# FAIR Assessment Tools:
An evaluation of assessment tools of data sets according to the FAIR principles

Cora Assmann, Roman Gerlach, Kevin Lang, Nadine Neute, Jessica Rex

FDM THÜRINGEN

This overview presents a selection of FAIR assessment tools. The categories also use the acronym FAIR but stand for: **F**ully Configurable Tools, **A**utomatic Tools, **I**mproved Survey Tools and **R**egular List Tools. The classification of tools into these categories is not always 100% clear, especially when certain properties overlap. In preparation for the poster, an evaluation of all known FAIR assessment tools was carried out between July 2022 and February 2023. Each tool was evaluated by three independent people. The categories used included the time taken to carry out the FAIR assessment, the presence of feedback and the required prior knowledge. With regard to the technical specifics, the research revealed that only generic tools are currently on offer. The poster and the table with the evaluated tools can be found on Zenodo: https://doi.org/10.5281/zenodo.7022038
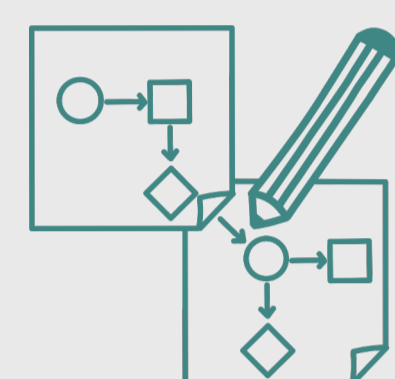
Fully Configurable Tools

The self-configuration of the tools requires certain programming and RDM knowledge as well as a longer training period. The tools are suitable for evaluating a series of data sets. Existing indicators can be used as well as new indicators and test series can be created. Fully configurable tools are particularly suitable for FDM communities and RDM service points that want to evaluate data sets according to their own guidelines. Ultimately, surveys from the "Improved Survey Tools" section can also be completely reconfigured if they have been put online as an open source project. Of course, advanced programming skills are required in this case.

Examples:

FAIRshake

FAIR Evaluation Services
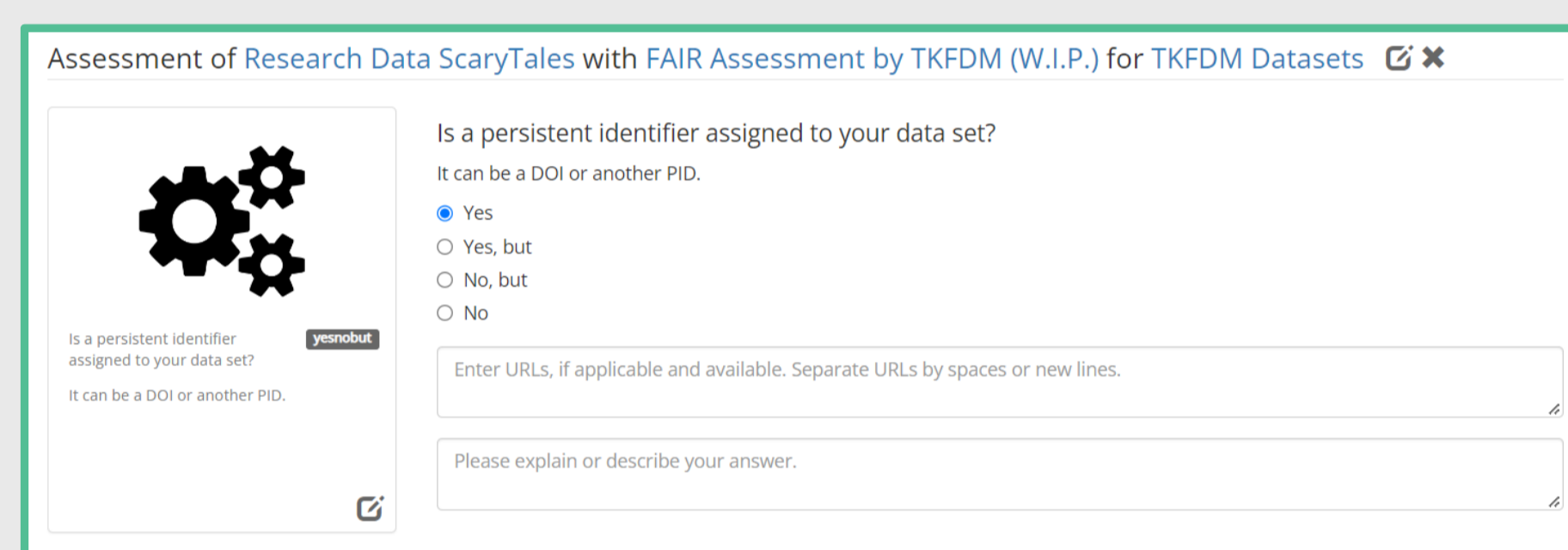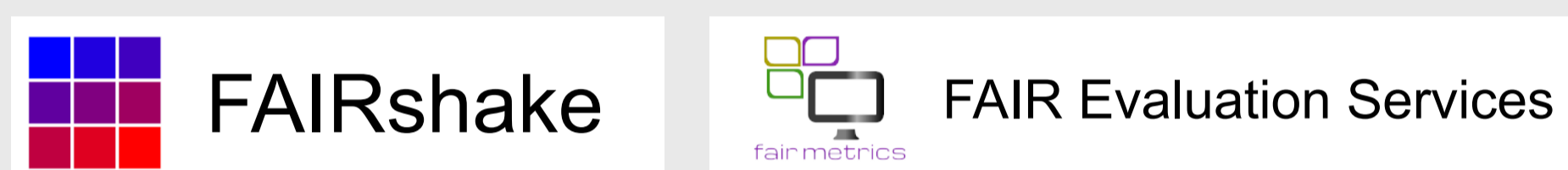
Screenshot of FAIRShake: created question for the data record
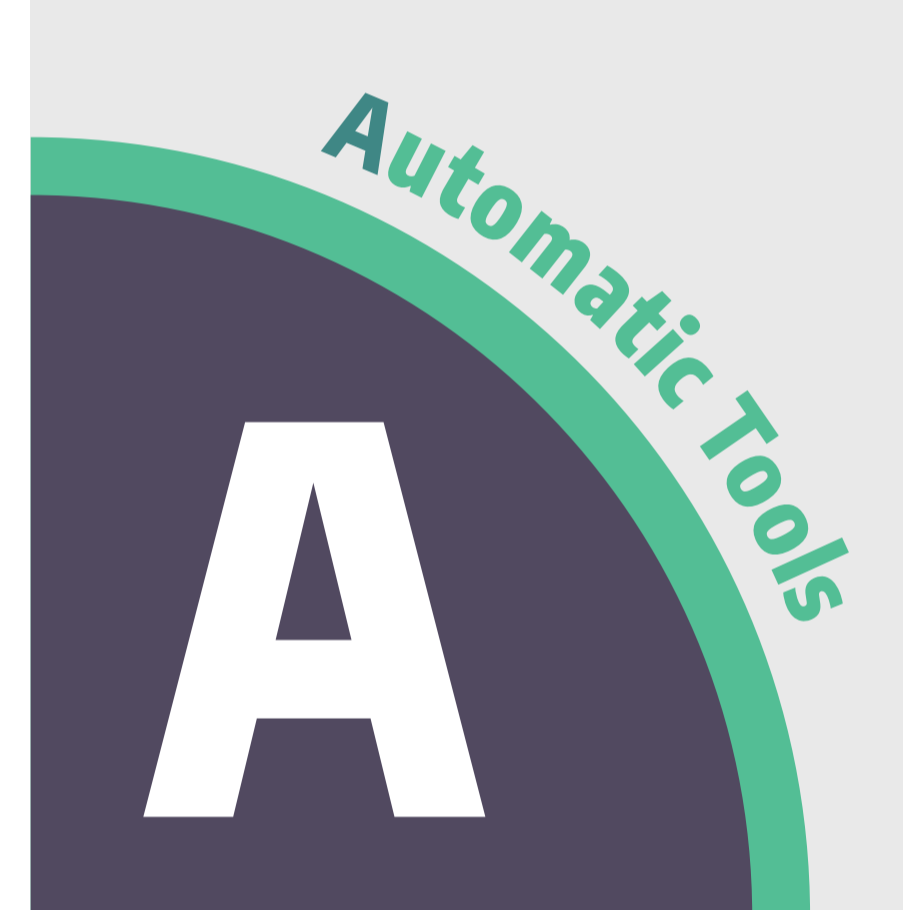
Automatic Tools

The automatic tools are suitable for quickly and consistently evaluating many data sets. The time required per data record is 1 to 5 minutes and is therefore comparatively low. The final report of the evaluation addresses the technical tests passed, such as whether the dataset has a persistent identifier or is licensed. The procedure uses machine-readable data from the repositories and is error-prone where the (meta) data schema of the repository used does not follow the expected interface. Annotations to the results require RDM and technical knowledge. The target group are researchers and data managers (e.g. data stewards, data curators).

Examples:

F-UJI Automated FAIR Data Assessment Tool

FAIR Enough
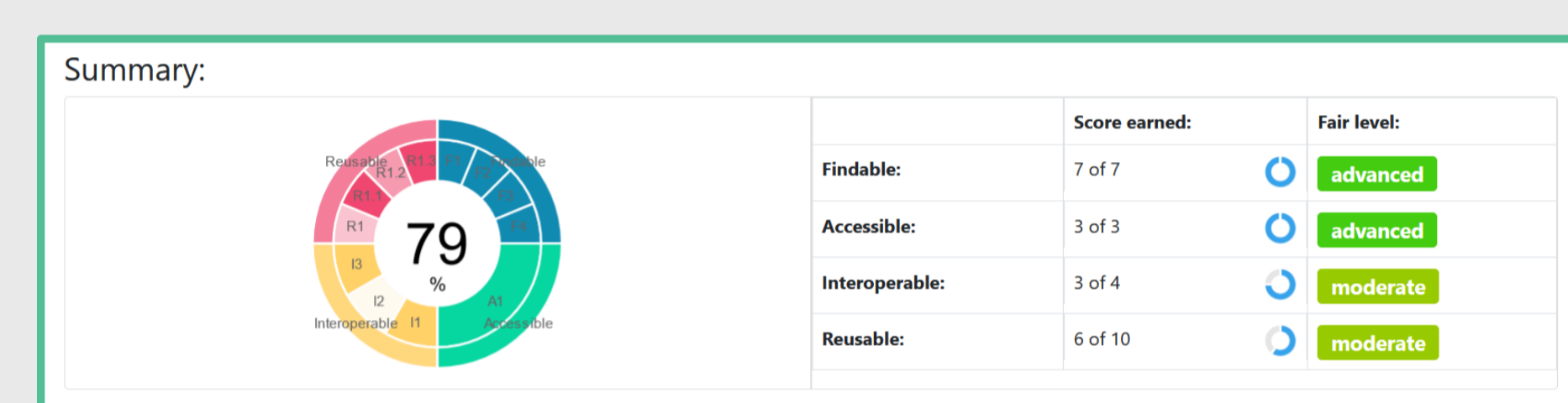
Screenshot of F-UJI: graphical summary of the results

Improved Survey Tools

These tools provide answers that users can select based on their data set. The FAIRness is evaluated automatically based on the answers. Some of the tools also offer background information on questions and individualized feedback to improve the data set examined. The processing time depends on the previous knowledge of the user and can be between 20 minutes and 1 hour. The target group are researchers and data managers.

Examples:

Australian Research Data Commons — FAIR self assessment tool

FairData UFPB
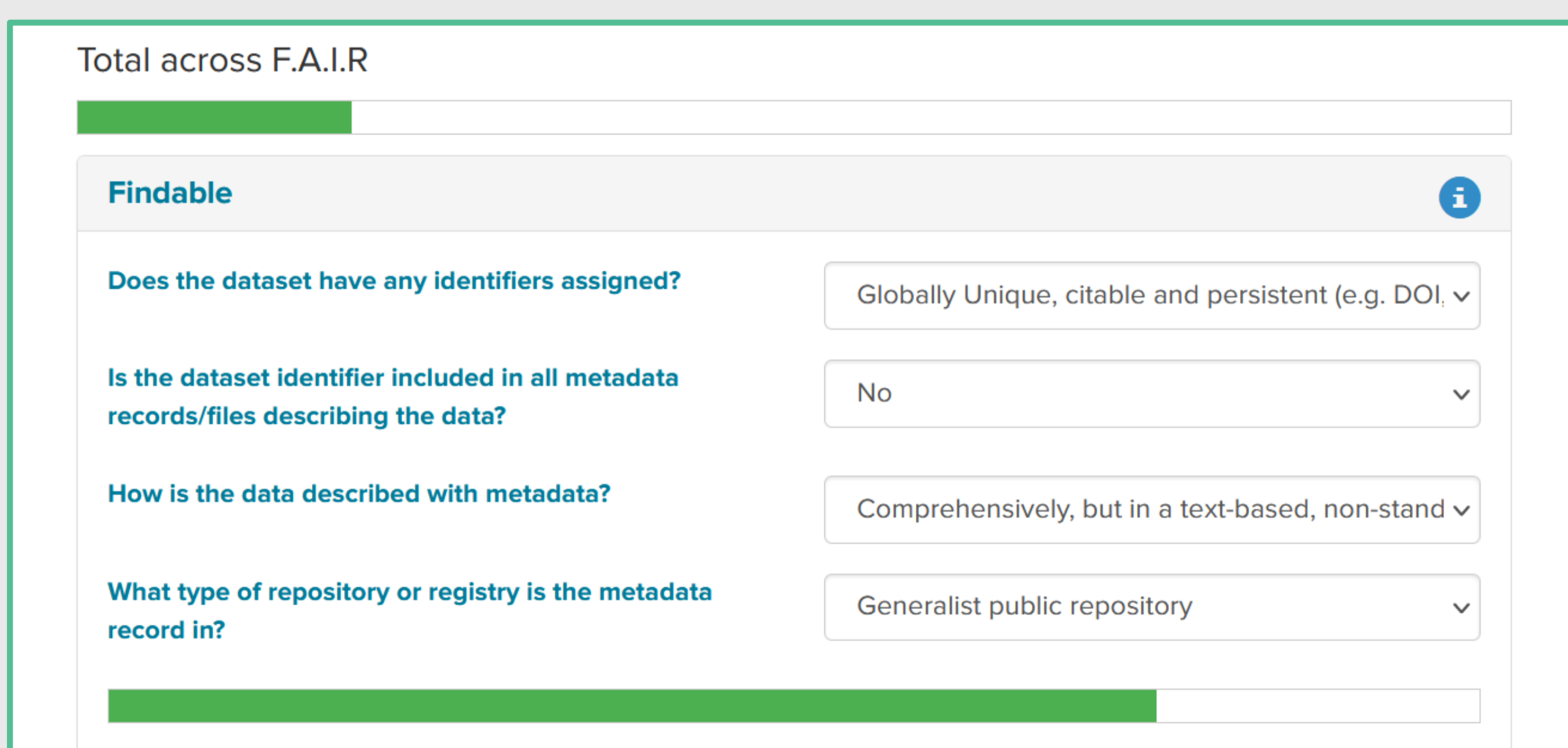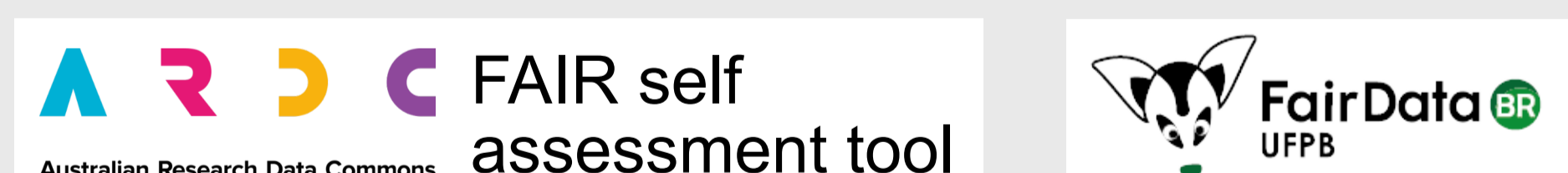
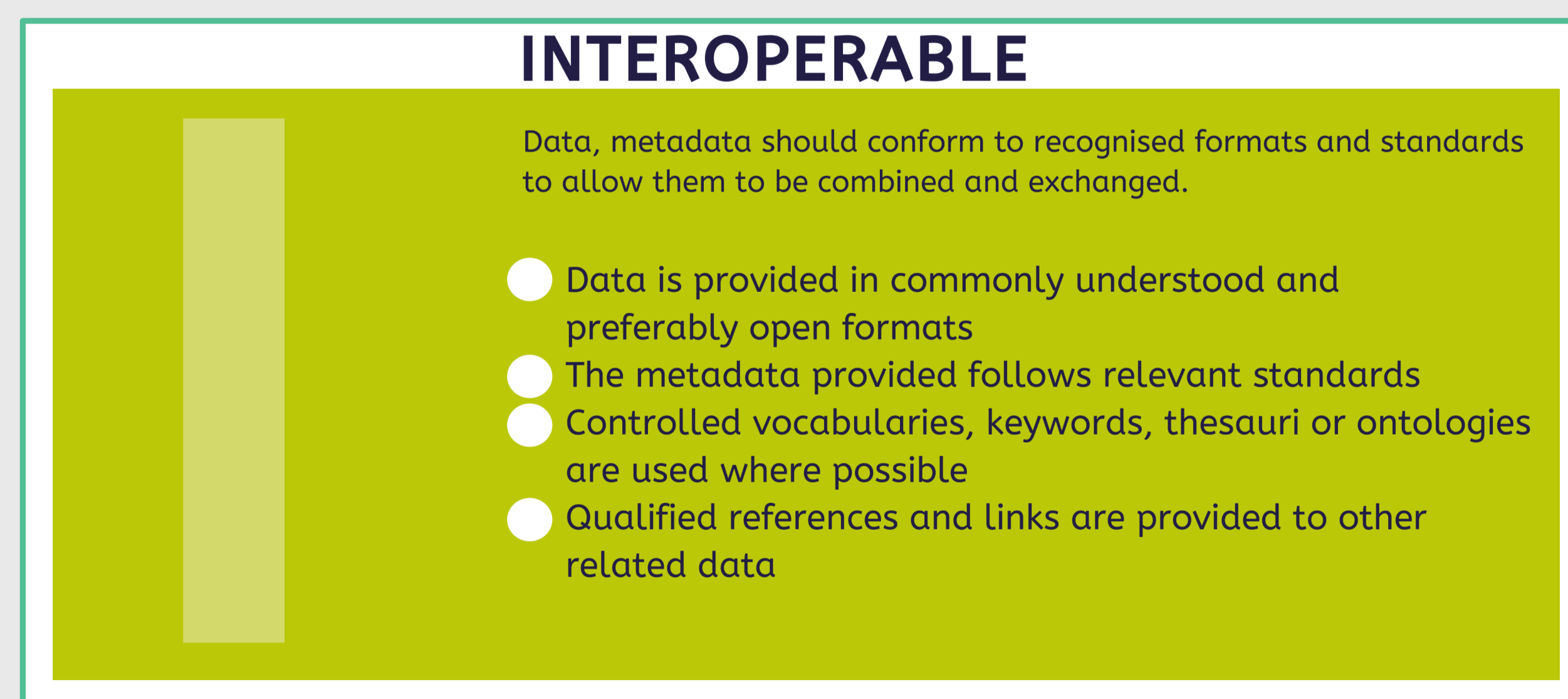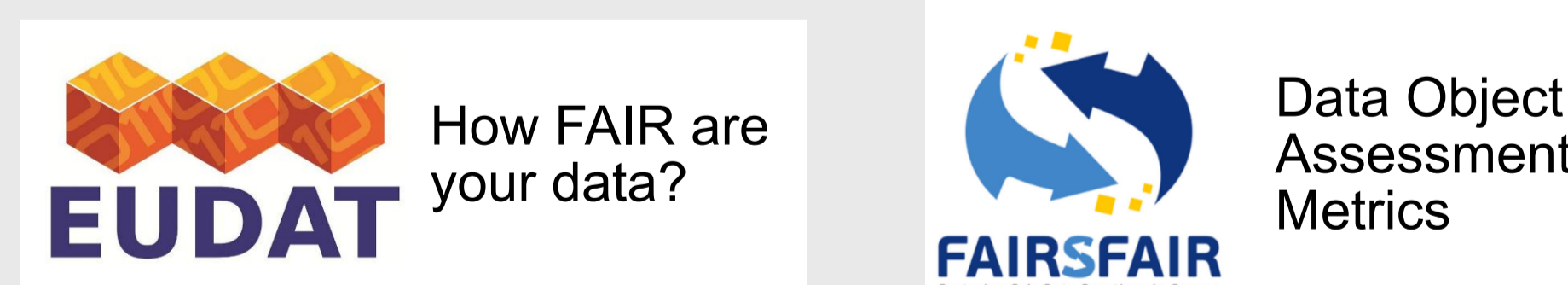Screenshot from ARDC's FAIR self assessment tool for the category "Findable"

Regular List Tools

These tools are simple lists of criteria for meeting the FAIR principles that can be checked on the dataset. Previous knowledge is sometimes necessary because often only limited explanations of the criteria are available. Since these are static documents in the form of a PDF file, there is no evaluation of the results and no individualized feedback. The target group are researchers.

Examples:

EUDAT — How FAIR are your data?

FAIRSFAIR — Data Object Assessment Metrics

INTEROPERABLE

Data, metadata should conform to recognised formats and standards to allow them to be combined and exchanged.

- Data is provided in commonly understood and preferably open formats
- The metadata provided follows relevant standards
- Controlled vocabularies, keywords, thesauri or ontologies are used where possible
- Qualified references and links are provided to other related data

Excerpt from checklist "How FAIR are your data?" for EUDAT summer school