Narrowing the gap between international FAIR Best Practices for Open Science and University implementation

Kristina Hettne | Centre for Digital Scholarship, Leiden University Library

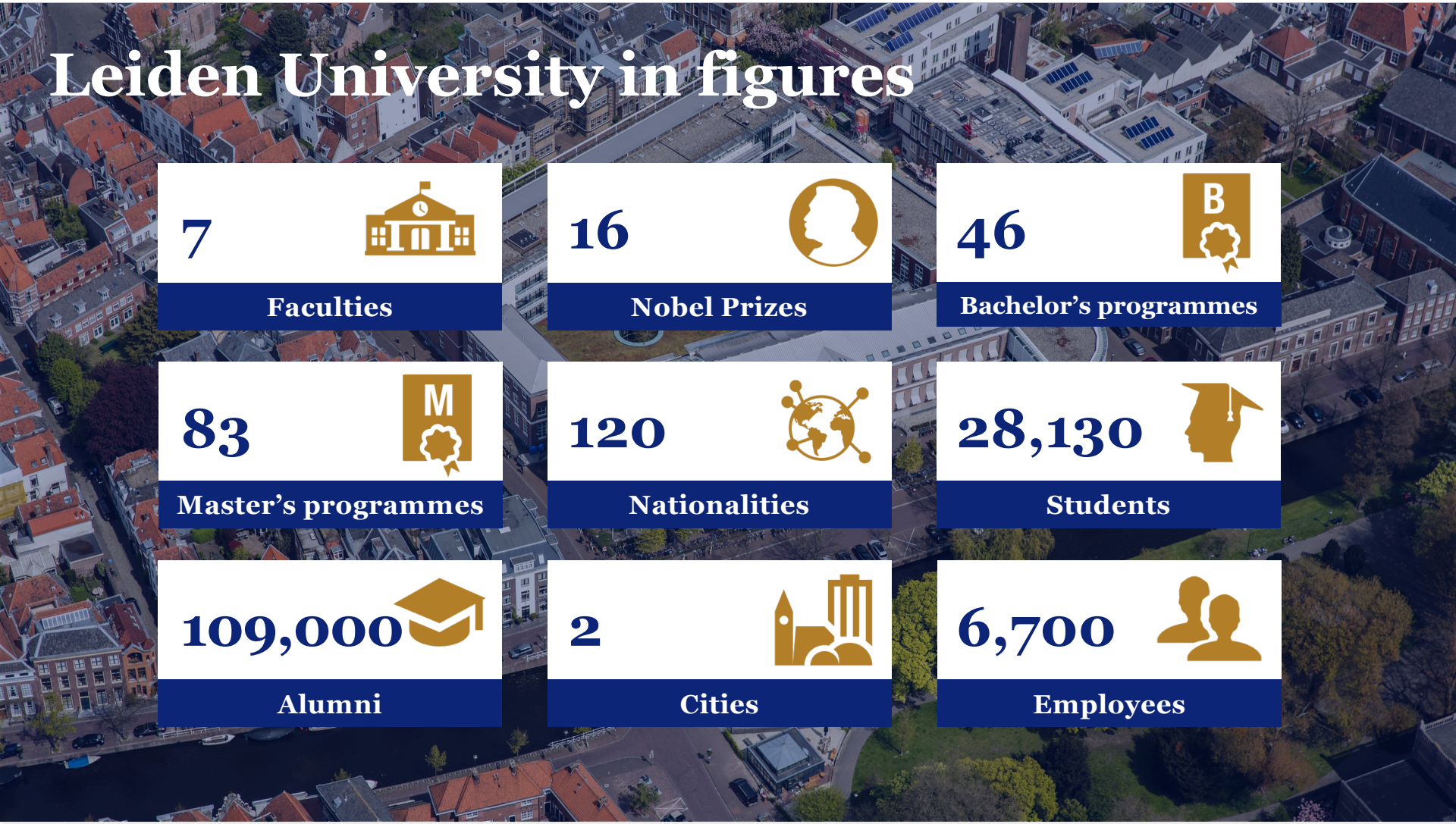
Focus on Open Science, Kaunas

29 Oct 2019

Universiteit Leiden
The Netherlands

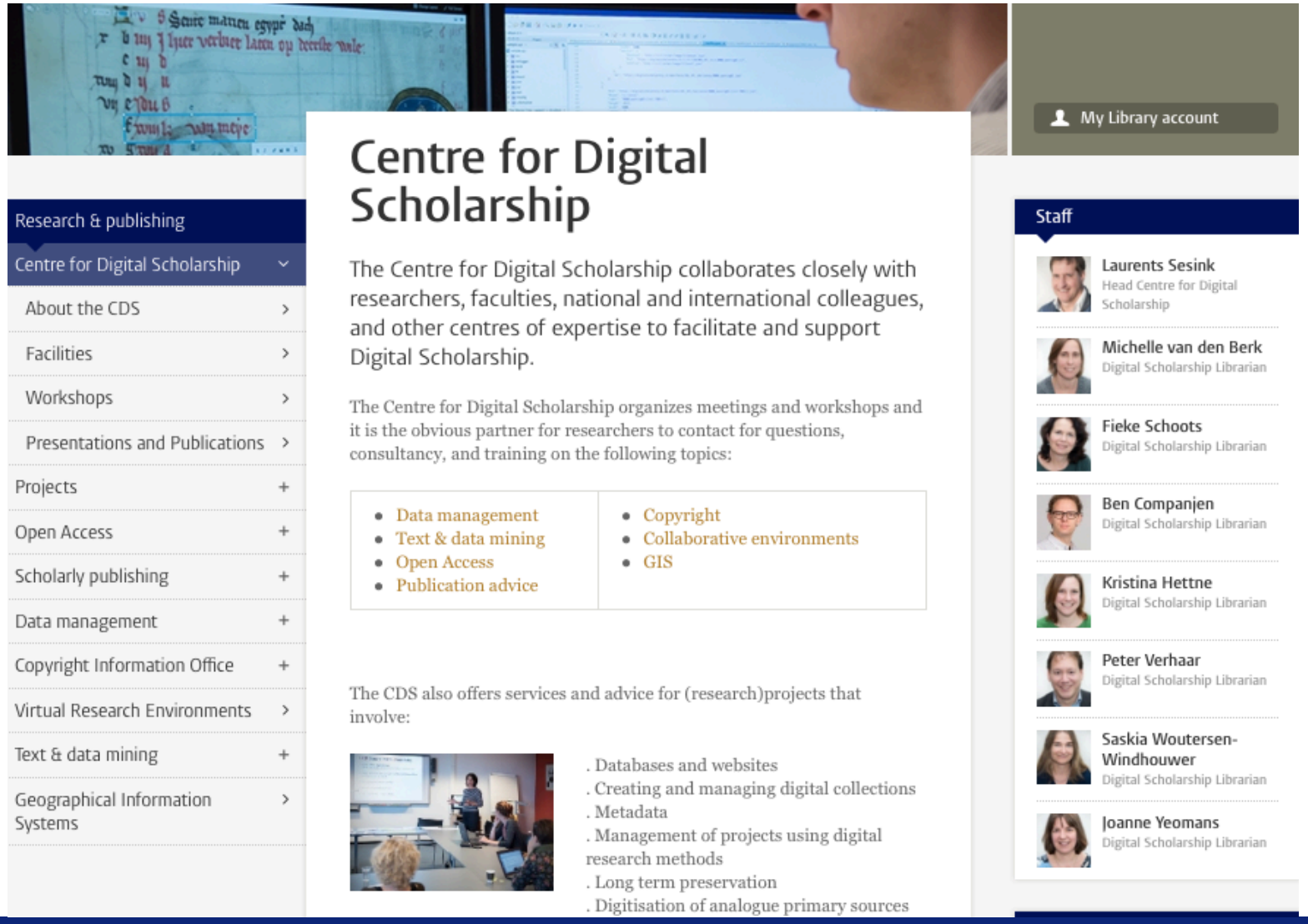
CC BY

Leiden University in figures



7	16	46	
Faculties	Nobel Prizes	Bachelor's programmes	
83	120	28,130	
Master's programmes	Nationalities	Students	
109,000	2	6,700	
Alumni	Cities	Employees	

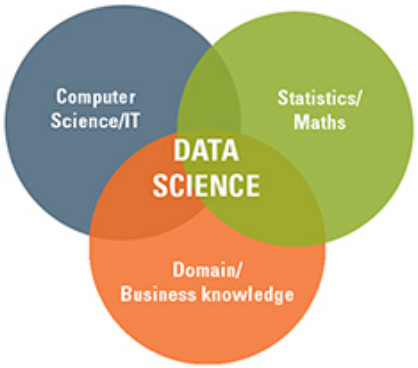
CDS provides supports for Open Access, data management, use of digital data and copyright



Leiden University Institutional Plan

2015-2020

Supporting academics in the transition to a more interactive academic environment: ***Open Science***



- Facilitate **sharing** and **collaboration**
- Provide **state-of-the-art** support and infrastructure to facilitate **Data Science**

Open Access policy @ Leiden University (December 2017)

- Implement the Open Access goals of the Dutch government: 100% 2020
- Stimulate Gold OA + upload publications to institutional repository

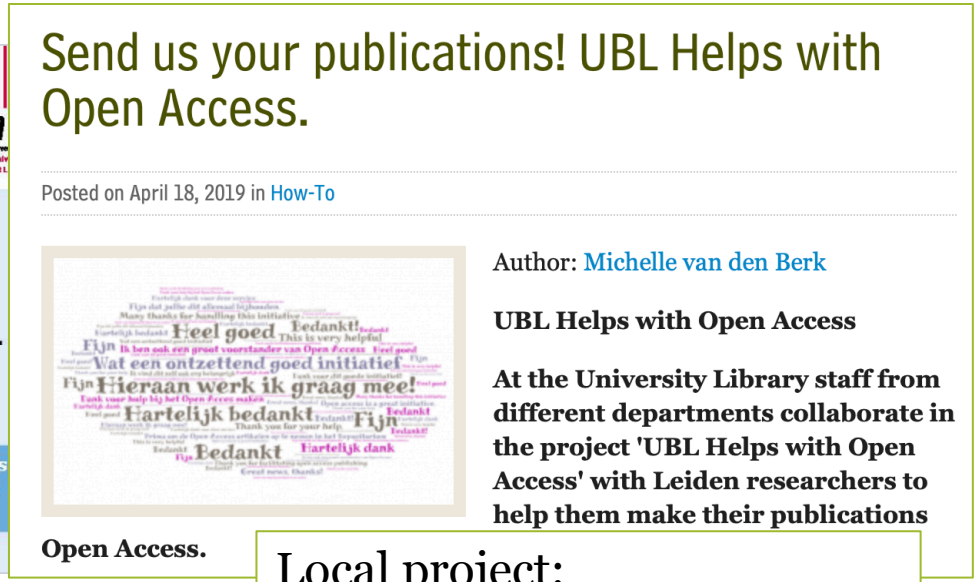
75% of Leiden University publications now Open Access available

Published on 21 October 2019

Researchers of Leiden University (excl. LUMC) are collectively using the Leiden Repository to make their publications Open Access available. They have, thereby, shown 75% Open Access publication to be an attainable goal for the 2017 and 2018 years. This has greatly increased the visibility and international accessibility of new Leiden University publications. The Centre for Digital Scholarship (CDS) at Leiden University Libraries (UBL) helps researchers with the Open Access publication process.

Projects to implement the Leiden Open Access policy

National project applying Dutch copyright law: upload publisher's version in institutional repository



Local project: upload accepted manuscripts in institutional repository

Open Access training @ Leiden University

- Open Access and Open Science workshops
- Information material

Graduate School of Science

Programme

Open Science Opens Careers

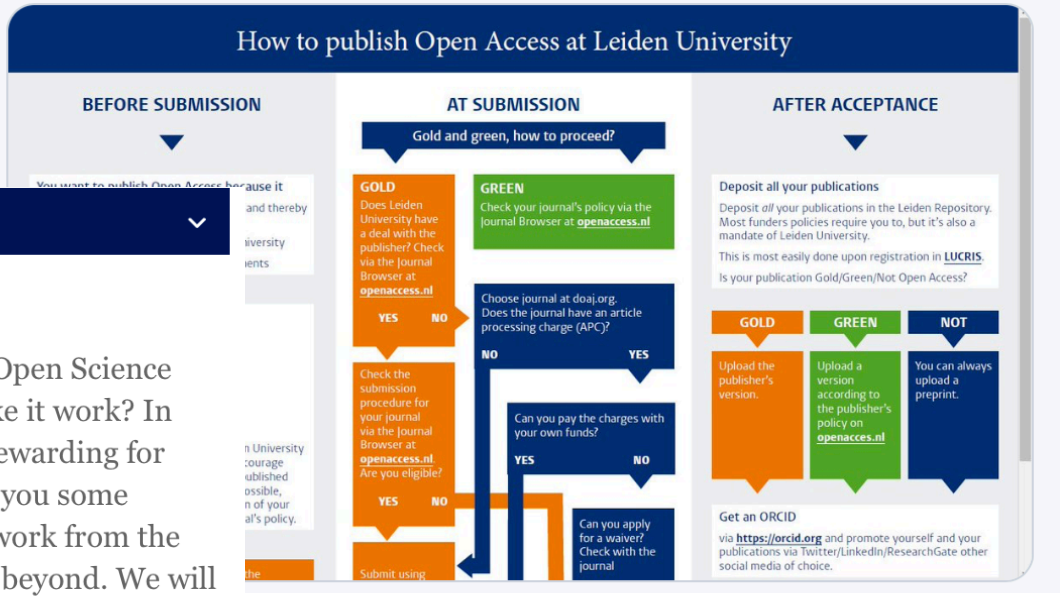
Centre for Digital Scholarship (CDS)

Open Science: just science done right. Right? What does Open Science actually mean? What is in it for you and how can you make it work? In this workshop we will explore how Open Science can be rewarding for researchers and the research community alike. We'll give you some practical guidance on tools and ways of opening up your work from the start of your research project to the final conclusions and beyond. We will explore opportunities to increase your impact and visibility. We highlight the debate on new reward systems in science and how they may open new career paths.

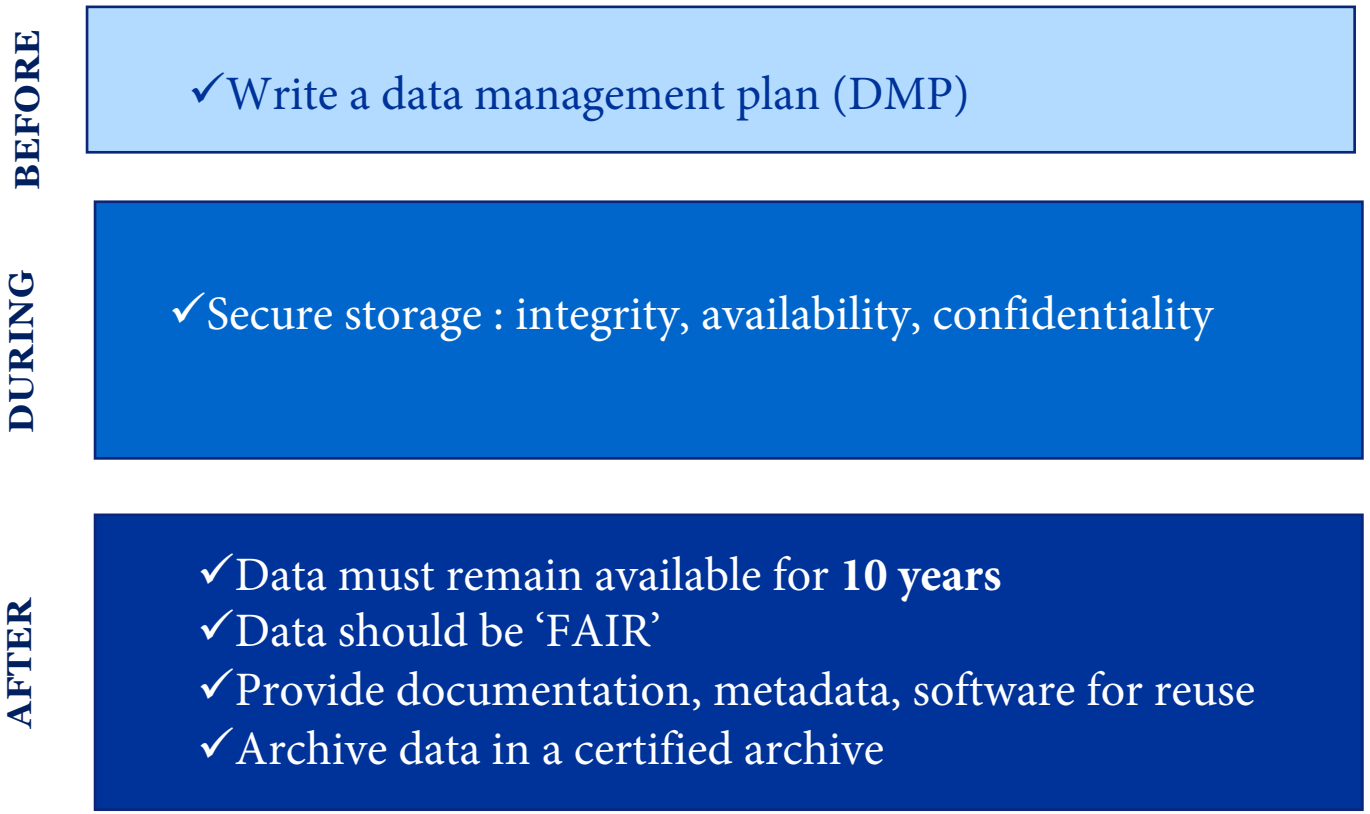
Slides: https://www.slideshare.net/UBL-CDS/open-science-opens-careers

Centre for Digital Scholarship @CDSLeiden · Oct 22

How to publish Open Access @UniLeiden. In this bit.ly/2VYXAEy flowchart you will find a step by step explanation on how you can publish your article Open Access. #OpenAccessWeek OA increases international visibility, accessibility and thereby the impact of research.



Research Data Management policy @ Leiden University (April 2016)



https://doi.org/10.2218/ijdc.v12i2.575

Box 2 | The FAIR Guiding Principles

To be Findable:
F1. (meta)data are assigned a globally unique and persistent identifier
F2. data are described with rich metadata (defined by R1 below)
F3. metadata clearly and explicitly include the identifier of the data it describes
F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:
A1. (meta)data are retrievable by their identifier using a standardized communications protocol
A1.1 the protocol is open, free, and universally implementable
A1.2 the protocol allows for an authentication and authorization procedure, where necessary
A2. metadata are accessible, even when the data are no longer available

To be Interoperable:
I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
I2. (meta)data use vocabularies that follow FAIR principles
I3. (meta)data include qualified references to other (meta)data

To be Reusable:
R1. meta(data) are richly described with a plurality of accurate and relevant attributes
R1.1. (meta)data are released with a clear and accessible data usage license
R1.2. (meta)data are associated with detailed provenance
R1.3. (meta)data meet domain-relevant community standards

doi:**10.1038/sdata.2016.18**

Research Data Management training @ Leiden University

- How to write a DMP workshops
- Workshops tailored to the needs of specific research institutes
- Hands-on training (Bring Your Own Data workshops)

Centre for Digital Scholarship (CDS)

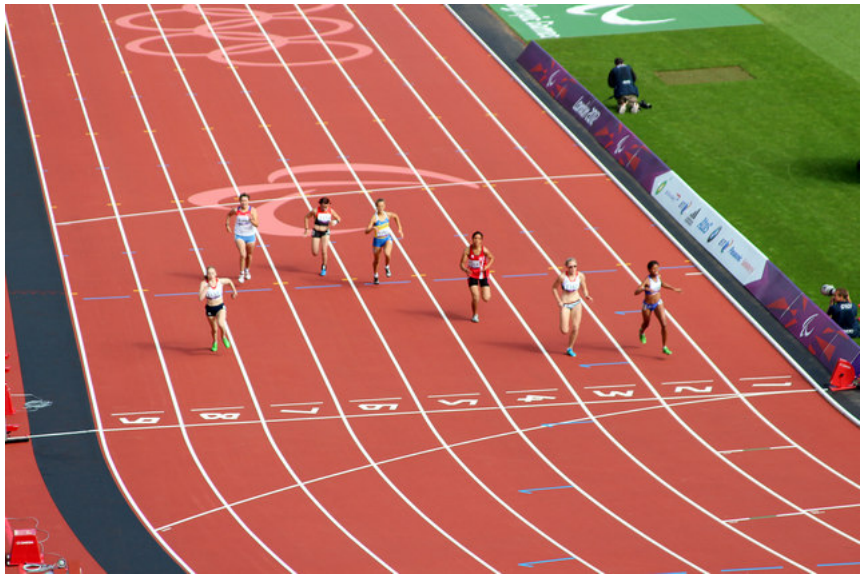
Your research is going well, you've collected lots of data, you've started writing your thesis, you're almost ready to publish, or you have already. What can go wrong? Good research data management can help you save time and stress during your research, help others (including you!) re-use your data in future, and bolster your reputation and career. We take a fun look at why good data management is important, and how to do it. Have you thought about making your data more FAIR? Sometimes it just takes a few steps.

Material: https://github.com/fair-ds-curriculum/byod-leidenuniv-lib

Slides: https://www.slideshare.net/UBL-CDS/let-your-researchbloom-practical-steps-for-fair-data

Challenge: keeping up to speed with national and international best practices

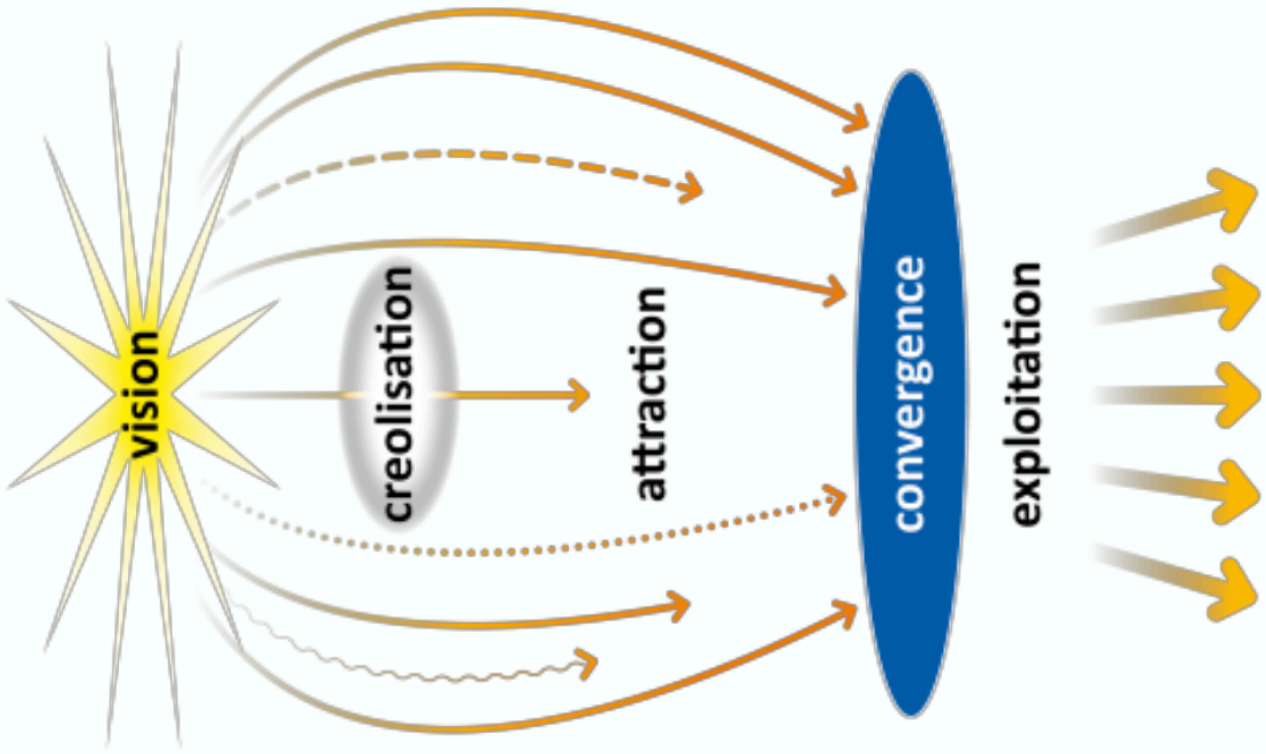
- Universities are complex organisations – need to translate best practices to the level of faculties, institutes, individual researchers
- Change takes time



This Photo by Unknown Author is licensed under CC BY-SA

Solution: the CDS actively engages with organisations to close the feedback loop

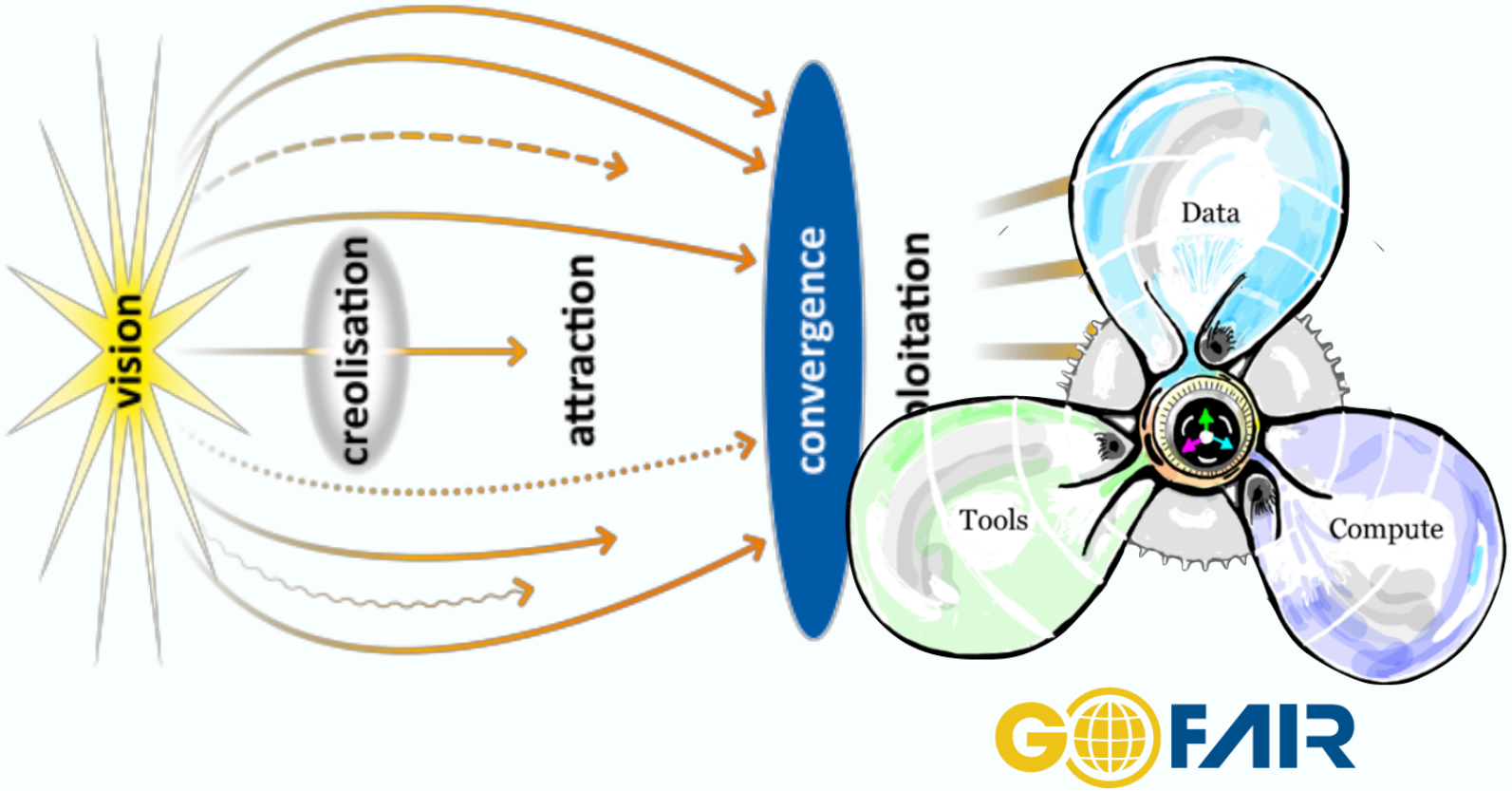
- At national and international level
- Goals:
  - Help drive change
  - Look after the interests of the Leiden researcher



Common Patterns in Revolutionary Infrastructures and Data

Peter Wittenburg, Max Planck Computing and Data Facility
George Strawn, US National Academy of Sciences
February 2018
https://www.rd-alliance.org/sites/default/files/Common_Patterns_in_Revolutionising_Infrastructures-final.pdf

8



Common Patterns in Revolutionary Infrastructures and Data

Peter Wittenburg, Max Planck Computing and Data Facility
George Strawn, US National Academy of Sciences
February 2018
https://www.rd-alliance.org/sites/default/files/Common_Patterns_in_Revolutionising_Infrastructures-final.pdf

9

Picture from Erik Schultes, GO FAIR

GO FAIR to drive convergence

Leiden University

The principles

Coordination

Implementation

GO FAIR Initiative

Home › GO FAIR Initiative

GO FAIR Initiative

- Vision and Strategy
- GO BUILD
  - FAIRification Process
- GO CHANGE
- GO TRAIN
- Countries
  - Guidelines for GO FAIR involvement on a country level
- Governance
  - Steering Committee
  - Executive Board
  - Stakeholder Forum
- GO FAIR Offices
- Contact

GO FAIR is a bottom-up, stakeholder-driven and self-governed initiative that aims to implement the **FAIR data principles**, making data Findable, Accessible, Interoperable and Reusable. It offers an open and inclusive ecosystem for individuals, institutions and organisations working together through **Implementation Networks** (INs). The INs are active in three activity pillars: **GO CHANGE**, **GO TRAIN** and **GO BUILD.**

Creolization Attractors Convergence IFDS

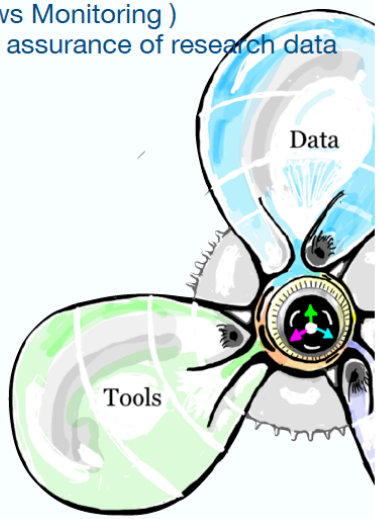
LUMC
UMC Utrect
UMCG
WUR
Maastricht University
BioSemantics Group
UCSD
BioCom
NDS
ANDS
NIH
FAIRdICT
DTL
LERU
CGIAR
DANS
RDA
Metrics Group
F1000
Force 11
Nerdalize
ODEX
Lorentz Center
Personal Health Train
ReproNIM
EOSC
EUDAT
OpenAIRE
FOSTER
CODATA
EDISON
BioSB
HRB
ZonMW
Elsevier
Springer-Nature

GO CHANGE

GO TRAIN

GO BUILD

ASTRON
Metrology
Chemistry
Nano Research
NOMAD
FAIR Funding
GO FAIR Brazil
CO-OPERAS
FAIR Journalism (Fake News Monitoring)
Reproducibility and quality assurance of research data
FAIR Funders
Neubias
Metabolomics
Vaccine IS
Rare disease
Training Frameworks
Training Curriculum
Seasons Schools
AGU Enabling FAIR Data
System Terre
Sea Data Net
FAIR Pointer
BiodiFAIRse
Agriculture & Food Systems
INOSIE
EcoSoc
PhenoMeNal
Discovery IN
GERDI
OPEDAS
C2CAMP
Personal Health Train
CBS (Economics)
Sustainability Research
Annotation



GO FAIR

2017 Q1 Q2 Q3 Q4 2019

DI Special Issue Oct 2019

FEATURED ARTICLES

MORE +

The FAIR Principles: First Generation Implementation Choices and Challenges

Author : Barend Mons; Erik Schultes; Fenghong Liu; Annika Jacobsen

Institution : Leiden University Medical Center, Leiden 2333 ZA, The Netherlands; GO FAIR Inter…

Doi : 10.1162/dint_e_00023

Abstract (views 93) | Full Text PDF | Full Text HTML

FAIR Principles: Interpretations and Implementation Considerations

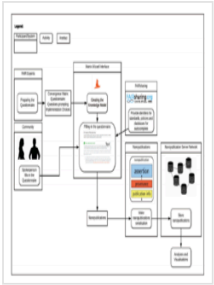
Author : Annika Jacobsen; Ricardo de Miranda Azevedo; Nick Juty; Dominique Batista; Simon C…

Institution : Leiden University Medical Center, Leiden 2333 ZA, The Netherlands; Institute of Dat…

Keywords : FAIR guiding principles; FAIR implementation; FAIR convergence; FAIR communitie…

Doi : 10.1162/dint_r_00024

Abstract (views 105) | Full Text PDF | Full Text HTML



FAIR Convergence Matrix: Optimizing the Reuse of Existing FAIR-Related Resources

Author : Hana Pergl Sustkova; Kristina Maria Hettne; Peter Wittenburg; Annika Jacobsen; Tobia…

Institution : GO FAIR International Support and Coordination Office, Leiden, The Netherlands; C…

Keywords : FAIR Implementation Choices and Challenges； Convergence； FAIR Communities

Doi : 10.1162/dint_a_00038

Abstract (views 114) | Full Text PDF | Full Text HTML

Data Intelligence

Host: National Science Library, Chinese Academy of Sciences

Publisher: National Science Library, Chinese Academy of Sciences

Co-Editors-in-Chief:

James Hendler , Huizhou Liu , Ying Ding

Executive Editors-in-Chief:

Guilin Qi , Yan Zhao

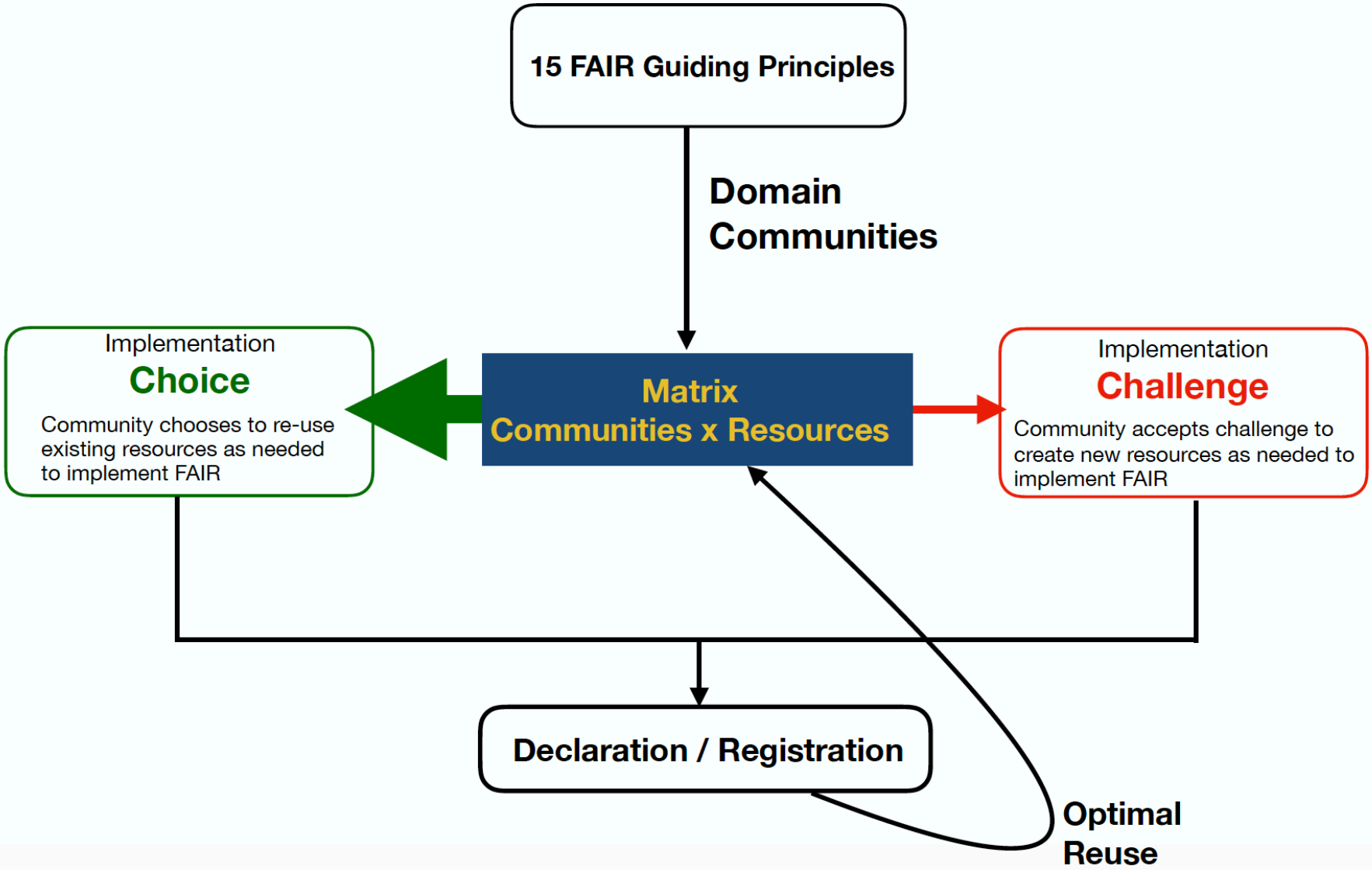
29 original papers

http://www.data-intelligence-journal.org

Example: From FAIR principles to implementation – the Convergence Matrix

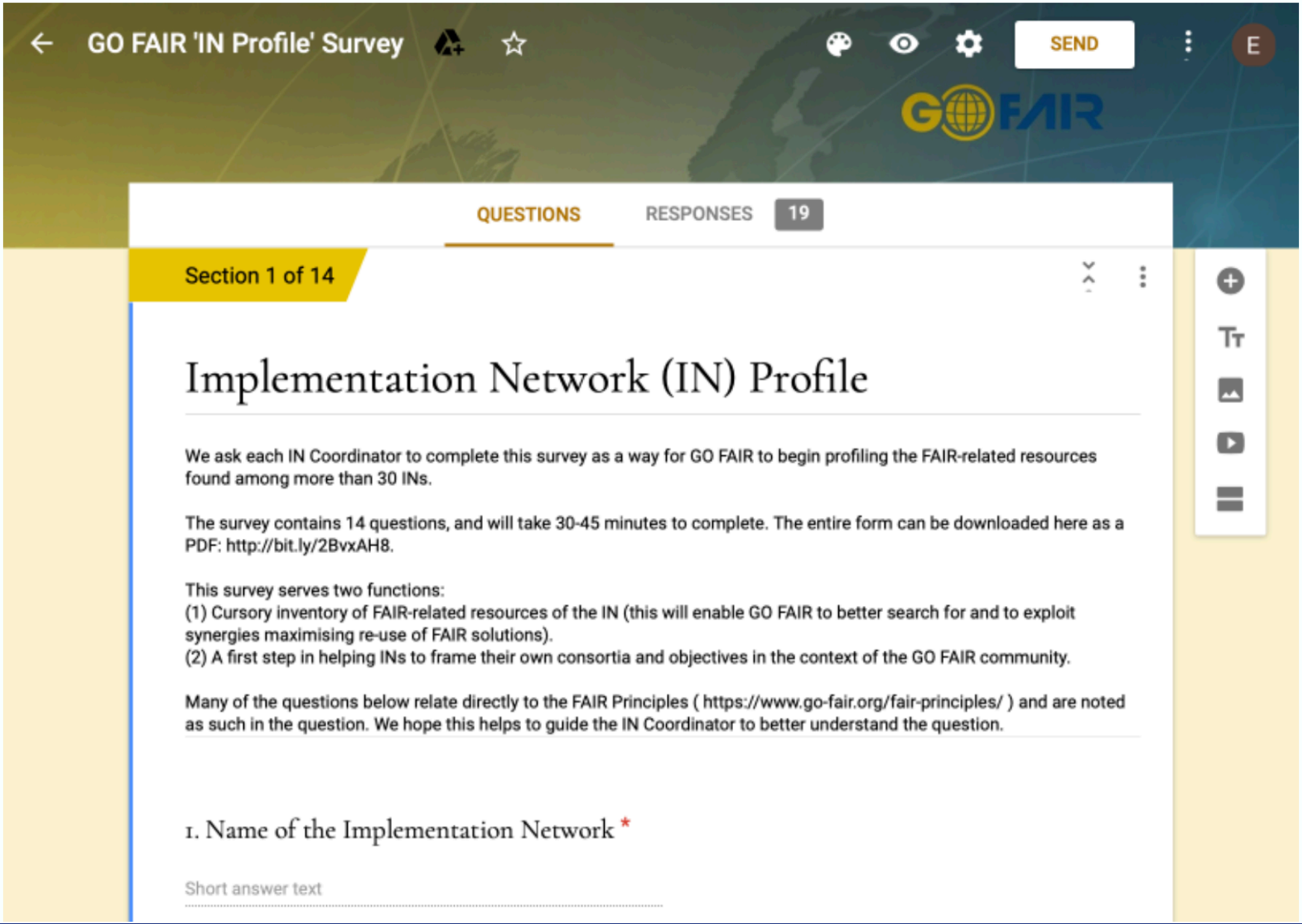
Picture from Erik Schultes, GO FAIR

Community choices and challenges



Survey https://docs.google.com/forms/d/1Oug6GowuG1jNZNsjkIXOeEvPbUrhyuS_F-d185SOy6A/edit

Kristina Hettne, CDS University Library



Survey https://docs.google.com/forms/d/1Oug6GowuG1jNZNsjkIXOeEvPbUrhyuS_F-d185SOy6A/edit

Kristina Hettne, CDS University Library

SUBJECT	PREDICATE	OBJECT	
name of IN (UPRI)	has-coordinator	ORCID	
name of IN (UPRI)	has-participant	ORCID	
name of IN (UPRI)	has-member-organisation	VIVO / CrossRef	
name of IN (UPRI)	uses-repository	CTS?	
name of IN (UPRI)	uses-registry-service	PW ?	F1
name of IN (UPRI)	provides-registry-service		F1
name of IN (UPRI)	uses-data-format	format-PID	F2
name of IN (UPRI)	provides-data-format	format-PID	F2
name of IN (UPRI)	provides-access-protocol	format-PID	A1
name of IN (UPRI)	uses-access-protocol	protocol-PID	A1
name of IN (UPRI)	has-persistence-policy	policy	F1 / A2
name of IN (UPRI)	is found by	Search engine	F4
name of IN (UPRI)	uses-term-system	Term System-PID	I
name of IN (UPRI)	provides-term-system	Term System-PID	I
name of IN (UPRI)	uses-license	MR-license ID	R1.1
name of IN (UPRI)	uses-metadata-format	format-PID	R1.2
name of IN (UPRI)	provides-meta-data-format	Format-PID	R1.2
name of IN (UPRI)	provides-training-material	Resource-ID	
name of IN (UPRI)	uses-uses-training-material	Resource-ID	
name of IN (UPRI)	provides-DS-tools	Resource-ID	
name of IN (UPRI)	uses-DS-tools	Resource-ID	
name of IN (UPRI)	uses-workspace-tool	Resource-ID	
name of IN (UPRI)	Provides-workspace-tool	Resource-ID	

GO FAIR IN Profile Matrix

January 15-16, Leiden

Survey https://docs.google.com/forms/d/1Oug6GowuG1jNZNsjkIXOeEvPbUrhyuS_F-d185SOy6A/edit
Matrix https://docs.google.com/spreadsheets/d/1MUZn7uh4x5YLPjqxi-V8XubsSEEonQWvx2jBlcyyNdU/edit#gid=0

IN Profile Matrix

File Edit View Insert Format Data Tools Add-ons Help All changes saved in Drive

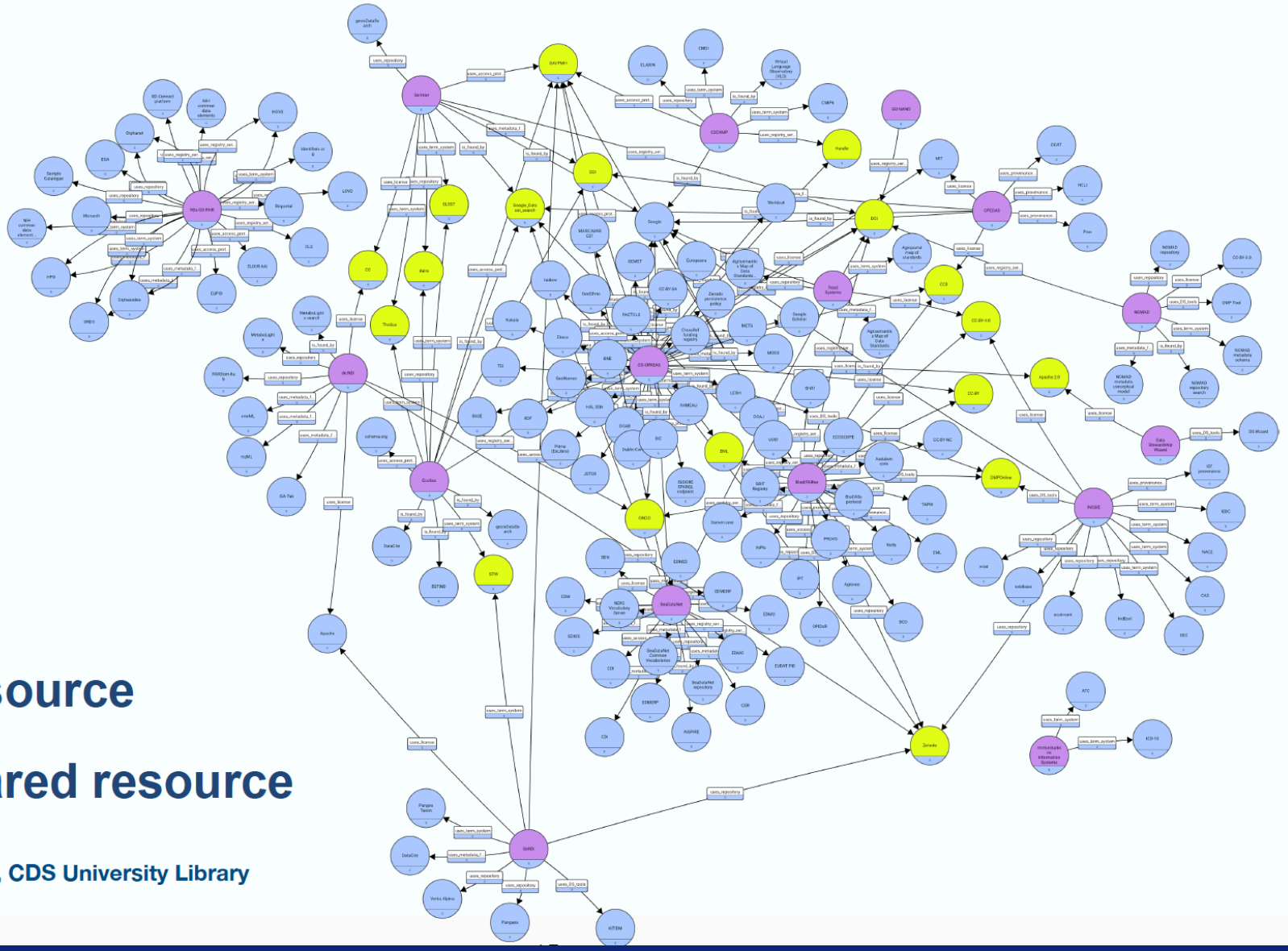
100% $ % .0 .00 123 Helvetica ... 10

	A	B	C	D	E	F	G	H	I
1	FAIR Implementation Matrix								
2	*On the OSF*	https://osf.io/n7uwp/							
3	*Red indicates waist of hourglass*								
4	*Blue is an Implementation **Choice***								
5	*Orange is Implementation **Challenge***								
6	*Green highlight indicates a service provided by the IN or spin-off*								
7	*Blank cell is not relevant for IN*								
8	**FAIR Principle**	**Services**	**Component**	**Most used**	**C2CAMP**	**OPEDAS**	**PHT**	**Rare-Diseases**	**GERI**
9		**central to all**	DOIP	DOIP	DOIP	DOIP	DOIP	DOIP	
10		**central to all**	Metadata format	RDF		RDF	RDF	RDF	
11		**central to all**	Metadata access protocol			LDP/FDP	LDP/FDP	LDP/FDP	
12		**central to all**	Metadata core elements	TBD on M4M		TBD on M4M	TBD on M4M	TBD on M4M	
13		**Technology**	Data Format			RDF for interop.	RDF for interop.	RDF for interop.	
14		**Technology**	Data Access Protocols (MR/A)			LDP/FDP	PHT-standard	PHT-standard	
15		**Technology**	Computer-actionable license description language			RDF	RDF	RDF	
16		**Tooling**	Repository (Data/Metadata)		DONA	IFDS Data Station	IFDS Data Station	ERN?	GERI
17		**Tooling(Repository)**	https://www.dataone.org						
18		**Tooling**	Registry Service		DONA	IFDS Station Registry	IFDS Station Registry	ERN?	
19		**tooling**	Metadata forms/creators			CEDAR/CASTOR			
20		**Tooling**	Search capability		DOIP	IFDS Station Registry	IFDS Station Registry	IFDS Station Registry	
21		**Policy**	Persistence Policy			TBD	TBD	TBD	
22		**Technology**	Computer-actionable policy description language			RDF	RDF	RDF	
23		**Tooling**	License protocols			TBD	TBD	TBD	

14

GO FAIR IN Profile Matrix

January 15-16, Leiden

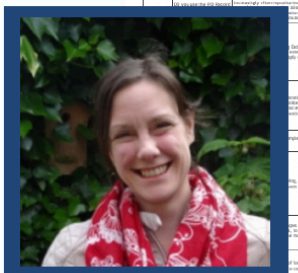


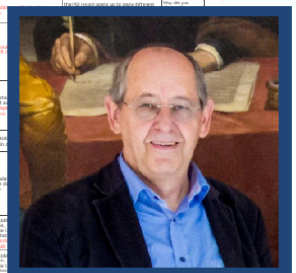
- INs
- Resource
- Shared resource

Kristina Hettne, CDS University Library

Matrix Development
June 13 2019

https://osf.io/n7uwp/


Kristina Hettne


Peter Wittenburg

Tobias Kuhn

Robert Pergl

Pete McQuilton

Universiteit Leiden

RDA RESEARCH DATA ALLIANCE

VU

CZECH TECHNICAL UNIVERSITY IN PRAGUE

UNIVERSITY OF OXFORD

MAX-PLANCK-GESELLSCHAFT

nanopublication
- assertion
- provenance
- publication info

DSW DATA STEWARDSHIP WIZARD

FAIRsharing.org standards, databases, policies

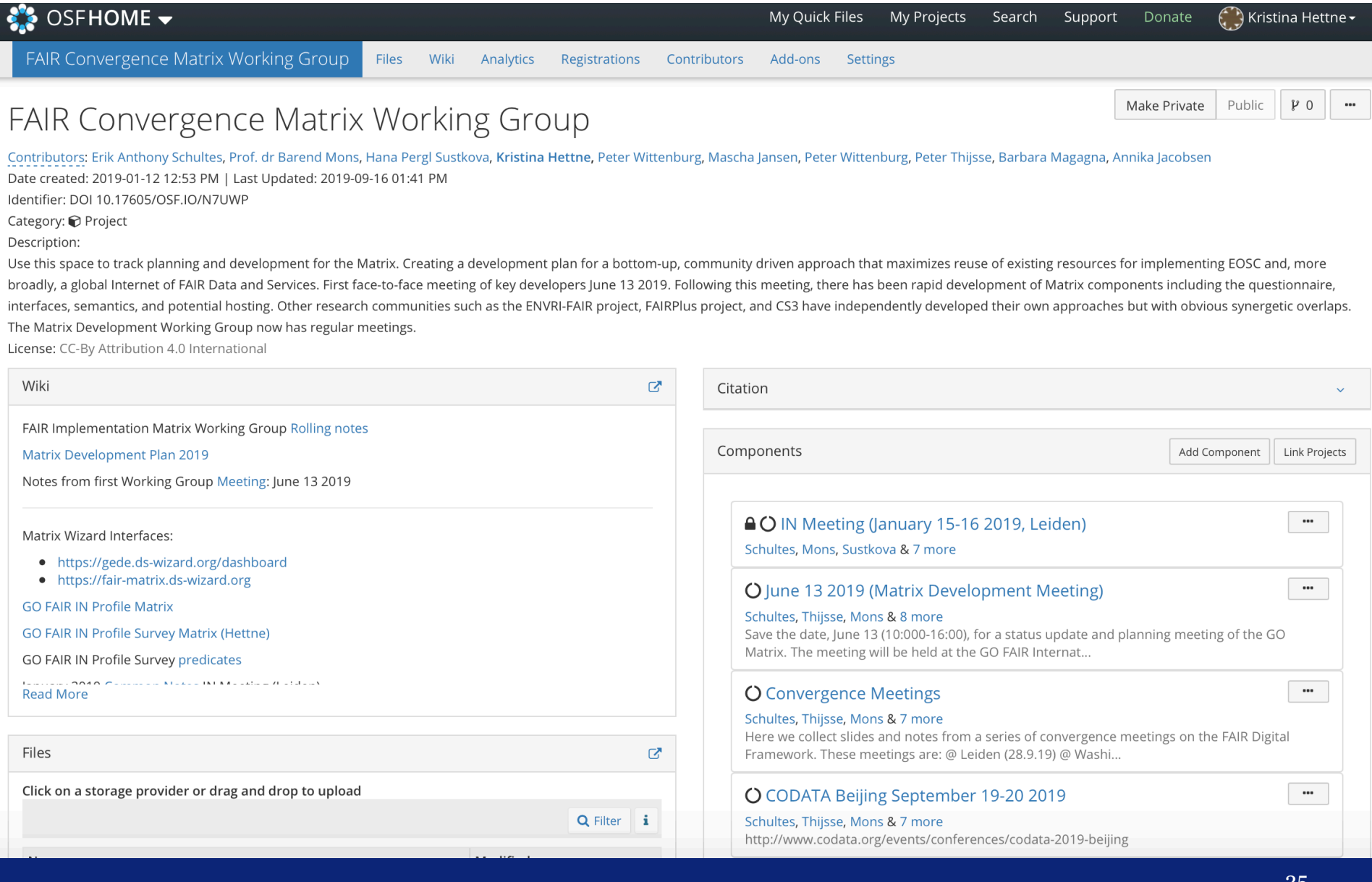
Barbara Magagna

umweltbundesamt ENVIRONMENT AGENCY AUSTRIA

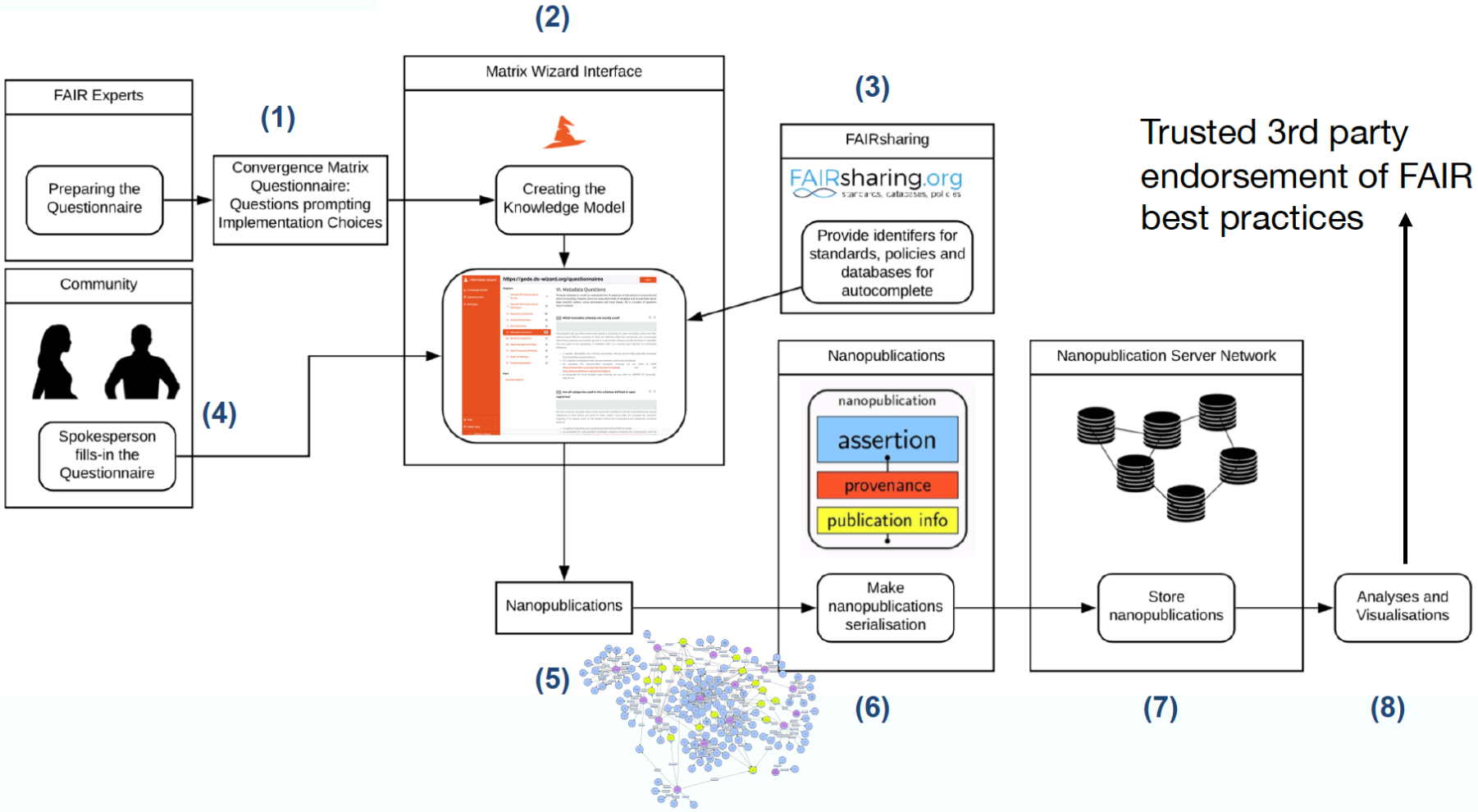
ENVRI-FAIR
KICK OFF MEETING
Hotel International Prague, Czech republic
January 14-16, 2019
Register at www.envri.eu

Slide from Erik Schultes, GO FAIR

Public project on Open Science Framework

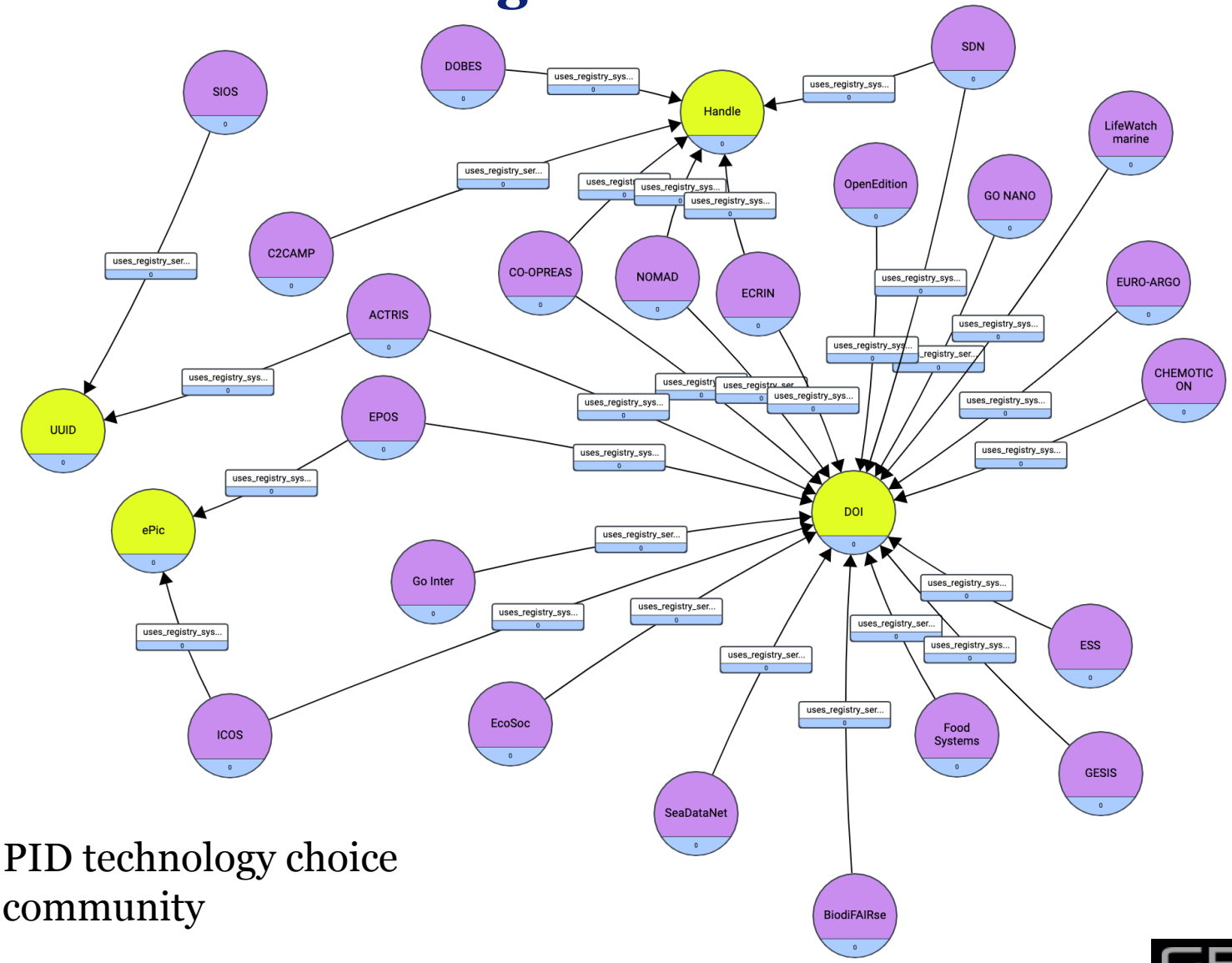


Convergence Matrix Implementation workflow



Picture from Erik Schultes, GO FAIR

Visual decision making



Yellow: PID technology choice
Purple: community

IN survey + ENVRI-FAIR + GEDE and 2nd IN survey: globally unique persistent identifier choice

Key factors for narrowing the gap

- Openness
  - Public project on Open Science Framework
  - Code on Github
  - All notes, presentations and results are open – i.e. we do not have to wait for the project deliverable or working group report to use the results
- Willingness to bridge and share
  - From all sides
- Common project of interest, directly useful for Leiden to:
  - Update Leiden Research Data Service Catalog
  - Advice researchers on current standards

A terminology for FAIR Stewardship Skills Workshop

CODATA, Paris, 20-21 May 2019

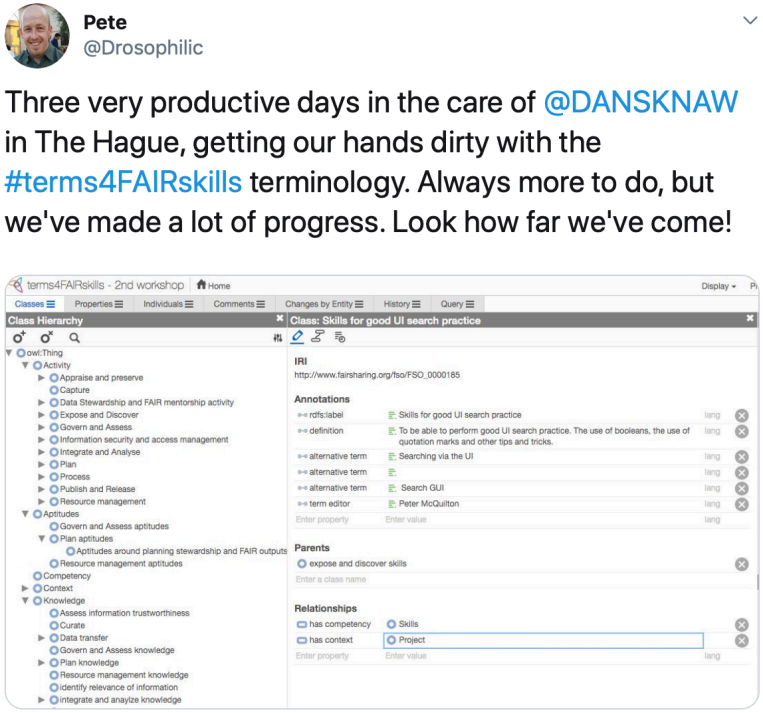
https://terms4fairskills.github.io

What did we do?

On 20-21 May 2019, representatives of the research data community met at the CODATA headquarters in Paris to take the first steps in building a terminology to describe FAIR stewardship skills (the skills necessary to make data FAIR and keep them FAIR): the tag #terms4FAIRskills was used on Twitter during the event.

This Photo by Unknown Author is licensed under CC BY-NC

Follow-up in October 2019, the Hague:



Thank you

Universiteit
Leiden
The Netherlands

CC BY