FAIR FOR PLAN

KARSTEN KRYGER HANSEN
CHIEF CONSULTANT
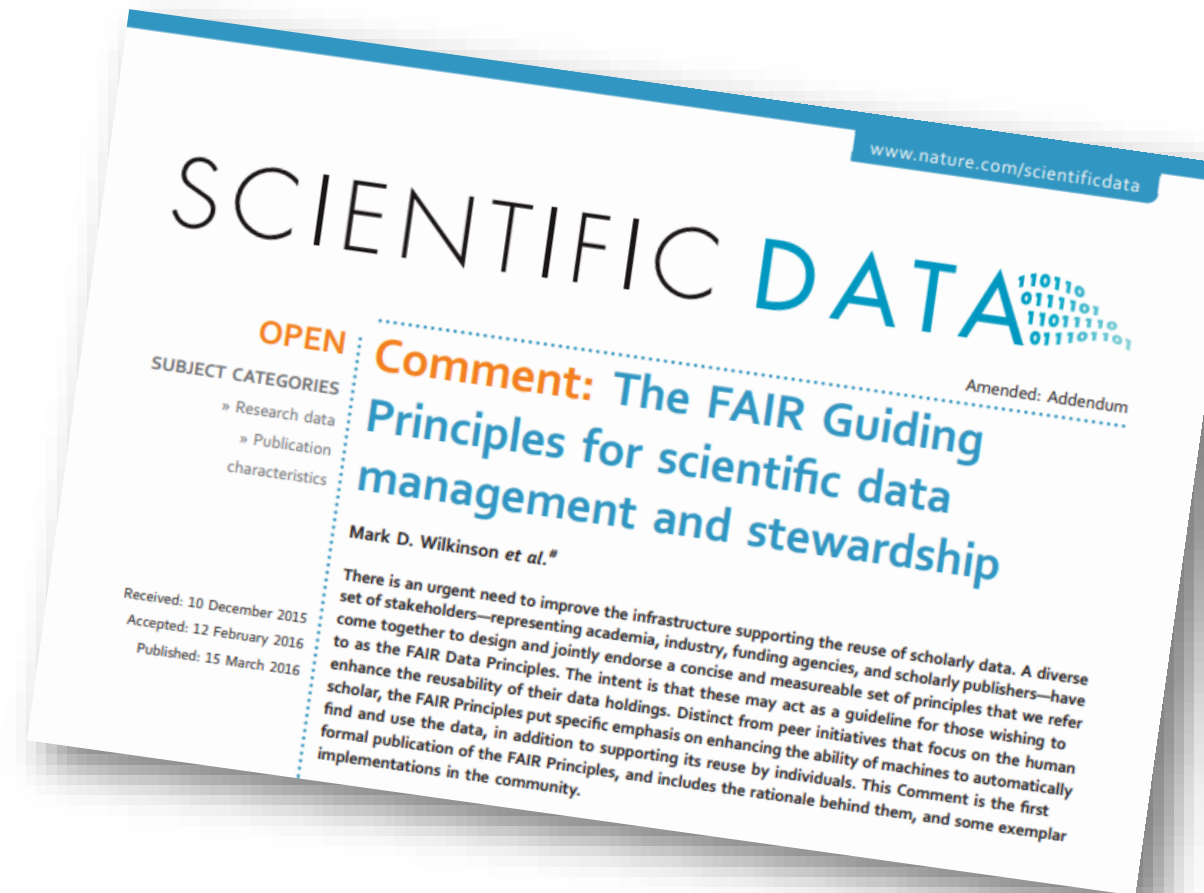
CL**AAU**DIA
RESEARCH DATA SERVICES

AALBORG UNIVERSITY
DENMARK

Findable **A**ccessible **I**nteroperable **R**eusable
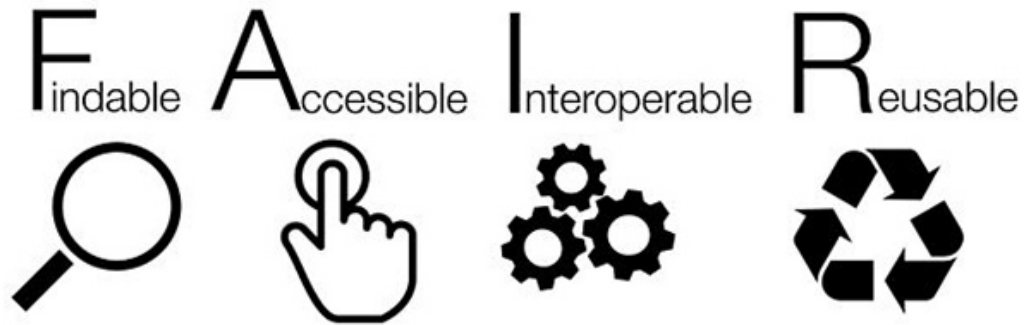
- Not only about public data

- Not only about restricted data

- Not a quality standard

- Applied different across domains

- A way to start – not a last minute "conversion"

www.nature.com/scientificdata

SCIENTIFIC DATA

OPEN

SUBJECT CATEGORIES

» Research data
» Publication characteristics

Amended: Addendum

Comment: The FAIR Guiding Principles for scientific data management and stewardship

Mark D. Wilkinson et al.#

Received: 10 December 2015
Accepted: 12 February 2016
Published: 15 March 2016

There is an urgent need to improve the infrastructure supporting the reuse of scholarly data. A diverse set of stakeholders—representing academia, industry, funding agencies, and scholarly publishers—have come together to design and jointly endorse a concise and measureable set of principles that we refer to as the FAIR Data Principles. The intent is that these may act as a guideline for those wishing to enhance the reusability of their data holdings. Distinct from peer initiatives that focus on the human scholar, the FAIR Principles put specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals. This Comment is the first formal publication of the FAIR Principles, and includes the rationale behind them, and some exemplar implementations in the community.

AALBORG UNIVERSITY
DENMARK

PRINCIPLES

noun

plural noun: **principles**

1. a fundamental truth or proposition that serves as the foundation for a system of belief or behaviour or for a chain of reasoning.
   "the basic principles of justice"

   Similar: truth   proposition   concept   idea   theory   postulate   ⌄

   - a rule or belief governing one's behaviour.
     "struggling to be true to their own principles"

     Similar: morals   morality   moral standards   moral values   ethics   ⌄

   - morally correct behaviour and attitudes.
     "a man of principle"

     Similar: morals   morality   moral standards   moral values   ethics   ⌄

2. a general scientific theorem or law that has numerous special applications across a wide field.

   - a natural law forming the basis for the construction or working of a machine.
     "these machines all operate on the same general principle"

3. a fundamental source or basis of something.
   "the first principle of all things was water"

   - a fundamental quality determining the nature of something.
     "the combination of male and female principles"

   - CHEMISTRY
     an active or characteristic constituent of a substance, obtained by simple analysis or separation.
     "the active principle of Spanish fly"

Definitions from Oxford Languages

AAU POLICY
– ON FAIR

"4.3. The management of research data takes place in accordance with the intentions of the FAIR principles that aim to make research data retrievable, (F: Findable), reachable, interpretable (I: Interoperable) and possible to reuse, hence FAIR. The FAIR principles always take the assumption of data confidentiality, thinking data into the process, including methodological information etc."

https://www.handbook.aau.dk/document... 2670

AALBORG UNIVERSITY
DENMARK

IMPACT ON PRACTICE

Concept and planning	Exploring and organising	Collection and analysis	Output and publications	Impact and engagement

AALBORG UNIVERSITY
DENMARK

MAN VS MACHINE

AALBORG UNIVERSITY
DENMARK
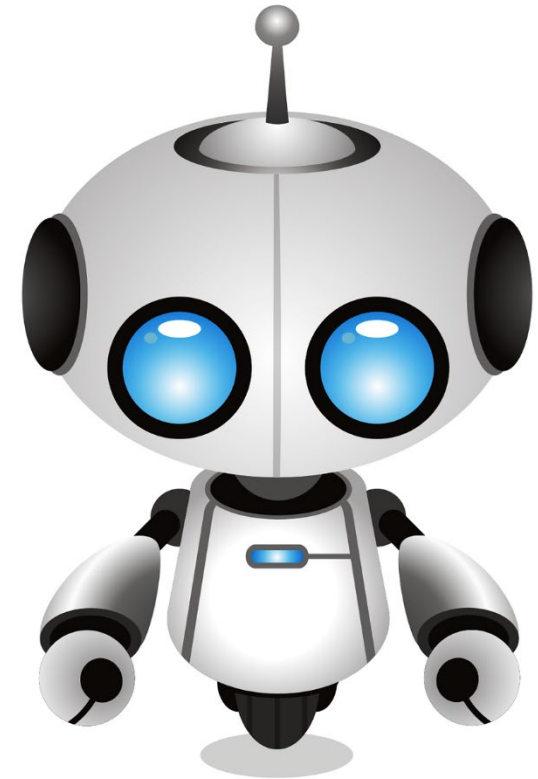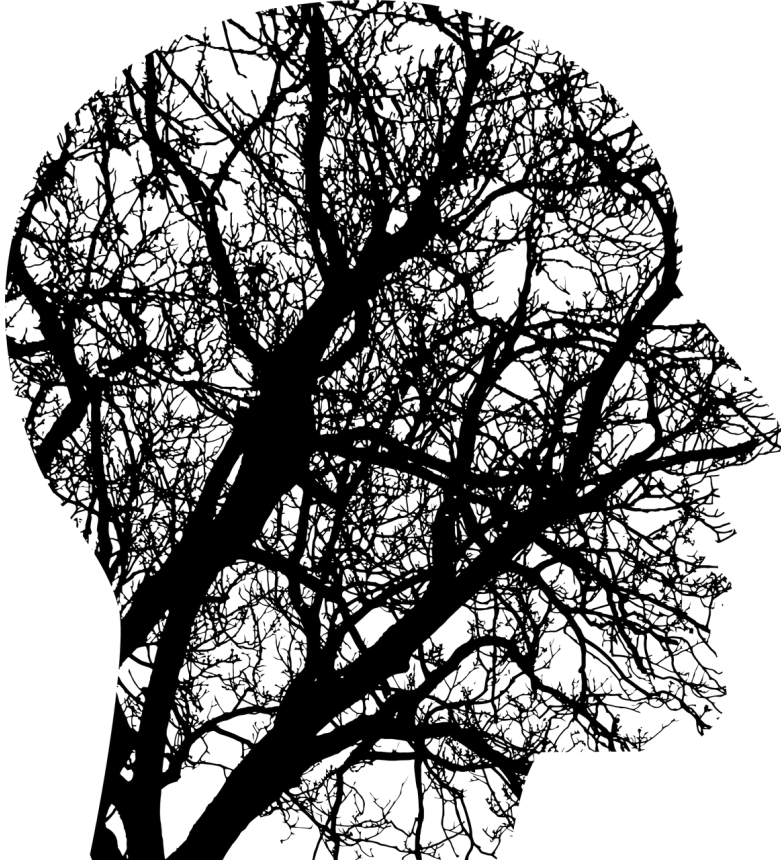
MACHINE ACTIONABLE

a)   identify the type of object (with respect to both structure and intent)

b)   determine if it is useful within the context of the agent's current task by interrogating metadata and/or data elements

c)   determine if it is usable, with respect to license, consent, or other accessibility or use constraints, and

d)   take appropriate action, in much the same manner that a human would.

Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* 3:160018 doi: 10.1038/sdata.2016.18 (2016)

AALBORG UNIVERSITY
DENMARK

F indable A ccessible I nteroperable R eusable
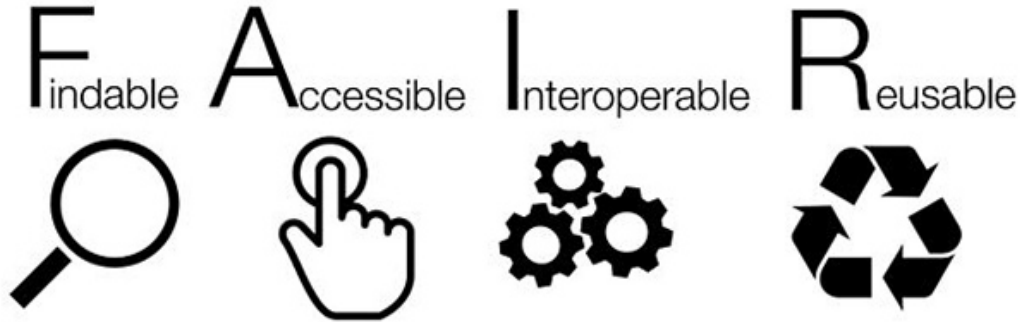
- Prerequisite: (Research) software / infrastructure / method support

- "It cannot be applied to us …."
  Might require a transition in software and/or method

AALBORG UNIVERSITY
DENMARK

THREE REPEATED TERMS
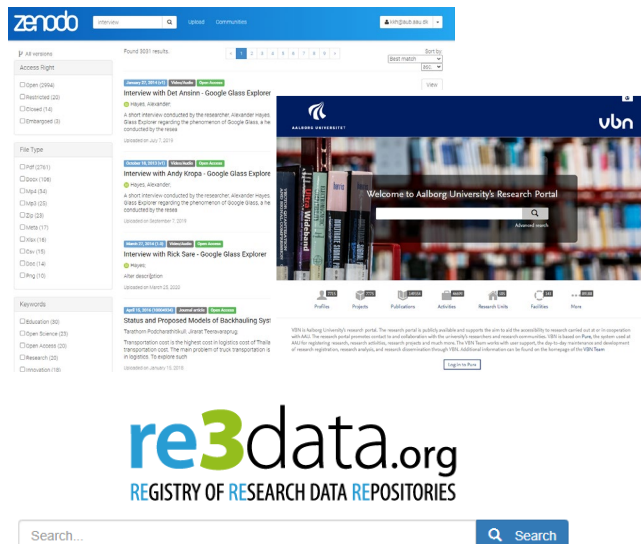
- Repository

- Standard

- (meta)data

AALBORG UNIVERSITY
DENMARK

REPOSITORY TYPES

A library
- the download analogy

Generic vs subject specific

An interface (API)
- the query analogy

```
SELECT ?item ?itemLabel 2
WHERE 3 { 4 ?item wdt:P31
wd:Q146. 5 SERVICE
wikibase:label {
bd:serviceParam
wikibase:language
"[AUTO_LANGUAGE],en". } 6 }
```


A black box service
- the visiting analogy
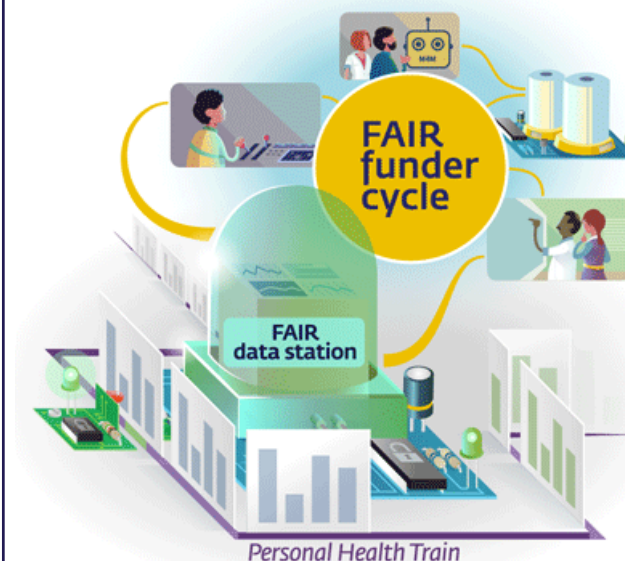
FAIR Data Point

AALBORG UNIVERSITY
DENMARK

METADATA – AND STANDARDS

METADATA

- Descriptive
- Structural
- Administrative
- Intrinsic
- User generated

METADATA

Create → Analyze → Publish

- Descriptive
- Structural
- Administrative
- Intrinsic
- User generated

METADATA

- ❯ Descriptive

- ❯ Structural

- ❯ Administrative

- ❯ Intrinsic

- ❯ User generated

AALBORG UNIVERSITY
DENMARK

METADATA

- Descriptive
- Structural
- Administrative
- Intrinsic
- User generated

- Good
- Bad
- Ugly

AALBORG UNIVERSITY
DENMARK

METADATA

- Descriptive
- Structural
- Administrative
- Intrinsic
- User generated

STANDARDS (META)DATA

❯ Agreed representation of data/information

❯ Allows for (some) common interpretation

❯ https://fairsharing.org/standards/?q=&selected_facets=status:Ready&selected_facets=recommended:true

AALBORG UNIVERSITY
DENMARK

STANDARDS
- FILE FORMATS

❯ Time to live

❯ Purpose

❯ Requirements

Balance
- Open or proprietary
- Software compliance
- Capabilities
- Standards and documentation

AALBORG UNIVERSITY
DENMARK

THREE REPEATED TERMS

- Repository
- (meta)data
- Standard

AALBORG UNIVERSITY
DENMARK
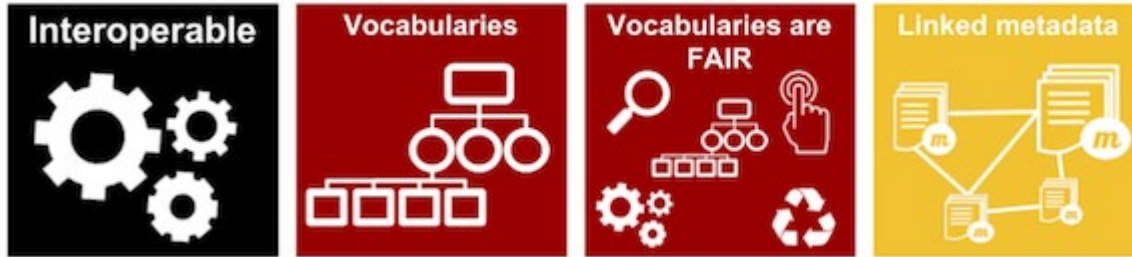
❯ I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

❯ I2. (Meta)data use vocabularies that follow FAIR principles

❯ I3. (Meta)data include qualified references to other (meta)data

https://asistdl.onlinelibrary.wiley.com/doi/epdf/10.1002/bult.2013.1720390211

AALBORG UNIVERSITY
DENMARK
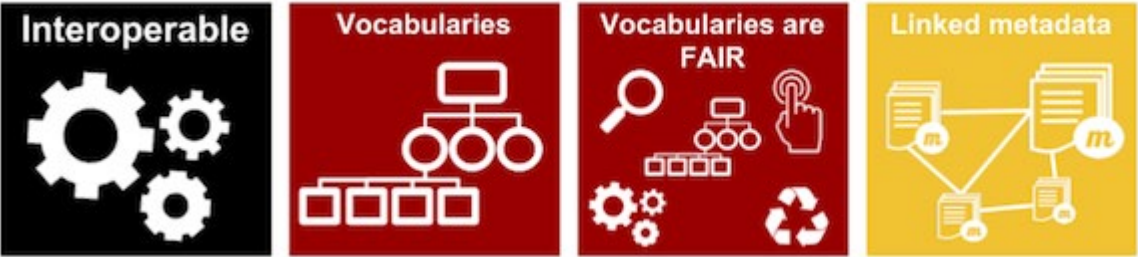
Language is (too) ambiguous

Open file formats, but …

Structure of controlled vocabularies

❯ https://github.com/tensorflow/tensorflow/blob/r2.0/tensorflow/core/example/example.proto
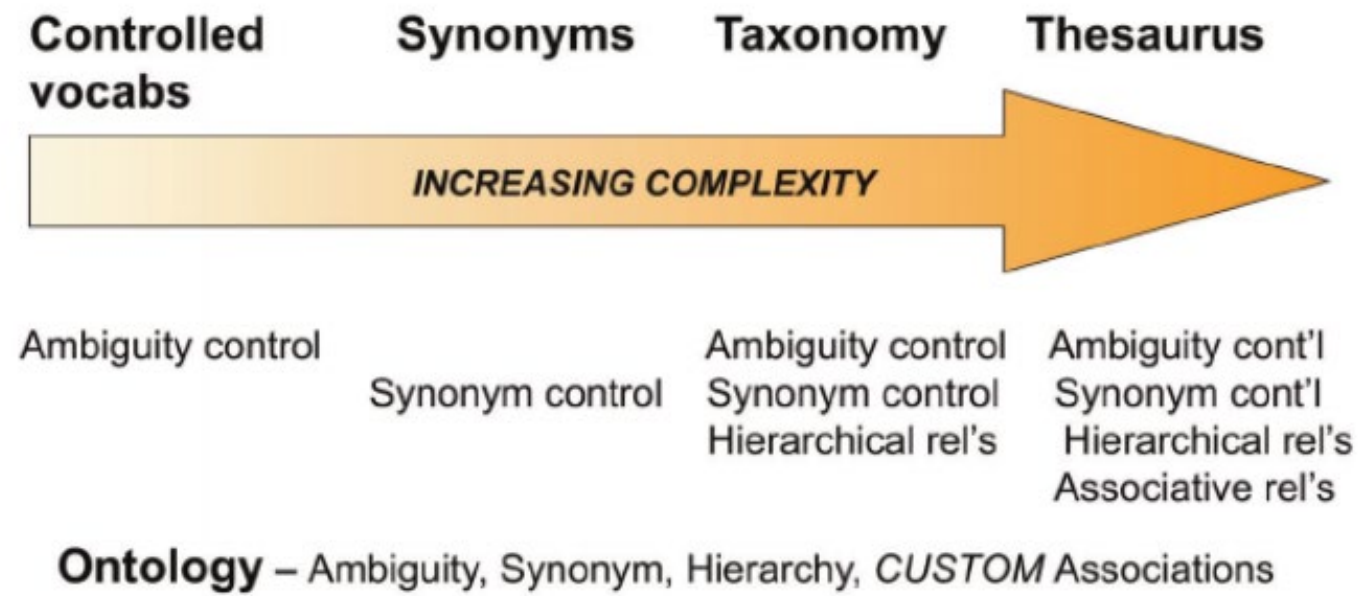
An Example for a movie recommendation application:
```
//   features {
//     feature {
//       key: "age"
//       value { float_list {
//         value: 29.0
//       }}
//     }
//     feature {
//       key: "movie"
//       value { bytes_list {
//         value: "The Shawshank Redemption"
//         value: "Fight Club"
//       }}
```

AALBORG UNIVERSITY
DENMARK

movie identifiers

About 51.400.000 results (0,82 seconds)

EIDR | EIDR – A universal unique identifier for movie and ...
https://eidr.org ▾

Media & Entertainment Services Alliance (MESA) advances new technologies that are enabling the digital transformation of media & entertainment.

EIDR – A universal unique identifier for movie and ... - EIDR
https://eidr.org › page ▾

EIDR is a universal unique **identifier** system for **movie** and television assets. From top level franchises, titles, edits, and collections, to series, seasons, episodes, ...

EIDR - Wikipedia
https://en.wikipedia.org › wiki › EIDR ▾

The identification system resolves an **identifier** to a metadata record that is associated with top-level titles, edits, DVDs, encodings, clips, and mash-ups. EIDR also provides **identifiers** for Video Service providers, such as broadcast and cable networks.

History · Content Model · Use in Standards ... · Architecture

[PDF] a short guide to identifiers for the digital supply ... - MovieLabs
https://movielabs.com › papers › A Short Guide to Identifiers for the Digita... ▾

Dec 13, 2017 - 9328-2" is the unique number assigned to the content in the registry.†. In the context of a **film** Avail, EIDR **IDs** are typically used for the following ...

EIDR

More images

EIDR, or the Entertainment Identifier Registry, is a global unique identifier system for a broad array of audio visual objects, including motion pictures, television, and radio programs. Wikipedia

Founded: 2010

Headquarters: Redwood City, California, United States

Executive Director: Will Kreth

Type of business: 501(c) organization

Profiles

LinkedIn   Twitter

Feedback

EIDR
Entertainment Identifier Registry

Home   Search   Register   Help

Content ID, Alternate ID, Party ID   LOOKUP

View XML

Content Results   Refine Content Search   Parties   Video Services

8 Results.   Page size: 8

Content ID	Title	Title Lang.	Ref. Type	Relationships	Original Lang.	Lang. Mode.	Structural Type	Release Date
10.5240/8EF1-8356-2873-6C7D-CBD2-7	Hope Springs Eternal: A Look Back at 'The Shawshank Redemption'	en	Short		en	Audio	Abstraction	2004
10.5240/8B4D-292D-5B4C-6ECA-7CFC-3	The Shawshank Redemption	en	Movie	isEditOf (10.5240/31CA-F962-3A40-18BD-BFA5-7)	en	Audio	Performance	1994-09-23
10.5240/7682-60D6-ADC5-ECB5-72B6-C	The Shawshank Redemption	en	Movie	isManifestationOf (10.5240/982A-B71D-8644-1160-7E63-0)	en	Audio	Digital	1994
10.5240/B929-CBA1-182B-901A-F0B4-N	The Shawshank Redemption	en	Movie	isManifestationOf (10.5240/982A-B71D-8644-1160-7E63-0)	en	Audio	Performance	1994
10.5240/982A-B71D-8644-1160-7E63-0	The Shawshank Redemption	en	Movie	isEditOf (10.5240/31CA-F962-3A40-18BD-BFA5-7)	en	Audio	Performance	1994
10.5240/31CA-F962-3A40-18BD-BFA5-7	The Shawshank Redemption	en	Movie		en	Audio	Abstraction	1994
10.5240/ECD3-E2D2-D597-6389-080F-H	The Shawshank Redemption/ Friendship Is Tragic	en	TV	isEpisodeOf (10.5240/18E5-2D94-15C5-6594-7037-V)	en	Audio	Abstraction	2018-12-31
10.5240/EA25-3249-B83F-35DA-975E-T	The Shawshank Redemption/ Friendship Is Tragic	en	TV	isEditOf (10.5240/ECD3-E2D2-D597-6389-080F-H)	en	Audio	Performance	2018-12-31

8 Results.   Page size: 8
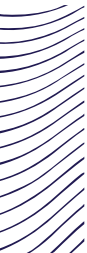
© 2019 Entertainment ID Registry Association   **doi** International DOI Foundation   5f14c98   resolve.eidr.org   v2.1

```
//   features {
//     feature {
//       key: "age"
//       value { float_list {
//         value: 29.0
//       }}
//     }
//     feature {
//       key: "movie"
//       value { bytes_list {
//         value: "EIDR:10.5240/982A-B71D…"
//         value: "Fight Club"
//       }}
```


EIDR — Entertainment Identifier Registry

Home   Search   Register

Content ID, Alternate ID, Pa

VIEW

Metadata   Relationships

BASE OBJECT DATA

EIDR ID	10.5240/982A-B71D-8644-1160-7E63-0		
Structural Type	Performance		
Mode	*AudioVisual*		
Referent Type	*Movie*		
Title	***The Shawshank Redemption***		
	Lang: **en** *Title Class:* **release**		
Original Language	**en**		
	Mode: **Audio** *Type:* **primary**		
Associated Org	Warner Bros		
	ID Type: EIDRPartyID Party ID: 10.5237/A929-C667 Role: producer		
Alternate Name	WB		
Release Date	1994		
Country of Origin	**US**		
Status	valid		
Approximate Length	PT2H22M0S		
Alternate ID	-v5myc1Pba4		
	Domain: play.google.com Relation: IsSourceOf Type: Proprietary		

- R1. Meta(data) are richly described with a plurality of accurate and relevant attributes

- R1.1. (Meta)data are released with a clear and accessible data usage license

- R1.2. (Meta)data are associated with detailed provenance

- R1.3. (Meta)data meet domain-relevant community standards

AALBORG UNIVERSITY
DENMARK

Reusable · **Metadata have multiple attributes** · **Usage license** · **Provenance** · **Community standards**

Publication date:
August 12, 2020

DOI:
DOI 10.5281/zenodo.3980491

Communities:
OpenAIRE
OpenAIRE Research Graph
Zenodo

License (for files):
Creative Commons Zero v1.0 Universal

Code

1,558 commits

2 days ago
18 hours ago
2 days ago
21 days ago
13 hours ago
2 days ago
13 hours ago
18 hours ago
8 days ago
16 hours ago

About

Project Connected Home over IP is a new Working Group within the Zigbee Alliance. This Working Group plans to develop and promote the adoption of a new connectivity standard to increase compatibility among smart home products, with security as a fundamental design tenet.

www.connectedhomeip.com

zigbee · standard · internet-of-things
iot · chip · connectedhomeip
connected-devices · connected-home

Readme

Apache-2.0 License

- ❯ F1. (Meta)data are assigned a globally unique and persistent identifier

- ❯ F2. Data are described with rich metadata (defined by R1 below)

- ❯ F3. Metadata clearly and explicitly include the identifier of the data they describe

- ❯ F4. (Meta)data are registered or indexed in a searchable resource

AALBORG UNIVERSITY
DENMARK
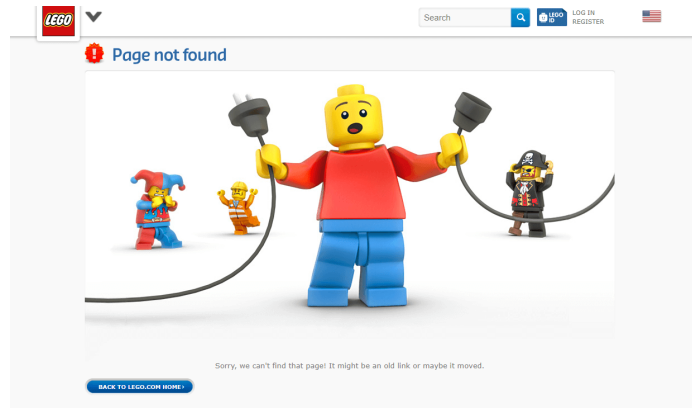
DOI example:
    10.5281/zenodo.2248200 =>
    https://doi.org/10.5281/zenodo.2248200
GRID example:
    grid.5117.2 => https://grid.ac/institutes/grid.5117.2

AALBORG UNIVERSITY
DENMARK

Findable	Persistent Identifiers (PIDs)	Rich metadata	Indexed data repositories	PIDs in metadata
Accessible	Standard communications protocol	Open, free protocol	Authentication, where necessary	Metadata is always available
Interoperable	Vocabularies	Vocabularies are FAIR	Linked metadata	
Reusable	Metadata have multiple attributes	Usage license	Provenance	Community standards

- A1. (Meta)data are retrievable by their identifier using a standardised communications protocol

- A1.1 The protocol is open, free, and universally implementable

- A1.2 The protocol allows for an authentication and authorisation procedure, where necessary

- A2. Metadata are accessible, even when the data are no longer available

AALBORG UNIVERSITY
DENMARK

Accessible

Standard communications protocol

Open, free protocol

Authentication, where necessary

Metadata is always available

Files (52.8 MB)

Name	Size	
COVID-19.json.gz	52.8 MB	⬇ Download

md5:c7ab7c8b97c2281146ea643cc5df6d90 ❓

University of Bristol Research Data Request

What will happen to the information which you provide?
We will use the information you provide to fulfil your data access request following the process outlined below. Your personal information will be processed in order to fulfil this contract. We will keep your information securely and will only share it beyond the University of Bristol if required to do so by law. We will retain your information for six years or, should your request be successful, for the duration of the Data Access Agreement, whichever is longer. After this time your information will be deleted.
For more information on how we use your data and to request a copy of the information you provide in this form please email the Research Data Service at data-bris@bristol.ac.uk. If you wish to make a complaint about how your personal information is being processed by the Research Data Service, contact data-protection@bristol.ac.uk.

How do we process applications for 'Restricted' data?
The Research Data Service review the application, and will check that the information you have provided is complete and in adherence with the following criteria:
- The dataset you have requested is available for sharing
- You have provided a verifiable institutional affiliation
- You have provided verifiable institutional contact details
- You have nominated an appropriate institutional signatory (see paragraph below)
- You have ethical approval in place (this may not be required depending on the nature of the requested dataset)
- If your research is funded, commissioned or sponsored, you should provide documentary evidence of this
If your application meets the criteria, a Data Access Agreement (example Data Access Agreement for a restricted dataset - http://bit.ly/2ITs3T3) is sent to your nominated institutional signatory. A suitable signatory may be from your contracts department, legal department or research office, depending on the nature of your institution. Once the agreement is signed and returned, the Research Data Service team will arrange secure access to the dataset.

How do we process applications for 'Controlled' data?
The assessment of requests for 'Controlled' data access is very similar to that described above, with an additional stage. The Research Data Service will perform the same initial checks as outlined under and if your application meets the criteria given above, the Research Data Service team will forward your request to the creator of the data (or a colleague nominated by them) to check that your planned research can be supported by the dataset. If it cannot, you may wish to amend your planned research. This stage can take some time and should be seen as a conversation between an

AALBORG UNIVERSITY
DENMARK

Findable	Persistent Identifiers (PIDs)	Rich metadata	Indexed data repositories	PIDs in metadata
Accessible	Standard communications protocol	Open, free protocol	Authentication, where necessary	Metadata is always available
Interoperable	Vocabularies	Vocabularies are FAIR	Linked metadata	
Reusable	Metadata have multiple attributes	Usage license	Provenance	Community standards

AALBORG UNIVERSITY
DENMARK

Highest reuse value?

Restricted access   Open access

Sensitive · Sensitive · Sensitive · Sensitive · Sensitive · Sensitive · Anonymous data

AALBORG UNIVERSITY
DENMARK

REASONS NOT TO PUBLISH

❯ Confidentiality   *adhere*

❯ Patent filing   *wait*

❯ Sensitive data   *alter or restrict*

❯ Licensed material   *seperate*
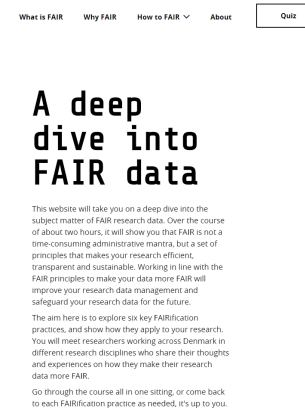
❯ Uncertainty   *ask*

FAIR is still possible

STEPWISE APPROACH TOWARDS MORE FAIR'ness

❯ **Investigate** - find any requirements, expectations, good practice …

❯ **Invest** - any changes needed in your current setup

❯ **Implement** - go do FAIR

❯ **Iterate** - follow up as reality overtakes the plan

AALBORG UNIVERSITY
DENMARK

A FAIRy tale

A FAKE STORY IN A TRUSTWORTHY GUIDE
TO THE FAIR PRINCIPLES FOR RESEARCH DATA.

10.5281/zenodo.2248200

HOW TO FAIR

What is FAIR   Why FAIR   How to FAIR ⌄   About   Quiz

A deep dive into FAIR data

This website will take you on a deep dive into the subject matter of FAIR research data. Over the course of about two hours, it will show you that FAIR is not a time-consuming administrative mantra, but a set of principles that makes your research efficient, transparent and sustainable. Working in line with the FAIR principles to make your data more FAIR will improve your research data management and safeguard your research data for the future.

The aim here is to explore six key FAIRification practices, and show how they apply to your research. You will meet researchers working across Denmark in different research disciplines who share their thoughts and experiences on how they make their research data more FAIR.

Go through the course all in one sitting, or come back to each FAIRification practice as needed, it's up to you.

howtofair.dk

research domain fair principles

Google Search   I'm Feeling Lucky

AALBORG UNIVERSITY
DENMARK