

Dados abertos sem fricção

*Documentando, validando e publicando
dados com datapackages*

Diretoria Central de Transparência Ativa
Controladoria Geral do Estado de Minas Gerais



O QUE?

Boas práticas de documentação e validação de dados e metadados

COMO?

Uso de aplicativos online e sua relação com conceitos e boas práticas de dados abertos

EXPECTATIVAS:

- Criar um arquivo que documente os metadados de um conjunto de dados;
- Iniciar o controle de versão e a validação automática do conjunto criado;
- Identificar e corrigir erros mais comuns do fluxo de validação de dados

Problemas comuns na publicação de dados

Nomes dos cabeçalhos

Valores dos campos

Padrões de valor dos campos

Formatos de número e data

descsitser	nmefet	tem_ost	desccomi	descinst	descunid	carga_hora	remuner	data
ATIVO	CABO	0		CBMMG	5 CIA IND(SETE LAGO	40	5692	01/12/2020
ATIVO	AGENTE DE SEGURANCA SOCIOEDUCATIVO			SECRETARIA DE JUSTICA E SE	UNIDADE PRISIONAL	40	4.778,50	01/12/2020
ATIVO	SD 1 CL			POLICIA MILITAR DE MINAS	RCAT/CPE	40	4508,26	01/12/2020
ATIVO	PROFESSOR DE EDUCACAO BASICA			SECRETARIA DE EDUCACAO	EE PE PICCININI	21	3124,36	01/12/2020
ATIVO	AUXILIAR DE SERVICOS DE EDUCACAO BASICA			SECRETARIA DE EDUCACAO	EE SAO JOSE	30	1091,13	01/12/2020
ATIVO	PROFESSOR DE EDUCACAO BASICA			SECRETARIA DE EDUCACAO	EE DELANO BROCHA	15	1463,3	01/12/2020
INATIVO	PROFESSOR DE EDUCACAO BASICA	Sim		SECRETARIA DE EDUCACAO	EE N SRA DA PIEDAD	0	3769,21	01/12/2020
ATIVO	PROFESSOR DE EDUCACAO BASICA			SECRETARIA DE EDUCACAO	EE N SRA DA PIEDAD	21	4.239,54	01/12/2020

Problemas comuns e documentação: metadados

- Upload de arquivos incorretos ou alteração no formato dos arquivos (csv -> xlsx);
- Alteração na ordem ou nos nomes das colunas
- Alteração de leiautes sem comunicação
- Documentação inexistente para dados de referência (descrições, interpretações e domínio das colunas) ou sem formato estruturado ou não disponibilizada
- **METADADOS: significado dos dados, sua estrutura, qualidade, atualizações, licenças e outros atributos; níveis de granularidade e taxonomias**
- **Metadados insuficientes: dados não serão descobertos ou não serão reutilizados.**

Introdução ao Datapackage

- qualidade: grau em que um conjunto de características inerentes [ie. metadados] de um objeto [ie. dados] atende a especificação [ie. padrão]

- Datapackage: padrão de metadados, com formato preciso de arquivo (json); facilmente compartilhado e que permite reuso dos dados

“Se os dados não estão disponíveis num formato aberto e legível por máquina, eles não podem ser reutilizados.” (David Eaves/Opendata Charter - principles)

Utilização de especificações da Frictionless Data

O atrito ou “fricção” ocorre quando os consumidores gastam tempo e recursos demais apenas para poder entender e trabalhar com os dados

Criando um Datapackage: mão na massa!

Passo 1: <https://github.com/andrelgamei/inova>

Passo 2: <https://create.frictionlessdata.io/>

a) No painel à esquerda: o conjunto de dados e seus metadados:

name, title, profile, description, home page, version, author, licence, keywords

b) No painel central, quando se clica no ícone da engrenagem: cada recurso do conjunto:

name, path, title, profile, format, encoding, description

c) No painel central, após carregar cada recurso: as colunas de cada recurso (arquivo ou URL) que contém o dicionário de dados (~ schema), com

name, title, description, type (texto, datas, números), format (emails, dia/mês/ano, números com vírgulas como separador decimal)

Controle de versão e validação contínua

VERSIONAMENTO:

• <https://github.com/>

- criação de um repositório;
- histórico das alterações no repositório;
- boas práticas de alterações;
- ***Documentação é uma conversa!***

VALIDAÇÃO CONTÍNUA:

• <http://goodtables.io/>

- autorização no github;
- sincronização do repositório;
- upload dos arquivos e resultado da validação;
- Inserção do *badge* de validação

Requisitos legais dos dados abertos

[Lei de Acesso à Informação \(12527/2011\) art. 8º](#)

§ 3º Os sítios [oficiais de governo] deverão atender aos requisitos (dentre outros):

II - possibilitar a gravação de relatórios em diversos formatos eletrônicos, inclusive abertos e não proprietários, tais como planilhas e texto, de modo a facilitar a análise das informações;

III - possibilitar o acesso automatizado por sistemas externos em formatos abertos, estruturados e legíveis por máquina;

IV - divulgar em detalhes os formatos utilizados para estruturação da informação;

Requisitos legais dos dados abertos

[Decreto Federal 8777/2016](#): art. 2º, inciso IV:

formato aberto implica que a especificação esteja documentada publicamente

Dados Abertos ([Resolução CGE 20/2014](#)), art. 2º, inciso VI:

dados públicos (1)

representados em meio digital (2),

estruturados em formato aberto (3),

processáveis por máquina (4)

referenciados na rede mundial de computadores (5),

disponibilizados sob licença aberta (6)

que permita sua livre reutilização, consumo ou cruzamento em aplicações digitais desenvolvidas pela sociedade.

O que fizemos?

- Conjunto de dados através do *datapackage creator* utilizando uma especificação
- Preparação do *datapackage* criado para controle de alteração no *github*
- Validação automática dos dados pela sincronização do repositório com o *goodtables*

O que isso significa? Promoção de qualidade e redução dos custos de acesso e compreensão dos dados

Para quê?

Aumentar a chance do uso dos dados e os benefícios desse uso como retornos em escala para a própria sociedade.

Indução de uma cultura de colaboração e governo aberto

Relacionamento com órgãos e entidades

- *Se você é um gestor público do Estado de Minas Gerais custodiante de dados e tem interesse em realizar a abertura dos dados da sua unidade, entre em contato pelo email . Estamos em busca de parcerias para construção e validação da infraestrutura e processo de publicação de dados abertos antes da expansão da Política de Dados Abertos no Estado de Minas Gerais.*
- Oportunidades: aprendizado mútuo, *data literacy*, incremento do controle social e participação ativa da comunidade, avaliações de usuários, parcerias

dados.mg.gov.br

transparencia@cge.mg.gov.br