Исследовательские инфраструктуры

И другие блага открытой науки

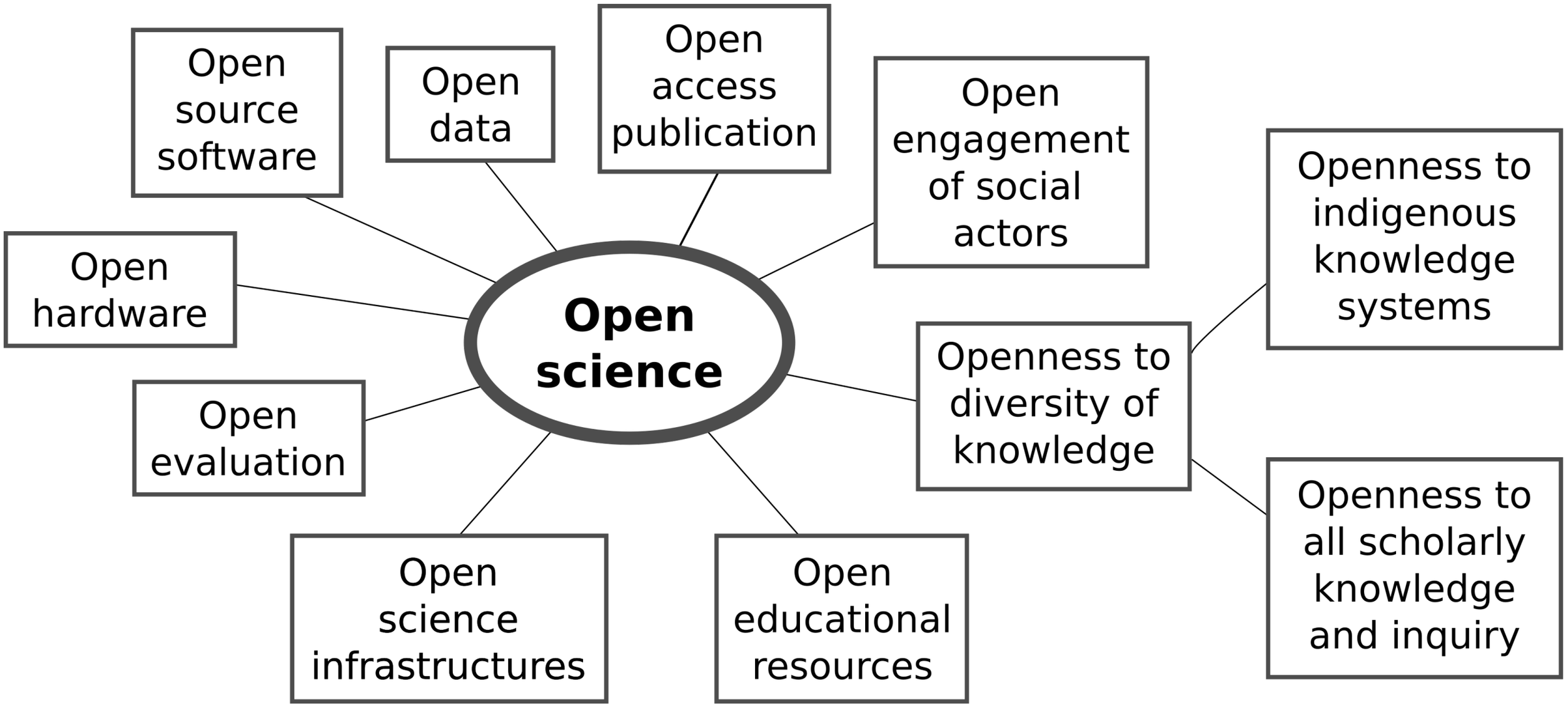
Калинин Николай,
Аспирант ФИЦ ИУ РАН
tg: @kalinidze

План

1. **Открытая наука и честная (FAIR) наука**
2. **Что такое исследовательские инфраструктуры**
3. **Научный репозиторий Zenodo**
4. **FAIRsharing**
5. **Свободный GO-FAIR**
6. **EOSC: великий и ужасный**
7. **Причем тут онтологии**
8. **Как много наук честны и открыты?**

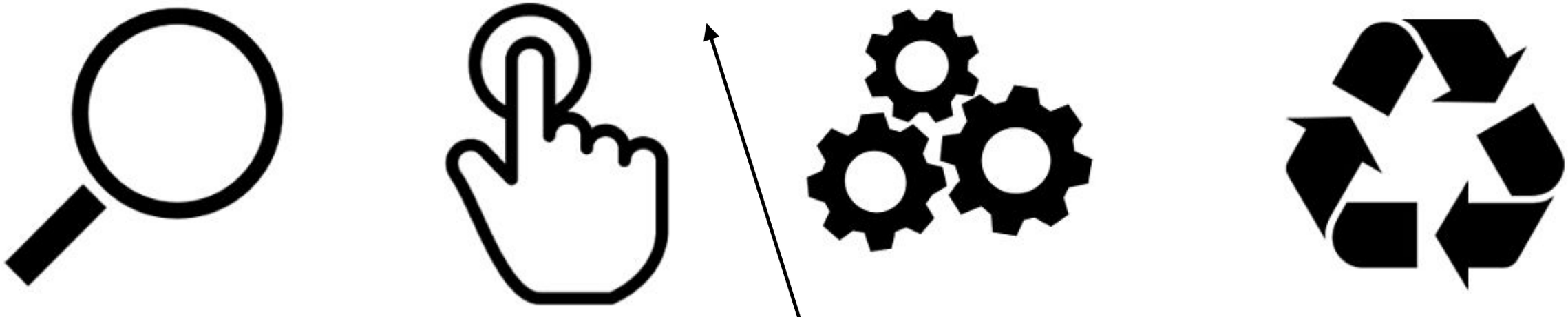
Open Science

Цель: сделать науку доступной для всех.



FAIR Science

Findable Accessible Interoperable Reusable



FAIR in depth

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

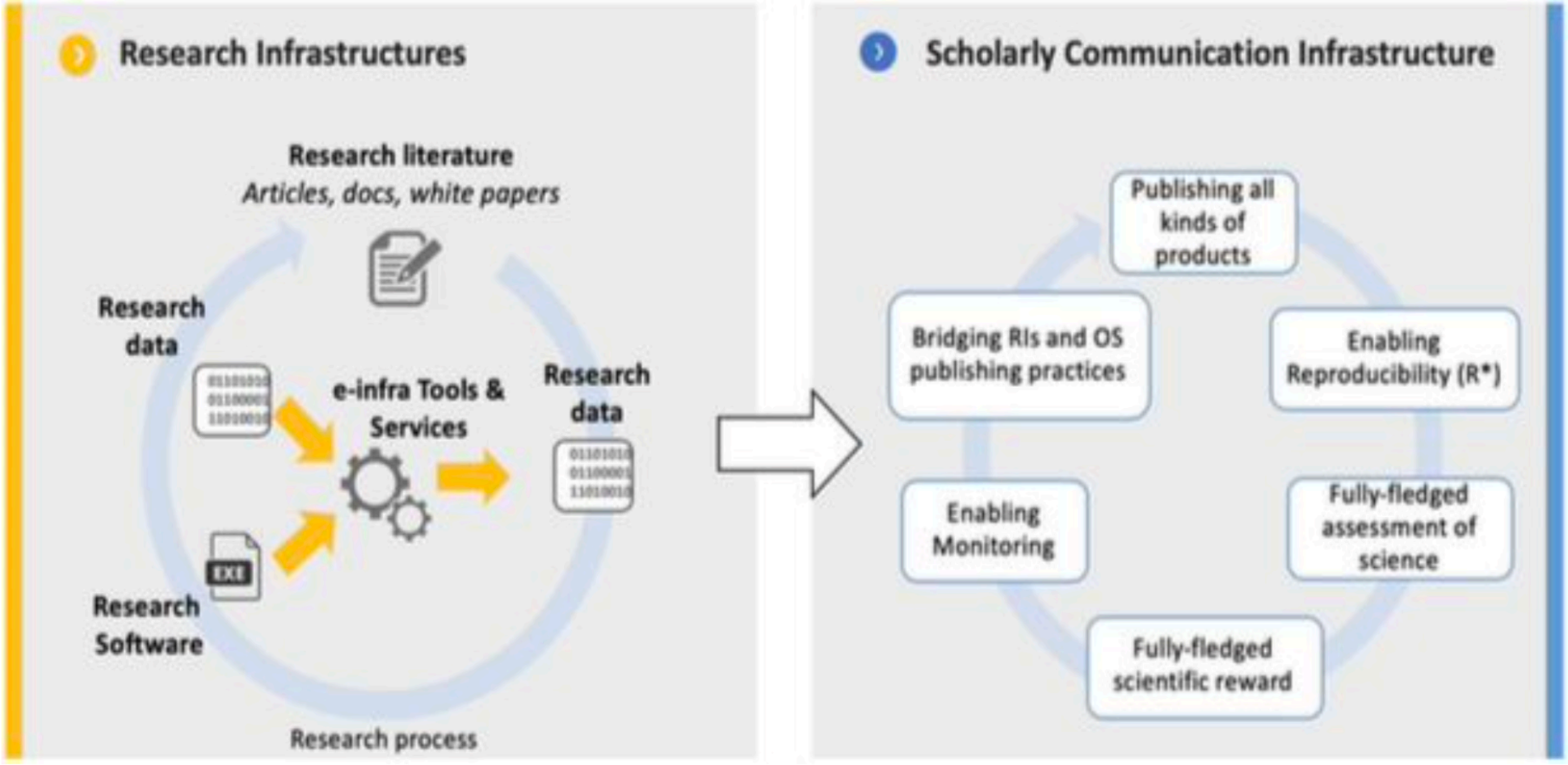
To be Accessible:
A1. (meta)data are retrievable by their identifier using a standardized communications protocol
A1.1 the protocol is open, free, and universally implementable
A1.2 the protocol allows for an authentication and authorization procedure, where necessary
A2. metadata are accessible, even 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

Исследовательские инфраструктуры

Исследовательские инфраструктуры —
это объекты, которые предоставляют ресурсы и услуги исследовательским сообществам для проведения исследований и содействия инновациям в своих областях.



Zenodo by CERN

Универсальный репозиторий с DOI и интеграцией с GitHub

All versions

Access Right

- Open (17685)
- Closed (2803)
- Restricted (26)
- Embargoed (4)

Type

- Publication (15058) +
- Image (4869) +
- Dataset (343)
- Presentation (82)
- Software (63)
- Poster (45)
- Video (31)
- Other (26)
- Lesson (1)

File Type

- Pdf (6976)
- Html (5303)
- Jpg (2621)
- Png (2303)
- Xml (773)
- Zip (147)
- Docx (76)
- Xlsx (68)
- Txt (56)
- Mp4 (41)

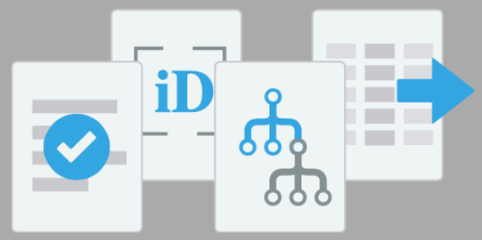
Zenodo

Under the hood

1. **CERN Servers: OpenStack, Puppet**
2. **Frontend: Python + Flask**
3. **Data Storage: EOS Service, 18 птб + CERN's CASTOR backups**
4. **Metada storage: JSON in Postgresql**

FAIR sharing

«Поддерживаемый, информативный и образовательный ресурс о стандартах данных и метаданных, взаимосвязанных с базами данных и политиками в отношении данных»



1591 Standards

Terminology Artifact	830
Model/Format	506
Reporting Guideline	233
Identifier Schema	22

VIEW ALL

1884 Databases

Repositories	966
Knowledgebases	794
Knowledgebase/Repositories	124

VIEW ALL

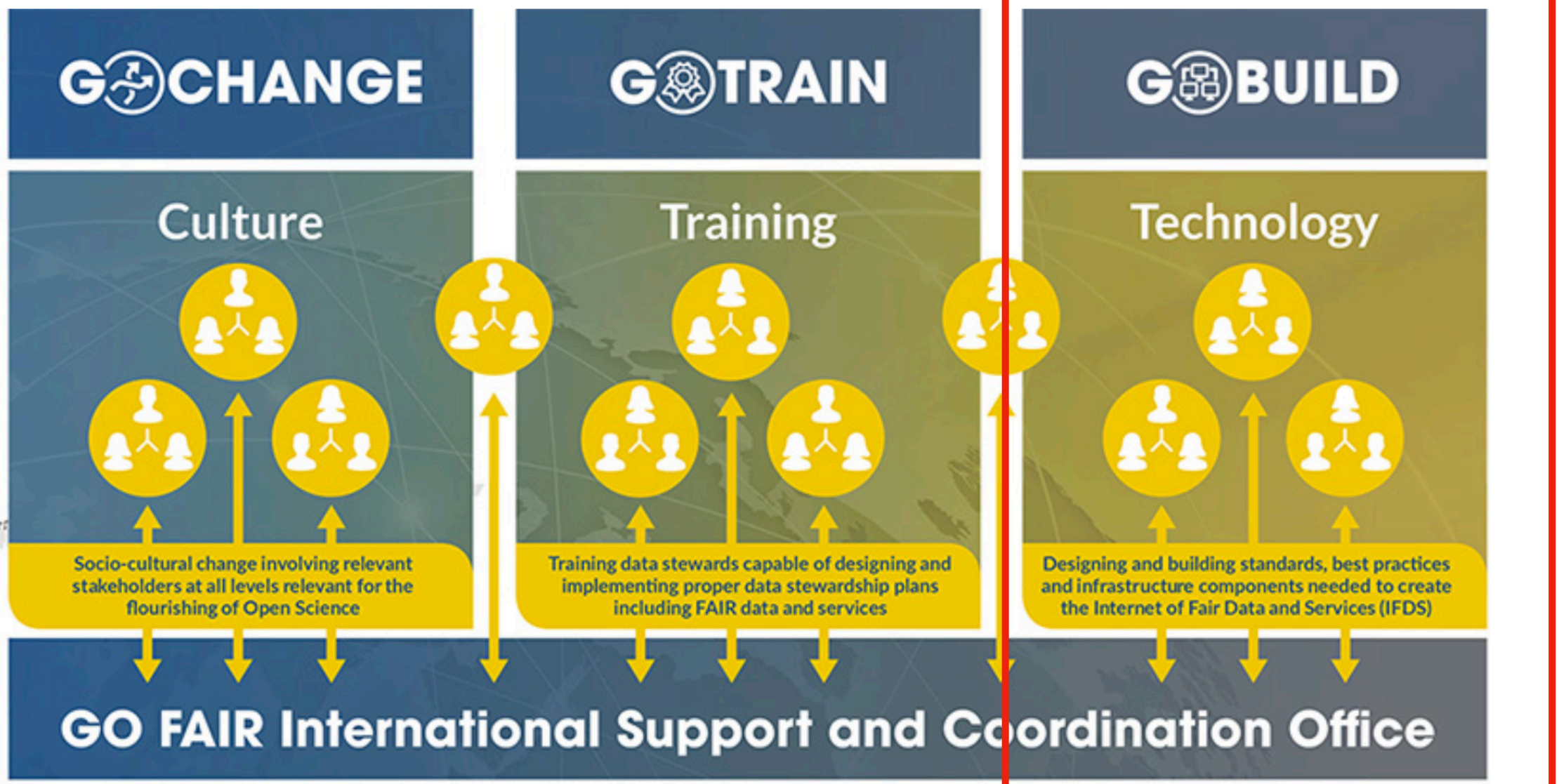
158 Policies

Journal	94
Funder	23
Society	13
Project	13

VIEW ALL

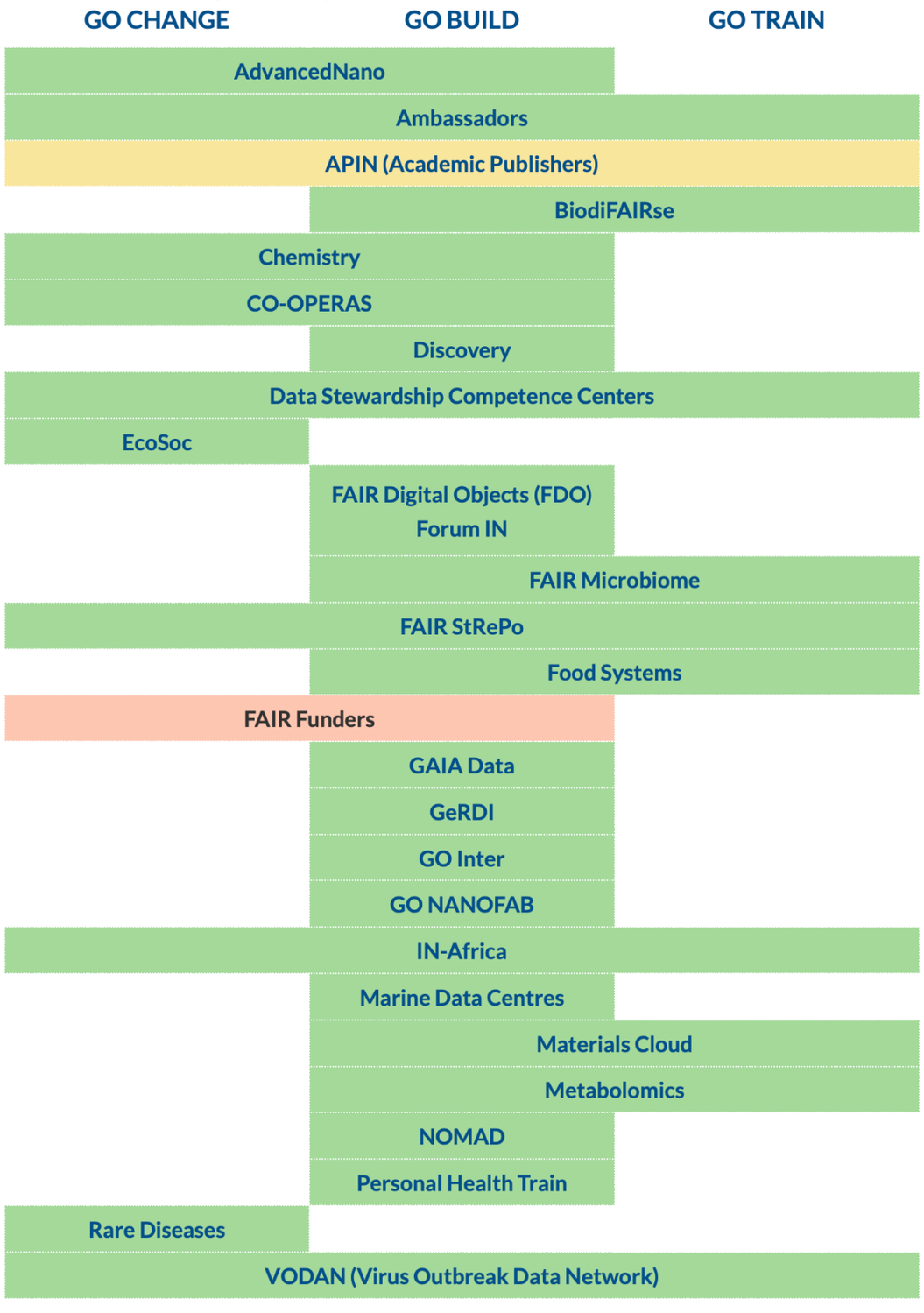
GO-FAIR

Американская свобода



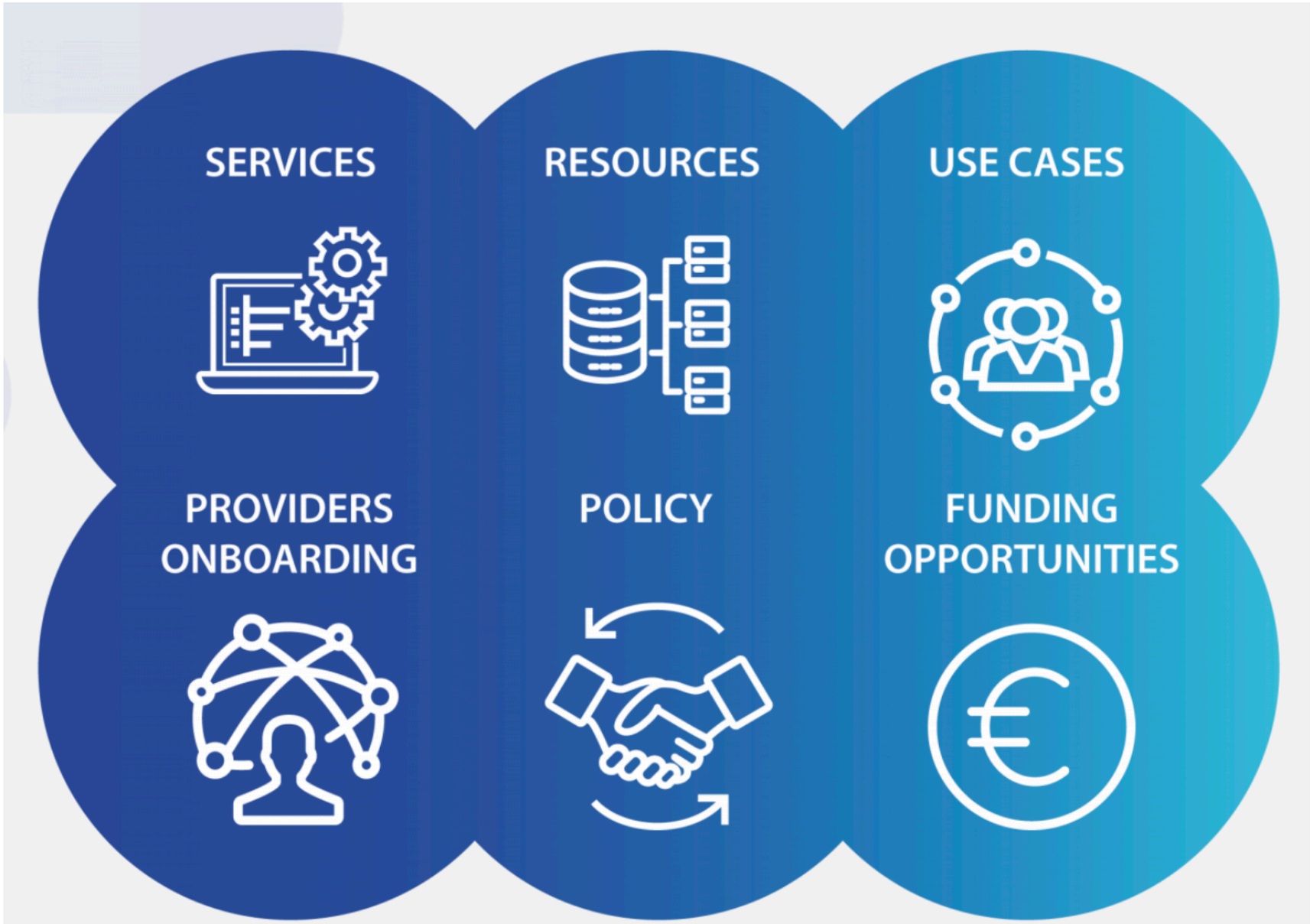
Implementation Networks B

GO-FAIR

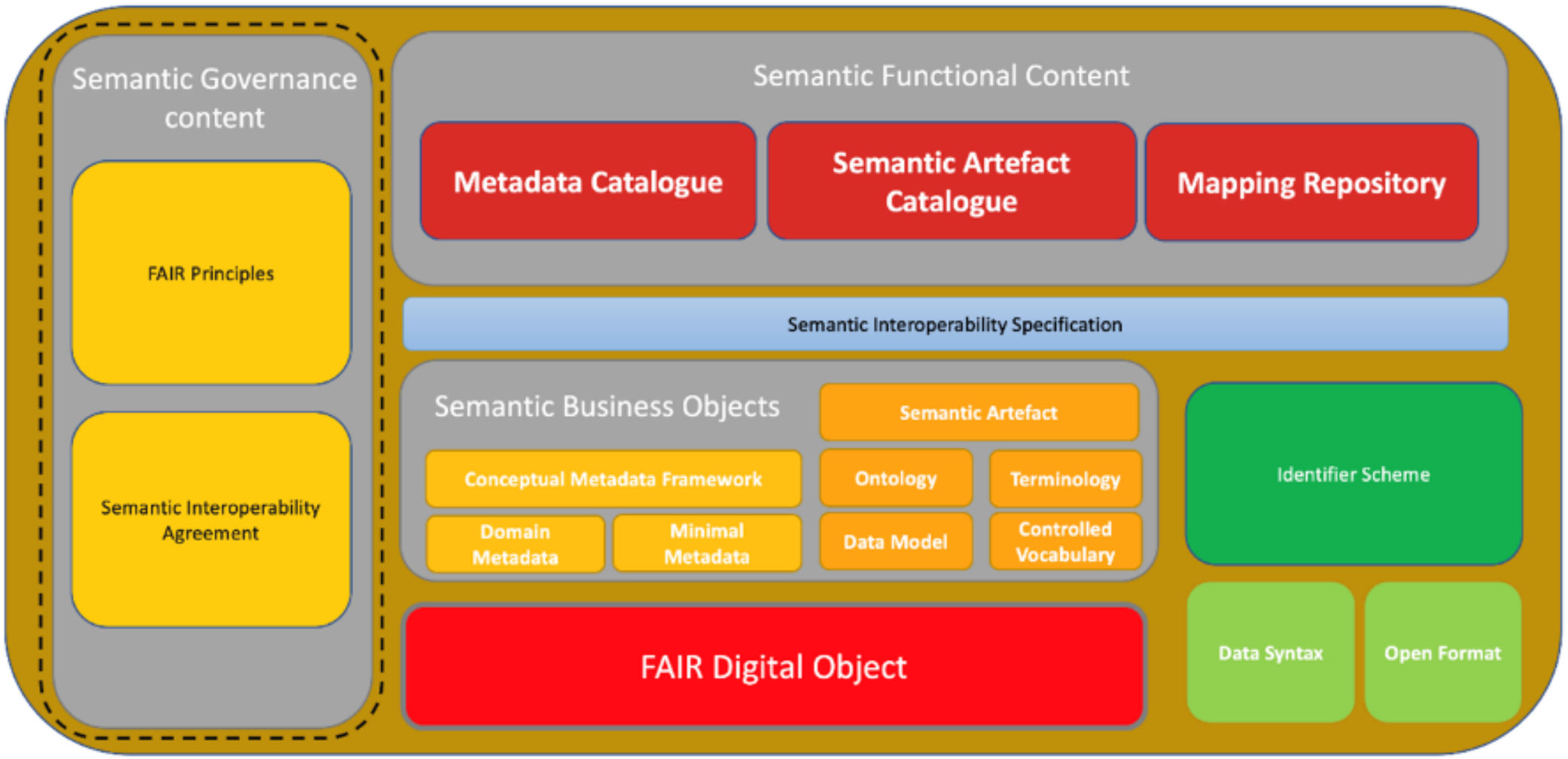


Европейское облако научных данных (EOSC) by EU council

Европейская бюрократия



EOSC: Архитектура



Семантическая интеграция в EOSC

Сервисы в EOSC

EOSC - сервис сервисов

EUROPEAN OPEN SCIENCE CLOUD

Find resource... | All resour... | My EOSC Marketplace

Resources › Processing & Analysis › Data Management › Anonymisation › AMNESIA

AMNESIA

Anonymize your datasets

Organisation: **OpenAIRE**

(0.0 /5) 0 reviews | Add to comparison | Add to favourites

Access the resource

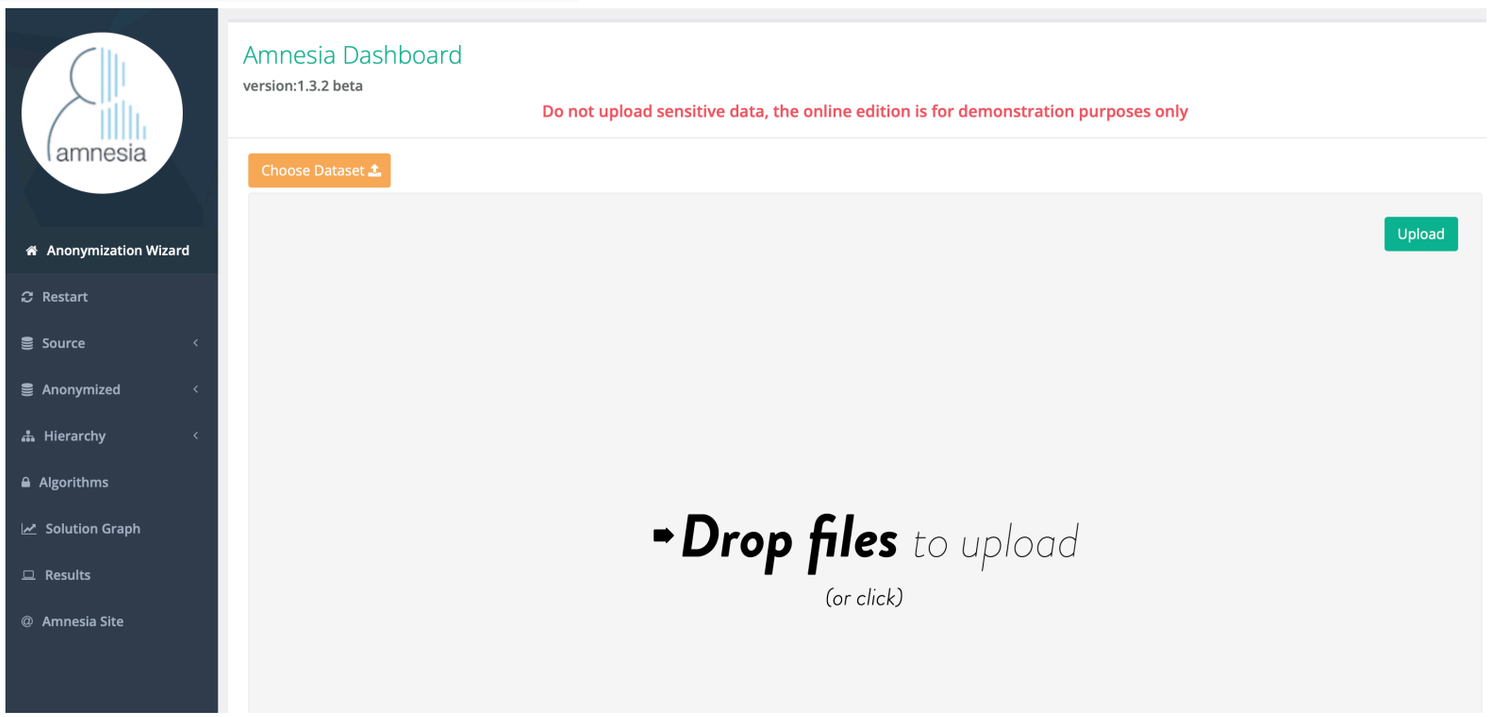
OPEN ACCESS

→ Webpage → Helpdesk → Helpdesk e-mail → Manual

→ Training information

Ask a question about this resource?

ABOUT | DETAILS | REVIEWS (0)



Причем тут онтологии?

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

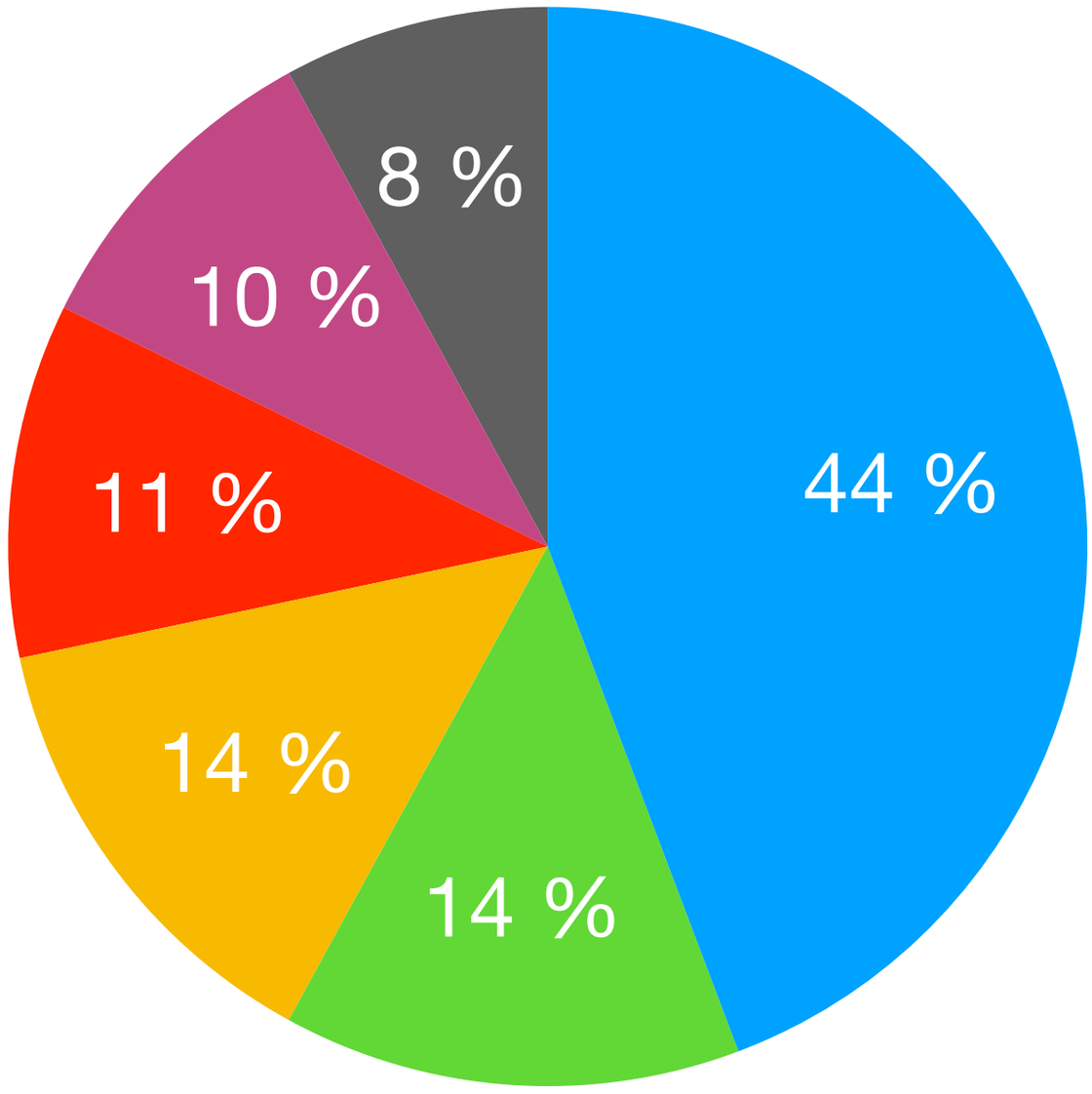
To be Interoperable:
I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation
I2. (meta)data use vocabularies that follow FAIR principles
I3. (meta)data include qualified references to other (meta)data

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

Как много наук честны и открыты?

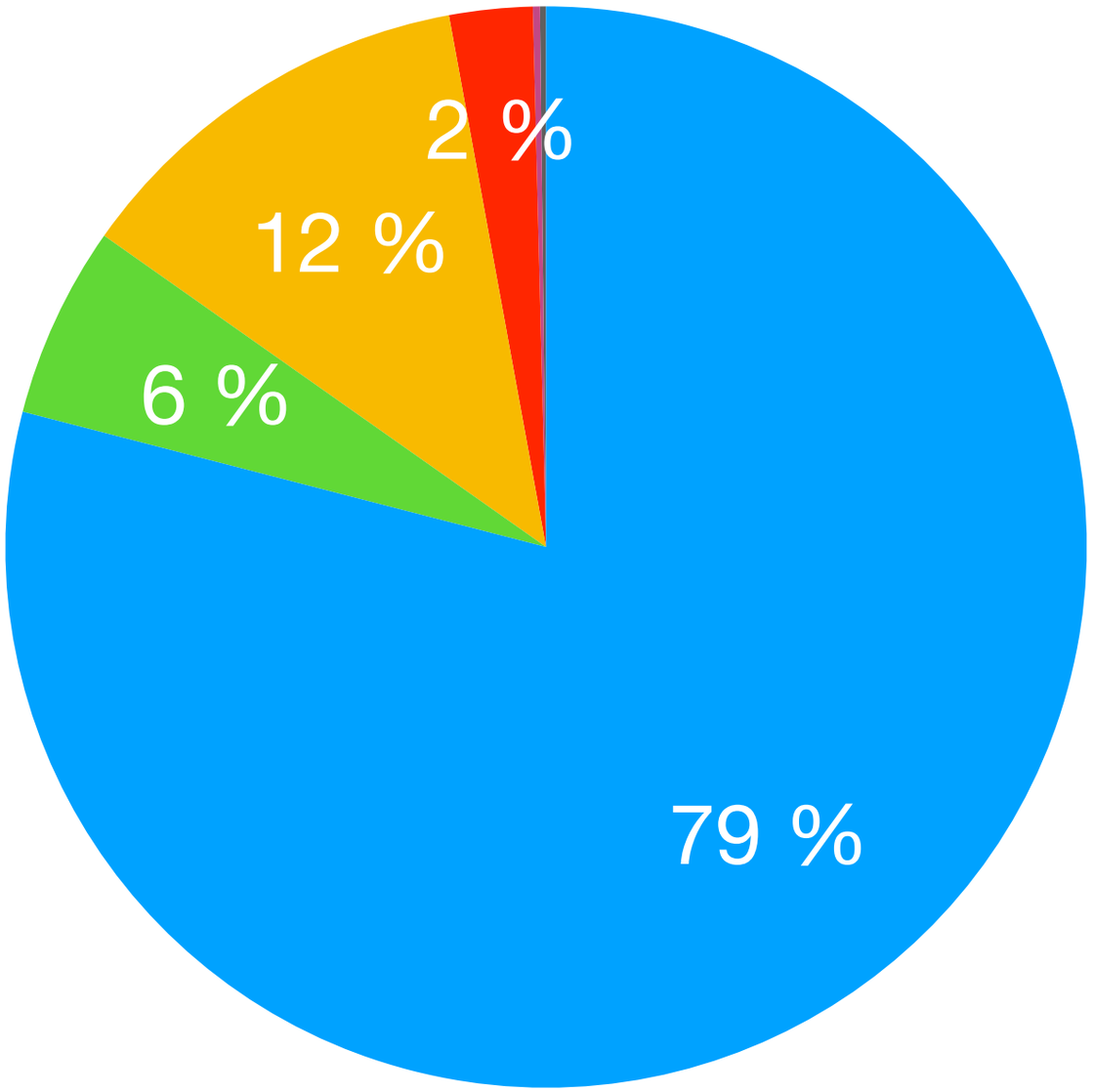
Спойлер: не много

EOSC Resources Domains



- Natural Sciences
- Humanities
- Medical & Health Sciences
- Engineering & Technology
- Social Sciences
- Other


Subjec...
Natural Science
Engineering Science
Humanities and Social Science
Humanities
Social and Behavioural Scien...
Physics
Life Scien...
Mathemat...
Earth Scie...
Chemistry
Aqua...
Paleo...
Philosophy
Ancient C...
Literary S...
Fine Arts
History
Theology
Linguistics
Anthropol...
Prote...
Rom...
Economics
Social Sci...
Educatio...
Psycholo...
Jurisprud...
Industrial...
Mechanic...
Aerospac...
Energy En...
Maritime ...
Materials ...
Electrical ...
Civil Engi...
Materials ...
Systems E...
Computer...
Bioengine...
Process E...
Synthetic ...
Biom...
Ther...
Thermal ...
Safety Sci...
Toxic...



79%

- Natural Science 7886
- Humanities and Social Science 573
- Subject Agnostic 248
- Engineering Science 1233