FAIR
Semantics, Interoperabiliy and Services
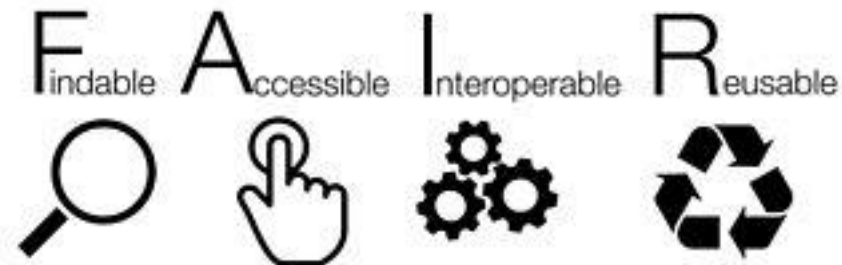
Jessica Parland-von Essen

27 October 2021

Virtual SciDataCon 2021

FAIR semantics

- Semantic artefacts are tools for creating semantic interoperability

- Semantic artefacts should be FAIR
  - Identifiers
  - Metadata
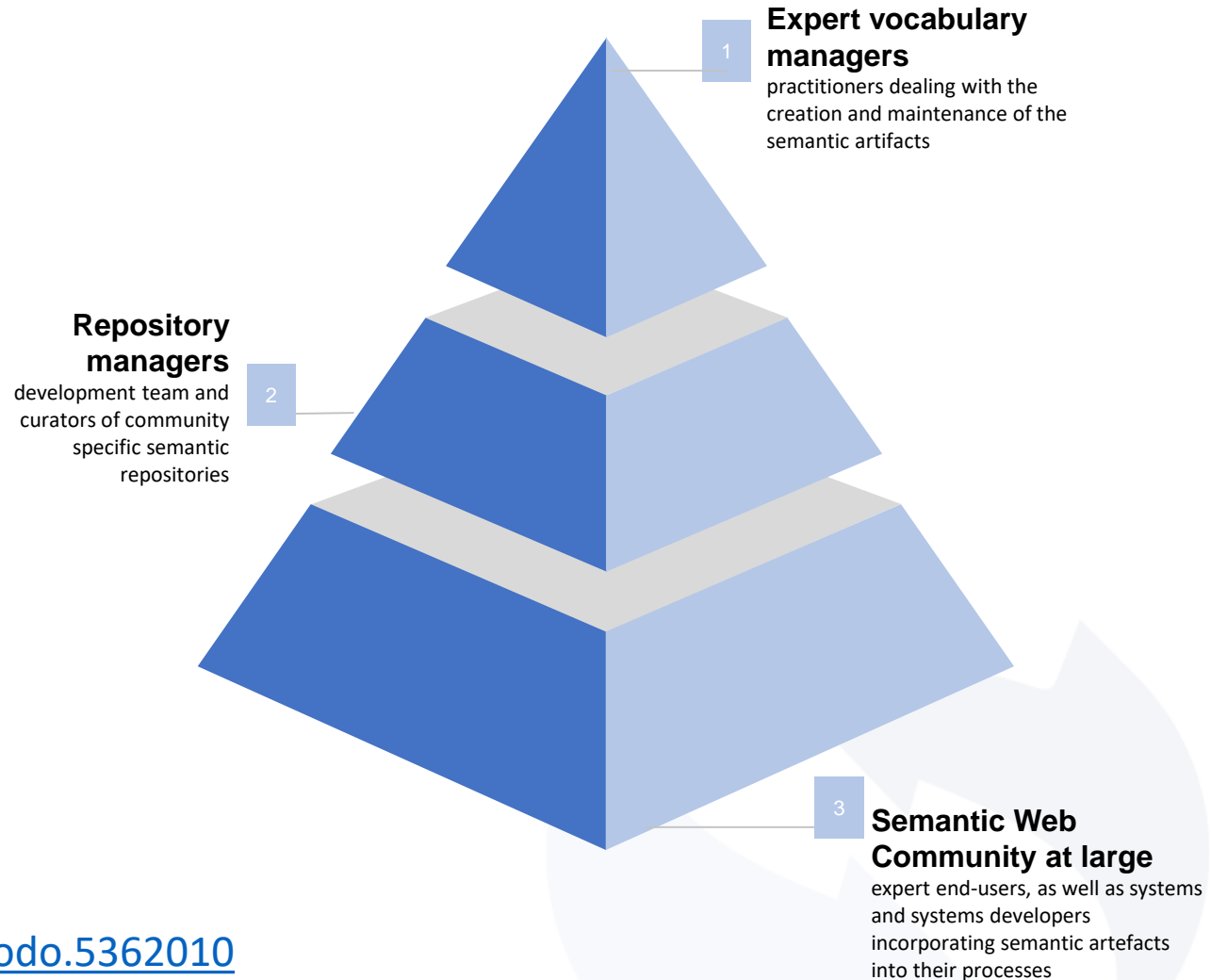 - Interoperability
  - Machine actionability

Findable Accessible Interoperable Reusable

FAIR semantic artefacts

17 recommendations

13 best practices

https://doi.org/10.5281/zenodo.5362010

1 **Expert vocabulary managers**
practitioners dealing with the creation and maintenance of the semantic artifacts

2 **Repository managers**
development team and curators of community specific semantic repositories

3 **Semantic Web Community at large**
expert end-users, as well as systems and systems developers incorporating semantic artefacts into their processes

Recommendations for FAIR semantics

- P-Rec. 1: Globally Unique, Persistent and Resolvable Identifiers must be used for Semantic Artefacts, their content (terms/ concepts/ classes and relations), and their versions

- P-Rec. 2: Globally Unique, Persistent, and Resolvable Identifiers must be used for Semantic Artefact Metadata Records. Metadata and data must be published separately, even if it is managed jointly

- P-Rec. 3: A common minimum metadata schema must be used to describe semantic artefacts and their content

- P-Rec. 5: Semantic repositories must offer access to Semantic Artefacts and their content using community standard APIs and serializations to support both use/ reuse and indexation by search engines

- P-Rec. 7: Repositories must offer a secure access protocol, and appropriate user access control functionalities

- P-Rec. 8: Human and machine-readable persistence policies for semantic artefacts metadata and data must be published

- P-Rec. 9: Semantic artefacts must be made available as a minimum portfolio of common serialization formats

- P-Rec. 16: The semantic artefact must be clearly licenced for use by machines and humans

- P-Rec. 17: Provenance must be clear for both humans and machines

Learn more: 2nd set of Recommendations for FAIR Semantics https://doi.org/10.5281/zenodo.4314321

Metadata for semantic artefacts

- MOD and DCAT2 offer a starting point

- DCAT extension for semantic artefacts under discussion

- SHACL and XML will be promoted

- Work done together with RDA VSSIG

WiP

RDF Property:	dct:title			
Definition:	A homepage of the catalog (a public Web document usually available in HTML).			
Range:	dcat: Resources			
Usage note:	dct:title is an inverse functional property (IFP) which means that it *MUST* be unique and precisely identify the Web-page for the resource. This property indicates the canonical Web-page, which might be helpful in cases where there is more than one Web-page about the resource.			
Voting Results	Total number of votes	Mandatory (%)	Recommended (%)	Optional (%)
	21	91.30%	4.35%	4.35%

Skärmavbild

Some best practices for FAIR semantics

- Interact with the designated community and manage user-centric development

- Provide a structured definition for each concept

- The underlying logic of semantic artefacts should be grounded in the domain it intends to be used in

- A standard architecture for semantic artefact management, services for content and metadata, protocols, and serialisations/ content negotiation should be proposed, and semantic repositories are urged to use this architecture as a design pattern

FAIR Software

- Software can be tools, research outputs or even research objects

- Software citation is an important aspect in all cases, but for the output is often most important for the researcher

- *reuse* and *interoperability*

Learn more: Assessment report on 'FAIRness of software'
https://doi.org/10.5281/zenodo.5472911

Challenges

- Software dependencies and environment - technical challenge

- Documentation

- Accessibility & Licensing

- Time and skill

- Quality control

- Software sustainability & management plan

FAIR - Aspects of assessing services

- **Technical & service provisioning aspects**
  - SAF-F FAIR enablement
  - SAF-Q Quality of service
  - SAF-O Open & Connected
- **Socially-oriented aspects**
  - SAF-U User centricity
  - SAF-T Transparency
  - SAF-L Longevity
  - SAF-E Ethical & Legal

Aspect: SAF-F FAIR enablement

- **Objective**: The service enables FAIR data by elevating the FAIRness of digital objects and/or supporting the FAIRification process. FAIR enablement is actively driven through the implementation of community-supported standards and interoperability frameworks.

- The following classification defines the levels in which a service affects the FAIRness of data on which it operates.

  - *Enable:*

    - *Augment*: The service provides elements improving FAIRness of the digital object for example automatically assigning a PID;

    - *Facilitate*: The service actively helps to realize a particular FAIR principle — for example by allowing the user to add metadata or enabling discoverability;

  - *Respect*: The service neither actively enables a particular FAIR principle nor interferes with it — it can be said to respect the "FAIR-in-FAIR-out" principle;

  - *Reduce*: The service actually makes data less FAIR — at least for a particular principle — for example by detaching metadata or a PID when it acts on a digital object.

F indable A ccessible I nteroperable R eusable
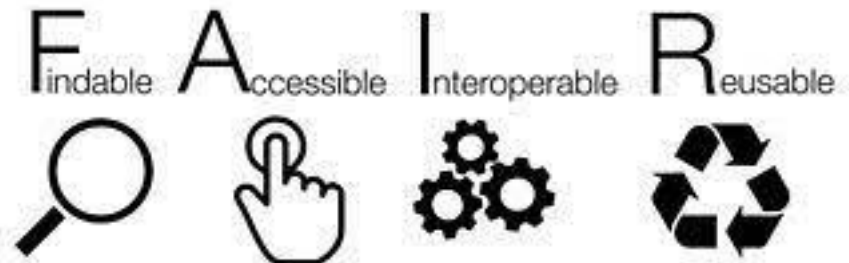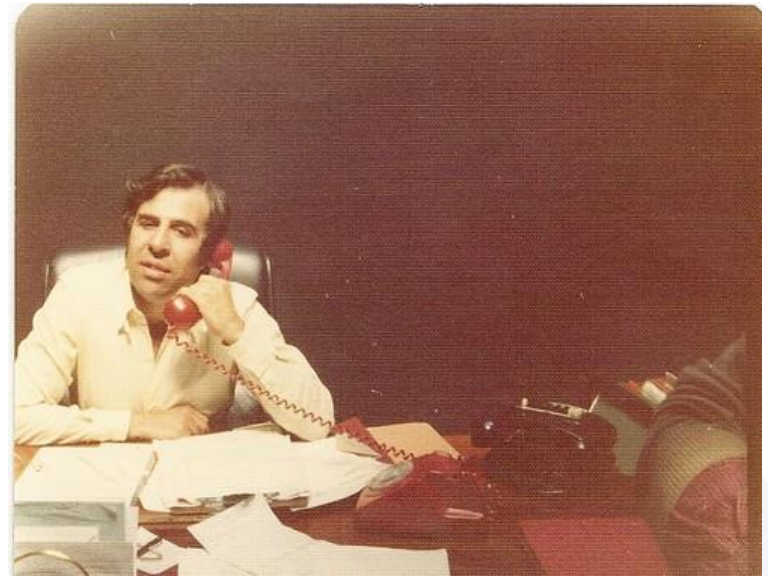
- 6 recommendations
  - 3 essential
  - 1 important
  - 1 useful

Aspect: SAF-Q Quality of service

Objective

The service is delivered in a reliable, secure, high-quality way, consistent with its specifications.

- 11 recommendations
  - 3 essential
  - 5 important
  - 3 useful

Aspect: SAF-O Open & Connected

Objective

The service is operated in a low-barrier and inclusive way, seeking integrations and connections with other services and championing principles of openness consistent with Open Science and Open Research.

■ **8 recommendations**
  ■ 4 essential
  ■ 4 important

SOCIAL

Aspect: SAF-U User centricity

Objective

The service is managed so that it serves the (possibly evolving) goals of the user community and maximises usability while minimizing burden.

A BARBERS SHOP

- **8 recommendations**
  - 4 essential
  - 2 important
  - 2 useful

Henry William Bunbury - Wellcome Collection, United Kingdom - CC BY.
https://www.europeana.eu/fi/item/9200579/tgzcvkc8

Aspect: SAF-T Transparency

SOCIAL

- **6 recommendations**
  - 2 essential
  - 2 important
  - 2 useful

Objective

The service provider communicates with its stakeholders in a transparent manner.

Aspect: SAF-L Longevity

SOCIAL

Objective

The service provider designs the service with a timeframe for the maintenance and sustainability of the service in mind and implements measures accordingly, considering the researchers' need for reproducible research.

- 4 recommendations
  - 2 essential
  - 2 important

SOCIAL

Aspect: SAF-E Ethical and Legal

Objective

The service complies with all applicable legal and ethical guidelines, in a transparent and auditable way.

◆ **6 recommendations**
  ◆ 3 essential
  ◆ 2 important
  ◆ 1 useful

Konstaapeli Matti Berg - Finnish Heritage Agency, Finland - CC BY.
https://www.europeana.eu/fi/item/2021009/_b93a5e7b_61f9_4347_8eb9_
cb93c83d4fc5

LEARN MORE: FAIRsFAIR: D2.7 Framework for assessing FAIR Services
https://doi.org/10.5281/zenodo.5336234

Resilient and FAIR data - Researchers

Aim at consistent use of PIDs, supported by good metadata. This will enable findability of your research outputs and make research data management easier in the long run.

In order to increase interoperability, try to imagine making sense of your data in 10 years: is there any "implicit" knowledge, e.g. about data types, that can be made explicit?

Before creating datasets, plan the use of PIDs, data formats, and metadata with help and guidance from your data stewards. In other words, don't try to do this alone - be sure to contact your data stewards for support in developing sustainable PIDs and metadata, thus increasing FAIRness.

Resilient and FAIR data – Data Stewards

It is important to think about interoperability and longevity.

Support researchers in determining the appropriate depth of FAIRness of the data and provide them with examples of what good FAIR PIDs and metadata should look like.

Educate researchers about reproducibility and semantic artefacts.

Think with the researchers about the understandability of the data in 10 years; making assumptions on explicit data documentation will really help interoperability.

Engaging researchers in related discussions, and decision-making processes would help to improve FAIRness from the grass roots.

There are many (good) solutions for developing FAIRness. We encourage data stewards to use existing services, instead of trying to implement their own services from scratch.

Resilient and FAIR data – Service Providers

Researchers and data stewards alike need the service providers' support in making appropriate use of the solutions for implementing FAIR.

Consider each need and use case by evaluating the FAIR principles and assessing the value of implementing them.

We recommend conducting a cost-benefit analysis on each principle with a sustainability perspective: what can be managed and curated over time?

Strive to support scientific reproducibility and data lifecycle management with well documented technologies, well managed services and workflows, and curated data.

27/10/2021

Learn more

D2.7 Framework for assessing FAIR Services
https://doi.org/10.5281/zenodo.5336234

D2.5 2nd set of Recommendations for FAIR Semantics
https://doi.org/10.5281/zenodo.4314321

M15 Assessment report on FAIRness of software
https://doi.org/10.5281/zenodo.5472911

D2.4 Report on FAIR requirements for persistence and interoperability
https://doi.org/10.5281/zenodo.5356517

27/10/2021