

SoBigData RI: Infrastructure, Governance, and FAIR Services for Social Data Science

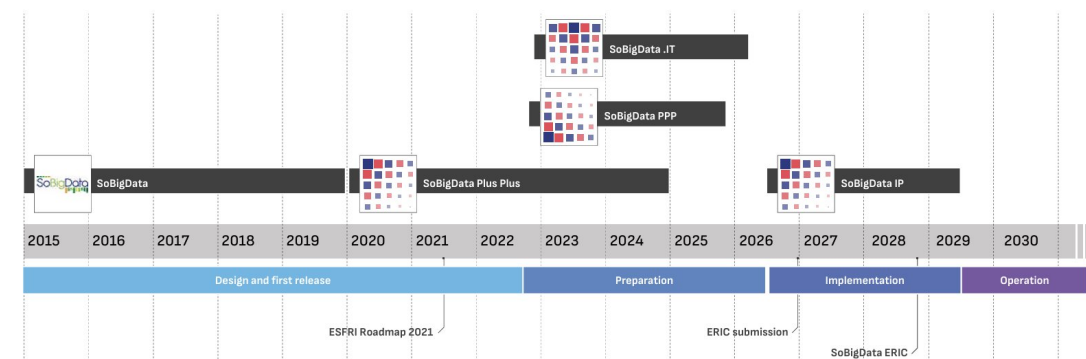
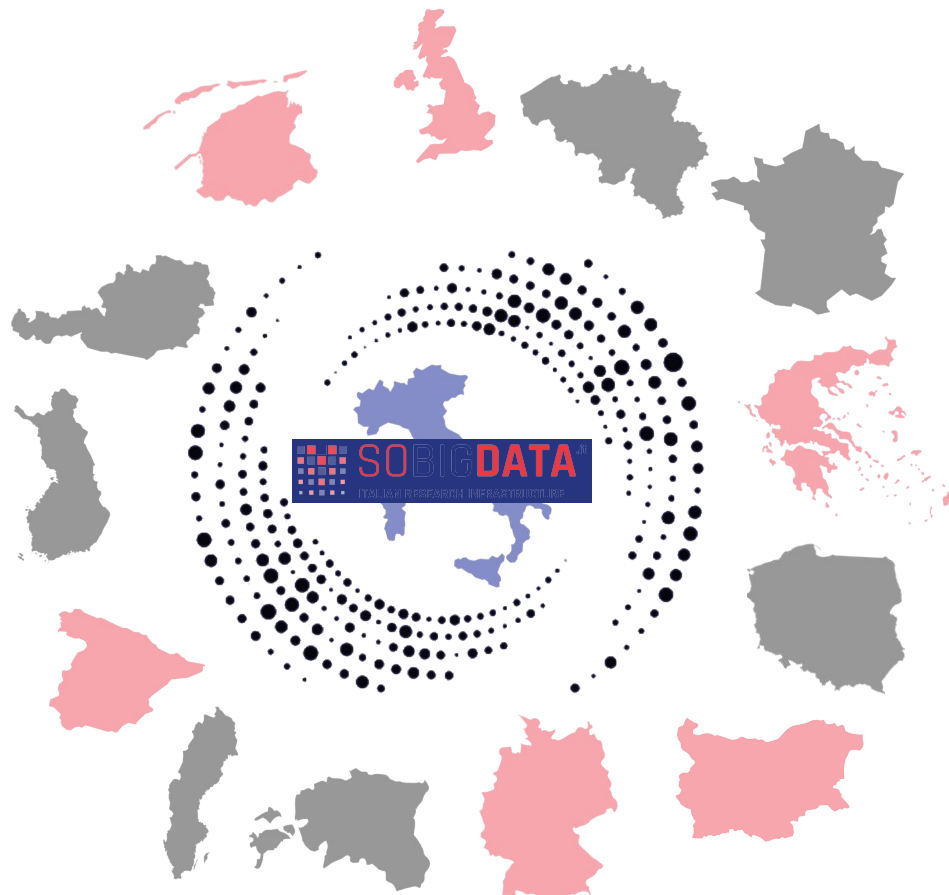
Michela Natilli

CNR – ISTI

SoBigData RI



Use social mining and big data to **understand** the **complexity** of our contemporary, globally interconnected society and offer **services** to researchers, industry, public bodies, and citizens through the creation of a **multidisciplinary** scientific community according to the European vision on **ethics, legality, and open science**.



Why a research infrastructure for social data?

Social data science needs scale, responsibility and reproducibility.

Digital traces from platforms, mobility systems and online interactions can help study collective behaviour – but they also bring hard constraints.

SoBigData RI creates a controlled environment where openness and protection can coexist.

Data quality

heterogeneous, noisy, changing sources

Privacy & legal compliance

sensitive data, consent, access rules

Algorithmic bias

transparent and accountable analytics

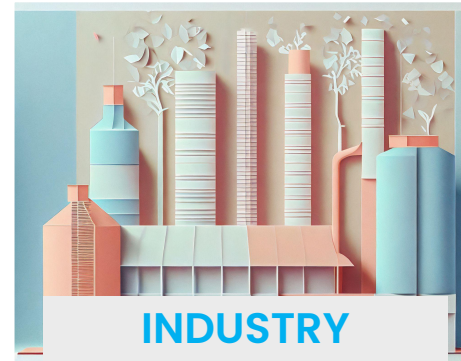
Reproducibility

traceable workflows and reusable resources

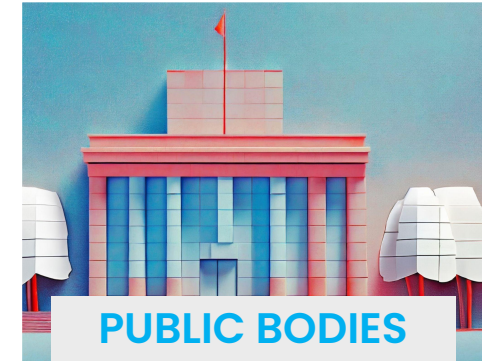
SoBigData RI Activities



- Creating a **multi-disciplinary** scientific community
- Boosting research **quality** among the nodes
- Increasing the opportunities in EU partnerships and calls
- Embracing the European **ethical** and **legal** principles



- Promoting the technological transfer
- Offering consulting and data analysis services
- **Tailored** training courses
- Facilitating the creation of **new** business **opportunities**

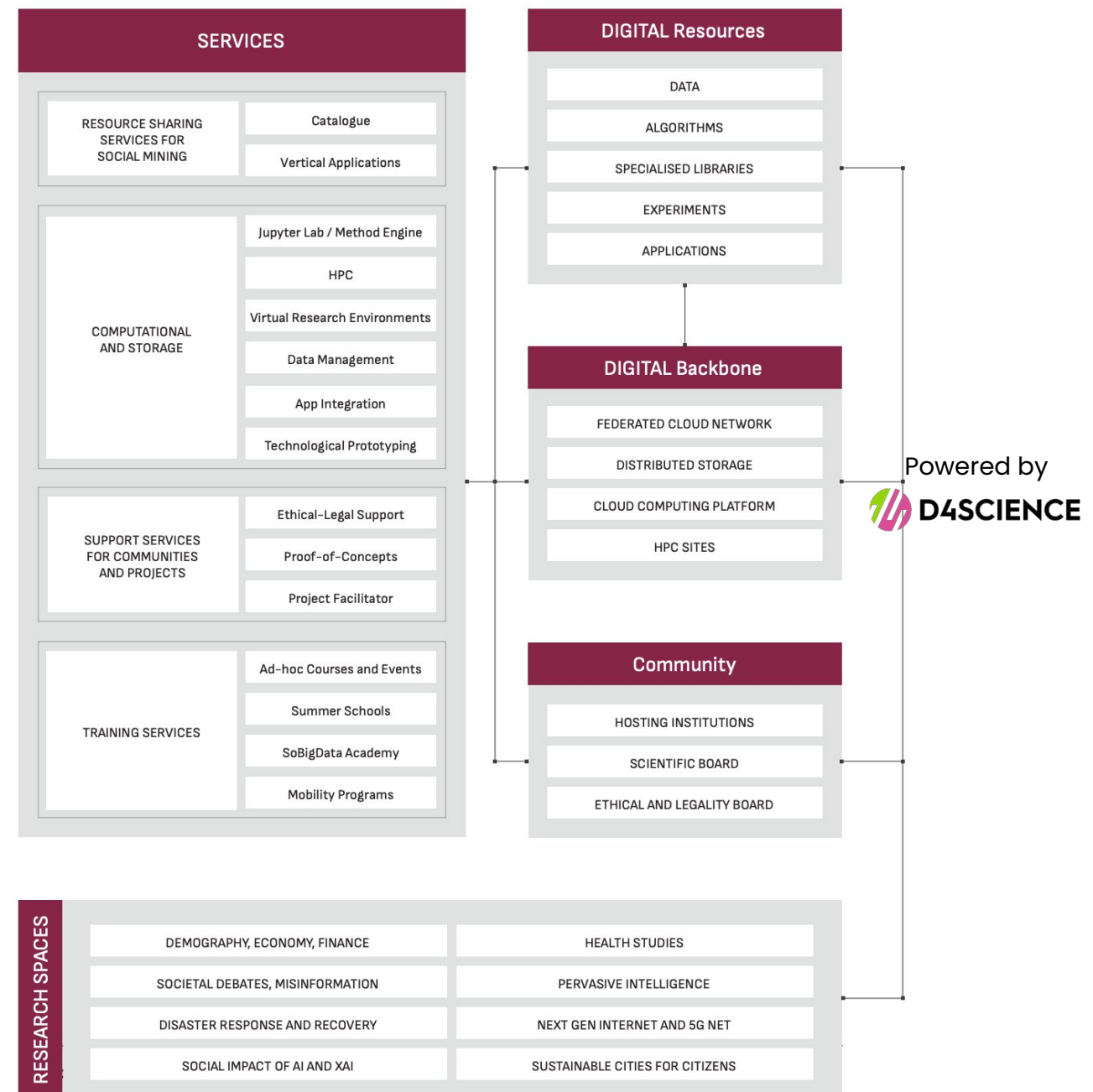


- Giving value to **public data** to understand the territory
- Creating indicators for policy makers
- Designing new services for citizens

Integrated architecture

Five layers:

- **Services**
catalogue, VREs, HPC, training, support
- **Digital resources**
data, algorithms, libraries, experiments, applications
- **Digital backbone**
federated cloud, storage, computing platform, HPC sites
- **Community**
Institutions, collaborations
- **Research spaces**
thematic contexts where resources become reusable science



Research Spaces

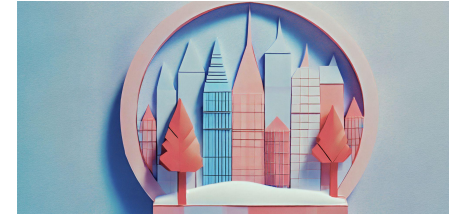
Vertical contexts fostering tangible progress toward societal challenges.



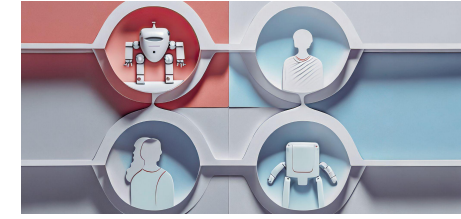
Societal Debates
and Misinformation
Analysis



Demography, Economy
and Finance 2.0



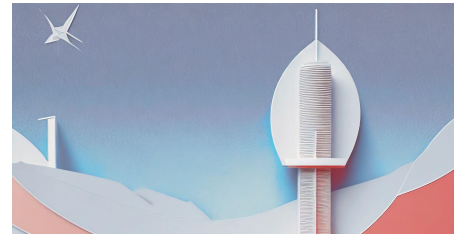
Sustainable
Cities for Citizens



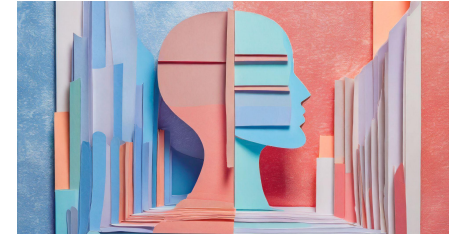
Social Impact of AI
and explainable machine
learning



Health Studies



Societal and Industrial
Impact of Next-Gen.
Internet & beyond 5G
Networks



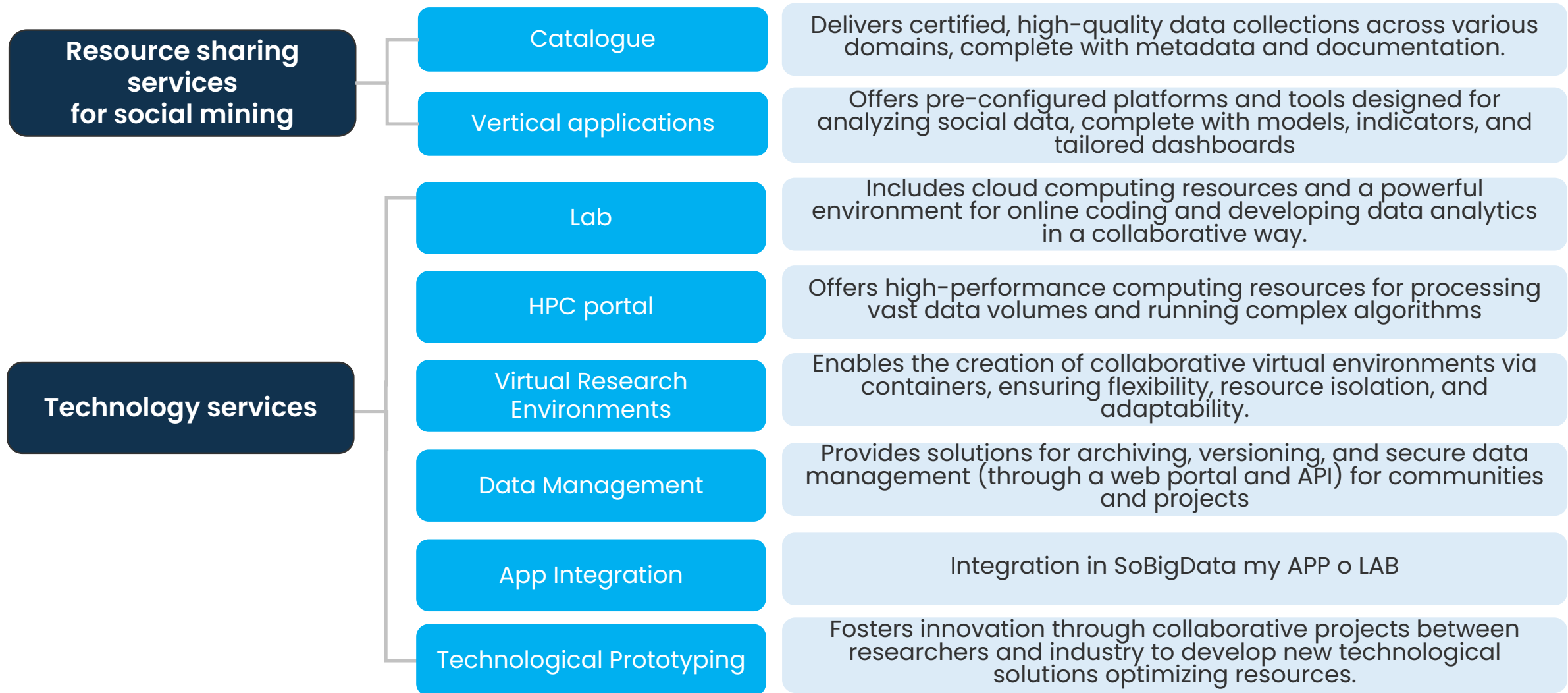
Pervasive Intelligence in
Cyber-Physical Systems
for Future Society



Disaster response
and recovery

More info at: https://www.sobigdata.eu/research_spaces

Services



Services

Support services for communities and projects

Ethical-Legal Support

Provides dedicated consultancy to help address the ethical and legal aspects of Big Data and AI usage

Proof-of-Concepts

Fosters innovation through collaborative experiments to develop data-driven solutions using cloud and technical testbeds

Project Facilitator

Assists organizations in forming strategic partnerships for European funding programs in the context of Social Data and Artificial Intelligence.

Training services

Ad-hoc Courses and Events

Covering topics such as technological innovation, resource optimization, data lifecycle management, and tools provided by SoBigData.

Masters and Summer Schools

Offers intensive, expert-led courses about emerging topics and the practical use of infrastructure tools.

SoBigData Academy

Provides a MOOC designed to train the next generation of responsible data scientists through self-paced, high-quality learning modules.

Mobility programs (TNA)

Facilitates cooperation by allowing international or national institutions and researchers to visit various nodes.

FAIR services: from data to reusable research resources

In SoBigData, FAIR applies not only to datasets, but also to methods, algorithms, experiments and applications.

F

Findable

indexed and discoverable through rich metadata

Public metadata provided

External Identifier	10.5281/zenodo.6367682
Field Scope of use	Any use
Group	Demography, Economy and Finance 2.0
License term	2022-03-01 (2048-12-31)
Manifestation Type	Original
Processing Degree	Secondary
Retention Period	2024-08-20 (2069-12-31)
SoBigData Mode	SoBigData EU
SoBigData Node	SoBigData IT
Sublicense rights	No
Territory of use	World Wide
Thematic Cluster	Human Mobility Analytics (HMA)
Time Coverage	2015-08-01 (2015-10-31)
xyelem:Type	Dataset

A

Accessible

clear conditions: open, restricted or conditional

Data and Resources

Link to data [Explore](#)

Link to scientific article. [Explore](#)

I

Interoperable

standardised descriptions and interfaces

connecting with other catalogues



R

Reusable

context, provenance, licensing and limitations documented

rich description of items



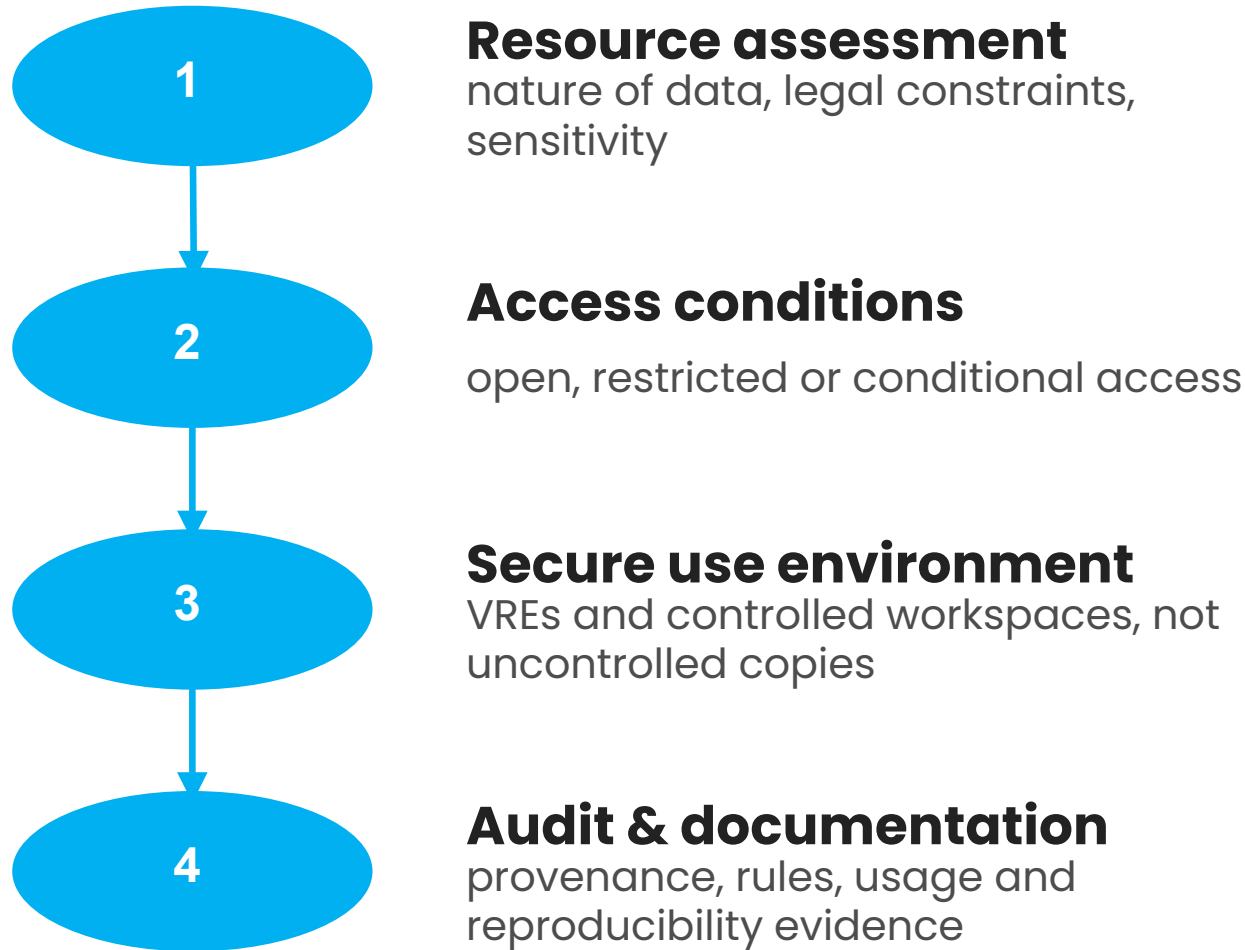
SoBigData Catalogue

main environment for publishing and discovering datasets, methods, tools, experiments and applications

Metadata is the connective tissue: provenance, ownership, licences, access restrictions, methodological context, and ethical/legal limits.

Governance by design

Sharing is controlled, documented and aligned with ethical and legal requirements.



BOEL

Board of Ethics and Legality

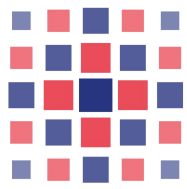
Oversight and assistance
are integrated into
procedures, not added at
the end.

SoBigData connections to EOSC



Primary link





SOBIGDATA

RESEARCH INFRASTRUCTURE

350+	researchers and PhDs
1100+	scientific publications
14,500+	RI users
6,200+	young researchers trained

320+	big datasets
240+	methods and apps
2 mln+	peak access to apps
200+	pilot projects

Updated on 01 April 2026

Thank you

michela.natilli@isti.cnr.it
info@sobigdata.eu