

Workshop

Práticas de reutilização de dados de investigação: das restrições ao potencial

GT - Formação e Competências para Gestão e Dados FAIR

12º Fórum GDI, Instituto Politécnico de Bragança, 28/11/2025

AGENDA

01

Introdução à reutilização de dados de investigação

- **André Vieira**, Universidade do Minho

02

Partilha de boas práticas de reutilização de dados

- **Ana Alice Baptista**,
Centro ALGORITMI, Universidade do Minho
- **Guilherme Schardong**,
Instituto de Sistemas e Robótica, Universidade de Coimbra

03

Grupos de discussão

- Análise de práticas que promovam a reutilização
- Discussão sobre as limitações e potencial da reutilização

04

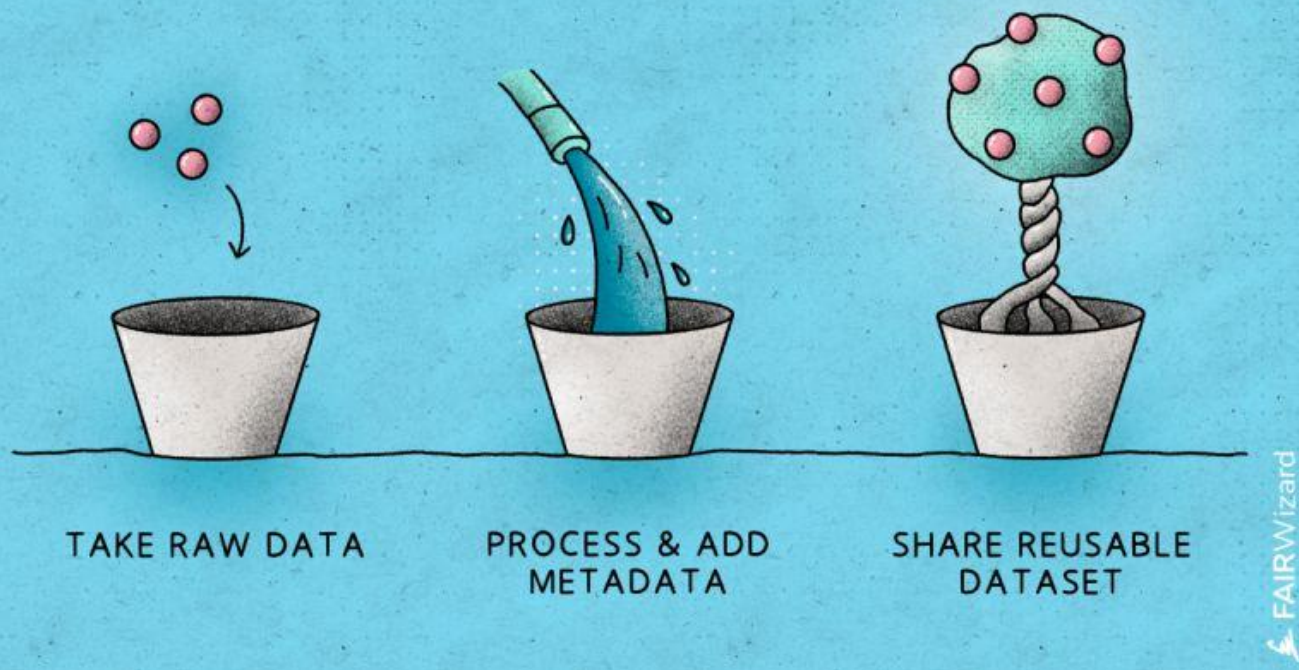
Conclusões

- Partilha das conclusões dos grupos de discussão

Introdução à reutilização de dados de investigação

Os dados em bruto não são suficientes!

GROW YOUR DATASET: FROM RAW TO REUSABLE



O valor dos dados só surge quando são enriquecidos com metadados, regidos por normas e preparados para reutilização.

É aí que deixam de ser um subproduto da investigação para se tornarem um ativo de longo prazo.

Uma boa gestão dos dados torna-os mais do que um simples registo. Torna-os a base para a descoberta, colaboração e inovação.

Os seus conjuntos de dados estão prontos para reutilização ou ainda estão à espera de crescer?

A sua experiência na reutilização de dados



partici.fi/97937483





Reutilização de dados de investigação

Significa usar dados para outros fins que não aqueles para os quais foram originalmente recolhidos.

A reutilização de dados é particularmente importante na ciência, pois permite que diferentes investigadores analisem e publiquem descobertas com base nos mesmos dados, independentemente uns dos outros.

Os dados que são bem descritos, organizados e partilhados sob termos e condições claros têm mais probabilidades de serem reutilizados.

A reutilização é um componente fundamental dos princípios dos dados FAIR.

Princípios de Reutilização (FAIR)

Os dados e os metadados têm potencial de reutilização quando são descritos com um nível de pormenor que permite a sua reprodução e aplicação em novos contextos.

R1

Os (meta)dados são descritos de forma rica, com uma pluralidade de atributos exatos e relevantes

R1.1

Os (meta)dados são divulgados com uma licença de utilização de dados clara e acessível

R1.2

Os (meta)dados têm associada, de forma detalhada, a sua proveniência

R1.3

Os (meta)dados cumprem as normas do domínio científico

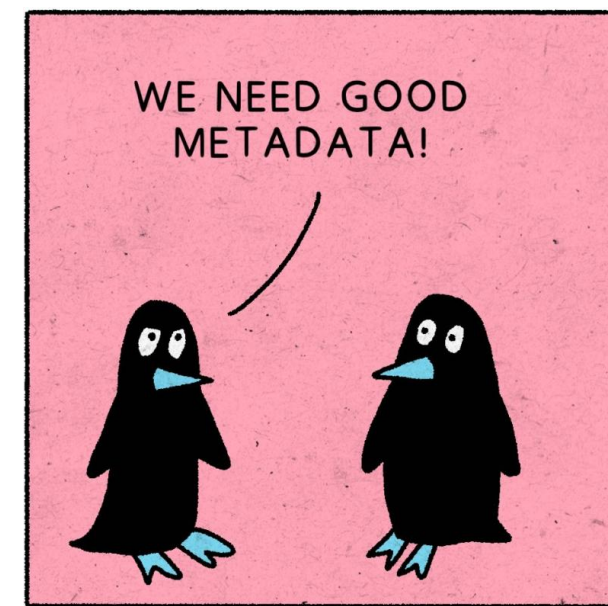


Imagem: FAIR Wizard,
<https://www.linkedin.com/company/fair-wizard>

Reutilização de dados: Qual a sua importância?

Ao reutilizar dados existentes, podemos:

- Obter dados de referência para a nossa investigação;
- Evitar duplicação de experiências;
- Verificar se as conclusões relatadas estão corretas, tornando assim as conclusões subsequentes mais robustas;
- Tornar a investigação mais robusta e transparente, agregando resultados obtidos a partir de diferentes métodos ou amostras;
- Obter novos contributos.

O que deve ser considerado para a reutilização de dados?

Reutilizar dados existentes implica verificar se as condições necessárias para a reutilização estão reunidas.

- **Procurar por dados relevantes** em bases de dados, repositórios, ou dados associados a publicações.
- **Verificar em que termos e condições os dados são partilhados.**
Que licença está associada?
- **Verificar se há metadados e documentação suficientes** para permitir a correta compreensão/reutilização dos dados.
- **Avaliar a qualidade dos dados:**
 - Os dados provêm de uma fonte confiável?
 - Estão em conformidade com normativos adequados?
- **Os dados foram recolhidos de forma ética? Possuem dados sensíveis?**
Devo seguir procedimentos adequados.
- **Os dados a ser reutilizados foram atualizados?**
Documente qual a versão que está a utilizar.
- **Cite os dados corretamente.**

DATA



METADATA



Partilha de boas práticas de reutilização de dados

Ana Alice Baptista,

Centro ALGORITMI, Universidade do Minho

Guilherme Schardong,

Instituto de Sistemas e Robótica, Universidade de Coimbra

Grupos de Discussão

- **Análise de práticas que promovem a reutilização**
- **Discussão sobre as limitações e potencial da reutilização**

Grupos de Discussão

I. Analisar práticas que promovem a reutilização

II. Refletir sobre as limitações e potencial da reutilização

- Duração: 20 minutos
- 4 Grupos
- 1 a 2 Porta voz por Grupo
 - Registo de anotações
 - Partilha das principais conclusões em plenário (2 min)

Atividade:

- I. Analisar dataset com base na grelha disponibilizada
- II. Refletir sobre as questões colocadas

I. PRINCÍPIOS DE REUTILIZAÇÃO DOS DADOS FAIR:

Significa que os dados estão bem documentados, são objeto de curadoria e fornecem informações ricas sobre o contexto da sua criação. Os dados devem estar em conformidade com os standards da comunidade e incluir termos e condições claros sobre a forma como os dados podem ser acedidos e reutilizados, de preferência através da aplicação de licenças normalizadas legíveis por máquina.

Grupo 1
Dataset para
avaliação:



Princípios de Reutilização	Sim	Não	Observações
<p>R1. Os (meta)dados são descritos de forma rica, com uma pluralidade de atributos exatos e relevantes. Devem ser fornecidos metadados que permitam, não só, a descoberta dos dados, mas também que descrevam de forma rica o contexto em que os dados foram gerados. Exemplos: protocolos experimentais usados; instrumentos, condições de operação; contextualização das variáveis; espécies usadas e em que condições; entre outros.</p>			
<p>R1.1. Os (meta)dados são divulgados com uma licença de utilização de dados clara e acessível. Os dados e metadados devem ter uma licença de uso associada que permita a humanos e máquinas a sua reutilização em conformidade com as disposições legais refletidas na licença. Exemplos: Licença Creative Commons ou MIT.</p>			
<p>R1.2. Os (meta)dados têm associada, de forma detalhada, a sua proveniência. A proveniência dos dados deve ser detalhada nos metadados, para garantir a sua devida reutilização por humanos e máquinas. Exemplo: São disponibilizados detalhes sobre como os dados foram gerados, porque foram gerados, por quem, em que condições, se foram processados, entre outros (alguns destes detalhes são disponibilizados no ficheiro Readme).</p>			
<p>R1.3. Os (meta)dados cumprem as normas do domínio científico. Seguir boas práticas existentes num determinado domínio científico, em relação ao arquivo e à partilha de dados, para maximizar a reutilização dos dados pela comunidade em questão. Existem domínios científicos com normas estabelecidas em relação aos requisitos mínimos que são necessários para a partilha de dados, que deverão ser seguidos para garantir que dados similares são descritos da mesma forma, com o mesmo vocabulário e que os mesmos formatos de ficheiros são usados. Exemplos:</p> <ul style="list-style-type: none"> • São aplicados normativos na descrição dos dados? • É documentado com exatidão todo o processo da análise dos dados, incluindo o software/código desenvolvido e utilizado, as ferramentas e todo o ambiente de análise? • Os ficheiros são disponibilizados em formatos abertos ou normalizados? 			

II. QUESTÕES DE DISCUSSÃO

1. Em que medida o cumprimento dos princípios de reutilização promove a correta reutilização dos dados?
2. Quais as principais limitações à reutilização de dados que identifica? (tanto no dataset analisado como com base na sua experiência).
3. O que considera mais relevante para potencializar uma efetiva reutilização dos dados?

Datasets para análise

GRUPO 1



GRUPO 2



GRUPO 3



GRUPO 4



Conclusões

Partilha das conclusões dos grupos de discussão



Obrigado!

Grupo de Trabalho de Competências e Formação para Gestão e Dados FAIR

<https://forumgdi.rcaap.pt/gt-formacao/>