EERAdata

TOWARDS A FAIR AND OPEN DATA ECOSYSTEM IN THE LOW-CARBON ENERGY RESEARCH COMMUNITY

What does machine-actionability mean?

An API perspective on the EERAdata platform

Overview

- Machine-Actionability
- Human to Machine POV
- API
- Conclusion

What is machine actionability?

- Machine actionable data "is structured in a consistent way so that machines, or computers, can be programmed against the structure." (https://ddialliance.org/taxonomy/term/198)
- More loosely, with machine actionable data, you can write a program to use the data without having to examine the individual values.
- FAIR needs machine actionability, but machine actionability isn't enough on its own to make data FAIR

Example: German coastal states

- Task: determine which German states have a coastal boundary and which do not
- Obtained some relevant data from Eurostat: https://ec.europa.eu/eurostat/statistics-explained/images/6/62/Focus_on_coastal_regions_RYB2012.xls

Eurostat data

NUTS	Region name	Value	Label
BE211	Arr. Antwerpen	1	North Sea
BE212	Arr. Mechelen	1	North Sea
BE231	Arr. Aalst	1	North Sea
BE232	Arr. Dendermonde	1	North Sea
BE233	Arr. Eeklo	1	North Sea
BE234	Arr. Gent	1	North Sea
BE236	Arr. Sint-Niklaas	1	North Sea
BE251	Arr. Brugge	1	North Sea
BE252	Arr. Diksmuide	1	North Sea
BE253	Arr. Ieper	1	North Sea
BE255	Arr. Oostende	1	North Sea
BE256	Arr. Roeselare	1	North Sea
BE257	Arr. Tielt	1	North Sea
BE258	Arr. Veurne	1	North Sea
BG331	Varna	2	Black Sea
BG332	Dobrich	2	Black Sea
BG341	Burgas	2	Black Sea
DK011	Byen København	4	Baltic Sea
DK012	Københavns omegn	4	Baltic Sea
DK013	Nordsjælland	4	Baltic Sea
DK014	Bornholm	4	Baltic Sea
DK021	Østsjælland	4	Baltic Sea
DK022	Vest- og Sydsjælland	4	Baltic Sea
DK031	Fyn	4	Baltic Sea
DK032	Sydjylland	1	North Sea
DK041	Vestjylland	1	North Sea
DK042	Østjylland	1	North Sea
DK050	Nordjylland	1	North Sea

Resolving the task

- Extract coastal NUTS3 regions for Germany
- Convert to NUTS1, corresponding to German states
- Use the list to partition German states by whether they are coastal or not
- Result: Bremen, Hamburg, Mecklenburg-Vorpommern, Lower Saxony, and Schleswig-Holstein are coastal; all other states are not

Challenges for a program

- Data underspecified: which NUTS revision?
- Need to augment the data: relation between NUTS statistical regions and German states, relation between different NUTS levels
- Need to provide our computer program with a formal representation of how the relevant knowledge is structured: we can do more when data sets are accompanied by ontologies and semantic models

Overview

- Machine-Actionability
- Human to Machine POV
- API
- Conclusion

FAIR Machine Actionability

"Capacity of a computational system to find, access, interoperate, and reuse data with none or minimal human intervention" (https://www.go-fair.org/fair-principles/)

Human to Machine PoV

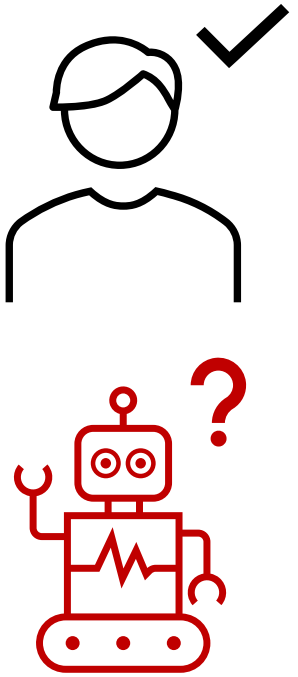
- Find, Access, Interoperate, Reuse
- Example

✓ Found

✓ Access

✓ Interoperate

✓ Reuse

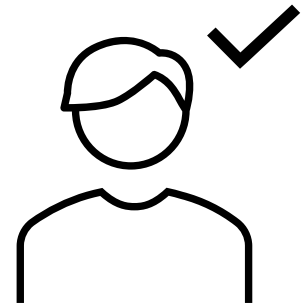


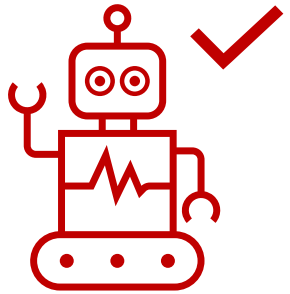
Human to Machine PoV

- Characters painted or drawn on a space
- Copy table into document
- Machine reads line by line, symbol by symbol
- Missing structure

"The representation of the data on disk does not represent the actual relationships present in the data."
(https://opendatahandbook.org/glossary/en/terms/human-readable/)

Human to Machine PoV





shutterstock.com · 1060912469

Principles

- Structured data
  - Structure within file
  - Structure within system (stored on disk)
- Standard
  - Many types of structure allowed
  - XML, JSON,...
- Semantically unique and shared among computer systems
- No PDF, plain text, Word, images, audio, video...

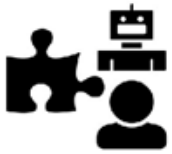
Principles

Miksa, Simm et al.
"Ten principles for machine-actionable data management plans"

1 Integrate DMPs with the workflows of all stakeholders in the research data ecosystem

2 Allow automated systems to act on behalf of stakeholders

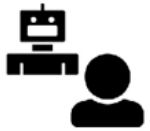
3 Make policies (also) for machines, not just for people



4 Describe—for both machines and humans—the components of the data management ecosystem

5 Use PIDs and controlled vocabularies

6 Follow a common data model for maDMPs



7 Make DMPs available for human and machine consumption

8 Support data management evaluation and monitoring

9 Make DMPs updatable, living, versioned documents

10 Make DMPs publicly available

Miksa T, Simms S, Mietchen D, Jones S (2019) Ten principles for machine-actionable data management plans. PLoS Comput Biol 15(3): e1006750. https://doi.org/10.1371/journal.pcbi.1006750

Overview

- Machine-Actionability
- Human to Machine POV
- API
- Conclusion

API

- Application Programming Interface
- Software interface
- Offers a service to other pieces of software
- Standard that defines how to use it

API

1) API receives a request

API

1) API receives a request

2) API collects and processes a reponse, then returns with that response

API

- Practiced in Workshops
- Automated Access
- Identification
- Followed by Implementation

Overview

- Machine-Actionability
- Human to Machine POV
- API
- Conclusion

Conclusion

- Afternoon Session
- UC Discussion
- At what points can platform help with machine-actionability?
- Integration

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 883823.

THANK YOU!

EERA data

TOWARDS A FAIR AND OPEN DATA ECOSYSTEM IN THE LOW-CARBON ENERGY RESEARCH COMMUNITY

Manfred Paier, Michael Barber, Astrid Unger
Center for Innovation Systems & Policy

AIT Austrian Institute of Technology GmbH
Giefinggasse 4 | 1210 Vienna | Austria
T +43 (0) 50550 4559 | M +43 (0) 664 8833 5494
manfred.paier@ait.ac.at | http://www.ait.ac.at