Let's do a thought experiment...

If a piece of information only exists with the person who generated it, and that person is not available; Does this information really exists?

Ps. You from 6 months ago is no longer available (paraphrasing @mtholder)

Context

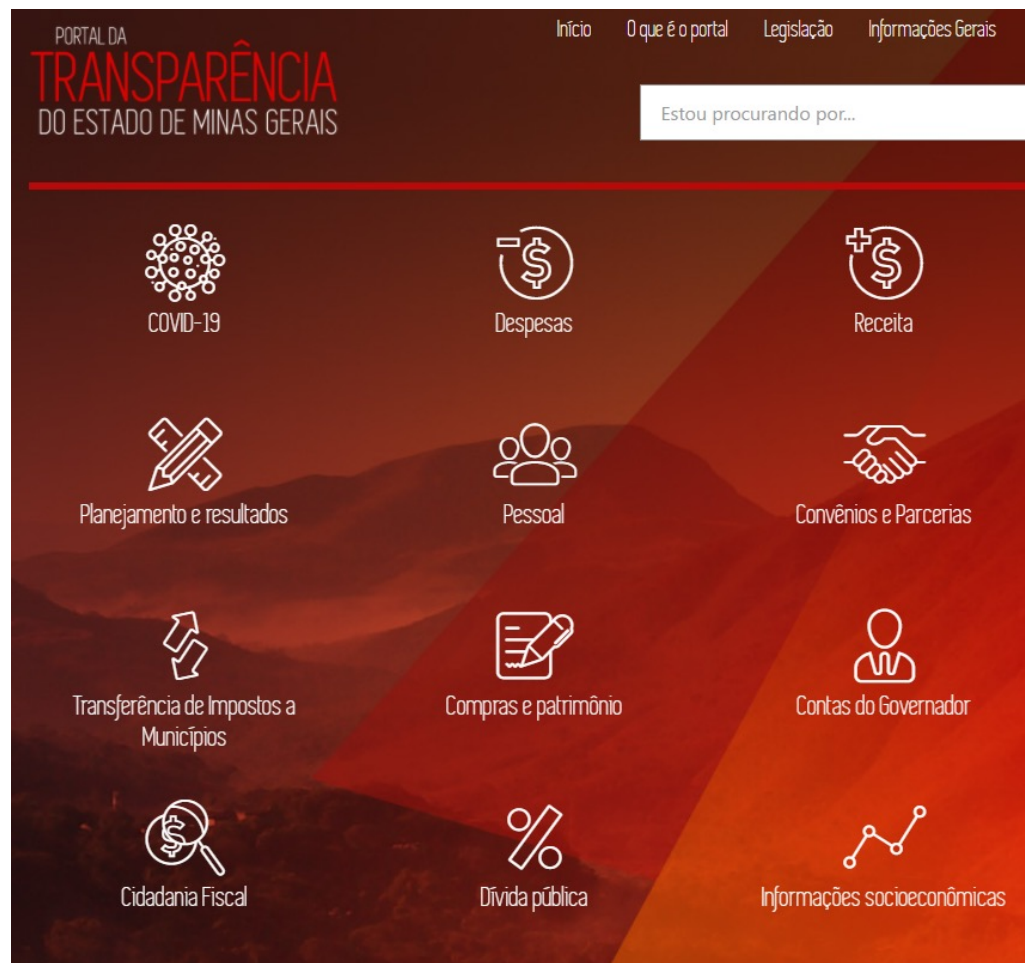# Proactive Transparency

> **Fiscal Transparency Portal**
> https://transparencia.mg.gov.br/

> **Open Data Portal**
> https://dados.mg.gov.br/

> **Agencies institutional websites**

Fiscal Transparency Portal
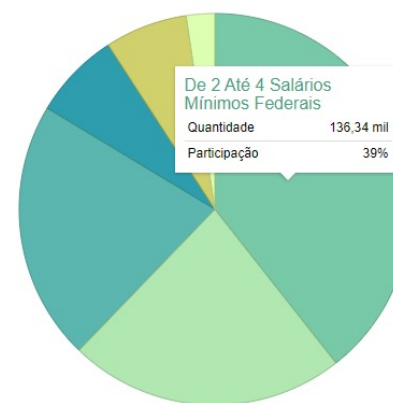
Docs as code

Docs as code

Publish

Publish the artifacts
without much
human intervention

Review

Ensure that a
trusted set of
reviewers
meticulously
reviews the docs

Build

Build the doc
artifacts
automatically

Write

Store the doc
source files in a
version control
system

Write

⟩ Doc source files

Doc source files stored in plain text formats (json + yaml + markdown) following Frictionless specs and standard practice in open source projects for documentation (README, CHANGELOG, CONTRIBUTING).

Using YAML multiline strings for long-form explanations in description fields allows for more readable and diffable documentation.

⟩ Version control system

A version control system allows the ability to group together a set of logically related changes and to permanently store the full history of changes.

Because the core operation in GitHub is 'diff' — what are the differences between two pieces of text — it elevates "Show me" as a social norm. Expressing, clearly and comparably, where you think a particular piece of text could or should be changed for the better becomes a core operation.

-- Clay Shirky, How the Internet will (one day) transform government

⌥ main ▾   **age7** / schemas / **dm-indicador-plan.yaml**

fjuniorr Fix block scalar notation in yaml …

👥 **2 contributors**

86 lines (86 sloc) 2.86 KB

```yaml
 1   fields:
 2     - name: id_indicador
 3       title: Indicador - Chave
 4       type: integer
 5       format: default
 6       bareNumber: true
 7       constraints:
 8         required: true
 9     - name: nome
10       title: Indicador
11       type: string
12       format: default
13       description: |
14         Nome descritivo do indicador do Programa.
15         É o título do indicador. Trata-se da forma pela qual o indicador será apresentado à sociedade.
16         O título deve ser congruente com a unidade de medida do indicador. Ou seja, se a unidade de medida
17         for percentual (%) sugere-se que o título do indicador se inicie com a expressão Percentual.
18       constraints:
19         required: true
```

Build

- Tech./Operational Metadata

- Human-readable docs

- Continous Validation

dados-mg / **termos-parceria-contratos-gestao** Public

<> Code ⊙ Issues 2 ⊓ Pull requests ▶ Actions ▦ Projects 📖 Wiki

✓ **atualiza readme e contributing** frictionless #22

⌂ Summary

Jobs

✓ **validate**

validate
succeeded 2 hours ago in 1m 5s

> ✓ Set up job

> ✓ Build frictionlessdata/repository@v0.8.0

> ✓ Checkout repository

> ✓ Validate data

> ✓ Post Checkout repository

> ✓ Complete job

Review

Public Repo (git)

A software engineer's enthusiasm for having high quality documentation varies inversely with their enthusiasm for being the person who provides it.

-- David Poole, [Documenting a SQL Server Data Catalog in HTML and Git Markdown](#)

Incentives to not let documentation go stale.

Collaboration

*Imagine if every time the government posted a dataset, rather than posting the data as a zip file or to a proprietary data portal, the agency treated the data as open source. All of a sudden data sets get a **running log of known issues**, and not just those known to the agency. Consumers of the data can **submit proposed changes** to do everything from normalizing columns to correcting errors to making the data itself more usable. Most importantly, as that data evolves over time, **there's a running log of exactly what's changed**, a critical feature in the regulatory context (e.g., what licenses were issued in the past week?).*

-- Ben Balter, [Treat Data As Code](#)

⑆ Merged **Adiciona dm_indicador_plan** #151

Andrelamor merged 3 commits into `main` from `dm_indicador_plan` ⧉ on Sep 22

○ 😎 Adiciona dm_indicador_plan   8ed187f

👤 🖼 **silviana-faria** self-assigned this on Sep 20

○ 🖼 Acrescenta título, descrição e atributo ao recurso dm_indicador_plan   4d1d10e

👤 🖼 **silviana-faria** assigned **fjuniorr** and unassigned **silviana-faria** on Sep 20

🖼 **Andrelamor** commented on Sep 21   [Member] ☺ ⋯

as colunas `nome` e `título` são idênticas; vide valores e descrição no schema; há uso possível para ambas?

➕ **New changes since you last viewed**   [View Changes]

○ 🖼 Merge branch 'main' into dm_indicador_plan   [Verified]   f70a52d

🖼 **silviana-faria** commented on Sep 22   [Member] ☺ ⋯

@Andrelamor realmente são iguais, mas no Portal elas estão em 'tabelas' diferentes. Acredito que seja devido a isso que as informações aparecem duas vezes.

Publish

Push-to-Deploy

Package managers for data

For those familiar with the Debian distribution system for Linux, the initial ideal is of "a debian of data". Through the 'apt' package management engine, when one installs a piece of software, all the libraries and other programs which it needs to run are walked through and downloaded with it.

        -- Jo Walsh and Rufus Pollock, Componentisation and Open Data, 2007

Is there a package manager for data? Not Big Data, just data.

        -- Kinrany, reddit, 2020

Portal de Dados Abertos do Estado de Minas Gerais

Conjuntos de dados   Organizações   Grupos   Documentação   Sobre   Pesquisar

🏠 / Organizações / Secretaria de Estado de ... / Contrato de Gestão e ...

Contrato de Gestão e Termos de Parceria

Seguidores
0

📄 **Organização**

Secretaria de Estado de Planejamento e Gestão - SEPLAG

A Secretaria de Estado de Planejamento e Gestão - SEPLAG - tem como competência: formular, propor, planejar e coordenar a ação governamental; promover a gestão estratégica e... Leia mais

📊 Conjunto de dados   👥 Grupos   ⊘ Fluxo de Atividades

Contrato de Gestão e Termos de Parceria

Documentos dos Termos de Parceria e Contratos de Gestão firmados pelos órgãos e entidades do Poder Executivo Estadual com OSCIPs e Organizações Sociais

Termos de Parceria e Contratos de Gestão

`frictionless passing`
`data valid`

Este conjunto de dados apresenta um extrato de informações dos Termos de Parceria e Contratos de Gestão firmados pelos órgãos e entidades do Poder Executivo Estadual com entidades privadas sem fins lucrativos, qualificadas como Organizações Sociais (OS) ou Organizações da Sociedade Civil de Interesse Público (Oscip), para a execução de políticas públicas.

Os instrumentos jurídicos que formalizam as parcerias são:

- Contrato de gestão, celebrado com Organização Social (OS); e
- Termo de parceria, celebrado com Organização da Sociedade Civil de Interesse Público (Oscip).

São disponibilizados os contratos de gestão e termos de parceria firmados com o Estado, acompanhados dos respectivos termos aditivos, os repasses previstos e realizados, bem como os relatórios das comissões de avaliação, os relatórios de monitoramento e os relatórios gerenciais.

Mais informações podem ser obtidas no sítio eletrônico da Secretaria de Estado de Planejamento e Gestão.

CHANGELOG

Arquivo para documentação das principais alterações sofridas por este conjunto de dados.

Dicionário de Dados

Coluna	Tipo	Rótulo	Descrição
ID_INSTRUMENTO	text	Número do Instrumento	Numeração do instrumento jurídico, estabelecida pela Secretaria de Estado de Planejamento e Gestão
TIPO_DOCUMENTO	text	Tipo de Documento	Informa o tipo de relatório a que se refere
DOCUMENTO	text	Documento	Informa o número do documento
PERIODO_AVALIATORIO	text	Período Avaliatório	Informa o período da avaliação do relatório
URL_DOCUMENTO	text	URL do Documento	Link para acessar o relatório
OBSERVACOES	text	Observações	Dados adicionais para esclarecer alguma informação sobre a linha

Demo

https://github.com/dados-mg/frictionless-hangout-jan2022

Next steps

⊘ dpckan dataset get

⊘ Respect table schema types in CKAN DataStore

⊘ Human readable viz of metadata (In static website and in CKAN)

- validation reports
- diagrams (mainly for PK and FK relationships)
- richer data dictionaries (missing values, enum, etc)

⊘ Contribute back upstream

- frictionless[ckan] plugin
- react components
- ckan / datapusher / xloader

Where's the fun in just knowing what the code is supposed to do?

Essential

Excuses for Not Writing Documentation

O RLY?

@ThePracticalDev

Thanks!