



RESEARCH DATA ALLIANCE

# FAIR Data Maturity Model

Workshop #3

18th June 2019

# Agenda

1. Welcome, objectives of the meeting
2. Roundtable
3. State of play
4. Development | First phase  
Presentation of the work conducted & **discussion**
5. Development | Second phase  
Presentation of an approach & **discussion**
6. Action items and next steps

# Welcome, objectives of the meeting



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**



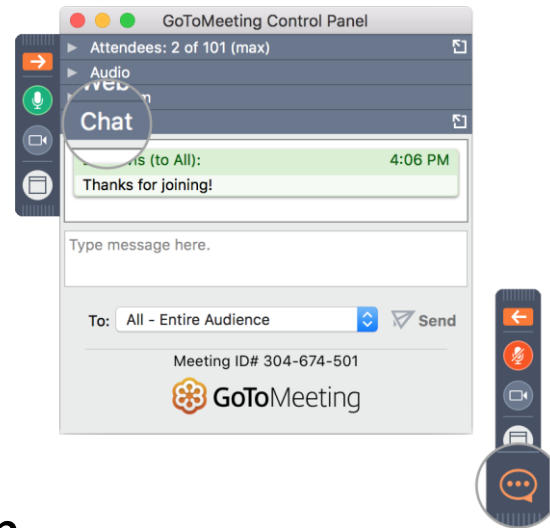
- Set of **core assessment criteria** for FAIRness
- FAIR **data maturity model & toolset**
- RDA recommendation
- FAIR data **checklist**

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

# Roundtable

Please type your name and affiliation in the chat window

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





# State of play

# State of play

1. Definition

DONE

2. Development

ONGOING

i) First phase

ONGOING

ii) Second phase

TO BE COMMENCED

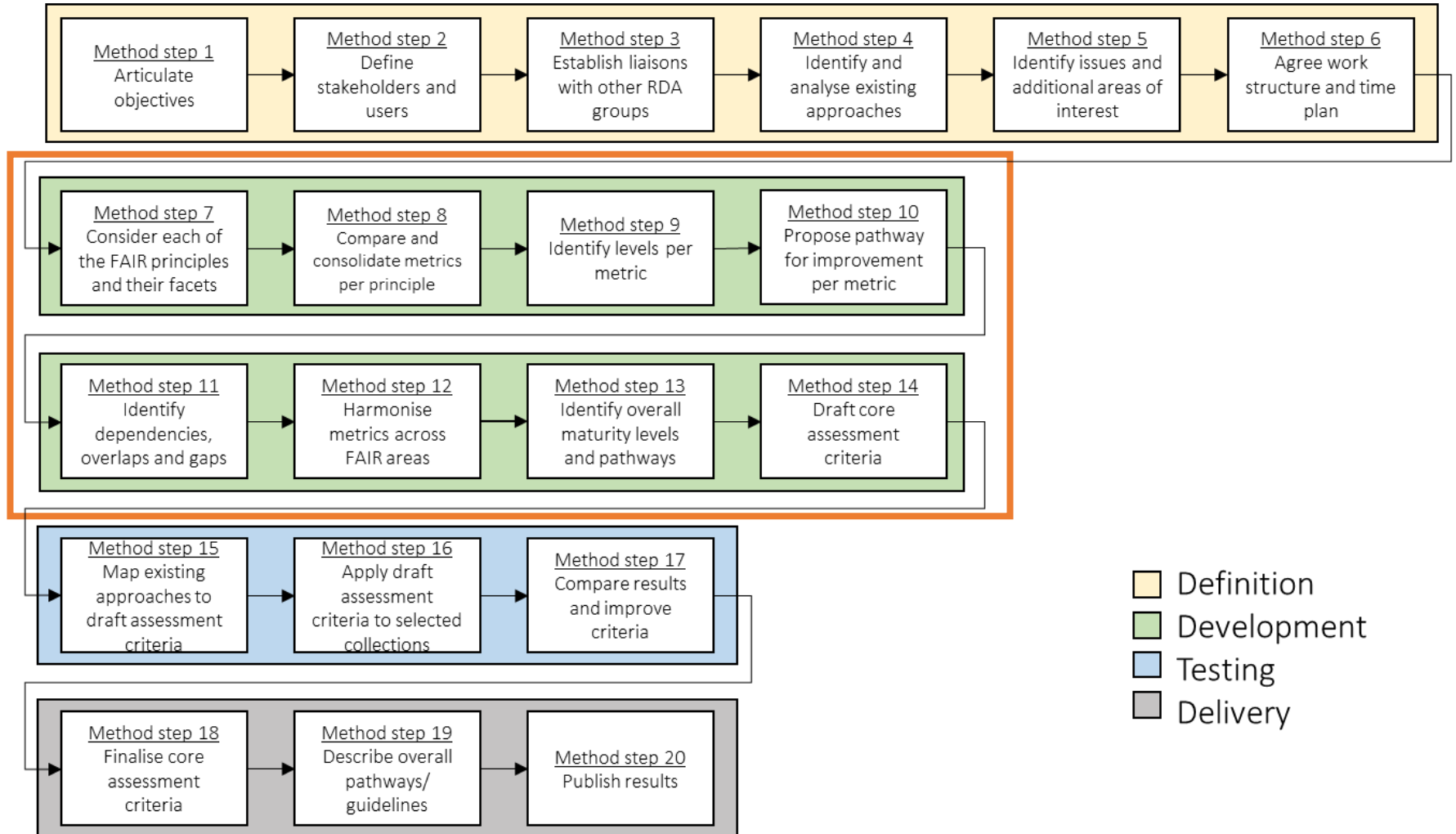
3. Testing

ON HOLD

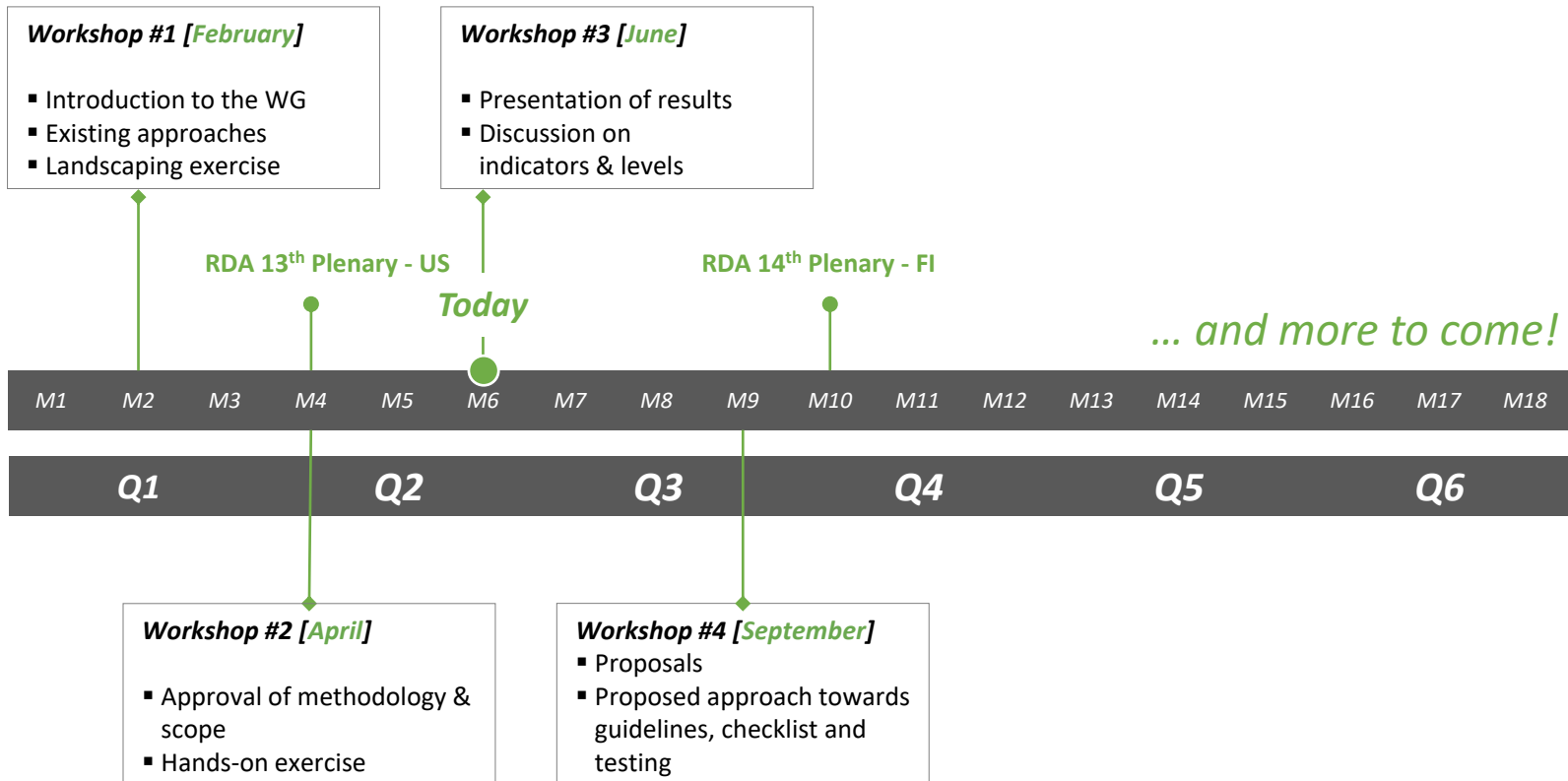
4. Delivery

ON HOLD

# Overview of the methodology



# Timeline



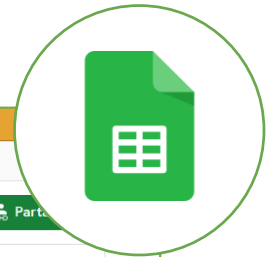
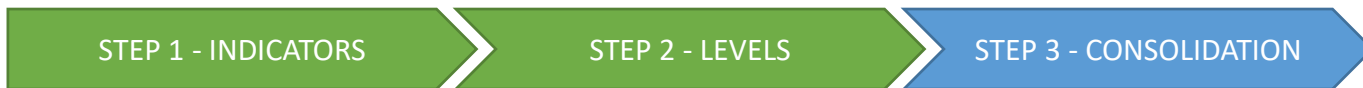




# Development

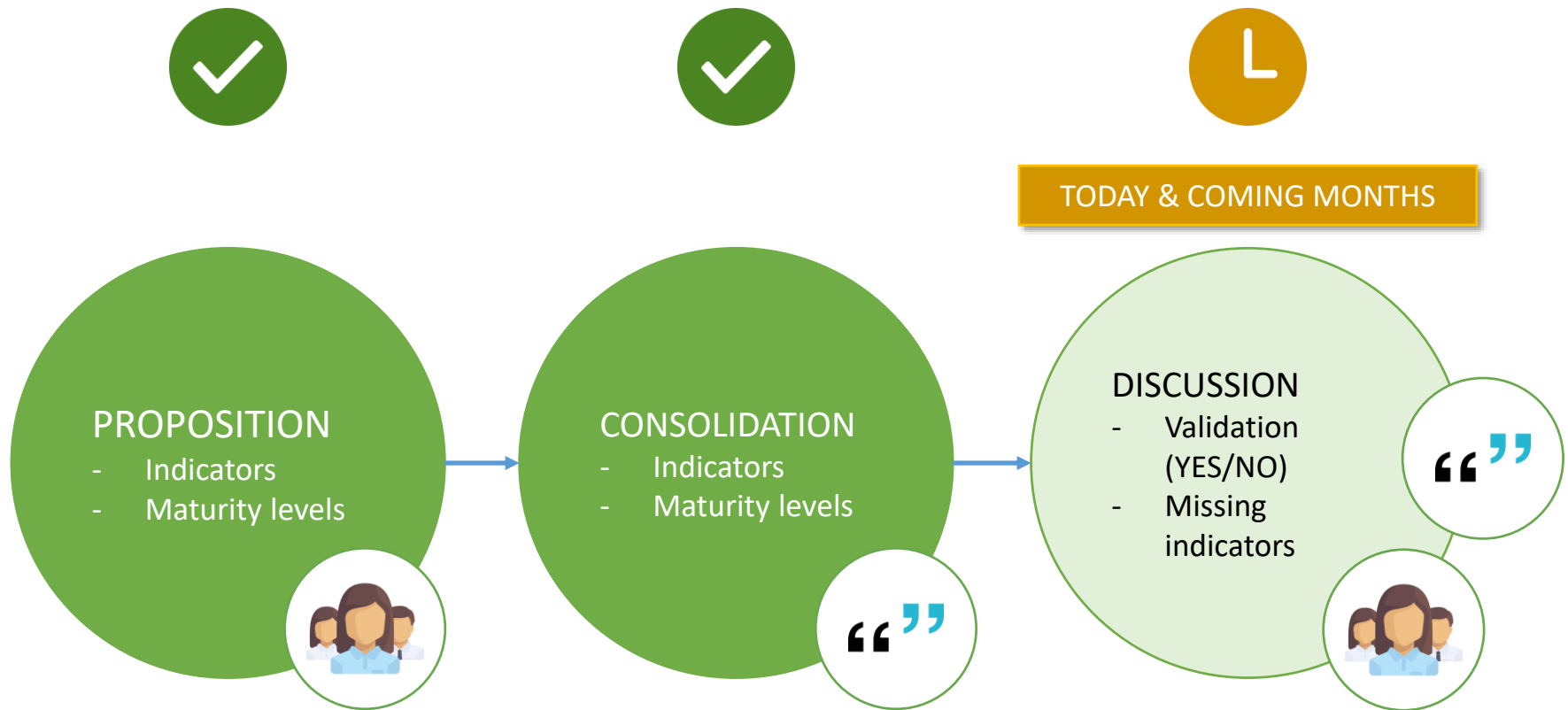
## *First Phase*

# Development | First phase

\* The indicators and levels later presented are derived from the contributions on the [GSheet](#)

# Development | First phase



# Development | Bottom-up approach

- ✓ Looking at all ‘**atomic**’ indicators and their ‘**binary**’ maturity levels [[Slide 20 Workshop #2](#)]

Indicator #1

- YES
- NO

Indicator #2

- YES
- NO

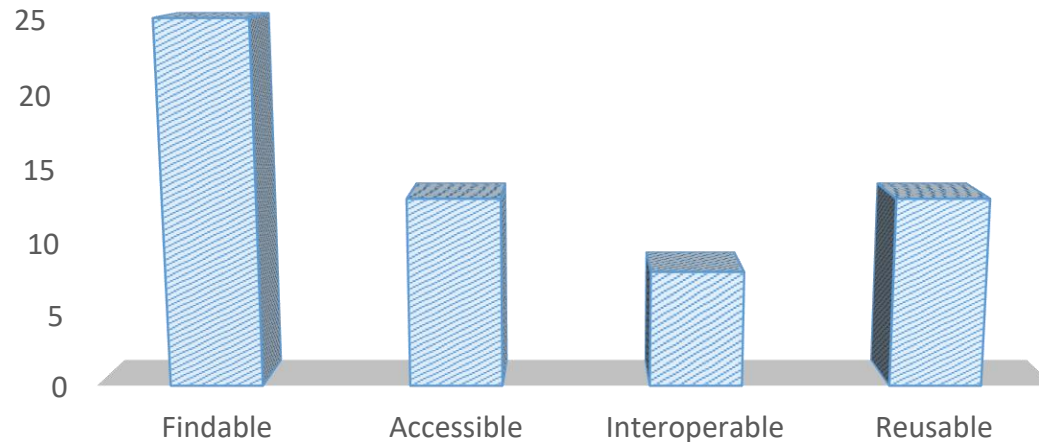
- L Looking at deriving a set of levels across indicators for a principle [[Slide 19 Workshop #2](#)]

Combination of Indicator #1 and Indicator #2

- Level 0
- Level 1
- Level 2

# Development | Statistics

## MEMBERS CONTRIBUTION



- More concrete contribution for F & A
- Most contribution are about metadata
- Complex versus simple principles [e.g. 11 indicators for F1 compared to 1 indicator for I3]

# Development | Indicators & levels

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

## Globally unique

- **NO** globally unique identifier
- Globally unique identifier

## Persistent

- **NO** persistent identifier
- Persistent identifier

## Resolve

- Identifier does **NOT** resolve
- Identifier resolve



# Development | Indicators & levels

F2. data are described with rich metadata  
(defined by R1 below)

## Metadata

- **NO** metadata
- Metadata

## Landing Page

- **NO** landing page
- Landing page

Providing descriptive information according to a formal metadata standard

- **NON**-standard metadata
- Standard metadata



# Development | Indicators & levels

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

Presence of globally unique and eternally persistent identifier in the metadata

- **NO** globally unique and eternally persistent identifier
- Globally unique and eternally persistent identifier (e.g. DOI)





# Development | Indicators & levels

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

## Harvested by search engine

- **NOT** harvested by a search engine
- Harvested by a search engine

## Providing metadata to specific portals

- Metadata **NOT** indexed in specific portals
- Metadata indexed in specific portals

## Institution repositories

- (meta)data **NOT** present in institution repositories
- Presence of the (meta)data in institution repositories



# Development | Indicators & levels

A1. (meta)data are retrievable by their identifier using a standardised communications protocol

## Access conditions

- **NO** access conditions
- Access conditions

## Manual access

- Data retrievable via the researcher
- Data retrievable via a repository

## Automated access

- Data retrievable via human interaction
- Data retrievable using a standard client software



# Development | Indicators & levels

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

## Free and open source protocol

- **NO** free and open source protocol
- Free and open source protocol



# Development | Indicators & levels

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

## Protocol authentication

- **NO** protocol authentication
- Protocol authentication



# Development | Indicators & levels

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

## Metadata persistence policy / guarantee

- **NO** persistence policy / guarantee
- Persistence policy / guarantee



# Development | Indicators & levels

I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation

## Format

- Proprietary format
- Community standard format

## Machine-readable [Knowledge representation]

- **NOT** machine readable
- Machine readable

## Self-describing

- **NOT** self-describing
- Self-describing



# Development | Indicators & levels

I2. (meta)data use vocabularies that follow FAIR principles

## Standard vocabularies

- **NO** standard vocabularies
- Standard vocabularies

## FAIR compliant vocabularies

- **NO** FAIR compliant vocabularies
- FAIR compliant vocabularies



# Development | Indicators & levels

13. (meta)data include qualified references to other (meta)data

## References to other metadata

- **NO** references to other metadata
- References to other metadata (e.g. ORCID for author)

## Sufficient level of qualification for other metadata

- General qualification (e.g. contributor)
- Specific qualification (e.g. author, reviewer, publisher, funder)





# Development | Indicators & levels

R1. meta(data) are richly described with a plurality of accurate and relevant attributes

Use of guidelines for relevant attributes accompanying metadata

- **NO** guidelines
- Guidelines



# Development | Indicators & levels

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

## User licence

- **NO** user licence
- Presence of a user licence

## Nature of the licence

- Local licence
- Standard licence (e.g. CC)

## Machine-readable licence

- **NON** machine-readable licence
- Machine-readable licence

## Consent for reuse

- **NO** information about the consent
- Explicit documented consent



# Development | Indicators & levels

R1.2. (meta)data are associated with detailed provenance

## Authorship included

- **NO** author
- Reference to an author

## Technical parameters [e.g. software's and instruments]

- **NO** information about the technical parameters
- Information about the technical parameters



# Development | Indicators & levels

R1.3. (meta)data meet domain-relevant community standards

Presence of a template for metadata following a community standard

- **NO** template
- Presence of a template



# Development | Indicators & levels

## X. Choices beyond FAIR

- Versioning of the identifier [F1]
- Versioning of the dataset [F1]
- Keywords for rich metadata [F2]
- Quality (i.e. referential and functional integrity) [F2]
- Access control [A1.2]
- Minimal metadata [R1.3]

Should they have **indicators**? If so,  
under which **principle**?





# Development

## *Second Phase*

# Development | Second phase

## Core assessment criteria to evaluate and compare FAIRness

- One to three indicators per FAIR principle [e.g. F1, A1.1]
- Bottom-up perspective

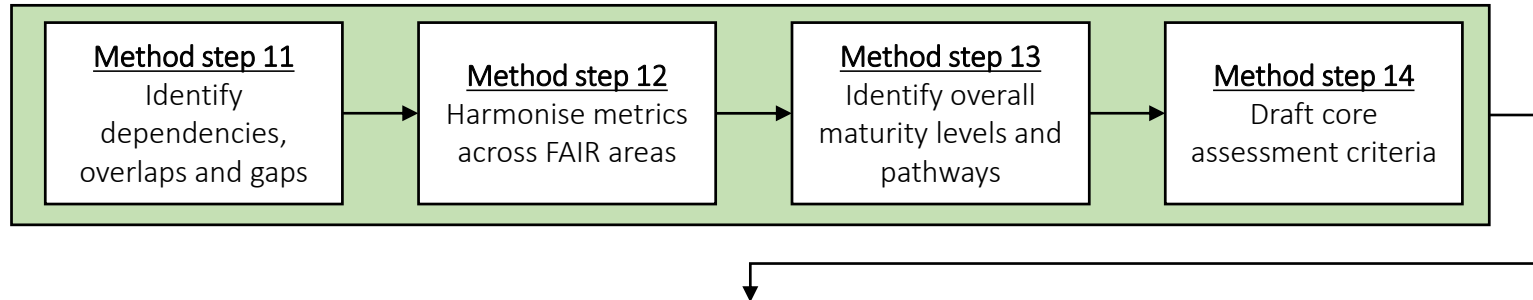
### FAIRness on a two | three level scale for the indicator 1

- Not FAIR
- FAIR
- Not FAIR
- Partly FAIR
- FAIR

### Weighing indicators? 2

- **Required** [Must be FAIR]
- **Recommended** [Should be FAIR, if possible]
- **Optional** [May be FAIR but not essentially so]

# Development | Second phase



## Core assessment criteria to evaluate and compare FAIRness

- FAIRness report for a resource under evaluation
  - Indicators classified per importance
- FAIRness score per principle [to which the indicator pertain]
- FAIRness score for the FAIR areas
- FAIRness score across the FAIR areas, possibly?
- Documentation of the results





Next steps

# Next steps

Provide feedback to the proposals presented at the meeting of today on the [GitHub](#), if at all possible, **by the 30<sup>th</sup> June**

Contribute more indicators and maturity levels on [Google Sheet](#), **until the 31<sup>st</sup> of August**

- Analysis of **all** the FAIR principles
  - **F**AIR – Findable [\[Link\]](#)
  - **A**IR – Accessible [\[Link\]](#)
  - **I**AIR – Interoperable [\[Link\]](#)
  - **R**AIR – Reusable [\[Link\]](#)

Share ideas about consolidation and weighting of indicators and maturity levels on the [GitHub](#)

## Online Workshop #4

- at 09:00 CEST on the 12 September 2019
- at 17:00 CEST on the 12 September 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 – **Collaborative document**

[https://docs.google.com/spreadsheets/d/1gvMfbw46oV1idztsr586aG6-teSn2cPWe\\_RJZG0U4Hg/edit#gid=0](https://docs.google.com/spreadsheets/d/1gvMfbw46oV1idztsr586aG6-teSn2cPWe_RJZG0U4Hg/edit#gid=0)

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

fair\_maturity@rda-groups.org



Thank you!