

EOSC-Nordic intro & highlights from WP4 FAIR Data

Andreas O Jaunsen (NeIC / WP4 lead)

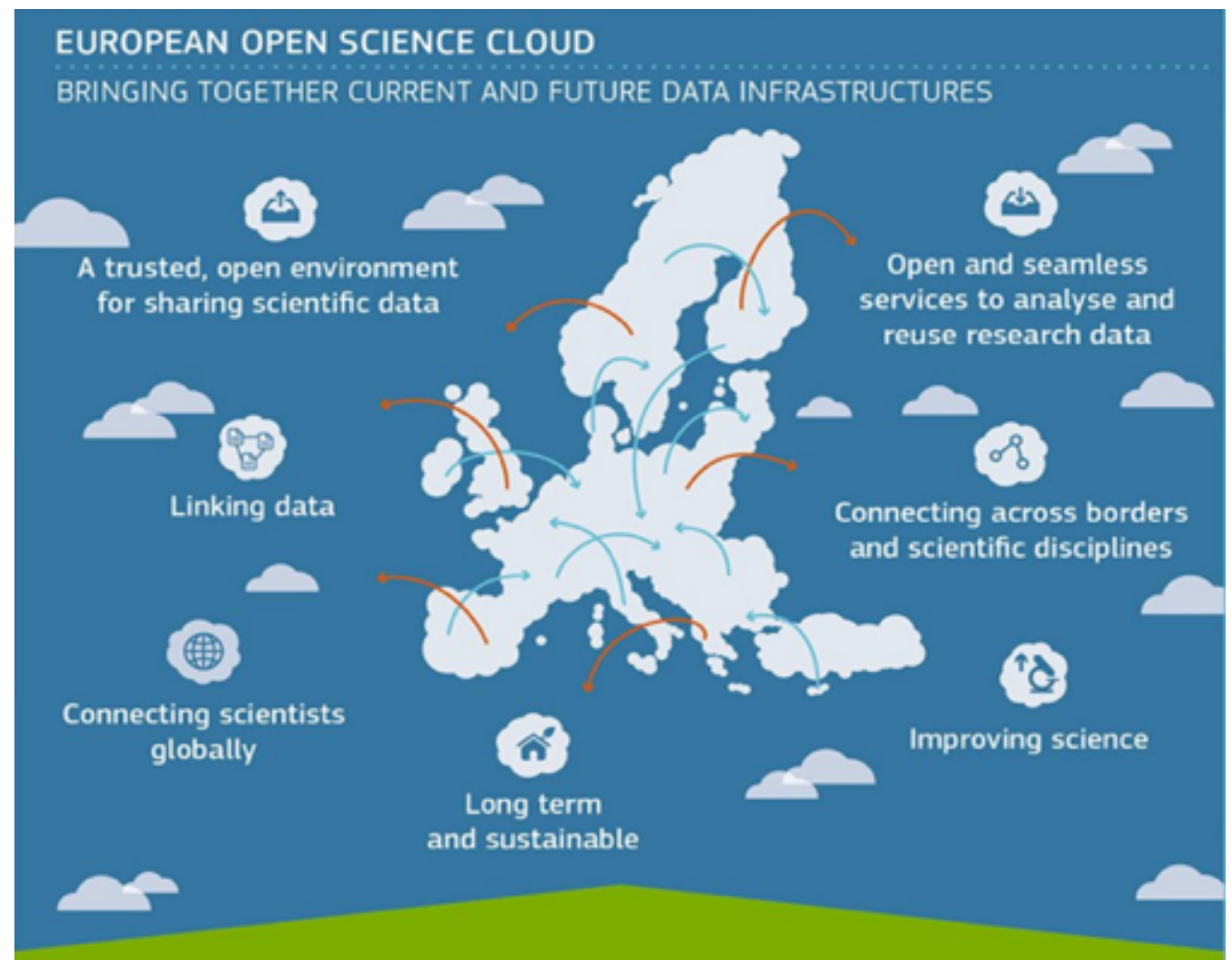
EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

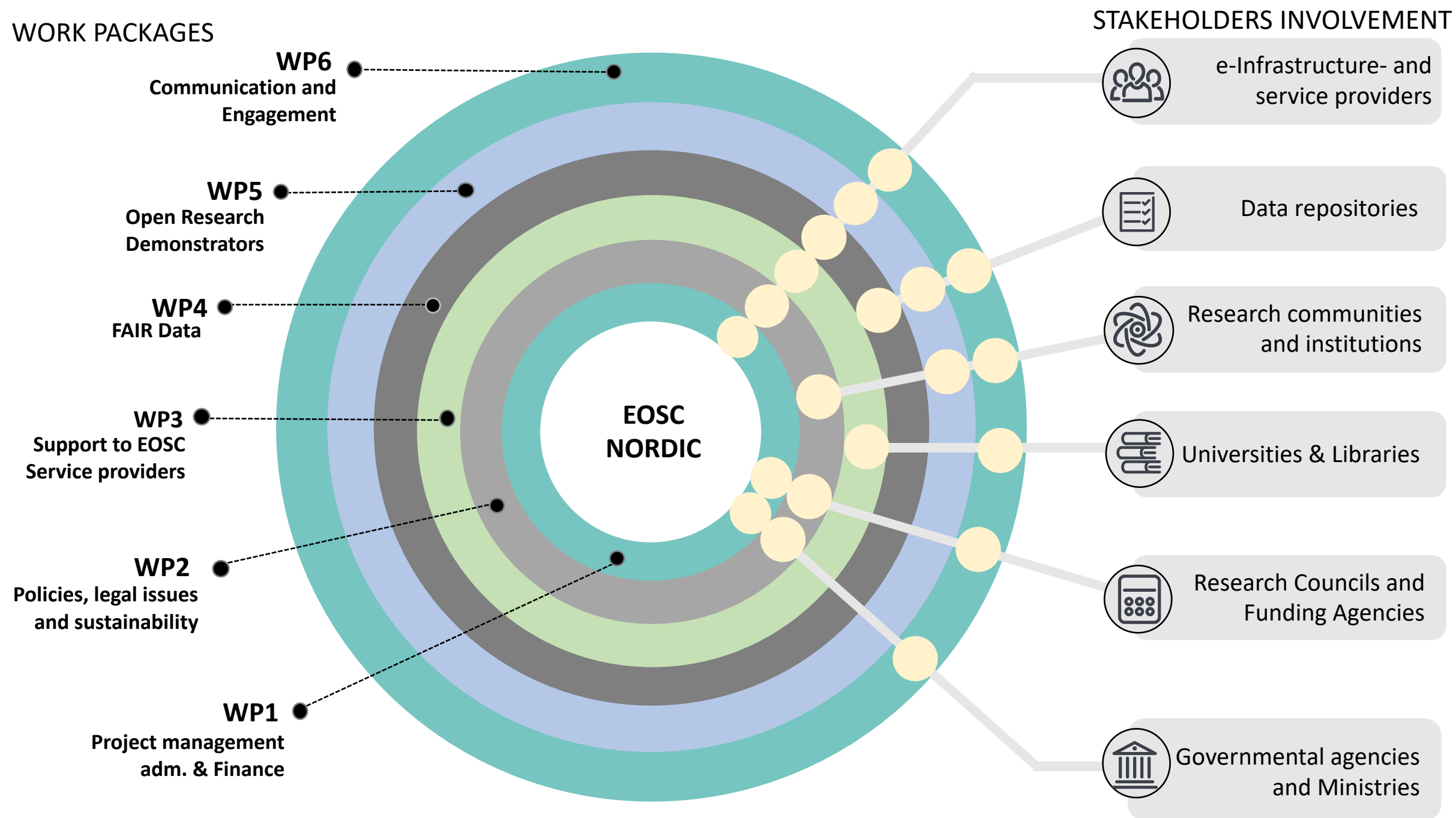


EOSC - European Commission's vision for Federating Data Infrastructures



- 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





24 Participants 10 countries

Iceland

- UNIVERSITY OF ICELAND

Norway

- NORDFORSK
- UNINETT SIGMA2 AS
- NORWEGIAN CENTER FOR RESEARCH DATA

Denmark

- DENMARK TECHNICAL UNIVERSITY
- UNIVERSITY OF SOUTHERN DENMARK
- DANISH NATIONAL ARCHIVES
- UNIVERSITY OF COPENHAGEN
- CAPITAL REGION OF DENMARK
- NORDUNET / AS

Netherlands

GoFair

Germany

DKRZ

Sweden

- UNIVERSITY OF UPPSALA
- SWEDISH RESEARCH COUNCIL
- UNIVERSITY OF GOTHENBORG

Finland

- CSC – IT CENTER OF SCIENCE
- UNIVERSITY OF HELSINKI
- UNIVERSITY OF TAMPERE
- UNIVERSITY OF EASTERN FINLAND
- FINNISH METEOROLOGICAL INSTITUTE

Estonia

- UNIVERSITY OF TARTU
- NATIONAL INSTITUTE OF CHEMICAL PHYSICS AND BIOPHYSICS

Latvia

- RIGA TECHNICAL UNIVERSITY

Lithuania

- UNIVERSITY OF VILNIUS

Your Gateway for European Services for Research

**These datasets may be available,
but are not necessarily findable,
accessible, interoperable or
reusable**

The catalogue gives access to:

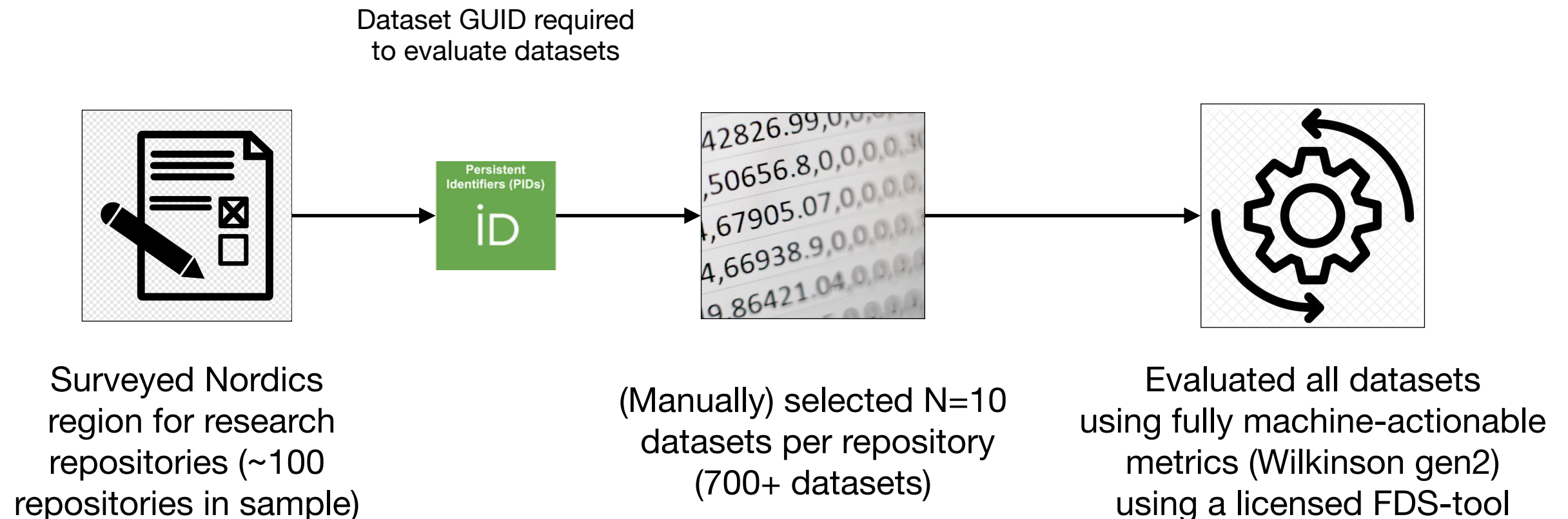
254 services, **4,4M** datasets, **7,1M** software and applications,

34,6M publications and **3M** other research products,

offered by **73** Service/Resource Providers and Aggregators, organized in **238** entries.



WP4 first year activities

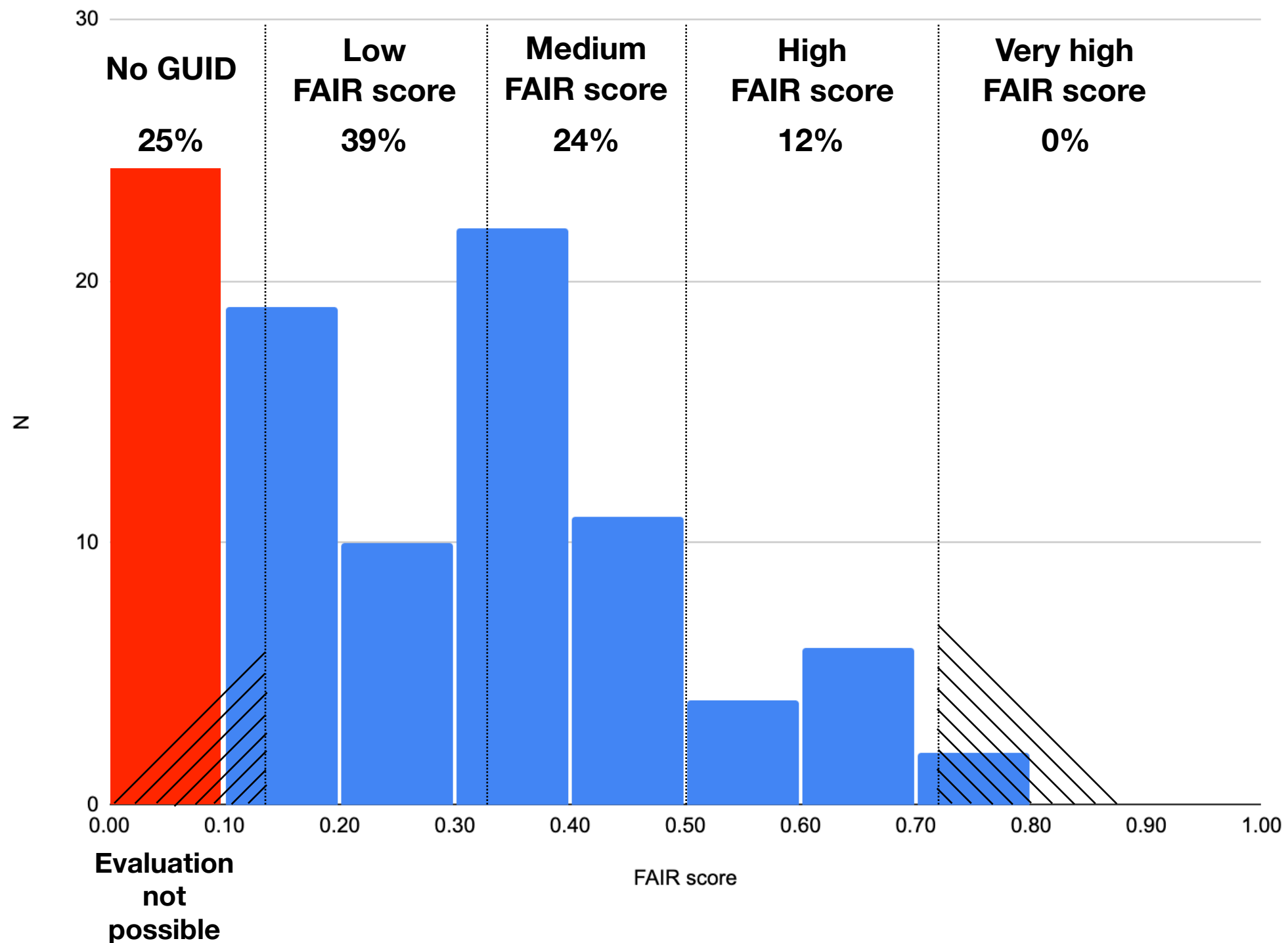


Machine-actionable FAIR Maturity indicators

	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

FAIR uptake

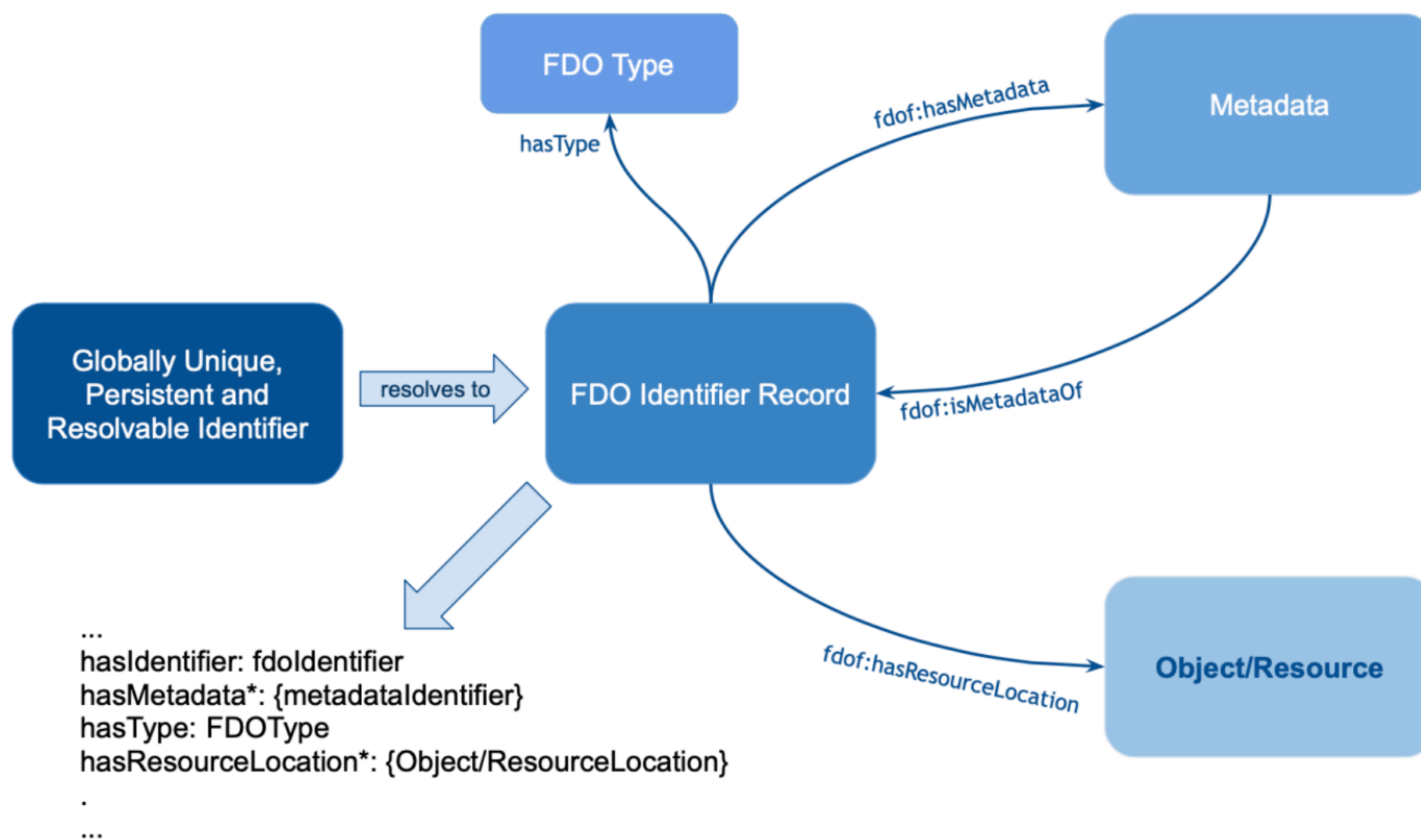
FAIR scores from 1018 PID+URI datasets



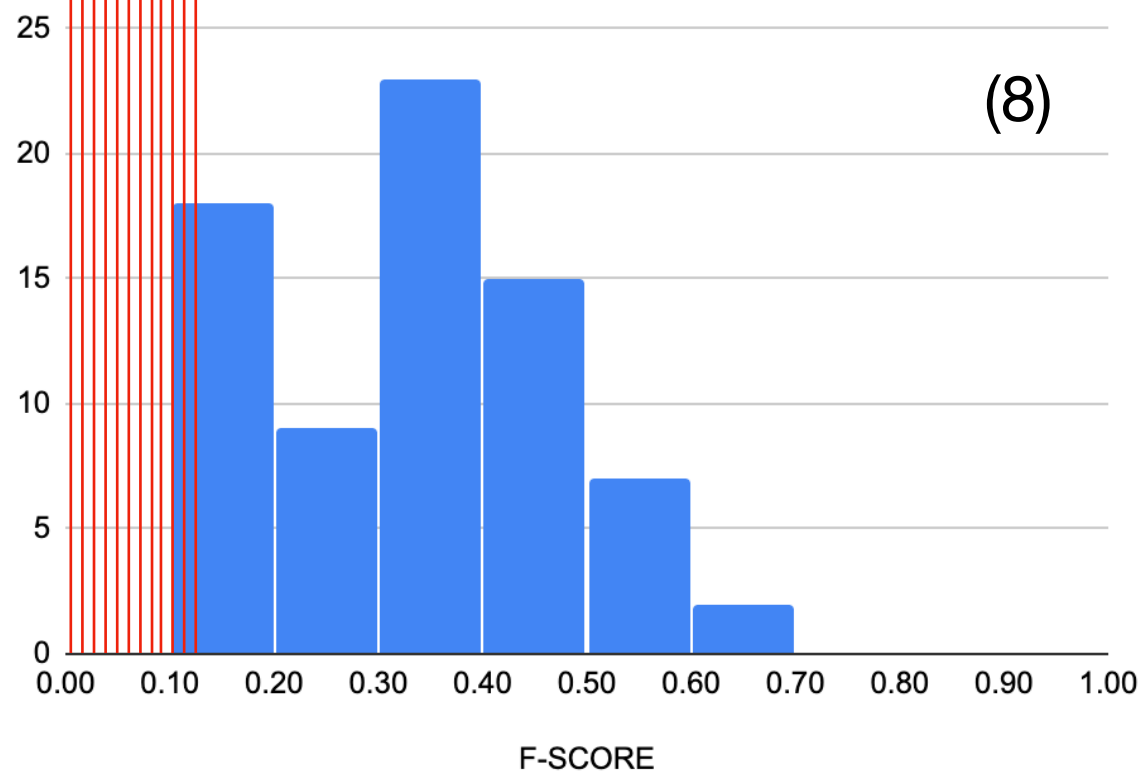
20 July 2020

98 repositories (74 evaluated)

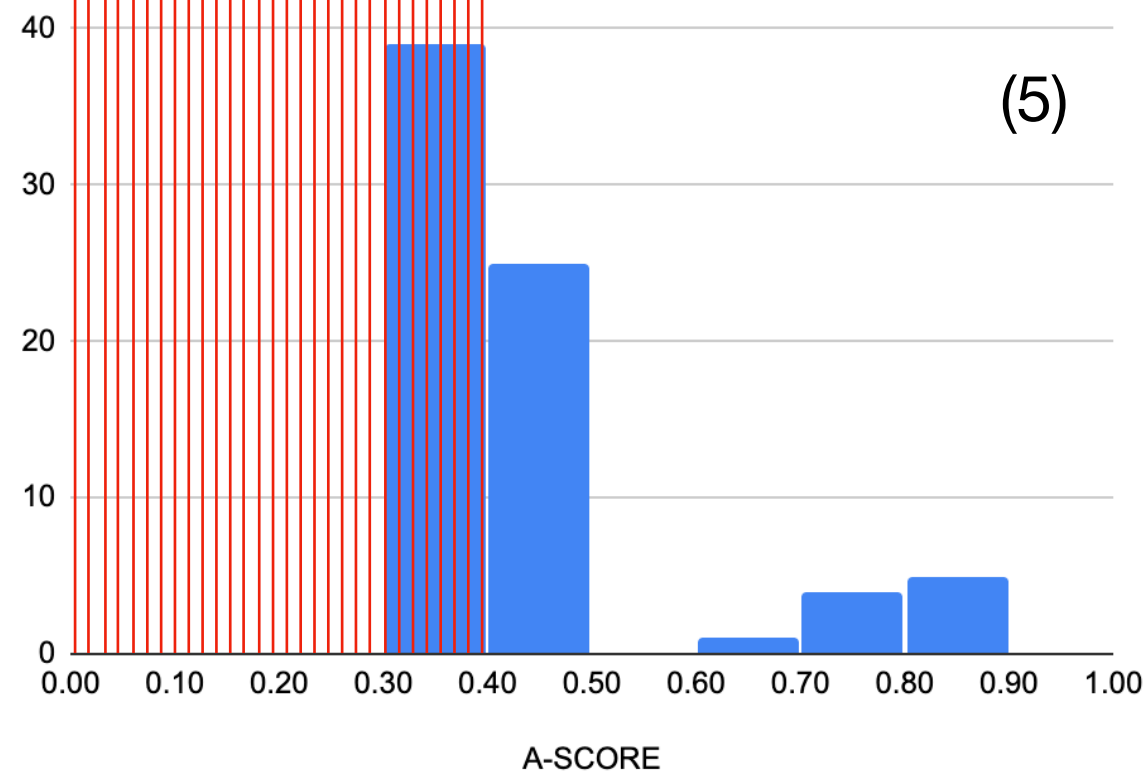
FAIR Digital Objects



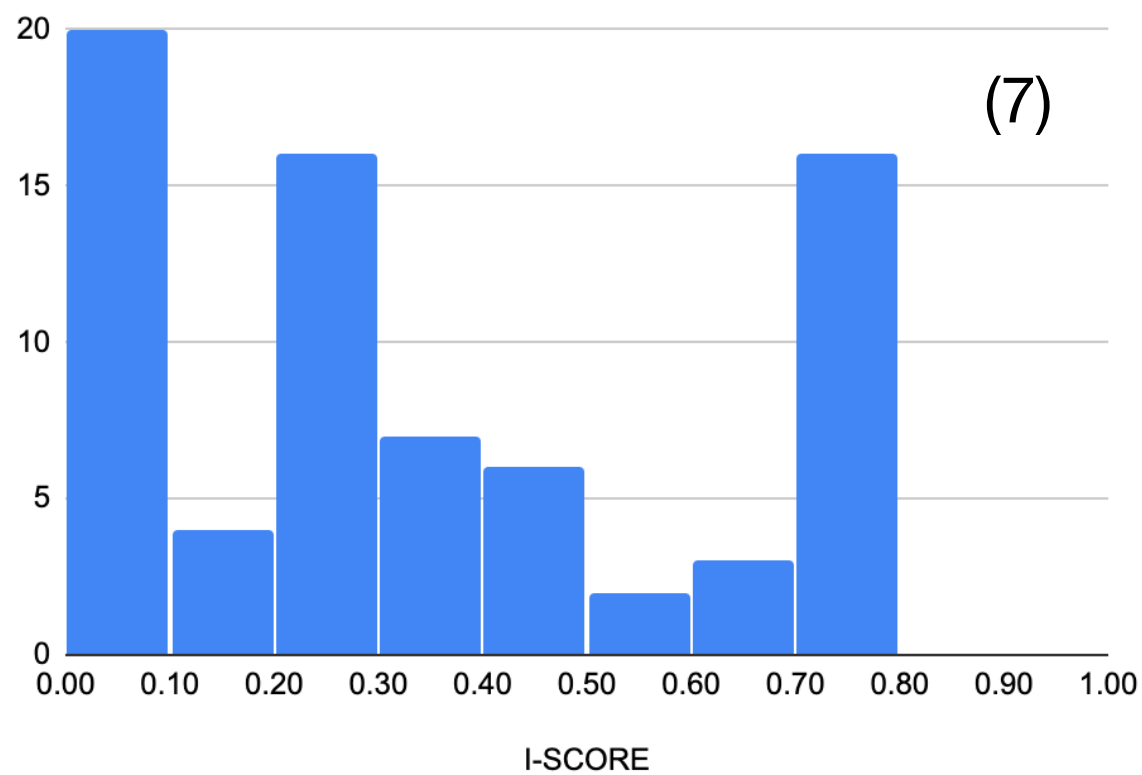
Histogram of F-SCORE



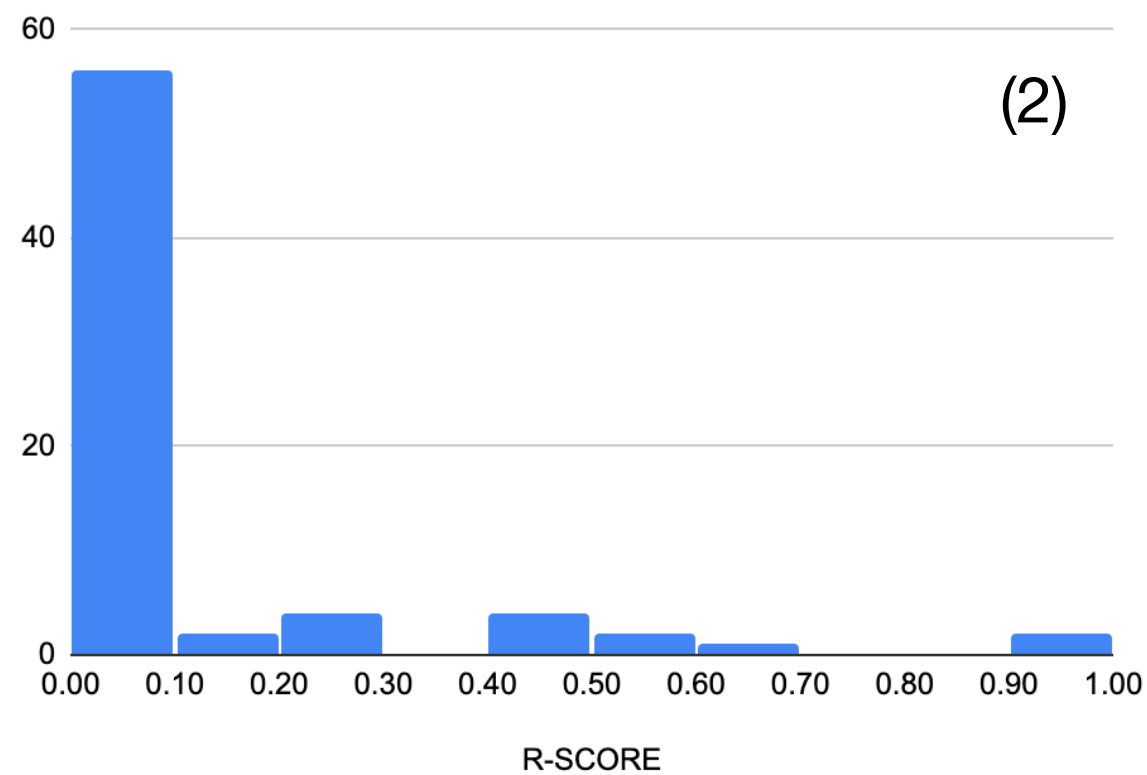
Histogram of A-SCORE



Histogram of I-SCORE



Histogram of R-SCORE



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