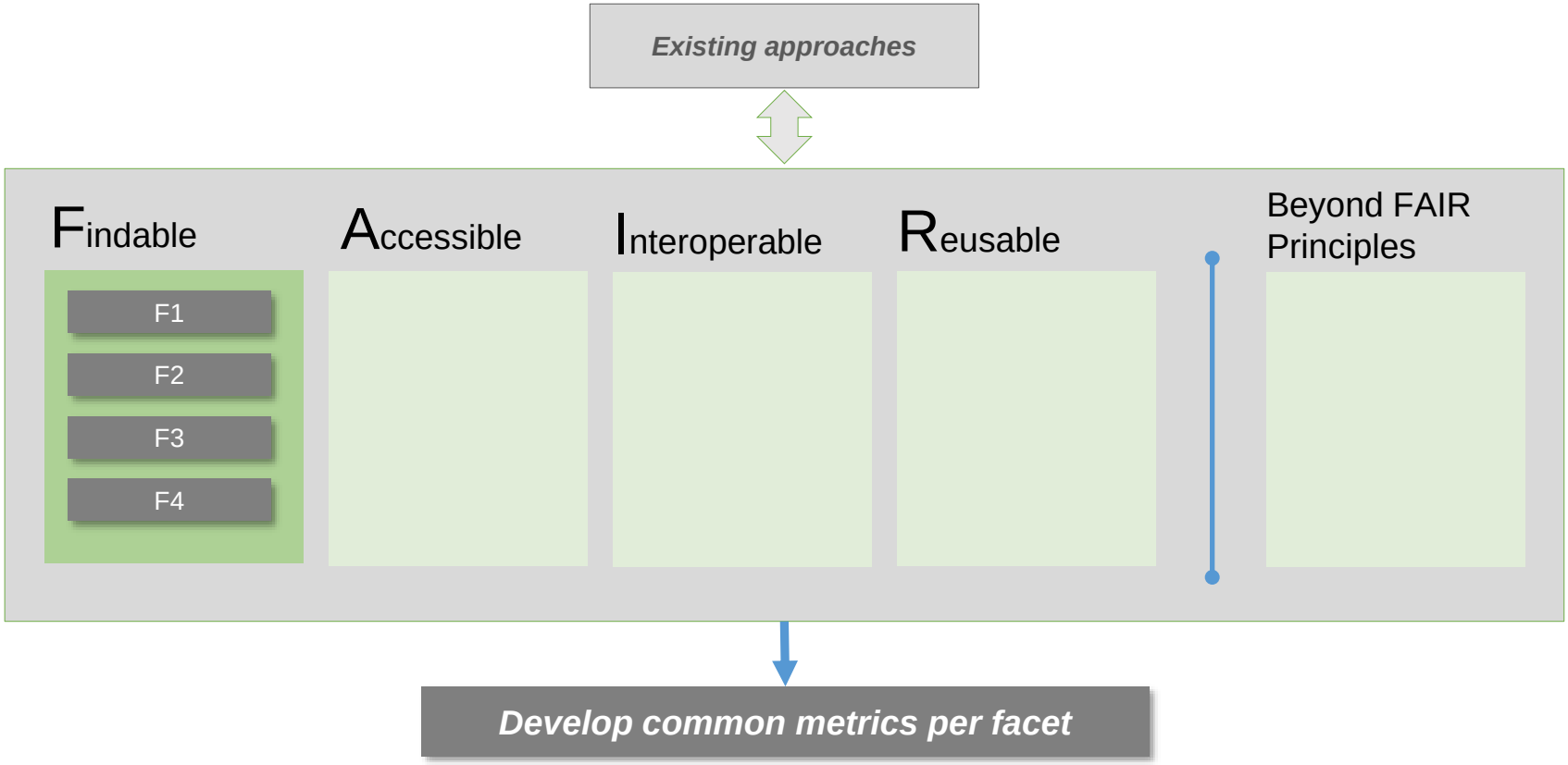
FAIR Principles

–

Findable

Analysis of existing approaches


Existing approaches
Findable
F1
F2
F3
F4
Accessible
Interoperable
Reusable
Beyond FAIR Principles
Develop common metrics per facet

FAIR 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

LEGEND

1. ANDS-NECTAR-RDS-FAIR data assessment tool
2. DANS-Fairdat
3. DANS-Fair enough?
4. The CSIRO 5-star Data Rating tool
5. FAIR Metrics Questionnaire
6. Big Data Readiness
7. FAIR Evaluator
8. Data Stewardship Wizard
9. Checklist for Evaluation of Dataset Fitness for Use
10. RDA-SHARC Evaluation

Principle

 # Facet

1 Question
- Option #1
- Option #2
- Option #3

Potential Overlap

F1 (meta)data is assigned a globally unique and eternally persistent identifier

1 Does the dataset have any identifiers assigned?
- No identifier
- Web adress (URL)
- Local identifier

2 Does the dataset have a persistent identifier (PID)?
- No
- Yes

3 Will your dataset have a Persistent Identifier after deposit?
- No
- Yes

4 Citable - denoted using a formal identifier
- Not citeable
- Local identifier
- Web address (URL - not guaranteed stable)
- Persistent web identifier (URI)

4 Citable - denoted using a formal identifier
- Not citeable
- Local identifier
- Web address (URL - not guaranteed stable)
- Persistent web identifier (URI)

5 Please provide the IRI for a registered identifier schema for your resource's IRI (e.g. DOI, HTTP)

5 Please provide the IRI to the document describing the persistence policy for the identifier of this (meta)data

7 Whether there is a scheme to uniquely identify the digital resource.

7 Whether there is a policy that describes what the provider will do in the event an identifier scheme becomes deprecated. *A2*

9 Citation exists, including authorship, year, comprehensive title, persistent identifier (e.g. DOI) *F2*

- No
- Somewhat
- Yes

9 Persistent identification of the dataset and related work (related literature and data, authors, projects, terms)

- No
- Somewhat
- Yes

9 Citation exists, including authorship, year, comprehensive title, persistent identifier (e.g. DOI) *R1.2*

- No
- Somewhat
- Yes

10 Are each data/dataset identified by an indexed and independant identifier ?

- Never /NA
- If mandatory
- Sometimes
- Always

10 Are the data identifiers unique, global and persistent ?

- Never /NA
- If mandatory
- Sometimes
- Always

10 Has any identifying schema been used for data (e.g. DOI)

- Never /NA
- If mandatory
- Sometimes
- Always

10 Are all datasets linked to an authority (legal entity) through a unique and persistent identifier over time (e.g. institution, association or established body)?

- Never /NA
- If mandatory
- Sometimes
- Always

10 Are the metadata of each dataset linked to a unique authority (responsible for the datasets at a given time)?

- Never /NA
- If mandatory
- Sometimes
- Always

F2 data is described with rich metadata

2 How is the data described with metadata? *R1*

- The data is not described
- Brief title and description
- Comprehensively, but in a text-based, non standard format
- Comprehensively, using a recognized formal machine readable metadata schema

2 Please provide the IRI to a document that contains machine-readable metadata for the digital resource

2 Are the metadata accessible? *F4, A2*

- No
- Yes

4 Described - tagged with metadata ?

- No metadata
- Abstract and keywords
- Basic metadata (e.g. Dublin Core)
- Specialized metadata (e.g. Darwin Core, ISO 19115/19139, schema.org scientific data profile)
- Rich metadata using multiple standard RDF vocabularies (e.g. DCAT, PROV, ADMS, GeoDCAT, FOAF, ORG, GeoSPARQL)

7 The availability of machine-readable metadata that describes a digital resource.

9 Dataset is provided in a widely-used or community-accepted machine-readable format and using standard terminologies for nominal data and available

- No
- Somewhat
- Yes

10 Are the types and formats of data generated / collected well described?

- Never /NA
- If mandatory
- Sometimes
- Always

3 Did you provide rich additional documentation?

- No
- Yes

9 Description of methods used to create this dataset are appropriate for the context and discipline

- No
- Somewhat
- Yes

10 Does the researcher use efficient and rich services to access data
(various formats, visualisations, practical tools and systems adapted to different types of use and users)

- Never /NA
- If mandatory
- Sometimes
- Always

F3 metadata clearly and explicitly includes the identifier of the data it describes

1 Is the dataset identifier included in all metadata records/files describing the data?

- No
- Yes

5 Please provide the IRI of the metadata

7 Whether the metadata document contains the globally unique and persistent identifier for the digital resource

10 Are the metadata linked to the dataset through a persistent identifier?

- Never /NA
- If mandatory
- Sometimes
- Always

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

1 How accessible is the data?

- No access to metadata or data
- Access to metadata only
- Unspecified conditional access (e.g contact the data custodian for access)
- Embargoed access after a specified date
- A de -identified / modified subset of the data is publicly accessible
- Publicly accessible
- Fully accessible to persons who meet explicity stated questions (e.g. ethics approval for sensitive data)

2 Are the metadata accessible?

- No
- Yes

2 Is the dataset available for public access? (i.e. the restriction is only registration on a website before the person has access the data) *A1.1*

- No
- Yes

3 Is the metadata publicly accessible?

- No
- Yes

4 Findable - Indexed in a discovery system

- No
- Local or internal system only
- Community wide or jurisdictional system
- Highly ranked in general purpose index (Google, Bing etc)

5 Please provide the URL to a search engine, and the query that will be executed to discover your RESOURCE ID

7 The degree to which the digital resource can be found using web-based search engines