

## Plano de Dados Abertos na Universidade Federal de Pernambuco – UFPE

### Open Data Plan at the Federal University of Pernambuco – UFPE

### Plan de Datos Abiertos en la Universidad Federal de Pernambuco - UFPE

*Bernadette Farias Lóscio* | [bfl@cin.ufpe.br](mailto:bfl@cin.ufpe.br)

Universidade Federal de Pernambuco. Recife, Brasil.

*Susimery Vila Nova* | [susimery.ufpe@gmail.com](mailto:susimery.ufpe@gmail.com)

Universidade Federal de Pernambuco. Recife, Brasil.

## Resumo

Esta comunicação se destina a gestores de informação que atuem na execução da LAI – Lei de Acesso à informação, bem como bibliotecários, dirigentes, pesquisadores, decisores políticos e interessados pelo tema. O objetivo é apresentar os pilares de criação do Plano de Dados Abertos (PDA) da Universidade Federal de Pernambuco (UFPE), articuladas de forma colaborativa com o Grupo de Trabalho em Dados Abertos composto pela Procit e Centro de Informática (CIn) da UFPE. Este, por sua vez, desenvolve e atua em todo o processo para garantia da abertura de dados na Instituição.

**Palavras-chave:** Acesso aberto; Dados abertos; Universidade Federal de Pernambuco.

## Abstract

This communication is intended for information managers who act in the implementation of the LAI, as well as librarians, managers, researchers, policy makers and interested parties. The objective is to present the pillars of creation of the Open Data Plan (PDA) of the Federal University of Pernambuco (UFPE), articulated in a collaborative way with the Working Group on Open Data composed by Procit and Center of Informatics (CIn) of UFPE. This, in turn, develops and operates throughout the process to guarantee the opening of data in the Institution.

**Keywords:** Open access; Data open; Federal University of Pernambuco.

## Resumen

Esta comunicación se dirige a gestores de información que actúen en la ejecución de LAI - Ley de Acceso a la información, así como bibliotecarios, dirigentes, investigadores, decisores políticos e interesados por el tema. El objetivo es presentar los pilares de creación del Plan de Datos Abiertos (PDA) de la Universidad Federal de Pernambuco (UFPE), articulados de forma colaborativa con el Grupo de Trabajo en Datos Abiertos compuesto por la Procit y Centro de Informática (CIn) de la UFPE. Este, a su vez, desarrolla y actúa en todo el proceso para garantizar la apertura de datos en la Institución.

**Palabras clave:** Acceso abierto; Datos abiertos; Universidad Federal de Pernambuco.

## Introdução

A Universidade Federal de Pernambuco apresenta um grande diferencial de inovação dentre as Instituições de Ensino Superior Federal no Brasil a partir da criação da primeira unidade voltada para o desenvolvimento de políticas integradas de Comunicação, Informação e Tecnologia da Informação, a Pró-Reitoria PROCIT, criada em 2014.

Como suas unidades integradas estão o Núcleo de Tecnologia da Informação (NTI), a Editora Universitária, o Setor de Inovação Educacional (Conecte), o Núcleo de TV e Rádios Universitárias (NTVRU), o Sistema Integrado de Bibliotecas (SIB), e a Assessoria de Comunicação Social (ASCOM).

Como umas das estratégias para modelagem da política de informação institucional, a coordenação de Padrões e Conteúdos, vinculada à Diretoria de Informação da PROCIT, desenvolve ações articuladoras junto aos dirigentes e órgãos suplementares da Universidade, no sentido de instigar e promover discussões acerca do acesso aberto ao conhecimento e à democratização da informação.

Dentre as ações desenvolvidas pela PROCIT, destaca-se a criação do Plano de Dados Abertos da UFPE. De acordo com o Manual para Elaboração de Plano de Dados Abertos, o PDA “é o documento que vai orientar as ações de implementação e promoção de abertura de dados, inclusive os geoespacializados, que deverão obedecer aos padrões mínimos de qualidade, de forma a facilitar o entendimento e a reutilização das informações. É ele quem organiza o planejamento referente à implantação e racionalização dos processos de publicação de dados abertos nas organizações públicas.”

Vale salientar que o PDA atende as exigências estabelecidas no Decreto nº 8.777 de 2016<sup>i</sup> e suas ações estão em consonância com o disposto na Lei de Acesso à Informação – LAI<sup>ii</sup>; no art. 48 da Lei Complementar no 101, de 4 de maio de 2000<sup>iii</sup>; na Instrução Normativa SLTI no 4, de 13 de abril de 2012 (que institui a Infraestrutura Nacional de Dados Abertos)<sup>iv</sup>; no Decreto Presidencial no 6.666, de 27 de novembro de 2008

i [http://www.planalto.gov. br/ccivil\\_03/ Ato2015- 2018/2016/Decreto/ D8777.htm](http://www.planalto.gov.br/ccivil_03/ Ato2015- 2018/2016/Decreto/ D8777.htm)

ii [http://www.planalto.gov. br/ccivil\\_03/ ato2011- 2014/2011/lei/l12527.htm](http://www.planalto.gov. br/ccivil_03/ ato2011- 2014/2011/lei/l12527.htm)

iii [http://www.planalto.gov. br/ccivil\\_03/leis/lcp/ lcp101.htm](http://www.planalto.gov. br/ccivil_03/leis/lcp/ lcp101.htm)

iv <http://dados.gov.br/ instrucao-normativa-dainda/>

(que institui a Infraestrutura Nacional de Dados Espaciais)<sup>v</sup>; assim como com os compromissos assumidos pelo Governo no âmbito da Parceria para Governo Aberto (Open Government Partnership – OGP)<sup>vi</sup>.

O objetivo desta comunicação, portanto, é apresentar os pilares de criação do Plano de Dados Abertos (PDA) da Universidade. Seu processo de criação foi articulado de forma colaborativa a partir de reuniões realizadas com o Grupo de Trabalho em Dados Abertos composto pela Procit e Centro de Informática (CIn) da UFPE. Este, por sua vez, desenvolve e atua em todo o processo para garantia da abertura de dados na Instituição.

Ao contrário da grande maioria dos Planos de Dados Abertos de outras instituições governamentais, o PDA da UFPE faz uso de alguns fundamentos que são essenciais para a garantia da sustentabilidade da iniciativa de dados abertos.

Vale ressaltar que um dos principais desafios enfrentados pelas iniciativas de dados abertos diz respeito à continuidade do processo de abertura e garantia da qualidade dos dados que são publicados. Sendo assim, atualmente, diversos portais de dados abertos oferecem dados pouco relevantes, desatualizados ou até mesmo difíceis de serem compreendidos e manipulados. Isso acontece porque não são consideradas ações ou fundamentos que contribuam diretamente para a sustentabilidade da iniciativa.

Nesse contexto, a fim de garantir o sucesso e a sustentabilidade da iniciativa de abertura de dados, o PDA da UFPE foi fundamentado nos seguintes pilares:

### *1. O Ciclo de vida dos dados abertos*

O processo de publicação e consumo dos dados abertos da UFPE deverá seguir o ciclo de vida dos dados na Web composto pelas seguintes etapas:

- i. **Preparação:** Esta fase se estende desde o momento em que surge a intenção de publicar os dados até a seleção dos dados que serão publicados;
- ii. **Criação:** momento em que os dados são criados, ou seja, compreende a etapa de extração dos dados de fontes de dados já existentes até a sua transformação para o formato adequado para publicação na Web. Também devem ser criados os metadados que irão descrever os dados;
- iii. **Avaliação:** Nesta fase, os dados já devem estar prontos para publicação e deverão ser avaliados por especialistas do domínio a fim de garantir a sua validade;
- iv. **Publicação:** Compreende o momento em que os dados serão disponibilizados de forma pública na Web. Para isso, podem ser usadas ferramentas de catalogação de dados, como CKAN e Socrata, e APIs;
- v. **Consumo:** Implica o momento em que os dados são usados para a criação de visualizações, como gráficos e mapas de calor, bem como para aplicações que permitem o cruzamento e a realização de análises sobre os dados;
- vi. **Feedback:** Esta fase compreende o momento em que os consumidores devem prover comentários sobre os dados e metadados previamente utilizados;
- vii. **Refinamento:** Esta fase compreende todas as atividades relacionadas a adições ou atualizações nos dados que já foram publicados.

Com o uso do ciclo de vida espera-se que os dados abertos sejam frequentemente revisados e atualizados. Além disso, um maior incentivo à publicação contínua de novos dados, prezando sempre pela sua qualidade e relevância, também é esperado.

v [http://www.planalto.gov.br/ccivil\\_03/ Ato2007- 2010/2008/Decreto/ D6666.htm](http://www.planalto.gov.br/ccivil_03/ Ato2007- 2010/2008/Decreto/ D6666.htm)

vi <http://www.opengovpartnership.org/>

## 2. Boas Práticas para publicação de dados na Web

Durante todas as etapas do ciclo de vida serão utilizadas as Boas Práticas para Dados na Web (DWBP) recomendadas pelo W3C. Com uso dessas Boas Práticas, espera-se a publicação de conjuntos de dados que possam ser facilmente reutilizados por diferentes grupos de consumidores. As Boas Práticas para Dados na Web foram propostas a fim de auxiliar publicadores e consumidores de dados a enfrentar os principais desafios encontrados na publicação/consumo de dados na Web. Dentre esses desafios, destacam-se:

- i. **Metadados:** os metadados podem ser definidos como “dados que descrevem os dados” e são usados para auxiliar a descoberta e a reutilização de dados e/ou conjunto de dados;
- ii. **Licença de Dados:** Uma licença é uma especificação ou contrato que determina as condições de utilização de algo. No caso de dados, as licenças podem ser usadas para explicitar as condições e as possíveis formas de utilização de um determinado conjunto de dados;
- iii. **Proveniência de Dados:** Proveniência de dados consiste em um conjunto de informações que oferece detalhes sobre a história dos dados aos seus usuários, possibilitando o rastreamento da origem dos dados;
- iv. **Versionamento:** Tendo em vista que os dados publicados na Web são compartilhados em grande escala, torna-se fundamental oferecer informações sobre o versionamento dos conjuntos de dados, bem como garantir que os conjuntos de dados estão sendo atualizados de acordo com a frequência de atualização previamente estabelecida;
- v. **Identificação:** A Web provê um sistema de identificação único baseado no conceito de URI. Uma URI é um identificador que pode ser usado para qualquer recurso, incluindo aqueles que não estão disponíveis na Web, tais como pessoas ou imóveis. Existem diversos esquemas de URI, porém nem todos podem levar a identificação de recursos que podem ser localizados na Internet;
- vi. **Mecanismos de Acesso:** A infraestrutura da Web oferece diferentes métodos de acesso aos dados por meio de protocolos de comunicação como HTTP. Um consumidor de dados deve ser capaz de requisitar a recuperação de arquivos específicos ou conjuntos de dados inteiros de um determinado domínio de informação, sendo assim, torna-se necessário o suporte de diferentes abordagens de acesso aos dados, como *download* simultâneo de múltiplos arquivos (também conhecido como *bulk download*) e as Web APIs;
- vii. **Formatos de Dados:** Os formatos nos quais os dados estão disponíveis para os consumidores de dados são de grande importância para garantir um bom nível de utilidade, ou seja, para garantir que de fato os dados são úteis para os consumidores. Atualmente, existe uma grande variedade de formatos de dados disponível para publicação e troca de dados. Contudo, nem todos os formatos oferecem uma estrutura adequada que facilite o uso e reúso;
- viii. **Vocabulários:** Em qualquer cenário no qual os dados são oriundos de fontes de dados diversas, a interoperabilidade semântica é um desafio que deverá ser enfrentado. A representação adequada dos conceitos que descrevem o domínio dos dados sendo publicados é de fundamental importância para garantir que diferentes provedores e consumidores de dados compartilhem a mesma visão da realidade; e
- ix. **Feedback:** Publicar dados na Web possibilita o compartilhamento de dados em grande escala com um amplo público, o qual pode possuir diferentes níveis de *expertise*. Nesse contexto, os provedores de dados desejam ter a garantia de que os dados publicados estão atendendo aos requisitos dos consumidores de dados e, para alcançar esse objetivo, a aquisição de *feedback* dos usuários torna-se fundamental.

### 3. Uma sólida estrutura de governança

A governança do PDA será exercida pelos seguintes atores:

- i. O Comitê Gestor de Tecnologia da Informação da UFPE (CGTI): tem a função de acompanhar a execução do PDA no nível estratégico;
- ii. O Laboratório de Dados e Informações (aLADIN): é responsável pela publicação de novos conjuntos de dados e refinamento de conjuntos de dados existentes, considerando a lista de prioridades estabelecida pelo PDA e as demandas recebidas pelas unidades setoriais da UFPE;
- iii. Os Responsáveis pelas Unidades Setoriais (RUS): são encarregados de coordenar os processos de abertura de dados das respectivas unidades de acordo com o plano de ação do PDA e
- iv. O Núcleo de Tecnologia da Informação (NTI): é responsável por oferecer e gerenciar toda a infraestrutura necessária para hospedar o portal de Dados Abertos da UFPE.

Um dos diferenciais da estrutura de governança do PDA da UFPE diz respeito à implantação do Laboratório de Dados e Informações. Dentre as responsabilidades do aLADIN, destacam-se:

- i. Atender às demandas para publicação de conjuntos de dados. Estas demandas serão identificadas a partir do processo de levantamento de dados, o qual deverá ser contínuo;
- ii. Solicitar aos setores responsáveis que forneçam informações complementares e realizem a validação dos dados antes da publicação dos dados;
- iii. Promover o consumo dos dados a partir do desenvolvimento de aplicações, análises e visualizações dos dados;
- iv. Coletar estatísticas sobre o consumo dos dados publicados; e
- v. Avaliar a maturidade do processo de publicação de dados abertos da UFPE.

### Considerações finais

O PDA pretende ser um marco geral de implantação da iniciativa de dados abertos da UFPE. Portanto, é um instrumento de planejamento e coordenação das ações de disponibilização de dados na Universidade Federal de Pernambuco válidas para o biênio 2017-2019.

Uma vez que os dados sejam disponibilizados em formato aberto, espera-se o seu uso no desenvolvimento de aplicativos que possam ser facilmente acessados pela sociedade, em especial, por servidores e discentes da UFPE. Tais aplicativos devem oferecer meios para análise dos dados, por meio de filtros, bem como permitir a visualização de dados de forma simples e criativa.

Espera-se que estes aplicativos contribuam para a melhoria dos serviços prestados pelos diferentes setores da UFPE, auxiliando na resolução de problemas enfrentados pela instituição, seus membros e sociedade em geral; bem como, sirva de referência e estímulo às discussões acerca da também abertura de dados científicos na Universidade, coadunado às práticas já iniciadas pelo Repositório Institucional da UFPE.

### Referências

1. Lóscio, B.F., Burle, C., e Calegari, N.: Data on the Web Best Practices. W3C Recommendation (2017). Disponível em: <https://www.w3.org/TR/dwbp/> . Acessado em: 1 de Março de 2017.
2. W3C Brasil. Manual dos Dados Abertos: Governo. Disponível em: [http://www.w3c.br/pub/Materiais/PublicacoesW3C/Manual\\_Dados\\_Abertos\\_WEB.pdf](http://www.w3c.br/pub/Materiais/PublicacoesW3C/Manual_Dados_Abertos_WEB.pdf) . Acessado em: 1 de Março de 2017.