GEDE Webinar on Maturity Indicators for FAIRness and Certification of Repositories

Generation 1 FAIR Maturity Indicators (formerly the FAIR Metrics)

Erik Schultes, PhD
International Science Coordinator
GO FAIR International Support and Coordination Office
erik.schultes@go-fair.org
go-fair.org

March 22, 2019
What is FAIR?

“What data and services that are findable, accessible, interoperable, re-usable both for machines and for people.”
What is FAIR?

“Data and services that are findable, accessible, interoperable, re-usable both for machines and for people.”
**FAIR Principles for machine-actionable F, A, I, & R**


---

**Findable:**

F1 (meta)data are assigned a globally unique and persistent identifier;

F2 data are described with rich metadata;

F3 metadata clearly and explicitly include the identifier of the data it describes;

F4 (meta)data are registered or indexed in a searchable resource;

---

**Accessible:**

A1 (meta)data are retrievable by their identifier using a standardized communications protocol;

  - A1.1 the protocol is open, free, and universally implementable;
  - A1.2 the protocol allows for an authentication and authorization procedure, where necessary;

A2 metadata are accessible, even when the data are no longer available;

---

**Interoperable:**

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;

---

**Reusable:**

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;
Objective targets to aim for…

FAIR

Maturity Indicators
Objective targets to aim for…

FAIR

Maturity Indicators

Repos

Core Trust Seal
**Findable:**

F1 (meta)data are assigned a globally unique and persistent identifier;

F2 data are described with rich metadata;

F3 metadata clearly and explicitly include the identifier of the data it describes;

F4 (meta)data are registered or indexed in a searchable resource;

**Accessible:**

A1 (meta)data are retrievable by their identifier using a standardized communications protocol;

A1.1 the protocol is open, free, and universally implementable;

A1.2 the protocol allows for an authentication and authorization procedure, where necessary;

A2 metadata are accessible, even when the data are no longer available;

**Interoperable:**

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;

**Reusable:**

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;
Comment: A design framework and exemplar metrics for FAIRness

Mark D. Wilkinson¹, Susanna-Assunta Sansone², Erik Schultes³, Peter Doom⁴, Luiz Olavo Bonino da Silva Santos⁵,⁶ & Michel Dumontier⁷

The FAIR Principles¹ (https://doi.org/10.25504/FAIRsharing.WW110U) provide guidelines for the publication of digital resources such as datasets, code, workflows, and research objects, in a manner that makes them Findable, Accessible, Interoperable, and Reusable (FAIR). The Principles have rapidly been adopted by publishers, funders, and pan-disciplinary infrastructure programmes and societies. The Principles are aspirational, in that they do not strictly define how to achieve a state of "FAIRness", but rather they describe a continuum of features, attributes, and behaviors that will move a digital resource closer to that goal. This ambiguity has led to a wide range of interpretations of FAIRness, with some resources even claiming to already 'be FAIR'! The increasing number of such statements, the emergence of subjective and self-assessments of FAIRness²,³, and the need of data and service providers, journals, funding agencies, and regulatory bodies to qualitatively or quantitatively evaluate such claims, led us to self-assemble and establish a FAIR Metrics group (http://fairmetrics.org) to pursue the goal of defining ways to measure FAIRness.

As co-authors of the FAIR Principles and its associated manuscript, founding this small focus group was a natural and timely step for us, and we foresee group membership expanding and broadening according to the needs and enthusiasm of the various stakeholder communities. Nevertheless, in this first
FAIR Metrics

- Community defined
- Objective
- Quantifiable
- Reproducible
- Automatic (scalable)
- Certifiable
14 Core FAIR Metrics

Findable:

F1 (meta)data are assigned a globally unique and persistent identifier;

F2 data are described with rich metadata;

F3 metadata clearly and explicitly include the identifier of the data it describes;

F4 (meta)data are registered or indexed in a searchable resource;

Accessible:

A1 (meta)data are retrievable by their identifier using a standardized communications protocol;

A1.1 the protocol is open, free, and universally implementable;

A1.2 the protocol allows for an authentication and authorization procedure, where necessary;

A2 metadata are accessible, even when the data are no longer available;

Interoperable:

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;

Reusable:

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;


http://fairmetrics.org

https://github.com/FAIRMetrics/Metrics/blob/master/MaturityIndicators/Gen1/ALL.pdf
# 14 Core FAIR Metrics

<table>
<thead>
<tr>
<th>FM</th>
<th>Question</th>
<th>Dataverse</th>
<th>Dryad</th>
<th>Nanopub</th>
<th>Zenodo</th>
<th>Yale iSPS</th>
<th>Figshare</th>
<th>Broad’s SCP</th>
<th>SeaDataNet</th>
<th>CDI</th>
<th>Wikidata</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRI Exists</td>
<td>1</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
</tr>
<tr>
<td>F1A</td>
<td>2</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
</tr>
<tr>
<td>F1B</td>
<td>3</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>NRP</td>
<td>none</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
</tr>
<tr>
<td>F2A</td>
<td>4A</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>none</td>
<td>none</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
</tr>
<tr>
<td>F2B</td>
<td>4B</td>
<td>IRI</td>
<td>none</td>
<td>IRI</td>
<td>IRI</td>
<td>“Multiple”</td>
<td>none</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
</tr>
<tr>
<td>F3</td>
<td>5A</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>none</td>
<td>NRP</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
</tr>
<tr>
<td>F3</td>
<td>5B</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
</tr>
<tr>
<td>F4</td>
<td>6A</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
</tr>
<tr>
<td>F4</td>
<td>6B</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
</tr>
<tr>
<td>A1.1</td>
<td>7A</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
</tr>
<tr>
<td>A1.1</td>
<td>7B</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
</tr>
<tr>
<td>A1.1</td>
<td>7C</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
</tr>
<tr>
<td>A1.2</td>
<td>8A</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>true</td>
<td>false</td>
</tr>
<tr>
<td>A1.2</td>
<td>8B</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>NRP</td>
<td>NRP</td>
<td>NRP</td>
<td>link</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>A2</td>
<td>9</td>
<td>IRI</td>
<td>IRI</td>
<td>none</td>
<td>IRI</td>
<td>none</td>
<td>IRI</td>
<td>none</td>
<td>IRI</td>
<td>NRP</td>
<td>IRI</td>
</tr>
<tr>
<td>I1</td>
<td>10</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>NRP</td>
<td>IRI</td>
<td>IRI</td>
</tr>
<tr>
<td>I2</td>
<td>11</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
</tr>
<tr>
<td>I3</td>
<td>12</td>
<td>NRP</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>NRP</td>
<td>NRP</td>
<td>IRI</td>
</tr>
<tr>
<td>R1.1</td>
<td>13</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>NRP</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
</tr>
<tr>
<td>R1.2</td>
<td>14A</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>NRP</td>
<td>NRP</td>
<td>IRI</td>
</tr>
<tr>
<td>R1.2</td>
<td>14B</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>NRP</td>
<td>NRP</td>
<td>IRI</td>
</tr>
<tr>
<td>R1.3</td>
<td>15</td>
<td>NRP</td>
<td>IRI</td>
<td>IRI</td>
<td>IRI</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>NRP</td>
<td>IRI</td>
<td>IRI</td>
</tr>
</tbody>
</table>
## FAIR Metrics Template

<table>
<thead>
<tr>
<th>FIELD</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metric Identifier</td>
<td>FM-F1B: <a href="https://purl.org/fair-metrics/FM_F1B">https://purl.org/fair-metrics/FM_F1B</a></td>
</tr>
<tr>
<td>Metric Name</td>
<td>Identifier persistence</td>
</tr>
<tr>
<td>To which principle does it apply?</td>
<td>F1</td>
</tr>
<tr>
<td>What is being measured?</td>
<td>Whether there is a policy that describes what the provider will do in the event an identifier scheme becomes deprecated.</td>
</tr>
<tr>
<td>Why should we measure it?</td>
<td>The change to an identifier scheme will have widespread implications for resource lookup, linking, and data sharing. Providers of digital resources must ensure that they have a policy to manage changes in their identifier scheme, with a specific emphasis on maintaining/redirecting previously generated identifiers.</td>
</tr>
<tr>
<td>What must be provided?</td>
<td>A URL that resolves to a document containing the relevant policy.</td>
</tr>
<tr>
<td>How do we measure it?</td>
<td>Use an HTTP GET on URL provided.</td>
</tr>
<tr>
<td>What is a valid result?</td>
<td>Present (a 200, 202, 203 or 206 HTTP response after resolving all and any prior redirects. e.g. 301 -&gt; 302 -&gt; 200 OK.) or Absent (any other HTTP code)</td>
</tr>
<tr>
<td>For which digital resource(s) is this relevant?</td>
<td>All</td>
</tr>
<tr>
<td>Comments</td>
<td>A first version of this metric would focus on just checking a URL that resolves to a document. We can’t verify that document. A second version would indicate how to structure the data policy document with a particular section (similar to how the CC license includes a formal structure in PDF).</td>
</tr>
</tbody>
</table>
FAIR Metrics Upgrades

Example: FM-F1B, Identifier Persistence

v1.0 checks for HTTP 200 return

v2.0 validates a standard RDF persistence policy

v3.0 scores multiple parameters of persistence policy
The “15th” FAIR Metric

Networkmeeting ZonMw
FAIR data and a new approach for data management
September 21 2018
Den Haag

**Nicoline Smit**  Project Manager at Netherland Heart Institute
**Mira van der Naald**  Department of Cardiology, UMC Utrecht

[https://preclinicaltrials.eu](https://preclinicaltrials.eu)
**Preclinicaltrials** aims to provide a comprehensive listing of preclinical animal study protocols.

Preferably registered at inception in order to **increase transparency**, help **avoid duplication**, and **reduce the risk of reporting bias** by enabling comparison of the completed study with what was planned in the protocol.

**Registration** of your study requires you to create an account that is

- Anonymous
- Free of charge
- Has an optional embargo period

This register is web-based, open to all types of animal studies and freely accessible and searchable to all with a preclinicaltrials.eu account.

The **registration form** is designed by experts on preclinical animal studies and preclinical evidence synthesis.

Please **join** us and create an user account, this will provide access to the database and enables you to register your preclinical trial.

Contact us at [info@preclinicaltrials.eu](mailto:info@preclinicaltrials.eu).
**Preclinicaltrials** aims to provide a comprehensive listing of preclinical animal study protocols.

Preferably registered at inception in order to **increase transparency**, help **avoid duplication**, and **reduce the risk of reporting bias** by enabling comparison of the completed study with what was planned in the protocol.

**Registration** of your study requires you to create an account that is

- Anonymous
- Free of charge
- Has an optional embargo period

This register is web-based, open to all types of animal studies and freely accessible and searchable to all with a preclinicaltrials.eu account.

The **registration form** is designed by experts on preclinical animal studies and preclinical evidence synthesis.

Please join us and create an user account, this will provide access to the database and enables you to register your preclinical trial.

Contact us at info@preclinicaltrials.eu.
Section 1. General information

1. * Title of the study
   Enter the full title of the study

2. Acronym/short title
   Enter optional acronym/short title for the study

3. * Contact details
   Give the name of the main administrative contact for the study
   
   Name

   Role
   What is the role the main contact in the study (e.g. executive researcher, research group supervisor)?

   Email address
   Provide the email address of the main contact

4. * Study centre details
   (This section is optional and includes information on the study centre's role and details if applicable)
## The “15th” FAIR Metric

<table>
<thead>
<tr>
<th>Metric Identifier</th>
<th>FM-PCT1 (FAIR Metric PreClinical Trail 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metric Name</td>
<td>Project registration</td>
</tr>
<tr>
<td>To which principle does it apply?</td>
<td>R1.2 (meta)data are associated with detailed provenance</td>
</tr>
<tr>
<td>What is being measured?</td>
<td>The existence of clinical trial registration</td>
</tr>
<tr>
<td>Why should we measure it?</td>
<td>Registration is important for Increased transparency and reduced risk of bias and help avoid duplication.</td>
</tr>
<tr>
<td>What must be provided?</td>
<td>A URL to the completed preclinical trial registration document</td>
</tr>
<tr>
<td>How do we measure it?</td>
<td>Use HTTP GET on URL provided.</td>
</tr>
<tr>
<td>What is a valid result?</td>
<td>HTTP 200 (now); Validted RDF file (later)</td>
</tr>
<tr>
<td>For which digital resource(s) is this relevant?</td>
<td>preclinicaltrails.eu</td>
</tr>
</tbody>
</table>
FAIR Metrics Upgrades

Example: FM-PCT1, Existence of project registration

v1.0 **checks** for HTTP 200 return

v2.0 **validates** a standard RDF project registrations form

v3.0 **scores** multiple parameters of project registration form
FAIR Metrics Upgrades

Example: FM-PCT1, Existence of project registration

- v1.0 checks for HTTP 200 return
- v2.0 validates a standard RDF project registrations form
- v3.0 scores multiple parameters of project registration form

https://www.go-fair.org/events/m4m-2-preclinical-trials-m4m-3-funders/

14 January 2019
FAIR Metrics Upgrades

Example: FM-PCT1, Existence of project registration

v1.0 **checks** for HTTP 200 return

v2.0 **validates** a standard RDF project registrations form

v3.0 **scores** multiple parameters of project registration form

14 January 2019

https://www.go-fair.org/events/m4m-2-preclinical-trials-m4m-3-funders/
M4M in the works

October 2018

- Inaugural M4M Workshop (October 15-16 2018)
  Schultes, Wittenburg, Mons & 7 more
  This project space is for collecting relevant resources used by or created during the inaugural M4M Workshop.

January 14

- #2 Preclinical Trials
  Schultes, Wittenburg, Mons & 7 more
  Community interested for animal research. In particular, advancing the use of research project registration forms, and making a FAIR Metric for this...

January 14

- #3 Funders
  Schultes, Wittenburg, Mons & 7 more
  International community of Science Funders who are interested in supporting FAIR Data outputs (tracking re-use of data)

June

- #4 Publishers
  Schultes, Wittenburg, Mons & 7 more
  Initiated by Myles Axton (now at Wiley). The idea is that a library of metadata statements will enable a publisher to (easily) compose metadata profiles.

Fall

- #5 Antimicrobial Resistance Surveillance
  Schultes, Wittenburg, Mons & 7 more
  Community trying to get real-time interoperability of clinical findings in Antimicrobial Resistance.

- #6 Integrating the Healthcare Enterprise (IHE)
  Schultes, Wittenburg, Mons & 7 more
  IHE is an initiative by healthcare professionals and industry to improve the way computer systems in healthcare share information. IHE promotes the co...

- #7 GO FAIR Chemistry IN
  Schultes, Wittenburg, Mons & 7 more
  Community (IUPAC and others) that is setting standards for FAIR Chemistry.

- #8 GO FAIR Personal Health Train IN
  Schultes, Wittenburg, Mons & 7 more
  PHT Community needs to decide on "red box" / "blue box" issues.

- #9 FAIR Data Point
  Schultes, Wittenburg, Mons & 9 more

- #10 Rare Diseases
  Schultes & Roos
Generation 2
FAIR Maturity Indicators

The FAIR Maturity Measurement Service
Test a (meta)data resource against the FAIR Maturity Indicators

Public Entry Points:
- Search for Tests and Test Collections
- Browse Maturity Indicator Tests
- Register a new Maturity Indicator Test
- Browse Indicator Collections
- Create New Indicator Collection
- Browse existing Maturity Evaluations
- Execute a new Maturity Evaluation

http://linkeddata.systems:3000/FAIR_Evaluator/
<table>
<thead>
<tr>
<th>Name</th>
<th>Creator</th>
<th>Email</th>
<th>Principle</th>
<th>Tests Maturity Indicator Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>FAIR Metrics Gen2- Unique Identifier</td>
<td>Mark D Wilkinson</td>
<td><a href="mailto:markw@illuminae.com">markw@illuminae.com</a></td>
<td>F1</td>
<td><a href="https://purl.org/fair-metrics/Gen2_FM_F1A">https://purl.org/fair-metrics/Gen2_FM_F1A</a></td>
</tr>
<tr>
<td>FAIR Metrics Gen2 - Identifier Persistence</td>
<td>Mark D Wilkinson</td>
<td><a href="mailto:markw@illuminae.com">markw@illuminae.com</a></td>
<td>F1</td>
<td><a href="https://purl.org/fair-metrics/Gen2_FM_F1B">https://purl.org/fair-metrics/Gen2_FM_F1B</a></td>
</tr>
<tr>
<td>FAIR Metrics Gen2 - Data Identifier Persistence</td>
<td>Mark D Wilkinson</td>
<td><a href="mailto:markw@illuminae.com">markw@illuminae.com</a></td>
<td>F1</td>
<td><a href="https://purl.org/fair-metrics/Gen2_FM_F1B">https://purl.org/fair-metrics/Gen2_FM_F1B</a></td>
</tr>
<tr>
<td>FAIR Metrics Gen2 - Structured Metadata</td>
<td>Mark D Wilkinson</td>
<td><a href="mailto:markw@illuminae.com">markw@illuminae.com</a></td>
<td>F2</td>
<td><a href="https://purl.org/fair-metrics/Gen2_FM_F2A">https://purl.org/fair-metrics/Gen2_FM_F2A</a></td>
</tr>
<tr>
<td>FAIR Metrics Gen2 - Grounded Metadata</td>
<td>Mark D Wilkinson</td>
<td><a href="mailto:markw@illuminae.com">markw@illuminae.com</a></td>
<td>F2</td>
<td><a href="https://purl.org/fair-metrics/Gen2_FM_F2B">https://purl.org/fair-metrics/Gen2_FM_F2B</a></td>
</tr>
<tr>
<td>FAIR Metrics Gen2- Data Identifier Explicitly In Metadata</td>
<td>Mark D Wilkinson</td>
<td><a href="mailto:markw@illuminae.com">markw@illuminae.com</a></td>
<td>F3</td>
<td><a href="https://purl.org/fair-metrics/Gen2_FM_F3">https://purl.org/fair-metrics/Gen2_FM_F3</a></td>
</tr>
<tr>
<td>FAIR Metrics Gen2- Metadata Identifier Explicitly In Metadata</td>
<td>Mark D Wilkinson</td>
<td><a href="mailto:markw@illuminae.com">markw@illuminae.com</a></td>
<td>F3</td>
<td><a href="https://purl.org/fair-metrics/Gen2_FM_F3">https://purl.org/fair-metrics/Gen2_FM_F3</a></td>
</tr>
<tr>
<td>FAIR Metrics Gen2 - Searchable in major search engine</td>
<td>Mark D Wilkinson</td>
<td><a href="mailto:markw@illuminae.com">markw@illuminae.com</a></td>
<td>F4</td>
<td><a href="https://purl.org/fair-metrics/Gen2_FM_F4">https://purl.org/fair-metrics/Gen2_FM_F4</a></td>
</tr>
<tr>
<td>FAIR Metrics Gen2 - Uses open free protocol for data retrieval</td>
<td>Mark D Wilkinson</td>
<td><a href="mailto:markw@illuminae.com">markw@illuminae.com</a></td>
<td>A1.1</td>
<td><a href="https://purl.org/fair-metrics/Gen2_FM_A1.1">https://purl.org/fair-metrics/Gen2_FM_A1.1</a></td>
</tr>
<tr>
<td>FAIR Metrics Gen2 - Uses open free protocol for metadata retrieval</td>
<td>Mark D Wilkinson</td>
<td><a href="mailto:markw@illuminae.com">markw@illuminae.com</a></td>
<td>A1.1</td>
<td><a href="https://purl.org/fair-metrics/Gen2_FM_A1.1">https://purl.org/fair-metrics/Gen2_FM_A1.1</a></td>
</tr>
<tr>
<td>FAIR Metrics Gen2 - data autentication and authorization</td>
<td>Mark D Wilkinson</td>
<td><a href="mailto:markw@illuminae.com">markw@illuminae.com</a></td>
<td>A1.2</td>
<td><a href="https://purl.org/fair-metrics/Gen2_FM_A1.2">https://purl.org/fair-metrics/Gen2_FM_A1.2</a></td>
</tr>
<tr>
<td>FAIR Metrics Gen2 - metadata autentication and authorization</td>
<td>Mark D Wilkinson</td>
<td><a href="mailto:markw@illuminae.com">markw@illuminae.com</a></td>
<td>A1.2</td>
<td><a href="https://purl.org/fair-metrics/Gen2_FM_A1.2">https://purl.org/fair-metrics/Gen2_FM_A1.2</a></td>
</tr>
<tr>
<td>FAIR Metrics Gen2 - Metadata Persistence</td>
<td>Mark D Wilkinson</td>
<td><a href="mailto:markw@illuminae.com">markw@illuminae.com</a></td>
<td>A2</td>
<td><a href="https://purl.org/fair-metrics/Gen2_FM_A2">https://purl.org/fair-metrics/Gen2_FM_A2</a></td>
</tr>
<tr>
<td>FAIR Metrics Gen2 - Metadata Knowledge Representation Language (weak)</td>
<td>Mark D Wilkinson</td>
<td><a href="mailto:markw@illuminae.com">markw@illuminae.com</a></td>
<td>I1</td>
<td><a href="https://purl.org/fair-metrics/Gen2_FM_I1A">https://purl.org/fair-metrics/Gen2_FM_I1A</a></td>
</tr>
<tr>
<td>FAIR Metrics Gen2 - Metadata Knowledge Representation Language (strong)</td>
<td>Mark D Wilkinson</td>
<td><a href="mailto:markw@illuminae.com">markw@illuminae.com</a></td>
<td>I1</td>
<td><a href="https://purl.org/fair-metrics/Gen2_FM_I1B">https://purl.org/fair-metrics/Gen2_FM_I1B</a></td>
</tr>
<tr>
<td>FAIR Metrics Gen2 - Data Knowledge Representation Language (weak)</td>
<td>Mark D Wilkinson</td>
<td><a href="mailto:markw@illuminae.com">markw@illuminae.com</a></td>
<td>I1</td>
<td><a href="https://purl.org/fair-metrics/Gen2_FM_I1A">https://purl.org/fair-metrics/Gen2_FM_I1A</a></td>
</tr>
<tr>
<td>FAIR Metrics Gen2 - Data Knowledge Representation Language (strong)</td>
<td>Mark D Wilkinson</td>
<td><a href="mailto:markw@illuminae.com">markw@illuminae.com</a></td>
<td>I1</td>
<td><a href="https://purl.org/fair-metrics/Gen2_FM_I1B">https://purl.org/fair-metrics/Gen2_FM_I1B</a></td>
</tr>
<tr>
<td>FAIR Metrics Gen2 - Metadata uses FAIR vocabularies (weak)</td>
<td>Mark D Wilkinson</td>
<td><a href="mailto:markw@illuminae.com">markw@illuminae.com</a></td>
<td>I2</td>
<td><a href="https://purl.org/fair-metrics/Gen2_FM_I2A">https://purl.org/fair-metrics/Gen2_FM_I2A</a></td>
</tr>
<tr>
<td>FAIR Metrics Gen2 - Metadata uses FAIR vocabularies (strong)</td>
<td>Mark D Wilkinson</td>
<td><a href="mailto:markw@illuminae.com">markw@illuminae.com</a></td>
<td>I2</td>
<td><a href="https://purl.org/fair-metrics/Gen2_FM_I2B">https://purl.org/fair-metrics/Gen2_FM_I2B</a></td>
</tr>
<tr>
<td>FAIR Metrics Gen2 - Metadata contains qualified outward references</td>
<td>Mark D Wilkinson</td>
<td><a href="mailto:markw@illuminae.com">markw@illuminae.com</a></td>
<td>I3</td>
<td><a href="https://purl.org/fair-metrics/Gen2_FM_I3A">https://purl.org/fair-metrics/Gen2_FM_I3A</a></td>
</tr>
<tr>
<td>FAIR Metrics Gen2 - Metadata Includes License (strong)</td>
<td>Mark D Wilkinson</td>
<td><a href="mailto:markw@illuminae.com">markw@illuminae.com</a></td>
<td>R1.1</td>
<td><a href="https://purl.org/fair-metrics/Gen2_FM_R1.1">https://purl.org/fair-metrics/Gen2_FM_R1.1</a></td>
</tr>
<tr>
<td>FAIR Metrics Gen2 - Metadata Includes License (weak)</td>
<td>Mark D Wilkinson</td>
<td><a href="mailto:markw@illuminae.com">markw@illuminae.com</a></td>
<td>R1.1</td>
<td><a href="https://purl.org/fair-metrics/Gen2_FM_R1.1">https://purl.org/fair-metrics/Gen2_FM_R1.1</a></td>
</tr>
</tbody>
</table>
Metadata for Machines

Making it easy for humans to make metadata for machines

https://www.go-fair.org/resources/go-fair-workshop-series/metadata-for-machines-workshops/
Metadata for Machines

First M4M: October 15-16, Leiden
Co-organizers: Wittenburg & Schultes

https://digitalscholarshipleiden.nl/articles/metadata-4-machines-help-you-find-and-reuse-relevant-research-data
M4M #2: Preclinical Trials
M4M #3: Funders

January 14, 2019
Leiden

OSF Project: https://osf.io/qe9fa/
Common Notes: http://bit.ly/2HbGrUr
Metadata for Machines

Vehicle for Community Decision Making
Metadata for Machines

Likely Stakeholders:

- Data steward
- Researcher / Research community
- University / organizations
- Funder
- Publisher
- Repository
- FAIR tools and services
- Instrumentation and other 3rd parties
M4M Workshop

Research Community:
F2
R1.1
R1.2
R1.3

Currently, not a lot of structure in this process

Research Community:
Decisions on domain-relevant community standards
M4M Workshop

Research Community: F2 R1.1 R1.2 R1.3

Research Community: Decisions on domain-relevant community standards
Research Community:
F2
R1.1
R1.2
R1.3

Re-use existing community standards

Create new community standards

Research Community:
Decisions on domain-relevant community standards

M4M Workshop
GO FAIR Supports and Coordinates INs in their implementation choices and challenges.

15 FAIR Guiding Principles

- EOSC
- NIH Data Commons
- Preclinical Trials
- Funders
- American Geophysical Union
- Bayer
- Journalists
- Financial industry

Self-Identified Community aiming to become more FAIR

Implementation Choice
Community chooses to re-use existing technology to implement FAIR

Implementation Challenge
Community accepts challenge to create new technology to implement FAIR

Inspect IN Profile Matrix

Reuse Sustainability plans Funded infrastructure PPP services providers
Community Implementation Choices & Challenges

15 FAIR Guiding Principles
- EOSC
- NIH Data Commons
- Preclinical Trials
- Funders
- American Geophysical Union
- Bayer
- Journalists
- Financial industry

GO FAIR Supports and Coordinates the community in their implementation choices and challenges.

Maximizing Metadata Reuse
- Tools & registries
- Sustainability plans
- Funded infrastructure
- PPP services providers

Implementation Choice
Community chooses to re-use existing metadata elements as needed to implement FAIR

Implementation Challenge
Community accepts challenge to create new metadata elements as needed to implement FAIR

Inspect IN Profile Matrix
Self-Identified Community aiming to become more FAIR
Semi-automatic

Would you like more information first? Please click here.

Select an earlier saved dataset evaluation or evaluate a new dataset
Erik's Manual Data Set

Findable

F1 Q0: Specify the URL of the resource to be evaluated.

F1 Q1: Provide a URL to a registered scheme that defines the globally-unique structure of the identifier(s) for your digital resource.

F1 Q2: Provide a URL to a document that defines the persistence policy of your identifier(s).

F2 Q3: Provide the URL to a document that contains machine-readable metadata for the digital resource.

F2 Q4: Provide the URL for the file format of this metadata.

FAIR Principle F2: Data are described with rich metadata.

Data should be machine-discoverable by the widest range of stakeholders possible. That is, you should not presume who will want to use your data, or for what purpose.

Resource providers should be generous and expansive with their metadata (see also R1). Exemplar rich metadata frameworks: DCAT; ISA framework.

Community challenge:
(4) Can you define a minimal set of metadata for your community?
(5) Can you make your metadata machine-readable?

Accessible

A1.1 Q9: Provide a URL to the description of the access protocol.

A1.1 Q10: The protocol is Open (Yes/No).

A1.1 Q11: The protocol is (royalty) free (Yes/No).
Hackathon Result Summary

56 users
32 datasets
25 metric evaluations
Hackathon Result Summary

<table>
<thead>
<tr>
<th>PPB FAIR Score</th>
<th>Before</th>
<th>After</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>cBioPortal</td>
<td>39</td>
<td>59</td>
<td>+20</td>
</tr>
<tr>
<td>JaxTeam</td>
<td>26</td>
<td>54</td>
<td>+28</td>
</tr>
<tr>
<td>Broad Single Cell</td>
<td>29</td>
<td>52</td>
<td>+23</td>
</tr>
<tr>
<td>Bioassay</td>
<td>48</td>
<td>75</td>
<td>+27</td>
</tr>
</tbody>
</table>

Metric evaluator: Manual check

Performance run at: 2018-06-15 10:07