



RESEARCH DATA ALLIANCE

FAIR Data Maturity Model

Workshop #8

24th March 2020

Agenda

| | |
|-----|---|
| 5' | Welcome, objectives of the meeting |
| 5' | Roundtable |
| 10' | State of play |
| 30' | General discussion (part I) – strategic to practical issues |
| 30' | General discussion (part II) – indicators revision |
| 10' | Action items and next steps |

Context

The principles are **NOT** strict

- **Ambiguity**
- Wide range of **interpretations** of FAIRness



Different **FAIR Assessment** Frameworks

- Different metrics
- No comparison of results
- No benchmark

FAIR

SOLUTION is to bring together **stakeholders** to build on **existing approaches** and **expertise**

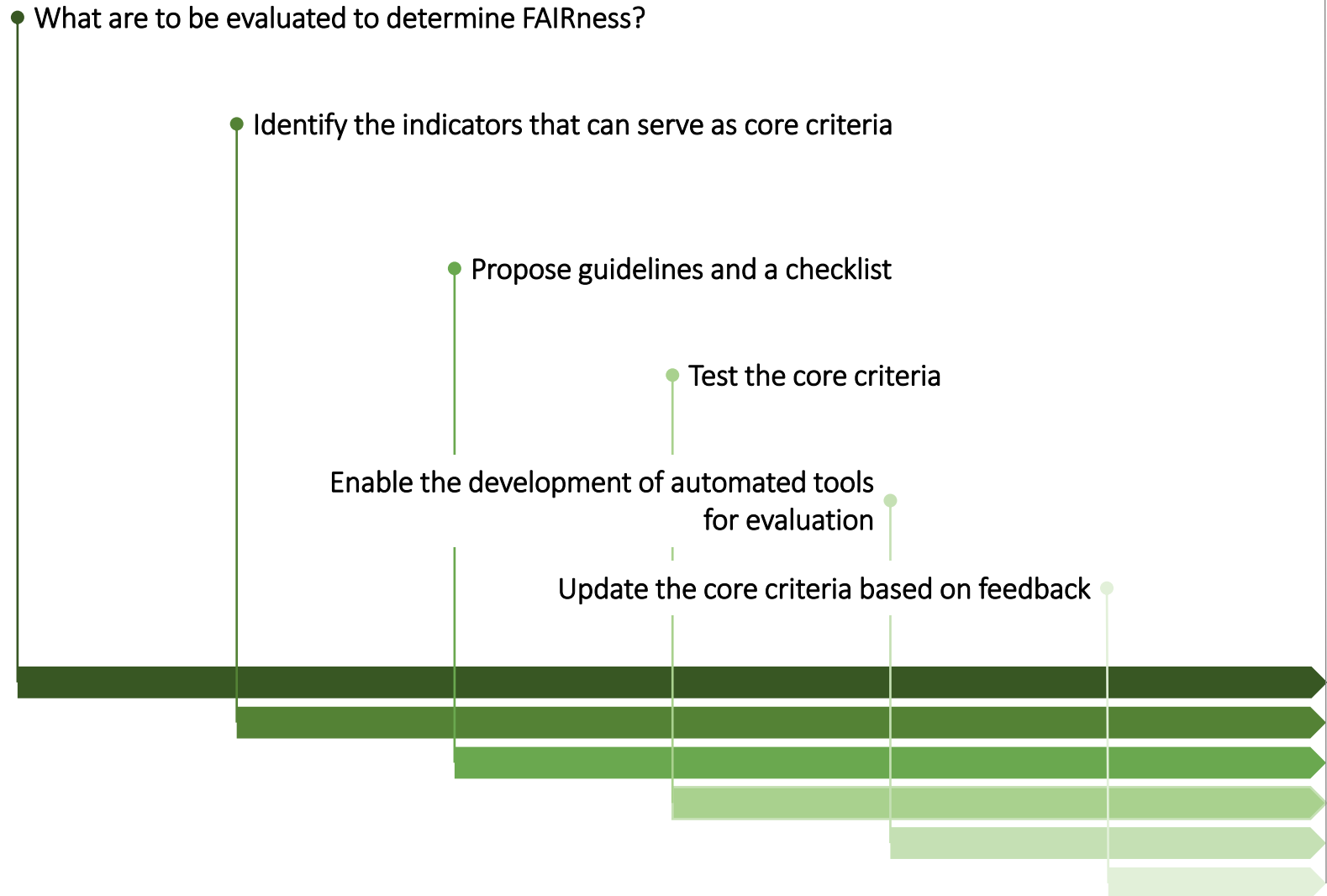


- Set of **core assessment criteria** for FAIRness
- FAIR **data maturity model & toolset**
- FAIR data **checklist**
- RDA recommendation

Join the **RDA Working Group**: [RDA WG web page](#) | [GitHub](#)

Objectives

FAIR data maturity model



Scope

BUT the Working Group does NOT have the purpose to ...

- ⊗ **develop yet-another-evaluation-method:** the core criteria are intended to provide a common ‘language’ across evaluation approaches, not to be applied directly to datasets.
- ⊗ **define how the core criteria need to be evaluated.** The exact way to evaluate data based on the core criteria is up to the owners of the evaluation approaches, taking into account the requirements of their community
- ⊗ **revise and re-design the FAIR principles**

Roundtable

In the chat window, please type...

- > Your name
- > Your affiliation
- > Your role
 - > Researcher
 - > Librarian
 - > Service provider
 - > Policy maker
 - > Funder

- > Introducing the editorial team

Join at slido.com
event code # 6405

Which topic
would you like to
focus on most
today?

[ACCESS THE POLL](#)





State of play

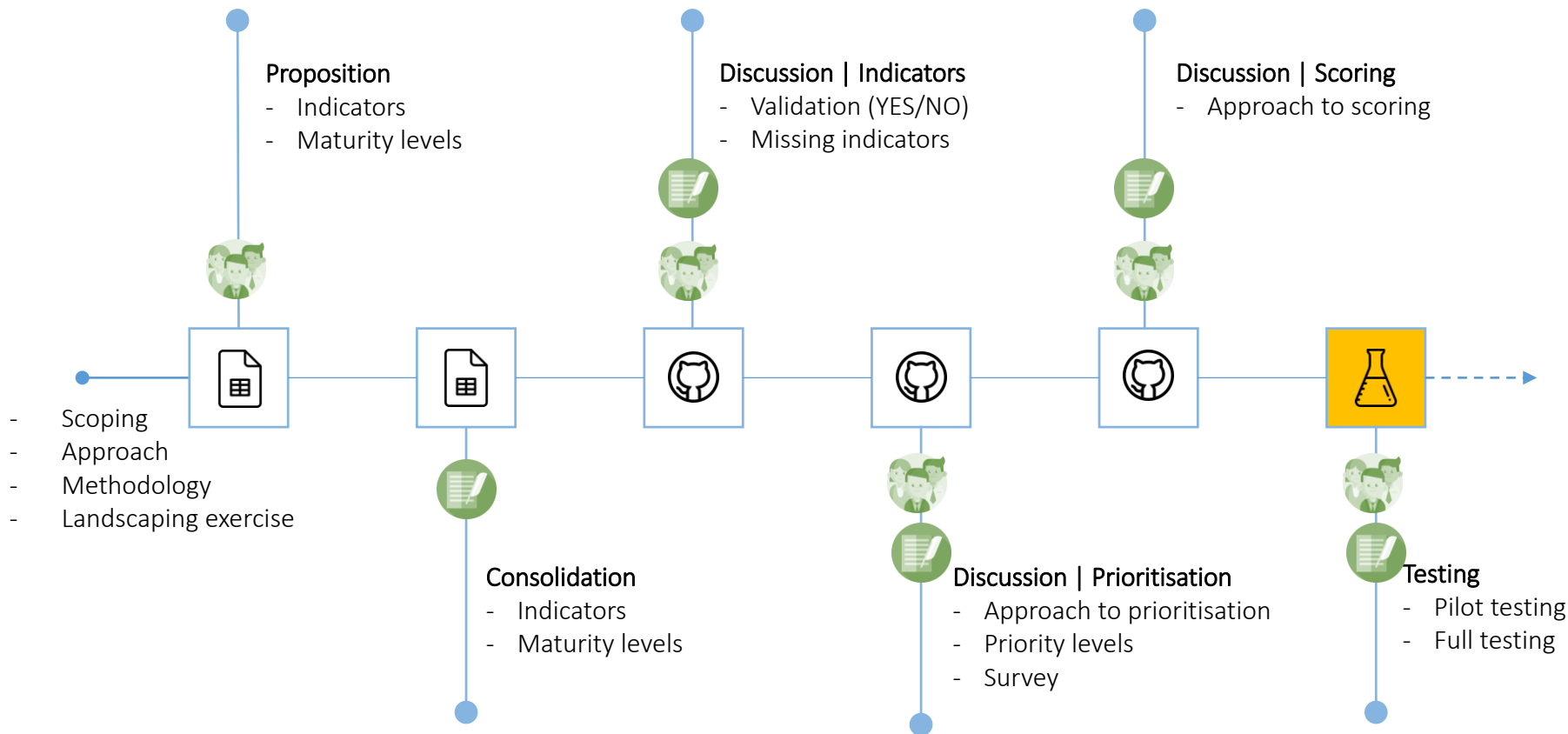
State of play

| | |
|------------------|----------|
| 1. Definition | DONE |
| 2. Development | DONE |
| i) First phase | DONE |
| ii) Second phase | DONE |
| 3. Testing | CLOSING |
| 4. Delivery | STARTING |

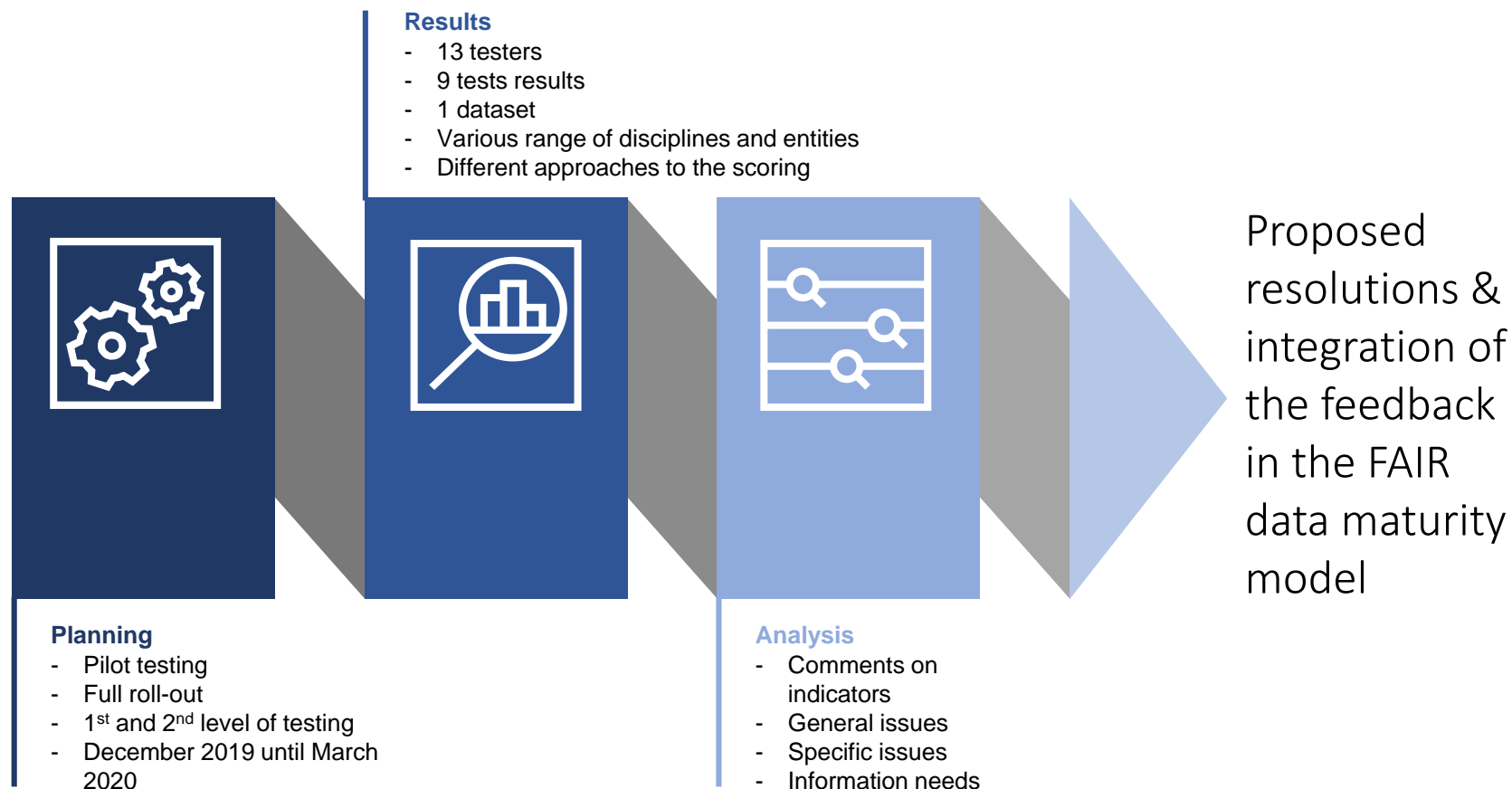
* Any comments are still welcomed with regards to the output produced during the first phase | [GitHub](#)

State of play

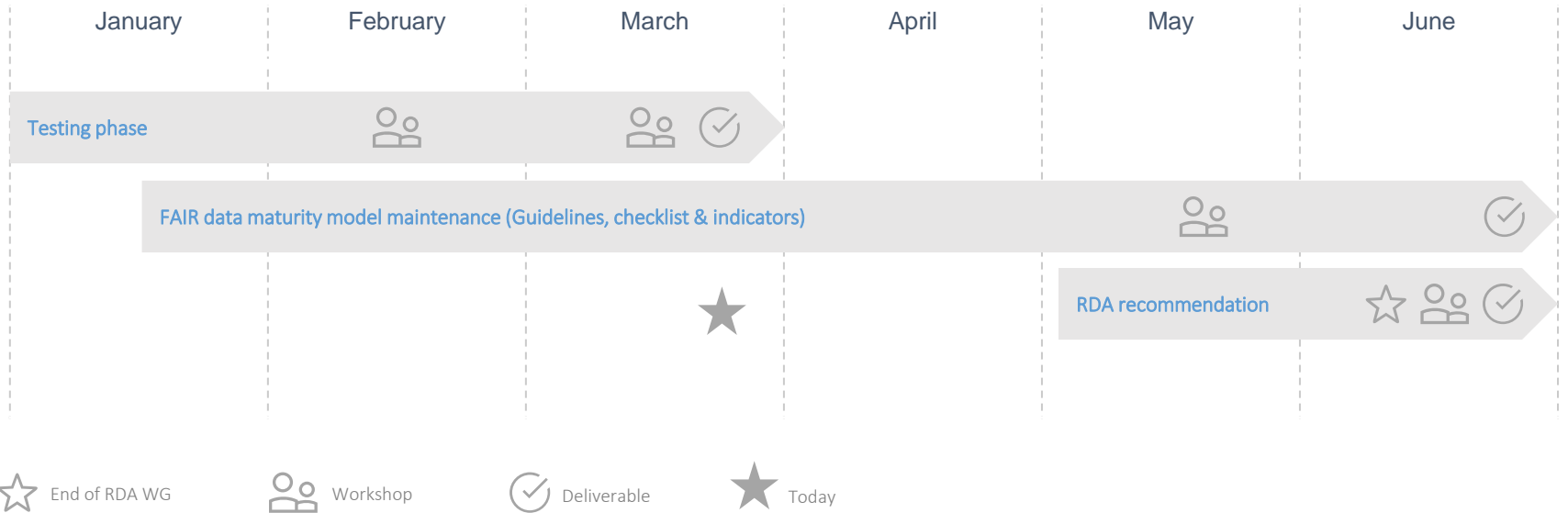
 Editorial team
 Working group



Testing phase | Overview & results

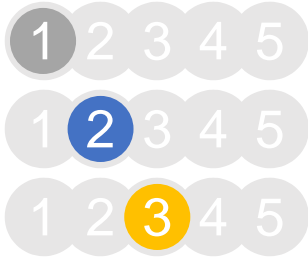


2020





5-level scale per indicator



- Five levels of compliance
- Per indicator – aggregated per FAIR area
- Non applicable or consideration/implementation as options
- Useful for giving credit for evolution and helping people to improve

SELECTED

FAIRNESS per area

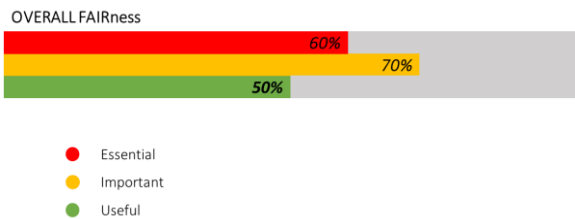
| | Essential | Important | Useful |
|---------|-----------|-----------|--------|
| Level 0 | ○ | | |
| Level 1 | ● | | |
| Level 2 | ● | ◐ | |
| Level 3 | ● | ● | |
| Level 4 | ● | ● | ◐ |
| Level 5 | ● | ● | ● |

- None of the indicators are satisfied
- ◐ Half of the indicators are satisfied
- All indicators are satisfied

- Measurement based on priorities
- Per indicator – aggregated per FAIR area
- Score determined based on the compliance to priorities
- Provides a ‘measure of FAIRness’

SELECTED

Overall FAIRNESS

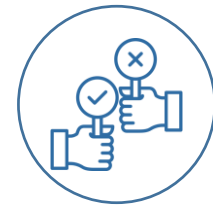


- Essential
- Important
- Useful

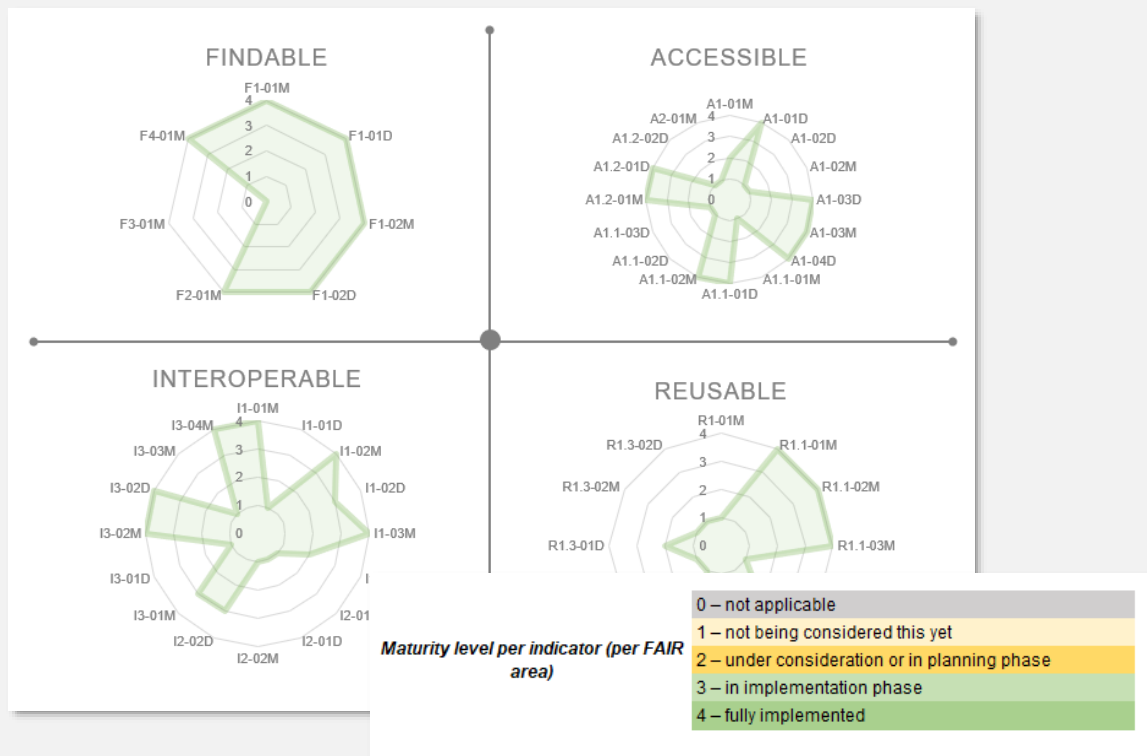
- Measurement based on priorities
- Per indicator – overall score
- Aggregated score
- Provides a quick view of how priorities are met -- but does not give detailed view

DISCARDED

Scoring mechanism | Proposition – 1



Two-layer scoring mechanism – *five-level scale per indicator*



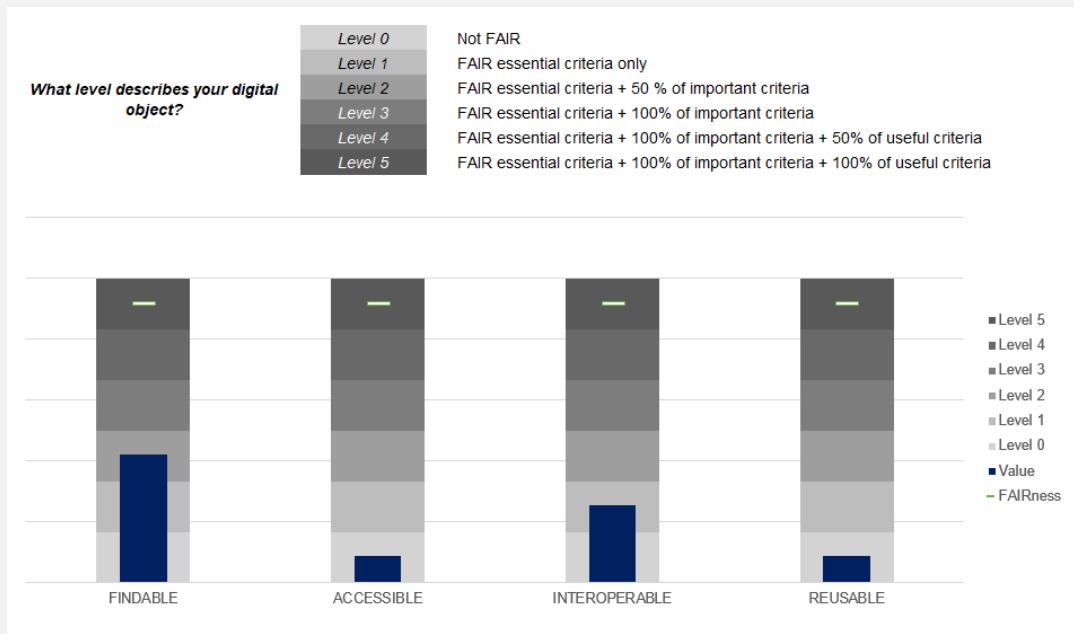
1. Useful for giving credit for evolution and helping people to improve
2. Non applicable or consideration/implementation as options

[DOWNLOAD](#)

Scoring mechanism | Proposition – 2



Two-layer scoring mechanism – *binary evaluation summarized per FAIR area*



1. Five levels determining how far along the journey towards FAIRness a resource under evaluation is
2. Binary evaluation of each indicator based on priorities (i.e. compliance to the indicator)

[DOWNLOAD](#)



Discussion items

FDMM | General discussion

Strategic issue

What kinds of metadata models do disciplines have (infrastructure, repository, collection, individual datasets, separate/embedded) and how does that match the view expressed in the FAIR principles (metadata as a digital object in itself)

How can communities come together to coordinate FAIR approaches and targets, and who would be in the best position to do this (e.g. in communities represented in the meeting)

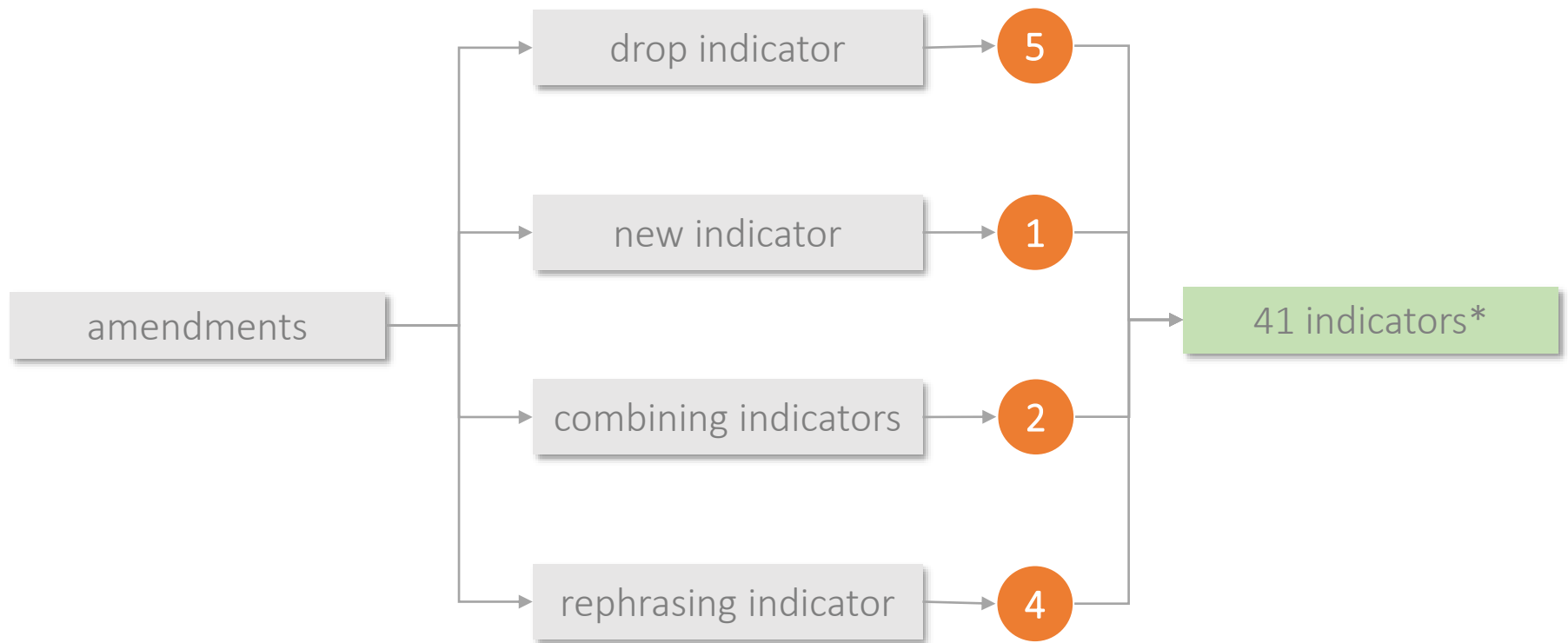
Practical issue

How do funders intend to apply FAIR evaluation? What do they see as the best way – the hard binary way, or the soft, progress-testing way. Or if there are not enough funders in the room, how would research data providers want funders to apply evaluations? Should results be public or confidential? How can they alleviate fear linked to FAIR evaluation?

How do the participants (funders and data providers) see the longer-term future of FAIR in achieving more and better reuse of research data? Are there aspects that are not covered that might be added in the future? Are there aspects that should be suppressed?

Indicators | amendments

Unless otherwise agreed, the **proposed changes** – derived from the testing phase – will be **put into action** [end of April]



* after revision, the final list of indicators will be composed of 41 indicators

Indicators | amendments – 1



If there are no objections, the **proposed changes** – derived from the testing phase – will be **put into action** [early April]

drop indicator

out of scope

[A1.1-02M+D] indicator tests for open-source protocol, which is not mentioned in FAIR principle A1.1.

[I2-01M+D] indicator tests for use of standard vocabularies, but principle I2 only refers to FAIR vocabularies and says nothing about 'standard'

[R1.1-03M] indicator tests that licence information is in the 'right' element, but this is basically a quality aspect; we don't test this for other metadata requirements either

[R1.1-05M] indicator tests information on consent for personal data; this is not mentioned in principle R1.1 which is about licensing.

[I1-03M+D] Principle I1 does not mention 'self-describing'. Furthermore, it has been noted that the term itself is not entirely clear and could be seen to be very close to the representation being machine-understandable.

Indicators | amendments – 2



If there are no objections, the **proposed changes** – derived from the testing phase – will be **put into action** [early April]

new indicator

[A1-0xM] new indicator alongside A1-01D for manual access to metadata; it was pointed out that there is no indicator for manual access to metadata while there is an indicator for manual access to data (A1-01D Data can be accessed manually (i.e. with human intervention))

Indicators | amendments – 3



If there are no objections, the **proposed changes** – derived from the testing phase – will be **put into action** [early April]

combining indicators

[A1-01M] merge closely related indicators A1-01M, A1.1-03D and A1.2-01M into A1-01M with definition “Metadata contains information to enable the user to get access to the data”; in guidelines include that this can be information about access control, e.g. need to register or provide username/password.

[A1.2-01D, A1.2-02D] Merge these two: having separate indicators for authentication and authorisation puts undue emphasis on these aspects; they will always be evaluated together.

Indicators | amendments – 4



If there are no objections, the **proposed changes** – derived from the testing phase – will be **put into action** [early April]

rephrase indicator

[R1-01M] align with wording of principle R1; not ‘sufficient metadata’ but ‘plurality of accurate and relevant attributes’.

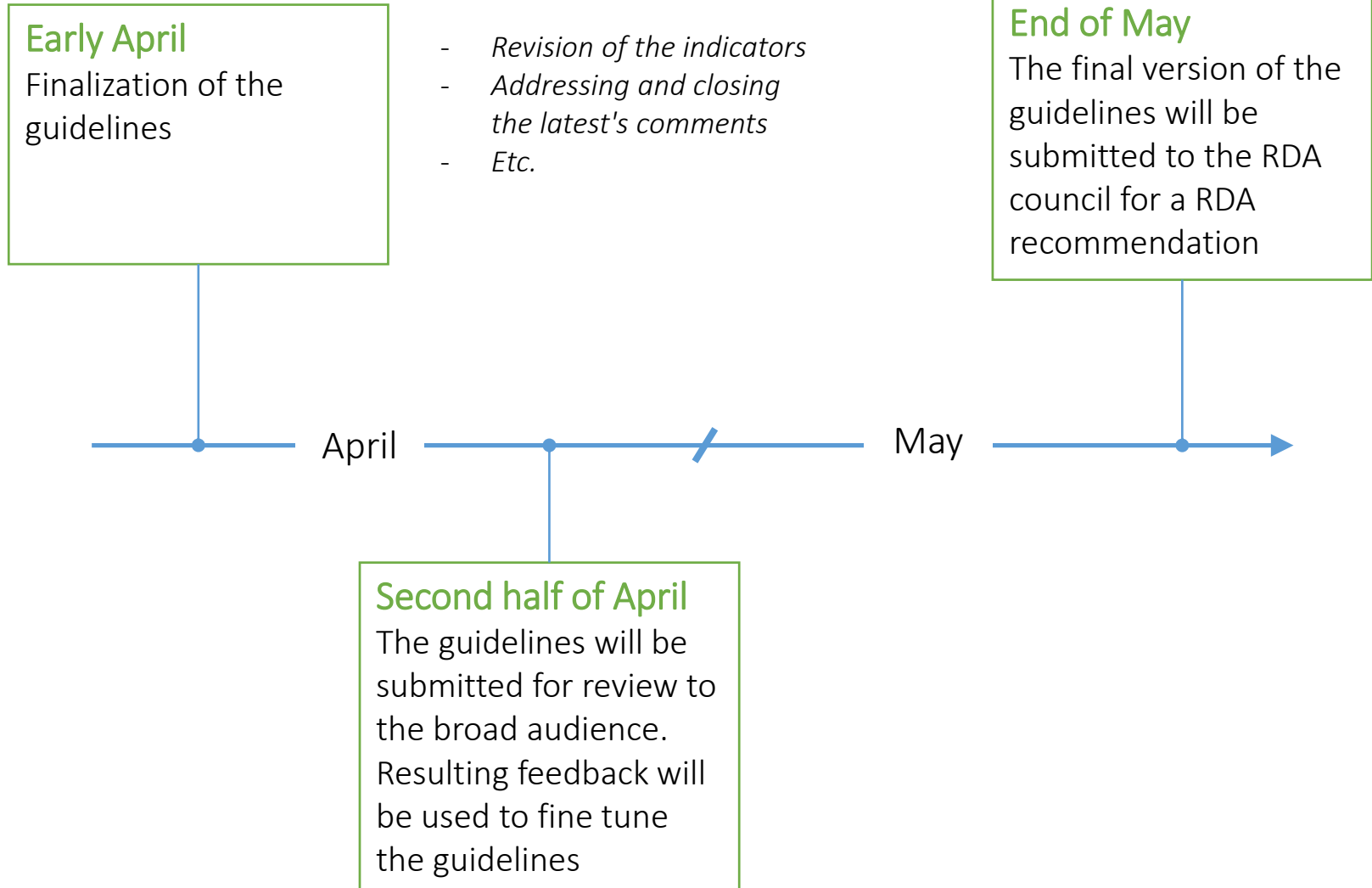
[F2-01M] delete the mention of standard in indicator for F2 – the standard aspect is covered in R1.3 – the indicators in F2 and R1 will just focus on the amount of metadata, and R1.3 will test that metadata is standard.

[F1-02M+D] use ‘globally unique’ instead of ‘universally unique’ – to align with the principle and because it might be confusing as UUID has a very specific meaning.

[I3-02D, I3-03M, I3-04M] drop ‘sufficiently’ qualified – that’s not in the principle, only ‘qualified’



Action items & Next steps



RDA FAIR DMM WG | Future outlook

Working Group

RDA FAIR data maturity
model Working Group
→
RDA recommendation

Interest Group

with different aim and
possibly broader (*i.e.*
platform to maintain and
agree indicators and services
to support FAIR data)



Turning the **Working Group**
into an **Interest Group**

Action item and next steps

Working Group members are invited to:

- Share feedback, comments & suggestions – on the [Guidelines](#)
- Contribute to GitHub discussion on [revising the indicators](#)
- Contribute to GitHub discussion on [scoring](#)
- Contribute to GitHub discussion on turning the WG into an IG (i.e. what do you think the scope and remit of this IG should be?) – *to be created*

WORKSHOP #9

20 May 2020

15.00 to 16.30 UTC

21 May 2020

07.00 to 08.30 UTC

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 – **Collaborative document**

https://docs.google.com/spreadsheets/d/1gvMfbw46oV1idztsr586aG6-teSn2cPWe_RJZG0U4Hg/edit#gid=0

- RDA FAIR data maturity model WG – **Indicators prioritisation**

<https://docs.google.com/spreadsheets/d/1mkjElFrTBPBH0QViODexNur0xNGhJqau0zkL4w8RRAw/edit>

- RDA FAIR data maturity model WG – **Indicators prioritisation survey results**

https://drive.google.com/open?id=11hyAYCKz_NV0Ob9-vlPqjN9LCarOFmc3

- RDA FAIR data maturity model WG – **Guidelines**

https://docs.google.com/document/d/1pDGGL3-BbBJu18KlfZUI3AizKLHXGXdli_mPtpEWmeg/

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

fair_maturity@rda-groups.org



Thank you!