

oooo



MTSR'24



ALIGNING DATA MANAGEMENT PLANS WITH COMMUNITY STANDARDS USING FAIR IMPLEMENTATION PROFILES

oooo

AUTHORS



Shuai Wang

Scientific engineer
VU Amsterdam
Data Steward



Navroop K. Singh

Bachelor student
VU Amsterdam



Angelica Maineri

Data manager
ODISSEI &
Erasmus University
Rotterdam



**Jacco van
Ossenbruggen**

Professor
VU Amsterdam



Tobias Kuhn

Assistant Professor
VU Amsterdam



Ronald Siebes

Assistant Professor
VU Amsterdam



**Stephanie van de
Sandt**

Research Data Steward
& RDM Trainer, VU
Amsterdam



Tycho Hofstra

Research Data Steward
& RDM Trainer, VU
Amsterdam

**Stephanie van de
Sandt**

Research Data
Steward, VU
Amsterdam

FAIR IMPLEMENTATION PROFILE

Definition

The FIP is a collection of FAIR implementation choices made by a community of practice for each of the FAIR Principles.

FIC

A group of people who are using/publishing data on the same topics

FER

A resource that enables the implementation of FAIR behaviors

Their relations

A FIC used a FER
A FIC is using a FER
A FIC will use a FER

FIPS FOR SOCIAL SCIENCE



ESS

European Social Survey (ESS)

AUSSI-ESS

Australian Social Survey International – European Social Survey

GESIS SSSR

GESIS Social Science Survey Research

SEH

SSHOC-NL Socio-Economic History

MCAL

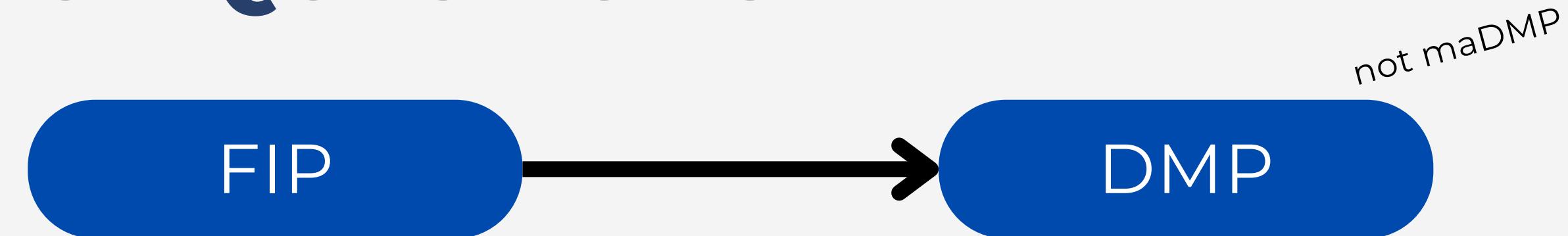
Media Content Analysis Lab scholars

LGBTQVoc

LGBTQ+ Linked Open Vocabulary

To create your own,
choose a topic that
Our MTSR paper last year: https://shuai.ai/static/files/paper/MTSR_paper.pdf
interests you.

RESEARCH QUESTIONS



Communities' FAIR Implementation decisions about their research datasets

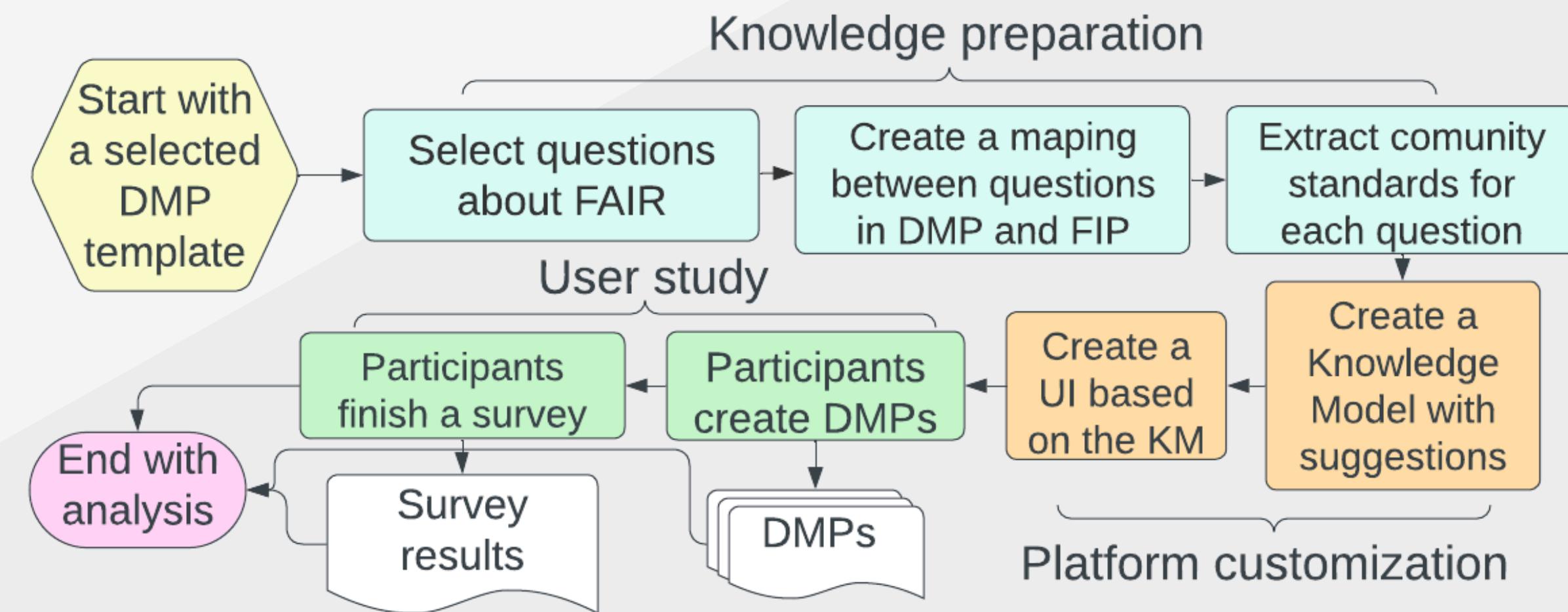
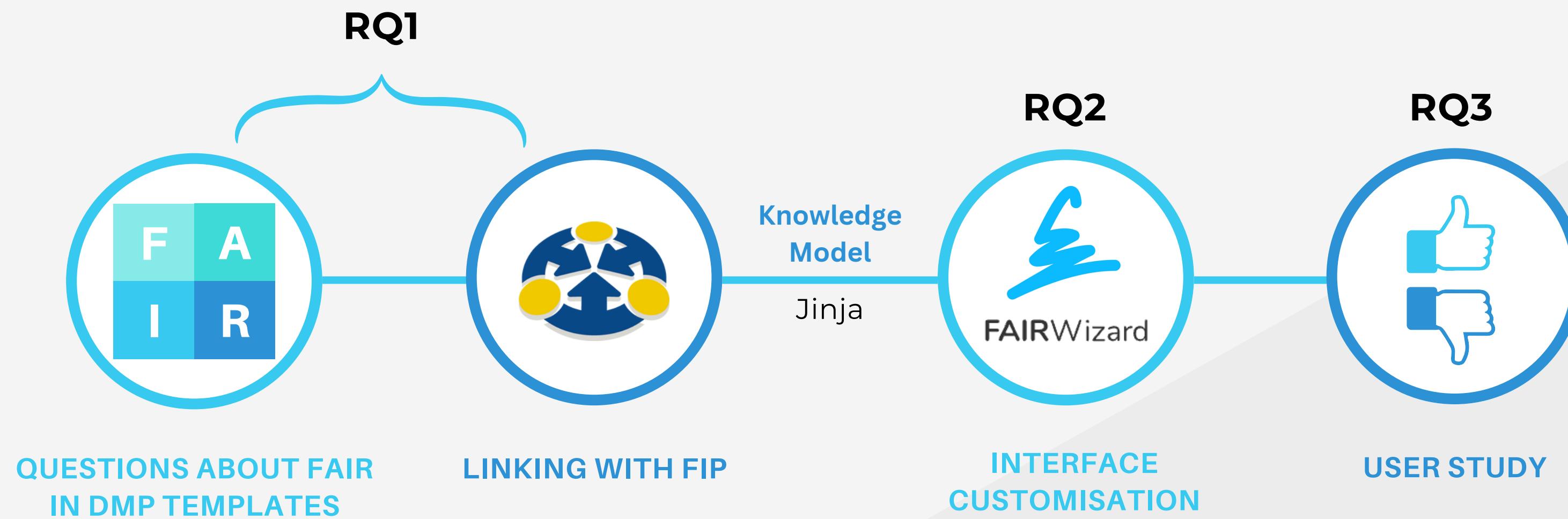
Researchers' decisions about the datasets of a project.

^{1 Related work}

RQ1: Which questions in the DMP template can take community standards in FIPs as suggestions?

RQ2: How can we build a user interface that takes community standards as suggestions?

RQ3: How do users take advantage of suggestions from FIPs while writing their DMPs?



DMP QUESTIONS → FAIR ← FIP QUESTIONS

TABLE II
MAPPING OF DMP QUESTIONS TO FIP QUESTIONS VIA THE FAIR PRINCIPLES.

ID	DMP Question	FAIR principle	FIP Question
4.6	Where will you publish your data assets?	F4 Data	In which search engines are your datasets indexed?
4.8	How will you ensure your data assets get a persistent identifier (e.g. a DOI-code)?	F1 Data	What globally unique, persistent, resolvable identifiers do you use for datasets?
4.9	Will you register your datasets in an online registry other than PURE? If yes, where?	F1 Data	What globally unique, persistent, resolvable identifiers do you use for datasets?
4.13	Please indicate the license and/ or terms of use under which you share your data.	R1.1 Data	Which usage license do you use for your datasets?
5.1	What metadata and documentation will accompany the project?	F2 & R1.2 Data	Which metadata schemas do you use for findability? & Which metadata schemas do you use for describing the provenance of your datasets?
5.2	What metadata and documentation will accompany the data assets?	F2	Which metadata schemas do you use for findability?
5.3	What methods, software, or hardware are needed to access and use your data?	R1.2 Data	Which metadata schemas do you use for describing the provenance of your datasets?

Around 63.6% of the questions about FAIR in VU's DMP template can be mapped to FIP questions

IMPLEMENTATION

- Build the knowledge model (Jinja template) for selected questions
- API of the nanopublication (FIPs)
- Load it on Data Steward Wizard to get the GUI

✓ II.1 4.6 Where will you publish your data assets?

Description

Data assets can only be shared in external repositories with which the VU has a processing agreement. Discuss with your Faculty's Privacy Champion whether or not you can share data in an external repository. Do this as early as possible in your research; don't wait until you are ready to share your data.

Suggestion from the FIPs you selected

You will first be given an option to describe your answer in text, and afterwards you can search using the search bar in case the FER is present there.

The following are the options suggested by the FIPs you selected:

FERs	FIC
The Dataverse Project	SEH
DANS Data Station SSH, The Dataverse Project, Figshare, Open Science Framework	MCAL

✓ II.1.a.1 In text, describe your answer

I will publish the data on the Dataverse Project.

Clear answer

Answered 17 days ago by Shuai Wang.

✓ II.1.a.2 Additionally, please specify the FER using by searching (ignore it if not found).

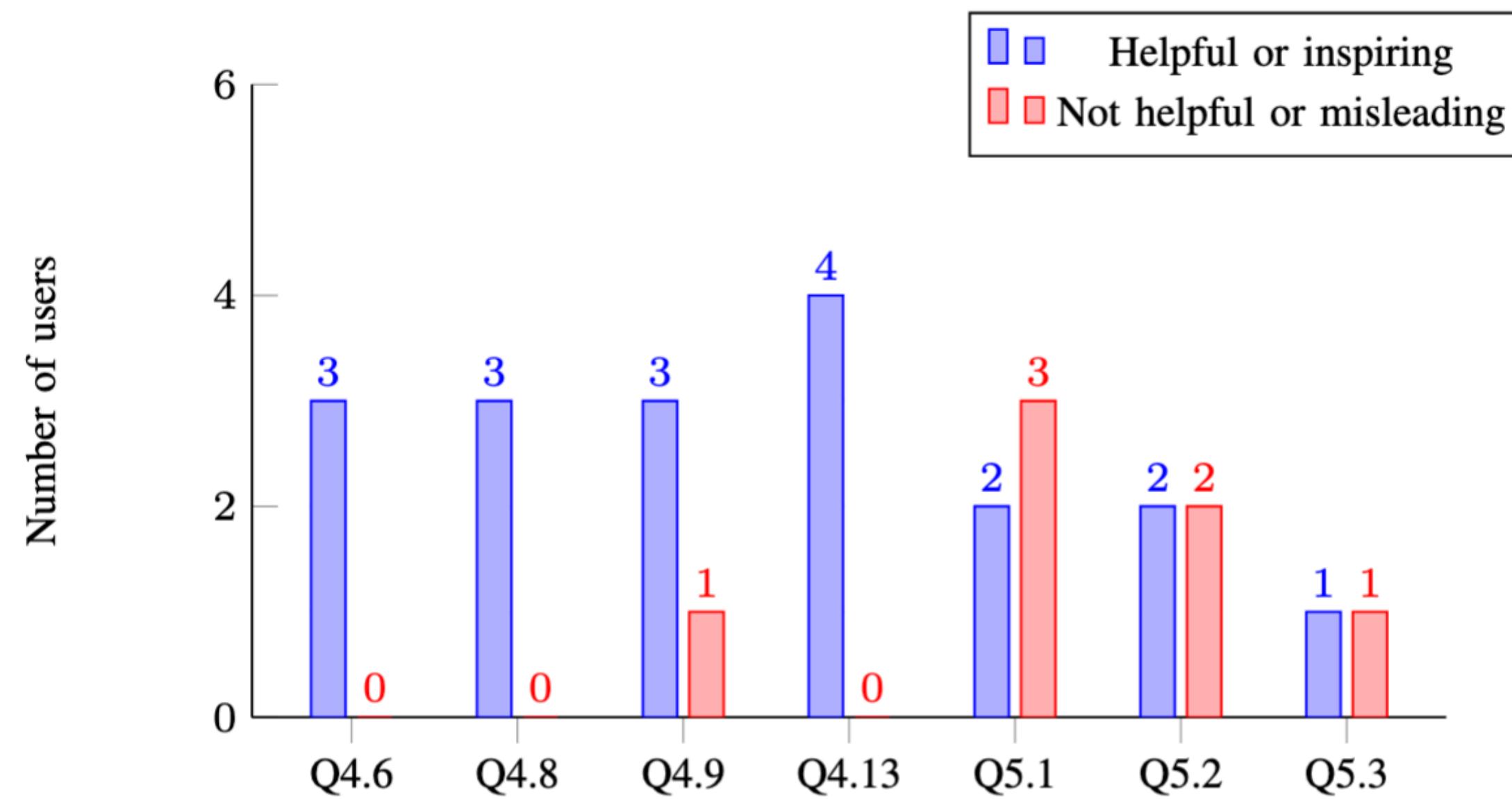
✓ II.1.a.2.a.1

dataverse

Australian Data Archive Dataverse FER
An FER for the ADA Dataverse service

ODISSEI Portal
The ODISSEI Portal is a metadata repository running on the open source DataVerse software developed by the Open Data Infrastructure for Social Science and

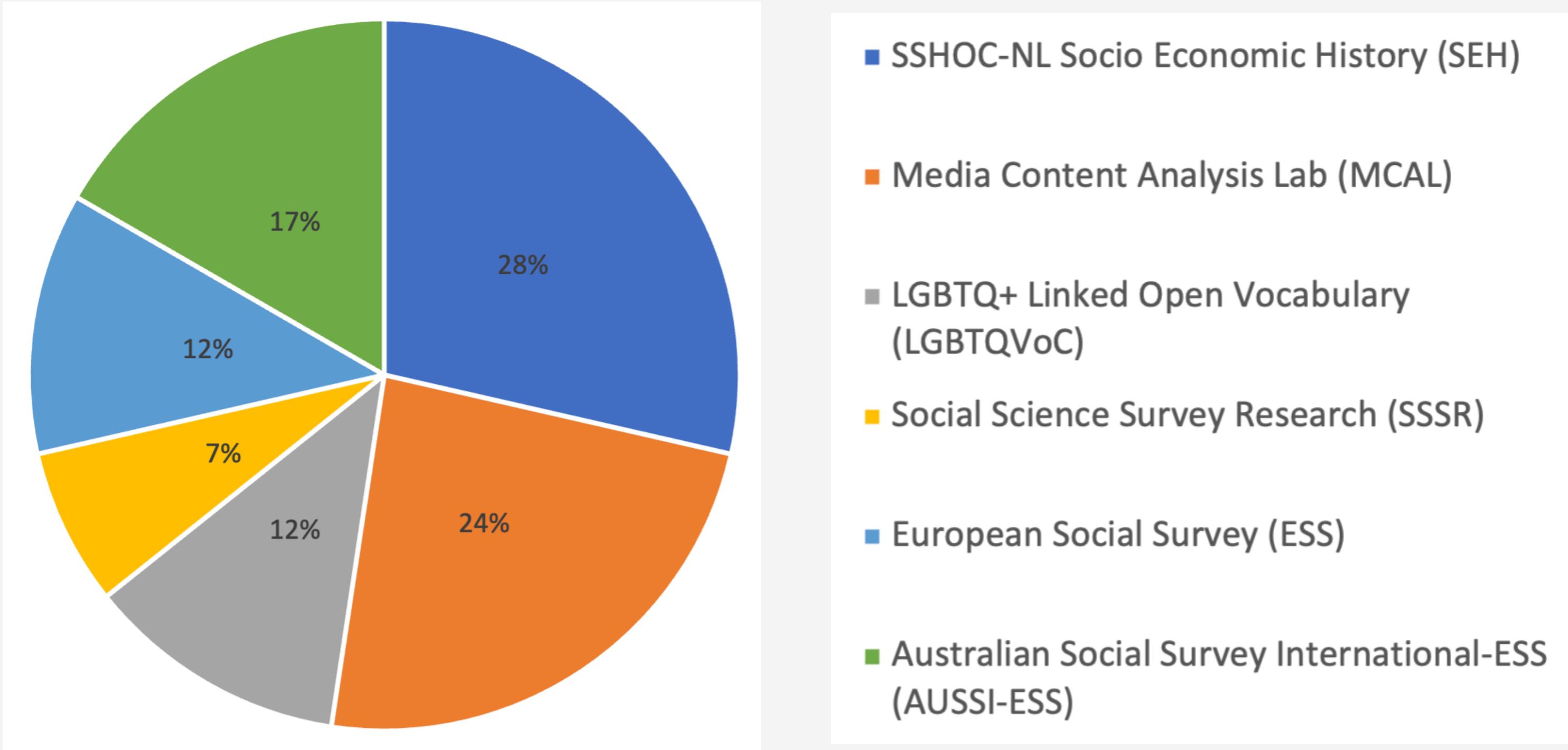
USER STUDY



6 replies after reaching out to 31 researchers in the social science dept.
Provided a sketchy mock DMP.

One was filtered out due to an unexpected error in the mock DMP

USER STUDY: DISTRIBUTION OF FERS AMONG RESEARCH COMMUNITIES



USER STUDY

Survey Question ID	Survey question	Mean	Median	Std
Q1	On a scale of 1 to 5, how relevant are communities for this DMP? 1 indicating that no community is relevant and 5 indicating that many communities are relevant.	3.4	3	0.89
Q2	On a scale of 1 to 5, please evaluate whether the suggestions provided in this DMP are helpful for the communities in answering their corresponding questions. 1 indicating that it is not helpful and 5 indicating that it is very helpful.	3.4	4	0.89
Q3	On a scale of 1 to 5, how much would you consider aligning the decisions in this DMP with those made by the relevant community? 1 indicating minimal alignment and 5 indicating complete alignment.	3.4	3	1.14
Q7	On a scale of 1 to 5, how easy was it for you to find the FAIR-Enabling Resource in the search bar? 1 indicating extremely difficult and 5 indicating extremely easy.	3.6	4	1.14
Q10	On a scale of 1 to 5, how clear was the goal of the study to you? 1 indicating not clear at all and 5 indicating very clear.	3.4	4	0.89

Table 2. Survey questions and their results together the range, mean, median, and standard deviation.

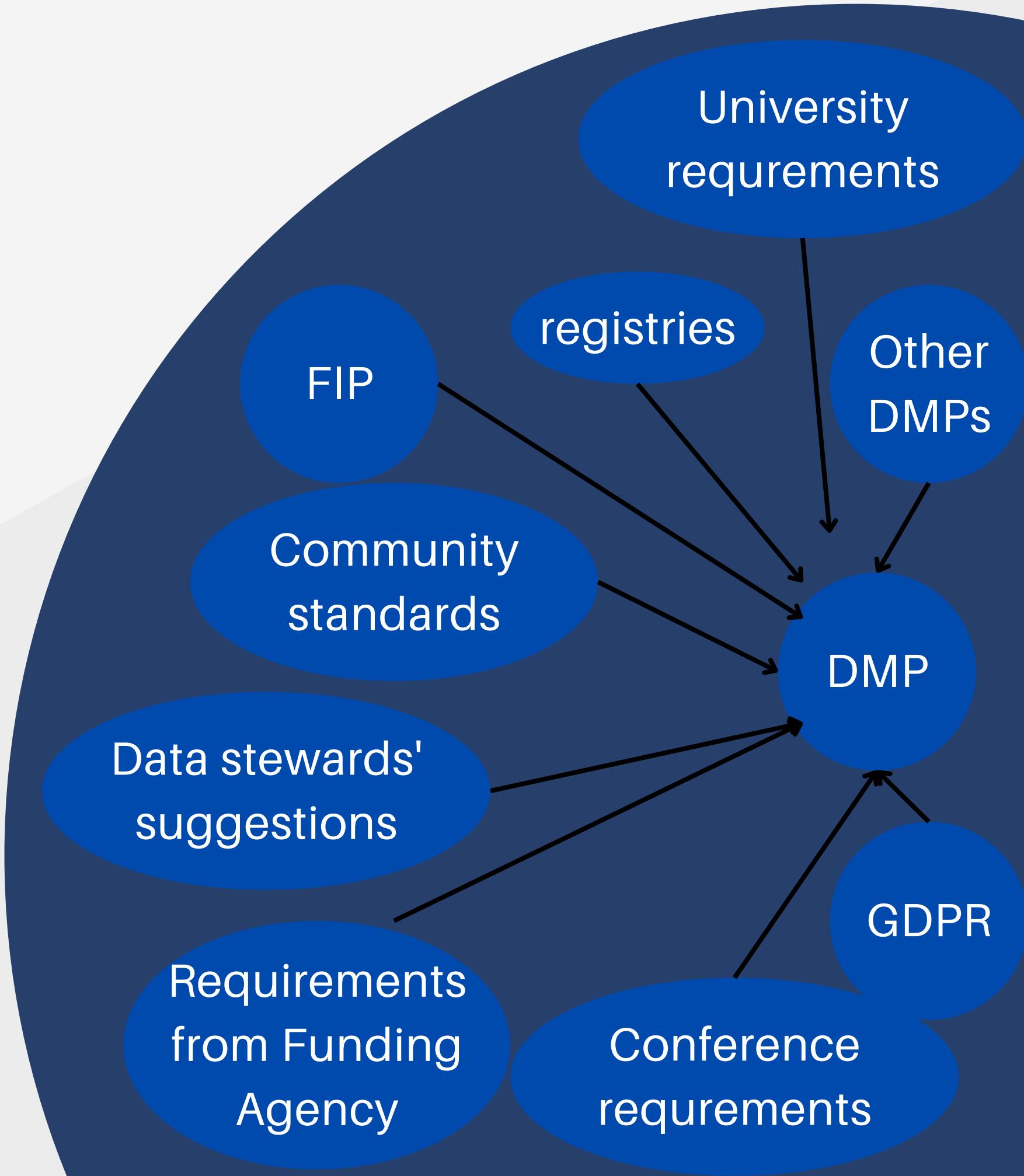
○ ○ ○ ○

Discussion

- A pioneering pipeline about reusing FIPs
- Users lack knowledge about FAIR
- Unspecified FERs
- Knowledge model for a specific template
- A small amount of participant

Future Work

- Extend it to the entire DMP
- Extend it to other domains beyond social science
- maDMP
- Apply the same approach to other sources





THANK YOU



Platform Digitale Infrastructuur
Social Sciences & Humanities

CONTACT



shuai.wang@vu.nl



www.fairexpertisehub.org



References

- Kristina Hettne et al. “FIP2DMP: Linking data management plans with FAIR implementation profiles”.
- Erik Schultes et al. “Reusable FAIR Implementation Profiles as Accelerators of FAIR Convergence”.
- Mark D Wilkinson et al. “The FAIR Guiding Principles for scientific data management and stewardship”. In: *Scientific data* 3.1 (2016), pp. 1–9
- Barbara Magagna et al. “FIPs and Practice”. In: *Research Ideas and Outcomes* 8 (2022). doi: 10.3897/rio.8.e94451.
- Hana Pergl Sustkova et al. “FAIR Convergence Matrix: Optimizing the Reuse of Existing FAIR-Related Resources”. In: *Data Intelligence* 2.1-2 (Jan. 2020), pp. 158–170. issn: 2641-435X

References (useful links)

FAIR Implementation Profile: 3-part video series:

<https://www.go-fair.org/2021/08/02/fair-implementation-profile-3-part-video-series/>

WorldFAIR: <https://docs.google.com/presentation/d/1vDgMCWW-RM806RnDul5ataGh4md5VX49u5NjcolVMU8/edit?usp=sharing>

FIP questionnaire and FER types: <https://bit.ly/WF-FIPs-Questions>

The FIP Wizard: <https://fip-wizard.ds-wizard.org/>