

GOVERNANÇA DIGITAL

MARCELO SOARES PIMENTA
DIEGO RAFAEL CANABARRO
ORGANIZADORES

GOVERNANÇ DIGITAL



UNIVERSIDADE
FEDERAL DO RIO
GRANDE DO SUL

Reitor

Rui Vicente Oppermann

Vice-Reitora e Pró-Reitora
de Coordenação Acadêmica

Jane Fraga Tutikian

EDITORA DA UFRGS

Diretor

Alex Niche Teixeira

Conselho Editorial

Álvaro Roberto Crespo Merlo

Augusto Jaeger Jr.

Carlos Pérez Bergmann

José Vicente Tavares dos Santos

Marcelo Antonio Conterato

Marcia Ivana Lima e Silva

Maria Stephanou

Regina Zilberman

Tânia Denise Miskinis Salgado

Temístocles Cezar

Alex Niche Teixeira, presidente

Centro de Estudos Internacionais sobre Governo (CEGOV)

Diretor

Marco Cepik

Vice Diretor

Ricardo Augusto Cassel

Conselho Superior CEGOV

Ana Maria Pellini, Ario Zimmermann, José
Henrique Paim Fernandes, José Jorge Ro-
drigues Branco, José Luis Duarte Ribeiro,
Paulo Gilberto Fagundes Visentini

Conselho Científico CEGOV

Cássio da Silva Calvete, Diogo Joel
Demarco, Fabiano Engelmann, Hélio
Henkin, Leandro Valiati, Lúcia Mury
Scalco, Luis Gustavo Mello Grohmann,
Marcelo Soares Pimenta, Marília Patta
Ramos, Vanessa Marx

Coordenação Coleção Editorial CEGOV

Cláudio José Muller, Gentil Corazza, Marco
Cepik

GOVERNANÇA DIGITAL

MARCELO SOARES PIMENTA
DIEGO RAFAEL CANABARRO
ORGANIZADORES

© dos autores
1ª edição: 2014

Direitos reservados desta edição:
Universidade Federal do Rio Grande do Sul

Coleção CEGOV Capacidade Estatal e Democracia

Revisão: Fernando Preusser de Mattos, Fernanda Lopes
Silva, Ricardo Fagundes Leães

Projeto Gráfico: Joana Oliveira de Oliveira, Liza Bastos
Bischoff, Henrique Pigozzo da Silva

Capa: Joana Oliveira de Oliveira

Foto da Capa: Joana Oliveira de Oliveira

Impressão: Gráfica UFRGS

Apoio: Reitoria UFRGS e Editora UFRGS

Os materiais publicados na Coleção CEGOV Capacidade Estatal e Democracia são de exclusiva responsabilidade dos autores. É permitida a reprodução parcial e total dos trabalhos, desde que citada a fonte.



G721 Governança Digital [recurso eletrônico] / organizadores Marcelo Soares Pimenta e Diego Rafael Canabarro. - dados eletrônicos. - Porto Alegre: Editora da UFRGS/CEGOV, 2014.
212 p. : pdf

(CEGOV Capacidade Estatal e Democracia)

1. Administração pública. 2. Tecnologia da informação. 3. Governança digital. 4. Segurança cibernética. 5. Economia da informação. I. Pimenta, Marcelo Soares. II. Canabarro, Diego Rafael. III. Série.

CDU – 681.3:35(81)

CIP-Brasil. Dados Internacionais de Catalogação na Publicação
(Jaqueline Trombin - Bibliotecária responsável CRB10/979)

ISBN 978-85-386-0478-5

BIG DATA, VISUALIZAÇÃO DE INFORMAÇÕES E *VISUAL ANALYTICS* EM SUPORTE A POLÍTICAS PÚBLICAS

CARLA DAL SASSO FREITAS

Possui graduação em Processamento de Dados pela Universidade Federal do Rio Grande do Sul (UFRGS) (1977), mestrado e doutorado em Ciência da Computação pela UFRGS (em 1981 e 1994, respectivamente), e estágio de pós-doutorado no International Computer Science Institute e E.O. Lawrence Berkeley National Laboratory, em Berkeley, CA, EUA (1996). É professora da UFRGS desde 1980, hoje como Professora Associada.

MARCELO SOARES PIMENTA

Doutor em Informática – Université Toulouse 1 (1997), Bacharel (1988) e Mestre (1991) em Ciências da Computação pela UFRGS, com pós-doutorado na Université Paul Sabatier, Toulouse, França (2002-2003). Atualmente é professor associado e pesquisador no Instituto de Informática (INF) da UFRGS.

INTRODUÇÃO

O uso disseminado da Internet na sociedade em geral, seja através de sistemas computacionais convencionais ou de aplicações em dispositivos móveis e a informatização das atividades nos diversos poderes (Legislativo, Judiciário e Executivo) observada no Brasil ao longo dos últimos quinze anos, criou um ambiente onde a geração de dados se dá de forma vertiginosa. Do ponto de vista computacional, o armazenamento e o tratamento de tais conjuntos de dados podem ser enquadrados no que se convencionou chamar de *Big Data*, e a disponibilização desses dados para a sociedade em geral, de *Open Data* (ou dados abertos).

Essas expressões correspondem a soluções tecnológicas que permitem lidar com esses dados diversos em grande volume (*Big Data*) e disponibilizar visões desses dados para o público em geral (*Open Data*). Considerando seu uso na esfera dos três poderes, tais soluções devem prover uma velocidade que seja compatível com o ritmo das tomadas de decisões. Além de consultar e manipular esse grande volume de informações, gestores de órgãos públicos, por exemplo, precisam também analisar e interpretar adequadamente essas informações disponíveis, de modo a ter condições e conhecimento para embasar racionalmente uma decisão. Por outro lado, cada vez mais os cidadãos querem ter acesso às informações geradas pelos diversos setores e sistemas públicos, tanto para informação como para participação consciente em momentos de discussão e decisão. Assim, a disponibilização desses dados deve se dar em uma linguagem e em formato compatíveis com as diversas necessidades e perfis de seus usuários.

Neste capítulo, são apresentados os conceitos essenciais relacionados a *Big Data* e *Open Data*, incluindo Visualização de Informações e *Visual Analytics*. Em seguida, são discutidos conjuntos de soluções integradas e atuais envolvendo esses conceitos e que podem ser usadas para agregar informações e permitir ações no setor público. Exemplos de algumas dessas soluções e uma discussão sobre aspectos de seu uso são também apresentados.

BIG DATA, OPEN DATA, VISUALIZAÇÃO DE INFORMAÇÕES E VISUAL ANALYTICS: CONCEITOS E FUNDAMENTOS

Expressão muito utilizada atualmente, *Big Data* se refere a conjuntos de dados cujo tamanho e complexidade tornam difícil, oneroso ou até impossível seu tratamento pelas ferramentas de *software* atuais. O conceito de *Big Data* é,

portanto, relativo à tecnologia do momento. Se pensarmos que o crescimento exponencial da densidade de transistores nos circuitos integrados foi previsto por Gordon Moore ainda na década de 1970, ficando conhecido como Lei de Moore, e foi constatado com o conseqüente declínio dos custos de hardware, progressão similar afetou o conceito de volumes de dados. Na década de 1980, um sistema de armazenamento de massa considerado grande tinha a capacidade de cerca de 100 GB, e nós convivíamos com computadores com apenas algumas dezenas de KBytes de memória RAM. Hoje, conjuntos de dados considerados *Big Data* ocupam terabytes (10¹² bytes), petabytes (10¹⁵ bytes), e exabytes (10¹⁸ bytes), dependendo de sua origem, características e uso.

Uma das definições mais aceitas de *Big Data* é do grupo Gartner (GARTNER INC., 2011) que caracteriza esse conceito como conjuntos de dados com grande volume, grande velocidade (de geração, alteração, processamento e utilização) e grande variedade (diversidade de elementos), no que é conhecido como “3 Vs”. Outras características como veracidade (confiabilidade das fontes de dados) e valor (valor econômico agregado aos dados) também são usadas (SATHI, 2012; CHAN, 2013), mas não são plenamente aceitas, prevalecendo a definição técnica de volume, velocidade e variedade.

O entendimento e a utilização de tais conjuntos de dados levam a duas questões essenciais: (i) como apresentar os dados (ou informações deles extraídas) e (ii) como analisar tais conjuntos de dados. A primeira questão remete ao conceito de visualização de dados (ou visualização de informações)¹. Por visualização de informações entendem-se técnicas de representação visual, geralmente disponibilizadas na forma de ferramentas interativas, que utilizam desde gráficos de linha simples até formas geométricas, símbolos e imagens mais complexas para representar visualmente dados e permitir sua “exploração” com vistas a apoiar o entendimento dos mesmos. Ferramentas de visualização de informações utilizam desde técnicas introduzidas nos primórdios da área (CARD; MACKINLAY; SHNEIDERMAN, 1999) até soluções recentes popularizadas por recursos computacionais na *web*, como, por exemplo, a biblioteca D3 (D3, 2014).

Apesar de visualizações terem como objetivo serem informativas, como muitas das apresentadas por Yau (2011), elas são em geral fruto de um processo de análise que envolve métodos estatísticos, técnicas de inteligência artificial e de mineração de dados, ou seja, métodos diversos dependentes da natureza, da fonte e do uso dos dados. Da integração de técnicas de visualização de informações com técnicas e métodos de análise de dados surgiu a área de “*visual analytics*” (THOMAS; COOK, 2005). Essa área trata, portanto, do desenvolvimento e da aplicação de técnicas de análise conjugadas a técnicas de visualização, onde os usuários podem ex-

(1) Aqui utilizaremos visualização de dados e visualização de informações como sinônimos.

plorar seus conjuntos de dados através de recursos interativos visuais e empregar, ao mesmo tempo, técnicas de análise, com o objetivo de entender o conjunto de dados, o que envolve interpretar e, frequentemente, descobrir padrões e tendências.

Finalmente, apesar do conceito de *Open Data* ser análogo aos de open source e open access, apenas recentemente difundiu-se, justamente após as iniciativas de disponibilização pública de dados de governo, em 2009, através dos projetos data.gov.uk e Data.gov, na Inglaterra e nos Estados Unidos, respectivamente. A ideia é que certos dados devem ser disponibilizados livremente para qualquer pessoa ou instituição, para uso e republicação, inclusive, sem restrições de *copyright*, patentes ou outros mecanismos de controle (AUER et al., 2007).

Esse movimento internacional em prol de abertura dos dados governamentais gerou a Parceria para Governo Aberto, ou OGP (do inglês *Open Government Partnership*). Lançada em 2011, “a OGP é uma iniciativa internacional que pretende difundir e incentivar globalmente práticas governamentais visando à transparência dos governos, ao acesso à informação pública e à participação social” (OGP, 2014). Os oito países fundadores da Parceria (África do Sul, Brasil, Estados Unidos, Filipinas, Indonésia, México, Noruega e Reino Unido) assinaram a Declaração de Governo Aberto e apresentaram seus Planos de Ação. A OGP, atualmente, agrega 63 países.

Há, assim, uma relação quase que indissociável dessas áreas: a geração constante e crescente de dados de natureza científica, econômica, social e política leva às necessidades computacionais de armazenamento e tratamento (*Big Data*); sua disponibilização (*Open Data*), que, por sua vez, leva à geração de novos dados, requer formas intuitivas de apresentação (Visualização de Informações) e torna necessária a adoção de métodos confiáveis de análise (*Visual Analytics*).

BIG DATA, VISUALIZAÇÃO DE INFORMAÇÕES E VISUAL ANALYTICS NO SETOR PÚBLICO

Muitas instituições públicas coletam e/ou produzem diferentes tipos de dados como insumos ou resultado de suas atividades. Essa grande quantidade de dados torna-os particularmente significativos para aumentar a oferta de vários serviços que, além de envolverem consultas (a esses dados) com os mais variados objetivos, podem contribuir para sua atualização (UBALDI, 2013) e manipulação. Conforme os objetivos, novos dados podem resultar e serem, por sua vez, disponibilizados.

A informatização crescente da sociedade aumenta a demanda por mais e melhores serviços públicos – sejam eles da esfera municipal, estadual ou federal.

Com o uso de tecnologias de informação e comunicação (TIC) espera-se que os governos deem respostas mais rápidas aos interesses da população, aumentem a transparência dos seus gastos e criem oportunidades mais flexíveis para a participação da sociedade nas grandes decisões. Além desses usos mais óbvios das tecnologias na esfera pública, a maturidade no seu uso pode permitir contemplar áreas incomuns, como a mitigação de riscos relativos a catástrofes naturais, defesa civil, segurança nacional e meio ambiente (IBM, 2013).

Independentemente de esfera (municipal, estadual, federal) ou do tipo de serviço disponibilizado, é uma tendência cada vez maior que as organizações públicas procurem formas de tornar-se mais eficientes operacionalmente e eficazes na resposta às necessidades da sociedade, reduzindo custos e esforços. É para contribuir exatamente em relação a esses desafios que podem ser (e estão sendo) adotados os conceitos e técnicas de *Big Data*, Visualização de Informações e *Visual Analytics*.

Big Data, como vimos na seção anterior, é um conceito definido em termos de “3 Vs”: grandes volumes e alta variedade de dados sendo gerados e/ou manipulados a grandes velocidades. Além das bases de dados existentes, novos dados estão sendo gerados cada vez mais em taxas crescentemente aceleradas. A habilidade de exibir todos esses dados em diferentes formas e perspectivas usando técnicas variadas de visualização de informações permite representar de forma direta ou indireta relações que não seriam tão evidentes de serem descobertas sem esta exibição. Além disso, a capacidade de realizar análise sobre os dados existentes habilita as organizações públicas a identificar pontos de melhoria nos seus processos de negócio e nas suas atividades-fim, algo que não seria possível antes.

Outro benefício para a sociedade em geral é que a disponibilização de dados motiva a geração de novos serviços, não necessariamente ofertados pelas instituições públicas, mas mediados por estas. Por exemplo, a iniciativa pioneira de *Open Data* no Reino Unido disponibiliza uma série de conjuntos de dados², que são utilizados não necessariamente por entidades públicas para a construção de aplicações computacionais, as quais são também disponibilizadas no mesmo repositório³ após uma fase de aprovação. Tais aplicações utilizam técnicas variadas de visualização e análise.

Os tipos de *Big Data* que o governo pode potencialmente analisar incluem informações sobre cidadãos (pessoas físicas), instituições e empresas (pessoas jurídicas), abarcando não só informações existentes em bases de dados e “*datawarehouses*” governamentais, mas eventualmente complementando-as com outras informações disponíveis via documentos sobre transações realizadas (incluindo contratos, procurações, uso de cartões bancários e operadoras de telecomunicações,

(2) Disponível em: <<http://data.gov.uk/dataset/>>. Acesso em: 09 set. 2014.

(3) Disponível em: <<http://data.gov.uk/apps/>>. Acesso em: 09 set. 2014.

etc.), e via redes sociais, blogs e e-mails, e abrangendo diferentes tipos de mídia, como vídeos, *broadcasts*, fotos e dados de localização (GPS). Todos estes dados e conjuntos de dados são grandes e complexos, com diferentes fontes e origens e múltiplas formas de representação, variando desde dados muito bem estruturados a dados desestruturados, com vários estágios intermediários entre esses extremos.

Sem uma capacidade de analisar esses dados, o governo torna-se cada vez mais rico em termos de dados, mas, ironicamente, pobre em termos de informações e, muito provavelmente, em termos de qualidade (e diversidade) de serviços que pode prestar aos cidadãos. Tecnologias e sistemas relacionados a *Big Data* e a *Visual Analytics* apresentam uma enorme oportunidade para as instituições públicas, e, conseqüentemente, para os governos, de entenderem os processos e fenômenos ocorrendo desde o nível de bairros e cidades até os níveis macrorregional, nacional e continental.

Vale a pena ressaltar que, muitas vezes, as referências a Big Data no âmbito governamental podem levar a uma interpretação de se tratar de disponibilização de dados econômicos e sociais, mas uma rápida revisão de iniciativas de alguns países permite observar a disponibilização de dados e aplicações também em áreas como clima, energia, ambiente, geologia, ciência e tecnologia, etc⁴. Áreas-chave onde as organizações governamentais têm alcançado sucesso com *Big Data* e *Visual Analytics* incluem a detecção de fraudes em programas sociais e coleta de impostos e a prevenção e previsão de riscos e crimes (IBM, 2013).

Muitas das aplicações disponibilizadas para a consulta aos dados utilizam técnicas de Visualização de Informações para apresentá-los e, assim, auxiliar o usuário na análise e compreensão das informações. Várias técnicas são desenvolvidas com esse intuito e são reconhecidamente importantes como ferramentas analíticas e de comunicação para lidar com volume e complexidade de dados. Nesse sentido, tem havido um crescente interesse das instituições públicas em usar tais técnicas. Esse interesse deve-se a várias razões (LINDQUIST, 2011):

1. Constatação da inerente complexidade dos desafios relacionados à informação pública;
2. Aumento da consciência dos cidadãos e dos servidores públicos sobre as várias alternativas (usando TICs) para manipular e disponibilizar serviços e informações públicas;
3. Maior conhecimento e experimentação no uso de técnicas de visualização em diferentes domínios têm levado instituições governamentais a desejar adotá-las em diferentes setores.

(4) Nos Estados Unidos, <<https://www.data.gov/>>; na Austrália, <<http://data.gov.au/>>; na Espanha, <<http://datos.gob.es>>, entre outros. Acesso em: 09 set. 2014.

Obviamente, apenas a existência, disponibilização ou aquisição de tecnologia não é suficiente para aumentar seu uso: é necessário também encorajar o funcionalismo (sobretudo o quadro técnico) a se capacitar através de treinamento específico para uso dessas tecnologias como apoio a suas atividades. A próxima seção ilustra usos dessas tecnologias no cenário nacional.

SOLUÇÕES EXISTENTES NO CENÁRIO NACIONAL

Dada a crescente agregação de tecnologia de informação e comunicação aos processos nas entidades governamentais como resultante dos esforços de modernização da gestão pública, não apenas a formulação e a implementação, mas também o monitoramento e a avaliação de políticas públicas se baseiam, cada vez mais, em compartilhamento de dados e fluxos de informação suportados por sistemas de dados e informações digitais. Esta não é uma iniciativa de cunho localizado, uma vez que o Brasil está internacionalmente inserido na Parceria para Governo Aberto, conforme já mencionado, o que levou ao estabelecimento de um Plano de Ação do Governo Brasileiro em relação a dados abertos (OGP, 2014a, 2014b, 2014c) e de uma Infraestrutura Nacional de Dados Abertos (INDA). Nesse contexto, foi criado o Portal Brasil⁵, onde estão disponíveis diversos canais que facilitam a comunicação entre o Estado e a Sociedade.

104

Entretanto, um elemento de complexidade adicional no ecossistema da administração pública decorre do emprego não uniforme de uma multiplicidade de tecnologias de informação e comunicação em suporte aos processos decisórios no campo das políticas públicas. Nesse caso, a criação de sistemas informatizados orientados à resolução de tarefas específicas, sem consideração à realidade multidimensional da ação do Estado, resulta na fragmentação setorial das bases de dados e da própria aplicação de tecnologia da informação (TI) no âmbito governamental. Porém, a tomada de decisão, a formulação, o controle e o monitoramento, bem como o processo de avaliação de políticas públicas requerem um complexo processo de articulação de diversos entes governamentais, em diferentes esferas de atuação, cada um responsável por missões distintas e dotado de um conjunto de recursos variáveis. Em um contexto de crescente interconectividade habilitada pelas TICs, essa necessidade se torna ainda mais premente na medida em que os próprios problemas sociais a serem enfrentados pelos órgãos governamentais se tornam complexos e interconectados. Isso requer a adoção de uma abordagem sistêmica para seu entendimento e resolução.

(5) Disponível em: <<http://www.brasil.gov.br/barra#acesso-informacao>>. Acesso em: 09 set. 2014.

Não obstante esse quadro e a implementação da plataforma de dados abertos⁶, a adoção de técnicas de Big Data, Visualização de Informações e Visual Analytics não é ainda muito disseminada no Brasil. Os benefícios do uso dessas tecnologias, entretanto, começam a ficar evidentes através de algumas soluções, brevemente apresentadas a seguir.

IGOV (INTEGRAÇÃO E INTELIGÊNCIA EM INFORMAÇÕES DE GOVERNO) E DADOS ABERTOS ESTADUAIS

O problema da Gestão da Informação na Administração Pública, o qual abarca uma variedade de funções, entidades componentes e soluções tecnológicas que permitem os relacionamentos entre essas entidades, pode ser visto como um sistema complexo⁷. Uma maneira de afrontar esse problema é promover a interoperabilidade de sistemas de informação governamentais através de técnicas de Ecossistemas Digitais e Sistemas Sociotécnicos, abordagens multidisciplinares que consideram tanto as necessidades técnicas quanto as necessidades do ciclo de serviços de cada um dos atores envolvidos naquele sistema complexo.

A questão da interoperabilidade se desdobra em três problemáticas fundamentais. Do ponto de vista da infraestrutura técnica, é necessário definir uma camada de intercâmbio de dados composta por padrões, linguagens e modelos que suportem e facilitem a difusão e o compartilhamento de dados entre os diversos sistemas de informação e bancos de dados governamentais, bem como permitam a aplicação de técnicas de visualização de dados que facilitem a tomada de decisão na Gestão Pública. Essa infraestrutura está relacionada a tópicos bem atuais de pesquisa em Computação, tais como *Big Data*, Mineração de Dados, Dados como Serviços e Visualização (e Análise) de Dados. Do ponto de vista da governança, é fundamental que o compartilhamento de dados e os processos e fluxos de informação estejam alinhados com os processos relativos ao monitoramento e à avaliação das políticas públicas. No intermeio desses dois âmbitos, do ponto de vista da gestão da informação, deve-se assegurar que os processos e fluxos de informação sejam coerentes entre os diversos entes governamentais envolvidos nos processos de formulação e monitoramento de políticas públicas.

No âmbito do Governo Federal, a Arquitetura e-Ping e a Arquitetura Referencial de Integração de Sistemas Informatizados de Governo (AR) objetivam, em conjunto, melhorar a integração de sistemas estruturantes e corporativos do Governo Federal, seja pela uniformização das plataformas tecnológicas empregadas em cada sistema, seja pela uniformização da estruturação e do acesso aos diversos sistemas

(6) Disponível em: <<http://dados.gov.br/>>. Acesso em: 09 set. 2014.

(7) Ver o capítulo “Sistemas de Governo Eletrônico são Ecossistemas Digitais” deste livro.

e aos dados e metadados correspondentes. O objetivo final é reduzir os custos de transação no acesso e compartilhamento de dados e informações úteis à tomada de decisão – seja pelo Estado, seja pelos cidadãos – no ciclo de políticas públicas.

A plataforma piloto i3Gov⁸ (Informação e Inteligência em Informações de Governo) – inicialmente posta em funcionamento para o monitoramento e a avaliação do Programa de Aceleração do Crescimento (PAC) – é o primeiro módulo integrante da AR. O processo de desenvolvimento institucional e tecnológico da plataforma i3Gov pode ter como consequência produzir conhecimento útil ao processo de desenvolvimento da Arquitetura Referencial. Por definição, i3Gov é uma iniciativa que tem como objetivo a implementação de uma arquitetura referencial de interoperabilidade de sistemas para integração de sistemas informatizados de Governo, troca de dados em larga escala e oferta de um Catálogo de serviços *web*.

É importante, para isso, que se promova o alinhamento com iniciativas estabelecidas – incluindo o Plano de Ação Nacional sobre Governo Aberto, o catálogo DadosGov, a Infraestrutura Nacional de Dados Abertos (INDA), a Infraestrutura Nacional de Dados Espaciais (INDE) –, bem como facilitar a integração de ferramentas e soluções que já se configuram como um início desse processo, tal como o projeto VisPublica (Visualização de Dados Públicos). Destaca-se que, enquanto uma infraestrutura que permite atender à Lei de Acesso à Informação tem por usuários os cidadãos (*front office*), esse tipo de infraestrutura tem por usuários os Gestores Estratégicos (*back office*). A proposta é viabilizar a comunicação e a disponibilização de conjunto de dados brutos e agregados, que permita sua visualização e interpretação, com vistas a atender às demandas por informação para a tomada de decisão mais eficiente e eficaz – beneficiando, por conseguinte, a sociedade como um todo. Para que isso seja possível em todos os níveis governamentais e em diferentes organizações, uma série de diretrizes e recomendações está publicada⁹.

Nos níveis estadual e municipal, encontramos iniciativas de dados abertos, em geral na forma de disponibilização de conjuntos de dados, os quais agregam indicadores sobre várias áreas como, por exemplo, educação, comércio, indústria, transportes, etc. Essas iniciativas têm em comum a possibilidade do usuário fazer *download* dos dados em diversos formatos. Exemplos de tal iniciativa no nível estadual são o portal de dados abertos do estado de Pernambuco¹⁰, o portal de dados abertos do estado do Rio Grande do Sul¹¹, responsabilidade dos respectivos estados, e o portal de dados abertos da Fundação de Economia e Estatística do Rio

(8) Disponível em: <<https://i3gov.planejamento.gov.br/>>. Acesso em: 09 set. 2014.

(9) Guia de Abertura de Dados e Cartilha de Publicação de Dados, disponíveis na seção “Manuais” em <<http://dados.gov.br>>. Acesso em: 10 set. 2014.

(10) Disponível em: <<http://www.dadosabertos.pe.gov.br/>>. Acesso em: 04 nov. 2014

(11) Disponível em: <<http://dados.rs.gov.br/>>. Acesso em: 04 nov. 2014.

Grande do Sul¹². Já no nível de município, podem-se citar como exemplos as iniciativas de São Paulo e de Porto Alegre. A disponibilização de aplicativos variados que fazem uso dos dados para informar os usuários é encontrada em poucos portais, como, por exemplo, do Rio de Janeiro¹³.

4.2 VISPUBLICA E LODZONE

Recentemente, um projeto conjunto entre o Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia da Universidade Federal do Rio de Janeiro (COPPE/UFRJ) e a Universidade Federal de Itajubá (UNIFEI), apoiado pelo Ministério do Planejamento, Orçamento e Gestão (MPOG) iniciou o estudo de técnicas de visualização que ampliem a transparência de dados públicos, facilitando o entendimento de informações e a tomada de decisão, além de utilizar de forma eficiente as técnicas para a massificação dos dados públicos.

O projeto é denominado VisPublica¹⁴ e oferece, através do site, documentos que analisam várias técnicas; algumas visualizações (painéis) onde o usuário pode realizar consultas a alguns indicadores, escolhendo a forma de apresentação gráfica; uma biblioteca para desenvolvedores de *software*; e um conjunto de técnicas *on-line* que podem ser utilizadas para a geração de visualizações com dados do próprio usuário. Dessa forma, embora os painéis contemplem apenas poucos conjuntos de dados abertos, a possibilidade de o próprio usuário alimentar as técnicas com seus próprios dados faz com que a plataforma seja uma iniciativa importante no sentido de experimentação e, sobretudo, de familiarização do público com técnicas de visualização de informações.

Trabalhos publicados ao longo do projeto permitiram testar diferentes técnicas de visualização e como elas são utilizadas por cidadãos de diferentes perfis (PAULA et al., 2011). A partir desse trabalho, foi implementado um conjunto inicial de técnicas (RIBEIRO et al., 2012a), o qual permitiu testar se a forma como os dados são apresentados exerce alguma influência na maneira como cidadãos comuns reconhecem a transparência dos dados: um experimento foi realizado com dados do Tribunal Superior Eleitoral (RIBEIRO et al., 2012b). Em outro trabalho do mesmo grupo (RIBEIRO et al., 2013), foram enumeradas algumas heurísticas que podem ser utilizadas para verificar a adequação da aplicação das técnicas de visualização no contexto governamental¹⁵.

(12) Disponível em: <<http://dados.fee.tche.br/>>. Acesso em 04 nov. 2014.

(13) Disponível em: <<http://rioapps.com.br/>>. Acesso em: 09 set. 2014.

(14) Disponível em: <<http://vispublica.gov.br/vispublica/>>. Acesso em: 09 set. 2014.

(15) Todos estes trabalhos estão disponíveis em: <<http://vispublica.gov.br/vispublica/publico/publicacoes.jsp>>. Acesso em: 09 set. 2014.

É importante observar que a própria plataforma i3Gov⁸ disponibiliza dados de séries históricas de diferentes temáticas. Tais séries históricas foram objeto de visualizações providas pelo projeto *Linked Open Data Zone*¹⁶, através de diversos gráficos com indicadores brasileiros e americanos. *Linked Open Data Zone* é basicamente um *website* baseado no conceito de *linked data* (WOOD, 2011), fornecendo formas de acessar dados através de uma interface de consultas padrão *web*, assim como publicar dados para uso público.

Considerando as poucas iniciativas existentes, conforme relatado, pode-se afirmar que há ainda muito espaço para desenvolvimento com o objetivo de tornar os dados disponíveis mais apropriados para “consumo” por parte da sociedade em geral.

CONSIDERAÇÕES FINAIS

Segundo David Eaves (2009), as seguintes “leis” regem os dados abertos governamentais e são aceitas pela comunidade e pelo *World Wide Web Consortium* (W3C)¹⁷:

1. se o dado não pode ser encontrado e indexado na web, ele não existe;
2. se não estiver aberto e disponível em formato compreensível por máquina, ele não pode ser reaproveitado; e
3. se algum dispositivo legal não permitir sua replicação, ele não é útil.

A essas “leis” podemos acrescentar também uma quarta lei, decorrente da adoção das tecnologias discutidas neste capítulo: “se o dado não pode ser entendido e analisado, ele não será utilizado”.

Nessa mesma linha de definição de princípios do que sejam dados abertos, o *Open Government Working Group* definiu o seguinte como características desses dados:

1. completos: todos os dados públicos estão disponíveis. Dado público é o dado que não está sujeito a limitações válidas de privacidade, segurança ou controle de acesso.
2. Primários: os dados são apresentados tais como os coletados na fonte, com o maior nível de granularidade e sem agregação ou modificação.

(16) Disponível em: <<http://lod2.inf.puc-rio.br/site/index.php?page=1>>. Acesso em: 09 set. 2014.

(17) World Wide Web Consortium é a comunidade internacional que desenvolve padrões abertos para garantir o crescimento adequado e contínuo da web. Ver <<http://www.w3.org/>>. Acesso em: 09 set. 2014.

3. Atuais: os dados são disponibilizados tão rapidamente quanto necessário à preservação do seu valor.
4. Acessíveis: os dados são disponibilizados para o maior alcance possível de usuários e para o maior conjunto possível de finalidades.
5. Compreensíveis por máquinas: os dados são razoavelmente estruturados de modo a possibilitar o processamento automatizado.
6. Não discriminatórios: os dados são disponíveis para todos, sem exigência de requerimento ou cadastro.
7. Não proprietários: os dados são disponíveis em formato sobre o qual nenhuma entidade detenha controle exclusivo.
8. Livres de licenças: os dados não estão sujeitos a nenhuma restrição de direito autoral, patente, propriedade intelectual ou segredo industrial. Restrições sensatas relacionadas à privacidade, segurança e privilégios de acesso devem ser permitidas.

Esses princípios confirmam a convicção sobre a necessidade de conhecermos e adotarmos tecnologias adequadas para manipular os dados existentes, sejam eles abertos ou não. Pode-se afirmar, não basta que os dados governamentais estejam abertos para terem efeito sobre as políticas públicas: é preciso que seja possível entendê-los, analisá-los e, principalmente, utilizá-los. É exatamente para isso que existem as soluções de *Big Data*, Visualização de Informações e *Visual Analytics*.

Assim, se os dados governamentais disponibilizados forem volumosos e complexos, o uso dessas técnicas pode permitir que esse volume e essa complexidade sejam tratáveis, talvez até aumentando seu impacto e sua importância, permitindo que os interessados nas informações (sejam cidadãos, sejam outras instituições, governamentais ou não) as usem de maneira adequada a seu propósito e oferecendo possivelmente vários pontos de vista sobre o desempenho dos governos no cumprimento de suas metas em políticas públicas.

Existem várias técnicas de visualização, e alguns trabalhos discutem e ilustram exemplos de sua adoção em setores públicos internacionalmente (GRAVES; HENDLER, 2013). As diversas experiências com o uso de visualização e *visual analytics* em setores públicos e os relatos sobre o impacto dessas experiências sinalizam que, para decidir quais tipos de técnicas adotar e como fazê-lo, é importante refletir sobre o uso potencial e pretendido das informações. Ou seja, visualizar para quê? As formas e técnicas de visualização são tão diversas e ricas que é importante ter clareza sobre as motivações para adotá-las: (i) entender algum domínio em particular e relacionamentos e correlações subjacentes, geralmente difíceis de identificar? (ii) integrar e sintetizar informações? (iii) desenvolver uma visão estratégica ou uma visão para um órgão ou instituição? (iv) obter informação para tomada de

decisão em situações de risco ou emergência? Na verdade, todas essas intenções são legítimas mas, provavelmente, induzirão à escolha de diferentes técnicas e influenciarão a forma de avaliarmos os benefícios e desvantagens da escolha.

Mesmo que haja um consenso a respeito de a visualização de informações ser um modo importante e criativo de expressar o significado dos dados, ainda se tem muitas dúvidas sobre como proceder para isso. Em um trabalho interessante de compilação, Stowers (2013) enumerou algumas recomendações para o uso de visualização de informações no governo. As recomendações, elaboradas na forma de seis passos compostos por questões a serem respondidas e uma recomendação final, podem efetivamente melhorar o uso dessas técnicas e são, resumidamente, as seguintes:

Passo 1 - Começar e terminar com os dados

- Que história se quer “contar” com os dados?
- Quais são os padrões identificados nos dados?
- Que conclusões podem ser extraídas dos dados?
- Determinar a história que os dados “contam” e, então, usar as visualizações para “contar” essa história.

Passo 2 - Criar um storyboard para mapear os pontos a que se quer chamar atenção com a visualização

Passo 3 - Responder às seguintes questões sobre audiência e objetivos:

- Quem é a audiência?
- Que nível de conhecimento e acesso a tecnologias de informação e comunicação a audiência tem?
- Qual nível de entendimento dos dados a audiência possui?
- O que a audiência quer saber sobre a organização e seus dados?
- O que a audiência quer saber sobre a organização e o aspecto (tópico) que está se tentando destacar?
- Quanto mais sofisticada a audiência, mais interativa deve ser a visualização a ser desenvolvida.

Passo 4 - Analisar seus recursos

- Entre os membros da equipe, quem demonstra talento ou experiência em gráficos, análise de dados e gráficos? Ou quem está aprendendo isso?
- Quanto tempo a equipe deseja dedicar a esse projeto?
- É possível usar pessoas de outras equipes?
- É possível colaborar com outras instituições ou órgãos que já estejam

trabalhando com visualização e poderiam (e desejariam) compartilhar conhecimento e recursos?

Passo 5 - Selecionar o software a ser usado

Experimentar várias das ferramentas disponíveis antes de fazer a escolha final do *software* a ser adotado. Baseado na análise dos recursos, escolher entre uma grande variedade das ferramentas disponíveis atualmente. Não é necessário usar uma ferramenta sofisticada para produzir um resultado profissional. Como parte da análise dos recursos, deve-se determinar se há verba disponível para comprar *software* ou se deverá ser usado *software* gratuito.

Passo 6 - Começar com visualização

- Ser criativo na sua produção assim como na visualização propriamente dita.
- Lembrar: deve-se “contar” uma história: Dados + Análise + Design = melhor entendimento dos dados para os usuários.
- Experimentar com poucas visualizações para se acostumar ao processo e, então, decidir quais opções podem ser usadas e quais devem fazer parte das visualizações finais.
- Verificar os recursos de visualização existentes – muitos são sofisticados demais, mas muitos podem ser usados com pouco treinamento e poucos recursos. Revisar exemplos de outras organizações para conhecer o processo seguido por eles.
- Não pensar que os primeiros esforços terão resultados perfeitos; começar de modo simples e aos poucos, incrementalmente, acrescentar mais e mais elementos às visualizações.
- Procurar *feedback* dos cidadãos e outros órgãos para identificar o que está funcionando bem e o que não está funcionando. Flexibilizar ao fazer mudanças.
- Certificar-se de incorporar cláusulas de retratação (disclaimers) em relação às fontes de dados, assim como muitas instituições governamentais fazendo com seus *websites*. A referência, autor das recomendações, possui em seu Apêndice II alguns exemplos de cláusulas deste tipo (STOWERS, 2013).
- Finalmente, uma vez que esta é uma área em rápida atualização, os membros da equipe devem registrar *websites* relevantes de visualização de dados e acompanhar constante e periodicamente o desenvolvimento. Deve-se seguir a literatura formal e as publicações pragmáticas da

área para aprender novas propriedades e tendências atuais.

Mesmo seguindo esses passos, é preciso reconhecer que nem todos usuários da informação são sofisticados a ponto de saberem interagir com visualizações ou interpretá-las de maneira direta. É fundamental que a visualização adotada *não* seja a única forma de exibição da informação, mas apenas um dos canais possíveis para isso, mantendo visões alternativas da mesma informação em diferentes formatos. Essa versatilidade é difícil de manter, pois implica dinamicidade de atualização e sincronismo entre todas as formas de exibição, que devem ser compatíveis (exportar ou converter facilmente de uma para outra) e consistentes (todas exibindo o mesmo conteúdo, embora em formas variadas).

Este capítulo apenas introduziu as tecnologias *Big Data*, Visualização de Informações e *Visual Analytics*. Estas devem ser mais bem conhecidas para serem concretamente aplicadas no suporte a políticas públicas, uma vez que servem para lidar com quantidade, diversidade e complexidade dos dados públicos. Embora não sejam uma solução “mágica”: é preciso investigar quais dessas tecnologias são mais apropriadas em determinados contextos e quais são mais adequadas para atingir certos objetivos e – obviamente – é preciso capacitar as pessoas que vão utilizá-las para se obterem resultados realmente úteis.

Fundamentalmente, este capítulo pretende – com seus conceitos, ideias e, principalmente, com a discussão – alimentar um processo de reflexão estratégica sobre como desenvolver e aplicar o enorme potencial dessas tecnologias no setor público.

REFERÊNCIAS

AUER, S. R.; BIZER, C.; KOBILAROV, G.; LEHMANN, J.; CYGANIAK, R.; IVES, Z. DBpedia: A Nucleus for a Web of Open Data. In: ABERER, K. et al. **The Semantic Web**. 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference, ISWC 2007 + ASWC 2007, Busan, Korea, November 11-15, 2007. Proceedings. Lecture Notes in Computer Science 4825. Berlin/Heidelberg: Springer, 2007, p. 722-735.

CARD, S.; MACKINLAY, J. SHNEIDERMAN, B. **Readings in Information Visualization** – Using Vision to Think. Burlington: Morgan Kaufmann, 1999.

CHAN, Joseph. An Architecture for Big Data Analytics. **Communications of the IIMA**, vol. 13, n. 2, p. 1-14, 2013. Disponível em: <http://www.iima.org/index.php?option=com_phocadownload&view=category&download=432:anarchitectureforbigda-

taanalytics&id=77:2013-volume-13-issue-2&Itemid=68>. Acesso em: 10 set. 2014.

D3.org. Data-Driven Documents. **Biblioteca D3**. Disponível em: <<http://d3js.org/>>. Acesso em: 30 abr. 2014.

DESROCHERS, P. Visualizing Open Government: Case Study of the Canadian Recordkeeping Approach. In: WOOD, D. (Ed.) **Linking Government Data**. Berlin/Heidelberg: Springer, 2011. pp. 155-180.

EAVES, David. The Three Laws of Open Government Data. **eaves.ca**, 30 set. 2009. Disponível em: <<http://eaves.ca/2009/09/30/three-law-of-open-government-data/>>. Acesso em: 30 abr. 2014.

EDWARDS, K. **Visualizing Data from Government Census and Surveys**: Plans for the Future. Report of Censuses and Surveys of Governments: A Workshop on the Research and Methodology Behind the Estimates. US Department of Commerce: United States Census Bureau, 2012. Disponível em: <http://www2.census.gov/govs/pubs/research_reports/>. Acesso em: 30 abr. 2014.

GARTNER INC. **Pattern-Based Strategy**: Getting Value from Big Data. Gartner Group press release. Junho 2011. Disponível em <<https://www.gartner.com/doc/1727419/patternbased-strategy-getting-value-big>>. Acesso em: 30 abr. 2014

GRAVES, A.; HENDLER, J. Visualization Tools for Open Government Data. In: DG.O 2013. **Proceedings of the 14th Annual International Conference on Digital Government Research**. New York: ACM, 2013. Disponível em: <<http://dl.acm.org/citation.cfm?id=2479746&dl=ACM&coll=DL&CFID=559886329&CF-TOKEN=91690446>>. Acesso em: 10 set. 2014.

IBM Corporation. International Business Machines Corporation. Addressing government challenges with big data analytics. **IBM Software White Paper**. Somers: IBM Corporation, 2013. Disponível em: <<http://www-01.ibm.com/common/ssi/cgi-bin/ssialias?infotype=SA&subtype=WH&htmlfid=IMW14699USEN>>. Acesso em: 10 set. 2014.

LINDQUIST, E. **Grappling With Complex Policy Challenges**: Exploring the potential of visualization for analysis, advising and engagement. Canberra: Australian National University, 2011. Disponível em: <https://www.anzsog.edu.au/media/upload/publication/108_Lindquist-Discussion-paper-from-roundtable-2.pdf>. Acesso em: 10 set. 2014. (HC Coombs Policy Forum Discussion Paper).

OGP. Open Government Partnership. Parceria para Governo Aberto. **O que é a iniciativa?** Disponível em: <<http://governoaberto.cgu.gov.br/a-ogp/iniciativa.asp>> Acesso em: 30 abr. 2014.

_____. Parceria para Governo Aberto. **Planos de Ação**. Disponível em: <http://governoaberto.cgu.gov.br/no_brasil/plano-brasileiro/index.asp>. Acesso em: 30 abr. 2014.

_____. **Parceria para Governo Aberto. 2º Plano de Ação Brasileiro para Governo Aberto.** Disponível em: <http://governoaberto.cgu.gov.br/no_brasil/plano-brasileiro/segundo-plano/index.asp>. Acesso em: 30 abr. 2014.

PAULA, M. M. V.; RIBEIRO, F. C.; CHAVES, M.; RODRIGUES, S. A.; SOUZA, J. M. de. A Visualização de Informação e a Transparência de Dados Públicos. In: SIMPÓSIO BRASILEIRO DE SISTEMAS DE INFORMAÇÃO, 7, 2011, Salvador. **Anais eletrônicos...** Biblioteca Digital Brasileira de Computação: UFMG/SBC, 2011. Disponível em: <<http://www.lbd.dcc.ufmg.br/colecoes/sbsi/2011/avizualizacao.pdf>>. Acesso em: 10 set. 2014.

SATHI, A. **Big Data Analytics: Disruptive Technologies for Changing the Game.** Boise, ID: MC Press, 2012.

STOWERS, G. **The Use of Data Visualization in Government.** Washington: IBM Center for The Business of Government, 2013. Disponível em: <<http://www.businessofgovernment.org/sites/default/files/The%20Use%20of%20Visualization%20in%20Government.pdf>>. Acesso em: 10 set. 2014.

RIBEIRO, F. C.; FERREIRA, T. P.; PAULA, M. M. V.; CHAVES, M.; RODRIGUES, S. A.; SOUZA, J. M de; FRANZOSI, E.; OLIVEIRA, L. F. VisPublica: uma proposta para aprimorar a transparência de dados públicos. In: SIMPÓSIO BRASILEIRO DE SISTEMAS DE INFORMAÇÃO, 8, 2012, São Paulo. **Anais...** São Paulo: SBSI, 2012a.

RIBEIRO, F. C.; PEREIRA, I. R. M.; MACEDO, F. J.; PAULA, M. M. V.; RODRIGUES, S. A.; SOUZA, J. de. Uma investigação das técnicas de visualização como mecanismo para apoiar a transparência de dados públicos. In: WORKSHOP MINEIRO DE SISTEMAS DE INFORMAÇÃO, 3, 2012, Juiz de Fora. **Anais do I WMSI 2012 - III Simpósio Mineiro de Computação.** Juiz de Fora: SMC, 2012b.

RIBEIRO, F.C.; CAETANO, B. P.; PAULA, M. M. V.; CHAVES, M.; SILVA, V. F.; RODRIGUES, S. A.; SOUZA, J. M. Heurísticas para Visualização de Dados. In: SIMPÓSIO BRASILEIRO DE SISTEMAS DE INFORMAÇÃO, 9, 2013, João Pessoa. **Anais...** João Pessoa: SBSI, 2013.

THOMAS, J.; COOK, K. **Illuminating the Path: Research and Development Agenda for Visual Analytics.** IEEE Computer Society Press, 2005.

UBALDI, B. Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives. **OECD Working Papers on Public Governance**, n. 22, OECD Publishing, 2013. Disponível em: <<http://www.oecd-ilibrary.org/docserver/download/5k46bj4f03s7.pdf?expires=1410360728&id=id&accname=guest&checksum=63719952ECA84A447843DE04984F5DF3>>. Acesso em: 10 set. 2014.

WOOD, D. (Ed.) **Linking Government Data.** Berlin/Heidelberg: Springer, 2011.

YAU, N. **Visualize This – The FlowingData Guide to Design, Visualization and Statistics.** Indianapolis: Wiley, 2011.