

Quero uma revolução.

Não política e, certamente não violenta, mas uma revolução.

Uma revolução na forma como as pessoas pensam sobre a maneira como os dados são publicados na Web, abertos ou não. Aqui é onde eu costumo começar a falar sobre as pessoas usando a Web como um pendrive glorificado. Ou seja, usar a Web para não fazer mais do que transferir dados de A para B de uma forma que poderia ser facilmente alcançada, colocando-a em um pendrive e enviando-o através do correio.



Crédito da foto: Rosie Sutton

A Web é muito mais do que isso. Para citar a [Arquitetura da World Wide Web](#), ela é: “.. um notável espaço de informação de recursos inter-relacionados, crescendo através de idiomas, culturas e mídia”. É a conectividade de ideias e fatos entre pessoas que são desconhecidas umas às outras, que é tão emocionante e que tem profundas implicações.

“Existem diversas maneiras de publicar dados na Web, porém até hoje não existia um padrão, um conjunto de práticas capazes de guiar e facilitar o trabalho, tanto para os publicadores, quanto para os consumidores de dados” – Newton Calegari.

Mas como fazê-lo direito? Como Rebecca Williams, da GovEx, que trabalhou no data.gov, [twittou recentemente](#): “olhar para ‘portais de dados abertos’ para reunir suas melhores práticas em metadados e licenciamento é andar para trás, pois quase todos estão fazendo errado”. Eu não diria que estão quase todos fazendo errado, mas é verdade que há necessidade de ter uma referência para saber como fazê-lo direito.

Que é a [Recomendação das Boas Práticas para Dados na Web](#) (DWBP).

Levou quatro anos para ser publicada, desde o planejamento do [Workshop](#), até a criação do [Grupo de Trabalho](#), para o que [realmente é o escopo](#) e firmando o relacionamento com o [Projeto Share-PSI](#) (financiado externamente), para aperfeiçoar um conjunto de 35 boas práticas, as quais são passíveis de ação sem serem excessivamente prescritivas.

“Um excelente conjunto de Boas Práticas bem pesquisado e claramente escrito” – Terence Eden, HM Government

O primeiro é o conceito mais básico de todos: fornecer metadados. Parece bobagem, e pode-se argumentar que se você está publicando dados na Web e não fornece metadados, então, você provavelmente está muito interessado que ninguém encontre os dados publicados, muito menos alguém vá usá-los.

A Boa Prática (BP) 9 diz “Use URIs persistentes como identificadores de conjuntos de dados” e a BP 10 diz “Use URIs persistentes como identificadores dentro de conjuntos de dados”. Na minha opinião, esses dois estão no centro da diferença entre usar a Web como um pendrive glorificado e usá-la como um espaço de informação global. O Relatório de Implementação mostra muitos exemplos disso, desde as [Compras Públicas do Governo Federal](#) até o Projeto de Drogas Conectadas ([Linked Drug’s project](#)), da Universidade Metodista da Macedônia St Cyril (Macedonia’s St Cyril and Methodius University), assim como a API do Museu da Guerra de Auckland ([Auckland War Museum](#)) e o Projeto [Acropolis](#), do Reino Unido.



Reuso (Reuse), Compreensão (Comprehension), Conexão (Linkability), Descobrimto (Discoverability), Confiança (Trust), Acesso (Access), Interoperabilidade (Interoperability), Processamento (Processability)

Cada uma das BPs é classificada de acordo com um ou mais dos benefícios acima.

Existem boas práticas sobre assuntos que você provavelmente esperaria, como proveniência e licenciamento de dados, e outras menos óbvias como enriquecimento e arquivamento de dados. Estes são assuntos cujo escopo é enorme, assim, o documento de Boas Práticas para Dados na Web deve servir como base para publicar e consumir dados na Web.

No W3C, outros trabalhos estão em andamento, por exemplo, a padronização do [ODRL](#) para [permissões e obrigações](#) dos dados legíveis por máquina, assim como o [Grupo de Trabalho de Boas Práticas para Dados Espaciais na Web](#), cujas [boas práticas](#) estão sendo construídas com base no DWBP. Há sempre mais a dizer – e há sempre maneiras diferentes de trabalhar.

“Grande trabalho – um recurso extremamente útil” – Jeremy Tandy, Met office

As Boas Práticas para Dados na Web não prescrevem o uso de qualquer tecnologia em particular, além dos princípios básicos da Web. Cada BP tem um resultado pretendido, como a BP 14: o máximo de usuários possível será capaz de usar os dados sem primeiro ter que transformá-lo em seu formato preferido. Ou a BP 23: os desenvolvedores terão acesso programático aos dados para uso nas suas próprias aplicações, os quais serão atualizados sem exigir esforço por parte dos consumidores. Os aplicativos da Web poderão obter dados específicos consultando uma interface programática. Assim, cada boa prática oferece possíveis abordagens para a implementação com alguns exemplos. Se você conseguir o mesmo resultado desejado com uma tecnologia diferente, vá em frente, você ainda está seguindo as melhores práticas.

“O entendimento entre os publicadores e consumidores de dados é fundamental. Sem esse acordo, os esforços dos publicadores de dados podem ser incompatíveis com o desejo dos consumidores” – Caroline Burle

O Grupo de Trabalho de Boas Práticas para Dados na Web foi concebido, no seu [Charter](#), não apenas para criar um conjunto de boas práticas, mas para ajudar a promover um ecossistema de compartilhamento de dados. Parte disso é abordada em dois vocabulários, um para descrever o uso de um conjunto de dados ([Dataset Usage Vocabulary](#)) – através do uso em uma aplicação, citação no trabalho de outra pessoa etc. – e outro para descrever a qualidade dos dados ([Data Quality Vocabulary](#)). A qualidade é raramente um fato objetivo, mas o vocabulário fornece um quadro em que as declarações sobre a qualidade podem ser feitas.

DWBP não é apenas sobre dados governamentais. A GS1, a instituição que está por trás dos códigos de barras de produtos do mundo, contribuiu para o trabalho e já o alavancou em sua proposta GS1 SmartSearch. No mundo da pesquisa científica, o Laboratório Nacional do Noroeste do Pacífico (Pacific Northwest National Laboratory) está colocando em prática o trabalho em sua publicação sobre conjuntos de dados de simulação climática na Federação da Rede do Sistema Terrestre (Earth System Grid Federation), no Arquivo de Dados Atmosfera para Elétrons (A2e) e em seu Portal (DAP).

“A Web se destaca como um meio de compartilhamento de dados, porém, nem sempre esses dados podem ser facilmente descobertos, acessados e processados. O uso das Boas Práticas para Dados na Web é fundamental para alavancar o compartilhamento de dados na Web, uma vez que garante o fácil acesso e a reutilização dos dados” – Bernadette Lóscio

É sempre encorajador quando você ouve outras pessoas se referindo ao seu trabalho e o DWBP teve várias menções no Workshop Descrições Inteligentes e Vocabulários Mais Inteligentes (Smart Descriptions e Smarter Vocabularies – [SDSVoc](#)), realizado no final do ano passado (o relatório sairá em breve, eu prometo). E nós tivemos elogios de muitas pessoas. Gostaria de terminar observando duas características incomuns do Grupo de Trabalho. Primeiro, as três chairs e duas dos três editores do grupo são mulheres. Em segundo lugar este foi o primeiro W3C WG que teve uma participação tão forte do Brasil.

Foi um privilégio trabalhar com um grupo tão extraordinário de revolucionários de todo o mundo.



Conheça os editores do DWBP & equipe de contato do W3C: (da esquerda para direita)
Newton Calegari, Caroline Burle, Phil Archer, Bernadette Lóscio

Esta é uma tradução do artigo de Phil Archer. [Veja aqui](#) a versão original.



SAIBA MAIS

Caroline Burle



5 Artigo(s)

É responsável pelas Relações Institucionais do Centro de Estudos sobre Tecnologias Web (Ceweb.br). Mestre em Relações Internacionais pelo San Tiago Dantas, possui especialização em negociações pela Fundação Getúlio Vargas e é bacharel em Relações Internacionais pela Fundação Armando Álvares Penteado. Pesquisa sobre governo aberto, relações internacionais e política externa, com experiência na atuação de governos subnacionais, na área de Web, dados abertos e governança da Internet.