

Projeto SPUK
Melhoria do Ambiente de Negócios
por meio da Transparência no Estado de São Paulo

Relatório Sintético

**World Wide Web Consortium – (W3C) – Escritório Brasil
Centro de Estudos sobre Tecnologias WEB – (CEWEB)**

São Paulo, 24 de março de 2015

Este relatório foi produzido por **Allan Souza Santos** e **Ricardo Matheus** para apresentar as ações e resultados da Reunião do Projeto SPUK, Melhoria do Ambiente de Negócios por meio da Transparência no Estado de São Paulo.

Este relatório foi preparado para o ***Centro de Estudos sobre Tecnologias Web (Ceweb.br)***, parte do **Núcleo de Informação e Coordenação do Ponto Br (NIC.br)** e **Comitê Gestor da Internet no Brasil (CGI.br)**, aos participantes do Projeto SPUK – Melhoria do Ambiente de Negócios por meio da Transparência no Estado de São Paulo e aos interessados na temática de dados abertos e web semântica.

Sumário

1. Sumário Executivo	5
2. Abertura do Evento	6
2.1. Mensagem de Abertura de Silvio Aquino	6
2.2. Mensagem de Abertura de Roberto Agune.....	7
3. Apresentação dos Guias de Abertura de Dados e Web Semântica	9
3.1. Apresentação Guia de Abertura de Dados – Marcos Túlio Pires.....	9
3.2. Apresentação Guia de Web Semântica – Carlos Laufer.....	11
3.2.1. Comentários da Apresentação de Marco Túlio Pires e Carlos Laufer	13
4. Apresentação das iniciativas e experiências de sucesso do Reino Unido.....	15
4.1. Apresentação Eleanor Stewart.....	15
4.1.1.Comentários da apresentação de Eleanor Stewart.....	20
4.2. Apresentação Antonio Acuña	21
4.2.1.Comentários da apresentação de Antonio Acuña.....	22
5. Apresentação dos Projetos Pilotos de Abertura de Dados do Governo do Estado de São Paulo.....	26
5.1. Piloto do Metrô de São Paulo.....	26
5.1.1.Comentários sobre Piloto do Metrô de São Paulo	27
5.2. Piloto Secretaria de Planejamento	27
5.2.1.Comentários sobre Piloto da Secretaria de Planejamento de São Paulo	28
5.3. Piloto do Tribunal de Contas do Estado (TCE)	29

5.4. Piloto da Secretaria da Saúde – Medicamentos	29
5.5. Piloto da Secretaria da Saúde – Doenças.....	30
5.6. Apresentações de Pilotos Extras.....	30
5.6.1. Professores Domingos Alves e Vinícius Mazieiro da Universidade de São Paulo de Ribeirão Preto.....	30
6. Discussão dos Projetos Pilotos	30
7. Metodologia de organização e trabalho para discussão dos Projetos Pilotos	31
7.1. Planilhas dos Pilotos dos Grupos de Discussão	32
7.1.1. Planilha do Piloto da Secretaria Estadual de Saúde	32
7.1.2. Planilha do Piloto da Secretaria Estadual de Planejamento	33
7.1.3. Planilha do Piloto do Tribunal de Contas do Estado de São Paulo.....	33
7.1.4. Planilha do Piloto do Metrô de São Paulo	35
8. Encerramento.....	36
9. Anexos	37
9.1. Lista de participantes.....	37
9.2. Apresentação Roberto Agune	40
9.3. Apresentação Guia de Abertura de Dados – Marco Túlio Pires.....	40
9.4. Apresentação Guia de Web Semântica – Carlos Laufer.....	40
9.5. Apresentação Eleanor Stewart.....	40
9.6. Apresentação Antonio Acuña	40
9.7. Planilhas Originais apresentadas pelos grupos	40

1. Sumário Executivo

Os participantes do Projeto **SPUK – Melhoria do Ambiente de Negócios por meio da Transparência no Estado de São Paulo** se reuniram no dia 24 de março de 2015 nas dependências do Núcleo de Informação e Coordenação do Ponto Br (NIC.br) e do Comitê Gestor da Internet no Brasil (CGI.br) para definir as estratégias e ações de abertura de dados de projetos pilotos do grupo de trabalho para o ano de 2015.

A oficina, coordenada pelas equipes do *Centro de Estudos sobre Tecnologias Web (Ceweb.br)*, do **Governo do Estado de São Paulo** e da **Embaixada Britânica**, teve a presença de 51 pessoas, entre representantes de entidades da sociedade civil, de empresas nacionais e multinacionais do setor de tecnologia, universidades e representantes do Governo do Estado de São Paulo, além da presença internacional de **Antonio Acuña**, *Chefe do Portal de Dados Abertos do Reino Unido* e **Eleanor Stewart**, do *Gabinete do Primeiro Ministro do Reino Unido*.

Os objetivos da reunião foram:

- Apresentar os guias de dados abertos e de web semântica aos participantes;
- Apresentar as iniciativas e experiências de sucesso do Reino Unido; e,
- Planejar as estratégias e ações para criar e viabilizar estratégias de abertura de dados nos projetos pilotos do Governo do Estado de São Paulo.

A lista de presença está anexada na seção 8.1. [Lista de participantes](#).

2. Abertura do Evento

Após a apresentação dos participantes, foi realizada uma breve introdução por José Carlos Vaz, moderador do debate. O moderador detalhou a agenda e introduziu os participantes as metodologias de moderação que seriam realizadas durante o evento, bem como os produtos esperados.

A abertura oficial da reunião foi realizada por **Vagner Diniz**, gerente Centro de Estudos sobre Tecnologias Web (Ceweb.br) e do Escritório brasileiro do W3C, e **Roberto Agune**, coordenador da Unidade de Inovação da Subsecretaria de Parcerias e Inovação da Secretaria de Governo do Estado de São Paulo.

Após a abertura foi dada a palavra a **Silvio Aquino**, gerente de projetos da embaixada Britânica.

2.1. Mensagem de Abertura de Silvio Aquino

Silvio Aquino se apresentou como gerente de projetos da embaixada Britânica e aborda a temática do *Prosperity Fund*¹. Explicou que o Prosperity Fund tem como objetivo promover a cooperação e financiamento de programas estabelecidos no Brasil fomentando estas iniciativas que estejam contempladas nas áreas centrais do programa. Destacou que as áreas centrais são: Sustentabilidade, Energia, Infraestrutura, Educação e Melhoria do ambiente de negócios.

O gerente de projetos da Embaixada Britânica, apontou que a iniciativa proposta pelo governo do Estado de São Paulo, co-financiada pelo Prosperity Fund, tem como objetivo melhorar o ambiente de negócios no Estado de São Paulo, por meio da disponibilização e reuso de dados abertos governamentais. Ele enfatizou que a embaixada britânica já teve outros parceiros na temática no nível federal, tais como o Tribunal de Contas da União (TCU) e Controladoria Geral da União (CGU). O Estado de São Paulo faz parte da estratégia de ampliação de parcerias do trabalho da embaixada na temática de dados abertos no Brasil, prevendo que sejam referências para futuras parcerias. Finalizou agradecendo o trabalho e a presença de todos.

¹ https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/378048/PUBLIC_Brazil_Prosperty_Fund_Programme_Strategy_FY2015-16.pdf .

2.2. Mensagem de Abertura de Roberto Agune

Roberto Agune, coordenação da Unidade de Inovação da Subsecretaria de Parcerias e Inovação da Secretaria de Governo do Estado de São Paulo, iniciou sua fala destacando que foi um prazer estar no evento e agradeceu a presença de todos. Apontou que o *workshop* foi um dos passos do Projeto de Cooperação entre o Governo de São Paulo e do Reino Unido, denominado SPUK. O objetivo do *workshop* é trocar experiências em diferentes setores, visando a melhoria no ambiente de negócios no Estado de São Paulo. Este projeto foi selecionado pelo *Prosperity Fund*, com investimentos de cerca de 185.460 mil libras esterlinas por intermédio do *Foreign & Commonwealth Office (FCO)*², departamento do governo britânico para relações exteriores. Por parte do Governo do Estado de São Paulo, será investida a quantia de 250 mil libras esterlinas, totalizando um montante de aproximadamente 436 mil libras esterlinas a serem aplicadas no projeto, que teve início em abril de 2014 e conclusão prevista para março de 2016.

Com isso, dois agentes implementadores participam como intermediadores: pelo lado do Brasil o Centro de Estudos sobre Tecnologias Web (Ceweb.br) e pelo lado britânico o GovRisk (The International Governance and Risk Institute), organização especializada em prevenção à corrupção, combate às crises financeiras e fiscais, além da promoção de práticas de governança para organizações públicas visando o aprimoramento dos ambientes de negócios.

Agune explicitou que há uma serie de secretarias trabalhando juntas no Estado de São Paulo, com o objetivo de melhorar o ambiente de negócios através da transparência e uso de dados abertos, empoderado a sociedade civil e combatendo a corrupção, baseado na experiência compartilhada pelos ingleses. Com isso, o projeto tem sido executado baseado nos seguintes princípios:

- Governo como plataforma;
- Dados em formato aberto;
- Linguagem cidadã;
- Combate à corrupção;
- Negócios sociais.

² <https://www.gov.uk/government/organisations/foreign-commonwealth-office>

Segundo o coordenador da Unidade de Inovação da Subsecretaria de Parcerias e Inovação da Secretaria de Governo do Estado de São Paulo, a contrapartida do Estado de São Paulo será de melhorar o portal de dados abertos, com o objetivo de ampliar as experiências hoje existentes: o Portal de Transparência do Estado de São Paulo³, iniciativa de responsabilidade da Corregedoria Geral da Administração⁴. Além dos princípios norteadores, o projeto também possui cinco temas de relevância:

- Abertura de bases;
- Web Semântica;
- Linguagem cidadã;
- Fomento a negócios sociais;
- Diálogo com a população.

Agune apontou que, no workshop, os dois primeiros temas (Abertura de Bases e Web Semântica) seriam os tópicos principais. Descreveu que os consultores contratados, **Marco Túlio Pires** e **Carlos Laufer**, identificaram o cenário britânico de abertura de bases e web semântica durante a visita realizada ao Reino Unido com o propósito de desenvolver guias envolvendo as temáticas para o governo do Estado de São Paulo, que seriam mais detalhados na sequência de sua fala.

Para Agune, as discussões sobre abertura de bases e web semântica tinha o objetivo de desenvolver formas de ampliar e disponibilizar maiores e melhores bases de dados abertos, tomando como plano de fundo os projetos pilotos das secretarias envolvidas, por meio da interlocução dos dois grupos presentes: os especialistas britânicos e os demais envolvidos nos projetos pilotos das secretarias estaduais do governo do Estado de São Paulo. Segundo Agune, com isso será possível criar uma melhor estrutura de fomento a novos negócios através dessas ações, buscando ampliar os níveis de combate à corrupção, como o caso da Lei Federal N° 12.846/2013⁵, popularmente conhecida como **Lei de Combate à Corrupção**, que na opinião de Roberto Agune, apesar de ser uma legislação recente no âmbito Federal, a sua implementação no Estado de São Paulo é destaque por ser um dos Estados pioneiros em sua aplicação.

³ <http://www.transparencia.sp.gov.br>

⁴ <http://www.corregedoria.sp.gov.br>

⁵ http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2013/lei/l12846.htm

Finalizando, Agune destacou as visitas e participações em diversos seminários envolvendo o tema no país que seu grupo tem realizado, bem como convidando outros representantes de cidades e estados para uma maior troca de experiência.

A Apresentação de Roberto Agune está anexada na seção 8.2. Apresentação Roberto Agune

3. Apresentação dos Guias de Abertura de Dados e Web Semântica

Após a apresentação de Roberto Agune, o moderador José Carlos Vaz passou a palavra para os consultores especialistas em Abertura de Dados e Web Semântica.

3.1. Apresentação Guia de Abertura de Dados – Marco Túlio Pires

Marco Túlio Pires, consultor especialista em abertura de dados, iniciou sua apresentação abordando como criou o conteúdo do guia de abertura de dados e as diretrizes que ele possuía. Frisou que o guia ainda é um trabalho em construção e justamente por esse motivo todas as críticas e sugestões apresentadas seriam bem-vindas.

Inicialmente, ao começar a construir o Guia de Abertura de Bases, Marco Túlio revelou que realizou uma visita ao Reino Unido, onde realizou entrevistas com representantes do governo, empresas e sociedade civil, com o objetivo de adquirir uma boa visão de como a temática Dados Abertos está inserida no Reino Unido, fazendo assim um mapeamento do estado-da-arte britânico e descobrindo melhores diretrizes para encaminhar o desenvolvimento do guia brasileiro.

Marco Túlio enfatizou que o Guia dialoga com diversos públicos, apresentando uma mistura de linguagem utilizada, hora mais técnica e hora mais cidadã, para constantemente dialogar com os envolvidos na concepção do projeto: Gestores Públicos - responsáveis pela idealização e provimento de serviços públicos, pelo uso e integração das informações geradas com bases em bancos de dados - e pelos especialistas-responsáveis pela gestão, tratamento e disponibilização de informações e dados das bases governamentais.

Marco Túlio contou que a primeira parte do Guia possui uma ambientação do tema, levantando exemplos ao redor do mundo, visando sensibilizar os envolvidos com exemplos positivos da abertura e utilização de dados abertos ao redor do mundo. Na sequência, foi feita uma contextualização do que são Dados Abertos. Envolvendo conceitos de:

- Disponibilidade e acesso;
- Reutilização e redistribuição;
- Participação universal.

Segundo Marco Túlio, os benefícios da abertura de dados são para:

- O próprio governo;
- O avanço técnico/científico
- O setor privado
- As Organizações da Sociedade Civil;
- O maior controle e participação social através de:
 - Transparência;
 - Combate à corrupção;
 - Legitimidade do estado.

O especialista ainda trouxe com destaque algumas iniciativas que utilizam Dados Abertos ao redor do mundo:

- *Citymapper App*⁶: aplicativo britânico que está em processo de expansão, que utiliza dados públicos de trânsito e faz atualizações em tempo real para o transporte público;
- *QEdu*⁷: ferramenta criada pela Fundação Lemann, que utiliza dados do Índice de Desenvolvimento da Educação Básica (Ideb) e da Prova Brasil para criar indicadores de educação no país;

⁶ <https://citymapper.com/london>

⁷ <http://www.qedu.org.br/>

- *Prescribe Analytics*⁸: sistema desenvolvido pelo Sistema Nacional de Saúde Britânico (National Health Service - NHS) que através da utilização de dados cria indicadores referentes ao sistema de saúde, tais como os medicamentos mais receitados, espaços para a utilização de genéricos, além de outras funcionalidades.
- *Sunlight Foundation*⁹: fundação norte-americana que atua em três principais áreas: *lobby* junto ao poder legislativo; produção de conteúdo técnico, e; desenvolvimento de tecnologia para dados abertos.

Ele apontou também que há iniciativas de Abertura de Dados no contexto legal, por meio da Declaração Universal dos Direitos Humanos, da Legislação nacional, no caso a Lei de Acesso à Informação e também através de legislações regionais, como a Legislação paulista, por meio do Governo Aberto de São Paulo. Destacou também a importância de planejar e estruturar a abertura de dados, por meio de estratégias de abertura, consultando atores envolvidos (Guardiões das bases de dados, pessoal encarregado pela Tecnologia da Informação, técnicos e sociedade civil organizada), bem como do uso de metodologias de abertura, como a dos Dados Abertos Conectados (Cinco Estrelas)¹⁰.

Para saber mais, a apresentação está na seção 8.3. Apresentação Guia de Abertura de Dados – Marco Túlio Pires

3.2. Apresentação Guia de Web Semântica – Carlos Laufer

Carlos Laufer, *consultor especialista em web semântica*, começou sua fala se apresentando e apontando que atualmente é difícil definir o que de fato é web semântica. Deste modo, o Guia de Web Semântica tem sido desenvolvido de forma simples e direta ao tema. Destacou que é importante entender o que significa semântica: estudo dos significados e interpretação de uma palavra. Com isso, quando se trata de Web Semântica, se refere a um novo modelo de web, onde é possível que toda a informação existente possa ser organizada de forma que seja possível transformar a informação em um modelo de fácil entendimento para máquinas e pessoas.

Laufer também elucidou que não se trata apenas de uma nova rede de informações, mas sim de um projeto para aplicar conceitos inteligentes na internet atual. Para ele, cada informação vem com um significado bem definido e não se encontra mais solta em um mar de conteúdo, permitindo uma melhor interação com o usuário. O especialista enfatizou que o guia trata sobre Web Semântica, mas explicou que vivemos imersos em semântica e que faz parte da nossa vida cotidiana, sendo a ideia principal definir escopos do

⁸ <http://www.prescribinganalytics.com/>

⁹ <http://sunlightfoundation.com/>

¹⁰ <http://5stardata.info>

universo da web, desde seu início, com emissão de documentos em formato de aplicativos de escritório, até a internet das coisas. Para tanto, o guia é estruturado da seguinte maneira:

- WWW e Semântica;
- Ecossistema de Dados na Web;
- Web Semântica;
- Dados Conectados;
- Vocabulários e Ontologias;
- Ambientes de Desenvolvimento de Aplicações.

Carlos Laufer destacou a importância do evento pois alguns dos principais usuários do Guia de Web Semântica estavam presentes e poderiam ter ajudado na elaboração. Para Laufer, a Secretaria de Gestão, responsável por definir a implantação de certas políticas, poderia ter ajudado na escolha de formato, verificação de como os dados a serem disponibilizados poderão ser utilizados ou se serão preservados. Em um segundo momento, o Guia de Web Semântica trata da própria Web Semântica e a definição sobre qual vocabulário utilizar para passar ao público. Ele alerta que é preciso entender este ponto como um processo de comunicação, e especialmente, que ela funcione. Dessa forma, foi criado um critério de publicação de dados para definir em que parâmetros a disponibilização desses dados e informações se encontra, conforme abaixo:

- Publique a informação na Web (qualquer que seja o formato) sob um tipo de licença de dados abertos;
- Publique a informação na forma de dados estruturados;
- Use formatos não proprietários;
- Use URIs para identificar as coisas;
- Faça a conexão dos seus dados com outros dados.

Laufer apontou que o conceito de cinco estrelas é remetido para a uma melhoria de algo. Ele é uma escala de publicação de dados, e não quer dizer se ele está publicado ou não de forma satisfatória. No entanto, Laufer pergunta: “O que significa dizer que um dado está bem publicado?”. Respondeu que é o mesmo que fazer uma comunicação excelente com diferentes tipos de máquinas que possam realizar sua leitura. Para ele, o mais importante é o dado ser bem publicado, pois o maior objetivo deve ser sempre referente a boa publicação dos dados.

Para saber mais, a apresentação está na seção 8.4. Apresentação Guia de Web Semântica – Carlos Laufer

3.2.1. Comentários da Apresentação de Marco Túlio Pires e Carlos Laufer

José Carlos Vaz oferta a palavra aos participantes para que considerações e perguntas sejam feitas.

Silvio Aquino (Embaixada Britânica) apontou que no Brasil ainda existe um grande dilema envolvendo dados públicos e privados. Perguntou aos participantes como está sendo feita essa diferenciação, em especial no que se remete à construção dos guias e como é vista a questão da venda de dados.

Marco Túlio Pires responde a **Silvio Aquino**, que consideraram como dados sigilosos [privados], tudo aquilo que está descrito de acordo com a Lei nº 12.527 (Lei de Acesso à Informação – LAI)¹¹. No Estado de São Paulo, o Decreto nº 58.052, que regulamenta a LAI, segue os mesmos preceitos de dados sigilosos, então seguimos a mesma definição de dados sigilosos. Com relação à venda de dados, [ele e Carlos Laufer, os consultores] não chegaram a contemplar essa área, pois ela não é apreciada pela legislação brasileira, assim não há nada descrito.

Carlos Laufer complementou que, do que foi citado por **Marco Túlio Pires**, destaca-se o conceito de **dado privado**. Laufer explicita que ele pode ser utilizado em dois contextos diferentes: o primeiro relativo a privacidade, o segundo referente à posse do dado. Concluiu dizendo que ainda é um tema que demanda um pouco mais de aprofundamento.

Bernadette Farias Lóscio (UFPE) perguntou como é vista a questão da continuidade dos dados. Para a professora, é essencial que ela permaneça viva, com o aumento das bases e atualizações da mesma. Finalizou questionando os consultores **Carlos Laufer e Marco Túlio Pires** se isso foi pensado no guia.

Marco Túlio Pires, respondendo a **Bernadette**, as *APIs (Application Programming Interface*, em português Interface de Programação de Aplicativos), sabe-se que são as principais responsáveis por esse incremento nas bases. Sabe-se também que a

¹¹ http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm

criação e a manutenção de uma API é ainda algo muito custoso para o estado, no entanto é essencial para a continuidade da política.

Carlos Laufer complementou a resposta a **Bernadette**, somente o funcionamento de *APIs* não garante que uma base de dados esteja sempre correta e atualizada, por isso é importante prezar pela qualidade.

Vagner Diniz (W3C / CEWEB) perguntou a **Carlos Laufer**, sobre uma diferenciação entre a qualidade e os níveis de abertura do dado, tomando como base a classificação de uma a cinco estrelas. Vagner destacou que pelo *que* pôde entender, é melhor publicar um dado de boa qualidade em três estrelas, do que um de má qualidade em quatro ou cinco. Enfatizou que há ainda hoje uma grande discussão com relação qualidade dos dados publicados. Finalizou questionando quais caminhos os consultores (Carlos Laufer e Marco Túlio Pires) trilharam com relação a publicação nos guias e por quê?

Carlos Laufer, respondeu a **Vagner Diniz** que participa de um grupo do W3C Brasil sobre publicação de dados na *web*, e essa escala de cinco estrelas, para ele é uma forma que é possível mensurar o trabalho ao longo do processo. Carlos explicitou que esse é um mundo completamente novo. Explicou que se puder abrir o dado e fazer da melhor forma é um caminho, e não apenas publicar dados sem qualidade. Destacou que quando falou de semântica, falou de qualidade do dado, para que ele seja interpretado da melhor forma possível.

Marco Túlio Pires, responde a **Vagner Diniz** apontando que não foi impressa a opinião pessoal dos dois consultores no Guia, então se sabe que é importante tomar a iniciativa e dar incentivo à publicação, como estímulo para melhorias sequenciais aos dados. Ele disse acreditar que é mais valioso publicar, e ao longo do processo aplicar melhorias, pois do contrário, se esperarmos para fazer tudo corretamente, as coisas se tornam demasiadamente lentas.

José Carlos Vaz inaugurou o intervalo de 20 minutos, e no retorno iniciou a fala dos convidados especialistas em dados abertos do governo britânico,

4. Apresentação das iniciativas e experiências de sucesso do Reino Unido

José Carlos Vaz iniciou a oficina após todos retornarem do intervalo. Concedeu a palavra a **Eleanor Stewart**, *diretora de Transparência do Gabinete do Primeiro Ministro Britânico*.

4.1. Apresentação Eleanor Stewart

Eleanor Stewart, *diretora de Transparência do Gabinete do Primeiro Ministro Britânico*, iniciou sua fala se apresentando e comentou estar envolvida a tempo com o tema de Dados Abertos. Por isso, apresentaria aos presentes a experiência do Reino Unido nesse processo, mostrando lições aprendidas, obstáculos enfrentados e diretrizes futuras com relação à temática em seus país.

Segundo Eleanor Stewart, atualmente o Reino Unido é considerado um país extremamente aberto no que diz respeito ao acesso de dados e informações públicas. Há uma grande comunidade que faz uso das informações disponibilizadas, para as mais variadas formas e temáticas:

- Melhoria de serviços públicos;
- Oportunidades de negócios;
- Desenvolvimento social;

No entanto, Stewart destaca que isso nem sempre foi assim. No passado, o Reino Unido ficou marcado como um país muito ruim no que diz respeito a guarda de dados e informações. A falta de registro de prisioneiros durante a 2ª Guerra Mundial e a queima de informações e dados a respeito de incursões britânicas em países africanos ao longo do século XXI são apenas alguns desses exemplos.

A diretora de transparência explicou que a corrupção no Reino Unido ainda é um tema evitado na agenda governamental. No entanto, entre 2007 e 2008, gastos excessivos realizados por parte do Parlamento inglês foram alvo de um grande escândalo, que somente pôde ser descoberto através da utilização massiva de dados divulgados pelo poder público. Essa foi uma mostra da mudança na direção dos rumos da mídia, e a apresentação do termo dados abertos como ferramenta de pressão por parte da mídia.

Stewart também lembrou que quando foi iniciado o processo de abertura de dados, no ano 2000, não havia nada de concreto em relação ao tema, com algumas poucas experiências em nível local e internacional servindo como parâmetros para o desenvolvimento do projeto. Contudo, somente após uma alteração na legislação em vigor, com relação à utilização de informação pública, percebeu-se que havia naquele momento um excesso de burocracia que impedia os cidadãos a terem acesso livre as informações produzidas pelo seu governo. Esse processo coincidiu diretamente com a fase de imersão da Web 2.0, onde junto à grande expansão de uso de redes sociais foi um pontapé para os cidadãos demonstrarem seu descontentamento com os serviços prestados pelo governo.

Para ela, só foi possível mapear, dialogar e mensurar parte do descontentamento somente a partir das redes sociais. A partir disso, foi possível desenvolver quatro principais pilares em relação ao processo de abertura:

- **Open Information (Abertura de Informação):** para ter uma voz eficaz, as pessoas precisam ser capazes de compreender o que está acontecendo em seus serviços públicos. Para tanto, o governo deve publicar informações sobre serviços públicos de forma simples de ser encontrada e reutilizada;
- **Open Innovation (Abertura de Inovação):** para promover a inovação em serviços públicos on-line de modo a responder às novas expectativas geradas pela população;
- **Open Discussion (Discussão Aberta):** para promover maior envolvimento com o público através de consultas on-line, o que gera maior interação e colaboração entre as partes envolvidas no processo;
- **Open Feedback (Retorno da População):** essa abertura trata da parte final do processo, onde o público deve ser capaz de ter uma palavra a dizer sobre os serviços utilizados, servindo como parâmetros para avaliar o que vem sendo realizado.

Assim, em janeiro de 2009, Stewart pontuou que foi lançado o portal de dados abertos do governo britânico, o data.gov.uk¹², e durante aquele período, perceberam que essa é uma ferramenta importante de comunicação com a população, uma vez que aquele cenário, pós Crise Econômica Mundial de 2008, demandava um maior nível de diálogo entre governo e sociedade, principalmente utilizando a informação pública como ferramenta desse elo. Com isso iniciaram uma nova abordagem, convidando a população a participar através da Web. Stewart contou que foi criado o projeto *Show us a Better Way*¹³, uma iniciativa que buscou aliar população, desenvolvedores e demais interessados a utilizar os dados públicos que estavam sendo disponibilizados de modo a

¹² <http://data.gov.uk/>

¹³ <http://www.showusabetterway.co.uk/call/2008/11/and-the-winners-are.html>

reutilizar as informações de uma forma diferente da utilizada inicialmente, visando a implementação de melhorias nos serviços públicos implementados.

Stewart de transparência britânica contou que já recebeu diversos tipos de abordagens e ideias com relação ao uso dos dados. Foram sugeridos diferentes tipos de mapeamento de informações, envolvendo desde a localização de banheiros públicos, matrículas em escolas públicas ou o uso dos dados dos sistemas de transporte público. O programa hoje oferece cerca de 20 mil libras esterlinas em premiações para as melhores inovações colocadas em prática. Eleanor adicionou que essa experiência foi o pontapé inicial para o primeiro *Government Hack Day*¹⁴, evento onde desenvolvedores, gestores e população podiam trabalhar juntos para mostrar como, através de aplicativos tecnológicos, problemas do cotidiano poderiam ser resolvidos. No entanto, ela alertou que nem tudo durante essas experiências saiu conforme o planejado. Foram encontradas muitas barreiras, em especial envolvendo licenças de uso das aplicações criadas, uma vez que precisavam de autorização do governo para usar qualquer tipo de informação, o que acabou atrasando inicialmente a utilização dos projetos desenvolvidos.

Stewart explicou que essa foi uma fase importante, pois conseguiram mensurar que a experiência que estavam implantando era de fato positivo, superior inclusive a que àquela altura estava sendo implementada nos EUA¹⁵. Ela adicionou que o que de fato tornava a experiência britânica positiva era o fato da reutilização da expertise, em especial por parte de pessoas que estavam fora do setor governamental. Ela enfatizou estar liberando, de fato, o tipo de conteúdo e informações que os usuários buscavam, e o mais importante, por toda a utilização de dados, foi feita através de uma abordagem interativa, em diálogo constante com a comunidade de dados abertos, observando o que estava ou não funcionando. Essa foi a principal lição que aprendemos no processo.

Relembrando o passado, Stewart descreveu que em 2009, foi percebido que a equipe de governo que tratava de dados abertos ainda não possuía uma boa infraestrutura para o uso de dados, uma vez que boa parte do que era disponibilizada ainda estava em formato .PDF (Portable Document Format). Ela confessou que o PDF não é de boa qualidade para reutilização das informações, ou seja, toda a informação que eles possuíam não era relevante para produção estatística, o que tornava as ações irrelevantes. Com isso, tiveram auxílio de diversos outros atores que participaram dessas discussões, desde organizações da comunidade de dados abertos, mídia, e a população em geral para definir as melhores formas de disponibilizarmos esses dados. Um dos exemplos na utilização estatística e mapeamento dos dados disponibilizados se encontra a seguir.

¹⁴ <http://opengovhackday.pbworks.com/w/page/12978145/FrontPage>

¹⁵ https://www.whitehouse.gov/the_press_office/TransparencyandOpenGovernment/

Stewart exemplificou o mapeamento de comportamentos antissociais em vizinhanças, o que mostra o grande grau de diversificação no uso de dados (Figura 1).

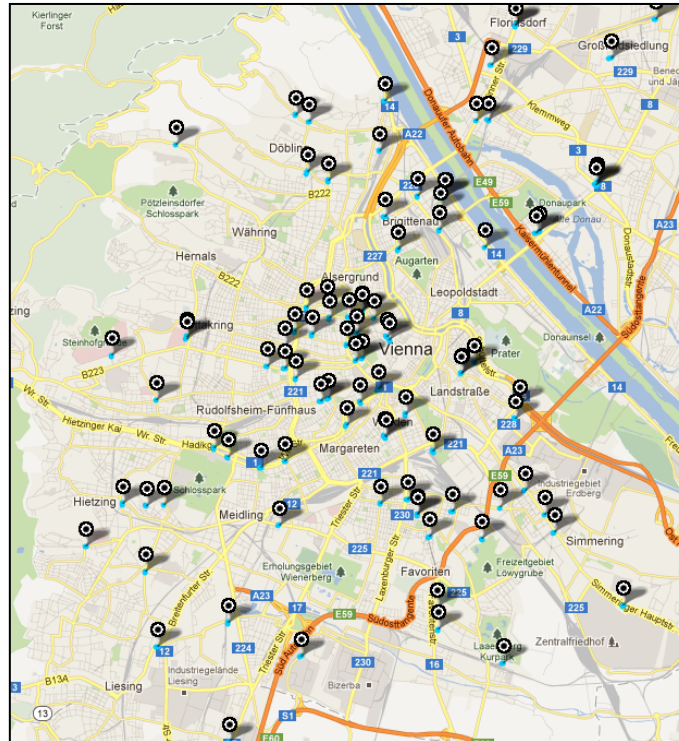


Figura 1 – Mapeamento de comportamentos antissociais. Fonte: Eleanor Stewart (2015).

Stewart explicou que após a eleição do atual primeiro-ministro, o conservador David Cameron, em 2010, houve uma diferenciação na cultura de dados abertos, fazendo com que o tema transparência entrasse em pauta de forma evidente na agenda governamental. Para ela, pode-se dizer que essa agenda ganhou um aspecto realmente fixo pelo fato da existência do que chamamos de “dados privados”, em outras palavras, dados relativos a perfis socioeconômicos de pessoas ligadas às atividades governamentais que antes tinham suas informações anônimas. O grande pontapé para essa onda foi o escândalo referente a

descoberta de 172 servidores¹⁶ de funções públicas que na época possuíam salários superiores aos recebidos pelo primeiro-ministro.

Com a ampliação do uso dos dados abertos, foi exponencial o nascimento de novos portais que passaram a fazer uso dessas informações nas mais diferentes áreas. Stewart comentou que as pessoas passaram a buscar saber como estavam alocados os investimentos públicos, como os impostos seriam investidos, e de certo modo passaram a junto ao governo se tornarem responsáveis por suas ações. Após isso, Stewart explicou que o governo passou a ter estatísticas mais precisas, especialmente relacionadas às quantidades de crimes por ruas (acidentes de carro, tentativas de assalto, etc.), e aumentaram o policiamento nas áreas necessárias, o que mostra as oportunidades geradas na alocação de recursos após o uso dos dados abertos.

Exemplificou que na área de saúde, por exemplo, passamos a realizar pesquisas com diferentes profissionais e verificamos que é importante entender de onde esses recursos são extraídos, para onde eles vão, se o paciente está em uma lista de espera, daqui a quanto tempo, então, há todo um monitoramento no setor da saúde também. No setor de transporte público, a cidade de Londres gastou menos de um milhão de libras esterlinas para liberar todos os dados existentes, onde cerca de 500 mil aplicativos (*apps*) foram desenvolvidos a partir dessas bases de dados, envolvendo os mais diferenciados mapeamentos da cidade. A resposta veio ao fim de 2013, pesquisas mostraram que desde o processo de abertura de dados, ainda em 2010, houve uma economia de tempo média de 58 minutos por parte dos usuários do sistema de transporte que passaram a utilizar esses aplicativos para economizar tempo. Isso mostrou que o processo de abertura de dados gerou impactos significantes, como ganho de tempo, dinheiro e qualidade de vida, e tem funcionado muito bem.

Stewart também deu como exemplo o próprio governo, o qual apontou ser também um grande usuário dos dados abertos. A cidade de Manchester calcula que está economizando cerca de 6,5 milhões de libras esterlinas por ano com a utilização desses dados em serviços oferecidos por *apps* a partir dos dados abertos. Ela alertou os participantes de que é preciso criar lideranças nas organizações, visando criar uma nova visão do uso de dados abertos. Também se deve visualizar as demandas dos cidadãos, focar nos dados de coisas que as pessoas se preocupam, tais como transporte público, localização de banheiros, serviços de saúde locais, são apenas alguns dos exemplos do que de fato a população em melhorias, fazendo com que as abordagens sejam feitas de maneira ser clara e apropriada.

Finalizando, ela argumentou que no Reino Unido temos uma lei de proteção de dados muito rigorosa, mas como indivíduos devemos ter ciência dessas responsabilidades, por isso gosta de ver como cada cidade ou país trabalham esse cenário. São

¹⁶ <http://www.dailymail.co.uk/news/article-1282999/172-civil-servants-paid-PM.html>

grandes deságios lidar e manipular as informações, pois as pessoas têm medo de liberar seus dados sem uma última análise, as pessoas evitam porque não querem ver seus dados sendo liberados. Tudo isso deve ser feito passo a passo, e não de uma vez, deve ser feito de maneira incremental. Isso ajuda a alavancar um ecossistema de dados abertos.

A apresentação está anexada na seção anexos 8.5. [Apresentação Eleanor Stewart](#).

4.1.1. Comentários da apresentação de Eleanor Stewart

Daphne Abreu Souza (Tribunal de Contas do Estado de São Paulo) questionou a dupla de especialistas do Governo Britânico (**Eleanor Stewart** e **Antonio Acuña**) qual era a metodologia utilizada para disponibilização de dados.

Eleanor Stewart, respondeu a questão de **Daphne Abreu Souza** dizendo que para promover o interesse dos cidadãos, o governo britânico possui uma mídia muito ativa, desde o início do processo foi lançado alguns dados que seriam significativos para os cidadãos. Stewart deu o exemplo do uso de jornais para avisar de sua utilização, algo que chamamos de jornalismo de dados. A britânica disse que trabalha fortemente com usuários das comunidades de dados. Ela descreveu que em Londres existe uma área em repleta de indústrias de tecnologia e lá foi criado o Instituto de Dados Abertos (*The Open Data Institute - ODI*)¹⁷, para que as pessoas possam criar negócios baseados em nossos dados.

Stewart também deu o exemplo em sua família, onde a avó utiliza os dados do Ministério da Saúde, buscando, através das aplicações criadas a partir das bases de dados que passaram a ser disponibilizadas, informações referentes a serviços que ela pode utilizar. Segundo ela, essa é uma forma de empoderar a população no processo de decisão através do uso de dados abertos. Apesar de ser uma mulher idosa e ser muito inteligente, ela jamais poderia tratar diretamente os dados da forma que disponibilizamos, e isso evidencia a importância do tratamento e reutilização de forma clara das informações que disponibilizamos.

Stewart também alertou que o Reino Unido é um caso interessante no que tange a utilização de dados abertos, pois os cidadãos se acham no direito de fazer parte. Ela contou que os britânicos veem a informação como uma forma de participação direta. Ela disse que passou muito tempo em países do leste europeu, como Croácia, Bósnia, e outros nos Balcãs, e vi que eles estão em uma fase insipiente, o que me lembra o Brasil. Vendo que a informação é uma forma de desafiar o governo.

¹⁷ <http://opendatainstitute.org>

Ela finalizou dizendo que dentro do Reino Unido, as autoridades têm ainda certo medo de liberar certos tipos de informação, mas essa é uma barreira que já foi transposta, então é um caminho sem volta. Diferente do que ocorre na Croácia, já que a população ainda teme utilizar essa informação, pois não sabe o que o governo poderia fazer.

Após a apresentação e comentários da apresentação de **Eleanor Stewart**, o moderador **José Carlos Vaz** concedeu a palavra a **Antonio Acuña**.

4.2. Apresentação Antonio Acuña

Antonio Acuña, diretor do data.gov.uk¹⁸, iniciou sua fala parabenizando os guias de abertura de dados e de Web Semântica que estão sendo desenvolvidos destacando a importância dessas ações. Sua apresentação foi fundamentada no processo de abertura de dados feito em dez passos. Para tanto, cada etapa segue padrões de desenvolvimento conforme abaixo:

1. **Necessidade de Orientação Prévia:** é necessário saber de antemão como, quando e por onde iniciar o processo de abertura, uma vez que as pessoas precisam de orientações claras e diretas nesse sentido para saber onde caminhar. Nesse sentido, para a melhor realização desse passo é preciso que três etapas sejam bem concluídas:
 - a. Definir o processo de publicação dos dados;
 - b. Desenvolver um modelo de maturidade para dados abertos;
 - c. Explicar os princípios básicos envolvendo o tema.
2. **Framework de Dados:** é necessário ter estrutura para organização dos dados, de acordo com os níveis de serviços. Sem uma estrutura adequada para a organização dos dados, os responsáveis por alimentarem os sistemas passam a fazê-lo da forma como preferem, gerando assim problemas pois as bases de dados não conseguem conversar entre si.
3. **Padronização:** a coleta e armazenamento de informações devem ser feitas de forma organizada e padronizada para que não ocorram conflitos.
4. **Qualidade dos Dados:** é importante que os dados respeitem certos níveis de qualidade para sua utilização. A falta de CSVs (Comma-Separated Values, em português para “Valores separados por vírgulas”) adequados, torna a informação irrelevante.

¹⁸ <http://data.gov.uk>

5. **Implementação de Princípios Semânticos na Fonte:** necessidade que exista uma definição dos princípios para utilização dos dados e códigos para que os mesmos possam ser utilizá-los consistentemente. Linked Data é uma forma estruturada de manipulação que só pode ser utilizada para dados já estruturados.
6. **Vocabulário:** o vocabulário envolvendo a temática deve ser introduzida no cotidiano das pessoas de forma fluída, e não de forma obrigatória.
7. **Portais Devem ser Utilizados Para Experimentação de Linked Data:** responsáveis por portais de dados portal tem o poder de criar semântica e para isso não devem ter medo de realizar experiências.
8. **Todos os Hotéis Não Precisam ser 5 Estrelas, Assim Como Dados:** É importante ter em mente que todos os dados devem ser brutos, viáveis e confiáveis, mas nem todos os dados precisam ser cinco estrelas, é preciso trabalhar com metas reais, para em seguida aperfeiçoar sua disponibilização.
9. **Estratégia Digital:** muitos de seus problemas relativos à publicação de dados é referente a disponibilização dos mesmos. Deve-se pensar em estratégias que englobem o sistema por inteiro.
10. **Não Tenham Medo de Tentar.**

A apresentação de **Antonio Acuña** está anexada nos anexos 8.6. [Apresentação Antonio Acuña](#).

Neste momento, o moderador, **José Carlos Vaz** convida a todos para o intervalo do café.

4.2.1. Comentários da apresentação de Antonio Acuña

Marco Túlio Pires, perguntou a **Antonio Acuña** sobre os desafios que enfrentou ao longo do processo de abertura de dados que ele liderou no Reino Unido.

Antonio Acuña, respondeu a **Marco Túlio Pires** que não poderia dizer que foram poucos desafios. Enfatizou que a parte especial foi da persuasão, em termos de estratégias, pois foi um processo um tanto quando complicado, segundo Acuña. Para ele, foi preciso utilizar ideias de inovação para mostrar os resultados positivos que poderiam ser atingidos para a expansão do projeto. Desta forma, Acuña direcionou os esforços basicamente para:

- Alocação de recursos e finanças;
- Tentamos trabalhar de forma direta e individual com gestores e diretores de diferentes departamentos para que pudessem entender as necessidades e benefícios do processo de abertura de dados;
- Existe até hoje um estigma do governo em abrir dados, em especial sob a alegação de autoproteção;
- Ainda sofremos certa resistência, em especial com integrantes de alto escalão do governo. Então, esse processo de convencimento ainda é algo existente, mesmo com a política de certa forma bem estabelecida.

Deu o testemunho de que ainda passaram a fazer parte de algumas iniciativas internacionais visando a maior abertura de dados e informações, algo que passou a dar maior reconhecimento aquilo que estavam fazendo, o que ajudou muito no processo de abertura e de persuasão.

Carlos Laufer agradeceu a Antonio Acuña pela explanação feita por ele e por Eleanor Stewart. Fez dois questionamentos em seguida: existe algum documento oficial a respeito dos dez passos para a abertura de dados utilizado [por Antonio Acuña e descrita no 4.2- Apresentação Antonio Acuña]? Se fosse positiva a resposta da questão acima a Acuña, Carlos Laufer perguntou qual a licença utilizada e também se isso poderia ser mais um Guia a servir para os projetos que estão sendo elaborados.

Laufer finalizou perguntando se em relação aos dez passos, você possui algum tipo de ferramenta que os apoie ao longo do processo de abertura?

Antonio Acuña respondeu a Carlos Laufer de que enquanto estavam falando, no Reino Unido, a equipe do data.gov.uk está trabalhando na forma como essa estrutura de apoio será lançada. Acredito que já na semana que vem teremos algo publicado. Ele explicou que esse modelo de maturidade de dados abertos mostrará a necessidade de se avaliar os estágios ao longo do processo e, buscando assim indicar o estágio que cada uma das iniciativas de abertura de dados se encontra. Segundo Acuña, sua equipe está trabalhando pesado nisso e na última quarta (18/03/2015), concluíram as diretrizes, e esperam em breve obter os registros

necessários para então trazeremos tudo isso para a estrutura de configuração de dados, e podemos fazer isso de maneira que os compilamos. Com isso possuem um conjunto de ferramentas e avaliações, que nos ajudam a fazer as publicações ou não. Temos certificado de publicação de dados, e relatórios que indicam o que está apropriado ou não os responsáveis por publicamos certos dados.

Tarcila Peres Santos perguntou aos especialistas do governo britânico (**Eleanor Stewart** e **Antonio Acuña**) se existe atualmente uma estratégia desse processo desenhada. Se sim, onde seria possível fazer um desenho desse plano e nos localizarmos ao longo das etapas, pois inicialmente seu projeto tinha uma ideia de promover capacitação dispersa ao longo de todo o estado. Depois pensaram na estratégia do Reino Unido que optou por descentralizar as ações. Então Tarcila disse que gostaria de saber como estruturar essa estratégia de abertura de dados.

Antonio Acuña respondeu a **Tarcila Peres Santos** que lembrou a inexistência de um consenso a respeito da melhor estratégia utilizada, pois isso ainda é algo que se busca. Para ele, de grosso modo, as pessoas ainda não sabem claramente do que estão falando quando se referem à abertura de dados. E elas não entendem porque ainda não tem conhecimento específico com relação ao tema. Questionou então como buscar o apoio para chegar onde se quer. Respondeu que essa foi uma pergunta difícil de chegar a uma solução e de que então se criou um grupo com os melhores profissionais envolvidos com a temática, e então, ainda como seus funcionários, os enviaram para diversos setores governamentais, como o Ministério da Saúde, Educação e Transportes por cerca de três meses. Então, a cerca de cada mês, eram realizadas reuniões-chave com esses funcionários para avaliarmos como o processo estava sendo desenvolvido nessas diferentes áreas. Acuña explicou que essa foi a forma que foi encontrada para mensurar o andamento dos processos fora do nosso eixo de ação, uma vez que ao final de cada mês poderia ser visualizada mudanças, dificuldades, ou onde se deveria focar as próximas ações.

Bernadette Farias (UFPE) perguntou a **Antonio Acuña**, se no Reino Unido existe alguma ferramenta para rastrear o uso dos dados disponibilizados.

Antonio Acuña respondeu a **Bernadette Farias (UFPE)** que sim, todos os conjuntos de dados disponibilizados pelo portal de dados abertos do governo britânico permitem que os usuários postem comentários. Adicionou que também existe uma sessão onde podem ser realizados fóruns de debates e as pessoas podem falar a respeito dos diferentes aspectos dos dados: suas funcionalidades, aplicações ou ainda pedir por mais dados. Acuña destacou que trabalham com grupos independentes para criar novos grupos de dados a serem publicados, visando a ampliação das bases. No entanto, alertou de que é preciso do registro, uma vez que não queremos abusos, o que impede o anonimato, mas somente para isso, do contrário não fazemos rastreamento. Para

ele, essas são as formas de comunicação com os usuários, buscando maior engajamento, muito mais no sentido de ouvir o que eles têm a dizer.

Bernadette Farias (UFPE) repete a pergunta a **Acuña**, explicando que quando perguntou a respeito do rastreamento, era referente à forma de cobertura da utilização do dado após ele ser disponibilizado.

Antonio Acuña respondeu que não rastreiam o uso dos dados dessa forma. O objetivo é dar a oportunidade das pessoas mostrarem como querem utilizar aquele dado. É algo voluntário, no momento que ele é baixado, ele se perde, não é mais rastreado.

Edson Sales Júnior (TCE/SP) perguntou aos especialistas do governo britânico **Eleanor Stewart** e **Antonio Acuña** se eles trabalharam com bases heterogêneas de dados nos mais diferentes órgãos e questionou se eles possuem CMS (Content Management System, em português Sistema de Gerenciamento de Conteúdo) padrões ou se cada base exige um modelo diferente de aplicação.

Antonio Acuña esclareceu a **Edson (TCE/SP)** de que é importante dizer que não estavam atentos aos dados, não os armazenam, apenas catalogam. O portal de dados abertos do governo do Reino Unido (data.gov.uk) é na maior parte do tempo apenas um provedor dos dados, pois não estão exatamente preocupados com o servidor do outro lado. Para ele, nesse modelo é mais fácil, pois pode-se focar na captação e indexação dos dados. Assim simplesmente, deixam os servidores e provedores serviços livres para utilizar da forma que melhor entenderem.

Carlos Laufer perguntou a **Antonio Acuña** se os participantes teriam acesso aos critérios para cada certificado, o que ele está avaliando e de que forma.

Antonio Acuña respondeu a **Carlos Laufer** que medição nesse caso é subjetiva, pois são um país democrático, e o Brasil também, mas a indexação de algo vai depender da subjetividade e de outros critérios mais específicos.

Marco Túlio Pires perguntou a **Antonio Acuña** qual a importância do papel da sua equipe na articulação com órgãos de governo visando ampliar o processo de abertura de dados.

Antonio Acuña replicou a **Marco Túlio Pires** de que todo o processo não seria viável caso estivessem em outro setor. O que fizeram depende da posição que estão atualmente, ou seja, um órgão independente e que pode transitar entre os demais setores, fazendo as bases conversarem.

5. Apresentação dos Projetos Pilotos de Abertura de Dados do Governo do Estado de São Paulo

5.1. Piloto do Metrô de São Paulo

Daniel Martim (Metrô SP) explicou que o Metrô de São Paulo já disponibiliza algumas informações sobre suas linhas, trajetos, número de usuários. Para o futuro, eles desejam ter esses bancos de dados e informações em formato de dado aberto. Os datasets escolhidos foram: números anuais de trens em operação, passageiros por estação, médias mensais de uso de passageiros, etc. Outra base de dados que Daniel apontou estar dentro do escopo do projeto-piloto é a utilização da **Pesquisa Origem e Destino**, que traz informações uma vasta gama de informações a respeito dos usuários do sistema de Metrô de São Paulo.

Daniel destacou que a Pesquisa Origem e Destino é realizada a cada dez anos, e é de suma importância para a realização do planejamento de transportes da cidade. Sua próxima edição está prevista para 2017, e a última tem seus resultados hoje disponibilizados no portal do metrô. Atualmente ela está em formato *.xls* (*o formato das planilhas eletrônicas como Microsoft Excel*), e sua conversão em formato aberto, em sua opinião, seria algo simples. Confessou que também estão estudando como fazer a divulgação da base de dados em tempo real do que acontece nas linhas do Metrô de São Paulo, ou seja, um serviço que em tempo real das informações das linhas 1, 2, 3 e 5. Hoje trabalhamos com outras empresas parceiras (CPTM, TV Minuto, etc.), mas pretendemos expandir. Adicionou que outro ponto de discussão do projeto-piloto é a respeito da forma como pode ser feita a publicação de outras informações coletadas através da pesquisa. Em especial relativo ao formato dos dados. Também temos dúvidas com relação ao portal de disponibilização dos dados. Devemos criar um portal próprio para divulgar essas informações, ou o fazemos através de outro portal governamental já existente? Nossa primeira ideia é termos um próprio, mas ainda estamos avaliando as possibilidades.

O representante do Metrô de São Paulo ainda enfatizou que não possuem prazos definidos, pois precisam de licenciamento, mas o Direto do Metrô-SP (sistema que informa sobre ocorrências operacionais, em tempo real, que afetam a circulação dos trens em uma linha) está atualmente licenciado por um termo restrito de seu uso.

Ele finalizou informando a todos de que um segundo passo do piloto tem a segurança da informação como principal objetivo, apontando que sua área, responsável pelo SIC (Serviços de Informação ao Cidadão), são alguns dos setores que estão sendo criadas interlocuções para o melhor andamento do projeto.

José Carlos Vaz abriu para perguntas ao projeto piloto apresentado pelo Metrô de São Paulo.

5.1.1. Comentários sobre Piloto do Metrô de São Paulo

Carlos Laufer explicitou a **Daniel Martim** que o Projeto-Piloto não possui problemas com a seleção de bases, pois elas já existem e estão publicadas. Perguntou então ao representante do Metrô de São Paulo, se foi considerada a qualidade destas bases de dados (*datasets*) já disponibilizados. A segunda pergunta foi sobre os metadados das bases de dados. Carlos perguntou se as pessoas já entendem do que se tratam os dados que estamos falando. Ele explicou que de acordo com o que Daniel disse, outras aplicações já estão fazendo uso dos dados em tempo real que o Metrô possui, então, Carlos apontou que há de se entender no que exatamente é pretendido avançar.

Daniel Martim (Metrô SP) respondeu a **Carlos Laufer** positivamente para a questão da qualidade, onde primeiramente se pretende ter um piloto como um padrão e ter no mínimo três estrelas de qualidade dos dados. Ele explicou que a primeira base de dados de demandas está publicada em formato HTML (*HyperText Markup Language* - Linguagem de Marcação de Hipertexto) no portal do metrô. Agora o objetivo é transformar estes dados em formato compatível com as diretrizes dos dados abertos e em um local único e específico. Deu o exemplo do “Direto do Metrô” que é direcionado apenas a algumas empresas.

Daniel adicionou que o Metrô de São Paulo já disponibiliza algumas informações sobre suas linhas, trajetos, número de usuários, a ideia é pegar essas informações e transforma-las em uma grande base de dados abertos.

José Carlos Vaz pediu na sequência que fosse apresentado o Projeto-Piloto da Secretaria de Planejamento.

5.2. Piloto Secretaria de Planejamento

Tarcila Peres Santos, representante da Secretaria de Planejamento do Estado de São Paulo, iniciou sua fala descrevendo o Plano Plurianual (PPA) onde são estabelecidos programas e ações para os próximos quatro anos (2016 a 2019). Ela descreveu que os arquivos atualmente estão em PDF desde 2011. Confessou que as mudanças do PPA atualmente não são incluídas no documento PDF divulgado pela secretaria, e, desta forma, infelizmente, a sociedade não pode acompanhar essas mudanças pela falta de ferramentas que promovessem essa interação.

Disto, o objetivo é desenvolver as informações do próximo PPA em um aplicativo (*app*), de forma que possa ser acessado por qualquer dispositivo móvel, onde qualquer cidadão poderá fazer buscas, de acordo com seu interesse, e ao longo de sua execução a ideia de que esse Plano Plurianual possa crescer e evoluir, e ainda assim ser acompanhado de perto pela população. Para isso o app precisará ser constantemente atualizado por dados abertos (estrutura do PPA e atualizações do PPA).

Tarcila explicou a todos que as bases de dados da Secretaria de Planejamento estão em um Banco de Dados Oracle e podem ser disponibilizados através do nosso BI (*Business Intelligence* - Inteligência de Negócios). Dessa forma, já possuem muitas coisas já disponíveis, mas é preciso crescer e abranger um formato mais geral, bem como ter informações em parceria com a Secretaria da Fazenda, as quais poderiam ser referenciadas com as informações do Plano Plurianual (PPA), da Lei de Diretrizes Orçamentárias (LDO) e da Lei Orçamentária Anual (LOA).

5.2.1. Comentários sobre Piloto da Secretaria de Planejamento de São Paulo

Maria Isabela Falleiros (Secretaria da Fazenda) acrescentou o que **Tarcila (Secretaria de Planejamento)** comentou, apontando que está ocorrendo atualmente uma modernização dos sistemas EPA (Revisão da Estrutura de Programa e Ações), SIMPA (Sistema de Monitoramento de Programas e Ações do PPA), além da revisão da proposta orçamentária. Questionou Tarcila como pretendem integrar essas informações nessa nova aplicação que está sendo desenvolvida.

Tarcila Peres Santos respondeu a **Maria Isabela (Secretaria da Fazenda)**, que esse é o primeiro ciclo de modernização e que o SIMPA deverá servir como base para esse novo portal aplicação que estão desenvolvendo. Como estão trabalhando em aplicações especialmente de BI, a tendência é que ele seja substituído por outras aplicações, e tudo isso está sendo desenvolvido em cima dessas novas estruturas. O objetivo é de que essas informações já estejam disponibilizadas até a aprovação do próximo Plano Plurianual (2016-2019).

Carlos Laufer comentou a **Tarcila (Secretaria de Planejamento)** que no projeto-piloto da Secretaria de Planejamento há um enfoque maior no sistema do que na abertura dos dados. Questionou a representante da Secretaria do Planejamento se há como objetivo desenvolver não apenas uma aplicação onde os dados podem ser acessados de forma unilateralmente.

Tarcila Peres Santos explicou a **Carlos Laufer** que os dados já estão públicos, mas estão desatualizados. Como estão em processo de elaboração, até setembro de 2015, existe o objetivo de novo modelo de monitoramento, e pensam em uma nova forma de comunicação com a sociedade.

José Carlos Vaz passou a palavra para o Projeto-Piloto do Tribunal de Contas do Estado de São Paulo (TCE/SP).

5.3. Piloto do Tribunal de Contas do Estado (TCE)

Edson Sales Júnior (TCE-SP) iniciou a descrição do Projeto-Piloto do Tribunal de Contas do Estado de São Paulo explicando o objetivo da criação de um portal da transparência com informações municipais para o Estado. Ele inclusive adiantou que na primeira semana de abril, teriam uma versão inicial do projeto disponibilizada online. Edson pontuou que atualmente, estão trabalhando com aplicações e formatos livres (Postgre¹⁹ e JSON²⁰). Ele disse não possuir bases de dados de fontes externas, desde portais institucionais até o SIAPENET²¹ (Sistema de Informações da Administração Pública), e pretendem integrar essas bases de dados a o sistema ODESP, para serem criados novos significados para esses dados. O objetivo é concluir o projeto até o final do ano, e estão testando alguns módulos do grupo, e esbarramos em um primeiro momento na questão do vocabulário. No entanto, para evitar complicações, iremos criar um novo vocabulário dentro do próprio sistema. Também estão agregando vários dados que hoje não existem no sistema, e transformando-os em dados mais primários para que possam ser manipulados livremente da forma que o usuário decidir.

José Carlos Vaz passou a palavra para os dois Projetos-Piloto na área da Saúde, um referente ao controle de medicamentos e outro de doenças contagiosas. O primeiro a se pronunciar foi o projeto-piloto de Medicamentos.

5.4. Piloto da Secretaria da Saúde – Medicamentos

Paula Opromolla iniciou sua fala descrevendo que possuem atualmente em formato PDF uma base de dados formada por informações coletadas nas farmácias que fazem distribuição de medicamentos populares. Acrescentou que também existem informações referentes à proporção de doenças por região que estão localizadas esses postos de distribuição. Exemplificou as dimensões de dados tais como doenças, idades dos usuários, medicamentos mais utilizados, etc. No entanto, ela admitiu que não possui ainda nenhum aprofundamento ou ideia de como trabalhar essas bases de dados. Para isso, está disposto a colaborar da melhor forma. Ela aprofundou a descrição ao citar que fala de diferentes segmentos, desde a atenção básica, até o grupo fundamental, o projeto-piloto escolhido, onde existe um alto valor agregado e temos um controle maior na distribuição. Já os de atenção básica temos uma maior dificuldade de controle e mensuração, pois ficam a cargo dos municípios.

¹⁹ <http://www.postgresql.org>

²⁰ <http://json.org>

²¹ <http://siapnet.tce.sp.gov.br>

5.5. Piloto da Secretaria da Saúde – Doenças

O representante da Secretaria da Saúde do Estado de São Paulo iniciou sua fala apontando que atualmente já fazem a publicação de certos tipos de dados, contudo, ainda enfrentam alguns problemas referentes à questão de licenciamento. Como o **Roberto Agune** havia comentado em sua apresentação mais cedo, a questão da abertura de dados não era algo tão novo, mas o tema “Web Semântica” é uma grande novidade e pretendem exercitar esse conceito no Projeto-Piloto. O representante disse conhecer a limitação técnica, mas citou o apoio da **PRODESP**²² (Companhia de Processamento de Dados do Estado de São Paulo – Tecnologia de Informação do Governo do Estado) para superar este obstáculo.

5.6. Apresentações de Pilotos Extras

5.6.1. Professores Vinícius e Domingos da Universidade de São Paulo de Ribeirão Preto

Os professores da Universidade de São Paulo comentaram sobre o projeto que possuem na cidade de Ribeirão Preto. Eles indicaram que possuem três micro bases referentes à análise de dados e diagnósticos da Tuberculose. Elas são provenientes de trinta e um hospitais pertencentes à região e recebem esses dados dos hospitais parceiros. Ele disse acreditar que essas bases de dados na USP podem servir de grande valor para a criação de uma cultura de uso desses dados públicos.

6. Discussão dos Projetos Pilotos

Vagner Diniz iniciou sua fala de análise chamando a atenção para alguns pontos que já foram abordados no período da manhã e que acreditou terem sido importantes em consideração para um melhor desenvolvimento dos projetos-pilotos. Vagner destacou que Antonio Acuña mostrou existirem 10 etapas a serem cumpridas antes de coletar um dado. Para ele, é óbvio que essa lista tem como objetivo fazer uma provocação aos participantes que irão realizar a abertura de dados dos projetos-pilotos, de modo que possam incluir as questões diferentes das que foram levantadas, bem como outros relevantes.

Primeiramente, Vagner chamou a atenção para que os projetos-pilotos tenham em mente que estão em estágios diferentes e saibam aonde querem chegar, quais os seus objetivos. Essa é a primeira pergunta que se deve ser feita. Mesmo tendo clareza de que querem publicar X, Y e Z bases de dados, deve-se ter certeza dos motivos em não publicar as bases de dados A, B e C.

Em segundo lugar, Vagner disse acreditar que a análise dos dados já existe e disso questiona os projetos-pilotos se as bases de dados serão publicadas em sua totalidade, visto que é importante definir um propósito de onde se quer chegar e ter clareza de

²² <http://www.prodesp.sp.gov.br>

quais dados serão disponibilizados. Ele ressaltou de que a definição do que é útil ou não é uma forma de fazer uma limpeza dessas bases de dados. Questionou se já foi feito este trabalho de análise de limpeza nos projetos-pilotos, verificando os excessos, se os dados já estão modelados e se estão, de qual forma. Ele deu o exemplo do projeto-piloto de doenças contagiosas, pois se não for realizada a modelagem que se deseja para publicar o conteúdo semântico, Vagner questiona se o usuário de fato sabe que aqueles dados são referentes a doenças contagiosas.

Vagner destacou outro ponto importante que é a definição de um vocabulário semântico, quanto para a publicação de dados de forma linkada e conectada a outros dados. Perguntou aos encarregados da abertura das bases de dados se os guardiões dessas bases de dados estão empoderados para a liberação dos mesmo. Outro ponto explicado por Vagner Diniz foi a intenção de se chegar a quatro ou cinco estrelas. Perguntou aos participantes se já se chegou a uma definição de quais formatos serão utilizados, se já foram discutidas as vantagens e desvantagens, pois seria interessante já no protótipo do Projeto-Piloto, a conexão com alguma base de dados do próprio governo. Deste modo, quem seriam o alvos dessas conexões, indagou Vagner.

Finalizando, chegou a conclusão de que os projetos-pilotos estão em diferentes estágios: definição de qual abertura de base de dados e alguns já abrindo as bases de dados. Pediu para que todos tenham isso registrado por escrito e que seja compartilhado e transformado em um cronograma, de modo a trabalhar de forma parceira até o final.

7. Metodologia de organização e trabalho para discussão dos Projetos Pilotos

José Carlos Vaz agradeceu a **Vagner Diniz**, e iniciou a etapa de grupos. Organizou em cinco grupos de projetos-pilotos e pediu para que preenchessem uma planilha disponibilizada por sua equipe para verificar desafios e oportunidades da abertura de bases de dados em cada projeto-piloto. As Planilhas originais estão anexadas na seção anexos 8.7. [Planilhas Originais apresentadas pelos grupos.](#)

7.1 Planilhas dos Pilotos dos Grupos de Discussão

7.1.1 Planilha do Piloto da Secretaria Estadual de Saúde

Tabela 1 – Projeto-Piloto da Saúde

Análise e definição dos dados a publicar	Quais dados serão publicados?	Doenças de notificação e medicamentos
	Quais são os potenciais usos e usuários dos dados?	controle, planejamento, vigilância população, organização da sociedade, gestores, academia
	Quais são potenciais conexões com outros dados ?	comunicação e integração entre medicamentos, notificação, Fundação SEADE, município
	Quem será o responsável pelos dados?	SES, USP, PRODESP, Seade , IBGE
Preparação dos dados	O que deverá ser feito para garantir a qualidade dos dados (limpeza)?	Ficha de metadados, modelagem, limpeza
	Qual será a melhor estratégia para modelar os dados?	levantamento do modelo estruturado e de relacionamento entre medicamentos e doenças
	Como escolher os vocabulários a serem utilizados?	diagnósticos=CID-10 ; Tabela Medicamentos
Publicação	Como definir as licenças a serem utilizadas?	Regras Governo Aberto (Guia) aval Jurídico, Ética
	Qual será a forma de publicação na Web? Interfaces, ferramentas etc.	Nova APP PRODESP

7.1.2 Planilha do Piloto da Secretaria Estadual de Planejamento

Tabela 2 - Projeto-Piloto do Planejamento

Análise e definição dos dados a publicar	Quais dados serão publicados?	Dados do PPA (Planejado dos 4 anos, objetivos, programas, metas, ações, público alvo e orçamento para os 4 anos)
	Quais são os potenciais usos e usuários dos dados?	Acompanhamento do planejamento, reconhecer as prioridades do governo, as metas e os indicadores, problemas e responsabilidades por parte do cidadão, pesquisador, imprensa, prefeitos e gestores públicos e interessados em geral.
	Quais são potenciais conexões com outros dados ?	Com os dados dos indicadores socioeconômicos e demográficos da Fundação SEAD.
	Quem será o responsável pelos dados?	Secretaria de Planejamento
Preparação dos dados	O que deverá ser feito para garantir a qualidade dos dados (limpeza)?	Aprofundar a documentação sobre o PPA e metadados com descrições mais ricas para atender aos diferentes usuários.
	Qual será a melhor estratégia para modelar os dados?	Parte dos dados serão publicados em CSV (3 estrelas) e parte em RDF (4 estrelas). O orçamento será em RDF
	Como escolher os vocabulários a serem utilizados?	Utilizar o vocabulário de orçamento do governo federal.
Publicação	Como definir as licenças a serem utilizadas?	Utilizar licença aberta uma vez que os dados do PPA são públicos.
	Qual será a forma de publicação na Web? Interfaces, ferramentas etc.	No portal de secretaria do planejamento por meio de ferramentas a serem definidas posteriormente.

7.1.3 Planilha do Piloto do Tribunal de Contas do Estado de São Paulo

Tabela 3 – Projeto-Piloto do Tribunal de Contas do Estado de São Paulo

Análise e definição dos dados a publicar	Quais dados serão publicados?	Dados Econômicos, Financeiros E Orçamentários Dos Municípios Do Estado De São Paulo Exceto Capital
	Quais são os potenciais usos e usuários dos dados?	Potenciais Usuários: Sociedade Em Geral, Gestores Públicos, Organizações Não-Governamentais, Imprensa Em Geral, Blogueiros Com Interesse Em Controle Social / Potenciais Usos: Controle Social, Fiscalização De Políticas Públicas, Estudos Comparativos, Desenvolvimentos De Aplicativos
	Quais são potenciais conexões com outros dados ?	IBGE, Administração Penitenciária, Receita Federal, Jurisdicionados Em Geral
	Quem será o responsável pelos dados?	Centro De Apoio Estratégico à Fiscalização - Caef/Sdg
Preparação dos dados	O que deverá ser feito para garantir a qualidade dos dados (limpeza)?	O tratamento dos dados é feito durante o Extraction Transform and Loading (ETL, em português Extrair, Transformar e Carregar) de modo que o dado publicado já encontra-se limpo
	Qual será a melhor estratégia para modelar os dados?	O principal quesito na modelagem de dados É a performance para acesso via Web, tendo em vista a quantidade de registro que detemos
	Como escolher os vocabulários a serem utilizados?	Vamos criar um Vocabulário próprio com estrutura taxonômica - baseado na sintaxe de termos do Pcas 2015
Publicação	Como definir as licenças a serem utilizadas?	O pressuposto atual é da não-necessidade de licenças, contudo nós citamos as fontes que utilizamos
	Qual será a forma de publicação na Web? Interfaces, ferramentas etc.	Já publicado em site próprio nos formatos Html, Xml, Csv, Jason E Rdf

7.1.4 Planilha do Piloto do Metrô de São Paulo

Tabela 4 – Projeto-Piloto do Metrô de São Paulo

Análise e definição dos dados a publicar	Quais dados serão publicados?	Pesquisa Origem e Destino
	Quais são os potenciais usos e usuários dos dados?	Planejamento urbano na região metropolitana de São Paulo
	Quais são potenciais conexões com outros dados ?	Necessária análise
	Quem será o responsável pelos dados?	Metrô de São Paulo
Preparação dos dados	O que deverá ser feito para garantir a qualidade dos dados (limpeza)?	Feedback do utilizador
	Qual será a melhor estratégia para modelar os dados?	Conversão dos dados, atualmente em .XLS
	Como escolher os vocabulários a serem utilizados?	Vocabulário já existente na Pesquisa Origem-Destino. Necessária análise.
Publicação	Como definir as licenças a serem utilizadas?	A definição das licenças deverá ser objeto de discussão interna,
	Qual será a forma de publicação na Web? Interfaces, ferramentas etc.	Portal CMS e arquivos .XLS

8. Encerramento

José Carlos Vaz iniciou o encerramento das atividades passando a palavra para os convidados internacionais para finalizar as discussões.

Eleanor Stewart agradeceu e destacou ter sido um dia interessante e encorajador. Avaliadas por ela, as diretrizes propostas apresentadas pelo Governo do Estado de São Paulo são muito boas, inclusive elogiou, sendo as melhores que ela já viu. Destacou a energia encontrada durante a oficina. Ela disse acreditar que seja importante manter o foco nos resultados pretendidos e na liberação dos dados. Ela destacou que depois que os dados são liberados devem se despreocupar-se, pois eles se tornam livres. Porém, lembra que é importante que eles tenham licenças claras, justamente para evitar eventuais problemas no futuro. Disse estar ansiosa para acompanhar os resultados das iniciativas propostas e para ler os manuais que estão sendo desenvolvidos.

Antonio Acuña iniciou seu discurso final agradecendo pelo dia de hoje. Disse estar fascinado pelo que deparou na oficina, mesmo com tanto problemas e dificuldades no processo de abertura de dados. Quando o dia começou, confessou estar bastante ansioso e agora sente-se aliviado, mas gostaria de enfatizar a importância das licenças. Organizem-se e liberem os dados seguindo os passos que foram informados e os resultados positivos acontecerão naturalmente, finalizou Acuña.

José Carlos Vaz moderador agradeceu pelo dia de hoje e enfatizou a maravilhosa contribuição dos participantes. Passou a palavra para Roberto Agune.

Roberto Agune finalizou agradecendo aos especialistas internacionais britânicos Eleanor Steawart e Antonio Acuña pela contribuição que puderam dar. Ficou feliz por saber que está sendo trilhado o caminho correto e a certeza de que atingirão os objetivos propostos. Em relação ao projeto-piloto, acredita ser fundamental a organização deste tipo de *workshop* (oficina) onde é possível conversar, enriquecer e trocar informações para que os projetos possam caminhar bem. Futuramente serão realizados outros *workshops* com temáticas diferentes, em especial de alguns temas interessantes que vieram à tona no dia da oficina, tais como licenças, pois acredita que são importantes para a definição de uma política de dados abertos do Estado de São Paulo. Agradeceu a PRODESP (Companhia de Processamento de Dados do Estado de São Paulo) por auxiliar no desenvolvimento dos projetos-pilotos, verificando as condições necessárias para que esses projetos possam ser realizados da melhor forma. Finalizou agradecendo a presença e o esforço de todos na Oficina.

9. Anexos

9.1 Lista de participantes

Nome do convidado	Empresa/organização
Alcione de Godoy	Secretaria de Governo do Estado de São Paulo
Allan Santos	Relatoria
Álvaro Gregório	Secretaria de Governo do Estado de São Paulo
Ana Marques	Ouvidoria do Estado de São Paulo
Ana Paula Conte	Núcleo de Informação e Coordenação do Ponto br (NIC.br)
Antonio Acuña	UK - Cabinet Office
Antonio Gonçalves	Fundação Sistema Estadual de Análise de Dados
Antônio Martins	Tribunal de Contas do Estado de São Paulo
Bernadette Farias Lóscio	Universidade Federal de Pernambuco
Breno Mazieiro	Universidade de São Paulo
Carlos Laufer	Pontifícia Universidade Católica do Rio de Janeiro
Caroline Burle	Centro de Estudos sobre Tecnologias Web (Cewe.br) / Núcleo de Informação e Coordenação do Ponto br (NIC.br)
Claudia Montero	Secretaria da Saúde Governo do Estado de São Paulo
Daniel Martim	Companhia do Metropolitano de São Paulo
Daphne Abreu Souza	Tribunal de Contas do Estado de São Paulo
Domingos Alves	Universidade de São Paulo - Ribeirão Preto

Edson Sales Jr.	Tribunal de Contas do Estado de São Paulo
Eduarda Giffoni	Embaixada Britânica
Edward Gerth	Secretaria de Governo do Estado de São Paulo
Eleanor Stewart	UK - Cabinet Office
Gabriela Lubies	Secretaria da Fazenda Governo do Estado de São Paulo
Gustavo Alberto	Companhia de Processamento de Dados do Estado de São Paulo
Helena Pchevuzinske	Secretaria de Governo do Estado de São Paulo
Heloisa Pait	Open Knowledge
José Carlos Vaz	Universidade de São Paulo
José Leomar Todesco	Universidade Federal de Santa Catarina
M. Isabela Falleiros	Secretaria da Fazenda Governo do Estado de São Paulo
Magali Valente	Fundação Sistema Estadual de Análise de Dados
Marcelo Ferraz	Secretaria da Fazenda Governo do Estado de São Paulo
Marcio Biczuk	Secretaria da Saúde Governo do Estado de São Paulo
Marco Túlio Pires	Escola de Dados
Marcos Prezato	Secretaria da Saúde Governo do Estado de São Paulo
Miguel Morel	Secretaria da Fazenda Governo do Estado de São Paulo
Miriam de Cássia Tomaz Canoas	Secretaria de Planejamento e Gestão do Estado de São Paulo
Nanci M. Padilla	Fundação Sistema Estadual de Análise de Dados

Newton Calegari	Núcleo de Informação e Coordenação do Ponto br (NIC.br)
Paula Opromolla	Secretaria da Saúde Governo do Estado de São Paulo
Paulus Magrini de Lima	Secretaria da Fazenda Governo do Estado de São Paulo
Ricardo Matheus	Delft University of Technology
Roberto Agune	Secretaria do Governo do Estado
Sergio Pinto Bolliger	Secretaria do Governo do Estado
Silvia Campanile	Companhia de Processamentos de Dados do Estado de São Paulo
Silvio Aquino	Embaixada Britânica
Silvio Mendonça	Secretaria da Fazenda Governo do Estado de São Paulo
Sulimara Takahashi	Assessoria Internacional Governo do Estado de São Paulo
Tarcila Peres Santos	Secretaria de Planejamento e Gestão do Estado de São Paulo
Vagner Diniz	Centro de Estudos sobre Tecnologias Web (Ceweb.br) / Núcleo de Informação e Coordenação do Ponto br (NIC.br)
Vivaldo Conti	Fundação Sistema Estadual de Análise de Dados
Wagner Meira	Universidade Federal de Minas Gerais
Yasodara códova	Centro de Estudos sobre Tecnologias Web (Ceweb.br) / Núcleo de Informação e Coordenação do Ponto br (NIC.br)
Zilda Souza	Secretaria da Saúde Governo do Estado de São Paulo

9.2 Apresentação Roberto Agune

9.3 Apresentação Guia de Abertura de Dados – Marco Túlio Pires

9.4 Apresentação Guia de Web Semântica – Carlos Laufer

9.5 Apresentação Eleanor Stewart

9.6 Apresentação Antonio Acuña

9.7 Planilhas Originais apresentadas pelos grupos