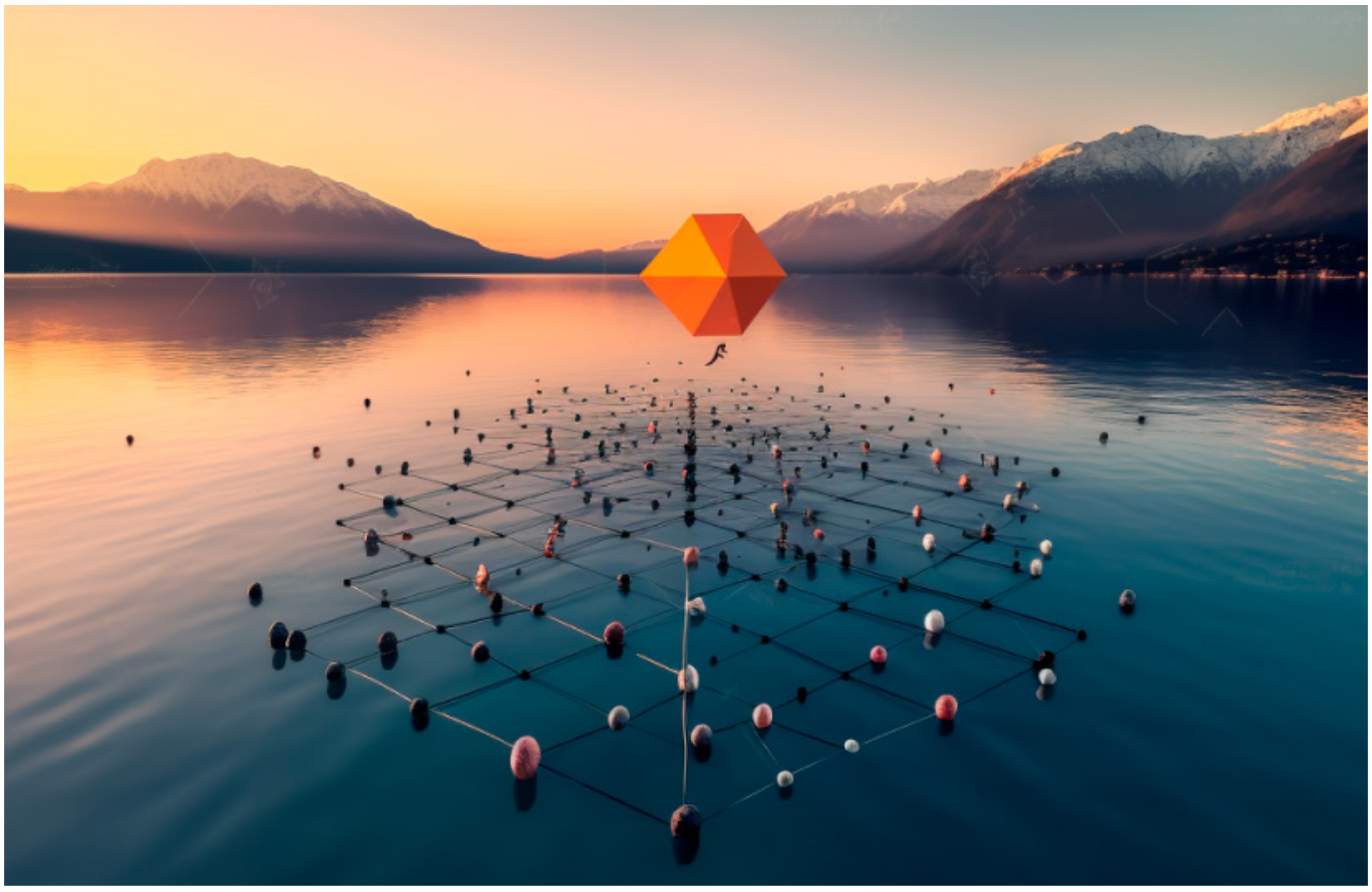
open research data National Strategy Council

FINAL REPORT
April 2025

ORD Task Force
Social Sciences and Humanities Data Report



MANDATOR Swiss National ORD Strategy Council

TASK FORCE Rainer Gabriel, ZHAW • Tobias Hodel, UniBE • Ben Jann, UniBE • Sylvia Jeney, SNSF • Béla Kapossy (Chair), Unil • Tabea Lurk, FHNW • Kurt Schmidheiny, Unibas

SCIENTIFIC COLLABORATOR Simon van Rekum, ZHAW

COORDINATION Rudolf Mumenthaler, UZH • Sarah Schlunegger, UniBE

NATURE OF PROJECT Final report to the attention of the Swiss National ORD Strategy Council

PLEASE CITE THIS PUBLICATION AS: Béla Kapossy, Simon van Rekum, Tabea Lurk, Sarah Schlunegger, Rudolf Mumenthaler, Silvia Jeney, Kurt Schmidheiny, Tobias Hodel, Rainer Gabriel, Ben Jann. (2025). ORD Task Force Social Sciences and Humanities Data Report. Swiss National ORD Strategy Council. https://doi.org/10.5281/zenodo.15268747

Executive Summary

Context and goals

The Swiss National Open Research Data Strategy Council (StraCo) mandated the Task Force for Social Sciences and Humanities (TF SSH) to analyze Cluster 2 within Switzerland's Open Research Data (ORD) ecosystem. This report provides the analytical, factual foundation for a strategic push to improve coordination, sustainability, and alignment of SSH research data infrastructures (RDIs) as part of the Blueprint process.

Social Sciences and Humanities are critical to addressing societal challenges such as digital transformation, cultural preservation, and evidence-based policymaking. Despite considerable progress, SSH RDIs often face fragmented governance, short-term funding models, and data access and reuse barriers. This report aims to map existing SSH RDIs and services and identify gaps, redundancies, and opportunities for coordination.

Key Findings

1. Strengths:

- Leading infrastructures, such as SWISSUbase and DaSCH, demonstrate maturity, technical expertise, and adherence to FAIR[1] principles.
- Emerging networks like SSHOC-CH and Swiss nodes in European Research Infrastructure Consortia (ERICs) (DARIAH, CLARIN, CESSDA) offer strong potential for national and international alignment.
- Plurality of bottom-up small-scale initiatives ensures methodological diversity and linkage to the different communities.

2. Challenges:

- Governance gaps: Lack of centralized coordination prevents strategic integration and reduces interoperability.
- Funding risks: Short-term funding models undermine the sustainability of critical SSH infrastructures.
- Data access barriers: Complex and restrictive legal, technical, and ethical frameworks limit research potential of data domains such as government microdata, free cultural assets, and social media.

3. Opportunities for Coordination

- Strengthen National Coordination:
  - Consolidate national SSH repositories under SWISSUbase, in close collaboration with DaSCH, to eliminate redundancies and improve interoperability.
  - Mandate SSHOC-CH to coordinate data activities across the SSH domain, ensuring consistent governance and technical alignment.
  - Improve coordination between longitudinal studies, with FORS playing a leading role in standardizing, developing and improving methodologies, aligning survey designs, exploiting synergies, and enabling FAIR data sharing.
- Enhance International Alignment:
  - Strengthen Swiss participation in the European Open Science Cloud (EOSC) by establishing nodes and integrating smaller initiatives to align with international standards.
  - Support Swiss ERIC nodes (DARIAH, CLARIN, CESSDA) to strengthen Switzerland's position in SSH data infrastructures and foster global interoperability.
  - Establish reciprocal data-sharing agreements with international research infrastructures (RIs) to ensure global access and alignment.
- Ensure Long-term Funding Sustainability:
  - Implement stable funding mechanisms for SWISSUbase, DaSCH, FORS, scientific activities of the Swiss Federal Statistical Office (FSO), longitudinal studies, and essential tools for data to ensure operational continuity.
  - Develop funding schemes to integrate short-term research outputs into sustainable platforms.
  - Prioritize long-term support for research data repositories to safeguard research outputs and ensure accessibility.
- Improve Data Access and Governance:
  - Develop transparent frameworks for accessing Swiss government microdata based on international best practices (e.g., the Five-Safes approach), promoting responsible and ethical data use.
  - Address legal frameworks to enable the controlled, ethical reuse of SSH data in Artificial Intelligence (AI) applications, ensuring provenance, transparency, and bias mitigation.
  - Integrate metasearch platforms into a unified system to streamline government, cultural, and research data access.

Next Steps

Embedded within the National ORD Strategy, this report provides the groundwork for shaping strategic options to enhance coordination and efficiency within the SSH Cluster. Collaborating with the TF, the StraCo's Coordination Group (CoG) will translate these insights into a strategic vision for the cluster. This vision should consider overarching objectives such as FAIR and open research data (ORD), open scholarly communication, inclusive and transparent processes, and societal engagement and participation. It will be discussed and refined with stakeholders throughout 2025 and 2026.

1 Findable, Accessible, Interoperable, Reusable

Table of contents

Executive Summary 2
Context and Goals 2
Key Findings 3
Next Steps 3
1 Introduction 8
1.1 The National ORD Strategy 9
1.2 The Blueprint Process 9
1.3 The Social Sciences and Humanities Data Cluster 10
1.4 National efforts and current state of discussion 10
1.5 Cluster analysis 11
1.6 Set up and composition of the Task Force ocial Sciences and Humanities 11
1.7 Initiatives as the entry point 12
1.8 The mandate 12
2 The Task Force Methodology 13
2.1 Identification of initiatives 15
2.2 Structuring of the SSH initiatives – sub-clusters 15
2.3 Data collection 17
2.4 Data analysis – The concept of perspectives 18
2.5 Stakeholders involvement in the process 18
3 Analysis: Perspectives on Cluster 2 19
3.1 Open Research Data vs. Open Data for Research 21
3.2 Financial dimension: Temporary vs. Long-term funding 22
3.3 Thematic perspective: Findability 25
3.4 Thematic perspective: Accessibility 27
3.5 Thematic perspective: Interoperability 29
3.6 Thematic perspective: Reusability 30
3.7 Thematic perspective: Identified professional and infrastructural emergent networks 32
3.8 Thematic perspective: International alignment 33
4 Discussion 36
4.1 Recommendations 37
4.2 Priority areas 40
4.3 Outlook and further steps 40
5 Conclusion 41
Annexes 45
Annex 1: (Long) list of initiatives 46
Annex 2: Participants of stakeholder event 47
Annex 3: Overview of conducted interviews 49
Annex 4: Dashboards of selected initiatives 50
Annex 5: Financial summary 61

List of figures

Figure 1. The Task Force Methodology for ORD Landscape analyses 14
Figure 2. Perspective on temporal dimension 22
Figure 3. Findability and visibility of (Open) Research Data 25
Figure 4. Forms of accessibility based on data type/source 28
Figure 5. Language data - Example of a domain-specific network 32
Figure 6. International alignment 35

List of tables

Table 1. Domains identified as sub-clusters of Cluster 2 when grouping initiatives 15
Table 2. Analytical perspectives for Cluster 2 18
Table 3. Funding structure for Competence Centers and Consortia, Data production tools for researchers, and Longitudinal studies and surveys 23
Table 4. Funding structure for cultural assets, metasearch services, and linked open data 24
Table 5. Future financial needs in the SSH Cluster 40

List of abbreviations

AI: Artificial Intelligence

CARE: Collective benefit, Authority to control, Responsibility, Ethics

CESSDA-CH: Consortium of European Social Science Data Archives Switzerland
CLARIN-CH: Common Language Resources and Technology Infrastructure Switzerland
CoG: Coordination Group of the National ORD Strategy Council

DaSCH: Swiss National Data and Service Center for the Humanities

DARIAH-CH: Digital Research Infrastructure for the Arts and Humanities Switzerland
DSU: Data Service Unit

ECCCH: European Collaborative Cloud for Cultural Heritage

ELEXIS: European Lexicographic Infrastructure Consortium

EOSC: European Open Science Cloud

EPFL: Ecole Polytechnique Fédérale de Lausanne

ERC: European Reseach Council

ERI : Education, research, and innovation

ERIC: European Research Infrastructure Consortium

ESFRI: European Strategy Forum on Research Infrastructures

ETHZ: Federal Institute of Technology Zurich

FAIR: Findability, Accessibility, Interoperability, Reusability

FHNW: Fachhochschule Nordwestschweiz

FOPH: Swiss Federal Office of Public Health

FORS: Swiss Center of Expertise in the Social Sciences

FSO: Swiss Federal Statistical Office

GLAM: Galleries, Libraries, Archives, Museums

HEI: Higher Education Institutions

HLS: Health and Life Sciences

IIIF: International Image Interoperability Framework

LiRI: Linguistic Research Infrastructure

NGO: Non-Governmental Organization

NLP: Natural Language Processing

OGD: Open Government Data

ORD: Open Research Data

OA: Open Access

OAI-PMH: Open Archives Initiative Protocol for Metadata Harvesting

ODN: Open Data Navigator

ORD: Open Research Data

OS: Open Science

PgB: projektgebundene Beiträge (swissuniversities) [project-bound contributions according to Art. 50 Higher Education Act]

PISA: Programme for International Student Assessment

PID: Persistent identifiers

RDF: Resource Description Framework

RDI: Research Data Infrastructure

RI: Research Infrastructure

SAGW: Swiss Academy of Social Sciences and Humanities

SARI: SARI - Swiss Art Research Infrastructure

SB SP: Sounding Board Service Providers

SB R: Sounding Board Researchers

SERI: State Secretariat for Education, Research and Innovation

SHP: Swiss Household Panel

SJMM: Swiss Job Market Monitor

SLFS: Swiss Labour Force Survey

SSA: Société Suisse des acteurs

SSH: Social Sciences and Humanities

SSHOC-CH: The Social Sciences & Humanities Open Science Cluster Switzerland

StraCo: National ORD Strategy Council

SUPSI: University of Applied Sciences and Arts of Southern Switzerland

TF: Task Force

TREE: Transitions from Education to Employment

Unibas: University of Basel

UniBE: University of Bern

Unil: University of Lausanne

ZHAW: Zurich University of Applied Sciences

1. Introduction

1.1 The National ORD Strategy 9
1.2 The Blueprint Process 9
1.3 The Social Sciences and Humanities Data Cluster 10
1.4 National efforts and current state of discussion 10
1.5 Cluster analysis 11
1.6 Set up and composition of the Task Force ocial Sciences and Humanities 11
1.7 Initiatives as the entry point 12
1.8 The mandate 12

1.1 The National ORD Strategy

In July 2021, the four national education, research, and innovation (ERI) actors in Switzerland (the ETH Domain, the Swiss Academies of Arts and Sciences, the Swiss National Science Foundation, and swissuniversities) published the Swiss National ORD Strategy[2], initiated by a mandate of the State Secretariat for Education, Research and Innovation (SERI).

The recommendations and principles of the Strategy were concretised through an Action Plan[3] published in January 2022. The Action Plan defined Action Areas and formalised the four ERI actors' partnership in establishing the StraCo as a new governance model. The StraCo is responsible for:

1. the strategic coordination of the ORD actions of ERI actors
2. the consolidation of the ORD Landscape, as part of the Action Plan's Action Area B and D[4]
3. the development and updating the ORD Action Plan
4. representing the national ORD vision and the interests of ERI actors.

In this framework, the StraCo is developing the Blueprint process as a tool for the strategic coordination of the ORD landscape, of which this landscape analysis report is an integral part.

1.2 The Blueprint Process

The starting point of the Blueprint process is the identification by the StraCo of disciplinary clusters within the ORD ecosystem where concrete action is needed[5]. Task Forces are then assembled with the mandate of conducting a landscape analysis of each cluster. Using the information from the cluster analysis, the StraCo develops strategic options to be integrated into its Blueprint. These options include directions for the development and coordination of the cluster (for instance, development of new areas, extensions, closure, or merger of infrastructures). The Blueprint is meant as a guidance instrument, providing a framework for long-term planning to which funding decisions can refer, but it is not a decision-making instrument.

StraCo defines a cluster as a data- or discipline-specific area where

- there is a highly dynamic ORD development;
- multiple actors are operating (national and international);
- strategic coordination is needed and/or desired;
- infrastructures of "foremost national relevance" are likely to exist.

The three initial domain clusters to be investigated and strategically apprehended in the first phase of the ORD landscape consolidation work are:

- Health and life sciences (HLS) (Cluster 1);
- Social sciences and the humanities (SSH) (Cluster 2)
- AI/Data science (Cluster 3)

As of 2024, the StraCo has established a TF for the first two of these clusters. The analysis for Cluster 1 (HLS) has been completed, allowing the StraCo to explore strategic options and formulate the Blueprint for further coordination and consolidation of this cluster and its ecosystem. The analysis for Cluster 2 (SSH) is the subject of the present report, while the analysis for Cluster 3 (AI/Data science) is still pending.

Cluster analyses are intended to be updated and extended on a regular basis to ensure they present an accurate representation of the field and incorporate new developments. Delineating clusters is inherently difficult due to multiple interfaces between them. In fields like social work, psychology, or education, ethical and legal dilemmas intersect between HLS and SSH domains. Data-heavy and data-driven disciplines like corpus and computational linguistics feature intersections between SSH and Data science. Similarly, in linguistics, the intersection of language and medicine introduces complexities in understanding disease perception. The StraCo's Sounding Board of Researchers (SB R) is working on better integrating these interfaces within the Blueprint process. This also underscores the broader scope of ORD challenges and requirements across clusters, confirming the relevance of a unified cluster analysis methodology across disciplines.

2 Swiss National ORD Strategy (July 2021) (https://www.swissuniversities.ch/fileadmin/swissuniversities/Dokumente/Hochschulpolitik/ORD/Swiss_National_ORD_Strategy_en.pdf).
3 Swiss National ORD Action Plan (January 2022) (https://www.swissuniversities.ch/fileadmin/swissuniversities/Dokumente/Hochschulpolitik/ORD/ActionPlanV1.0_Decembe r_2021_def.pdf).
4 *Ibid.*
5 In this context, a focus on ORD encompasses any activity involving research data that must be taken into consideration to reach the four objectives of the Swiss National ORD Strategy.

1.3 The Social Sciences and Humanities Data Cluster

Social sciences encompass a wide array of academic disciplines that study human behaviour and societies. Key subjects include anthropology, sociology, psychology, political science, economics, geography, law, communication, and cultural studies. Humanities focus on studying human culture and thought, including the arts, culture, and socially engaged stakeholders. Key subjects herein include history, philosophy, literature, linguistics, religion, art (including visual and performing arts), music and theatre. In addition, cultural, socially engaged, and political actors such as museums, archives, Linked Open Data communities, and public authorities need to be accounted for, as they contribute, produce, link, and provide research-relevant data in various ways. Unlike in the HLS cluster, the SSH landscape is structured around a few major stakeholders with strong advocacy capabilities. There are several a priori distributed, smaller scale but potentially important initiatives in a very heterogeneous and scattered landscape. Since both areas are traditionally literature-intensive and often deal with historical sources and documents, the academic libraries, archives and various cultural collections are part of the SSH environment (as in this report).

There are several specific challenges faced by Swiss SSH infrastructures as articulated in a position paper and a white paper by SSHOC-CH (Social Sciences & Humanities Open Cluster Switzerland):[6]

- Funding and prioritization: One of the main challenges is allocating funding and prioritizing SSH projects. The national infrastructure roadmap process has historically favoured natural sciences and medical RIs, often leaving SSH projects underfunded. This imbalance makes it difficult for SSH disciplines to develop and maintain the necessary infrastructure.
- Diversity and coordination: The SSH landscape in Switzerland is highly diverse, with many small-scale initiatives whose need for coordination differs. This diversity of data and research tools across the SSH domain, while warranting customization of the research communities' needs, runs the risk of leading to inefficiencies and duplication of efforts. From a strategic and national viewpoint, better coordination and integration of these initiatives are needed to create a more cohesive and effective RIs.
- Digital transformation: The shift towards digital research methods and the need for digital infrastructure pose significant challenges. All SSH researchers require access to advanced digital tools and platforms for data analysis, storage, and sharing. It is crucial to ensure that these digital infrastructures are adequately developed and maintained.
- Sustainability and long-term planning: Another challenge is developing sustainable research data infrastructures that can support long-term projects. SSH research often requires long-term data collection and archiving, necessitating stable and sustained funding and infrastructure support.

Assessing these challenges in-depth and outlining potential action areas for the StraCo requires concerted efforts from policymakers, funding agencies, and the research community to ensure that SSH disciplines have the necessary resources and infrastructure to thrive.

1.4 National efforts and current state of discussion

In Switzerland, current discussions and national efforts around SSH disciplines and their infrastructure focus on several key areas:

National Research Programs and Funding: The Swiss National Science Foundation (SNSF) plays a crucial role in funding SSH research. Additionally, programs like the National Research Programmes (NRPs) support interdisciplinary projects that address societal challenges. Another key actor is the Swiss Academy of Humanities and Social Sciences, which supports various initiatives and RIs.[7] These commitments aim to strengthen the SSH research landscape by providing substantial funding and fostering collaboration among researchers. The discussions in the TF and with stakeholders put another emphasis on the sector of cultural assets, such as academic libraries, archives,

6 Position paper: Social sciences and humanities research infrastructures in Switzerland," Lausanne 2022 (position-paper-ssh-infrastructures_24082022.pdf); White paper: Founding the Social Sciences and Humanities Open Cluster (SSHOC-CH) in Switzerland (https://sshoc.ch/_media/sshoc-ch_white_paper_v03.04.2024.pdf).

7 https://www.sagw.ch/sagw/forschungsinfrastrukturen

museums and collections as well as stakeholders from the field of linked open data. They maintain an important infrastructure, especially for research in the Digital Humanities. Many of these infrastructures received initial funding by swissuniversities with its current Open Science (OS) program or the former program P5 Scientific Information.[8]

Digital Infrastructure and OS: There is a strong emphasis on enhancing digital infrastructure to support SSH research. The Swiss government promotes OS policies, encouraging the digitization of cultural and historical resources, and the development of OA platforms.[9] This effort aims to make research outputs more accessible and to facilitate data sharing among SSH researchers.

Interdisciplinary Collaboration: Efforts are being made to bridge the gap between SSH and other scientific disciplines. Initiatives like the Swiss ORD strategy emphasize the importance of interdisciplinary collaboration in addressing complex societal issues. This includes integrating SSH perspectives into broader research agendas and ensuring that SSH researchers have access to the necessary infrastructure and resources.

International Integration: Swiss researchers actively participate, whenever politically possible, in European research programs such as Horizon Europe. There is a concerted effort to ensure that Swiss SSH researchers are well-represented in these programs and that their contributions are integrated into international research initiatives.

These efforts highlight the importance of SSH disciplines in addressing societal challenges and the need for robust infrastructure to support their research and integration into broader scientific and policy frameworks.

8 https://www.swissuniversities.ch/themen/open-science/programm-open-science

9 https://www.sbfi.admin.ch/en/open-science-en

1.5 Cluster analysis

The present analysis of Cluster 2 aims to offer a comprehensive view on:

- an initial list of initiatives active in the cluster (including the services and infrastructures they provide to researchers along the research data lifecycle and regarding ORD) selected by the StraCo and carefully supplemented by the TF SSH.
- the overall cluster dynamic, to understand more substantially how users navigate the services and infrastructures offered, how services and infrastructures interact with one another, how the existing facilities meet their needs, and to identify potential gaps and overlaps as well as bottlenecks for the coordination and development of the cluster.

This landscape analysis aims to offer a presentation of the SSH ORD and data ecosystem that, while not exhaustive, is representative of researchers and service providers active in the cluster.

The cluster analysis methodology, outlined in Chapter 2, was designed to consider the specificities of RDI: it uses data and disciplinary clusters as a unit for analysis instead of individual RDIs, considers infrastructures and services across the whole ORD value chain (*see Fig. 1*), and focuses on interconnections between RDI as an inherent part of their services.

1.6 Set up and composition of the Task Force Social Sciences and Humanities

The TF comprises individuals who can effectively contribute to its mandate while maintaining sufficient independence from any initiative on the list. This approach aims to facilitate a more comprehensive understanding of the current situation and enable the cluster analysis to focus on the existing services and infrastructures, rather than the initiatives' outlook.

The TF consists of the following members:

External experts

- Béla Kapossy (chair), Professor of Early Modern History at the University of Lausanne (Unil) and former Director of the College of Humanities at the École Polytechnique Fédérale de Lausanne (EPFL).
- Rainer Gabriel (member), Researcher and lecturer at the Zurich University of Applied Sciences (ZHAW) School of Social Work, data steward for the department and member of the advisory board of the ZHAW center for research data.
- Tobias Hodel (member), Assistant Professor in Digital Humanities at the University of Bern (UniBE). Board member of the Academy of Humanities and Social Sciences.

- Ben Jann (member), Professor of Social Stratification at the Institute of Sociology at the UniBE.
- Sylvia Jeney (member), Head of ORD at SNSF, co-responsible for the life-time management of RIs in SSH funded by the SNSF. EOSC delegate for the SNSF.
- Tabea Lurk (member; co-lead author), Head of the media library at the Academy of Art and Design Basel Fachhochschule Nordwestschweiz (FHNW), its digital archive and member of the library board at the FHNW.
- Kurt Schmidheiny (member), Professor of Economics and Applied Econometrics at the University of Basel (Unibas). President of Swiss Society of Economics and Statistics.
- Simon van Rekum (scientific collaborator; co-lead author), Staff Member at the ZHAW University Library, Research & Infrastructure area.

Members of the Coordination Group

- Rudolf Mumenthaler, Director of University Library Zurich (UZH, CoG)
- Sarah Schlunegger, Member of Scientific Staff at Vice-Rectorate Research and Innovation (UniBE, CoG)
- Support: Antonio Giannico, Scientific Assistant at Vice-Rectorate Research and Innovation (UniBE)

Former members of the Task Force SSH

- Iolanda Pensa (member), Senior Researcher and Head of Culture and Territory, Design Institute at the University of Applied Sciences and Arts of Southern Switzerland (SUPSI)

1.7 Initiatives as the entry point

Initiatives are the entry point for the TF in investigating the cluster. They are defined as any organisation or network providing RDIs or services to researchers or academic institutions. The term initiative is versatile enough to accommodate the diversity of governance and financial structures of the services and infrastructure providers examined by the TF.

Following StraCo's definition of a cluster – characterized by high ORD dynamics, active international and multinational actors, existence of infrastructures of national relevance, and a need for strategic coordination – the set of initiatives presented in this report should be seen as representative, but not exhaustive. The goal is not to provide a comprehensive inventory of services and RDIs in the cluster but to select key initiatives that exemplify its ORD dynamics. This approach allows for a flexible understanding of the cluster's current dynamics and aligns well with the subsequent Blueprint process.

In a first step, the StraCo proposed a minimum list of initiatives to be considered by the TF: FORS, SWISSUbase, DaSCH, LiRI, SSH ERICs, Webservices (e.g. metagrid), platforms (e.g. nodegoat), and smaller RDIs like SARI, e-codices, e-manuscripta, e-rara.

In addition, the TF was mandated to identify further initiatives of national importance or with the potential to be of national importance. The TF consequently supplemented the initiatives proposed by the StraCo following the methodological approach described in Chapter 2. The TF aimed to introduce more initiatives of national relevance through that process and gain insights into their contribution to the Cluster 2, as defined by the StraCo. With this strategy in mind, the TF aimed to be inclusive of the nationally relevant initiatives that represented the SSH Cluster and the needs of the SSH community.

1.8 The mandate

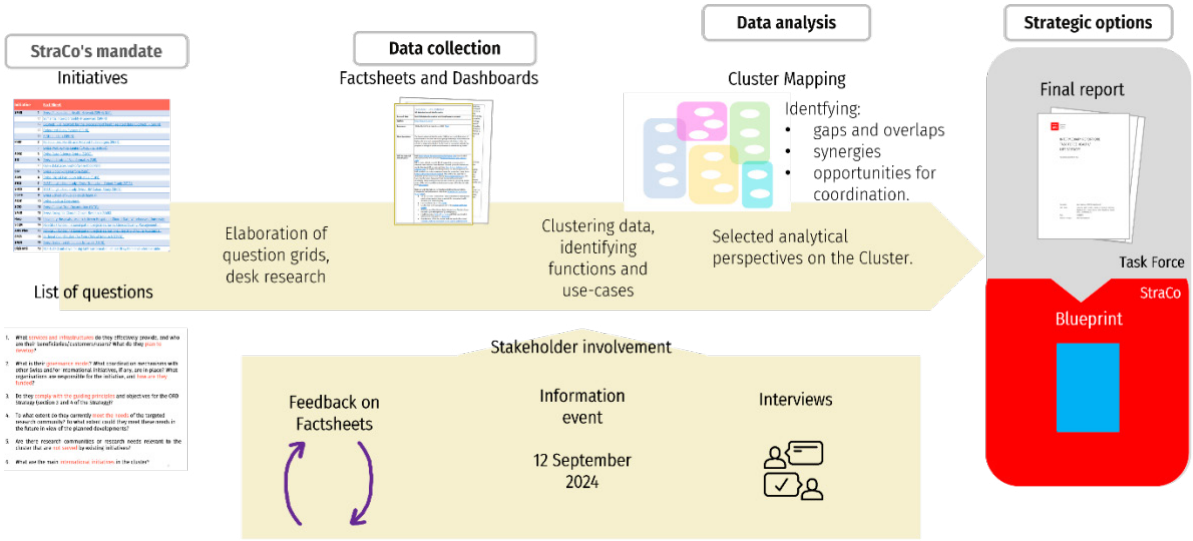
The StraCo's mandate to the TF lists questions to be answered for each initiative to investigate its position and interaction in the cluster. The TF has applied this list of questions to all initiatives in the list above (*see 1.7*) and to the added initiatives selected for the landscape analysis.

1. What services and infrastructures do they effectively provide, and who are their beneficiaries/customers/users? What do they plan to develop?
2. What is their governance model? What coordination mechanisms with other Swiss and/or international initiatives, if any, are in place? What organisations are responsible for the initiative, and how are they funded?
3. Do they comply with the guiding principles and objectives for the ORD Strategy (*sections 3 and 4 of the Strategy*)?
4. To what extent do they currently meet the needs of the targeted research community? To what extent could they meet these needs in the future in view of the planned developments?
5. Are there research communities or research needs relevant to the cluster that are not served by existing initiatives?
6. What are the main international initiatives in the cluster?

2. The Task Force Methodology

2.1 Identification of initiatives 15
2.2 Structuring of the SSH initiatives – sub-clusters 15
2.3 Data collection 17
2.4 Data analysis – The concept of perspectives 18
2.5 Stakeholders involvement in the process 18

FIGURE 1
The Task Force Methodology for ORD Landscape analyses



The analysis starts with identifying key initiatives active in the cluster (*see sections 1.7 and 2.1*) and compiling information along the questions developed in the mandate (*section 1.8*), such as the infrastructures and services they provide and their governance and funding models. Data is compiled in Factsheets (*section 2.3.1*) that serve as factual references and are used to feed into initiative Dashboards (*section 2.3.1*). These Dashboards summarise the Factsheets in a succinct yet comprehensive overview, avoiding conclusive content. Factsheets and Dashboards together with information received from stakeholders in informal interviews and exchanges for communicative validation and 3 semi-structured interviews (*section 2.3.2*) are the basis for the TF's mapping exercise to understandthe internal dynamics of the cluster and its relationships with the ORD ecosystem.

2.1 Identification of initiatives

As a first step, the TF worked on expanding the list of initiatives in the mandate to accurately represent the landscape. The approach has been for TF members to think of initiatives active in Cluster 2 while discussing criteria for selection in parallel. Such selection criteria encompassed: (1) national importance or potential for national importance; (2) potentially valuable and open to SSH field; (3) data accessible for research community within SSH; (4) the adoption of FAIR principles; and (5) EOSC compatibility. This resulted in a "long" list of initiatives (*see annex 1*), which still does not provide a complete coverage of the landscape. In a subsequent step, and based on common characteristics, the "long" list of initiatives was further organised into sub-clusters that aggregated and consolidated the collected information on individual initiatives. Section 2.2 takes a deeper look into this process of sub-clustering.

2.2 Structuring of the SSH initiatives – sub-clusters

After identifying initiatives that were not included in the original mandate, the TF members recognized the importance to: (1) cluster the initiatives based on commonalities and (2) acknowledge the list's limitations in fully covering all relevant initiatives. The sub clustering aided the TF in gaining a common understanding of the SSH Cluster regarding common characteristics and thus similarities and potential parallel efforts. The following identified sets of initiatives were defined as sub-clusters (*see Table 1*), and they are shortly defined and described[10]:

TABLE 1
Domains identified as sub-clusters of Cluster 2 when grouping initiatives

Institutional Networks	Institutional networks are collaborative entities, such as consortia and asasociations, that bring together various institutions, researchers, and resources to address complex scientific and societal challenges. These entities often focus on interdisciplinary research, pooling expertise and infrastructures to achieve common goals. Two notable examples of such initiatives are the SSHOC-CH initiative and the Zentrum für Digitale Editionen [Center for digital editions].
Competence Centers	Competence Centers are specialized institutions that provide expertise, resources, and support for specific research domains. These centers often serve as hubs for innovation, collaboration, and knowledge dissemination, helping to advance their respective fields. Here are some notable initiatives within this domain: FORS, DaSCH and LiRI. Note that competence centers may establish around and coincide with infrastructural services as is the case with both DaSCH and LiRI.
Cultural assets, metasearch services & linked open data	The core mission of institutions such as academic libraries, archives, museums, and data collections is to collect, preserve and process data of historical, cultural and scientific value in a FAIR way so that they are accessible and (re-)usable. Therefore, a) eight databases with focus on 'presentation of digitized collections' and b) six 'metasearch' initiatives were initially examined. It appears that many of the a) data providers merge into the b) metasearch initiatives. Thus, naming was adapted. Considering the ecosystem of linked open data brings federal authorities with open government data (OGD)[11] and stakeholders from the Wikiverse as well as scientifically relevant, communities-based RIs such as GitHub or Hugging Face[12] into focus. However, stakeholders outside academia could only indirectly be included in the data analysis for structural reasons.

10 Alphabetically ordered
11 C.f. the research reports of the Federal Statistical Office, Federal Social Insurance Office and others. See also: https://opendata.swiss
12 https://huggingface.co/

Data production tools for researchers	Data production tools are essential for researchers across various disciplines, enabling them to collect, manage, analyse, and share data efficiently. These tools or research environments often provide specialized functionalities tailored to the needs of specific research domains. Here are some notable initiatives within this domain: surveys, nodegoat, geovistory, transcriptiones, etc.
Data repositories	Data repositories are centralized locations where research data resulting from scientific research is stored and made accessible by researchers. Three types of repositories can be distinguished: generalist repositories (e.g. Zenodo or OSF), which accommodate data across disciplines; domain-specific repositories (e.g. DaSCH), which often provide discipline-tailored curation, and institutional repositories (e.g. ETH Research Collection, BORIS). This report focuses on the national FAIR-compliant repositories in the SSH domain: primarilyas SWISSUbase, - and DaSCH.
Digital data platforms	Digital data platforms in the SSH field encompass several key initiatives, including Dodis, Année Politique Suisse, and the Historisches Lexikon der Schweiz or haller.net / République des Lettres. These platforms are designed to aggregate comprehensive, accessible, and high-quality data on specific topics to researchers, educators, and the public. They support a wide range of academic and public interests, contribute to foundational research and foster a deeper understanding of Swiss history, politics, and international relations. Typically, these data platforms are long-term project, which, in addition to data provision, offer professional expertise on the respective research domain.
ERICs with Swiss participation[13]	ERICs are legal entities established by the European Union to facilitate the development and operation of RIs of European interest. Switzerland actively participates in several ERICs, contributing to the advancement of research and innovation across Europe. Here are some key initiatives for the SSH domain: CESSDA (Consortium of European Social Science Data Archives), CLARIN (Common Language Resources and Technology Infrastructure), and DARIAH (Digital Research Infrastructure for the Arts and Humanities).
Longitudinal studies and surveys	Longitudinal studies and surveys are long-term projects that involve recurring surveys on the same variables over long periods. These studies are crucial for understanding changes and developments in various areas of society, such as health, education, and the labour market. Because of its comparatively high socio-economic and political stability, the sense of duty and trust in science of its population, Switzerland is particularly suitable for such long-term studies. Some examples are SHARE-Survey of Health, Ageing and Retirement in Europe - Switzerland; Swiss Household Panel (SHP) of FORS; Swiss Job Market Monitor (SJMM); Swiss Labour Force Survey (SLFS) of the FSO; the Swiss participation in the Programme for International Student Assessment (PISA); Transitions from Education to Employment (TREE) or Science Barometer Switzerland.

13 https://www.sbfi.admin.ch/en/switzerlands-participation-in-european-research-infrastructure-networks

Scholarly digital editions	Scholarly digital editions are curated collections that serve as critical resources in the humanities. They provide researchers with access to meticulously edited and annotated texts in digital formats. Digital editions often include transcriptions, facsimiles, and extensive metadata, making primary sources more accessible and usable for academic research. Some examples are Lavater edition, Bernoulli-Briefwechsel, Nachlass Karl Barth, Online Edition of the Paippalada Recension of the Atharvaveda.
Sensitive government data	The Swiss government collects data on individuals and firms in administrative processes at all levels of the federal system.[14] Such sensitive individual-level data is pivotal for cutting-edge empirical research in the social sciences. It is also the basis for impact assessment, policy evaluation, and evidence-based public policy in general. The FSO has exclusive authority to link and pseudonymize data from various government and other sources, constituting a key data infrastructure for social science research in Switzerland.

Again, the above-listed domains helped the TF structure the landscape prior to the analysis. Regarding division of labour, the TF members distributed information and data collection along the different sub-clusters. It is important to note that this remains a methodological tool. Grouping initiatives did not result in clear-cut sub-clusters but overlaps became more evident.

2.3 Data collection

Data collection was a heavy workload of the TF methodology and was distributed across the TF. TF members took over responsibility for data collection in their field of expertise. As the composition of the TF was centrally founded on representation from different disciplines and sectors of the SSH landscape, this ensured multi-perspective analysis of the landscape.

2.3.1 Factsheets and Dashboards

The questions raised by the mandate (*see Section 1.8*) are answered mainly through secondary data such as the initiatives' websites, and their yearly reports. For some initiatives, the information is supplemented by data from interviews conducted with representatives or experts (*stakeholders; see section 2.3.2*).

The TF has compiled a Factsheet for each investigated initiative that provides essential details and key facts about the initiative. The factsheets serve as a reference guide for the TF throughout its analysis of the Cluster and contain the following information: type of research data handled, ORD services and infrastructures, access policy, beneficiaries/users, development plan, governance model, funding, alignment with the National ORD strategy, involvement of the targeted research communities, and coordination with national and international initiatives.

The information gathered from initiatives in Factsheets is showcased in the final report through Dashboards and only for initiatives that show a prominent character and / or are presented as case studies in the analysis at this stage (*Chapter 3*). Examples of such Dashboards can be found in *Annex 4*. Dashboards facilitate cluster analysis and comparison between initiatives by offering a condensed overview of initiatives' main characteristics.

2.3.2 Interviews

As already sketched above, interviews were conducted at two different stages of the work: on the one hand, during the data collection phase, when desk research provided insufficient information (via the website or secondary documentation), on the other, to address transversal questions regarding future developments, need for coordination, role of AI by service providers.

Firstly, each TF member was responsible for compiling the necessary information about the selected initiatives. Moreover, they were to obtain validation from the relevant stakeholders and supplementary information requested in informal interviews and exchanges. As stated, this was a continual process, taking place not only during the data collection phase but also during and after the stakeholder event in September 2024 (*see section 2.4*).

Secondly, provided that the aggregated information left gaps regarding aspects pertaining to the scope of the mandate, semi-structured interviews were conducted (*see Annex 3*). These included an interview focused on

14 There is also much other important government data that are not sensitive such as voting results, geographic information, meteorological data, or highly aggregated data. Such non-sensitive government data can be openly shared without privacy concerns and FAIR access is successfully approached by initiatives like opendata.swiss.

This report focuses solely on access to *sensitive* government data because it is of particular importance for research in the SSH domain.

repositories conducted with representatives of SWISSUbase, DaSCH, and OLOS; an interview focused on the role of the national nodes of CLARIN and DARIAH and their coordinative role in advancing alignment and integration of key infrastructures; and an interview focused on the role of university libraries.

2.4 Data analysis – The concept of perspectives

After collecting data on the various selected initiatives and first interpretations on the level of the methodological tool of the "sub-clustered" landscape (*see section 2.2 and table 1*) resulting in Factsheets and Dashboards, this report aims to make sense of this information at the level of Cluster 2: What are the positions of these initiatives, and what does the provision of ORD infrastructures and services look like within the cluster? The TF used perspective-relevant guiding questions for that goal, closely linked to the perspectives that the TF HLS applied. Such perspectives serve as a viewpoint, a "lens" among others on Cluster 2.

TABLE 2
Analytical perspectives for Cluster 2

Perspectives
Open Research Data vs. Open Data for Research
Financial dimension: temporary vs. long-term funding
Findability
Accessibility
Interoperability
Reusability
Identified professional and infrastructural emergent networks (national coordination)
International alignment (coordination; EOSC & initiatives outside Europe) (international cooperation)

All perspectives were identified based on key questions from the mandate, which are present in the Factsheets and Dashboards. Additional perspectives emerged from discussions with stakeholders, highlighting activities that significantly impact the entire cluster. This fact emphasizes that the perspectives employed by the TF do not represent exhaustive, definitive overviews of the cluster; rather, they are used as tools to guide discussions and improve collective understanding of the cluster.

It is important to note that perspectives, like any modelling exercise, provide a schematic and simplified view of reality. Nonetheless, they remain a valuable analytical tool for the StraCo and the stakeholder community in a landscape analysis of clusters.

Unlike the report on Cluster 1, HLS, the present report does not provide exhaustive visualizations for each perspective. The main reason is that the FAIR principles, as the main analytical perspectives, are each discussed from different vantage points (e.g., technical vs. cultural vs. financial aspects) depending on data source (*see section 3.1*), subdomain or discipline, and the many RDIs unique characteristics. While some observations are illustrated graphically, the TF holds that reducing the perspectives to single visualizations runs the risk of unjustifiably distorting the complexity of the landscape.

2.5 Stakeholders involvement in the process

Stakeholders are involved in the process through three different activities:

1. **Feedback on Factsheets.** To ensure the quality and actuality of information, each Factsheet underwent review by the corresponding initiative representative and was then validated by the TF member who assumed responsibility for the individual factsheet[15].
2. **Interviews.** To better understand the dynamics and relationships within the cluster, and to discuss specific topics (for example the use of AI by repositories), the TF has conducted three interviews with representatives from initiatives and individual experts, either because they are major players in the cluster or because their roles required clarification. There have been two interview phases (*see Annex 3*).
3. **Stakeholder event.** On September 12, 2024, an event convened 40 stakeholders (*see list of participants in Annex 2*), to present the TF's work and solicit feedback on the proposed cluster perspectives. The stakeholders actively engaged with the TF, with preliminary results and perceived potentials and challenges structured along the sub-clusters (*see 2.2*), which served as a valuable tool to stimulate discussions. After the stakeholder event, the TF sustained the discussion with the stakeholders through bilateral exchanges. In addition, a collaborative and shared online form was set up and curated by the TF to receive written feedback.

15 Consequently, Factsheets also serve as a tool to see how initiatives portray themselves. The Task Force acknowledges that Factsheets may be biased due to the involvement of initiatives, presenting them more favourably. This bias is especially noticeable when assessing how initiatives align with the ORD principles and objectives of the National Strategy. To counter this, the task force has monitored the original Factsheets (before initiatives' reviews) throughout the analysis.

3. Analysis: Perspectives on Cluster 2

3.1 Open Research Data vs. Open Data for Research 21
3.2 Financial dimension: Temporary vs. Long-term funding 22
3.3 Thematic perspective: Findability 25
3.4 Thematic perspective: Accessibility 27
3.5 Thematic perspective: Interoperability 29
3.6 Thematic perspective: Reusability 30
3.7 Thematic perspective: Identified professional and infrastructural emergent networks 32
3.8 Thematic perspective: International alignment 33

The analysis of the subclusters shows that significant progress in digitalization has been made in the field of SSH. Numerous initiatives, research activities, and networks contribute important efforts. Repositories and data collections have been established, tools and instruments, methodologies and workflows, data platforms and infrastructures with a wide variety of services have emerged. Most of these initiatives show an awareness of the needs and principles of OS. As shown in sections 3.7 and 3.8 below, the relevant stakeholders carter their networking activities and national and international efforts toward alignment with the global scientific communities.

At the same time, various disciplinary, data-related, infrastructural, and, not least, governance-specific characteristics of the SSH domain and the Swiss research and funding landscape somewhat impede the development of a coherent and integrated data ecosystem. To date, there is no dedicated SSH governance body, which provides a legal, organizational and technical framework, that enables the shared use and reuse of data as is typically done within shared data spaces.[16] Such spaces facilitate data exchange and sharing in a wide variety of formats, enable accessibility through a variety of different systems and allow the use of different applications (Bacco, Kocian, Chessa et al., 2024). Like in many other scientific domains, trustworthy conditions and pathways of data sharing and exchange, participation and alignment still need to be developed. At the same time, data ecosystems that span the whole research cycle could allow researchers – and possibly also organizations and other stakeholders (authorities, government services, social groups, non-governmental organizations (NGOs)) – to manage, collaborate and exchange data in a secure and controlled environment, and compliant with the FAIR principles.

Alongside the FAIR principles, four more sections are added into the analysis chapter to characterize the cluster.

A first section addresses the differentiation between "research data" generated during and/or as a result of research, and "data for research". The latter consists of data and resources typically produced outside academic research environments.[17] Examples include ephemeral data from social media platforms, language resources from traditional media, or complex networking structures of scientifically recognized and semantically linked community platforms such as Wikidata. Moreover, digitized cultural assets and data from governmental organizations play an important role in the SSH domain. Challenges appear to be increasing when considering the types of data (including aggregated and synthetically generated data), accessibility conditions (such as access/login requirements and proprietary formats), ethical and legal constraints, and other specific factors.

Closely related to this are obstacles which focus on the discrepancy between the short-term/start-up funding logics (SNSF, Innosuisse, and swissuniversities) and the need for long-term planning and development stability. A second section addresses these issues in-depth.

The four subsequent perspectives describe and analyse the cluster through the lens of the FAIR principles. Like in the HLS cluster, in the SSH Cluster too, some areas and initiatives are already highly advanced in terms of FAIRness, while other areas show room for improvement. Here, achievements and consistency of FAIRness are assessed along the whole research data life cycle. The different perspectives provide another foundation for initiatives and RIs represented in different sub-clusters. It is questionable to evaluate a service provider such as FORS in its totality, given that various products, services, and functionalities are part of the portfolio. The same holds true with competence centers, such as DaSCH or LiRI.

The analysis concludes with two perspectives on national coordination and international alignment. Here, questions of governance structures and consortium models occur that consolidate the requirements of various stakeholders with one another. Even here, a gradual alignment of Swiss SSH RDIs with the EOSC as well as the role of the European Strategy Forum on Research Infrastructures (ESFRI) affiliations gain once again importance.

16 Since the Swiss Open Science Cloud was just recently founded, it is not considered in this report.

17 External data carries moments of fragility. It fosters dependencies, which may be difficult to control, and may imply the need to work with heterogeneous data structures on a day-to-day routine.

3.1 Open Research Data vs. Open Data for Research

SSH face the challenge of handling a vast range of different data types and formats, each tied to specific methodological, technical, and legal requirements. In addition, contextual, context-sensitive data, and those from external, partly non-academic stakeholders, follow their own rules and regulations. While FAIR principles can be promoted through research funding and evaluations guidelines, their effectiveness diminishes outside academia. Therefore, a more detailed look seems worthwhile:

Data from comparatively harmonized providers such as academic libraries and archives follow well-documented, mostly commonly accepted standards. However, varying access conditions even within certain collection segments inhibit automatized reuse. The complexity in accessing and reusing data increases with smaller or non-academic infrastructures: smaller libraries, archives, museums, media portals, social and/or cultural communities, NGOs, and not least internationally assembled data hubs provide data with their own conventions and access barriers. Missing persistent identifiers,[18] defective link resolvers, and/or customized, contextual, historical, or cultural circumstances can amplify the risk that research outcomes cannot be reused or published. While some data, such as language data from traditional media, committed to large-scale information extraction or Natural Language Processing (NLP), or online published content may be web-crawled, others need to be purchased (e.g., from rights holders, through agencies such as Swissdox, image or other third-party databases). Data pipeline architecture and processes can become increasingly complex, depending on the corpus, legal barriers and restrictions, and context.

In addition, sensitive government data, critical for basic social science research and empirical evidence in public policy decisions, face legal and ethical constraints. Even if data can be accessed, they are rarely fit for immediate processing by researchers. Their evaluation is nevertheless central to public policies after implementation, besides their use for humanities research (e.g. administrative history). Social media and community-based data (e.g. Wikiverse, data networks of defined interest groups) pose new demands in terms of data handling, quality control, and longevity, while media integration faces technical and legal hurdles, including copyright and data protection laws. At an international level, the EU has created the Digital Service Act, especially for very large online providers (e.g. Meta, LinkedIn, TikTok, etc.).

While Big Data has become more central in academic research,[19] centralized systems that manage computational resources are missing. Currently, Swiss infrastructural solutions for media and textual data are mainly developed within the CLARIN framework and will be reviewed in this report. Due to the computational resources required, centralized cross-cluster frameworks are needed to facilitate effective use of Big Data in research.

Even though AI and text processing advances offer new opportunities, they also present growing challenges. Ways must be found to integrate AI into the entire research process in a legally and ethically compliant manner: from automated search (e.g. systematic reviews), data collection (including the reuse of data for training) and data analysis to validation, publication and revision of research data and outcomes. Even when certain aspects or elements cannot be published due to legal or technical constraints (e.g., capacity limitations), they must still be documented and described in sufficient detail to ensure transparency, reproducibility, and compliance with OS principles. Hence, regulations and solutions must address and apply to AI as a (working) tool just as they apply to the use of research data, for example, as training data.

Gap analysis

The diversity of "data for research" illustrates that almost any topic and its related data can become the subject of SSH research. More importantly, stakeholders and data actors, environments and formats need to be identified to ensure connectivity in both directions: from researchers to external stakeholders and vice versa. An interesting future question might be what form of openness will be achieved through data exchange in the different domains. Beyond technical challenges in providing and handling interfaces, access regulations inhibit FAIRness. At the same time, limitations are necessary to enable data routines of trust and security.

Conflicts between copyright protection, economic interests, knowledge security and the FAIRness of data and methodologies of research are commonplace.

18 While both humanities and practice-based disciplines recognize works of art as data, a standardized PID system are still missing. This issue extends to other fields as well. However, any born-digital resource without a recognized PID is a missed opportunity.

19 Regarding the NRP 75 program see: https://www.nfp75.ch/en/j8UeVFxtxCpBGccS/news/the-programme-resume-of-nrp-75-big-data-is-now-available

In this context, societal and cultural interests, such as participation and inclusion, can act as drivers for greater transparency, openness, sharing and collaboration. The SSH domain may offer qualified contributions to scientifically based data economies. While the EU has created the Digital Service Act, which allows end-users to gain information about their private data at large online providers such as Meta, LinkedIn, TikTok, etc., the European Research Council (ERC) is currently envisaging tools to promote the use of such data in a research-related environment. The role of the Swiss SSH domain may be discussed in this field.

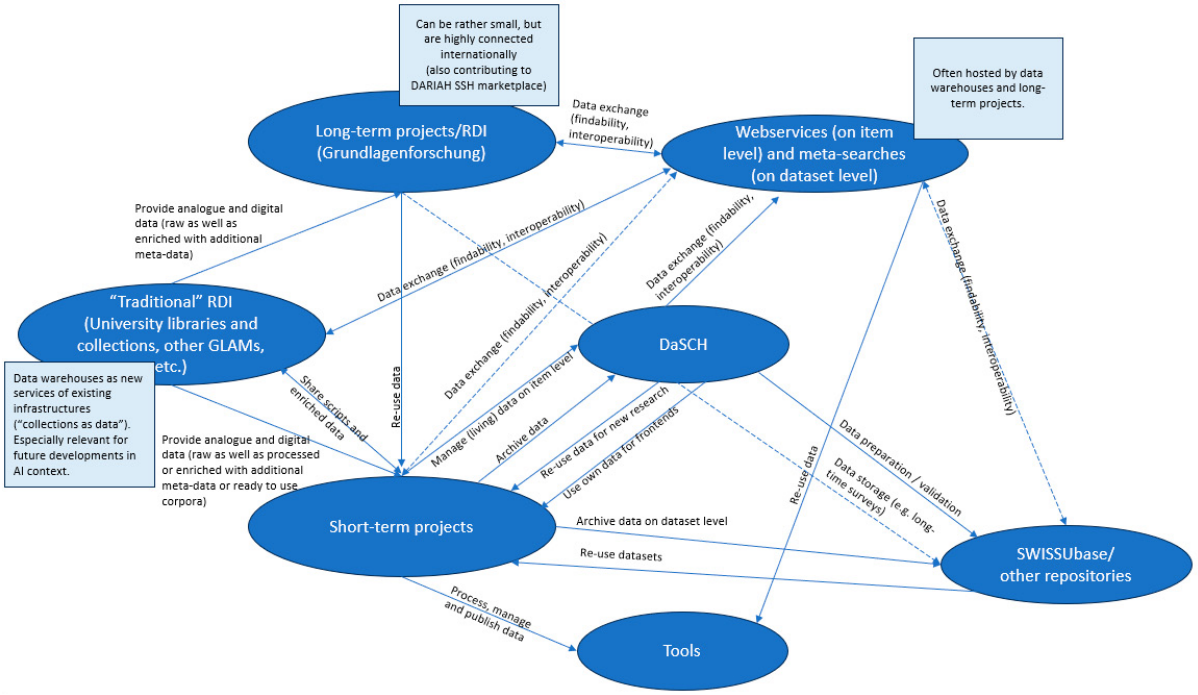
3.2 Financial dimension: Temporary vs. Long-term funding

Challenges arise from opposing logics of small- versus large scale projects, short versus long term involvement, and project-focused versus permanent infrastructural needs. These challenges encompass two distinct but closely connected aspects. On the one hand, the disparity between small and large initiatives, along with their core functions and networking structures, is visualised in *Figure 2* below. On the other hand, the imbalance becomes apparent when looking at financial resources. Taking a closer look at the financial and temporal scope, one can state that short-term, data-driven research projects are typically funded by SNSF. They usually rely on external, robust, shared data infrastructures.

By contrast, specialized long-term projects, such as Dodis (Diplomatic Documents of Switzerland), Inventar der Fundmünzen, the surveys conducted by FORS, and long-term editions, develop their own RDI throughout the entire research cycle. SERI traditionally funds these latter projects through the Swiss Academy of Social Sciences and Humanities (SAGW) or the SNSF and notably combines both research expertise and RDI development. Many of these long-term projects also sustainably convert raw data from libraries and archives (e.g., Historisches Lexikon der Schweiz) institutions or governmental sources (e.g., Année Politique Suisse, Dodis) into research data. Similarly, data production and analysis tools (such as nodegoat, Geovsitory, or heurist) provide the necessary infrastructure for short-term, data-driven, or data-producing projects, enabling them to create publishable and reusable data. These tools function as service frameworks rather than research projects in the narrow sense. Consequently, they are often only indirectly funded through SSH research projects or other small-scale funding instruments like PgB (e.g., in the case of the corpus infrastructure, Swiss-AL or Geovistory).

Funding constraints or instability present a broader challenge for long-term projects. While basic maintenance is sometimes provided by SAGW or host institutions, strategic advancements become difficult. Likewise, short-term SNSF-funded projects that developed specialized data platforms may struggle to secure a sustainable, long-term perspective. Some of these SSH platforms or databases potentially could be ingested into DaSCH. While this may be the right solution

FIGURE 2
Perspective on temporal dimension



in many cases, it does not meet the needs of all platforms and their users and might involve undue efforts, e.g., where platforms feature specialized user frontends.

Stakeholders curating cultural assets, such as libraries and archives, are tasked with the long-term preservation of their data. Aside from this core task, which is not part of the current report, they offer special services dedicated to research and have research communities as a prior target group. Their development is often initially funded, e.g., through PgB programs, but then continues to be developed. Therefore, dedicated repositories and metasearch services of individual institutions such as the Federal Institute of Technology Zurich (ETHZ) or the National Library and consortia of these organizations need to be mentioned. Due to their mixed financial resources (combination of core funding, institutional investment, occasional seed funding for innovation), it seems difficult for stakeholders to provide exact numbers. Ongoing costs and core funding are not included in the table below. Table 1 focuses on special or additional services and infrastructures. The same applies to the number of staff involved in the work. As people often work on different products or services, it can be challenging to identify surplus actions. Additionally, some stakeholders drew attention to the difficulty in clearly delineating the scope of individual initiatives, which makes it challenging to provide precise funding information. Overhead costs allocated to initiatives often lack transparency. One could work with the following estimates (*see Tables 3-4*), although one should use the numbers with caution.

TABLE 3

Funding structure for Competence Centers and Consortia, Data production tools for researchers, and Longitudinal studies and surveys

	2021–2024			
Subcluster Name	**FTE per year**	**Total budget (Mio CHF)**	**Hosting Institutions**	**Funding Institutions**
Competence Centers and Consortia				
FORS (whole activity)	47.3 (36.4 by SNF)	54.34	UniL	SNSF (64%), UNIL
DaSCH	17	12.93	UniBas	SNSF (62%), UniBas
LiRI	17	10.2	UZH	UZH, Users, SNSF
SWISSUbase	7.5	3.4	FORS, UniL	SNSF through FORS, UniL, UniNe, UZH
LaRS	6	0.8	UZH	UZH, University Library Zurich
Data production tools for researchers				
nodegoat		0.1	UniBe, UniBas, DaSCH	UniBe, UniBas, DaSCH
geovistory		0.35	UniBe	UniBe
transcriptiones		0.1	UniBas, ETH Library Lab	Fonds zur Förderung von Lehre und Forschung der Freiwilligen Akademischen Gesellschaft Basel (FAG), Max Geldner-Stiftung, Dr. H. A. Vögelin-Bienz-Stiftung
OMEKA-S		0.08	UniBe	UniBe
Heurist		0.07		
Longitudinal studies and surveys				
FORS (only Surveys)	30	20.284	University of Lausanne	SNSF (70%), UNIL
Survey of Health, Ageing and Retirement in Europe - Switzerland	4.2	4.3	University of Lausanne	SNSF (85%), FORS
Swiss Human-Relations Barometer	1.5	0.662	University of Lucerne	SNSF (100%)
Swiss Job Market Monitor (SJMM)	3	1.2	University of Zurich	SNSF (80%)
The Zurich Project on the Social Development from Childhood to Adulthood: Phase VI	2.5	2	University of Zurich	SNSF (50%)
Transitions from Education to Employment	5.5	6.4	University of Bern	SNSF (90%)
ZEPPELIN Longitudinal Study - Early Childhood Intervention, 2. Follow-up	2.2	1.783	University of Teacher Education in Special Needs Zurich	SNSF (50%), HfH, Foundations

TABLE 4
Funding structure for cultural assets, metasearch services, and linked open data

Subcluster Name	FTE per year (2021–2024)	Total budget (Mio CHF) (2021–2024)	Hosting Institutions	Funding Institutions
Cultural assets, metasearch services& linked open data				
Repositories				
Single Institution: ETH Image Archive; University library archives, institutional archives of regional or national importance (e.g. National Library, Swiss Social Archives, etc.) [12]	26	10.4	ETHZ, All University Libraries, NB, ZB, Swiss social Archive, ZHB Luzern	Federal Funding, SNSF Funding, Host Institution, User/ member fees, Third Party, Partnerships, private Sponsoring
Consortia: e-manuscripta, e-rara, e-periodica, e-codices, Zentralgut [5]	16	8.6		
Meta-catalogue / Metasearch Systems				
Single Institution: memobase, metagrid, Open Data Navigator, archives-online	7.2	5.62	Memoriav, Switch, UB Basel, Partner Institutions in consortia	Federal Funding, SNSF Funding, Host Institution, User/ member fees, Third Party, Partnerships, private Sponsoring
Consortia: Swisscollections	8.4	2.33		
Linked Open Data: Wikidata, Wikimedia Foundation, OpenData.ch, Github, Hugging Face	unknown		Community	
Total	57.6	26.95		

Table 4 shows that challenges in cultural assets domains include the rapid growth of data volume due to digitization, the demands of OS and research, the rise of AI, threats posed by cybercrime, as well as shifting usability and participation expectations from a diverse global audience. To address these swiftly changing demands, stable funding for key GLAM institutions (Galleries, Libraries, Archives, and Museums) is needed.

Gap Analysis

Despite their critical role and long-term responsibilities, libraries and other heritage institutions curating cultural assets – due to their institutional embedding outside the SERI context – often lack the necessary support to meet their research and operational needs. Research funding does not adequately address their requirements for continuity, maintenance, and coordination. Cultural agendas further fail to allocate resources for developing, networking, FAIRification and maintaining existing cultural stocks and services. As the demands of OS and the challenges of digitization grow, these gaps place additional strain on institutions. To balance the tension between large-scale infrastructures and smaller, external stakeholders' databases, which guarantee flexibility and dynamism in the field of SSH, technical, cultural, and financial incentive schemes are required to bridge the gap.

RDIs such as Dodis or Inventar der Fundmünzen are long-term funded to cover maintenance and operational tasks. However, depending on their institutional setting, applying for SNSF funding can be challenging. Therefore, these RDIs often face challenges combining their long-term tasks with project-based research on specific topics (Grundlagenforschung, Research & Development (R&D)). Other infrastructures, data platforms, such as Swiss-AL, or software like heuristic or nodegoat face sustainability challenges as they currently often hinge on short-term small-scale funding instruments.

In the case of the FSO, research access of government data remains an endeavour that is currently underfunded and understaffed, with no dedicated personnel to handle the task. Access mechanisms are neither streamlined nor transparent, failing to meet the FAIR principles and secure data sharing. Substantial investments are required to ensure research access becomes efficient and sustainable.

3.3 Thematic perspective: Findability

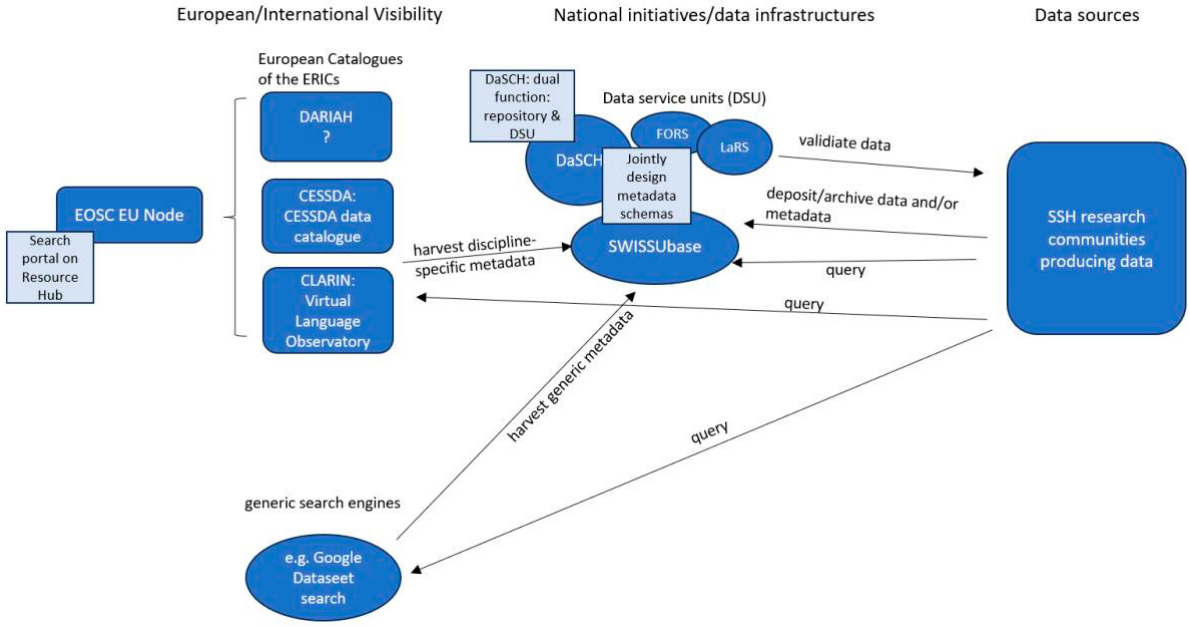
Findability, a key characteristic of FAIRness, is achieved at multiple levels in the SSH domain. At the metadata level, the analyzed databases and RDIs ensure findability through well-known, unique, and persistent identifiers (PIDs).[20] In addition, the RDIs emphasize the need for rich and standardized metadata.

As indicated in *Figure 3*, metadata is shared among various stakeholders, third-party services, and registries, which improves findability in broader contexts. SWISSUbase ensures the findability of most, if not all, longitudinal studies and surveys. It also relies on joint efforts with data service units such as LaRS (language data), DaSCH (humanities), and FORS (social sciences) to design metadata schemas tailored to domain-specific needs and standards. This approach enables metadata harvesting by European catalogues and their connection to the respective ERICs.

For language data, metadata from SWISSUbase is harvested via LaRS and distributed through the Virtual Language Observatory, a key component of the CLARIN-EU infrastructure. CLARIN-EU's centralized content search systems enable searches across all connected language corpora. At the national level, an ongoing swissuniversities PgB-funded project,[21] under the umbrella of CLARIN-CH, seeks to integrate the Swiss language resources of Swiss-AL and LiRI into this tool, and to implement a national federated content search system for Swiss language corpora.

Furthermore, like language data, SWISSUbase enables the findability of social sciences projects and data deposits by registering them in the CESSDA Data Catalogue.

FIGURE 3
Findability and visibility of (Open) Research Data



20 In addition to Handle, DOI, URN, and Archival Resource Keys (ARKs), as probably the most widespread, global accepted linkage keys which are used by most government agencies at municipal, cantonal, and federal level (e.g., social security number, Enterprise Identification Number) are applied.

21 https://www.liri.uzh.ch/en/projects/Completed-Projects/FAIR-FI-LD.html

It is important to note that, due to the SSH domain's disciplinary diversity, not all ORD will find a home in the European catalogues. The SWISSUbase/ERICs framework may be a good fit for data-heavy/data-driven SSH disciplines that benefit from data harmonization and integration into catalogues that facilitate mapping and clustering. Generic repositories, such as Zenodo and the Open Science Framework,[22] offer low-threshold alternatives for simpler datasets.[23] All these repositories – if FAIR-compliant – are typically listed in re3data, openDOAR, or similar registries, which allocate scientific repositories and data infrastructures through globally discoverable databases.

Metasearch systems and cross-domain services, such as metagrid.ch and the prospective SWITCH Open Data Navigator (ODN),[24] play an important role in improving findability. Metagrid.ch[25] provides an interlinking infrastructure that makes shared personal entities machine-actionable and visualizes references. Switch's ODN, meanwhile, is designed to focus on cross-cluster solutions. Both services integrate external data, including normed and sometimes nested datasets from Wikidata and other standardized data sources (geodata, authority data, etc.).

In terms of cultural heritage data, specialized metasearch systems exist. Examples include swisscollections[26] (an umbrella portal encompassing resources like e-manuscripta, e-rara, and digitized collections from the participating libraries); archivesonline.org,[27] which streamlines search requests to assigned assets from mainly cantonal archives; Memobase,[28] which specializes in audiovisual cultural heritage; and Swisscovery,[29] a portal for bibliographic records and sources, managed by academic libraries[30]. While the above-mentioned metasearch services connect data pools from across Switzerland, the following cover more regional or institutional data pools: ZentralGut,[31] which connects mainly cultural heritage resources from academic libraries, public archives, and cultural institutions in Central Switzerland; the Swiss Social Archive in Zurich,[32] which allocates its various collections and thematic clusters; and, lastly, e-pics,[33] ETH's institutional cross-catalogue portal for images.

Another type of data portal is offered by the FSO,[34] which presents data from various administrative data providers. Various types of OGD, not limited to the SSH domain, are available through the opendata.swiss portal[35] and the SNSF Data Portal.[36] However, part of the FSO's data cannot be found directly online. Detailed lists of available variables and their quality are not publicly available, nor are many individual-level datasets from the FSO and other federal offices, which have potential for research.

In contrast to the situation with metadata, the findability (and discoverability) of data remains a significant obstacle in many cases due to missing access information. Many datasets that are not freely accessible are not even publicly mentioned. Furthermore, not all metadata for closed data that could legally be shared are published properly.

Generally, the lack of overarching search portals and the limited connection of SSH subdomains emphasize the fragmented appearance of the SSH data landscape. Unless data are integrated into a cohesive framework, communication tends to be siloed within individual platforms.

Gap Analysis

A cohesive platform for searching for the diverse types of SSH data does not yet exist, hampering the possibility of interdisciplinary research. Identifying and scraping resources remains time-consuming for SSH researchers due to diverse analogization, dispersed datasets, individualized description forms, and heterogeneous data types.

While the metadata findability within the cluster generally shows maturity, the findability of distinct data types can still be improved. Individual-level datasets from the FSO and other federal offices are difficult to track

22 https://osf.io/

23 SNSF has recently published a survey of used data repositories in the SSH domain based on self-declarations (see: https://data.snf.ch/stories/open-research-data-2023-en.html). The report admittedly considers data numerically and brackets out volume, complexity, and demand for curation.

24 https://opendatanavigator.switch.ch/

25 https://metagrid.ch/

26 https://swisscollections.ch/

27 https://www.archives-online.org/Search. Dedicated collection subset can also be accessed via pre-customized search filter at https://sources-online.org/.

28 https://memobase.ch/de/start.

29 https://swisscovery.slsp.ch/.

30 SLSP's Swisscovery still harms accessibility of OA due to operational reasons: libraries can hardly publish their records to the entire Network Zone (NZ), from where other libraries could transfer metadata. And even if OA-records are available at the NZ, they must actively take records over into their own institutional zone (IZ), if they want users to immediately access the related publications. This hinders the acceptance and spreading of OA in Switzerland.

31 https://zentralgut.ch/index/.

32 https://www.sozialarchiv.ch/spezialseiten/recherche/.

33 https://e-pics3.ethz.ch/.

34 https://www.bfs.admin.ch/bfs/en/home.html

35 https://opendata.swiss/de

36 https://data.snf.ch/.

down and assess for accessibility. A similar issue arises with digital tools and workflows, which are often only published in the context of specific research projects (e.g., via project websites).

In addition to assessment scales and quality assurance, examining the technical persistency of PIDs seems appropriate. While PIDs promise to remain accessible to the assigned data (or at least their metadata), it should be assessed whether these PIDs still resolve to the correct data over time.

3.4 Thematic perspective: Accessibility

Regarding accessibility, it is useful to distinguish between access to research tools, access to metadata, and access to data (sets). While the stakeholders surveyed rated the accessibility of digital tools and metadata as largely good, the relatively low number of published (raw) datasets prompts an examination of the obstacles that arise throughout the research data life cycle.

Accessibility (Wilkinson et al. 2016) is considered widely implemented in data production and analysis tools within the Swiss SSH landscape. Many tools support or mandate the publication of their produced or analyzed (meta)data under open licenses. Numerous data providers, platforms, and networks specify access conditions that align with the FAIR principles. For example, SWISSUbase uses SWITCH edu-ID with a two-step authentication process to secure access to the platform. However, access control for individual datasets rests with data owners, who determine the usage license for each dataset. While open datasets can be downloaded by non-authenticated users, restricted datasets require authentication, authorization, and an approval workflow to enable access. The procedure is similar with DaSCH, which focuses on curation processes customized to metadata schemas and models, and with FORS, where data owners regulate access and control the editing or modification of their datasets. Accessibility varies considerably within the subcluster of "longitudinal studies and surveys." Due to the sensitivity of much of this data, access is often restricted to authorized users, or the data is anonymized before being made available. Some studies still maintain embargo periods of up to three years to first serve the immediate research community. For example, cooperative platforms such as swisscollections offer added value through their long-term duration and sampling services. As a result, data in these cases is easily findable and can often be accessed in bulk. However, depending on the context, automated mass downloads require agreement from peers (often members of supporting associations). This can present a challenge even when dealing with copyright-free and/or public domain content, such as historical resources and digitized cultural assets.

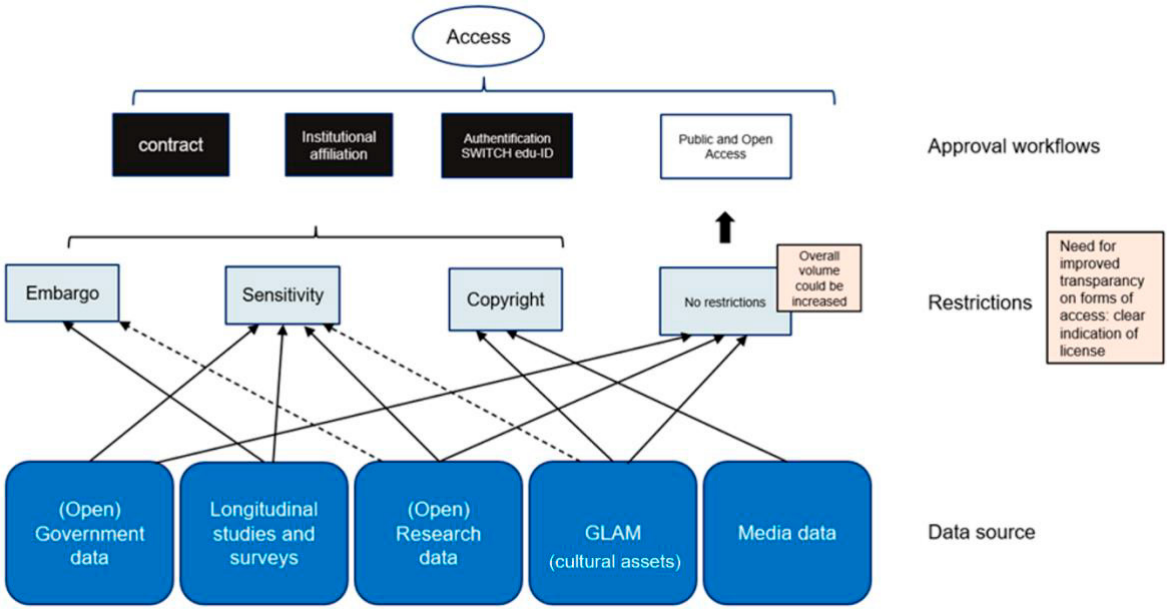
The situation differs when considering access to other types of data. For government data, access regulations and procedures vary by platform, stakeholder, and affiliation. Access is often linked to institutional affiliation.[37] Upon approval, data linkage and protection contracts are issued by the FSO. Once signed by both parties, data transmission and linkage take place. A new pilot project at the FSO aims to enable remote access to analyze sensitive data on a collaboration platform. Additionally, a project under the DigiSanté program[38] aims to foster the secondary use of government data; pre-accreditation and fast-track approval processes are anticipated to speed up data accessibility in times of crisis. However, there remains a distinct lack of transparency regarding accessibility when it comes to government data. Data is often insufficiently described in terms of whether it is accessible and to whom.

The question also arises as to whether data is accessible in bulk. If mass download is prohibited, researchers must scrape the required data in complex and time-consuming ways. Often, data has to be gathered from various platforms (sometimes manually) before the research-driven curation work of analytical consolidation can begin. The reasons for technical restrictions and thresholds are manifold.

37 Applicants from recognized research institutions (e.g., universities) or federal, cantonal, or municipal organizations can request access to linked data for projects related to research, planning, or statistics.

38 https://www.digisante.admin.ch/de

FIGURE 4
Forms of accessibility based on data type/source



Licensing and copyright issues present another challenge. While non-sensitive research data, government data, and historical resources are often openly accessible, the situation is individualized and opaque when it comes to collections of contemporary art, culture and videographic resources of research, where humans were involved in the production in terms of authorship. Here access is complicated due to copyright (intellectual property) or physical ownership,[39] but also legal aspects: For audio-visual resources, with the exception of commercials, there is a statutory right for remuneration. Right of usage is inalienable for film and video resources as long as the author's rights (ending 70 years after death) are active.[40] This means that data providers such as libraries, but also repository and RI providers must pay the Société Suisse des Auteurs (SSA) annual exploitation fees related to published video minutes, when making free video-on-demand accessible, even if the originators/authors (want to) enable free, cost-neutral use.[41] The continuously increasing volume of video resources and the rising life expectancy (duration) of the people involved are currently leading to immense, difficult-to-predict costs for anyone who wants to make third-party videos (including those of researchers) freely and publicly accessible. This hinders OS.

Access restrictions or the cost of clarifying usage rights increase with the growing distance between data-providing stakeholders and academic institutions. Data from providers with unclear or vague organizational structures, community-based entities, international organizations such as NGOs, low-funded galleries, museums, or cultural heritage organizations are often difficult to access directly. Open data portals such as GitHub, Wikidata, and especially Wikimedia Commons have emerged as important data hubs for SSH research. These platforms provide structured access to data via defined interfaces, with clearly declared access modes and licensing information.

Access to rich metadata descriptions can sometimes be another challenge. These descriptions often result from intensive research and are thus regarded as outputs or data in their own right. Unlike basic metadata, rich metadata descriptions are sometimes placed under copyright. While an open license with attribution is ideal, more restrictive licenses are often chosen, preventing the data from being offered via automated data access interfaces like Open Archives Initiative Protocol for

39 Copyright becomes especially prevalent for third-party data if the author has already passed away. Additional complexities may arise from privacy laws, political factors, and cultural or national concerns. As regards Swiss copyright law, it is important to refer to the statutory right to remuneration. Authors' right is inalienable, with regard to current video resources - also beyond the arts

40 https://www.fedlex.admin.ch/eli/cc/1993/1798_1798_1798/de, see Art. 35.

41 https://ssa.ch/de/werke-nutzen/video-on-demand/.

Metadata Harvesting (OAI-PMH). The accessibility of data can also be affected by platform-dependent usage policies and access restrictions that prohibit mass downloads.

Finally, making research data accessible requires considerable effort from researchers. If legally required documents are missing or the effort required by the storing RDI is deemed too great (due to factors like pseudonymization, anonymization, or protection of vulnerable groups), data is often not published. In case of platforms not clearly indicating whether a metadata set refers to missing data or blocked data, and explanations for such, it remains unclear to third parties whether the research data is available. Until now, closed data storage or archiving verification has largely relied on self-declaration.

Gap Analysis

Even though accessibility is provided on multiple levels within the SSH sub-clusters, data accessibility needs to be improved. It varies in terms of handling, quality and clarity in the presentation of the chosen license and data accessibility. Restriction and control mechanisms stay opaque. Conditions and the utilization of legally defined scientific exception[42] seem hardly known in the community of researchers. This is particularly evident with datasets (Perspective Accessibility) compared to metadata (Perspective Findability). A significant gap persists between those for whom OA and OS are standard practices and those who still struggle with them, avoid them, or actively limit their research outcomes. Restricted access or exclusive access for defined researchers or discipline members risks hindering interdisciplinary or cross-disciplinary research, innovation and social responsibility of research.

Improving transparency and standardization in communicating access conditions could be achieved through clear license indications of attached files, even if they are closed. Sensitive data evidently must be protected, and even references to them require careful consideration. However, this does not conflict with an open scientific data culture.

Regarding larger datasets, collections, or bulk downloads, it must be emphasized that sustainable storage and archiving practices are essential to maintaining reliable accessibility. Potential needs and financial considerations are discussed in *section 3.2*. However, complex interdependencies remain between different stakeholders. While strictly technical access restrictions appear largely solvable through scientifically accepted interfaces (*see section 3.5*), content-related barriers seem more challenging.

Other aspects, such as barrier-free access and sensitive topics like semantic drift and changing meanings, remain under-addressed.

3.5 Thematic perspective: Interoperability

Interoperability requires that data can be integrated with other data. In addition, data need to interoperate with applications or workflows for analysis, storage, and processing.[43] This is a significant challenge in a heterogenous landscape that needs to be addressed through discipline-specific initiatives.

The EOSC EU Node technical launch event[44] in October 2024 illustrates the transformative potential of interoperability. By harmonizing workflows, metadata, and data handling, interoperability ensures continuous research quality and innovation. It is developed and implemented across multiple levels of interaction, ranging from local research groups to national research clusters (*see section 3.7*) to global entities like the ERICs (DARIAH, CESSDA, CLARIN, and soon OPERAS) and other specialized interest groups, such as the International Image Interoperability Framework (IIIF).

In the Swiss SSH Cluster, fundamental requirements for interoperability are well-established. Metadata standards are generally robust, interfaces are clearly defined, and widely accepted communication protocols are in place. Higer Education Institutions (HEIs) along with entities like Switch, provide cloud-capable systems and digital environments, which are continuously enhanced in collaborations with the SSH disciplines. Moreover, there are several cross-infrastructural collaborative activities to lever interoperability. Examples include the planned integration of SWISSUbase's metadata catalogue into I14Y Interoperability Platform by the FSO;[45] the common design of metadata schemes by SWISSUbase and the DSUs; or domain-specific projects, such as the FAIR-FI-LD project in language data, involving LiRI and Swiss-AL. At an international level, efforts are widely aligned with the

42 Scientific exceptions intend to prevent copyright law from unnecessarily hindering scientific research. See: https://www.newsd.admin.ch/newsd/message/attachments/69777.pdf
43 www.go-fair.org
44 https://eosc.eu/events/eosc-eu-node-technical-launch-event/
45 https://www.i14y.admin.ch/en/home

EOSC through ERICs, but also through interest groups like the IIIF, of which DaSCH is a member – involved in a working group to define standards for annotation of 3D and suitable long-term archiving formats.

Interoperability also means that (meta)data adhere to a formal, accessible, shared and broadly applicable language for knowledge representation. The goal is to establish a common understanding of digital objects by means of standardized representation languages. Data should be readable by humans and machines without requiring specialised algorithms, translators, or mappings.[46] A notable example is the Resource Description Framework (RDF) extensible knowledge representation model. Some initiatives in Switzerland, such as SARI and DaSCH are actively working with RDF schemas. However, the implementation of such a framework can be time consuming and is often not a primary focus for researchers.

While significant progress has been made in metadata interoperability, challenges remain in terms of data interoperability[47] and workflows. Collaborative research environments such as SARI, and digital workspace-based tools TEI-Publisher, OMEKA-S, Heurist, HEDERA and many others play critical roles here. They are fully integrated within research processes and research production/publication workflows. Their importance cannot be overstated, especially when viewed alongside EOSC developments, where workflows and tools are increasingly integrated to enable global standards while addressing specialized in cutting-edge research.[48]

Gap analysis

Despite the relative maturity of the Swiss SSH landscape, barriers to data interoperability remain evident across the entire research life cycle – from production to publication and/or archiving, and from data management to digital tools and workflows development. Inconsistent workflows and varied data sharing practices pose challenges, depending on context and use case. Efforts to build trust, foster and establish sustainable structures require further attention. There are, as mentioned above, persisting technological issues. Data, if accessible, is often incompatible with required formats, standards, or granularity. Additionally, many databases lack defined interfaces for automated processing. Similar issues and tendencies can be observed within the European OS context, where the transition from development to implementation and consolidation is ongoing but slow.

A significant issue is also that most of the SSH-RI are designed as monolithic systems. However, interoperability can happen at different levels: shared usage of storage infrastructure, metadata systems, computational facilities, and frontends, as offered by IT principles such as MACH[49] architectures or Zero-Trust[50] systems, are largely lacking in Switzerland. The services from Switch are an exception here. Conversely, major centralized providers like SWISSUbase and DaSCH should evolve into multi-service providers to enhance connectivity and adaptability. Nevertheless, this process seems hampered in Switzerland by the absence of permanent financial commitments and cohesive strategies (*see section 3.2*).

Analysing the current situation seems difficult due to a lack of binding metrics for evaluation. Before interoperability can be assessed across its various dimensions, stakeholders must first agree on clear goals and objectives.

3.6 Thematic perspective: Reusability

Technically, reusability requires accurate and richly described (meta)data, as well as clear licensing information. Metadata should adhere to well-known standards and meet community needs.[51] Missing licenses information hinders both human (individual researcher) and machine-based (automated) reuse. At the same time, traceability of data reuse remains an unresolved aspect. While referencing texts through citations and bibliographies has a long tradition, systematic approaches for documenting citations and efficient methodologies for monitoring and qualifying the reuse of (raw) data and datasets are underdeveloped. Qualifying and quantifying reuse of data becomes particularly challenging at the dataset level. Here questions arise about where and how referencing should take place.

46 www.go-fair.org/fair-principles/

47 The technical limitations in terms of data interoperability are highlighted by the report mandated through the Swiss National ORD Strategy Council and conducted by the ORD Sounding Board of Service Providers: Enhancing Open Research Data in Switzerland: Analysis and recommendations from the ORD Sounding Board of Service Providers on "Data Archiving & Sharing", "User Access", "Technical Interoperability" and "Reuse".

48 https://open-science-cloud.ec.europa.eu/services

49 MACH is an acronym for microservice based, api-first, cloud-native, and headless.

50 https://en.wikipedia.org/wiki/Zero_trust_architecture

51 Wilkinson et al. (2016) define accessibility by claiming that "(meta)data "are retrievable by their identifier using a standardized communications protocol" and "are accessible, even when the data are no longer available" (see: Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018. https://doi.org/10.1038/sdata.2016.18).

The reuse of data in AI applications, for example, gives a good impression of the complexity of these requirements: quantifying and qualifying reuse, here, not only means identifying which datasets were generally used to train a model but also how (often) and for what purpose.[52] A simple reference to a dataset or algorithm source is insufficient. In addition, standards seem important to ensure that data reuse can be consistently and effectively traced at the dataset level. Monitoring systems need to be capable of detecting and tracking reuse references to enable automated identification and analysis. Furthermore, proper citations not only improve traceability but encourage higher citation rates and thus incentivize sharing.

Another topic is the integration of RDIs and other initiatives into defined research communities' ecosystems. FORS's longitudinal studies are generally well-known and integrated into scientific communities. For example, SHARE (Survey of Health, Ageing and Retirement in Europe – Switzerland) as of 2024 features almost 500 registered users, with more than 18'000 users at the European level. Users include academics, policymakers, and government bodies like the Swiss Federal Office of Public Health (FOPH). High reuse potential and proven reuse is also evident in the specialized data platforms funded by . For instance, Dodis and *Année Politique Suisse* show high reuse potential, serving both academia and professional fields. Platforms combining data with expertise see growing user benefits.

In terms of the repositories, DaSCH provides a promising reuse approach by offering an environment where users can edit and work with data directly on the platform. SWISSUbase validates data to ensure quality, boosting trust and usability.

Regardless of the earlier discussion, AI tools raise complex questions about data reuse. AI relies on large datasets, including SSH data, yet concerns about ownership, misuse, and bias complicate reuse. An ever-increasing level of data literacy is required to understand the structure and potential of large data corpora, their verifiability, authenticity, susceptibility to bias, as well as privacy, data security and potential impact.

Until now, copyright law in Switzerland makes a clear distinction between human beings and machines regarding both the protection of the works created as well as the subsequent re-use. While some countries have tried to restrict access and the usage of AI systems in a first wave of nervousness, other stakeholders are already considering how they can legally protect the output of their AI systems (copyright, patents) to better commercialize them. From a research perspective, both developments are problematic: On the one hand, research needs access to data and permission to use AI-based tools.[53] Hiding research data from reuse by AI-technologies, systems, and/or services also encourages the codification of bias: Data that is blocked for AI-systems, cannot be used to train the related models and systems. This then creates intentional blind spots in the systems and thus disturbs the DNA of next-generation models, tools, and services. Explicitly prohibiting the re-use of legally published research data by AI systems therefore promotes bias. On the other hand, there remains a certain skepticism regarding AI tools and systems.

Due to the ambivalent attitude towards data reuse with AI tools, the current disciplinary discussion lacks structure and uniformity. RDIs such as SWISSUbase and DaSCH currently do not prioritize AI-based (re)use. Cultural assets, metasearch services, and linked open datashow a wide range of attitudes towards AI from caution to experimentation, especially with historical versus protected data: some prefer a wait-and-see approach, others tend to prohibit access, while some engage in enthusiastic experimentation. The gap between the application of AI procedures to historical, public domain, non-protected data and current or protected data is once again evident. A stronger focus in the SSH domain lies on creatingenvironments for working with data without download needs.

Concerning scholarly digital editions, the training of (Large) Language Models by researchers and commercial entities was noted in the Factsheets as an unintended user group. For sensitive government data, researchers can apply for access to reproduce research. In principle, data from previous studies can be accessed under the same conditions as the original research. However, neither linked data nor the code and keys for data linking are stored long-term or archived, making reproduction more costly than necessary. Until now, there has been no requirement to store such research code in public repositories (or to store it at all).

52 A project for analysing images from historical newspapers may serve as an example. For training a model to identify bounding boxes, a test set of a different language area might be used, because it is not about semantic understanding of texts but about formally recognising images.

53 Just think of the integration of speech and character recognition systems in OCR software, the use of voice-to-text software or SQL or other support mechanisms that researchers can already understand and use with AI-based tools. Big data is difficult to imagine without intelligent digital support.

Gap analysis

Assessing data reuse remains challenging. Looking at the classical characteristics such as the availability of license information or the quality of metadata show tendencies but lack reliability. Data and measures such as key performance indicators for evaluation achievements are missing, and current information relies on self-declarations. Furthermore, traceability seems a task of growing importance.

AI technologies and systems are adding pressure on SSH fields and reusability. Although discussions have been limited, open dialogue among disciplines seems urgent. The topic is too significant to delay. Application, usage, or restrictions of AI systems and mechanisms must be conducted carefully, as there is a risk that research will exclude itself through hasty regulations. Swiss SSH research is unlikely to operate without external AI models and data. Thus, making Swiss research data reusable in a thoughtful way is essential to reduce bias, ensure context sensitivity, and maintain research quality. Or to put it another way: at some point, it will be about assessing the conversion rate of reused research data in AI systems and environments.

It can be helpful to create incentive systems that promote data skills, data ethics, and data protection. These should support the documentation, tagging, publishing, and archiving of research data across the entire data lifecycle, bridging generational, technological, and cultural boundaries.

3.7 Thematic perspective: Identified professional and infrastructural emergent networks

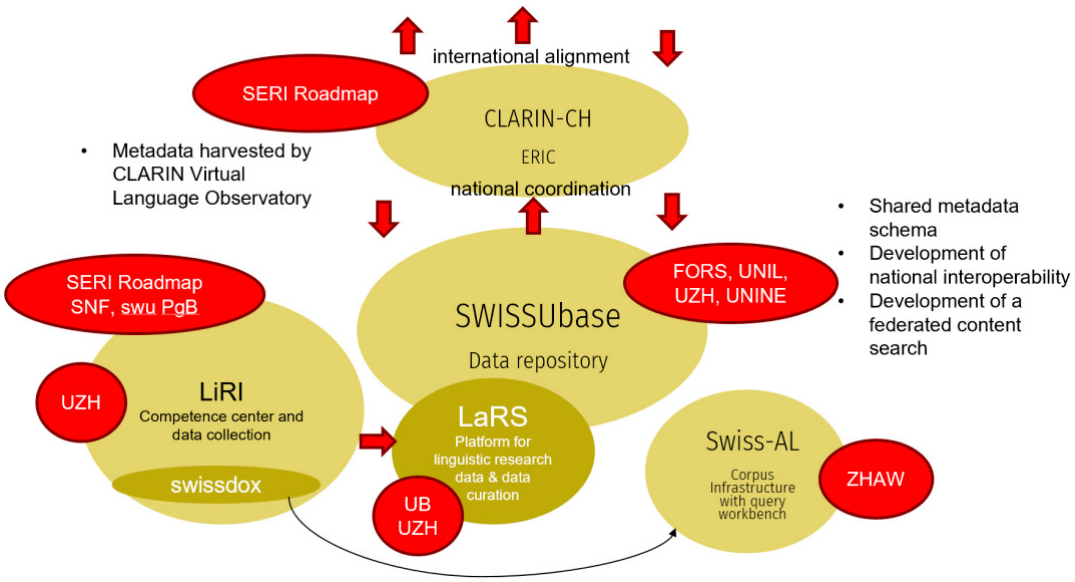
The perspective on professional and infrastructural emergent networks in Switzerland reflects the dynamics of the SSH landscape. It shows the field's heterogeneity alongside ongoing efforts in cooperation, networking and structuring, while highlighting hierarchical patterns essential for international connections (*see section 3.8*).

The newly established umbrella organization SSHOC-CH plays a crucial role in coordinating infrastructure providers and researchers across Switzerland.[54] Founded in 2024, it has already established connections with the ERICs and individual projects such as SARI, Dodis or Inventar der Fundmünzen der Schweiz, which are represented on its board. On a secondary level, the ERICs strengthen exchange and networking of researchers across universities and other communities.

Among the national ERIC nodes in the SSH domain, CLARIN-CH stands out for its well-developed infrastructural ecosystem. Under its umbrella, a robust network centered on language data has emerged. Closely related to CLARIN-CH areLiRI[55] and Swiss-AL,[56] both offering infrastructural services and research-oriented services (LiRI: data acquisition laboratory; Swiss-AL: media corpus infrastructure with integrated workbench). Language data (where legally permitted) and metadata are shared and archived through LaRS on SWISSUbase

FIGURE 5

Language data - Example of a domain-specific network



54 SSHOC-CH is organized as an association https://sshoc.ch/start
55 https://www.liri.uzh.ch/en.html
56 https://www.zhaw.ch/en/linguistics/research/swiss-al/

and harvested by the European Virtual Language Observatory. A contributing factor to the consolidation of this network were three swissuniversities projects during the Open Science program 1: UpLORD (UZH), Swiss-AL (ZHAW), and a project on interactional data (USI). Focused on enhancing adherence to ORD principles in several types of language data, the personnel from these projects formed a working group to ensure continual exchange and alignment.

Two national repositories stand out in the SSH domain: SWISSUbase and DaSCH. While SWISSUbase's scope principally goes beyond the SSH domain, it has a traditional focus on social sciences, more recently also on language data, and data from the humanities. DaSCH is intended for publishing and archiving of data that require more complex database structure, both from arts and the humanities. Data on DaSCH remain editable, meaning that DaSCH provides an environment that goes beyond mere sharing and archiving. The two repositories complement each other and are in close collaboration: In 2024, both have signed an agreement on how to work together and how to share the workload; DaSCH is a collaborating partner of the SWISSUbase consortium; DaSCH acts as a DSU (data service unit) for humanities data of lower complexity shared on SWISSUbase.

Networking activities in the field of libraries, archives, cultural, and historical collections and linked open data relate to the organizations behind the above-mentioned metasearch systems, RIs, and services, which are either managed by ETH library Zurich (e-periodica, e-rara, e-pics), University Libraries (Basel: swisscollections; Luzern: ZentralGut), or associations such as Memoriav (Memobase)[57]. Finally, potential consolidating factors, such as working groups, swissuniversities projects, as well as the role of university libraries, respectively institutions deserve being mentioned for their effort in developing structures for networking.

Gap analysis

At the level of individual initiatives, such as the longitudinal studies funded by SNSF, coordination is lacking. Currently, there are big disparities in sample sizes, budget and data FAIR readiness level. Most initiatives are embedded in a specific faculty of their host institution and are dependent on cantonal budget planning and on the engagement of key individuals, making inter-/cross-disciplinary collaborations beyond faculty and cantonal borders extremely challenging. In addition, many studies have been conducted at the cantonal level (most of which have not been captured and analysed in the present report).

Obviously, all studies should produce FAIR data and follow similar governance rules to ensure highest scientific value and methodological standards and stimulate interdisciplinary research and innovation. However, there are gaps in data access and publication and obstacles that impede immediate networking and/or reuse.

The same observation applies also to scholarly digital editions and the RSIM Digital Center, which all present high scientific value. While collaboration with DaSCH might ensure the implementation of FAIR principles and sustainable data archiving, concrete plans are missing for developing their digital platform.

A sustainable data preservation and user environment must be funded for all beyond 2028. Currently, there are first steps undertaken to coordinate digital editions on a local level (for example at UniBE or at UZH in the Zentrum für Digitale Editionen).

3.8 Thematic perspective: International alignment

The analysis shows that there is already a well-established and effective cooperation between Swiss nodes and the ERIC in the SSH field. These include DARIAH, CLARIN and CESSDA. While the organization and status of the respective Swiss nodes differ, they share the goal of coordinating activities in their respective fields and offering standards and services to ensure international interoperability and findability of research data. These ERICs are already operational, while the overarching EOSC is still work in progress (but gained momentum in the second half of 2024).

DARIAH-CH is hosted by DaSCH.[58] The current DARIAH-CH consortium was constituted in 2021. The National DARIAH Coordination Office Switzerland officially joined DARIAH ERIC as observer in 2021 and joined as full member in 2023. The DARIAH ERIC is a RI designed to enhance and support digitally enabled research and teaching across the Arts and Humanities. It operates as a network with a small number of jointly managed infrastructures, such as the DHCourse Registry, DARIAH Campus, and various working groups.

57 Swiss Library Service Platform as host of the central cataloguing and discovery system swisscovery for Swiss academic libraries, is organized legally as a corporation (AG)

58 https://www.dariah.ch/

CLARIN-CH is hosted by LiZZ (Zurich Center for Linguistics).[59] It is a pan-European RI that aims to make all digital language resources and tools from all over Europe accessible through a single sign-on online environment. The consortium CLARIN-CH was founded in 2020 and joined the CLARIN ERIC as an observer in 2023. Full membership is aspired in 2025. Following the SNSF's recommendation, SERI decided to include CLARIN and DARIAH in the 2023 Swiss Roadmap for RIs.

CESSDA-CH is hosted by FORS[60] and has been granted full membership in 2023. FORS is actively involved in different CESSDA ERIC projects related to archiving and RI as a task member.

The European Commission launched the first EOSC EU Node services in October 2024. Currently, 27 million researchers in theory have automated access to 127+ million scientific resources, cloud storage, interactive notebooks, virtual machines, software, workflows, personal working spaces, interactive spaces for research teams, secure data sharing, large file transfer, supercomputing resources. The resources hosted by the ERICs –and hence of SWISSUbase, DaSCH, LaRS, FORS – are already integrated into the EOSC EU reference node.

Besides these three national ERIC nodes, several specialized RDIs (often funded by the SAGW) have established successful cooperation with international partners. Examples include Dodis, which is a member of the International Committee of Editors of Diplomatic Documents and *Inventar der Fundmünzen der Schweiz* (Swiss Inventory of Coin Finds), which aligns with the nomisma.org project, an international collaboration outlining formalized ontologies to provide stable digital representation of numismatic concepts. Another example is the *Schweizerisches Idiotikon*, which maintains bilateral cooperations with German lexicographic projects and currently holds observer status in the European Lexicographic Infrastructure consortium (ELEXIS), which, in turn, is linked to the CLARIN and DARIAH ERICs. Notably, many of these specialized infrastructures are aligned with domain-specific infrastructural initiatives of an international scope not limited to Europe. This ensures a more diverse international connectedness of data spaces and should be further encouraged. Through the European Collaborative Cloud for Cultural Heritage (ECCCH)[61], libraries, archives, and other cultural heritage institutions are related to data spaces such as European Data Space for Cultural Heritage/Europeana[62], Archives Portal Europe[63], and the EOSC.

Regarding sensitive government data, many countries around the world offer access to individual-level administrative data through one or several micro data centers at either national statistical offices, national science foundations or dedicated government agencies: Austrian Micro Data Center (AMDC), 41 Research German Data Centers (FDZ) accredited by KonsortSWD, CASD France (Secure Data Access Center), Microdata Services of Statistics Netherlands (CBS), Forskningsservice of Statistics Denmark, Statistics Norway Microdata, MONA (Microdata Online Access) of Statistics Sweden, Finnish Social and Health Data Permit Authority Findata, Administrative Data Research (ADR) UK, U.S. National Secure Data Service (NSDS). There is no formal collaboration of the FSO with micro data centers in other countries.

International networks are particularly important where the research community in Switzerland is too small and specialized to be able to network meaningfully and where other international players are better positioned in terms of networking for various reasons. Examples for music are RIs such as Musiconn[64], for theology Index Theologicus,[65] or for legal sciences, where OS topics seem particularly difficult to perceive, OpenRewi.[66] In this context, the national research infrastructure clusters from Germany,[67] for example, should also be mentioned, as there are often internal disciplinary networks that enable connection to the ERICS.

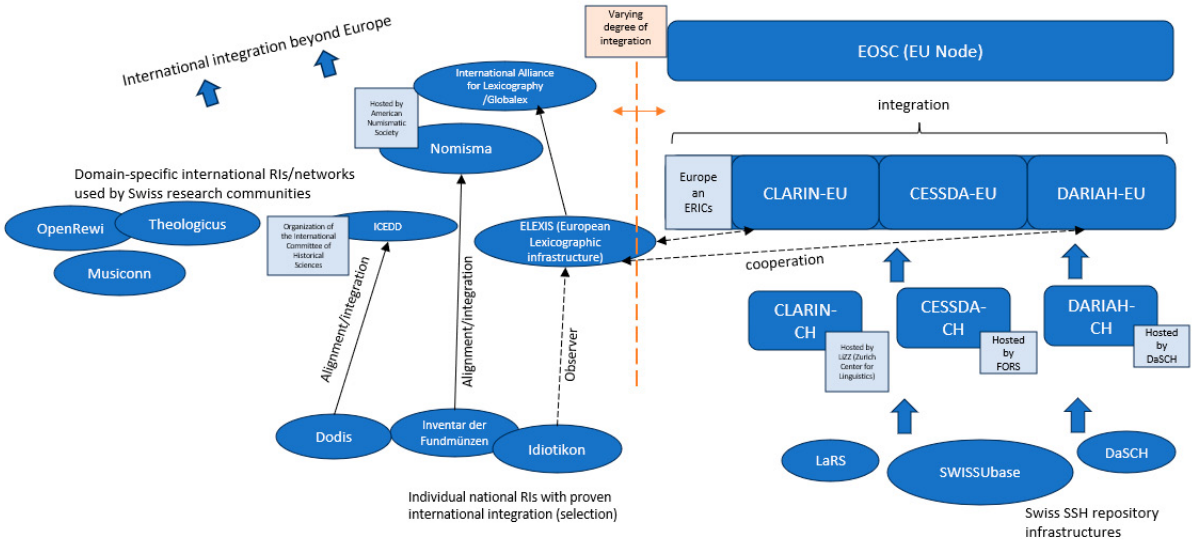
59 https://clarin-ch.ch/
60 https://www.cessda.eu/About/Consortium-and-Partners/List-of-Service-Providers/Switzerland-spid1925
61 https://research-and-innovation.ec.europa.eu/research-area/social-sciences-and-humanities/cultural-heritage-and-cultural-and-creative-industries-ccis/cultural-heritage-cloud_en
62 https://pro.europeana.eu/page/common-european-data-space-for-cultural-heritage
63 https://www.archivesportaleurope.net/
64 https://www.musiconn.de/
65 https://ixtheo.de/
66 https://openrewi.org/
67 https://www.dfg.de/en/research-funding/funding-initiative/nfdi/funded-consortia

Gap Analysis

Despite strong cooperation between Swiss nodes and the ERICs in the SSH, significant gaps remain. The integration of Swiss resources into the EOSC is still incomplete. As EOSC continues to develop, the role of Swiss nodes remains undefined. This limits the seamless access and utilization of Swiss contributions within European frameworks.

In terms of sensitive government data, the FSO currently lacks the formal collaborations seen in other countries, such as Austria's Micro Data Center or Sweden's MONA. Without secure and efficient cross-border collaborations that regulate and facilitate access to individual-level administrative data, cross-border studies leveraging Swiss resources may be hampered.

FIGURE 6
International alignment



4. Discussion

4.1 Recommendations 37
4.2 Priority areas 40
4.3 Outlook and further steps 40

The strengths of the current infrastructures, such as SWISSUbase and DaSCH, lie in their maturity, technical expertise, and adherence to FAIR principles. Emerging networks like SSHOC-CH and Swiss nodes in ERICs (DARIAH, CLARIN, CESSDA) also show strong potential for both national and international alignment. However, there are notable challenges, including governance gaps due to a lack of centralized coordination, funding risks from short-term funding models and data access barriers caused by complex legal, technical, and ethical frameworks. These issues restrict access to government microdata and cultural assets, limiting their research potential.

Several coordination opportunities exist to address these challenges. Strengthening national coordination involves consolidating national SSH repositories under SWISSUbase in collaboration with DaSCH to eliminate redundancies and improve interoperability. Mandating SSHOC-CH to coordinate FAIR data activities across the SSH domain can ensure consistent governance and technical alignment. Better coordination between longitudinal studies can standardize methodologies, align survey designs, and enable FAIR data sharing.

Enhancing international alignment includes strengthening Swiss participation in the EOSC by establishing nodes and integrating smaller initiatives to align with international standards. Supporting Swiss ERIC nodes (DARIAH, CLARIN, CESSDA) can help maintain Switzerland's leadership in SSH data infrastructures and foster global interoperability. Establishing reciprocal data-sharing agreements with international RIs can ensure global access and alignment.

Ensuring long-term funding sustainability involves implementing stable funding mechanisms for FORS, SWISSUbase, DaSCH, and for the scientific activities of the FSO and essential tools for data production to ensure operational continuity. Developing funding schemes to integrate exceptional short-term research outputs into sustainable platforms and prioritizing long-term support for digital scholarly editions and cultural assets repositories can safeguard research outputs and ensure accessibility.

"Donec aliquam risus id pellentesque lacinia. Sed vel commodo nunc. Curabitur vel porta lorem, eu viverra erat. In iaculis velit eu sapien vulputate vehicu vivamus elits."

Improving data access and governance requires developing transparent frameworks for accessing Swiss government microdata based on international best practices, such as the Five-Safes approach, to promote responsible and ethical data use. Addressing legal frameworks to enable controlled, ethical reuse of SSH data in AI-related applications can ensure provenance, transparency, and mitigation of bias. Integrating metasearch platforms into a unified system can streamline access to government, cultural, and research data.

4.1 Recommendations

The following recommendations address the most critical identified gaps on a general level, on the level of the clustered initiatives, as well as on the level of the analytical perspectives. The recommendations center on the following aspects: 1) National coordination and integration 2) International alignment 3) Legal frameworks 4) Funding.

4.1.1 National coordination and integration

The analysis has identified specific areas of initiatives that show a lack of coordination, which can lead to inefficiencies and redundancy. Such areas include (1) repositories, (2) longitudinal studies, (3) digital scholarly editions, (4) cultural assets, metasearch services & linked open data and (6) sensitive government data. In addition (5) SSHOC-CH is identified as a key player in the SSH Cluster with a great potential for national coordination. To fortify the national data ecosystem, targeted coordination without creating new overhead is essential. Such coordination is described in the following recommendations.

Repositories

1. Consolidate national SSH Repositories: SWISSUbase and DaSCH should be supported in their development to national repositories in the SSH domain and to foster their synergies accordingly.
2. Review the business models of repositories such as SWISSUbase and OLOS to avoid potential negative impact of a "competition" for institutional memberships and intransparent cross-financing of services between HEIs.

Longitudinal studies

1. Improve the coordination between longitudinal studies including those from the FSO and cantonal authorities, with the FORS survey department playing a leading role in ensuring methodological consistency, efficient use of resources, and standardized data sharing. This may include integrating more of these studies into FORS.
2. Make the data of all longitudinal studies available through SWISSUbase.
3. Align survey designs and methodologies across all studies to improve cross-disciplinary collaboration.

Digital scholarly editions

1. Create a specific repository-driven approach for data publication of scholarly editions (e.g., by DaSCH).
2. Ensure that exchange, connectivity, interoperability, and reuse across different platforms and tools are guaranteed. Make sure that the data remains discoverable and accessible beyond the platforms themselves.

Cultural assets, metasearch services & linked open data

1. Ensure that stakeholders that provide services to the scientific community at a national level are (technically, financially) effectively equipped to uphold the OS principles and make content openly accessible.
2. Ensure that nationally significant RDIs remain accessible to external stakeholders and community-organized networks like linked open data platforms, c.f. by developing appropriate policies and interfaces.
3. Integrate the various customized catalogues and metasearch platforms into centralized systems that harvest and share metadata of diverse types of resources (both research data, cultural and government data, resources from archives and libraries, and media corpora).
4. Ensure that the security and quality of the digital collections remain permanent

SSHOC-CH

1. Mandate and support SSHOC-CH to coordinate data activities across the SSH field to ensure consistent governance, accessibility, and interoperability.
2. Support national infrastructures to become national multi-service providers at all levels, including different software architectures, interoperability, AI and regulatory compliance.
3. Support efforts to be able to verify the objectives of the OS and good scientific data practice more sustainably.

Future financial expectations for SSHOC-CH to coordinate FAIR activities in the SSH Cluster sum up to CHF 250 000 (*see Table 6*).

Sensitive government data

1. Establish a Swiss Microdata Center at the FSO as a secure and transparent single point of entry for research access to sensitive government data from all levels of government (confederation, cantons, municipalities).
2. Implement a transparent governance structure with members of government, science community and data protection commissioner.
3. Ensure international best practices such as the Five-Safes approach, to promote responsible and ethical data use.
4. Long-term storage facility for data (linkage keys) and code to enable reproduction of research.

Future financial expectations for the suggestion of establishing a Swiss Microdata Center at the FSO sum up to CHF 2.75 Mio.

4.1.2 International Alignment

In the SSH domain Swiss nodes of the ERICs include DARIAH, CLARIN, and CESSDA. While the Swiss nodes vary in organization and status, they all aim to coordinate activities and enhance international interoperability and data findability/discoverability. Although the development of EOSC and Switzerland's contribution to it are important, the focus should not be solely narrowed on the European context. The following recommendations target specific options for international alignment.

1. Continue supporting and strengthening the Swiss ERIC nodes in the SSH domain.
2. Consider coordinating Swiss participation in EOSC by establishing nodes (with SWISSUbase being a suitable candidate due to its present position as infrastructural node of three SSH ERICs) and/or for a national EOSC node.
3. Evaluate smaller initiatives for inclusion in EOSC to ensure they align with international data models and create sustainable business models for such integrations.
4. Prioritize, support, and acknowledge international alignment of domain-specific RDIs (such as Dodis, Inventar der Fundmünzen, or the Idiotikon) in proven infrastructural and professional global networks beyond the European context.
5. Work on establishing reciprocal data-sharing agreements with other European and global RIs.
6. Technical collaboration of the FSO with data infrastructures from other countries to facilitate access to sensitive government data from other countries for Swiss researchers and vice-versa.

4.1.3 Legal frameworks

Research needs a safe and reliable legal environment that is robust against day-to-day political affectations and guarantees the objectivity and neutrality of research, its sustainability, FAIRness, and non-discriminatory openness (adhering to CARE[68] principles). Jurisprudence is correspondingly important for the development of scientific research in general, but especially in the field of SSH, where, in addition to technological improvement, political, social, and cultural developments have a direct impact on research. The following recommendations provide guidance for identified gaps in the context of legal frameworks.

1. Promote closer collaboration between government, academic institutions, and data protection agencies to safeguard privacy and responsible data use.
2. Keep the interests of science, such as limitations and exceptions, in mind when discussing and adopting new legal regulations and remove financial barriers that hinder the legally compliant availability of OA content.
3. Monitor international legal developments c.f. in copyright law, AI regulations, etc. to assess whether new provisions are needed.
4. Ensure that the accountability of under-represented groups, as well as social, political, and cultural responsibilities is reflected and carefully applied in the development of legislation.

4.1.4 Funding

To guarantee mid- and long-term sustenance and advancement of national data infrastructures and initiatives stable funding is required. Many current initiatives lack long-term financial sustainability, which jeopardizes continuity and planning security of services. The following recommendations mitigate such lacking long-term financial situations for the SSH Cluster.

1. Provide stable and long-term funding and a governance for identified leading national SSH RI like SWISSUbase, FORS, DaSCH, LiRI to ensure they can continue providing high-quality services, data curation, and preservation.
2. Recognize the tension between supporting large, centralized infrastructures and smaller, independent RDIs that offer flexibility for smaller initiatives. Promote technical, cultural, and financial incentives to bridge these gaps and become more inclusive. Consider developing funding schemes for further developing exceptional RDIs that emerge as byproducts of short-term research projects and whose incorporation into DaSCH, or other platforms is not implementable.
3. Review funding schemes of RDIs (such as Swiss-AL) that presently rely on short and mid-term programs such as swissuniversities' OS programs and improve transparency regarding overhead costs, for example.
4. Support funding models for the sustainability of digital scholarly editions and data repositories of libraries, archives, cultural assets (including linked open data), ensuring the long-term preservation of data and research outputs.
5. Consider developing funding mechanisms to support data production tools (e.g., heurist, nodegoat, Geovistory) that foster data entry, modeling, and sharing in accordance with the FAIR principles within the SSH community.
6. Guarantee sufficient resources for secure, transparent and timely research access to sensitive government data at the FSO. In the latter case, guarantee sufficient funding for junior researchers within existing or through new funding schemes.

68 Collective benefit, Authority to control, Responsibility, Ethics.

Summarized and estimated future financial needs are presented in the table (*table 6*) below.

Sub-clusters	Current funding agencies	Future needs (Mio.; estimates)
Competence Centers and Consortia	SNSF, UZH	54.9 SSHOC-CH: 0.25
Data repositories	SNSF, Institutions	16.8
Data production tools for researchers	Institutions	0.5
Digital data platforms	SAGW	8.74 (2.5 (RISM), 1.94 (SARI), 4.3 (HLS, IFS, Dodis, Infoclio))
ERICs with Swiss participation	SERI	N/A
Cultural assets, metasearch services & linked open data	SERI, SNSF, Host Institution, User/ member fees, Third Party, Partnerships, private Sponsoring	33.2
Longitudinal studies and surveys	SNSF	35
Scholarly digital editions	SNSF, SAGW	6.4
Sensitive government data	FSO	11
Total		**166.79**

TABLE 6
Future financial needs in the SSH Cluster(estimates)

4.2 Priority areas

The identified priority areas are heavily based on the recommendations formulated and presented in section 4.1. The priority areas therefore conclude the recommendations made further and along the same structure: (1) National coordination and integration, (2) International alignment, (3) Legal frameworks, and (4) Funding.

These areas highlight the need for improved coordination, sustainable funding, and enhanced international collaboration to address the identified gaps and strengthen the overall data ecosystem.

4.3 Outlook and further steps

Embedded within the National ORD Strategy, this report provides the groundwork for shaping strategic orientations to enhance coordination and efficiency within the SSH Cluster. Collaborating with the TF F, the StraCo's Coordination Group will translate these insights into a strategic vision for the cluster. This vision will be discussed and refined with stakeholders throughout 2025 and 2026.

5. Conclusion

This report investigates the challenges and opportunities in the SSH Cluster in Switzerland in the framework of the National ORD Strategy with the ambition of advancing the national agenda towards enhanced FAIR practices and an efficient, future-driven RDIs ecosystem. By conducting an in-depth examination of the current disposition of the main infrastructures and services, this landscape analysis aims to offer policymakers and stakeholders a well-rounded view on the cluster, which would serve as a basis to formulate strategic options for the StraCo's Blueprint process. However, due to its internal dynamics and heterogeneity, the landscape analysis presents a clearly framed possibly fragmented snapshot which – even if it has been carried out systematically and carefully – recognises the high likelihood of not being complete. The TF's methodology includes a thorough review of initiatives presenting the data infrastructures and services available to research communities, and how these infrastructures and services interact at different levels. This comprehensive assessment involved stakeholders at different stages in the process to ensure that the recommendations are both evidence-based and attuned to the specific needs and challenges of research communities across the SSH Cluster.

Among the key findings the report observes that while significant progress in digitization in the SSH Cluster has been made, various disciplinary, data-related, infrastructural, and, not at least, governance-specific characteristics of the Cluster and the Swiss research and funding landscape impede the development of a coherent and integrated data ecosystem. The Cluster can be characterized as highly diverse. It consists of key players such as SWISSUbase and DaSCH and many smaller, customized discipline-specific initiatives. Many of these latter predate the discussion around ORD by long years and played a relevant role in pioneering developments of FAIR and data sharing. It is pivotal to balance the benefit that arises from this multiplicity of approaches and the potential of their mutual integration. To meet this demand, involvement of all relevant stakeholders is inevitable. In this context, it should be noted that there is currently no dedicated SSH governance body, which provides a legal, organizational and technical framework, that enables the shared use and reuse of data as is typically done within shared data spaces. Mandating SSHOC-CH to coordinate FAIR data activities could ensure governance alignment.

Shared standards and technical frameworks are derived from internationally aligning the Swiss SSH Cluster with ERICs and other international initiatives. Despite Swiss nodes strongly collaborating with ERICs, there still seem to be significant gaps. And while the integration of Swiss resources into EOSC is still incomplete, also the role of Swiss nodes remains undefined, which limits the seamless access and utilization of Swiss contributions within European frameworks. In this process, supporting Swiss ERIC nodes helps maintaining Switzerland's strong position in SSH data infrastructures and foster global interoperability. Establishing reciprocal data-sharing agreements with international RIs further ensures global access and alignment.

Transparent frameworks for accessing Swiss government microdata should be developed based on international best practices to promote responsible and ethical data use. Legal frameworks need to be addressed to enable the controlled, ethical, but also financeable reuse of research data produced, generated and reused in the SSH Cluster. Integrating metasearch platforms into a unified system will streamline access to sensitive government data, cultural assets, linked open data, and other research relevant data communities to the SSH domain. The role of AI in the overall research process and the transparency of provenance, operationalization, and mitigation of bias are crucial. Upcoming challenges in AI technologies and its compliance in the whole research process highlight the need for risk classification (data protection, data security, ethical, ecological, financial etc.) to ensure the continuity of data quality, accuracy, and cybersecurity.

Stable funding mechanisms should be implemented to guarantee the operational continuity of SWISSUbase, DaSCH, FORS, longitudinal studies, and essential data tools, thus safeguarding research outputs. In addition, it is crucial to develop funding schemes that allow for the transformation of emergent data platforms, which are byproducts of short-term research projects, into sustainable platforms. Long-term support for research data repositories, RIs, and affiliated cultural assets that support research actions should be prioritized to safeguard the relevance and continuity of high-quality research outputs and their continuous accessibility.

It seems important to better understand the evolving ORD needs and introduce effective additional measures for optimizing the SSH ORD landscape development, with a special focus on cross-RI findability and accessibility. The TF further suggests conducting a longitudinal analysis of the initiatives' output and environment. It recommends analyzing impact and comprehensive international benchmarking of success factors and failures in other countries.

Embedded within the National ORD strategy, this report establishes the groundwork for shaping strategic orientations that enhance coordination, efficiency, and transparency within the SSH Cluster as part of the StraCo's Blueprint . Collaborating with the TF, the StraCo's Coordination Group is tasked to translate these insights into actionable propositions for the cluster. These propositions will be assessed and refined by the StraCo in consultation with stakeholders. One of the report's significant contributions is its emphasis on the collective value of findability and accessibility that can be generated through better data management practices. However, it highlights how improving data FAIRness in all mentioned dimensions can facilitate scientific recognition and innovation. It shows how far Switzerland's position as a leader in SSH research can be strengthened by achieving these goals and by ensuring compatibility with the diversity of its various SSH initiatives and cultures. The insights provided are intended to catalyze action among various stakeholders, including legal and government bodies, research institutions, funding agencies, and the broader scientific community.

Annex 1: (Long) list of initiatives 46
Annex 2: Participants of stakeholder event 47
Annex 3: Overview of conducted interviews 49
Annex 4: Dashboards of selected initiatives 50
Annex 5: Financial summary 61

Annexes

ANNEX 1

(Long) list of initiatives

Task Force Cluster 2 – Social Sciences and Humanities
Initiatives on the SSH landscape
Initiatives as listed in the mandate, incl. additional initiatives added by the TF SSH

Nr.	Initiative
1	FORS
2	SwissUBase
3	DaSCH
4	LiRI
5	SARI – Swiss Art Research Infrastructure
6	DARIAH
7	CESSDA
8	CLARIN
9	nodegoat
10	geovistory
11	e-codices
12	e-manuscripta
13	e-rara
14	e-periodica
15	SSHOC-CH initiative
16	Annee Politique
17	PISA
18	ÜGK
19	Lavater Edition
20	Zentrum für Digitale Editionen (UZH)
21	Connectome (switch.ch)
22	Répertoire International des Sources Musicales – RISM Digital Center; Hinrichsen Hans-Joachim
23	e-newspaperarchives
24	OPERAS
25	memobase
26	e-pics
27	scriptorium
28	swisscollections
29	archives-online
30	museums-online
31	hls-dhs-dss
32	dodis
33	ifs (Inventar der Fundmünzen Schweiz)
34	transcriptiones
35	openglam
36	Online-Edition der Paippalada-Rezension des Atharvaveda; Widmer Paul
37	Gottfried Semper: Der Stil. Kritische und kommentierte Ausgabe – Teil 2; Ursprung Philip
38	Edition der St.Galler Missiven in Text und Bild, 1400-1650; Sonderegger Stefan
39	Der Neue Ameis-Hentze. Ein Gesamtkommentar zu Homers Ilias; Bierl Anton
40	Die Kirchen Roms im Mittelalter 1050-1300. Liturgische Ausstattung und Architektur (Corpus Cosmatorum II); Mondini Daniela.
41	Gotthelf Edition
42	Haller.net / République des Lettres
43	Schweizerische Rechtsquellenstiftung
44	Anton Webern Gesamtausgabe
45	Nachlass Karl Barth
46	Bernoulli-Briefwechsel
47	Surveys, which are part of FORS (includes SHARE,
48	SHARE-Survey of Health, Ageing and Retirement in Europe - Switzerland – 2021-2024; Maurer Jürgen
49	Swiss Human-Relations Barometer; Staffelbach Bruno
50	Swiss Job Market Monitor (SJMM); Rost Katja
51	The Zurich Project on the Social Development from Childhood to Adulthood: Phase VI; Ribeaud Denis
52	Transitions from Education to Employment; Jann Ben
53	ZEPPELIN Longitudinal Study - Early Childhood Intervention, 2. Follow-up; Klaver Peter
54	Cantonal initiatives: "Offene Behördendaten" (ZH)
55	Swiss National Cohort
56	SFSO's efforts of providing an open platform for government data
57	Career Tracker Cohort Study
58	Längsschnittanalysen im Bildungsbereich (LABB)
59	PIAAC Schweiz - Grundkompetenzen von Erwachsenen
60	Checks und Mindsteps
61	e-editiones
62	TEI-Publisher
63	OMEKA-S
64	Heurist
65	HEGERA
66	I14Y Interoperability platform
67	opendata.swiss
68	metagrid
69	EOSC
70	Zenodo
71	SIKART (Lexikon zur Kunst in der Schweiz)
72	Zentralgut
73	Sozialarchiv

ANNEX 2

Participants of stakeholder event

Participants of the stakeholder event organized by the TF SSH on 12 September 2024
ATTENDED

Name	Initiative/RI
Cristina Grisot	DARIAH-CH, CLARIN-CH, SSHOC
Béatrice Joyeux	
Elena Chestnova	SSHOC
Georg Lutz	FORS, SWISSUBase; SSH ERICs
Christiane Sibille	SSHOC / metagrid
Samuel Schütz	Federal Statistical Office
Noémi Villars-Amberg	DaSCH
Sacha Zala	SSHOC, DODIS
Yann Stricker	Digital Editions, ZDE
Beat Immenhauser	CoG / SAGW
Bojana Tasic	CESSDA (Data catalogue)
Enrico Natale	infoclio
Thomas Hänsli	SARI
Muriel Heisch	SARI
Martin Stuber	hallerNet
Christian Forney	hallerNet
Pascale Sutter	SSRQ
Moritz Mähr	UNIBAS/UNIBE
Sylvia Jeney	TF SSH
Tabea Lurk	TF SSH
Kurt Schmidheiny	TF SSH
Rainer Gabriel	TF SSH
Simon van Rekum	TF SSH
Ben Jann	TF SSH
Béla Kapossy	TF SSH
Sarah Schlunegger	TF SSH / CoG
Rudolf Mumenthaler	TF SSH / CoG
Elena Simukovic	CoG
Magali Mari	swissuniversities
Christoph Moor	CoG
Sarah Platte	FHNW
Yvonne Fuchs	Uni Luzern
Klaus Rothenhäusler	Digital Linguistics Gruppe der ZHAW
Regine Maritz	UniFr
Simon Birrer	UniLuzern
Lukas Burkhart	UniBasel
Eric Decker	UniBasel
Christina Besmer	UniBasel
Isabelle Marthot-Santaniello	UniBasel
Levyn Bürki	UniBE
Lea Bühlmann	SBFI

ANNEX 2

Participants of stakeholder event

Participants of the stakeholder event organized by the TF SSH on 12 September 2024

EXCUSED

Name	Initiative/RI
Simon Gabay	UniGe
William Duba	e-codices
Nicole Graf	E-Pics (platform for image catalogues of ETHZ
Samantha Foulger	ETHZ Library
Katrin Keller	museums-online.ch
Monika Studer	swisscollections
Christian Erlinger	Zentralgut
Andrea Bertino	Connectome
Noah Bubenhofer	SSHOC, CLARIN, LiRI
Julia Krasselt	Swiss-AL
Yves Debardi	CoG
Gerard Bagnoud	CoG
Ignaz Strebel	CoG
Ariane Studer	CoG
Beat Immenhauser	CoG
Iolanda Pensa	SB Researchers
Michael Baudis	SB Researchers
Evie Vergauwe	SB Researchers
Jean-Paul Calbimonte	SB Researchers
Peter Klaver	SB Researchers
Silke Fürst	SB Researchers
Makhlouf Shabou Basma	OLOS
Jo Vergeat	Grimm Kulturmanagement
Christine Baur	ZHDK Archiv
Daria Janusa	OLOS
Rita Gautschy	DaSCH
Andrea Malits	LaRS, ZORA
Michael Gasser	e-rara, e-pics, e-periodica, GLAM
Marc Bühlmann	Année Politique Suisse
Christian Weibel	Nationale Wörterbücher
Rahel Ackermann	Inventar der Fundmünzen der Schweiz
Sonja Matter	Historisches Lexikon der Schweiz
Kaspar Gubler	Repertorium Academicum / nodegoat
Roger Favet	SIKART
Radu Suciu	Bodmerlab
Laurent Pugin	RISM
N.N. (Tobias Hodel)	Geovistory
Matthias Schmidt	Anton Webern Edition
Andreas Kränzli	e-editiones / TEI-Publisher
Ursula Caflisch-Schnetzler	Lavater Briefwechsel
Michéle Ernst Stähli	ESS (FORS is partner for CH)
Cécile Vilas	Memoriav, MemoBase
Beat Estermann	OpenGLAM, OpenData.ch

Name	Initiative/RI
Denis Ribeaud	The Zurich Project on the Social Development from Childhood to Adulthood: Phase VI
Jürgen Maurer	Survey of Health, Ageing and Retirement in Europe – Switzerland – 2025–2028
Andrea Lanfranchi	ZEPPELIN Longitudinal Study – Early Childhood Intervention, 3. Follow-up
Marlis Buchmann	Swiss Job Market Monitor (SJMM)
Bruno Staffelbach	Swiss Human-Relations Barometer
Daniela Mondini	Die Kirchen Roms im Mittelalter 1050–1300. Liturgische Ausstattung und Architektur (Corpus Cosmatorum II)
Stefan Sonderegger	Edition der St.Galler Missiven in Text und Bild, 1400
Paul Widmer	Online-Edition der Paippalada-Rezension des Atharvaveda
Philip Ursprung	Gottfried Semper: Der Stil. Kritische und kommentierte Ausgabe – Teil 2
Marion Rivoal	Resp. pôle valorisation données, Ci/DCSR
Thierry Lombardot	Resp. pôle web databses, Ci/DCSR
Anita Auer	Prés. UNIL dans DARIAH-CH
Patrick Michel	Prés. NUCLEUS, Lettres
Boris Beaude	SSP, membre comité scient. Master hn et STSlab
Daniel Oesch	Co-Dir. Centre LIVES, SSP
Lucas Burkart	UNIBAS, Digitales Schaudepot
Peter Fornaro	UNIBAS, Digitales Schaudepot, DH
Vera Chiquet	UNIBAS
Moritz Feichtinger	UNIBAS
Michael Piotrowski	UNIL
Elena Spadini	UNIBE/UNIBAS RISE
Sebastian Flick	UNIBE Data Science Lab
Peter Dängeli	UNIBe/Data Science Lab
Laura Bitterli	UZH Ad fontes
Christine Grundig	UZH Digital History Lab
Tobias Hodel	TF SSH
Madeleine Hamel	CoG
Arnold Olympio	CoG
Gilles Dubochet	CoG
Jean-Luc Barras	CoG

ANNEX 3

Overview of conducted interviews

Interview phase	Contact person	Date	Goal(s)
Data collection: validation	Representatives of initiatives selected for the analysis of Cluster 2	Continuous process; March – November 2024	1. Compilation of information for factsheets 2. Validation of compiled information
Data collection: problematisation	Representatives of national repositories	30.08.2024	1. Discussion of national repositories 2. Role and future development of national repositories
Data collection: validation	Representatives of the ETH library	19.09.2024	1. Discussion of preliminary results presented on 12.09.2024 to SSH stakeholders
Data collection: problematisation	Responsible person for SSH ERICs (CLARIN-CH, DARIAH-CH) and board member of SSHOCH-CH	20.09.2024	1. SSH RDI landscape in Switzerland and Europe 2. Future developments 3. Emergent networks

ANNEX 4

Dashboards of selected initiatives DARIAH-CH

Key role

The National DARIAH Coordination Office Switzerland officially joined DARIAH ERIC as Observer on January 1, 2021. Switzerland joined as full member on June 1, 2023. DARIAH ERIC: A research infrastructure that enhances and supports digitally enabled research and teaching across the Arts and Humanities. DARIAH is to be understood as a network with a small number of joint run infrastructures (among them the DHCourse Registry, DARIAH Campus and working groups).

Development plan

strategic focus:

- To become a nationally recognized research infrastructure and enhance European and international integration and cooperation
- Expand to disciplines outside SSH (current pilot: geosciences)
- Increase volume of deposited data and download score
- Participation in joint efforts for creating infrastructure(s) to securely manage sensitive data

Main ORD Services and infrastructure

Strong proponent of FAIR (and CARE) principles. As a network, DARIAH does not provide a dedicated infrastructure (in the narrow sense of the word) but serves as a framework coordinative node of various infrastructure initiatives.

Funding

- SERI: Annual fees for DARIAH ERIC
- SNF: Position of National Coordination Officer (through contract with DaSCH)

Governance model

DARIAH ERIC is organized through the board of directors and the general assembly (consisting of representatives of the member states). DARIAH-CH is member of the DARIAH-ERIC and represented by national representative and the National Coordinator. DARIAH's activities in Switzerland are coordinated by the Swiss National Data and Service Center for the Humanities DaSCH. The current DARIAH-CH consortium was constituted on 30 November 2021 by the Universities of Basel, Bern, Geneva, Lausanne, Neuchâtel and Zurich, the Università della Svizzera italiana, the EPFL and the SAGW/SAHSS. Later, three universities of applied sciences – SUPSI, the FHNW and the HKB – also joined the consortium.

Beneficiaries and Users

- Humanities scholars and students (except for linguistics, which is regulated by the ERIC CLARIN)
- GLAM sector

International cooperation

- Main task of DARIAH-CH: national coordination in the humanities and with DARIAH-EU. DARIAH-ERIC: Serves as international standard-setter/discussion platform.
- Coordination with other SSH domains (esp. Social sciences and linguistics) through SSHOC-CH.
- Connection to humanities research and digital approaches on the European level. Coordination of Swiss (digital) humanities centers.

ANNEX 4

Dashboards of selected initiatives
DaSCH

Key role

DaSCH provides a repository tailored to community needs, as well as an archive and virtual research environment for humanities data. Moreover, it offers workshops, training, and consulting on research data management, standards, FAIR data, and open science. As a national node of DARIAH ERIC, DaSCH collaborates internationally and fosters community building.

Development plan

Has been submitted to SNSF on April 1st 2024 for the period 2025-28

Funding

- In 2017, DaSCH was established as a national facility operated by SAGW with a mandate set out in the ERI Dispatch 2017–2020
- Since 2021, DaSCH has operated as a national research data infrastructure primarily funded by the Swiss National Science Foundation (SNSF) which is continuously further developed.
- If help by DaSCH staff is required for project specific tasks such as data modelling, data cleaning, or scientific programming, then DaSCH's standard rates for services apply

Main ORD Services and infrastructure

- DaSCH as a repository with focus on open research data. All software produced by DaSCH is open source and freely available on GitHub. The DaSCH platform is compliant with the FAIR principles.
- Besides the possibility of taking over research data of varying complexities towards the end of a research project and to model and import them into its own system, DaSCH also provides the possibility of creating the data directly on its platform during the lifetime of a research project. The data are collected and processed by researchers funded by national funding agencies or foundations.
- Current focus areas are digital editions, 3D, and interoperability

Governance model

The association consists of the following bodies:

- Assembly of Delegates
- Board of the Association
- Scientific Advisory Board
- Executive Board.

Beneficiaries and Users

- Researchers in humanities
- Members of all Swiss Higher Education Institutions.

Data types and flow

Members of all Swiss Higher Education Institutions can archive their data in DaSCH free of charge if the topic of the research project is in accordance with DaSCH's mission and if the amount of data does not exceed 500 GB. The DaSCH platform provides:

- descriptive metadata and persistent identifiers
- web-application and API, thus accessible by humans and machines
- use of controlled vocabularies, API is based on widely used open standards (Interoperable)
- licences and provenance information

National and international cooperation

- DaSCH is the Swiss national node for DARIAH and the hosting institution of the national coordination office
- Furthermore, DaSCH is member of the International Image Interoperability Framework (IIIF) and currently engaged in the respective 3D working groups to define a standard for annotation of 3D data and suitable long-term archiving formats. DaSCH is also participating in Swiss EOSC nodes
- DaSCH hosts the Bernoulli-Euler Online (BEOL) project
- Since April 2024, DaSCH is a collaborating partner of the SWISSUbase consortium. As such, it acts as a local data service unit (DSU) and is jointly responsible for operating and developing the platform.
- In June 2024, DaSCH signed a collaboration agreement with FORS.

ANNEX 4

Dashboards of selected initiatives
FORS – Swiss Centre of Expertise in the Social Sciences

Key role

FORS is an umbrella institution that
- acts as national node of **two ERICS** (CESSDA ERIC Service Provider & ESS ERIC National Coordinator)
- coordinates the Swiss participation in several **international social science programs** such as the European Social Survey (ESS), the International Social Survey Programme (ISSP) as well as national surveys like Swiss Household Panel (SHP), the Swiss election study (Selects), and MOSAiCH (Measurement and Observations of Social Attitudes in Switzerland) in combination with the ISSP
- hosts **SWISSUbase**
- provides **service and support** for data producers and data consumers.

Development plan

- Collecting data for national projects (SHP, Selects), MOSAiCH
- Data archiving and data management support (active curation of several hundred datasets into SWISSUbase)
- Further development of SWISSUbase and surrounding services
- Up international and national collaborations

Main ORD Services and infrastructure

FORS prioritizes effective data management for the social science community. To this end, they offer comprehensive support, including tailored data management strategies, trainings, and promote FAIR principles. FORS assists researchers in meeting ethical and legal standards, assuring data security, and enhancing research visibility.

Funding

- SNSF
- University in-kind contributions
- Funds from swu ORD programme (co-funded by Universities of Lausanne, Neuchâtel and Zurich) to launch SWISSUbase

Data types and flow

SWISSUbase

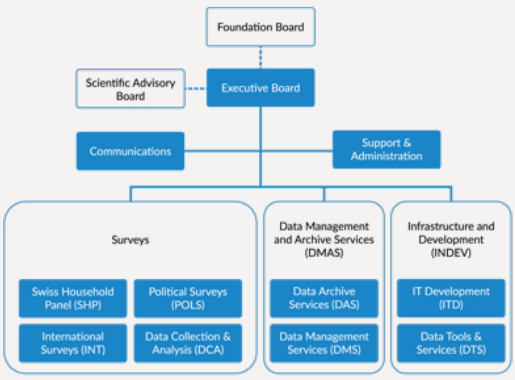
Beneficiaries and Users

- Data producers
- Data consumers
- Provides services to other actors in the national and international research ecosystem (journals, libraries, funders, infrastructures, NGOs and government agencies)

National and international cooperation

- In collaboration with Universities of Lausanne, Neuchâtel, and Zurich: responsible for **SWISSUbase**
- Cooperation Agreement between **DaSCH** and FORS was concluded in spring 2024.
- Data provider in the SWITCH-led **Connectome** project
- Collaboration and exchange with the **Federal Statistical Office**
- CESSDA ERIC
- EES ERIC
- International social sciences programmes: ISSP, EES, Comparative Studies of Electoral Programmes, et. al

Governance model



ANNEX 4

Dashboards of selected initiatives

Memobase – The portal to Switzerland's audiovisual heritage

Key role

Memobase is the national research portal of Memoriav, the competence center for Switzerland's audiovisual cultural heritage. In cooperation with the partner institutions, it provides a metasearch thought their content/materials and collections. Archives, libraries and museums from all national and linguistic regions of Switzerland provide their content or references to their collections via Memobase as online portal and thus enable the research of professionally archived photographs, films, audio and video documents for education, research and the interested public.
The dissemination and valorization of Switzerland's audiovisual cultural heritage in the fields of photography, sound, film and video is the main aim.

Development plan

- Constantly expanded with additional data and new institutions partners
- Continuous technical development e.g. further improvements in data normalization and enrichment (Linked Open Data)
- Promotion of persistent identifiers (ARK)

Main ORD Services and infrastructure

- Multilingual search with autocomplete feature
- Provenience-based access
- Faceted and advanced search
- Machine access to metadata via a RESTful API
- Ensure of access to high quality, cross-media and multilingual metadata from Swiss institutions with a long-term preservation perspective.
- Promote the recognition of the intangible (cultural, historical, political, documentary) and material value of audiovisual cultural assets as part of the historical heritage.
- Provide access to audiovisual documents (content) for education, teaching and research.
- Enable the connection to other information portals in Switzerland and abroad
- As an association, Memoriav supports a broad portfolio of data-related services that support the long-term preservation of and the access to cultural heritage.

Funding

- Memoriav Association
- Swiss Federal Office of Culture

Governance model

Memobase is provided by Memoriav – Association for the preservation of the audiovisual heritage of Switzerland
Technical realisation and operation: Universitätsbibliothek Basel, IT Abteilung

Data types and flow

Aggregated metadata of audiovisual objects held by Swiss GLAM institutions. Digital access to data content whenever possible.

- photographs
- Film, audiovideo, diaporama
- computer games
- recent cultural goods (in future)

Beneficiaries and Users

All disciplines with mainly historical focus.

- broad public
- researchers
- collections and archiving (GLAM) specialists
- interested (professional) public

National and international cooperation

- Partner for the SRG Archive Opening
- Europeana

ANNEX 4

Dashboards of selected initiatives
Scholarly Digital Editions

Key role

Scholarly editions are the basis of work for many disciplines in the humanities. Accordingly, scholarly editions are comparable to data platforms, maintained by experts. Scholarly editions provide access to professionally elaborated and scientifically attributed, annotated texts.

Development potential

- Need for coordination for technical infrastructures
- Ensuring interoperability of data (XML dialects)
- Ensuring linking to norm data (usage of identifiers)
- Community-building for training and exchange

National and international coordination

DARIAH-CH Working Group Scholarly Editions.
International initiatives mainly through DH conferences (including the annual TEI conference).
- Adho.org
- TEI (Text-Encoding-Initiative) annual conference:

ORD aspects

- Findability usually provided through community outreach. Specific registries exist
- Interoperability: On level of metadata, editions can be treated similar to publications. Standardization (and accordingly interoperability) on the level of the data, based on community standards. Mostly TEI XML or at least statements with relation to TEI XML: tei-c.org. In rare cases, graph-based approaches are taken (see Semper, Stil).
- Reusability: Due to low access barriers and publication as data dump (on Zenodo, Swissubase, etc.) the data can be easily used in digital form or for close reading.

Beneficiaries and Users

- Scholars and broader public
- Teaching on the level of high schools and universities
- Unintended users: Training of (Large) Language Models by researcher and commercial players

Funding

- Diversity of funding opportunities:
- Project funding (SNSF)
- Long-term project funding (SNSF & ASSH)
- Private funding

Generally, it's often very costly and requires specialized personnel. Often not combined with qualifications, esp. of Ph.Ds.

Coordination need

No redundancies regarding provided data identified. Partial cooperation with regard to digital infrastructures (via e-editiones/TEI-Publisher).
Editions are not produced en masse. In Switzerland per year about 5-8 edition projects are started (including new volumes of larger entities such as DODIS and Rechtsquellenstiftung). It is therefore not so difficult to reach the community (also possible through DARIAH-CH working group "Editions" and the ZDE at UZH).
Most cited **problems** by scholarly editions:
- **Long term availability** (compared to long term archiving)
- **Complex visual/front-end developments**, mostly custom-made (with low sustainability)

Selection of Initiatives

- Lavater Edition
- Online-Edition der Paippalada-Rezension des Atharvaveda
- Gottfried Semper: Der Stil
- Edition der St.Galler Missiven in Text und Bild, 1400-1650
- Der Neue Ameis-Hentze. Ein Gesamtkommentar zu Homers Ilias
- Die Kirchen Roms im Mittelalter 1050-1300. Liturgische Ausstattung und Architektur (Corpus Cosmatorum II)
- Gotthelf Edition
- Haller.net
- Schweizerische Rechtsquellenstiftung
- Anton Webern Gesamtausgabe
- Nachlass Karl Barth
- Bernoulli-Briefwechsel
- Bullinger digital
- Robert Walser-Ausgabe

ANNEX 4

Dashboards of selected initiatives Swiss Art Research Infrastructure Technology Platform (SARI TPF)

Key role

The SARI TPF of the Philosophical Faculty the University of Zurich is a service-oriented platform providing technological support for art research/history and digital humanities. It offers various services such as data hosting, integration and visualization to both academic and non-academic users/clients.

Development plan

SARI development consists of project funding and, until 2024, a basic contribution from the UZH, which allow to, both, cover running costs and periodical technology leaps. SARI's recognition as Technology Platform (TPF) of the UZH is based on an officially approved Business Plan. Please consult section 11 below (Funding) and the attached Business Plan for further information about SARI's future development, and please note that, after SARI's 2024 evaluation by the UZH, an updated Business Plan will be issued for 2025 and following.

Main ORD Services and infrastructure

SARI provides research data and visual research resources through semantic network based on acknowledged standards. SARI's services are designed to comply with ORD practices and FAIR principles the linked-open-data framework (LOD), in particular enhances interoperability and accessibility. SARI offers clients services in data preparation and the provision of customizable research environments, workflows, and visualizations.

Data types and flow

research data and visual research resources through semantic network based on acknowledged standards.

Funding

- Basic operation costs are guaranteed by UZH funding until 2024: after 2024 SARI Center expects the necessity for additional funds to cover the fixed costs
- Revenues is generated by the mandates: the development of the overall revenues has shown continuous growth and is expected to grow in 2025

Beneficiaries and Users

SARI holds both academic and non-academic mandates. Beneficiaries are

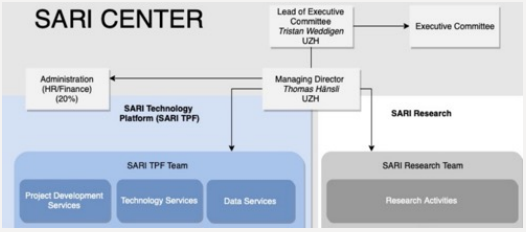
- Individual researchers
- Collaborative research projects
- GLAM institutions (both local, cantonal, and national)
- Public sector
- Public (citizen science grassroots projects)

National and international cooperation

- Targeted cooperation with LiRI and TCC as complementary services (with a focus on textual data processing, while SARI focuses on image and digital representant)
- On a national level, SARI is associated with SWITCH Connectome, DaSCH
- GLAM

Governance model

The SARI Center encompasses the SARI TPF and SARI Research (hosting research projects)



ANNEX 4

Dashboards of selected initiatives
Schweizerisches Sozialarchiv

Key role

The Social Archives' strategy focuses on the three objectives of (1) providing infrastructure services (safeguarding analogue and digital, written and audiovisual primary sources), (2) contributing to the preservation of collective memory and (3) promoting research. As part of the Swiss Inventory of Cultural Assets of National Importance, it has a unique collection of photographs, prints, posters, films, videos, sound documents and objects, which include specific audiovisual holdings and a collection of genuinely digital holdings.The infrastructure for the acquisition, storage, cataloguing and use of digital archive submissions, which have been put into regular operation in 2022.

Development plan

- New strategy of digital transformation is currently being developed and will be communicated in due course of time.

Main ORD Services and infrastructure

- Online metadata and catalogues of all collections
- Audiovisual online database (including photographs, prints and posters, movies, videos, audios, oral history interviews, 'Digitale Schriften', presentation of 'grey' digital-born journals, retro-digitised resources on e-newspaperarchives and e-periodica
- Participation in 'Webarchiv Schweiz', meta-catalogues archivesonline.ch, swisscovery, and memobase.

Funding

- Staatssekretariat für Bildung, Forschung und Innovation
- Bildungsdirektion des Kantons Zürich
- Schul- und Sportdepartement der Stadt Zürich
- Project-related external funding from various sources

Beneficiaries and Users

Usage is on a par with much larger federal memory institutions. Beneficiaries are

- National and international researchers
- Historians
- Social scientists
- Journalists
- Students of Universities, Universities of Applied Sciences and Grammar Schools
- Interested public

Governance model

The institution is run by a non-partisan association. It consists of the following bodies

- Verein Schweizerisches Sozialarchiv, based in Zurich, founded in 1906

Data types and flow

- photographs, prints, posters, films, videos, sound documents
- audiovisual legacy of the organisations documented in the Archive department
- Increasing collection of genuinely digital holdings such as 'Digitale Schriften'

Access is principally free (within the limits of personal privacy, data protection and copyright legislation)

National and international cooperation

Participation in national initiatives and infrastructures:

- swisscovery, archivesonline.ch, e-newspaperarchives, e-periodica, Memobase, Webarchiv Schweiz,
- swissvotes.ch, arbeiterbewegung.ch, unseregeschichte.ch, sportshistory.ch, corona-memory.ch etc.

Participation in international initiatives and infrastructures:

- Archival IIIF, Europeana, socialhistoryportal.org
- Archives Portal Europe, Online-Portal of the Archives of Rural History (ARH) and the European Rural History Film Association (ERHFA) etc.

ANNEX 4

Dashboards of selected initiatives
Swiss-AL

Key role

Swiss-AL is a multilingual language data resource of Swiss public communication that allows for data-driven analyses and simulations of social and public discourse. Swiss-AL is available to researchers and the public as an open research data resource and operable through an integrated workbench. The corpora are based on public language data, more specifically on the data stock of SMD Schweizer Mediendatenbank and webcrawled texts (e.g., from blogs of governmental institutions, political parties, etc.) Due to copyright reasons the corpora cannot be accessed by users but analyzed through queries on co-occurrences and trending topics (topic modeling).

Development plan

- Level up standards of sustainable documentation, legal issues/ privacy, data circulation, and research data management in accordance with FAIR principles.
- Interoperability framework with LiRI under the umbrella of CLARIN
- AI interface

Main ORD Services and infrastructure

- A family of readily available multilingual, linguistically annotated **corpora**, comprising journalistic media (via Swissdox) and webcrawled content
- A **nlp-pipeline** to process language data into linguistically annotated corpora. The pipeline was developed in-house, is based on open-source components and contains several modules, e.g. for duplicate detection, webscraping and linguistic annotation. Swiss-AL uses common data and metadata standards
- A **browser-based workbench** to access and analyze the corpora (currently in Beta mode)

Funding

- Basic funding for minimal maintenance by the host institution (ZHAW)
- Terminated PgB funds from swissuniversities projects
- Through projects with praxis partners (delineation of topic-specific corpora)

Governance model

The operating body of Swiss-AL is the ZHAW Digital Discourse Lab

Data types and flow

- Corpora consist of Journalistic media data and webcrawled data from various public players and institutions
- Due to copyright reasons, access on full text of the corpora is not possible, however, metadata of corpora are publicly accessible and harvestable on the Virtual Language Observatory
- Through a browser-based workbench/ user interface, users can identify discourse linguistic patterns within the selected language corpora (e.g., topic modelling, cooccurrences/collocations)

Beneficiaries and Users

- Specialized user group: applied linguistics (corpus-, discourse-, and media linguistics)
- Inter- and transdisciplinary research, and research in any field interested in (national) public discourses such as public health, social welfare, sustainability, climate/environment, geopolitics
- Non-academic parties (journalism, educational institutions, administration, citizen science)

National and international cooperation

- CLARIN-CH: community embedment
- LiRI: access to Swissdox media stock (retrieved from SMD)
- LaRS/SWISSUBase: Metadata of corpora are published on SWISSUBase
- In a current PgB project (FAIR-FI-LD), Swiss-AL and LiRI lever the interoperability and FAIR-compliance of different national language data initiatives
- metadata can be harvested by CLARIN Virtual Language Observatory

ANNEX 4

Dashboards of selected initiatives
SWISSUbase

Key role

SWISSUbase is an online platform that supports the Swiss scientific community by facilitating the **sharing** and **long-term preservation** of research data. It ensures data quality and compliance with FAIR principles and long-term preservation (ISO 14721) standards. Local Data Service Units (DSU) validate and archive data, while the platform offers a secure repository for open and restricted data sharing. It also provides a portal for searching, discovering, and promoting research projects and associated researchers including their affiliations.

Development plan

strategic focus:
- To become a nationally recognized research infrastructure and enhance European and international integration and cooperation
- Expand to disciplines outside SSH (current pilot: geosciences)
- Increase volume of deposited data and download score
- Participation in joint efforts for creating infrastructure(s) to securely manage sensitive data

Main ORD Services and infrastructure

- Repository
- Long-term preservation
- Data validation/quality check
- Tailored metadata profiles

Funding

- SNSF (through) FORS targeting the provision of various services to the social science community, incl. mandate to provide research data infrastructure
- Yearly fee of SWISSUbase Consortium members: funds salaries of FORS staff
- The following tasks are performed by local DSU: Validation of submitted projects (quality assurance on data and meta data), archiving, publishing deposited data, instructions/training on how to use the platform

Data types and flow

- various data types
- metadata (in some cases exclusively, e.g. Swiss-AL)
- Data are submitted by research communities and subsequently validated, published and archived by local DSUs

Governance model

The governance of the SWISSUbase consortium is assured by the three governance bodies which are:
- the **oversight board** (composed of high-ranking officials such as rectors and vice-rectors),
- the **steering committee** (composed of high-level infrastructure and data management specialists),
- the **SWISSUbase management team** (operated by **FORS** staff) manages the communications and coordination between all partners, the outreach and promotion of the platform; the planning of future development of SWISSUbase and the overall strategy and policy development

Beneficiaries and Users

Swiss research communities with a focus on SSH (targets a national scope for social sciences, linguistics, and humanities).
UNIL and UNINE use SWISSUbase as institutional repository, i.e., also deposite data from domains outside SSH
Meta data profiles of other disciplines can be developed

International cooperation

SWISSUBase serves as the national node for
- CESSDA
- CLARIN
- DARIAH (from 2025)

Data deposited on SWISSUBase is harvested by the ERICs' respective common catalogues
- CESSDA Data Catalogue
- Virtual Langauge Observatory (end 2024)
- DARIAH (from 2025)

For every new research discipline, the same approach is intended

ANNEX 4

Dashboards of selected initiatives

swisscollections – gateway to historical and modern collections in Swiss libraries and archives

Key role

swisscollections contains archival materials (including unpublished papers), images, old prints and rare books, document collections, film material, manuscripts, maps, music recordings, sheet music, text recordings as well as listings in bibliographies of cantons. It is a metacatalogue that draws on various data pools.The swisscollections Association was formed in December 2020 as successor of the HAN (Handschriften – Archivbestände – Nachlässe) association founded in March 2011.

Development plan

- Soon to come: Export of a result list on swisscollections.ch as a list or as a data package, including bibliographic metadata and URLs for images and fulltexts of digitised resources published on e-rara or e-manuscripta in various formats. This will enable researchers to create an individual data package on swisscollections, download the metadata and use the URLs to download digitised content from e-rara and e-manuscripta.

Main ORD Services and infrastructure

- Online metadata and catalogues of participating collections
- Merging and mapping of content
- Maintenance and operation of the data processing pipeline, full-text index and search interfaces
- Metadata services: defined interfaces, special filters, special searches

Preservation:

- The data security and preservation of the digitized data is the responsibility of the respective sponsoring institutions
- metadata is temporarily stored in the swisscollections full-text index in addition to the data source according to the current version

Funding

- Verein swisscollections and partner institutions

Beneficiaries and Users

All disciplines with mainly historical focus.

- historians, researchers
- collection specialists and
- maintainer
- interested (professional) public
- students

Governance model

Driven by the Verein swisscollections (Host: Universitätsbibliothek Basel) with the following members: Cantonal Library of Aargau, the Cantonal Library of Appenzell Ausserrhoden, the Cantonal Library Vadiana of St. Gallen, the Stiftsbibliothek St. Gallen, the University Library of Basel, the University Library of Bern, the Central and University Library of Lucerne, the Central Library of Solothurn, and the Central Library of Zurich.

Executive committee of the association: Stefan Wiederkehr, Central Library of Zurich, President; Heidi Eisenhut, Cantonal Library of Appenzell Ausserrhoden, Vicepresident, Daniel Tschirren, Central and University Library of Lucerne, committee member

National and internat. cooperations

The purpose of the association is to maintain and further develop a meta-catalog for special collections.

- It has evolved from the association HAN in order to meet the new requirements as of 1.1.2021 under the Swiss Library Service Platform.
- Import of metadata from SLSP and ZBCollections
- Soon: Import of structural metadata from e-rara and e-manuscripta plus Data export to Archivportal Europa, Archives Online, Kalliope, KVK

Connex to Europeana

Data types and flow

- Bibliographic records and digitized versions of them
- semantic/descriptive metadata of sources
- archival materials (including unpublished papers)
- manuscriptsold prints (until 1900) and rare books
- document collections, images, maps, film
- material, sheet music, music recordings,
- text recordings, bibliographic listings (in particular in bibliographies of cantons)

ANNEX 4

Dashboards of selected initiatives

ZentralGut – Plattform für das digitalisierte Kulturgut der Zentralschweiz

Key role

ZentralGut is a diverse regional cultural heritage platform that – like our cultural heritage itself – is constantly being expanded and updated. The texts, images, audio and video documents can be researched and existing objects can be commented on, downloaded and reused. The portal for the digitized cultural heritage of Central Switzerland is open to interested memory and cultural institutions from the region to publish images of objects, photographs and texts from their collections, as well as to private individuals who are invited to share their knowledge of the region and contribute their own treasures.

Development plan

- ZentralGut collects as an ongoing project materials from institutions or even private persons of Central Switzerland which could be part of the collection.

Main ORD Services and infrastructure

- ZentralGut can be retrieved by OAI-PMH and SRU standardized APIS for retrieval.
- To ingest data ZentralGut offers different use-case dependend APIS or opportunities to store, describe and publish data.
- Everyone is encouraged to give feedback both on the platform in general or on different documents.
- ZentralGut allows a direct interaction for every registered user to upload its own material, which will go through an internal review before publishing.
- Certain aspects of meta data generation will be done with a participatory approach like the geo-referencing of the graphic materials of ZHB Luzern Sondersammlung.

Funding

- ZentralGut is a service of ZHB Luzern and retrieves no external funding

Beneficiaries and Users

The plattform is dedicated to all users both researchers and citizens.

Governance model

- ZentralGut ist hosted / published by ZHB Luzern
- Technical realisation and operation: ZHB Luzern on behalf of Intranda Goobi

Data types and flow

- digitzed material from GLAM institutions of all six cantons of Central Switzerland (Lucerne, Nid- and Obwalden, Schwyz, Uri and Zug)
- Digitized and “born digital” documents
- photographs and other 2-d visual artworks such as paintings, graphics or postcards
- videos
- magazines and journals
- daily and weekly newspapers
- books
- manuscripts
- archival materials
- photographs of museal objects
- maps

National and international cooperation

- ZentralGut currently publishes its material to Wikimedia Commons:
- ZentralGut plans to publish data to Europeana, register metadata as Open Data on opendata.swiss portal and articles of journals published in ZentralGut on Zenodo.
- Currently an idea-exchange with the project “Connectome” of Switch is ongoing.
- ZentralGut is part of the GLAM-Wiki group Switzerland and the “Netzwerk offene Kulturdaten” supported by Wikimedia Germany

ANNEX 5

Financial summary

The financial summary is currently under development and will be continued by the StraCo office. The financial summary and its data base can be accessed via SharePoint:

https://snsf.sharepoint.com/:x:/r/sites/ORDStraCo/Shared%20Documents/General/7.%20Cluster%20Analyses/2.%20SSH%20-%20Social%20Sciences%20and%20Humanities/Landscape%20analysis%20-%20Task%20Force%20SSH/Reporting/SSH-Initiatives-FinancialOverview.


open
research
data
National
Strategy
Council