



Universidade Federal de Pernambuco - UFPE  
Centro de Informática - CIn  
Graduação em Engenharia da Computação

## **Especificação e Criação do Vocabulário para o Ecosistema de Dados Abertos Governamentais**

Glória de Fátima Andrade Barros Lima - GFABL

Proposta de Trabalho de Graduação

Orientador: Bernadette Farias Lóscio

Recife  
Setembro de 2016

## Resumo

Nos últimos anos com a publicação em massa dos dados abertos governamentais muitas soluções têm sido desenvolvidas com o intuito de suprir as demandas dos usuários. Iniciativas privadas e governamentais estimulam o desenvolvimento de pesquisas e soluções que utilizem os dados abertos, mas tais aplicações acabam não sendo propriamente utilizadas ou terminam dispersas por falta de comunicação entre as partes interessadas. Este trabalho de graduação visa especificar e criar um vocabulário para o ecossistema de dados abertos governamentais através do mapeamento dos atores que compõem o ecossistema, os papéis que cada um exerce e os relacionamentos existentes no processo de produção de soluções inovadoras. Além disso, uma base de conhecimento será gerada disponibilizando informações dos membros dos ecossistemas, resultados das parcerias entre eles, áreas de interesse, demandas, gerando inferências, recomendações e novos dados sobre ecossistemas.

**Palavras-chaves:** dados abertos, ecossistema, vocabulário, aplicações.

# 1. Contextualização

Hoje em dia, os portais de dados abertos podem ser encontrados na maioria dos países desenvolvidos (por exemplo, Reino Unido, França, Alemanha), bem como nas economias em desenvolvimento (por exemplo, Brasil, Rússia, Índia). O conceito de dados abertos é tipicamente apresentado através de conjuntos brutos de dados, em formatos que máquinas conseguem ler, que podem ser usados por cidadãos e instituições como eles quiserem (por exemplo, no desenvolvimento de sistemas e aplicações, análise de dados utilizando ferramentas de software, etc). [1]

A consolidação dos dados e a criação dos portais que oferecem conjuntos de dados brutos em grande quantidade já seguem normas ditadas pelas iniciativas de governos abertos. Esta estratégia de publicação de dados permite terceiros construírem aplicações que utilizam as fontes dos dados para relacioná-los, analisá-los e extrair informações relevantes para o objetivo de cada uso. No entanto, o acesso a essas fontes de dados é tipicamente feito manualmente. Aqueles que desejam utilizar esses dados precisam acessar o portal, copiar os volumes de dados, em seguida, extrair o que é relevante e, finalmente, integrar os dados em suas aplicações que devem acessar os bancos de dados locais com cópias completas ou parciais dos dados abertos adquiridos.

Apesar dos vários portais de dados abertos que têm sido lançados nos últimos anos a identificação de dados relevantes, integração de dados distribuídos, falta de metadados e padrões, bem como a qualidade dos dados divulgados são bons exemplos de desafios a serem enfrentados quando se inicia uma iniciativa de dados abertos. De acordo com o relatório McKinsey algumas questões-chave de alto nível têm que ser abordadas a fim de compreender o potencial total do valor dos dados abertos como a criação de um local de dados abertos que fornece canais claros de compartilhamento de dados e constrói uma comunidade com regras. [2]

Esse canal de compartilhamento e integração entre as partes de um ecossistema de dados abertos é descrito por Ding et al. como " um sistema baseado em Linked Data em que as partes interessadas de diferentes tamanhos e funções encontram , gerenciam, arquivam, publicam, reutilizam, integram, misturam, geram e consomem dados governamentais abertos por meio de ferramentas on-line, serviços e conexões com a sociedade". [3]

Neste contexto, novas soluções de dados abertos são necessárias para lidar com as limitações de abordagens existentes e para ajudar a alavancar o valor econômico dos dados abertos. Por exemplo, serviços que podem ser facilmente combinados para construir aplicações, incluindo dados de serviços de publicação e análise de dados; uma infraestrutura que pode fornecer escalabilidade, disponibilidade e confiabilidade nos serviços de dados disponíveis, ligação transparente entre as fontes dos dados e integração de dados a partir de diferentes fontes.[4]

## 2. Objetivos

Este trabalho de graduação tem como objetivo a especificação e criação de um vocabulário para descrever o ecossistema de dados abertos governamentais. O vocabulário promoverá a integração dos diferentes atores, permitindo que o desenvolvimento de produtos e serviços baseados nos dados abertos governamentais proporcionem benefícios práticos para os usuários. Como resultado será gerada uma base de conhecimento que permitirá a produção de inferências, recomendações, relatórios sobre os consumos dos dados abertos e geração de novos conjuntos de dados sobre ecossistemas.

Como objetivos específicos deste trabalho destacam-se:

- Mapeamento dos atores e relacionamentos entre eles.
- Especificação e criação do vocabulário para o ecossistema de dados abertos governamentais
- Instanciamento dos atores e relacionamentos
- Geração da base de conhecimento
- Visualização e consumo dos dados

### 3. Cronograma

Atividades	Agosto				Setembro				Outubro				Novembro				Dezembro			
Estudo da literatura	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		
Mapeamento dos atores do ecossistema de dados abertos	x	x	x	x																
Especificação do ecossistema de dados abertos			x	x																
Criação do vocabulário					x	x	x	x												
Instância dos atores e relacionamentos								x	x	x										
Processo de consumo dos dados e resultados										x	x	x	x							
Elaboração do relatório final													x	x	x	x				
Preparação para apresentação e defesa																	x	x		

Tabela 1: Cronograma de Atividades

## **4. Possíveis Avaliadores**

Os possíveis avaliadores deste trabalho de graduação serão:

- Bernadette Farias Lóscio
- Kiev Santos da Gama

## 5. Referências

- [1] Dietrich, D., Gray, J., McNamara, T., Poikola, A., Pollock, P., Tait, J., & Zijlstra, T. (2009). Open data handbook. 2013-01-29]. <http://opendatahandbook.org>.
- [2] Manyika, J., Chui, M., Groves, P., Farrell, D., Van Kuiken, S., & Doshi, E. A. (2013). Open data: Unlocking innovation and performance with liquid information. *McKinsey Global Institute*, 21.
- [3] Ding, L., Lebo, T., Erickson, J. S., DiFranzo, D., Williams, G. T., Li, X., ... & Flores, J. (2011). TWC LOGD: A portal for linked open government data ecosystems. *Web Semantics: Science, Services and Agents on the World Wide Web*, 9(3), 325-333.
- [4] Gama, K., & Lóscio, B. F. (2014). Towards Ecosystems based on Open Data as a Service. In *ICEIS* (2) (pp. 659-664).

## 6. Assinaturas

---

Glória de Fátima Andrade Barros Lima  
**Orientando**

---

Bernadette Farias Lóscio  
**Orientador**