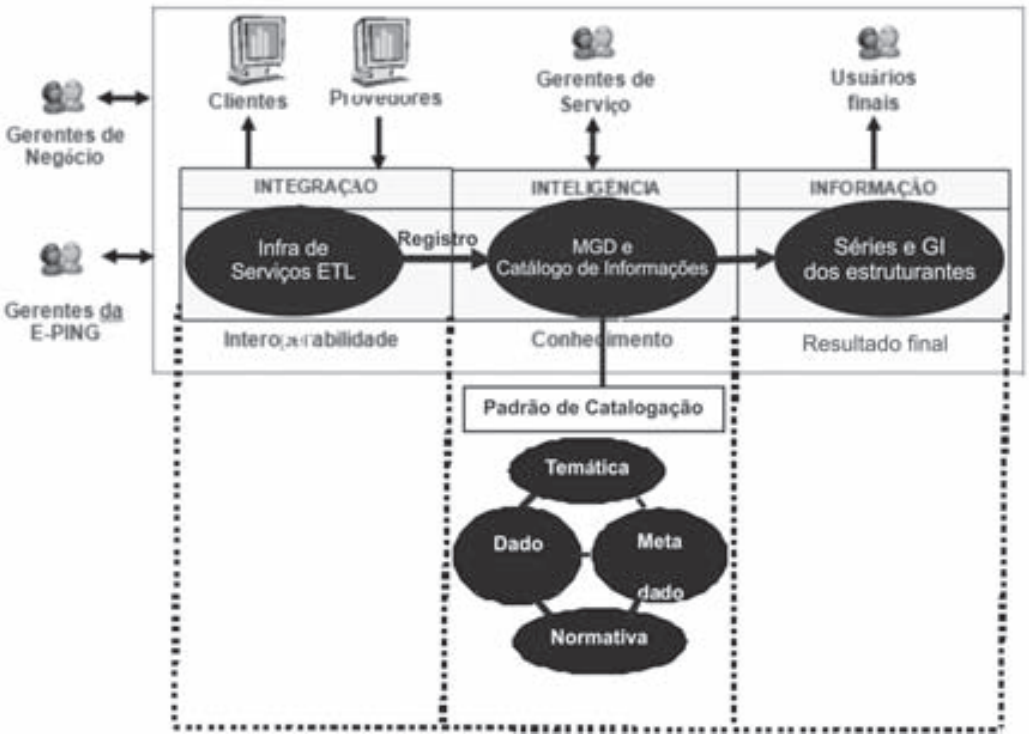
Integração	Inteligência	Informação
Upload dos dados em PHP Atualização dos metadados em PHP ETL, *Talend open source*, de automação origem/destino	Camada de metadados, qualidade dos dados, proveniência dos dados e outros acessórios	Recepção e atualização de dados e metadados dos Grupos de Informação

Bancos de dados *Postgres* *Software* de *data* *mart* para apresentação Ontologia VCGE	Camada de ajuste às regras padrão de catalogação de dados e ao trabalho colaborativo em rede Busca semântica	Organização dos dados das séries históricas no formato padrão de catalogação de informações Publicação das séries históricas em árvores temáticas
Crawler de notícias em *sites* oficiais Editor de texto com acesso aos dados das séries históricas no banco de dados Orquestração dos serviços em BPM	Camada de administração dos contratos, políticas de uso e acompanhamento das atividades automáticas de atualização	Notícias Relacionadas e Fichas de destaques Validação e publicação das séries históricas pelos gestores Busca de dados de serviços *web* em sistemas estruturantes e corporativos
Formatos *linked* data Controle de acesso	Camada de administração da extração dos dados e metadados das séries históricas	Disponibilização de dados para gestores e sociedade
Portal PHP de administração	Camada de administração e capacitação	Gestão dos serviços e Portais Mediação tutorial de especialistas

Componentes físicos da AR podem ser vistos abaixo na figura 3 com destaque para o Catálogo de Informações do projeto DadosGov, o MGD e

o padrão de catalogação de dados apresentados em detalhe nos itens 5, 6 e 7 respectivamente.

Figura 3. i3Gov - Camadas Lógicas e componentes físicos da AR



5. O projeto DadosGov

A Presidência da República é assessorada por gestores, especializados em temáticas diversas de Governo, responsáveis por coletar, analisar e prover informações de ações governamentais ao Gabinete do Presidente.

Até meados de 2006, o cenário era exclusivamente *ad hoc* para a produção deste grande volume de informação gerada nos mais diversos formatos. Os dados coletados sobre as ações de Governo podem ser fornecidos por setores diferentes do mesmo órgão, correndo o risco de não serem os mesmos. Por outro lado, áreas

que atendem a Presidência e os órgãos têm grande demanda de informação.

O Comitê de Organização de Informações da Presidência da República (COI-PR), criado em 2006, vem desenvolvendo ferramentas para facilitar a coleta, o armazenamento, validação e utilização de informações sobre a ação governamental. O projeto DadosGov, iniciado em 2009 (**https://i3gov.planejamento.gov.br/dadosgov/**), tem o objetivo principal de criar um Catálogo de Informações Aberto para aprimorar a gestão pública e facilitar o acompanhamento pela sociedade das ações governamentais em suas comunidades.

Alguns princípios importantes na modelagem do DadosGov precisam ser explicitados, pois viabilizam uma estratégica que pode ser vitoriosa e diferenciada das outras práticas de abertura de dados pelo mundo.

Princípios da modelagem do DadosGov:

- Dispor informações em árvores temáticas

O primeiro princípio na modelagem do DadosGov foi a construção de árvores temáticas, em qualquer quantidade, que referenciam o mesmo objeto, Grupo de Informação (GI), definido como um conjunto de informações de um domínio de negócio. Estes Grupos de Informação são organizados em diferentes ramos de temáticas, construídas livremente por gestores e usuários em seus domínios de conhecimento. As mesmas séries históricas do Grupo de Informação podem se repetir em diferentes formas de organização de modo a respeitar gestores, usuários e cidadãos que assim são e foram acostumados a perceber estas informações.

- Utilizar Grupos de Informação como objeto de aquisição e responsabilização pela qualidade dos dados

O Grupo de Informação é utilizado como canal de comunicação e aquisição dos dados e metadados da fonte gestora desde a primeira até outras atualizações subsequentes para ampliação das séries históricas ou acréscimo de novas. A unidade Grupo de Informação pode se materializar a partir de uma planilha ou vir a ser alimentada por um esquema XML de envio e recepção automática dos dados. Depende da evolução dos platôs tecnológicos e de conteúdo.

A publicação das séries históricas é de responsabilidade do órgão gestor que valida a corretude dos dados na *Web*.

- Abordar requisitos do Geral para o Específico. Do sintético para o analítico

Como forma de disciplinar o acesso aos dados brutos, o glossário temático parte da apresentação de séries históricas em nível agregado com *link* para extração dos dados brutos que deram origem aos totais exibidos. Parece aconselhável esta forma de Catálogo de Informações aos moldes de qualquer catálogo que exista como facilitador na obtenção de detalhes. Não se trata de "engessar" os dados para a sociedade, no caso de dados abertos, mas sim organizar as informações em camadas e tratá-las, desta forma, insumo importante na gestão do conhecimento.

- Organizar o armazenamento dos dados em formato padrão de catalogação

No caso particular do DadosGov, foi estudado e elaborado um padrão de organização de dados que contempla as dimensões Tempo: mês e ano e Local: Brasil, UF e Município. Esta granularidade obedeceu a uma definição do projeto de organizar dados para apoio às decisões estratégicas no âmbito de Governo. Posteriormente, com o crescente interesse por maiores detalhes e abertura dos dados brutos para a sociedade, este padrão se mostrou propício e eficiente para estas finalidades, pois contextualiza em temáticas governamentais dados e metadados além de poder ser estendido para menores granularidades como agora, em fase implementação para os empreendimentos do PAC.

Os Grupos de Informação apresentados na figura 4 estão classificados em árvores temáticas entre outras as do Balanço de Governo 2003-2010, com dois níveis, a dos Programas de Governo (PPA) em um nível, a dos Ministérios provedores dos dados, a do VCGE, Lista de Assuntos de Governo da e-PING com 6 níveis:

Figura 4 – Exemplo de Grupos de Informação disponíveis no DadosGov [12]

- Agentes Comunitários de Saúde
- Biocombustível
- Bolsa Família
- Brasil Alfabetizado

- C&T Arranjo Produtivo Local
- C&T Fundo Setorial
- C&T Instituto de Pesquisa
- Cisternas Construídas
- Crédito Fundiário
- Equipes de Saúde Bucal
- Equipes de Saúde da Família
- Idosos
- Luz Para Todos
- Mulher
- Organização Agrária
- Pessoas com Deficiência
- Programa de Documentação da Mulher Trabalhadora
- Programa Garantia Safra
- ProInfância
- PRONAF
- SAMU
- Seguro da Agricultura Familiar (SEAF)

Cada um destes Grupos de Informação contém um conjunto de séries históricas apresentadas em forma de tabela, gráfico e mapa com início e fim das séries históricas.

Um conjunto de fichas "Destaque", figura 5, pode ser editado no DadosGov com análises quantitativas e qualitativas dos Grupos de Informação. Este texto, pela característica da modelagem do banco de dados, com formato padrão de catalogação de dados, permite que a variável apresentada seja calculada em tempo de publicação, constituindo assim, análises dinâmicas e atualizadas da conjuntura.

Figura 5 – Destaque analítico de Grupo de Informação [12]

O que é o Bolsa Família

Criado em out/03, o programa de transferência de renda atende a 12 milhões de famílias pobres. Os benefícios, com valores entre R$ 22 e máximo de R$ 200 por família, são repassados diretamente às mulheres, por meio de cartão eletrônico pessoal, atendendo às famílias com renda *per capita* de até R$ 140. Seu recebimento é vinculado ao cumprimento de metas de frequência escolar dos filhos e cuidados de saúde. O programa promove a segurança alimentar e nutricional, a conquista da cidadania e o desenvolvimento das potencialidades da parcela mais vulnerável da população.

Benefício do Bolsa Família tem ganho real de 4%

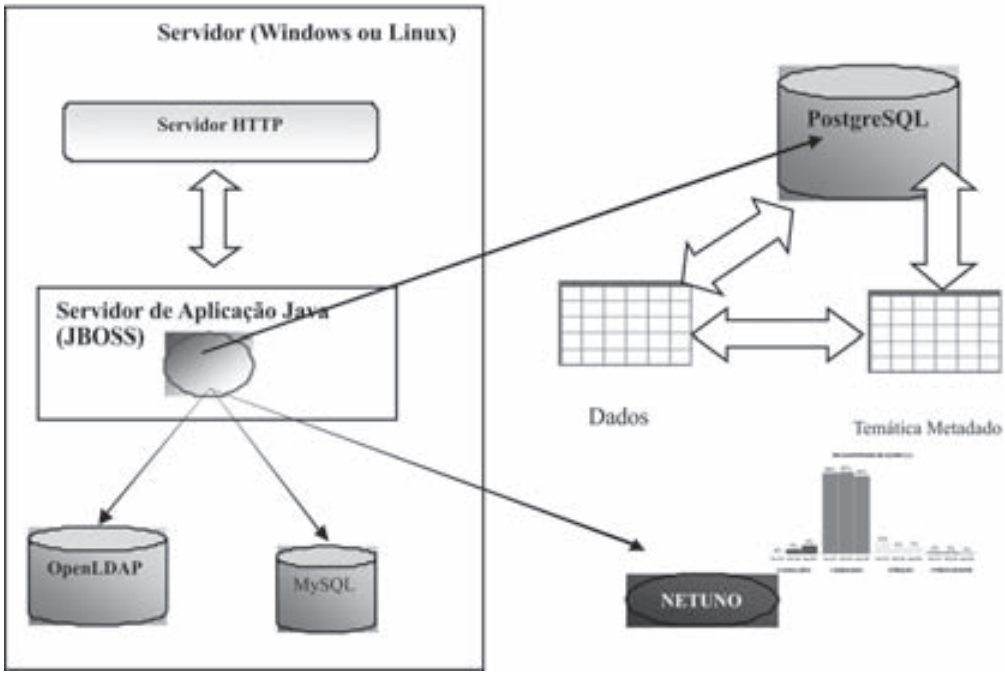
Os benefícios foram reajustados em set/09, pela variação do INPC, passando para R$ 68 o básico e para R$ 22 o variável por criança e R$ 33 o variável por adolescente. Um ganho real de cerca de 4%.

Bolsa Família em Roraima

Em Roraima, em 2008, o programa atingiu 33.622 famílias beneficiadas destinando mais de 37,1 milhões de reais para os beneficiários (...).

Segue a arquitetura do Projeto DadosGov em operação

Figura 6– Arquitetura DadosGov



Ainda são funcionalidades do projeto DadosGov:

- Fichas destaques. Possibilidade de construção de textos analíticos através de editor HTML referenciando séries históricas, períodos e outros metadados, por *tags*, e recuperando valores atualizados nas tuplas do banco de dados. Combinação de texto e dados atualizados;
- RSS de notícias construído através de *crouler*, a *sites* oficiais, parametrizado com sinônimos das temáticas que organizam os dados. Combinação de notícias e dados.
- *Links* das fontes gestora de dados e gráficos diversos.
- Indicadores das ações de Governo nos municípios. O sistema de agenda dos prefeitos disponibiliza os dados periodicamente

através de *download* no formato padrão estabelecido. Funcionalidade em implementação.

- Busca semântica. Esta busca é realizada sobre o repositório de metadados usando uma ontologia do domínio de Serviços Públicos da Administração Federal que foi construída a partir do Vocabulário Controlado do Governo Eletrônico (VCGE).
- Rede de Colaboração instalada no SERPRO composta de 6 servidores em ambiente operacional Linux/Apache/*Postgres*/PHP. A Rede de Colaboração se destina a gestores e usuários que participam da cadeia de atualização dos Grupos de Informação e de outras parcerias com fornecedores, Universidades e desenvolvedores de órgãos diversos que contribuem com a aplicação DadosGov.
- Link RDF para extração de dados brutos máquina a máquina. Em estudo.

5.1. Processo produtivo do DadosGov

A participação de parceiros, universidades e fornecedores no processo produtivo do Projeto DadosGov, através da Rede de Desenvolvimento Colaborativo 189, apresentado na figura 7 abaixo, considera três etapas de produção:

- As atividades dos parceiros em manter atualizados os componentes do DadosGov;
- As atividades do Cliente Central, e Clientes Setoriais, fornecedores de dados, que atualizam os dados e metadados; e
- As rotinas de publicação dos dados e metadados na *Web* para homologação, rede 189 de Desenvolvimento Colaborativo e rede de produção.

Este ciclo produtivo é controlado pelo indicador de situação das séries históricas, conforme o seguinte fluxo:

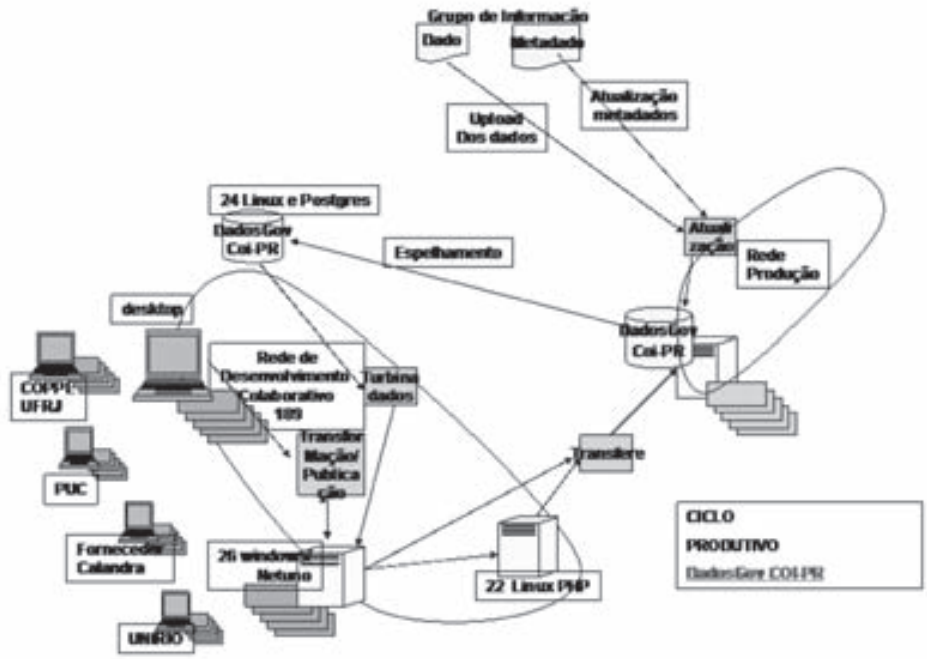
1. O Cliente central do DadosGov solicita aos Clientes Setoriais o envio de dados e metadados de um determinado Grupo de Informações considerado como dados para comporem série

histórica relevante e estratégica para o acompanhamento das ações de Governo. O indicador de situação neste caso é ligado em Dados Solicitados;

2. Os CS enviam planilha, com dados e metadados, via funcionalidade de *upload* para atualização automática do banco de dados do DadosGov;
3. Os dados são armazenados e o indicador de situação é ligado em Dados no Formato Padrão;
4. Rotinas automáticas de extração, tratamento e movimentação de dados da ferramenta Netuno publicam as séries históricas na Rede de Desenvolvimento Colaborativo 189;
5. Valores, metadados e gráficos são visualizados pelos Clientes e homologados;
6. Rotinas automáticas de movimentação de dados transferem as séries históricas da Rede de Desenvolvimento Colaborativo para a Rede de Produção do SERPRO;
7. O CS valida os dados em produção e muda a situação para Dados Validados;
8. O Cliente, autorizando o acesso à série histórica pela sociedade, altera o indicador de situação para Dados Publicados.

Até a situação Dados Validados, as séries históricas só podem ser consultadas em área restrita de administração do DadosGov.

Figura 7. Ciclo produtivo DadosGov



6. Modelo Global de Dados, MGD

Em 2009, o SERPRO criou um projeto intitulado "Projeto de Integração do Macroprocesso de Planejamento, Orçamento e Finanças – PRINT", como necessidade identificada no âmbito do Comitê do Macroprocesso Orçamentário-Financeiro constituído por representantes do Ministério da Fazenda, do Ministério do Planejamento, Orçamento e Gestão e do Serpro e encabeçado pelos Secretários-Executivos dos dois Ministérios [x]. A proposta de trabalho está focada em solucionar a oferta de informações necessárias aos processos decisórios e aprimorar os sistemas que apoiam o Macroprocesso. Segundo o próprio Comitê,

> *os sistemas informatizados de Governo e as demandas por informação vêm se tornando mais urgentes e complexas, as necessidades de informações para apoio à decisão mais freqüentes e novas solicitações de acesso e troca de dados estão se tornando urgentes. [6]*

Os resultados alcançados em relação à atual falta de integração das informações do Macroprocesso e sua incorporação pela Arquitetura e-PING de Interoperabilidade do Governo Brasileiro incorreram em sua expansão para outros domínios, contextos e escopos, além do macroprocesso de Planejamento, Orçamento e Finanças configurando-se um modelo de dados estratégico de todo o Governo servindo de planta mestra de localização e proveniência dos dados utilizados por gestores, interessados em dados abertos à sociedade.

Para criação do MGD foi desenvolvida a Modelagem Essencial de Dados e Processos, metodologia elaborada com o objetivo de realizar um levantamento em alto nível das informações e dos processos de um escopo abrangente de negócios do governo num curto espaço de tempo. A modelagem essencial considera apenas as informações relevantes para o entendimento do negócio, não se atendo aos detalhes operacionais ou tecnológicos [13].

Como resultado dessa modelagem, tem-se um conjunto de artefatos que descreve uma visão geral e integrada, sob o ponto de vista de dados e processos, dos vários negócios mapeados composta por três fases: modelagem de dados para geração do Modelo Global de Dados; refinamento do Modelo Global de Dados; e atualização do Modelo Global de Dados com visão de negócio. Ressalta-se que a metodologia prevê que essas duas últimas etapas possam acontecer em paralelo ou separadamente, de acordo com a complexidade do escopo em estudo.

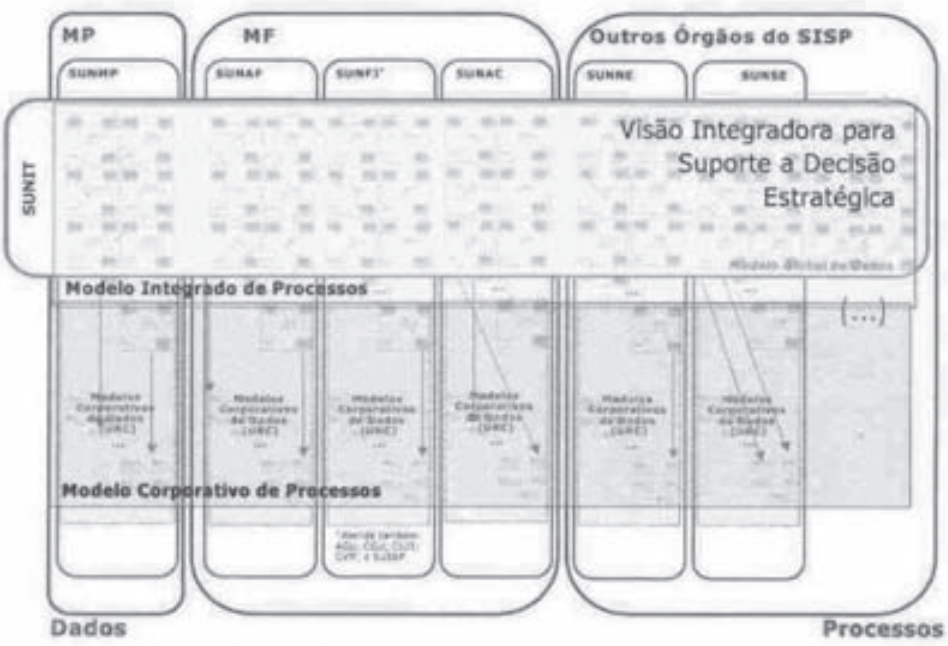
Em cada uma dessas etapas são gerados resultados – ou artefatos – que vão dando visibilidade às partes interessadas quanto aos sistemas, ambientes e contextos, interações, entidades de dados e relacionamentos, semântica, convergência entre as entidades e as tabelas físicas dos sistemas, donos e usuários dos dados e a relação entre as entidades e os processos de negócios, entre outros.

Para tanto, o Modelo Global de Dados disponibiliza ao todo dez artefatos: Diagrama de Contexto, Diagrama de Entidade e Relacionamento, Dicionário Único de Dados, Matriz de Convergência entre as Principais Entidades, Matriz de Entidades *versus* Áreas de Negócios, Diagrama de Processo, Descrição de Atividades, Matriz de Processos *versus* Dados, Matriz de Gestor da Informação *versus* Dados e Registro de Alterações no MGD.

O MGD atua como um barramento transversal de metadados, permitindo que a modernização dos sistemas estruturantes possa ser feita gradualmente com compartilhamento de conceitos e de dados, permitindo ainda, a partir de um mapa de informações de alto nível, a tomada de decisões estratégicas de

governo a partir de uma visão transversal das informações, independente dos sistemas que dão suporte a esses.

Figura 8 - Modelo Global de Dados



Desde o seu processo de criação, o MGD começou a ser utilizado para promover decisões quanto a demandas por atualização, inovação ou integração de sistemas. Entre elas uma das soluções que merece destaque está relacionada à identificação dos possíveis pontos de integração de dados, antes restritos aqueles que detinham o conhecimento da estrutura física dos dados ou responsáveis por realizar extrações, ou Apurações Especiais, AESP.

As AESPs são relatórios feitos sob demanda para fornecer informações de Sistemas. Essa forma de obter a informação onera os Sistemas de tal forma que é necessário fazer o agendamento das execuções de extrações de dados para períodos noturnos, com o intuito de evitar a concorrência com as atividades operacionais. Além disso, os sistemas de tomada de decisão são prejudicados devido à dificuldade de integração das informações no processo

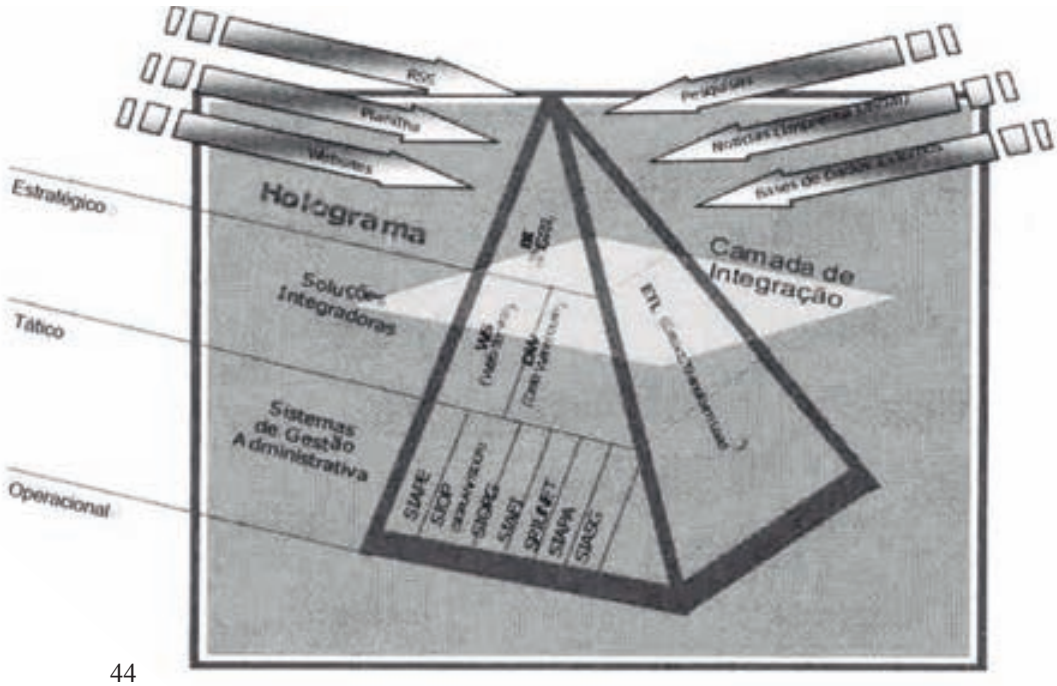
de carga, pois os técnicos responsáveis pelo processo de extração precisam deter o conhecimento das informações no âmbito da sua instituição e das demais que são contempladas nesses sistemas, com informações provenientes de diversos órgãos ou entidades.

A Camada de Integração torna-se fonte de referência para informações estratégicas de governo que proverá aos seus diversos sistemas gerenciais informações de qualidade, de forma ágil e sem redundância, a fim de se evitar o "re-trabalho", garantindo o reúso através:

Da interoperabilidade, respondendo como cada parte do Governo poderá ter "visibilidade" do processo da outra parte, promovendo sua integração, evitando o re-trabalho e gerando sinergia para prestação de serviços de melhor qualidade ao cidadão, a partir do reconhecimento público dos "donos de processos"; e

Da aquisição de conhecimento, considerando que os dados servem aos processos de transformação. Cada parte poderá ter acesso aos dados existentes nos diversos sistemas da Administração Pública Federal, evitando redundâncias, promovendo o reúso e viabilizando a construção de soluções integradas para operação e gestão dos processos, com foco na evolução constante da qualidade [14].

Figura 9. Camada de Integração



À medida que a modelagem de dados se expande no Governo, aumenta a complexidade, visto que não envolve uma intervenção somente nos sistemas e soluções de TI, mas também na forma de trabalho de centenas de pessoas que, no seu dia a dia, atuam nos processos que compõem os vários macroprocessos da Administração Pública Federal.

Para possibilitar controle efetivo de um processo tão amplo e complexo e garantir integridade e perenidade das informações que constituem o MGD garantindo seu uso pleno foi desenvolvido, adicionalmente, um Modelo de Governança e Gestão, ou MGG, que se caracteriza como um conjunto de regras que norteiam a utilização do MGD. Define papéis e responsabilidades dos atores que interagem no processo sejam eles grupos de gestores da informação, grupos de analistas de TI ou consumidores das informações.

Ao disciplinar o processo de utilização do MGD, o MGG promove o alinhamento com as estratégias de TI do Governo Federal, mudando a forma de trabalhar das pessoas, simplificando e conferindo critérios de segurança, de forma que demandas por informações, novas soluções e manutenções de TI, passam a requerer consulta ao MGD requerendo, eventualmente, prospecções e atualizações desse instrumento.

Tais características de sistematização promoveram sua incorporação como padrão, em 2011, à Arquitetura e-PING Padrões de Interoperabilidade do Governo Eletrônico como plataforma para interoperabilidade do Governo.

7. Padrão de Catalogação de Dados

Como os sistemas e a quantidade de informação disponível crescem de forma vertiginosa pessoas e organizações desenvolvem processos de gestão de conhecimento não só acessando informações estruturadas em bancos de dados, mas também em uma categoria de informação conhecida como não estruturada, tais quais textos, planilhas, vídeos, e-mails, multimídia etc. Os dados continuam dispersos em diversos repositórios e os sistemas funcionam em diferentes tecnologias e são administrados por distintos sistemas operacionais.

Combinado a isso, o surgimento da *Web* introduziu uma nova perspectiva de tempo na dinâmica de negócios alinhada a uma nova noção de espaço. Logo, informações são geradas e disponibilizadas quase instantaneamente, organizadas ao modo jornalístico com demanda por respostas mais rápidas aos fatos, por parte de clientes e gerentes de negócio. [15]

Para isso, é necessário que a informação relevante seja recuperada onde estiver independente do formato e das fontes de informação.

O conceito de Grupo de Informação, GI, como um objeto que organiza um conjunto de dados, em qualquer formato, classificado em temáticas de contexto de uso dos usuários se apresenta como elo entre o fornecedor do dado e todos outros usuários do cinturão de dados e da Rede de Desenvolvimento Colaborativo.

A análise de um conjunto grande de planilhas e pequenos bancos de dados utilizados pelo Comitê de Organização de Informações da Presidência da República mostrou que a quase totalidade das informações utilizadas no apoio às ações estratégicas de Governo são apresentadas em três dimensões: a de tempo; ano ou mês/ano, a de localidade geográfica; nacional, estadual ou municipal e a de localidade institucional, um órgão, um empreendimento, ou outra forma institucional de estrutura organizacional. A constatação de que as informações para o apoio à decisão no nível estratégico estavam neste formato possibilitou a estruturação de um conjunto de conceitos e formas denominado de padrão de catalogação de informações para apoio às ações governamentais.

O padrão de catalogação de informações é composto por três componentes técnicos: os dados, os metadados e a temática e um componente administrativo, que trata da política de uso do padrão e dos direitos e deveres de gerentes e provedores de dados.

O componente de dados tem estrutura única com registros de mesmo tamanho composta pelo *id* da variável, o código geográfico do IBGE, o ano ou ano/mês e o valor do dado ou um número *link* para informações em outros formatos tipo texto, foto, vídeo etc.

Os metadados informam a periodicidade, a base territorial, a fonte do dado e outros atributos, enquanto a temática contextualiza em domínios do negócio o Grupo de Informação catalogado.

A figura 10 apresenta os componentes do padrão de catalogação de informações.

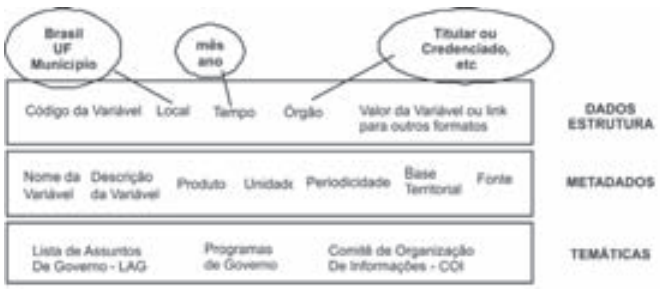


Figura 10 - Padrão de Catalogação de Informações

Os dados e metadados ficam disponíveis em grupos de informação, cabendo aos usuários receber estes dados classificados ou não por uma temática que representa uma visão de negócio catalogada pelo fornecedor responsável pelo dado. Esta opção de classificar ou não os dados brutos segundo um domínio é considerada uma opção no padrão uma vez que o usuário de posse dos Grupos de Informação pode classificá-los à sua maneira.

Os dados coletados e tratados estão organizados em seis estruturas: Brasil/ano, Brasil/mês; UF/ano, UF/mês; Município/ano, Município/mês. Por exemplo, a variável *Valor do PIB* é um dado nacional anual e pertence ao Grupo de Informação Brasil/ano. A variável *Quantidade de Agentes do Programa Bolsa Família* é um dado estadual, mensal e pertence ao Grupo de Informação UF/mês e assim por diante compondo os seis Grupos de Informação.

Outros conjuntos de seis Grupos de Informação ditas operacionais podem ocorrer com a inclusão da dimensão localidade institucional, que situa os dados em uma estrutura de organização da UF ou do Município como é o caso, por exemplo, da lista de empreendimentos do Programa de Aceleração do Crescimento, PAC.

O dado é armazenado na granularidade do seu fornecimento. Qualquer consolidação não faz parte do padrão. Assim, uma informação bruta armazenada por município e mês pertence ao componente Município/mês e não irá ser consolidada nos componente Brasil, UF e Ano. Qualquer consolidação é de responsabilidade dos usuários.

Esta e outras regras de organização dos componentes e questões mais gerais são abordadas: que dados podem ser considerados públicos? Que critérios de privacidade podem ser considerados? Que indicadores podem avaliar o grau de tratamento e de qualidade do dado disponibilizado? Este é realmente o dado bruto? Qual política de uso para um determinado dado bruto? Estas são questões a serem tratadas no componente administrativo do padrão em que fornecedores dos dados e gerentes tutores do padrão se comprometem com a política de uso e garantem o nível de serviço acordado no padrão bem como a qualidade e consistência dos Grupos de Informação fornecidos.

Os metadados por sua vez estão organizados em componentes e acompanham os Grupos de Informações.

O Grupo de Informação, referente às tabelas seguintes, ilustra a apresentação de duas variáveis, valor repassado para saneamento nas escolas e valor repassado para o Programa Água para Todos que se referem à UF, ano e/ou órgão.

Estratégico (visão ano por UF)

Id	CoUF	Ano	Valor
3	16	2005	1075
3	16	2006	1111
3	16	2007	1098
3	16	2008	1139
4	16	2005	527429
4	16	2006	588451
4	16	2007	588767
4	16	2008	585652

Tabela 1. Arquivo “tripado” estratégico

Operacional, visão ano por órgãos da UF

Id	CoUF	Ano	Órgão	Valor
3	16	2005	10101	75
3	16	2005	10102	1000
3	16	2006	10101	111
3	16	2006	10102	1000
3	16	2007	10101	98
3	16	2007	10102	1000
3	16	2008	10101	138
3	16	2008	10102	1001
4	16	2005	10101	527000
4	16	2005	10102	429

Tabela 2. arquivo “tripado” operacional

Ocorrência de valores das variáveis *id* 3 e *id* 4 (valores fictícios) com:

Id	Variável	Descrição	Produto	Unidade	Periodicidade	Base Territorial	Fonte
3	Saneamento nas escolas	Ações que visam a melhoria...	Estado atendido	Unidade	Anual	Território da Cidadania	MIN CODEVASF
4	Água para todos	Revitalização do São Francisco	Família atendida	Família	Anual	Território da Cidadania	MIN CODEVASF

Tabela 3. Metadados

Os seguintes metadados classificados na:

Id	Área	Tema	Subtema
3	Saúde	Água para todos	Saneamento e acesso a água
4	Saúde	Água para todos	Revitalização do são Francisco

Tabela 4. Temática saúde. Temática DadosGov.

8. Conclusão

Fornecer dados de Governo para a sociedade organizada, de forma padrão, credenciado por entidade reconhecida, é a maneira de estabelecer uma política permanente de democratização da informação onde sua veiculação se dá de forma compartilhada, correta, tempestiva e com qualidade. Combinar padrões de organização de dados catalogados de forma pública e massiva em um catálogo de informações e de serviços de informações aberto e livre é melhorar a gestão pública e incentivar a sociedade civil organizada a conhecer e acompanhar as ações de melhoria da qualidade de vida de suas comunidades.

Os esforços atualmente empreendidos no âmbito do Governo Federal ganham visibilidade e provocam a adesão da sociedade civil, interessada na transparência do Estado brasileiro. Estes esforços devem ser vistos como concomitantes, porém não como ações isoladas ou desintegradas, garantindo a obtenção de resultados em tempos diferentes e de forma escalável, interoperando às soluções à medida que forem sendo implantadas.

Este paradigma baseado em proposição do consórcio W3C, de *Open Data*: disponibilizar dados de Governo através de componente padrão de catalogação de informações pode revolucionar a forma de exercitar a cidadania se apoiada em uma arquitetura referencial de interoperabilidade orientada a informações e serviços.

Referências

[1] PEREIRA P. C. (2009) Tecnologia da Informação a serviço do Governo, http://www.profissionaisti.com.br/

[2] FRESNEDA, P. S. V., Goulart S. M. G., Papa M. e Fonseca A. F. (2006) *Diagnóstico da Gestão do Conhecimento nas Organizações Públicas utilizando o método Organizational Knowledge Assessment (OKA).*

[3] i3Gov (2008) "i3Gov: Integração e Inteligência em Informações de Governo". Disponível em http://i3gov.softwarepublico.gov.br/wikigestor, acesso em Março/2010.

[4] SERPRO (2009) Subgrupo Mapeamento e Modelagem de Dados, Relatoria SLTI e SERPRO.

[5] SLTI/MP, (2010) A integração de Dados no Âmbito do macroprocesso de Planejamento, Orçamento e Finanças, Panorama da Interoperabilidade no Brasil.

[6] HOLANDA, Victor Branco e outros (2010). *Sistema de Informação de Custos na Administração Pública Federal: uma política de Estado,* FGV, Rio de Janeiro.

[7] OMG (2003) MDA Guide Version 1.0.1. Disponível em http://www.omg.org/docs/omg/03-06-01.pdf

[8] ERDER, M. e PUREUR P., (2006) *Transitional Architectures for Enterprise Evolution*, IT Pro May.

[9] e-PING (2008) "e-PING: Padrões de Interoperabilidade de Governo Eletrônico", Comitê Executivo de Governo Eletrônico. Versão 4.0, Dez. 2008. Disponível em www.eping.e.gov.br, Acesso em Março/2011.

[10] FRANZOSI, E., FALCÃO, V. (2007), "AR - Arquitetura Referencial", *In: Tecnologia da Informação e da Comunicação: a busca de uma visão ampla e estruturada*, Pearson Prentice Hall: Fundação Getúlio Vargas, pp. 224-230.

[11] FRANZOSI, E.M, GARCIA, A., RODRIGUES, S.A, BLASCHEK, J.R, SOUZA, J.M (2009) *Uma Proposta de Arquitetura Referencial SOA para Desenvolvimento de Sistemas para o Governo*, WCGE/SBC/2009, Bento Gonçalves, Brasil.

[12] CHAVES M., MAYA T., POJO R., S. M. CUGAN O. J, SOUZA D. D., COSTA L. J. (2009) Comitê de Organização de Informações.

[13] SERPRO (2009) MGD – Modelo Conceitual Disponível em http://modeloglobaldados.serpro.gov.br/.

[14] COSTA, Marcus Vinicius (2010) *O desafio da Interoperabilidade para a Carta de Serviços ao Cidadão: como integrar processos com processos, processos com dados e dados com dados.*

[15] Calandra Soluções (2008) *Gestão da Informação em Contextos.*

Segurança Eletrônica em *Web* Semântica

Jose Maria Leocadio[1]

Resumo

A arquitetura da *Web* Semântica pode ser representada como uma pilha de camadas em que características de segurança permeiam todas as camadas criando uma segurança em profundidade. Atualmente os padrões de segurança que estão com maior maturidade de desenvolvimento são os relativos às camadas URI, XML e RDF. O desafio de se escolher padrões de segurança está em garantir que a interoperabilidade de dados não seja afetada pelas características de segurança obrigatórias. Estas características de segurança podem tornar dados inacessíveis para tratamento por máquinas, ou mesmo para visualização de pessoas, trazendo impactos negativos a um dos princípios fundamentais da *Web* Semântica que é a possibilidade de se gerar conhecimento a partir de dados vinculados (*linked data*).

1. Introdução a *Web* Semântica

A *Web* Semântica é a nova geração da arquitetura *Web* e tem como objetivo principal criar uma representação de conhecimento a partir

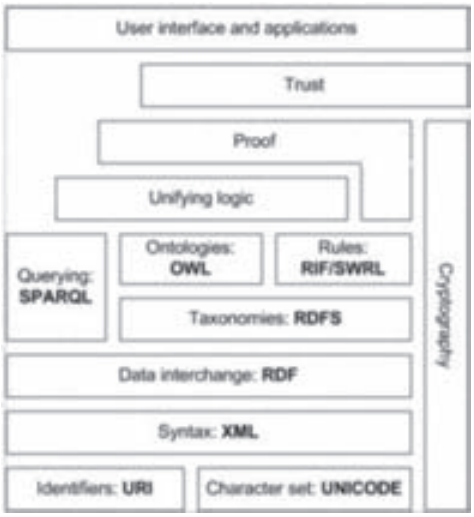
[1] Coordenação Estratégica de Tecnologia CETGC/ SERPRO.

de dados vinculados (*linked data*) que permitam o processamento por máquinas em escala. A *Web* Semântica também é conhecida como *Web* dos dados e nesse sentido duas tecnologias se apresentam como catalisadoras: *eXtensible Markup Language* (XML) e *Resource Description Framework* (RDF). Resumidamente, a XML trata da sintaxe enquanto o RDF trata da semântica dos dados.

Uma visão em camadas auxilia a identificar onde padrões de segurança podem contribuir para a segurança em *Web* Semântica. A Figura 1 ilustra a pilha de camadas para a *Web* Semântica.

A definição de *Web* Semântica – disponível no sítio do W3C[2] – a ilustra como um *framework* que permite a interoperabilidade de dados entre aplicações, empresas e comunidades: "*The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries. It is a collaborative effort led by W3C. It is based on the Resource Description Framework (RDF).*" Estas características determinam que máquinas podem entender de forma automatizada o significado dos dados facilitando a interoperabilidade.

Figura 1 – Pilha de camadas da *Web Semântica*



Fonte: http://www.w3.org/2007/03/layerCake-small.png

[2] www.w3.org/2001/sw

Quando os dados disponíveis na *Web* são vinculados é possível, a partir de um determinado dado, encontrar outro. Diferentemente da *Web* de hipertextos com referências a documentos escritos em HTML tem-se referência a documentos escritos em RDF. Em um artigo sobre *linked data* escrito por Tim Bernes-Lee, em 2010, ele propôs um *ranking* com cinco gradações que serve de guia para classificar dados que estão publicados[3]:

★	Dado publicado, em qualquer formato, com licenciamento livre
★★	Dado publicado em uma estrutura que permita a leitura por máquinas (por exemplo, uma planilha eletrônica ao invés de uma imagem digitalizada)
★★★	Igual ao item anterior (duas estrelas) adicionado o formato não proprietário (por exemplo, CSV ao invés de Excel)
★★★★	Todos os acima, porém usando os padrões W3C (RDF e SPARQL)
★★★★★	Todos os acima adicionando a vinculação de seus dados aos dados de outras fontes para identificação de contexto (*linked data*).

2. Segurança para *Web* Semântica

Não há dúvidas de que a palavra "segurança" tenha uma grande sobrecarga de definições o que é um terreno fértil para o aparecimento de listas de propriedades de segurança, não exaustivas, heuristicamente identificadas a partir de experimentos em um determinado contexto ou, até mesmo, a partir da intuição de indivíduos. O que normalmente se vê na literatura são conjuntos de requisitos desejáveis, tais como confidencialidade, integridade e autenticidade. Shafi Goldwasser e Silvio Micali propuseram uma definição formal de segurança, restrita ao âmbito da criptografia de chaves públicas, relacionada à indistinguibilidade polinomial no tempo e à segurança semântica. A **segurança semântica** não deve ser confundida com a segurança em *Web* Semântica e a intuição

[3] http://www.w3.org/DesignIssues/LinkedData.html

por trás da definição formal é que um agente malicioso não obtém nenhuma informação adicional a partir da observação de um texto que foi criptografado[4].

Existem três grandes áreas relativas à segurança na literatura que são importantes para o foco de estudo deste artigo: ataques, mecanismos de defesa e serviços de defesa[5].

Ataques: qualquer ação que comprometa a confiabilidade de uma determinada informação. Há dois tipos principais, os ataques passivos que são aqueles em que o ataque é feito apenas observando-se a informação que está sendo transmitida e os ataques ativos em que o ataque envolve alterações na informação que está sendo transmitida.

Mecanismos de defesa: quaisquer processos para se detectar um ataque; prevenir-se de um ou recuperar-se após outro. A Tabela 1 traz um resumo da recomendação X.800 para os mecanismos de defesa dividindo-os em mecanismos de defesa específicos para determinada camada do protocolo *Open Systems Interconnection* (OSI) e em mecanismos de defesa *pervasivos* por mais de uma camada.

Tabela 1 – Mecanismos de defesa do padrão X.800

MECANISMOS DE DEFESA ESPECÍFICOS
Ciframento: é o mecanismo de transformação da informação, via criptografia, de modo a se produzir um objeto diferente do objeto original.
Assinatura digital: é o mecanismo que permite ao ator receptor da informação verificar a identidade do ator emissário da informação.
Controle de acesso: é o mecanismo que previne o uso não autorizado de um determinado recurso incluindo a prevenção do seu uso de maneira inadequada.

[4] ANDERSON, R. *Security Engineering. A Guide to Building Dependable Distributed Systems. Second Edition*. Wiley Publishing, Inc. 2008.

[5] *Security architecture for Open Systems Interconnection (OSI) for CCITT applications. Recommendation X.800*. Geneva. 1991.

Integridade da informação: é o mecanismo usado para garantir que a informação não foi alterada ou destruída de maneira não autorizada.
Troca de autenticação: é o mecanismo usado para garantir a identidade de um ator por meio da troca de informações entre atores.
Preenchimento de tráfego: é o mecanismo de inserção de *bits* em trechos que apareçam no fluxo da informação de modo a evitar-se tentativa de análise de tráfego.
Controle de roteamento: é o mecanismo usado para permitir que determinadas informações classificadas possam ter diferentes rotas especialmente quando se desconfia de falha na segurança em alguma rota.
Cartório: é o apoio em uma terceira parte confiável para a garantia da acurácia de certas características da informação.
MECANISMOS DE DEFESA PERVASIVOS
Funcionalidade confiável: é o mecanismo que permite que uma funcionalidade seja percebida como correta, por exemplo, de acordo com a política de segurança.
Selo de segurança: é o mecanismo de aplicação de uma marca ou selo a um determinado recurso que tipifica os atributos de segurança dele.
Detecção de evento: é o mecanismo disponível para detecção de eventos relevantes de segurança.
Trilha de auditoria: é o mecanismo para se coletar informação de modo a permitir uma inspeção em registros e atividades a ser realizada por um ator independente segundo uma política de segurança.
Recuperação de segurança: é o mecanismo que lida com requisições de atores e gera ações de recuperação.

Fonte: *Security architecture for Open Systems Interconnection (OSI) for CCITT applications. Recommendation X.800*. Geneva. 1991.

Serviços de defesa: quaisquer serviços que são providos por um sistema para dar um determinado tipo de proteção aos recursos do sistema, estes serviços usam políticas de segurança e são implementados por mecanismos de segurança. A Tabela 2 traz um resumo da recomendação X.800 para os serviços de defesa.

Tabela 2 – Serviços de defesa do padrão X.800

Autenticidade	Serviço para garantia de que os atores no processo de comunicação são, de fato, aqueles que eles dizem ser.
Confidencialidade	Serviço para garantia de que a informação não ficará disponível ou revelada para atores não autorizados.
Integridade	Serviço para garantia de que a informação recebida é a mesma que foi enviada.
Não repúdio	Serviço de proteção contra a negação de participação, por qualquer ator, em um processo de troca de informações.
Controle de acesso	Serviço de proteção contra o acesso não autorizado a um determinado recurso.

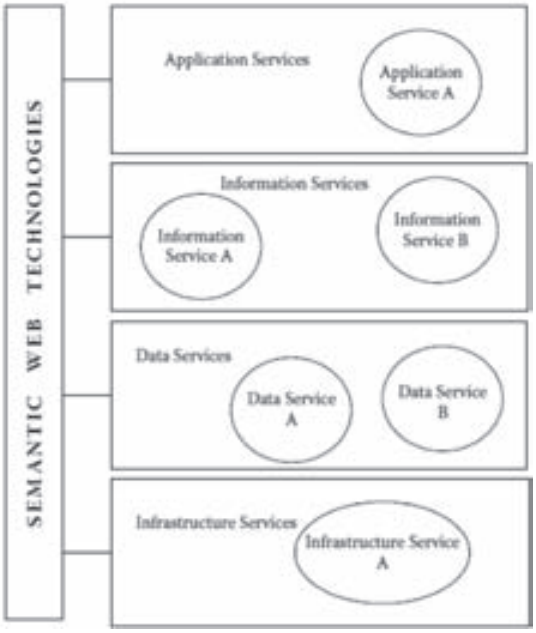
Fonte: *Security architecture for Open Systems Interconnection (OSI) for CCITT applications. Recommendation X.800*. Geneva. 1991.

Para várias aplicações disponíveis na rede Internet no formato de acesso Web, a arquitetura funcional da *Web* Semântica[6] é composta por serviços de infraestrutura, serviços de dados, serviços de informação e serviços de aplicação, conforme ilustrado na Figura 2. Para algumas soluções pode ser exigido que a *Web* Semântica tenha que apresentar níveis de segurança elevados, principalmente naquelas aplicações conhecidas como C4ISR (*Command, Control, Communications, Computers, Intelligence, Surveillance and Reconnaissance* – Comando, Controle, Comunicações, Computadores, Inteligência, Vigilância e Reconhecimento[7]).

A pilha que forma a *Web* Semântica é ilustrada nas Figuras 1 e 3 e constitui-se de 5 camadas horizontais e 1 camada vertical. Para o foco deste estudo, o item segurança será medido separadamente em cada camada, assim se uma camada apresentar uma falha de segurança ainda haverá outras camadas para proteção de toda a pilha (princípio da Segurança em Profundidade). Atualmente somente as 3 primeiras camadas têm mecanismos de segurança maduros.

[6] THURAISINGHAM, B. *Building Trustworty Semantic Webs*. Auerbach Publications. 2008.
[7] http://www.defesabr.com/glossario.htm

Figura 2 – Arquitetura funcional para Web semântica

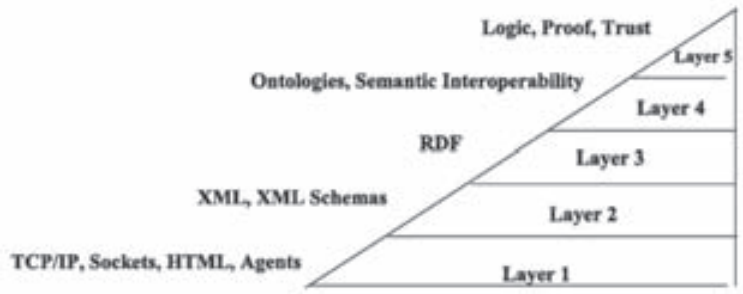


Fonte: THURAISINGHAM, B. *Building Trustworty Semantic Web*. Auerbach Publications. 2008.

A camada mais baixa (URI) é a mais importante para o transporte de dados sendo os mecanismos de segurança aplicados em TCP/IP seguro, HTTP seguro e *socket* seguro. As próximas camadas são as que tratam XML e RDF as quais são usadas em *web services* e *web services* semânticos. Os mecanismos de segurança incluem XML assinado, criptografia do XML, gerenciamento do controle de acesso no XML e classificação de partes de um documento RDF sobre determinados contextos. Atualmente as camadas superiores ilustradas na Figura 3 (Camadas 4 e 5) dependem dos mecanismos de segurança das camadas de baixo.

A camada de Criptografia permeia as outras camadas e fornece, para as camadas horizontais, mecanismos de criptografia para garantia da segurança e privacidade.

Figura 3 – Agrupamento de camadas da Web Semântica em função da segurança



Fonte: THURAISINGHAM, B. *Security Standards for the Semantic Web*. 2004.

Camada URI

A segurança desta camada permeia o item transporte de dados e é considerada uma camada já madura em termos de segurança pois tem padrões já estabelecidos tal como o HTTPS.

Camada XML e RDF

Documentos XML são grafos e os padrões de segurança existentes se aplicam a controle de acesso, autenticação, assinatura digital e criptografia. Já documentos RDF são basicamente documentos XML em que se define o significado do conteúdo do documento, assim os padrões de segurança válidos para XML também se aplicam para RDF. A partir do entendimento do significado, pode-se tornar viável a classificação da informação de partes do documento RDF, em um determinado contexto, considerando-se papéis e restrições como, por exemplo, em esquemas *Role-Based Access Control* (RBAC).

Camada Ontologia e Interoperabilidade semântica

Não há mecanismos de segurança específicos para esta camada, no entanto, para o foco deste artigo, alguns itens de segurança ainda não padronizados são ilustrados abaixo, baseados na proposta de padrões de segurança para Ontologia descritas por Fratemi *et al*[8]:

[8] FRATEMI, M. *et al*. *A proposal for Ontology Security Standards*. 2008.

- Como controlar quem pode acessar determinadas seções de documentos?
- Quais *queries* devem ser respondidas e com que profundidade?
- Como se garantir a identidade de quem explora as ontologias?

Camada Lógica, Prova e Confiabilidade

A camada Lógica é parcialmente disponível nas camadas inferiores quando se fala de papéis e restrições. As camadas Prova e Confiabilidade são camadas da *Web* Semântica que ainda estão em padronização. Para o foco deste artigo, alguns itens de segurança ainda não completamente padronizados são ilustrados abaixo:

- Como confiar na informação?
- Quanto confiar na informação?
- Que padrões devem ser negociados com outras partes durante a comunicação?

3. Conclusão

A *Web* Semântica é construída de forma participativa como citado no próprio sítio do W3C. A pilha de camadas que a compõe possui níveis de segurança diferenciados e aproveita-se construtivamente do conceito de segurança em profundidade. Pode-se perceber que atualmente existem diferentes níveis de maturidade entre as camadas da *Web* Semântica. As camadas inferiores (URI, RDF e XML) apresentam um nível de padronização disponível maior e mais abrangente do que as camadas superiores (Ontologias e Lógica, Prova e Confiabilidade). *Frameworks* de segurança complementares estão sendo pesquisados mundialmente e a ideia é que estes possam futuramente fazer parte da *Web* Semântica conforme a construção colaborativa do W3C.

A Disseminação de Dados Governamentais como Serviço Público – Os Dados Abertos Governamentais e a Experiência Brasileira

Christian Moryah Contiero Miranda[1]
christian.miranda@planejamento.gov.br

1. Introdução

O Governo Eletrônico brasileiro, instituído no ano 2000, tem uma série de princípios e diretrizes dos quais dois merecem destaque: "A Prioridade do Governo Eletrônico é a promoção da cidadania" e "A Inclusão Digital é indissociável do Governo Eletrônico". Essas premissas visam a uma mudança da relação do cidadão com o Estado, preocupando-se com a participação do cidadão e o controle social das ações governamentais, reduzindo a assimetria de informações entre o governo e a sociedade.

A colaboração dos cidadãos, através da inteligência e conhecimento coletivo, proporciona o desenvolvimento de serviços inovadores para a população a baixos custos, beneficiando diretamente o governo. Esse modelo de mobilização vem sendo cunhado pelo termo *Crowdsourcing*.

Uma das formas de implementar essa mudança é a publicação de dados e informações mantidos nos sistemas do governo. Assim, o cidadão tem acesso a informações para o exercício efetivo do controle social sobre o Estado.

[1] Ministério do Planejamento, Orçamento e Gestão.

Com a publicação do Decreto 6.932/2009, o governo atribuiu aos órgãos da Administração pública federal a obrigação de publicar informações sobre os serviços prestados, bem como forma de atendimento e forma de prestação desses serviços, instituindo a carta de serviço ao cidadão.

"A Carta de Serviços é um documento elaborado por uma organização pública que visa informar aos cidadãos quais os serviços prestados por ela, como acessar e obter esses serviços e quais são os compromissos de atendimento estabelecidos." [Ministério do Planejamento, Orçamento e Gestão, 2008].

Existe no contexto internacional um conjunto de padrões para a publicação de dados de governo, esses padrões atendem a todas as expectativas das diretrizes de governo eletrônico, da carta de serviços ao cidadão e ainda vão além com os chamados Dados Abertos Governamentais.

1.1. O problema e a solução

O Brasil já possui diversos casos práticos de publicação de dados na web: IBGE, Datasus, IPEA, INEP, Portal da Transparência, e muitos outros. Só que esses casos não estão aderentes às premissas dos dados abertos e da reusabilidade dos dados publicados, além de cada um fazer suas publicações em formatos e padrões diferentes.

Temos adiante duas indagações:

1. Como os órgãos da Administração Pública podem prover seus dados como um serviço à sociedade?
2. Os Dados Abertos Governamentais são a metodologia que vai atender às necessidades e expectativas na aproximação entre ele e cidadão?

Pretendemos analisar o contexto que favoreceu a concretização do movimento no âmbito internacional, absorver as lições aprendidas nas experiências existentes e responder a essas duas perguntas.

2. Analisando nosso universo

Nosso escopo é o contexto político e sociocultural dos casos de dados abertos existentes. Portanto as tecnologias que dão suporte a esses

casos não serão estudadas, limitando-se o trabalho a declarar que essas tecnologias e padrões de mercado estão consolidados e já atendem as necessidades dos publicadores.

A interoperabilidade, principal característica visada pela e-Ping e resultado natural da publicação de dados abertos, pode ser definida de forma simplificada como a habilidade de sistemas distintos trabalharem em conjunto, interoperar. Cabe lembrar que sistema não se limita à ideia de sistema de informação. Sistema é de uma forma geral um todo composto por partes. Podemos então nos referir a sistema como: sistemas de informação, comunidades civis, sociedade em geral, administração pública e governos de outros países. Daí a visão da interoperabilidade governo-cidadão, também citada na plataforma de colaboração, mais adiante.

Outro conceito importante é a interoperabilidade semântica. A interoperabilidade semântica, descendo ao escopo dos sistemas de informação, é a habilidade de dois sistemas se comunicarem, tendo entre eles a mesma interpretação e, portanto, as mesmas implicações lógicas sobre as informações trocadas. A importância disso é que o movimento dos Dados Abertos Governamentais pretende, num cenário presente ou posterior, a utilização de padrões de arquivo ou de comunicação que promovam interoperabilidade semântica entre sistemas, permitindo interpretações automatizadas por máquina e agregação de valor às informações apresentadas ao cidadão.

O livro Open Government – Collaboration, Transparency Participation in Practice (Lathrop, D. e Ruma, L., 2010) provê uma questão anterior à simples publicação dos dados abertos: o Governo Aberto no contexto americano. Basicamente, o conceito Governo Aberto defende a ideia de que o governo deveria ser totalmente transparente para poder ser acompanhado pelo cidadão. Essa visão, originária do iluminismo europeu, tem como principal ponto de interseção com os princípios do governo eletrônico, a participação e o controle social.

Segundo o livro (Lathrop, D. e Ruma, L.,2010), o governo deve prover ao cidadão uma Plataforma de Colaboração baseada em serviços e informações para que o cidadão possa, a partir dela, contribuir com o governo da maneira que puder ou desejar.

Entre os desenvolvimentos mais recentes (do conceito, e não do livro) está a teoria da governança do código aberto, que defende a aplicação das

filosofias do movimento do software livre aos princípios democráticos para permitir que os cidadãos mais interessados se envolvam mais diretamente com o processo legislativo.

2.1. Conceito de Dados Abertos Governamentais

Segundo o World Wide Web Consortium (W3C Brasil, 2009),

> *Dados Abertos Governamentais são a publicação e disseminação das informações do setor público na Web, compartilhadas em formato bruto e aberto, compreensíveis logicamente, de modo a permitir sua reutilização em aplicações digitais desenvolvidas pela sociedade.*

Em síntese, a intenção é publicar dados de governo na web, da maneira mais granular possível, para que a sociedade possa aproveitá-los da maneira que desejar, trazendo benefícios também para o governo.

2.2. O Movimento Dados Abertos Governamentais

O movimento dos Dados Abertos Governamentais diz respeito às iniciativas e aos atores que participam dos acontecimentos relativos à publicação de dados abertos, sendo os atores divididos em entusiastas, ativistas, hackers e atores de governo.

Ações formais e coordenadas de governo foram iniciadas em 2009, com os portais data.gov e data.gov.uk. Ações ativistas da sociedade organizada são difíceis de rastrear, sendo a mais antiga encontrada datada em 2007: a publicação dos oito princípios dos Dados Abertos Governamentais demonstrada mais adiante.

O conceito dos Dados Abertos Governamentais vem sendo promovido pela sociedade organizada, pelos governos e pelo W3C, entidade envolvida na padronização e provimento de tecnologias para atender suas necessidades.

2.3. Dados Abertos e a e-Ping

Antes do Decreto 6932/2009, o governo criou a e-Ping, Padrões de Interoperabilidade de Governo Eletrônico, um conjunto de políticas

e padrões abertos para intercomunicação entre entidades de governo e entre governo e cidadão. Teve sua primeira versão publicada em 2004, e já proveria insumos para o atendimento das necessidades do governo eletrônico.

A e-Ping prevê em seu sítio oficial o estabelecimento das condições de interação entre governos e com a sociedade. Frente ao grande sucesso observado nos casos internacionais em conjunção à já existente pressão pela sociedade brasileira para a publicação de dados em formatos reutilizáveis, o Governo Federal criou uma Política de Disseminação de Dados e Informações em 2010, visando estimular e coordenar ações de publicação de dados.

Após a criação da política, foi iniciado um subgrupo de trabalho na e-Ping e definida uma equipe para dar início às pesquisas e desenvolvimento das ações previstas. Trabalhos já realizados pela equipe:

- Apoio ao projeto piloto da Dataprev para publicação de dados abertos da Previdência, estimando a publicação de no mínimo um Dataset até o final de 2010.
- Criação de um roteiro que orienta a publicação de dados em formato aberto, orientando o gestor público sobre quais as atividades gerenciais a serem feitas antes do procedimento técnico de publicação.
- Concepção de uma Infraestrutura distribuída para acesso centralizado aos dados abertos publicados pela Administração Pública.
- Em parceria ao W3C, provimento de um curso técnico sobre publicação de dados abertos no padrão Linked Data e realização de uma oficina para apresentação e discussão do Roteiro de publicação de Dados Abertos.
- Publicação piloto de dados sobre compras governamentais.
- Adoção de padrões para publicação de dados abertos de forma aderente à Web Semântica na e-Ping 2011: RDF e OWL.

2.4. Breve histórico sobre os Dados Abertos

As primeiras iniciativas conhecidas por parte do setor público são os portais de dados abertos “data.gov” e “data.gov.uk”, portais onde

são indexadas milhares de séries de dados em formato bruto para utilização da sociedade, lançados em Maio de 2009 e Setembro de 2009, respectivamente.

Em seguida, Canadá, Austrália e Nova Zelândia começaram a publicar dados seguindo as diretrizes do Open Government Data nos mesmos moldes do Reino Unido e Estados Unidos, com portais oficiais de governo.

No caso da sociedade ativista, a pressão começou muito antes:

> *Groups from the Guardian to the Sunlight Foundation had already been pushing for it for a long time. People like Watchdog.net, mysociety.org, and govtrack.us had been pushing by publishing government data themselves in various formats, including Linked Data. (Berners Lee, T., 2009. "Putting Government Data Online").*

Frases como "Nossos impostos financiam a coleta de dados públicos – no entanto, temos que pagar novamente para acessá-los. Tornando-os livremente disponíveis podemos estimular a inovação." e "Embasando todos os nossos esforços há uma crença fundamental de que o aumento da transparência irá melhorar a confiança do público no governo" são encontradas nos sítios dessas comunidades, demonstrando o interesse e o estímulo para que o governo publique esses dados.

Em meio aos acontecimentos, em Setembro de 2009, David Eaves, um ativista canadense dos Dados Abertos Governamentais, publicou as chamadas Três Leis dos Dados Abertos Governamentais, hoje consideradas diretrizes para a publicação de dados e citadas pelo W3C como importantes. As leis são um reflexo dos oito princípios dos Dados Abertos Governamentais, um conjunto de características que os dados devem ter para serem considerados Abertos, definidos pelo Open Government Working Group em Dezembro de 2007.

Embora milhares de datasets tenham sido publicados nesses portais, relativamente, poucas aplicações foram feitas com os dados, por causa disso, esses governos e governos municipais dos Estados Unidos e do Canadá, começaram a promover campeonatos de elaboração de aplicações, com prêmios para as mais úteis, estimulando o uso desses dados publicados.

O momento internacional atual é o da continuidade da publicação e a expectativa de que, no médio e longo prazo os frutos sejam mais numerosos em 2011 do que em 2009 e 2010.

No Brasil, destaques para comunidades – como a Transparência Hacker Day que nasceu também em 2009 em acompanhamento aos acontecimentos internacionais –, que surgiram de forma tímida, mas aos poucos se desenvolveram, mantendo a postura de pressionar o governo em eventos cívicos e governamentais.

2.5. Marcos Políticos dos Dados Abertos

Direcionando agora para as origens do movimento: quais foram as ações, ou alicerces, para que tudo isso fosse viabilizado? A resposta já foi dada no estudo de Tim Berners-Lee (Berners-Lee, 2009). A sociedade pressionou por muito tempo os governos dos EUA e Reino Unido (em especial este último) para publicar os dados nesse formato, publicando dados até mesmo em formatos aderentes à Web Semântica para demonstrar a vontade de fazer e a força do movimento.

O primeiro manifesto formal de governo, o memorando "Transparency and Open Government", foi publicado pelo Presidente dos EUA, Barack Obama, em 21 de Janeiro de 2009, apresentando as motivações de criar um nível de abertura sem precedentes no governo norte-americano. O manifesto versa sobre os princípios da Transparência, Colaboração e Participação e orienta a publicação, em 120 dias, de uma diretriz sobre como os órgãos executivos tinham que exercer esses princípios.

Em 8 de Dezembro de 2009, a diretriz "Open Government Directive" é publicada, orientando a publicação de dados do governo na rede, o aumento da qualidade das informações de governo, a criação e institucionalização da cultura do Governo Aberto e a criação de um framework de habilitação de políticas para o Governo Aberto. Por último, a diretriz orienta cada agência (como os órgãos são chamados nos Estados Unidos) a criar um plano de Governo Aberto, explicando como pretendem atingir os objetivos do memorando.

No caso do Reino Unido, o Primeiro-Ministro Gordon Brown falou, em sua declaração de renovação constitucional de 10 de Junho de 2009, sobre a necessidade de espalhar a cultura e prática da liberdade de informação pois segundo ele, o público deve saber como seu dinheiro

é gasto e sobre a intenção de tornar o parlamento mais confiável ao público, tornando as informações de governo acessíveis e úteis para a maior quantidade de pessoas o possível, na web. Para isso, ele pediu auxílio ao Sir Tim Berners Lee, pretendendo levar essas informações ao público já no mês seguinte.

No Brasil, as ações pela transparência vêm acontecendo de forma descentralizada. O Portal da Transparência do Governo Federal, por exemplo, é uma iniciativa da Controladoria-Geral da União (CGU) de 2004, que pretende assegurar o controle popular da aplicação dos recursos públicos. "O objetivo é aumentar a transparência da gestão pública, permitindo que o cidadão acompanhe como o dinheiro público está sendo utilizado e ajude a fiscalizar." [Portal da transparência, 2010].

A lei complementar 131/2009, a lei da transparência ou lei Capiberibe, foi um marco na obrigatoriedade da publicação da execução orçamentária e financeira, em tempo real, em meios eletrônicos de acesso (ou seja, internet). Outro marco normativo é a lei de acesso à informação, esta já foi aprovada pela Câmara dos Deputados e está em tramitação no Senado. Entre suas garantias, estão o acesso imediato e gratuito a todas as informações públicas e o estabelecimento de critérios para proteção das informações pessoais e sigilosas.

2.6. Avaliação de aspectos socioculturais do Reino Unido e do Brasil

No Reino Unido, a cultura de acompanhamento e participação nas decisões do Estado já era desenvolvida desde a Reforma Inglesa e sua constituição de 1793. Segundo Andrew Narato (Narato, 2002),

> *No Reino Unido é o poder da sociedade civil que previne que o Parlamento aprove uma versão moderna da lei septenal, que não seria ilegal ou inconstitucional mesmo nos dias de hoje. Em muitas democracias recentes, poderes legais deste tipo inibiriam drasticamente iniciativas civis e públicas que no Reino Unido controlam o abuso destes poderes.*

Esse desenvolvimento naturalmente prosseguiu para a era da informação e do governo eletrônico, onde o ativismo manifestou-se através das citadas comunidades e do reúso das informações do governo.

No caso do Brasil, o nível de confiança que o cidadão tem no governo nunca foi dos grandes, ocasiões como a ditadura, onde não existia transparência de governo, refletiram em legislações e movimentos atuais pela transparência ao longo dos anos.

Escândalos de corrupção recorrentes, governo após governo, foram outro fator motivador dessas causas, esses eventos aconteciam com a maioria dos partidos que assumiam posições de poder no país.

A seguir, um ranking de IDH do Programa das Nações Unidas para o Desenvolvimento, publicado em Novembro de 2010.

Tabela 1. *Ranking* IDH 2010 dos países que publicam oficialmente Dados Abertos

	Ranking IDH		Valor IDH	Expectativa de vida ao nascer	Média de anos de escolaridade	Expectativa de anos de escolaridade	PIB *per capita*
Desenvolvimento muito alto	2	Austrália	0.937	81.9	12.0	20.5	38,692
	3	Nova Zelândia	0.907	80.6	12.5	19.7	25,438
	4	Estados Unidos	0.902	79.6	12.4	15.7	47,094
	8	Canadá	0.888	81.0	11.5	16.0	38,668
	26	Reino Unido	0.849	79.8	9.5	15.9	35,087
Des. alto	73	Brasil	0.699	72.9	7.2	13.8	10,607

Verificando os dados apresentados na tabela, atentemos à parte da análise social do Reino Unido por Andrew Narato (Narato, 2002), que versa sobre a relação entre a classe social e controle social do Estado:

> *A questão não é simplesmente que a eleição enquanto tal é um princípio de distinção, mas sim que um eleitorado sem educação formal, relativamente pobre e trabalhador tem muito mais dificuldades de observar, criticar e controlar os representantes eleitos do que os estratos sociais mais privilegiados de regimes representativos pré-democráticos.*

Considerando então, baseado nos dados da tabela, que a maior parte do eleitorado brasileiro possui pouca educação formal, é relativamente pobre e trabalhador, não há cultura de monitoramento do governo por parte da maioria dos brasileiros. Por outro lado, a minoria privilegiada, por vivermos num país com uma população de 190 milhões, ainda

é numerosa e tem um papel presente na cobrança de resultados e de informações do governo.

No caso dos outros países, todos Desenvolvidos, o PIB (Produto Interno Bruto) e a escolaridade são maiores, justificando em parte porque a cobrança social por transparência é mais presente.

3. Conclusões

A visão internacional que motivou a utilização das tecnologias e metodologias relativas aos Dados Abertos Governamentais foi baseada em duas ideias: prover dados ao cidadão é um serviço público, portanto os dados públicos devem ser disponibilizados; e a participação por parte dos cidadãos numa rede social colaborativa traz grandes resultados para a sociedade e para o próprio governo.

Embora possamos afirmar que utilizar ferramentas inovadoras e esforço para uma sociedade pouco culta não traga retorno imediato, devemos observar que a publicação de dados públicos do governo será legalmente exigida e que a publicação num formato reutilizável é investimento para a promoção da transparência e eficiência do estado. O governo brasileiro está construindo um marco normativo para padronizar e dar mais efetividade às praticas atuais de disseminação de dados e informações.

O avanço da Política de Disseminação de Dados e Informações permitirá que o provimento destes dados para a sociedade seja caracterizado como um serviço público em harmonia com o que preconiza o Decreto 6.932/2009.

Segundo Becky Hogge em seu relatório sobre Dados Abertos (Becky Hogge, 2010),

> *(...) Tanto nos Estados Unidos quanto o Reino Unido havia uma abordagem de três partes. Os três grupos de atores que foram cruciais para o sucesso dos projetos:*
> - *A sociedade civil, em particular um motivado grupo de "hackers cívicos";*
> - *Uma comprometida e bem equipada "camada do meio" de habilidosos burocratas do governo; e*

- *Um "alto escalão", motivado ora por uma força de fora (no caso do Reino Unido), ora por uma administração política nova e faminta por mudança (nos Estados Unidos).*

Como Tim Berners-Lee observou em sua entrevista: "Deve começar em cima, deve começar no meio e deve começar embaixo". A conclusão desse relatório fortalece essa assertiva e adverte aqueles tentando se espelhar no sucesso dos projetos do Reino Unido e dos Estados Unidos para não negligenciar nenhuma dessas três camadas de influência.

Coordenar a publicação dos dados com ações para estimular sua utilização por parte do cidadão é desejável, mas essas ações tocam uma realidade cultural que vai além da simples inclusão digital. Antes das deficiências tecnológicas, a sociedade brasileira necessita ser educada para desenvolver uma cultura de controle e fiscalização do governo.

Os órgãos da Administração Pública podem prover seus dados como um serviço à sociedade através da publicação de dados na internet, utilizando padrões abertos e processáveis por máquina. Os Dados Abertos Governamentais podem ser a solução, mas é necessário que governo, atores públicos intermediários e sociedade estejam preparados a colaborar para que a metodologia funcione.

A Infraestrutura Nacional de Dados Espaciais, um projeto colaborativo para interoperabilidade e disseminação de dados abertos no Governo Brasileiro

Hesley da Silva Py[1]

Resumo

No final do ano de 2009 o Governo Brasileiro publicou o decreto 6.666 que institui e dá as diretrizes para a Infraestrutura Nacional de Dados Espaciais (INDE). Para atender ao decreto, foi criado no âmbito da Comissão Nacional de Cartografia (CONCAR) um comitê responsável pela definição de um plano de ação para implantação da INDE (CINDE). Além das complexidades inerentes à definição de uma Infraestrutura de Dados Espaciais (IDE) e seus componentes, a tarefa recebida pelo comitê agrega ainda complexidades adicionais, a saber, as concernentes às dimensões do território brasileiro, a necessidade de realização de um trabalho colaborativo com pessoas distribuídas em diversos órgãos espalhados no território nacional, a inexistência de políticas e padrões definidos para interoperabilidade de dados espaciais e a falta de uma cultura de dados abertos. Frente a esse desafio, contando com a colaboração de cento e quinze pessoas de vinte e três órgãos, foi realizado o trabalho do CINDE, que em maio de 2010 entregou à CONCAR o plano de ação da INDE. O plano de ação é um documento dividido em oito capítulos, com duzentos e três páginas, contendo as diretrizes para a implantação da INDE, o que

[1] Instituto Brasileiro de Geografia e Estatística – IBGE.

envolve, entre outras questões, os padrões e políticas para interoperabilidade dos dados e informações e as premissas para a utilização de dados abertos.

O cenário do trabalho

Escrever um plano de ação que aborde os componentes necessários para a implantação de uma Infraestrutura de Dados Espaciais Brasileira, tendo em vista as dimensões e complexidades do país, já seria um trabalho árduo, soma-se a isto a necessidade de o trabalho ser feito de forma colaborativa, contando com cento e quinze colaboradores, distribuídos em vinte e três instituições espalhadas no território nacional, mais a inexistência de políticas e padrões definidos para interoperabilidade de dados espaciais e a falta da cultura de dados abertos. Esse era o cenário inicial para os trabalhos do CINDE, comitê responsável pela criação do plano de ação da INDE.

A Infraestrutura Nacional de Dados Espaciais

A Infraestrutura Nacional de Dados Espaciais (INDE) é um projeto do Governo Brasileiro, sob a responsabilidade da Comissão Nacional de Cartografia (CONCAR), cujo objetivo é propiciar um ambiente integrado, a partir do qual as instituições e sociedade em geral possam localizar, conhecer e acessar os dados e informações espaciais produzidos pelos diversos órgãos, em um primeiro momento federais, do Governo Brasileiro. O decreto 6.666, que institui a Infraestrutura Nacional de Dados Espaciais, a define como um conjunto integrado de tecnologias; políticas; mecanismos e procedimentos de coordenação e monitoramento; padrões e acordos; necessário para facilitar e ordenar a geração, o armazenamento, o acesso, o compartilhamento, a disseminação e o uso dos dados espaciais de origem federal, estadual, distrital e municipal.

Tais acordos de compartilhamento de bases de dados espaciais, com vistas à integração, compatibilização (harmonização) e disponibilização daquelas consideradas de uso comum, foi um fator que impulsionou a evolução das IDEs em todo o mundo. Nesse sentido, é possível observar em diversos países que esses acordos, estabelecidos inicialmente entre órgãos públicos, atualmente abrangem todos os atores da sociedade [Plano de ação da INDE].

Ainda com respeito às IDEs em outros países e à necessidade de compartilhamento e disponibilização do dado espacial, cita-se a definição de IDE segundo o Instituto Nacional Geográfico da Espanha:

> *partindo-se da premissa de que os processos relacionados com a informação geográfica (IG) devam ser unificados, que a IG deva ser amplamente acessível, e que deva existir um consenso entre instituições para compartilhar informação, o termo Infraestrutura de Dados Espaciais é utilizado para se nomear o conjunto de tecnologias, políticas, estruturas e arranjos institucionais que facilitam a disponibilidade e o acesso à informação espacial. (IGN/ IDEE, chamada de IDE Nacional, 2008).*

As propostas para as IDEs identificadas em diversos países foram analisadas e contribuíram para a definição da INDE. Além dessas, constituiu-se condição essencial para os trabalhos realizados a necessidade de "devolver" para a sociedade um conjunto de informações úteis que, na realidade, pertence à própria sociedade. Tal necessidade verifica-se na crescente demanda por dados espaciais, o que pode ser identificado a partir das diversas aplicações que surgem a cada dia explorando esses dados. Entretanto, para que esses dados e informações disponibilizados possam atender de forma ampla esse apelo social não basta que os dados estejam disponíveis, mas sim, que estejam disponíveis padronizados e consequentemente interoperáveis e abertos. Pois, somente dessa forma esses dados poderão ser utilizados de forma plena, constituindo-se elemento fundamental para o desenvolvimento tecnológico e permitindo a geração de novas informações e aplicações até então impensadas. Ainda nesse sentido, identifica-se a necessidade, também apontada no plano de ação da INDE, de investimento em *software* livre para a exploração desses dados.

O trabalho colaborativo

Para a realização do trabalho de escrita do plano de ação, dividiu-se o documento em oito capítulos, para cada capítulo foram definidos dois coordenadores de órgãos diferentes. Os coordenadores estabeleceram, juntamente com os participantes, a estrutura de tópicos do capítulo e

identificaram entre esses mesmos participantes aqueles que eram mais competentes, levando-se em consideração também as competências do órgão ao qual estava vinculado, para preencher o texto de cada um dos tópicos.

Os capítulos foram planejados para abordarem as dimensões de implementação de uma Infraestrutura de Dados Espaciais (IDE), a saber: a dimensão organizacional, a dimensão técnica e a dimensão humana. O primeiro capítulo trata de conceitos gerais e diretrizes para a implantação da INDE; os Capítulos 2 a 7 endereçam questões organizacionais, técnicas e humanas essenciais A consolidação dá-se no Capítulo 8, que corresponde ao plano de ação propriamente dito [Plano de ação da INDE, 2009].

Tendo em vista a incerteza dos participantes no sucesso de um trabalho envolvendo tantas pessoas de tão diferentes órgãos, por falta de casos similares, os trabalhos iniciais do comitê foram direcionados para alcance a confiança e o engajamento dos participantes com o projeto. O alcance dessa credibilidade e o consequente comprometimento da maior parte dos participantes foi um ponto essencial para a realização do trabalho. Tal confiança foi alcançada na medida em que os participantes entendiam que tinham toda autonomia, como técnicos competentes indicados por seus órgãos, para definir o conteúdo de seu capítulo, sem interferências.

Para o trabalho de escrita do plano foi disponibilizado na *Web* um ambiente colaborativo (*wiki*) que continha além do espaço para o texto propriamente dito uma área de fórum onde as dúvidas e divergências eram tratadas. Além do ambiente disponibilizado, durante o período proposto para a realização dos trabalhos foram realizadas algumas reuniões presenciais entre os participantes de cada capítulo e para a finalização e consistência dos textos aconteceu um encontro com duração de dois dias, com todos os participantes, mantendo o espírito colaborativo, em que as maiores pendências foram sanadas.

Interoperabilidade e dados abertos

Interoperabilidade de dados é um termo utilizado para, basicamente, informar a possibilidade de dados advindos de fontes distintas serem combinados para produzir um outro dado ou informação. Para tanto, alguns requisitos mínimos devem ser respeitados. Dentre esses requisitos

estão: o estabelecimento de padrões mínimos, a utilização de formatos abertos e a descrição de metadados.

Padrões são condições essenciais para a interoperabilidade. Para os trabalhos da INDE foram definidos uma série de padrões, sejam eles estruturais – tratam da forma como o dado está organizado estruturalmente – ou de comunicação – tratam de protocolos para comunicação entre os diversos sistemas para compartilhamento de dados.

Tão importante quanto os padrões são os metadados. Metadados são informações que permitem o entendimento dos dados com relação à sua localização e uso. Ou seja, a partir dos metadados é possível identificar se o dado descrito possui as características necessárias para ser utilizado em um determinado trabalho, além disso, permite identificar onde e como o dado está disponível para ser utilizado. Dessa forma, os metadados fornecem uma visão de quais dados existem, de onde eles podem ser acessados, além de transferir para o usuário a responsabilidade de identificar se aquele dado é o mais adequado para a realização dos seus trabalhos.

É também propósito da INDE a disponibilização, tanto quanto possível, dos dados segundo o conceito de dados abertos.

Nesse sentido, para efeitos da INDE, dados abertos podem ser compreendidos como aqueles estão disponíveis na *Web* através de endereços persistentes, de forma o mais desagregada possível, que fazem uso de formatos abertos e conhecidos e que não possuem restrições quanto ao seu uso.

Destaca-se ainda em todo esse esforço para o compartilhamento de dados, em formato que atenda aos anseios da sociedade, a importância da utilização e desenvolvimento de softwares livres que permitam o consumo desses dados e informações disponíveis.

Conclusão

A INDE é um marco tanto na forma de compartilhamento dos dados e informações espaciais produzidas no Brasil, quanto na forma colaborativa como o projeto foi construído no âmbito do Governo brasileiro. A justificativa principal para o projeto é a necessidade do acesso aos dados geográficos existentes ocorrer de modo fácil, cômodo e eficaz para a sociedade e Governo em geral.

Além dos benefícios da disseminação dos dados propriamente dita, a INDE colabora ainda com outras iniciativas do Governo disseminando a cultura de dados abertos e a importância do trabalho colaborativo, visando evitar o esforço duplicado, otimizando assim a utilização do recurso público. Os trabalhos do comitê da INDE colaboram também com a iniciativa e-PING – programa de Governo eletrônico responsável pela definição dos padrões de interoperabilidade do Governo eletrônico – definindo e propondo os padrões identificados, fornecendo casos de sua utilização, para que possam ser adotados.

Os dados disponíveis na INDE, padronizados e em formato aberto, fornecem à sociedade a possibilidade de geração de novas informações até então não pensadas. Tal possibilidade somada à existência de um conjunto de *softwares* livres para manipulação desses dados colabora para o desenvolvimento tecnológico da nação. Da mesma forma, a possibilidade do cruzamento dos dados disponíveis na INDE, auxilia o Governo no estabelecimento das políticas públicas na medida em que dá mais subsídios para as tomadas de decisão.

Formato	*15,5 x 22,5 cm*
Mancha gráfica	*12 x 18,3cm*
Papel	*pólen soft 80g (miolo), duo design 250g (capa)*
Fontes	*Times New Roman 17/20,4 (títulos),*
	12/14 (textos)