FAIR Maturity Indicators & Evaluation

Work done by:

Mark D. Wilkinson, Susanna-Assunta Sansone, Erik Schultes,
Peter Doorn, Luiz Olavo Bonino da Silva Santos, Michel Dumontier

Do you find any of the FAIR Principles

… weird?

The principles

To be Findable:

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 it describes

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

To be 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

To be 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

To be 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

What about this one:

F3: Metadata clearly and explicitly include the identifier of the data it describes

That's an unusually specific rule, given the generality of the other Principles! …why?

Understanding the Objective/Purpose of each Principle helps clarify the basis of a "Maturity Indicator" we designed to measure it

FAIR Principles Explained:
https://www.go-fair.org/fair-principles/

What makes a measurement "good"?

- **Clear:** so that anybody can understand what is meant.

- **Realistic:** so that anybody can report on what is being asked of them.

- **Discriminating:** so that we can distinguish the degree to which a resource meets a specific FAIR principle, and can provide instruction as to what would maximize that value.

- **Measurable:** The assessment can be made in an objective, quantitative, machine-interpretable, scalable and reproducible manner → transparency of what is being measured, and how.

- **Universality:** The extent to which the MIis applicable to all digital resources.

We designed a set of parameters that must be considered for every MI.

The parameters are designed to help ensure that the MI you are designing is "good".

The List

MI Identifier	FAIR MIs should, themselves, be FAIR objects, and thus should have globally unique identifiers.
MI Name	human-readable name for the MI
To which principle does it apply?	MIs should address only one sub-principle, since each FAIR principle is particular to one feature of a digital resource; MIs that address multiple principles are likely to be measuring multiple features, and those should be separated whenever possible.
What is being measured?	A precise description of the aspect of that digital resource that is going to be evaluated
Why should we measure it?	Describe why it is relevant to measure this aspect
What must be provided?	What information is required to make this measurement?
How do we measure it?	In what way will that information be evaluated?
What is a valid result?	What outcome represents "success" versus "failure"
For which digital resource(s) is this relevant?	If possible, a MI should apply to all digital resources; however, some MIs may be applicable only to a subset. In this case, it is necessary to specify the range of resources to which the MI is reasonably applicable.

There have been two iterations of MI
building

"Generation 1" MIs
are used as the basis of a questionnaire

"Generation 2" - MIs
are used for fully-automated evaluations

Caveat

We decided that we would only design MIs that test the FAIRness of a resource **from the perspective of a machine**

The FAIR principles emphasise that data must be FAIR both for humans, and for machines

i.e. a machine should be able to replace a human with respect to:

- Discovery of the data of interest
- Discovery of how to access the data (both technological and "contractual")
- Identification of the data format, and the ability to parse the data into a "sensible" in-memory representation
- Discovery of linked information
- Discovery (and download of) of contextual information relevant to the interpretation of the data
- Discovery of the license associated with that data.

Generation 1 MIs

F1: (meta) data are assigned globally unique and persistent identifiers

How should we measure this in a way that is clear, realistic, measurable, discriminating, and universal?

MI Identifier: FM-F1A: https://purl.org/fair-MIs/FM_F1A

MI Name: Identifier **Uniqueness**

To which principle does it apply? F1

What is being measured? Whether there is a scheme to uniquely identify the digital resource.

Why should we measure it? The uniqueness of an identifier is a necessary condition to unambiguously refer that resource, and that resource alone. Otherwise, an identifier shared by multiple resources will confound efforts to describe that resource, or to use the identifier to retrieve it. Examples of identifier schemes include, but are not limited to URN, IRI,DOI, Handle, trustyURI, LSID, etc. For an in-depth understanding of the issues around identifiers, please see http://dx.plos.org/10.1371/journal.pbio.2001414

What must be provided?

How do we measure it?

What must be provided? URL to a registered identifier scheme.

How do we measure it? An identifier scheme is valid if and only if it is described in a repository that can register and present such identifier schemes (e.g. **fairsharing.org**).

What is a valid result? Present or Absent

FM-F1A: Identifier Uniqueness

FIELD	DESCRIPTION
Metric Identifier	FM-F1A: https://purl.org/fair-metrics/FM_F1A
Metric Name	Identifier Uniqueness
To which principle does it apply?	F1
What is being measured?	Whether there is a scheme to uniquely identify the digital resource.
Why should we measure it?	The uniqueness of an identifier is a necessary condition to unambiguously refer that resource, and that resource alone. Otherwise, an identifier shared by multiple resources will confound efforts to describe that resource, or to use the identifier to retrieve it. Examples of identifier schemes include, but are not limited to URN, IRI, DOI, Handle, trustyURI, LSID, etc. For an in-depth understanding of the issues around identifiers, please see http://dx.plos.org/10.1371/journal.pbio.2001414
What must be provided?	URL to a registered identifier scheme.
How do we measure it?	An identifier scheme is valid if and only if it is described in a repository that can register and present such identifier schemes (e.g. fairsharing.org). Information about the identifier scheme must be presented with a machine-readable document containing the FM1 attribute with the URL to where the scheme is described. see specification for implementation.
What is a valid result?	Present or Absent
For which digital resource(s) is this relevant?	All

FM-F1B: Identifier Persistence

FIELD	DESCRIPTION
Metric Identifier	FM-F1B: https://purl.org/fair-metrics/FM_F1B
Metric Name	Identifier persistence
To which principle does it apply?	F1
What is being measured?	Whether there is a policy that describes what the provider will do in the event an identifier scheme becomes deprecated.
Why should we measure it?	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.
What must be provided?	A URL that resolves to a document containing the relevant policy.
How do we measure it?	Use an HTTP GET on URL provided.
What is a valid result?	Present (a 200,202,203 or 206 HTTP response after resolving all and any prior redirects. e.g. 301 -> 302 -> 200 OK.) or Absent (any other HTTP code)
For which digital resource(s) is this relevant?	All
Comments	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 licenses now have a formal structure in RDF). A third version would insist that that document and section is signed by an approved organization and made available in an appropriate repository.

FM-F2: Machine Readability of Metadata

FIELD	DESCRIPTION
Metric Identifier	FM-F2: https://purl.org/fair-metrics/FM_F2
Metric Name	Machine-readability of metadata
To which principle does it apply?	F2 - Data are described with rich metadata
What is being measured?	The availability of machine-readable metadata that describes a digital resource.
Why should we measure it?	Richness of metadata can refer to many different aspects. One aspect is that the machine readability of metadata makes it possible to optimize their discovery. For instance, Web search engines suggest the use of particular structured metadata elements to optimize search. Thus, the machine-readability aspect can help people and machines find a digital resource of interest. Here, we focus on metadata being sufficiently rich in this sense - that the metadata document and the metadata elements are machine readable. Otherwise, it will also be difficult to understand what the digital resource is and what information is being provided about it.
What must be provided?	A URL to a document that contains machine-readable metadata for the digital resource. Furthermore, the file format must be specified.
How do we measure it?	HTTP GET on the metadata URL. A response of [a 200,202,203 or 206 HTTP response after resolving all and any prior redirects. e.g. 301 -> 302 -> 200 OK.] indicates that there is indeed a document. The second URL should resolve to the record of a registered file format (e.g. DCAT, DICOM, schema.org etc.) in a registry like FAIRsharing.
What is a valid result?	Machine-readable or Machine-not-readable
For which digital resource(s) is this relevant?	All
Comments	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 licenses now have a formal structure in RDF). A third version would insist that that document and section is signed by an approved organization and made available in an appropriate repository.

FM-F3: Resource Identifier in Metadata

FIELD	DESCRIPTION
Metric Identifier	FM-F3: https://purl.org/fair-metrics/FM_F3
Metric Name	Resource Identifier in Metadata
To which principle does it apply?	F3 - metadata clearly and explicitly include the identifier of the data it describes
What is being measured?	Whether the metadata document contains the globally unique and persistent identifier for the digital resource.
Why should we measure it?	The discovery of digital object should be possible from its metadata. For this to happen, the metadata must explicitly contain the identifier for the digital resource it describes. A metadata document should also not result in ambiguity about the digital object it is describing. This can be assured if the metadata document explicitly refers to the digital object by its IRI.
What must be provided?	The URL of the metadata and the IRI of the digital resource it describes.
How do we measure it?	Parsing the metadata for the given digital resource IRI.

Generation 2: FAIR Maturity Indicators

SWITCH HERE

1. They don't scale to the entire world!
2. They are time-consuming for busy people
3. They cannot be executed by "anyone" (only by the person who knows the resource deeply)
4. They are (potentially) biased
5. They don't adequately test one of the main point of FAIR, because a human is not capable of evaluating this:

Can a machine find and (re)use the data?

History:

- Erik and I did an in-depth review of the answers to the original questionnaire-based FAIRness assessments (~11 resources responded)
- I took-note of what people were doing, in-practice, that they felt was "FAIR"
- I also took-note of the complaints from key data repositories about what they thought was "unfair" in our initial evaluations
- I compiled a catalogue of various approaches to, in particular, the provision of metadata
- I used that to build a "metadata harvesting" library that attempts to be "exhaustive"
  - That is, it pursues paths that I (personally) don't consider to be "FAIR in-spirit"
  - Trying not to be "prescriptive"!

Why do objective evaluations?

- Organizations want to know if they are FAIR
- Organizations want ADVICE on what they can do better
  - What is required for FAIR compliance?
  - How difficult is it?
  - How much will it cost?
  - What expertise do I need?

- Objective evaluations (with narrative feedback on failures) provide these answers!

Meta Data Harvester Workflow
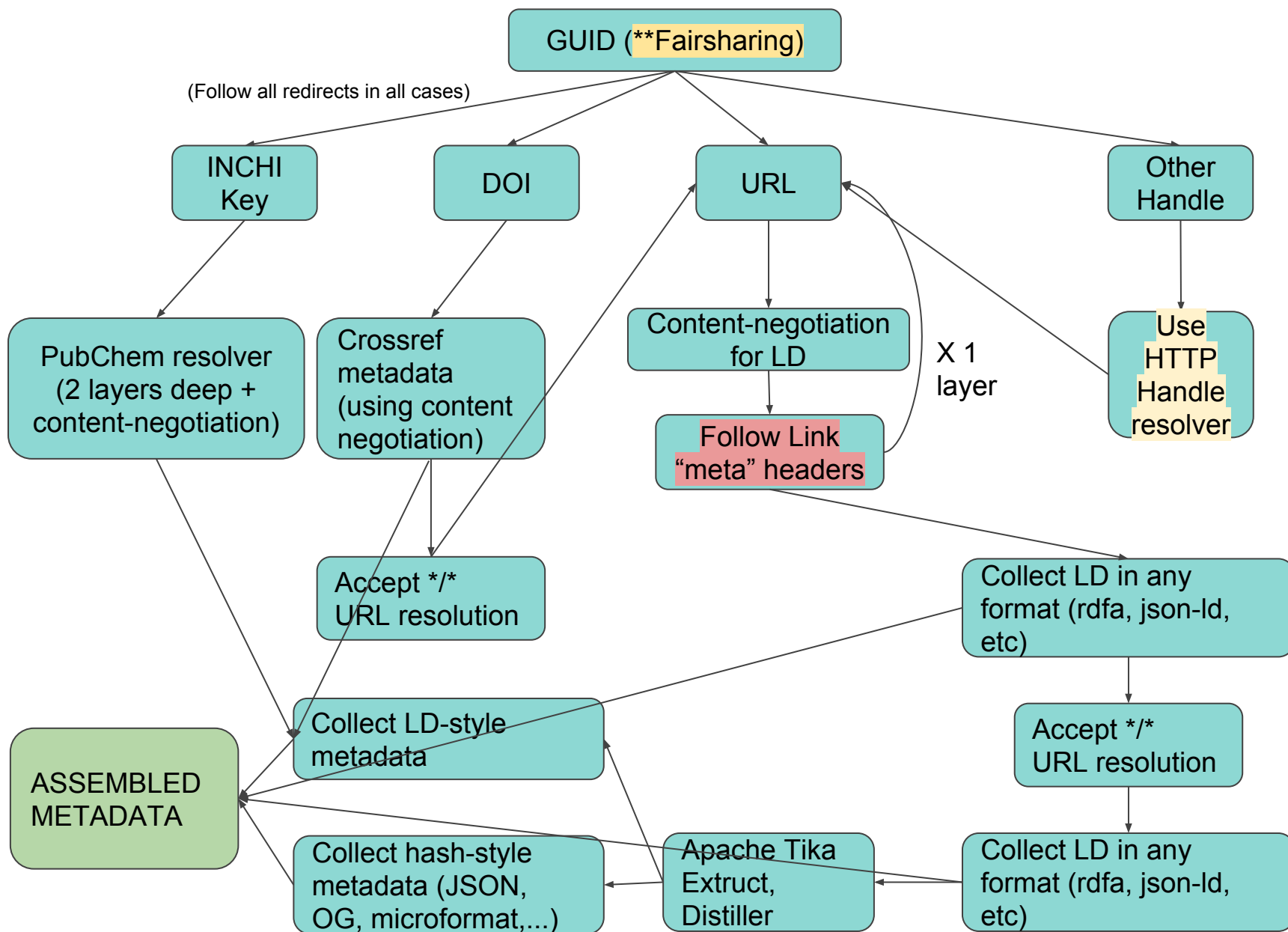
This mass of Key/value and LD-style metadata

is provided to each MI test

Gen2 Maturity Indicators

- Are "standalone" Web interfaces that can be written by anyone
- Consume ONLY the GUID of the metadata (they invoke the harvester with that)
- Their interface metadata is recorded using smartAPI (or openAPI)
- They are registered in the smartAPI registry for discovery (not required)
- They are registered in The Evaluator to be used by others (not required)
- They test the "assembled metadata" for various features
- Some may attempt to interrogate the data also

For example, the Gen2 "data identifier in metadata" looks for a hash key, or a LD property, from a list of widely-used properties that are intended to point at data, including:

foaf:primaryTopic, dcat:distribution, ldp:contains, schema:mainEntity....

"Data identifier explicitly in metadata"

To locate the data identifier, hash data is tested for the keys:

- codeRepository
- mainEntity
- primaryTopic
- IAO:0000136 (is about)
- IAO_0000136
- SIO:000332 (is about)
- SIO_000332
- distribution
- contains

Graph data is tested for the properties:

- schema:codeRepository
- schema:mainEntity
- foaf:primaryTopic
- IAO:0000136 (information artifact ontology 'is about')
- SIO:000332 (SemanticScience Integrated Ontology 'is about')
- schema:distribution
- DCAT:distribution (Data Catalogue vocabulary)
- ldp:contains (Linked Data Platform)

Maturity Indicators return binary - pass/fail

- We decided that trying to assign a partial score to a test was too arbitrary

- Gen2 Maturity Indicators (with one exception, that I need to re-code!) return binary pass/fail

- They attempt to log everything they do, so that the output contains a record of why the test passed/failed

  - The POINT of the test is to encourage incremental improvements of FAIRness, so this feedback is important not only for transparency, but to be informative/instructive

(Human-readable representation)

Completed Evaluation of https://www.lifelines.nl

Title: Lifelines cohort

F1: FAIR Metrics Gen2- Unique Identifier

★★★★★

Date: 2019-02-20T16:08:02+00:00
Score: 1
Comment: Found a URI - pass

F3: FAIR Metrics Gen2- Data Identifier Explicitly In Metadata

★

Date: 2019-02-20T16:08:06+00:00
Score: 0
Comment: Found html text/html type of file by resolving GUID. Using 'extruct' to try to extract metadata from return value (message body) of https://www.lifelines.nl. Was unable to locate the data identifier in the metadata using any (common) property/predicate reserved for this purpose. Tested SIO:is-about, SIO:0003323, schema:mainEntity, IAO:0000136, schema:codeRepository, and foaf:primaryTopic. Sorry!

F3: FAIR Metrics Gen2- Metadata Identifier Explicitly In Metadata

★★★★★

Date: 2019-02-20T16:08:09+00:00
Score: 1.0
Comment: Found html text/html type of file by resolving GUID. Using 'extruct' to try to extract metadata from return value (message body) of https://www.lifelines.nl. Found pattern-match in metadata https://www.lifelines.nl http://ogp.me/ns#url https://www.lifelines.nl/. This provides a partial success score. Found pattern-match in metadata https://www.lifelines.nl http://ogp.me/ns#description The Lifelines biobank is an international resource for health research. Using a large-scale cohort study all data and samples are collected.. This provides a partial success score. Found exact-match of https://www.lifelines.nl (e.g. as a string) in RDF output. found matching metadata https://www.lifelines.nl http://ogp.me/ns#description The Lifelines biobank is an international resource for health research. Using a large-scale cohort study all data and samples are collected.. Success!

The Evaluator - designed for bottom-up!

- The Evaluator provides both a human and a machine-accessible (JSON) interface for:
  - Registering new Maturity Indicators **designed by the community**
  - Registering **community-specific collections** of Maturity Indicators
  - Executing Evaluations (the application of a collection of MIs to a specific GUID) by anyone
  - Searching for MIs and Collections based on keywords

The API is documented both in human-readable form (https://github.com/FAIRMIs/MIs/blob/master/MIsEvaluatorCode/Ruby/fairMIs/README.md) and as a Swagger-enabled registration in smartAPI

Community Participation

- Anyone can create a new Maturity Indicator, and submit it for open discussion
- Anyone can suggest edits to existing Maturity Indicators, if we're not "fair"

https://github.com/FAIRMIs/MIs/blob/master/MaturityIndicators/README.md

HOW TO CREATE A NEW MATURITY INDICATOR

FAIR Maturity Indicators are created, initially, as a narrative document, following a template extablished by the FAIR Metrics Authoring Group. A MarkDown version of this template is available above, and should be used for Maturity Indicator submissions by the public. Guidance for how to complete this document is found in the authoring framework overview.

A Template Markdown file is provided for you in the MaturityIndicator folder. Once a Maturity Indicator has been designed, the document should be submitted via 'pull request' to this repository, at which time it becomes available for community discussion. The author of the Maturity Indicator should publicize their submission as widely as possible, to encourage maximal community input.

At this time, there is no formal process for adoption of Maturity Indicators (incuding those that the Authoring Group have designed themselves!), as there is no official body that can recognize or "stamp" a Maturity Indicator as being "valid". Nevertheless, authors should consider the comments and criticisms they receive, and modify the submission accordingly if the criticisms are justified.

HOW TO CREATE A NEW MATURITY INDICATOR TEST

With the goal of providing an objective, automated way of testing (meta)data resources against the Maturity Indicator, the

(Current) Workflow for registering a new MI

- Using the provided template, the community member writes a human-readable Maturity Indicator description
- Pull request on GitHub
- The proposed Indicator is "scraped" by FAIRSharing, and registered as "under consideration"
- A (yet to be defined) process, including discussion and advice from FAIR evaluation experts, will ensue
- The Maturity Indicator will be approved (yet to be defined)
- FAIRSharing will update their record to show that this is an approved Maturity Indicator
- An associated MI Test should then also be written (at the moment, muggins is doing this for everyone, only because the metadata harvester is a Ruby module that must therefore be imported by a Ruby test)

Registering a new Evaluation Collection

- Communities can decide which Maturity Indicators are relevant to them
- These are registered in the Evaluator as a "Collection", with some documentation about what MIs are included, and to what communities the Collection would be relevant (for the purpose of re-use)
- Evaluations are executed by POSTing the GUID to the URI of the Collection that the community thinks is relevant

- Anyone can execute an evaluation on any GUID
- Anyone can select the Collection they wish to apply (e.g. Journals may select different collections than funding agencies, or researchers)

The point of objective, community-driven evaluations is to give control to community governance organizations

THEY choose what is evaluated, based on what THEY care about!

THEY can choose to check their resources against a more "core/global" standard set of tests

https://w3id.org/FAIR_Evaluator/

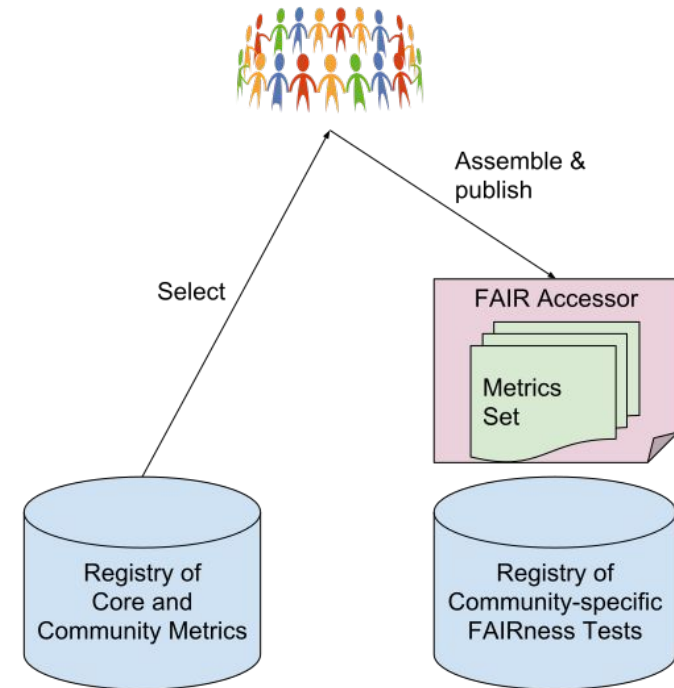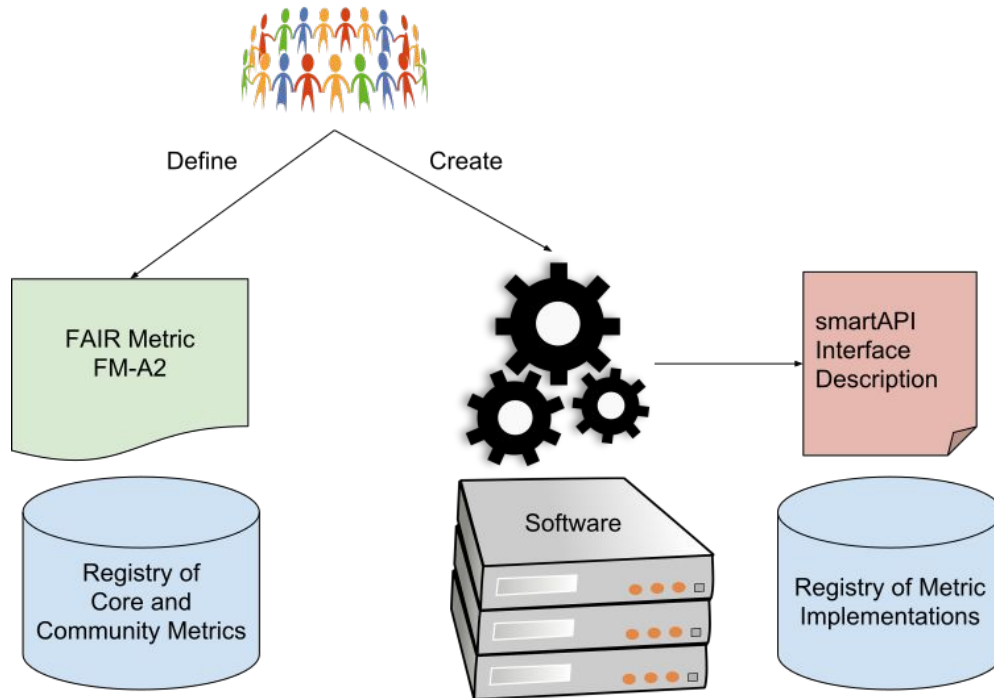The human interface is…. Ugly! Sorry, I am NOT a Web designer! ;-)

I welcome anyone to create front-ends to The Evaluator - the API is fully documented (at the URLs mentioned earlier)

Pull requests to GitHub that make improvements to the interface are also welcome!
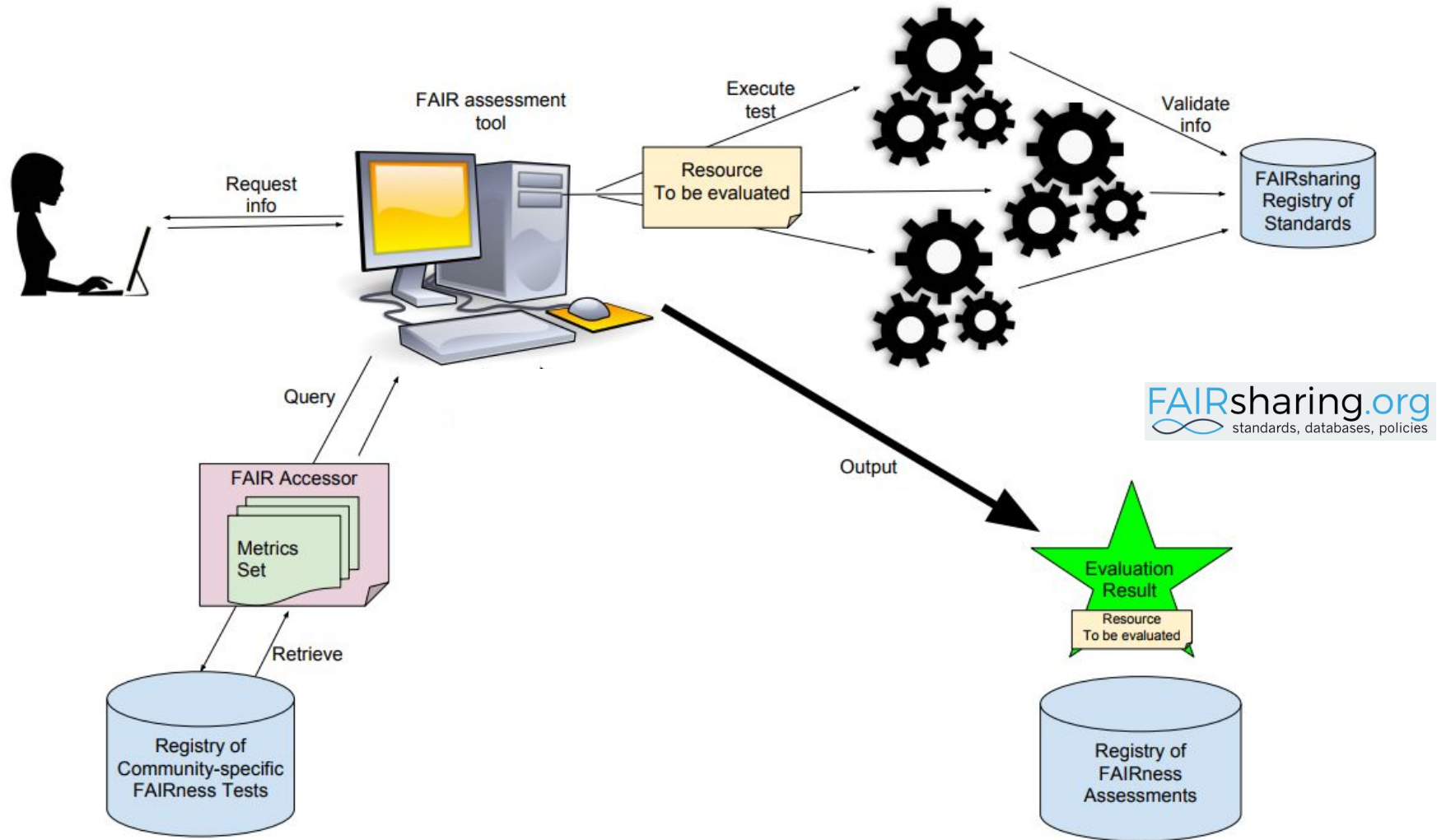
Desiderata for a FAIR MIs testing framework

1) All components of the framework should themselves be FAIR
2) Tests should be modular - mirroring (as much as possible) the MIs themselves
3) Tests should be as objective as possible
4) All stakeholders should be able to define their own MI Tests
5) All stakeholders should be able to define their own "Evaluations" (combinations of Tests relevant to that stakeholder)
6) The Evaluation system should evolve to accept new standards, without re-coding
7) Anyone should be able to evaluate anything
8) The system should be aspirational - provide feedback for improvement

FAIRsharing.org
standards, databases, policies

FAIR assessment tool

Request info

Execute test

Resource To be evaluated

Validate info

FAIRsharing Registry of Standards

FAIRsharing.org
standards, databases, policies

Query

FAIR Accessor

Metrics Set

Output

Retrieve

Registry of Community-specific FAIRness Tests

Evaluation Result

Resource To be evaluated

Registry of FAIRness Assessments

Automated, Objective, Aspirational!

https://w3id.org/FAIR_Evaluator/

Overview

1. Every MI is associated with a Web-based interface that can evaluate compliance with that MI

2. New MIs can be registered simply by pointing to the URL for their smartAPI

3. Collection of MIs can be assembled by anyone, to represent the aspects of FAIRness that they care about (e.g. a journal vs. funding agency vs researcher)

4. You can execute an evaluation by providing an IRI to be tested, and a collection of MIs to be applied to it

5. Evaluations can be recovered (see previous input data) and/or re-executed, either through the Web interface, or by direct connection to the Evaluator from software (i.e. the Web page is only for people)

Mark Wilkinson

mark.wilkinson@upm.es