health RI

*enabling data driven health*

# Introductie FAIR data

Jan-Willem Boiten; 9 maart 2022

Met veel slides van Rob Hooft

janwillem.boiten@health-ri.nl

# De FAIR principes – het begon met een workshop in Leiden



scientific **data**

Explore content | About the journal | Publish with us

nature › scientific data › comment › article

Open Access | Published: 15 March 2016

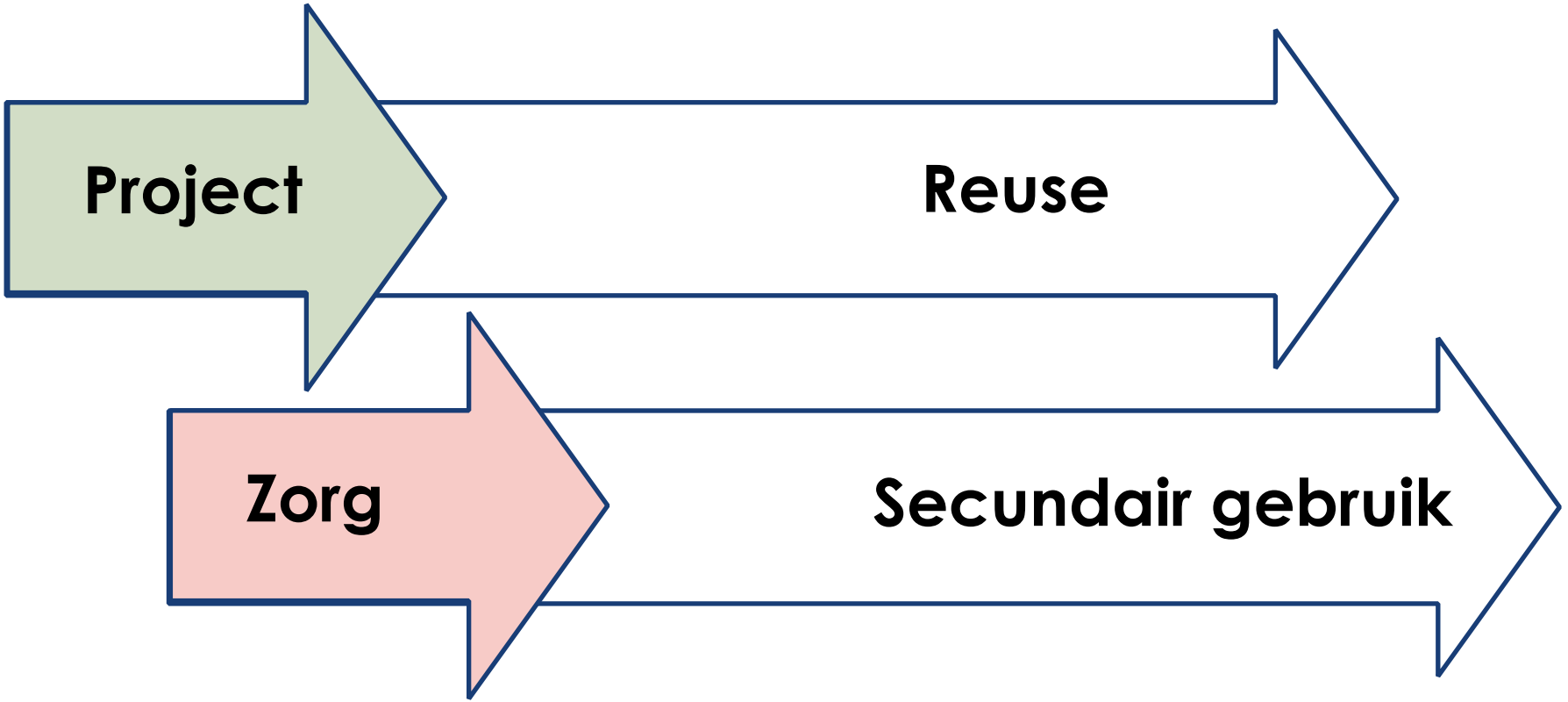
**The FAIR Guiding Principles for scientific data management and stewardship**

Mark D. Wilkinson, Michel Dumontier, ... Barend Mons + Show authors

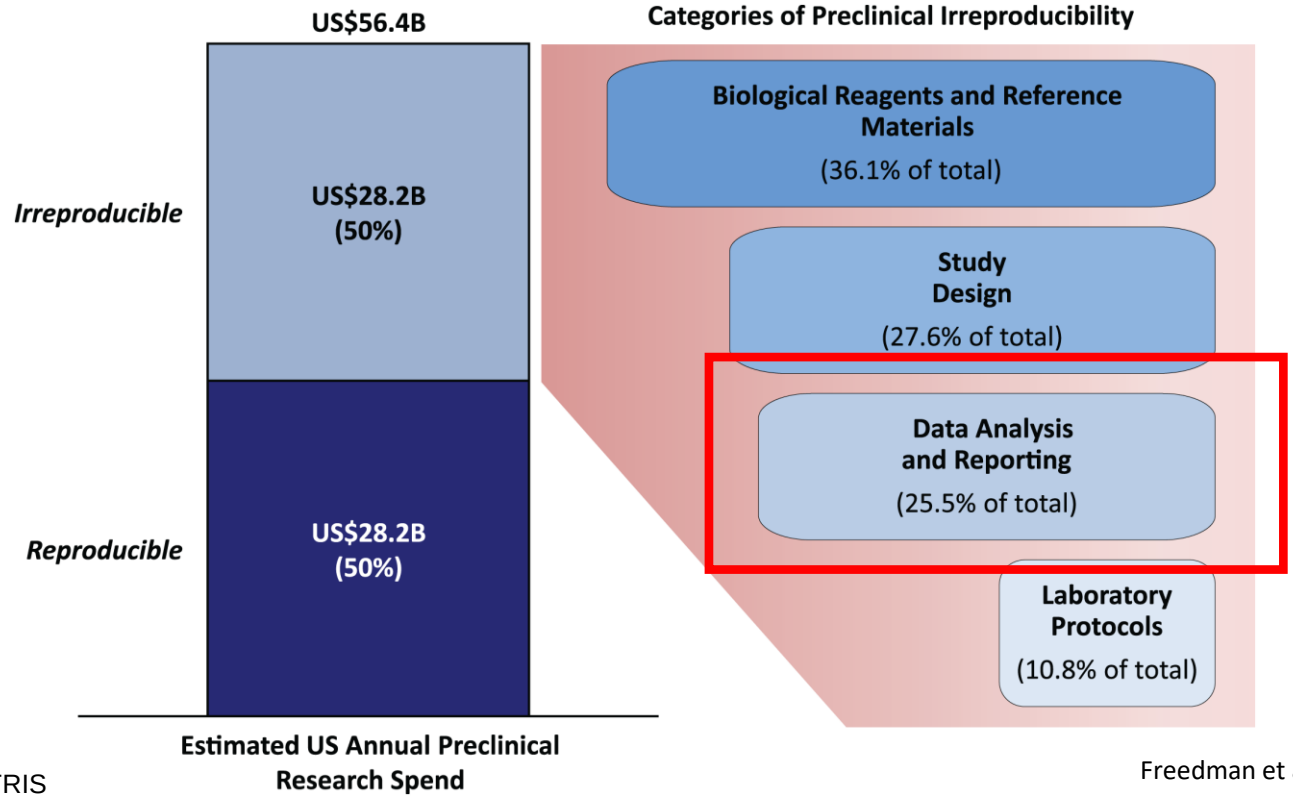
*Scientific Data* **3**, Article number: 160018 (2016) | Cite this article

**428k** Accesses | **3642** Citations | **1937** Altmetric | Metrics

# Data is van oudsher beperkt in hergebruik



# FAIR data – ook essentieel voor reproduceerbaarheid




Slide: Anton Ussi; EATRIS

Freedman et al PLOS Biology (2015)

THEME ARTICLE: SPECIAL SECTION ON THE 75TH ANNIVERSARY OF THE IEEE COMPUTER SOCIETY

# 75 Years of Astonishing Evolution of IT: 1946–2021

George Strawn, *US National Academy of Sciences, Engineering, and Medicine*

prescription not an implementation. As mentioned in the previous paragraph, currently data is "munged" manually for analysis rather than by automated processes. This results in data preparation/munging taking 80% of the time leaving only 20% for analytics (science or business). I hope/believe that in the decade of the 2020s, we will learn how to automate the process of making data FAIR, which will go a long way to making data *Fully AI-Ready* much more rapidly.

https://doi.org/10.1109/MITP.2021.3070963

I would like to exploit common genotype-phenotype relations between Alzheimer's Disease and Huntington's Disease…
I need to combine AD and HD data…

I can help with that!

Source: Marco Roos

???

DOES NOT COMPUTE

Here's my data, have fun!

米当局は、あなたの国籍故に、在米日本領事代表にあなたが逮捕又は拘禁されたことを通報する必要があります。領事官は通報を受けた後、あなたに電話を掛けたり、あるいはあなたを訪問することができます。あなたは領事官の援助を受ける必要はありませんが、あ
あなたの家族との連絡、身
置を取ってくれるかも知れ
事官に通報します。

Очи чёрные, очи жгучие,
Очи страстные и прекрасные,
Как люблю я вас, как боюсь я вас,
Знать увидел вас я не в добрый час.

Очи чёрные, очи пламенны
И монят они в страны дальные,
Где царит любовь, где царит покой,
Где страданья нет, где вражды запрет.

Here's my data, have fun!

Source: Marco Roos

???

DOES NOT COMPUTE

Here's my data, have fun!

## It is very hard to make data FAIR/interoperable afterwards:

- We need to use data standards from the start
- We need to supply sufficient meta data for others to interpret our data

## We need to be FAIR-by-design

Source: Marco Roos

# Findable

- Choose Repository
  - Domain specific
  - Institutional
  - National
  - Special
- Make sure to get and use Persistent Identifiers
- Describe what the data **is**, not only why you collected it
- Register the data in a catalogue, where others would look for it

# Accessible

- Assure Longevity
- Check Legal Conditions
- Limit the Embargo
- Use standard protocols also for machine actionability
- Ensure proper ICT procedures

# Interoperable

- Terminology
- Format
- Sufficient metadata

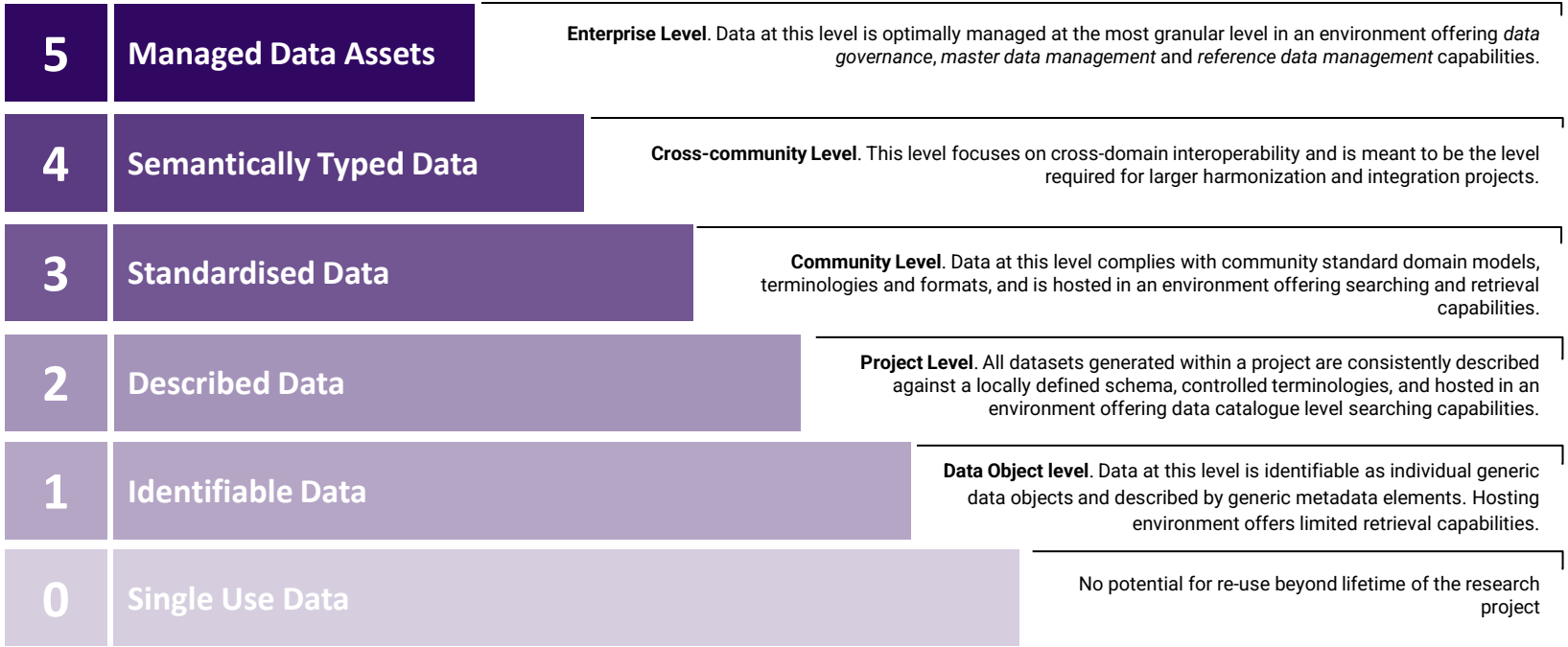
**Brussels ← → Bruxelles**

**Cancer ← → Malignant Neoplasm**

# Reusable

- Document provenance
- Use *minimal metadata* standards
- Choose a liberal license
- Make sure results are not only narrated

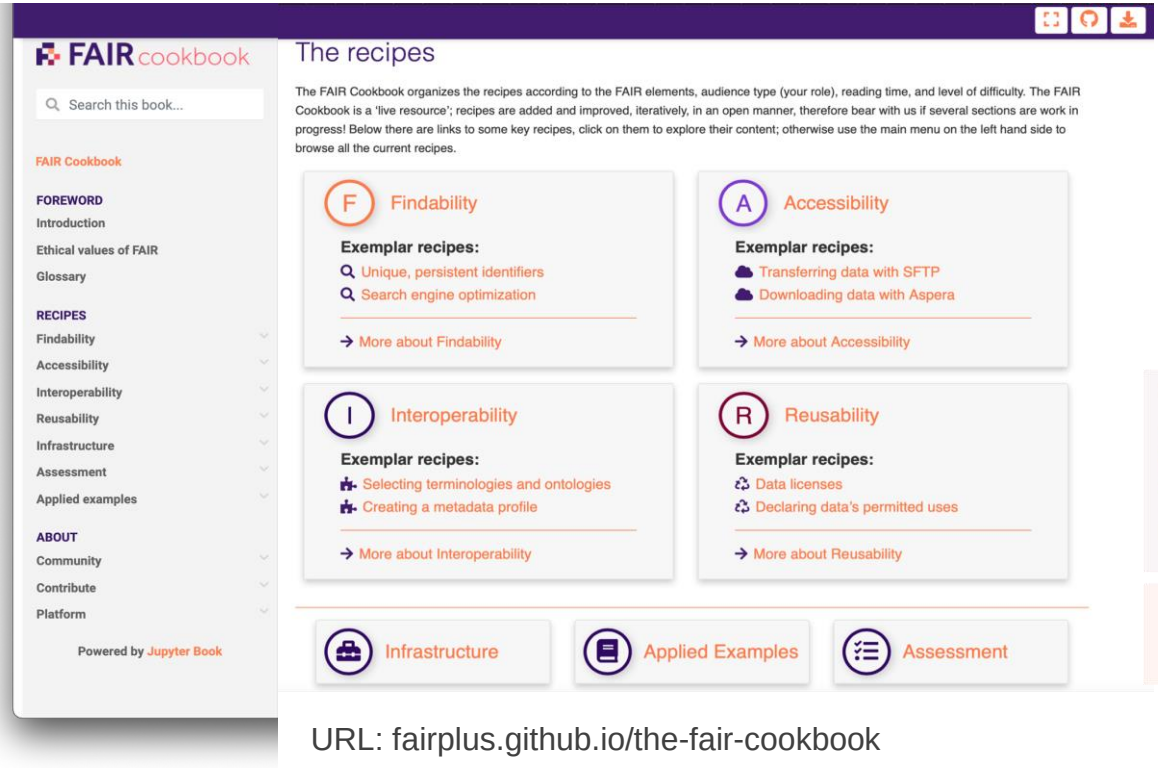
# FAIRplus Dataset Maturity Levels




Model - Ibrahim Emam

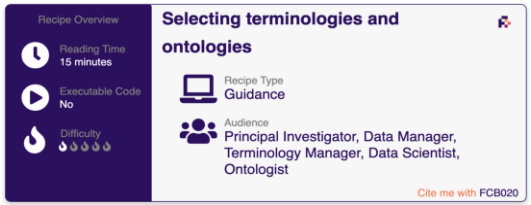
# FAIR cookbook: from knowledge to recipes

FAIRplus




URL: fairplus.github.io/the-fair-cookbook

**Examples:**




**Authors:**


…..and open to other contributions, because FAIR is a 'team sport'!
fairplus-cookbook@elixir-europe.org

WEBINAR: elixir-europe.org/events/fairplus-webinar-discovering-fair-cookbook

Slide – Susanna Sansone

# Tot slot….

FAIR ≠ Open Science