Introducing the FAIR Expertise Hub

Enacting the FAIR principles in the social sciences

VU RDM Symposium, Winter Training Days, VU
22nd November 2022

Shuai Wang |shuai.wang@vu.nl | https://orcid.org/0000-0002-1261-9930
User-centric Data Science (UCDS) group, Computer Science Department, VU

Based on slides of Angelica Maineri

PDI SSH

Erasmus University Rotterdam

VU

Data Archiving and Networked Services
DANS

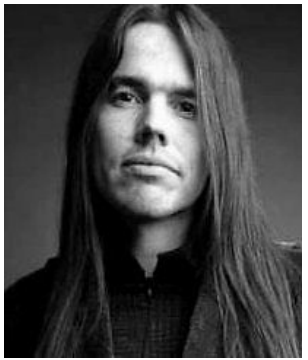
The Core Team

Angelica M. Maineri | angelica@odissei-data.nl |
https://orcid.org/0000-0002-6978-5278
EUR-ESSB + ODISSEI

Elena Beretta | elena.beretta@vu.nl |
https://orcid.org/0000-0002-6090-2765

Tobias Kuhn | t.kuhn@vu.nl |
https://orcid.org/0000-0002-1267-0234

Ronald Siebes | r.m.siebes@vu.nl |
https://orcid.org/0000-0001-8772-7904

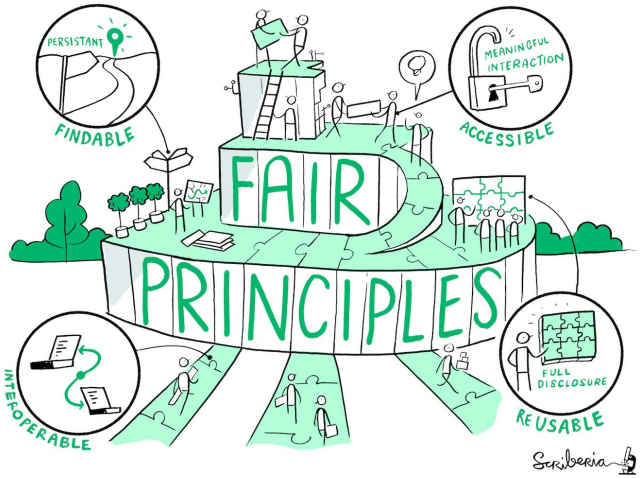


FAIR

FAIR Implementation requires <u>specific technical knowledge</u>

FA

- Data described with custom metadata schemas
- Many datasets not hosted in open repositories
- Restricted access data



IR

- Limited availability of relevant semantic artifacts (e.g. ontologies)

FAIR: From Principles to Implementation

FAIR allows freedom of choice in the implementation strategy

However

- Many resources are not FAIR
- Diversity of stakeholders
- Limited interoperability and reusability within and across disciplines

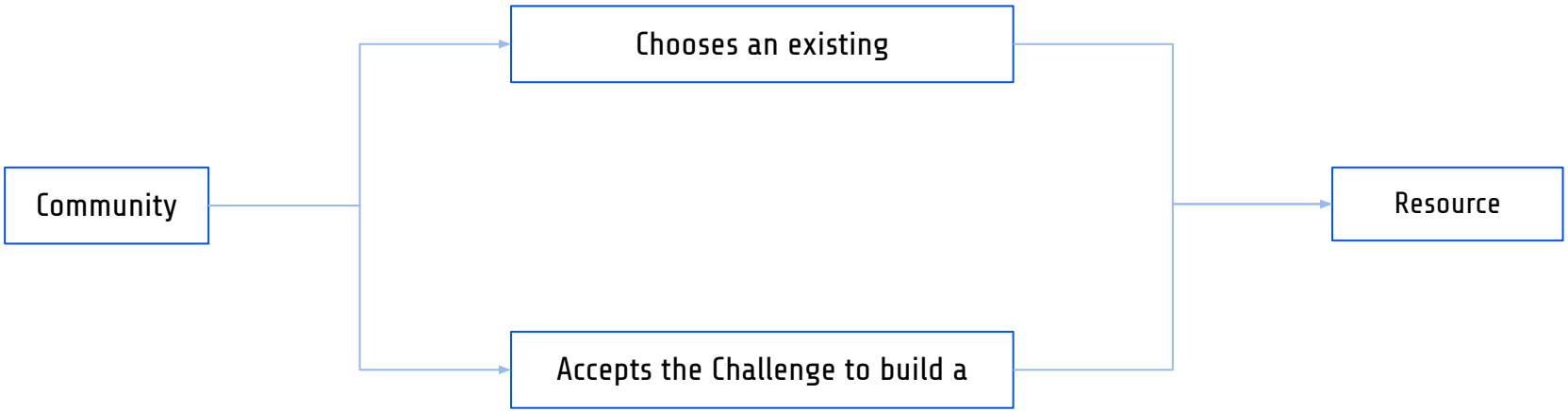
⇒ To foster synergy: FAIR Implementation Profile (FIP)

FIP = a collection of decision a community made about their data

a.k.a. Implementation choices

FAIR Implementation Profiles (FIPs)

= collection of FAIR implementation choices



The FAIR Expertise Hub for the Social Sciences

Problem statement:

Relevant data for the social science comes from communities that typically lack the knowledge, skills and incentives to actually implement FAIR

Our proposed solution:

Make expertise & support available to the communities

Use FIPs to foster reusability and convergence

Warning: too many acronym terms about FAIR

FIP: FAIR Implementation Profile

FER: FAIR Enabling Resources

FIC: FAIR Implementation Community

FDO: FAIR Digital Object

…

FAIR Implementation Community (FIC)

"self-identified organization (composed of more than one person) sharing a common interest that aspires to the creation of FAIR data and services." (Schultes et al. 2020)

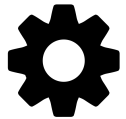
E.G. administrative data community, media content data community; ODISSEI community

Providers

whoever produces data and wants to increase FAIRness (can be organization or individual)

Users

whoever consumes data and has FAIR needs (can be organization or individual)

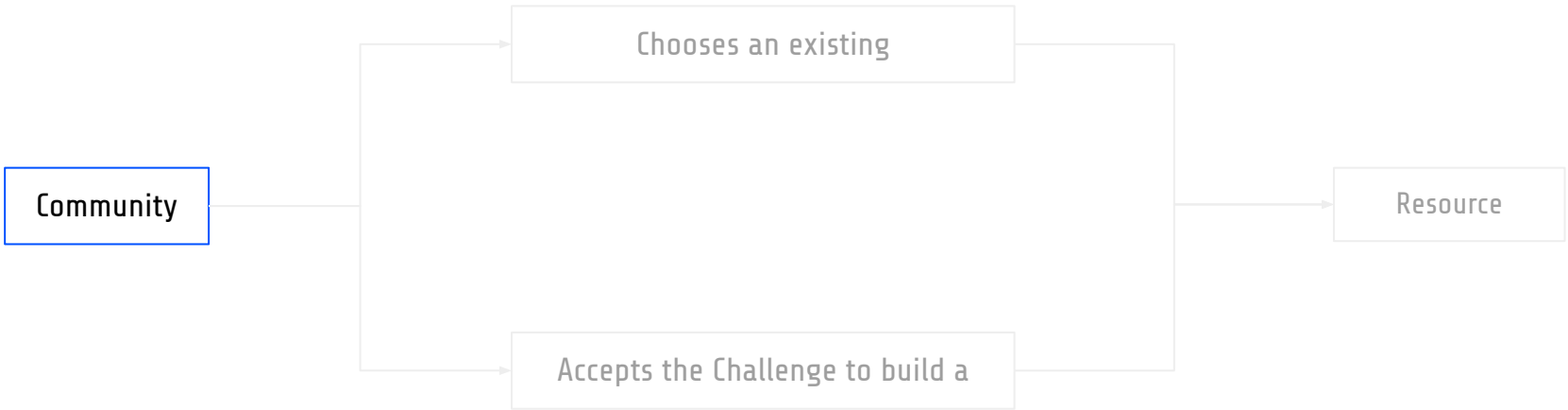


Data supporters

whoever supports FAIR uptake (can be organization or individual)

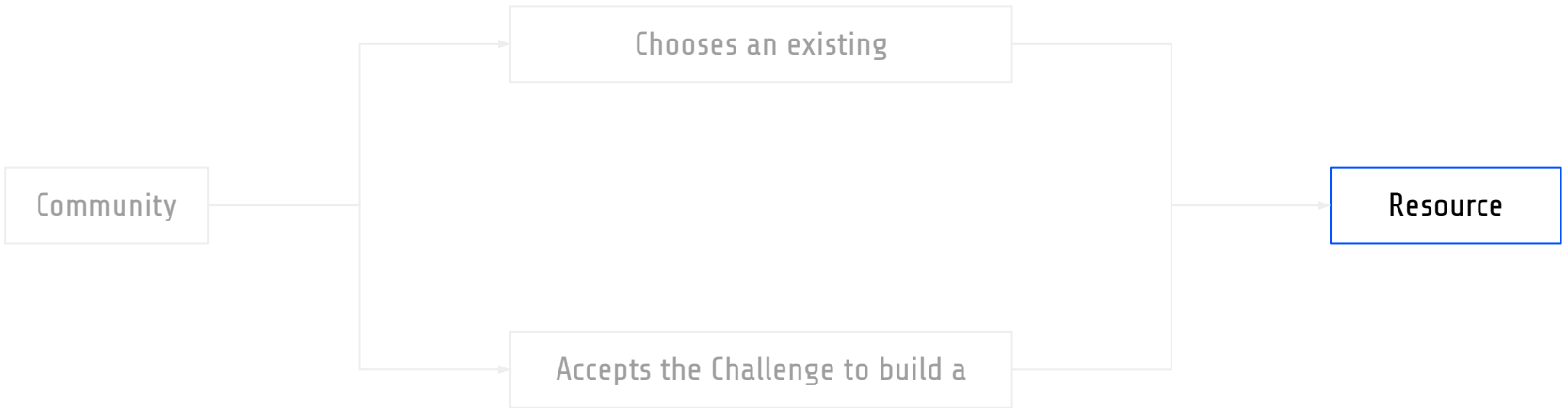
FAIR Implementation Profiles (FIPs)

= collection of FAIR implementation choices, which is FAIR by itself



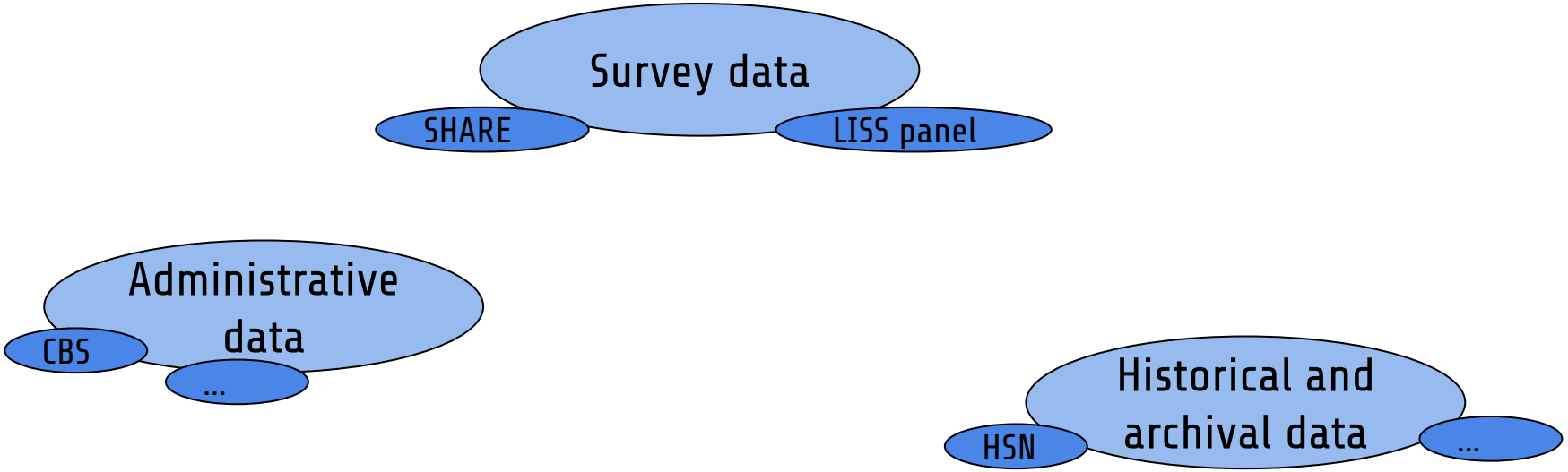
FAIR Implementation Profiles (FIPs)

= collection of FAIR implementation choices, which is FAIR by itself

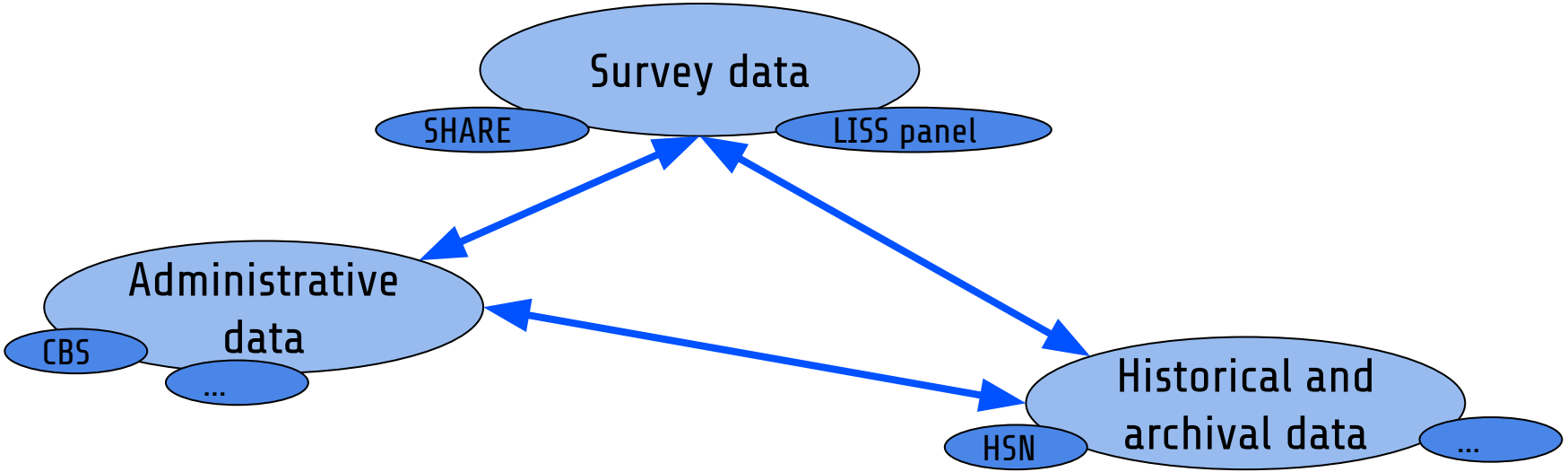


FAIR Implementation Profile (FIP) for Social Science?

Diversity of relevant data for the Social Sciences



Diversity of relevant data for the Social Sciences



FAIR Enabling Resources (FERs)

FER = "any digital object that provides a function needed to achieve some aspect of FAIRness and is explicitly linked to one or more FAIR Principles"

E.G.

Purpose	#	FAIR principle (see Wilkinson et al. 2016)	FER type (Magagna's slides)
Findability	F1	(M)D are assigned a globally unique and persistent identifier	Identifier type

Identifier types: DOI, handle, ePIC, ORCID, ...

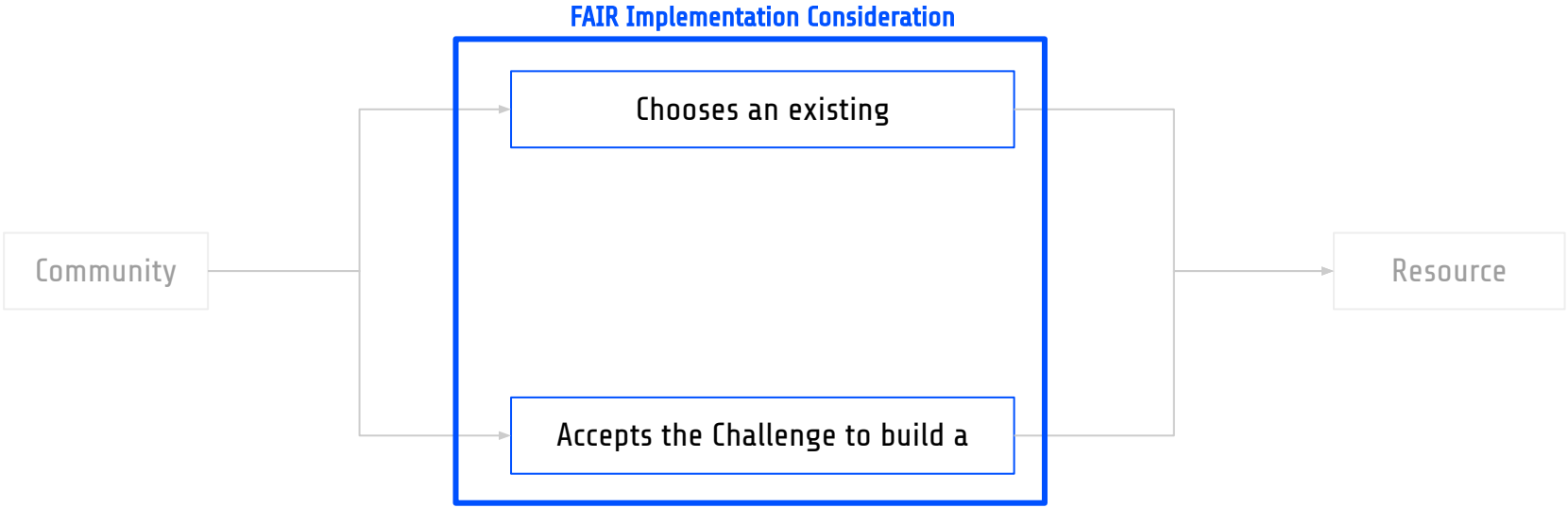
Purpose	#	FAIR principle (see Wilkinson et al. 2016)	FER type (Magagna's slides)
Interoperability	I2	(M)D use vocabularies that follow FAIR principles	Structured vocabularies

Structured vocabularies: ELSST, DDI vocabularies, CESSDA topic classification, ...

See also short article

FAIR Implementation Profiles (FIPs)

= collection of FAIR implementation choices, which is FAIR by itself



FIP → FAIR Convergence Matrix [1]

		Community A	*Community B*	*Community C*	...
F1	DOI	1	1	0	
F1	Handle	0	0	1	
I2	ELSST	0	0	1	
I2	DCAT	0	1	0	
I2	New resource X	0	0	2	

1 = Choice
2 = Challenge

FIP → FAIR Convergence Matrix [2]

		Community A	*Community B*	*Community C*	…
F1	DOI	1	1	0	
F1	Handle	0	0	1	
I2	ELSST	0	0	1	
I2	DCAT	0	1	0	
I2	New resource X	0	0	2	

1 = Choice
2 = Challenge

FIP → FAIR Convergence Matrix [3]

		Community A	*Community B*	*Community C*	...
F1	DOI	1	1	0	
F1	Handle	0	0	1	
I2	ELSST	0	0	1	
I2	DCAT	0	1	0	
I2	New resource X	0	0	2	

1 = Choice
2 = Challenge

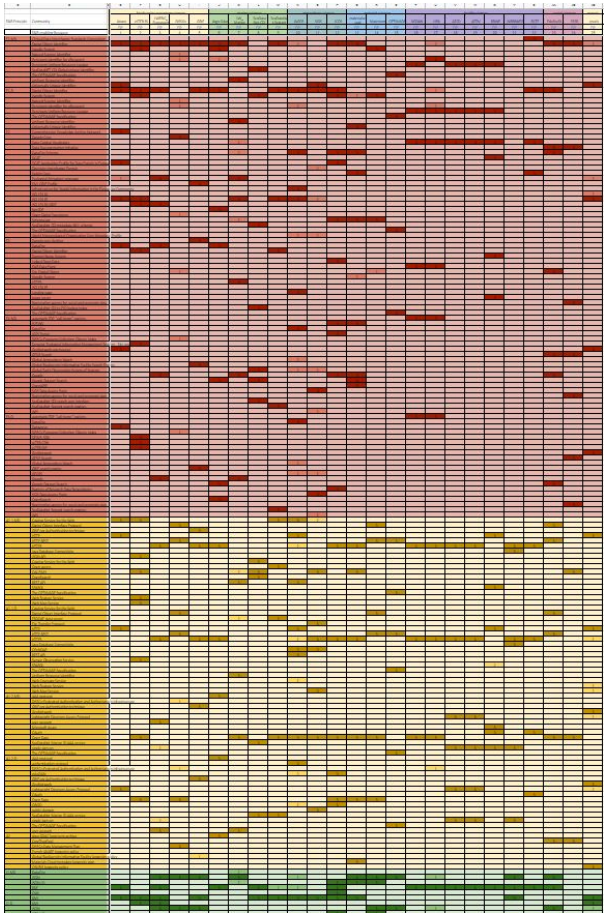
And Community D?

A new community can:

- Declare their profile from scratch
- Reuse an existing one
- Propose a new resource to be added

Why FIPs

- Identify gaps
- Create blueprint
    - Assessment purposes
    - Policy compliance
- Fosters convergence & interoperability
    - Of communities within domains
    - Across domains



The ODISSEI FIP

~ FIP mini-questionnaire ~
Build your FAIR Implementation Profile

Angelica, Shuai, Elena, Tobias

Community description	
Name of Community	*ODISSEI*
Description of Community	ODISSEI
Supporting Links	http://odissei-data.nl/
Research Domain	*Social Sciences, Social and Behavioural Science*
Data Steward	*orcid:0000-0002-6978-5278*
Date of FIP creation	15th Nov 2022

FAIR principle	Question	FAIR enabling resource types	Your answers
F1	What globally unique, persistent, resolvable identifiers do you use for metadata records?	Identifier type	DOI, ORCID,
F1	What globally unique, persistent, resolvable identifiers do you use for datasets?	Identifier type	DOI
F2	Which metadata schemas do you use for findability?	Metadata schema	DDI, DCAT
F3	What is the technology that links the persistent identifiers of your data to the metadata description?	Metadata-Data linking mechanism	DOI
F4	In which search engines are your metadata records indexed?	Search engines	Zenodo,
F4	In which search engines are your datasets indexed?	Search engines	Google Dataset Searech
A1.1	Which standardized communication protocol do you use for metadata records?	Communication protocol	HTTPS, OAI-PMH Schema
A1.1	Which standardized communication protocol do you use for datasets?	Communication protocol	HTTPS
A1.2	Which authentication & authorisation technique do you use for metadata records?	Authentication & authorisation technique	None
A1.2	Which authentication & authorisation technique do you use for datasets?	Authentication & authorisation technique	TODO
A2	Which metadata longevity plan do you use?	Metadata longevity	DataCite DOI Policy
I1	Which knowledge representation languages (allowing machine interoperation) do you use for metadata records?	Knowledge representation language	JSON-LD, OWL, RDFS
I1	Which knowledge representation languages (allowing machine interoperation) do you use for datasets?	Knowledge representation language	None
I2	Which structured vocabularies do you use to annotate your metadata records?	Structured vocabularies	ELSST (Social Science Thesaurus), CESSDA Topic Classication
I2	Which structured vocabularies do you use to encode your datasets?	Structured vocabularies	TODO
I3	Which models, schema(s) do you use for your metadata records?	Metadata schema	DC
I3	Which models, schema(s) do you use for your datasets?	Data schema	TODO
R1.1	Which usage license do you use for your metadata records?	Data usage license	Creative Cooms (CC0)
R1.1	Which usage license do you use for your datasets?	Data usage license	TODO
R1.2	Which metadata schemas do you use for describing the provenance of your metadata records?	Provenance model	TODO
R1.2	Which metadata schemas do you use for describing the provenance of your datasets?	Provenance model	TODO

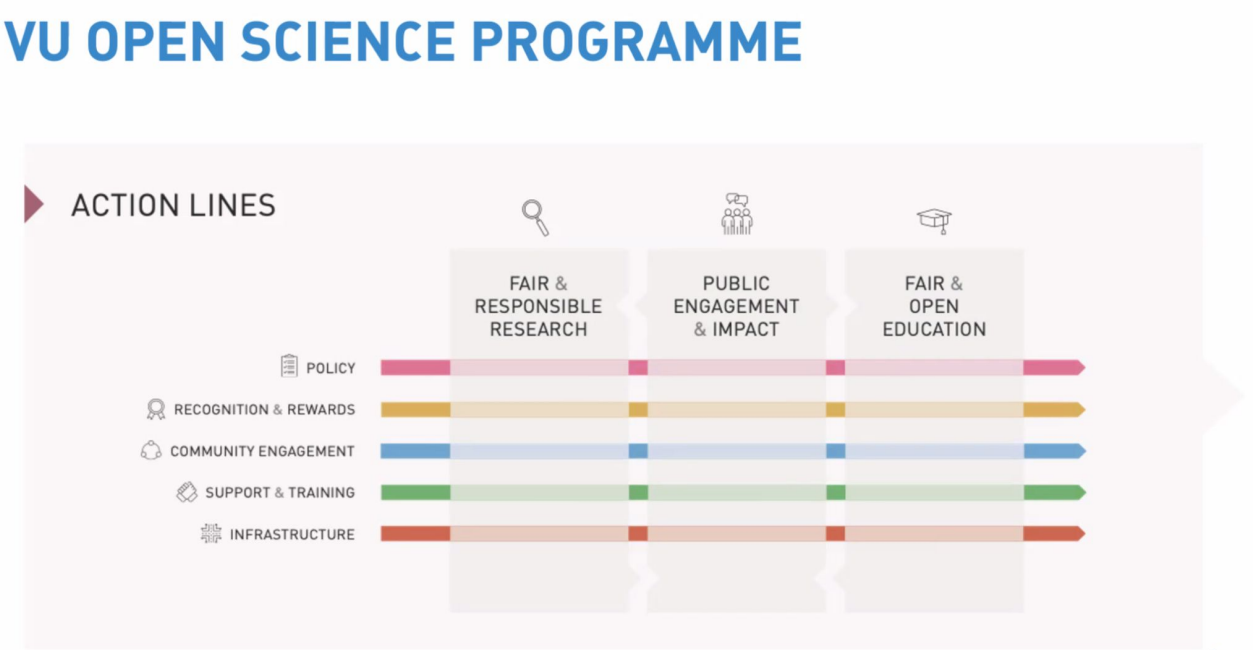
- First attempt creating a FIP from scratch
- Missing resources and communities
- Unclear use of FER
- Lack of connections
- Data analysis
- Convergent analysis

Why FIPs @ VU

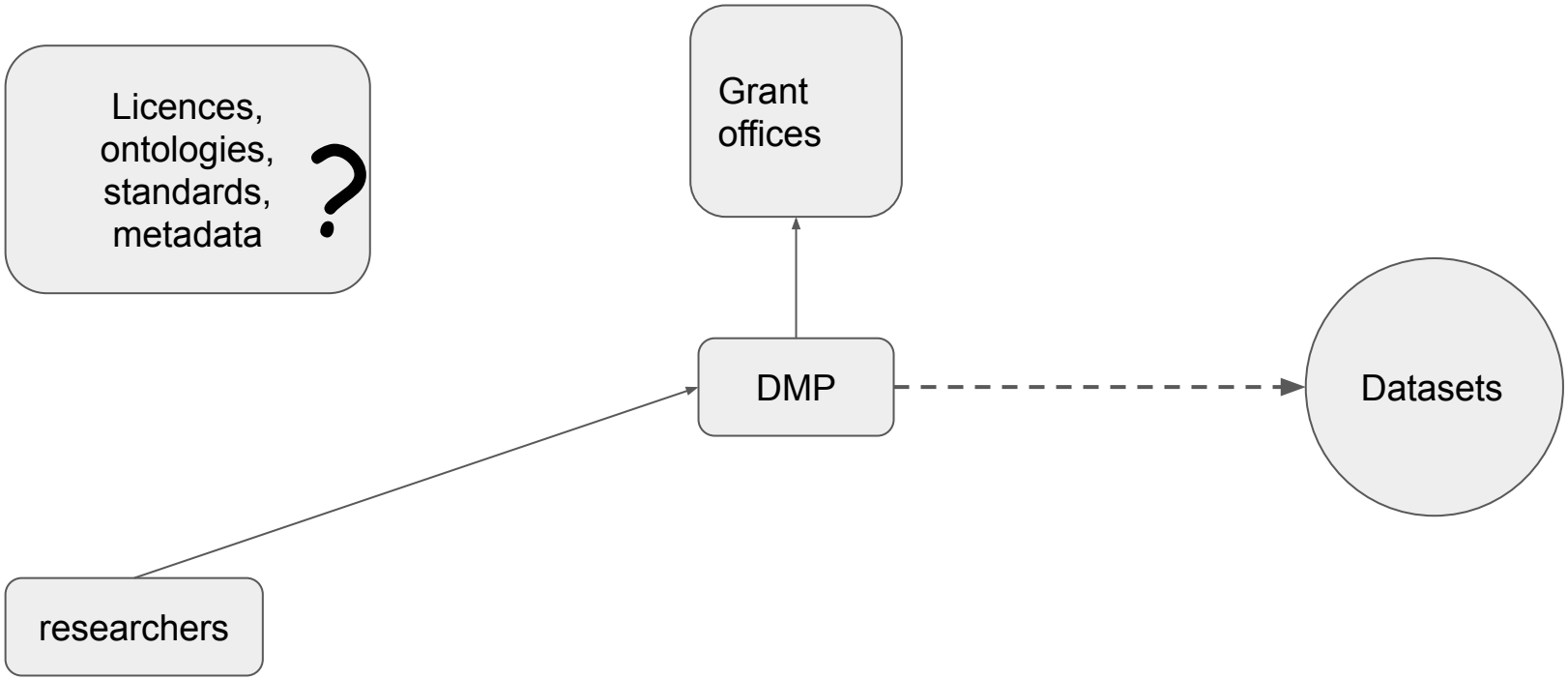
- Connections:
  - Open Science team
  - CS department
  - Institutes and groups in social science @ VU
- Projects
  - maDMP (machine-actionable data management plans) with Lena, Tycho and Stephanie.
    - A small student group project in Janurary
    - An webapp/wizard for maDMPs using FIPs
  - FAIR assessment of datasets in social science (two bachelor students)
- Training (from next year)
  - FIP workshops, community training, website, etc.

Stealing slides from Sander :)

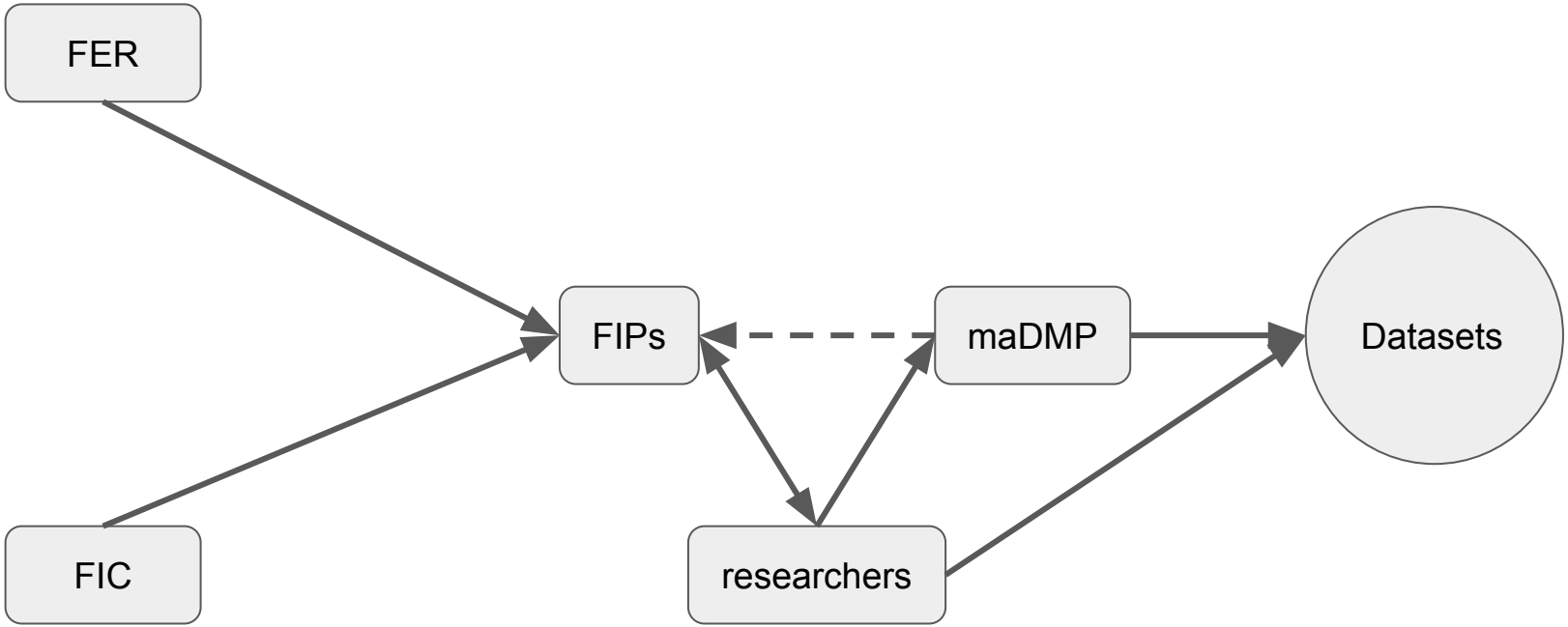
- Community Engagement
- Support and Training



Zooming out for the big picture



Zooming out for the big picture



Twitter @FAIRexperthub

Get in touch!

shuai.wang@vu.nl

UCDS, Computer Science Department, VU

10A-07, NU building