Luciana Taddei
Mario Paolucci *Editors*

Longitudinal Data Infrastructures in Europe

Tools for Open Science in Social Science Research

Springer

Longitudinal Data Infrastructures in Europe

Luciana Taddei • Mario Paolucci
Editors

Longitudinal Data Infrastructures in Europe

Tools for Open Science in Social Science Research

Springer

Editors
Luciana Taddei
Institute for Research on Population and Social Policies
National Research Council
Fisciano (SA), Italy

Mario Paolucci
Institute for Research on Population and Social Policies
National Research Council
Rome, Italy

ISBN 978-3-032-07004-3 ISBN 978-3-032-07005-0 (eBook)
https://doi.org/10.1007/978-3-032-07005-0

This work was supported by Next Generation Foundation (CUP B83C22003950001).

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Acknowledgments This open access book was funded by National Recovery and Resilience Plan (NRRP)—NextGenerationEU, grant number MUR IR0000008 within the FOSSR project "Fostering Open Science in Social Science Research."

We would like to express our sincere gratitude to all those without whose support this work would not have been possible.

We are particularly grateful to *Giuseppe Sindoni* from *ISTAT* for his timely insights and continued engagement. Special thanks also go to the *CONSIP* team, who went above and beyond in assisting us with complex procurement procedures; to the *Direzione Centrale per i Servizi Demografici* of the *Ministero dell'Interno*; and to the staff at the *Autorità Garante per la Protezione dei Dati Personali*, whose collaboration and clarity were crucial throughout.

At the *CNR*, many colleagues have contributed time, expertise, and dedication to this project. While we cannot name everyone, we must acknowledge *Emanuela Reale* for her unwavering commitment, as well as the current leadership team of the *FOSSR* infrastructure—*Giovanni Cerulli* and *Marco Sprocati*—whose support has been central to the realization of this volume, and the *Procurement Office* for practical and legal advice.

To those we have inadvertently overlooked, we ask for understanding—and to the readers of this volume, we ask for the same patience and generosity we have tried to embody in the work presented here.

Contents

Part I Open Science and Research Infrastructures for Social Sciences

1 Open Science and the Role of Social Sciences Research Infrastructures and Data ... 3
Francesca Di Donato

2 Secondary, Longitudinal, and Panel Data in Social Science Research ... 19
Ilaria Primerano, Nicolò Marchesini, Francesco Santelli, Luciana Taddei, and Loredana Cerbara

3 Open Cloud Platform ... 37
Mario Ciampi, Emanuele Damiano, Giovanni Massafra, Pier Giuseppe Meo, and Mario Sicuranza

4 Ethical Implications in Building Longitudinal Data Infrastructures ... 57
Loredana Cerbara, Dario Germani, and Michele Santurro

Part II European Infrastructures and Surveys

5 The Rise of Social Science Infrastructures in Europe and Italy ... 69
Luciana Taddei and Mario Paolucci

6 SHARE-ERIC: A European Infrastructure for Research on Ageing and Retirement ... 81
Agar Brugiavini, Stefano Castaldo, Guglielmo Weber, and Nancy Zambon

7 The Consortium of European Social Science Data Archives (CESSDA) and the Data Archive for Social Sciences in Italy (DASSI) 91
Filippo Accordino, Fabrizio Pecoraro, Daniela Luzi, Carlo Pisano, and Domingo Scisci

8 Retirement, Health, and Digital Gaps: Studying European Ageing with SHARE 107
Chiara Dal Bianco, Guglielmo Weber, and Nancy Zambon

9 GUIDE: Innovations and Challenges to Survey Child Well-Being in Italy 117
Emilio Maria Colella, Giulio Ecchia, Dario Germani, Ilaria Primerano, Michele Santurro, Francesca Tosi, Massimo Ventrucci, and Matthew John Wakefield

10 GGS: Generations and Gender Survey 131
Letizia Mercarini, Nicolò Cavalli, Elena Marseglia, Matilde Perotti, Ilaria Primerano, Michele Santurro, and Nicolò Marchesini

11 The Italian Way to an Online Probability Panel 141
Luciana Taddei, Ferruccio Biolcati Rinaldi, Frank Heins, Nicolò Marchesini, Angela Paparusso, Claudia Pennacchiotti, Francesco Piacentini, Ilaria Primerano, Cristiano Vezzoni, and Michele Santurro

12 Synthetic Populations in Research Infrastructures 153
Rocco Paolillo, Nicholas Roxburgh, Alice Sbrana, Gary Polhill, Evelina Carmen Sabatella, and Mario Paolucci

Open Conclusions 165

Part I
Open Science and Research Infrastructures for Social Sciences

Chapter 1
Open Science and the Role of Social Sciences Research Infrastructures and Data

Francesca Di Donato

1.1 Open Science Definition

Although Open Science is rooted in an ancient idea, the first widely shared definition is relatively recent and can be found in UNESCO's, 2021 Recommendation on Open Science. The Recommendation was prepared through an inclusive, transparent, and regionally balanced consultation process and guided by the Open Science Advisory Committee. The committee involved stakeholders from 193 countries and gathered input collected over two years of consultation (UNESCO, 2023).

In the Recommendation, Open Science is defined as "an inclusive construct that combines various movements and practices aiming to make multilingual scientific knowledge openly available, accessible and reusable for everyone, to increase scientific collaborations and sharing of information for the benefits of science and society, and to open the processes of scientific knowledge creation, evaluation and communication to societal actors beyond the traditional scientific community. It comprises all scientific disciplines and aspects of scholarly practices, including basic and applied sciences, natural and social sciences and the humanities, and it builds on the following key pillars: open scientific knowledge, open science infrastructures, science communication, open engagement of societal actors and open dialogue with other knowledge systems." (UNESCO, 2021, 7).

The proposed definition is rich, composite and complex, as it encompasses and implies multiple aspects that are worth to be broken down and analysed in order to understand the definition terms.

Firstly, Open Science is defined as an inclusive construct, combining movements and practices. The movements referred to are those that gave rise to the media revolution of the past sixty years—particularly, the Internet and the Web -, and to

F. Di Donato (✉)
Institute of Computational Linguistics "A. Zampolli" (CNR-ILC), Pisa, Italy
e-mail: francesca.didonato@cnr.it

L. Taddei, M. Paolucci (eds.), *Longitudinal Data Infrastructures in Europe*,
https://doi.org/10.1007/978-3-032-07005-0_1

the development of free software; these different communities not only share a way of working, but also the principles and values that underpin their research practices (Leiner et al., 2009), along with a philosophy that defines the network (understood as infrastructure, as architecture, and as a set of communities) as a space for collective intelligence (Lévy, 1994, 1997):

1. the early sharing of ideas, grounded in a paradigm of scientific collaboration based on Requests for Comments (RFCs)—formal standards-track documents developed by working groups within the Internet Engineering Task Force (IETF). RFCs are freely available to download, copy, publish, display and distribute, in a variety of formats under a license granted by the IETF Trust, and constitute forms of open peer reviewing (IETF, 2025);
2. the choice to release the Internet and World Wide Web protocols into the public domain, along with the definition of software licences such as the GNU-GPL (Free Software Foundation, 2022) to encourage maximum sharing;
3. the architectural openness of the network, based on a distributed, scalable and flexible topology (Leiner et al., 2009).

Born in response to the so-called serial price crisis (Guédon, 2001), the Open Access movement is the forerunner par excellence of Open Science. Open Access translated the above mentioned principles into policies and practices aimed at the free circulation of scientific knowledge. As the 2003 Berlin Declaration states:

"The Internet has fundamentally changed the practical and economic realities of distributing scientific knowledge and cultural heritage. For the first time ever, the Internet now offers the chance to constitute a global and interactive representation of human knowledge, including cultural heritage and the guarantee of worldwide access."

From then onwards, new possibilities for knowledge dissemination emerged through the Open Access (OA) paradigm via the Internet. The Berlin declaration defines OA"as a comprehensive source of human knowledge and cultural heritage that has been approved by the scientific community", which includes original scientific research results, raw data and metadata, source materials, digital representations of pictorial and graphical materials, and scholarly multimedia material. To provide full access to scientific knowledge, the movement has equipped itself with tools (trusted repositories and platforms for the creation and management of journals, which share common standards such as the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)[1]) and common practices, endorsed by several funders, including the European Commission, and international coalitions, such as COAlition S.[2] Other important tools are Creative Commons[3] (CC) licenses, which have provided an essential framework, borrowed from the experience of the Free Software Foundation's GNU free licenses.[4] CC licenses make it possible to share

[1] https://www.openarchives.org/pmh/

[2] https://www.coalition-s.org/about/

[3] https://creativecommons.org/

and circulate knowledge through a legal infrastructure that relies on existing laws and bends them to allow the free circulation of texts, with a view to publicising public research (Lessig, 2004).

Significantly, the Open Data movement spread in the years immediately following the Open Access movement, which, beyond its official statements, has focused mainly on opening access to the main outputs of research, namely scientific articles (and, to a lesser extent, monographs). The 2012 Royal Society report *Science as an open enterprise* recognises data openness as an essential element of the scientific method, enabling science to evolve and improve. "Open Science is here defined as open data (available, intelligible, assessable and usable data) combined with open access to scientific publications, and effective communication of their contents" (Royal Society, 2012, 16). In scientific terms, this leads to better, higher-quality science, whose effects also bring important social benefits. The four principles mentioned above—availability, intelligibility, assessability and usability—can be considered precursors to the FAIR data principles published three years later (Wilkinson et al., 2016), which are now among the core pillars of Open Science practices.

FAIR is the acronym of Findable, Accessible, Interoperable and Reusable (Di Donato & Provost, 2025):

1. Findable: Humans and machines must be able to search and discover research data. The automatic discovery of datasets and research products is enabled by complete, accurate, machine-readable metadata following existing standards. Persistent Identifiers (PIDs) must be assigned to data, metadata and research objects for long-term identification. Datasets and metadata must be stored in a trusted, publicly accessible repository for long-term preservation.
2. Accessible: Humans and machines should be able to access data and know who can access it. The data and metadata should be accessible through their PID using a standard communication protocol. Accessible data does not imply openness; it means clear access conditions for humans and machines. The data access protocol should be open, free and universal, with authentication and authorisation when required. Metadata should remain accessible even when the data is unavailable.
3. Interoperable: Humans and machines must be able to understand, interpret and integrate data. The data and metadata should be described following recognised community standards for formats, specifications and vocabularies. Context should be provided through qualified references to relevant datasets and metadata. Data is interoperable when both humans and machines can interpret data exchanged between different systems or organisations, supporting interdisciplinary research.
4. Reusable: To reuse data, humans and machines need to be able to do so for future research. This allows experiments to be reproduced, scientific findings

[4] https://www.gnu.org/licenses/licenses.en.html

to be verified and work to be built on previous analyses. To facilitate this, data and metadata should be described according to community standards and accompanied by rich documentation. This documentation should include contextual information such as data provenance and processing, and a licence that indicates to machines and humans the conditions under which the data can be used.

These principles differ from the concept of Open Data used in the Royal Society report, focusing on the transparency and interoperability of machine-readable data through ontologies, rather than on access and availability for use (Mons et al., 2017; Ramachandran et al., 2021). For this purpose, as we will see later, research infrastructures that support open science and serve the needs of different communities are essential.

1.2 Open Science Concepts: Values and Principles

The UNESCO Recommendation definition contains key concepts which refer to an ancient tradition that is recalled in the Budapest Open Access Initiative (BOAI) declaration (2002), and opens as follows: "an old tradition and a new technology have converged to make possible an unprecedented public good".

The above-mentioned old tradition embodies values and principles rooted in Western philosophical thought since its origins. First, collaboration, which is the foundation of the Socratic-Platonic dialectic. In the Meno dialogue, Socrates contrasts the eristic method, based on a competitive form of refutation, with the dialectic method. Dialectic, unlike the strategy adopted by the Sophists, is based on shared premises and on a collaborative methodology (Plato, 2005). Second, philosophy is understood as a way of life (Hadot, 1995), grounded in the existence of communities of practice, i.e. communities of people who critically discuss the facts of nature in the broadest sense in a collaborative manner, as "friends have all things in common" (Plato, 2002, 279b-c).

Following this tradition, openness, intended as the abandonment of secrecy, is one of the fundamental steps in the birth of modern science. While the idea that essential knowledge was secret prevailed in European culture for many centuries, it was in the modern age that the communication and dissemination of knowledge, including the public discussion of ideas, became values, alongside the emergence of linguistic rigor and the non-allusive nature of terminology. (Rossi, 2000). In this frame, availability, accessibility and reusability are fundamental to both the ethics and methodology of research. Science is in fact based on the idea of open, publicly verifiable debate grounded in the foundations of knowledge. It also rests on respect for rules: starting from the observation of nature in order to formulate hypotheses and explanatory models. These rules were codified as the Galilean method.

As early as 1610, Galileo Galilei understood the necessity to make everything known to everyone and claimed that his discoveries should be subject to free scrutiny by the majority of people (Galilei, 1610). To do this, Galilei had telescopes built so that others could verify his discoveries with their own eyes (Greco, 2010). The aim of this collective and public endeavour is to establish a scientific community whose goal is to build a rational consensus on the facts of nature in a transparent way. The communication of the results of this work to the entire community ensures that everyone can not only know them, but also verify them, thereby contributing to the construction of scientific consensus. However, there is another important consequence of abandoning the paradigm of secrecy: The awareness that science is a social enterprise whose development is governed by precise rules shared by experts is affirmed, but it is also acknowledged that it has enormous social effects, and that confrontation and dialogue with society as a whole cannot be shied away from. On the contrary, society should encourage it (Greco, 2010).

The English expression 'Open Science' appeared in the second half of the Nineteenth century to refer to modern science, which a century later would be defined as the science of the scientific revolution in that it was open, and opposed to occult or initiatory science. By the end of the century, all this translated into a great effort to disseminate scientific knowledge and into the principle of communism, as defined by Merton in his 1942 essay *The normative structure of science* as one of the four institutional imperatives constituting the ethos of modern science, here understood "in the nontechnical and extended sense of the common ownership of goods" (Merton, 1942, 273).

According to this fundamental principle, foundational discoveries belong to the community and constitute a common heritage in which the rights of the individual producer are severely limited. This principle is linked to the imperative of communicating results. "Secrecy is the antithesis of this norm; full and open communication its enactment" (Merton, 1942, 274). Those who do not embrace it are selfish or anti-social. The well-known remark attributed to Newton, "If I have seen farther, it is by standing on the shoulders of giants," (273–74) expresses both a sense of debt to the pre-existing common heritage and a recognition of the essentially collaborative and cumulative nature of scientific results. The scientific method itself is part of this heritage. Merton also anticipates something that Paul David will later observe, namely, that the communism inherent to the scientific ethos is incompatible with the capitalist conception of technology as private property.

The technical and political meaning of Open Science that we commonly refer to dates back to the late twentieth century. The term was first used by Chubin, who observes how the imperative of openness, made explicit through the principle of communism, is in practice challenged by opposing tendencies that make closure and secrecy the usual behavior (Chubin, 1985).

This distinction and tension was taken up in the early 2000s by Paul David (David, 2000, 2014),[5] who defined the two opposing models, that of Open Science and that of commercial science, which embody and represent the tension between the principles of collaboration and competition. The first model — that of Open Science — is based on the principle of collaboration. According to David, it was in

[5] His work, which is more economic than philosophical in nature, is cited in numerous papers in the literature.

the 1980s and 1990s that international organizations established principles, (OECD, 2015) and developed guidelines to protect easy and broad access to scientific data and high-quality information generated by publicly funded entities, and that recommendations began to circulate through implementations.

In his 2011 work *Reinventing Discovery: The New Era of Networked Science*, Michael Nielsen provides a thorough discussion of how digital technologies and online tools are changing the way scientists work. He gives many examples of the practical applications of collective intelligence, open source, Open Science and citizen science. Referring explicitly to Vannevar Bush (1945), Ted Nelson (1987), Pierre Lévy (1994), Douglas Engelbart (1962), Tim Berners-Lee (1999), Eric Raymond (2001), Yochai Benkler (2006) and Clay Shirky (2010), Nielsen describes the transition to open and collaborative science.

It is from this composite framework that the principles and values of Open Science emerge. The UNESCO recommendation lists four Open Science main values (UNESCO, 2021, 17). These values arise from the various effects of making science available to society and applying the rules of openness to every stage of the scientific research process:

1. *Quality and integrity*. Research must be subject to rigorous scrutiny and support academic freedom and human rights. High-quality research is achieved through collaboration and widespread accessibility of scientific methods and results.
2. *Collective benefit*. Science is a global public good, therefore its benefits must be shared universally, and scientific knowledge must be available to all. To achieve this, science must be inclusive, sustainable and equitable.
3. *Equity and fairness*. Scientific knowledge should be accessible to all, regardless of background. No one should be excluded from science or subjected to differential treatment. Producers and consumers of scientific knowledge should have equal opportunities to participate in research.
4. *Diversity and inclusiveness*. Science must be inclusive of all communities. This means encouraging and supporting a wide variety of knowledge practices, workflows, languages and research outputs.

Values are translated into a methodology through the following six guiding principles (UNESCO, 2021, 18–19), which provide a framework for the conditions and practices necessary to uphold these values and make the Open Science vision a reality:

1. *Transparency, scrutiny, critique and reproducibility*. To enhance the impact of science on society and tackle global challenges, we must maximise openness at all stages of the research lifecycle. Greater transparency in scientific data leads to greater openness, which in turn improves trust.
2. *Equality of opportunities*. All people, irrespective of background, should equally access, contribute to and benefit from scientific knowledge. This principle ensures science is open to all, so that all voices can contribute to progress.
3. *Responsibility, respect and accountability*. Researchers and actors involved in the scientific process must conduct research with integrity and be aware of its

wider impacts. This principle highlights the importance of public accountability, conflict of interest avoidance, research integrity and ethics.

4. *Collaboration, participation and inclusion*. To tackle large, complex, societal problems, scientists must collaborate, overcoming barriers related to geography, language, age, resources. Cross-disciplinary collaboration and the effective involvement of diverse knowledge systems, particularly marginalised communities, is essential.
5. *Flexibility*. Research contexts and capabilities vary worldwide, so there is no one-size-fits-all approach to science. This principle promotes flexibility, allowing for various strategies to achieve Open Science goals while upholding core values.
6. *Sustainability.* Ensuring long-term efficiency and impact of science requires sustainable practices, infrastructures and financial models. These should guarantee that researchers from all organisations and countries can take part in knowledge production. The concept of sustainability shows this is best done through not-for-profit, long-term infrastructures which fund Open Science practices, ensuring access to all.

As shown by Nielsen, digital technologies are instrumental to scientific collaboration and to implement Open Science values and principles. In both Nielsen's and the UNESCO definition, the old tradition and the new technologies mentioned in the BOAI are closely linked.

1.3 Open Science Infrastructures

The UNESCO Recommendation identifies specific areas for action, focusing on the necessary steps for governments and communities to implement it. These include Open Science infrastructures, defined as shared research infrastructures that support Open Science and serve the needs of different communities. Along with Open Science knowledge, open engagement of societal actors and open dialog with other knowledge systems, Open Science infrastructures are one of the four key pillars of Open Science referred to in the definition, and represented in the left part of Fig. 1.1.

They are also mentioned in the sustainability principle, "Open science infrastructures are often the result of community-building efforts, which are crucial for their long-term sustainability and therefore should be not-for-profit and guarantee permanent and unrestricted access to all public to the largest extent possible" (UNESCO, 2021, 12).

Fecher and Friesike identified five Open Science schools of thought, one of which focuses on infrastructure. This school of thought "regards Open Science as a technological challenge" (Fecher & Friesike, 2013, 36), with two specific trends. The first is distributed computing, a concept and practice closely linked to the development of the internet and related technologies. The second is the creation of social and collaboration networks for researchers — tools that enable them to

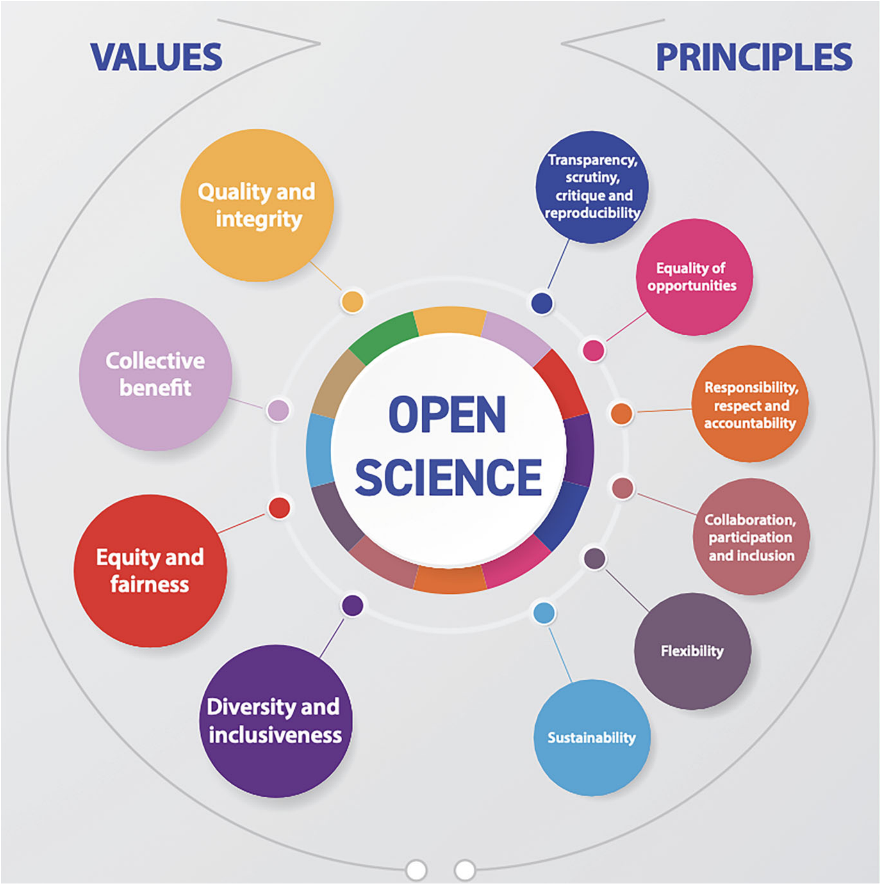

Fig. 1.1 Open Science core values and principles. © UNESCO, 2022. https://doi.org/10.54677/DLYW1405. Attribution-ShareAlike 3.0 IGO (CC-BY-SA 3.0 IGO) license

interact and collaborate. The technological element is thus combined with social aspects of fundamental importance.

Cameron Neylon (2013) recognised the lack of shared infrastructure as a major problem for Open Science uptake. In 2015, the *Principles for Open Scholarly Infrastructure* (POSI) were created to deal with the fact that scientific publications and datasets were becoming more accessible, but the infrastructure for sharing them was not keeping up. The Principles are three (Bilder et al., 2020):

- *Coverage across the scholarly enterprise*—research transcends disciplines, geography, institutions, and stakeholders. Organisations and the infrastructure they run need to reflect this.

- *Stakeholder Governed*—a board-governed organisation drawn from the stakeholder community builds confidence that the organisation will take decisions driven by community consensus and a balance of interests.
- *Non-discriminatory participation or membership*—we see the best option as an "opt-in" approach with principles of non-discrimination and inclusivity where any stakeholder group may express an interest and should be welcome. Representation in governance must reflect the character of the community or membership.
- *Transparent governance*—to achieve trust, the processes and policies for selecting representatives to governance groups should be transparent (within the constraints of privacy laws).
- *Cannot lobby*—infrastructure organisations should not lobby for regulatory change to cement their own positions or narrow self-interest. However, an infrastructure organisation's role is to support its community, and this can include advocating for policy changes.
- *Living will*—a powerful way to create trust is to publicly describe a plan addressing the conditions under which an organisation or service would be wound down. It should include how this would happen and how any assets could be archived and preserved when passed to a successor organisation or service. Any such organisation or service must adopt POSI and honour the POSI principles.
- *Formal incentives to fulfil mission & wind-down*—infrastructures exist for a specific purpose, and that purpose can be radically simplified or even rendered unnecessary by technological or social change. Organisations and services should regularly review community support and the need for their activities. If it is possible, the organisation or service (and staff) should have direct incentives to deliver on the mission and wind down.
- *Time-limited funds are used only for time-limited activities*—operations are supported by sustainable revenue sources—whereas time-limited funds are used only for time-limited activities. Depending on grants to fund ongoing and/or long-term infrastructure operations fully makes them fragile and distracts from building core infrastructure.
- *Goal to generate surplus*—organisations (or services) that define sustainability based merely on recovering costs are brittle and stagnant. It is not enough to merely survive; organisations and services have to be able to adapt and change. To weather economic, social and technological volatility, they need financial resources beyond immediate operating costs.
- *Goal to create financial reserves*—a high priority should be having ring-fenced financial reserves, separate from operating funds, that can support implementing living will plans, including a complete, orderly wind down or transition to a successor organisation, or major unexpected events.
- *Mission-consistent revenue generation*—revenue sources should be evaluated against the infrastructure's mission and not run counter to the aims of the organisation or service.

- *Revenue based on services, not data*—data related to the running of the scholarly infrastructure should be community property. Appropriate revenue sources might include value-added services, consulting, API Service Level Agreements or membership fees.
- *Open source*—all software and assets required to run the infrastructure should be available under an open-source licence. This does not include other software that may be involved with running the organisation.
- *Open data (within constraints of privacy laws)*—for an infrastructure to be forked (reproduced), it will be necessary to replicate all relevant data. The CC0 waiver[6] is the best practice in making data openly and legally available. Privacy and data protection laws will limit the extent to which this is possible.
- *Available data (within constraints of privacy laws)*—it is not enough that the data be "open" if there is no practical way to obtain it. Underlying data should be made easily available via periodic open data dumps.
- *Patent non-assertion*—the organisation should commit to a patent non-assertion policy or covenant. The organisation may obtain patents to protect its own operations but not use them to prevent the community from replicating the infrastructure."

Subsequent to being revised in 2020, 2023 and 2025, respectively, these principles have exerted a considerable influence and continue to serve as a point of reference. Several definitions, including the UNESCO Recommendation, have been developed on this basis.

According to them it is worth underlining that Open Science infrastructures should not be regarded as a mere technical product; rather, they should be considered as a complex system that incorporates a range of tools, institutions and social norms. It is evident that the notion of openness cannot be regarded solely as a technical attribute; rather, it is a fundamental value and a guiding principle that exerts a profound influence on the objectives, governance and management of the infrastructure (Fecher et al., 2021).

More in general, in recent years the development of open scientific infrastructure has become a topic of debate regarding the future of online scientific research. In January 2021, a group of researchers made a call for a Plan Infrastructure (Plan I), aiming at integrating all research outputs on large interoperable infrastructures. In fact, the dependency of research and scholarship on an information infrastructure that treats all scholarly outputs equally, and that is based on open standards and open markets, is paramount. (Brembs et al., 2021, 4).

More recently, the Barcelona Declaration on Open Research Information[7] was launched, whose commitments 2 ('We will work with services and systems that

[6] Thanks to which creators are allowed to waive all copyright and related rights in their works, making them freely available for any use without restrictions. Unlike other Creative Commons licenses, CC0 does not require attribution.

[7] https://barcelona-declaration.org/

support and enable open research information') and 3 ('We will support the sustainability of infrastructures for open research information') insist on the need for open, public, sustainable infrastructures, recognising and highlighting their central role as enabling factors in the creation of an open research information space.

It has been observed that the majority of landscape reports on Open Infrastructure have been carried out in Europe—with a lesser number of reports conducted in Latin America (Langlais, 2023).

Of particular note are the efforts made since 2016 to establish the European Open Science Cloud (EOSC),[8] an ecosystem of research infrastructures based on open procedures and standards, which allows data, resources, and knowledge to be shared—thus enabling Open Scholarship to become a reality (European Commission, 2016; Burgelman et al., 2019). This project is being implemented through specific funding programmes (first Horizon 2020 and now Horizon Europe) and its development will continue in the next framework programme, FP10.

Since early 2019, five European Strategy Forum on Research Infrastructures (ESFRI) Cluster projects—including SSHOC, the Social Science and Humanities Open Cloud project—have been initiated with the objective of establishing a connection to the European Open Science Cloud. The projects were invited to contribute to the development and implementation process, with the aim of working together to implement interfaces, integrate computer and data management solutions, create cross-border and open cooperation spaces, and promote clouds via the EOSC portal for a larger user community. (Gotz et al., 2020).

The 2021 ESFRI Roadmap has explicitly embraced Open Science principles: "Most of the Research Infrastructures on the ESFRI Roadmap are at the forefront of the Open Science movement and make important contributions to the digital transformation by transforming the whole research process according to the Open Science paradigm" (ESFRI, 2021, 159).

In Italy, the National Research Programme (PNR 2021–2027) provides strategic guidance for the country's research policies and refers to two strategic documents: the National Plan for Open Science (PNSA) to implement Open Science policies, (MUR, 2022) and the National Research Infrastructure Plan (PNIR) for specific aspects related to research infrastructures. "RIs—as stated in the second document—are crucial to our ability to make scientific progress and promote innovation, and play an enabling role in research and innovation to achieve the most challenging goals set at European and national level" (DM n. 1082, 2021, 3). The document aims to provide more detail on the technical and strategic plan for Research Infrastructures, by defining and updating national priorities. Because of this unique potential, which is inherent to RIs, the document includes both the European and national plans, which are considered closely connected. At the European level, in particular, the document takes its cue from the European framework programme Horizon Europe, which supports the cross-cutting actions

[8] https://research-and-innovation.ec.europa.eu/strategy/strategy-research-and-innovation/our-digital-future/open-science/european-open-science-cloud-eosc_en

of the new European Research Area (ERA), emphasising the fundamental role that RIs can play in this context, especially considering the significant contributions the ESFRI has made to advancing the sector in Europe. At national level, in addition to the PNR, the strategic and forward-looking contribution of the National Recovery and Resilience Plan[9] is fundamental, as it considers RIs as a key factor for the development of the country. Through the infrastructure and projects that build it, including FOSSR—Fostering Open Science in Social Science Research, a space is created to develop tools and services to meet the real needs of researchers.

1.4 Conclusions

In recent years, it has become clear that data on social behaviour and cultural practices (past and present) is essential for addressing key societal issues such as climate change, environmental sustainability, energy transition, migration management, health promotion, and disease prevention. This data must be considered alongside a clear acknowledgement of the importance of ethical, legal, and societal aspects (ESFRI, 2021).

The areas enabled by the FOSSR project are varied and include, among others, the development of research services, which range from data collection, curation, analysis and preservation, online longitudinal panels and survey maker, virtual research and simulation environments, applications for policy evaluation and forecasting, ontogeny-semantics and text analysis, as well as training and documentary and consultancy support. Structuring research as Open Science practices promises to enhance transparency, accountability and inclusiveness.

Social science infrastructures enable the observation and comparison of contemporary societies over time—both synchronically and diachronically. More specifically, panel data—which refers to information collected by the same researchers repeatedly over a period—and longitudinal data—which tracks changes within a sample population over time, regardless of whether the same individuals are observed—are two essential components in the toolkit for social scientists, and are fundamental to research in general. All these efforts underscore a commitment to generate high-quality and robust datasets that can drive insightful social research and inform policy decisions on a national and international scale—and are essential in fostering informed policy-making and active citizenship. SHARE ERIC,[10] CESSDA[11] and its Italian national service DASSI,[12] the GUIDE ESFRI research

[9] https://www.italiadomani.gov.it/content/sogei-ng/it/en/home.html

[10] https://share-eric.eu/

[11] https://www.cessda.eu/

[12] https://www.cessda.eu/About/Consortium-and-Partners/List-of-Service-Providers/Italy-spid1908

infrastructure,[13] and the Generation and Gender survey (GGS),[14] offer a rich environment for both social science and multidisciplinary research. Some of these infrastructures will be detailed in the next chapters of this book.

Despite the central role of Social Sciences data, much remains to be done in the design and development of tools to improve collaboration, data sharing and reuse in Social Sciences, and thus overcome the fragmentation that characterises them, as highlighted in the ESFRI Roadmap 2021. Rafols (2025) examines three well-known models that show the marginalisation of the Social Sciences and Humanities in priority setting in relation to societal challenges, and argues that processes for assessing these challenges should include SSH research to ensure they are sensitive and responsive to society's needs—and that, as a result, certain issues should be prioritised over others. This is an important proposal that deserves further exploration. What is important to emphasise here is that Open Science, and specifically its Social Science infrastructures, data, and tools have a great deal to offer if we are to respond to the global challenges we face today.

References

Budapest Open Access Initiative (BOAI). (2002). https://www.budapestopenaccessinitiative.org/read/

Bilder, G., Lin, J., & Neylon, C. (2020). *The principles of open scholarly infrastructure*. Retrieved June 2025, from https://doi.org/10.24343/C34W2H

Brembs, B., Förstner, K., Goedicke, M., Konrad, U., Wannemacher, K., & Kett, J. (2021). Plan I – Towards a sustainable research information infrastructure. *Zenodo*. https://doi.org/10.5281/zenodo.4454640

Burgelman, J.-C., Pascu, C., Szkuta, K., Von Schomberg, R., Karalopoulos, A., Repanas, K., & Schouppe, M. (2019). Open Science, open data, and open scholarship: European policies to make science fit for the twenty-first century. *Frontiers in Big Data, 2*, 43. https://doi.org/10.3389/fdata.2019.00043

Benkler, I. (2006). *The wealth of networks*. Yale University Press.

Berners-Lee, T. (1999). *Weaving the web: The original design and ultimate destiny of the world wide web*. Harper.

Bush, V. (1945). As we may think. *The Atlantic Monthly Review*, pp. 101–108. https://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/

Chubin, D. E. (1985). Open Science and closed science: Tradeoff in a democracy. *Science Technology and Human Values, 1985, 10*(2), 73–81.

CIPE. *PNR* 2021*–2027 - Programma Nazionale della Ricerca*. https://www.mur.gov.it/sites/default/files/2021-01/Pnr2021-27.pdf

David, P. (2000). The historical origins of 'open science', an essay on patronage, reputation and common agency contracting in the scientific revolution. *MERIT Working Papers* 2014-082. United Nations University - Maastricht Economic and Social Research Institute on Innovation and Technology (MERIT). https://ideas.repec.org/p/sip/dpaper/06-038.html.

[13] https://roadmap2021.esfri.eu/projects-and-landmarks/browse-the-catalogue/guide/

[14] https://www.ggp-i.org/generations-and-gender-survey/

David, P. (2014). The republic of open science. The institution's historical origins and prospects for continued vitality. *Contributi del Centro Linceo Interdisciplinare 'Beniamino Segre'*.

Decreto Ministeriale n.1082. (2021) del 10-09-2021 - Piano Nazionale Infrastrutture di Ricerca (PNIR) 2021–2027. https://www.mur.gov.it/sites/default/files/2021-10/Decreto%20Ministeriale%20n.1082%20del%2010-09-2021%20-%20PNIR%202021%20-%202027.pdf

Di Donato, F., & Provost, L. (2025). Why isn't FAIR enough? Bringing together methods and values for Open Science uptake. *Umanistica Digitale, 9*(19), 17–46. https://doi.org/10.6092/issn.2532-8816/20976

Engelbart, D. C. (1962), *Augmenting human intellect: A conceptual framework*. Stanford Research Institute Report.

European Commission. (2016). *Open innovation, open science, open to the world - a vision for Europe*. https://ec.europa.eu/digital-single-market/en/news/open-innovation-open-science-open-world-vision-europe

European Strategy Forum on Research Infrastructures (ESFRI) (2021). Roadmap 2021. Strategy Report on Research Infrastructures. https://roadmap2021.esfri.eu/media/1295/esfri-roadmap-2021.pdf

Fecher, B., & Friesike, S. (2013). *Open science: One term, five schools of thought*. Id., Opening science: The evolving guide on how the internet is changing research, collaboration and scholarly publishing. Springer International Publishing.

Fecher, B., Kahn, R., Sokolovska, N., Völker, T., & Nebe, P. (2021). Making a research infrastructure: Conditions and strategies to transform a service into an infrastructure. *Science and Public Policy, 48*(4), 499–507. https://doi.org/10.1093/scipol/scab026. ISSN 0302-3427

Free Software Foundation. (2022). *Licenses*. https://www.gnu.org/licenses/licenses.html.en

Galilei, G. (1610). *Letter to Belisario Vinta*. https://brunelleschi.imss.fi.it/galileopalazzostrozzi/object/GalileoGalileiLetterToBelisarioVinta.html

Gotz, A., Petzold, A., Asmi, A., Blomberg, N., Lamanna, G., & Dekker, R. (2020). ESFRI cluster projects - Position papers on expectations and planned contributions to the EOSC. *Zenodo*. https://doi.org/10.5281/zenodo.3675081.

Greco, P. (2010). Il Sidereus Nuncius e l'origine della comunicazione pubblica della scienza. *S&F_scienzaefilosofia.it*. https://www.scienzaefilosofia.com/2018/03/26/il-sidereus-nuncius-e-lorigine-della-comunicazione-pubblica-della-scienza/

Guédon, J.-C. (2001). *In Oldenburg's long shadow*. Librarians, Research Scientists, Publishers, and the Control of Scientific Publishing, Association of Research Libraries, ISBN 0-918006-81-3.

Hadot, P. (1995). *Qu'est-ce que la philosophie antique?* Gallimard.

IETF. (2025). *About RFCs*. IETF. https://www.ietf.org/process/rfcs/

Langlais, P.-C. (2023). Open science infrastructure. *Petite encyclopédie de la science ouverte*. https://encyclo.ouvrirlascience.fr/articles/open-science-infrastructure/.

Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., Postel, J., Roberts, L. G., & Wolff, S. (2009). Brief history of the internet. *ACM SIGCOMM Computer Communication Review, 39*(5), 22–31.

Lessig, L. (2004). The creative commons. *Montana Law Review, 65*(1) https://scholarworks.umt.edu/mlr/vol65/iss1/1

Lévy, P. (1994). *L'intelligence collective*. La Découverte.

Lévy, P. (1997). *Cyberculture*. Rapport au Conseil de l'Europe. Editions Odile Jacob/Editions du Conseil de l'Europe.

Merton, R. K. (1973) [1942]. The normative structure of science, in Merton R. K., *The Sociology of Science: Theoretical and Empirical Investigations*, Chicago: University of Chicago Press, pp. 267–278, ISBN 978-0-226-52091-9, OCLC 755754. Originally published as "Science and Technology in a Democratic Order," Journal of Legal and Political Sociology 1 (1942): 115–26; later published as "Science and Democratic Social Structure," in Robert K. Merton, Social Theory and Social Structure.

Mons, B., Neylon, C., Velterop, J., Dumontier, M., da Silva Santos, L. O. B., & Wilkinson, M. D. (2017). Cloudy, increasingly FAIR; revisiting the FAIR data guiding principles for the European open science cloud. *Information Services and Use, 37*(1), 49–56. https://doi.org/10.3233/ISU-170824

MUR. (2022). *Piano nazionale per la scienza aperta (PNSA) 2021–2027*. https://www.mur.gov.it/it/atti-e-normativa/decreto-ministeriale-n-268-del-28-02-2022

Nelson, T. H. (1987). *Literary machines*. Mindful Press.

Neylon, C. (2013). Architecting the future of research communication: Building the models and analytics for an open access future. *PLoS Biology, 11*(10). https://doi.org/10.1371/journal.pbio.1001691

Nielsen, M. (2011). *Reinventing discovery: The new era of networked science*. Princeton University Press.

OECD. (2015). *Making open science a reality*. OECD Science, Technology and Industry Policy Papers, No. 25. OECD Publishing. https://doi.org/10.1787/5jrs2f963zs1-en

Plato. (2002). *Oxford world's classics* (Phaedrus, R. W. Ed.), Oxford University Press. ISBN: 9780199554027. https://doi.org/10.1093/actrade/9780199554027.book.1.

Plato. (2005). Meno and other dialogues: Charmides, Laches, Lysis, Meno, Robin Waterfield (ed.), Oxford University Press; *Oxford World's Classics*. ISBN: 9780199555666. https://doi.org/10.1093/actrade/9780199555666.book.1.

Rafols, I. (2025). *The relative marginalisation of SSH in research for societal challenges and the need for system level assessment to foster epistemic diversity*. RESSH 2025 book of Abstracts, pp. 66–69. https://vastuullinentiede.fi/sites/default/files/2025-05/RESSH2025_Conference_Book_of_Abstracts.pdf

Ramachandran, R., Bugbee, K., & Murphy, K. (2021). From open data to open science. *Earth and Space Science, 8*, e2020EA001562. https://doi.org/10.1029/2020EA001562

Raymond, E. S. (2001). The cathedral and the bazaar. In *The cathedral and the bazaar: Musings on Linux and open source by an accidental revolutionary*. O'Reilly Media.

Rossi, P. (2000). *La nascita della scienza moderna in Europa, Roma-Bari, Laterza.*

Shirky, C. (2010). *Cognitive surplus: Creativity and generosity in a connected age*. Penguin.

The Royal Society Science Policy Centre. (2012). *Science as an open enterprise*. https://royalsociety.org/-/media/policy/projects/sape/2012-06-20-saoe.pdf

UNESCO. (2021). *Recommendation on open science*. https://doi.org/10.54677/MNMH8546.

UNESCO. (2022). Open science toolkit. Building capacity for open science. https://doi.org/10.54677/DLYW1405

UNESCO. (2023). Development of the UNESCO recommendation on open science. https://www.unesco.org/en/open-science/development?hub=686

Wilkinson, M., Dumontier, M., Aalbersberg, I., et al. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data, 3*, 160018. https://doi.org/10.1038/sdata.2016.18