eosc

Winter School 2024

29 January - 1 February 2024 / Thessaloniki, Greece

eosc.eu | #eoscwinterschool2024

Supported by
eosc | RAISE

eosc | Focus
Funded by
the European Union

Task Forces and projects synergy work around opportunity areas

01.02.2024

Funded by
the European Union

Opportunity Area 3: FAIR Assessment and Alignment

Chris Schubert
TF FAIR Metrics and Data Quality, TU Wien
 on behalf of OA 3 participants

OA 3 Participants

- Karl Presser, Premotec - FNS Cloud & TF FAIR Metrics & Data Quality
- Munaza Andrabi, University of Manchester - EOSC4Cancer
- Daniel Garijo, Universidad Politécnica de Madrid - FAIR-IMPACT
- Apostolos Ampatzoglou, U Macedonia
- Alexandros Chatzigeorgiou, U Macedonia
- Ruda Miroslav, cesnet - CRAFT-OA & EOSC Beyond
- Elli Papadopoulou , Openaire - RAISE & OSTrails
- Fernando Alguilar Gomez, CSIC - AI4EOSC & SIESTA
- Marine Vernet, IFREMER - FAIR-EASE
- Andrea Bertino, Switch - TF FAIR Metrics & Data Quality

- Chris Schubert, TU Wien - TF FAIR Metrics & Data Quality
- Mark Wilkinson Centre for Plant Biotechnology and Genomics UPM INIA -TF FAIR Metrics & Data Quality
- Roxanne Wyns KULeuven - TF Long-term Data Preservation
- Hervé L'hours, UK Data Service - TF Long-term Data Preservation
- Chris De Loof, Belnet - EOSC Focus
- Marthe Bierens, TU Graz - EOSC Focus
- Eleni Bolieraki, U Thessoloniki (AUTH) - RAISE (communication)
- Nana Anastasopoulou, GRNET - Skills4EOSC (communication)
- Alexandra Delipalta, RDA - RDA Tiger (communication)

OA 3 Objectives and Approach

Enhance knowledge on FAIR assessment tools, methods and FAIR implementation profiles and provide practical lessons on how to include FAIR assessment in project specific use cases and communities

Integrate the diverse landscape on FAIR assessment and bringing them into focus for HE EOSC-related projects and EOSC-A Task Forces

How to align Data Long Term Preservation & Data Quality recommendations

Funded by
the European Union

eosc | Focus

OA 3 Session – Tuesday Afternoon

Task Force Outcomes FAIR Metrics and Data Quality & Long Term Preservation

*hands on session* - Comparison of specific assessment tools:
- FAIR Evaluator https://w3id.org/AmIFAIR
- F-UJI - https://www.f-uji.net
- Assessment of "non-data" digital objects
  - Semantic Artifacts (FOOPS) https://w3id.org/foops/
  - Research Software (phyton) https://github.com/fair-software/howfairis/
  - Research Objects (FAIRO) https://www.rohub.org/
- EVA Tool https://doi.org/10.1038/s41597-023-02652-8

**Funded by
the European Union**

OA 3 Session – Wednesday Morning

Round table FAIR assessment implementation

- FAIR data vs Data Quality vs Data value
  - metadata vs data → clarifying the difference to minimize the risk that FAIR data is interpreted as the qualifier for good data
- Users needing more guidance and explanation: Risk of understanding FAIR metadata as good data sets
  - creating exemplars and advice on improvement
  - including the human factor (help desk)
- Signposting - as a unifying strategy
  - Clear recommendation for the community in not just stick but also carrots:: The reuse of your data increases citations
- Governance: how to 'control' the use of tests in communities (survey)
  - increasing trust in tools → moving away from self-assessment
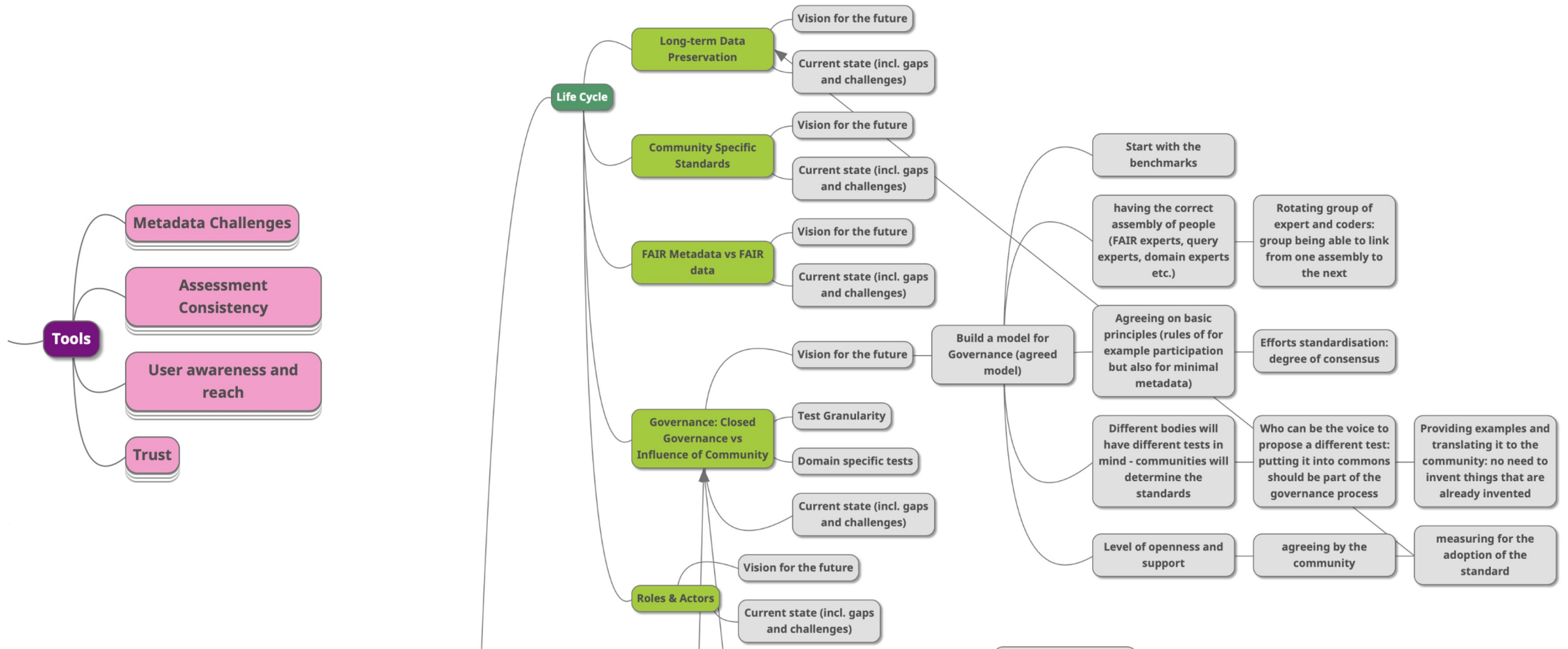 - creating a governing body (OS Trails)
  - benchmarking tools in communities
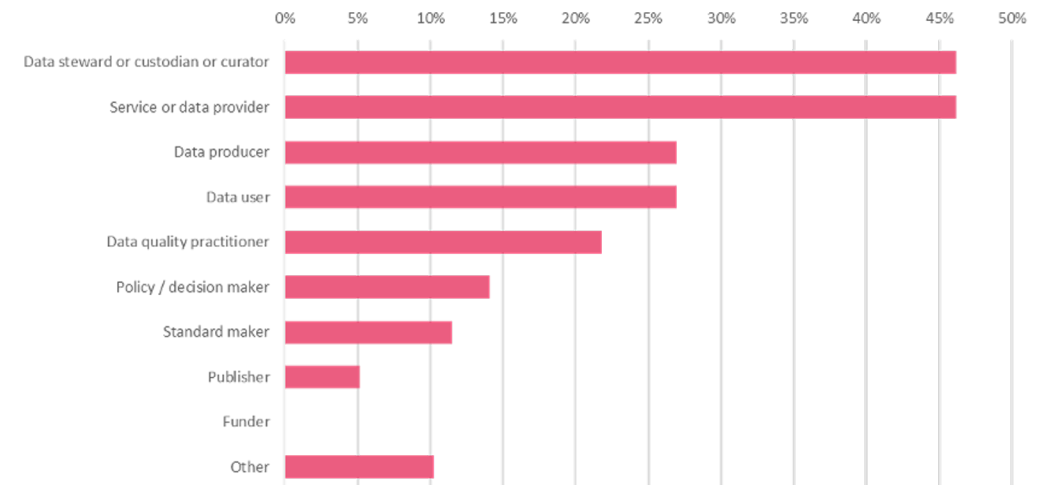
Survey Results

Community Perspectives on Assessment

OA 3 Session – Wednesday Afternoon

OA 3 Recommendations

Recommendation: Clear differentiation (including the landing of it) between <u>FAIR data</u>, <u>data quality</u> and <u>value assessment</u> and applying this in different environments

- Rationale: FAIR data is not the same as data of high quality but often used interchangeable, which creates the risk of using FAIR Assessment tools for the wrong ends.

- Key players: Researchers and repositories

- Expected Outcome: By clarifying and disseminating this clarification in a way that it is clear to the average user, and test how this clarification is received in different environments, the research community becomes more aware of

Recommendation: General applicable and specific recommendations for data quality

- Rationale: Data quality involves many different topics, in order to be able to grasp it and guide data quality in a successful way we need agreed or steered recommendations for the EOSC Community

- Key players: EOSC Environment – possible topic for future TF

- Expected Outcome: Agreed and unified understanding and implementation of Data Quality

Recommendation: Objective research into added value/impact of FAIR

- Rationale: Objective research on the added value and impact of FAIR will serve as an <u>evidence base</u> for our actions in EOSC and convince EOSC users and beyond of its importance

- Key players: within EOSC environment and beyond - possible topic for EOSC observatory?

- Expected Outcome: Objective measurement on benefits of FAIR

Recommendation: Development metadata criteria for transparency levels

- Rationale: Assessment tools differ internally based on the choices made by the test developers- Transparency of what the tool is exactly testing is important to understand the use of the tool and for it to reach its target user group.

- Key players: Researchers, repositories and any one who uses FAIR Assessment Tools

- Expected Outcome: FAIR Assessment tools become transparent to its users.

OA 3 Recommendations

Recommendation: Choosing digital object interchanging mechanisms like RO-Crate

- Rationale: Making a decision on which interchanging mechanisms to embed into the community

- Key players: EOSC environment

- Expected Outcome: Step towards RO-Crate becoming the most used method for interchanging digital objects

Recommendation: Making general and specific recommendations for data quality

- Rationale: FAIR in AI but also AI in FAIR - what can we use AI for to make FAIRification easier and less time consuming

- Key players: Within EOSC Environment

- Expected Outcome: The FAIRification of metadata can be very time consuming, using a developed training model in AI that provides direct integrations can make this process less time consuming

**Recommendation: Providing a unified strategy for signposting**

- Rationale: A solution for harmonisation is Signposting that is capturing the traversal of a FAIR Record. Making this an integral part of EOSC with a unified strategy will help the uptake of signposting.

- Key players: Researchers, repositories, within EOSC environment - taken up by new Task Force and future projects

- Expected Outcome: More harmonisation through a bigger uptake of signposting in the EOSC Community

Recommendation: Need for a human component to make FAIR Assessment to help the researcher with FAIR Assessment tools and advice : helpdesk & dissemination

- Rationale: FAIR Assessment tools are including different tests and have different outcomes. To translate the FAIR Assessment tools into not just scoring tools but make them insightful enough for researchers to implement the improvements behind the scores, a humon component like a help desk would be of great value.

- Key players: Researchers, repositories and any one who uses FAIR Assessment Tools

- Expected Outcome: The FAIR Assessment tool user will be able to implement the improvements indicated by the FAIR Assessment tool and improve the FAIRness of their metadata

OA 3 Next Steps
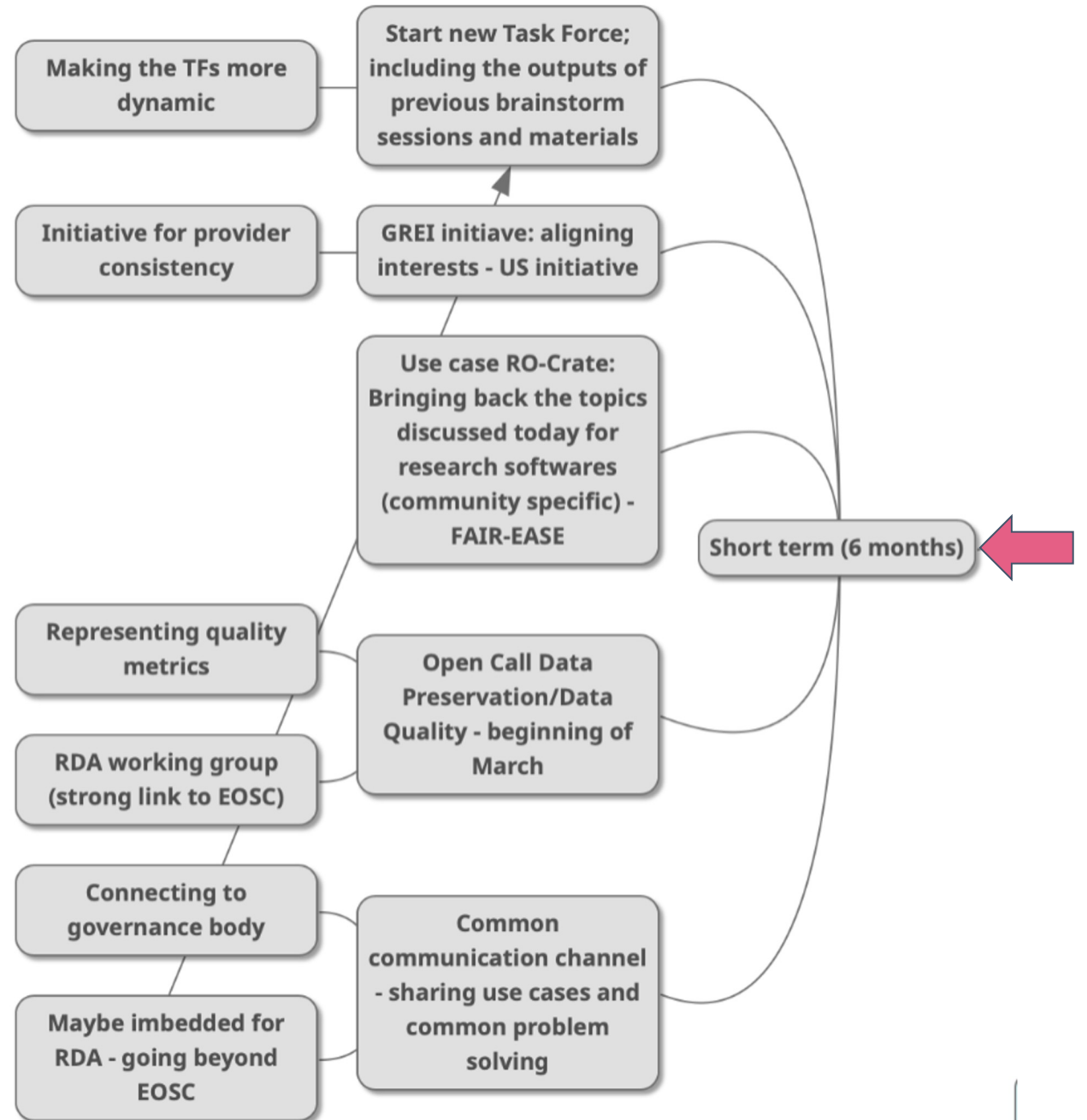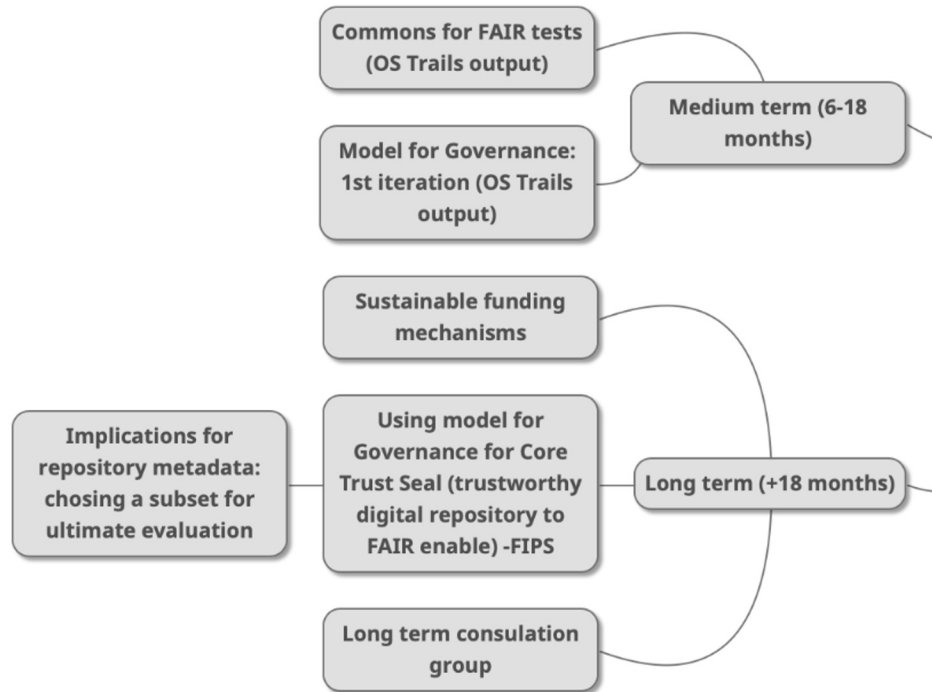
OA 3 Commitment

key message: FAIR data (Scoring) is not Data Quality (Scoring)

- A risky misconception is that a good FAIR data scoring automatically means that the data is of good quality, leading to the possibility of data of bad quality being used more because of their FAIR score instead of their content
- OA3 is committed to carrying out that message towards their own communities, into the new Task Forces and into their (future) projects through several actions:
  - All project members will be advocators for the topics discussed in the OA3 session, taking home lessons on the importance of FAIR metadata and data quality and preservation;
  - In this message the importance and benefits of FAIR data and data Quality for researchers and repositories is carried out, as FAIR data can lead to data being reused which comes with more citations but also with the importance of guaranteeing data quality (on the long term).
  - Mechanisms as RO-Crate and Signposting are being recognized as important steps to making data more FAIR;
  - In order to make FAIR Assessment tools more transparent the FAIR-IMPACT and OSTrails project work together to make the output format of FOOPS, FUJI and the FAIR Evaluator more similar, working towards making the tests and their scores and the improvements they entail more accessible for the average researcher;
  - A Governance model for FAIR testing will be developed in the OS Trails project.