

EOSC-Nordic WP4 FAIR Data

Andreas O Jaunsen (NeIC / WP4 lead)

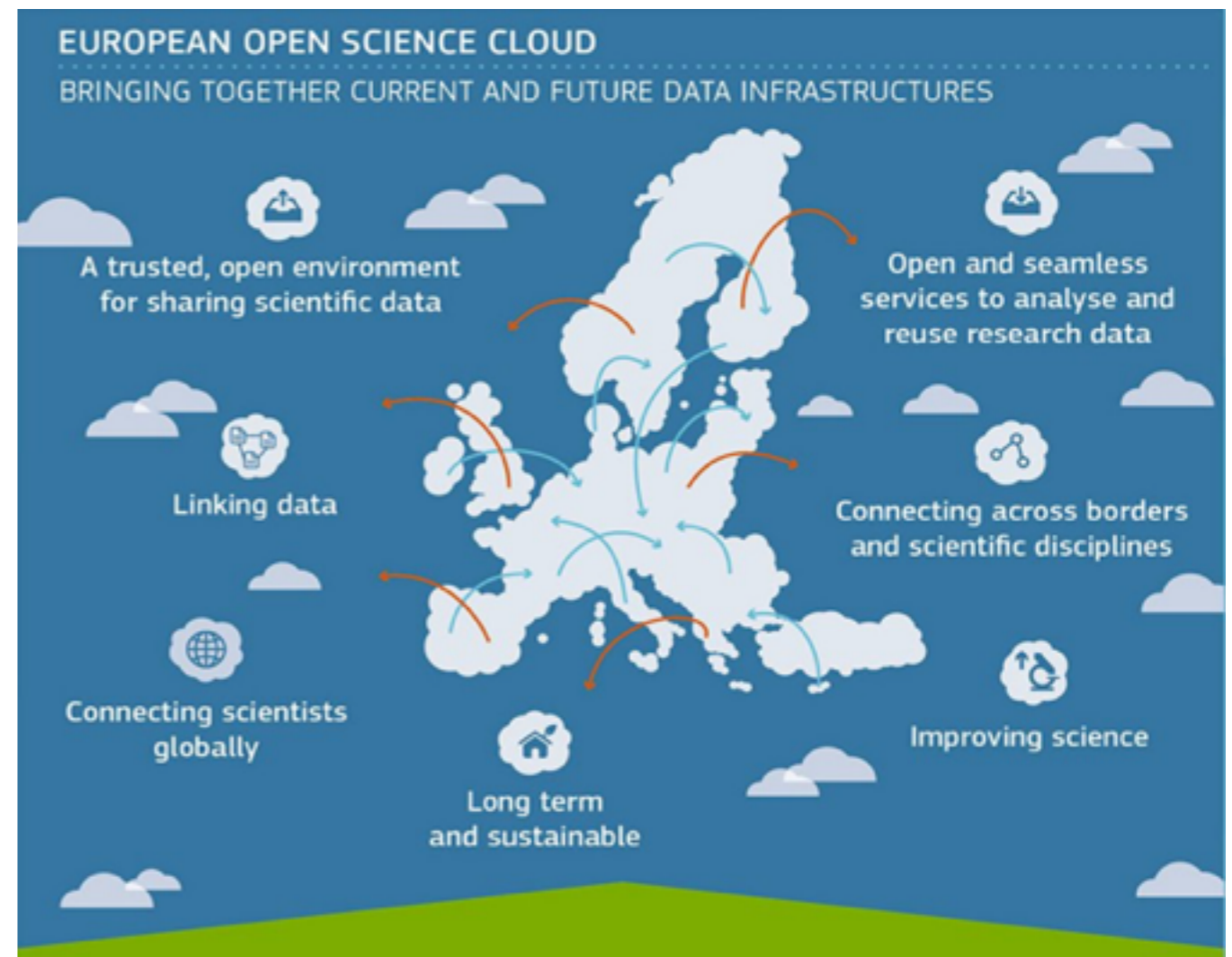
EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652



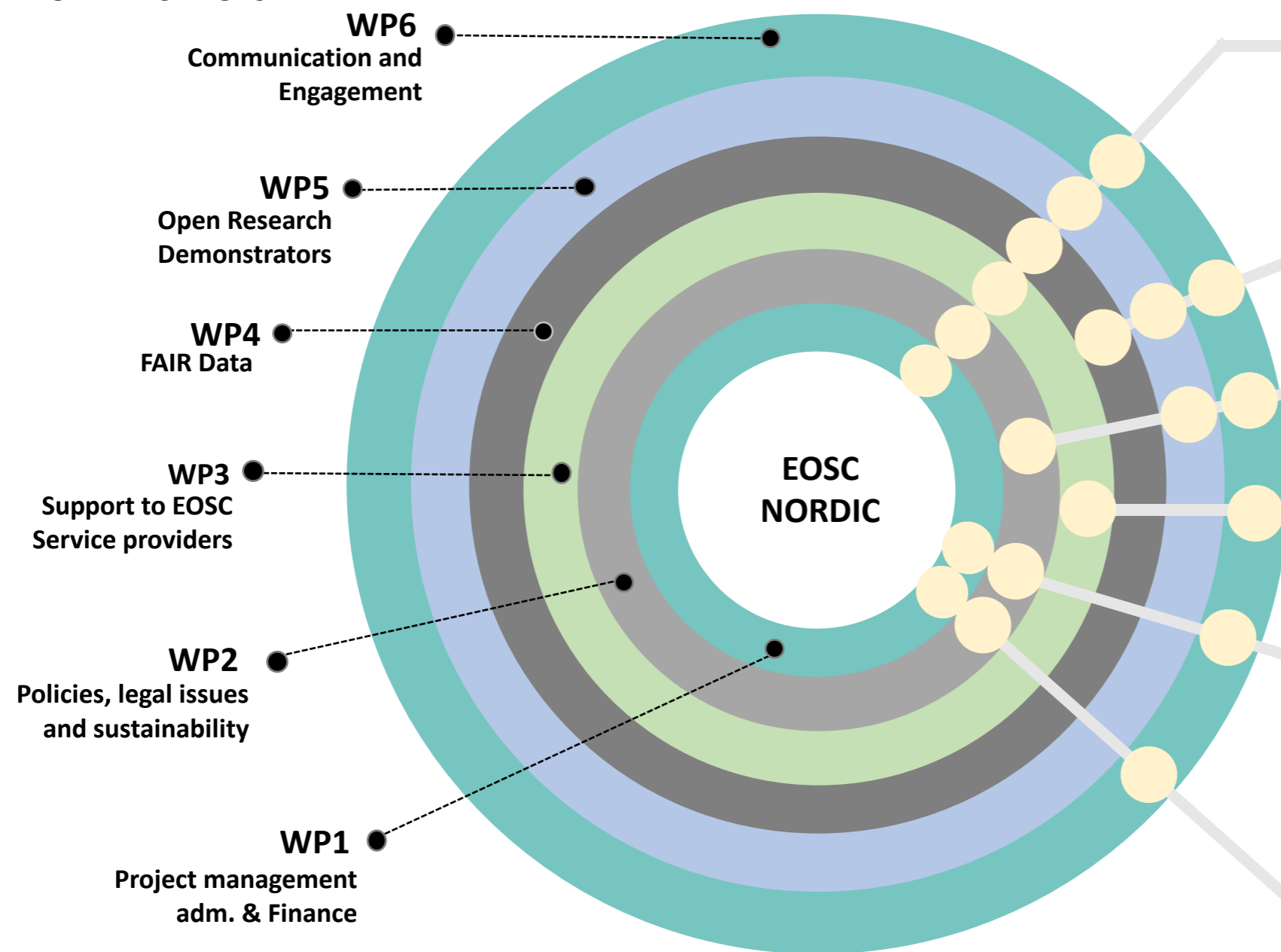
EOSC - European Commission's vision for Federating Data Infrastructures



- Vision for the European Open Science Cloud (EOSC) presented in the Commission communication on the 'European Cloud Initiative', as a part of the Digital Single Market Strategy [April 2016]
 - "A seamless environment enabling interdisciplinary research, an environment to foster data-intensive innovation. The EOSC will allow for universal access to data and a new level playing field for EU researchers." [EOSC Strategic Implementation Roadmap 2018-2020]
- From 2021 fully operational



WORK PACKAGES



WP6
Communication and
Engagement

WP5
Open Research
Demonstrators

WP4
FAIR Data

WP3
Support to EOSC
Service providers

WP2
Policies, legal issues
and sustainability

WP1
Project management
adm. & Finance

**EOSC
NORDIC**

STAKEHOLDERS INVOLVEMENT

-  e-Infrastructure- and service providers
-  Data repositories
-  Research communities and institutions
-  Universities & Libraries
-  Research Councils and Funding Agencies
-  Governmental agencies and Ministries

WP4 members

21 participants

Iceland

Guðbjörg A Jónsdóttir (HI)

Norway

Adil Hasan (Sigma2)
Trond Kvamme (NSD)
Andreas Jaunsen (NeIC)

Denmark

Henrik Jakobsen (DNA)
Troels Rasmussen (DeiC)

Netherlands

Bert Meermans (GFF)

Sweden

Birger Jerlehag (SND)
Iris Alfredsson (SND)
Monica Lassi (SNIC/LU)

Finland

Heikki Lehväslaiho (CSC)
Josefine Nordling (CSC)
Mari Elisa Kuusniemi (UHEL)
Mari Kleemola (UTA)
Pauli Assinen (UHEL)
Tuomas Aleterä (FSD)

Estonia

Liisi Lembinen (UTartu)

Latvia

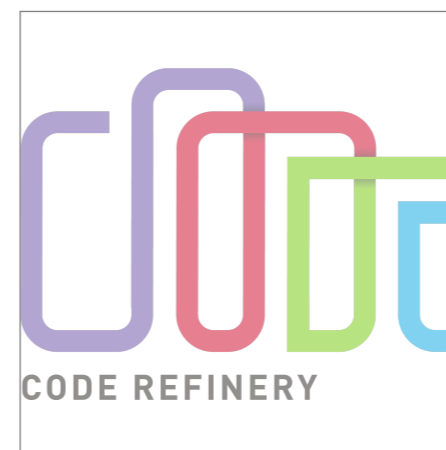
Janis Kampars (RTU)
Lauris Cikovskis (RTU)



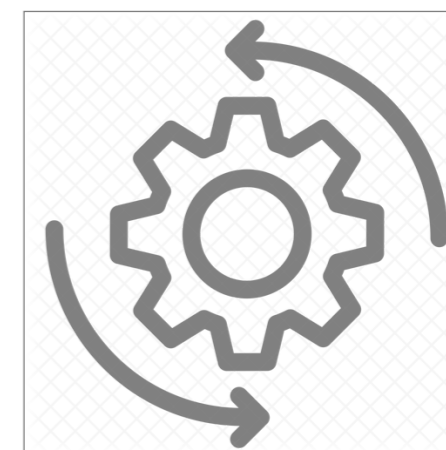
TRAINING



ENGAGE



SUPPORT



IMPLEMENT



EVALUATE

Nordic FAIR data stewardship: an introduction course



Sponsored by



October 5-9, 2020

Uppsala, Sweden

<http://bit.ly/FAIRds-Nordic-SE>

Nordic FAIR data stewardship course

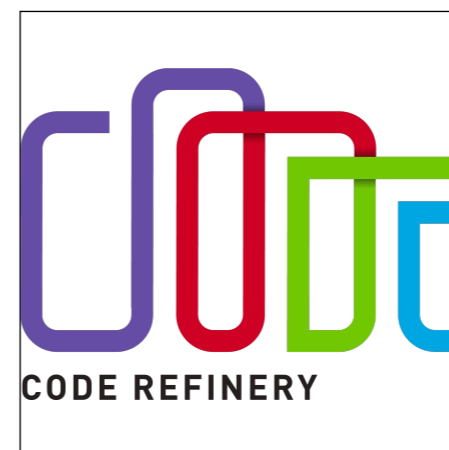
- FAIRds-Nordic Norway – 36 participants
- FAIRds-Nordic Denmark – 31 participants
- FAIRds-Nordic Sweden – 39 participants
- FAIRds-Nordic Finland – ? participants



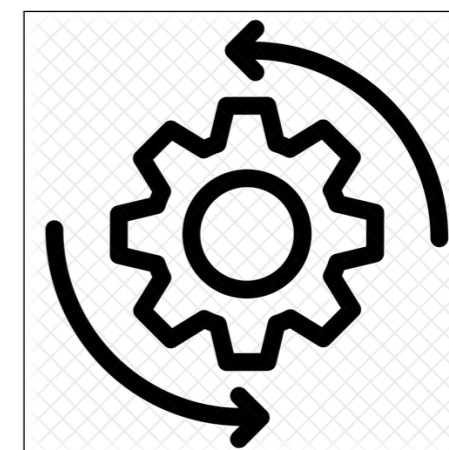
TRAINING



ENGAGE



SUPPORT



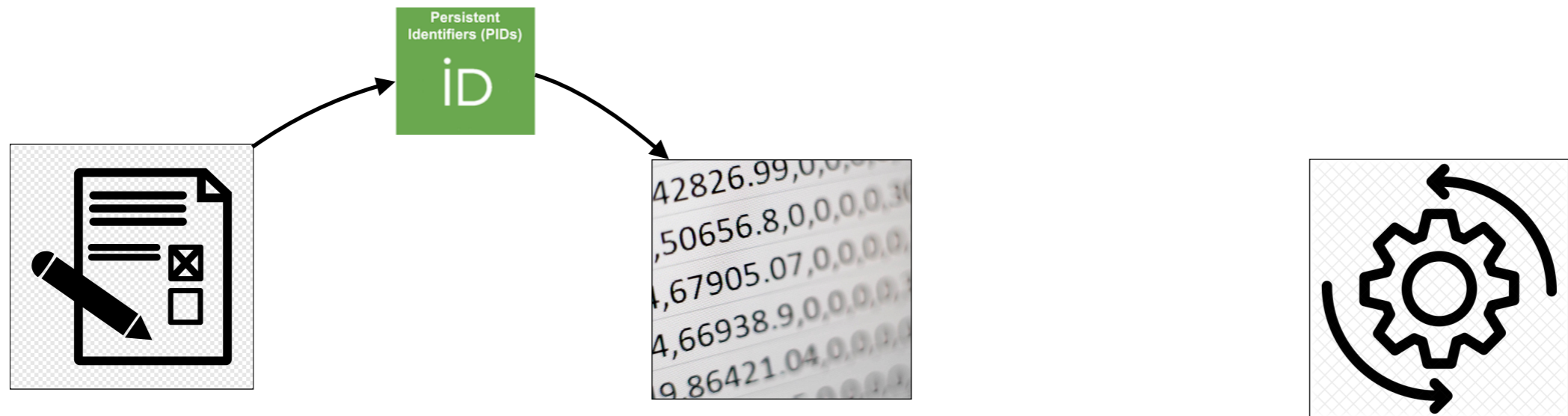
IMPLEMENT



EVALUATE

Task activities so far...

Dataset GUID required
(30% repos do NOT provide this)

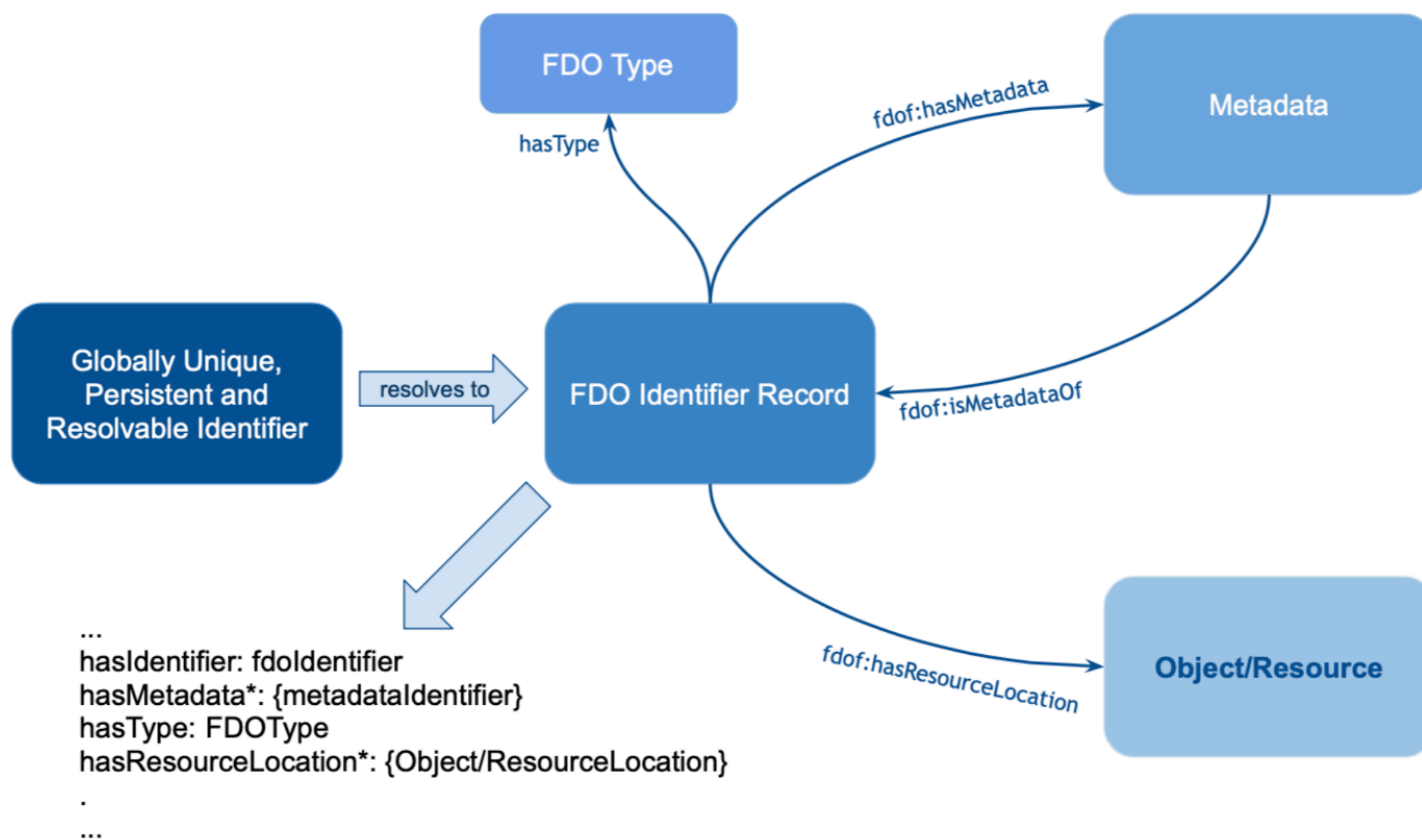


Surveyed Nordics region for research repositories (~100 repositories in sample)

(Manually) selected N=10 datasets per repository (700+ datasets)

Evaluated all datasets using fully machine-actionable metrics (Wilkinson gen2)

FAIR Digital Objects



Machine-actionable FAIR Maturity indicators



	Metric name	Principle association	Principle description
1	UNIQUE IDENTIFIER	F1	(Meta)data are assigned a globally unique and persistent identifier
2	IDENTIFIER PERSISTENCE	F1	(Meta)data are assigned a globally unique and persistent identifier
3	DATA IDENTIFIER PERSISTENCE	F1	(Meta)data are assigned a globally unique and persistent identifier
4	STRUCTURED METADATA	F2	Data are described with rich metadata (defined by R1 below)
5	GROUNDING METADATA	F2	Data are described with rich metadata (defined by R1 below)
6	DATA IDENTIFIER EXPLICITLY IN METADATA	F3	Metadata clearly and explicitly include the identifier of the data they describe
7	METADATA IDENTIFIER EXPLICITLY IN METADATA	F3	Metadata clearly and explicitly include the identifier of the data they describe
8	SEARCHABLE IN MAJOR SEARCH ENGINE	F4	(Meta)data are registered or indexed in a searchable resource
9	USES OPEN FREE PROTOCOL FOR DATA RETRIEVAL	A1.1	The protocol is open, free, and universally implementable
10	USES OPEN FREE PROTOCOL FOR METADATA RETRIEVAL	A1.1	The protocol is open, free, and universally implementable
11	DATA AUTHENTICATION AND AUTHORIZATION	A1.2	The protocol allows for an authentication and authorisation procedure, where necessary
12	METADATA AUTHENTICATION AND AUTHORIZATION	A1.2	The protocol allows for an authentication and authorisation procedure, where necessary
13	METADATA PERSISTENCE	A2	Metadata are accessible, even when the data are no longer available
14	METADATA KNOWLEDGE REPRESENTATION LANGUAGE (WEAK)	I1	(Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
15	METADATA KNOWLEDGE REPRESENTATION LANGUAGE (STRONG)	I1	(Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
16	DATA KNOWLEDGE REPRESENTATION LANGUAGE (WEAK)	I1	(Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
17	DATA KNOWLEDGE REPRESENTATION LANGUAGE (STRONG)	I1	(Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
18	METADATA USES FAIR VOCABULARIES (WEAK)	I2	(Meta)data use vocabularies that follow FAIR principles
19	METADATA USES FAIR VOCABULARIES (STRONG)	I2	(Meta)data use vocabularies that follow FAIR principles
20	METADATA CONTAINS QUALIFIED OUTWARD REFERENCES	I3	(Meta)data include qualified references to other (meta)data
21	METADATA INCLUDES LICENSE (STRONG)	R1.1	(Meta)data are released with a clear and accessible data usage license
22	METADATA INCLUDES LICENSE (WEAK)	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

Machine-actionable metadata for a DO/dataset



EXPLORE

PROVIDE

CONNECT

MONITOR

DEVELOP



SEARCH

DEPOSIT

LINK

CONTENT PROVIDERS

SIGN IN

Community Health Workers and Mobile Technology: A Systematic Review of the Literature

DATASET UNKNOWN

Braun, Rebecca; Catalani, Caricia; Wimbush, Julian; Israelski, Dennis; (2015)

Publisher: Figshare

Identifiers: [doi: 10.1371/journal.pone.0065772](https://doi.org/10.1371/journal.pone.0065772)

Subject: Medicine | Information and Computing Sciences | Information technology | Global health | Non-clinical medicine | Health care policy | Health education and awareness | Health systems strengthening | Quality of care | Communication in health care | Health care providers | Health care quality | Health informatics | Public health | workers | systematic

Introduction
In low-resource settings, community health workers are frontline providers who shoulder the health service delivery burden. Increasingly, mobile technologies are developed, tested, and deployed with community health workers to facilitate tasks... [View more](#)

[Related Research Results \(56\)](#)



[Metrics](#)



[LINK THIS RESEARCH DATA TO...](#)

Share - Bookmark



Download from

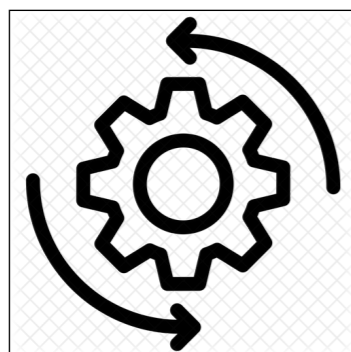
[figshare via figshare](#) (Dataset, 2015) ?

[Figshare](#) ?

Cite this research data

select a citation style

Milestones reached...



MS28: Evaluated
all datasets for
FAIR maturity



Provided FAIRification
recommendations
based on
results offered to community



MS26: workshop
executed
(April 22)

No GUID.
Evaluation
not possible

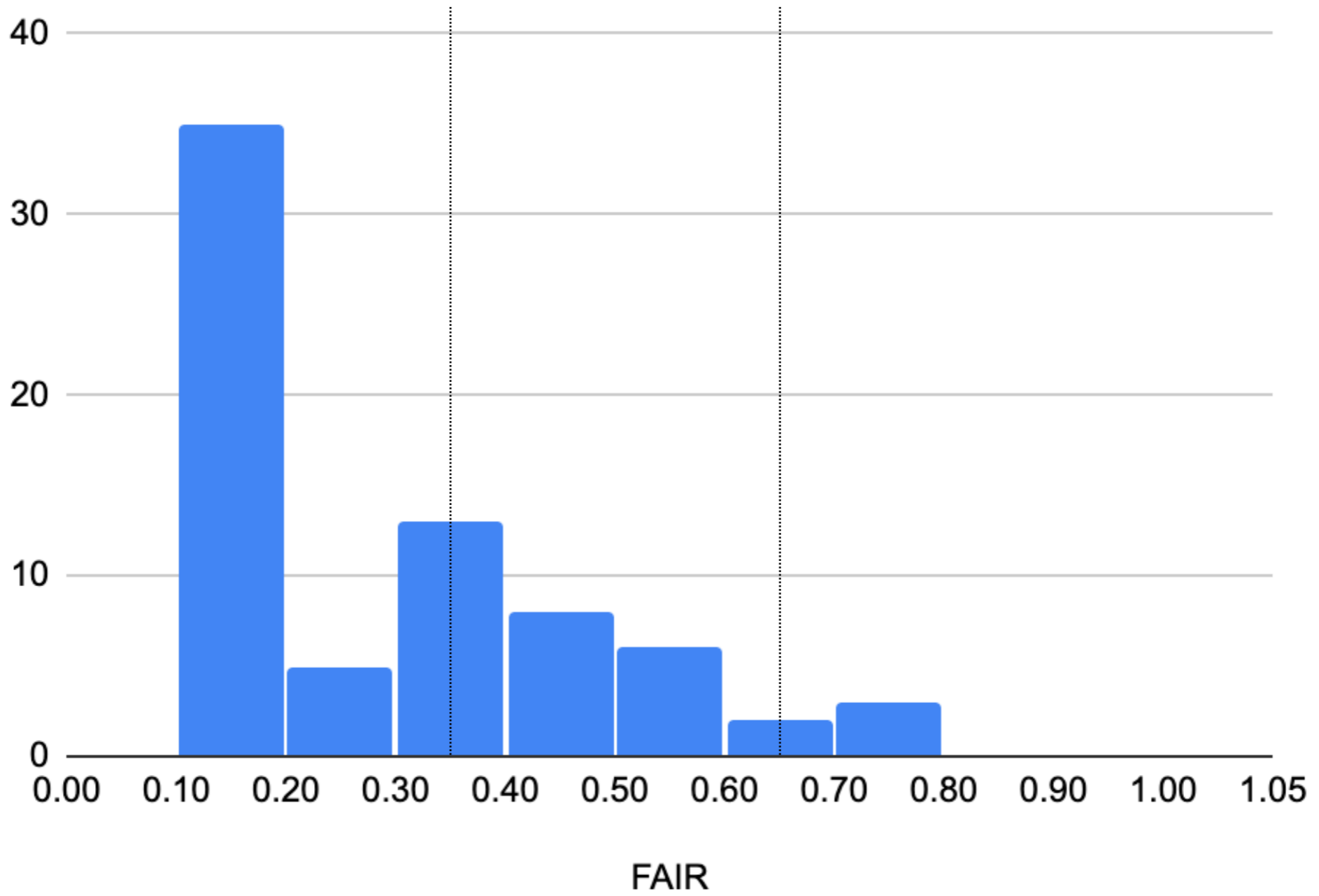
Low FAIR
machine
actionability

Medium FAIR
machine
actionability

High FAIR
machine
actionability



Histogram of FAIR score



100 repositories, 72 evaluated

Disclaimer: PRELIMINARY RESULTS

No GUID.
Evaluation
not possible

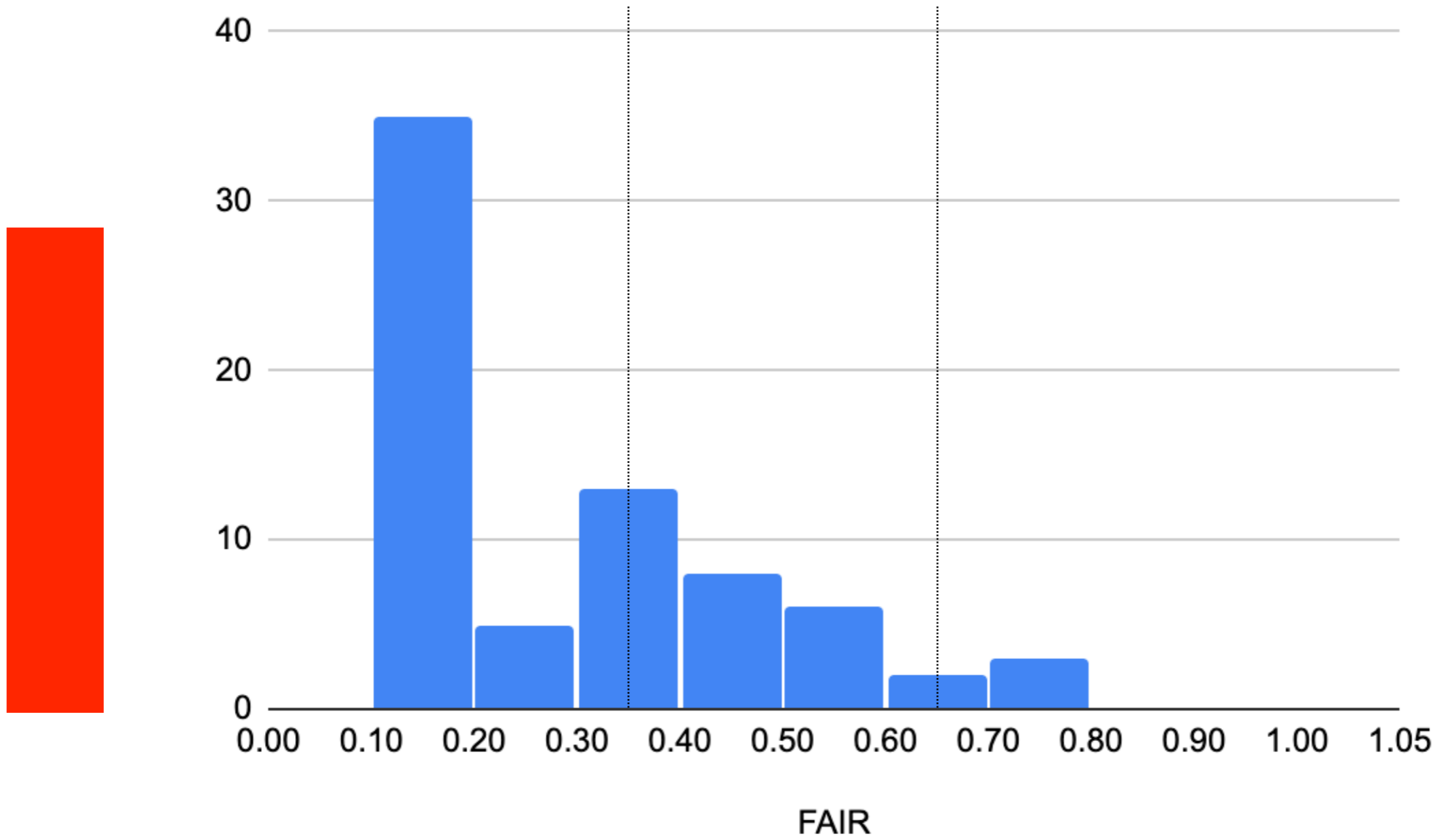
Low FAIR
machine
actionability

Medium FAIR
machine
actionability

High FAIR
machine
actionability



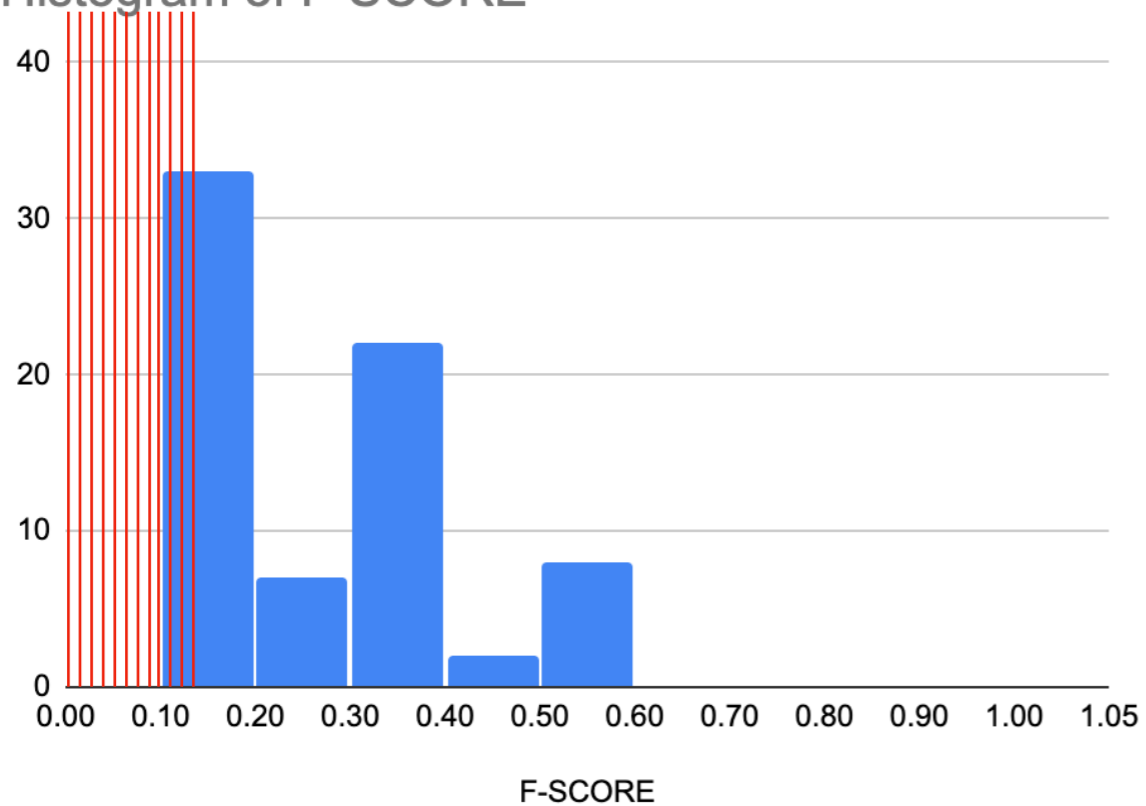
Histogram of FAIR score



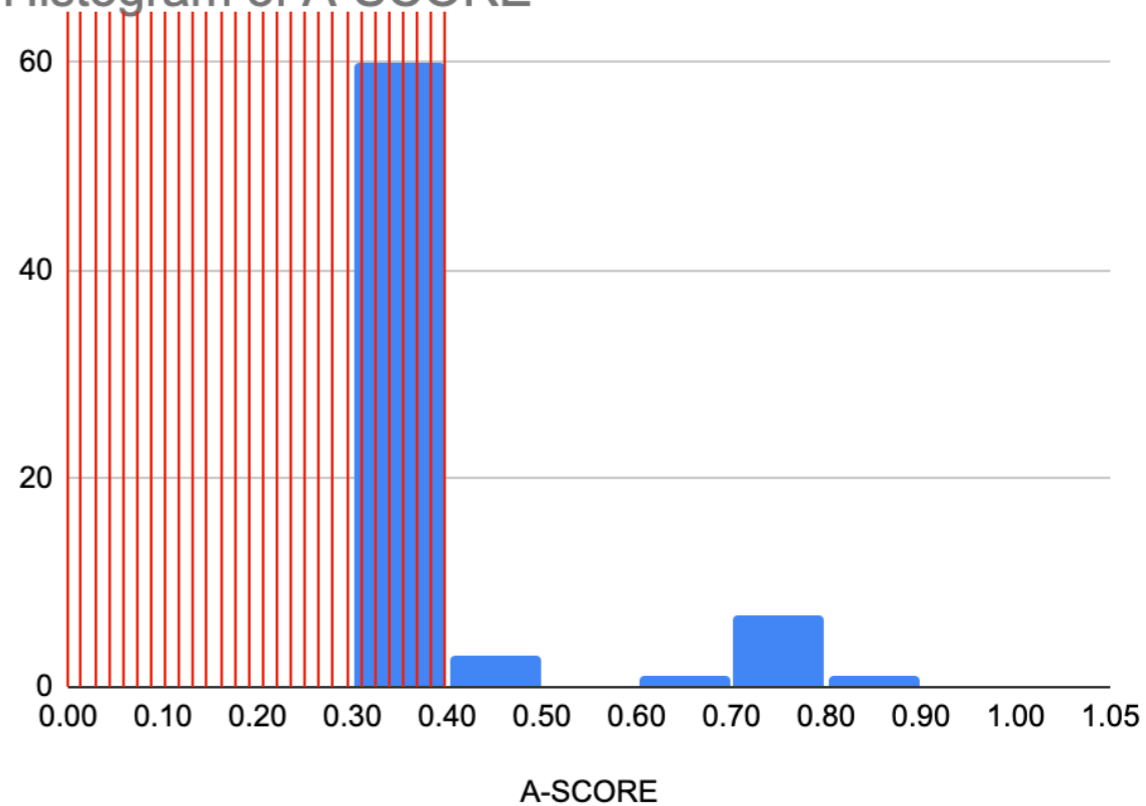
100 repositories, 72 evaluated

Disclaimer: PRELIMINARY RESULTS

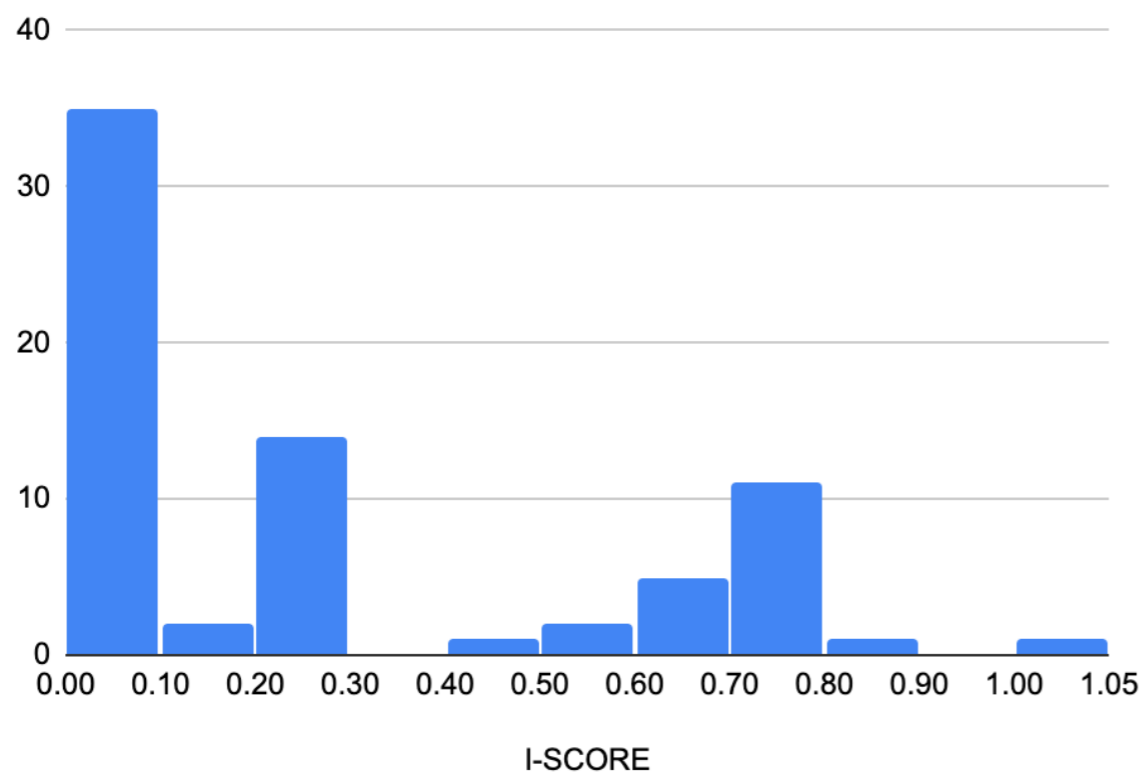
Histogram of F-SCORE



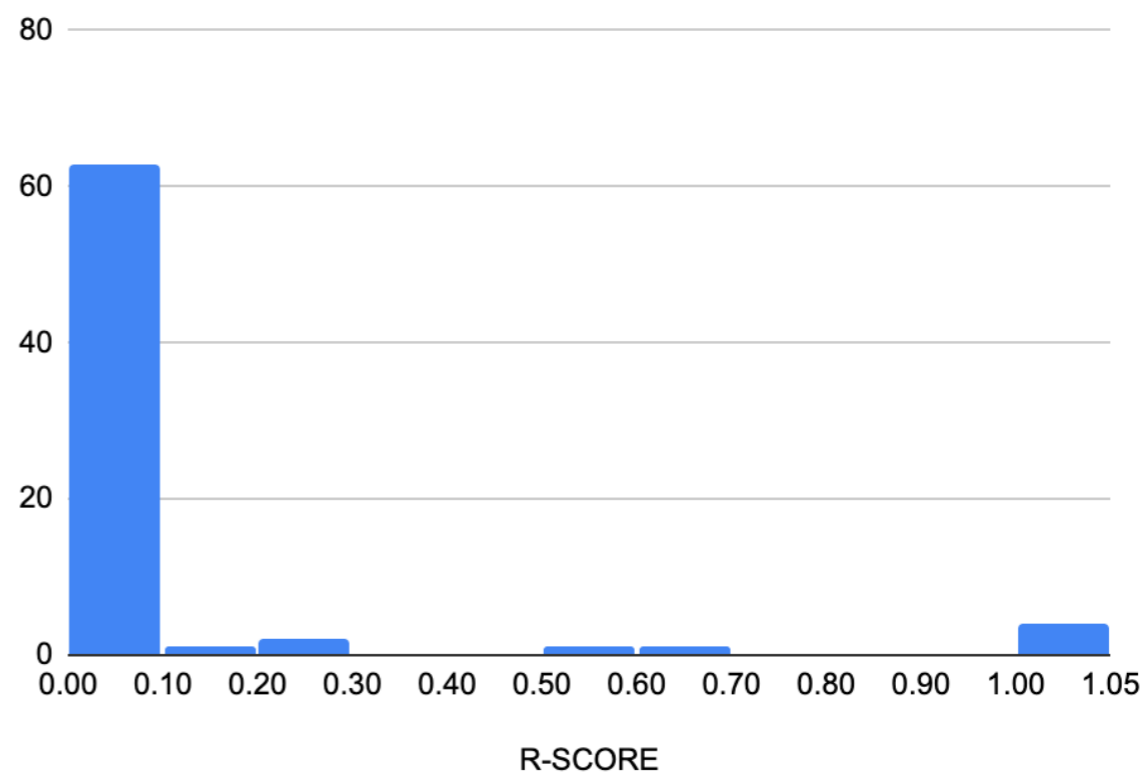
Histogram of A-SCORE



Histogram of I-SCORE

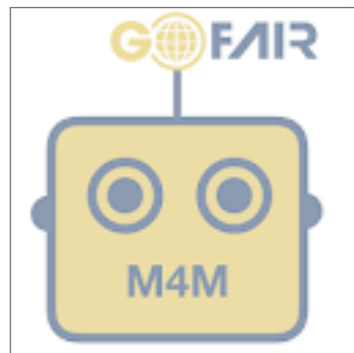


Histogram of R-SCORE



Disclaimer: PRELIMINARY RESULTS

Coming tasks



M4M events to support community FAIRification process



Certification support for repositories that wish to go this route



FAIR data standards and semantic attributes



FAIR incentives and stakeholder liaison

Recommendations

- All datasets should be identified by a globally unique identifier (GUID), preferably a persistent identifier (PID)
- Repositories should register on re3data.org to increase discoverability
- Employ the concept of FAIR digital object for published datasets (cf. “Metadata Identifier Explicitly in Metadata” and “Data Identifier Explicitly in Metadata”)
- State under what license agreement the dataset is provided, using one of the standard “license” predicates/keys

The End