

Dados Abertos e Conectados: Uma breve descrição

André Luiz G. Larrosa, Rafael S. Silva, Késsia Rita da Costa Marchi

Campus Paranavaí - Instituto Federal do Paraná (IFPR)

Paranavaí– PR – Brasil

{ aluiz8822, rafael.sanches.s35}@gmail.com, kessia.marchi@ifpr.edu.br

A explosão informacional oriunda da utilização massiva de dispositivos tecnológicos tem ocasionado crescimento exponencial de dados gerados pela sociedade. Os dados são criados ao utilizar um celular, navegar na internet, realizar uma compra, interagir em redes sociais, entre outras ações. Essa grande quantidade de dados leva a necessidade de habilidades que proporcione lidar com estes dados que, na maioria das vezes, não são estruturados. Isotani (2015) afirma que este excesso impacta na obtenção de informações ou produção de conhecimentos, reduzindo a agilidade e a eficácia necessárias para lidar com questões diversas da sociedade contemporânea.

Diante deste contexto, este resumo tem por objetivo apresentar uma introdução aos conceitos, classificações, benefícios, normas e leis acerca da utilização de dados abertos e dados abertos governamentais. Para a elaboração deste resumo foi realizado uma breve revisão de literatura.

Dados abertos, segundo a *Open Definition* (2014), são dados que podem ser livremente utilizados, reutilizados e redistribuídos por qualquer pessoa com a exigência máxima de atribuir à fonte original e ao compartilhamento pelas mesmas licenças em que as informações foram apresentadas. De acordo com essa definição, existem três normas fundamentais para a disponibilização dos dados abertos, sendo elas:

- Disponibilidade: os dados devem estar acessíveis de maneira completa;
- Reutilização: os dados não devem ter restrições para reuso e compartilhamento;
- Participação Universal: Não haver restrição para determinados grupos, indivíduos ou campos de atuação;

Respeitando essas normas, é possível que diversas pessoas trabalhem colaborativamente, bastando para isto, manter os dados acessíveis, legíveis por homens e máquinas e ter formato aberto. Esta necessidade leva a preocupação com a conexão e a publicação dos dados. Neste contexto, Tim Berners-Lee propõe os princípios conhecidos como “As 5 estrelas dos dados abertos”. Seu objetivo é facilitar a caracterização de dados publicados, além de evidenciar a utilidade e aproveitamento dos mesmos. Esse esquema consiste na atribuição de estrelas aos dados analisados, podendo ser atribuído um número entre 1 a 5, este esquema é exibido na Tabela 1 [ISOTANI; BITTENCOURT, 2015; PEDROSO; TANAKA; CAPPELLI, 2013].

Para que um dado seja considerado efetivamente aberto, ele deve atingir a marca mínima de 3 estrelas, sendo que ao atingir um marco maior, apenas proporcionará uma melhor distribuição e acessibilidade desses dados para as pessoas.

Para a disponibilização desses dados, pode-se usar um *Web Service* (Serviço Web), mas, é importante saber algumas tecnologias, tais como JSON (*JavaScript Object Notation*), YAML (*Yet Another Markup Language*), XML (*Extensible Markup Language*) e saber ainda sobre RDF (*Resource Description Framework*), essas tecnologias, unidas a um framework como o Mojolicious formarão o Web Service, que se comunicará com o cidadão que quiser ver esses dados publicados. Em questão dos formatos JSON, YAML e XML, a escolha pode ser feita

pela sua necessidade, já que a maioria dos frameworks para desenvolvimento web tem um suporte à todos [ISOTANI; BITTENCOURT, 2015; PEDROSO; TANAKA; CAPPELLI, 2013]. O RDF, por sua vez, é um modelo padrão para a troca de informação na rede que facilita algumas tarefas tais como teste de schema diferentes e mescla de dados. O RDF é descrito como: um conjunto de dados e seus relacionamentos em forma de triplas.

Tabela 1: Esquema das 5 Estrelas de Dados Conectados. Fonte: Adaptado de Isotani (2015)

★ OP (Open License)	Os dados devem estar disponibilizados na <i>web</i> , independente do formato, necessitando apenas estar em um formato aberto.
★★ RE (Reused)	Além da condição anterior, os dados devem ser legível por máquina (formato estruturado).
★★★ OF (Open Format)	Além das condições anteriores, os dados devem estar disponíveis em formato estruturado não proprietário.
★★★★ URI (Uniform Resource Identifier)	Além das condições anteriores, os dados devem estar na <i>web</i> com URI que os descreve.
★★★★★ LD (Linked Data)	Além das condições anteriores, os dados devem ser capazes de se conectar a outras bases.

A partir destes recursos tecnológicos acessíveis, em 2009, foi sancionada a Lei Complementar 131 [Lcp 131], também conhecida como Lei da Transparência. Essa lei brasileira determina que todos os entes devem liberar, em tempo real, informações pormenorizadas sobre a execução orçamentária e financeira para a sociedade em meios eletrônicos de acesso público, dessa forma, dando origem aos dados abertos governamentais. Esses dados fazem parte de uma ação política de acesso à informação do Governo Federal (INDA – Infraestrutura Nacional de Dados Abertos). Essa categoria de dados abertos segue basicamente 8 princípios, sendo:

- Completos: Todos os dados devem ser disponibilizados.
- Primários: Os dados publicados não podem ser alterados ou modificados e devem conter o maior detalhamento possível.
- Atuais: Novos dados devem ser lançados constantemente, sendo que dados atuais possuem mais valor.
- Acessíveis: Os dados devem ser acessíveis ao maior público possível, além de serem capazes de atender os mais variados propósitos.
- Processáveis por máquinas: Os dados devem estar estruturados para que seu processamento possa ser automatizado.
- Acesso não discriminatório: Os dados devem estar disponíveis para qualquer pessoa, sem que haja a necessidade de identificação ou registro por parte da pessoa.
- Formatos não proprietários: Os dados devem estar disponíveis em um formato sobre o qual nenhum ente tenha controle exclusivo.

- Livres de licença: Os dados não devem estar sujeitos a restrições por regulações de direitos autorais, marcas, patentes ou segredo industrial. Restrições razoáveis de privacidade, segurança e controle de acesso são permitidas na forma regulada do estatuto.

A disponibilização desses dados para a sociedade, além de gerar um valor para a mesma, caso se aproprie deles, torna-se uma ferramenta essencial para um Estado Transparente e para o fortalecimento dos processos democráticos, gerando assim muitos benefícios, tais como: controle social para as ações do governo, geração de novos negócios, transparência na prestação de contas e ações da gestão pública, entre outros.

Além dos dados abertos governamentais, outro exemplo de utilização de dados abertos é a DBPedia, que possui uma estrutura RDF/OWL (*Ontology Web Language*). A DBPedia consiste quase como um nó central da web de dados. Esse projeto foi iniciado em 2007, e atualmente conta com mais de 30 milhões de recursos descritos em 120 idiomas. Além disso, a DBPedia reusa recursos de quase 40 catálogos de dados diferentes e foi conectada por mais de 40 milhões de links a uma centena de catálogos de dados [ISOTANI; BITTENCOURT, 2015].

Neste resumo foram abordados os principais conceitos sobre dados abertos e dados abertos conectados e a sua utilização, assim como, a estruturação dos dados de acordo com as recomendações da W3C, abordando brevemente suas principais características e tecnologias utilizadas. Finalmente, citando a DBPedia como um exemplo de aplicação que contempla as normas fundamentais para a disponibilização de dados abertos.

Referências

Isotani, S. and Bittencourt, I. I. (2015) "Dados Abertos Conectados". São Paulo: Editora Novat. doi: 10.1017/CBO9781107415324.004.

Lcp 131 (no date). Disponível em: https://www.planalto.gov.br/ccivil_03/leis/lcp/lcp131.htm (Acessado em: 17 de Agosto de 2020).

Open Definition (2014) "The Open Definition.", on line. Disponível em: <http://opendefinition.org/> (Acessado em: 17 de Agosto de 2020).

Pedroso, L., Tanaka, A. and Cappelli, C. (2013) "A Lei de Acesso à Informação brasileira e os desafios tecnológicos dos dados abertos governamentais", Simpósio Brasileiro de Sistemas de Informação, pp. 523–528. doi: 10.21118/apgs.v12i2.5872.