ESFRI LANDSCAPE ANALYSIS 2024 - SECTION 1

SOCIAL SCIENCES AND HUMANITIES

CONTENTS

SOCIAL SCIENCES AND HUMANITIES

CURRENT STATUS IN THE DOMAIN 00

The Social Sciences & Humanities (SSH) landscape covers a wide range of academic disciplines, which help us to understand ourselves, others, and the human societies around us. These disciplines can be grouped in the following main areas[1]: Individuals, Markets and Organisations, Institutions, Governance and Legal Systems, the Social World and its Interactions, the Human Mind and its Complexity, Texts and Concepts, the Study of the Human Past, Human Mobility, Environment, and Space, Studies of Cultures and Arts.

1. ERC panel structure (2024) https://erc.europa.eu/sites/default/files/2023-03/ERC_panel_structure_2024_calls.pdf

Moreover, SSH contributes to multidisciplinary research, inter alia in the field of Health, Security, Urban Studies, Science and Technology Studies.

The Domain of Social and Cultural Innovation (SCI) in the ESFRI Landscape includes all those infrastructures projects and European Research Infrastructures legal entities (ERICs) that support our understanding of people and society and the research in the disciplines under the acronym of SSH. Given this acronym is used outside of ESFRI, it should again be used instead of SCI.

SSH disciplines address the digitalisation paradigm by exploiting and adopting new methods, data types, and tools. Human behaviour and its consequences are complex, and affect how culture is generated and shared, how society is organised, and how technology and policies should be developed and implemented. Understanding the mechanisms behind these processes requires Research Infrastructures (RIs) that enable the cross-disciplinary study of development over time, locations, and interlinkages. It requires facilities capable of dealing with the growing amounts of complex, multi-layered data, which needs proper contextualisation and evaluation. The recent diffusion of systems, which generate information on the basis of Large Language Models (LLM), has added the need for evaluating the generated 'knowledge' and account for the impact of these systems on societies.

SSH is at the forefront of establishing RIs that efficiently implement sharing of FAIR data resources[2] as a core driver for research development. At the same time this data driven SSH research creates a vital platform for development of fair social policies in Europe to advance citizens' well-being, welfare, and cultural dialogue. Moreover, SSH researchers address the various societal challenges in the forms of transformations, crises, and divides. In this line, the SSH provide underpinning research for the European pillar of social rights and European policies enabling an understanding of local, national, and supra national realities. Even though most of the SSH RIs primarily address researchers supporting the highest standards of research quality, they also serve policy makers and the public at large.

RIs are long-term enterprises and represent strategic investments, which are indispensable for enabling and developing research. In addition to fostering scientific knowledge for the benefit of SSH researchers, they also impact the research environment socially and economically. The Council's conclusions on RIs in December 2022 recognised RIs' central role in the development of the European Research Area (ERA)[3].

Data analysis on social transformation provides crucial insight into people's capacity to adapt to fast environmental and technological changes. The Council's 2022 conclusions affirm that RIs "provide knowledge-based solutions to societal challenges and help to deliver the EU's green and digital transitions". Finally, SSH RIs support the preservation and reuse of cultural data which are essential to understand the development and variety of cultures, and consequently to the resilience of Europe as a project of Member States (MS), alongside the creation of economic value in multiple industries, spanning from the creative arts to tourism.

Providing an in-depth understanding of the social and human dimension, the contribution of SSH RIs to different challenges, including the Sustainable Development Goals (SDGs) and policy making in general, is more visible than ever before. At the core

2. FAIR data principles are a central concept of Open Science. Their implementation means making the data Findable, Accessible, Interoperable and Reusable. See, e.g., GO-FAIR initiative https://www.go-fair.org/

3. Council of the European Union (2022), Council Conclusions on Research Infrastructures https://data.consilium.europa.eu/doc/document/ST-15429-2022-INIT/en/pdf

of these research domains lies the understanding of societal upheavals and changes and cultural transitions; it involves engaging citizens to modify behaviour to help alleviate the consequences of climate, societal and technological changes, exploring political debates in different languages, and understanding the impact of events such as wars, migration, and pandemics on people's well-being and health. Complementary to the approaches of the natural and physical sciences, SSH research provides methods, pre-requisites, sources, tools, and evidence to improve our societies' well-being and resilience. This research does not solely focus on behavioural change or social acceptance of the future; it also yields hard data and solutions that support policies and programmes, catering to the needs and desires of citizens.

SSH RIs support research that significantly contributes to European scientific excellence. A recent EC publication claims that Humanities is a focus of expertise in Europe[4] and SSH research plays a crucial role in scientific competitiveness. Focusing on Europe, the US National Science Foundation's report on publications (2022) claims that the European authored Social Sciences publications are 17% of world share versus the 8.4% in the US[5]. Furthermore, the most recent report by the International Association of Scientific, Technical, and Medical Publishers (STM) shows that the global share of SSH publishing in 2020 was $4.5 billion, and Europe benefits by 32% of total value (US 40%, Asia 20%)[6].

SSH RIs identify important themes for European research, contributing towards scientific excellence in these domains. They support the creation, collection, and curation of essential data to support disciplinary areas in exploring new approaches to fundamental questions. SSH RIs also play a key role in the innovation of methods and research processes both at disciplinary and interdisciplinary levels. RIs in this domain also draw in non-academic actors who accelerate the realisation of the benefits of research for society. These RIs are ideally positioned to support current research and nurture the next generation of scientists across Europe using resources from SSH data archives to cross-national surveys along with new methods and processes. In some areas, SSH RIs are at the forefront of current developments, as relevant practices have been established for a long time; for instance, sharing data for the purpose of historical

4. EC (2022), "Science, Research and Innovation Performance of the EU 2022. Building A Sustainable Future in Uncertain Times", pp. 7 & 54.

5. NSF Science and Engineering indicators 2022 https://ncses.nsf.gov/pubs/nsb20214/publication-output-by-country-region-or-economy-and-scientific-field#; Gingras Y. and Mosbah-Natanso S. (2010), "Where are Social Sciences produced?", World Social Science Report, Chapter 4, p. 150.

6. STM Global Brief 2021 – Economics and Market Size https://www.stmassoc.org/2022_08_24_STM_White_Report_a4_v15.pdf, pp. 10-12

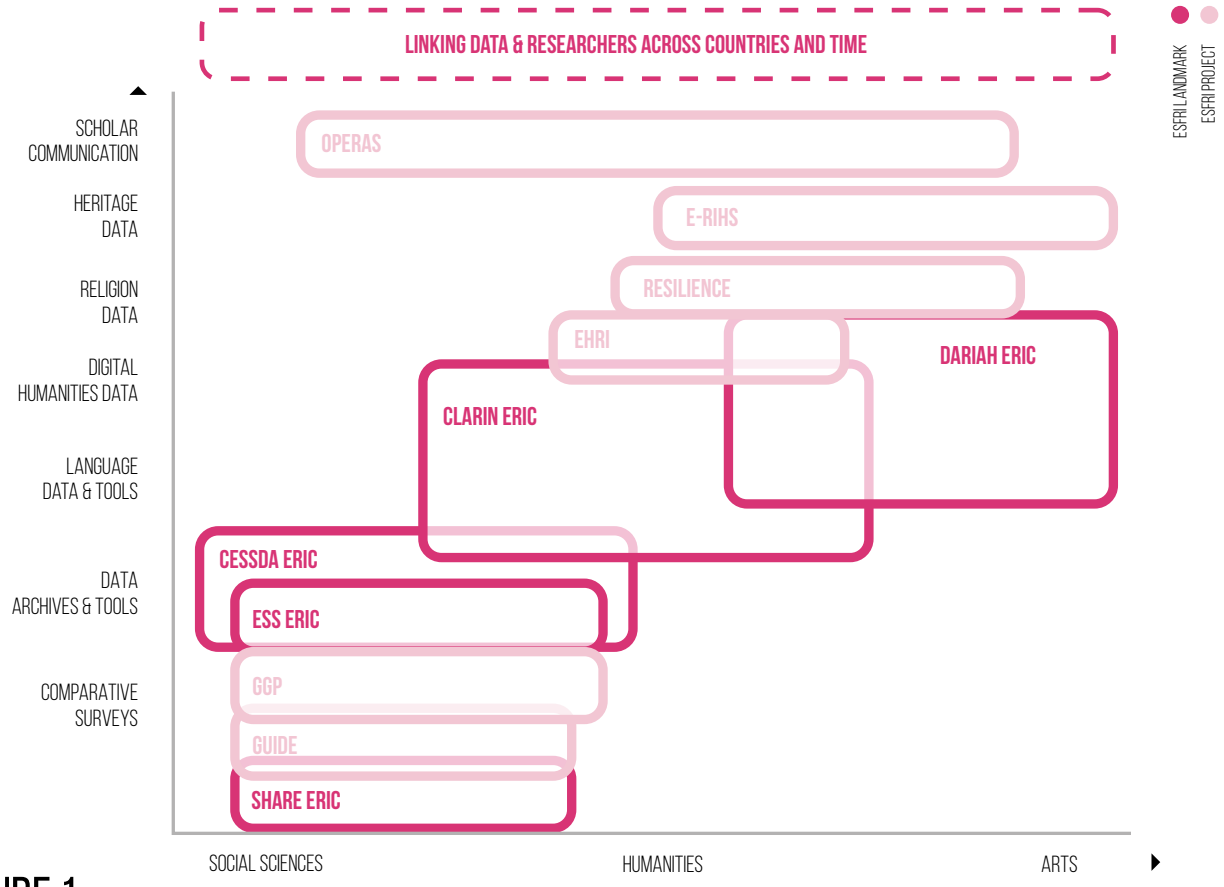

FIGURE 1.

The indicative position of ESFRI RIs in relation to the types of data and tools employed in the Social Sciences & Humanities domain

comparison between cultures and countries and study of trends in social developments.

Furthermore, working with experts from other domains encourages collaboration and exchanging best practices. RIs may also assist the coordination within or between disciplines and support research with provision of tools. More importantly, they consider the FAIR principles and encourage the use of standards in data collections. Namely, RIs can help ensure that researchers have the options for re-use of existing data before creating new datasets; where possible, new data are designed to be interoperable with existing high-quality, comparable data, creating new knowledge. RIs deliver more than access to data, they facilitate shared data collection on a grand scale and provide access to expertise and guidance to researchers early in the design phase of new research.

The first five large SSH ERICs (**CESSDA ERIC**, **CLARIN ERIC**, **DARIAH ERIC**, **ESS ERIC** and **SHARE ERIC**) were included in the ESFRI roadmap in 2006 and awarded Landmark status in 2016; some of them built upon collaborations started decades earlier. Two more (**E-RIHS**, **EHRI**) were added in 2018, followed by another four that joined in 2021 (**GGP**, **GUIDE**, **OPERAS** and **RESILIENCE**).

The current SSH RIs are displayed in ***Figure 1*** and ***Figure 2*** and are summarised below, following a thematic division and with the Landmarks (ERICs) first, followed by the Projects (in alphabetical order).

SOCIAL SCIENCES – SURVEYS, DATA TOOLS AND SERVICES

The **ESFRI Landmark ESS ERIC** (European Social Survey)[7] is an academically driven long-term cross-national survey in Social Sciences that has been conducted across Europe since 2001. It assesses the attitudes, beliefs, and behaviour patterns of diverse populations in more than 30 nations, measuring change over time within and between European countries in their living conditions, social structure, public opinion, and attitudes.

The **ESFRI Landmark SHARE ERIC** (Survey of Health, Ageing and Retirement in Europe)[8] is a Research Infrastructure for studying health, social, economic, and environmental questions over the life-course of European citizens and beyond. SHARE collects multidisciplinary and cross-national survey panel data on health, socio-economic status and social and family networks of individuals aged 50 or older (biennial survey waves).

The **ESFRI Project GUIDE** (Growing Up in Digital Europe)[9] is a pan-European comparative birth cohort survey including a sample of newborn infants and a sample of school age children. Both cohorts are surveyed using a common questionnaire and data collection methodology at regular intervals until the age of 24. The Research Infrastructure is a source of high quality longitudinal statistical evidence to support the development of social policies to enhance the well-being of children, young people, and their families across Europe.

The **ESFRI Project GGP** (Generations and Gender Programme)[10] provides scientists and policy makers with high-quality and timely data about families and life course trajectories of individuals to enable researchers to contribute insights and answers to current societal and public policy challenges. GGP survey focuses on inter-generational and gender relations between people, expressed in care arrangements and the organisation of paid and unpaid work.

The **ESFRI Landmark CESSDA ERIC** (Consortium of European Social Science Data Archives)[11] provides large-scale, integrated, and sustainable data services to the Social Sciences. It brings together Social Sciences data archives across Europe, aiming at promoting the results of Social Science research and supporting national and international research and cooperation. CESSDA provides a central data catalogue and a platform to jointly develop user-friendly tools and services.

HUMANITIES – SURVEYS, DATA TOOLS AND SERVICES

The **ESFRI Landmark CLARIN ERIC** (Common Language Resources and Technology Infrastructure)[12] provides data, tools, and services to support research based on language resources, and it is available to all disciplines and not exclusively to SSH. CLARIN's Virtual Language Observatory (VLO) provides easy and sustainable access to digital language data and advanced tools to discover, explore, exploit, annotate, analyse, or combine them, wherever they are located.

The **ESFRI Landmark DARIAH ERIC** (Digital Research Infrastructure for the Arts and Humanities)[13] provides support to digitally enabled research and teaching across the Arts and Humanities. DARIAH is a network of people, expertise, information, content, methods, tools, and technologies. It develops, maintains, and operates an infrastructure in support of digital research practices and sustains researchers in using them to build, analyse and interpret digital resources.

CLARIN ERIC and DARIAH ERIC jointly manage the Digital Humanities Course Registry[14] and in some countries, they have joined their efforts.

The **ESFRI Project EHRI** (European Holocaust Research Infrastructure)[15] enables online and physical access to Holocaust sources and expertise dispersed across many institutions in Europe and beyond. EHRI provides innovative tools and training

7. ESS ERIC https://www.europeansocialsurvey.org/
8. SHARE ERIC https://share-eric.eu/
9. GUIDE https://www.guidecohort.eu/
10. GGP https://www.ggp-i.org/
11. CESSDA https://www.cessda.eu/
12. CLARIN https://www.clarin.eu/
13. DARIAH ERIC https://www.dariah.eu/
14. Digital Humanities Course Registry https://dhcr.clarin-dariah.eu/
15. EHRI https://www.ehri-project.eu/

that advance the digital transformation of Holocaust research, remembrance, and education. The vision of EHRI is to secure seamless access to all sources and expertise that are relevant to the study of the Holocaust.

The **ESFRI Project E-RIHS** (European Research Infrastructure for Heritage Science)[16] is dedicated to heritage science, an interdisciplinary domain merging STEM and SSH to enhance knowledge, conservation, and appreciation of cultural heritage. E-RIHS aims to preserve heritage's accessibility and significance in a changing world by unravelling its cultural and historical layers, understanding material change, and utilising it for socioeconomic and environmental sustainability.

The **ESFRI Project RESILIENCE** (REligious Studies Infrastructure: tooLs, Innovation, Experts, conNections and Centres in Europe)[17] focuses on Religious Studies, building a high-performance platform, supplying tools and access to physical and digital data to scholars from all scientific disciplines. RESILIENCE primarily serves the academic community, but its impact extends to the non-academic community. It gives physical and digital access to major relevant data archives for Religious Studies.

Open scholarly communication of Social Sciences & Humanities in the ERA is supported by the **ESFRI Project OPERAS** (OPen scholarly communication in the European Research Area for Social Sciences and Humanities)[18]. Its mission is to coordinate and federate resources in Europe to efficiently address the scholarly communication needs of European researchers in the field of SSH. OPERAS fills a gap in the European landscape, between generic e-infrastructures and RIs dedicated to research data in specific disciplines or topics.

16. E-RIHS
https://www.e-rihs.eu/

17. RESILIENCE
https://www.resilience-ri.eu/

18. OPERAS
https://operas-eu.org/

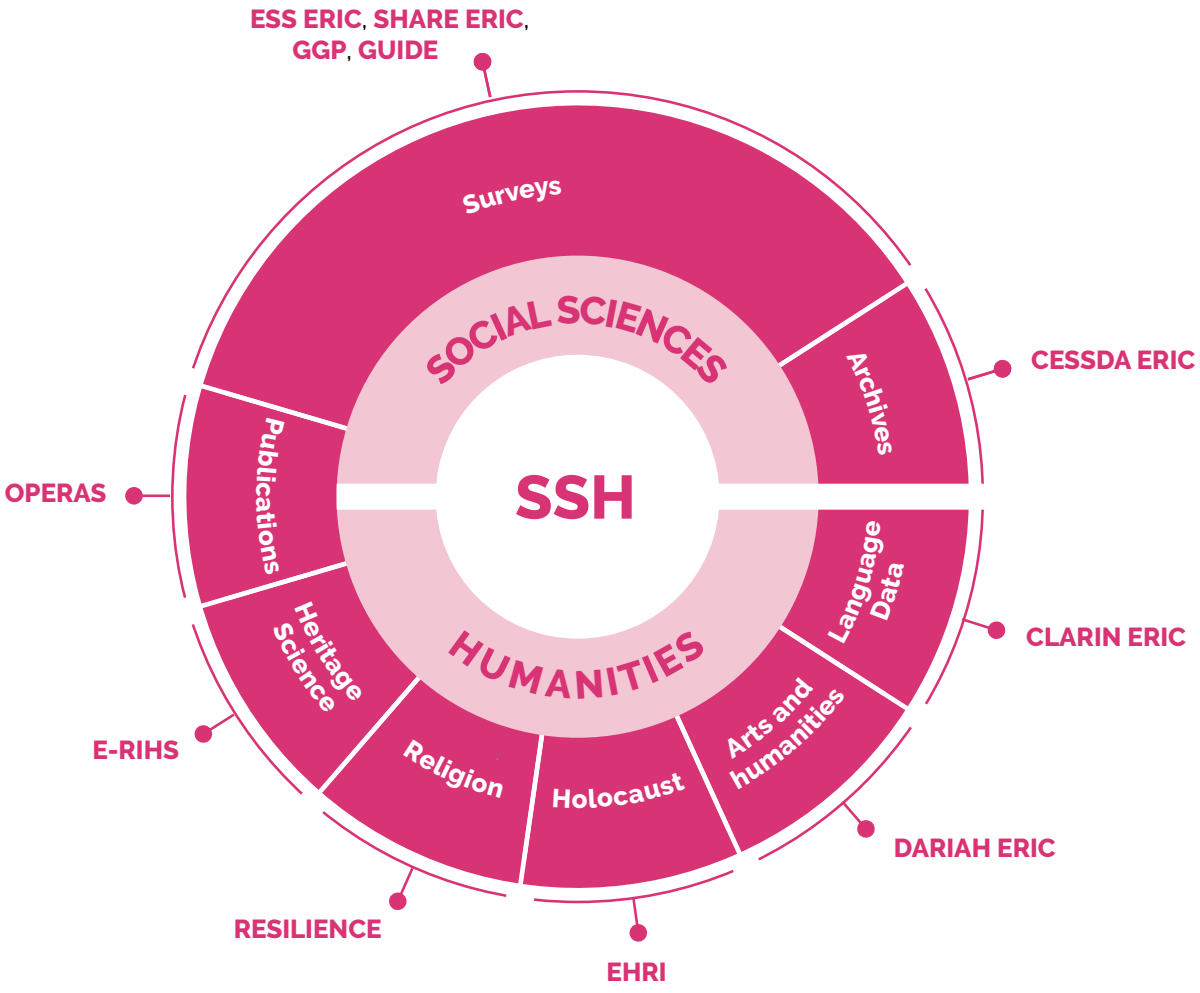

FIGURE 2.
The Landscape of the
Social Sciences & Humanities domain

SERVICES IN THE DOMAIN

The main user group for the SSH RIs are researchers, both individuals as well as institutions at the world, European, and national level. Furthermore, a growing number of organisations are using data provided in the SSH domain. **DARIAH ERIC**, **EHRI** and **CLARIN ERIC**, for instance, collaborate with non-academic organisations, in particular in the cultural heritage sector of galleries, libraries, archives, and museums (GLAM). Similarly, non-governmental organisations focusing on migration and population studies utilise **SHARE ERIC** data; Holocaust memorialisation, anti-discrimination, and Human Rights organisations are supported by EHRI governmental organisations use data from **ESS ERIC** and SHARE to inform public policies[19], while CLARIN and DARIAH help address societal awareness of the ethical issues related to the use of Artificial Intelligence and large language models. More importantly, the number of civil servants and industry users is growing. National governments' research units and social enterprises, publishers and the health industry seek access to datasets for their business planning and implementation. The conservation industry, for example, is targeted in the innovative approaches developed via **E-RIHS** for the preservation of cultural heritage. User statistics have been tracked, sometimes using ESFRI KPI for scientific excellence.

Many SSH RIs are distributed and provide access to large volumes of complimentary digital resources and offer federated services. Many of the large datasets being distributed by SSH RIs are open and free at the point of use, being based on Open Science and FAIR principles. Most of the shared datasets have not been created and annotated automatically, but they have been collected and/or curated by surveys or by human work and expertise in many countries worldwide. RIs' coordinating offices ensure support for the development of standards as well as of common metadata, implement shared platforms to make resources and federated services available, and blend user training and support. Heritage sciences also require instruments and techniques to study tangible heritage objects. Some RIs, like EHRI, **RESILIENCE**, DARIAH and E-RIHS, deploy instruments for Transnational Access (TNA) that facilitate physical access to their resources. For instance, results of surveys provide a wide range of information on societies and European citizens offering key insights on the impact of different social policies, while standardised datasets allow for the analysis of phenomena across languages, time periods and cultures, testing of methods and tools, reusability, and extension.

Data archives respond to the current demand of large data and research resource availability. Scattered resources and scarce accessibility of original collections, especially those that require preservation, are just two of the obstacles that researchers can overcome with RIs. Studying languages, religions or heritage, financial data or historical phenomena, and public opinion via social media still heavily rely on the capacity of individual researchers to identify and access specific original resources. Many types of socially related data have different owners and their use implies a series of ethical issues. Here the CARE principles[20] can complement the FAIR principles. RIs' activities therefore also encourage the development of tools, methodologies and standards for their harmonious collection and set up pathways for collaboration among different entities responsible for such data.

Some RIs offer observatories, blogs, spin-off projects and fellowships and in-person services for face-to-face training across different facilities, both for researchers and practitioners. EHRI, for instance, offers in-person access to distributed facilities, CLARIN and E-RIHS provide state-of-the-art tools and services to cross-disciplinary users and communities. Some RIs, like ESS, provide data analysis tools with visualisation.

All infrastructures in the Social Sciences & Humanities domain offer critical services with different characteristics. Among these services are data consultation, and opportunity to contribute to the datasets.

Some SSH data-driven projects need independently collected datasets. In these cases, RIs provide standards for data collection and annotation. Users are also encouraged to deposit their original datasets and to compile new datasets according to existing criteria and standards for metadata or according to new research questions. Users then become creators and contributors to the RI and can benefit from RI resources. Notable examples are CLARIN resources, enabling researchers to reinforce their area of expertise and develop novel methodological frameworks for the exploration of multilingual data; and E-RIHS, which encourages the user-provider co-creation process.

Working with common standards also informs and supports researchers from other domains, like computer scientists for example, and reinforces their contribution in all aspects of data management and curation.

Yet, the number and involvement of researchers at national level varies greatly for each infrastructure, which hampers the possibilities for scholars to understand mutual methods and resources. Infrastructures will need to continue to develop their 'service provision' capability.

Services provided by all SSH RIs have proven particularly relevant in times of crises or societal upheavals. For example, the SHARE Corona Survey allowed the in-depth examination of how individuals were coping with the health-related and socio-economic impact of COVID-19. DARIAH launched a theme call dedicated to artistic responses to the pandemic[21]. At the same time, CLARIN was collecting parliamentary debates across Member States supporting the

19. Reports on happiness, well-being and economic challenges by France Stratégie, part of the French Prime Minister's Department, have used ESS and SHARE data.

20. CARE Principles
https://www.gida-global.org/care

21. DARIAH Theme Call: Arts Research during Covid-19
https://www.dariah.eu/activities/impact-case-studies/arts-research-during-covid-19/

comparison of how the pandemic has been addressed. Micro-data analysis of society makes researchers and policy makers far more aware of unintended consequences of certain policies and functions as an indicator for social and psychological behaviour, which can be utilised to adapt future political interventions.

More recently, some ESFRI SSH RIs started or renewed activities with Ukraine and/or are exploring ways to cooperate with the country. The ESS ERIC is preparing a data collection in Ukraine as part of the coming 11th wave of data collection in Europe; Ukraine has also participated in earlier rounds. SHARE ERIC and **GGP** work on collaborations with researchers and organisations in Ukraine to organise future data collections. **CESSDA ERIC** supports Ukrainian data experts. They work on saving data assets in Ukraine as well as on developing tools and services to provide better access to data, especially from researchers outside of the country. CLARIN ERIC collaborates with researchers from Ukraine in different ways, for instance establishing an expertise centre (K-Centre) for Ukrainian and helping to develop language and text corpora or translation systems. Some of these services are already exploited for investigations in war crimes. EHRI has an active presence in Ukraine since 2010 and is currently facilitating a permanent association of the country to the RI. In the case of Ukraine, as well as of other countries of the Eastern Partnership of the European Union, it is not only important to reconstruct RIs, but also to allow participation and access of the produced data for the present and future of a country's identity and sovereignty.

The RIs also offer the possibility to raise the visibility of data and services they provide. For example, the Social Sciences & Humanities Open Cluster (SSHOC) is the thematic cluster[22] for the SSH infrastructures and has developed the SSH Open Marketplace[23], a central hub to find detailed information about software tools, publications, datasets, training materials, and workflows supporting research across the many facets within the SSH domain.

The emergence of citizen science is encouraging a fast-growing audience of citizens that for varied reasons demand access to SSH data facilities. Therefore, some infrastructures made various resources available, including books, policy reports, summary materials, data dictionaries, data collection, data analysis and visualisation tools, often supported by the Zenodo community[24]. Services like VERA offered by **OPERAS** allow researchers to collaborate on citizen science projects and provide support throughout the preparation of the project[25].

While the RIs present a significant opportunity for SSH research across Europe, there is potential for further development, both at European headquarter level and at national level. One of the issues is interoperability, but there are also issues of visibility, training, and access. More effort could be dedicated to train students, scholars at all levels, data stewards, and research managers so that they could harvest the full benefits of data offered by the RIs. Even though some RIs already adopt these instruments, more efforts can be made to disseminate best practice for researchers and support the endeavours that RIs are doing in terms of training and education. Such efforts could build further on existing training materials such as those developed in the context of the SSH Open Marketplace[26] and DARIAH Campus[27].

22. Overview of the thematic Science Clusters
https://science-clusters.eu/

23. SSH Open Marketplace
https://marketplace.sshopencloud.eu

24. Zenodo
https://about.zenodo.org/

25. VERA Service by OPERAS
https://operas-eu.org/?s=VERA

26. Training Material – SSH Open Marketplace
https://marketplace.sshopencloud.eu/search?categories=training-material

27. DARIAH Campus
https://campus.dariah.eu

IMPACT IN THE DOMAIN

This section outlines the progress in measuring the scientific, economic and societal impact of SSH RIs, as well as the challenges in determining the diverse types of impact in the Social Sciences & Humanities domains. Several best practices and SSH impact case studies are provided.

> *RIs represent an increasingly large share of research investment, and both national and European funders are expected to develop systematic and transparent procedures underlying their investment choices. Consequently, there's a growing consensus regarding the importance of assessing the value of RIs beyond research itself and for society at large. This should not be difficult for SSH RIs given that they are the most proper candidates among RIs to have an effect on the wellbeing and future of society.*

Indeed, there are outstanding examples of SSH RIs undertaking ground-breaking work in analysing their scientific, economic and societal impact. **ESS ERIC** stands out among RIs in all domains regarding study, analysis, and reflection of the various categories of impact of RIs and use of assessment methodologies. Soon after becoming an ERIC in 2013, ESS has started with impact studies, tracking its academic and non-academic impact. A recent impact case study noted a set of common general types of non-academic ESS impacts, including general intelligence and insights for NGOs or government ministries, agencies, or advisory bodies. Such impact drives agenda setting, where ESS data highlights a particular problem or challenge, triggering various types of policy action. ESS data or ESS-based findings can also influence public debate when presented in the news media. Using ESS data as indicators enables the tracking of certain aspects of societal progress, for instance aiding in the assessment of policy effectiveness according to desired outcomes. This includes instances where ESS methodology or questions have been integrated into other surveys run for such monitoring processes. In addition, in 2022, ESS has contracted a study to map ESS topics to European Commission policy priorities. This will help to identify future opportunities for ESS data to be more visible at the European Commission level.

At the same time, several studies, policy papers and statements note that assessing impact of RIs beyond the scientific impact is challenging and that a unified framework for impact assessment of investment in RIs is questionable. This was underlined again recently, by OECD in 2019 and by ESFRI in 2023.

In a recent policy paper, the OECD noted that assessing the impact of RIs remains a challenging endeavour.[28] The main reasons lie in the fact that, although RIs deliver quality services that facilitate excellent science, outcomes are indirect and not produced by the RIs themselves but rather by its users. In many cases, RIs do not know the users, given also that the data are freely obtainable for non-commercial purposes.

These findings have also been supported by a recent Policy Brief (June 2023) by ESFRI, Assessment of Impact of RIs[29], which underlines that impact measurement of RIs is feasible and necessary but should be undergone with a relative amount of caution. The report emphasises the importance of scientific impact while recognising the socio-economic impacts of RIs. Quality over quantity of impact assessment should be the guiding principle. Soft evidence such as impact cases and anecdotal evidence as opposed to hard evidence is the recommended way forward for showcasing impact.

Most of these findings and frameworks address all RI domains and do not concentrate on Social Sciences & Humanities specifically. The specificity of SSH Research Infrastructures stems from their influence on the vision of decision makers and social actors.

SSH RIs have all individually started to monitor and assess their scientific, societal and economic impact with the KPIs developed centrally by ESFRI. However, for this to remain a scientifically grounded and recognised assessment, more depth and time will need to be invested at all levels. The individual uniqueness of RIs also means that their impact can have many different forms.

As regards scientific impact, SSH RIs score high because they have revolutionised the way in which SSH research in Europe is conducted, as explained above. A push to more data-intensive SSH research has brought about a shift in European SSH research, slowly moving it away from nationally focused and context-dependent approaches. Not only has this shift made SSH research more international (through the availability of cross-country data), but it has also encouraged a fruitful exchange and further development of research methods.

In terms of societal impact, Humanities RIs contribute significantly to cultural impact: digitalisation of cultural artefacts (e.g., archival records, books, newspapers, manuscripts, museum collections) makes them more accessible and appreciated, boosting community integration and identity. Benefits for society include contributions to art, movies, and books that may ultimately lead to more societal awareness about the benefits of science. Other examples for societal impacts include providing open access publications, data and software for societal use and contributing to social inclusion, for example by hiring people from under-represented groups or by adopting practices

28. "Reference framework for assessing the scientific and socio-economic impact of research infrastructures", OECD (March 2019) https://read.oecd-ilibrary.org/science-and-technology/reference-framework-for-assessing-the-scientific-and-socio-economic-impact-of-research-infrastructures_3ffee43b-en#page1

29. "Assessment of Impact of RIs", ESFRI (June 2023) https://www.esfri.eu/latest-esfri-news/esfri-publishes-policy-brief-impact

promoting gender equality. There are also broader societal impacts on national research systems in the European Research Area (ERA): European RIs are a cornerstone of the ERA and impact directly on national governance in its structuring of the national research and funding systems.

Moreover, several of the SSH RIs contribute fundamental facts about the current state of European societies and policies, which is particularly important in order to fight disinformation and organised campaigns aimed to undermine public trust and legitimacy threatening European democracies.

Assessing economic impacts is more difficult, as they are solely indirect: they are to be seen as a long-term outcome of policies adopted on the basis of information and insight provided by data generated at the level of SSH RIs.

For example, in 2020, **DARIAH ERIC** launched an initiative of producing impact case studies to showcase the depth and richness of the impact achieved over the year into research communities, national consortia, and the practices and knowledge base of individual researchers in the area of (Digital) Arts and Humanities. This collection of case studies is enriched annually. The DARIAH Impact Case Studies are inspired by the UK's Research Excellence Framework[30] which included more narrative and qualitative Impact assessment approaches, more suited to the Humanities & Social Sciences. Furthermore, some RIs have a truly global reach: **SHARE ERIC** does not only cover all EU Member States in a harmonised way, but is also embedded in a network of sister studies all over the world, from the Americas to Eastern Asia. This has to do with the comparability of its data with other international ageing surveys such as HRS in USand ELSA in UK. Such comparability has allowed for a worldwide understanding of how health care policies have coped with the COVID-19 pandemic, prompting important lessons for the future. In the same vein, during the COVID-19 pandemic, ESFRI showcased many stories on its website where SSH RIs directly contributed to addressing the crisis[31].

In summary, it can be noted that societal and economic impacts can be assessed, in addition to scientific impacts. However, especially in the SSH domain, societal and economic impacts are indirect, very long-term and dispersed due to the large variety of user groups.

30. UK's Research Excellence Framework
https://www.ref.ac.uk/about-the-ref/what-is-the-ref/

31. ESFRI RIs against COVID-19 pandemic
https://www.esfri.eu/covid-19

TRENDS IN THE DOMAIN

> *RIs in the SSH domain are evolving in a fast-changing context, where the impact of technology has major effects. Europe is a knowledge society. Within a generation, more than half of the European population will have completed a tertiary level education, and the number will continue growing. Knowledge is no longer the driver of the economy alone; it will also shape social and political processes in Europe and allow citizens to be more informed. Digital technologies change the way research is done across all scientific fields, and the greatest change might be seen in SSH.*

Increasingly, data relevant for research comes from multiple sources, such as administrative documents, statistical collections, businesses, electronic devices, online transactions, social media. The competitiveness of the ERA might hinge upon the ability to harness and integrate this data alongside data generated by researchers more than ever before. However, these types of data present many challenges. They often align less with FAIR principles than traditional research data. The existence of specific collections might be obscured, accessibility issues arise due to commercial interests, data protection or copyright concerns. Data is often not well-documented, and there might be a lack of standards facilitating its interoperability.

While it might be impossible to freely disseminate all data types, it is still important to secure their access to researchers. This holds not only for data from public organisations, like statistical offices or public administrations, but also for data from the private sector including publishers, especially when there is a significant public interest at stake. In these cases, making data available may be an obligation, as outlined in the European Data Act. Unlike industrial data, data within SSH RIs is in most cases free and open, with some exceptions regarding access to cultural heritage data.

The evolution of AI, with for instance its generative systems based on Large Language Models (LLM) like ChatGPT, is determining a paradigm change, and access to high performance computing is required to researchers and students. Moreover, the evaluation of the results generated by LLM, including ethical aspects, is an important future task to be addressed by SSH RIs. Not only digitalisations have changed the scale of the results, but they have also transformed methodologies and training of all scholars. Among the actors of this transformations are new forms of data and data analytics, the use of MRI for behavioural studies, satellites used

to map and measure economic growth, laser scanner data from archaeology and art. Just like any field, researchers in the Social Sciences & Humanities domain find innovative tools to conduct research, of which many are new and expensive.

RIs like **CLARIN ERIC** and **DARIAH ERIC** must reflect on their distinctive role in this process. They must engage further showing their capacities, distinctiveness, and unique role under the light of these new developments.

Developing and enhancing the usage and impact of data about societies in Europe from various sources should be a priority. Vast amounts of scientific data about our European societies are being archived and distributed in data repositories within the network of **CESSDA ERIC** and have a large potential. Legislative reforms like the Data Governance Act (DGA) and the Digital Services Act (DSA), enabling better data access for research, should be brought to their full potential by supporting research endeavours that build on analysing large data sets. The DGA invites Member States to establish safe and secure platforms to give researchers access to data that originally stems from administrative acts. Legal and administrative obstacles that still hinder sharing of, for instance, administrative data for research should be removed also at the national level. The technological means for storing and sharing such data in an ethically responsible way are already in place.

More and more MS are using these new possibilities. The DSA enables access to large commercial platforms, giving access to social media data for research from Very Large Online Platforms (VLOPs) and Very Large Online Search Engines (VLOSEs).

More established data assets are produced and distributed by National Statistical Institutes and Eurostat within the framework of the European Statistical System (ESS). There is large capacity to enhance their impact for research. The interplay between academic research and NSIs has untapped potential that could be brought to use to gain new knowledge about climate behaviour, mobility and economic questions. Furthermore, the major innovation in public policy over the next 10 years will be the personalisation of public services, from health to education. People will rely on new knowledge to help guide and inform their decisions and choices throughout their life from birth, through their working life and on into retirement. The issues that concern and challenge EU citizens the most are largely 'social'.

Green and Digital transitions are strong European priorities and they have become the lens to look at European policies for the future. SSH RIs are moving towards focusing on the twin transition. In relation to federated and connected databases, it will be possible to understand the impact of the twin transition in a more geographically focused way – local and regional – in addition to national and European. Large cities and small communities are impacted differently and respond in dissimilar ways to the transformation implied by the digital and green transitions. Locally held social data can offer key insights about how to adapt to European diversity, especially in crucial areas such as labour relations and labour markets, market behaviour and social standards. Social policy could benefit greatly from a better understanding of a diverse European society with locally sourced data augmenting the larger national and pan-European datasets.

It is therefore key to explore trends for this domain in the significant investments made in large-scale Research and Innovation (R&I) projects through the EU Framework programmes and with attention to multidisciplinary approaches.

SSH research started to be funded under the 4th Framework Programme with a limited amount and a small number of projects. In Horizon Europe, the cluster dedicated to SSH research became far more relevant. For instance, Horizon 2020 funded projects around Migration and Democracy (under the Governance topics), many of which had significant data collections and analysis components. A summary report on **migration studies**[32] identified that the resources collected could lead to an important cluster of data for emerging infrastructures in support of population studies, mobility and, above all, migration flows.

Similarly, Horizon Europe devotes special attention to the intervention area of democracy. One important field of application will be research on our European democracies, which are the cornerstones upon which the European Union is built. Since their establishment, European democracies have faced challenges due to societal developments, and they will most likely continue to do so in the future, being confronted with climate change, energy supply, novel technologies, and the like. Thus, to strengthen the resilience of European democracies and render them fit for the future, it will be crucial and inevitable to monitor how they navigate these challenges.

Democracy research is, as of now, very fragmented. It is in fact characterised by many separate data collections spread across prominent research institutions all over Europe, each one employing different data collection methods, coding practices, and archival standards. This currently impedes a comprehensive and efficient analysis of democracy research data. The role of SSH RIs in the standardisation and harmonisation of different data objects in order to make them interoperable and linkable, and to a greater extent reusable, is therefore central to ERA.

Evolving RIs like Monitoring Electoral Democracy (MEDem) or projects like Reconstructing Democracy in Times of Crisis, A Voter-Centred Perspective (REDEM) provide data that will enable better, more comprehensive, and highly innovative comparative research on electoral democracies and electoral systems.

The research community encouraged participation in large consortia and inter- and multidisciplinary approaches in large-scale projects. Over time, cross national collaboration has also increased the presence and, only to some extent, funding for SSH and a better visibility of its research results at both European and National levels. Horizon Europe surpasses Horizon 2020 as for Cluster 2, a dedicated SSH research-driven cluster focusing on three intervention areas: democracy, cultural heritage, and social transformation. The programme also includes Humanities and Social Sciences as a cross item element, and a collaboration to be taken into account in a number of flagged topics across different clusters. The Missions launched in Horizon Europe are another example of how crucial it is to align national and European efforts for wider and more ambitious achievements in science. SSH sector data could be particularly important for addressing these missions as they can lead to Data Spaces combining academic, governmental, administrative and company data. SSH RIs are actively participating in data space initiatives including cultural heritage, language and skills. Furthermore,

32. WP10 — Deliverable 10.2 Strategic Research Agenda on Migration. Proposal number: GA 7701121, Horizon 2020: Call: H2020-SC6-REV-INEQUAL-2016-2017

a strong collaboration between SSH RIs and the emerging European Collaborative Cloud for Cultural Heritage (ECCCH) will be crucial.

More trends can be identified in other areas of development. In the area of language, both spoken and written, commercial actors are developing platforms and services at an unprecedented speed. Translation tools such as dictionaries (e.g., dict.leo.org), or translation machines like Google Translate or Deepl have become common tools for citizens as well as researchers. Spoken language tools and voice translations are developing at a similar pace, already offering or poised to offer spoken interpretation and translation services for many languages.

The interest of creating new infrastructures or extending the scope of existing ones lies in the fact that infrastructures, which could be created out of the Europe wide collection of smaller facilities, can easily become 'greater than the sum of the parts' for researchers if brought together under the coordination of a RI. For example, the number and spread of 'behavioural labs' (rather than living labs) which exist across Europe, if coordinated under a RI, could provide a significant resource to European Social and Behavioural Sciences. Research subject availability, research subject diversity, and ethics standards are all critical in appropriate subject sampling for research design innovation; at present, these and other resources remain fragmented. Structural support of such 'core facilities' could provide a significant boost to their sustainability.

Social and behavioural sciences miss a single point of observation about emerging large consortia and initiatives. They also miss a representation within the EIROforum[33], an entity that plays a crucial role in simplifying and facilitating interactions with the European Commission, the organs of the European Union, national governments, industry and so on.

All RIs in this domain offer critical services but with different characteristics. A concentrated effort should be made to see whether services can be 'unified' at least up to a certain extent, or whether connectedness in this respect can be achieved between certain RIs. It should be noted that some efforts are being tested in this regard.

An emerging generation of RIs is the one which combines physical labs with data analysis of tangible outputs; an exemplary case being **E-RIHS**.

33. EIROforum https://www.eiroforum.org/about-eiroforum/

GAPS AND NEEDS IN THE DOMAIN

The levels of funding and support in research programmes is low for the SSH domain across most Member States and Associated Countries, given the great number of disciplines it covers. This has two consequences: SSH research, especially at national level, is often characterised by small size projects, which can generate scattered data collections, and a variety of methods and scientific approaches. In some cases, scientific collections – but usually not basic research – are funded and supported by donations, small local authorities, and foundations. In the field of SSH, researchers have often been trained as individual single scholars rather than team scientists, although the impact of technology and the opportunities offered by large datasets are profoundly changing the methods and practice of these disciplines, especially in the last 15-20 years. However, this trend mainly focuses on the employment of those technologies rather than on genuine collaboration. The importance of RIs to contribute to these changes, for instance through harmonisation and support of the creation and sharing of FAIR resources and the development of new methodologies and tools, is evident; but support to RIs from the various national funding policies is not always adequate. This is true especially if considering the new technological requirements imposed by very large data and models. This includes access to high performance computing to researchers and students.

There is then a fundamental hurdle about how to finance such efforts if they lack large national investment. A separate analysis of existing datasets with such characteristics could indicate how pressing this issue may be. The aim will be to make timely provisions to support the emergence of scattered, local, and differently funded infrastructures.

EU level programmes, in particular Framework Programmes, and COST actions have had a significant influence on shaping some areas of research within the European SSH research domain. Even so, the current R&I landscape is rather fragmented and incoherent across the different countries, and this will remain a significant barrier to finding a place for RIs in the research ecosystems. Such alignment is required both thematically – within specific scientific domains such as SSH – as well as at the level of optimising national and European R&I priorities. Many RIs tend to feature in national strategies, but their founding focus is often as a national resource; the task to make them aligned to European aims and purpose comes with high costs in time and effort. In 2021, the European Research Area launched action 8, which aims to share more information to align better national and European funding policies to foster common approaches across Member States. This action, bringing together all MS interested in joining the dialogue, tries to tackle a variety of problems: sustainability of funding, equal access to RIs throughout MS, the economic and societal impact of RIs in Europe, and priority setting with regard to the focus on specific scientific and political needs[34].

Within the emergence of the European Open Science Cloud (EOSC)

34. WP10 – Deliverable 10.2 Strategic Research Agenda on Migration. Proposal number: GA 7701121 Horizon 2020: Call: H2020-SC6-REV-INEQUAL-2016-2017

in the R&I ecosystem, this positioning becomes even more urgent as Member States and Associated Countries should also report on their EOSC contribution; the EOSC partnership assumes in fact that the €500 million of EC budget will be matched by the countries. The SSH Thematic Science Cluster SSHOC, which brings together the SSH RIs with other key research organisations in the domain, will facilitate a stronger SSH voice within the EOSC.

However, the integration or collaboration of SSH research with other domains to address European priorities is still very weak. For instance, at the European Mission level – as the E-RIHS report highlights – none of the five Missions identified has properly focused or included Humanities and Social Sciences effectively, making it harder for these RIs to leverage the data that could be particularly valuable for the Missions' objectives. This situation can be exemplified by a case from the cultural heritage perspective: in the Ocean Mission, opportunities can be found to connect coastal and maritime communities with their cultural heritage and the adaptation to Climate Change to increase awareness of the value of preserving cultural landscape. In the Soil Mission, art, culture, and creative industries have been used to stimulate citizens' consensus. As for Climate-Neutral and Smart Cities Mission, there is unexplored potential for data from **E-RIHS**, **GUIDE** and **SHARE ERIC**, which can substantially contribute to new social policies on cultural heritage, ageing societies and children wellbeing.

Even more cooperation between existing RIs should be pursued along with a tentative expansion of their activities to cover fields related to the thematic areas they already serve. This also includes cooperation with RIs in other domains, which requires establishing stronger, sufficient, and stable financial support to the RIs.

CROSS-DOMAIN ASPECTS

The SSH RIs sector is a particularly versatile area: social, linguistic, and cultural data serve large demands across all science domains. Some SSH infrastructures have already demonstrated high levels of interdisciplinarity. As an example, **SHARE**'s future development will be driven by interdisciplinary studies on the interaction of health and socio-economic living conditions. Over the life course of European citizens, health and socio-economic policies determine individuals' standards and demand for their social and health care. Similarly, **E-RIHS** will enable the provision of state-of-the-art tools and services to cross-disciplinary users and communities. Among these are STEM and SSH researchers, curators and cultural heritage professionals, PhD students and technical staff: this variety encourages applications from user teams with a strong interdisciplinary character. So far, 90 successful user group leaders belonging to main academic backgrounds in the disciplines of Chemistry (20%), Humanities (50%) and Conservation Science (25%), have already been involved in working with E-RIHS data; the remaining figure includes users with a background in Engineering Physics, Earth and Life Sciences.

Metadata schemes and classification criteria either have been already advanced (as is the case for **CLARIN ERIC**), or they are gradually being established. The best practices developed can be shared with other domains. Besides metadata and standards, **CESSDA ERIC**, as an example, also promotes the need for data to be more interoperable within and beyond disciplines. Developments in the data space shall focus on the easier (machine-readable) processing of data from different disciplines to tackle societal challenges. A special focus is given to restricted and sensitive data.

SSH can also make important contributions to Destination Earth, the new digital twin of the Earth, which will help tackle climate change and protect nature[35].

RIs have already achieved great results. Now the challenge is to ensure they become front-line resources in research to contribute to the major research challenges prioritised in a programme like Horizon Europe. Recently, **ESS ERIC** has been collaborating with **ICOS ERIC** on a H2020 project to improve carbon emission monitoring at the city level and collect data from residents in those cities in the same period, with the aim to create a more complete picture and help address the Green Deal. ESS has also been linking its data to data coming from the **Environment** domain in the context of the EOSC Future project, again underlining the power of cross-domain collaboration to address the Missions and other key areas. Collaboration between the various Science Clusters including SSHOC will become increasingly important, but this should not imply duplication of work, including metadata and tools.

Advanced studies to join information about health and socio-economic conditions are well developed. Understanding the impacts of environmental changes on communities and new industrial and biological threats are all part of what future science scholars will explore and where big data approaches will be dominant. A demand for more priority focused data sets, understood through the lens of Sustainable Development Goals or science advancements in specific fields, could pave the way to faster breakthroughs.

Moreover, more comparative data focused on special groups (e.g., societal elites or vulnerable groups) along with a number of challenges coming out of them would be welcomed and useful for understanding how European societies and democracy evolve as well as for future developments.

35. Destination Earth (EC)
https://digital-strategy.ec.europa.eu/en/policies/destination-earth