



RESEARCH DATA ALLIANCE

FAIR Data Maturity Model

Workshop #2 [RDA 13th Plenary session]

3rd April 2019

Agenda

- › Welcome, objectives of the meeting 5'
- › Round table 5'
- › Approval of methodology and scope 10'
- › Report on discussions on GitHub 10'
- › Hands-on exercise 45'
- › Action items and next steps 15'

Case statement of the WG

Challenge

- › Ambiguity and wide range of interpretations of FAIRness
- › Lack of a common set of core assessment criteria and a minimum set of shared guidelines

Approach

- › Bring together stakeholders
- › Build on existing approaches and expertise

Intended results

- › RDA Recommendation of core assessment criteria
- › Generic and expandable self-assessment model
- › Self-assessment toolset
- › FAIR data checklist

Case statement of the WG

Target audiences

- › Researchers, data stewards, other data professionals
- › Data service owners, e.g. infrastructure, repositories
- › Organisations that manage research data
- › Policymakers

Connections

- › RDA Disciplinary Framework Interest Group
- › RDA Domain Repositories Interest Group
- › Other RDA groups

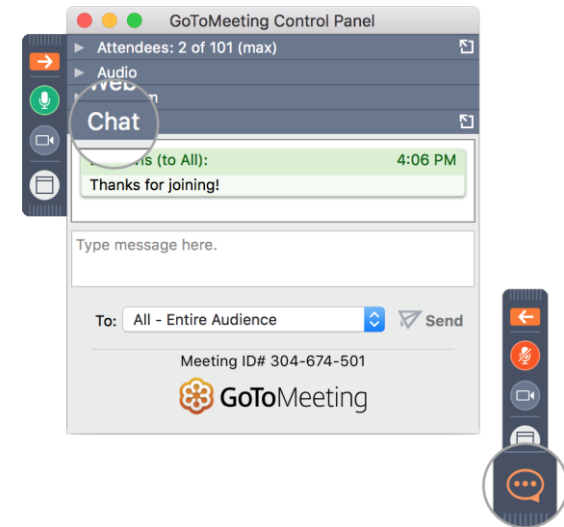
Scope of the assessment

- › Datasets
- › Data-related aspects (e.g. algorithms, tools, workflows)

Roundtable

- › Which region?
- › Your role
 - › Researcher
 - › Librarian
 - › Infrastructure manager
 - › Policy developer
 - › Research funder
- › Introducing the editorial team
- › First adopters: EOSC & European Commission
- › Owners of FAIR assessment methodologies

If you are dialling in,
please type your name and affiliation in the chat window





WG

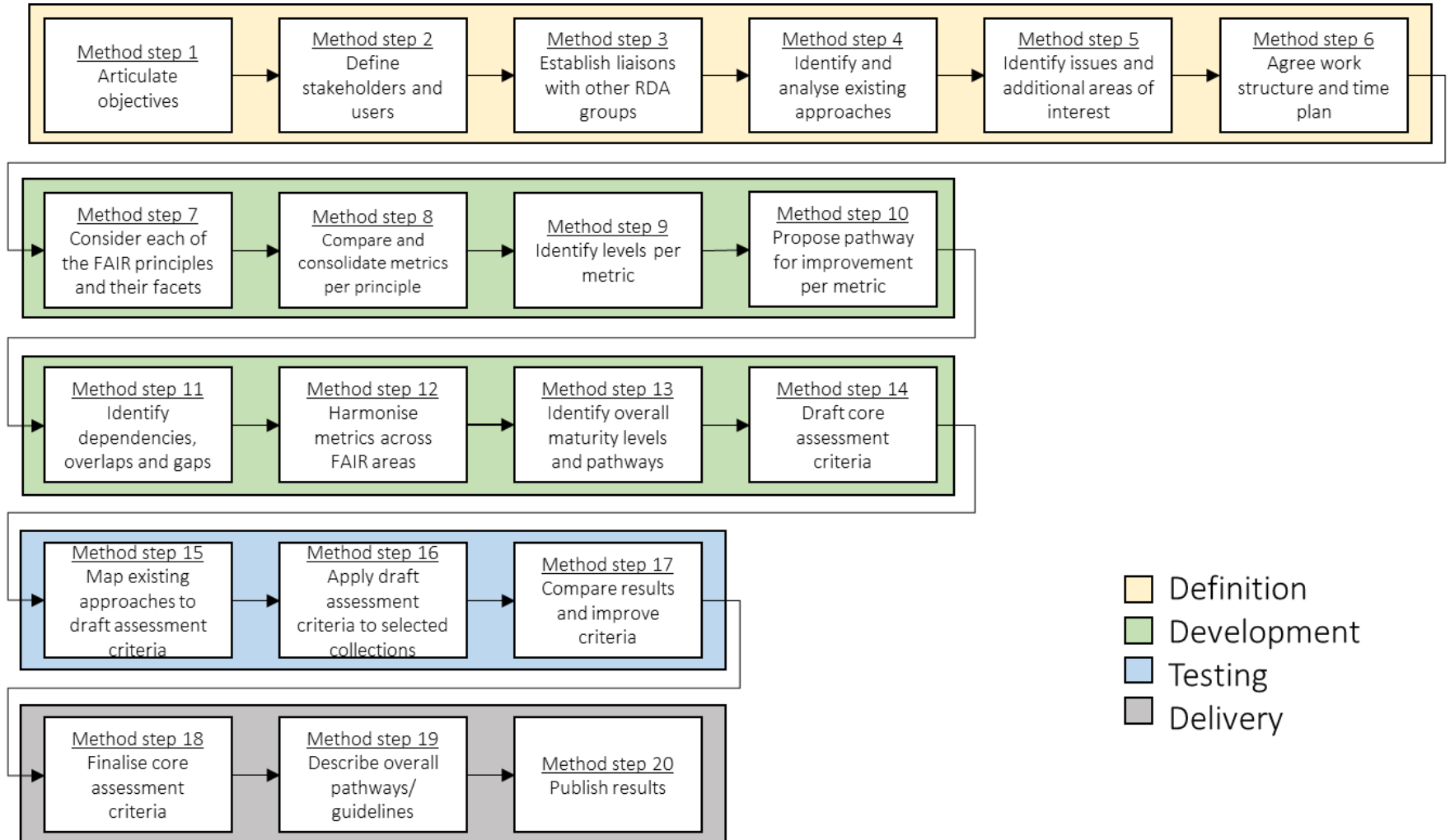
methodology,
timeline &
scope

Proposed development methodology

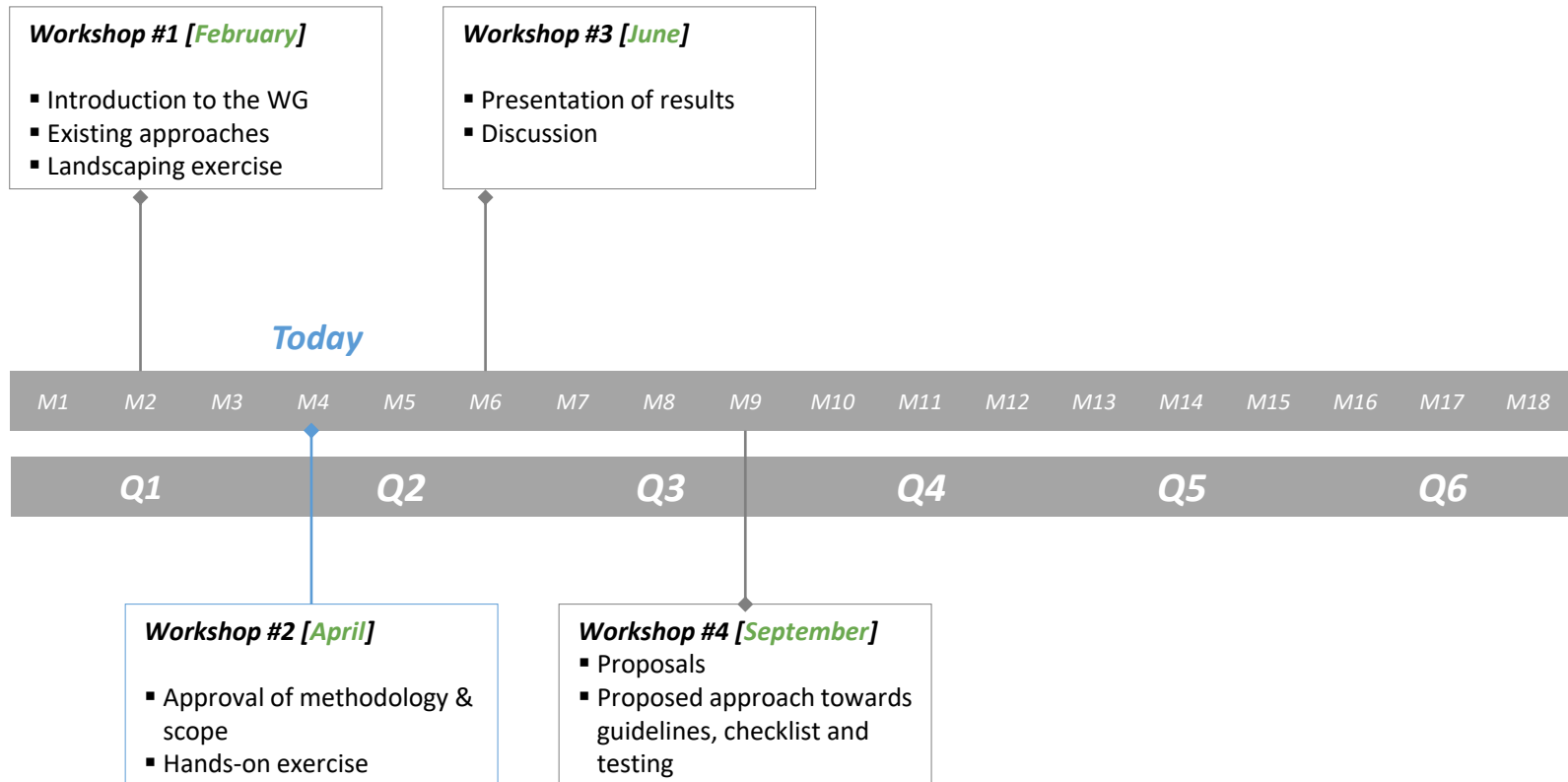
Bottom-up approach comprising 4 phases

- Definition
- Development
 - Assessment of the four FAIR principles in four 'strands'
 - Fifth 'strand': beyond the FAIR principles
- Testing
- Delivery

Overview of the methodology



Proposed timeline



Proposed scope

	Proposed resolutions
ENTITY	<u>Dataset</u> and <u>data-related aspects</u> (e.g. algorithms, tools and workflows)
NATURE	<u>Generic</u> assessment (i.e. cross-disciplines)
FORMAT	<u>Manual</u> assessment
TIME	<u>Periodically</u> throughout the lifecycle of the data
RESPONDENT	People with <u>data literacy</u> (e.g. researchers, data librarians, data stewards)
AUDIENCE	Researchers, data stewards, data professionals, data service owners, organisations involved in research data and policy makers

Overview of discussions on GitHub

Findable: What does it mean?

[\[GitHub\]](#)

- Human Findable
- Machine Findable
- Meaning of 'rich metadata'

'Flows' beyond the FAIR assessment

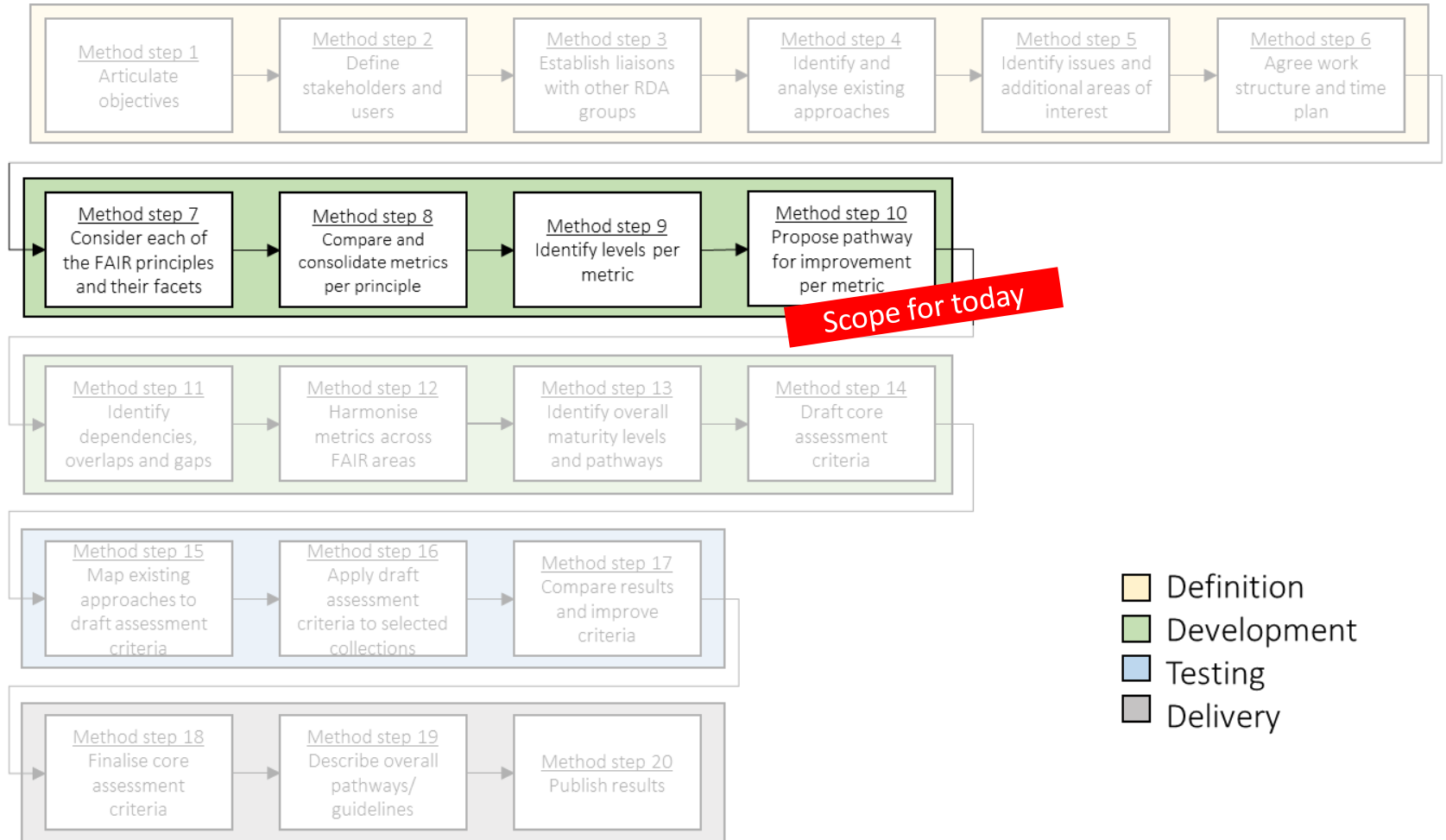
[\[GitHub\]](#)

- Data flow
- Data flow legal issues
- People flow
- Financial flow
- Hardware infrastructure



Hands-on Exercise

Scope of the exercise



- Definition
- Development
- Testing
- Delivery

Scope of the exercise

- Analysis of one of the FAIR principles

- Findable

- Accessible

- Interoperable

- Reusable

- R1

- **R1.1 (Meta)data are released with a clear and accessible data usage licence**

- R1.2

- R1.3

Method step 7

- Comparison and consolidation of the metric

Method step 8

- Identification of levels per metric

Method step 9

- Pathways for increasing maturity per metric

Method step 10

Deep dive into R1.1

“ [...] R1.1 is about **legal interoperability**. What **usage rights** do you attach to your data? This should be described clearly. Ambiguity could severely limit the reuse of your data by organisations that struggle to comply with licensing **restrictions**. **Clarity of licensing** status will become more important with automated searches involving more licensing considerations. The conditions under which the data can be used should be clear to machines and humans.

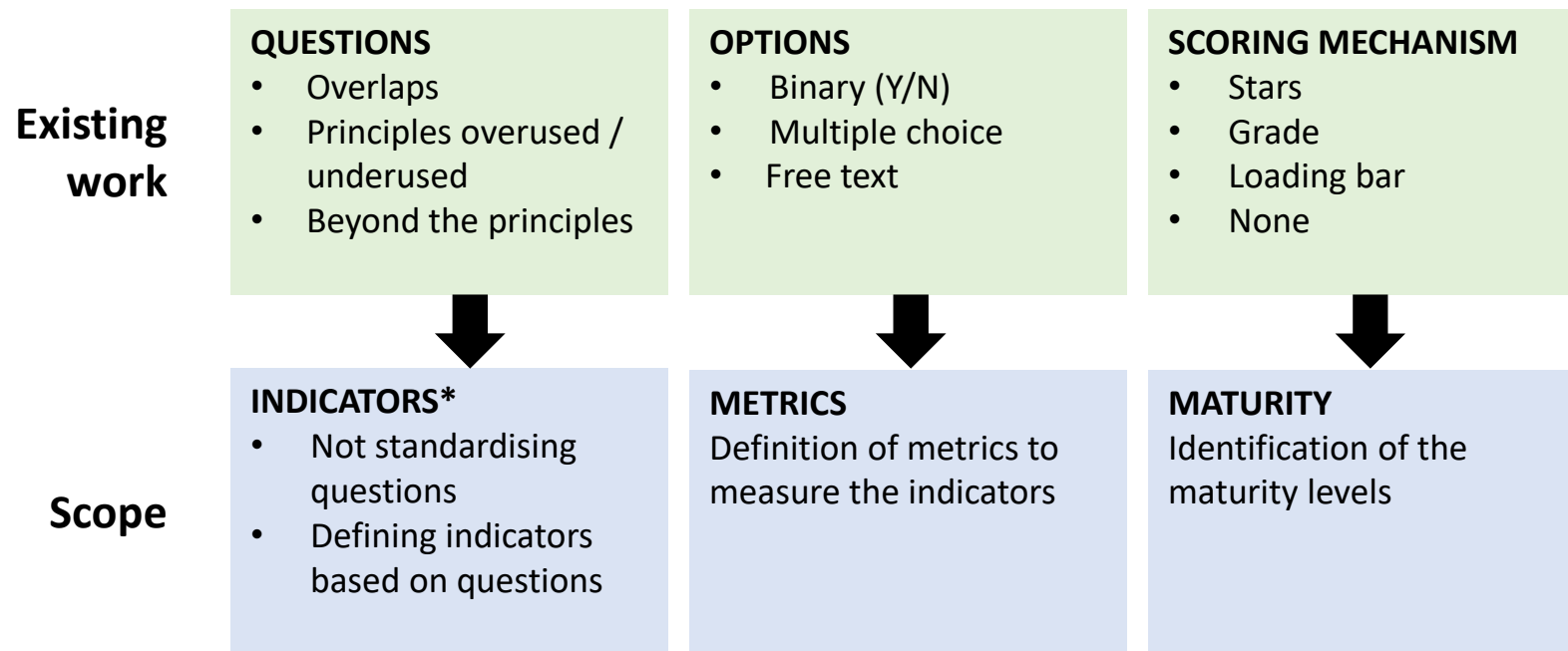
GO-FAIR [definition](#) of R1.1

How do existing methodologies assess R1.1

R1.1 (meta)data is released with a clear and accessible data usage licence

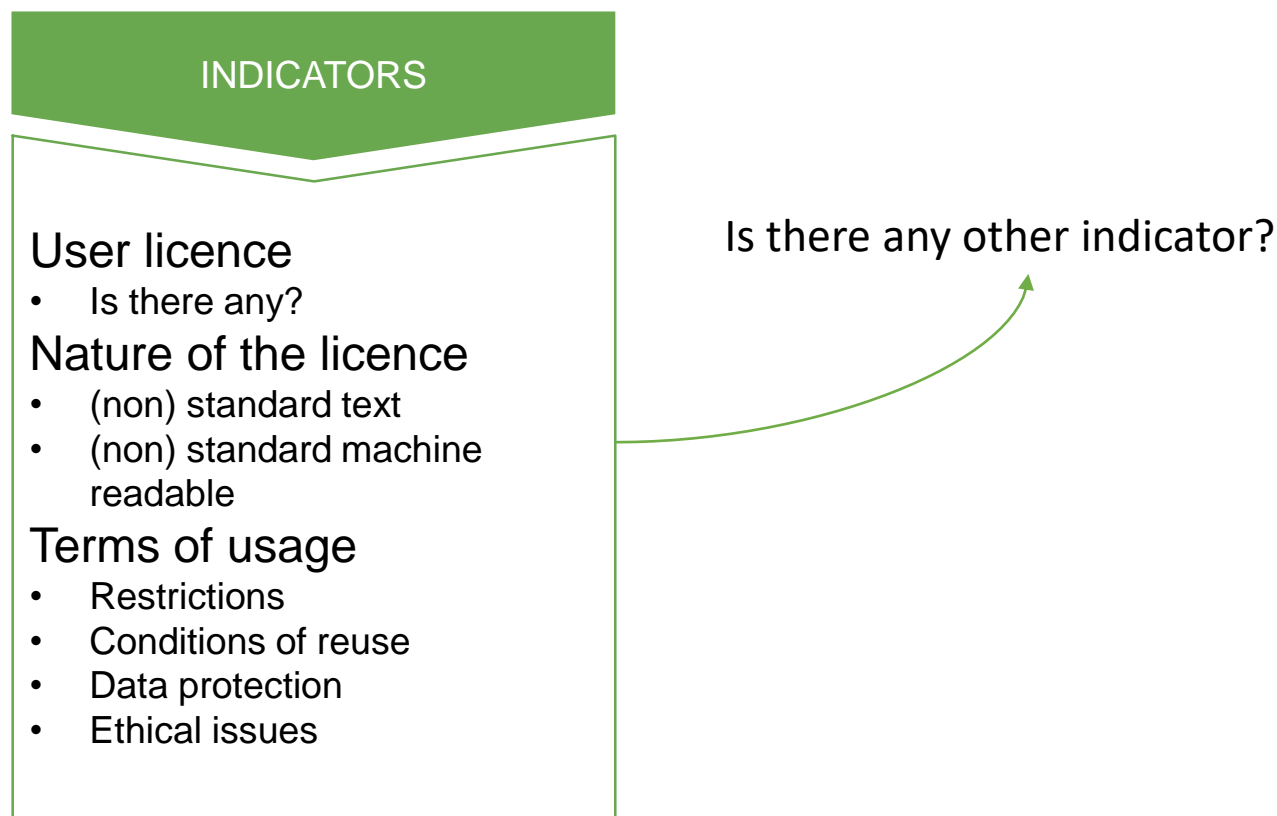
- 1 Which of the following best describes the license/usage rights attached to the data?
 - No licence
 - Non standard text based licence
 - Non standard machine readable licence (e.g. clearly indicating under what conditions the data may be used)
 - Standard text based licence
 - Standard machine-readable licence (e.g. Creative Commons)
- 2 Does the user license have any user restrictions for accessing the data?
 - No
 - Yes
- 2 Does the dataset have a user license?
 - No
 - Yes
- 3 Does the dataset have a usage licence?
 - No
 - Yes
- 4 Licensed - conditions for re-use are available and clearly expressed
 - No
 - License described in text
 - Link to a standard license (e.g. Creative Commons)
- 5 Please provide the IRI for you usage license regarding the content returned from RESOURCE ID
- 7 The existence of a license document, for BOTH (independently) the data and its associated metadata, and the ability to retrieve those documents
- 9 Terms of usage (licenses, other conditions of reuse, data protection, ethical issues)
 - No
 - Somewhat
 - Yes

from the [landscaping exercise](#)



* an indicator can be seen as a component of a Principle (e.g. F1, R1)

R1.1 (Meta)data are released with a clear and accessible data usage licence



Measuring maturity

Alternative #1

Five options for R1.1 [metadata/data]

- Level 0: no licence
- Level 1: non standard licence in a human-readable format allowing access
- Level 2: standard licence in a human-readable format allowing access
- Level 3: standard open licence in a human-readable format allowing reuse
- Level 4: standard open licence in a machine-readable format allowing reuse
- Level 5: standard open licence in a machine-readable format with clear criteria allowing reuse

Each option is defining a maturity level

FAIRness compliance for R1.1



Measuring maturity Alternative #2

Two options per indicator for R1.1 [metadata/data]

- > (1) Is there a *licence*? [YES/NO]
- > (2) Is the licence a *standard* licence? [YES/NO]
- > (3) Is the licence *machine* readable? [YES/NO]
- > (4) Does the licence permit *access*? [YES/NO]
- > (5) Does the licence permit *reuse*? [YES/NO]

The options aggregated correspond to a maturity level



FAIRness compliance for R1.1

Measuring maturity

Your feedback

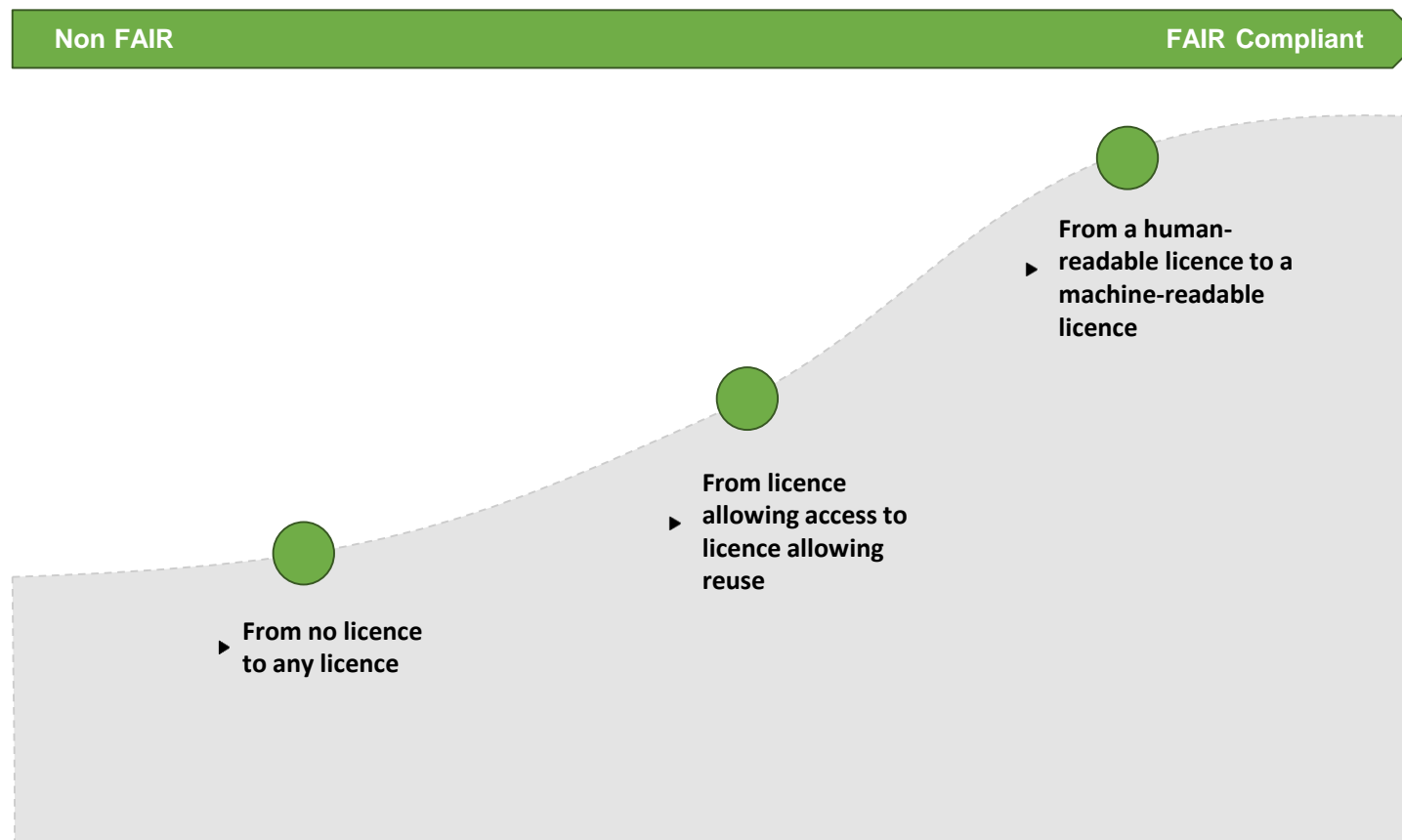
Do you agree with the proposed alternatives?

Which alternative is more suitable?

Do you think of other alternatives?

...

Pathways for increasing maturity





Actions items & next steps

Discussion

- **Nature** of RDA recommendations & outputs
- How to keep you involved?

Action items

- › Call for volunteers
- › Development of the core assessment criteria on [GitHub](#)
 - › Analysis of all the FAIR principles Method step 7
 - › FAIR – Findable [\[Link\]](#)
 - › FAIR – Accessible [\[Link\]](#)
 - › FAIR – Interoperable [\[Link\]](#)
 - › FAIR – Reusable [\[Link\]](#) Method step 8
 - › Comparison and consolidation of the metrics per principle Method step 8
 - › Identification of levels per metric Method step 9
 - › Pathways of improvement per metric Method step 10
- › Online workshop #3
 - › at 09:00 CEST on the 18 June 2019
 - › at 17:00 CEST on the 18 June 2019

Resources

- › RDA FAIR data maturity model WG

<https://www.rd-alliance.org/groups/fair-data-maturity-model-wg>

- › RDA FAIR data maturity model WG – **Case Statement**

<https://www.rd-alliance.org/group/fair-data-maturity-model-wg/case-statement/fair-data-maturity-model-wg-case-statement>

- › RDA FAIR data maturity model WG – **GitHub**

<https://github.com/RDA-FAIR/FAIR-data-maturity-model-WG>

- › RDA FAIR data maturity model WG – **Mailing list**

fair_maturity@rda-groups.org



Thank you!