Connecting Data & People in a Changing Landscape

CESSDA Strategy
2023-2027

cessda

Table of Contents

Preamble

Europe faces deep societal, ecological, and economic challenges, first aggravated by the coronavirus crisis and now by the Russian invasion of Ukraine. State-of-the-art digital technologies and research infrastructures, like CESSDA, remain crucial to support social scientists in their quest to provide answers to global challenges. It is, therefore, vital that national service providers continue to do what they do best: provide metadata expertise, develop standards and training to enable interoperability and make data FAIR, support data deposits, enable access and reuse of data, therefore ensuring that digital assets retain their value with time.

Vision

CESSDA is a renowned, sustainable, and trusted European social science infrastructure, that enables excellent services to its members, research for all social sciences disciplines, and beyond for societal benefit.

Mission

1. The mission of CESSDA, as outlined in the statutes, is:
   - To provide a distributed and sustainable research infrastructure.
   - To enable the research community to conduct high-quality research in the social sciences, contribute to the production of effective solutions to the major challenges facing society today and tomorrow.
2. To facilitate teaching and learning in the social sciences.

Five-Year Vision (2023-2027)

In the next five years CESSDA will enable long-term access to documented and curated high-quality social science data to researchers. CESSDA will build upon its **data** and knowledge assets and engage with **people** and the surrounding policy and research data **landscape** for the benefit of research and society.

Planning & Consultation

CESSDA's strategic planning aims to support and advise the SSH community and the Service Providers in an increasingly challenging and fast-moving environment. It is an opportunity and a must for social sciences to participate in the shaping of the European and global research landscape. The strategy for the period 2023-2027 takes stock of the achievements of CESSDA as an ERIC since its creation in 2017 and the rapid evolution of the role of data in modern society in recent years.

Fig.1 Map with member countries

The new strategy looks ahead to future developments in the research landscape, the European Research Area (ERA) at the global level. The analysis lays the foundation to anticipate and provide services to the research community and stakeholders from academia, industry, and policy.

The CESSDA Strategy was co-designed by the Director, Dr Bonnie Wolff-Boenisch, and all CESSDA's constituencies. These include the Strategic Committee, the Service Provider Forum and Directors, the General Assembly and the Chairs, the CESSDA Main Office staff and the Chairs of the working groups 'Tools', 'Training', 'Trust' and 'Widening and Outreach'.

CESSDA's Unique Expertise for its Services Providers and the Researchers. Furthering Interconnectivity in an Increasingly Digital Age.

Key values of CESSDA ERIC are to provide interconnectivity to its Service Providers by:

- Sharing knowledge and providing learning opportunities on technical, operational, scientific, and educational, and political matters.
- Understanding the diversity of the European landscape and connecting with other research infrastructures in SSH and other disciplines at European (see By-COVID on page 7) and international level.
- Speaking with one voice on behalf of its members.

The alignment of knowledge, practice, and support across CESSDA directly benefits researchers by providing them with a pan-European framework with data and metadata. The cooperation and knowledge exchange enabled by CESSDA helps Service Providers to offer their data depositors a strong level of consistent and efficient service. CESSDA is a 'hub' space where best practice and guidance can be shared.

CESSDA's collective expertise in managing, curating and archiving research data provides social scientists with the best conditions to undertake excellent research to provide answers to societal questions. CESSDA is an important partner for a wide range of social science communities and increasingly other research communities such as from the humanities, life and environmental sciences that either use CESSDA's tools and services or access information via their national data service. CESSDA Secretariat (Main Office) and its consortium members, the Service Providers, are also increasingly solicited by third parties that either lack the knowledge or require advice on how to preserve and harness data.

CESSDA Today

Since the `start-up' phase in 2017 as an European Research Infrastructure Consortium (ERIC), CESSDA has been transitioning into a more mature European research infrastructure organisation with 22 member countries.

In 2022 a Memorandum of Understanding (MoU) between the SSH Landmarks (SSH ESFRI ERICs) was signed[1]. This MoU attracts emerging projects and communities highly interested in formally entering CESSDA. The SSH thematic communities cover topics such as migration, elections, historical financial data, religious studies, family studies, and youth.

Through SSHOC, CESSDA is involved in the collaboration between the five ESFRI Clusters (SSHOC, ENVRI-FAIR for environmental research, PaNOSC for multidisciplinary scientific analysis, ESCAPE for astronomy and particle physics and EOSC-Life for life sciences providing a gathering point for various ESFRI projects and landmarks to connect to the EOSC.

---

1. www.sshopencloud.eu/news/sshoc-signs-mou-bolster-cloud-services-social-science-and-humanities

CESSDA is also engaged in EC-projects and pilots that address data handling matters with non-SSH communities for example in the eRimote project discussing best practices in remote solutions for digital and remote service provision across RI domains with stakeholders from the physics, life sciences and geosciences), in EOSC Future projects (e.g. working on data related to political and social trust, health and health inequality, attitudes towards climate change and energy) or in the BY-COVID project (made up of a wide range of partners from the life sciences.[2][3]

BY-COVID - an Interdisciplinary Approach

In an unprecedented interdisciplinary effort, BeYond-COVID (BY-COVID)[4] brings together 53 partners from 19 countries from the biomedical field, hospitals, public health, social sciences, and humanities.

The project aims at integrating open data on SARS-CoV-2 and other infectious diseases from national and European infrastructures such as ELIXIR, BBMRI, ECRIN, PHIRI and CESSDA.

BY-COVID will enable federated data analysis compliant with data protection regulations, harmonise and manage metadata and sample identifiers, and facilitate long-term cataloguing to ensure interoperability of national and global efforts.

The project will enable the linking of FAIR data and metadata on SARS-CoV-2 and COVID-19, other infectious diseases, and related data, and ultimately increase the potential for collaboration and exploitation of data.

BY-COVID

The BeYond-COVID project aims to make **COVID-19 data** accessible to everyone, scientists in laboratories, medical staff in hospitals or government officials.

Going beyond SARS-CoV-2 data, the project will provide a framework for making data from other infectious diseases **open and accessible** to everyone, and build a bridge between public health, clinical sciences, social sciences, and molecular sciences.
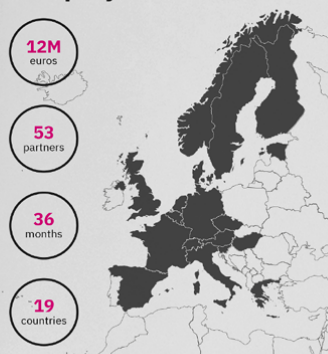
Subscribe to the newsletter

by-covid.org

The project in numbers

12M euros

53 partners

36 months

19 countries

The challenge

The world has generated vast amounts of data in response to the COVID-19 pandemic. This data comes from many different sources, and identifying, connecting and integrating these for effective analysis remains a challenge.

Public Health Laboratories | Hospital | Medical centres | Consortia | Research Infrastructures | Universities | Research Centres

Data from SARS-CoV-2, other pathogens, their hosts, and the associated diseases come from many different sources.

What we are doing

1. Mobilising data
2. Connecting data
3. Standardising data
4. Exposing and analysing data

COVID-19 Data Portal

The outcomes

Make SARS-CoV-2 and other infectious disease data easier to access, aggregate and analyse to:

- Enable **scientists in academia and industry** to respond faster to new SARS-CoV-2 strains or other pathogens causing infectious diseases.
- Help **policy-makers** assess the impact of COVID-19 and take appropriate measures to protect people.
- Facilitate **open innovation** in academia and industry across disciplines.

BY-COVID receives funding from the European Union's Horizon Europe Research and Innovation Programme under grant agreement number 101046203.

2. https://erimote.eu/home
3. https://eoscfuture.eu/data/climate-neutral-and-smart-cities/
4. https://by-covid.org/

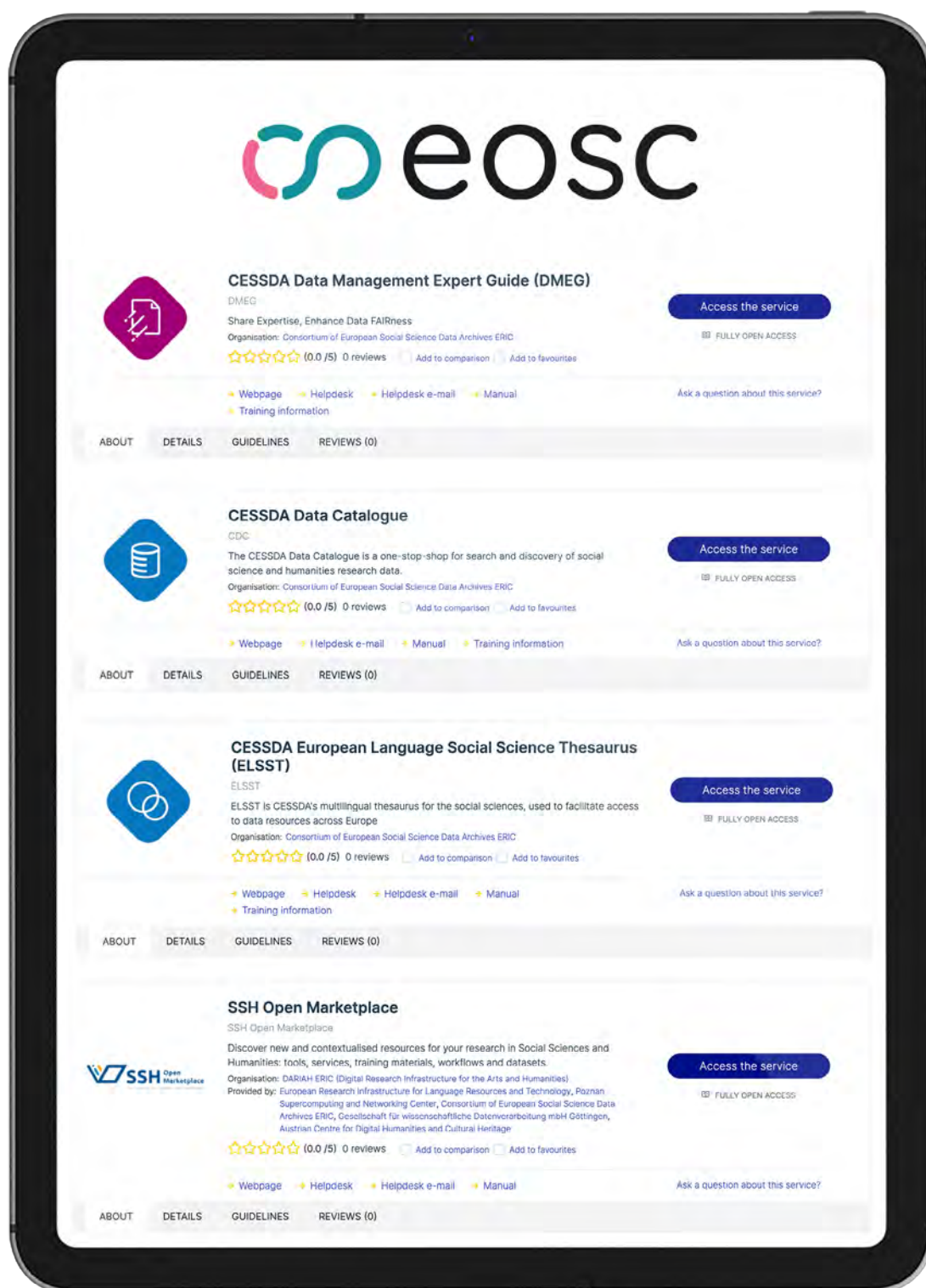Fig.3 CESSDA's EOSC-onboarded services

CESSDA is also a member of the EOSC association, strongly engaged in different EOSC projects and task forces and recognised as an early adopter of EOSC research infrastructure. Currently three of CESSDA's services have been onboarded by EOSC: The CESSDA Data Catalogue, the Data Management Expert Catalogue (DMEG) and the European Language Social Sciences Thesaurus (ELSST).

Fig.4 SSHOC League of Data

CESSDA Achievements for the Benefit of Research:

- The adoption of and contribution to best practice requirements in trustworthy digital repositories (CoreTrustSeal) to create compatibility and cooperation within CESSDA's Service Providers and to increase the quality of data and services for the research community.[5]

- The establishment of tools (e.g. CESSDA Training, DMEG) to increase the culture of metadata sharing and responsible data management within the social science community to enhance the quality, efficiency, and transparency of research.[6]

- The enhancement of collaborative cooperation between Service Providers across Europe and beyond CESSDA to increase coverage of European social science data resources.

- The collaborative development of effective tools, policies, and information resources for the support of archiving and sharing social science data in Europe.

---

5. https://www.coretrustseal.org/
6. https://www.cessda.eu/Training/DMEG.

Fig.5 A word cloud generated from CESSDA Data Catalogue topics

Technical & Data Management Achievements

- The CESSDA Data Catalogue, one of the central services offered through the CESSDA cloud platform, today enables access to about 35,000 social sciences data sets from the Service Providers across Europe, bridging national borders.

- The coordination and standardisation efforts within CESSDA to make such a large-scale data collection possible involved the Service Providers. The key enablers for metadata exchange are the DDI metadata standards and the use of multilingual controlled vocabularies and thesauri. Both were adopted across the CESSDA consortium to ensure metadata can be automatically exchanged and updated as necessary.

- A wider standardisation effort, specifically in the SSHOC project, has been undertaken to enhance the interoperability across the social sciences and humanities by connecting existing services and standards, now provided through a common SSH Open Marketplace.

- CESSDA together with CLARIN, DARIAH and OPERAS, developed the EURISE Technical Reference. All four research infrastructures implemented this as their underlying standard for the quality assurance of all technical development and to ensure the software components remain sustainable and adaptable to future needs of the users.

With the knowledge gained through the harmonisation work, CESSDA can contribute its metadata, software, and standards expertise and deliver interoperability within EOSC, for SSH and across disciplines.

CESSDA's Environment

Social Sciences & Service Providers in a Changing World

Among its political agenda the EC and the 27 Member States have emphasised the importance of state-of-the-art digital technologies and research infrastructures, the importance of academic freedom, the need for resilience of the economy and the society, and the acceleration of the green/digital transition.

National Service Providers and European research infrastructures such as CESSDA ERIC play a critical role by supporting social scientists in addressing societal challenges by:

- Facilitating easy, trusted and seamless access to infrastructures to scientists, who then can provide advice to policy makers to cope with and prepare for crises (e.g., health, climate, war, migration) via their research outcomes.
- Improved cooperation among members and Member States, and European research infrastructures through data sharing (e.g. via the COVID-19 platform).
- Identifying opportunities for new governance and funding mechanisms (e.g. Horizon Europe Missions).
- Equipping scientists with the skills they need in a fast-changing global world (training and data literacy).

The European Research Area and the Role of Research Infrastructures

The European Research Area (ERA) aims at building a common scientific and technological area for the EU in which researchers, scientific knowledge, innovation (and data to a certain extent) move freely among its members.
Among many initiatives and actions, the Open Science Initiative, the European Strategy Forum on Research Infrastructures (ESFRI), the development and implementation of European Research Infrastructures across all fields of science and the European Open Science Cloud (EOSC), a web of FAIR data and services for research data in Europe, are seen as major contributors to this vision.[7]

CESSDA as an ERIC contributes directly to the evolution and implementation of ERA.

---

7. https://research-and-innovation.ec.europa.eu/system/files/2021-11/ec_rtd_era-policy-agenda-2021.pdf

Changing Research Landscape

The digital transformation and increasing interdisciplinary collaborations cause changes in the research system, and the research data life cycle. Such changes are manifested for example by the open science and the open access to publication movement.

The access and sharing of (research) data and open-source software and the blending of new forms of data has become a focus of political science discussions, due to the potential for digital innovation in a near future (research) data economy and provision of solutions for societal challenges.

New forms of research assessments and how to incentivise a collaborative research culture to leverage the full potential of research data have been discussed at Member State level (Action 3 of the ERA policy agenda 2022-2024: 'Reform the Assessment System for research, researchers, and institutions') and resulted in the Coalition for Advancing Research Assessment (CoARA).[8][9] The vision of CoARA is that assessment of research, researchers and research organisations recognises the diverse outputs, practices and activities that maximise the quality and impact of research.

Linked to this initiative, social scientists who use CESSDA's tools and service via their national service providers, may ask CESSDA or their national providers to validate or certify demonstration of data sharing and responsible data management efforts for research assessment.

Trusted Environments

CESSDA's Service Providers offer trustworthy repositories, certified with CoreTrustSeal, to ensure the long-term sustainability of data deposited with them and fulfilling the requirements of the EC, for Horizon Europe activities[10]. Simultaneously, the Service Providers provide both technical and legal expertise to process and store sensitive and restricted data in compliant and secure environments.

8. https://coara.eu/
9. https://research-and-innovation.ec.europa.eu/system/files/2021-11/ec_rtd_era-policy-agenda-2021.pdf
10. https://www.coretrustseal.org

Fairification of Research

The requirements to make both metadata (including information about data access) and data more interoperable and understandable by machines will grow to harness new forms or combine existing data in innovative ways.

The role of domain data services in this process is irreplaceable because they increase the understanding of data, make data `FAIR'er and thus create the conditions for effective data reuse. The increasing interoperability will make RIs more resilient and sustainable, hence more cost-efficient. It will also pave the way for AI-based applications.

At the same time SSH data may include sensitive data that must be appropriately protected through information security and access management measures.

There is a lot of potential for CESSDA collectively (MO and Service Providers) to work together for synergies and cost efficiency, instead of Service Providers working individually.

Interoperability Across Disciplines

Through the SSH Cluster 'SSHOC', CESSDA and the Service Providers have been crucial in shaping the design of EOSC, and indirectly the ERA Policy Agenda 2022-2024 where working to 'Enable Open Science, including through the European Open Science Cloud (EOSC)' is listed as the first top priority of the Council conclusions of the ERA Governance.[11]

Together with the other four Science Clusters, CESSDA contributed to the refinement and transformation of the EOSC services towards specific interdisciplinary services and platforms for research infrastructures and for scientists and (potentially) EU member state institutions.[12]

This was done by enriching existing services with additional information, tailored for scientific use, translating EOSC services into more specific disciplinary science services.

11. https://research-and-innovation.ec.europa.eu/system/files/2021-11/ec_rtd_era-policy-agenda-2021.pdf

12. https://sshopencloud.eu/news/33-shocing-key-exploitable-results-support-your-research-ssh-and-beyond

*CESSDA will build upon its **data** and knowledge assets and engage with **people** and the surrounding policy and research data **landscape** for the benefit of research and society.*

CESSDA General Assembly

Priority 1
Seek to Further the Interoperability of Metadata and Data

Goal 1 - Enable Data Discovery

Documentation and long-term availability of data are key in ensuring reusability of research. With changing data types relevant to researchers, including social media data, the needs for their description and provision are changing. CESSDA will address researcher's needs and support the crossing of boundaries of scientific domains and languages by exploring ways of furthering the interoperability of metadata and data.

By 2027 CESSDA will be:
- The leading European social sciences research infrastructure following the FAIR data principles, enabling data discovery and reuse (including reproducibility of research outcomes).[13]
- A forerunner of new and additional SSH and cross-domain data sources including social media and statistical data and data from other SSH disciplines, and/or other domains (e.g. behavioural data).

How?
- Provide input to best practice in FAIR data principles and skills.
- Addressing how survey and statistical data may be best aligned with social media, and behavioural data or new SSH and cross-domain outcomes.
- Refine and implement crosswalks (i.e. translation of one into another meta standard) to enlarge user audiences and incorporate new data archives.
- Engage with publishers and funders (e.g. on data citation requirements by the provision of new forms of outputs (digital objects) and incentivisation of scientists for the provision of metadata.
- Tackle legal, ethical and procedural challenges (i.e. confidentiality, GDPR, copyright) in a data sharing context.
- Align with the European Open Science Agenda and the EOSC.

Goal 2 - Working on the Interoperability of Services

Services provided to researchers and shared between infrastructure providers should be transparent, trustworthy and accessible to both humans and machines. Interoperability at the technical, data and metadata levels provides for seamless transfers between service providers, including the EOSC and towards the researchers. Standards must be discussed, defined, implemented, communicated, and maintained.

By 2027 CESSDA will:
- Be a default provider of professional data archiving solutions for social sciences (i.e. trusted repositories) for long-term preservation.
- Explore the provision of collaborative analysis tools (for scientists) from different research disciplines and/or processing of (e.g. restricted, commercial limited) data.
- Be a forerunner in professional and legal expertise regarding sensitive data in the social sciences and in concert with research infrastructures (e.g. from the health and biomedical domain, humanities etc.) dealing with sensitive data, too.
- Improve the technology readiness levels for selected CESSDA products and potential new services.

How?
- Continue to explore and address the interoperability of services to ensure that standards and interfaces are available to directly connect data, processing services and publication and archiving platforms across Europe.
- Explore new technologies where appropriate such as AI (enhancing data and metadata) to enable researchers to perform large-scale analysis using cloud computing resources.

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

Priority 2
Connecting People

Goal 3 - Facilitate Mutual Exchange among CESSDA's Service Providers and Support within their National Landscape

The service providers and the MO that form CESSDA encapsulate an immense pool of knowledge. To leverage this resource CESSDA must mobilise and combine its internal resources and services and increase their visibility to guide the social scientist community through the ongoing digital and societal transformation.

By 2027 CESSDA will:

- Strengthen relationships within CESSDA.
- Leverage the common pool of knowledge among Service Providers and GA members.
- Support Service Providers in promoting their work and the work of researchers in their national context.

How?

- Develop mechanisms and opportunities for enhancing mutual understanding among CESSDA and the different needs.
- Diversify services for different types of CESSDA's Service Providers and their needs (e.g. provide different learning and capacity building opportunities).
- Explore different formats to channel information from the European Research area to CESSDA members (e.g. provide regular reports from European events).
- Support consistent national messaging for impact and efficiency (e.g. on deposit in disciplinary TDRs or what is the added-value of CESSDA for the Service Providers and the Member States).

Goal 4 - Reach Out to the SSH Researcher at National, European and Global Level

CESSDA expertise in managing research data is a valuable and sought-after asset. This equips CESSDA to serve other communities in the social sciences domains.

By 2027 CESSDA will:

- Enlarge the pool of users from the research community to enrich CESSDA's thematic catalogue and services.
- Establish strategic partnerships within the SSH community and beyond.

How?

- Identify and reach out to new social sciences researchers' communities and beyond that seek support and expertise in research data management.
- Identify and develop together with social sciences research communities innovative science driven types of services. Explore how services of service providers can be pooled and leveraged for researchers and public good.
- Co-develop pilots with other research communities from or outside the SSH community with common needs and interests to address societal goals (e.g. EC mission-type or and/or SDGs related interdisciplinary projects).
- Provide training in research data management.

Priority 3
Shaping the European Research Arena

Goal 5 - Establishing Strategic Partnership with ERA Actors

CESSDA with all its Service Providers will continue to build a strong interlinked science-driven infrastructure for social science data for the benefit of the social sciences research communities and beyond.

By 2027 CESSDA will:

- Shape the European landscape through participation in relevant science policy arenas and thematically linked international and global organisations, for example CODATA, the Committee on Data of the International Science Council and the WorldFAIR project[14][15].

How?

- Continue to collaborate with other RIs (domain and cross) on topics of mutual interests such as the development and alignment of standards (e.g. CoreTrustSeal, Data Documentation Initiative)[16][17].
- Support the alignment of research infrastructures at European and national level.
- Provide data for policy advice.

Goal 6 - Establishing Strategic Partnership for the Benefit of Society

CESSDA on behalf of its members will engage with national, European and global organisations for raising its external visibility, for mutual learning and for alignment of standards and policies (e.g. EOSC, CODATA, ADA in Australia, ICPSR in the US, SSJDA in Japan and other worldwide data archives)[18].

By 2027 CESSDA will:

- Be a visible part of the data-led/driven global landscape.

How?

- Engage data-driven umbrella organisations in Europe and beyond for mutual learning, alignment of standards and policies.
- Organise joint sessions in selected European and national workshops.
- Represent CESSDA or position CESSDA members in strategic outlets.

---

14. https://codata.org
15. https://codata.org/worldfair-global-cooperation-on-fair-data-policy-and-practice-kick-off-meeting-introduces-major-new -initiative-to-advance-implementation-of-the-fair-data-principles
16. https://www.coretrustseal.org
17. https://ddialliance.org
18. https://ukdataservice.ac.uk/help/other-data-providers/data-archives/other-worldwide-data-archives

CESSDA Strategy

Implementing the Strategy

This CESSDA strategy begins in 2023 and runs through to the end of 2027. Granular identification of needs and teams, alongside more detailed engagement and planning already started in 2023. The implementation of the work plan will take place from 2024 to 2026 before the results are consolidated and refined.

Conclusion

The new strategy captures the latest developments in the research data landscape and in the European Research Area (ERA), such as the digital transformation, the increasing cross-fertilisation and FAIRification of research and the way it is conducted and assessed.

The three strategic strands (data, people, and landscape) organised in six goals, will increase the interoperability of services within the social science community and beyond, to enable data discoverability for advancing research. This will be achieved by exchange among CESSDA's Service Providers, for mutual learning and transfer of knowledge to their national research landscape.

Strategic partnership with ERA Actors and global, will ensure that social science researchers at national and international level can provide evidence-based research for the benefit of society.

CESSDA will continue where it is best to 'support CESSDA's mission to provide long-term access to documented and curated high-quality social science data to researchers and beyond, for societal benefit'.
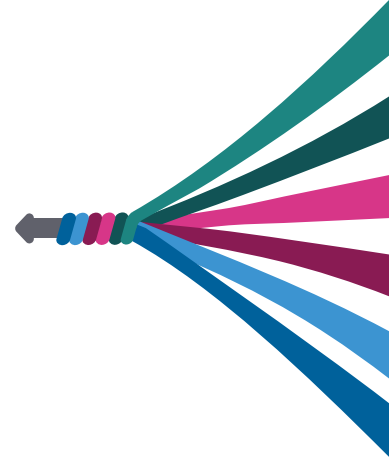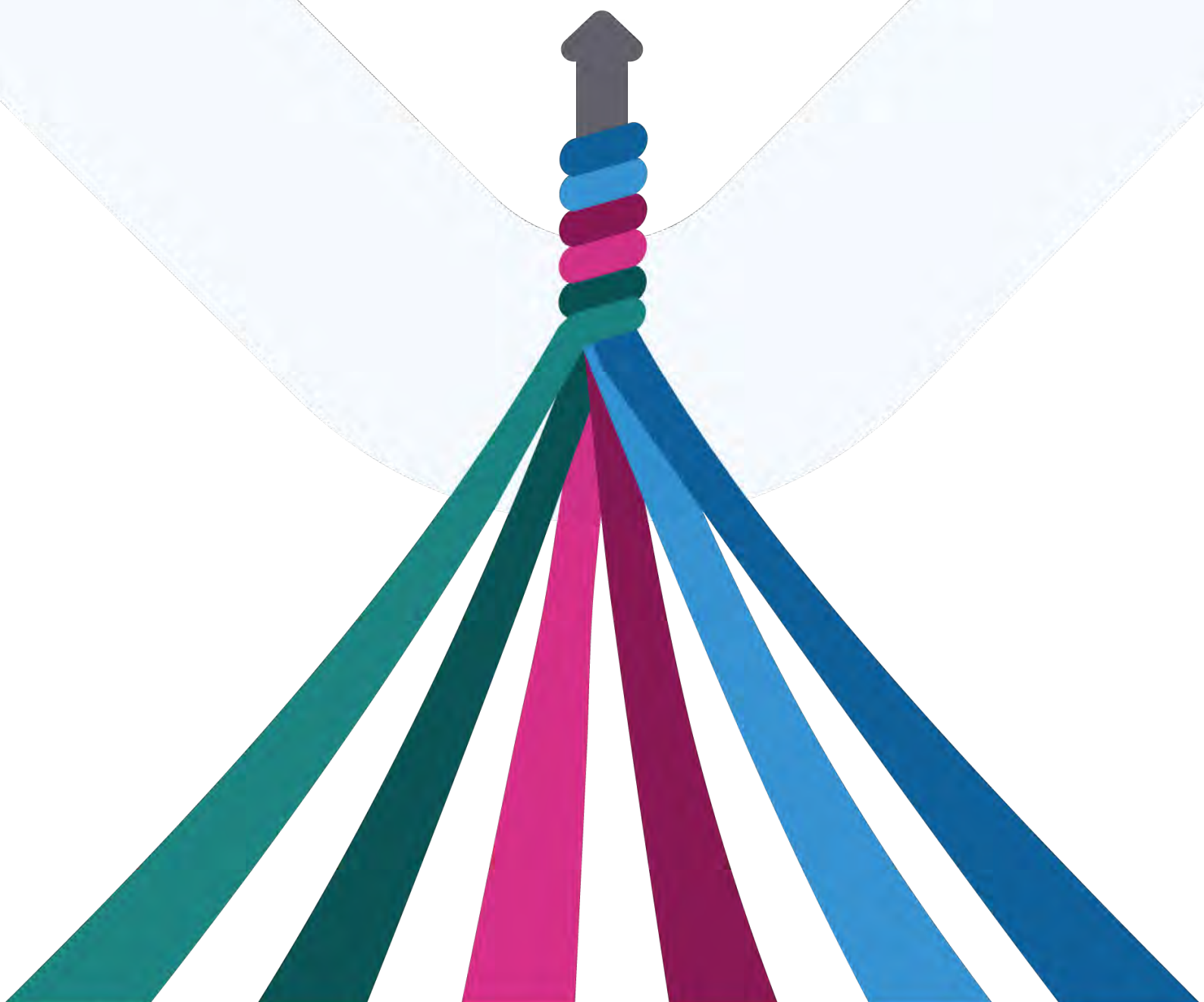
For more information visit our website:
cessda.eu

CESSDA Governance

Governance

The governance of the CESSDA ERIC Consortium consists of the General Assembly, the Service Providers' Forum (SPF) and the Director who oversees the Main Office. CESSDA´s activities are supported by the CESSDA´s Internal Advisory bodies and CESSDA´s External Advisory Groups.

The General Assembly

CESSDA's General Assembly is a CESSDA decision-making body as outlined in CESSDA's statutes (available on the web page CESSDA Statutes). CESSDA currently has 22 members and 12 partner countries (Albania, Bosnia and Herzegovina, Bulgaria, Estonia, Kosovo, Latvia, Lithuania, Luxembourg, Montenegro, Poland, Romania and Ukraine). CESSDA is supporting the Ukraine National Data Bank of Sociological Data "Kyiv Archive" through the Czech Data Archive with two grants for two research positions.

Service Providers and the Service Provider Forum

Member countries of CESSDA nominate a national Service Provider (SP) to be responsible for delivering the relevant services. **These SPs (mostly data archives)** have a primary responsibility to provide data services to their own country, but they are also funded internally to provide pan-European activities. The SPs are the main resource for CESSDA, and CESSDA is responsible for integrating the work of the SPs by establishing a one-stop shop for data location, access, analysis and delivery. Each SP has different overall objectives, but in general they have a responsibility for acquiring data from data creators – government, researchers, commerce, etc. – and preparing that data for long-term access. SPs also carry out curation functions, which means that data is always fit for contemporary use and is available for discovery and reuse. In essence, each SP ensures that data are always available for social science research purposes (see Q1.2 Asset 2 or under CESSDA DAG).

The SPs meet twice a year at the Service Provider's Forum (SPF). The SPF advises the Director on technical and national developments, needs of data archives and national researchers, suggests **bottom-up key topic working groups (WG)** and discusses CESSDA services, projects and special measures.

Current key topic working groups work on data citation, sensitive data and dataverse topics. Workshops on AI are planned for June 2024.

The Main Office and Director

CESSDA's Main Office (MO) is the operational hub (node) of the European Research Infrastructure. It collaborates and coordinates activities with the SPs and supports meetings of the General Assembly, the SPF and internal and external advisory bodies (see below). The MO sets up and manages the technological backbone of its products and services through Service Level Agreements (SLAs), runs joint EU and internal projects (internal funding schemes of CESSDA), and communication activities.

The Director is the Chief Executive Officer (CEO), Chief Scientific Officer and legal representative of CESSDA. The Director sets the strategic vision for CESSDA, reports to the GA and implements the policies and procedures.

CESSDA's Internal Advisory Bodies

CESSDA's current internal advisory bodies are two CESSDA Core Business Working Groups and the informal Directors Forum.

The Core Business WGs work, as the name insinuates, deal with CESSDAs core businesses. The interoperability (board) working group represents all Service Owners and Technical Maintainers of CESSDA services (see list of CESSDA services). The groups exchange best practices and brief each other on technical developments of the different CESSDA services. The interoperability working group provides advice on future technical research and development needs in line with state-of-the-art.

The Working Group Trust and Landscape covers topics on the development of the **CoreTrustSeal label** (for trusted depositories), oversees the **CESSDA-KPIs** and scrutinises the changing **Open Science and FAIR landscape**. They are also active working group members of EOSC task forces.

The Directors Forum is an informal meeting with the Directors of CESSDAs Service Providers to exchange matters of common interest. While the Directors Forum is informal, it also serves as a **sounding board on strategic decisions** in preparation for the General Assembly. The SPs Directors advise the CESSDA Director and, in most (not all) cases, brief their member country representative on content matters ahead of GA meetings. By doing so, discussions and decisions at the GA level are facilitated.

CESSDA's External Advisory Groups

CESSDAs external advisory bodies consist of different scientific advisory bodies such as SSHOC (the Social Sciences and Humanities Open Cluster), the ERIC Forum and other informal scientific advisory groups.

CESSDA also runs the SSH Open Science Cluster (**SSHOC**) secretariat on behalf of the five SSH ERICs and the six SSH projects on the ESFRI preparatory list. As the current SSHOC Chair, CESSDA is also currently in the **Executive Board of the ERIC Forum**.

Scientific advisory groups to CESSDA comprise the Research Data Alliance, DDI, CODATA and national research networks in different countries.
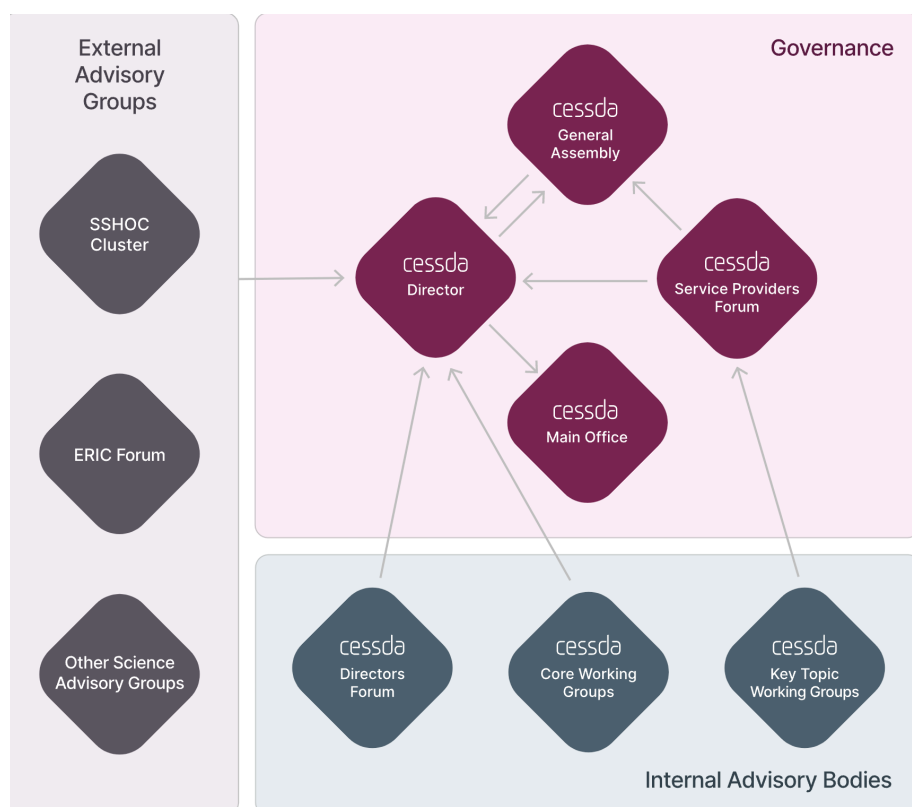
CESSDA Statutes

CESSDA's Statutes are currently being revisited for updating to CESSDA's decision making reality. The role of the working groups has been adjusted to the needs of CESSDA: Working Groups that deal with CESSDA's Core Business and that advise the Director directly and the bottom-up working groups working on a few selected topics (according to the strategic plan) which will report directly to the Service Provider Forum.

In the current CESSDA Statutes, the role of the Scientific Advisory Board is not well defined, has overlaps with the CESSDA core business and risks interfering with the prerogatives of SPs and the national institutions. For this reason, the roles of the internal advisory group (SP and technological development) and the external advisory group (science advice) have been separated to better capture both research and technological developments. It is planned to finalise the update of CESSDA's Statutes in 2024.

Supporting Document CESSDAs Tools & Services Supporting Research and Data Managers

CESSDA Data Catalogue (CDC)

The CESSDA Data Catalogue (CDC) is the flagship service of CESSDA. The CDC harnesses data and metadata of all domains of the SoS ranging from demography to immigration, from public opinions and political votes to behavioural sciences, and increasingly public health. The CDC is a one-stop shop for searching and finding European social science data.

It is the largest social sciences data catalogue in the world with ca. 37.000 datasets from 19814 studies in 15 languages. The provision of social science data in 15 European languages is unique in the world. The data described may be quantitative, qualitative or mixed-modes data, cross-sectional or longitudinal, recently collected or historical data. The metadata (study descriptions) are available in the language they were provided in. Some provide study descriptions both in English and in the local language, some only in either English or the local language. Currently, about 75% of study descriptions are available in English.

The CDC User Guide presents an overview of how to use the data catalogue for searching, covering basic search operations, application of filters and advanced search functionalities. The Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) is a low-barrier mechanism for repository interoperability. Data Providers are repositories that expose structured metadata via OAI-PMH. SPs then make OAI-PMH service requests to harvest that metadata.

Compliant endpoint enables data aggregators such as OpenAIRE, B2Find, BASE and GoTriple to easily harvest the entire contents of the CESSDA data catalogue. This provides a route for SP's metadata to find its way to a much wider audience than just Social Science researchers without any additional effort on their part.

There is also a search API. SPs can use it in their own catalogues to show users the results of their search from CDC as well as the local catalogue. Documentation is available on the public 'REST APIs for CESSDA Services' page.

An OAI-PMH compliant endpoint enables data aggregators such as OpenAIRE, B2Find, BASE and GoTriple to easily harvest the entire contents of the CESSDA data catalogue. This provides a route for SP's metadata to find its way to a much wider audience than just Social Science researchers without any additional effort on their part.

There is also a search API[1]. SPs can use it in their own catalogues to show users the results of their search from CDC as well as the local catalogue. Documentation is available on the public 'REST APIs for CESSDA Services' page[2].

CDC itself may have a small number of users. Its strength lies in the integration of national data services providing universal access to the world of European data, that are curated and maintained for longevity.

https://www.cessda.eu/Tools/Data-Catalogue

CDC and EOSC

The CESSDA Data Catalogue is harvested by organisations such as OpenAIRE, B2Find, BASE, GoTriple and Google Dataset Search, making its contents available to a wider audience of researchers.

---

[1] Search APIs are software components that allow developers to seamlessly introduce search capabilities to websites and applications. They provide backend tools for indexing documents, querying various types of data, managing cluster configurations, viewing search analytics, and more.
[2] https://vocabularies.cessda.eu/documentation/rest-api.html.

CESSDA will combine elements of the CESSDA Data Catalogue, CESSDA Metadata Validator, CESSDA Vocabulary Service and CESSDA ELSST to provide enhanced knowledge representation, richer Scientific Knowledge Graphs, linkages to Data Management Plans and large-scale FAIR assessment of metadata.

The CESSDA Vocabulary Service and DDI

The CESSDA Vocabulary Service enables users to discover, browse, and download Controlled Vocabularies (CVs) in a variety of languages. The service also contains an Editor where authorised users create, manage and translate the vocabularies. Access rights to the Editor are granted by the Vocabulary Service Service Owner.   The tool can also be used for curating and publishing in-house vocabularies - i.e. an SP may have some CVs that only it uses. These can be kept apart from the public CVs.

Many of the source (English) vocabularies have been created by the DDI Alliance. The Data Documentation Initiative (DDI) is an international standard for describing data produced by surveys and other observational methods in the social, behavioural, economic, and health sciences. The non-English language versions of the DDI vocabularies are provided by CESSDA members and associated organisations in different countries.

As DDI vocabularies are used worldwide, organisations which are willing to produce and maintain a non-CESSDA language version are welcome to contact CESSDA.

https://www.cessda.eu/Tools/Vocabulary-Service

The European Language Social Science Thesaurus

The European Language Social Science Thesaurus (ELSST) is a broad-based, multilingual thesaurus for the social sciences, owned and published by CESSDA and its national Service Providers. The thesaurus consists of over 3,300 concepts and covers the core social science disciplines: politics, sociology, economics, education, law, crime, demography, health, employment, information and communication technology, and environmental science.

ELSST supports data discovery within CESSDA and facilitates access to data resources across Europe, independent of domain, resource, language or vocabulary. ELSST is available in 15 languages: Dutch, Czech, English, Finnish, French, German, Greek, Hungarian, Icelandic, Lithuanian, Norwegian, Romanian, Slovenian, Spanish, and Swedish.

https://www.cessda.eu/Tools/ELSST-Thesaurus

DMEG

The CESSDA Data Management Expert Guide (DMEG) is designed by European experts to help social science researchers make their research data Findable, Accessible, Interoperable and Reusable (FAIR) using the following structure.

1. **Plan**: An introductory tour on what data management and a data management plan (DMP) are and why they are important. General concepts such as social science data and FAIR data are explained. The aim is that scientists to be able to write their own DMP.

2. **Organise & Document:** The chapter showcases good practices in designing an appropriate data file structure, naming, documenting and organising data files within suitable folder structures.

3. **Process**: This chapter deals with topics of data entry and coding as the first steps of proper data management.

4. The **Store** chapter explains how to plan a storage and backup strategy, advantages, disadvantages and measures to protect data from unauthorised access with strong passwords and encryption.

5. **Protect**. This chapter highlights legal and ethical obligations and shows how a combination of gaining consent, anonymising data, gaining clarity over who owns the copyright to data and controlling access can enable the ethical and legal sharing of data.

6. **Archive & Publish** recapitulates available data publication services and how to promote your data.

7. **Discover** explains how one can discover and reuse existing or previously collected datasets.

"The full DMEG is downloadable for personal study offline and PDFs for every chapter are available to be used as handouts for training. A pilot interactive game version of the guide is available.

In addition, data professionals can reuse a series of train-the-trainer materials like slides, workshop outlines, and exercises to be used in training sessions."

https://dmeg.cessda.eu/

See Training Resource below.

The CESSDA Data Archiving Guide

The CESSDA Data Archiving Guide (DAG) is designed to provide new employees at social science data archives with a general understanding of the work a data archive performs.

It has an introductory chapter suitable for the layman that addresses questions on what data archives are, why they exist, and what their mission, function, and way of operating is. Some sections contain expert tips pointing to useful materials or approaches related to the topic.

https://dag.cessda.eu/

CESSDA Training Resources

The CESSDA Training Resource Directory (TRD) disseminates relevant information necessary to build a sustainable and mature data archive and support the development of new services and features within existing data archives. Information on relevant documents, training materials, tools and support services are collected, selected and reviewed, making the TRD a curated inventory of existing resources. The TRD gathers not only resources from past and current CESSDA projects but also resources from CESSDA Service Providers and CESSDA Partners (i.e., non-member Service Providers) as well as resources beyond CESSDA that are important for data professionals and data archives. The RD is thus a central place where all CESSDA Service Providers and CESSDA Partners can turn to when looking for resources useful for building, developing and improving their services and practices.

In addition to communication and support information the resources are organised under categories that reflect the structure of the CoreTrustSeal Requirements: organisational infrastructure, digital object management and technical infrastructure.

Moreover, many materials are relevant beyond CESSDA and social science data services. The CESSDA Resource Directory may serve as a vital resource of knowledge for developing data services in Open Science in general.

https://www.cessda.eu/Training; https://www.cessda.eu/Training-Resource

A pilot interactive game version of the guide is available:

https://lod.sshopencloud.eu/

CESSDA Metadata Model and Metadata Validator

The CESSDA Metadata model (CMM) defines a set of metadata elements supported by CESSDA and guides the CESSDA Service Providers in making their data more discoverable and understandable to users both now and in the future. The CMM is based on the DDI Lifecycle metadata standard but reduces the complexity of the whole specification.

The CESSDA Metadata Validator (CMV) validates DDI metadata records used by CESSDA's tools and services against a given DDI Profile. These DDI Profiles are developed and maintained by the CESSDA Metadata Office. They contain the rules and constraints that metadata records need to follow to be compatible with the CESSDA Data Catalogue and the European Question Bank (see below European Question Bank). The CMV checks if the metadata record is valid according to the rules defined in the selected profile. If there are violations, a list of validation errors specifying which rules have been broken is produced. The CMV is an essential tool for quality assurance of the metadata used by some of CESSDA's services, providing feedback to the SPs about any defects so they can be fixed at the source. This type of activity is not practical to do by hand as there are many thousands of records that need to be checked on a regular basis.

https://www.cessda.eu/Tools/Metadata-Validator

CESSDA Dataverse

The CESSDA consortium contributes to advancing the capabilities and adoption of Dataverse, a key open-source repository software in the global research community. Through projects like DataverseEU, CESSDA helps equip new and aspiring SPs with essential tools to manage, preserve, and provide access to research data efficiently, particularly those with limited technical resources.

Dataverse serves as an invaluable tool for data archiving and discovery, offering a robust platform for researchers and institutions worldwide to access, preserve, explore, and analyse data. The software supports high data reproducibility and long-term usability, underpinned by a diverse and extensive global development community. Notably, CESSDA's involvement includes significant enhancements such as the integration of Docker container technologies, the development of specific plugins for persistent identifiers and controlled vocabularies, and the translation of the user interface and metadata to meet diverse linguistic and national requirements.

These adaptations are crucial for aligning Dataverse with the European research infrastructure landscape, further supported by the SSHOC project's efforts to develop a mature software pipeline. Additionally, improvements such as the integration of Apache Taverna for workflow management, the development of data migration solutions, and custom metadata support highlight the strategic importance of Dataverse in enhancing data management practices across disciplines.

Overall, CESSDA's contributions to Dataverse development not only enhance the software's functionality and accessibility but also reinforce the importance of open-source solutions in the scientific and academic communities for effective data management and collaboration.

https://www.cessda.eu/News/CESSDA-Newsitem-nid3350

Future CESSDA activities 2024-2026

The European Question Bank

Of high interest to scientists, questionnaire designers, and survey methodologists is the possibility to describe the set of questions conceived for the respective surveys. CESSDA is working on a pilot with a commercial company to set up a European Question Bank (EQB). The CESSDA EQB is a cross-national question bank that integrates and displays metadata from several European archives, mostly CESSDA SPs.

The platform allows users to search, browse, compare, and download multilingual survey questions and study-related metadata. Users can search question texts, response alternatives and keywords. Questions can be filtered by data collection dates or countries, question languages, and by the methodology used for data collection. Researchers can find relevant questions for their research topics in one place. They can compare different versions and different translations of questions. They can find out about question use and get direct links to the data collected using these questions.

The EQB can also help questionnaire designers and translators interested in high-quality, tested questions or translations of questions. They can get a sense of how translators handled translation issues for particular questions by comparing questions side-by-side.

The EQB is also useful for survey methodologists. They can see how questions on a particular topic have changed over time. Not only can they access and export survey questions, they can also find study-related information and documentation.

The Euro Question Bank is based on the DDI-Lifecycle metadata standard. CESSDA SPs contribute documentation from their holdings to the EQB using DDI standardised metadata, preferably in DDI Lifecycle version 3.2 or DDI Codebook. The DDI-based documentation is provided by OAI-PMH.

The CESSDA Metadata Validator (CMV) and the EQB DDI profile are used to validate the DDI metadata. The EQB metadata schema is compliant with the CESSDA Metadata Model (CMM).

https://www.cessda.eu/Tools/EQB