"Data stewardship in library

for sustainable and open science"

Damiano Orrù [1][0000-0002-7493-5648]

Librarian at Vilfredo Pareto Library
University of Rome "Tor Vergata"

orru@torvergataeconomia.it
http://orcid.org/0000-0002-7493-5648
https://it.linkedin.com/in/orrudamiano

Damiano is librarian at University of Rome Tor Vergata since 1999 as a computer technician, aimed at reducing difficulties and barriers to access to free and shared knowledge. He is fostering use of online tools for academic research and open to citizenship of every draw. He is involved in problem solving framework, working throught social / chat / blog / forum, for constructive and constant interaction with students and teachers to facilitate the use of the services offered by the Vilfredo Pareto Library, open to all citizens.

Damiano is librarian blogger since 2017 for blog BiblioVerifica, supporting citizens in digital reference [for free and on demand] and strategies supporting verification of data and facts through open contents (journals, books, data). This blog purpose to consistently offer tools and sources for fact checking in places at points in time that are relevant to the citizen, based on open science for data, books, articles, education resources freely usable for desired knowledge.

In memory of

MARIA GRAZIA GARGIOLI

(20 November 1946 – 18 April 2020)

JONATHAN TENNANT

(6 May 1988 – 09 April 2020)

A crowning achievement of the close twenty-year library collaboration with PAOLA COPPOLA, Director of the Library of economic area "*VILFREDO PARETO*" of the University of Rome *TOR VERGATA*:

- ✓ "*«#BiblioVerifica», il blog a supporto del cittadino fact-checker*"- year 2019 http://www.bibliotecheoggi.it/trends/article/view/1005
- ✓ *"How librarians can engage citizens to use open access contents and open data as source for fact-checking"- year 2019* https://informationliteracy.eu/conference/
- ✓ "*International Staff Week Open Science at UC3M*"- year 2019 https://www.yerun.eu/events/international-staff-week-open-science-at-uc3m/
- ✓ *"OPEN DATA – OPEN ACCESS: New Frontiers for Archives and Digital Platforms dedicated to the Performing Arts"- year 2019* http://opendataspa.uniroma2.it/
- ✓ *"Lo spazio social per la media literacy: #biblioVerifica vs #fakenews"- year 2018* https://www.convegnostelline.it/stelline2018/relatori1.php?IdUnivoco=84
- ✓ *"In-formare in biblioteca: fonti, strumenti e strategie sostenibili per smascherare le fake news"- year 2018* https://festivalsviluppososteni bile.it/2018/cal/525/in-formare-in-biblioteca-fonti-strumenti-e-strategie-sostenibili-per-smascherare-le-fake-news-#.XqsPDJ4zbIU
- ✓ *"Decennale della Dichiarazione di Messina: la via italiana all'Accesso Aperto"- year 2014* http://decennale.unime.it/?page_id=98
- ✓ *"Il futuro della comunicazione scientifica tra e-science e open access"- year 2011* https://economia.uniroma2.it/biblioteca/open-access-2011/

Abstract

The research analyses practices, tools, stakeholders, initiatives in support of open and sustainable science in the academic library.

The first part of the thesis frames practices and social benefits related to the library's citizen engagement in open access to academic scientific production, contributing to information literacy in the pursuit of the target 4.6: " By 2030, ensure that all youth and a substantial proportion of adults, both men and women, achieve literacy and numeracy" (SDG 4 Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all).

The heart of the work deals with data stewardship as a support for SA8000 ethical certification.

Subsequently, the paper deals with the topic of academic research data stewardship, evaluating the practices and tools in the library aimed at the target group 16.10: "ensure public access to information and protect fundamental freedoms, in accordance with national legislation and international agreements"

(SDG 16 Peace, justice and strong institutions).

KEYWORS: Open Science, data stewardship, open data, open access, citizen science, social accountability 8000, stewardship, data management plan

Index

Introduction _____ *4*

Chapter 1
Media, information and data literacy as an activity for the sustainable development in the library _____ *10*

Chapter 2
Citizens' access to open data and to the scientific publications _____ 22

Chapter 3
Portals and tools for citizen engagement via open data _____ 29

Chapter 4
Proposal for Data Stewardship as a support to the SA8000 ethical certification _____ 41

Chapter 5
Data stewardship in the public engagement library of researchers _____ 48

Chapter 6
Data stewardship for monitoring and verification of the SA8000 application _____ 56

Chapter 7
Date repository for open science _____ 63

Conclusions _____ 74

Bibliography _____ 78

Websites _____ 81

Introduction

Since 2018 IFLA (*International Federation of Library Associations and Institutions*)[1] promoted the global campaign in support of the SDGs (Sustainable Development Goals) of the Agenda 2030[2] in libraries.

Citizens are no longer simply users of institutional services, but can find a constructive and formative environment for a renewed process of change, based on Article 19 of the UN Declaration of Human Rights.:

> "*Everyone has the right to freedom of opinion and expression, including the right not to be harassed for his or her opinion, and the right to seek, receive and impart information and ideas through any media and frontier.*"[3]

Since 2014 by "Lyon Declaration" [4] The library world has reaffirmed and established that access to information is the basis for the sustainable development of society. A new idea of progress centred on economic prosperity and shared well-being is central to this vision. Libraries can help to increase administrative transparency and civic participation by training citizens in finding and evaluating reliable and freely accessible information. In this context, "universal literacy" (point 3 of the Declaration) based on social inclusion and technology as a driver for new forms of collaboration between science and citizens (point 5 of the Declaration) has started to be discussed.

Libraries supporting academic and specialist research can become a hub for open and participatory science in order to increase the visibility of scientific products and actively involve potential stakeholders from civil society and the production system.

[1] IFLA International Federation of Library Associations and Institutions https://www.ifla.org/publications/node/91777

[2] Sustainable Development Goals (Agenda 2030) https://sustainabledevelopment.un.org/?menu=1300

[3] Dichiarazione Universale dei Diritti Umani https://www.ohchr.org/en/udhr/pages/Language.aspx?LangID=itn

[4] Dichiarazione di Lione sull'Accesso all'Informazione e lo Sviluppo https://www.lyondeclaration.org/content/pages/lyon-declaration-it.pdf

The open science movement has its roots in the "*Bermuda Principles*"[5]: in 1996 the scientific community imposed rules to publicly share the results of research on the human genome.

Two years later the movement was born "*Open Source Initiative*"[6], from which derive today's myriad of open Apps and Software, including Linux/Unix. In these footsteps the scientific community founded the OA (open access) Initiative with the manifesto signed in Budapest in 2002[7], guidelines are disseminated to open up scientific content in medicine and health to the Internet community. The cornerstone of OA is to provide the public with free and unlimited access to academic research, especially if publicly funded, multiplying the possibilities for dissemination and scientific collaboration.

The following year, the "Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities"[8] is signed, where they define the cornerstones of OA:

> *"Open access contributions must meet two conditions: the author(s) and rights holder(s) of such contributions grant all users a free, irrevocable, worldwide right of access and a license to copy, use, distribute , transmit and display the work publicly and make and distribute derivative works, on any digital medium for any responsible purpose, subject to proper attribution of authorship (community standards will continue to provide the mechanism to enforce proper attribution and responsible use of published works, as they do now), as well as the right to make a limited number of printed copies for their personal use. A full version of the work and all additional materials, including a copy of the authorization as stated above, in an appropriate standard electronic format is deposited (and then published) in at least one online repository using appropriate technical standards (such as Open Archive definitions) that is supported and maintained by an academic institution, academic society, government agency, or other established organization that seeks to enable open access, unrestricted distribution, interoperability, and long-term archiving."*

[5] Bermuda Principles https://www.bermudaprinciples.org/
[6] Open Source Definition https://opensource.org/osd-annotated
[7] Budapest Open Access Initiative http://www.budapestopenaccessinitiative.org
[8] Berlin Declaration http://openaccess.mpg.de/Berlin-Declaration

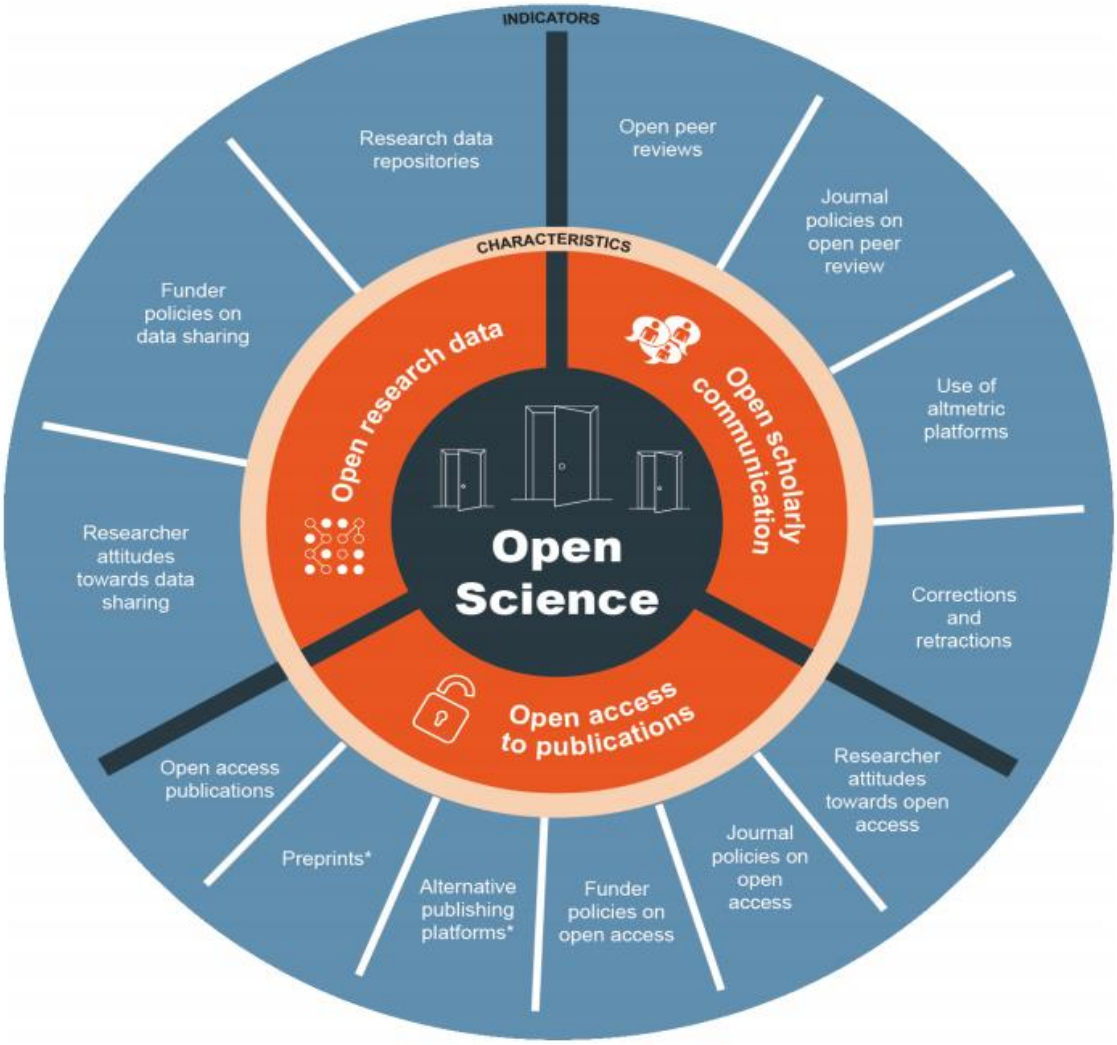
In Italy the OA movement affirms itself with the "Messina Declaration: Italian universities for open access to research literature"[9], on 5 November 2004 the signatory universities reaffirm the principles and objectives of the Berlin Declaration.

Since 2007, the European Commission has also embraced open access policies for funded research as a tool to support innovation and competitiveness. The "*mandatory policy*" enshrined in *Horizon 2020*[10] requires the opening of EU-funded publications from 2012, while the sharing of open data underlying research is necessary from 2020. This process of openness and sharing of knowledge and scientific production is based on levels of sharing that concern:

- Open access to scientific publications OA (*open access*)
- Open research data, *open data*
- Freely shared teaching materials, OER (*open educational resources*)
- Shared Bibliographies, *open bibliography / open metadata*
- *App* and open processing software, OSS (*open source software*)
- Open methodologies, ODM (*open development method*)
- Open training practices, OEP (*Open educational practices*)
- Open scientific review of publications, OPR (*open peer review*)
- Shared Innovation, *open innovation*
- Common open standards, *open standards*

The European Commission established OpenAIRE on 1 December 2009, with the aim of supporting OA in Europe through a Helpdesk service in the 27 EU countries, together with the technological infrastructure for the deposit and collection of open access publications.

[9] Dichiarazione di Messina http://cab.unime.it/decennale/wp-content/uploads/2014/03/Dich_MessinaITA.pdf
[10] Horizon2020 https://ec.europa.eu/programmes/horizon2020/en



SOURCE: Open Science Monitor 2015 [11]

Reaffirming this orientation, the European Council endorsed the transition to open science as a driver for innovation and competitiveness with conclusions "*The transition towards an Open Science system*" 27 may 2016[12].

Simultaneously LIBER (*Ligue des Bibliothèques Européennes de Recherche – Association of European Research Libraries*) has responded to the European call [13] deploying strategies in favour of *Citizen Science*[14], urging libraries to be proactive and open, as a neutral space

[11] Science "Wheel" - Open Science Monitor 2015 http://ec.europa.eu/research/openscience/index.cfm?pg=home&section=monitor

[12] "The transition towards an Open Science system" del 27 maggio 2016 http://data.consilium.europa.eu/doc/document/ST-9526-2016-INIT/en/pdf

[13] LIBER Response to the Amsterdam Call for Action http://libereurope.eu/blog/2016/05/17/liber-response-amsterdam-call-action

[14] Scienza in Rete: Citizen Science https://www.scienzainrete.it/articolo/citizen-science-scienza-di-tutti/valentina-meschia/2016-03-10

and traditional provider of access to knowledge in order to achieve an open scientific communication market coexisting with a sustainable publishing ecosystem.
In this framework the European platform was born OSPP (*Open Science Policy Platform*)[15] it disseminates the practices of Open Science based on eight "*pillars*":

- *Rewards and Incentives*
- *Research Indicators and Next-Generation Metrics*
- *Future of Scholarly Communication*
- *European Open Science Cloud*
- *FAIR (findability, accessibility, interoperability, and reusability)[16] data*
- *Research Integrity*
- *Skills and Education*
- *Citizen Science*

In recent years the European Union has launched EOSC (*European Open Science Cloud*)[17], the European scientific research data cloud, with the aim of sharing access to research data in the European scientific community through 1.7 million European researchers.
This work deals mainly with the ethical management of data stewardship of library research:

> *Data stewardship[18] in the scientific field is defined as "the set of activities to preserve and implement the information content, accessibility and usability of data. This research concept emphasises practices and functions aimed at creating appropriate supervision to ensure the quality and consistent use of the product and/or to provide value-added information."*

The figure of the steward is typically associated with caring for the needs of others:
the passengers of a ship, a plane or a train. In recent decades, in the private or public organisational sphere, Stewardship has been defined as the responsible management of goods or services shared and usable by others.

[15] OSPP https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-policy-platform
[16] FAIR Data https://www.force11.org/group/fairgroup/fairprinciples
[17] EOSC https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud
[18] Peng, G., Ritchey, N. A., Casey, K. S., Kearns, E. J., A Privette, J. L., Saunders, D., & Ansari, S. (2016). Scientific Stewardship in the Open Data and Big Data Era-Roles and Responsibilities of Stewards and Other Major Product Stakeholders

Today's market pays more and more attention to corporate social responsibility/entity, whose certification becomes a strong competitive advantage, in order to increase the trust of customers/users, financiers, local control bodies and stakeholders in general.

The Standard SA (*Social Accountability*) 8000 was first published in 1997 by *Council on Economic Priorities Accreditation Agency* (CEPAA)[19], today SAI (*Social Accountability International*)[20], supervised by SAAS (*Social Accountability Accreditation Services*)[21]. The requirements have been regularly updated: in 2001, 2008 and 2014 (whose transition ended on 30 June 2017).

This is not a legal obligation or a multinational lobby, but a voluntary and verifiable vignette, applicable on a voluntary and verifiable basis, to certify social and sustainable performance starting from the organization of the research center, research products, researchers and technicians, relations with all academic stakeholders.

[19] CEPAA (Council of Economic Priorities Accreditation Agency) US institute founded in 1969 to provide investors and consumers with information tools to analyze the social performance of companies. The primary goal is to empower organizations to be socially responsible by bringing together key stakeholders to develop voluntary consensus-based standards, accrediting qualified organizations to verify compliance, promoting knowledge and understanding of the standard, and encouraging its implementation worldwide.

[20] Social Accountability International (SAI) http://www.sa-intl.org/

[21] Social Accountability Accreditation Services (SAAS) http://www.saasaccreditation.org/

CHAPTER I

Media, information and data literacy as activities for sustainable development in the library

Libraries, through their literacy services to search and analyze information and data, contribute directly to the achievement of the targets:

- 4.6 By 2030, ensure that all youth and a substantial proportion of adults, both men and women, achieve literacy and numeracy
  (SDG 4 Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all)
- 10.2 by 2030 empower and promote the social, economic and political inclusion of all irrespective of age, sex, disability, race, ethnicity, origin, religion or economic or other status
  (SDG 10: Reduced inequalities)
- 16.6 Develop effective, accountable and transparent institutions at all levels
- 16.10 ensure public access to information and protect fundamental freedoms, in accordance with national legislation and international agreements"
  (SDG 16 Peace, justice and strong institutions)

Education and lifelong learning are the golden thread running through the implementation of all 17 SDGs. Expectations of education are very high: it is a prerequisite for poverty and also a prerequisite for rewarding employment and decent jobs.
It is also crucial for sustainable growth, building social cohesion, achieving prosperity and promoting human rights and equality.

In addition, today the sharing and openness of science and data are acting as a deterrent to the North-South digital divide in the world, expanding the use of information and communication technologies, through online collaboration, with insignificant Internet connection costs. The basis of shared knowledge is access to information, knowledge and certified data. In recent years there has been much discussion about reliable data and sources for a rational decision-making process, aimed at expanding statistical capabilities and the use of data by "ordinary people", such as food, health, civic, social and environmental data.

SOURCE: IFLA "This Library Supports the SDGs"

On the other hand, citizens as taxpayers of the state coffers are the first funders of Italian and European research: open science allows them to be materially repaid for the research they finance.

The Italian Constitution itself encourages scientific research to be free from restrictions in order to open up to society through training in the cultural growth of the primary stakeholder, the citizen:

- "Art and science are free and free is its teaching" (Article 9)
- "The Republic promotes the development of culture and scientific and technical research. It protects the landscape and the historical and artistic heritage of the Nation" (Article 33)

The citizen today is a user of open science conceived as an approach aimed at making research results freely available, to reach people outside the academic world.[22] The main objective is the accessibility of contributions (monographs, papers, articles, reports, historical series, etc.), maximizing transparency and ensuring the reproducibility of research. At the basis of everything there are shared data and open research tools, followed by open peer review and open access to the final publication.

Libraries are a free and open training point, where every citizen can find tools and directions to access shared knowledge, through the reference service.

Today the diffusion of news through traditional media is outclassed by the spread and speed of social media, with what is called "Truth Decay"[23], result of 4 evolutions:

- *growing disagreement over objective facts and analytical interpretations of data;*
- *a blurring of the line between facts and opinions;*
- *a growing relative volume and the resulting influence of opinion over facts;*
- *a decrease in confidence in key sources of information that were once considered to be factual sources of information, such as government and the media.*

Libraries respond to these information challenges with Media and Information Literacy and Data Literacy strategies.

IFLA launched in 2017 the infographics "*How to Spot Fake News*"[24] to push library users to use critical thinking, especially to counter misinformation with eight simple steps to verify news and data, overcoming personal and social prejudices.

SOURCE: IFLA - How To Spot Fake News

[22] What is Open Science? Open Science Community Utrecht https://openscience-utrecht.com/what-is-open-science/
[23] Kavanagh J. and Michael D. Rich, Truth Decay: An Initial Exploration of the Diminishing Role of Facts and Analysis in American Public Life, Santa Monica, Calif.: RAND Corporation, RR-2324-RC, 2018.
[24] How To Spot Fake News – IFLA https://www.ifla.org/publications/node/11174

The aim of this information education is the development of a lifelong learning of the citizen, evolving from a taxable person (user) to an active player, producer and megaphone of reliable information: the *prosumer*.[25]

Since 2003 l'UNESCO (*United Nations Educational, Scientific and Cultural Organization*) launched the principles of information literacy in the "Prague Declaration, towards an Information Society"[26]:

- *information literacy includes the knowledge of one's own concerns and needs, the ability to identify, locate, assess, organise and create, use and effectively use data and news to address current issues or problems*
- *information literacy is a prerequisite for effective participation in the Information Society, and is part of*
- *the fundamental human right to lifelong learning*
- *information literacy, together with access to essential information and the effective use of information and communication technologies, plays a major role in reducing barriers and injustices locally and between countries or peoples*
- *information literacy is fundamental in promoting tolerance and mutual understanding through the use of information in multicultural and multilingual contexts*
- *Information literacy should be an integral part of Education for All, to contribute critically to the achievement of the UN Millennium Development Goals and respect for the Universal Declaration of Human Rights.*

[25] Lugli G., Una panoramica sulla biblioteca pubblica d'oggi, Temperino Rosso, 2019.
[26] Dichiarazione di Praga UNESCO
http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/PragueDeclaration.pdf

L'AIB (associazione professionale dei bibliotecari italiani) promoted in 2014 the "Manifesto for *Information Literacy*"[27] for the creation of an open, inclusive and welcoming community of information specialists. The foundations of this project are the notions of:

- *Media Information Literacy*[28] *understood as "media and information competence consists of knowledge, attitudes, the set of skills needed to recognize when and what kind of information is needed; where and how to obtain that information; how to critically evaluate and organize it once found; and how to use it in an ethical way".*
- *Information literacy*[29] *defined as "the set of skills, competences, knowledge and attitudes that lead the individual to mature over time, throughout life, a complex and diversified relationship with the information sources: the documents and information they contain. These sources must be understood independently of the medium through which the information is conveyed. The ultimate goal of their use must be the creation of new knowledge for oneself and others, acting critically with respect to information. In summary, information competence includes the ability to recognize an information need, search, evaluate, use information in a conscious way to create new knowledge".*

The information is evaluated as an asset, produced and exchanged between individuals and private and public institutions, in the form of textual/graphical/multimedia data or documents. In particular, the theme of data focuses on the concept of "Granularity of information", extrapolated and available in separate contexts, complicating the possibility of reconstruction from the original content.

In order to face these criticalities, the search for open and freely accessible sources is suggested, critically evaluating the information, starting from the authoritativeness of authors and sources to arrive at fact-checking practices.

[27] Manifesto per Information Literacy AIB https://www.aib.it/struttura/commissioni-e-gruppi/gruppo-literacy/ilmanifesto/
[28] IFLA Media and Information Literacy Recommendations 2011 http://www.ifla.org/publications/ifla-media-and-information-literacy-recommendations?og=81
[29] AGID Agenzia per l'Italia digitale. Programma nazionale per la cultura, la formazione e le competenze digitali, LINEE GUIDA Indicazioni strategiche e operative 2014 https://www.agid.gov.it/sites/default/files/repository_files/documenti_indirizzo/programma_nazionale_cultura_formazione_competenze_digitali_-_linee_guida_indicazioni_strategiche_operative_0.pdf

The network consultation can be conceived as a "structured survey" based on data and news retrieval, verification strategies, source selection tools.
In the society of social networks and ultra-fast connections, the time needed to analyse and understand data and its use is reduced. The maturation of skills for access and analysis of the digital environment becomes indispensable.

Data literacy is the "ability to read, process, analyze and discuss data. Just like literacy as a general concept, data literacy focuses on the skills needed to work with data".[30]
In this vision the citizen participates as "metaliterate"[31], in a constant interaction with the librarians and the network, completing the personal information expertise, in four areas:

- *Behavioural,what students should be able to do upon successful completion of learning activities—skills, competencies*
- *Cognitive, what students should know upon successful completion of learning activities—comprehension, organization, application, evaluation*
- *Affective, changes in learners' emotions or attitudes through engagement with learning activities*
- *Metacognitive, what learners think about their own thinking—a reflective understanding of how and why they learn, what they do and do not know, their preconceptions, and how to continue to learn*

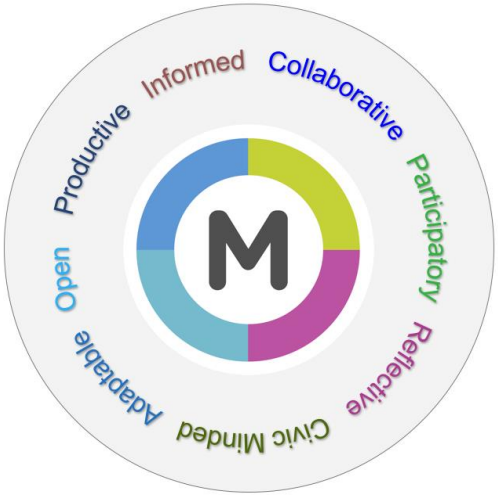
This theory defined in the blog "*Metaliteracy*" provides in the process for the citizen different learning objectives with the aim of:

- *actively evaluate the contents, also evaluating one's own prejudices*
- *take action in defence of intellectual property in an ethical and responsible manner*

[30] 'Data Literacy', Wikipedia https://en.wikipedia.org/wiki/Data_literacy
[31] Metaliteracy Goals and Learning Objectives https://metaliteracy.org/learning-objectives/goals-and-learning-objectives-translated/

- produce and share information in collaborative and participatory environments

- *develop learning strategies to achieve personal and professional lifelong goals*



SOURCE: Mackey & Jacobson, Metaliterate Learner Characteristics, 2018

Several universities and training centres have developed open repositories and online self-learning environments, based on interactive educational materials (OER, Open Educational Resources), oriented towards reuse and free distribution to citizens. One of the most widely used forms today is MOOC (Massive Open Online Courses).

An example of these self-learning tools is the course run by Thomas P. Mackey and Trudi E. Jacobson "Empowering Yourself in a Post-Truth World" [32], usable from the free Coursera platform, through short text and video content, self-assessment quizzes integrated by a discussion forum with tutors.

For data management you can use the self-learning platform MANTRA[33], managed by the Research Data Service in Information Services of the University of Edinburgh[34] to take advantage of several interactive online units that allow the acquisition of terminology, key concepts and practices to manage research data.

[32] MOOC "Empowering Yourself in a Post-Truth World" https://www.coursera.org/learn/empowering-yourself-post-truth-world gestito da Tom Mackey, Professore al dipartimento " Arts and Media" presso University of New York - Trudi E. Jacobson, Bibliotecaria responsabile del "Information Literacy Department at the University Libraries" presso University of New York

[33] Blog MANTRA https://mantra.edina.ac.uk/

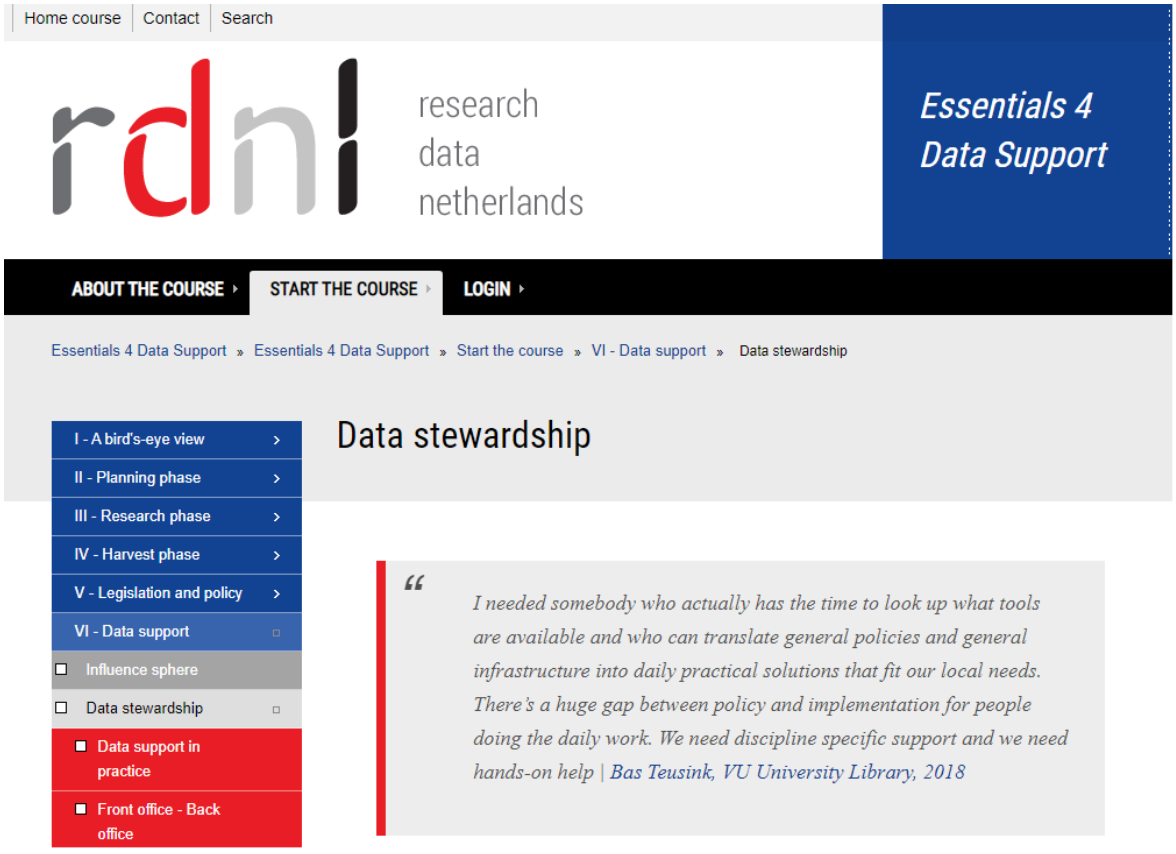
[34] Edinburgh Research Data http://datablog.is.ed.ac.uk/

https://mantra.edina.ac.uk/

You can also learn the first basics for software environments such as R, SPSS, NVivo or ArcGIS. The platform also integrates external readings, online exercises, examples, video clips, tutorials on data management. To obtain the certificate anyone can attend the MOOC on Coursera "Management and sharing of research data"[35], in collaboration with the University of North Carolina-Chapel Hill. In addition, a training kit is being disseminated for information professionals to train students.[36]

A similar and much more technical course is "*Essentials 4 Data Support*" [37] aimed at PhD students and researchers looking for solutions to store, manage, preserve and share research data.

[35] Research Data Management and Sharing https://www.coursera.org/learn/data-management
[36] Mantra KIT https://mantra.edina.ac.uk/libtraining.html
[37] Essentials 4 Data Support https://datasupport.researchdata.nl/en/



https://datasupport.researchdata.nl/en/

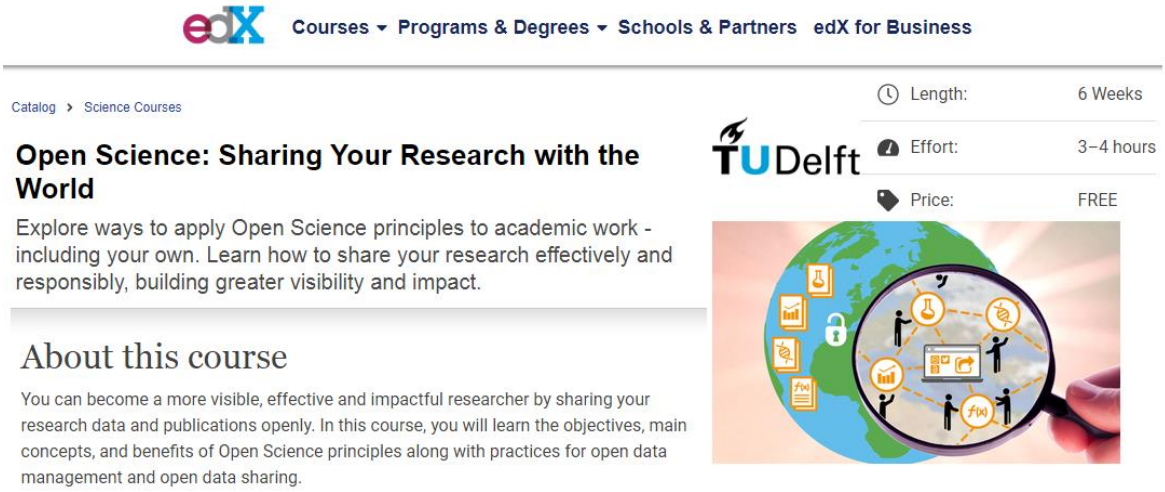
PhD students and researchers are offered the possibility to become "*Eurodoc Open Science Ambassador Training*"[38] through 5 modules structured in webinar on the platform managed by the European Council of PhD students and young researchers (international federation of 28 national organisations in 26 EU countries.

http://eurodoc.net/ambassadors

This course designed by Gareth O'Neill and Ivo Grigorov aims to train researchers in the key practices of Open Science, available free of charge to all interested researchers and policy makers.

[38] Eurodoc Open Science Ambassador Training http://eurodoc.net/ambassadors

A third very complete MOOC is the free course "*Open Science: Sharing Your Research with the World*"[39], managed on the edX platform by the library TU Delft[40] within the project "4TU.ResearchData"[41], aimed at disseminating open data management and sharing practices through FAIR practices, implementing links between publications, data, software and methods.



https://www.edx.org/course/open-science-sharing-your-research-with-the-world

The best known of the MOOCs on open science is "*Open Science MOOC*"[42]. This course is designed to help students and researchers develop skills for participatory and shared research. The MOOC consists of 10 modules, consisting of videos, research articles, fictional data sets and code, as well as individual and group tutorials.

It is a platform designed and founded by Jonathan Tennant[43], paleontologist and independent researcher at *Institute for Globally Distributed Open Research and Education*, died prematurely on 9 April 2020.

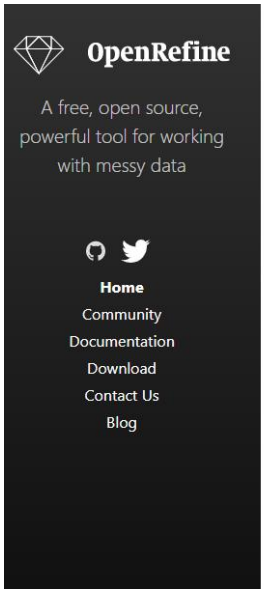
https://opensciencemooc.eu/

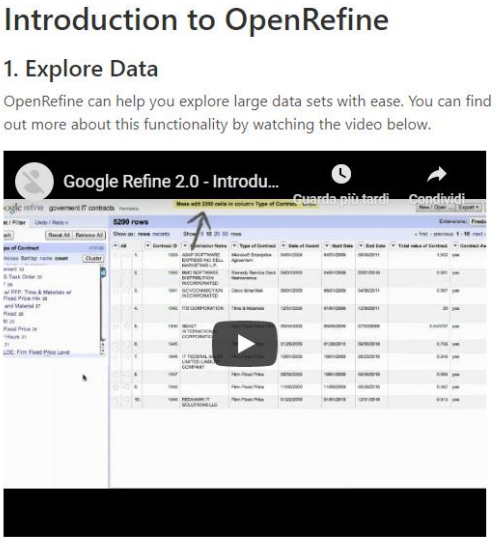
[39] Open Science: Sharing Your Research with the World https://www.edx.org/course/open-science-sharing-your-research-with-the-world
[40] biblioteca TU Delft https://www.tudelft.nl/library
[41] 4TU.ResearchData https://researchdata.4tu.nl/en/about-4turesearchdata/organisation
[42] Open Science MOOC https://opensciencemooc.eu/
[43] Jon Tennant e l'insostenibile virus della scienza chiusa http://biblioverifica.altervista.org/jon-tennant/

For advanced training in data management targeted at librarians, PhD students and researchers the community is active *Data Carpentry*[44] fruit of initiatives *Library Carpentry*[45] and *Software Carpentry*[46], training spaces to expand IT skills aimed at data management, cleaning and processing, with exercises that can be replicated independently by the individual user.

https://software-carpentry.org/lessons/

Data Carpentry does not require computational or statistical experience, while the exercises are targeted at specific areas of research to be quickly applied by departments or individual researchers developing experiments with standard protocols. The training is based on CC-BY contents usable through the web that are displayed in workshops through laboratories on data management with step-by-step simulations in the use of open source applications such as: R software[47], RStudio[48], Open Refine[49].





http://openrefine.org/

[44] Data Carpentry https://datacarpentry.org/
[45] Library Carpentry https://librarycarpentry.org/blog/
[46] Software Carpentry https://software-carpentry.org/
[47] R software https://cran.r-project.org/
[48] RStudio https://www.rstudio.com/products/rstudio/
[49] Open Refine http://openrefine.org/

There is also a form of "direct on-demand support" of librarians based on the specific research of the citizen: the reference in the library or through the network, such as *Ask at Librarian*. [50] It is an interactive collaborative space between users and librarians, in which every citizen proposes research questions by e-mail, social, chat, web forms.

https://twitter.com/Ask_Librarian - https://fb.com/asklibrarian

The librarian provides assistance, as a specialist in online research, not only bibliographical research, establishing a direct, immediate and focused conversation on the specific need.

A similar initiative is developed at the Library of economic area "Vilfredo Pareto" of the University of Rome "Tor Vergata", it is the blog "B*iblioVerifica*"[51], a social space of interaction between citizens and Volunteer Librarians, aimed at promoting "the fact-checking and use of open data, certified sources, free search tools, search strategies to facilitate access to sustainable information in terms of accuracy, traceability, independence, legality, impartiality".[52]

http://www.biblioverifica.cloud

[50] *Ask at Librarian* https://twitter.com/Ask_Librarian https://www.facebook.com/asklibrarian
[51] Blog BiblioVerifica http://www.biblioverifica.cloud/
[52] Orru, D., & Coppola, P. (2019). «# BiblioVerifica», il blog a supporto del cittadino fact-checker. Biblioteche oggi Trends.

CHAPTER II

Citizen access to open data and scientific publications

The verification of the facts is based on reliable sources, transparent institutions but above all on the availability of open data[53] "*which may be freely used, reused and redistributed by anyone, subject to the possible need to cite their source and share them with the same type of licence under which they were originally issued:*

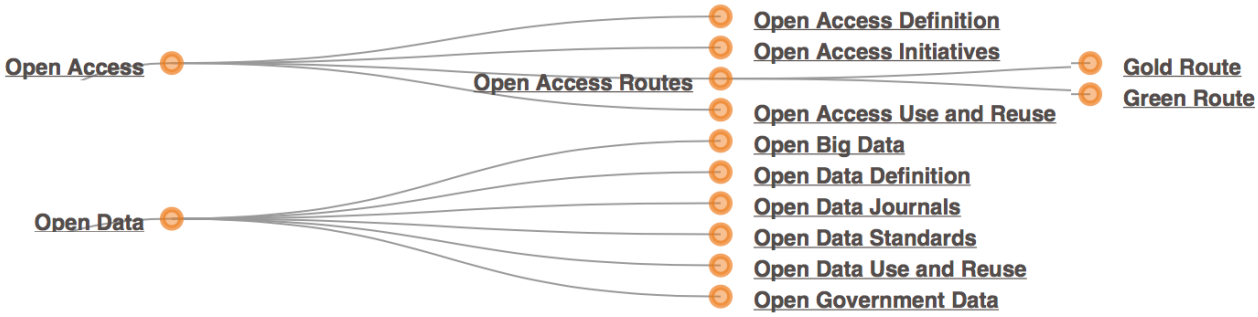
- *Availability and access: the data must be available as a whole, for a price not exceeding a reasonable cost of reproduction, preferably by downloading from the Internet. The data must be available in a useful and editable format.*
- *Re-use and redistribution: data must be provided under conditions allowing reuse and redistribution. This includes the possibility to combine them with other databases.*
- *Universal participation: Everyone must be able to use, reuse and redistribute data. There must be no discrimination either on initiative or against individuals or groups. For example, the 'non-commercial' clause, which prohibits use for commercial purposes or restricts use for certain purposes only (e.g. education) is not allowed."*

The citizen can access different types of data[54] classifiable in:

- *"Observation data is acquired in real time, usually unique and irreplaceable. (For example, brain images, survey data.)*
- *Experimental data comes from laboratory equipment, often reproducible, but can be expensive. (For example, chromatograms, microassays.)*
- *Simulation or model data are generated from test models where the model and metadata may be more important than the model output data. (For example, economic or climate models.)*

[53] Open Data Handbook http://opendatahandbook.org/guide/it/
[54] FONTE Active GO FAIR Implementation Network https://www.go-fair.org/implementation-networks/overview/dscc/

- *Data derived or compiled result from the processing or combination of "raw" data, often reproducible but expensive. (For example, compiled databases, text extraction.).*
- *Reference or canonical data are a conglomerate (static or organic) or a collection of smaller data sets (peer reviewed), most likely published and edited. (For example, genetic databases, crystallographic databases.)"*



SOURCE: FOSTER Open Science Taxonomy [55]

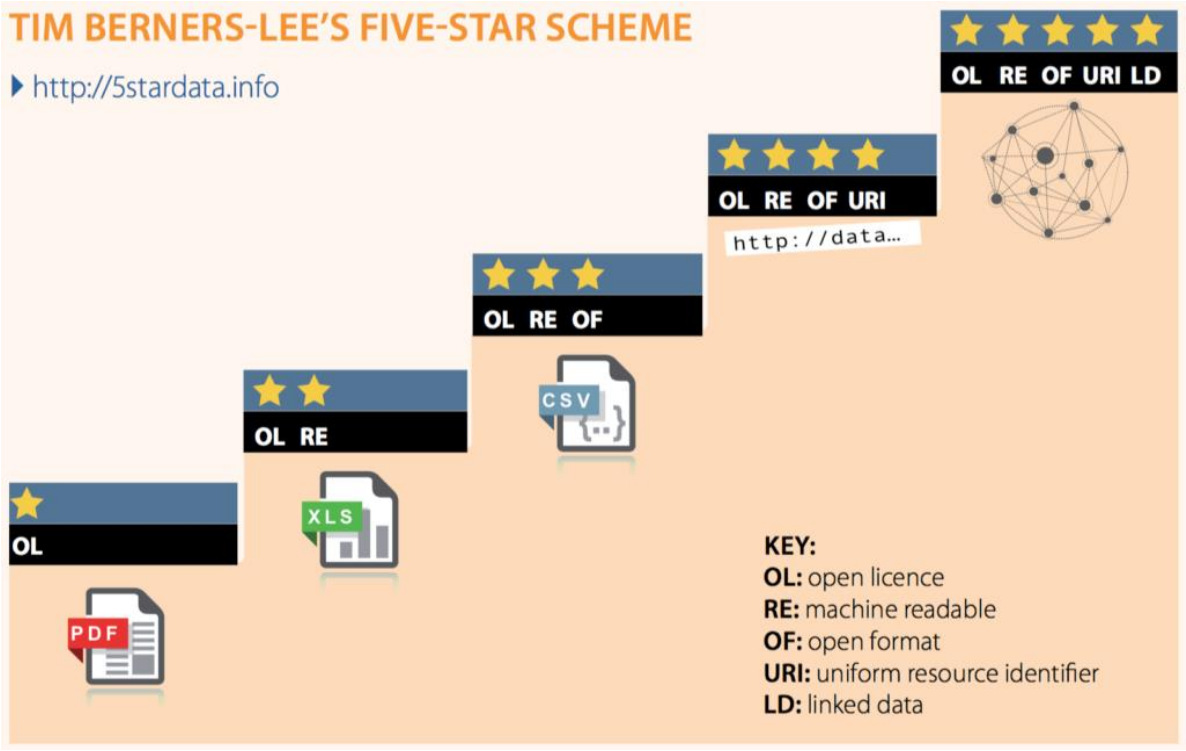
In this information flow, the role of the data steward, responsible for supervising the life cycle of open data, becomes fundamental, verifying compliance with the following FAIR[56] (*findability*, *accessibility*, *interoperability*, *and reusability*) principles.

It is also possible to recognize the degree of openness of data found online using the classification designed by Tim Berners-Lee[57], inventor of the http protocol of the World Wide Web, promoter of Linked Data.

The degree of openness of the data is measured in 5 levels identifiable on an increasing scale (as for hotels) with different benefits structured by the Open Data Institute.[58]

[55] FOSTER Open Science Taxonomy https://www.fosteropenscience.eu/taxonomy/term/7?page=11
[56] Mons, Barend. data stewardship for open science: Implementing FAIR principles. Chapman and Hall/CRC, 2018
[57] 5 star data https://5stardata.info/en/
[58] Open Data Institute http://theodi.org/

The data, which are put online in any format, but with open license[59], get a star of approval: as a citizen you can consult, print, store and share the data.



To get the second star requires the data to be structured (as in Microsoft Excel folders or other closed proprietary software), so it can be processed / aggregated / exported easily.
The third star can be obtained by making the CSV (Comma-Separated Values) open non-proprietary format available, so any manipulation is available without having to pay for software packages.
With the URI (Uniform Resource Identifier) format, a sequence of characters that universally and univocally identifies a resource, the data is open with 4 evaluation stars, allowing the citizen to use it from any web-connected environment, including the storage of the link or favorite bookmarks, as well as partial reuse.
An example of this encoding is RDF (Resource Description Framework) which is the basis of interoperability between applications, allowing the transition to the maximum level.

[59] Licenza CC0 https://creativecommons.org/publicdomain/zero/1.0/deed.it

The maximum rating of 5 stars is attributed to those data that are connected with other external data sets, allowing the simultaneous use of joint related data, also sharing the structured scheme.

Citizens have access to a vast repository of open data, which can be understood and re-used if the publications using or re-using the research data are available. The openness of scientific publications is crucial for this.

For the consultation of open access journals, the DOAJ (Directory of Open Access Journals) portal, created in 2003 by Lund University, is available as an online repository of high-quality, peer-reviewed, open access journals.

https://doaj.org/

"DOAJ's mission is to increase the visibility, accessibility, reputation, use and impact of quality, peer-reviewed, open-access academic research journals globally, regardless of discipline, geography or language. DOAJ will work with publishers, editors and journal owners to help them understand the value of publishing and best practice standards and apply them to their operations. DOAJ is committed to being 100% independent and to keeping all of its services and metadata free to use or reuse for everyone."[60]

[60] DOAJ https://doaj.org/about

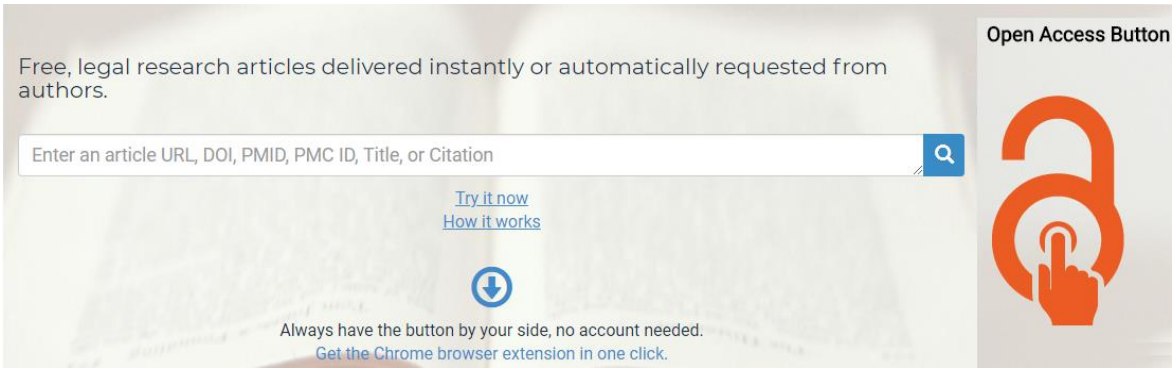
For open access to scientific e-books with peer review, DOAB (Directory of Open Access Books) is available, with the possibility to freely download chapters or full text monographs, through open access academic publishers. Among the disciplines most present are the humanities, legal and social sciences. The portal was created by OAPEN (Open Access Publishing in European Networks), currently managed with OpenEdition, CNRS (Centre national de la recherche scientifique) and Aix-Marseille Université.

https://www.doabooks.org/

> *" The main objective of DOAB is to increase the detectability of Open Access books. Academic publishers are invited to provide DOAB with the metadata of their Open Access books. The metadata will be collected to maximize dissemination, visibility and impact. Aggregators can integrate records into their commercial services and libraries can integrate the directory into their online catalogues, helping scholars and students discover books. The directory is open to all publishers who publish peer-reviewed academic books in Open Access and must contain as many books as possible, provided that such publications are in Open Access and meet academic standards. "* [61]

If, on the other hand, the citizen already knows the title or DOI (*Digital Object Identifier*)[62] of a publication can check the availability of the full content in open access with the free service "Open Access Button", an initiative supported by SPARC (Scholarly Publishing and Academic Resources Coalition)[63] global network of libraries for open access to research and training.

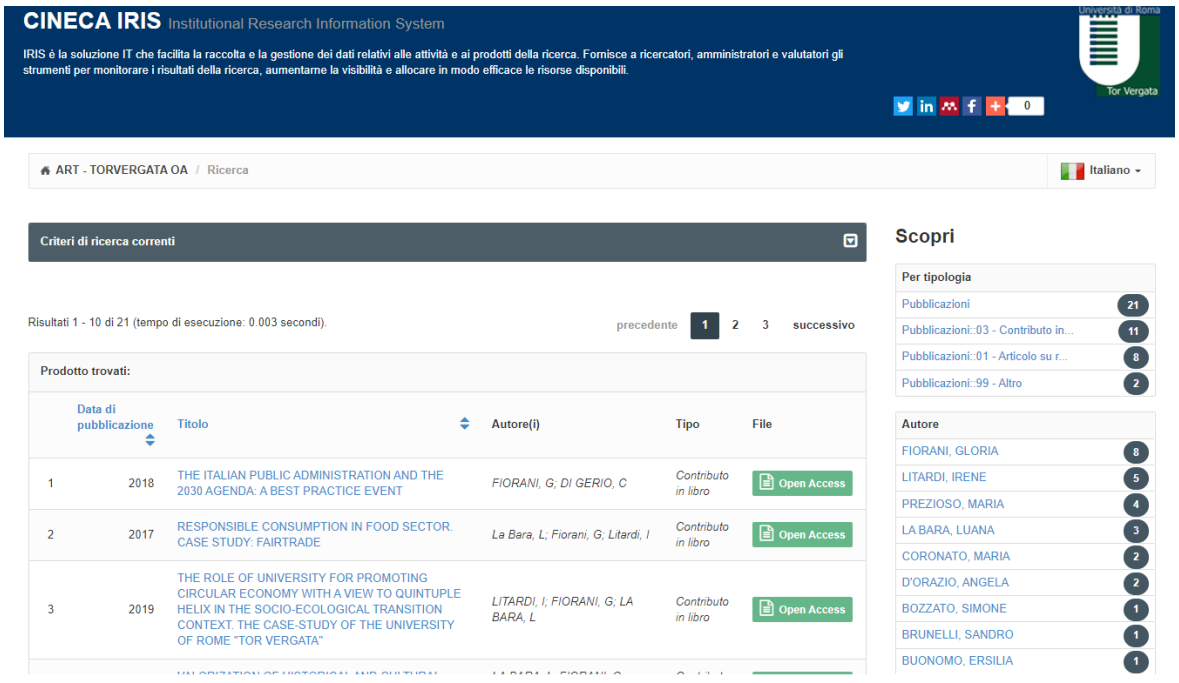
[61] DOAB https://www.doabooks.org/doab?func=about&uiLanguage=en
[62] DOI - Digital Object Identifier https://www.doi.org/
[63] SPARC - Scholarly Publishing and Academic Resources Coalition https://www.sparcopen.org/



https://openaccessbutton.org/

Italian universities in recent decades have traced academic research through the Institutional Research Archives, where part of the scientific publications are disseminated with open access to the entire Internet community.

An example is the Anagrafe della Ricerca of the University of Rome Tor Vergata.[64]



https://art.torvergata.it/

[64] Anagrafe della Ricerca dell'Università Tor Vergata https://art.torvergata.it/

Beyond these platforms for the protection of intellectual property there are pirate projects, which go beyond international copyright rules, the best known case is Sci-Hub [65].

https://twitter.com/sci_hub

The supporters of this boundless archive of publications justify their work with what is stated in the first sentence of Article 27 of the Universal Declaration of Human Rights [66]:

> *" Everyone has the right to participate freely in the cultural life of the community, to enjoy the arts and to share in scientific progress and its benefits."*

This philosophy, however, ignores what is stated in the second sentence of the same article:

> *" Everyone has the right to the protection of the moral and material interests arising from any scientific, literary and artistic production of which he is the author."*

[65] Sci-Hub https://twitter.com/sci_hub
[66] Dichiarazione universale dei diritti umani
http://www.senato.it/documenti/repository/relazioni/libreria/fascicolo_diritti_umani.pdf

CHAPTER III

Portals and tools for citizen engagement through open data

Research institutions, like the whole public administration, are at the service of citizens and businesses that need open data, in compliance with the European GDPR (General Data Protection Regulation).

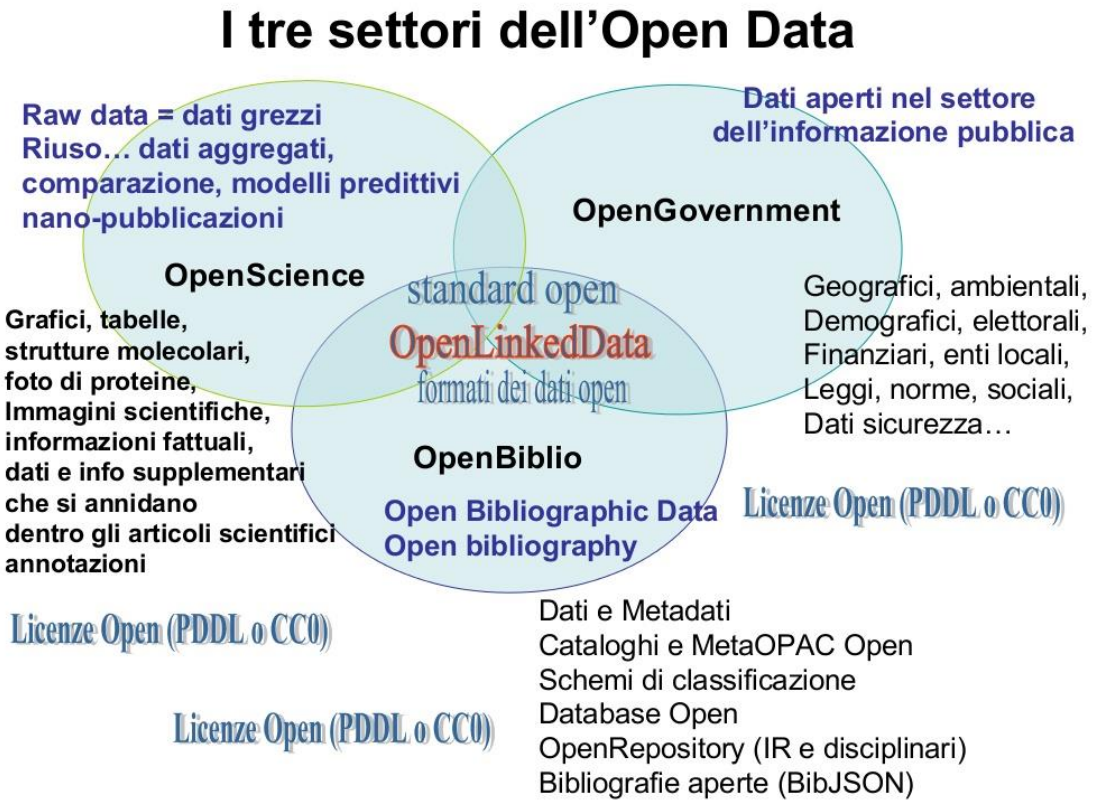
The services of the future impose new standards and interoperability procedures on the institutions to improve existing services, but above all to develop integrations with the productive world. The aim is to create the "Digital Society", in which citizens and businesses interact with the Public Administration for integrated, advanced and timely services. For these purposes, entities are opening up to agreements aimed at "data sharing", for application cooperation solutions in data access with a selected or indistinct audience.
According to the OECD "open science includes unhindered access to scientific articles, access to public research data and collaborative research made possible by ICT (Information and Communication Technology) tools and incentives. The sustainability of access to publicly funded research data implies an administrative responsibility of the research organisation to ensure permanent access to data in long-term preservation.
Regardless of project duration, data retention and perpetual access are a prerequisite for:

- strengthen the open scientific investigation;
- encourage diversity of analysis and opinion;
- promote new research;
- make it possible to test new or alternative hypotheses and methods of analysis;
- support studies on data collection and measurement methods;
- facilitate the training of new researchers;
- allow the exploration of topics not foreseen by the original researchers;

- allow the creation of new data sets when data from multiple sources are combined." [67]

At the beginning of the research projects, the types of data that will be acquired and the potential results of the experiments, the methodology of the experimentation management, the descriptive metadata, the techniques of analysis and conservation, the limitations in sharing and reuse are defined: the data management plan.

In order to make these data sets well structured and indexed, the figures of Data Scientist and data stewards become indispensable to facilitate the active participation of society and to develop "intelligent" applications, promoting scientific research. Today Public Administrations have the technology to optimize the diffusion of Open Data, in data access and reuse.



SOURCE: Antonella De Robbio, "Open Data?" giornata di studio "Il futuro della comunicazione scientifica tra e-science e open access" 27 ottobre 2011, Università degli Studi di Roma Tor Vergata[68]

[67] OECD *Principles and Guidelines for Access to Research Data from Public Funding* http://www.oecd.org/sti/inno/38500813.pdf

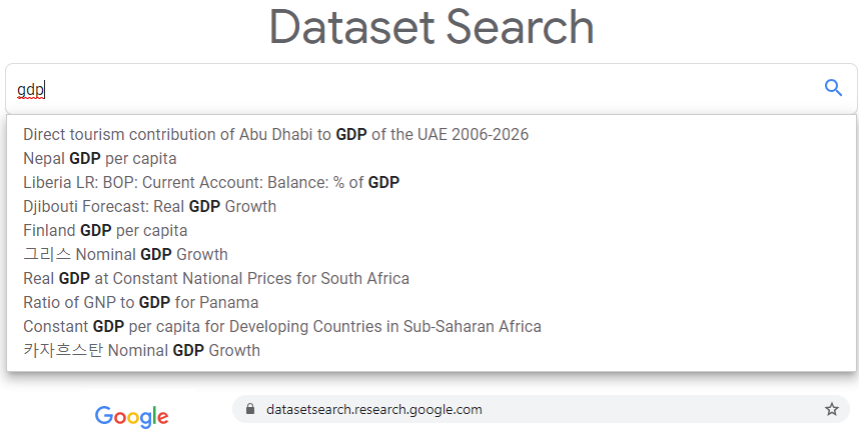
[68] Antonella De Robbio, "Open Data?" giornata di studio "Il futuro della comunicazione scientifica tra e-science e open access", 27 ottobre 2011, Università degli Studi di Roma Tor Vergata, Open Access Week 2011 https://www.slideshare.net/derobbio/il-futuro-della-comunicazione-scientifica-tra-escience-e-open-access

AGID (Agenzia per l'Italia digitale) has put in place since 2017 twenty actions[69], partly aimed at inclusive and sustainable development, interoperability and exchange of public data with stakeholders. The initiative aims to significantly increase the value of the Public Administration's information assets, facilitating analysis and knowledge tools for decision makers, interacting with citizens and businesses.

An Italian case of application of open data is the PDND (National Digital Data Platform managed by AGID, Agency for Digital Italy):

> " a single interface, within which individual administrations can communicate and share data and APIs in a free and open manner, allowing the emergence of new and previously unthinkable services and data applications tailored to the needs of the citizen." [70]

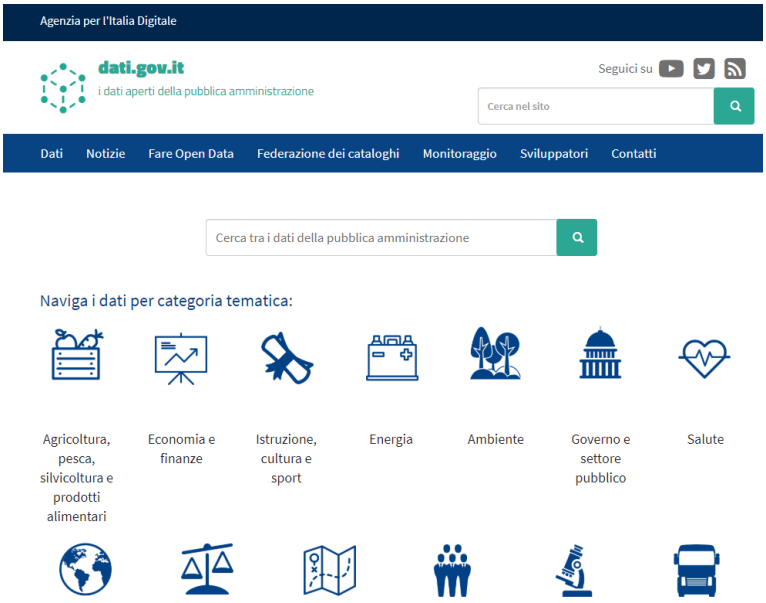
The basis for data availability is interoperability and indexing via specific tags in the Schema.org[71] structured vocabulary, compatible with IoT (Internet of Things) applications, pushing towards metadata standardization, which in the future may support AI (Artificial Intelligence) algorithms. Today, it is not necessary to be a specialist, journalist or researcher to consult, process and combine data sets. The Internet offers free search tools such as Google Dataset Search[72], allowing the download of open and shared historical series by government agencies, scientific publishers, research institutes and individual researchers through institutional repositories or thematic portals.



[69] Strategia digitale 2025, IA e Open data: come vincere l'inerzia dell'Italia https://www.agendadigitale.eu/cittadinanza-digitale/strategia-digitale-2025-ia-e-open-data-come-vincere-linerzia-dellitalia/
[70] PDND (Piattaforma Digitale Nazionale Dati) https://innovazione.gov.it/it/progetti/pdnd/
[71] IoT and Schema.org https://schema.org/docs/iot-gettingstarted.html
[72] Google Scholar and Dataset Search https://www.nature.com/articles/d41586-018-06201-x?fbclid=IwAR0BmCQ1q4F_vQTXiElzmqMCjFNzfNo5ngHApV_Tpq8MsjK9ISFZL5HEzXc

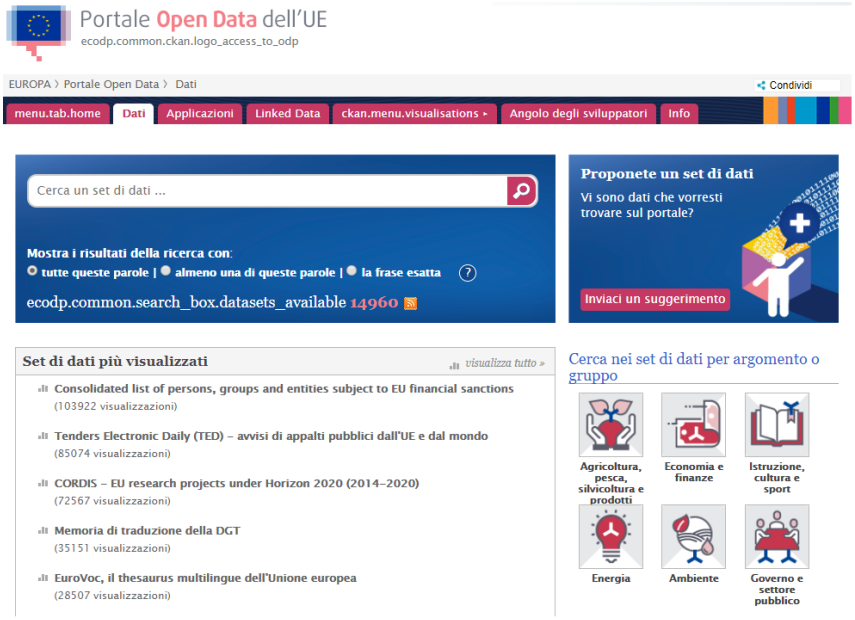
The DATI.GOV project is managed and developed by AgID (Agenzia per l'Italia Digitale) as part of the Open Government strategies with the aim of disseminating the culture of open data, increasing the transparency of the PA, facilitating businesses and citizens through the reuse of open data for value-added services/applications.

Technically the metadata are managed according to the standard DCAT-AP_IT [73], ensuring interoperability as defined at international and European level.



https://www.dati.gov.it

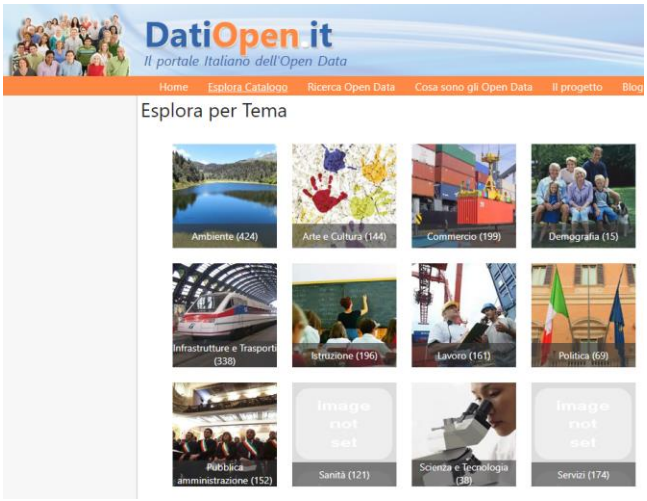
The EU Open Data portal was set up in 2012 to share data from countries and European institutions in compliance with national and European regulations, so that it can be reused free of charge by the Internet community. The data are mainly geographical, financial, statistical, electoral, legal, health, environmental, electoral, statistical and legal.



https://data.europa.eu

[73] Metadati DCAT-AP_IT https://www.dati.gov.it/content/dcat-ap-it-v10-profilo-italiano-dcat-ap-0

The Open Data portal indexes external data sources to convert them into RDF format, with which you can interact at the Linked Open Data level. Each record is accessible via URL, reaching the "4 star" level of open data. The initiative aims to collect and catalogue as much Italian open data as possible to allow uniform access to data in tables, graphs or maps.



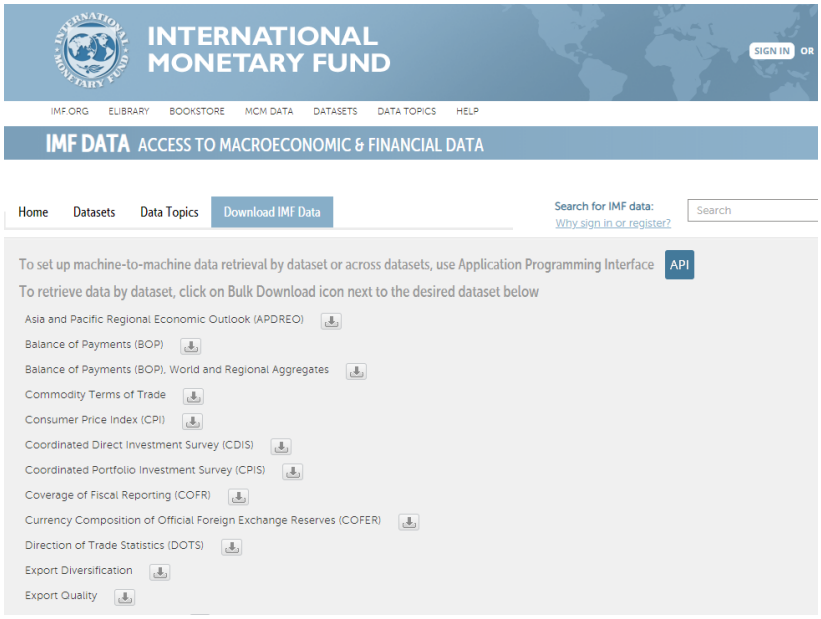
http://www.datiopen.it

The platform is open to participation as a source of data open to citizens, professionals, institutions, society, etc.

UNdata was launched in 2005 as part of a project called "Statistics as a Public Good" with the aim of providing free access to global statistics, literacy to statistics for web users. This United Nations Statistical Service is a service open to the entire network community to consult and download data from over 60 million time series on agriculture, crime, communication, development assistance, education, energy, environment, finance, health, labour, national accounts, population and migration, science and technology, tourism, transport and trade.

https://data.un.org

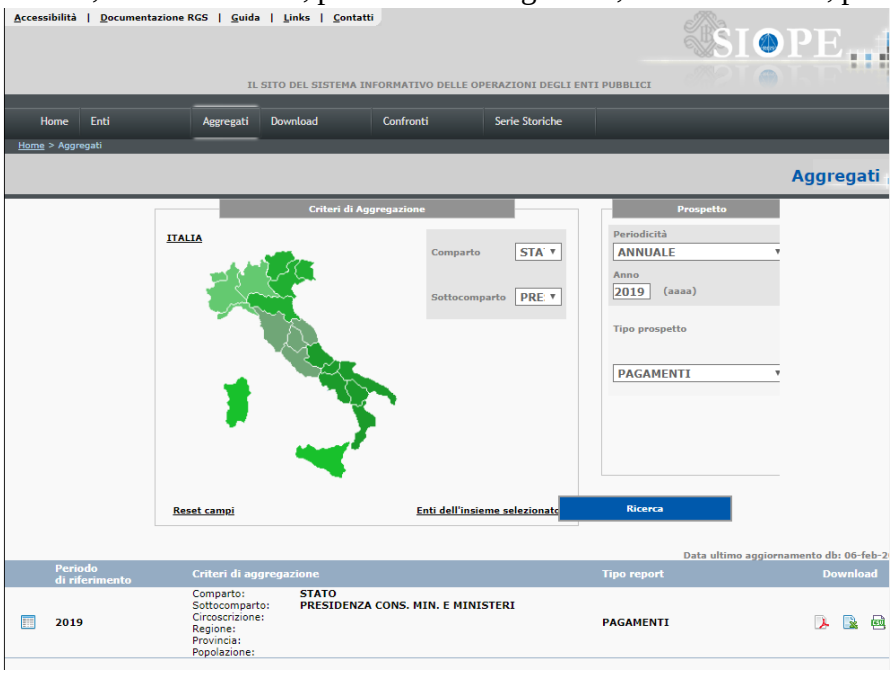
The IMF (International Monetary Fund) offers access to the statistics produced through the "IMF Data Portal", together with graphical statistical tools, to facilitate the analysis of customized data. Among the data sets you can access International Financial Statistics (IFS), Balance of Payments Statistics (BOPS), Directorate of Trade Statistics (DOTS), Public Finance Statistics (GFS).



http://data.imf.org

The SIOPE platform (Information System on the operations of public bodies) tracks collections and payments made by all public administrations, through the synergy between the General Accounting Office of the State, the Bank of Italy and ISTAT.

Since 2003, the system allows monthly monitoring of expenditure and receipts of Regions, local authorities, universities, research bodies, public welfare agencies, health facilities, park and marine protected area managers, Chambers of Commerce, Industry, Crafts and Agriculture.



https://www.siope.it

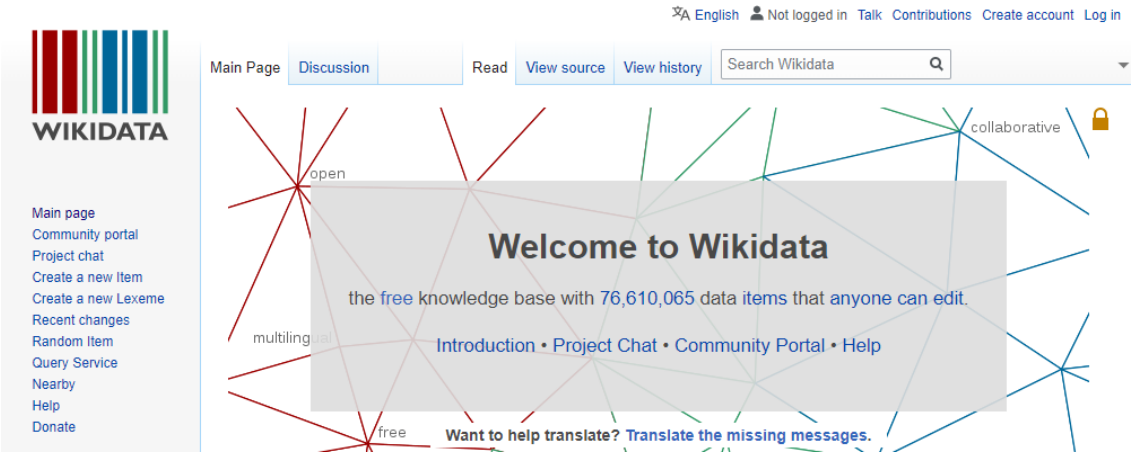
ISTAT (National Institute of Statistics) is the largest producer of open data in Italy, downloadable in CSV format and freely reworkable all. The Institute also makes the data of the corporate datawarehouse available through an SDMX web service, thanks to a machine-to-machine exchange in real time. Since 2015, ISTAT has offered the Linked Open Data (LOD) platform to access and navigate interoperable data directly via the RDF-Data Cube vocabulary.

https://www.istat.it

Wikidata is a secondary, free, collaborative and multilingual knowledge base.
More specifically:

- *"SECONDARY. Wikidata contains, in addition to the structured data mentioned in the 'properties: value' format, also the sources from which they are taken, so as to reflect the diversity of the same in the description of a given topic and to reiterate the centrality that the notion of verifiability has in all Wikimedia projects.*
- *FREE. All data is published under a Creative commons zero (CC0) license, i.e. released into the public domain, so as to ensure the widest possible reuse. In particular, this means that anyone can freely copy, modify, distribute and use Wikidata data in any form, even for commercial use, with the sole obligation to cite the source from which they were taken, just as it happens in all other Wikimedi projects.*
- *MULTILINGUA. Navigation, use, modification and reuse of data is guaranteed in the 285 official languages recognized by the Wikimedia Foundation, as well as in machine-readable formats.*

- *COLLABORATIVE. Data entry and management is supervised by the Wikidata community, which also decides on general project guidelines and specific content guidelines. In particular, the community manages the import of material from reliable sources through automated programs (called bots).. ”*[74]



https://www.wikidata.org/

Google Public Data Explorer platform since 2011 shares data shared by network users and institutions, facilitating the exploration and visualization process.

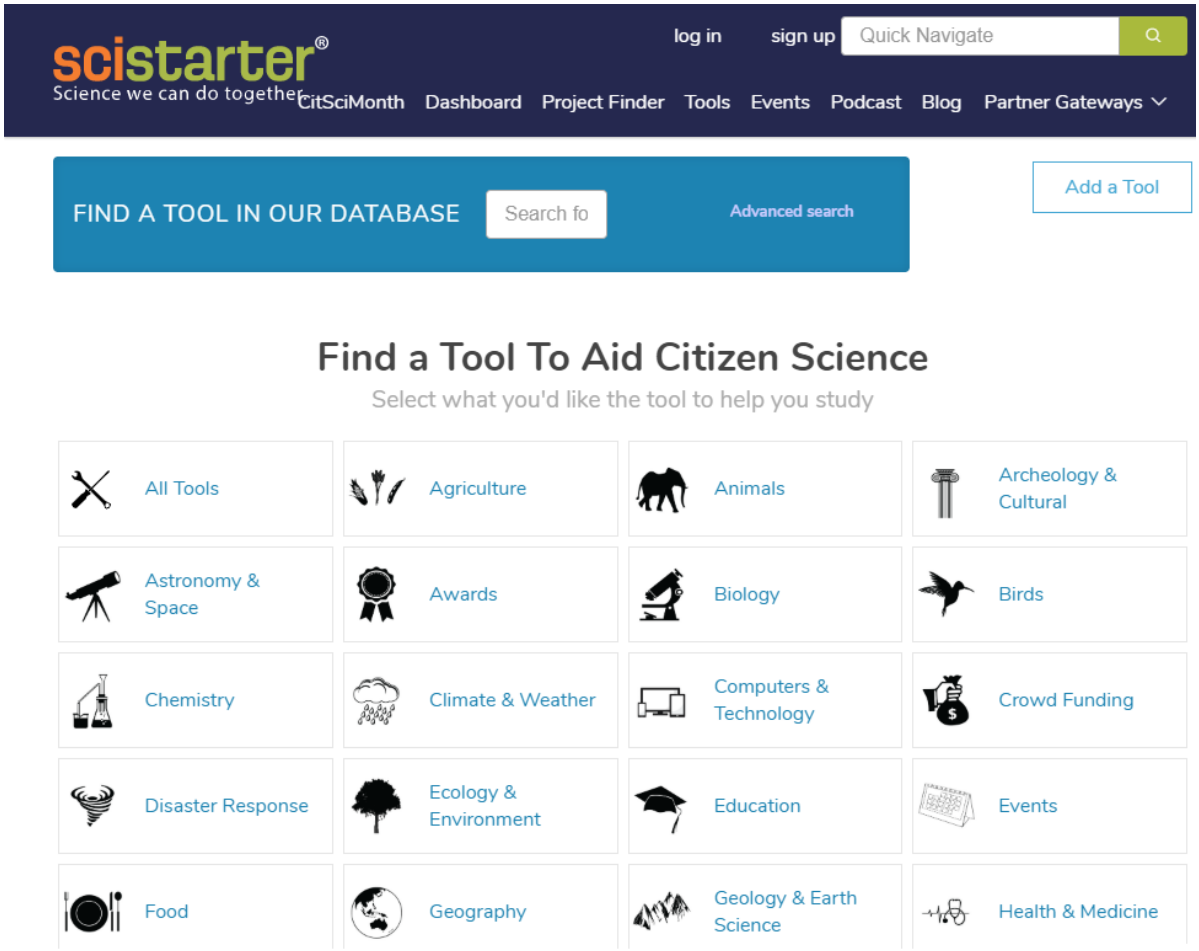
https://www.google.com/publicdata/directory

All datasets are implemented through the interchange format JSON-LD[75] (*JavaScript Object Notation for Linked Data*, object-notation JavaScript for *Linked data*). The user has tools to

[74] Wikidata: la soluzione wikimediana ai *linked open data* https://aibstudi.aib.it/article/view/11434/10655
[75] JSON-LD https://developers.google.com/search/docs/guides/intro-structured-data?hl=it

filter data by categories, changing sorting and graph variables. The system also allows interoperability on databases of international organizations, national statistical offices, non-governmental organizations and research institutes.
In recent years, initiatives have emerged involving citizens as scientists to involve them in data experiments with the aim of facilitating the dissemination of scientific research, participation and understanding of science and technology among the general public.

One example is SciStarter, an online community dedicated to improving citizens' scientific experience with researchers and scientific partners. SciStarter was established at the University of Pennsylvania, in 2014, in partnership with the University of Arizona's School for the Future of Innovation in Society, creating a system to manage and identify the users involved through different platforms and disciplines.



https://scistarter.org

GIS (*Geographic Information System*)[76] tracking allows interaction between scientists and citizens in the same place or city. Participants are autonomous in contributing voluntarily to the projects, with guarantees to protect the privacy of personal profiles.

The association has been active in Europe since 2001 EUSEA (European Science Engagement Association)[77] which uses a knowledge sharing platform involving experts and citizens to design, organize and implement activities with an international and innovative scope. The project is one of the initiatives funded by the European Commission to encourage changes in scientific institutions and civil society, especially with collaborative events aimed at building bridges - between science and society, researchers and citizens, policy makers and innovators.

The main purpose of this initiative is the application of the 10 principles of citizen science disseminated by the European Commission and ECSA (*European Citizen Science Association*)[78]:

[76] GIS (Geographic Information System) http://www.opengeospatial.org/
[77] EUSEA (European Science Engagement Association) https://eusea.info/platform/about-this-platform/about/
[78] ECSA (European Citizen Science Association) 10 Principles of Citizen Science https://ecsa.citizen-science.net/documents

"1) citizens must be able to participate as collaborators or leaders of the project and have a significant role in the project

2) citizens' scientific projects have a genuine scientific result. For example, responding to a research question, informing conservation action or facilitating policy decisions

3) Citizen science provides benefits to both science and society. Benefits can include learning opportunities, personal enjoyment, social benefits, publication of research results, contribution to scientific evidence that can influence policy on many scales (local, national and international) and the connection of the wider community with science.

4) Citizens scientists can participate in the various stages of the scientific process. This may include developing research questions, designing methods, collecting and analysing data and reporting the results.

5) Citizens scientists receive feedback from the project. For example, how their data and the results of research, politics or society are used.

6) Citizen science, like all forms of scientific investigation, has limitations and prejudices that should be considered and controlled. However, unlike traditional approaches to research, citizen science offers greater opportunities for public involvement and participation, increasing the accessibility of science in society.

7) Where possible and appropriate, project data and metadata of citizen science projects shall be made publicly available and the results published in an open access format. Data sharing may take place during or after the project, unless there are security or privacy issues that prevent this from happening.

8) the city scientists are adequately recognised by the projects. This may include recognition in project communications, communication of results and publications.

9) Citizens' scientific programmes offer a number of advantages and results which should be recognised and taken into account in the evaluation of the project.

10) Citizens' scientific project leaders take into account legal and ethical considerations of the project. These considerations include copyright, intellectual property, data sharing agreements, confidentiality, attribution, security and well-being of participants, traditional consultation of owners and the environmental impact of any activity."

CHAPTER IV

Proposal for Data Stewardship as a support to the ethical certification SA8000

In recent years, many private and public organizations have added a social report to their financial statements to certify the sensitivity of management to issues such as sustainability and the ethical management of economic, human, environmental and financial resources.

At the same time, several international organizations, such as the World Bank, the International Monetary Fund, the World Trade Organization and the Organization for Economic Cooperation and Development have promoted the development and recognition of standards that go alongside social norms and conventions.

Transnational non-governmental organizations such as ISO (International Standard Organization), SAI (Social Accountability International), ISEAL (International Social and Environmental Accreditation and Labelling Alliance) have been established to define principles to resolve disagreements and create convergence among the various stakeholders. Although they do not enjoy local or international regulatory power, standardization procedures strongly influence global market mechanisms with peculiarities similar to governmental ones. In different contexts, standards are agreed rules, the result of negotiation between several actors, who voluntarily decide to adhere to them. In other contexts the standard is a set of technical specifications shared by producers, tacitly or through formal agreements.

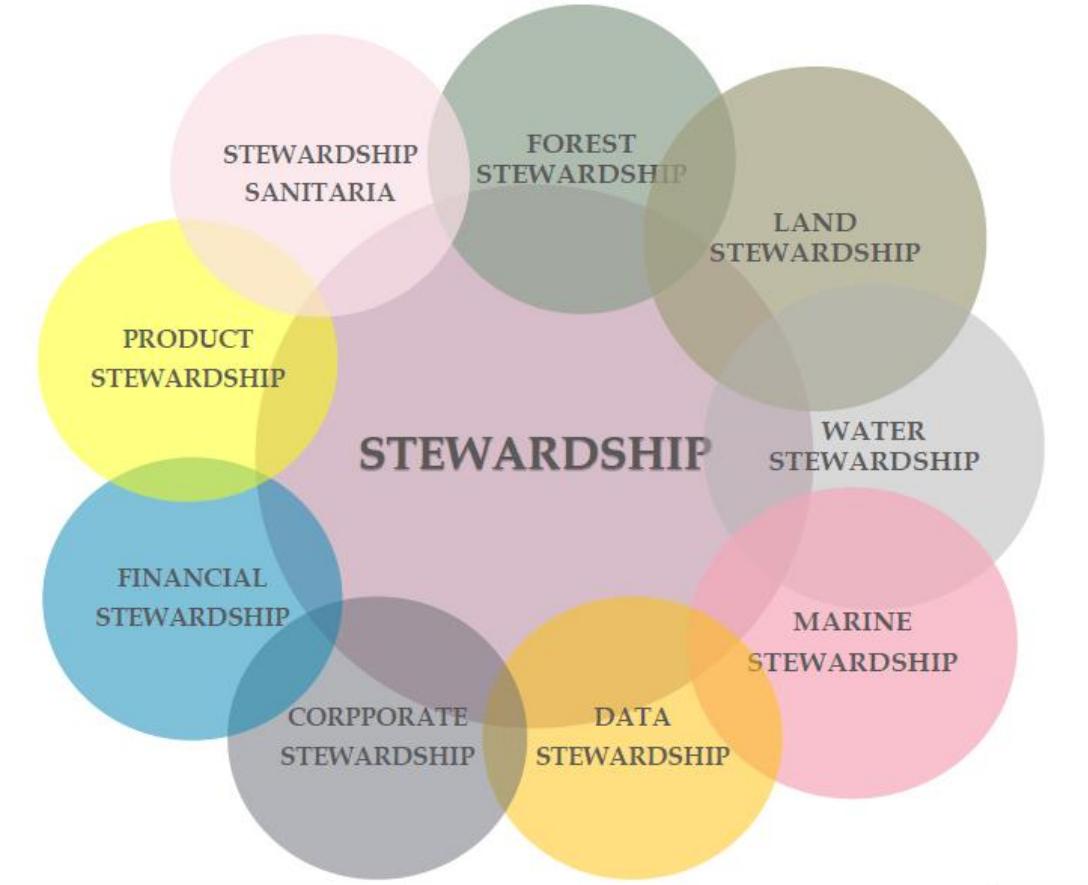
The standards do not have the tax/sanctifying power of laws and governments. The basis of this paradigm is voluntary adherence and the absence of sanctions for those who do not comply with the standard.

The set of international standards makes up the system of good management certification, monitoring and accrediting the adaptation and application of criteria, principles and procedures defined by the relevant body.

Today SA8000 certification is an international standard, with a three-year duration and six-monthly verification, aimed at certifying certain aspects of company management related to social responsibility, with particular reference to ethical work. This SA (Social

Accountability) focuses on improving working conditions and safety in the workplace, based on the ILO (International Labour Organization) Conventions, the Universal Declaration of Human Rights, the International Convention on the Rights of the Child, the United Nations Convention to eliminate all forms of discrimination against women.

Stewardship in the organization focuses on the importance of cooperation, accountability and stakeholder sharing, emphasizing their influence through engagement. Some international examples of environmental stewardship are the following schemes ASC (Aquaculture Stewardship Council)[79], CSA (*Canadian Standards Association*)[80], SFI (*Sustainable Forestry Initiative*)[81], FSC (*Forest Stewardship Council*)[82] and PEFC (*Programme for Endorsement of Forest Certification*)[83].
The fields of application are different:



SOURCE: Casino A., Stewardship. Politiche e pratiche per una gestione etica delle risorse

[79] ASC (Aquaculture Stewardship Council) https://www.asc-aqua.org/about-us/about-the-asc/
[80] CSA (Canadian Standards Association) https://www.csagroup.org/standards
[81] SFI (Sustainable Forestry Initiative) https://www.sfiprogram.org/sfistandards/
[82] FSC (Forest Stewardship Council) https://it.fsc.org/it-it/certificazioni
[83] PEFC (Programme for Endorsement of Forest Certification) https://www.pefc.org/discover-pefc/what-is-pefc

The Stewardship[84] theory of community values are emphasized through cooperative strategies, to the detriment of utilitarian strategies. It speaks of models of governance (or corporate governance) where leadership is based on close collaboration between stakeholders, but also on mutual trust and ethics, reducing the weight of regulatory obligations and decision-making hierarchies. A vision based on transparency, accountability and collaborative partnership is born. In some production contexts, Stewardship means the guarantee of eco-efficiency of the product, in relation to its sustainable characteristics for industry and the environment.

The international standard SA8000 establishes the requirements for an ethically correct behaviour of companies and the production chain towards workers. At the base there is the orientation to improve the competitiveness of the market on a voluntary basis of stakeholders, through the detection of ethicality of its production chain and its production cycle. Fundamental issues such as: respect for human rights, respect for workers' rights, protection against exploitation of minors, guarantees of safety and health in the workplace become fundamental[85]. In this "social balance sheet analysis" the actors involved are suppliers, customers, consumers and their associations, environmental and humanitarian groups, trade unions, the media, the local/national catchment area.

Today 4,483 organizations, for 2,092,945 workers, in over 60 countries and 57 production sectors[86] have chosen SA8000 to build a sustainable working environment by analyzing:

1. child labour
2. labor
3. health and safety
4. freedom of association, right to collective bargaining
5. discrimination
6. disciplinary practices
7. working hours
8. salaries
9. management system

---

[84] Davis J., Schoorman D. and Donaldson L., *Toward a Stewardship Theory of Management*, in The Academy of Management Review Vol. 22, No. 1 (Jan., 1997), pp. 20-47

[85] La responsabilita' sociale e etica http://www.sa8000.info

[86] Dati al 20/2/2020 forniti da Social Accountability Accreditation Services http://www.saasaccreditation.org/certfacilitieslist

This standard bases the evaluation of the organisation on the involvement of stakeholders (internal and external stakeholders), with the contribution of the staff in service at each level/location. The technical references are summarised in the annexes: Performance Indicators (document defining the minimum performance expectations of an organisation) and "SA8000 Guidance Document" (interpretation of the standard to clarify how to apply it during audits).

The approach is based on interviews that maximise management improvement through prevention at the expense of ex-post control, in order to achieve the competitive advantages linked to:

- *"an increase in credibility, transparency and corporate image in the reference market, a reputational benefit;*
- *increased consumer confidence, improved relations with social institutions and organisations;*
- *a control of ethicality and social fairness in the supplier chain;*
- *an improvement in the company climate: workers feel protected by the company in which they operate and more involved in achieving their objectives;*
- *an improvement in internal and external communication, through reports made publicly available."*[87]

Companies are part of the data stakeholders, those who need data available quickly and with interoperable platforms based on policies and standards of entry and exit in the information management flow: data governance. On the other hand, data stewardship deals with the ethical and responsible management of the data or data set, focusing on the database architecture and related automation processes, in order to "*foster:*

- *consistent use of resource management data;*
- *easier data mapping between computer systems and document exchange;*
- *a reduction in costs associated with the migration of service architectures."*

[87] CSQA Certificazioni https://www.csqa.it/CSQA/Norme/Responsabilita-Sociale/SA-8000

Therefore, data stewardship tools are used in data quality and master data management projects with the aim of eliminating duplicate entities and improving the reusability and quality of corporate data."[88]

In Italy there is no certification or accreditation system that can issue data stewardship certification. The SA8000 needs an update of the transversal requirements to be able to decline data stewardship in companies and research institutions, integrating the sharing and openness guaranteed by the principles of FAIR (*Findable, Accessible, Interoperable, Reusable*)[89] data. This version of the research data allows anyone to verify the methods, software, content, processing and dissemination processes of the scientific community. The European Commission has adopted and clearly defined the technical requirements for FAIR data repository[90] and limits on the use of open data in the Regulation GDPR *(General Data Protection Regulation)*[91].

In the near future the open science movement may involve realities such as AIB (professional association of Italian librarians), RUS (Network of Universities for Sustainable Development), IFLA (International Federation of Library Associations and Institutions) and ASVIS (Italian Alliance for Sustainable Development), to propose to SAI (Social Accountability International) updates of technical requirements, emphasizing how open and shared data can facilitate the pursuit within the UN Agenda 2030 for the targets:

- 4.6 ensure that all young people and a substantial proportion of adults, both men and women, achieve literacy and numeracy skills
  (in Goal 4: Providing quality, fair and inclusive education and learning opportunities for all)
- 10.2 By 2030, to enhance and promote the social, economic and political inclusion of all, regardless of age, gender, disability, race, ethnicity, origin, religion, economic status or otherwise
  (in Goal 10: Reducing inequalities within and between Nations)
- 16.6 Developing effective, accountable and transparent institutions at all levels

[88] Casino A., *Stewardship. Politiche e pratiche per una gestione etica delle risorse*, Aracne, 2015
[89] Dati FAIR (Findable, Accessible, Interoperable, Reusable) https://www.force11.org/group/fairgroup/fairprinciples
[90] Data repository FAIR https://ec.europa.eu/info/sites/info/files/turning_fair_into_reality_0.pdf
https://www.openaire.eu/how-to-make-your-data-fair
[91] GDPR https://www.europeandataportal.eu/it/highlights/protecting-data-and-opening-data

- 16.10 Ensuring public access to information and protecting fundamental freedoms, in accordance with national legislation and international agreements
  (in Goal 16: Promoting peaceful and more inclusive societies for sustainable development; providing access to justice for all and creating effective, responsible and inclusive bodies at all levels)
- 17.6 Improving North-South, South-South and triangular cooperation at regional and international level and access to science, technology and innovation and improving knowledge sharing on mutually agreed conditions, including through greater coordination between existing mechanisms, in particular at UN level, and through a global technology facilitation mechanism
- 17.7 Promoting the development, transfer, dissemination and diffusion of environmentally friendly technologies to developing countries on favourable conditions, including at concessional and preferential terms, as mutually agreed
- 17.9 Strengthening international support for the implementation of an effective and targeted capacity building system in developing countries to support national plans to implement all sustainable development objectives, including through North-South, South-South and triangular cooperation
  (in Goal 17: Strengthening Means of Implementation and Renewing the Global Partnership for Sustainable Development)

Within the scientific community, research institutions and universities use the social budget to respond to: strategic needs (disciplinary leadership), reputational needs (towards students and citizens), attractive needs (for partnerships and sponsorships), distinctive needs compared to national and international competition. These bodies, as well as individual departments, consortia or research centres, could benefit from the future "SA8000 data stewardship", in management reporting both for the management of workers and suppliers, and for research data. The latter in fact can become a competitive factor of stakeholder engagement: an element of involvement of the main stakeholders, to activate participation, increase transparency, create opportunities for innovative partnerships, increase stakeholder trust.

All Horizon2020 projects funded by the European Commission require open access dissemination of the publication but also of research data (excluding patents). In the new

Horizon Europe programme, research is driven by policies in line with the Sustainable Development Goals (UN Agenda 2030) in order to develop more competitive research infrastructures in the European Research Area to promote open science and broaden the audience through open access to scientific publications and research data[92]. The European Commission requires that such data and publications be deposited in open institutional or disciplinary archives or in Zenodo[93].

The free availability and interoperability of research data also contributes to the provisions of Article 19 of the UN Declaration of Human Rights:

> *"Everyone has the right to freedom of opinion and expression, including the right not to be harassed for his or her opinion, and the right to seek, receive and impart information and ideas through any media and frontier".*

[92] Horizon Europe - the next research and innovation framework programme https://ec.europa.eu/info/horizon-europe-next-research-and-innovation-framework-programme_en
[93] Progetto OpenAIRE - https://zenodo.org/communities/openaire/?page=1&size=20

CHAPTER V

Data stewardship in the library in the public engagement of researchers

In order to have a measure of the degree of importance of efficient data management, it is possible to evaluate the amount of the Data Management tender, published in CONSIP in February 2020 for a value of 222 million euro with "the objective of making available to administrations application and professional services for the implementation of IT solutions for the valorisation of information assets and data models to support decision-making processes"[94]. This appropriation is part of the Three-Year IT Plan in the PA 2019-2021.

The field of open data management is one of the new frontiers of library services, especially in universities and research centres. It is a point of contact and support for the library world to participatory research and open science, based on the "FAIR principles for scientific data management"[95].

The FAIR research data principles:

FINDABLE

F1. (meta)data are assigned a globally unique and eternally persistent identifier.

F2. data are described with rich metadata.

F3. (meta)data are registered or indexed in a searchable resource.

F4. metadata specify the data identifier.

[94] PA digitale, nuova gara "data management" con 222 mln https://www.01net.it/consip-gara-data-management/
[95] The FAIR Guiding Principles for scientific data management and stewardship https://www.nature.com/articles/sdata201618

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.

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.

RE-USABLE:

R1. meta(data) have 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 their provenance.

R1.3. (meta)data meet domain-relevant community standards.

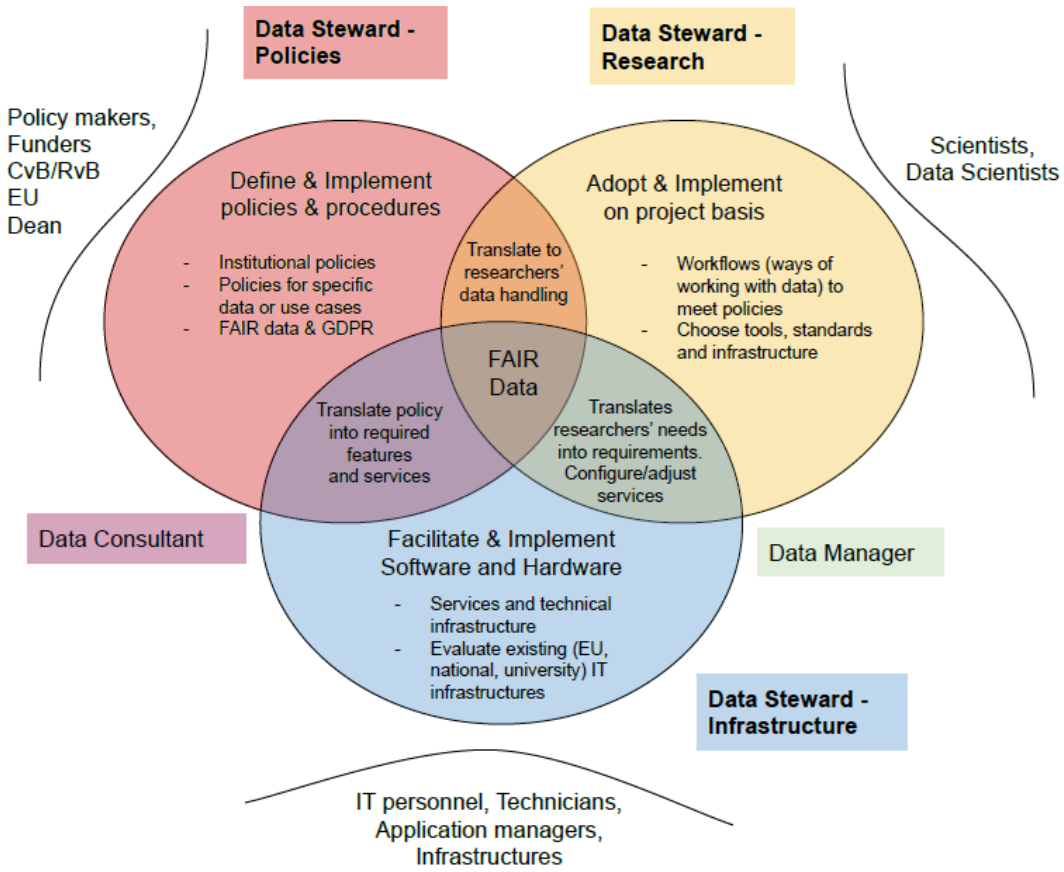
The librarian evolves from data curator to data steward in order to improve data management for open science. This implies the need to adapt skills and professionalism to the needs of the institution, discipline, country and region where the research is conducted.

The researcher is the main "data stakeholder" within the scientific community, linked to partnerships between universities (especially in European projects) and between competing research organizations such as ministries, companies, consortia, foundations, non-profit organizations.

It becomes essential to establish pragmatic cooperation, alongside the sharing of expertise, requirements, services and protocols, formats, models and standards that transcend geographical and interdisciplinary boundaries.

We are talking about a new concept of data stewardship called "FAIR Research data stewardship" aimed at sharing:[96]

- *"FAIR good practices, optimizing technologies (apps and open source code) and repositories*
- *FAIR metadata models (generic or discipline specific)*
- *strategies for managing data privacy, licensing, security and copyright*
- *testing and feedback of funded instruments*
- *necessary skills and data management requirements*
- *e-learning tools, courses and training programmes*
- *incentives approved to support FAIR research*
- *strategies to make this process sustainable in the long term"*



SOURCE: data stewards function landscape and its stakeholders

[96] DSCC (data stewardshiphip Competence Centres) https://www.go-fair.org/implementation-networks/overview/dscc/

The librarian as *data steward*[97] has the task of promoting cultural change towards shared research, acting on those who produce and reuse the data.

An effort is also required to facilitate the technology transfer necessary for open science, through synergy at all operational, financial and decision-making levels of the institution.

A critical mass needs to be achieved within a reasonable timeframe, supported by a proactive role of researchers and ICT (Information and Communication Technology) sectors.

In addition, the data steward will have to support platforms for the exchange of practical information, frequently asked questions, replicable solutions, and mutually supportive discussion groups.

In this perspective active communities in the scientific community can be particularly useful such as "*Research Data Alliance*"[98], *European Open Science Cloud*[99], *Open Science Support Group*[100].

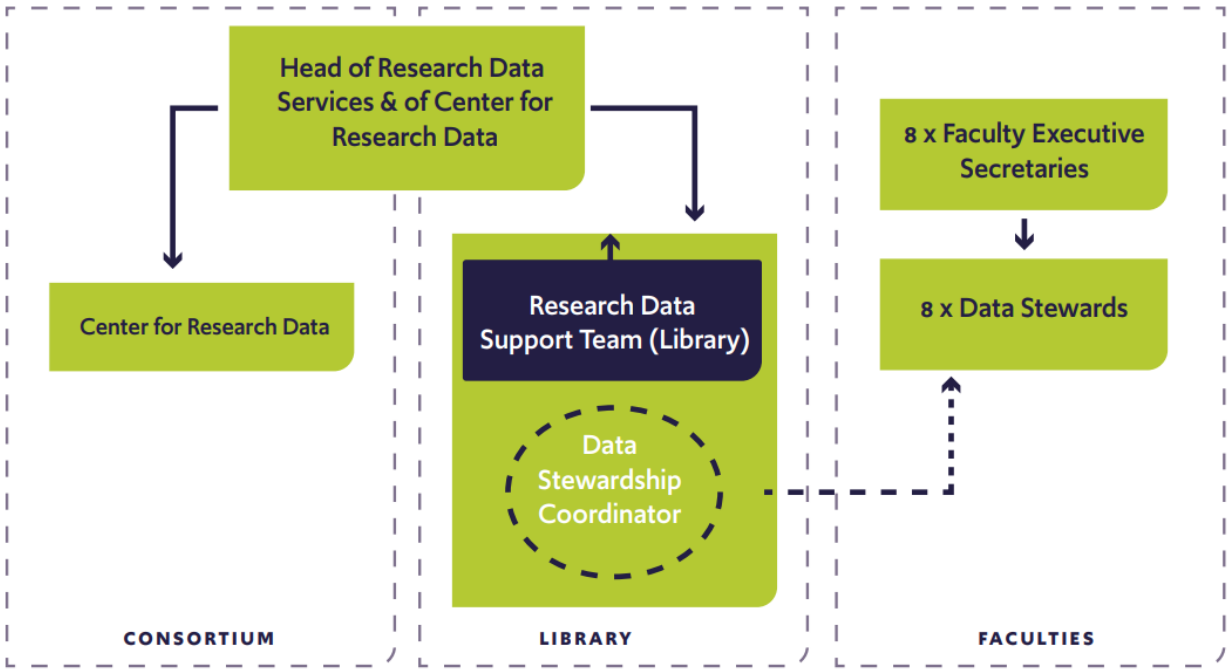
The data steward responds with his skills to the typical needs of DMPs (data management plans), providing solutions on how to*:*

- data are generated and/or retrieved
- data are used, processed and organised
- data and persons concerned are protected
- data, code and auxiliary elements are described and documented
- data are stored and protected and for how long they will be kept
- authorship and credit of data are assigned
- data are stored
- results of the research are shared

If there are changes in the archiving structure, formats, software, it is necessary to accurately track how the process is modified and how it can be remedied in the future for the replicability of the processing / experiment. In this view, in addition to the data, it is essential to clearly detail the methodology, storage and reuse conditions.

[97] data stewards function landscape and its stakeholders https://zenodo.org/record/3460552#.XkBO52hKhhk
[98] Research Data Alliance https://www.rd-alliance.org
[99] European Open Science Cloud https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud
[100] IOSSG - Italian Open Science Support Group https://sites.google.com/view/iossg/home

The working group "*National network research data management*"[101] has developed the following scheme to frame tasks and roles in Dutch research institutions [102].



SOURCE: Dutch institutes: National Coordination Point Research Data Management facilitates a national strategy for research data management in the Netherlands

Two teams provide support for research data: the support team in the library and the disciplinary data managers deployed at the different faculties. The data stewards act as primary contacts for researchers thanks to their specific research know-how.

The central team acts as a contact point for more general issues, such as the use of the central data archive. The data management coordinator, who is based in the library for organisational reasons, is the linchpin between the central support team and the data managers. A coordinator manages the data steward team.

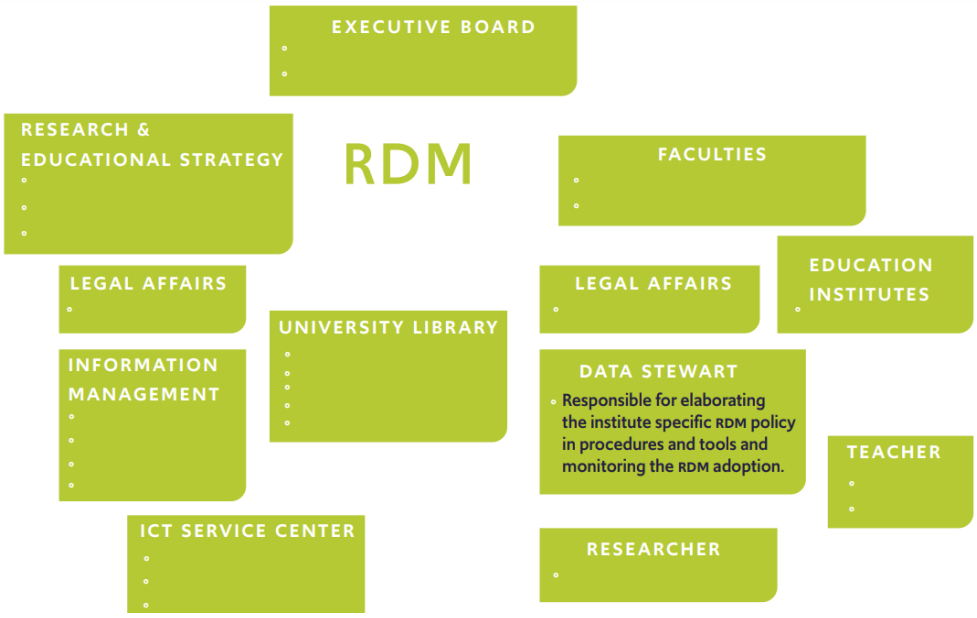
At the heart of the entire DMP are the metadata (data on the data) that allow appropriate indexing inside and outside the institutional archives:

- name of the data creator (data manager for consortia) tracked with id (ORCID Open Researcher and Contributor ID [103])
- contact email(s)
- title of the dataset

[101] National network research data management https://www.lcrdm.nl/en
[102] Dutch institutes: National Coordination Point Research Data Management facilitates a national strategy for research data management in the Netherlands https://zenodo.org/record/2669150#.XkLyh1VKjIU
[103] ORCID (Open Researcher and Contributor ID) https://orcid.org

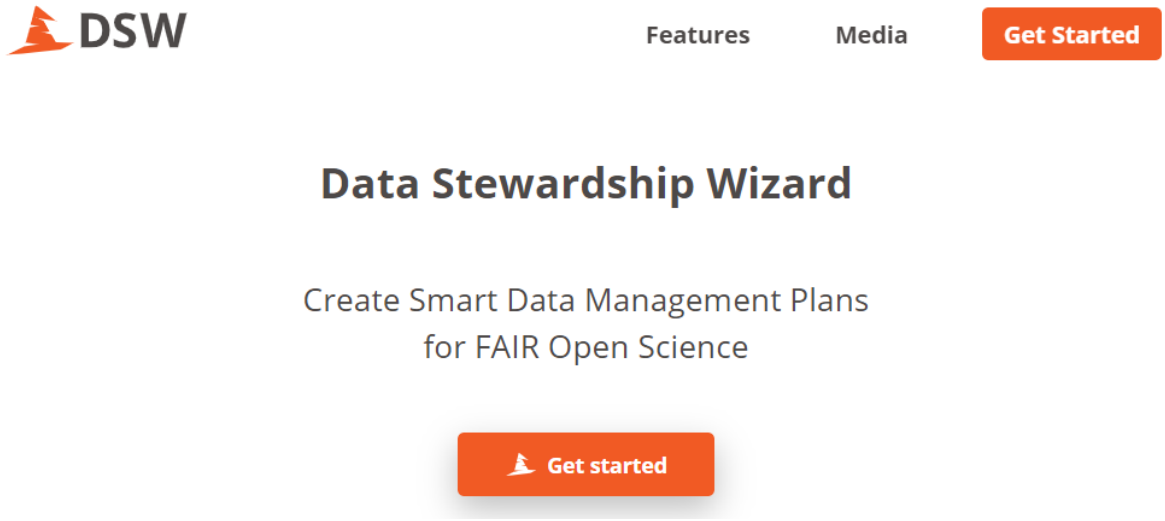
- description of the data (what data, the purpose of the research project and the methodology)
- data source (newly generated data set / data collection methods / third party sources)
- type of data (statistical / textual / observational / computational / experimental...)
- year of completion of data
- dataset coverage range
- geographic coverage of the dataset
- data format (e.g. state 15, .csv, excel, .txt, &c.).
- codeboook / supporting documentation
- access status (opening data by default or embargo)
- CC (Creative Commons) license
- language
- funding statement
- related publications
- future implementation projects
- citation of the dataset (short citation of the data in future publications)

In this context, interaction with all internal and external stakeholders is fundamental:



SOURCE: Dutch institutes: National Coordination Point Research Data Management facilitates a national strategy for research data management in the Netherlands

In order to facilitate librarians and researchers in the elaboration and updating of the information flow DMP, verify compliance with FAIR principles for data, there is an ad hoc tool that can be customized online free of charge: "*data stewardship Wizard*"[104].



https://docs.ds-wizard.org/en/latest/

We start from the EDITOR phase to define the data set and the research context.
The MIGRATION function enables the integration phase or modifies other stewards. With KNOWLEDGE MODEL you choose the questions with their options and advanced customizations. The QUESTIONAIRE phase involves researchers in compiling the metadata for the data set. Finally, based on the feedback, you choose whether to close the DMP or to proceed with additional investigations.

SOURCE: data stewardship Wizard Documentation

[104] data stewardship Wizard is an international project to help serious researchers and data stewards with building smart Data Management Plans for FAIR Open Science. https://docs.ds-wizard.org/en/latest/

The functioning of the workflow is based on on-line responses that report the opening requirements defined by the data manager, in order to create a precise picture of all shared data sets. The platform also offers many schemes prepared for customization and reuse.

CHAPTER VI

Data stewardship for monitoring and verifying the SA8000 application

The SA8000 Certification in its application of the 9 requirements of the *Performance Indicator Annex*[105] produces a set of data related to the internal management of the organization, relations with suppliers and workers. Generally, certified companies produce, manage and are obliged to maintain data and information to ensure the company's compliance with the certification requirements, while certification bodies acquire and store this data in order to draw up the relevant documents typical of surveillance or certification processes.

Among the data and information that are collected during the six-monthly surveillance or certification stage are:

1. first interviews with company department heads and inspections at all company sites
2. management review
3. system documentation for corporate social responsibility, company roles and functions
4. supply chain
5. supporting documentation for the management system, which varies according to company activities, personal data and employee statistics
6. system documentation for the monitoring and assessment of suppliers' risks
7. system documentation relating to employee data

[105] Performance Indicator Annex (2014)
http://www.sa-intl.org/_data/n_0001/resources/live/Performance%20Indicator%20Annex_11_17_2014.pdf

8. records relating to SA8000 policy communications with suppliers, customers and stakeholders, training of personnel involved and training on their management
9. forms and records relating to complaints and/or reports from employees, customers, suppliers and stakeholders all regarding the company's compliance with SA8000 requirements
10. plans and audit reports at the company's suppliers
11. scheme for improving the system
12. documents and data for the activation of the certification process
13. assistance until certification is obtained (documentary verification and body visits to assess the system)
14. SA8000 ethical certification

In applying data stewardship to these quantitative and qualitative flows it is proposed an adjustment (further to what is indicated in Chapter IV) by SA8000 to implement monitoring, transparency and verifiability by applying the FAIR principles (findability, accessibility, interoperability, and reusability) and the dictates of the European GDPR (General Data Protection Regulation).

In perspective, data governance could also integrate other international data protection regulations such as the CCPA (*California Consumer Protection Act*)[106], l'HIPAA (*Health InsurancePortability and Accountability Act*)[107], and industry requirements such as PCI DSS (*Payment Card Industry Data Security Standards*)[108].

This process will require a single global data platform for certification companies to make the tracked data accessible and interoperable directly on the site *Social Accountability Accreditation Services* **http://www.saasaccreditation.org**

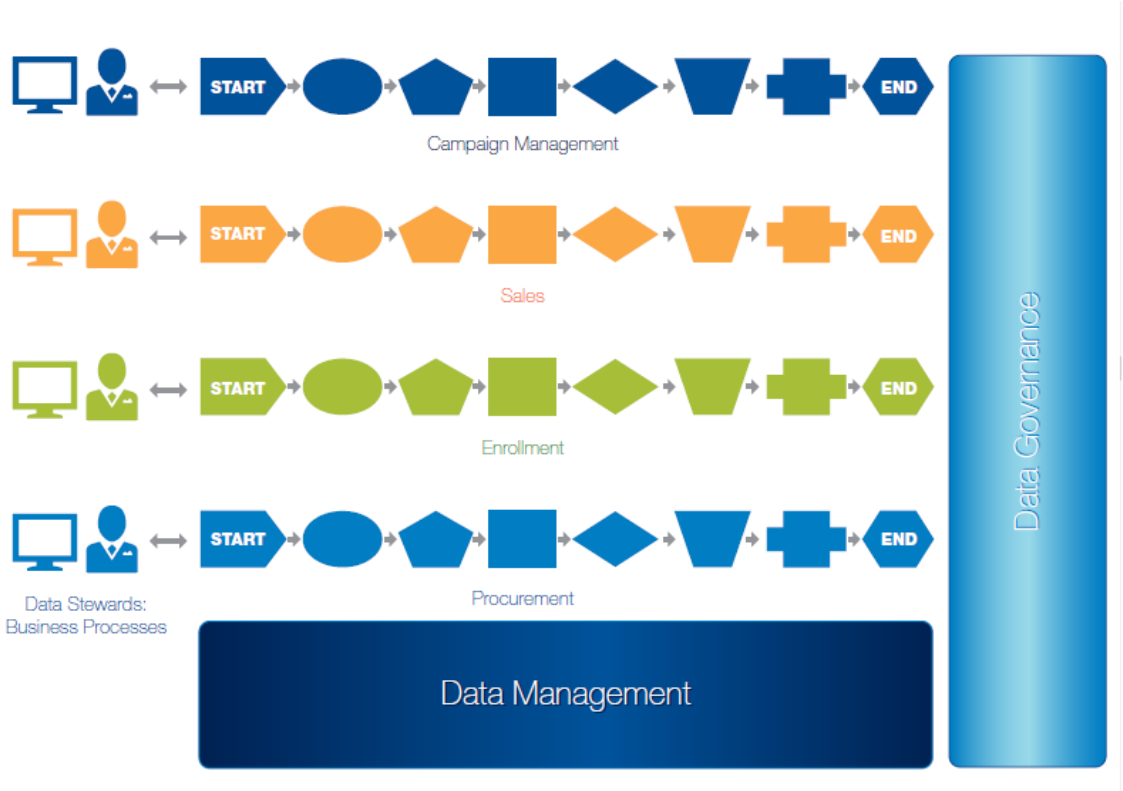
The platform would allow the sharing in real time, in a standardized and complete form, of the information detected by the entire certification process. Within data governance, data stewardship is the process of managing the data lifecycle, ensuring the maintenance of data models, traceability, processing, according to a priori defined policies.

[106] California Consumer Protection Act https://www.caprivacy.org/
[107] Health InsurancePortability and Accountability Act
[108] Payment Card Industry Data Security Standards https://www.pcisecuritystandards.org/pci_security/

Once fully operational, this system would bring improvements in data governance processes, especially in monitoring, refining, de-duplication and aggregation activities. With the necessary limitations, interoperability protocols could be implemented for third party applications to allow the verifiability of the activities carried out and data integrity.

The document flow should allow sharing without the support of data experts, in order to facilitate the participation of different consultants of certification companies. In this context, data stewardship provides for the generation of a collaborative system of data fixing, merging and certification. This also involves a system of permissions on which figures can send or modify data, delegating actors able to send reliable information in the necessary technical time.

To achieve this degree of organization it is necessary to define a "data map" to ensure the accuracy, completeness and consistency of the data. The data map in the SA8000 allows to understand the insertion phases, the completeness of the indexed entities and the status of data integration.



SOURCE: data steward by Business Process [109]

[109] Dyché J., Polsky A., 5 Models for data stewardship https://www.sas.com/content/dam/SAS/en_au/doc/whitepaper1/5%20Models%20Data%20Stewardship%20White%20Paper.pdf

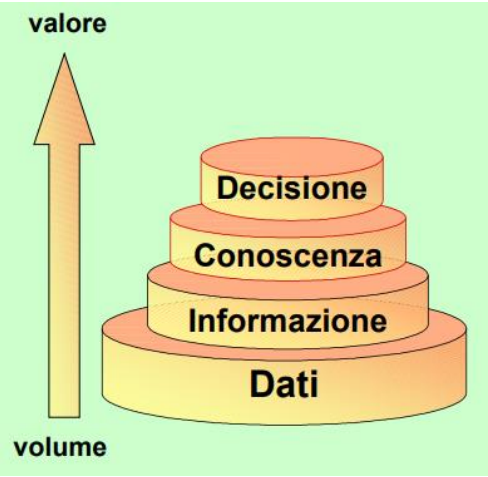
In this scheme the data steward is responsible for his specific data set, available in the certification processes.

Since the steps and which data to collect are well defined, the data stewards are responsible for data domains or applications/systems of the same process, with the possibility to be supported in case of large amounts of data in a tight time frame.

The linear management model of the certification steps would allow several advantages:

- the companies that are certified would have a complete overview of the certification status
- traceability allows clear, consistent, timely visibility of the status of the detection phases
- Automation of the data sharing process can improve data management consistent with certification.
- ease of access can increase the degree of reliability, retention and availability of data

The biggest risk, as in all processing systems, is "garbage in - garbage out": even if the model is advanced and refined, it will never make up for the poor correctness/distortion of the inserted data.



SOURCE: Data Mining [111]

The final result of the certification process would become the "*Knowledge Discovery in Database*"[110]: a process of extraction of knowledge from a database, organizing the results by certification phase, by sectors, by geographical areas, by types of companies.

These elaborations can also be structured by recurring objectives.

[110] Knowledge Discovery in Database https://www.techopedia.com/definition/25827/knowledge-discovery-in-databases-kdd

[111] Data Mining – blog Lorenzo Govoni https://lorenzogovoni.com/processo-estrazione-dei-dati-data-mining/

Data processing and extraction procedures are defined as data mining: data processing strategies and algorithms in a dynamic and integrated environment. Each data processing and extraction session corresponds to a pattern[112]*: "a structure, a model, or, in general, a synthetic representation of the data.*

The important thing is that the pattern is:

- *understandable, from a semantic and syntactic point of view, so that the user can interpret it;*
- *potentially useful, so that the user can understand it;*
- *valid on data with a certain degree of confidence;*
- *previously unknown."*

In particular, with the descriptive models of data mining, similarities or shared groupings are found in the data collected to determine groupings by sectors or objectives:

- clusters group the elements of a set, depending on their characteristics
- classification models make it possible to derive a model based on a set of classes assigned a priori (e.g. decision trees)
- association rules allow you to determine affinity groups between objects

Operationally, once fully operational, the platform would work as a Data Warehouse: a unique, complete and consistent "container" of data obtained from heterogeneous sources, built for end users so that they can carry out analyses aimed at achieving their business objectives.

This type of tool would allow for certification process assessments focused on advanced statistical processing in a "*data environment:*

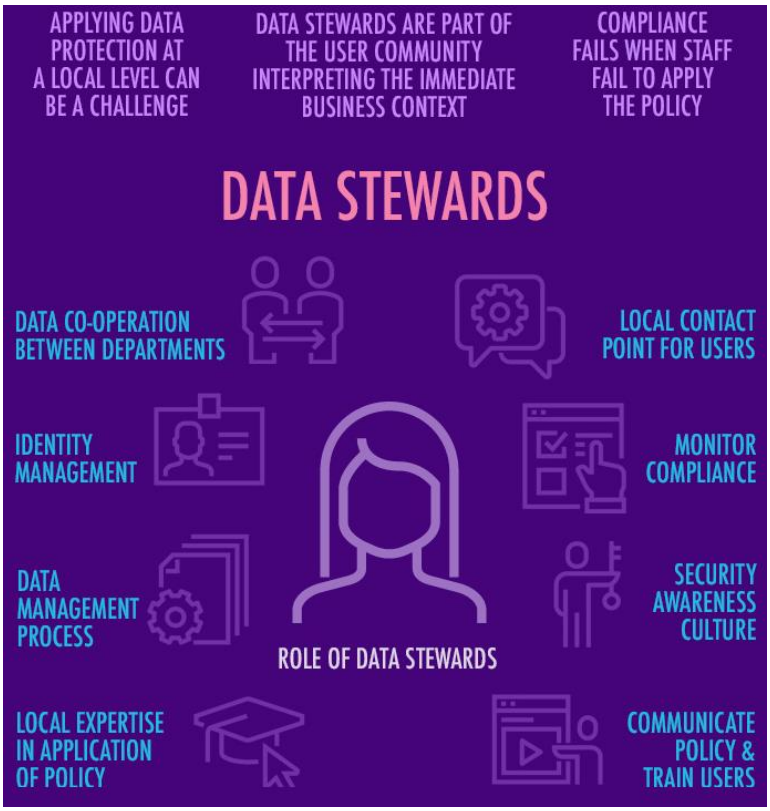
- *Integrated: the different sources flow into a single homogeneous "container" within it*
- *Subject oriented: the Data Warehouse focuses on the main topics of corporate interest (customers, products, channels, etc.) and not on individual applications/processes (sales, loans, outbound traffic, etc.).*

[112] Data Mining https://lorenzogovoni.com/processo-estrazione-dei-dati-data-mining/

- *Non-volatile: the data contained in the Data Warehouse are subject to periodic updates (generally monthly), which determine their continuous growth, but in fact they are "static" data and cannot be modified by end users (read-only access).*
- *Variable over time: the data stored in the Data Warehouse represent a periodic "snapshot" of the company's situation and cover a time horizon of several years (extended history)"*[113]

Data stewardship theory applied to the SA8000 certification procedure becomes a tool to support strategic decisions, based on certified data that allow to create shared knowledge.

The data Steward[114] becomes "the person responsible for Data Quality and usability, with the task of making them available to anyone who has to develop questions, analyses, transactional applications. A logical and organizational role, free from specific technologies, but closely related to the procedures for safeguarding data and their integrity."



SOURCE: DATAVAULT[115] data stewards GDPR ambassadors

[113] Inmon WH, Building the data Warehouse http://fit.hcmute.edu.vn/Resources/Docs/SubDomain/fit/ThayTuan/DataWH/Bulding%20the%20Data%20Warehouse%204%20Edition.pdf

[114] Giacchino A, Data Steward per Mestiere: più Semplice che Data Analyst, più di Valore per l'Impresa! https://www.itware.com/blog-itware/big-data-analytics-data-management/item/1311-data-steward-per-mestiere-piu-semplice-che-data-analyst-piu-di-valore-per-l-impresa

[115] Datavault: data stewards gdpr ambassadors – infographic https://www.data-vault.co.uk/data-stewards-gdpr-ambassadors-infographic/

One of the main interlocutors in this new paradigm will be the "Business Data Steward", an indispensable pivot in the communication between corporate governance and society, understood as a set of different stakeholders, for whom the implementation of the information flow becomes crucial, within the regulatory framework of the European GDPR regulation for privacy and data protection.

CHAPTER VII

Date repositories for open science

Academic research uses institutional archives to share scientific products with the Internet community: articles, books, papers, conference contributions, teaching materials, doctoral theses.
Some universities and consortia have developed ad hoc platforms to share also data sets, software and methods for the reproducibility of studies and research in compliance with FAIR principles (findability, accessibility, interoperability, and reusability) and the limitations of the European GDPR (General Data Protection Regulation).

For researchers the main advantages of archiving their own data may be:

- long-term storage of data in the necessary formats and metadata
- sharing with the whole Internet beyond any geographical limit
- data traceability via DOI (Digital Object Identifier)
- the visibility of content for disciplinary and civil society collaborations
- the possibility of receiving feedback from specialists in the field

For the architecture (hardware/software server) of the repositories there is an international certification that attests the technical compliance with FAIR requirements: CoreTrustSeal[116] certification. In Italy the platform "Phaidra"[117], by Library System of the University of Padua, has obtained this certificate by verifying the management of the archive in ensuring its sustainability for digital preservation in the long term.

[116] certificazione CoreTrustSeal https://www.coretrustseal.org/why-certification/certified-repositories/
[117] piattaforma "Phaidra" del Sistema bibliotecario dell'Università di Padova
https://bibliotecadigitale.cab.unipd.it/news/phaidra.-ultime-novita

In the open data universe for research some archives are multidisciplinary, others restrict the field of research to specific subjects. Today to find data from different repositories of interoperable data is available Re3data.org (*Registry of Research Data Repositories*)[118] by *DataCite*, since 2012, funded by GRF (German Research Foundation)[119], in partnership with academic libraries and publishers.

The platform is among the tools recommended by the "Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020"[120].

https://www.re3data.org/

[118] Re3data.org (Registry of Research Data Repositories) http://www.re3data.org/
[119] GRF (German Research Foundation) http://www.dfg.de/
[120] Linee guida sull'accesso aperto alle pubblicazioni scientifiche e ai dati della ricerca in Orizzonte 2020 http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

In the Netherlands, the Public Technical University of the city of Delft has given a central role to the figure of the data steward in each faculty, in order to technically support the researchers, among which are the "*Data Champions*". [121]

It is a community that enables an open disciplinary approach to systemic change in collaboration with academic and civil society stakeholders.

Research management at every level is oriented towards data sharing, training and stimulating new generations of researchers to the new standards. The data management coordinator is structured in the university library, collaborating with the faculty data stewards, but also with the communication team, the PhD School, the IT technical support, the Ethics Committee for Human Research, other central offices.

SOURCE: data stewardship Addressing Disciplinary Data Management Needs [122]

The result of this synergy is the "4TU.ResearchData" portal set up in 2010 in partnership with Eindhoven and Twente universities.

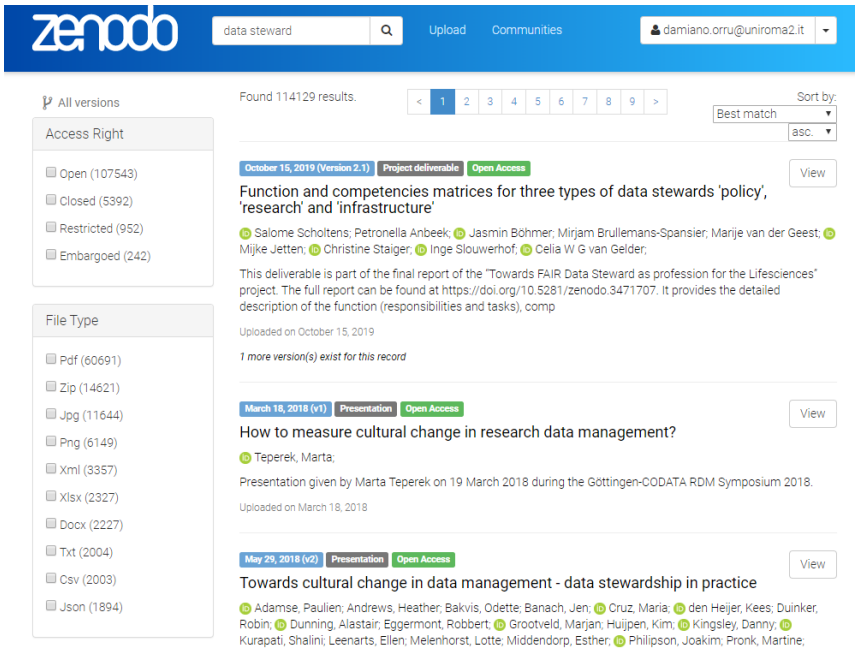
[121] Writing articles about Data Champions https://zenodo.org/record/3386060#.XkLhIVVKjIU
[122] data stewardship Addressing Disciplinary Data Management Needs http://www.ijdc.net/article/view/604

The archive allows access and preservation of research data sets, focused on open science, engineering and technology data. It started in 2010 as a service for researchers from all over the world, in an open and interoperable way. Recently, the University of Delft has launched the strategy 2020-2023[123] to define priorities for the development and expansion of services aimed at increasing the reliability of partnerships and optimizing efficiency in terms of reducing data retrieval and cleaning times.

There are several platforms that facilitate the dissemination and interoperability of research data; the best known is Zenodo, chosen in the Horizon Europe programme by the European Commission as one of the platforms for the deposit of data and publications.

The Zenodo platform was born in 2013, managed by CERN (*Conseil européen pour la recherche nucléaire*)[124] for OPENAIRE (*Open Access Infrastructure for Research in Europe*) [125] to facilitate self-archiving for researchers with open data publication rights.

Zenodo is currently a fundamental pivot of Open Science, with a particular focus on open data in Europe, without disciplinary or geographical limits, for sharing with the Network community, which also includes software, metadata and methodologies performed by researchers. The possibility of hosting closed and limited material is also offered, in order to protect collaborations on content stored securely before disclosure.



https://zenodo.org/

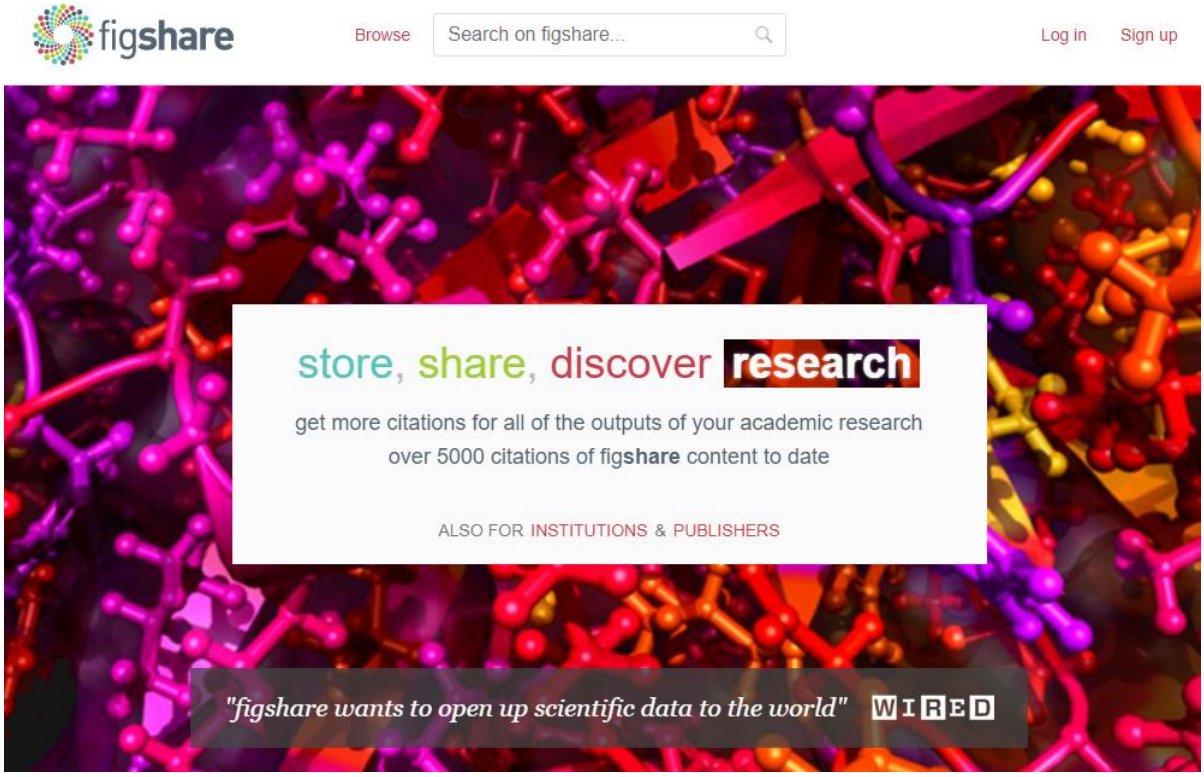
[123] Strategic Framework for data stewardship at TU Delft 2020 to 2024 https://www.tudelft.nl/en/library/current-topics/research-data-management/r/support/data-stewardship/support/strategic-framework-for-data-stewardship/

[124] CERN (Conseil européen pour la recherche nucléaire) http://home.web.cern.ch/

[125] OPENAIRE (Open Access Infrastructure for Research in Europe) http://www.openaire.eu/it

The web interface is complemented by a rich range of APIs that integrate third-party tools and services, using Zenodo as a backend in their workflow. Users are identified by Mendeley, ORCID (*Open Researcher and Contributor ID*), OpenAIRE. Each contribution is identified by a specific DOI (Digital Object Identifier), with bibliographic references that can be exported in BibTeX, DataCite, DC, EndNote, NLM, RefWorks MARC and MARCXML format. The impact measurement of the publication is offered via Altmetric. The platform allows the deposit of raw data at any stage of research, for sharing with consortium laboratories in separate experiments. Indexing is guaranteed by the standard Oai-Pmh protocol.

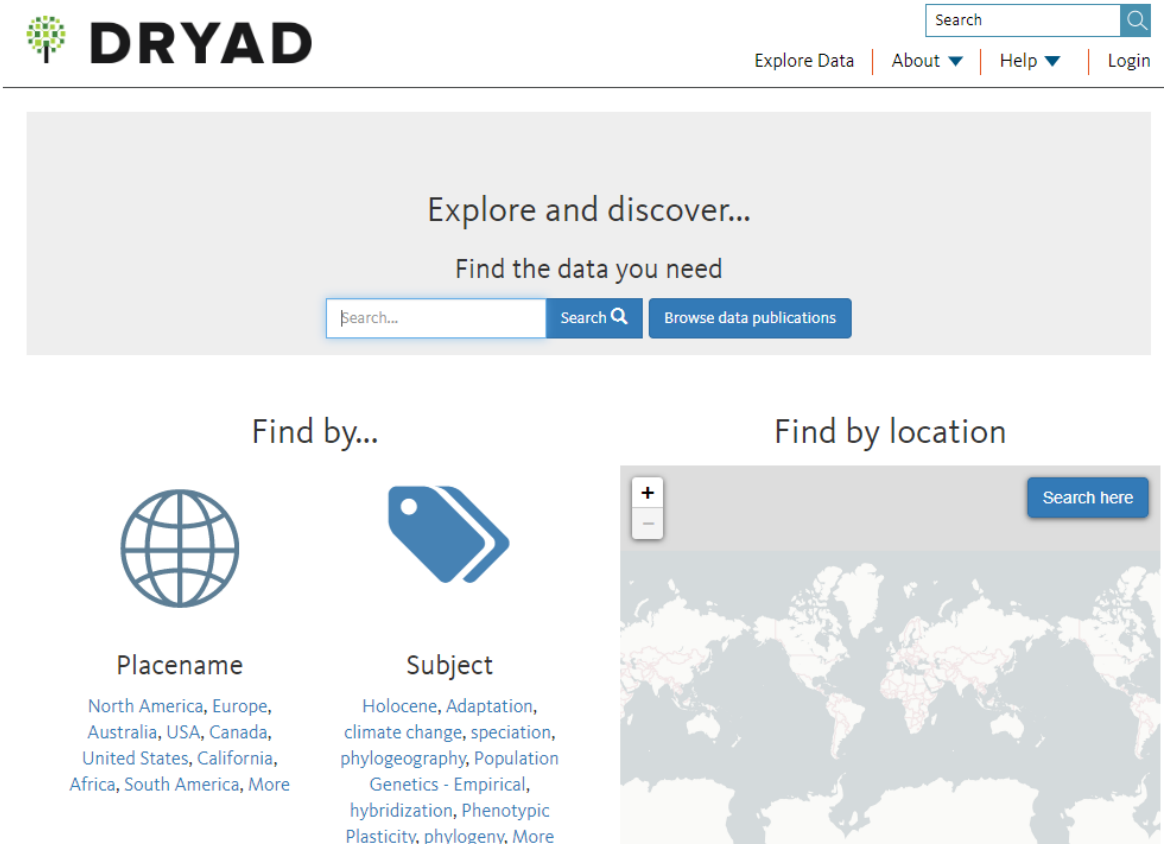
A second platform widely used to archive data is Figshare, with the possibility to upload data sets, tracking authors via ORCID and publications via DOI, with visualization and sharing statistics from data sets. The system also allows the export of quotations in RefWorks, BibTeX, Endnote, DataCite, NLM, DC and RefMan formats.



http://figshare.com/

It also offers the possibility to create and manage projects, with free storage space of up to 20 GB with 5 other employees. This portal from 2012 to March 2020 has exceeded 26,000,000 views, with 7,500,000 data accesses, 800,000 data set shares, 2,000,000 items viewed, in 500,000 collections and 5,000 projects.[126]

In the biomedical field DRYAD is widely used for free access to published data. Loading from data sets is subject to a fee, for a cost per author starting from $80.00, which includes DOI assignment, monitoring of citations of their works, limited embargo periods for their data in the peer review phase.

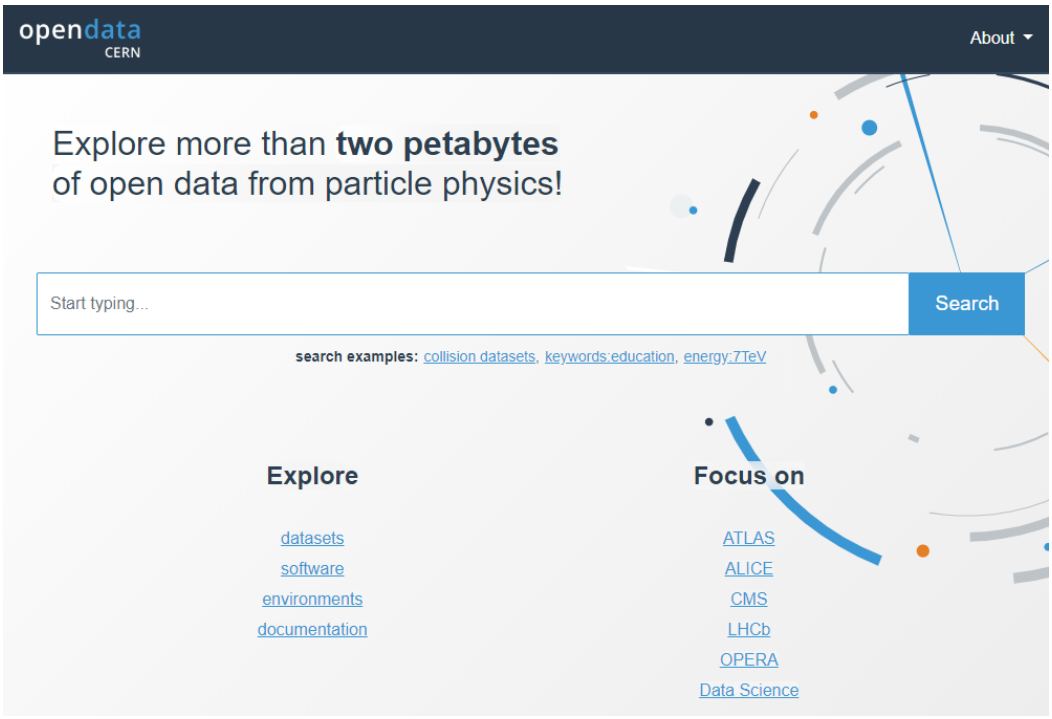


https://datadryad.org

CERN has shared on the Open Data portal most of the data produced in the various research activities, including the software and documentation necessary to understand and analyze the data shared under open licenses, identified with DOIs, with the possibility of embargo periods on some experiments.

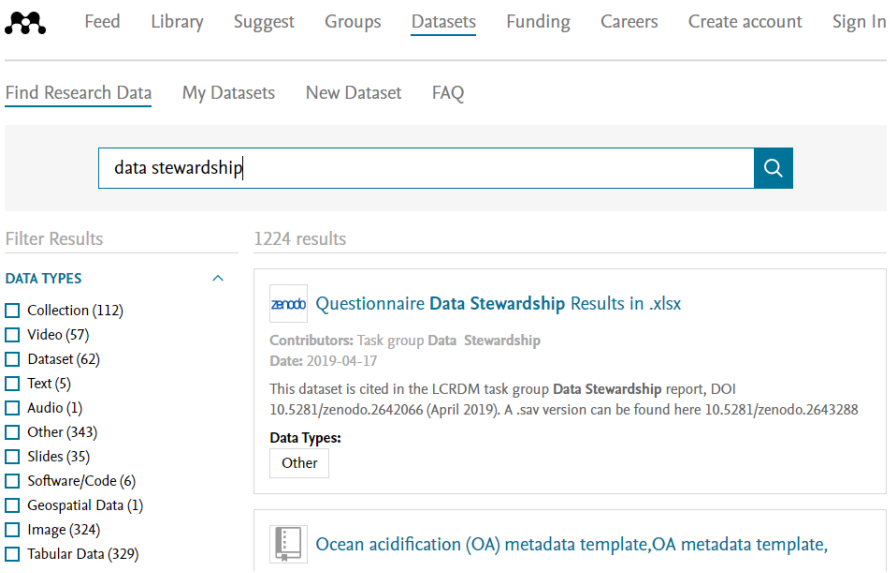
[126] Statistiche Figshare https://figshare.com/about

CERN's open data portal adheres to the following principles FORCE 11[127] for data citations, to track their use on interoperable platforms such as INSPIRE[128] to measure its impact.



http://opendata.cern.ch/

As an extension of the social network of the same name, Mendeley Data allows the sharing of open search data. The open and free repository was born in 2015, with the aim to increase the exposure of publications and monitor the use of their data.

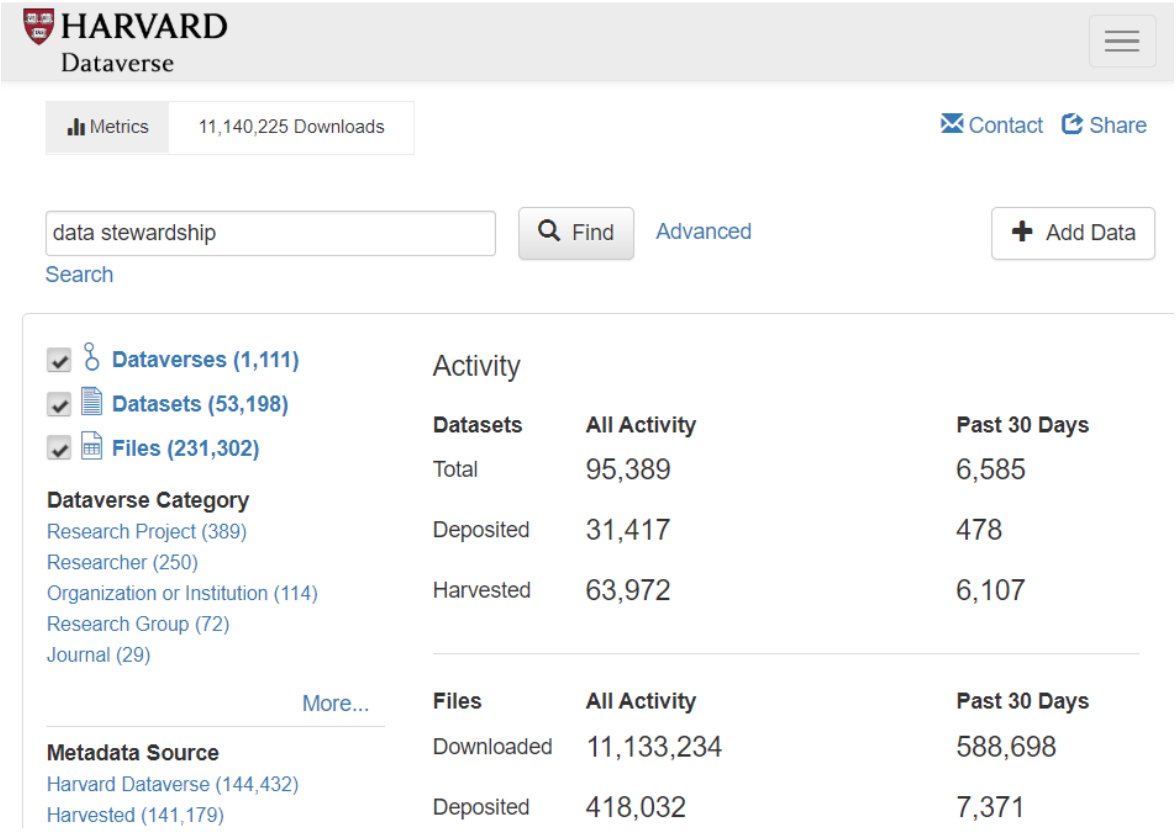


https://data.mendeley.com

[127] FORCE 11 declaration of data citation principles http://www.force11.org/datacitation
[128] INSPIRE http://inspirehep.net/?ln=en

Each product is registered through DOI (Digital Object Identifier), with indexing in OpenAIRE for perpetual preservation guaranteed by the partnership with DANS (*Data Archiving and Networked Services*)[129], including what is necessary to reproduce the research: raw and processed data, protocols and methods, software settings and algorithms. Mendeley Data has obtained CoreTrustSeal certification for the secure archiving process, as well as for data set verification by qualified auditors.

Harvard University has also created research data repositories: Dataverse.



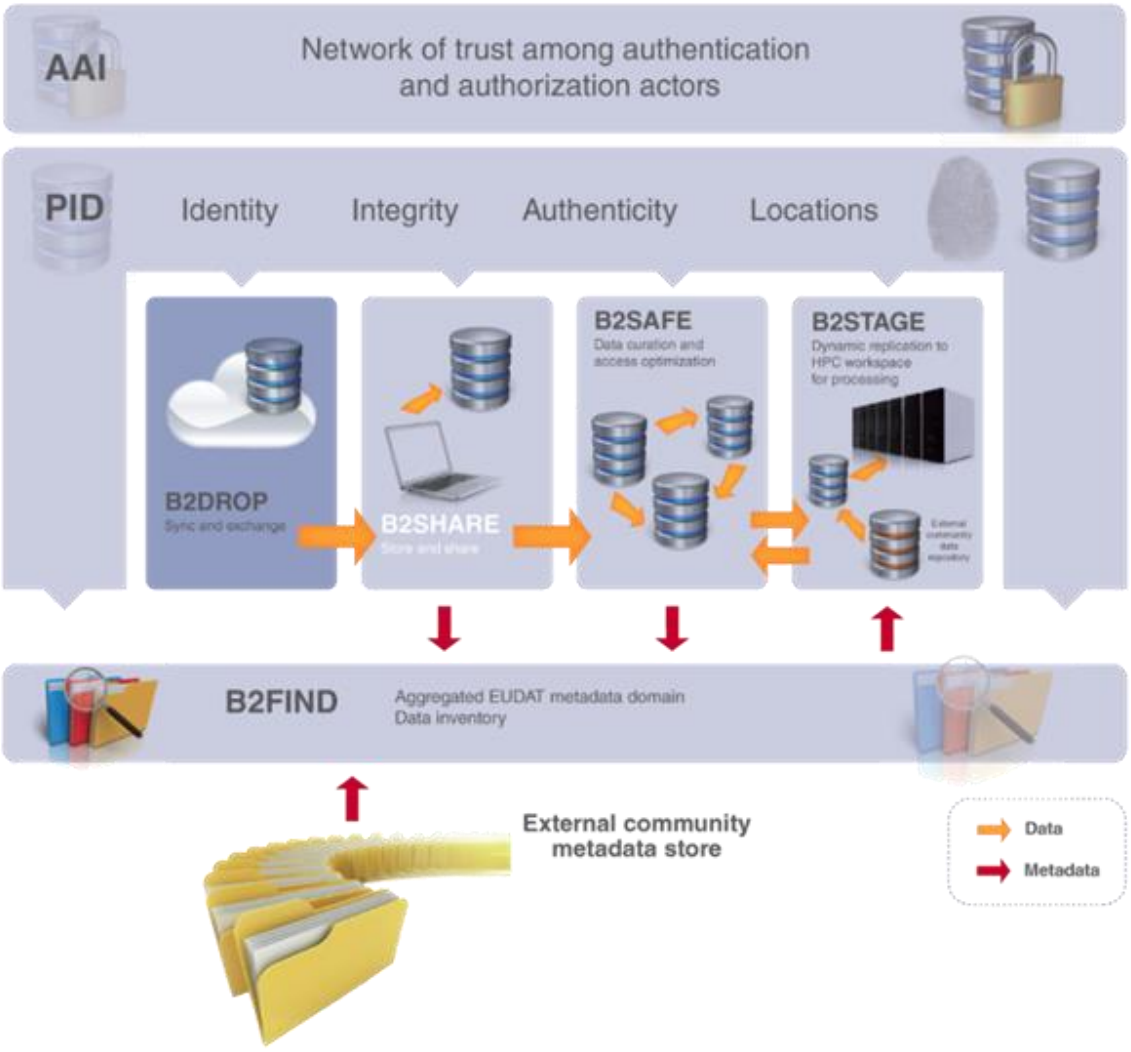
https://dataverse.harvard.edu/

The web application, developed in open source at the Institute for Quantitative Social Science, allows the upload and consultation of data from any discipline. As of March 2020, it offers access to over 95,000 datasets, 31,000 of which are locally stored, with 418,00 archived files that generated 11,133,000 downloads.

[129] DANS (Data Archiving and Networked Services) https://dans.knaw.nl/en

For doctoral candidates and researchers, the EUDAT Community is available to support access to and re-use of data, in synergy with different European realities to optimise the management of research data, developing value-added services such as:

- Data storage: B2SHARE (very large data sets) and B2DROP
- Data search: B2FIND (search based on data metadata)
- Authentication and Authorization (AAI) of users: B2ACCESS (upload/download permissions management).

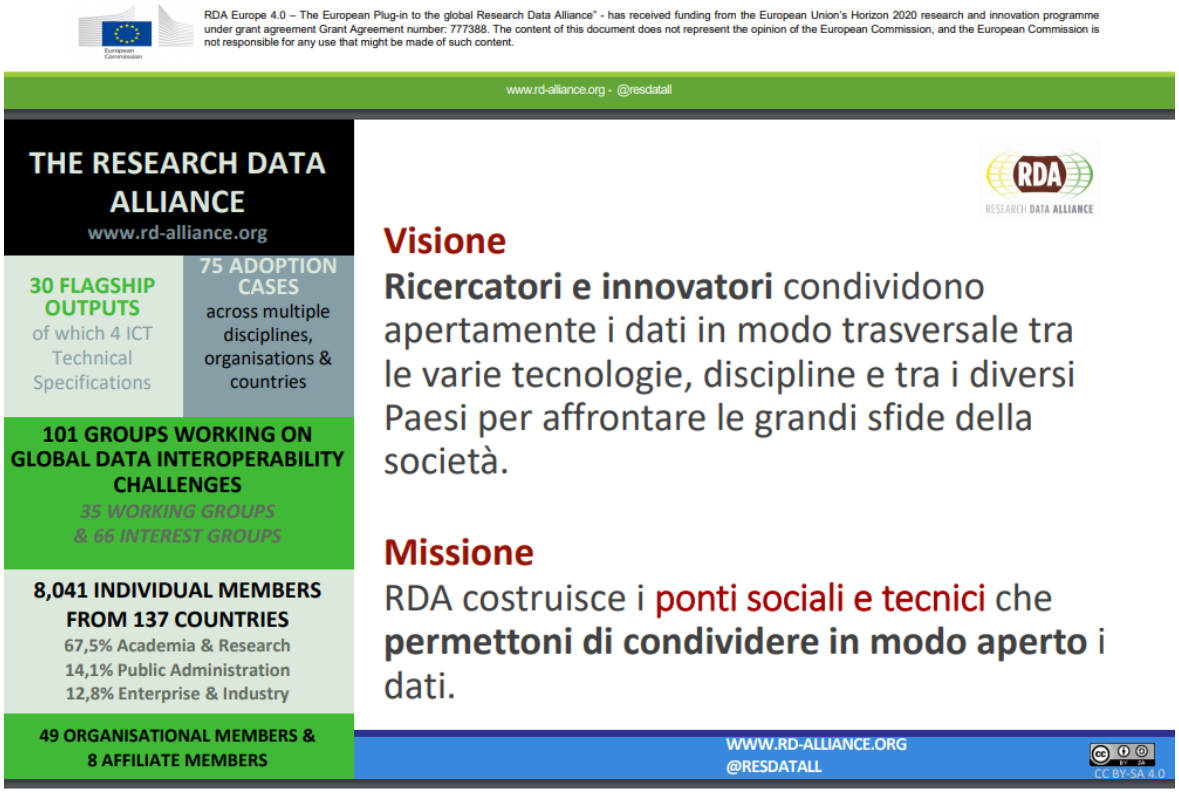


SOURCE: EUDAT B2DROP[130]

[130] EUDAT B2DROP https://www.eudat.eu/services/b2drop

The aim is to achieve a common model and service infrastructure for data management covering several European research centres and community data repositories: Collaborative Data Infrastructure (CDI), aimed at collaboration and cooperation between European generic and specific scientists and data repositories (15 countries currently).

Every citizen can start to take an interest in the research data through "*Research Data Alliance*" [131], international organisation, whose members collaborate in the development of infrastructure and activities, aimed at reducing existing barriers to data sharing and exchange, in order to accelerate data-driven innovation worldwide.



https://www.rd-alliance.org/

Today this network has more than 7,000 members from 137 countries, with the participation of researchers, scientists and data professionals from all disciplines, fields and themes.

The aim of the Alliance is to build social and technical bridges for open data sharing, through network technologies and with the support of researchers and innovators to address major societal challenges.

[131] Research Data Alliance
https://www.rd-alliance.org/professionalizing-data-stewardship-rda-community

The focus of civic engagement is confrontation through focused Working Groups and exploratory Interest Groups. Participation comes spontaneously from the community, on a voluntary basis, to exchange opinions and knowledge, obstacles and solutions, to improve and facilitate the sharing and reuse of data globally.

The international collaboration aims to provide answers to the challenges associated with open science for aspects related to reproducibility, data retention, best practices for thematic repositories, legal interoperability, data citation, data log types, metadata.

CONCLUSIONS

We all hear about IoT (Internet of things), but few know that this form of artificial intelligence is all based on data-powered algorithms. So we talk about IoD (Internet of Data) as a source for the experimental development of integrated and innovative applications in the environmental, health, energy fields, with implementations also in Smart Cities.

For a correct and transparent management of the microsystemic and macrosystemic information flow in sharing and aggregating from various sources. If these apps get the necessary data in real time and in a structured and interoperable way not based on advanced software or even a super fast 5G connection, but need data specialists, for an "ethical management" (Stewardship) aimed at new intelligent and sustainable services.

Stewardship[132] means:

"*1: the office, duties and obligations of a steward*

2: the running, supervision or management of something, in particular the careful and responsible management of something entrusted to one's own care"

According to Barend Mons, chairman of the High Level Expert Group on EOSC (European Open Science Cloud), 500,000 data stewards are needed by 2026:

"*For some jobs, you need all the data in one place. For many other jobs, even very large ones, the data is simply too big to move, even with high-speed internet, or it cannot be moved legally. We need to define the ways to do distributed computing to achieve the same results that you probably get now in large centralized computing*."[133]

Data stewardship[134] relates to:

" *1: the roles, functions and responsibilities of a given steward*

2: the systematic, sustainable and responsible management of data for the benefit of the public"

[132] Stewardship https://www.merriam-webster.com/dictionary/stewardship
[133] E-Infrastructure Reflection Group: We need 500.000 respected data stewards to operate the European Open Science Cloud https://www.youtube.com/watch?v=sJPFR7bkz4Q
[134] data stewardship https://www.merriam-webster.com/dictionary/steward

Academic research is at the service of citizens, first funders and first stakeholders of open data with social impact, pushing for reuse and collaboration in a fast and agile way with businesses and civil society.

In a vision of sustainable and ethical research, partnerships are born in which data stewards can play a leading role:

- in choosing partners to involve the community
- in defining roles and tasks within the organisation
- in defining interviews to collect data
- in the dissemination and communication of results
- in collaborations for the sustainability of projects in the long term

For success in organizations based on openness and data sharing it is critical to initiate, facilitate and proactively coordinate collaborations in a systematic, sustainable and responsible manner.

The actions that each data steward should take, in both public and private organizations, should help to bring out the intrinsic value of data, responsibly protecting the rights of individuals and businesses, ensuring interactions that maximize data reuse, while respecting regulations and contractual obligations.

Data stewards make data sustainable through the Data Management Plan (research institutions) or by articulating the Business Case (business partnerships) on data, applying measurements and indicators of social impact.

In the measurement of the Stewardship for the architecture (hardware/software server) of the repositories (data archives) there is the CoreTrustSeal certificate, an international certification that attests the technical compliance with FAIR requirements (Findable, Accessible, Interoperable, Reusable) in the verification of sustainability for long-term digital preservation.

It was analyzed how the SA8000 certification needs an update of the transversal requirements in order to decline data stewardship in companies and research institutions, integrating not only the FAIR data principles, but also the GDPR (General Data Protection Regulation), CCPA (California Consumer Protection Act), HIPAA (Health InsurancePortability and Accountability Act) of the United States, requirements such as PCI DSS (Payment Card Industry Data Security Standards).

This new horizon sees open and shared data as leverage within Agenda 2030 for targets:

- ✓ 4.6 By 2030, ensure that all youth and a substantial proportion of adults, both men and women, achieve literacy and numeracy
  (SDG 4 Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all)
- ✓ 10.2 by 2030 empower and promote the social, economic and political inclusion of all irrespective of age, sex, disability, race, ethnicity, origin, religion or economic or other status
  (SDG 10: Reduced inequalities)
- ✓ 16.6 Develop effective, accountable and transparent institutions at all levels
- ✓ 16.10 ensure public access to information and protect fundamental freedoms, in accordance with national legislation and international agreements
  (SDG 16 Peace, justice and strong institutions)
- ✓ 17.6 Enhance North-South, South-South and triangular regional and international cooperation on and access to science, technology and innovation and enhance knowledge-sharing on mutually agreed terms, including through improved coordination among existing mechanisms, in particular at the United Nations level, and through a global technology facilitation mechanism.
- ✓ 17.7 Promote the development, transfer, dissemination and diffusion of environmentally sound technologies to developing countries on favorable terms, including on concessional and preferential terms, as mutually agreed
- ✓ 17.9 Enhance international support for implementing effective and targeted capacity-building in developing countries to support national plans to implement all the Sustainable Development Goals, including through North-South, South-South and triangular cooperation
  (SDG 17: Strengthen the means of implementation and revitalize the global partnership for sustainable development)

The new Horizon Europe programme directs research towards the Sustainable Development Goals (UN Agenda 2030) in order to develop competitive infrastructures in the European Research Area.

Data stewardship can also be used in the management of quantitative and qualitative flows from the SA8000 to implement monitoring, transparency and verifiability by applying FAIR principles and privacy protection dictates.

Through an ad hoc platform it would be possible to share in real time, in a standardized and complete form, the information detected by the entire certification process. Data stewardship would become the process of data life cycle management, guaranteeing transparency, traceability and processing according to shared policies. In this way the SA8000 certification procedure could evolve into a tool to further support strategic decisions, based on certified data that allow to create shared knowledge.

BIBLIOGRAPHY

aa.vv. Open data, open curation. Sci Data 5, 180204 (2018)

Abrams, M., Abrams, J., Cullen, P., & Goldstein, L. (2019). Artificial intelligence, ethics, and enhanced data stewardship. IEEE Security & Privacy

Allen, R., & Hartland, D. (2018). FAIR in practice-Jisc report on the Findable Accessible Interoperable and Reuseable Data Principles. JISC: Bristol, UK.

Ayris, P. (2017). The importance of European collaborations to support Open Science and digital library development. Liber Quarterly

Ayris, P., & Research Data Working Group. (2013). LERU Roadmap for Research Data.

Ayris, P., López de San Román, A., Maes, K., & Labastida, I. (2018). Open science and its role in universities: A roadmap for cultural change. Leuven: LERU Office. 13/12/2019.

Baker, E. N. (2017). Data archiving and availability in an era of open science

Baker, J., Moore, C., Priego, E., Alegre, R., Cope, J., Price, L., ... & Wilson, G. (2016). Library Carpentry: software skills training for library professionals. Liber Quarterly: The Journal of European Research Libraries

Baker, K. S. (2009). data stewardship: Environmental data curation and a web of repositories. Digital Discourse: The E-volution of Scholarly Communication, 1(1).

Bangert, D. et al, Recommendations for Services in a FAIR data ecosystem.

Bianco, S., Patrizii, L. (2019), Plan S per un accesso aperto – equo e sostenibile – alla conoscenza. Bibloteche oggi trends.

Boeckhout, M., Zielhuis, G. A., & Bredenoord, A. L. (2018). The FAIR guiding principles for data stewardship: fair enough?. European journal of human genetics

Borgman, C. L. (2017). Syllabus for Data Management and Practice, Part I, Winter 2017, UCLA Information Studies

Borgman, C. L. (2018). Open data, grey data, and stewardship: Universities at the privacy frontier. Berkeley Tech. LJ

BorrellI, G., Sparano, F. (2019) Perché bisogna garantire a tutti l'accesso alle riviste scientifiche. Internazionale.

Boulton, G., Campbell, P., Collins, B., Elias, P., Hall, W., Laurie, G., ... & Walport, M. (2012). Science as an open enterprise. The Royal Society.

Bradley, F. (2016). 'A world with universal literacy' The role of libraries and access to information in the UN 2030 Agenda. IFLA journal

Casino A., Stewardship. Politiche e pratiche per una gestione etica delle risorse, Aracne, 2015

Casino A., Stewardship: la nuova sostenibilità? Aracne, 2018

Constantinescu, N. (2018). Data Librarian, the Steward. Revista Română de Biblioteconomie şi Ştiinţa Informării= Romanian Journal of Library and Information Science

Cruz, M. J., Kurapati, S., & Turkyilmaz-van der Velden, Y. (2018, October). The Role of data stewardship in Software Sustainability and Reproducibility. In 2018 IEEE 14th International Conference on e-Science (e-Science) (pp. 1-8). IEEE.

Davis J., Schoorman D. and Donaldson L., Toward a Stewardship Theory of Management, in The Academy of Management Review Vol. 22, No. 1 (Jan., 1997)

Deville, J., Sondervan, J., Stone, G., & Wennström, S. (2019). Rebels with a Cause? Supporting Library and Academic-led Open Access Publishing. LIBER Quarterly, 29(1).

European Commission, Guidelines on FAIR Data Management in Horizon 2020. 26/07/2016

Farnham, A., Kurz, C., Öztürk, M. A., Solbiati, M., Myllyntaus, O., Meekes, J., ... & Kanninen, L. (2017). Early career researchers want Open Science. Genome biology

Giovannini, E. (2018). L'utopia sostenibile. Gius. Laterza & Figli Spa.

Greco, P. (2017). Per Una scienza aperta. Scienza e Filosofia.

Hanisch, R., & Choudhury, S. (2009). The data conservancy: Building a sustainable system for interdisciplinary scientific data curation and preservation.

Hecker, S., Haklay, M., Bowser, A., Makuch, Z., & Vogel, J. (Eds.). (2018). Citizen science: innovation in open science, society and policy. UCL Press.

Higman, R., Bangert, D., & Jones, S. 2019. Three camps, one destination: the intersections of research data management, FAIR and Open

Jansen, P., van den Berg, L., van Overveld, P., & Boiten, J. W. 2019. Research data stewardship for Healthcare Professionals. In Fundamentals of Clinical Data Science

Johnston, L. R., Carlson, J., Hudson-Vitale, C., Imker, H., Kozlowski, W., Olendorf, R., & Stewart, C. (2018). How important are data curation activities to researchers? Gaps and opportunities for academic libraries.

Jones, S., Pergl, R., Hooft, R., Miksa, T., Samors, R., Ungvari, J., ... & Lee, T. (2020). Data management planning: How requirements and solutions are beginning to converge. Data Intelligence

Kavanagh J. and Michael D. Rich, Truth Decay: An Initial Exploration of the Diminishing Role of Facts and Analysis in American Public Life, RAND Corporation, 2018.

Lepore, V., & Vellani, S. (2017). FAIR Data Management: Workshop di Research Data Alliance (RDA) a Firenze. Bibelot: notizie dalle biblioteche toscane

Li, X., Nan, Z., Cheng, G., Ding, Y., Wu, L., Wang, L., ... & Zhu, Z. (2011). Toward an improved data stewardship and service for environmental and ecological science data in West China. International Journal of Digital Earth

Lugli G., Una panoramica sulla biblioteca pubblica d'oggi, Temperino Rosso, 2019.

Lyon, L. (2016). Transparency: the emerging third dimension of Open Science and Open Data. Liber quarterly, 25(4).

Mackey, T. P., & Jacobson, T. E., Metaliterate learning for the post-truth world. American Library Association, 2019

Mons, B. (2018). data stewardship for open science: Implementing FAIR principles. Chapman and Hall/CRC.

Orru, D., & Coppola, P. (2019). «# BiblioVerifica», il blog a supporto del cittadino fact-checker. Biblioteche oggi Trends, 5(2)

Orth, A. 2014. FOSTER–Open Science Training. In Open Science Days.

Parsons, M. A. (2013). Data policy. Data Science Journal, GRDI-008.

Peng, G., Ritchey, N. A., Casey, K. S., Kearns, E. J., A Privette, J. L., Saunders, D., ... & Ansari, S. (2016). Scientific Stewardship in the Open Data and Big Data Era-Roles and Responsibilities of Stewards and Other Major Product Stakeholders.

Pergl, R., Hooft, R., Suchánek, M., Knaisl, V., & Slifka, J. (2019). "data stewardship Wizard": A Tool Bringing Together Researchers, data stewards, and Data Experts around Data Management Planning. Data Science Journal

Shen, Y. (2016). Research data sharing and reuse practices of academic faculty researchers: A Study of the Virginia Tech Data Landscape. International Journal of Digital Curation

Suchánek, M., & Pergl, R. (2018). data stewardship wizard for open science. Brno, Czech Republic.

Teperek, M., Cruz, M. J., Verbakel, E., Böhmer, J. K., & Dunning, A. (2018). data stewardship–addressing disciplinary data management needs.

Verhaar, P., Schoots, F., Sesink, L., & Frederiks, F. (2017). Fostering effective data management practices at Leiden University. Liber Quarterly

Verheul, I., Imming, M., Ringersma, J., Mordant, A., van der Ploeg, J. L., & Pronk, M. (2019). data stewardship on the map: A study of tasks and roles in Dutch research institutes. LCRDM.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... & Bouwman, J. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific data.

WEBSITES

4TU.ResearchData

https://researchdata.4tu.nl/en/about-4turesearchdata/organisation/ (last visit march 15th 2020)

5 star data

https://5stardata.info/en/ (last visit march 15th 2020)

AGID Agenzia per l'Italia digitale. Programma nazionale per la cultura, la formazione e le competenze digitali, LINEE GUIDA Indicazioni strategiche e operative 2014

https://www.agid.gov.it/sites/default/files/repository_files/documenti_indirizzo/programma_nazionale_cultura_formazione_competenze_digitali_-_linee_guida_indicazioni_strategiche_operative_0.pdf (last visit march 15th 2020)

Anagrafe della Ricerca dell'Università Tor Vergata

https://art.torvergata.it/ (last visit march 15th 2020)

ASC (Aquaculture Stewardship Council)

https://www.asc-aqua.org/about-us/about-the-asc/ (last visit march 15th 2020)

Bermuda Principles

https://www.bermudaprinciples.org/ (last visit march 15th 2020)

biblioteca TU Delft

https://www.tudelft.nl/library/ (last visit march 15th 2020)

Blog MANTRA

https://mantra.edina.ac.uk/ (last visit march 15th 2020)

Budapest Open Access Initiative

http://www.budapestopenaccessinitiative.org (last visit march 15th 2020)

California Consumer Protection Act

https://www.caprivacy.org/ (last visit march 15th 2020)

CEPAA (Council of Economic Priorities Accreditation Agency)

http://www.sa-intl.org/ (last visit march 15th 2020)

CERN (Conseil européen pour la recherche nucléaire)

http://home.web.cern.ch/ (last visit march 15th 2020)

Certificazione CoreTrustSeal
https://www.coretrustseal.org/why-certification/certified-repositories/ (last visit march 15th 2020)
CSA (Canadian Standards Association)
https://www.csagroup.org/standards/ (last visit march 15th 2020)
CSQA Certificazioni
https://www.csqa.it/CSQA/Norme/Responsabilita-Sociale/SA-8000 (last visit march 15th 2020)
DANS (Data Archiving and Networked Services)
https://dans.knaw.nl/en (last visit march 15th 2020)
Data Carpentry
https://datacarpentry.org/ (last visit march 15th 2020)
Data Champions - writing articles
https://zenodo.org/record/3386060#.XkLhIVVKjIU (last visit march 15th 2020)
Data Literacy, Wikipedia
https://en.wikipedia.org/wiki/Data_literacy (last visit march 15th 2020)
Data Mining, Lorenzo Govoni
https://lorenzogovoni.com/processo-estrazione-dei-dati-data-mining/ (last visit march 15th 2020)
Data Steward per Mestiere: più Semplice che Data Analyst, più di Valore per l'Impresa!
https://www.itware.com/blog-itware/big-data-analytics-data-management/item/1311-data-steward-per-mestiere-piu-semplice-che-data-analyst-piu-di-valore-per-l-impresa (last visit march 15th 2020)
Data stewards function landscape and its stakeholders
https://zenodo.org/record/3460552#.XkBO52hKhhk (last visit march 15th 2020)
Data stewardship Addressing Disciplinary Data Management Needs
http://www.ijdc.net/article/view/604 (last visit march 15th 2020)
Data stewardship Wizard - Guida
https://docs.ds-wizard.org/en/latest/ (last visit march 15th 2020)
Data stewardship Wizard progetto internazionale a supporto di ricercatori e data stewards nel costruire Data Management Plans per la scienza aperta
https://ds-wizard.org/ (last visit march 15th 2020)
Datavault: data stewards gdpr ambassadors – infographic https://www.data-vault.co.uk/data-stewards-gdpr-ambassadors-infographic (last visit march 15th 2020)

DCAT-AP_IT
https://www.dati.gov.it/content/dcat-ap-it-v10-profilo-italiano-dcat-ap-0 (last visit march 15th 2020)
Dichiarazione di Berlino
http://openaccess.mpg.de/Berlin-Declaration (last visit march 15th 2020)
Dichiarazione di Lione sull'Accesso all'Informazione e lo Sviluppo
https://www.lyondeclaration.org/content/pages/lyon-declaration-it.pdf (last visit march 15th 2020)
Dichiarazione di Messina
http://cab.unime.it/decennale/wp-content/uploads/2014/03/Dich_MessinaITA.pdf (last visit march 15th 2020)
Dichiarazione di Praga UNESCO
http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/PragueDeclaration.pdf (last visit march 15th 2020)
Dichiarazione Universale dei Diritti Umani
https://www.ohchr.org/en/udhr/pages/Language.aspx?LangID=itn (last visit march 15th 2020)
Dichiarazione universale dei diritti umani
http://www.senato.it/documenti/repository/relazioni/libreria/fascicolo_diritti_umani.pdf (last visit march 15th 2020)
DOAB - Directory of Open Access Books
https://www.doabooks.org/doab?func=about&uiLanguage=en (last visit march 15th 2020)
DOAJ - Directory of Open Access Journals
https://doaj.org/about (last visit march 15th 2020)
DOI - Digital Object Identifier
https://www.doi.org/ (last visit march 15th 2020)
DSCC (data stewardshiphip Competence Centres)
https://www.go-fair.org/implementation-networks/overview/dscc/ (last visit march 15th 2020)
Dyché J., Polsky A., 5 Models for data stewardship
https://www.sas.com/content/dam/SAS/en_au/doc/whitepaper1/5%20Models%20Data%20Stewardship%20White%20Paper.pdf (last visit march 15th 2020)
Edinburgh Research Data
http://datablog.is.ed.ac.uk/ (last visit march 15th 2020)

E-Infrastructure Reflection Group: We need 500.000 respected data stewards to operate the European Open Science Cloud
https://www.youtube.com/watch?v=sJPFR7bkz4Q (last visit march 15th 2020)
Essentials 4 Data Support
https://datasupport.researchdata.nl/en/ (last visit march 15th 2020)
EUDAT B2DROP
https://www.eudat.eu/services/b2drop (last visit march 15th 2020)
Eurodoc Open Science Ambassador Training
http://eurodoc.net/ambassadors (last visit march 15th 2020)
European Open Science Cloud
https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud (last visit march 15th 2020)
FAIR Data Force11
https://www.force11.org/group/fairgroup/fairprinciples (last visit march 15th 2020)
FAIR Guiding Principles for scientific data management and stewardship
https://www.nature.com/articles/sdata201618 (last visit march 15th 2020)
FORCE 11 JOINT DECLARATION OF DATA CITATION PRINCIPLES - FINAL
http://www.force11.org/datacitation (last visit march 15th 2020)
FOSTER Open Science Taxonomy
https://www.fosteropenscience.eu/taxonomy/term/7?page=11 (last visit march 15th 2020)
FSC (Forest Stewardship Council)
https://it.fsc.org/it-it/certificazioni (last visit march 15th 2020)
GIS (Geographic Information System)
http://www.opengeospatial.org/ (last visit march 15th 2020)
GRF (German Research Foundation) http://www.dfg.de/
http://www.dfg.de/ (last visit march 15th 2020)
Health InsurancePortability and Accountability Act
https://www.cdc.gov/phlp/publications/topic/hipaa.html (last visit march 15th 2020)
Horizon Europe - the next research and innovation framework programme
https://ec.europa.eu/info/horizon-europe-next-research-and-innovation-framework-programme_en (last visit march 15th 2020)
Horizon2020
https://ec.europa.eu/programmes/horizon2020/en (last visit march 15th 2020)

IFLA - organizzazione internazionale che promuove le biblioteche e i servizi d'informazione nel mondo
https://www.ifla.org/publications/node/91777 (last visit march 15th 2020)
IFLA infografica How To Spot Fake News
https://www.ifla.org/publications/node/11174 (last visit march 15th 2020)
IFLA Media and Information Literacy Recommendations 2011
http://www.ifla.org/publications/ifla-media-and-information-literacy-recommendations?og=81 (last visit march 15th 2020)
INMON, Building the data Warehouse
http://fit.hcmute.edu.vn/Resources/Docs/SubDomain/fit/ThayTuan/DataWH/Bulding%20the%20Data%20Warehouse%204%20Edition.pdf (last visit march 15th 2020)
INSPIRE
http://inspirehep.net/?ln=en (last visit march 15th 2020)
IOSSG - Italian Open Science Support Group
https://sites.google.com/view/iossg/home (last visit march 15th 2020)
Knowledge Discovery in Database
https://www.techopedia.com/definition/25827/knowledge-discovery-in-databases-kdd (last visit march 15th 2020)
LIBER Response to the Amsterdam Call for Action
http://libereurope.eu/blog/2016/05/17/liber-response-amsterdam-call-action (last visit march 15th 2020)
Library Carpentry
https://librarycarpentry.org/blog/ (last visit march 15th 2020)
Licenza CC0
https://creativecommons.org/publicdomain/zero/1.0/deed.it (last visit march 15th 2020)
Linee guida sull'accesso aperto alle pubblicazioni scientifiche e ai dati della ricerca in Orizzonte 2020"
http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf (last visit march 15th 2020)
Manifesto per Information Literacy AIB
https://www.aib.it/struttura/commissioni-e-gruppi/gruppo-literacy/ilmanifesto/ (last visit march 15th 2020)
Mantra KIT
https://mantra.edina.ac.uk/libtraining.html (last visit march 15th 2020)

Metaliteracy Goals and Learning Objectives
https://metaliteracy.org/learning-objectives/goals-and-learning-objectives-translated/ (last visit march 15th 2020)
MOOC "Empowering Yourself in a Post-Truth World"
https://www.coursera.org/learn/empowering-yourself-post-truth-worldgestitodaTomMackey,Professorealdipartimento"ArtsandMedia"pressoUniversityofNewYork-TrudiE.Jacobson,Bibliotecariaresponsabiledel"InformationLiteracyDepartmentattheUniversityLibraries"pressoUniversityofNewYork (last visit march 15th 2020)
MOOC Open Science: Sharing Your Research with the World
https://www.edx.org/course/open-science-sharing-your-research-with-the-world (last visit march 15th 2020)
National network research data management
https://www.lcrdm.nl/en (last visit march 15th 2020)
OECD Principles and Guidelines for Access to Research Data from Public Funding
http://www.oecd.org/sti/inno/38500813.pdf (last visit march 15th 2020)
Open Data Handbook
http://opendatahandbook.org/guide/it/ (last visit march 15th 2020)
Open Data Institute
http://theodi.org/ (last visit march 15th 2020)
Open Refine
http://openrefine.org/ (last visit march 15th 2020)
Open Science “Wheel” - Open Science Monitor 2015
http://ec.europa.eu/research/openscience/index.cfm?pg=home&section=monitor (last visit march 15th 2020)
Open Science Community Utrecht
https://openscience-utrecht.com/what-is-open-science/ (last visit march 15th 2020)
Open Science MOOC https://opensciencemooc.eu/ (last visit march 15th 2020)
Open Science: Sharing Your Research with the World
https://www.edx.org/course/open-science-sharing-your-research-with-the-world (last visit march 15th 2020)
Open Source Definition
https://opensource.org/osd-annotated (last visit march 15th 2020)
OPENAIRE (Open Access Infrastructure for Research in Europe)
http://www.openaire.eu/it (last visit march 15th 2020)

ORCID (Open Researcher and Contributor ID)
https://orcid.org/ (last visit march 15th 2020)
OSPP (Open Science Policy Platform)
https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-policy-platform (last visit march 15th 2020)
PA digitale, nuova gara "data management" con 222 mln
https://www.01net.it/consip-gara-data-management/ (last visit march 15th 2020)
Payment Card Industry Data Security Standards
https://www.pcisecuritystandards.org/pci_security/ (last visit march 15th 2020)
PDND (Piattaforma Digitale Nazionale Dati)
https://innovazione.gov.it/it/progetti/pdnd/ (last visit march 15th 2020)
PEFC (Programme for Endorsement of Forest Certification)
https://www.pefc.org/discover-pefc/what-is-pefc (last visit march 15th 2020)
Performance Indicator Annex (edizione 2014)
http://www.sa-intl.org/_data/n_0001/resources/live/Performance%20Indicator%20Annex_11_17_2014.pdf (last visit march 15th 2020)
Phaidra - Sistema bibliotecario dell'Università di Padova
https://bibliotecadigitale.cab.unipd.it/news/phaidra.-ultime-novita (last visit march 15th 2020)
Principi dei dati FAIR – Commissione Europea
https://ec.europa.eu/info/sites/info/files/turning_fair_into_reality_0.pdf (last visit march 15th 2020)
Principi dei dati FAIR – OpenAIRE
https://www.openaire.eu/how-to-make-your-data-fair (last visit march 15th 2020)
R software
https://cran.r-project.org/ (last visit march 15th 2020)
Re3data.org (Registry of Research Data Repositories)
http://www.re3data.org/ (last visit march 15th 2020)
Requisiti del regolamento GDPR per gli open data
https://www.europeandataportal.eu/it/highlights/protecting-data-and-opening-data (last visit march 15th 2020)
Research Data Alliance
https://www.rd-alliance.org (last visit march 15th 2020)

Research Data Management and Sharing
https://www.coursera.org/learn/data-management (last visit march 15th 2020)
Responsabilita' sociale e etica
http://www.sa8000.info/ (last visit march 15th 2020)
Rstudio
https://www.rstudio.com/products/rstudio/ (last visit march 15th 2020)
Scienza in Rete: Citizen Science
https://www.scienzainrete.it/articolo/citizen-science-scienza-di-tutti/valentina-meschia/2016-03-10 (last visit march 15th 2020)
Sci-Hub profilo twitter
https://twitter.com/sci_hub (last visit march 15th 2020)
SciStarter Real science we can do together
https://scistarter.org (last visit march 15th 2020)
SFI (Sustainable Forestry Initiative)
https://www.sfiprogram.org/sfistandards/ (last visit march 15th 2020)
Social Accountability Accreditation Services (SAAS)
http://www.saasaccreditation.org/ (last visit march 15th 2020)
Social Accountability Accreditation Services statistiche al 3/3/2020
http://www.saasaccreditation.org/certfacilitieslist (last visit march 15th 2020)
Social Accountability International (SAI)
http://www.sa-intl.org/ (last visit march 15th 2020)
Software Carpentry
https://software-carpentry.org/ (last visit march 15th 2020)
SPARC - Scholarly Publishing and Academic Resources Coalition
https://www.sparcopen.org/ (last visit march 15th 2020)
Statistiche Figshare
https://figshare.com/about (last visit march 15th 2020)
Strategia digitale 2025, IA e Open data: come vincere l'inerzia dell'Italia
https://www.agendadigitale.eu/cittadinanza-digitale/strategia-digitale-2025-ia-e-open-data-come-vincere-linerzia-dellitalia/ (last visit march 15th 2020)
Sustainable Development Goals (Agenda 2030)
https://sustainabledevelopment.un.org/?menu=1300 (last visit march 15th 2020)
Transition towards an Open Science system del 27 maggio 2016
http://data.consilium.europa.eu/doc/document/ST-9526-2016-INIT/en/pdf (last visit march 15th 2020)

TU Delft Strategic Framework for data stewardship 2020 to 2024
https://www.tudelft.nl/en/library/current-topics/research-data-management/r/support/data-stewardship/support/strategic-framework-for-data-stewardship/ (last visit march 15th 2020)
Wikidata: la soluzione wikimediana ai linked open data
https://aibstudi.aib.it/article/view/11434/10655 (last visit march 15th 2020)
Zenodo (Progetto OpenAIRE)
https://zenodo.org/communities/openaire/?page=1&size=20 (last visit march 15th 2020)