



Mapping International Research Infrastructures for the Humanities, Arts and Social Sciences

A REPORT FOR THE COMMONWEALTH DEPARTMENT OF EDUCATION

FEBRUARY 2020



This report has been prepared by the Australian Academy of the Humanities using multiple sources of data. The analysis and findings are subject to the limitations of the data used. While every effort has been made to ensure the accuracy of that information, the Australian Academy of the Humanities does not make any warranty, express or implied, regarding it.

ADVISORY GROUP

Professor Duncan Ivison FAHA (Chair), Deputy Vice-Chancellor Research, University of Sydney

Professor Peter Anderson, Director, National Indigenous Research and Knowledges Network

Dr Marie-Louise Ayres, Director General, National Library of Australia

Anne-Marie Lansdown, Deputy CEO, Universities Australia

Dr Merran Smith, Director, Population Health Research Network

Professor Mark Western FASSA, Director, Institute for Social Science Research, University of Queensland

PROJECT MANAGER

Dr Kylie Brass, Director Policy and Research, Australian Academy of the Humanities

CONTRIBUTING AUTHORS

Kylie Brass

Peter Elford

Ingrid Mason

Christina Parolin

Contents

1	Executive Summary	4
1.1	Introduction.....	4
1.2	Scope	4
1.3	Key findings – International Mapping	5
1.4	Key lessons – International Mapping	5
1.5	Findings from the Gap Analysis	6
1.6	Models to consider for HASS scoping study	6
1.7	Potential pathways and priorities for Australia.....	7
2	International HASS Research Infrastructures	8
2.1	Introduction.....	8
2.2	International HASS infrastructures selected	8
2.3	Selection rationale.....	10
2.3	National profile: The Netherlands	16
2.4	Findings	19
2.5	Lessons	22
2.6	Opportunities for collaboration	30
2.7	Models for further analysis for HASS scoping	32
2.8	Conclusion	34
3	Gap Analysis	35
3.1	Introduction.....	35
3.2	The Australian landscape	35
3.3	Findings from the gap analysis	38
4	Potential Priorities	41
4.1	Australian HASS research infrastructure entity.....	41
4.2	HASS Research Data Commons	43
4.3	Social Sciences and Languages Data Hub	43
4.4	Indigenous Data Framework	44
4.5	National Digitisation Capability	44
4.6	Digital HASS Peak Capability.....	45
4.7	Implementation.....	45
	Appendix A – Research phases	48
	Appendix B: HASS Research	49
	Appendix C: HASS Data	51
	References	54

1 Executive Summary

1.1 Introduction

Australian humanities, arts and social sciences (HASS) researchers are working on pressing social and cultural challenges: from maximising the benefits of digital technologies for young people to providing the evidence base for decisions about education, employment, public health, and social and cultural policy.

The *2016 National Research Infrastructure Roadmap* (the Roadmap) identified the need for national-scale infrastructure for humanities, arts and social science (HASS) research to “drive transformations in the way researchers discover, access, curate and analyse social and cultural data”.

The Australian Academy of the Humanities was engaged by the Department of Education to map key international HASS research infrastructures – across humanities, arts, social sciences and Indigenous research – to inform the development of Australian HASS national research infrastructure (NRI).

This report outlines key findings from the international mapping and Australian gap analysis, drawing on lessons from international investments and specifically to identify common patterns in nationally or regionally significant research infrastructures, timeframes and investment levels required to establish, maintain and scale up national research infrastructures, and what may be required to collaborate and compete in data and technology-intensive HASS research globally.

Although the research infrastructures mapped for the project may fall outside what would traditionally be defined as ‘national research infrastructure’ as part of Australia’s National Collaborative Research Infrastructure Strategy (NCRIS) program, they are nevertheless national or pan-national scale infrastructures. This provides an opportunity to reconsider the definition applied in Australia in terms of international comparators.

1.2 Scope

The analysis in this report is primarily informed by a detailed mapping of research infrastructures in Europe, the UK, the Netherlands, USA, Canada and New Zealand – with the Netherlands serving as a case-study of best-practice national approach to research infrastructure development across the HASS community, including stakeholder organisations.

The infrastructures highlighted in the report are diverse in terms of scope, type and organisation. The infrastructures surveyed meet a variety of research needs and development capabilities – from large-scale, integrated and sustainable data services – including storage and use; digitisation of printed text for analysis purposes; digital platforms for discovery; high performance computing; interoperability across platforms and/or data sets; data sovereignty and privacy.

1.3 Key findings – International Mapping

International investments in HASS research infrastructure have shown pathways for infrastructure investment in:

- > heritage research, data and technologies – building capabilities in characterisation technology and pattern recognition;
- > language research, data and technologies – building capabilities in informatics, semantics and AI; and
- > social research, data and technologies – building capabilities in internet of things, civic technologies and precision services.

HASS national research infrastructures mapped for this project show staged investment building on existing research practices, peer networks, investments in research and infrastructures, and capabilities, and are heavily oriented to improving existing data assets and processes.

The international mapping identifies the emergence of new and unique HASS research infrastructures that are heavily data, information and computer science enabled and provide different arrangements of data custody, new data and large scales of data that lead to radically new research techniques.

In last ten years, international HASS research infrastructures have focused on equipping HASS researchers with the tools and techniques to take on big data challenges. In the social sciences, there have been major investments to align data archives, survey instruments and methods. Humanities and arts investments have focused on experimentation, building capacity, and platforms and digitisation technologies for cultural heritage.

In Europe, there are new large-scale text and audio-visual data collections curated and co-located with cloud and high-performance computing at a national level. They are oriented to enabling novel exploitation of new, curated, diverse and large data assets, utilising cloud and peak computing facility specialists and technologies.

1.4 Key lessons – International Mapping

International mapping tells a story of what can be achieved through national research infrastructure stimulus, indicating where Australia might seek to ‘leapfrog’ and what pitfalls to avoid. The report identified seven broad lessons for the development of HASS NRI in Australia:

1. **Laying the foundations:** There is evidence of staged investment and pathways for the development of infrastructure over time. Infrastructures have transitioned from program funding to ‘landmark’ status.
2. **Capacity-building and community-building:** supporting and encouraging researchers undertaking digitally enhanced research is a fundamental part of the process of research infrastructure development and operation.
3. **Consolidation:** funding programs are moving from seeding infrastructure build to focus on cross-cutting initiatives designed to accelerate collaboration and common infrastructure at scale, advancing interoperability, comparative research and support for big data research.

4. **Fragmentation is an issue:** this has been identified as an area to address in humanities, arts and heritage, with programs funded to address the need for stronger coordination.
5. **Balancing domain-specific and opportunities for common infrastructure development:** In Europe domain-specific development is giving way to a trend for consolidation and common infrastructure development. Yet there are indications that while some of the challenges will be solved by common approaches, there is still a need for specialised, domain-specific infrastructure build.
6. **Developing complex governance models and partnerships:** the strong trend towards consolidation across the HASS research infrastructures is leading to increasingly sophisticated governance and co-investment models.
7. **Alignment of strategies:** National strategies also recognise the critical importance of participating in, and contributing to, international and regional research infrastructures.

1.5 Findings from the Gap Analysis

Australia does not have any nationally comparable HASS research infrastructures, and this represents a significant gap in national research capacity. Consequently, research data assets currently enabling HASS research either do not exist in Australia, or are institutionally hosted and uncoordinated, and largely not FAIR (findable, accessible, interoperable and reusable).

Patterns in the Australian landscape look like the pre-stimulus phase in Europe in which Framework Program funding seeded key HASS strengths. Australia has made some investments via ANDS/NeCTAR/RDS and through the Australian Research Council's Linkage Infrastructure, Equipment, and Facilities scheme and there are Australian comparators to many of the European research infrastructures operating at an institutional level that are well placed to be evolved into national capabilities.

Much of Australia's public sector data – both social and cultural – is locked up and/or underutilised in HASS research, and in other disciplines. It is clear from the mapping survey that in Europe many of the cross-government jurisdiction challenges that are typical of the Australian federation have been addressed through formal Memoranda of Understanding (MoUs) and EU incentives (fiscal and regulatory).

1.6 Models to consider for HASS scoping study

The report suggests seven key national-scale infrastructures warrant further analysis to address gaps in Australian national research infrastructure, as part of a HASS scoping study. These include:

1. IMPACT Centre of Competence, providing state-of-the-art tools, services and facilities in document imaging, language technology and the processing of historical text.
2. Europeana, Europe's digital platform for cultural heritage.
3. SAIL Databank – providing robust secure storage and use of anonymised person-based data for research to improve health, well-being and services.

4. CLOSER, hosting longitudinal research and brings together eight world-leading social longitudinal studies.
5. HathiTrust, digitisation, text and data mining services.
6. CLARIAH, providing access to large collections of digital data and to innovative applications for the processing of data.
7. ODISSEI, a platform for the survey and collection of social science data, designed to draw the efforts of research more closely to feeding into social policy directions and outcomes.

1.7 Potential pathways and priorities for Australia

Australia is in a position to plan out the HASS NRI ecosystem based on the lessons identified in this international mapping report. The following priorities for Australia have emerged from the mapping and gap analysis:

1. Given the strong trend internationally towards consolidation across the HASS research infrastructures, Australia should consider a coordinating entity to create focus, clarify responsibility and reduce complexity for the development of requirements and roadmaps for Australian HASS research infrastructures.
2. Adopt a portfolio approach to program and infrastructure delivery to foster the capacity and capabilities of the Australian HASS sector over the next decade:
 - a. Deliver a HASS Research Data Commons
 - b. Establish a Social Science and Language Data Hub
 - c. Develop an Indigenous Data Framework
 - d. Establish a National Digitisation Capability
 - e. Develop a Digital HASS Peak capability

2 International HASS Research Infrastructures

2.1 Introduction

Internationally, governments have recognised both the potential and need for intensive digital research infrastructure for humanities, arts and social science researchers for almost two decades. The mapping undertaken for this project sought to identify key international infrastructures to inform Australia's planning for HASS research infrastructure. The aims of the project were twofold: to provide evidence-based advice on the scale and impact of public investment in international HASS research infrastructure; and an understanding of the structure and approaches to international HASS research infrastructure.

The analysis in this report is primarily informed by a detailed mapping of research infrastructures in Europe, the UK, the Netherlands, USA, Canada and New Zealand. A range of infrastructure stimuli have enabled research fields to build capability via targeted calls, such as through successive European Framework programs. Europe, Canada, and the Netherlands have also engaged in strategic roadmapping exercises, which have scoped HASS capabilities, and the UK is currently undergoing the first strategic national roadmapping of its own – which includes HASS.

An in-depth case-study of the approach to HASS research infrastructure development in the Netherlands is provided as an exemplar of a national approach to investment.

2.2 International HASS infrastructures selected

Exemplar infrastructures in these jurisdictions were identified with the assistance of a project Advisory Group – experts drawn from key stakeholder communities – HASS research, Indigenous research, the cultural and collecting sector, and universities.

The project adopted a broad definition used by the European Strategy Forum on Research Infrastructures (ESFRI) and the EU Framework Program:

Facilities, resources and services that are used by the research and innovation communities to conduct research and foster innovation in their fields. They include: major scientific equipment (or sets of instruments), knowledge-based resources such as collections, archives and scientific data, e-infrastructures, such as data and computing systems and communication networks and any other tools that are essential to achieve excellence in research and innovation.

The project was also informed by the UK's first national roadmapping exercise (in progress) which takes an 'ecosystem' view to scope the broadest range of 'research and innovation' infrastructure.

The research infrastructures mapped for the project may fall outside what would traditionally be in scope as part of Australia's National Collaborative Research Infrastructure Strategy (NCRIS) program. Infrastructure development in HASS, specifically the question of whether to develop existing infrastructures or build new capability, will require strategic investment decisions from a range of organisations including the government.

The infrastructures selected for analysis span 'humanities, arts and social sciences' domains; 'heritage' infrastructures which have successfully brought together research and cultural institutions; and 'Indigenous' research infrastructures.

Some of infrastructures mapped for this study are relatively small, institutionally based programs, others are pan-European landmark efforts. The mapping and gap analysis were carried out contiguously with the Advisory Group asked to identify key areas of Australia's strengths, gaps and prospects at the outset of the project. The mapping work was not intended to be comprehensive; it was necessarily selective.

2.2.1. Humanities & Arts/ Heritage

- > CLARIN – Common Language Resources and Technology Infrastructure (Pan-national/ European)
- > DARIAH – Digital Research Infrastructure for the Arts and Humanities (Pan-national/ European)
- > CLARIAH – Common Lab Research Infrastructure for the Arts and Humanities (Netherlands)
- > EU E-RIHS – European Research Infrastructure for Heritage Science (Pan-national/ European)
- > Europeana (Pan-national/ European)
- > Impact Centre of Competence (Digitisation) (Pan-national/ European)
- > HathiTrust (USA)

2.2.2. Social Sciences

- > CLOSER – Cohort and Longitudinal Studies Enhancement Resources (Pan-national/ European)
- > CESSDA Consortium of European Social Science Data Archives (Pan-national/ European)
- > ESS – European Social Science Survey (Pan-national/ European)
- > SAIL Databank (Pan-national/ European)
- > SHARE – Survey of Health, Ageing and Retirement in Europe (Pan-national/ European)
- > ODDISEI – Open Data Infrastructure for Social Science and Economic Innovations (Netherlands)

2.2.3 Indigenous

- > FNIGC – First Nations Information Governance Centre (Canada)
- > Te Mana Raraunga – Māori Data Sovereignty Network (New Zealand)

2.2.4 HPC Facility

- > Compute Canada (Canada)

2.3 Selection rationale

All the infrastructures selected for this mapping exercise are focussed on transforming access to and analysis of cultural and social data for research (Table 1).

The work we have undertaken in this report ranges over infrastructures with very different funding sources and involves a range of government stakeholder departments and agencies. We have sought to understand how other jurisdictions have solved problems relating to:

- > Building relationships and partnerships across sectors
- > Enabling research access to government data
- > Bringing together of cultural and research infrastructures to achieve efficiencies for respective stakeholders, particularly around digitisation challenges
- > Timing, staging and scale of investments – specifically, whether capability should lead investment, or whether investment should build capability.
- > The types of investment (is it institutional, program funding, or strategic?) to best meet research needs and achieve impact.
- > Working towards solutions for sensitive data storage and access.

The infrastructures mapped are diverse in terms of scope, type and organisation. Selections were based on success in terms of longevity (indicating recurrent funding based on evidence of impact) and meeting research needs of the respective domain communities.

The infrastructures surveyed meet a variety of research needs and development capabilities – from large-scale, integrated and sustainable data services – including storage and use; digitisation of printed text for analysis purposes; digital platforms for discovery; high performance computing; interoperability across platforms and/or data sets; data sovereignty and privacy.

All infrastructures are operational and well-established. The only infrastructures listed as not being at the operational phase is the new large-scale European Research Infrastructure for Heritage Science (EU E-RIHS) heritage science program.

Table 1: International infrastructures mapped for this project – at a glance

<i>Domains</i>	Research Infrastructure	Scope	Infrastructure type	Start	Duration	Investment
<i>Social Sciences</i>	European Social Survey (ESS)	A research-driven cross-national survey measuring the attitudes, beliefs and behaviour patterns of diverse populations of Europe across more than thirty nations. Conducted across Europe since its establishment in 2001. Through the participating countries, ESS operates as a pan-national data discovery service. Targeted at survey methodology and interoperability of data infrastructure to support comparative analysis. http://www.europeansocialsurvey.org/about/	Archive or repository Data service Discovery platform Capability building Community building Information sources Research practices and methods Research tools/ platforms	2002	16-20 Years	\$16-20M
	Survey of Health, Ageing and Retirement in Europe (SHARE)	A multidisciplinary and cross-national panel database of micro data on health, socio-economic status and social and family networks of about 140,000 individuals aged 50 or older (around 380,000 interviews). Covers 27 European countries and Israel. Pan-national cooperation around survey methods and data collection of ageing populations in Europe. Support for cross-cutting infrastructure to advance interoperability, comparative research and support for big data research. http://www.share-project.org/home0.html	Archive or repository Data service Discovery platform Capability building Community building Information sources Research practices/ methods Research tools/ platforms	2004	11-15 Years	\$21-50M
	Consortium of European Social Science Data Archives (CESSDA)	Provides large-scale, integrated and sustainable data services to the social sciences. It brings together social science data archives across Europe, with the aim of promoting the results of social science research and supporting national and international research and cooperation. https://www.cessda.eu/About	Archive or repository Data service, Discovery platform Capability building Community building Research practices and methods	2006	11-15 Years	\$101-150M

*Social
Sciences*

SAIL Databank	<p>SAIL (Secure Anonymised Information Linkage) provides secure access to demographic, health, social and education data from the resident population of Wales – a flagship for the robust secure storage and use of anonymised person-based data for research to improve health, well-being and services. Now powered by the UK Secure e-Research Platform (UKSeRP).</p> <p>https://saildatabank.com/about-us/overview/</p>	<p>Archive or repository Data service Discovery platform Capability building Community building Software service Information sources Research practices and methods HPC service Cloud storage/compute Research tools/ platforms</p>	2007	11-15 Years	\$21-50M
Cohort and Longitudinal Studies Enhancement Resources (CLOSER)	<p>Hosts longitudinal research, brings together eight world-leading longitudinal studies with participants born throughout the 20th and 21st centuries. Its work maximises the use, value and impact of these, and other, longitudinal studies to help improve understanding of key social and biomedical challenges.</p> <p>https://www.closer.ac.uk/home/what-we-do/</p>	<p>Archive or repository Data service Discovery platform Capability building Community building Information sources Research practices and methods Research libraries</p>	2012	6-10 Years	\$11-15M

Humanities & Arts	Common Language Resources and Technology Infrastructure (CLARIN)	Provides access to digital language resources for researchers, students, and citizen-scientists, especially in the humanities and social sciences, through single sign-on access. Offers solutions and technology services for deploying, connecting, analysing and sustaining digital language data and tools. Enables advanced analytics and new types of services and technologies. Links between heritage and language infrastructures, and new industries. https://www.clarin.eu/content/vision-and-strategy	Archive or repository Data service Discovery platform Capability building Community building Information sources Research practices and methods Research tools/ platforms Language data and tools	2006	11-15 Years	\$16-20M
	Digital Research Infrastructure for the Arts and Humanities (DARIAH)	Enhances and supports digitally enabled research and teaching across the arts and humanities. DARIAH is a network of researchers, expertise, information, knowledge, content, methods, tools and technologies from its member countries. It develops, maintains and operates an infrastructure in support of ICT-based research practices and sustains researchers in using them to build, analyse and interpret digital resources. https://www.dariah.eu/about/dariah-in-nutshell/	Discovery platform Capability building Community building Information sources Research practices and methods Research libraries Research tools and platforms	2006	11-15 Years	\$11-15M
	Common Lab Research Infrastructure for the Arts and Humanities (CLARIAH)	CLARIAH is a distributed infrastructure for the humanities and social sciences. It builds on various infrastructure and research projects carried out both nationally and internationally. Its focus areas are linguistics, social and economic history media studies, and text. These areas function as precursors for other disciplines and together comprise all forms of data: text, image, audio-visual material and structured data (databases). https://www.clariah.nl/en/	Discovery platform Capability building Community building Information sources Research practices and methods Research libraries Research tools and platforms Research pilots Dissemination	2012-2024	12 Years	\$42.3M

Heritage

IMPACT Centre of Competence	<p>A not for profit organisation with the mission to make the digitisation of historical printed text “better, faster, cheaper”. It provides tools, services and facilities to further advance the state-of-the-art in the field of document imaging, language technology and the processing of historical text.</p> <p>https://www.digitisation.eu/</p>	<p>Discovery platform Capability building Community building Software service Information sources Digitisation service</p>	2008	11-15 Years	\$21-50M
Europeana	<p>Europe’s digital platform for cultural heritage, Europeana is an exemplar in the advancement of cultural heritage interoperability. An aggregator service, and heavily supported by the European Library, a hub of 48 national and research libraries in Europe. It is one of the European Commission’s Digital Service Infrastructures (DSI). “As a DSI, Europeana’s objectives are to innovate the aggregation infrastructure, boost the distribution infrastructure and work towards long-term financial stability through business model innovation. All of this helps make sure that Europe’s businesses and people reap the full benefits of the technological revolution in digital services in culture.”</p> <p>https://pro.europeana.eu/our-mission/history</p>	<p>Data service Discovery platform Capability building Community building Software service Information sources Research tools and platforms</p>	2008	11-15 Years	\$101-150M
HathiTrust	<p>Text and data mining infrastructure partnership of academic and research institutions, offering a collection of millions of titles digitised from libraries around the world. Capacity for digitisation, text and data mining services to be shared across academic libraries.</p> <p>https://www.hathitrust.org/</p>	<p>Archive or repository Data service Discovery platform Capability building Community building Software/ Information Research practices and methods Research libraries Digitisation service Cloud storage/compute Research tools/ platforms</p>	2008	11-15 Years	Co-investment through membership and investment through Google deal undisclosed

Indigenous	European Research Infrastructure for Heritage Science (E-RIHS)	In pre-operational phase. Aims to deliver integrated access to expertise, data and technologies through four platforms (heritage interpretation, preservation, documentation and management) to support heritage science. Mission to deliver integrated access to expertise, data and technologies through a standardized, coordinated approach. http://www.e-rihs.eu/	Discovery platform Capability building Community building Information sources Research practices and methods Digitisation service Research tools and platforms	2016	3-5 Years	\$6-10M
	First Nations Information Governance Centre	Supports data sovereignty and the development of information governance and management at the community level through regional and national partnerships. Focused surveys and data capture for a range of health and wellbeing objectives where they impact on First Nations peoples in Canada. It adheres to free, prior and informed consent, respects nation-to-nation relationships, and recognises the distinct customs of nations. https://fnigc.ca/about-fnigc/mission.html	Data service Discovery platform Capability building Community building Information sources Research practices and methods	2010	6-10 Years	No national policy driven investment, survey project funding.
	Te Mana Raraunga - Māori Data Sovereignty Network	Established to advocate for Māori rights and interests in data to be protected as the world moves into an increasingly open data environment” https://www.temanararaunga.maori.nz/	Capability building Community building Information sources	2015	3-5 Years	No national policy driven investment, short-term research project funding.
HPC	Compute Canada	Provides essential ARC services and infrastructure for Canadian researchers and their collaborators in all academic and industrial sectors. Specific support for humanities and social sciences built into existing national computational services. https://www.computecanada.ca/about/	Capability building HPC service Cloud storage/compute service Local storage Research tools and platforms	2016	3-5 Years	\$3-5M

2.3 National profile: The Netherlands

The Netherlands offers an instructive national model for Australia. Although the Netherlands and Australia are not direct comparators in terms of size, there are factors that make the Netherlands a useful template for Australia.

The Dutch national research infrastructure roadmap (2016-2020) developed by the Netherlands Organisation for Scientific Research (NWO) includes two new entries: CLARIAH (Common Lab Infrastructure for the Arts and Humanities – which brings together CLARIN and DARIAH) and ODISSEI, a data facility, observatory, laboratory and hub led by CBS (national statistics agency), SURFSara (NREN), CentERdata (data archive, institutionally based), NWO (national research infrastructure funder) – with 30 participating organisations.

These large-scale infrastructures are intended to meet the needs of humanities and arts, and social science researchers (respectively). In the Netherlands, there is a drive to strengthen and integrate national social science and humanities research infrastructure. In the broader EU context, a Social Science and Humanities Open Cloud (SSHOC) (as part of the European Open Science Cloud initiative) is in development and NWO collaborates internationally and is connected into European research infrastructures, research programs, peer networks and joint programming initiatives.

The national profile for the Netherlands indicates EU subsidies operating prior to HASS research infrastructure investment appearing on the NWO national research infrastructure roadmap; and identifies Data Archiving and Networked Services (DANS) – a data archiving service provider for researchers in the Netherlands – support for HASS (within the national landscape) as a coordinator of EU subsidy and provider of social science data infrastructure services.

Analysis of Dutch spending on digital infrastructure (enabling research) indicates sizeable EU investments, that are leveraged and over time increased structural spending nationally to: include two new HASS research infrastructures the national roadmap and investment plan, and to fund DANS as critical part of that agenda. NWO allocated funds to support existing HASS research practices in two funding streams (as innovation programs) in 2015 and 2016, as funding for HASS, once the EU subsidies wane.

By 2017 more national investment was required to sustain digital research infrastructure for a much wider array of research domains with funding from both NWO and the Royal Netherlands Academy of Arts and Sciences (KNAW). DANS played a key role in drawing in and managing EU research infrastructure funds for HASS research and common/interoperability research and data infrastructures (nationally, regionally, and internationally). The expenditure for the Netherlands from 2012-2024 on HASS research infrastructure equates to: 73.1M AUD (42.3M CLARIAH + 30.8M ODISSEI) (Table 3).

Table 3: The Netherlands investment in HASS research infrastructure 2012-2024

CLARIAH – humanities and arts	ODISSEI – social sciences	DANS – humanities and social sciences support
Total: 2012-2024 (12 years) @42.3M AUD	Total: 2016-2024 (8 years) @30.8M AUD	Total: 2012-2019 (8 years) @60M AUD
2012-2014, 1M Euro (1.6M AUD)	2016, 0.5M Euro (0.8M AUD)	2012 ongoing ~4.5M pa, 36M Euro (60M AUD)
2015-2018, 12M Euro (17.8M AUD)	2019, 18M Euro (30M AUD)	
2019-2024, 13.8M Euro (22.9M AUD)		[2012-2015 ~4.5M Euro pa, 18M Euro, 2017 5.2M Euro pa, 2018 4.9M Euro pa]*

* See KNAW annual reports for DANS annual funding: <https://www.nidi.knaw.nl/en/about/documentst>

2.3.1. Humanities & Arts/ Heritage

The Netherlands plays an important regional role in Europe as the lead for CLARIAH. The Netherlands has developed innovative partnerships and governance models (national and global nodes) to facilitate a research infrastructure ecosystem which achieves regional impact through a combination of institutional centres of excellence, government linkages, and cross-cutting investments.

The key infrastructures supported in the Netherlands are CLARIN, DARIAH, CLARIAH, Europeana, Time Machines, DASISH (Data Service Infrastructure for the Social Sciences and Humanities – a cluster project that brings together all five ESFRI research infrastructures), the National Library of the Netherlands' KB Lab, and the eScience Center (clustered in association with CLARIAH-PLUS).

The National Library of the Netherlands, Koninklijke Bibliotheek (KB), participates in Europeana projects, and has developed services specifically for digital humanities research. CLARIAH provides datalegend, a structured data service. Time Machines in Amsterdam and Utrecht are exemplars of institutional centres of competence and national networks required to enable humanities and arts research based on common data infrastructure, i.e. linked open data, technologies, and expertise, digitised historical heritage materials (2D and 3D). Local Time Machines also operate as part of a European wide Time Machine initiative.

National networks include DEN (Dutch national knowledge institute for culture and digitalisation and a national digital heritage strategy) and CLARIAH.

CLARIAH centres in the Netherlands are: DANS (Data Archiving and Networked Services), International Institute of Social History, Dutch Language Institute, National Library of the Netherlands, Huygens ING, Meertens Institute, Max-Planck Institute, Netherlands Institute for Sound and Vision. The CLARIAH network includes Austria and Germany (where the consolidation of infrastructures has also occurred) and connected into a national eHumanities platform.

eHumanities.nl is the national platform and has partnerships with research platforms, universities, cultural institutions around expertise, research, technology, and collections.

The Netherlands features cross-sector and cross-institutional partnerships via the Centre for Digital Humanities – a partnership between Amsterdam and VU University, and the Royal Netherlands Academy of Arts and Sciences (KNAW), with support from the Netherlands eScience Center.

A major success factor for the Netherlands is the coordination and integration of research data and technologies with knowledge infrastructures across higher education and heritage to support humanities and arts research. This involves national data centres e.g. National Library and Netherlands Institute for Sound and Vision as a critical strategic partnership along with centres of competence in higher education e.g. Max Planck Institute for Psycholinguistics, Meertens Institute, Institute for Dutch Lexicology, Data Archiving and Networked Services (DANS) and Huygens ING.

The National Library of the Netherlands is a partner in Impact Centre of Competence on digitisation. As part of CLARIAH-CORE proposal, major cultural heritage institutions will provide their collections as digitised resources and are referred to as “national data centres”.

2.3.2. Social Sciences

Over time there has been cross-sector stakeholder investment around establishing large scale infrastructure building upon existing institution and national capacity in the survey and collection of social science data, and as a national node for global collaboration. Building on institutional competencies and long-term commitments in pan-national and global initiatives in social surveys and data archiving.

The collaboration around surveys commenced very early on within Europe and in global spheres (2002). Stakeholder relationships at the national level were formed through collaborative survey data collection (longitudinal) and then in the establishment of national data infrastructure and more recently the large-scale platform (ODISSEI).

ODISSEI builds upon a pre-existing consortium of multiple research organisations from across the university and public sector. The major aim of that consortium (and ODISSEI) is to increase infrastructure system efficiencies (reduce fragmentation and overlap) and coordinate effort. Australian equivalents of this infrastructure in universities are the Australian Data Archive (ADA), based at the Australian National University, and longitudinal surveys and datasets such as Household, Income and Labour Dynamics in Australia (HILDA) Survey. Public federal research agency equivalents would be the Australian Bureau of Statistics (ABS), Australian Institute for Health and Welfare (AIHW), and departments of Social Services and Education.

A shared aim in ODISSEI that operates across jurisdictions in the Netherlands is to draw the efforts of research more closely to feeding into social policy directions and outcomes. From a practices perspective the coordination of effort is designed to leverage more effectively the data and expertise in that community and share common infrastructure such as surveys and panels.

2.4 Findings

2.4.1. Investment approaches and timeframes

There are broadly two investment approaches internationally and evidence of a balance between:

1. Augmentation: building upon existing capacity and research practices, in which research domains and capability levels have matured to the point where they drive the need for collaborative infrastructure; and
2. Transformation: building new capacity and establishing new research practices -- more visionary, future-focused infrastructures which are achieved through strategic investments.

HASS national research infrastructures mapped for this project show staged investment, particularly in Europe, over the last two decades. In the main those changes have been widespread and small to medium scale investments in data infrastructure and research technologies, for example European Social Survey (ESS) (established in 2002), Survey of Health, Ageing and Retirement in Europe (SHARE) (2004), and Common Language Resources and Technology Infrastructure (CLARIN) (2004).

These investments, dating back 15 to 20 years, build on existing research practices, peer networks, investments in research and infrastructures, and capabilities, and are heavily oriented to improving existing data assets and processes. Complementary and relevant research technologies, information and expertise, to maintain and extend research skills are provided, and increase researchers' capacity to take up new research techniques.

In the last ten years, international HASS research infrastructures have focused on equipping HASS researchers with the tools and techniques to take on big data challenges.

In the social sciences, there have been major investments to align data archives, survey instruments and methods. This is evident in the cooperative relationships between CESSDA with other infrastructure projects e.g. SSHOC and SERISS and other social science-related infrastructures e.g. ESS and SHARE.

Humanities and arts investments have focused on experimentation, building capacity, and platforms and digitisation technologies for cultural heritage. In Europe, there are new large-scale text and audio-visual data collections curated and co-located with cloud and high-performance computing at a national level. These larger scale infrastructures operate on two dimensions: platforms for use by many in the cloud, and platforms connected to peak facilities for highly specialised research. They are oriented to enabling novel exploitation of new, curated, diverse and large data assets, utilising cloud and peak computing facility specialists and technologies.

There is evidence of 'top down' and 'bottom up' approaches to development. A range of infrastructure stimuli have enabled research fields to build capability via targeted calls, such as through successive European Framework programs. Europe, Canada, and the Netherlands have engaged in strategic roadmapping exercises, which have scoped HASS capabilities. The UK is currently undergoing the first strategic national roadmapping of its own, which includes a map for the 'social sciences, arts and humanities' sector.

2.4.2. Organisational structure

Very few research infrastructures examined through this mapping exercise operate as a shared facility with a centralising role. All infrastructures are of necessity operating in multi-stakeholder networks and they are supported by institutional centres of competence as key partners that provide support, and in most cases critical enabling expertise and infrastructure.

The research infrastructures included in the mapping survey are mostly distributed ‘expert networks’ composed of institutional competence and institutional infrastructures, coordinated by an entity that represents their collective, often national or European, interests. HASS infrastructures therefore tend to place a greater emphasis on community and capability building. To that extent, they can be said to be distinct from research infrastructures serving the Science, Technology, Engineering and Mathematics (STEM) sector tend to be more single-sited, compared with those in the HASS sector.

2.4.3. Partnerships

For new types of HASS research to be enabled, data needs to be unlocked, coordinated, and curated at much larger scales in new data sourcing arrangements and co-developed infrastructure partnerships with government organisations or industry.

In Europe, critical stakeholder relationships around researchers’ access to data have been effectively negotiated over time. For example, the Europeana and IMPACT cultural data infrastructure programs operate in a parallel stream to that of European research infrastructure.

Recent funding partnerships and investments of interest to this exercise are:

- > DARIAH and CLARIN programs spent two years (2016-2017) liaising with LIBER (the European Research Libraries) to establish a memorandum of understanding. It is on that basis, that the partners and members that constitute DARIAH assert: “We bring the world of cultural data with us” to the European Open Science Cloud (EOSC) initiative.
- > DARIAH and CLARIN partners and members, and collaborators have a place within the EOSC funded Social Sciences and Humanities Open Cloud program (SSHOC) – and – on the E-RIHS (an emerging ERIC).
- > CESSDA has MoU in place with other ERIC social science research infrastructures, and in 2017 tackled a “strengthening and widening” activity and reported on the state-of-the-art, obstacles, models and roadmaps for widening the data perimeter of the data services, through planning and engagement (see also their role in the RAIRD project for access to register data). The expected outcome being to “address new data sources and new actors” including statistics agencies and producers of web, transactional, administrative and historical data and to establish “establish agreements with other institutions/organizations and data producers in order to keep up with researchers’ needs, which change over time”.

2.4.4. Trends and emerging priorities

Internationally, new HASS national research infrastructures have been created in the last decade, and this has affected a kind of HASS research ‘disruption’. In the main those investments have been focused on developing entirely new types of infrastructure cloud computing and platforms and large-scale types of research infrastructure.

The new HASS research infrastructures enable entirely new research practices to be fostered, and they require new peer networks to be created in the infrastructure design and research pilots (informing those designs). New investments are directed toward constructing new data infrastructure, i.e. new types of data are drawn together (administrative, social media, and digital heritage) with cloud and high-performance computing infrastructures.

These infrastructures are oriented to enabling novel exploitation from new curated, diverse and large data assets, utilising cloud and peak computing facility specialists and technologies, the skills and research interests of data, information and computer scientists.

These types of HASS research infrastructures afford the capacity to tackle research, that currently it is not possible to undertake without that scale of infrastructure or investment. Key success factors in their development is balancing the HASS research requirements driving the design, and the data/information/computer scientists, operating as both infrastructure designers and research enabling partners.

The emergence of new and unique humanities research infrastructures is evident in the Time Machine initiatives in Europe (below), the ‘Living With Machines’ initiative in the UK, the HathiTrust in the USA, and global collaboration around digitised cultural heritage materials maintained by large national research and heritage libraries e.g. Global Dataset of Digitised Texts Network (GDD Network) (Arts and Humanities Research Council (AHRC), 2019).

In the case of the GDD Network, the fact that the UK has a separate research council for arts and humanities research has enabled strategic investment in domain-specific infrastructure. The GDD Network is funded by the AHRC, led by the University of Glasgow, in close collaboration with the HathiTrust, based in the USA, and involves key library partners as follows: the British Library, National Library of Scotland, National Library of Wales, and Research Libraries UK (RLUK).

The AHRC is also leading the development of the ‘Living with Machines’ project, which is a collaboration between the Alan Turing Institute and the British Library, together with researchers from a range of universities. Funding of £9.2 million has been awarded from the UK Research and Innovation (UKRI) Strategic Priorities Fund for the infrastructure initiative which “will take place over five years and is set to be one of the biggest and most ambitious humanities and science research initiatives ever to launch in the UK” (Alan Turing Institute, 2019).

In Europe, a pattern is emerging of new large-scale text and audio-visual data collections curated and co-located with cloud and high-performance computing at a national level. e.g. Denmark’s Cultural Heritage Cluster – a collaboration between DeIC (Danish e-Infrastructure Cooperation) and the Royal Danish Library. These larger scale infrastructures operate on two levels: platforms for use by many in the cloud, and platforms connected to

peak facility for highly specialised research. Australia is missing this national domain focused capability to support HASS research.

2.5 Lessons

International mapping tells a story of what can be achieved through national research infrastructure stimulus, indicating where Australia might seek to ‘leapfrog’ and what pitfalls to avoid.

Key lessons for Australia are in Table 4, which also includes potential priorities and identifies comparative Australian infrastructures (at project, program or institutional level). Overall, the following patterns have emerged from the research and are significant for Australia:

1. **Laying the foundations.** There is evidence of staged investment and pathways for the development of infrastructure over time. Infrastructures have transitioned from program funding to ‘landmark’ status. For example, DARIAH, which was funded in 2006 through the European Strategy Framework for Research Infrastructure (ESFRI) and reached European Research Infrastructure Consortium (ERIC) status in 2014. In the social sciences, longitudinal survey infrastructure has developed over decades, and is now undergoing further transformation through possibilities offered by technology, compute, and data linkage.
2. **Capacity-building and community-building** is a fundamental part of the process. For example, DARIAH’s national approach to boosting support for arts and humanities in a regional network (member and partner organisations) and strong programs for community building and coordination (ground up first).
3. In Europe, **infrastructure stimulus programs** (Framework Program funding have seeded infrastructure build and are now focusing (via Horizon 2020) on cross-cutting initiatives designed to accelerate collaboration and common infrastructure at scale. For example, the cross-cutting program Synergies for Europe’s Research Infrastructures in the Social Sciences (SERISS), which brings together ESS, SHARE, CESSDA, among others to advance interoperability, comparative research and support for big data research.
4. **Fragmentation in infrastructure investment** has been identified as an area to address in humanities, arts and heritage. Two Horizon 2020 programs have been funded to address the need for stronger coordination: PARTHENOS (cross-cutting to improve data practices across programs); and E-RIHS-PP (ESFRI-funded).
5. **Balancing domain-specific and opportunities for common infrastructure development.** In Europe domain-specific development is giving way to a trend for consolidation and common infrastructure development. The alignment of CLARIN and DARIAH as CLARIAH is one example of this trend, as is the new Social Science and Humanities Open Cloud (SSHOC), which is part of the general shift in data and technology intensive research to using cloud infrastructure and working in virtual environments. There are indications in the new SSHOC that while some of the challenges will be solved by common approaches, there is still a need for specialised, domain-specific infrastructure build. The lesson for Australia is that the planning for this should happen at the outset of NRI development. There is an opportunity to ensure the HASS, Indigenous and Heritage domains can be brought together as part of

the planning for a comprehensive assessment of common needs and to identify where specialist infrastructure will be required.

This type of infrastructure development work could align with the strategic policy agenda in Australia (through the Australian Research Data Commons) for stronger research data management for data enabled and driven research, improvements in the production and reuse of quality research data and software assets, and the development of domain oriented virtual environments, and cloud enabled infrastructures.

6. **Developing complex governance models and partnerships to align the needs of different sectors.** Across Europe there is a strong trend towards consolidation across the HASS research infrastructures, leading to increasingly sophisticated governance and co-investment models. Significant effort has been invested in EU legislation to address data privacy and data rights. There has been a focus also on aligning culture/heritage and research infrastructure. The IMPACT Centre of Competence is an exemplar of aligning investments in digitisation to meet the needs of European culture and society and to meet the needs of research (access to heritage data), and also the capacity to leverage collaborative networks and consolidate investments (for both sectors and institutional members).
7. **National strategies also recognise the critical importance of participating in, and contributing to, international and regional research infrastructures.** Here the Netherlands is particularly instructive. The CLARIAH network, includes Austria and Germany (where the consolidation of infrastructures has also occurred) and connected into a national eHumanities platform. The Netherlands Organisation for Scientific Research (NWO) collaborates with CESSDA, ESS, SHARE as part of the internationalisation of social sciences.

Table 4: Key lessons for Australia from the International mapping

Research Infrastructure	Lessons	Trends and priorities for Australia
European Social Survey (ESS)	<p>At a national level through the participating countries, ESS operates as a pan-national data discovery service. Targeted at survey methodology and interoperability of data infrastructure to support comparative analysis (closer to the research process).</p> <p>The impact of different funding levels across time as the mode appears to be €1.5M but ramped up in design phases (€4.4M & €6.4M). This may be an investment trend across ESFRI projects (front loaded during these phases). Investments in cross-cutting program SERISS (ESS, SHARE, CESSDA, GGP, EVS) are designed to advance interoperability, comparative research and support for big data research.</p>	<p>Informing potential Australian priorities: coordinating entity and data hub to provide national focus for identifying and facilitating access to government datasets for research.</p> <p>Australian comparators include: Longitudinal datasets e.g. HILDA, LSAC, LSIC, BNLA (via the Department of Social Services)</p>
Survey of Health, Ageing and Retirement in Europe (SHARE)	<p>Pan-national cooperation around survey methods and data collection of ageing populations in Europe. Notable growth in partners participating in surveys from 2011 onward and a jump in 2017.</p> <p>Support for cross-cutting infrastructure - SERISS (and alignment with work undertaken in ESS and CESSDA) and DASISH (another cross-cutting infrastructure). Cross-cutting program SERISS (ESS, SHARE, CESSDA, GGP, EVS) to advance interoperability, comparative research and support for big data research.</p>	<p>Informing potential Australian priorities: coordinating entity and data hub to provide national focus for identifying and facilitating access to government datasets for research.</p> <p>Australian comparators include: Longitudinal datasets e.g. ALSA (Australian Longitudinal Study of Ageing)</p>
Common Language Resources and Technology Infrastructure (CLARIN)	<p>Enables advanced analytics and new types of services and technologies. Links between heritage and language infrastructures, and new industries.</p> <p>Digitisation is a 'knowledge centre' in the CLARIN network – MoU with LIBER on digital collections (with DARIAH).</p> <p>Has not maximised user uptake and community engagement in design of data infrastructure.</p>	<p>Informing potential Australian priorities: coordinating entity, data hub, digitisation capability, digital HASS peak capability.</p> <p>Australian comparators include:</p> <p>Alveo virtual laboratory, Australian National Corpus, PARADISEC, Centre of Excellence for the Dynamics of Language, AIATSIS Collections.</p> <p>Updates to Trove recently with AUSTLANG vocabulary (partnership with AIATSIS).</p>

	<p>Centre certification process to identify useful pathways for infrastructure maturity, investment, growth and scaling up and cost-benefit for participation in CLARIN network.</p> <p>Move to blend CLARIN/DARIAH in Germany as per Netherlands as a CLARIAH.</p>	
Digital Research Infrastructure for the Arts and Humanities (DARIAH)	<p>National approach to boosting support for arts and humanities in regional network (member and partner organisations). Community building and coordination (ground up first).</p> <p>Longer time to get DARIAH operational (may indicate a need to build community capability and capacity all in parallel). By comparison language institutes established earlier and CLARIN became operational quickly (with potentially more data and technology capability in the research system).</p> <p>Strong collaborations built with digitisation.eu and MoU arrangements with LIBER (Association of European Research Libraries).</p>	<p>Informing potential Australian priorities: coordinating entity, digitisation capability, digital HASS peak capability.</p> <p>Australian comparators include: Institutional infrastructures arising from LIEF investment e.g. AusStage, AustLit, AustLII, Design and Art Australia Online.</p>
Consortium of European Social Science Data Archives (CESSDA)	<p>Early and tight coupling with national statistics bodies and university data archives. Comprehensive coverage of research data and tool requirements over time.</p> <p>Consortia approach for longitudinal, demographic data and data linkage of benefit for government, higher education and industry. Collaboration around discovery of social science data with government, e.g. Denmark cross-ministry support (culture and higher education and research) for data discoverability and access. Early linkages with national statistics bodies and university data archives.</p> <p>CESSDA 2018-2022 strategy indicates a push to global partnerships and to find "third parties" e.g. with ICSPR (USA) and Research Data Alliance. Emphasis on the breadth and depth (50 years) of experience. Cross over of CESSDA with other infrastructure projects e.g. SSHOC & SERISS; Collaboration within ERICs across domain boundaries e.g. Social Sciences and Humanities cluster; Bundling of CESSDA, ESS, SHARE</p>	<p>Informing potential Australian priorities: coordinating entity and data hub to provide national focus for identifying and facilitating access to government datasets for research.</p> <p>Australian comparators include:</p> <p>Australian Data Archive (an institutional service, that delivers national services).</p>

	together as PROGEDO at national level; Cross-cutting program SERISS (ESS, SHARE, CESSDA, GGP, EVS) to advance interoperability, comparative research and support for big data research.	
SAIL Databank	<p>Secure access to demographic, health, social and education data from the resident population of Wales. Data Linkage services. Supports medical and health research primarily. Could be re-usable for HASS research. Identifies Canada (Ministry of Health and Long-Term Care) and Australia (PHRN) as global equivalents for data linkage.</p> <p>Has a successful trajectory of combined infrastructure services (data linkage, a safe haven databank, tools) and national networks to support local, national and international collaborations. Movement from pilot stage (and design) to establish a shared national facility with formal partnerships with access controls, that serves local and national research infrastructure requirements (whether that approach could be generalised and used to meet the needs of research to access to indigenous heritage and research data in government and higher education collections).</p>	<p>Informing potential Australian priorities: coordinating entity and data hub.</p> <p>Australian comparators include:</p> <p>Population Health Research Network (NCRIS).</p> <p>Australian Bureau of Statistics (ABS Datalab remote or onsite).</p> <p>Bioplatforms Australia (NCRIS).</p> <p>Australian Data Archive (ANU)</p> <p>National Centre for Indigenous Genomics (ANU)</p>
IMPACT Centre of Competence	<p>The alignment of resources, expertise and technology (around digitisation of text and language-based heritage resources) and the link to research, and expertise around text and language-based heritage resources (the alignment of the investments in IMPACT - for heritage and in CLARIN - for research).</p> <p>Is a node of CLARIN (as a knowledge centre) also serves as a centre of competence for the heritage community to become a member of and build up a network of members and partners. Notably both heritage and higher education institutions are members of IMPACT to gain the benefits of a centre of competence.</p> <p>Exemplar of alignment of investments in digitisation to meet the needs of European cultural and society and to meet the needs of research (access to heritage data), and the capacity to leverage collaborative</p>	<p>Informing potential Australian priorities: coordinating entity and digitisation capability.</p> <p>Australian comparators include:</p> <p>Major cultural institutions (e.g. National Library of Australia, State Libraries), science and research institutions (e.g. CSIRO, University of Sydney and Melbourne) have digitisation facilities</p>

	networks and consolidate investments (for both sectors and institutional members).	
Europeana	<p>World leading in the advancement of cultural heritage interoperability. Aggregates using standards from diverse heritage organisations (and approaches to informatics) and uses linked open data practices to support data enhancement and richer semantic discovery. Community engagement through partnerships with research infrastructure programs e.g. DARIAH and Time Machine.</p> <p>Approaches to digitisation: advantages of EU funding, establishing economies of scale, and standards and quality challenges.</p>	<p>Informing potential Australian priorities: digitisation capability and digital HASS peak capability.</p> <p>Australian comparators include:</p> <p>Trove Australia (provided by the National Library of Australia, with a contributor holdings from 1000 libraries and 300 organisations with heritage collections).</p>
HathiTrust	<p>Text and data mining infrastructure arising through consortia, to meet a common research infrastructure requirement (across domains) for scholarly discovery phase (literature reviews), digital humanities, library and information science and computer science.</p> <p>Capacity for digitisation, text and data mining services to be shared across academic libraries.</p> <p>Established HaithiTrust Research Center to enable computational analysis and help meet the technical challenges researchers face when dealing with massive amounts of digital text.</p>	<p>Informing potential Australian priorities: digitisation capability and digital HASS peak capability.</p> <p>Australian comparators include:</p> <p>Trove Australia (provided by the National Library of Australia)</p> <p>Digitisation partnership.</p>
First Nations Information Governance Centre	<p>Membership comprises First Nation organisations with a focus on health and wellbeing.</p> <p>Provides services across government, higher education, industry, and community.</p> <p>Focused surveys and data capture for a range of health and wellbeing objectives where they impact on First Nations peoples in Canada.</p> <p>Funding at national level for surveys by Health Canada, and Indigenous and Northern Affairs Canada, and Employment and Social Development Canada.</p> <p>The integration of Indigenous knowledge in the research system is part</p>	<p>Informing potential Australian priorities: coordinating entity and Indigenous Data Framework.</p> <p>Australian comparators include:</p> <p>National Aboriginal and Torres Strait Islander Health Survey (NATSIHS), National Aboriginal and Torres Strait Islander Social Survey (NATSISS), National Aboriginal and Torres Strait Islander Nutrition and Physical Activity Survey (NATSINPAS), Australian Bureau of Statistics. Mayi Kuwayu Survey (National Study of Aboriginal and Torres Strait Islander Wellbeing).</p> <p>National Indigenous Languages Survey, Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS).</p>

	of the Canadian science and technology plan (2020-25).	Other organisations involved: National Aboriginal Community Controlled Health Organisation (NACCHO), Australian Institute of Health and Welfare (AIHW), Departments of Social Services (HILDA).
Cohort and Longitudinal Studies Enhancement Resources (CLOSER)	<p>Institutional centres and partnerships with universities, British Library and UK Data Service.</p> <p>Has evolved out of established survey mechanisms and data collection to a centre of competence as a hub in the research network and linking to other key stakeholders outside of higher education. CLOSER is embedded in the university context whereas in Australia, the National Centre for Longitudinal Data (NCLD) operates out of the Department of Social Services, this shifts the way the infrastructure is coordinated, and data can be made accessible for research</p>	<p>Informing potential Australian priorities: coordinating entity and data hub to provide national focus for identifying and facilitating access to government datasets for research.</p> <p>Australian comparators include:</p> <p>National Centre for Longitudinal Data (NCLD), Department of Social Services.</p> <p>Australian Data Archive.</p> <p>Australian Bureau of Statistics.</p>
Te Mana Raraunga - Māori Data Sovereignty Network	<p>Ensures data for and about Maori can be safeguarded and protected. Advocates for Maori involvement in the governance of data repositories. Supports the development of Maori data infrastructure.</p> <p>Strategic digitisation agenda (in relation to Māori) is led by the National Library of New Zealand and Archives New Zealand, in partnership with the Ministry for Culture and Heritage (Internal Affairs).</p> <p>Issues in common with Australia is the lack of Indigenous input and oversight into the use of administrative data and integration of cultural and privacy concerns.</p>	<p>Informing potential Australian priorities: coordinating entity and Indigenous Data Framework.</p> <p>Australian comparators include:</p> <p>Maia Māori Wingara (Aboriginal and Torres Strait Islander Data Sovereignty Collective), Indigenous Data Network (University of Melbourne)</p>
Compute Canada	<p>Specific support for humanities and social sciences built into existing national computational services.</p> <p>Progress on national HPC research infrastructure with publicly funded research organisations e.g. Statistics Canada. Working relationship between Compute Canada and CARL (Canadian Association of Research Libraries).</p> <p>Mix of community building (INKE), HASS projects needing cloud based research infrastructure versus HPC, and capability building (DHSI) as an</p>	<p>Informing potential Australian priorities: coordinating entity for HASS to enable access and support to existing NCRIS facilities.</p> <p>Australian comparators include:</p> <p>Not yet an Australian equivalent (through e.g. NCI or Pawsey).</p>

	indicators of growth in research uptake. Gap between higher education and government in the enabling of advanced research and the provision of national research infrastructure.	
European Research Infrastructure for Heritage Science (E-RIHS)	<p>In pre-operational phase. Designed for trans-disciplinary and interdisciplinary use.</p> <p>Aims to deliver integrated access to expertise, data and technologies through four platforms to support heritage science: ARCHLAB (physical collections), DIGILAB (digital tools and FAIR data), FIXLAB (material science tools), MOBLAB (mobile digital tools).</p> <p>Problem addressing: A need to address data infrastructure and the enabling language and image processing technologies and techniques (standardisation) across the range of disciplines in humanities and arts (and the range of heritage collection types).</p> <p>Fragmentation in infrastructure investment has been identified as an area to address in humanities, arts and heritage. Two H2020 programs: PARTHENOS (cross-cutting to improve data practices across programs); and E-RIHS-PP (ESFRI), reference the need for stronger coordination. Both infrastructures combine training, virtual environments, standardisation, community building; leveraging of a range of data, technologies, and techniques. PARTHENOS appears to focus on FAIR data and standardisation of research workflows for humanities research, as an improvement exercise across infrastructures, whereas E-RIHS is focused on all disciplinary dimensions of heritage sciences and global leadership.</p>	<p>Informing potential Australian priorities: digitisation capability and digital HASS peak capability.</p> <p>Australian comparators include:</p> <p>NCRIS Synchrotron (Melbourne) and National Imaging Facility (Brisbane). ARC funded scanning equipment: AustLii and Australian Policy Observatory). Facilities in major cultural heritage institutions (e.g. National Library digitisation) and equipment in universities (e.g. ANU CT Lab and University of Melbourne Digitisation Centre) and scientific organisations with heritage collections (e.g. CSIRO and Geoscience Australia).</p>

2.6 Opportunities for collaboration

Opportunities exist for information and expertise sharing, competency benchmarking, and formal linkages with international infrastructure. Potential opportunities identified through this project are listed in the following table alongside Australian comparators (NCRIS, national or institutional).

Table 5: Opportunities for global collaboration

AU Comparators	Potential Global Partner	Opportunity
Alveo virtual laboratory, Australian National Corpus, PARADISEC, Centre of Excellence for the Dynamics of Language. AIATSIS Collections.	Common Language Resources and Technology Infrastructure (CLARIN)	Affiliation: "third party" status to connect into global research linkages with Europe, USA and South Africa.
Institutional infrastructures arising from LIEF investment e.g. AusStage, AustLit, AustLII, Design and Art Australia Online.	Digital Research Infrastructure for the Arts and Humanities (DARIAH)	Affiliation: "third party" status to connect into global research linkages with Europe and USA.
Australian Data Archive (an institutional service, that delivers national services) .	Consortium of European Social Science Data Archives (CESSDA)	Affiliation: participate in a global community of practice. Leverage specialised sensitive data technologies.
Longitudinal datasets e.g. HILDA, LSAC, LSIC, BNLA (via the Department of Social Services)	European Social Survey (ESS)	Affiliation: connect into global research linkages with Asia, South Africa, USA, Latin America.
As above: longitudinal datasets.	Survey of Health, Ageing and Retirement in Europe (SHARE)	Affiliation: connect into global research linkages with Asia, USA, Latin America, and Israel.
Major cultural institutions (e.g. National Library of Australia, State Libraries), science and research institutions (e.g. CSIRO, University of Sydney and Melbourne) have digitisation facilities.	IMPACT Centre of Competence	Membership: participate in community of practice. Leverage specialised language technologies and corpora or lexicon.
NCRIS Synchrotron (Melbourne) and National Imaging Facility (Brisbane). ARC funded scanning equipment: AustLii and Australian Policy Observatory). Facilities in major cultural heritage institutions (e.g. National Library digitisation) and equipment in universities (e.g. ANU CT Lab and University of Melbourne Digitisation Centre) and scientific organisations with heritage collections (e.g. CSIRO and Geoscience Australia).	European Research Infrastructure for Heritage Science (E-RIHS)	Partnership: participate in community of practice and connect into global research linkages with Europe, Israel, USA, Latin America.

Trove Australia (provided by the National Library of Australia, with a contributor holdings from 1000 libraries and 300 organisations with heritage collections).	Europeana	Partnership: leverage existing global community of practice with Europe, USA and New Zealand.
MaiaM nayri Wingara (Aboriginal and Torres Strait Islander Data Sovereignty Collective), Indigenous Data Network (University of Melbourne)	Te Mana Raraunga - Māori Data Sovereignty Network	Leadership: establish global community of practice and connect into research linkages with Canada and New Zealand.
National Aboriginal and Torres Strait Islander Health Survey (NATSIHS), National Aboriginal and Torres Strait Islander Social Survey (NATSISS), National Aboriginal and Torres Strait Islander Nutrition and Physical Activity Survey (NATSINPAS), Australian Bureau of Statistics. Mayi Kuwayu Survey (National Study of Aboriginal and Torres Strait Islander Wellbeing). National Indigenous Languages Survey, Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS). Other organisations involved: National Aboriginal Community Controlled Health Organisation (NACCHO), Australian Institute of Health and Welfare (AIHW), Departments of Social Services (HILDA).	First Nations Information Governance Centre	Leadership: establish global community of practice and connect into research linkages with Canada and New Zealand.
Population Health Research Network (NCRIS). Australian Bureau of Statistics (ABS Datalab remote or onsite). Bioplatforms Australia (NCRIS). Australian Data Archive (ANU) National Centre for Indigenous Genomics (ANU)	SAIL Databank	Partnership: broaden out engagement around data linkage community of practice and connect into research linkages in UK, Europe, Canada beyond population health and into social sciences.
National Centre for Longitudinal Data (NCLD), Department of Social Services. Australian Data Archive. Australian Bureau of Statistics.	Cohort and Longitudinal Studies Enhancement Resources (CLOSER)	Affiliation: leverage existing global community of practice and connect into research linkages in UK and Europe.
Trove digitisation partnership (National Library of Australia).	HathiTrust	Affiliation: leverage existing global community of practice in USA.

In the social sciences, it is worth exploring in more detail CESSDA, ESS and SHARE; and for language-oriented humanities research, CLARIN. These three EU programs have strengths in data infrastructure and expertise, communities of practice, and their technologies are potentially portable into the Australian setting. The longitudinal surveying infrastructures i.e. SHARE and ESS also may have portable technologies.

Regarding wider humanities and arts research DARIAH is an obvious candidate, based on the strength of their communities of practice and their role the emerging ERIC E-RIHS. In addition to this, CESSDA, ESS, SHARE, DARIAH and CLARIN are all participants in the Social Sciences and Humanities Open Cloud (SSHOC) development (part of the European Open Science Cloud).

2.7 Models for further analysis for HASS scoping

In terms of models for further analysis, potentially as part of the scoping exercise to aid with addressing gaps in Australian national research infrastructure, it would be fruitful to investigate seven infrastructures further (Table 6): IMPACT, Europeana, SAIL, CLOSER, HathiTrust, CLARIAH and ODESSEI to understand the kind of research being enabled with a range of infrastructure types and capabilities. IMPACT and Europeana represent decentralised data collections, and SAIL, CLOSER, and HathiTrust represent centralised data collections. CLARIAH and ODESSEI build on pre-existing research infrastructure and consortia.

Table 6: Models for further analysis

Infrastructure	Trends	Further Analysis Areas
IMPACT	<p>Specialisation of digitisation technologies based on heritage object types and contexts, i.e. IMPACT (text) versus 3D (objects) and heritage institution type.</p> <p>An alignment of investments in digitisation to meet the needs of society and research (access to heritage data), and capacity to leverage collaborative networks and consolidate investments.</p> <p>An alignment of heritage and research resources, expertise and technology – around text and language-based heritage resources, IMPACT and CLARIN.</p>	<p>Connection between Europeana (as a pan-national broker) and other leading heritage institutions in orientation to research for expertise and to delivery heritage data (for research reuse).</p>
Europeana	<p>Co-investment compared to EU contribution has three modes: none, 1/4 or 1/2 of the funds co-contributed to the budget. This may indicate investment for national outcomes. Where the co-investment is negligible this may indicate cross-cutting interests with a pan-national outcome.</p>	<p>Examine Framework funding for humanities and arts against the funding for Europeana to ascertain any crossover. Does this impact the availability of digital material (supply) for research reuse?.</p>
Sail Databank	<p>Clustered infrastructures data, expertise, tools, compute operates as a highly effective means to attract research grants and establish local, national and global networks.</p> <p>Building diverse data processing techniques into platform and suite of services to support medical and health research. Are the tools and techniques to extract domain specific data from unstructured text and characterisation tools and techniques for image</p>	<p>Movement from pilot stage to establish a shared national facility with formal partnerships with access controls, that serves local and national research infrastructure requirements. Can this approach be generalised and used to meet the needs of research to access to Indigenous</p>

	<p>data generalisable and reusable for HASS research?</p> <p>A successful trajectory of combined infrastructure services (data linkage, a safe haven databank, tools) and national networks to support local, national and international collaborations.</p>	<p>heritage and research data in government and higher education collections?</p>
CLOSER	<p>Evolution of research infrastructure out of establishing survey mechanisms and data collection to a centre of competence as a hub in the research network and linking to key stakeholders outside of higher education.</p> <p>CLOSER is embedded in the university context whereas National Centre for Longitudinal Data (NCLD) operates out of the Department of Social Services, this shifts the way the infrastructure is coordinated, and how data can be made accessible for research.</p>	<p>How CLOSER liaises with the Australian DSS and ABS equivalents in the UK, to ascertain what network links are established and collaborations enabled.</p>
HathiTrust	<p>Text and data mining infrastructure arising through consortia, to meet a common research infrastructure requirement (across domains) for scholarly discovery phase (literature reviews), digital humanities, library and information science and computer science.</p> <p>Capacity for digitisation, text and data mining services to be shared across academic libraries, e.g. CADRE, California Digital Library.</p> <p>HathiTrust Research Center Advanced Collaborative Support Program is used as a means of addressing computational analysis challenges of the corpus.</p>	<p>What role has computer science played in large complex unstructured and structured data architectures for humanities?</p> <p>What research needs drove the delivery of digital library material as datasets for research. What factors relating to industry partnership and copyright have constrained data accessibility (for research)?</p>
CLARIAH	<p>CLARIAH consolidates existing infrastructure around language based research in HASS in the Netherlands and leverages research technologies already developed. Provides coordination platform and expertise in common infrastructure development, interoperability, as well as domain informatics. CLARIAH community of practice has broadened to include more universities and heritage and public institutions, including parliament and university libraries as partners.</p>	<p>Australia's opportunity to achieve the right clustering, cooperation and coordination from the outset of HASS NRI planning. Achieving interoperability with international infrastructure to link Australian-based records with the world is major area for further investigation</p>
ODISSEI	<p>ODISSEI builds upon a preexisting consortium of multiple research organisations from across the university and public sector. The major aim of that consortium (and ODESSEI) is to increase infrastructure system efficiencies (reduce fragmentation and overlap) and coordinate effort.</p>	<p>Achieving sustainability for social sciences infrastructure. How the 'open data infrastructure' model has enabled researchers to answer new, cross-disciplinary research questions or investigate existing questions in new ways.</p>

2.8 Conclusion

In Australia there is an opportunity to align the development of HASS NRI with the 2016 Roadmap and with key policy agendas for digital infrastructure, data sharing and citizen/consumer data rights.

International investments in HASS research infrastructure have shown pathways for infrastructure investment in:

- > heritage research, data and technologies – building capabilities in characterisation technology and pattern recognition;
- > language research, data and technologies – building capabilities in informatics, semantics and AI; and
- > social research, data and technologies – building capabilities in internet of things, civic technologies and precision services.

In general, the ‘use-value’ for Australia from these programs are in the following areas:

- > Identifying known ‘good’ (and known ‘bad’) approaches to establishing research infrastructures in HASS, which can be integrated with the existing Australia programs of collaborative infrastructure (e.g. Australian Research Data Commons – ARDC) and evolved from institutional capacity (e.g. Australian Data Archive – ADA).
- > Gaining access to technical expertise and frameworks that benefit from global adoption (e.g. IMPACT for digitisation).
- > Leveraging governance and funding (co-investment) models, some of which are at large scale (e.g. European) which could be applied in an Australian context (e.g. federated states in the case of government data, bring together research and collecting institutions).
- > Avoiding duplication of effort in newer areas through partnering and sharing of best practises, notably in indigenous research infrastructures and digital and data-intensive humanities.
- > Where new types of HASS research are to be enabled, this requires looking closely at where data needs to be unlocked, coordinated, and curated at much larger scales in new data sourcing arrangements and co-developed infrastructure partnerships with government organisations or industry.

3 Gap Analysis

3.1 Introduction

In contrast to the countries and regions surveyed for this study, there are no nationally comparable HASS research infrastructures in Australia, nor a national domain-focused capability to support computationally transformative HASS research. This represents a significant gap in Australia's national research capacity relative to other countries.

In Australia, existing HASS infrastructures are largely operating at an institutional or project-based level, and the state of play may be characterised as uncoordinated. Existing platforms separately work to standardise, harmonise and provide single points of access to data. There is not a combined set of tools to power innovation in the way researchers analyse combined datasets, and support for new research methodologies. Australia does not have a coordinated HASS NRI system which can connect data hubs, facilities and the wider research enabling ecosystem, including NCRIS funded entities and centres of excellence.

No other nations examined for this project explicitly highlight research into Indigenous health, wellbeing and culture as a national priority or key enabling feature of their large-scale research infrastructure. This is a gap in the international landscape that Australia could seek to lead and fill.

An opportunity exists for Australia to take a significant regional leadership or partnership role in a potential Indo-Pacific wide approach to research infrastructure.

3.2 The Australian landscape

Patterns in the Australian landscape look like the pre-stimulus phase in Europe in which Framework Program funding seeded key HASS strengths. Australia has made some investments via ANDS/NeCTAR/RDS and through the Australian Research Council's Linkage Infrastructure, Equipment, and Facilities scheme into institutional infrastructure and is in a good position to take advantage of strategic NRI investment.

Over a twenty-year period and particularly over the last decade, national systems within Europe have achieved transformations in HASS research, generating new skills, industries, and technology. In a country the size of Australia, with a federated system, we can take lessons from European investments in particular, which have now moved into a period of consolidation, at domain level (e.g. the collaborative agenda for CLARIN and DARIAH – (CLARIAH) and with underpinning infrastructures – the Social Sciences and Humanities Open Cloud (SSHOC) development (which is part of the European Open Science Cloud).

The project mapped select research projects and programs in Australia against international exemplars to identify areas of alignment and potential opportunity or collaboration (Table 7).

Table 7: A closer look at Australian comparators

Infrastructure Name	Associated HASS Projects	AU Comparators
Consortium of European Social Science Data Archives (CESSDA)	See UK Data Archive Data Impact Blog and Impact and Innovation Lab .	Australian Data Archive (an institutional service, that delivers national services)
European Social Survey (ESS)	See bibliography of publications based on ESS data .	Longitudinal datasets e.g. HILDA, LSAC, LSIC, BNLA (via the Department of Social Services)
Survey of Health, Ageing and Retirement in Europe (SHARE)	See publications based on SHARE data .	Longitudinal datasets e.g. ALSA (Australian Longitudinal Study of Ageing).
Compute Canada	Canadian Writing Research Collaboratory Web Archives for Longitudinal Knowledge	--
IMPACT Centre of Competence (digitisation)	See CLARIN Knowledge Centre for Denmark – publications from the Dept of Nordic Studies and Linguistics (NorS)	Major cultural institutions (e.g. National Library of Australia, State Libraries), science and research institutions (e.g. CSIRO, University of Sydney and Melbourne) have digitisation facilities. None specialise in language technologies or provide corpora or lexicon.
E-RIHS	See scientific publications from IPERION-CH (precursor to E-RIHS).	NCRIS Synchrotron (Melbourne) and National Imaging Facility (Brisbane). ARC funded scanning equipment: AustLii and Australian Policy Observatory). Facilities in major cultural heritage institutions (e.g. National Library digitisation) and equipment in universities (e.g. ANU CT Lab and University of Melbourne Digitisation Centre) and scientific organisations with heritage collections (e.g. CSIRO and Geoscience Australia).
Europeana	Venice Time Machine , Naturalis Biodiversity Center – Butterfly Species Identification , Golden Agents , Life of Newspapers	Trove Australia (National Library of Australia, with contributor holdings from 1000 libraries and 300 organisations with heritage collections.
CLARIN	See Danish node for peer reviewed papers and Swedish node for Tilltal	Alveo virtual laboratory, Australian National Corpus, PARADISEC, Centre of Excellence for the Dynamics of Language. AIATSIS Collections.
DARIAH	See CLARIAH (Netherlands node) for pilot research projects .	At an institutional level there are comparators arising from LIEF investment e.g. AusStage, AustLit, AustLII, DAAO, CoEDL.

Te Mana Raraunga	See research activities .	MaiaM nayri Wingara (Aboriginal and Torres Strait Islander Data Sovereignty Collective), Indigenous Data Network (University of Melbourne).
First Nations Information Governance Centre	See publications .	<p>National Aboriginal and Torres Strait Islander Health Survey (NATSIHS), National Aboriginal and Torres Strait Islander Social Survey (NATSISS), National Aboriginal and Torres Strait Islander Nutrition and Physical Activity Survey (NATSINPAS), Australian Bureau of Statistics.</p> <p>Mayi Kuwayu Survey (National Study of Aboriginal and Torres Strait Islander Wellbeing, National Indigenous Languages Survey, Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS).</p> <p>Other organisations involved: National Aboriginal Community Controlled Health Organisation (NACCHO), Australian Institute of Health and Welfare (AIHW), Departments of Social Services (HILDA).</p>
SAIL Databank	See uses for SAIL data and projects using SAIL data .	Australia Population Health Research Network (NCRIS), Australian Bureau of Statistics (ABS Datalab remote or onsite), Bioplatforms Australia, Australian Data Archive (ANU), National Centre for Indigenous Genomics (ANU).
CLOSER	See blog for research news .	<p>National Centre for Longitudinal Data (NCLD), Department of Social Services. Longitudinal data is maintained by survey managers for HILDA, LSAY etc. Longitudinal datasets are accessible for reuse from the Australian Data Archive and the Australian Bureau of Statistics.</p> <p>Education and training around longitudinal data analysis provided by the Australian Consortium for Social and Political Research Inc.</p> <p>Commonwealth Accredited Integrating Authorities have expertise, advice provided via the National Statistical Service on data integration.</p>
HathiTrust	See HathiTrust Research Center Advanced Collaborative Support Projects awarded in 2019.	<p>Trove Digitisation partnership. Large aggregation collection APIs: Trove</p> <p>Access to corpora: Trove Australia; Tinker (Humanities, Arts and Social Science Data Enhanced) & Alveo (Human and Communication Science) Virtual Labs, ARDC funded.</p> <p>Institutional collection APIs: State Library of New South Wales, ACMI, Museum Victoria, National Museum of Australia, State Library of Queensland</p>

3.3 Findings from the gap analysis

3.3.1. Consolidation and leadership

The mapping survey highlighted a strong consolidation trend across the international HASS research infrastructures initially established as small projects with a discipline focus, towards larger facilities with a broader Humanities or Social Science focus, or indeed a ‘HASS wide’ focus.

This is particularly evident in Europe where the role of the EU as a regional facilitator and catalyst ensures that a critical mass of expertise and resources from larger nations can enable smaller nations to gain access to research infrastructures that they otherwise might have been unable to sustain themselves.

Data curation expertise and skills and services play a significant role in the development of national (and international) capability, which in turn underpins the integration of research enabling capabilities, such as data mining, analysis (including text analysis), informatics, modelling and visualisation for the HASS community.

This trend has led to increasingly sophisticated governance models being established to support these consolidated research infrastructures exemplified by the European Research Infrastructure Consortium (ERIC), a specific legal structure that facilitates the establishment and operation of Research Infrastructures with European interest (2009).

Funding models have evolved in parallel with these governance structures, often based on a flat fee plus subscription (as a percentage of GDP) from participating nations. These co-investment models are key to ensuring such research infrastructures are sustainable. Some infrastructures, recognising the importance of broad participation, provide guidelines and costs estimates to join, e.g. CLARIN.

3.3.2. Systematic HASS research capacity

Many of the European HASS research infrastructures included in the mapping survey have been developed through project and expansion funding to build upon institutional capabilities in the first instance, and then on aligning and aggregating national capabilities.

This evolution extends back more than 15 years in the social science domains, and over a decade in the arts and humanities domains. Significantly, these research infrastructures have received central (EU) and/or co-contribution (national) funding over that entire period.

Building on existing research practises and peer networks, they are heavily oriented toward improving existing data sets and research methods, particularly when complemented by programs to raise awareness and augment researcher skills. At the European level, five of these research infrastructures are considered landmark facilities (out of a total of 37).

Australia does not have any nationally comparable HASS research infrastructures, and this represents a significant gap in national research capacity. Consequently, research data assets currently enabling HASS research either do not exist in Australia, or are institutionally hosted and uncoordinated, and largely not FAIR (findable, accessible, interoperable and reusable).

The mapping highlights that there are Australian comparators to many of the European research infrastructures operating at an institutional level (such as ADA, the Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC), and APO), or as a project within a larger program or institution (Trove, or Household, Income and Labour Dynamics in Australia (HILDA)) that are well placed to be evolved into national capabilities.

In Europe there have been two separate streams: humanities and arts, and social sciences. Europe is now embarking on how these can be better connected, which is something Australia has an opportunity to address from the outset. One of the major lessons from Europe is that disconnection between investments has necessitated a later phase ‘harmonisation’ of research infrastructures, both within and across HASS. Harmonisation is now a priority between the current ERIC infrastructures in the social sciences (European Social Survey (ESS), Survey of Health, Ageing and Retirement in Europe (SHARE)) and humanities (CLARIN – European Research Infrastructure for Language Resources and Technology) and DARIAH – Digital Research Infrastructure for the Arts and Humanities) as well as the ESFRI Landmark CESSDA (Consortium of European Social Science data Archives).

3.3.3. Access to government data

Much of Australia’s public sector data is locked up and underutilised in HASS research, and in other disciplines. It is clear from the mapping survey that in Europe many of the cross-government jurisdiction challenges that are typical of the Australian federation have been addressed through formal Memoranda of Understanding (MoUs) and EU incentives (fiscal and regulatory). Significant effort has been invested in EU legislation to address data privacy and data rights, which could usefully inform similar considerations in Australia, for example, the broad General Data Protection Regulation (GDPR).

Within our region, the New Zealand Integrated Data Infrastructure (IDI) database is a world-leading exemplar of a research infrastructure that value-adds government data for research. In the context of the current Australian activities of the Office of the National Data Commissioner and the recommendations of the Productivity Commissioner cited above, the NZ IDI appears relevant.

3.3.4. Transformative digital capability for humanities

In addition to the broad-based impact across the HASS sector of systemic infrastructure improvement activities outlined above, the international mapping identifies the emergence of new and unique HASS research infrastructures that are heavily data, information and computer science enabled and provide different arrangements of data custody, new data and large scales of data that lead to radically new research techniques.

The examples identified are mostly at the national scale with a focus on long timelines and unlocking large scale cultural heritage datasets (e.g. UK, Netherlands, Denmark, Sweden). Although there are Australian research exemplars that demonstrate this capability (e.g. Virtual Archaeology Unravels Historic Shipwreck Mystery (Pawsey, 2018)), there is no national domain focussed capability to support computationally intensive HASS research of this transformative nature.

3.3.5. Regional approach

National strategies and roadmaps for research infrastructure are primarily focussed on providing national outcomes. Significantly, the mapping for this project reveals that these national strategies also recognise the critical importance of participating in, and contributing to, international and regional research infrastructures.

As noted above the regional consolidation and leadership role the EU plays is crucial in bringing together the large community of disparate stakeholder countries and institutions from across Europe to co-fund capabilities that deliver outcomes for all of Europe.

There are lessons here for Australia at a national scale where the Federal government has the capacity to incentivise and catalyse a national federated approach to HASS research infrastructure across the broad stakeholder community. Proven collaborative mechanisms such as the National Collaborative Research Infrastructure Strategy (NCRIS) are well suited to such an opportunity.

At a larger scale, an opportunity exists for Australia to take a significant regional leadership or partnership role in a potential Indo-Pacific wide approach to research infrastructure. This might include, as it does in Europe, expertise and resources from larger nations enabling smaller nations to gain access to research infrastructures that they otherwise might not have been able to sustain themselves, as well as building stronger collaborative ties with HASS research communities within larger nations. Australia's close research links to New Zealand, and New Zealand's success with the IDI (for example), suggest such a program could be readily initiated.

The mapping identifies an international leadership opportunity gap that could be addressed through an investment focussed on Indigenous health, wellbeing and culture research, with several other countries having a related interest in indigenous research (New Zealand, Norway, Canada). Such an approach is also consistent with significant portions of Australia's HASS research interests as well as aligning strongly with Australia's reconciliation agenda, with Australia's Pacific Step-Up and Australia's broader geo-political interests as a 'soft power' (Department of Foreign Affairs and Trade).

4 Potential Priorities

Potential priority areas for Australian HASS National Research Infrastructure (NRI) development and international collaboration/partnership, are intended to address both the gaps identified and the unique characteristics of the Australian NRI landscape.

The approach recognises that this work will include initiatives that are unique to the HASS community and hence for which the HASS community needs to be primarily responsible, and initiatives that are not unique to HASS, but which need to be guided by HASS and be HASS-relevant.

In summary, Table 9 shows alignment between the international research infrastructures mapped and the proposed programs.

Table 9: International research infrastructures mapped to Australian priorities

	Coordinating	Data Hub	Digitisation	Peak	Indigenous
ESS	✓	✓			
SHARE	✓	✓			
CLARIN	✓	✓	✓	✓	
DARIAH	✓		✓	✓	
CESSDA	✓	✓			
SAIL	✓	✓			
IMPACT	✓		✓		
EUROPEANA			✓	✓	
HathiTrust			✓	✓	
FNIGC	✓				✓
CLOSER	✓	✓			
MDSN	✓				✓
COMPUTE CA	✓				
R-RIHS			✓	✓	

4.1 Australian HASS research infrastructure entity

Based on the lessons from the international mapping survey, one of the highest priorities for Australian HASS national research infrastructure development and investment is the establishment of an entity to create focus, clarify responsibility, maximise collaboration and reduce complexity regarding HASS research infrastructures. This recommendation is a direct response to the findings from the mapping exercise with regards the strong consolidation trend across HASS infrastructures in Europe outlined above and is a means to avoid fragmentation and duplication in the Australian system.

Such an entity would lead requirements gathering and roadmap development, and coordinate, facilitate and partner to deliver tools and systems to support HASS research.

This ‘focal point’ organisation, Australian HASS Research Infrastructure (AHRI; a working title), will simplify the existing mesh of interactions across the national HASS research stakeholder landscape and provide a national point of contact for international collaboration, as indicated below.

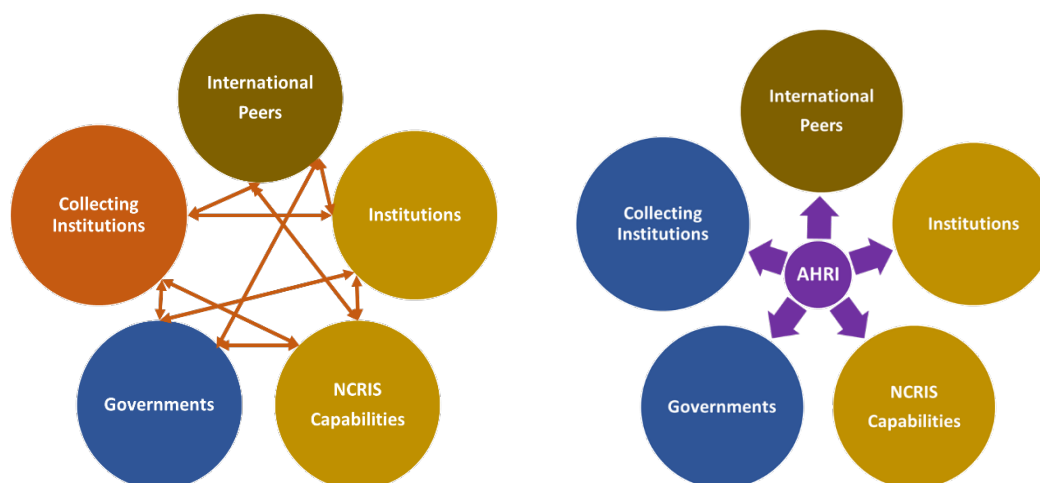


FIG.2 A single entity, Australian HASS Research Infrastructure (AHRI) will create focus, clarify responsibility and reduce complexity

AHRI is envisaged as having a mission and structure analogous to existing NCRIS capabilities such as Astronomy Australia Limited (AAL) and BioPlatforms Australia (BPA). A membership model that included both research and collecting institutions would be one mechanism to support a co-investment partnership with the Commonwealth.

In addition to the leadership and coordination role AHRI would undertake programs to foster the capacity and capabilities of the Australian HASS sector to boost its contribution to Australia’s health, environmental, economic and social wellbeing. The specific goals of these programs would be to:

- > Bridge the systemic HASS research infrastructure capacity gap through an awareness and upskilling initiative.
- > Liaise with Federal and State government departments to ensure HASS-relevant government datasets are made readily accessible to researchers.
- > Improve the consistency and coherence of access of digital data and physical objects held by collecting institutions by formally engaging the collecting institutions as partners in Australia’s national HASS research infrastructure. Collections will continue to be maintained by the collecting institutions, but it will be the responsibility of AHRI to deliver a layer of access and linkage capability across all these collections to best serve the HASS community’s research needs by addressing issues of HASS data sensitivity and access to government data.
- > Initiate work to identify humanities disciplines that would benefit from a peak HASS ‘compute+data’ capability to support transformative research methods (perhaps related to Indigenous research).

The following five programs are proposed to achieve these goals (described in more detail below):

- > HASS Research Data Commons (HRDC)
- > Social Science Data Hub (SSDH)
- > Indigenous Data Framework (IDF)
- > National Digitisation Capability (NDC)
- > Digital HASS Peak Capability (DHPC).

4.2 HASS Research Data Commons

Deliver a HASS Research Data Commons (HRDC) leveraging the experience, expertise and resources of relevant Australian programs (e.g. ARDC, Population Health Research Network (PHRN), and AURIN), relevant HASS stakeholder institutions (e.g. collecting institutions, institutional programs like the Australian Data Archive) and informed by and aligned with international programs. This would be a ‘catch up exercise’ focussed on improving awareness and expertise around some ‘minimally viable’ or ‘common denominator’ data commons tools for researchers and related professional staff (e.g. librarians).

The critical success factor would be to do this quickly and deliver something for as many HASS researchers as possible. The key elements might include the most common categories and capabilities of the programs included in the mapping (Table 10).

Table 10: Most common categories and capabilities on research infrastructure identified in mapping

Infrastructure Categories	Infrastructure Capabilities
Capability Building	Data Curation
Community Building	Data Reuse
Information Services	Domain Expertise
Discovery Platform	
Research Practises and Methods	
Data Service	
Research Tools and Platforms	

4.3 Social Sciences and Languages Data Hub

Establish a social science and languages data hub (SSDH) to provide a national focus for identifying and facilitating access to government datasets for research and the necessary tools to support research with them. This will include addressing cross-jurisdictional inconsistencies regarding data formats, access, open data policy, and working with ARDC’s sensitive data program on issues related to Indigenous data sovereignty, providence, privacy, etc. in partnership with, for example, PHRN and AURIN.

Social science research infrastructure, specifically longitudinal survey data, has been operational for a longer period than for humanities and arts in Europe and in the UK. Consequently, the global networks around this data infrastructure and associated tools and skills, are more evolved. Three of the five Social and Cultural Innovation (SCI, or HASS) landmark research infrastructures identified in the European Strategy Forum on Research Infrastructures (ESFRI) Roadmap (2018) are social science based – Consortium of European Social Science Data Archives (CESSDA), European Social Survey (ESS) and Survey of

Health, Ageing and Retirement in Europe (SHARE). Given the large number of countries participating in these programs, some from outside Europe, the Australia SSDH should strongly consider joining these programs at the European level, or some of the national programs (e.g. UK's Cohort and Longitudinal Studies Enhancement Resources (CLOSER)).

The HASS community needs to lead the SSDH program in partnership with others.

4.4 Indigenous Data Framework

The 2016 Roadmap identifies a number of existing platforms that support research into Indigenous health, social well-being, culture, language and history, noting that “creating a cohesive platform that harvests information [data], that is interoperable and that provides appropriate levels of accessibility for communities and researchers alike is required” (p.36).

The value of such a platform is well articulated in the Indigenous Data Network (IDN) roadmap (2019):

Addressing the complex issues of disadvantage for Aboriginal and Torres Strait Islander people requires relevant high-quality data. Data provides the evidence-base for the development, implementation and evaluation of effective policy solutions at local, state/territory and national levels. As the Prime Minister's 2019 Closing the Gap Report emphasises, building an evidence base is key to success. However, the lack of reliable and consistent data for Indigenous Australians results in a paucity of evidence-based Indigenous policy-making (p.3).

The Indigenous Data Framework (IDF) program proposed would address barriers to realising this opportunity, by aligning and integrating the challenges of Indigenous data sovereignty, access and repatriation with related efforts to improve and facilitate access to sensitive research data being undertaken by other national and international bodies (including in the fields of health and medical research).

The HASS community needs to be a partner in the IDF program.

4.5 National Digitisation Capability

Digitisation as a national capability is not an explicit goal of most of the programs included in the mapping but is an intrinsic requirement for many programs as it enables access to relevant data sources including historic government records, heritage collections and clinical records. It is recognised as a core capability for many of Australia's collecting institutions including the NFSA (Digitisation Strategy 2018-2015) and NLA (Digitisation, including through Trove). Integrating and aligning these institutional programs to create a national HASS research infrastructure capability is highlighted in the 2016 Roadmap (and 2011 Research Infrastructure Roadmap).

This National Digitisation Capacity (NDC) would be well suited to the ‘distributed node’ model of research infrastructure, and one that can draw on and leverage international programs, notably European Research Infrastructure for Heritage Science (E-RIHS).

The HASS community needs to be a lead in the NDC program.

4.6 Digital HASS Peak Capability

The research techniques enabled by the integration of modern computation technologies such as cloud computing and High-Performance Computing (HPC) with existing and newly generated and created big data are already transforming research techniques across many HASS disciplines.

This program would seek to extend and deepen the accessibility, usability and adoption of computationally intense methods across the broader HASS research community in area such as machine learning, artificial intelligence and text mining and analysis. The Digital HASS Peak Capability (DHPC) would necessarily leverage big data and HPC facilitates and other technology platform programs including NCI, Pawsey, and ARDC informed by similar international programs such as those identified in the mapping.

Unlike the other potential programs, which seek to achieve international parity with other national research infrastructures, one of the desired goals of the HASS Peak capability program would be to identify a field of research of significance to Australia where these techniques could be applied in a world-leading capacity. Further analysis is required to identify the specific fields of research in Australian histories and cultures, or Indigenous studies.

The HASS community needs to lead the Digital HASS Peak program.

4.7 Implementation

The table overleaf (Table 11) provides an assessment of the overall impact of the potential priorities.

EXPLANATORY NOTES:

Complexity – Is the solution/approach well understood and can be well defined? Are there exemplars at national or supra-national scale internationally? Is there an existing body of expertise and experience in Australia that can be leveraged?

Note – timeline is addressed below.

Impact – ‘Overall’ indicates scope of impact across all stakeholders – is a proxy for the estimated size of the research community that will directly benefit; ‘Researchers’ implicitly includes impact for research institutions; ‘Collecting Institutions’ includes public funded galleries, libraries, archives and museums; ‘Governments’ include state and Federal governments (policy) and agencies (service delivery).

Table 11: Program assessment: cost, complexity, impact

Program	Complexity	Impact			
		OVERALL	Researchers/Research Institutions	Collecting Institutions	Governments
Australian HASS Research Infrastructure	Low (Small Entity)	Very Broad	Consistent best practise research across sector. Avoid duplication of effort, and drive economies of scale (expertise/infrastructure).	Recognition of unique role in HASS research infrastructure. A framework to guide engagement with research institutions.	Improved focus, clearer responsibility and less complex HASS engagement. Clear Alignment between HASS RI investments and policy agenda.
HASS Research Data Commons	Low (Leverage Existing Programs)	Very Broad	Increased awareness and adoption of digital tools and resources; Builds a common framework for data initiatives and services	Improved collection curation. Platform to contribute data augmentation (interdependent value)	Aligns with broader data sharing and release and digital transformation policy agendas. Maximises sharing and interoperability of data across sectors.
Social Science Data Hub	Medium (Leverage Existing, Government Engagement)	Broad	Better access to relevant datasets; links to data networks within Australia (e.g. PHRN) and international (e.g. CESSDA); standardises safe, secure environments and outputs.	Longitudinal/linking of historic datasets (census, criminal trial records etc) for research and community use.	Improved understanding of the digital transformation of society. Better informed evidence-based policy.
Indigenous Data Framework	High (Complex Multi-shareholder issues)	Focussed	Facilitate access to Indigenous data with strong emphasis on community protocols.	Better sharing and linking of knowledge and collections, including vast community-based collections	Better informed evidence-based Indigenous policy, community-led.
National Digitisation Capability	Medium (Existing Institutional Capacity to be Coordinated)	Broad	Improved access to collections underpinned by data as enabler	Digitally preserved collections, improved curation; Formally engaged as research infrastructure	Leverages investments made in collecting institutions. Avoid duplication of effort, and drive economies of scale in expertise and infrastructure.
Digital HASS Peak Capability	High (Scoping, Mapping, Analysis)	Focussed	New research techniques enabled by large scale infrastructures	New techniques for engaging public	Social, economic and educational resource. Soft power asset.

The proposed programs are expected to commence in a staggered manner over the next five years with the objectives of delivering aggregate research impact over a decade within available resources. An indicate view of this staged approach is shown in the figure below (Fig.4). Note that the shading is suggestive of when greater intensity and progress may be required, or where existing programs can be leveraged, e.g. the ARDC already has a program of work underway regarding a HASS Research Data Commons.

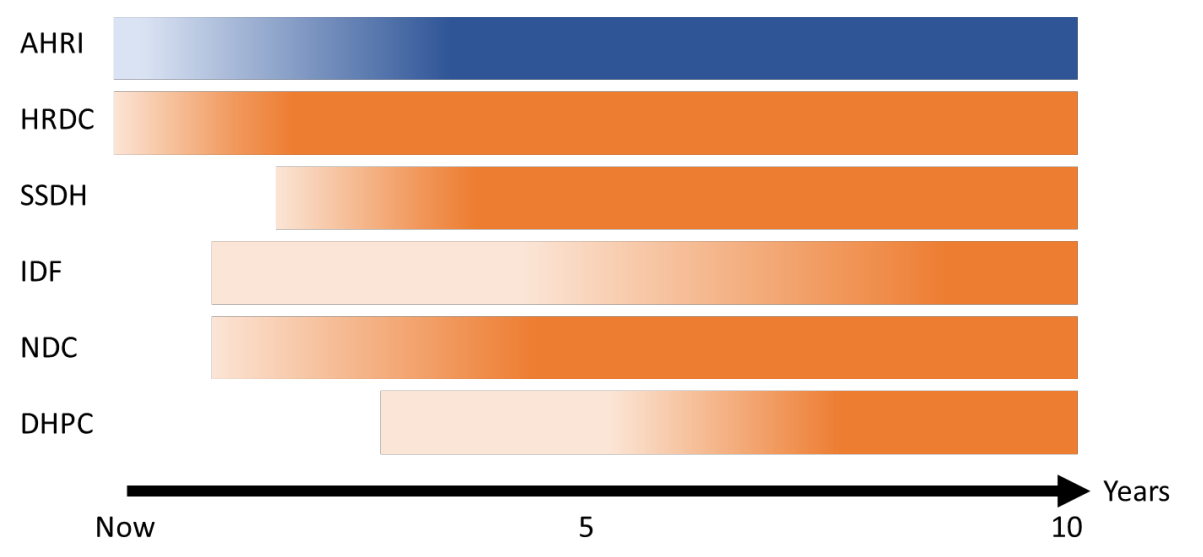


FIG.4 Proposed timelines for staged investment

There is some overlap and mutually supportive interaction between these programs, as highlighted in the table below (Table 12). Darker shading indicates a program directly addresses an AHRI goal; lighter shading indicates the program supports a goal.

Table 12: Goals and Priorities

Program (Priority)	Goals			
	Systemic HASS Capacity	Government Data Access for HASS	Integrating Collecting Institutions	Peak Digital HASS
HASS Research Data Commons				
Social Science and Languages Data Hub				
Indigenous Data Framework				
National Digitisation Capability				
Digital HASS Peak Capability				

Appendix A – Research phases

The research undertaken for this project was essentially desktop in design. The project did not have a mandate to undertake consultation with the HASS sector in Australia.

An Advisory Group was established to provide advice on international models and to guide the gap analysis. The breadth of expertise of the group spans key stakeholder communities – HASS research, Indigenous research, the cultural and collecting sector, and universities. The Advisory Group provided feedback at each stage of the research and on the final report.

Based on advice from the Advisory Group the project focused on select international models or exemplars. The analysis in this report is primarily informed by a detailed mapping of research infrastructures in Europe, the UK, the Netherlands, USA, Canada and New Zealand.

The key phases of research and focus questions were as follows:

1. MAPPING

Identify and map key international infrastructures in the HASS domain at national and pan national scale.

This phase of work addressed the following questions:

1. What international examples can Australia learn from?
2. How do they operate?
3. Are there international best practice models?

The mapping produced a database of international and pan national HASS infrastructure categorised according to type, scale, funding, costs, organisational/governance structure, communities of users. Where possible the project was able to draw out some data on levels of internationalisation, outputs, impact and sustainability.

2. GAP ANALYSIS

Identify gaps in Australian HASS infrastructure as informed by the analysis of the mapping of international models. This phase of work addressed the following questions:

4. Are there international exemplars which could apply or be adapted to Australian HASS NRI development?
5. What are the conditions required for Australia to operate similar models to those international exemplars?

The gap analysis drew on the project's Advisory Group for its:

- > Evaluation framework for assessing international HASS infrastructures and use-value for Australia.
 - > Analysis of the gaps in RI in Australia vs international.
 - > Identification of potential priorities and areas of focus, such as digitisation, international interoperability and access and data management infrastructures.
-

3. OPPORTUNITIES

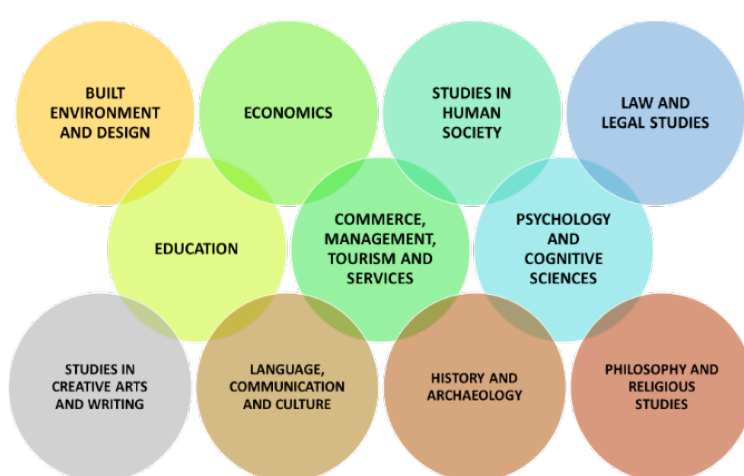
The final report to inform consideration of research infrastructure investment and delivery for HASS NRI in Australia, includes an analysis of lessons and opportunities from international HASS research infrastructure investments, a list of potential areas of cooperation/collaboration and options for HASS NRI into the future, and identifies where further work needs to be undertaken.

Appendix B: HASS Research

Humanities, Arts and Social Science (HASS) research enhances understanding about Australia, the Australian people, and Australia's place in the world.

More broadly, HASS research provides new frameworks for the analysis of people in the world and the products and outcomes of human activity. The disciplines that make up the HASS sector are fundamental to the development of a knowledge-based economy and in developing inter-disciplinary solutions to a broad range of complex problems and issues such as climate change, social cohesion, the impact and relevance of technological development, future workforce solutions, resource management, health and welfare.

The humanities investigate how people experience, understand and describe the world and their place in it. The humanities examine human cultures, values and beliefs (Australian Academy of the Humanities. 2019). The social sciences work on the systematic development of logic and evidence to understand human behaviour in its social setting, including the nature of economic, political, and community activity and institutions (Academy of the Social Sciences in Australia, 2019).



HASS comprises more than 50 disciplines at the granular four-digit) field of research level (Fig.1) (Appendix A).

Australian HASS researchers comprise 41 per cent (16,488 FTE) of the university-based system, based on the latest Excellence in Research Australia (ERA) audit (Australian Research Council, 2019).

FIG.1 HASS FIELDS OF RESEARCH

SOURCE: Australian Bureau of Statistics (2008), ANZSRC Fields of Research.

HASS research is undertaken at every university in Australia, and HASS is responsible for teaching 61 percent (606,721 students) of the university population in Australia (derived from Department of Education higher education statistics collection, where humanities, arts and social sciences (HASS) comprises, at the broad Field of Education level 12 to 22).

Humanities, Arts and Social Sciences (HASS) fields of research, two- and four-digit level

Social, Behavioural and Economic Sciences (SBE)	Humanities and Creative Arts (HCA)
13 EDUCATION	12 BUILT ENVIRONMENT AND DESIGN
1301 Education Systems	1201 Architecture
1302 Curriculum and Pedagogy	1202 Building
1303 Specialist Studies in Education	1203 Design Practice and Management
1399 Other Education	1204 Engineering Design
14 ECONOMICS	1205 Urban and Regional Planning
1401 Economic Theory	1299 Other Built Environment and Design
1402 Applied Economics	18 LAW AND LEGAL STUDIES
1403 Econometrics	1801 Law
1499 Other Economics	1802 Maori Law
15 COMMERCE, MANAGEMENT, TOURISM AND SERVICES	1899 Other Law and Legal Studies
1501 Accounting, Auditing and Accountability	19 STUDIES IN CREATIVE ARTS AND WRITING
1502 Banking, Finance and Investment	1901 Art Theory and Criticism
1503 Business and Management	1902 Film, Television and Digital Media
1504 Commercial Services	1903 Journalism and Professional Writing
1505 Marketing	1904 Performing Arts and Creative Writing
1506 Tourism	1905 Visual Arts and Crafts
1507 Transportation and Freight Services	1999 Other Studies in Creative Arts and Writing
1599 Other Commerce, Management, Tourism and Services	20 LANGUAGE, COMMUNICATION AND CULTURE
16 STUDIES IN HUMAN SOCIETY	2001 Communication and Media Studies
1601 Anthropology	2002 Cultural Studies
1602 Criminology	2003 Language Studies
1603 Demography	2004 Linguistics
1604 Human Geography	2005 Literary Studies
1605 Policy and Administration	2099 Other Language, Communication and Culture
1606 Political Science	21 HISTORY AND ARCHAEOLOGY
1607 Social Work	2101 Archaeology
1608 Sociology	2102 Curatorial and Related Studies
1699 Other Studies in Human Society	2103 Historical Studies
17 PSYCHOLOGY AND COGNITIVE SCIENCES	2199 Other History and Archaeology
1701 Psychology	22 PHILOSOPHY AND RELIGIOUS STUDIES
1702 Cognitive Science	2201 Applied Ethics
1799 Other Psychology and Cognitive Sciences	2202 History and Philosophy of Specific Fields
	2203 Philosophy
	2204 Religion and Religious Studies
	2299 Other Philosophy and Religious Studies

Source: Australian Bureau of Statistics (2008) 'Australian and New Zealand Standard Research Classification (ANZSRC)', cat. no. 297.0. Available from

<http://www.abs.gov.au/Ausstats/abs@.nsf/Latestproducts/4AE1B46AE2048A28CA25741800044242?opendocument>

Appendix C: HASS Data

HASS Research Infrastructures

Research infrastructures in HASS comprise:

- > Physical collections (including artefacts and larger physical structures such as archaeological sites) and their storage facilities
- > Galleries, Libraries, Archives and Museums (including both physical objects and digital artefacts)
- > Digital infrastructure (including digital record creation, digital data storage and tool sets)
- > Laboratory facilities for heritage science and archaeology
- > Infrastructures for data collection, services, linkage and analysis, such as public statistics and longitudinal surveys. Data collection may span topic areas such as election studies, involve long-term multi-generational population studies and surveys, and more recently social media data

The assemblages of data – data files, datasets, databases and data streams – collected, generated and curated by these infrastructures have all the characteristics of ‘big data’ in that their volume, their variety and the velocity of their creation pose severe challenges for many conventional analytical and computational methods. As a result, HASS research infrastructures, both internationally and in Australia, are increasingly focused on equipping HASS researchers with the tools and techniques to take on these challenges.

Government data

The Australian HASS community, like other research communities, makes extensive use of data held by federal and state agencies, including Departments of Health, Departments of Education, Departments of Social Services, Departments of Innovation and Industry, the Australian Bureau of Statistics (ABS), the Australian Taxation Office (ATO) and many others. There are significant challenges in achieving effective and consistent access to data across all jurisdictions of Australia’s federation, that are not unique to the HASS sector, as identified by the Productivity Commission’s report into *Data Availability and Use* (2017):

Governments across Australia hold enormous amounts of data, but mostly lag behind other comparable economies in beneficially using data beyond the purposes for which it was initially collected, or allowing others access to do so (p.24).

The Government’s response proposed reforms to:

... empower Australian citizens, governments, industries and researchers to use and share data, while maintaining the strict privacy, security and transparency safeguards essential to maintain trust in the system.

With the following goal:

These advances to Australia's data system will mean we can harness the power of data to drive innovation and opportunity for the Australian economy (p.1).

The HASS research sector is fundamental to realising this goal through the development of a social license to use data consistent with current activities with the Office of the National Data Commissioner and the Data Sharing and Release reforms. Using mechanisms like the EU Responsible Research and Innovation Toolkit (which helps to “align ... research and innovation processes to societal needs and challenges” (Responsible Research and Innovation Project, n.d)) alongside the Five-Safes Framework (“a multi-dimensional approach to managing disclosure risk” (Australian Bureau of Statistics, 2017)) will provide a way to secure and maintain the social license, which will be fundamental to using integrated data for research.

Indigenous data

Improving the lived experience of Australia's first nations peoples requires that data is made accessible for research. Advancing cultural connection and social cohesion of Australians and the wider Indo-Pacific region and globally requires data access and sharing, and services and processing tools that are co-located. Improving social and economic outcomes for Australians will be underwritten by the quality of data and data linkage that will bring together longitudinal social survey data with health and environmental data.

Research related to Australia's Aboriginal and Torres Strait Islander people (Indigenous research) faces challenges relating to data sovereignty and access.

Indigenous data needs a variety of community consent and access controls to allow Indigenous communities to access and selectively and safely release their data to individual researchers. A similar process also allows digitised materials to be repatriated to Indigenous communities.

Responses to the 2016 Roadmap, from a range of stakeholders, emphasised the need for Indigenous leaders and organisations to be at the heart of determining capabilities required to support Indigenous aspirations and needs. Indigenous co-participation will be essential in any project involving Indigenous communities directly and consideration given to cultural sensitivities and associated rights.

Several organisations in Australia are working on Indigenous data governance and access policies and strategies, including the Australian Institute for Aboriginal and Torres Strait Islander Studies (AIATSIS), the Indigenous Data Network (based at the University of Melbourne), and the National Aboriginal and Torres Strait Islander Consortium.

Collecting institutions

Galleries, Libraries, Archives and Museums (GLAMs), the collecting institutions, are therefore important stakeholders in Australian national HASS research infrastructure, particularly those that are federally funded with an explicit national collection mandate, e.g. National Library of Australia (NLA), National Archives of Australia (NAA), National Museum of Australia (NMA), and the National Film and Sound Archive (NFSA).

HASS researchers rely heavily and opportunistically upon the digitisation and digital access agendas of these public sector institutions. The value of these digital services is highlighted

by, for example, the frequent reference by individual researchers and bodies within the HASS sector to the usefulness and importance of the Trove service provided by the NLA.

Although nationally funded collecting institutions typically have a national mandate to collect, it is *not* for the sole purpose of supporting research, as the ‘mission to share’ is only a portion of their typical Collect-Preserve-Share charter. Consequently, a key aspect of the relationship between the HASS research community and the collecting and cultural institutions is that there is a high level of interdependent value. Research infrastructures that improve access to objects within collections facilitates research that creates new knowledge and data that can be attached to the objects. This process improves the collection, enables the linking of data and objects between collections and improves the value of the collection to the public.

Existing programs that operate across many collecting institutions in support of their activities enable and facilitate access to collections for researchers, and hence can also be considered components of national research infrastructure, e.g. the National eDeposit system is a national system operating across nine jurisdictions (and related legislation, accounting standards, Crown Solicitors, collecting and description policies).

References

- Academy of Social Sciences in Australia (2019), 'About the Academy', <https://www.assa.edu.au/about-the-academy/>
- Alan Turing Institute (2019), 'Living With Machines', <https://www.turing.ac.uk/research/research-projects/living-machines>
- Arts and Humanities Research Council (AHRC) (2019), AHRC Network for a Global Dataset of Digitised Texts, <https://gddnetwork.arts.gla.ac.uk/>
- Astronomy Australia Limited (n.d), <http://www.astronomyaustralia.org.au/>
- Australian Academy of the Humanities (2019), 'What are the Humanities?', <https://www.humanities.org.au/about-us/what-are-the-humanities/>
- Australian Bureau of Statistics (2017) 'Managing the Risks of Disclosure: the Five Safes Framework', 1160.0 - ABS Confidentiality Series, August 2017, <https://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/1160.0Main%20Features4Aug%202017>
- Australian Bureau of Statistics (2008) 'Australian and New Zealand Standard Research Classification (ANZSRC)', cat. no. 297.0, <http://www.abs.gov.au/Ausstats/abs@.nsf/Latestproducts/4AE1B46AE2048A28CA25741800044242?opendocument>
- Australian Government (2018), *2016 National Research Infrastructure Roadmap* https://docs.education.gov.au/system/files/doc/other/ed16-0269_national_research_infrastructure_roadmap_report_internals_acc.pdf
- Australian Government (2011), *2011 Strategic Roadmap for Australian Research Infrastructure*, https://docs.education.gov.au/system/files/doc/other/national_collaborative_research_infrastructure_strategic_roadmap_2011.pdf
- Australian Government (2008), *Strategic Roadmap for Australian Research Infrastructure* https://docs.education.gov.au/system/files/doc/other/national_collaborative_research_infrastructure_strategic_roadmap_2008.pdf
- AustLit (n.d), <https://www.austlit.edu.au/>
- Analysis and Policy Observatory (n.d), <https://apo.org.au/>
- Australian Research Council (2019a), *ERA National Report: State of University Research 2018-19*, <https://dataportal.arc.gov.au/ERA/NationalReport/2018/>
- Australian Research Council (2019b), *Engagement and Impact Assessment 2018*, <https://dataportal.arc.gov.au/EI/NationalReport/2018/>

- Australian Research Council (2019c) Impact Study: ‘Young and Resilient: Transforming research, policy and practice to maximise the benefits of digital technologies for young people’, <https://dataportal.arc.gov.au/EI/Web/Impact/ImpactStudy/800>
- Australian Research Council (2019d) Impact Study: ‘The Household, Income and Labour Dynamics in Australia (HILDA) Survey: Changing the Way We Think About Life in Australia’, <https://dataportal.arc.gov.au/EI/Web/Impact/ImpactStudy/872>
- Australian Research Council (2019e), Media Release: Centre of Excellence for Automated Decision-making and Society, 9 October, <https://www.arc.gov.au/news-publications/media/media-releases/improving-automated-decision-making>
- Australian Research Council (2019f), Media Release: Centre of Excellence for the Digital Child, 27 September, <https://www.arc.gov.au/news-publications/media/media-releases/supporting-children-digital-age>
- BioPlatforms Australia (n.d), <https://www.bioplatforms.com/what-we-do/>
- CLARIN (n.d) Information for Potential Members’, <https://www.clarin.eu/content/information-potential-new-members>
- Commonwealth of Australia, Department of the Prime Minister and Cabinet, (2018), *The Australian Government’s response to the Productivity Commission Data Availability and Use Inquiry*, <https://www.pc.gov.au/inquiries/completed/data-access/data-availability-use-government-response.pdf>
- Department of Foreign Affairs and Trade (n.d) ‘Soft Power Review’, <https://dfat.gov.au/people-to-people/soft-power-review/Pages/soft-power-review.aspx>
- European Commission (2009) *Community Legal Framework for a European Research Infrastructure Consortium (ERIC)* https://ec.europa.eu/research/infrastructures/pdf/council_regulation_eric.pdf
- European Commission (n.d), ‘What Are Research Infrastructures?’, https://ec.europa.eu/info/research-and-innovation/strategy/european-research-infrastructures_en
- European Commission (n.d), ‘What is an ERIC?’, https://ec.europa.eu/info/research-and-innovation/strategy/european-research-infrastructures/eric_en
- European Strategy Forum on Research Infrastructures (2018), *Strategy Report on Research Infrastructures Roadmap Part 2: Landscape Analysis* <http://roadmap2018.esfri.eu/media/1050/roadmap18-part2.pdf>
- European Strategy Forum on Research Infrastructures (n.d), ESFRI Glossary <https://www.esfri.eu/glossary>
- European Union (n.d) General Data Protection Legislation, <https://eugdpr.org/>
- Indigenous Data Network (n.d) <https://mbspgh.unimelb.edu.au/centres-institutes/centre-for-health-equity/research-group/indigenous-data-network>
- Indigenous Data Network (2019) *Indigenous Data Network Roadmap*, https://mbspgh.unimelb.edu.au/_data/assets/pdf_file/0011/3070487/NIDN-2019-03-08-Indigenous-Data-Network-Roadmap-190301A.pdf

- National Collaborative Research Infrastructure Strategy (NCRIS) Network (2017), 'Facilitating Real Research Impact: NCRIS Supports Prestigious ARC Centres of Excellence', <https://www.ncris-network.org.au/single-post/2017/11/06/Facilitating-real-research-impact>
- National Film and Sound Archive (2018), *Digitisation Strategy 2018-2025*, <https://www.nfsa.gov.au/footer/corporate-information/publications/digitisation-strategy>
- National and State Libraries Australasia (2019), National eDeposit, <https://ned.gov.au/portal/>
- National Library of Australia (n.d), Trove, <https://trove.nla.gov.au/>
- Office of the National Data Commissioner (2019), Data Sharing Principles, <https://www.datacommissioner.gov.au/safeguards/sharing-principles>
- Responsible Research and Innovation Project (n.d), 'The PRISMA Responsible, Research & Innovation Toolkit', <https://www.rrr-tools.eu/-/the-prisma-responsible-research-innovation-toolkit>
- Pawsey Supercomputing Centre (2016), 'Virtual Archaeology Unravels Historic Shipwreck Mystery', https://pawsey.org.au/wp-content/uploads/2018/10/AndrewWoods_HMASSydneyJuly16web.pdf
- Productivity Commission (2017), *Data Availability and Use*, <https://www.pc.gov.au/inquiries/completed/data-access/report/data-access.pdf>
- Securities Industry Research Centre of Asia-Pacific (n.d), <https://www.sirca.org.au/about-sirca/>
- Statistics New Zealand (2019), Integrated Data Infrastructure, http://archive.stats.govt.nz/browse_for_stats/snapshots-of-nz/integrated-data-infrastructure.aspx
- Statistics New Zealand (n.d), Integrated Data Tools, <https://www.digital.govt.nz/showcase/integrated-data-tools/>
- UK Research and Innovation (2019), UKRI Infrastructure Roadmap Progress Report, <https://www.ukri.org/files/infrastructure/progress-report-final-march-2019-low-res-pdf/>



ATTACHMENT: INFRASTRUCTURE PROFILES

MAPPING INTERNATIONAL RESEARCH
INFRASTRUCTURES FOR THE HUMANITIES, ARTS AND
SOCIAL SCIENCES

FEBRUARY 2020

A NOTE ON CONTRIBUTING RESEARCH

This contributing research has been prepared by the Australian Academy of the Humanities using multiple sources of data. The analysis and findings are subject to the limitations of the data used. While every effort has been made to ensure the accuracy of that information, the Australian Academy of the Humanities does not make any warranty, express or implied, regarding it.

ADVISORY GROUP

Professor Duncan Ivison FAHA (Chair), Deputy Vice-Chancellor Research, University of Sydney

Professor Peter Anderson, Director, National Indigenous Research and Knowledges Network

Dr Marie-Louise Ayres, Director General, National Library of Australia

Anne-Marie Lansdown, Deputy CEO, Universities Australia

Dr Merran Smith, Director, Population Health Research Network

Professor Mark Western FASSA, Director, Institute for Social Science Research, University of Queensland

PROJECT MANAGER

Dr Kylie Brass, Director Policy and Research, Australian Academy of the Humanities

CONTRIBUTING AUTHORS

Ingrid Mason

Kylie Brass

Contents

1	Infrastructure Profiles	4
1.1	Common Language Resources and Technology Infrastructure (CLARIN).....	4
1.2	Digital Research Infrastructure for the Arts and Humanities (DARIAH)	9
1.3	Compute Canada	13
1.4	Consortium of European Social Science Data Archives (CESSDA)	15
1.5	European Social Survey (ESS).....	19
1.6	Survey of Health, Ageing and Retirement in Europe (SHARE).....	23
1.7	IMPACT Centre of Competence.....	26
1.8	European Research Infrastructure for Heritage Science (E-RIHS)	29
1.9	Europeana.....	32
1.10	Te Mana Raraunga - Māori Data Sovereignty Network.....	37
1.11	First Nations Information Governance Centre	42
1.12	SAIL Databank.....	46
1.13	Cohort and Longitudinal Studies Enhancement Resources (CLOSER)	51
1.14	HathiTrust	54
1.15	Country Profile: The Netherlands.....	58
	The Netherlands: Humanities and Arts.....	58
	The Netherlands – Social Sciences.....	64
1.16	Additional Infrastructures.....	70
	Appendix B – Checklist / Evaluation Criteria	72

1 Infrastructure Profiles

1.1 Common Language Resources and Technology Infrastructure (CLARIN)

“CLARIN makes digital language resources available to scholars, researchers, students, and citizen-scientists from all disciplines, especially in the humanities and social sciences, through single sign-on access. CLARIN offers long-term solutions and technology services for deploying, connecting, analysing and sustaining digital language data and tools. CLARIN supports scholars who want to engage in cutting edge data-driven research, contributing to a truly multilingual European Research Area”, <https://www.clarin.eu/content/vision-and-strategy>

Administration

Funding sources	National, H2020, Institutional, FP (EU Framework Programme)
Jurisdiction	Europe
Leader	Coordinating country the Netherlands (Utrecht University) via CLARIAH.
Domains	Humanities, Arts, Social Science, Linguistics
Start date	2006
Timescale	6-10 Years
Investment	\$16-20M
Phases	Pre-operational (design/plan/construct), Operational (maintain/consolidate/grow)

Structure

Features	Regional impact, Institutional centres of excellence, National nodes, Global nodes
Type	Archive or repository service, Data service, Discovery platform, Capability building, Community building, Information sources, Research practices and methods, Research tools and platforms, Language data and tools
Capabilities	Data curation, Data reuse, Data mining and analysis, Informatics and data modelling, Domain expertise
Organisation	Distributed, Virtual
Governance	Coordinating entity (hub), Network (participating nodes), Decentralised (institutional), Governing body, Executive body, Operational function, Outreach and liaison function, Administration function, Technologies function, Informatics function, Policy function, Expert advisors
User community	In 2016 CLARIN language resources were accessed by 60 unique visitors/day. See CLARIN Value Proposition Report (2016).
Institutional centres	2017 Annual Report (CLARIN) record 10 knowledge centres, the latest count is 11 (representing Denmark, Spain, Austria, Sweden, Czech Republic, USA, Norway, Poland, Germany interests and capacity) and this requires CLARIN certification.

	<p>~57 centres, certification for centre status (B, C, E, T K)</p> <p>CLARIN-DK for Denmark as an example:</p> <p>University of Copenhagen</p> <p>Aalborg University</p> <p>Aarhus University</p> <p>Southern Danish University</p> <p>The Royal Danish Library (a national heritage library)</p> <p>CLARIN-D for Germany as an example:</p> <p>Bavarian Archive for Speech Signals, Ludwig-Maximilian University of Munich</p> <p>Berlin-Brandenburg Academy of Sciences and Humanities, Berlin</p> <p>Institute for the German Language, Mannheim</p> <p>Max Planck Institute for Psycholinguistics, Nijmegen</p> <p>Department of Linguistics, University of Tübingen</p> <p>Hamburg Centre for Language Corpora, University of Hamburg</p> <p>Department of Computer Science, University of Leipzig</p> <p>English Linguistics and Translation Science, Saarland University, Saarbrücken</p> <p>Institute for Natural Language Processing, University of Stuttgart</p>
Government links	<p>CLARIN-SWE knowledge centres have government mandate and are connected to the National Language Bank. CLARIN-IS highlights the need for government recognition of language technologies as a national priority.</p> <p>Countries as examples:</p> <p>Germany: Federal Ministry for Education and Research and States</p> <p>Austria: Austrian National Library</p> <p>Czech Republic: Ministry of Education, Sports and Youth of the Czech Republic</p> <p>United Kingdom: British Library & NACTEM</p>

Key Aspects

Global links	Federation includes a "third party" (Carnegie Mellon University) from the USA and South Africa as an observer.
Differentiator	<p>Specialisation in language resources (as corpora) and tools for processing diverse language data.</p> <p>Linked to existing language institutes and leveraging existing capacity with data and technology.</p> <p>ESFRI landmark. Legal status: ERIC, 2012.</p>
AU comparator	Alveo virtual laboratory, Australian National Corpus, PARADISEC, Centre of Excellence for the Dynamics of Language. AIATSIS Collections.

Data Infrastructure

Digitisation	digitisation.eu is a knowledge centre in the CLARIN network and MoU with LIBER on digital collections (along with DARIAH).
Government services	Partnership on data reuse (WP3) for Europeana-DSI project (2015-2016) € 8 900 000 - Connecting Europe Facility (CEF) funding.

Longitudinal	For linguistic research, additionally as a dimension of cultural and social research.
--------------	---

Research Leads

Trends	Connecting up of data infrastructure and existing areas of preexisting knowledge centres (as the infrastructure) as a challenge (for CLARIN) versus the challenge of building of community and user engagement as the data infrastructure emerges (for DARIAH). Integration of data infrastructures e.g. CLARIN-Europeana 2015-2018 and maturity level (of existing data infrastructure and community practices of data enabled or driven research).
Further analysis	<p>Critical analysis of CLARIN as lacking in user uptake and building data infrastructure without requisite community engagement. 2017 Annual Report for CLARIN records 10 Knowledge centres, 17 Data centres, 18 Metadata centres and members but doesn't count stakeholders or users. Lessons learned with CLARIN virtual language observatory on user engagement and uptake. Government involvement in mandating or supporting the creation advancement of language technologies (national definition) as a cross-cutting effort and system wide benefit. Centre certification process to identify useful pathways for infrastructure maturity, investment, growth and scaling up and cost-benefit for participation in CLARIN network. Move to blend CLARIN/DARIAH in Germany as per Netherlands as a CLARIAH.</p> <p>No official support for Time Machine, however there are project links.</p>

Additional Information

Relevant Info	<p>CLARIN</p> <p>Meril.eu - CLARIN</p> <p>ESFRI 2018 roadmap - CLARIN € 4 000 000</p> <p>H2020 2015-2017 CLARIN-PLUS € 1 495 300</p> <p>H2020 2015-2016 LT_Observatory € 982 563</p> <p>FP7 2013-2020 CLARIN_2013 € 2 664 710,40</p> <p>FP7 2011-2014 INNETH € 681 159</p> <p>FP7 2008-2011 CLARIN € 4 100 000</p> <p>Total funding: ~9.8M Euro = ~15.87M AUD</p> <p>CLARIN 2017 annual report</p> <p>The Shape of CLARIN (preparatory phase) 2010</p> <p>Usage Scenarios and Interoperability Case Studies CLARIN (preparatory phase) in 2009</p> <p>Annual Status Report CLARIN (preparatory phase) 2010</p> <p>Users of CLARIN - who are they? 2015</p> <p>CLARIN Value Proposition 2016</p> <p>Number game - Experience of a European research infrastructure (CLARIN) for the analysis of web traffic</p> <p>CLARIN's Virtual Language Observatory (VLO) under scrutiny -- The VLO taskforce of the CLARIN-D centres</p> <p>Bridging the Europeana and CLARIN infrastructures</p>
---------------	--

	<p>Tour de CLARIN: CLARIN Knowledge Centre for the Languages of Sweden</p> <p>About CLARIN-IS (Iceland)</p> <p>CLARIN Participating Consortia</p> <p>CLARIN governance</p> <p>Transition to CLARIAH infrastructure</p> <p>DASISH integrating infrastructure involves: CESSDA, CLARIN, DARIAH, ESS and SHARE.</p>
Notes	<p>No figures for national or institutional co-investment.</p> <p>No assumptions around 1:1 funding matches (to EU grants).</p> <p>Funding totalled across all projects and sources.</p> <p>CLARIN_2013 overall budget € 6 661 776 and FP7 € 1 2 664 710, 2/3 co-investment.</p> <p>CLARIN overall budget € 5 632 085 and FP7 € 4 100 000, 1/4 co-investment.</p> <p>INNET overall budget € 923 932,80 and FP7 € 681 159, 1/3 co-investment.</p> <p>DASISH overall budget € 8 304 039 and FP7 € 5 991 052, 1/4 co-investment.</p>

References

User community	https://www.clarin.eu/content/participating-consortia
Leadership	https://www.clariah.nl/en/about/international
Institutional centres of competence	https://www.clarin.eu/content/clarin-centres https://info.clarin.dk/en/ https://www.clarin-d.net/en/
Government links	https://www.clarin-d.net/en/2-uncategorised/482-data-transfer http://digital-humanities.at/en/dha/s-partners/oenb-austrian-national-library https://lindat.mff.cuni.cz/en/about-lindat-clarin https://www.clarin.ac.uk/centres http://www.nactem.ac.uk/
Global links	https://www.clarin.eu/news/cmu-talkbank-sign-third-party-agreement-clarin
Differentiator	https://www.clarin.eu/content/services
Comparator	http://alveo.edu.au/ http://www.ausnc.org.au http://www.paradisec.org.au/ https://www.dynamicsoflanguage.edu.au/ https://aiatsis.gov.au/collection
Further analysis	https://www.clarin.eu/sites/default/files/CLARIN-Annual-Report-2017.pdf
Digitisation	https://www.clarin.eu/news/impact-clarin-k-centre-digitisation-impact-ckc https://libereurope.eu/blog/2017/03/23/new-partnerships-clarin-eric-dariah/
Government services	https://www.clarin.eu/content/factsheet-clarin-europeana-dsi
Longitudinal	https://www.sciencedirect.com/science/article/pii/S0024384115002375
Relevant info	https://www.clarin.eu/

	https://portal.meril.eu/meril/view/facilitys/15360 http://roadmap2018.esfri.eu/projects-and-landmarks/browse-the-catalogue/clarin-eric/ https://cordis.europa.eu/project/rcn/197995/factsheet/en https://cordis.europa.eu/project/rcn/194219/factsheet/en https://cordis.europa.eu/project/rcn/109667/factsheet/en https://cordis.europa.eu/project/rcn/100344/reporting/fr https://cordis.europa.eu/project/rcn/87298/factsheet/en https://www.clarin.eu/sites/default/files/CLARIN-Annual-Report-2017.pdf https://office.clarin.eu/pp/D8S-1.2a.pdf https://office.clarin.eu/pp/D5R-2.pdf https://office.clarin.eu/pp/D1M-2.pdf https://www.clarin.eu/blog/users-clarin-who-are-they https://office.clarin.eu/v/CE-2016-0847-CLARINPLUS-D5_4.pdf https://arxiv.org/pdf/1706.05089.pdf https://www.researchgate.net/publication/267447727_CLARIN's_Virtual_Language_Observatory_VLO_under_scrutiny_-_The_VLO_taskforce_of_the_CLARIN-D_centres https://www.clarin.eu/blog/bridging-europeana-and-clarin-infrastructures https://www.clarin.eu/blog/tour-de-clarin-clarin-knowledge-centre-languages-sweden http://clarin.is/en/about/ https://www.clarin.eu/content/participating-consortia https://www.clarin.eu/content/governance https://www.clarin-d.net/en/current-issues/490-successful-start-to-clariah-de https://dasish.eu/about_dasish/ https://cordis.europa.eu/project/rcn/101702/factsheet/en
--	---

1.2 Digital Research Infrastructure for the Arts and Humanities (DARIAH)

“The Digital Research Infrastructure for the Arts and Humanities (DARIAH) aims to enhance and support digitally-enabled research and teaching across the arts and humanities. DARIAH is a network of people, expertise, information, knowledge, content, methods, tools and technologies from its member countries. It develops, maintains and operates an infrastructure in support of ICT-based research practices and sustains researchers in using them to build, analyse and interpret digital resources”,

<https://www.dariah.eu/about/dariah-in-nutshell/>

Administration

Funding sources	National, H2020, Research council, Institutional, FP (EU Framework Programme)
Jurisdiction	Europe
Leader	Coordinating country France (CNRS).
Domains	Humanities, Arts, Information Science & Technology
Start date	2006
Timescale	11-15 Years
Investment	\$11-15M
Phases	Pre-operational (design/plan/construct), Operational (maintain/consolidate/grow)

Structure

Features	Regional impact, Institutional centres of excellence, Government linkages, National nodes, Global nodes
Type	Discovery platform, Capability building, Community building, Information sources, Research practices and methods, Research libraries, Research tools and platforms
Capabilities	Data curation, Data mining and analysis, Domain expertise
Organisation	Distributed, Virtual
Governance	Coordinating entity (hub), Network (participating nodes), Decentralised (institutional), Governing body, Executive body, Expert advisors
User community	100,000+ stakeholders in research, policy and cultural heritage. 182 national institutions. 87 project partners. (DARIAH 2017 Annual report)
Institutional centres	Germany as an example: Partner Institutions: Academy of Sciences and Literature Mainz Berlin-Brandenburg Academy of Sciences and Humanities DAASI International German Archaeological Institute Herzog August Library Wolfenbüttel Karlsruhe Institute of Technology Leibniz Institute of European History Mainz Marbach Weimar Wolfenbüttel Research Association Max Planck Computing and Data Facility Max Planck Institute for the History of Science

	Max Weber Foundation Open Knowledge Foundation Research Centre Jülich - Jülich Supercomputing Centre Society for Scientific Data Processing (GWDG) Steinheim Institute for German-Jewish Studies at University of Duisburg-Essen Technical University of Darmstadt - Digital Humanities Team University of Bamberg (Chair of Media informatics) University of Göttingen - Göttingen State and University Library (SUB) University of Würzburg - Chair of Computer Philology and Modern German Literary History
Government links	Countries as exemplars France: Ministry for Higher Education and Research Germany: Federal Ministry of Education and Research Ireland: Irish Research Council Netherlands: Netherlands Organisation for Scientific Research (NWO)

Key Aspects

Global links	Within the rules of ERIC, it appears "third countries" or "specialised intergovernmental agencies" can become members.
Differentiator	Community building and coordination (ground up first). Linked to national or institutional digital humanities labs or centres. Relationship with LIBER (European Research Libraries) around access to cultural heritage data. ESFRI landmark. Legal status: ERIC, 2014.
AU comparator	At an institutional level there are comparators arising from LIEF investment e.g. AusStage, AustLit, AustLII, DAAO, CoEDL

Data Infrastructure

Digitisation	Collaboration with digitisation.eu collaboration around DaTeCH conference, and MoU with LIBER MoU with LIBER on digital collections DARIAH listed on the digitisation landscape (Impact Centre of Competence) Guidelines on collaboration with the cultural heritage sector (around access to collections and digitisation).
Government services	Pan-European level partnership on data reuse (WP3) for Europeana-DSI project (2015-2016) Euro 8 900 000 - Connecting Europe Facility (CEF) funding (see DARIAH past projects). National partnerships with European Research Libraries (see also CERL) evident through project work e.g. Time Machine (Antwerp)
Longitudinal	DARIAH Digital Methods and Practices Observatory working group (DiMPO)
Linkage	DASISH integrating infrastructure involves: CESSDA, CLARIN, DARIAH, ESS, SHARE.

Research Leads

Trends	<p>Networks of research and cultural heritage is indicated in DARIAH member organisations, less so in cooperating partner countries.</p> <p>Longer time to get DARIAH operational (may indicate a need to build community capability and capacity all in parallel). Look into the DH labs/hubs appearing in universities over the last 5 years as an indicator of capacity. By comparison language institutes established earlier and CLARIN became operational quickly (with potentially more data and technology capability in the research system).</p>
Further analysis	<p>National approach to boosting support for arts and humanities in regional network (member and partner organisations).</p> <p>Analyse the mix of universities, faculties, institutes, national and state GLAMs, are listed as partners under country profiles.</p> <p>Rationale for the support for Time Machine (along with E-RIHS).</p>

Additional Information

Relevant Info	<p>DARIAH</p> <p>DESIR H2020 2017-2019 € 2 717 320 EU\$</p> <p>HaS H2020 2015-2017 € 1 941 576,25 EU\$</p> <p>DARIAH-ERIC ESFRI 2006-2019 € 7 M EU\$ (3.7M + 3.3M in prep and construction phase).</p> <p>PREPARINGDARIAH FP7 2008-2011 € 2 500 000</p> <p>Total funding: ~ 7.1M Euro = ~11.5M AUD</p> <p>Meril.eu - DARIAH</p> <p>ESFRI 2018 roadmap - DARIAH</p> <p>DARIAH governance structure</p>
Notes	<p>No figures for national or institutional co-investment.</p> <p>No assumptions around 1:1 funding matches (to EU grants) with national or institutional contributions of funds or effort.</p> <p>Note co-investment in PREPARINGDARIAH for 2008-2011 of overall budget € 3 695 262,60 vs FP7 funding: € 2 500 000 (roughly 1/3 co-investment).</p> <p>Funding totalled across all projects and sources (into DARIAH).</p>

References

User community	https://www.dariah.eu/network/partners-countries/
Leadership	http://www.eassh.eu/i/dariah-digital-research-infrastructure-arts-and-humanities
Institutional centres of competence	https://www.dariah.eu/network/partners-countries/germany/
Government links	https://www.dariah.eu/network/partners-countries/france/ https://www.dariah.eu/network/partners-countries/germany/ https://www.dariah.eu/network/partners-countries/ireland/ https://www.dariah.eu/network/partners-countries/netherlands/

Global links	https://www.dariah.eu/wp-content/uploads/2019/06/DARIAH-ERIC_Statutes_May2019.pdf
Differentiator	http://darthcrimson.org/portfolio/a-review-of-selected-digital-humanities-centers-and-initiatives-march-2016/ https://www.dariah.eu/2017/03/30/liber-and-dariah-eu-signed-cooperation-agreement-advancing-digital-research-is-a-joint-effort/ http://roadmap2018.esfri.eu/projects-and-landmarks/browse-the-catalogue/dariah-eric/
Comparator	https://www.ausstage.edu.au/pages/browse/ https://www.austlit.edu.au/ https://www.austlii.edu.au/ https://www.daa0.org.au/ https://www.dynamicsoflanguage.edu.au/
Further analysis	https://www.dariah.eu/network/partners-countries/ http://www.e-rihs.eu/time-machine-fet-flagship-proposal/
Digitisation	https://libereurope.eu/blog/2017/03/23/new-partnerships-clarin-eric-dariah/ https://www.digitisation.eu/community/map-of-the-digitisation-landscape/?pag=6&pag=9 https://www.dariah.eu/wp-content/uploads/2019/04/DARIAH_Theme_Guidelines_How_to_Facilitate_Final_26March2019.pdf
Government services	https://www.dariah.eu/activities/projects-and-affiliations/dariah-in-past-eu-projects/
Longitudinal	https://www.dariah.eu/activities/working-groups/wg-digital-methods-and-practices-observatory-dimpo/
Data linkage	https://dasish.eu/about_dasish/
Relevant info	https://www.dariah.eu/ https://cordis.europa.eu/project/rcn/207190/results/en https://cordis.europa.eu/project/rcn/198110/factsheet/en http://roadmap2018.esfri.eu/projects-and-landmarks/browse-the-catalogue/dariah-eric/ https://cordis.europa.eu/project/rcn/88504/factsheet/en https://portal.meril.eu/meril/view/facilitys/15456 http://roadmap2018.esfri.eu/projects-and-landmarks/browse-the-catalogue/dariah-eric/ https://www.dariah.eu/about/organisation-and-governance/

1.3 Compute Canada

“Compute Canada, in partnership with regional organizations [ACENET](#), [Calcul Québec](#), [Compute Ontario](#) and [WestGrid](#), leads the acceleration of research and innovation by deploying state-of-the-art advanced research computing (ARC) systems, storage and software solutions. Together we provide essential ARC services and infrastructure for Canadian researchers and their collaborators in all academic and industrial sectors”, <https://www.computeCanada.ca/about/>

Administration

Funding sources	National, Provided by Canada Foundation for Innovation (CFI)
Jurisdiction	Canada
Domains	Humanities, Social Science
Start date	2016
Timescale	3-5 Years
Investment	\$3-5M
Phases	Pre-operational (design/plan/construct), Operational (maintain/consolidate/grow)

Structure

Features	Institutional centres of excellence, National nodes
Type	Capability building, HPC service, Cloud storage/compute service, Local storage, Research tools and platforms
Capabilities	Project support, Technology expertise, Domain expertise
Organisation	Shared facility
Governance	Centralised (nexus)
User community	Literary studies, textual studies (4 projects)
Institutional centres	University of Guelph University of Victoria Simon Fraser Uni PKP https://inke.ca/projects/members/#partners https://etcl.uvic.ca/inke/
Government links	Partnership with Statistics Canada

Key Aspects

Differentiator	Built into existing national computational services. Project and technology support.
AU comparator	No equivalent service provided by NCI or Pawsey.

Data Infrastructure

Digitisation	Canadiana Heritage program
Longitudinal	Web Archives for Longitudinal Knowledge (WALK) example for HPC allocation

Research Leads

Trends	Mix of community building (INKE), HASS projects needing cloud based research infrastructure versus HPC, and capability building (DHSI) as an indicators of growth
--------	---

	in research uptake. Gap between higher education and government in the enabling of advanced research and the provision of national research infrastructure.
Further analysis	Progress on national HPC research infrastructure with publicly funded research organisations e.g. Statistics Canada. Working relationship between Compute Canada and CARL (Canadian Association of Research Libraries).

Additional Information

Relevant Info	CFI - Canadian Foundation for Innovation Compute Canada support for HASS CFI 2016-2019 report \$CAN 4.3M Digital Research Canada: The National Advanced Research Computing and Data Management Service (2016) Federated Research Data Repository CARL (Canadian Academic and Research Libraries) partnerships CARL Dataverse - Portage Network
Notes	HPC support is focused on digital humanities (no links to microdata or survey instruments etc or national statistics agency). Compute Canada (as a shared HPC facility) commenced in 2006

References

Institutional centres of competence	https://www.computeCanada.ca/research/literary-criticism/ http://www.dhsi.org/ https://pkp.sfu.ca/2018/09/27/coalition-publi-ca-a-national-partnership-providing-sustainable-support-for-journals-transitioning-to-open-access/ https://inke.ca/projects/members/#partners https://etcl.uvic.ca/inke/
Government links	https://www.statcan.gc.ca/eng/about/yir2018
Digitisation	http://heritage.canadiana.ca/?usrlang=en
Longitudinal	https://www.computeCanada.ca/featured/compute-canada-announces-results-of-its-2016-resource-allocation-competitions/
Relevant info	https://www.innovation.ca/ https://www.computeCanada.ca/research-portal/humanities-and-social-sciences/ https://www.calj-acrs.ca/news/43-m-investment-create-canadian-cyberinfrastructure-humanities-and-social-sciences-research https://www.computeCanada.ca/wp-content/uploads/2016/02/Digital-Research-Canada-A-proposal-for-the-delivery-of-advanced-research-computing-in-Canada.pdf https://www.frdr.ca/repo/ http://www.carl-abrc.ca/about-carl/partners-relations/ https://portagenetwork.ca/network-of-experts/dataverse-north-working-group/
Notes	https://www.westgrid.ca/about_westgrid/history

1.4 Consortium of European Social Science Data Archives (CESSDA)

“CESSDA provides large-scale, integrated and sustainable data services to the social sciences. It brings together social science data archives across Europe, with the aim of promoting the results of social science research and supporting national and international research and cooperation”, <https://www.cessda.eu/About>

Administration

Funding sources	National, H2020, Institutional, FP (EU Framework Programme), National research and statistics agencies
Jurisdiction	Europe
Leader	Coordinating country Norway via the Norwegian Centre for Research Data
Domains	Humanities, Social Science
Start date	2006
Timescale	11-15 Years
Investment	\$101-\$150M
Phases	Pre-operational (design/plan/construct), Operational (maintain/consolidate/grow)

Structure

Features	Regional impact, Institutional centres of excellence, Government linkages, National nodes, Global nodes
Type	Archive or repository service, Data service, Discovery platform, Capability building, Community building, Research practices and methods
Capabilities	Data curation, Data preservation, Data reuse, Technology expertise, Domain expertise, Support for setting up national data services.
Organisation	Distributed, Virtual
Governance	Coordinating entity (hub), Network (participating nodes), Decentralised (institutional), Governing body, Executive body, Operational function, Outreach and liaison function, Administration function, Technologies function, Informatics function, Policy function, Expert advisors
User community	CESSDA identify their main stakeholders as: members, service providers, data producers, data users. (CESSDA 2018-2022 strategy).
Institutional centres	CESSDA has 17 members and 19 partners in the consortium. Austria as an example: AUSSDA is funded by Federal Ministry of Education, Science and Research, AUSSDA nodes located at Vienna, Graz, Linz and universities. France as an example: PROGEDO is a joint initiative supported by the Ministry of National Education, Higher Education and Research and National centre for scientific research (CNRS).
Government links	Data without boundaries project (29 partners) emphasises the crossover between national statistics institutes or departments (10 partners), data archives (14 partners), higher education researchers (7 partners) and industry (1 partner).

Key Aspects

Global links	CESSDA 2018-2022 strategy indicates a push to global partnerships and to find "third parties" e.g. with ICSPR (USA) and Research Data Alliance. Emphasis on the breadth and depth (50 years) of experience.
Differentiator	Constellation of government, higher education and industry interests in longitudinal and comparable demographic data, and data linkage (and adherence to standards to support interoperability). ESFRI landmark. Legal status: ERIC, 2017.
AU comparator	Australian Data Archive (an institutional service, that delivers national services)

Data Infrastructure

Digitisation	Historical census data (retrieval from analogue or legacy data formats) and transcription (reference points UK Data Service).
Government services	Pan-European level data discovery via European University Institute Research Library. Collaboration around discovery of social science data with government. See Denmark as an example of the cross-ministry support (culture and higher education and science) to make social science data discoverable. See Sweden as an example of cross institutional support to make social science data discoverable.
Longitudinal	CESSDA 2018 Annual Report - a focus on cross-national, longitudinal data harmonisation, and data linkage.
Linkage	CESSDA 2018 Annual Report

Research Leads

Trends	Early and tight coupling with national statistics bodies and university data archives. Comprehensive coverage of research data and tool requirements over time (useful to see whether data infrastructure and quality operates as a foundation to build in as a corollary outreach, tools and standards upon) parallels with CLARIN.
Further analysis	Cross over of CESSDA with other infrastructure projects e.g. SSHOC & SERISS Cross over of CESSDA with other SS related infrastructures e.g. ESS and SHARE Collaboration within ERICs across domain boundaries e.g. Social Sciences and Humanities cluster. Bundling of CESSDA, ESS, SHARE together as PROGEDO at national level. Cross-cutting program SERISS (ESS, SHARE, CESSDA, GGP, EVS) to advance interoperability, comparative research and support for big data research.

Additional Information

Relevant Info	MERIL - CESSDA ESFRI 2018 Roadmap - CESSDA € 84.7M EU\$ H2020 (2015-2017) CESSDA SaW € 2 498 187 FP7 Infrastructures (2012-2014) SERSCIDA project € 625 573 FP7 Infrastructures (2011-2015) DwB project € 6 493 017 Total grant funding = ~ 84.7M Euro = ~ 137M AUD CESSDA past project funding
---------------	---

	CESSDA governance CESSDA 2018 Annual Report
Notes	<p>No figures for national or institutional co-investment.</p> <p>No assumptions around 1:1 funding matches (to EU grants) with national or institutional contributions of funds or effort.</p> <p>Notable difference in co-investment in DwB: overall budget: € 8 686 425 vs FP7 funds: € 6 493 017 a €2M co-investment.</p> <p>Funding totalled across all projects and sources (into CESSDA).</p>

References

User community	https://www.cessda.eu/About/Consortium/CESSDA-Countries/CESSDA-Members
Leadership	https://www.cessda.eu/About/Consortium/CESSDA-Countries/CESSDA-Members/Norway
Institutional centres of competence	https://www.cessda.eu/About/Consortium https://www.cessda.eu/About/Consortium/CESSDA-Countries/CESSDA-Members/Austria https://www.cessda.eu/About/Consortium/CESSDA-Countries/CESSDA-Members/France
Government links	https://www.cessda.eu/About/Projects/Past-projects/DwB
Global links	https://www.cessda.eu/content/download/4260/48215/file/CESSDA%20Strategy%202018-2022.pdf
Differentiator	https://www.cessda.eu/Tools-Services
Comparator	https://ada.edu.au/
Further analysis	https://www.cessda.eu/About/Projects https://seriss.eu/
Digitisation	http://blog.ukdataservice.ac.uk/1961-census-digitisation/ http://blog.ukdataservice.ac.uk/transcribing-for-the-future-clarin-transcription-chain-tools-for-turning-recorded-human-speech-into-textual-representation/ https://www.ukdataservice.ac.uk/about-us/our-rd/historic-census-microdata.aspx
Government services	https://www.eui.eu/Research/Library/ResearchGuides/Economics/Statistics/DataPortal https://www.sa.dk/en/ https://www.vr.se/english/analysis-and-assignments/research-infrastructure/list-of-research-infrastructure/research-infrastructure/2018-10-18-swedpop---swedish-population-databases-for-research.html http://swedpop.se/
Longitudinal	https://www.cessda.eu/content/download/4731/53242/file/CESSDA_Annual_Report_2018.pdf
Data linkage	https://www.cessda.eu/content/download/4731/53242/file/CESSDA_Annual_Report_2018.pdf
Relevant info	https://portal.meril.eu/meril/view/facilitys/15577

	http://roadmap2018.esfri.eu/projects-and-landmarks/browse-the-catalogue/cessda-eric/ https://cordis.europa.eu/project/rcn/198258/factsheet/en http://www.serscida.eu/en/ https://cordis.europa.eu/project/rcn/100560/reporting/en http://www.dwbproject.org/ https://cordis.europa.eu/project/rcn/100056/reporting/en https://www.cessda.eu/About/Projects/Past-projects https://www.cessda.eu/About/Governance https://www.cessda.eu/content/download/4731/53242/file/CESSDA_Annual_Report_2018.pdf
--	---

1.5 European Social Survey (ESS)

“The European Social Survey (ESS) is an academically driven cross-national survey that has been conducted across Europe since its establishment in 2001. Every two years, face-to-face interviews are conducted with newly selected, cross-sectional samples. The survey measures the attitudes, beliefs and behaviour patterns of diverse populations in more than thirty nations”, <http://www.europeansocialsurvey.org/about/>

Administration

Funding sources	National, H2020, Institutional, FP (EU Framework Programme)
Jurisdiction	Europe
Leader	ESS ERIC headquarters is in the UK
Domains	Social Science
Start date	2002
Timescale	16-20 Years
Investment	\$16-20M
Phases	Pre-operational (design/plan/construct), Operational (maintain/consolidate/grow)

Structure

Features	Regional impact, Institutional centres of excellence, Government linkages, National nodes, Global nodes
Type	Archive or repository service, Data service, Discovery platform, Capability building, Community building, Information sources, Research practices and methods, Research tools and platforms
Capabilities	Data curation, Data reuse, Informatics and data modelling, Survey design and instrument, Technology expertise, Domain expertise
Organisation	Distributed, Virtual
Governance	Coordinating entity (hub), Network (participating nodes), Decentralised (institutional), Governing body, Executive body
User community	ESS usage statistics for 2019 by registered users https://www.europeansocialsurvey.org/about/user_statistics.html
Institutional centres	30 countries participate in ESS (27 have joined SHARE). 8 partner institutions: City University of London, UK Norwegian Social Science Data Services (NSD), Norway GESIS, Germany The Netherlands Institute for Social Research (SCP), Netherlands Universitat Pompeu Fabra, Spain University of Essex, UK University of Leuven, Belgium University of Ljubljana, Slovenia Norway as example: Norwegian Centre for Research Data operates from within the Ministry of

	Education Research
Government links	<p>National representatives in general assembly include key government agencies</p> <p>Partner countries as examples:</p> <p>Estonia: Analysis and Statistics Department, Ministry of Social Affairs</p> <p>Iceland: University of Iceland</p> <p>Poland: Ministry of Science and Higher Education</p> <p>Sweden: Swedish Research Council</p>

Key Aspects

Global links	<p>Representative from ISR Michigan on the methods advisory board.</p> <p>Global links workshop in 2017, included representatives/discussions on research links with and surveys in Australia, Asia, South Africa, USA, Latin America.</p>
Differentiator	<p>Targeted at survey methodology and interoperability of data infrastructure to support comparative analysis (closer to the research process).</p> <p>ESFRI landmark. Legal status: ERIC, 2013.</p>
AU comparator	Longitudinal datasets e.g. HILDA, LSAC, LSIC, BNLA (via the Department of Social Services)

Data Infrastructure

Government services	At a national level through the participating countries. ESS operates as a pan-national data discovery service.
Longitudinal	The surveys are biennial, 2002 onward, as rounds and include cumulative statistics. See cumulative variables listing across rounds and participating countries.
Linkage	<p>See SERISS project for efforts on data linkage, for ESS, CESSDA and SHARE. Data linkage operating at national level through participants.</p> <p>Switzerland as an example:</p> <p>FORS (Swiss Foundation for Research in Social Sciences) and its relationship to linkhub.ch</p>

Research Leads

Trends	Country capacity to fund social surveys and ability to participate in coordinated pan-national survey initiatives (consistent participation may correlate with strength of the economy to support research).
Further analysis	<p>The impact of different funding levels across time as the mode appears to be €1.5M, but ramped up in design/prep phases (€4.4M & €6.4M). This may be an investment trend across ESFRI projects (front loaded during these phases).</p> <p>Annual report 2017-2018 includes contribution to central costs by country e.g. France €278,671, UK €764,909, Netherlands €101,340. Useful to understand why there are proportional differences in co-contribution factored into the calculations by the General Assembly.</p> <p>Cross-cutting program SERISS (ESS, SHARE, CESSDA, GGP, EVS) to advance interoperability, comparative research and support for big data research.</p>

Additional Information

Relevant Info	<p>MERIL - ESS</p> <p>ESFRI 2018 Roadmap - ESS covers € 6.4M + € 2.3M for setup</p> <p>FP4 TSER (1994-1998)</p> <p>FP5 HUMAN-POTENTIAL ESS2 project (2003-2005) € 1 780 000</p> <p>FP6 SOCIETY ESS3 project (2005-2007) € 1 444 799</p> <p>FP6 CITIZENS ESS4 project (2007-2010) € 1 702 854</p> <p>FP7 ESS-DACE project (2010-2014) € 4 999 987,17</p> <p>Total funding = ~9.8 M Euro = 15.9 M AUD</p> <p>ESS 2017-2018 Annual Report</p>
Notes	<p>MERIL start date operationally is 2002 (4 years prior to entry in 2006 into ESFRI roadmap), has additional phase (interim/transition) in development.</p> <p>Notable 25% co-contribution to ESS2 (overall budget: € 2 348 353 compared to FP5 investment € 1 780 000)</p> <p>Notable 25% co-contribution to ESS-DACE (overall budget: € 6 724 936,66 compared to FP7 investment € 4 999 987,17).</p> <p>Investment in socio-economic research early in FP program. Total funding count covers ESS from 2003 onward.</p>

References

User community	https://www.europeansocialsurvey.org/about/participating_countries.html
Leadership	https://www.europeansocialsurvey.org/about/structure_and_governance.html
Institutional centres of competence	https://www.europeansocialsurvey.org/about/ https://www.europeansocialsurvey.org/about/participating_countries.html https://nsd.no/nsd/english/index.html
Government links	https://www.europeansocialsurvey.org/about/structure_and_governance.html
Global links	http://www.europeansocialsurvey.org/about/news/essnews0034.html
Differentiator	https://www.europeansocialsurvey.org/methodology/ https://www.europeansocialsurvey.org/data/
Comparator	https://www.dss.gov.au/national-centre-for-longitudinal-data-nclid/access-to-dss-longitudinal-datasets
Further analysis	https://seriss.eu/
Longitudinal	https://www.europeansocialsurvey.org/docs/cumulative/ESS_cumulative_variable_list.pdf
Data linkage	https://forscenter.ch/data-services/data-collection/ https://forscenter.ch/projects/projects-list/?showcontent=12265#projects_12265
Relevant info	https://portal.meril.eu/meril/view/facilitys/15200 http://www.roadmap2018.esfri.eu/projects-and-landmarks/browse-the-catalogue/ess-eric/ covers https://cordis.europa.eu/project/rcn/47522/factsheet/en

	https://cordis.europa.eu/programme/rcn/465/en https://cordis.europa.eu/project/rcn/67573/factsheet/en https://cordis.europa.eu/project/rcn/89088/factsheet/en https://cordis.europa.eu/project/rcn/79914/reporting/en https://cordis.europa.eu/project/rcn/97288/reporting/en https://www.europeansocialsurvey.org/docs/about/annualreports/ESS_ERIC_annual_activity_report_2017-2018.pdf
--	--

1.6 Survey of Health, Ageing and Retirement in Europe (SHARE)

“The Survey of Health, Ageing and Retirement in Europe (SHARE) is a multidisciplinary and cross-national panel database of micro data on health, socio-economic status and social and family networks of about 140,000 individuals aged 50 or older (around 380,000 interviews). SHARE covers 27 European countries and Israel”, <http://www.share-project.org/home0.html>

Administration

Funding sources	National, H2020, Institutional, FP (EU Framework Programme)
Jurisdiction	Europe
Leader	Coordinating country is Germany via the Munich Center for the Economics of Aging (MEA), Max Plank Institute for Social Law and Social Policy.
Domains	Social Science, Health
Start date	2004
Timescale	11-15 Years
Investment	\$21-50M
Phases	Pre-operational (design/plan/construct), Operational (maintain/consolidate/grow)

Structure

Features	Regional impact, Institutional centres of excellence, Government linkages, National nodes, Global nodes
Type	Archive or repository service, Data service, Discovery platform, Capability building, Community building, Information sources, Research practices and methods, Research tools and platforms
Capabilities	Data curation, Data reuse, Informatics and data modelling, Survey design and instrument, Data linkage, Technology expertise, Domain expertise
Organisation	Distributed, Virtual
Governance	Coordinating entity (hub), Network (participating nodes), Decentralised (institutional), Governing body, Executive body, Operational function, Outreach and liaison function, Administration function, Technologies function, Informatics function, Policy function, Expert advisors
User community	9000 registered users showing global usage (see 2018-2019 Annual Report).
Institutional centres	SHARE has 28 country teams Switzerland for example: Supported by FORS (gov), UNIL (higher education) and FNS (national science foundation).
Government links	SHARE ERIC Council has government representatives Participating countries and delegates as examples: Austria: Federal Ministry for Science and Research (BMWF) Hungary: National Research, Development and Innovation Office Israel: Ministry for Social Equality & National Insurance Institute Switzerland: FORS

Key Aspects

Global links	Global links to Asia, USA, Latin America. Partnership with WHO; and inclusion of Israel in data collection.
Differentiator	Pan-national cooperation around survey methods and data collection of ageing populations in Europe. ESFRI landmark. Legal status: ERIC, 2011.

Data Infrastructure

Digitisation	--
Government services	At a national level through the participating countries. SHARE operates as a pan-national data discovery service.
Longitudinal	Survey waves from 2004 annually.
Linkage	No linkage services, but linkages to administrative data included in waves 3-7. Provision of survey data with linkages to administrative data supplied by partners, see Germany's SHARE-RV project.

Research Leads

Trends	Notable growth in partners participating in surveys from 2011 onward and a jump in 2017 (see waves overview).
Further analysis	Cross-cutting infrastructure - SERISS (and alignment with work undertaken in ESS and CESSDA) and DASISH (another cross-cutting infrastructure). Cross-cutting program SERISS (ESS, SHARE, CESSDA, GGP, EVS) to advance interoperability, comparative research and support for big data research.

Additional Information

Relevant Info	SHARE MERIL - SHARE ESFRI 2018 Roadmap - SHARE FP5 LIFE QUALITY 2002-2004 € 2 758 630 overall budget € 3 085 269 FP5 LIFE QUALITY 2003-2006 € 2 999 612 overall budget € 3 232 591 FP6 CITIZENS 2006-2008 € 1 489 999 overall budget FP7 SHARE-PREP 2008-2009 € 2 499 842,08 overall budget € 3 743 601,60 FP7 SHARE_LEAP 2009-2010 € 2 999 999,27 overall budget € 4 141 233,58 FP7 HEALTH 2009-2011 € 1 751 791 overall budget € 1 964 885,60 FP7 SHARE_M4 2011-2014 € 5 499 991,81 overall budget € 7 663 671,88 Total funding = ~20M Euro = 32.4M AUD SHARE Dates and Facts SHARE 2018-2019 Annual Report covers registered users and publication counts
Notes	Full accounting of funding from EU and national co-investment.

References

User community	http://www.share-project.org/organisation/share-country-teams.html
Leadership	http://www.share-project.org/organisation/coordination.html

Institutional centres of competence	http://www.share-project.org/organisation/share-country-teams.html https://www.unil.ch/share/home/menuinst/share.html
Government links	http://www.share-project.org/organisation/share-eric/eric-council.html
Global links	http://www.share-project.org/organisation/international-cooperations.html?L=0 http://igdc.huji.ac.il/home/share/introduction.aspx
Differentiator	http://www.share-project.org/data-documentation.html
Further analysis	https://seriss.eu/
Trends	http://www.share-project.org/data-documentation/waves-overview.html
Government services	http://www.share-project.org/data-documentation/data-documentation-tool.html
Longitudinal	http://www.share-project.org/data-documentation/waves-overview.html
Data linkage	http://www.share-project.org/special-data-sets/record-linkage-project/share-rv.html
Relevant info	http://www.share-project.org/home0.html https://portal.meril.eu/meril/view/facilitys/15274 http://www.roadmap2018.esfri.eu/projects-and-landmarks/browse-the-catalogue/share-eric/ https://cordis.europa.eu/project/rcn/60432/factsheet/en https://cordis.europa.eu/project/rcn/67140/factsheet/en https://cordis.europa.eu/project/rcn/78667/factsheet/en https://cordis.europa.eu/project/rcn/87494/factsheet/en https://cordis.europa.eu/project/rcn/89479/factsheet/en https://cordis.europa.eu/project/rcn/92038/brief/en https://cordis.europa.eu/project/rcn/97261/reporting/en http://www.share-project.org/organisation/dates-facts.html http://www.share-project.org/fileadmin/pdf_documentation/SHARE_AnnualActivityReport_2018-19.pdf
Notes	http://www.share-project.org/organisation/funding.html

1.7 IMPACT Centre of Competence

“The Impact Centre of Competence in Digitisation is a not for profit organisation, comprised of public and private institutions, with the mission to make the digitisation of historical printed text “better, faster, cheaper”. It provides tools, services and facilities to further advance the state-of-the-art in the field of document imaging, language technology and the processing of historical text”, <https://www.digitisation.eu/>

Administration

Funding sources	National, Institutional, FP (EU Framework Programme)
Jurisdiction	Europe
Leader	Coordinating institution is Fundaci3n Biblioteca Virtual Miguel de Cervantes (Spain) - also a CLARIN K-Centre in Digitisation
Domains	Information Science & Technology, Text Digitisation & Language Technologies
Start date	2008
Timescale	11-15 Years
Investment	\$21-50M
Phases	Pre-operational (design/plan/construct), Operational (maintain/consolidate/grow)

Structure

Features	Cross-cutting, Regional impact, Institutional centres of excellence, Government linkages, National nodes, Global nodes
Type	Discovery platform, Capability building, Community building, Software service, Information sources, Digitisation service
Capabilities	Data curation, Data preservation, Data reuse, Data mining and analysis, Informatics and data modelling, Project support, Technology expertise, Domain expertise
Organisation	Virtual, Shared facility
Governance	Coordinating entity (hub), Network (participating nodes), Centralised (nexus), Governing body, Expert advisors
User community	Institutional membership by heritage collection holders.
Institutional centres	Ostensibly the premium members are those seeking to advance capacity and to "implement and further development the IMPACT, become a member of the IMPACT Board, obtain access to the full range of IMPACT facilities". Premium (8), standard (24) and affiliate members (7).
Government links	National libraries (see Netherlands and Finland) and state libraries (see Berlin) are members.

Key Aspects

Global links	Institutions from Mexico (premium) e.g. Universidad de Guadalajara (Mexico) and USA (affiliate) are members e.g. California Digital Library.
Differentiator	A node of CLARIN (as a knowledge centre) also serves as a centre of competence for the heritage community to become a member of and build up a network of members and partners. Notably both heritage and higher education institutions are members of IMPACT to gain the benefits of a centre of competence.

AU comparator	Major cultural institutions (e.g. National Library of Australia, State Libraries), science and research institutions (e.g. CSIRO, University of Sydney and Melbourne) have digitisation facilities. None specialise in language technologies or provide corpora or lexicon.
---------------	---

Data Infrastructure

Digitisation	Focus of the centre of competence is on digitisation of textual heritage and use of language technologies.
Longitudinal	Historical archives provide access to longitudinal data e.g. apprentice records in Venetian archives, or court records in Old Bailey.

Research Leads

Trends	<p>Specialisation of digitisation technologies based on heritage object types and contexts, i.e. IMPACT (text) versus 3D (objects) and heritage institution type (e.g. gallery versus science museum).</p> <p>Alignment of investments in digitisation to meet the needs of European cultural and society and to meet the needs of research (access to heritage data), and also the capacity to leverage collaborative networks and consolidate investments (for both sectors and institutional members).</p>
Further analysis	<p>The alignment of resources, expertise and technology (around digitisation of text and language based heritage resources) and the link to research, and expertise around text and language based heritage resources (the alignment of the investments in IMPACT - for heritage and in CLARIN - for research).</p> <p>Connection between Europeana (as a pan-national broker) and other leading heritage institutions in orientation to research for expertise and to delivery heritage data (for research reuse).</p>

Additional Information

Relevant Info	<p>IMPACT Centre of Competence</p> <p>FP7 IMPACT 2008-2012 € 12 163 911</p> <p>FP7 SUCCEED 2013-2014 € 1 793 809</p> <p>SUCCEED program fed into IMPACT Centre of Competence</p> <p>Crossover with CLARIN - operates as a certified knowledge centre for the CLARIN community (IMPACT-CKC)</p>
Notes	<p>Emergence of centres of competence began in FP6 and FP7 funding program e.g. SUCCEED and IMPACT.</p> <p>Unclear about overlap with ESFRI-RIHS entering the roadmap at 2016 with ~ €49M 2017-2025 (arose from IPERION-CH, CHARISMA, EU-ARTECH, ARIADNE)</p> <p>Unclear about overlap with PARTHENOS (that links to DARIAH and CLARIN)</p> <p>Unclear about overlap with SYNTHESIS & SYNTHESIS3 in FP program (that leads to DiSSCo (Distributed System of Scientific Collections))</p> <p>Unclear about roles of possible precursor projects PLANETS FP6, MINERVA FP5 & MINERVAPLUS FP6</p> <p>ESFRI initiative - SYNTHESIS</p>

References

User community	https://www.digitisation.eu/about/members/
Leadership	https://www.digitisation.eu/helpdesk/
Institutional centres of competence	https://www.digitisation.eu/about/members/
Government links	https://www.digitisation.eu/about/members/koninklijke-bibliotheek/ https://www.digitisation.eu/about/members/national-library-of-finland/
Global links	https://www.digitisation.eu/about/members/
Relevant info	https://www.digitisation.eu/ https://cordis.europa.eu/project/rcn/85383/factsheet/en https://cordis.europa.eu/project/rcn/106343/factsheet/en http://succeed-project.eu/?q=node/1 https://www.digitisation.eu/helpdesk/
Notes	" http://www.succeed-project.eu/sites/default/files/deliverables/Succeed_600555_WP7_D7.6_BestPracticesAndManagementProceduresForCentresOfCompetence_v1.0_D.pdf http://dissco.eu http://www.roadmap2018.esfri.eu/projects-and-landmarks/browse-the-catalogue/e-rihs/ http://roadmap2018.esfri.eu/projects-and-landmarks/browse-the-catalogue/dissco/ http://jpi-ch.eu/wp-content/uploads/IPERION-CH_presentation.pdf http://www.parthenos-project.eu/ https://cordis.europa.eu/project/rcn/194932/factsheet/en https://cordis.europa.eu/project/rcn/99184/factsheet/en (2006-2010) https://cordis.europa.eu/project/rcn/61819/factsheet/en (2002-2005) https://cordis.europa.eu/project/rcn/71391/factsheet/en (2004-2006) https://cordis.europa.eu/project/rcn/108204/factsheet/en https://cordis.europa.eu/docs/report/docs-results-312-312253-final1-synthesys3-final-publishable-summary.pdf "

1.8 European Research Infrastructure for Heritage Science (E-RIHS)

"E-RIHS is the European Research Infrastructure for Heritage Science that supports research on heritage interpretation, preservation, documentation and management. E-RIHS mission is to deliver integrated access to expertise, data and technologies through a standardized approach, and to integrate world-leading European facilities into an organisation with a clear identity and a strong cohesive role within the global heritage science community", <http://www.e-rihs.eu/>

Administration

Funding sources	National, H2020
Jurisdiction	Europe
Leader	Coordinating country is Italy.
Domains	Humanities, Information Science & Technology, Heritage science
Start date	2016
Timescale	3-5 Years
Investment	\$6-10M
Phases	Pre-operational (design/plan/construct)

Structure

Features	Regional impact, Institutional centres of excellence, Government linkages, National nodes, Global nodes
Type	Discovery platform, Capability building, Community building, Information sources, Research practices and methods, Digitisation service, Research tools and platforms
Capabilities	Data curation, Data preservation, Data reuse, Technology expertise, Domain expertise, Material science
Organisation	Distributed, Virtual, Shared facility, Institutional facility
Governance	Coordinating entity (hub), Network (participating nodes), Decentralised (institutional), Expert advisors, Governance and management TBD in preparation phase.
User community	Trans-disciplinary and interdisciplinary
Institutional centres	National hubs in 11 nations (hubs). UK as an example: University College, London is the national coordinator, and 13 institutional partners (cultural heritage, science facilities, universities) have joined.
Government links	National hubs join the E-RIHS. Spain as an example, in 2015:"an expression of support from the Ministry of Economy and Competitiveness (MINECO) to the E-RIHS proposal signed by the General Secretary of Science, Technology and Innovation. In the same, the support to the proposal is expressed and it is indicated that "the national node will be coordinated by CSIC and will count, among others, with the Institute of the Cultural Patrimony of Spain of the Ministry of Education, Culture and Sport (IPCE-MECD) and the National Research Center on Human Evolution (CENIEH)""

Key Aspects

Global links	Collaborations with institutions from Israel, USA (2) and Latin America (3). Presented as a European infrastructure (E-RIHS) seeking to be a global infrastructure (G-RIHS) and worldwide membership. Proposal for global infrastructure with partners from EU (18) and world (10) includes Australia.
Differentiator	Integration of access to expertise, data and technologies to support heritage science across all types of material culture. Provision of four platforms to support heritage science across all areas of heritage collecting and advanced technologies: ARCHLAB (physical collections), DIGILAB (digital tools and FAIR data), FIXLAB (material science tools), MOLAB (mobile digital tools).
AU comparator	NCRIS Synchrotron (Melbourne) and National Imaging Facility (Brisbane). ARC funded scanning equipment: AustLii and Analysis and Policy Observatory). Facilities in major cultural heritage institutions (e.g. National Library digitisation) and equipment in universities (e.g. ANU CT Lab and University of Melbourne Digitisation Centre) and scientific organisations with heritage collections (e.g. CSIRO and Geoscience Australia).

Data Infrastructure

Digitisation	EU declaration in May 2019 on digitisation as a priority area for cooperation.
Longitudinal	Includes material and natural collection items will cover ~1000 years of documentary record, archaeological record of occupation ~16,500 years, and flora, faunal, and geoscience records ~4 billion years of earth history.

Research Leads

Trends	A need to address data infrastructure and the enabling language and image processing technologies and techniques (standardisation) across the range of disciplines in humanities and arts (and the range of heritage collection types).
Further analysis	Fragmentation in infrastructure investment has been identified as an area to address in humanities, arts and heritage. Two H2020 programs: PARTHENOS (cross-cutting to improve data practices across programs); and E-RIHS-PP (ESFRI), reference the need for stronger coordination. Both infrastructures appear to combine training, virtual environments, standardisation, community building; leveraging of a range of data, technologies, and techniques. PARTHENOS appears to focus on FAIR data and standardisation of research workflows for humanities research, as an improvement exercise across infrastructures, whereas E-RIHS is focused on all disciplinary dimensions of heritage sciences and global leadership. Rationale for the support for Time Machine (along with DARIAH).

Additional Information

Relevant Info	E-RIHS H2020 E-RIHS PP (2017-2020) € 3 999 449 preparation phase ESFRI 2018 Roadmap, entry 2016 2017-2025 € 49M Unclear about the relationship between PARTHENOS (2015-2019) as a cross-cutting exercise between DARIAH and CLARIN and E-RIHS Progression on from IPERION-CH program (to E-RIHS)
---------------	--

Notes	<p>Ditta Zizi (Dept Ed) represented Australia in GRI meetings, refers to Australia's underground labs. E-RIHS is one of the candidates for GRI. See: Group of Senior Officials on Global Research Infrastructures, Progress Report 2017.</p> <p>Key focus is interdisciplinary research enabled through heritage science</p> <p>Potted history of EU heritage funding</p>
-------	---

References

User community	http://www.e-rihs.eu/about/e-rihs-national-hubs/
Leadership	http://www.e-rihs.eu/partners/
Institutional centres of competence	http://www.e-rihs.eu/about/e-rihs-national-hubs/ http://www.e-rihs.eu/partners/ucl-uk/ https://e-rihs.ac.uk/index.php/news/
Government links	http://e-rihs.es/inicio/
Global links	http://www.e-rihs.eu/international-collaborations/
Differentiator	http://www.e-rihs.eu/about/about/ https://www.ntu.ac.uk/about-us/events/events/2019/03/european-research-infrastructure-for-heritage-science-e-rihs-dissemination-day
Comparator	http://archive.synchrotron.org.au/ https://anif.org.au/ https://www.nla.gov.au/digitisation https://www.csiro.au/en/Research/Collections/Digitisation https://library.unimelb.edu.au/digitisation https://ctlab.anu.edu.au/
Further analysis	http://www.e-rihs.eu/time-machine-fet-flagship-proposal/
Digitisation	http://www.e-rihs.eu/wp-content/uploads/2019/05/190503_declaration_on_european_heritage_clean_cle424a38.pdf
Relevant info	http://www.e-rihs.eu/ https://cordis.europa.eu/project/rcn/209507/factsheet/en http://www.roadmap2018.esfri.eu/projects-and-landmarks/browse-the-catalogue/e-rihs/ https://cordis.europa.eu/project/rcn/194932/factsheet/en http://www.parthenos-project.eu/
Notes	<p>"http://www.e-rihs.eu/wp-content/uploads/2018/11/gso_progress_report_2017.pdf</p> <p>https://www.slideshare.net/E-RIHS/gilberto-corbellini-cnr</p> <p>https://e-rihs.ac.uk/index.php/milestones/"</p>

1.9 Europeana

Europeana is Europe's digital platform for cultural heritage. It is one of the European Commission's [Digital Service Infrastructures \(DSI\)](#). "As a DSI, Europeana's objectives are to innovate the aggregation infrastructure, boost the distribution infrastructure and work towards long-term financial stability through business model innovation. All of this helps make sure that Europe's businesses and people reap the full benefits of the technological revolution in digital services in culture." <https://pro.europeana.eu/our-mission/history>

Administration

Funding sources	National, Institutional, CIP (EU Competitiveness and Innovation Framework Programme) and CEF (Connecting Europe Facility)
Jurisdiction	Europe
Leader	Coordinating country is the Netherlands. Legal status: Foundation.
Domains	Heritage
Start date	2008
Timescale	11-15 Years
Investment	\$101-\$150M
Phases	Pre-operational (design/plan/construct), Operational (maintain/consolidate/grow)

Structure

Features	Regional impact
Type	Data service, Discovery platform, Capability building, Community building, Software service, Information sources, Research tools and platforms
Capabilities	Data curation, Data reuse, Data mining and analysis, Informatics and data modelling, Data linkage, Project support, Technology expertise, Domain expertise
Organisation	Distributed, Virtual, Shared facility
Governance	Coordinating entity (hub), Network (participating nodes), Centralised (nexus), Governing body, Executive body, Operational function, Outreach and liaison function, Administration function, Technologies function, Informatics function, Policy function, Expert advisors
User community	Vast usage levels unclear about research reuse. Direct link made through Europeana-DSI collaboration with DARIAH.
Institutional centres	Lead agencies on multiple Europeana CIP grants indicate that key areas of collection expertise, technical leadership provided by the Netherlands with national libraries in the Netherlands, France, Finland, and Germany. Key alignment with the advancement of The European Library. Aggregation network (accredited) of hubs: 11 thematic, 20 national and regional.
Government links	All data contributors to Europeana operate in national contexts and manage cultural heritage collections and provide access to physical and digital collections for research.

Key Aspects

Global links	Partnerships with national aggregator services in USA (Digital Public Library of America), Australia (Trove Australia), NZ (DigitalNZ) on shared issues with
--------------	--

	standardisation and rights for data reuse.
Differentiator	World leading in the advancement of cultural heritage interoperability. Aggregates using standards from diverse heritage organisations (and approaches to informatics) and uses linked open data practices to support data enhancement and richer semantic discovery. Community engagement through partnerships with research infrastructure programs e.g. DARIAH and Time Machine.
AU comparator	Trove Australia (provided by the National Library of Australia, with a contributor holdings from 1000 libraries and 300 organisations with heritage collections).

Data Infrastructure

Digitisation	Investment around material culture types and themes suggest that funds for digitisation have been provided as a EU contribution and a co-investment. European Commission report on the state of digitisation, online accessibility and digital preservation highlights: advantages of EU funding, establishing economies of scale, and standards and quality challenges.
Government services	Europeana is an aggregator service, and heavily supported by the European Library, a hub of 48 national and research libraries in Europe.
Longitudinal	Descriptive metadata and digital content will cover heritage collections in the documentary history timeline ~1000 years, archaeological timeline ~16,500 years and flora, faunal, and geoscience records ~4 billion years of earth history.
Linkage	Data enhancement (using linked open data) is one of the value add services that Europeana provides as well as vocabularies (for data integration).

Research Leads

Trends	Co-investment compared to EU contribution has three modes: none, 1/4 or 1/2 of the funds co-contributed to the budget. This may be an indicator of investment for national outcomes where there is 1/4 or 1/2 of the budget is co-investment. Equally where the co-investment is negligible this may indicate cross-cutting interests with a pan-national outcome. Examine trend in FP funding for humanities and arts against the funding for Europeana to ascertain any crossover (to see if there is an assumption about the availability of digital material (supply) for research reuse as a result. Note Natural Europe and the emergence of DiSSCo.
Further analysis	The Europeana Awareness programme and Europeana outreach to the research community.

Additional Information

Relevant Info	CIP Europeana Libraries 2011-2012 € 3 094 765, overall budget € 3 868 459 CIP EU-INSIDE 2012-2014 € 3 198 161 CIP ASSETS 2010-2012 € 4 250 000, overall budget € 5 312 502 CIP Europeana Regia 2010-2012 € 1 700 000, overall budget € 3 404 790 CIP Europeana Collections 1914-1918 2011-2014 € 2 693 954, overall budget € 5 387 911 CIP Natural Europe 2010-2013 € 2 350 000, overall budget € 4 700 000 CIP APEX 2012-2015 € 4 449 947, overall budget € 5 562 441
---------------	--

	<p>CIP 3D-ICONS 2012-2015 € 2 650 000, overall budget € 5 300 002</p> <p>CIP EURO-Photo 2010-2012 € 2 300 000, overall budget € 4 600 002</p> <p>CIP CARARE 2010-2013 € 4 300 000, overall budget € 4 300 000</p> <p>CIP Daguerreobase 2012-2015 € 1 499 963, overall budget € 1 874 962</p> <p>CIP Europeana Space 2014-2017 € 3 964 676, overall budget € 4 955 849</p> <p>CIP Ambrosia € 3 613 345, overall budget € 4 516 693</p> <p>CIP eCloud 2013-2016 € 3 799 661, overall budget € 4 749 583</p> <p>CIP EFG1914 2012-2014 € 2 100 000, overall budget € 4 200 005</p> <p>CIP ThinkMOTION 2010-2013 € 2 198 778, overall budget € 4 397 557</p> <p>CIP Europeana Fashion 2012-2015 € 2 649 969, overall budget € 3 312 464</p> <p>CIP OpenUp! 2011-2014 € 3 499 720, overall budget € 4 374 656</p> <p>CIP D2ME 2012-2015 € 2 099 955, overall budget € 2 624 944</p> <p>CIP Europeana Creative 2013-2015 € 4 250 000, overall budget: € 5 312 515</p> <p>CIP Europeana v3.0 2014-2015 € 4 300 000</p> <p>CIP LoCloud 2013-2016 € 3 400 000, overall budget € 4 250 003</p> <p>CIP EAwareness 2012-2014 € 3 909 118, overall budget € 4 886 405</p> <p>CIP EAGLE 2013-2016 € 2 999 812, overall budget € 3 749 765</p> <p>CIP PARTAGE PLUS 2012-2014 € 3 453 523, overall budget € 6 908 056</p> <p>Total funding = 78.25M Euro = 126M \$AUD (does not include 2007-2009 period)</p> <p>Allied projects that launched Europeana:</p> <p>ATHENA Access to Cultural Heritage Networks Across Europe</p> <p>FP6 ATHENA 2004-2007 € 14 399 999, overall budget € 26 207 361 (unclear how much went to Europeana)</p> <p>CIP ATHENAPLUS 2013-2015 € 4 117 204, overall budget: € 5 146 510</p> <p>Early digital libraries funding 2005-2008 led to Europeana v1.0 and v2.0 CIP investments</p> <p>Europeana Foundation Governance</p>
Notes	<p>Possible precursor to Europeana - advancement of European research libraries capacity (national libraries). FP6 TEL-ME-MOR 2005-2007 € 1 399 919</p> <p>Europeana Dashboard</p> <p>10 year anniversary (2018 Annual Report) outlines</p> <p>Europeana-DSI collaboration with DARIAH and CLARIN. 2015-2016 € 8 900 000</p> <p>Connecting European Facility (CEF) fund</p> <p>"The platform currently provides access to over 53 million cultural heritage items, (including image, text, sound, video and 3D material) from the collections of over 3,700 libraries, archives, museums, galleries and audio."</p> <p>Of the 10 CEF projects are all targeted at collection accessibility, enabling research and education through digitisation, curation, presentation and discovery. EU contribution (2016-2017) is with ~75% funds towards the overall budget: € 3.7M</p>

References

Leadership	https://pro.europeana.eu/contact-us
------------	---

Institutional centres of competence	https://en.wikipedia.org/wiki/European_Library https://pro.europeana.eu/services/data-publication-services/aggregators
Global links	https://pro.europeana.eu/files/Europeana_Professional/Publications/europeana-benchmark-report-sep-2016.pdf
Differentiator	https://pro.europeana.eu/page/next-generation-research-with-europeana-the-humanities-and-cultural-heritage-in-a-digital-perspective
Comparator	https://help.nla.gov.au/sites/default/files/StandardTrove_LargePresentationSlides_0.pdf
Further analysis	https://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_DSI/Deliverables/europeana-dsi-d3.2-europeana-research-distribution-plan.pdf https://pro.europeana.eu/project/research-requirements
Government services	http://www.theeuropeanlibrary.org/tel4/
Relevant info	https://cordis.europa.eu/project/rcn/191890/factsheet/en https://cordis.europa.eu/project/rcn/191926/factsheet/en https://cordis.europa.eu/project/rcn/191879/factsheet/en https://cordis.europa.eu/project/rcn/191728/factsheet/en https://cordis.europa.eu/project/rcn/191735/factsheet/en https://cordis.europa.eu/project/rcn/191881/factsheet/en https://cordis.europa.eu/project/rcn/191929/factsheet/en https://cordis.europa.eu/project/rcn/191908/factsheet/en https://cordis.europa.eu/project/rcn/191873/factsheet/en https://cordis.europa.eu/project/rcn/191807/factsheet/en https://cordis.europa.eu/project/rcn/191914/en https://cordis.europa.eu/project/rcn/191827/factsheet/en https://cordis.europa.eu/project/rcn/191822/factsheet/en https://cordis.europa.eu/project/rcn/191935/factsheet/en https://cordis.europa.eu/project/rcn/191752/factsheet/en https://cordis.europa.eu/project/rcn/191878/factsheet/en https://cordis.europa.eu/project/rcn/191748/factsheet/en https://cordis.europa.eu/project/rcn/191883/factsheet/en https://cordis.europa.eu/project/rcn/191921/factsheet/en https://cordis.europa.eu/project/rcn/191785/factsheet/en https://cordis.europa.eu/project/rcn/191518/factsheet/en https://cordis.europa.eu/project/rcn/191940/factsheet/en https://cordis.europa.eu/project/rcn/191917/factsheet/en https://cordis.europa.eu/project/rcn/191780/factsheet/en https://cordis.europa.eu/project/rcn/191933/factsheet/en https://pro.europeana.eu/post/european-commission-report-confirms-continued-member-state-support-for-europeana

	https://cordis.europa.eu/project/rcn/72762/factsheet/en https://www.promoter.it/athena (2008-2011) https://cordis.europa.eu/project/rcn/191937/factsheet/en https://pro.europeana.eu/project/athena http://ec.europa.eu/information_society/activities/econtentplus/projects/index_en.htm https://en.wikipedia.org/wiki/Europeana https://pro.europeana.eu/our-mission/foundation-governing-board
Notes	https://cordis.europa.eu/project/rcn/80564/factsheet/en https://pro.europeana.eu/post/introducing-the-europeana-statistics-dashboard https://www.dariah.eu/activities/projects-and-affiliations/dariah-in-past-eu-projects/ https://ec.europa.eu/inea/en/news-events/newsroom/european-commission-approves-report-europeana https://ec.europa.eu/inea/en/connecting-europe-facility/cef-telecom/projects-by-dsi/europeana

1.10 Te Mana Raraunga - Māori Data Sovereignty Network

Māori Data Sovereignty Network was established to “advocate for Māori rights and interests in data to be protected as the world moves into an increasingly open data environment”, <https://www.temanararaunga.maori.nz/>

Administration

Funding sources	Institutional, Community group
Jurisdiction	New Zealand
Leader	Three supporting organisations: National Hauora Coalition, Whakathea Māori Trust Board, Waiora Pacific (now Takiwā) with expertise in Māori community health, legislation, government policy, data management and reuse.
Domains	Indigenous
Start date	2015
Timescale	3-5 Years
Investment	No national policy driven investment, short-term research project funding.
Phases	Pre-operational (design/plan/construct), Operational (maintain/consolidate/grow)

Structure

Features	Cross-cutting, National nodes, Global nodes, Community affiliations and advocacy
Type	Capability building, Community building, Information sources
Capabilities	Data curation, Data reuse, Data linkage, Domain expertise, Indigenous knowledge, cultural knowledge and protocols
Organisation	Distributed, Virtual, Community driven
Governance	Coordinating entity (hub), Network (participating nodes), Decentralised (institutional), Operational function, Outreach and liaison function, Policy function, Expert advisors
User community	Te Mana Raraunga provide advocacy across the government, higher education, and industry sectors; and, the wider community.
Institutional centres	Working group members have institutional roles in: National Institute of Demographic and Economic Analysis & Faculty of Māori and Indigenous Studies at University of Waikato, Whakathea Māori Trust Board, Department of Statistics, University of Auckland, Te Rōpū Rangahau Hauora A Eru Pōmare, University of Otago (Wellington), and Te Kupenga Hauora Māori, University of Auckland. Working group members have diverse iwi affiliations and lead national initiatives e.g. Planet Maori (technological sovereignty) and Figure.nz. Te Mana Raraunga listed on University of Auckland School of Psychology as part of advice to researchers on Māori and Pacific research ethics.
Government links	The aliations of the working group members include institutional roles and memberships on advisory or expert boards advocating for and representing community interests. Note: New Zealand universities are governed under national legislation. Working group members have diverse government aliations e.g. Māori ICT Development Fund Expert Advisory Group, Māori Digital Technology Development

	<p>Fund (Te Puni Kōākiri - Ministry of Māori Development), Te Māngai Pāho (Māori Broadcasting Funding Agency), Tuhono Trust (Māori Aliation Service).</p> <p>Te Mana Raraunga in partnerships with Statistics New Zealand, involved in the development of a data stewardship framework through the government digital transformation agenda (digital.govt.nz) and the Ministry of Business, Innovation and Employment to: establish a Māori data governance framework, advise on the New Zealand Research Information System and operate as a member of the NSOG (NZRIS Stewardship and Oversight Group).</p>
--	---

Key Aspects

Global links	<p>Working group members have diverse international affiliations e.g. Te Ha (The Cultural Conservancy) and Te Ohu Whai Ao (Maori Indigenous Business Development Trust).</p> <p>Participants in the Research Data Alliance - Indigenous Data Sovereignty interest group.</p>
Differentiator	<p>Advocacy for Maori Data Sovereignty across all sectors and spheres.</p> <p>The purpose of Te Mana Raraunga is to enable Māori Data Sovereignty and to advance Māori aspirations for collective and individual wellbeing by: Asserting Māori rights and interests in relation to data; Ensuring data for and about Māori can be safeguarded and protected; Requiring the quality and integrity of Maori data and its collection; Advocating for Māori involvement in the governance of data repositories; Supporting the development of Māori data infrastructure and security systems; Supporting the development of sustainable Māori digital businesses and innovations</p>
AU comparator	<p>Maia m nāyri Wingara (Aboriginal and Torres Strait Islander Data Sovereignty Collective), Indigenous Data Network (University of Melbourne)</p>

Data Infrastructure

Digitisation	<p>Strategic digitisation agenda (in relation to Māori) is led by the National Library of New Zealand and Archives New Zealand, in partnership with the Ministry for Culture and Heritage (Internal Affairs). See report on the impact of digitisation on te reo (Māori language) archival collections.</p>
Longitudinal	<p>Ministry for Culture and Heritage work programme references digitisation and long term records retained in National Library of New Zealand and Archives New Zealand. https://mch.govt.nz/strengthening-our-national-archives-and-libraries that may impact upon the availability of longitudinal data (that relates to colonisation and early census records).</p>
Linkage	<p>Expert advice provided on the approaches to and impacts of administrative data linkage.</p>

Research Leads

Trends	<p>Issues in common with other nations is the lack of indigenous input and oversight into the use of administrative data (largely impacting indigenous communities), and the integration of cultural and privacy concerns.</p> <p>Consultation around management of administrative data, data linkage, particularly with regard to health and wellbeing challenges.</p>
--------	---

Further analysis	<p>The global network already established representing the interests of first nation peoples in the global south (links between Australia and New Zealand) and the global north (links to Canada, USA) established through common colonialisation history versus Sami in Europe (and first nations in European territories e.g. Kanak (New Caledonia).</p> <p>Progress underway and Australian involvement at a global level through Research Data Alliance international interest group and the introduction of the CARE (Collective ownership, Authority to control, Responsibility, Ethics) as core data stewardship principles.</p> <p>Whether there are moves (as for New Zealand) at federal level to integrate expertise in indigenous data stewardship into multiple agencies, especially with regard to data linkage of microdata and administrative data (as a form of in common critical research data infrastructure).</p>
------------------	--

Additional Information

Relevant Info	<p>Government digital transformation.</p> <p>Access to administrative data and indigenous heritage</p> <p>National research information system strategy</p>
Notes	<p>Project work (with some institutional funding attached)</p> <p>Te Mana Raraunga involvement with Statistics New Zealand - see reference to pilot projects and delays in 2017 Annual Report and other partnerships with MaŀNori and iwi</p> <p>Statistics New Zealand engagement with Te Mana Raraunga on census data quality</p> <p>The position of New Zealand's Government Chief Data Steward (GCDS) is held by the CEO of Statistics New Zealand (Liz MacPherson).</p> <p>A national government data stewardship framework is being established and has the following representatives: Chief Archivist, Government Chief Digital Officer, Government Chief Information and Security Officer, Government Chief Privacy Officer, Government chief executives, Government Statistician, Commissioner of Inland Revenue, Iwi Leaders Forum, sector leads, and Te Mana Raraunga.</p> <p>New Zealand's NOSG (NZRIS Oversight and Stewardship Group) terms of reference outlines that expert input is intended to:</p> <ol style="list-style-type: none"> "1. Provide a system-wide view of research, science and innovation information 2. Ensure open data which is easily accessible and widely used 3. Protect personal and commercially sensitive data 4. Enable the reuse of data 5. Reduce collection and reporting burden 6. Ensure data is trusted, authoritative and well-managed 7. Enable easy and automatic movement of data between systems" <p>Te Mana Raraunga represented at hui (meeting) in 2017 "Te Ritorito" to discuss research, policies and programmes that support whānau, hapuā and iwi (family, tribe, community) wellbeing.</p> <p>Notable that Te Mana Raraunga have representatives with expertise in: data, demographics, statistics, technology, cultural heritage, health and wellbeing.</p> <p>Advocacy around low census response (and decision-making around census design)</p>

	<p>and in particular from Māori.</p> <p>NZRI operates through partnerships. See also the four themes of MBIE "Vision Mātauranga" and purpose:</p> <p>"To use the science and innovation system to help unlock the potential of Māori knowledge, people and resources for the benefit of New Zealand.</p> <p>To recognise Māori as important partners in science and innovation; both as inter-generational guardians of significant natural resources and indigenous knowledge, and owners and managers of commercial assets.</p> <p>To build the capability of Māori individuals, businesses, incorporations, rūnanga, trusts, iwi, hapū, and marae to engage with science and innovation.</p> <p>To maximise the quality of the relationship between Māori and the Crown through science and innovation through the Treaty of Waitangi."</p>
--	--

References

Leadership	https://www.temanararaunga.maori.nz/whakapapa
Institutional centres of competence	https://www.temanararaunga.maori.nz/tangata https://www.nhc.maori.nz/ http://www.whakatohea.co.nz/ http://planetmaori.com/ http://www.nativeland.org/te-ha https://www.tpk.govt.nz/en/whakamahia/it-and-innovation/ka-hao-maori-digital-technology-development-fund https://www.taiuru.maori.nz/publications/analysis-maori-ict-groups/maori-ict-development-fund-expert-advisory-group/ https://www.tuhono.net/ https://www.tpk.govt.nz/en/whakamahia/it-and-innovation/ka-hao-maori-digital-technology-development-fund https://gure.nz/
Government links	https://www.stats.govt.nz/integrated-data/integrated-data-infrastructure/
Global links	http://www.teohuwhaiao.org/about.html https://www.rd-alliance.org/advancing-indigenous-data-sovereignty-care-principles-indigenous-data-governance
Comparator	https://www.maiamnayriwingara.org/ https://mspgh.unimelb.edu.au/centres-institutes/centre-for-health-equity/research-group/indigenous-data-network
Data linkage	https://www.psych.auckland.ac.nz/en/about/our-research/MaoriandPacificresearchethicswithintheSchoolofPsychology.html https://www.dss.gov.au/policies-legislation/data-integration
Relevant info	https://www.digital.govt.nz/ https://www.mbie.govt.nz/assets/3109d33a46/maori-engagement-on-data-conceptual-model.pdf http://www.mbie.govt.nz/info-services/science-innovation/research-and-data/nris/document-image-library/NRIS-overview.pdf http://www.mbie.govt.nz/info-services/science-innovation/research-and-data/pdf-library/research-science-and-innovation-domain-plan.pdf

	https://natlib.govt.nz/librarians/reports-and-research/korero-kitea
Notes	<p>"https://www.temanararaunga.maori.nz/ng-mahi</p> <p>http://archive.stats.govt.nz/about_us/what-we-do/our-publications/annual-reports/annual-report-2017.aspx</p> <p>https://www.stats.govt.nz/assets/Uploads/Data-leadership-fact-sheets/fact-sheet-partnering-with-maori-apr18.pdf</p> <p>https://www.stats.govt.nz/news/2018-census-external-data-quality-panel-conrmed</p> <p>https://www.digital.govt.nz/digital-government/leadership-and-governance/government-chief-data-steward-gcds/</p> <p>https://www.data.govt.nz/manage-data/data-stewardship/a-draft-data-stewardship-framework-for-nz/</p> <p>https://www.tpk.govt.nz/en/whakamahia/whanau-ora/te-ritorito-2017</p> <p>https://healthcentral.nz/failed-census-a-preventable-public-health-policy-disaster/</p> <p>https://www.mbie.govt.nz/science-and-technology/science-and-innovation/research-and-data/nzris/nzris-development/</p> <p>https://www.mbie.govt.nz/science-and-technology/science-and-innovation/agencies-policies-and-budget-initiatives/vision-matauranga-policy/"</p>

1.11 First Nations Information Governance Centre

“With First Nations, we assert data sovereignty and support the development of information governance and management at the community level through regional and national partnerships. We adhere to free, prior and informed consent, respect nation-to-nation relationships, and recognize the distinct customs of nations”,

<https://fnigc.ca/about-fnigc/mission.html>

Administration

Funding sources	National, Institutional, Community group
Jurisdiction	Canada
Leader	Membership comprises First Nation governing organisations, and with a focus on health and wellbeing: Union of Nova Scotia Indians; Union of New Brunswick Indians; First Nations of Quebec and Labrador Health and Social Services Commission; First Nations Health and Social Secretariat of Manitoba (Nanaandawewigamig); Chiefs of Ontario; Federation of Sovereign Indigenous Nations; Alberta First Nations Information Governance Centre; Dene Nation; Council of Yukon First Nations; First Nations Health Authority.
Domains	Indigenous
Start date	2010
Timescale	6-10 Years
Investment	No national policy driven investment, survey project funding.
Phases	Pre-operational (design/plan/construct), Operational (maintain/consolidate/grow)

Structure

Features	Cross-cutting, Government linkages, National nodes
Type	Data service, Discovery platform, Capability building, Community building, Information sources, Research practices and methods
Capabilities	Data curation, Data reuse, Informatics and data modelling, Survey design and instrument, Project support, Domain expertise
Organisation	Distributed, Shared facility
Governance	Coordinating entity (hub), Centralised (nexus), Operational function, Outreach and liaison function, Informatics function
User community	FNIGC provides services and advocacy across the government, higher education, and industry sectors; and, the wider community.
Institutional centres	Alberta First Nations Information Governance Centre appears to be a node of the FNIGC.
Government links	Funding at national for surveys by: Health Canada, and Indigenous and Northern Affairs Canada, and Employment and Social Development Canada. Funding at regional level for surveys by: Alberta First Nations Information Governance Centre, the First Nations Health and Social Secretariat of Manitoba (Nanaandawewigamig), the Chiefs of Ontario, the Union of New Brunswick Indians, the Dene Nation (Northwest Territories), the Union of Nova Scotia Indians, the First Nations of Quebec and Labrador Health and Social Service Commission, the

	<p>Federation of Sovereign Indigenous Nations, the Council of Yukon First Nations , and the First Nations Health Authority (British Columbia).</p> <p>See: Canada's Third Biennial Plan to the Open Government Partnership 2016-2018, End-of-Term Self-Assessment Report, November 2018. References integration of indigenous knowledge in research and consultation with FNIGC.</p>
--	--

Key Aspects

Global links	Participation in the Research Data Alliance (unclear if FNIGC representatives participate in the Indigenous Data Sovereignty interest group).
Differentiator	Focused surveys and data capture for a range of health and wellbeing reasons where they impact on First Nations (indigenous peoples) in Canada.
AU comparator	<p>National Aboriginal and Torres Strait Islander Health Survey (NATSIHS), National Aboriginal and Torres Strait Islander Social Survey (NATSISS), National Aboriginal and Torres Strait Islander Nutrition and Physical Activity Survey (NATSINPAS), Australian Bureau of Statistics.</p> <p>Mayi Kuwayu Survey (National Study of Aboriginal and Torres Strait Islander Wellbeing).</p> <p>National Indigenous Languages Survey, Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS).</p> <p>Other organisations involved: National Aboriginal Community Controlled Health Organisation (NACCHO), Australian Institute of Health and Welfare (AIHW), Departments of Social Services (HILDA).</p>

Data Infrastructure

Digitisation	See: The First Nations Information Governance Centre. Ownership, Control, Access and Possession (OCAP,™): The Path to First Nations Information Governance. May 2014. (Ottawa: The First Nations Information Governance Centre, May 2014) regarding issues of indigenous data sovereignty and administrative records and archives or documentary heritage retained by Libraries and Archives Canada.
Longitudinal	FNIGC specialises in the capture and management of longitudinal data through surveys: the First Nations Regional Health Survey (FNRHS), First Nations Regional Early Childhood, Education and Employment Survey (FNREEES), First Nations Labor and Employment Development Survey, First Nations Oral Health Survey, and the First Nations Community Survey.
Linkage	FNIGC survey data deemed a source of data to aid with government decision-making and reporting. See report "Strengthening the Availability of First Nations Data" prepared for the Indigenous Services Canada & The Assembly of First Nations by Shelley Trevethan, January 30, 2019. Report covers data relating to the lived experience of Inuit, Métis and First Nations (indigenous) peoples.

Research Leads

Trends	<p>The combination of survey design and instruments, national longitudinal dataset generation, with multiple surveys relating to health and wellbeing dimensions.</p> <p>Unclear about the existence of surveys of indigenous populations specifically targeted at cultural connection and language skills.</p>
Further analysis	Whether there are moves (as for Canada) at federal level to integrate expertise in

	indigenous data stewardship into multiple agencies, especially with regard to data linkage of microdata and administrative data (as a form of in common critical research data infrastructure). See SSHRC funded research on indigenous data stewardship.
--	---

Additional Information

Relevant Info	FNIGC as cross-cutting data infrastructure provider and community advocacy on data sovereignty
Notes	<p>FNIGC was incorporated in 2010 but their role in providing data infrastructure arose from a First Nations Regional Health Survey back in 1997 (longitudinal). Reference to the integration of indigenous knowledge in the research system as part of the Canadian science and technology plan (2020-2025).</p> <p>See Social Sciences and Humanities Research Council grant allocations (January 2019) for indigenous research capacity and reconciliation</p> <p>Canada's Science Vision</p> <p>See SSHRC support information</p> <p>Innovation Canada - lists one project on distributed data management for indigenous knowledge research with Inuit and First Nation communities \$516,323</p> <p>Libraries and Archives Canada strategic initiative to preserve indigenous culture and language (in documentary history collections) as part of the institution's "Indigenous Documentary Heritage Initiatives" (IDHI).</p> <p>Report: Data Resources and Challenges for First Nation Communities, Document Review and Position Paper, Prepared for the Alberta First Nations Information Governance Centre, by Kate McBride</p>

References

Leadership	https://fnigc.ca/about-fnigc/member-organizations.html
Institutional centres of competence	http://www.afnigc.ca/main/index.php?id=home
Government links	https://fnigc.ca/our-work/what-we-do.html
Differentiator	https://fnigc.ca/our-work/regional-health-survey/about-rhs.html
Comparator	https://meteor.aihw.gov.au/content/index.phtml/itemId/611753 https://www.abs.gov.au/ausstats/abs@.nsf/mf/4714.0 https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/4727.0.55.005main+features1 2012-13 https://mkstudy.com.au/the-survey/
Data linkage	https://www.afn.ca/wp-content/uploads/2019/05/NCR-11176060-v1-STRENGTHENING_THE_AVAILABILITY_OF_FIRST_NATIONS_DATA-MAR_25_2019-FINAL_E.pdf
Relevant info	https://fnigc.ca/about-fnigc/frequently-asked-questions.html https://fnigc.ca/news/introducing-first-nations-data-centre-new-way-access-first-nations-data.html https://fnigc.ca/sites/default/files/docs/ocap_path_to_fn_information_governance

	<p>_en_final.pdf</p> <p>http://vpri-irsi.sites.olt.ubc.ca/2019/04/18/dealing-with-indigenous-data-information-and-records-the-big-picture-part-1-of-2/</p> <p>http://vpri-irsi.sites.olt.ubc.ca/2019/06/13/dealing-with-indigenous-data-information-records-towards-data-sovereignty-part-2-of-2/</p> <p>https://research.ubc.ca/sites/research.ubc.ca/files/u9680/SSHRC_PositionPaper_Report_SUMMARY.pdf</p> <p>http://www.afnigc.ca/main/index.php?id=home&content=home</p>
Notes	<p>"https://fnigc.ca/first-nations-regional-health-survey.html</p> <p>https://www.canada.ca/en/polar-knowledge/science-and-technology-plan.html</p> <p>https://www.canada.ca/en/social-sciences-humanities-research/news/2019/01/government-of-canada-supports-indigenous-research-capacity-and-reconciliation.html</p> <p>http://www.ic.gc.ca/eic/site/131.nsf/eng/h_00000.html</p> <p>http://www.sshrc-crsh.gc.ca/society-societe/community-communite/indigenous_research-recherche_autochtone/index-eng.aspx</p> <p>https://www.innovation.ca/about/press-release/government-canada-invests-research-driven-knowledge-economy-516323</p> <p>https://www.canada.ca/en/library-archives/news/2019/04/library-and-archives-canada-launches-funding-and-services-to-help-preserve-indigenous-culture-and-language-recordings.html</p> <p>http://www.bac-lac.gc.ca/eng/discover/aboriginal-heritage/initiatives/Pages/default.aspx</p> <p>https://www.opengovpartnership.org/wp-content/uploads/2018/12/Canada_End-term_Self-Assessment-Report_2016-2018_EN.pdf</p> <p>http://www.afnigc.ca/main/includes/media/pdf/digital%20reports/Data_Resources_Report.pdf"</p>

1.12 SAIL Databank

“SAIL stands for Secure Anonymised Information Linkage. The SAIL Databank is a world-class flagship for the robust secure storage and use of anonymised person-based data for research to improve health, well-being and services. Its databank of anonymised data about the population of Wales is internationally recognised. Backed and endorsed by the Government, the SAIL Databank receives core funding from the Welsh Government’s Health and Care Research Wales”

“The SAIL Databank is now powered by the UK Secure e-Research Platform (UKSeRP), developed by the Health Informatics Group at Swansea University, with support from the Farr Institute of Health Informatics Research funded by Medical Research Council”, <https://saildatabank.com/about-us/overview/>

Administration

Funding sources	National, Research council, Institutional, Core funding provided by Welsh Government's Health and Care Research Wales to 2020.
Jurisdiction	UK
Leader	SAIL Databank is based at Swansea University in Wales
Domains	Social Science, Health and medicine
Start date	2007
Timescale	11-15 Years
Investment	\$21-50M
Phases	Pre-operational (design/plan/construct), Operational (maintain/consolidate/grow)

Structure

Features	Cross-cutting, Regional impact, Institutional centres of excellence, Government linkages, National nodes, Global nodes
Type	Archive or repository service, Data service, Discovery platform, Capability building, Community building, Software service, Information sources, Research practices and methods, HPC service, Cloud storage/compute service, Research tools and platforms
Capabilities	Data curation, Data preservation, Data reuse, Data mining and analysis, Informatics and data modelling, Data linkage, Project support, Technology expertise, Domain expertise
Organisation	Distributed, Virtual, Shared facility, Institutional facility, The Welsh node in a UK network of eHealth centres that form the Farr Institute
Governance	Coordinating entity (hub), Network (participating nodes), Centralised (nexus), Operational function, Technologies function, Expert advisors
User community	SAIL databank history report (2007-2017) references >370 users and 78 organisations from across the blog (mainly UK usage).
Institutional centres	SAIL Databank is one of four eHealth research centres of excellence in England and Scotland in the UK (other CoE: in London, Manchester and Dundee), that operate as the Farr Institute for Health Informatics Research. See 2016-2017 Annual Report.
Government links	Key working relationships with the Welsh government and Health and Care Wales on Welsh, Genomics England (CLIMB), GW4 Alliance, UK Biobank, Supercomputing

	<p>Wales, on research infrastructure and funding.</p> <p>Welsh node of UK Administrative Data Research Centre.</p> <p>Collaborations with the UK government agencies: National Health Service.</p> <p>Examples of government links: Department of Education, Medical Research Council, Wales Community Rehabilitation, Economic and Social Research Council, Department of Transport, National Institute for Health and Care Excellence, Public Health Wales.</p>
--	---

Key Aspects

Global links	<p>Partnerships with: Europe (European Commission, Euromedcat, EU Social Fund) and Australia (University of Western Australia, Curtin University, Monash University).</p> <p>Identifies Canada (Ministry of health and Long-Term Care) and Australia (PHRN) as global equivalents for data linkage.</p>
Differentiator	<p>Secure Anonymised Information Linkage (SAIL) is a national data safe haven of anonymised health and administrative datasets about the populations of Wales, in support of population data science and data linkage. In 2015 the SAIL Databank became ISO 27001 certified for its security (UKSeRP - UK Secure eResearch Platform) and operates in a new data science building, and is establishing cloud infrastructure for microbial bioinformatics (CLIMB).</p>
AU comparator	<p>Identifies Australia Population Health Research Network (NCRIS).</p> <p>Australian Bureau of Statistics (ABS Datalab remote or onsite).</p> <p>Bioplatforms Australia (NCRIS).</p> <p>Australian Data Archive (ANU)</p> <p>National Centre for Indigenous Genomics (ANU)</p>

Data Infrastructure

Digitisation	<p>Move to include text data, includes the text analysis of patient records as part of CLIMB project (SAIL Databank, UKSeRP).</p>
Government services	<p>Office for National Statistics, Secure Research Service and Government Statistical Service.</p>
Longitudinal	<p>Provides secure access to demographic, health, social and education data from the resident population of Wales. By example, provides the "NHS Welsh Demographic Service (WDS) dataset containing address information linked anonymously at individual level." Public Health Research, No. 4.3 D Fone et al, March 2016.</p> <p>Datasets types: health, social services, education and national survey data, based on populations of Wales and the UK.</p>
Linkage	<p>Provides secure data linkage services in support of population research in Wales (and supports comparative population research), strong focus on health and wellbeing data.</p>

Research Leads

Trends	<p>Clustered infrastructures data, expertise, tools, compute operates as a highly effective means to attract research grants and establish local, national and global networks.</p>
--------	---

Further analysis	<p>SAIL databank is building diverse data processing techniques into platform and suite of services to support medical and health research. Are the tools and techniques to extract domain specific data from unstructured text and characterisation tools and techniques for image data generalisable and reusable for HASS research, or are they too domain specific?</p> <p>The successful trajectory of combined infrastructure services (data linkage, a safe haven databank, tools) and national networks to support local, national and international collaborations.</p> <p>Movement from pilot stage (and design) to establish a shared national facility with formal partnerships with access controls, that serves local and national research infrastructure requirements (whether that approach could be generalised and used to meet the needs of research to access to indigenous heritage and research data in government and higher education collections).</p>
------------------	--

Additional Information

Relevant Info	<p>SAIL Databank</p> <p>Farr Institute</p> <p>Administrative Data Research Network</p> <p>Health and Care Research Wales (funder/partner)</p> <p>Total fund calculation based on figures specific to Swansea investment i.e. pilot, platform and building 17.7M pounds ~ 31.39M AUD. Rather than national funding for Farr Institute "With a ~£17.5M research award from the 10-funder consortium, plus an additional ~£20M of capital funds from the Medical Research Council"</p> <p>These are national figures for funding the funding across four sites, including Swansea.</p> <p>Pilot funding information on funding from Health and Care Research Wales is ~£399K.</p> <p>CIPHER project - ~£ 9.3M at Swansea to Farr Institute, ~£ 8M for Data Science Building to host ADRC Wales</p> <p>Administrative Data Research Centre (ADRC) Wales ~£ 13M from Economic and Social Research Council "big data initiative"</p> <p>Equivalent secure services provided by the ABS</p>
Notes	<p>Examples of sector partnerships: Biobank, AsthmaUK, Cedar, Arthritis Research UK, MS Society, Wellcome Trust.</p> <p>Bioplatfrom Australia (cloud infrastructure forms part of the Biocommons pathfinder project - BPA, ARDC, AARNet). Aims to expand capacity to add data processing capacities to include: textual data (natural language processing), MRI data and actigraphy data.</p> <p>SAIL Databank is colocated in a cluster at Swansea University, Health Informatics Group: the Administrative Data Research Centre, the Farr Institute, and the National Centre for Population health and Wellbeing Research.</p> <p>Swansea cluster is a partner in CLIMB (cloud infrastructure for microbial research) along with three other sites: Warwick, Birmingham, Cardiff.</p>

References

Leadership	https://saildatabank.com/contact/
Institutional centres of competence	http://farrinstitute.org/wp-content/uploads/2018/09/The-Farr-Institute-Annual-Report-2016-2017.pdf
Global links	https://saildatabank.com/the-sail-databank-10-years-of-spearheading-data-privacy-and-research-utility/
Differentiator	https://saildatabank.com/saildata/data-privacy-security/
Comparator	https://www.abs.gov.au/websitedbs/D3310114.nsf/89a5f3d8684682b6ca256de4002c809b/55c077b4d87d2fbdca2572090006fbcf!OpenDocument https://www.bioplatforms.com/ https://www.phrn.org.au/ http://www.ada.edu.au/ http://ncig.anu.edu.au/
Digitisation	https://saildatabank.com/wp-content/uploads/UKSeRP_Brochure_v1.5.pdf
Government services	https://www.statisticsauthority.gov.uk/joining-up-data-for-better-statistics/
Longitudinal	https://saildatabank.com/saildata/sail-datasets/#core
Data linkage	https://saildatabank.com/about-us/data-linkage/
Relevant info	https://saildatabank.com/about-us/overview/ https://saildatabank.com/category/publication/ https://saildatabank.com/wp-content/uploads/SAIL_10_year_anniversary_brochure.pdf https://saildatabank.com/about-us/history/ https://adrn.ac.uk https://www.farrinstitute.org https://www.healthandcarerresearch.gov.wales/research-infrastructure-map/ https://www.ncbi.nlm.nih.gov/books/NBK350762/ http://www.wales.nhs.uk/sites3/documents/952/nischr_ahsc_3_years_on.pdf https://www.abs.gov.au/websitedbs/D3310114.nsf/home/ABS+Data+Access+Arrangements https://mrc.ukri.org/documents/pdf/mapping-the-landscape-of-uk-health-data-research-and-innovation-report/ https://www.healthandcarerresearch.gov.wales/secure-anonymised-information-linkage-databank/ https://www.healthandcarerresearch.gov.wales/uploads/SAIL_Annual_Report_17-18_eng.pdf https://www.healthandcarerresearch.gov.wales/uploads/Policy%20%26%20Strategy/Use%20of%20routine%20data/Maximising%20the%20Use%20of%20Routine%20Data%20for%20Research%20in%20Wales.pdf https://www.healthandcarerresearch.gov.wales/uploads/About/Health_Care_Resea

	<p>rch_Wales_annual_report_17-18_eng.pdf</p> <p>https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2744675/</p> <p>https://royalsociety.org/~media/policy/projects/data-governance/data-governance-call-collated-evidence.pdf</p> <p>https://informatics.bmj.com/content/24/1/1</p> <p>https://www.ons.gov.uk/aboutus/whatwedo/paidservices/virtualmicrodatalaboratoryvml</p> <p>https://gss.civilservice.gov.uk/policy-store/data-linking/</p> <p>https://www.sciencedirect.com/science/article/pii/S014067361632298X?via%3Dihub</p> <p>https://www.bristol.ac.uk/media-library/sites/jean-golding-institute/documents/07-ChrisOrton.pdf</p> <p>https://www.hdruk.ac.uk/people/professor-david-ford/</p>
--	--

1.13 Cohort and Longitudinal Studies Enhancement Resources (CLOSER)

“CLOSER, the home of longitudinal research, brings together eight world-leading longitudinal studies with participants born throughout the 20th and 21st centuries. Our work maximises the use, value and impact of these, and other, longitudinal studies to help improve our understanding of key social and biomedical challenges”, <https://www.closer.ac.uk/home/what-we-do/>

Administration

Funding sources	National, Research council
Jurisdiction	UK
Leader	CLOSER is based at University College London, Department of Social Science
Domains	Social Science, Health and medical
Start date	2012
Timescale	6-10 Years
Investment	\$11-15M
Phases	Pre-operational (design/plan/construct), Operational (maintain/consolidate/grow)

Structure

Features	Cross-cutting, Institutional centres of excellence, Government linkages, National nodes, Global nodes
Type	Archive or repository service, Data service, Discovery platform, Capability building, Community building, Information sources, Research practices and methods, Research libraries
Capabilities	Data curation, Data reuse, Data mining and analysis, Informatics and data modelling, Survey design and instrument, Data linkage, Project support, Domain expertise
Organisation	Distributed, Virtual, Shared facility, Institutional facility
Governance	Coordinating entity (hub), Network (participating nodes), Centralised (nexus), Executive body, Operational function, Outreach and liaison function, Administration function, Informatics function, Expert advisors
User community	Social science and health and medical researchers
Institutional centres	British Library, UK Data Service, UCL Institute of Education, University of Southampton, University of Bristol and University of Essex.
Government links	British Library is a partner to assist with the Mass Observation data collection. Regular use of datasets from the Office for National Statistics, National Records agencies (UK) and referral to administrative data maintained in government departments and agencies.

Key Aspects

Global links	CLOSER International project extending reach to enabling international comparative research, around economies in developing countries. Research links with European genomics research.
Differentiator	Foundational strengths with cohort surveys and longitudinal data collection

	practices, commencing concertedly post-WWII and increasing in number and diversified in 1990s.
AU comparator	<p>National Centre for Longitudinal Data (NCLD), Department of Social Services.</p> <p>Longitudinal data is maintained by survey managers for HILDA, LSAY etc.</p> <p>Longitudinal datasets are accessible for reuse from the Australian Data Archive and the Australian Bureau of Statistics.</p> <p>Longitudinal surveys funded and overseen by Department of Social Services.</p> <p>Education and training around longitudinal data analysis provided by the Australian Consortium for Social and Political Research Inc.</p> <p>Commonwealth Accredited Integrating Authorities have expertise, advice provided via the National Statistical Service on data integration.</p>

Data Infrastructure

Digitisation	A range of references to digitisation of data in analogue formats (paper or microform) as means of making earlier collections of data accessible for reuse, see the Mass Observation data in particular (and the role of the British Library and the University of Sussex Library).
Government services	UK Data Service (funded by ESRC) provide access to social and economic data to research, teaching and policy (hosted at University of Essex).
Longitudinal	A range of social science-oriented cohort studies: millennial generation, households, women, parents and children, child development,
Linkage	Linkage with genetic data maintained as part of the National Survey of Health and Development (MRC) at UCL and University of Bristol, Avon Longitudinal Study of Parents and Children (ALSPAC) in the UK; and, the European Genome-Phenome Archive.

Research Leads

Trends	Evolution of research infrastructure out of establishing survey mechanisms and data collection to a centre of competence as a hub in the research network and linking to other key stakeholders outside of higher education.
Further analysis	<p>CLOSER is embedded in the university context whereas National Centre for Longitudinal Data (NCLD) operates out of the Department of Social Services, this shifts the way the infrastructure is coordinated, and data can be made accessible for research.</p> <p>How CLOSER liaises with the Australian DSS and ABS equivalents in the UK, to ascertain what network links are established and collaborations enabled.</p>

Additional Information

Relevant Info	<p>CLOSER funding information</p> <p>CLOSER data accessible via UK Data Service</p>
Notes	<p>Funding from the Economic and Social Research Council (ESRC) and the Medical Research Council (MRC) 2012-2017, and ESRC from 2017-2021.</p> <p>Collaboration with CESSDA around informatics, see CLOSER profile in CESSDA documentation on DDI.</p> <p>CLOSER has points researchers to ICSPR and World Bank</p>

	UK Data Archive has links to Office for National Statistics and ICSPR (Inter-University Consortium for Political and Social Research).
--	--

References

User community	https://www.closer.ac.uk/partners/
Leadership	https://gtr.ukri.org/projects?ref=ES/K000357/1 https://www.closer.ac.uk/home/what-we-do/ https://www.closer.ac.uk/people/team/
Institutional centres	2019/07/22 9:16:06 PM GMT+10
Government links	https://www.bl.uk/business-and-management/editorials/resource-highlight-using-the-mass-observation-archive-package https://www.closer.ac.uk/wp-content/uploads/CLOSER-resource-NHS-Numbers-and-their-management-systems.pdf https://learning.closer.ac.uk/glossary/administrative-data/
Global links	https://www.closer.ac.uk/news-opinion/news/closer-international-project-launch/ https://ega-archive.org/
Differentiator	https://www.closer.ac.uk/closer/explore-the-studies/
Comparator	https://www.dss.gov.au/sites/default/files/documents/03_2013/guide-to-australian-longitudinal-studies.pdf https://www.abs.gov.au/websitedbs/D3310114.nsf/Home/Statistical+Data+Integration https://www.dss.gov.au/about-the-department/national-centre-for-longitudinal-data
Government services	https://ukdataservice.ac.uk/
Longitudinal	https://gtr.ukri.org/projects?ref=ES/K000357/1 https://www.closer.ac.uk/wp-content/uploads/CLSSlides.pdf https://esrc.ukri.org/news-events-and-publications/publications/corporate-publications/longitudinal-studies-review-2017/the-future-of-longitudinal-data/ https://esrc.ukri.org/news-events-and-publications/publications/corporate-publications/longitudinal-studies-review-2017/background-to-the-longitudinal-studies-review/
Data linkage	https://www.nshd.mrc.ac.uk/nshd/about-nshd/ https://esrc.ukri.org/files/news-events-and-publications/publications/annexes-to-the-longitudinal-studies-review-2017/
Relevant info	https://www.closer.ac.uk/funders/ https://www.closer.ac.uk/news-opinion/news/esrc-longitudinal-review-2017/ https://www.closer.ac.uk/news-opinion/news/closers-funding-extended-2021/ https://www.closer.ac.uk/how-to-access-the-data
Notes	https://www.cessda.eu/content/download/834/7776/file/CMM_ServiceProvidersMetadataPractices_2016.pdf https://www.icpsr.umich.edu/icpsrweb/

1.14 HathiTrust

“HathiTrust is a partnership of academic & research institutions, offering a collection of millions of titles digitized from libraries around the world”, <https://www.hathitrust.org/>

Administration

Funding sources	Industry, Institutional
Jurisdiction	USA
Leader	HathiTrust Digital Library is administered by the University of Michigan
Domains	Humanities, Arts, Social Science, Information Science & Technology, Computer science
Start date	2008
Timescale	11-15 Years
Investment	Co-investment through membership and investment through Google deal undisclosed.
Phases	Pre-operational (design/plan/construct), Operational (maintain/consolidate/grow)

Structure

Features	Cross-cutting, Regional impact, Institutional centres of excellence, National nodes, Global nodes
Type	Archive or repository service, Data service, Discovery platform, Capability building, Community building, Software service, Information sources, Research practices and methods, Research libraries, Digitisation service, Cloud storage/compute service, Research tools and platforms
Capabilities	Data curation, Data preservation, Data reuse, Data mining and analysis, Informatics and data modelling, Project support, Technology expertise, Domain expertise
Organisation	Distributed, Virtual, Shared facility
Governance	Coordinating entity (hub), Centralised (nexus), Governing body, Executive body, Operational function, Outreach and liaison function, Administration function, Technologies function, Informatics function, Expert advisors
User community	Global membership (commonly English speaking) and dominated by North American universities in Canada and the USA.
Institutional centres	Content contributors and partners as examples: University of California, University of Michigan (biggest suppliers of digitised volumes); University of Illinois (Urbana-Champaign), Harvard University, Indiana University, New York Public Library, Cornell University, Library of Congress, Princeton University, University of Texas,
Government links	Collaboration program around the Federal Documents program and Copyright review.

Key Aspects

Global links	Digitised works supplied to industry partners e.g EBSCO and OCLC that supply
--------------	--

	discovery metadata and full text content globally.
Differentiator	Collaboration around common research infrastructure requirements for access to scholarly resources for text and data mining, by academics and university libraries for the purposes of literature reviews and bibliometrics.
AU comparator	<p>Trove Digitisation partnership.</p> <p>Access to corpora: Trove Australia; Tinker (Humanities, Arts and Social Science Data Enhanced) & Alveo (Human and Communication Science) Virtual Labs, ARDC funded.</p> <p>Large aggregation collection APIs: Trove</p> <p>Institutional collection APIs: State Library of New South Wales, ACMI, Museum Victoria, National Museum of Australia, State Library of Queensland</p>

Data Infrastructure

Digitisation	<p>Partnership through Google Books to establish the digital library. More digitised works are being added to ongoing by partners.</p> <p>Statistics as at July 2019:</p> <p>17,044,213 total volumes</p> <p>8,257,117 book titles</p> <p>455,483 serial titles</p> <p>5,965,474,550 pages</p> <p>764 terabytes</p> <p>202 miles</p> <p>13,848 tons</p> <p>6,490,415 volumes(~38% of total) in the public domain</p>
Government services	Digital Public Library of America
Longitudinal	The HathiTrust Digital Library could be construed as a longitudinal dataset, in that it has digitised books published over multiple centuries, predominantly in the nineteenth century.

Research Leads

Trends	Text and data mining infrastructure arising through consortia, to meet a common research infrastructure requirement (across domains) for scholarly discovery phase (literature reviews), digital humanities, library and information science and computer science.
Further analysis	<p>Capacity for digitisation and text and data mining services to be shared across academic libraries, e.g. CADRE and California Digital Library.</p> <p>HathiTrust Research Center Advanced Collaborative Support Program is used as a means of addressing computational analysis challenges of the corpus. What role must computer science play in the development of large and complex unstructured and structured data architectures to support humanities versus cross-domain requirements?</p> <p>What research requirements drove the decision to deliver digital library material as datasets for "research purposes", and what factors relating to industry partnership and copyright have constrained data accessibility (for research).</p>

Additional Information

Relevant Info	<p>HathiTrust</p> <p>California Digital Library</p> <p>CADRE project</p> <p>Digital Public Library of America (DPLA)</p> <p>Trove digitisation partnership</p>
Notes	<p>Longevity of HathiTrust and partnership with Google to form the basis of the corpora in the digital library</p> <p>Google Books Library Project commenced in 2004 in collaboration with research libraries in the UK and the USA, ~30 million books were digitised. Predominantly English language ~4.6M and a long tail of language groups are represented</p> <p>NEH awards research funds to digital humanities researchers undertaking computational methods using the HathiTrust digital library</p> <p>Advanced Collaborative Support Program (HTRC)</p> <p>Internet Archive mass digitised books (2005)</p> <p>CADRE led by Indiana University, Network Science Institute</p> <p>CADRE co-funding: IMLS USD \$850,000, total project budget USD \$2M</p> <p>Corpora for CADRE from Web of Knowledge will include datasets ranging in timescale from 1 to 25 years.</p> <p>Big Ten Academic Alliance led project, includes MidWest universities, coordinated through University of Illinois, Urbana-Champaign</p> <p>Trove aggregated collections API (open data)</p> <p>https://help.nla.gov.au/trove/building-with-trove</p> <p>Example of state agency collection API (open data)</p> <p>Trove bulk data download</p> <p>Tinker virtual lab and reference datasets</p> <p>Alveo virtual lab and reference datasets</p> <p>Emergence of HathiTrust and Digital Public Library of America</p> <p>HathiTrust as content hub for DPLA</p>

References

User community	https://www.hathitrust.org/community
Leadership	https://www.hathitrust.org/governance
Institutional centres of competence	https://www.hathitrust.org/partnership https://www.hathitrust.org/community https://www.hathitrust.org/collaborative-programs https://www.hathitrust.org/hathitrust-celebrates-10-years-2008-2018
Government links	https://www.hathitrust.org/usgovdocs https://www.hathitrust.org/copyright-review
Global links	https://www.hathitrust.org/data
Differentiator	https://www.hathitrust.org/digital_library

	https://www.hathitrust.org/htrc
Comparator	https://help.nla.gov.au/trove/digitisation-partner-guidelines#anchor-0 https://tinker.edu.au/data/available-datasets/ https://austalk-query.apps.alveo.edu.au/
Digitisation	https://www.hathitrust.org/data https://www.hathitrust.org/visualizations_deposited_volumes_current https://www.hathitrust.org/datasets
Government services	https://dp.la/news/hathitrust-to-partner-with-dpla/
Longitudinal	https://www.hathitrust.org/visualizations_dates
Relevant info	https://www.hathitrust.org/ https://www.cdlib.org/ http://iuni.iu.edu/resources/cadre https://dp.la/
Notes	https://www.neh.gov/blog/neh-supported-digital-humanities-projects-pave-way-comprehensive-computational-analysis https://www.hathitrust.org/htrc_sp19acs_rfp https://www.sciencedirect.com/topics/computer-science/google-book-project https://blogs.libraries.indiana.edu/sbd-gateway/ https://news.iu.edu/stories/2018/10/iu/releases/18-shared-bigdata-gateway-for-research-networks.html?fbclid=IwAROpCD-goSa8AXLx1Fny_ai6fpT-RZCzltOLYxBQz57wy94YrHqSoVBRU98 https://help.nla.gov.au/trove/building-with-trove https://data.nsw.gov.au/data/dataset/ww1-diary-and-letter-transcripts https://help.nla.gov.au/trove/building-with-trove/direct-bulk-download https://tinker.edu.au/ http://alveo.edu.au/ http://www.infotoday.com/OnlineSearcher/Articles/Features/HathiTrust-and-Digital-Public-Library-of-America-as-the-future-88089.shtml https://www.hathitrust.org/blogs/perspectives-from-hathitrust/getting-to-5-million-hathitrust039s-collection-of-open-books

1.15 Country Profile: The Netherlands

The Netherlands: Humanities and Arts

Administration

Funding sources	National, H2020, Research council, Industry, Institutional, FP (EU Framework Programme), CIP (EU Competitiveness and Innovation Framework Programme)
Infrastructures	CLARIN, DARIAH, CLARIAH, Europeana, Time Machines, DASISH, KB Lab, eScience Center (clustered in association with CLARIAH-PLUS)
Leader	<p>CLARIAH-PLUS listed in national roadmap, runs from 2019 onward</p> <p>CLARIN-NL coordinated by Utrecht University, see governance.</p> <p>Positioning of CLARIAH-NL as part of NL national research infrastructure roadmap in the context of ESFRI roadmap and DARIAH (page 10 of Interim Evaluation Factbook 2017)</p> <p>CLARIN-NL mooted for the development of CLARIAH in 2011, the proposal was seen as too ambitious, and instead seed money provided in 2013, and a new proposal created in 2015 (CLARIAH-CORE) funded by NWO.</p> <p>CLARIAH-SEED 2012-2014 funded by NWO € 1.0 M</p> <p>CLARIAH-CORE 2015-2018 funded by NWO € 12.6 M</p> <p>CLARIAH-PLUS 2019-2024 funded by NWO € 13.8 M</p> <p>Total funding for CLARIAH from NWO is 24.4 M Euro ~ AUD 39.13 M</p>
Domains	Humanities, Arts, Social Science, Information Science & Technology, Heritage, Computer science
Start date	2015
Timescale	3-5 Years
Investment	\$21-50M
Phases	Pre-operational (design/plan/construct), Operational (maintain/consolidate/grow)

Structure

Features	Cross-cutting, Regional impact, Institutional centres of excellence, Government linkages, National nodes, Global nodes
Type	Archive or repository service, Data service, Discovery platform, Capability building, Community building, Software service, Information sources, Research practices and methods, Research libraries, HPC service, Digitisation service, Cloud storage/compute service, Local storage, Research tools and platforms
Capabilities	Data curation, Data preservation, Data reuse, Data mining and analysis, Informatics and data modelling, Project support, Technology expertise, Domain expertise
Organisation	Distributed, Virtual, Shared facility, Institutional facility
Governance	Coordinating entity (hub), Network (participating nodes), Decentralised (institutional), Governing body, Executive body, Operational group, Outreach and liaison group, Administration group, Technologies group, Informatics group, International advisory board
User community	CLARIAH-CORE targeted to: computers and the humanities (main field of research),

	software for humanities, databases for humanities, social and economic history, linguistics, media studies.
Institutional centres	<p>KB (national library) have involvement in Europeana projects, have developed services specifically for digital humanities research.</p> <p>CLARIAH provides datalegend (structured data service)</p> <p>Time Machines in Amsterdam and Utrecht as examples of institutional centres of competence and national networks required to enable humanities and arts research based on common data infrastructure, i.e. linked open data, technologies, and expertise, digitised historical heritage materials (2D and 3D).</p> <p>National networks include: DEN (Dutch national knowledge institute for culture and digitalisation and a national digital heritage strategy) and CLARIAH</p> <p>Local Time Machines also operate as part of a European wide Time Machine initiative</p> <p>CLARIAH centres in the Netherlands: DANS (Data Archiving and Networked Services), International Institute of Social History, Dutch Language Institute, National Library of the Netherlands, Huygens ING, Meertens Institute, Max-Planck Institute, Netherlands Institute for Sound and Vision</p> <p>Centre for Digital Humanities a partnership between: Amsterdam & VU universities, and KNAW (Royal Academy)</p>
Government links	<p>eHumanities.nl is the national platform and has partnerships with research platforms, universities, cultural institutions around expertise, research, technology, and collections.</p> <p>Partners: CLARIAH, CLICKNL, CWI - Research in Mathematics and Computer Science, DANS - Data Archiving and Networked Services, Delft University of Technology, Dutch Language Institute, Erasmus University Rotterdam, Fryske Akademy, HuygensING Institute, International Institute of Social History, KITLV, Koninklijke Bibliotheek, Leiden University Centre for Digital Humanities, Meertens Institute, National Demographic Research Institute of the Netherlands, Netherlands eScience Centre, Netherlands Institute for Sound and Vision, Netherlands Institute for War, Holocaust and Genocide Studies, Radboud University, University of Twente, Tilburg University, University of Amsterdam, University of Groningen, Utrecht University, Vrije Universiteit Amsterdam.</p>

Key Aspects

Global links	CLARIAH network, includes Austria and Germany (where the consolidation of infrastructures has also occurred) and connected into a national eHumanities platform.
Differentiator	The coordination and integration of research data and technologies with knowledge infrastructures across higher education and heritage to support humanities and arts research. Presumes the involvement of national data centres e.g. National Library and Netherlands Institute for Sound and Vision as a critical strategic partnership along with centres of competence in higher education e.g. Max Planck Institute for Psycholinguistics, Meertens Institute, Institute for Dutch Lexicology, Data Archiving and Networked Services (DANS) and Huygens ING.
AU comparator	PARADISEC, Australian National Corpus, AURIN, Tinker (HASS), Alveo and HuNi virtual labs, Trove Australia, UNSW Epic and Curtin VisLab, SLNSW DX lab, ACMI X

Data Infrastructure

Digitisation	<p>DEN - National Digitisation Strategy, Section 5: Who does what? Outlines the effort required at institutional and sector levels and cross-sector to support a well-functioning digital heritage infrastructure. (p11)</p> <p>See KB statement about involvement strategic partnerships in their strategic plan 2015-2018 with: KNAW, DEN, Europeana, CLARIAH, DARIAH, Nederlab and the Dutch Digital Heritage Network (with reference to a national strategy for digitisation as part of the digital agenda for the Netherlands, and CLICKNL for knowledge and innovation agenda - creative industries).</p> <p>KB (national library) is a partner in Impact Centre of Competence on digitisation. KB and university libraries provided DH Clinics (2017) to engage librarians and others in universities around digitisation, databases, and software skills. The presumption that major cultural heritage institutions will provide their collections as digitised resources, and referred to as "national data centres" as part of CLARIAH-CORE proposal.</p> <p>See Enumerate report on digitisation (2012)</p>
Government services	<p>Memory of the Netherlands with contribution of materials from 92 institutions</p> <p>Unclear who the national geo-spatial data service provider is</p>
Longitudinal	<p>Descriptive metadata and digital content will cover heritage collections in the documentary history timeline ~1000 years, archaeological timeline ~16,500 years and flora, faunal, and geoscience records ~4 billion years of earth history.</p>
Linkage	<p>Focus is on linked data for text based resources, as opposed to administrative data linkage. See CLARIAH datalegend service with tools to support researcher production of linked data.</p>

Research Leads

Trends	<p>Consolidation of EU investment in the full spectrum of infrastructure: data, compute, skills across different domains in humanities and arts, to establish a national competence, that leads in regional networks and has the capacity for global collaboration (and competition).</p>
Further analysis	<p>National investments into CLARIAH infrastructures in Netherlands, Austria and Germany (similarities and differences in approach to support humanities and arts and connections with heritage institutions). Relationship between CLARIAH infrastructures and SSHOC (as part of the general shift in data and technology intensive research to using cloud infrastructure and working in virtual environments).</p> <p>How key prior investments in research and data infrastructure is to enabling research into the digital transformation (of society) - see NWO 6M Euro investment into SSH research. Whether the investment in CLARIN and DARIAH preparatory phases (by NWO for the Netherlands) provided sufficient funds and foundations (2006-2009) for CLARIAH.</p> <p>Whether any funds have been provided by NWO or the relevant Ministry to support the collaboration and research driven digitisation agenda, and the results of the evaluation of CLARIAH in early 2019. NDE (National Digital Heritage) is an initiative of the Ministry of Education, Culture and Science in cooperation with the National Institute for Sound and Vision, the National Archives, the National Cultural</p>

	<p>Heritage Agency, the National Library and the Royal Academy of Arts and Sciences (KNAW).</p> <p>How has DANS involvement enabled the development of social science versus humanities and arts research infrastructures? What is the difference here between the role of ARDC in Australia, and what's proposed as a large scale platform for HASS (an NCRIS facility)?</p>
--	---

Additional Information

Relevant info	<p>eHumanities.nl governance represents the interests of four research areas: humanities, social science, and computer and information sciences</p> <p>NWO funding for CLARIAH 12M Euro (2015-2018) ~ 17.8M AUD</p> <p>CLARIAH in the Netherlands - analysis</p> <p>Europeana 2018 Annual Report. Europeana DSI 2016-2019 funding.</p> <p>Europeana-DSI collaboration with DARIAH and CLARIN, 2015-2016 € 8 900 000</p> <p>Of the 10 CEF projects are all targeted at collection accessibility, enabling research and education through digitisation, curation, presentation and discovery. EU contribution (2016-2017) is with ~75% funds towards the overall budget: € 3.7M</p> <p>Example contribution by KB to meet research needs is the digitisation of early books and newspapers in alignment with Europeana initiatives.</p> <p>The Databank Digital Newspapers project was setup with funding from NWO (Dutch Research Council) - see KB 2011 Annual Report. See IFLA paper on the collaborative strategies for constructing digital libraries and co-funding arrangements</p> <p>CLARIN-NL Tools and Services and Centres of Competence</p> <p>With CLARIN seal: Data Archiving and Networked Services (C), Huygens ING (C), Meertens Instituut/HuC (B), MPI for Psycholinguistics (B)</p> <p>Without CLARIN seal: Beeld en Geluid (T), Centre for Language and Speech Technology (C & T), Utrecht Institute of Linguistics OTS (T)</p> <p>C = metadata, K = knowledge, T = trust, E = external</p> <p>CLARIN-NL funded by NWO and launched in 2009 (as part of the preparatory phase of CLARIN-EU) and collaborates with DARIAH and CLARIAH</p> <p>See CLARIN-NL 2014 Annual Report, CLARIN-NL created in three phases: preparation (2009-2010), construction and initial operation (2011-2012) and operation and further development (2013-2014). The budget for the EU collaboration is: EU-line is 1.35M€, for the NL-line 7.65M€. National investment is 7 times that of regional investment.</p> <p>CLARIAH-CORE proposal outlines target users for the infrastructure in 2017 interim evaluation factbook</p>
Notes	<p>CLARIAH-AT - there is a CLARIAH in Austria launched in 2017.</p> <p>Digital Humanities Austria</p> <p>CLARIAH-DE being launched in Germany.</p> <p>NWO funding SSH research into the digital transformation of society 6M Euro</p> <p>Ahead on trans-national studies on landscape with diverse data sources (national spatial data infrastructure: archaeological, historical, architectural, geological), see NWO investments in heritage and archaeological research (on the basis that data infrastructure is available), known as SDI (spatial data infrastructure)</p>

	<p>KB entered into a partnership with Proquest for the digitisation of early European</p> <p>Why did DARIAH-NL fail to get funding in 2008? Joint funding of CLARIN-NL & DARIAH-NL proposal in 2011 as CLARIAH (complementary technical and knowledge infrastructures). CLARIAH-SEED 2012-2014 funding provided to help progress the proposal.</p> <p>Relevant LIEF infrastructures as examples of project investment: AusStage, AustLit, AustLII, DAAO.</p>
--	--

References

Leadership	<p>https://www.nwo.nl/en/documents/nwo/permanent-commission/roadmap-large-scale-scientific-infrastructure</p> <p>https://www.clariah.nl/</p> <p>https://www.clariah.nl/over/bestanden/categories/send/0-/168-interim-evaluation-factbook</p> <p>https://www.clariah.nl/en/dissemination/clariah-core-book/clariah-a-digital-research-infrastructure-for-humanities-researchers-in-the-netherlands</p> <p>https://clariah.nl/boekje/clariah-digitaal.pdf</p> <p>https://www.clariah.nl/over/geschiedenis/zaaigeld</p> <p>https://www.clariah.nl/en/projects/finished/seed-money/seedmoney</p> <p>https://www.clariah.nl/en/new/news/clariah-plus-granted</p> <p>http://www.clarin.nl/sites/default/files/CLARIN%20NL%20Final%20Report%202016-06-08%20FINAL.pdf</p> <p>https://www.clariah.nl/en/about/organisation</p>
Institutional centres of competence	<p>https://www.kb.nl/en/organisation/research-expertise/digital-humanities</p> <p>http://www.datalegend.net/</p> <p>https://www.den.nl/over-ons/english</p> <p>https://amsterdamtimemachine.nl/category/data/</p> <p>http://utrechttimemachine.nl/</p> <p>https://amsterdamtimemachine.nl/category/consortium/institutional-partners/</p> <p>https://www.timemachine.eu/time-machines/</p> <p>https://www.clariah.nl/en/about/organisation/centres</p> <p>http://www.centrefordigitalhumanities.nl/</p>
Government links	<p>https://www.ehumanities.nl/</p>
Further analysis	<p>https://www.clariah.nl/en/events/toog-day-2018</p> <p>https://dans.knaw.nl/en/about/organisation-and-policy/collaboration</p>
Digitisation	<p>https://www.den.nl/uploads/5c9b818974b6fbf9d4f768cd0fe96f8d14e3d504324a7.pdf</p> <p>https://www.government.nl/documents/reports/2017/04/11/digital-agenda-for-the-netherlands-innovation-trust-acceleration</p> <p>https://www.netwerkdigitaalerfgoed.nl/en/</p> <p>https://www.clicknl.nl/en/about-clicknl/</p> <p>https://pro.europeana.eu/post/fostering-digital-humanities-projects-in-the-dutch-heritage-sector</p> <p>http://cultivate-cier.nl/wp-content/uploads/2012/03/ENUMERATE-Digitisation-</p>

	Survey-2012-1.pdf
Government services	https://www.geheugenvannederland.nl/en https://www.geheugenvannederland.nl/en/geheugen/browse/institutions
Relevant info	https://www.ehumanities.nl/governance/ https://www.clarin.eu/news/dutch-funding-clarin-and-dariah-infrastructures https://www.aclweb.org/anthology/L16-1394 https://pro.europeana.eu/files/Europeana_Professional/Publications/Europeana%20Annual%20Report%20and%20Accounts%202018.pdf https://www.dariah.eu/activities/projects-and-affiliations/dariah-in-past-eu-projects/ https://ec.europa.eu/inea/en/connecting-europe-facility/cef-telecom/projects-by-dsi/europeana https://www.kb.nl/en/organisation/research-expertise/digitization-projects-in-the-kb/databank-of-digital-daily-newspapers/the-ddd-project https://www.kb.nl/en/organisation/research-expertise/digitization-projects-in-the-kb/databank-of-digital-daily-newspapers/technical-information https://www.kb.nl/sites/default/files/docs/kbannualreport2011.pdf https://archive.ifla.org/IV/ifla64/001-98e.htm https://portal.clarin.nl/clariah-tools-fs?f%5B0%5D=field_clarin_national_project%3A417 https://centres.clarin.eu/ https://tla.mpi.nl/projects/past-projects/clarin-nl/ https://pdfs.semanticscholar.org/ae25/3b3bf80c57825d51b7a4e14398127b1a5e27.pdf https://www.clarin.nl/sites/default/files/448_Paper.pdf https://www.clariah.nl/en/dissemination/clariah-core-book/clariah-a-digital-research-infrastructure-for-humanities-researchers-in-the-netherlands http://www.clarin.nl/sites/default/files/CLARIN%20Annual%20Report%202014%20FINAL.pdf http://www.clarin.nl/sites/default/files/CLARIN-NL%20Multiyear%20Programme%20090409-2.pdf https://www.clariah.nl/over/bestanden/categories/send/0-/168-interim-evaluation-factbook
Notes	https://www.clarin.eu/blog/tour-de-clarin-austria http://digital-humanities.at/ https://www.dariah.eu/2019/04/30/successful-launch-of-clariah-de/ https://www.nwo.nl/en/news-and-events/news/2018/12/6-million-euros-for-digitalisation-in-the-social-sciences-and-humanities.html https://www.nwo.nl/en/research-and-results/research-projects/i/55/5955.html https://www.nwo.nl/en/news-and-events/news/2018/09/4.5-million-euros-for-international-heritage-research.html https://spinlab.vu.nl/wp-content/uploads/2016/09/The_construction_of_an_SDI_for_cultural_heritage.pdf https://www.kb.nl/en/organisation/research-expertise/digitization-projects-in-the-kb/early-european-books

The Netherlands – Social Sciences

Administration

Funding sources	National, H2020, Research council, Institutional, FP (EU Framework Programme)
Infrastructures	ESS, CESSDA, GGP, EVS, SHARE, DASISH, SERISS (clustered in association with ODISSEI)
Leader	<p>ODISSEI (a data facility, observatory, laboratory and hub) proposal (2016-2020) led by CBS (national statistics agency), SURFSara (NREN), CentERdata (data archive, institutionally based), NWO (national research infrastructure funder) - with 30 participating organisations.</p> <p>ODISSEI referenced in 2016-2020 national roadmap for large scale research infrastructure (NWO). In that roadmap ODISSEI comprises the ESFRI facilities SHARE, ESS, and the forthcoming GPP (all multi-national longitudinal surveys).</p>
Domains	Social Science
Start date	2016
Timescale	3-5 Years
Investment	\$21-50M
Phases	Pre-operational (design/plan/construct), Operational (maintain/consolidate/grow)

Structure

Features	Cross-cutting, Regional impact, Institutional centres of excellence, Government linkages, National nodes, Global nodes
Type	Archive or repository service, Data service, Discovery platform, Capability building, Community building, Software service, Information sources, Research practices and methods, Research libraries, HPC service, Cloud storage/compute service, Local storage, Research tools and platforms
Capabilities	Data curation, Data preservation, Data reuse, Informatics and data modelling, Survey design and instrument, Data linkage, Project support, Technology expertise, Domain expertise
Organisation	Distributed, Virtual, Shared facility, Institutional facility
Governance	Coordinating entity (hub), Network (participating nodes), Decentralised (institutional), Governing body, Operational group
User community	LISS references scientific, policy and societal studies (within social sciences) and multidisciplinary use of data. ODISSEI secure access to microdata for 700 researchers
Institutional centres	<p>ODISSEI membership - mostly universities and research institutes</p> <p>ODISSEI partnership - Statistics Netherlands, CentERdata, Data Archiving and Networked Services (DANS), SURFSara, eScienceCenter</p> <p>CentERdata Institute for Data Collection and Research, Tilburg University (partner in European Values Study)</p> <p>CentERdata has expertise in machine learning and text and data mining</p> <p>LISS is the data archive for EVS and core element of MESS (multidisciplinary facility</p>

	<p>for measurement and experimentation in the social sciences). 9 years worth of data collected supports: scientific, policy and societal studies (nationally and internationally).</p> <p>DANS provider of DataverseNL (repository) and NARCIS (discovery service)</p> <p>SHARE country team for the Netherlands is based at University of Utrecht</p> <p>NIDI (Netherlands Interdisciplinary Demographic Institute) - a research institute of KNAW is the Dutch partner in GGP</p>
Government links	<p>Collaboration via ODISSEI with CBS (Statistics Netherlands) and SURF (Australian equivalent to AARNet and AAF) on secure linking and analysis facility</p> <p>LISS (Longitudinal Internet Studies for the Social Sciences) panel was developed in partnership via CentERdata and Statistics Netherlands</p> <p>LISS collaboration with Netherlands Institute for Social Research</p> <p>See CentERdata list of partners and clients across national research infrastructure, universities and government, pan-national and international e.g. GESIS, University of Michigan</p>

Key Aspects

Global links	<p>NWO international collaboration with CESSDA, ESS, SHARE & MESS as part of the "internationalisation of social sciences"</p> <p>International Social Survey programme - DANS partnership with GESIS Leibniz Institute for the Social Sciences, Cologne, Germany. Australia is a member of this programme via the Australian Demographic and Social Research Institute (ANU).</p> <p>DANS operates as the national node for CESSDA</p> <p>ESS contact points in the Netherlands are from Radboud University Nijmegen, Department of Sociology</p> <p>SHARE data is distributed via CentERdata, Tilburg University</p> <p>DANS maintains a key position in the ecosystem e.g. the national membership to (international) ICPSR and the secretariat for (national) NPSO (Dutch Platform for Survey Research)</p>
Differentiator	<p>Cross-sector stakeholder investment around establishing large scale infrastructure building upon existing institution and national capacity in the survey and collection of social science data, and as a national node for global collaboration. Building on institutional competencies and long term commitments in pan-national and global initiatives in social surveys and data archiving.</p>
AU comparator	<p>Australian Data Archive, Population Health Research Network, NSW Centre for Health Record Linkage (ChReL), National Computational Infrastructure, AARNet, Australian Institute of Health and Welfare, Department of Social Services (HILDA survey etc), Australian Bureau of Statistics (statistics, secure environment and MADIP).</p>

Data Infrastructure

Digitisation	<p>CEDAR (2012-2016) as an example of data arising from the digitisation of historic census (1795-1971)</p> <p>CEDAR partners are International Institute of Social History (IISH) in KNAW Humanities Cluster, and DANS, this work crosses over into CLARIAH (around access</p>
--------------	---

	to data locked up in documentary heritage collections. CEDAR has generated linked open data and an a data service (API)
Government services	CB (Netherlands Statistics) provides custom, microdata, big data and open data services
Longitudinal	See ESS, EVS, SHARE, GPP as as examples of longitudinal datasets that involve national and global collaborations (with Dutch participation). Mix of national, regional and global longitudinal survey data
Linkage	Data linkage (using administrative data) enabled through remote access to Statistics Netherlands (see: Experimental and Longitudinal Data for Scientific and Policy Research:Open Access to Data Collected in the Longitudinal Internet Studies for the Social Sciences (LISS) Panel by Marcel Das and Marike Knoef, 2019) LISS data linkages with other agencies: Dutch National Institute for Public Health and the Environment and Royal Netherlands Meteorological Institute.

Research Leads

Trends	The collaboration around surveys commenced very early on within Europe and in global spheres (2002). Stakeholder relationships at the national level were formed through collaborative survey data collection (longitudinal) and then in the establishment of national data infrastructure and more recently the large-scale platform (ODISSEI). By example: the European Values Survey has been running since 1981. https://europeanvaluesstudy.eu/about-evs/history/
Further analysis	What role has EU funding through CESSDA, ESS, SERISS, SHARE had in establishing and boosting existing institutional infrastructure and competencies? What effect has prior institutional investment over several decades had on establishing national competency and infrastructure? See global collaboration around EVS. How has the collaboration between DANS and GESIS (Germany) around managing social science data informed the development of the ODISSEI infrastructure? How has the collaboration between DANS, 4TU.Federation (TU Delft Library) and SURFSara around long term data archiving (Research Data Netherlands) informed and fed into the ODISSEI proposal? What global collaborations around infrastructure are emerging through joint initiatives such as MESS and EVS with the US and Germany? How has DANS involvement enabled the development of social science versus humanities and arts research infrastructures? What is the difference here between the role of ARDC in Australia, and what's proposed as a large scale platform for HASS (an NCRIS facility)?

Additional Information

Relevant info	ODISSEI on national roadmap in 2016 (large scale research infrastructure) - bridging financing of € 0.5 M ODISSEI 2019 Roadmap requesting € 18 M from NWO Total funding € 18 M = AUD 29.59 M from NWO Does not include funding from H2020 via CESSDA, ESS etc Main activities of ODISSEI are: microdata services (secure), national lab for high
---------------	--

	<p>quality methodological and scientific innovation, international panels and fieldwork around SHARE, ESS, EVS and GGP, promoting open data, HPC access for survey and register data analysis, data standards and community development (national and international) - see</p> <p>MESS project funded by NWO 2006-2014, operational 2014 onward</p> <p>https://www.lissdata.nl/organization</p> <p>MESS project involved Tilburg University, CentERdata, and US partners (RAND, UC Berkeley, University of Michigan)</p> <p>LISS (Longitudinal Internet Studies for the Social Sciences) panel was developed in partnership via CentERdata and Statistics Netherlands</p> <p>ESSNeth (alternative name for Netherlands node of ESS)</p> <p>DANS key partner in CESSDA project activity</p> <p>MADIP Multi-Agency Data Integration Project (Australian Bureau of Statistics)</p>
Notes	<p>SERISS project has key people from the Netherlands listed from: Tilburg University, CentERdata, NIDI (national demographic research institute), and AIAS (Amsterdam Institute for Advanced Labour Studies) University of Amsterdam</p> <p>European Values Study (EVS) collaboration between the Netherlands (CentERdata, Tilburg University) and Germany (GESIS), informed the World Values Survey and a collaboration formed with University of Michigan (ICSPR)</p> <p>Generations and Gender Program</p> <p>DANS is supported by KNAW (Royal Netherlands Academy of Arts and Sciences) and NWO (Netherlands Organisation for Scientific Research)</p> <p>Relevant LIEF funded infrastructures as examples as project investment: QUT Digital Media Research Centre; UMELB Founders and Survivors; Griffith Prosecutions Project</p> <p>NWO 2016 Annual Report (English) references a request to support more women and HASS research in Gravitation programme.</p> <p>NWO has social sciences and humanities division</p>

References

Leadership	<p>https://odissei-data.nl/en/</p> <p>https://www.nwo.nl/en/documents/nwo/permanent-commission/roadmap-large-scale-scientific-infrastructure</p> <p>https://odissei-data.nl/en/en-odissei/</p>
Institutional centres of competence	<p>https://odissei-data.nl/en/en-odissei/member-organisations/</p> <p>https://odissei-data.nl/en/en-odissei/partners/</p> <p>https://www.centerdata.nl/en/node/4571</p> <p>https://www.centerdata.nl/en/projects-by-centerdata/machine-learning-and-deep-learning & https://www.centerdata.nl/en/projects-by-centerdata/text-analytics-data-mining-and-semantic-analyses</p> <p>https://www.lissdata.nl/</p> <p>https://www.dataarchive.lissdata.nl/study_units/view/31</p> <p>https://www.lissdata.nl/sites/default/files/bestanden/mess_proposal.pdf</p>

	https://dataverse.nl/ https://www.narcis.nl/?Language=en http://www.share-project.org/organisation/share-country-teams/netherlands.html https://www.ggp-i.org/about/organisation/
Government links	https://www.surf.nl/en/news/securely-linking-and-analysing-scientific-data-with-odissei-data-facility https://www.centerdata.nl/sites/default/files/experimental_and_longitudinal_data_for_scientific_and_policy_research.pdf https://www.scp.nl/english/ https://www.centerdata.nl/en/about-us/partners-and-clients
Global links	https://www.nwo.nl/en/about-nwo/organisation/nwo-domains/magw/internationalisation+of+NWO+Social+Sciences https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:39913 http://www.issp.org/menu-top/home/ http://www.issp.org/members/member-states/australia/ https://www.cessda.eu/About/Consortium/CESSDA-Countries/CESSDA-Members/Netherlands https://www.europeansocialsurvey.org/about/country/netherlands/ https://www.eui.eu/Research/Library/ResearchGuides/Economics/Statistics/DataPortal/SHARE https://dans.knaw.nl/en/about/organisation-and-policy/collaboration/icpsr
Further analysis	https://www.gesis.org/en/en/home/ https://researchdata.nl/en/ https://dans.knaw.nl/en/about/organisation-and-policy/collaboration
Trends	https://europeanvaluesstudy.eu/about-evs/history/
Digitisation	https://research.vu.nl/ws/portalfiles/portal/1161738/Abstract+A.pdf https://pure.knaw.nl/portal/en/publications/cedar-harmonization-of-historical-dutch-census-data(7d733d05-6e7f-4577-bc01-a6b6ba77060c).html https://research.vu.nl/en/publications/cedar-the-dutch-historical-censuses-as-linked-open-data https://research.vu.nl/en/publications/cedar-the-dutch-historical-censuses-as-linked-open-data http://lod.cedar-project.nl/cedar/data.html
Government services	https://www.cbs.nl/en-gb/our-services
Longitudinal	https://odissei-data.nl/en/en-odissei/surveys-panels-en/
Relevant info	https://odissei-data.nl/en/en-odissei/roadmap/ https://odissei-data.nl/en/2019/06/roadmap-proposal-submitted/ https://www.ru.nl/nsm/imr/vm/imr-academy-seminars/@1213997/imr-academy-seminar-odissei/ https://www.lissdata.nl/organization

	https://www.nwo.nl/en/research-and-results/research-projects/i/45/2445.html https://www.centerdata.nl/sites/default/files/experimental_and_longitudinal_data_for_scientific_and_policy_research.pdf https://www.europeansocialsurvey.org/data/country.html?c=netherlands http://www.onderzoeksfaciliteiten.nl/facility/european-social-survey-netherlands https://www.cessda.eu/News-Events/News/CESSDA/Open-if-possible-protected-if-needed-Research-data-via-DANS https://www.abs.gov.au/ausstats/abs@.nsf/mf/1700.0 https://csrc.cass.anu.edu.au/research/publications/introducing-longitudinal-madip-and-its-role-understanding-income-dynamics
Notes	https://seriss.eu/who-is-involved/key-people/ https://europeanvaluesstudy.eu/ https://www.gesis.org/en/services/data-analysis/international-survey-programs/european-values-study/data-access/ https://www.ggp-i.org/ https://dans.knaw.nl/en https://www.knaw.nl/en/homepage https://www.nwo.nl/en https://research.qut.edu.au/dmrc/ http://foundersandsurvivors.org/ Prosecutions Project https://prosecutionproject.griffith.edu.au/ https://www.nwo.nl/en/about-nwo/organisation/nwo-domains/sgw

1.16 Additional Infrastructures

The project also developed a list of additional infrastructures of note as follows:

UK Data Archive – maintained by the University of Essex to meet the needs of social science researchers.

UK Archaeology Data Service – maintained by the University of York to meet the needs of archaeology, classics and ancient history researchers.

DiSSCO – Distributed System of Scientific Collections maintained by Naturalis Biodiversity Centre (Netherlands) to support historical science research based on heritage natural history collections.

The Portage Network – a network of social science data archives maintained by university libraries in Canada (CARL) to support social science research

Pelagios – a network of experts and shared research infrastructure enabling ancient history research and maintained by multiple partners, original participants: the Open University and University of Southampton (UK) and the Austrian Institute of Technology.

Perseus Digital Library – a digital library of digitised classical works and objects maintained by Tufts University

DANS – Data Archiving and Networked Services data archiving service provider for researchers in the Netherlands.

Generations and Gender Program – a pan European longitudinal survey.

European Values Survey – a pan European longitudinal survey.

GESIS – Leibniz Institute for the Social Sciences.

SSHOC – Social Sciences and Humanities Open Cloud as part of the European Open Cloud program to support research.

Sami Archives: a collaboration between multiple universities and heritage organisations to make heritage resources available, led by the University of Lapland and associated with the EU Artic Information Centre. <https://digisamiarchives.com/>

Cultural Heritage Cluster: a collaboration between DeIC (Danish e-Infrastructure Cooperation) and the Royal Danish Library to provide access to radio, TV, websites and historical newspapers as data sources. “The agreement between DeIC and the Royal Danish Library has a total financial framework of DKK 7.2 million.” ~AUD \$1.54 million, <http://en.statsbiblioteket.dk/kulturarvscluster/>

Tilltal: a collaboration between Digisam (Swedish National Archives), The Institute for Language and Folklore, and the Royal Institute of Technology (KTH) a CLARIN-SWE node, to make historic speech recordings more accessible for research in language and folklore. Funded by Riksbankens Jubileumsfond for three years 2017-2020 at ~€1M. ~AUD\$1.59 million.
https://sweclarin.se/sites/sweclarin.se/files/event_atachements/berg-et-al-Swe-Clarín_2016_paper_7.pdf

Venice Time Machine: a collaboration between Ca’Foscari (University of Venice) and EPFL (Switzerland), and the State Archives of Venice, Marchiana Library, Istituto Veneto, and the Cini Foundation 2012-2016. Funding through H2020 and Europeana (ICARUS).
<https://ec.europa.eu/futurium/en/content/venice-time-machine-flagship>

Germany

Germany has co-invested in major programs in the financial reporting of the 2013 national roadmap¹. However, the figures provided are patchy, and the financial data inconsistently recorded i.e. a mix of

¹ Roadmap for research infrastructures: a pilot project of the Federal Ministry of Education and Research (BMBF), 2013 https://www.bmbf.de/upload_filestore/pub/roadmap_research_infrastructures.pdf last accessed 15 August 2019

overall and per annum figures. Set up costs (in millions in Euros) of the research infrastructures are listed on page 40 (effective 2013). Three columns list: Construction costs for overall project; German share; BMBF funds (German federal ministry of education and research) and 27 infrastructure programs of various types are listed. Five of the pan-national EU funded programs are listed as follows:

Construction costs for overall project €	German share €	BMBF funds €
CESSDA: 9.5M setup	Not specified	n/a -- Leibnitz Institute budget for coordination costs not specified
CLARIN: 104M setup	German share 14M	BMBF funds 14M
DARIAH: 20M setup	German share 10M	BMBF funds 10M
ESS-Social: 2.2M pa	German share 0.4M pa	n/a -- Leibnitz Institute budget for coordination costs not specified
SHARE: 10M pa	German share 2M pa	BMBF: 2M pa 2014, 0.7M pa 2018

EU pan-national funding operates in effect as de facto national funding in Germany – see “German share”. BMBF and Leibnitz Institute funding is very clearly targeted (as co-investment) to German HASS research infrastructure. There is some co-investment on a 1:1 level (matching national funds to EU subsidy), but it is not clear that there is a standard ratio applied across all programs. It is however an indicator that matching funding has occurred during and after the EU pan-national programs commenced.

All other programs list the following funding sources: BMBF only (x13), unspecified project and institutional funding in addition to BMBF(x2), unspecified (x3), or Helmholtz Association (x4).

EU subsidy percentages and national co-investment ratios in Germany in 2013:

- > CESSDA: -, -
- > CLARIN: 13% of EU total, 1:1 matched investment from BMBF funds
- > DARIAH: 50% of EU total, 1:1 matched investment from BMBF funds
- > ESS-Social: 18% of EU total, -
- > SHARE: 25% of EU total, 1:1 and 1:3 matched investments from BMBF funds
- > Germany’s various roles in supporting these pan-national infrastructures reveals early involvement, research strengths, and notable institutional representation (GESIS for its reputation in social science research and data curation and participation in CESSDA, ESS and SHARE, and six knowledge centres for language research and 19 partners for humanities and arts) and participation in DARIAH.
- > CESSDA: GESIS (Leibnitz Institute for the Social Sciences) in Germany is a member.
- > CLARIN: Germany certified as a knowledge centre (distributed across six universities).
- > DARIAH: coordinating institution is the Göttingen State and University Library on behalf of BMBF and 19 research institutions, libraries, and societies are listed as partner institutions.
- > ESS-Social: GESIS (Leibnitz Institute for the Social Sciences) in Germany is a partner and has participated in every round of surveying from 2002 onward.
- > SHARE: Germany is the coordinating country for the pan-national program, BMBF representative is the chair of the SHARE-ERIC council.

Appendix B – Checklist / Evaluation Criteria

1. What is the capability being sought?
2. What gap does this capability address?
3. Is this capability being sought to achieve parity (with other nations) or to strive for international leadership?
4. Who would be the Australian entity that would manage the relationship and deliver the capability?
5. Does this capability exist in any form in Australia? If so, who provides it, and will this capability grow or replace it?
6. Are the capabilities being sought unique? If not, what are the other options? Are there commercial alternatives?
7. What government policy/program(s) is this capability directly or indirectly supported by or relevant to?
8. Who is the entity offering the capability?
9. How much does the capability cost?
10. Is the offering entity's governance structure appropriate for the Australian entity and the Australian HASS environment?
 - a. Predominantly public funded research
 - b. Strong community of public funded collecting institutions
 - c. Federated state governments
 - d. Strong use of Federated Identity (AAF)
 - e. Will support unique needs of Australian HASS sector.
11. What is the sustainability of the offering entity?
12. Does the entity have the international scale to support the offer?
13. How will researchers be made aware of this capability and be provided with the skills to take advantage of it?
14. How many disciplines and researchers will this capability be relevant to?
15. Is the capability based on international standards?
16. Does the offering entity have other international relationships? With any other entities in Australia?