

# **Blueprint for the European Research Observatory for the Humanities and Social Sciences – EROHS**

Report Compiled for the European Strategy Forum for Research Infrastructure  
(ESFRI) by the Ad Hoc Working Group on Research Infrastructure in the  
Humanities and Social Sciences (RISSH), May 2004

# Content

1. Executive Summary	4
2. Preface	6
3. Setting the Context	8
4. The European Strategy	14

**Blueprint for the European Research Observatory  
for the Humanities and Social Sciences – EROHS**

Published by the Danish Research Council for the  
Humanities and Danish Social Science Research  
Council, The Danish Research Agency, Artillerivej 88,  
DK-2300 Copenhagen S, Tel. (+45) 3544 6200,  
Fax (+45) 3544 6201, [www.forsk.dk](http://www.forsk.dk), [forsk@forsk.dk](mailto:forsk@forsk.dk)

Graphic design: Marianne Dunker

Printed by: Clichéfa Tryk AS

Printed number: 700

ISBN: 87-90201-72-8

## 1. Executive summery

The humanities and social sciences are hampered by a raft of problems caused by the current state of research infrastructures at European level. This has severe consequences for the possibilities of conducting comparative European empirical research. Against this background a European strategy is proposed that addresses the current problems.

In outline, the central proposal is to establish a European Research Observatory for the Humanities and Social Sciences (EROHS), which will build upon existing resources, and both actively and systematically promote synergy and coherency. EROHS will be guided by four main principles:

- The facilitation of access to and sharing of existing European and national data, thereby more efficiently and effectively linking data resources already available.
- The development of improved standards and documentation relating to existing European and national data in order to enhance the scientific quality of data and their potential for interoperability.
- The generation of new and genuinely European data. This will involve both the collection of new data and the digitalisation of materials not currently computerised.
- The provision of research training programmes for the next generation of researchers.

EROHS promises to strengthen interdisciplinary and cross-border collaboration and comparative research on a European dimension. Further, it will enhance the building of research infrastructure capacity in today's the less

resourced European countries. Finally, it will increase the opportunities for improving knowledge on social processes and thus holds great potential in terms of advising European and national policy-makers on how to manage the challenges currently being faced by the societies of Europe.

## 2. Preface

In 2002 the EU Member States established the European Strategy Forum for Research Infrastructures (ESFRI), following an initiative from the European Commissioner for Research, Philippe Busquin. ESFRI brings together representatives appointed by the research ministers of member states with the objective to:

- Support a coherent and strategy-led approach to policy making on research infrastructures in Europe;
- Facilitate multilateral initiatives leading to a better development, construction and use of research infrastructures in Europe.

As one of its main working instruments ESFRI has the capacity to set up ad hoc working groups formed by participants from a given scientific community for in depth exploration of specific topics. At ESFRI's meeting in July 2003 it was decided to set up such a working group on the humanities and social sciences. This has subsequently functioned under the acronym RISSH (Research Infrastructures in Social Sciences and Humanities).

The broad remit of RISSH has been to identify the precise strategic issues that need to be tackled at European level within the humanities and social sciences. At the end of May 2004 the working group forwarded this report to ESFRI.

With the report RISSH enters contested territory. Many documents are currently being circulated which touch upon the significance of research infrastructures in general, on the position of humanities and social sciences in the European Research Area, or on the two subjects combined. The group

is pleased to note that its thoughts and recommendations are generally in line with the reports from the European Research Advisory Board (EURAB) and the European Commission's Communication on Europe and Basic Research.

### 3. Setting the context:

#### The need for a European strategy

The humanities and social sciences work to improve understanding and knowledge by focusing on the dynamics and social processes behind the societal, economic and cultural developments of society. In so doing, they contribute to contemporary and future societies. Research is conducted on the basis of many different units – including systems of welfare, markets, organisations, value systems, policies, knowledge, systems, political systems, social and ethnic groups and languages – and is conducted at multiple levels of analysis – from small localities to global systems.

Evidently, these themes and issues cover a variety of very important facets of modern society, and as such the humanities and social sciences also offer the promise of addressing a whole range of pertinent research questions directly related to the challenges facing Europe of today and tomorrow. Thus the humanities and social sciences actively contribute to our understanding of the profound social, political and economic changes occurring in Europe as well as the process of European cohesion.

In effect the disciplines make important contributions to attaining the strategic goals set up by the European Council in the Lisbon declaration, to strengthen employment, modernise social welfare and education systems, and to secure economic reform and social cohesion as part of a knowledge-based economy.

Indeed, the humanities and social sciences play a significant role in serving this need.

**3.1 THE PROBLEMS WITH DATA AND INFRASTRUCTURE** For the humanities and social sciences data are the fuel of the empirical research enterprise – just as they are for the natural sciences. All scientists need scientifically driven data that are accurate if they are to understand, improve and critically test and validate empirical knowledge about a given subject.

Within the scientific fields of the humanities and social sciences data consist of human artefacts in the broadest sense of the term. As in other sciences, data are provided through research infrastructures. The humanities and social sciences are thus currently not without research infrastructures. Indeed, the disciplines are supported by a wide array of very different research infrastructures, such as cross-sectional or cohort databases based on surveys or administrative registers, museum collections, libraries, linguistic corpora and citation indexes. Such research infrastructures also come in all shapes and sizes and on different geographic levels – local, national, regional, European, global – and can be both private and public, large and small, systematic and ad hoc.

In fact the last five to ten years has seen an upsurge in the development of the collection of instruments available within the humanities and social sciences at European level – some partly financed through the framework programmes. Among the most notable are the European Value Study, the International Social Survey Programme, the Luxembourg Income and Employment Study, European Social Survey (ESS), the Text Encoding Initiative, European Advisory Group on Language Engineering Standards. Collaboration between data archives in Europe has demonstrated the value of common accessible infrastructures for several decades. This goes especially for the Council of European Social Science Data Archives (CESSDA). To this can be added a number of outstanding national research infrastructures that have been established via specific research programmes and projects. Finally, the official statistical systems, which provide many scientists with relevant data, can also be considered part of the existing infrastructure.

Thus Europe has the potential to become a natural laboratory for the

humanities and social sciences, providing data with an optimal combination of diversity and homogeneity – and certainly much more than is the case within each nation-state. Europe could be constituting a world leading “best case” for research in the humanities and social sciences. Equally both disciplines could further thrive on the massive wealth of data on demographic, societal, constitutional, institutional, political, legal, economical, cultural, linguistic, religious, and historical variations in Europe.

However, the reality presently facing European researchers within the humanities and social sciences is somewhat bleaker. Despite the positive developments in research infrastructure within the last decade there are severe systematic shortcomings in European research infrastructures, which leave the humanities and social sciences under-equipped and failing to reach their full potential. In general terms, and irrespective of specific fields and disciplines, the state of European research infrastructure is hampered by the following problems, which can be categorized into three groups.

**3.1.1 SHORT ON COHERENCY AND FUNDING** Traditionally, data within the humanities and social sciences have primarily been collected and stored within the confines of the nation-state and often through specific isolated research projects with little or no follow-up. Alternatively, data are gathered through the agency of international administrative bodies with little focus on the needs of researchers. The absence of coordination at a European level leads to suboptimality and even duplication of efforts and incommensurable local solutions.

The funding of research infrastructures is volatile, often only secured year by year. Funding is most often secured nationally or regionally, but research infrastructures at European level are beyond the capabilities of single countries. Indeed, the European Union (EU) increasingly finances research infrastructures – and especially access – but funding is allocated on the level of specific research projects, consequently only adding to a patchwork of

research infrastructures. The lack of systematic, integrated long-term funding at European level is evident.

**3.1.2 SHORT ON ACCESSIBILITY** Data and research infrastructures are generally a public good, which calls for public investment. Specifically, this is more true within the humanities and social sciences than in other disciplines owing to a perceived general absence of commercial value for the use of data. However, whilst data are not a scarce resource in Europe, they are not as translational available for secondary analysis as they could be. Access is restricted due to many reasons: legality, privacy, confidentiality, ownership rights, linguistic barriers, financial hindrances, pricing systems, institutional impediments, lack of online availability, variety of storage formats and so on.

**3.1.3 SHORT ON STANDARDISATION AND QUALITY** In Europe research in the humanities and social sciences is currently often carried out within national contexts and based on nationally generated data, with a large variation in quality. This fragmentation and compartmentalisation has severe implications for the quality of the data in a European perspective, as crossing borders within the humanities and social sciences comes at a cost. European research based at European level often falls below the standards applied at national level as data are not immediately comparable due to differences in standards and documentation – sampling, collection, variables, size, formats. In conducting European research, one has to frequently rely on post harmonisation of national data at the level the lowest common denominator. As a result, quality and detail are both compromised.

**3.2 THE IMPLICATIONS** To sum up, the humanities and social sciences are

in a situation in which there is an astonishing discrepancy between the potency and research ambitions across Europe and the current state of the European infrastructure backbone, where European-level data either do not exist, are not available or are not comparable.

In other words the possible existence of Europe as a natural laboratory for the humanities and social sciences is currently left unexploited and uncaptured.

This diagnosis brings with it serious consequences for the conduct of the humanities and social sciences. Europe is in a situation where the potential for research in the humanities and social sciences is in many ways better than, for example, in the United States and Japan. But the actual possibilities for turning this potential into a competitive advantage for the European research communities is hampered by the lack of research infrastructures.

First, the deficiencies of European research infrastructures effectively prevent the conduct of high-quality research with a clear European agenda and perspective. As a consequence, such research is a scarce commodity and a situation of widespread ignorance therefore prevails. Researchers initially interested in engaging in comparative European research find it both a costly and cumbersome affair – if not impossible from the outset – in which case they invariably abstain from such work. And the research that is actually carried out with a European perspective is most often based on imperfect and inadequate data, with the research questions generating sometimes precarious answers and fragile comparisons. The resultant research drawn from a European perspective is all too frequently based on non-comparable national data with diverse sets of documentation, and consequently entail the risk of not comparing like for like, or even of comparing apples with pears.

Second, the situation is worse in some parts of Europe than others. Research infrastructures and data are not democratically distributed, as resources and capabilities are spread unevenly across Europe. Today research infrastructure facilities are generally concentrated in northern Europe, leaving southern regions less well equipped. This situation will be

accentuated with the enlargement of the EU. Normally, these concerns are couched in terms of the European Research Area.

However, the implications reach further than the scientific community. Third, the wider socio-economic impact of research in the humanities and social sciences is that certain critically important European research questions simply cannot be asked at all and that many of those that can be asked will be answered by empirical findings based on second-rate data. Clearly, this has severe implications for the overall European competitiveness and cohesion.

## 4. The European Strategy

The provision of first-class research infrastructure is paramount to the humanities and social sciences. The deplorable discrepancy between the scientific ambitions and potency of the disciplines and the available data needs to be remedied. Thus a European strategy is much needed.

Nevertheless, such an effort does not start from scratch. The ground is well pre-pared to launch a concerted effort within the humanities and social sciences. As already mentioned, the current experiences and collaborations – for example, CESSDA and ESS – clearly illustrate the readiness and determination of the scientific communities to cooperate. Indeed, an unprecedented opportunity exists today to take advantage of these enterprises by taking research infrastructure collaboration to a level that could not have been foreseen as little as ten years ago.

However, formulating a strategy for the humanities and social sciences is a daunting task. Not only must it tackle the variety of problems and shortcomings addressed above before Europe can be realised as a natural laboratory, but the strategy also needs to safeguard the plethora of theories, methods, data, subjects and fields subsumed under the disciplines of humanities and social sciences. The aim, therefore, will not be to set up scientific constraints and limitations but rather to aspire to harness and manage the diversity.

The strategy then is to progress from the current situation characterized by isolated patches of research infrastructure to a situation where research infrastructures and data are knit together to form a coherent and integrated system.

The vision is rather straightforward, yet nonetheless highly ambitious: to provide the optimal conditions for creating and circulating data and services within European humanities and social scientific research at European level and thereby promote scientific excellence.

This vision is to be materialised into a concrete proposal to create a coherent, comprehensive and integrated observatory unrivalled in the world – “the dream machine” of the humanities and social sciences – with a view to guaranteeing the existence, accessibility and comparability of data at European level. This facility will be termed the European Research Observatory for the Humanities and Social Sciences and will do duty under the acronym EROHS.

**4.1 BLUEPRINT FOR EROHS** EROHS is based on the philosophies underlying the recent developments in Grid technologies in order to ensure the coordination and quality of geographically dispersed data resources in existence at European level. It will build upon resources as an umbrella-organisation that actively and systematically addresses the current patchwork of research infrastructures for the humanities and social sciences in Europe with an eye to promoting synergy and coherency. This will reduce the risk of duplication and, in effect, will also address the joint issues of economy and transparency.

The four guiding objectives of EROHS – each progressing in terms of European ambition – are as follows:

- The facilitation of access to and sharing of existing European and national data, thereby more efficiently and effectively linking data resources already available.
- The development of improved standards and documentation relating to existing European and national data in order to enhance the scientific quality of data and their potential for interoperability.

Taken together these first two objectives represent a huge step forward for

the humanities and social sciences. But while they are necessary, they are in no way sufficient by themselves, due to the inherent weaknesses of the post harmonisation of existing data already alluded to. Consequently, as a third objective EROHS will have the:

- The generation of new and genuinely European data. This will involve both the collection of new data and the digitalisation of existing data currently in non-computerised materials.

Finally, in order to derive maximum benefit from the investment in the future, a fourth objective will be:

- The provision of research training programmes for the next generation of researchers.

**4.1.1 TYPOLOGIES OF INSTRUMENTS** The objectives spill over into the instruments deployed by EROHS in terms of data access, collection and generation in the humanities and social sciences.

At the forefront there will be four typologies of instruments. In the background there will be a suite of axes that cut across and supplement those at the forefront. As such, the activities can be visualised as a matrix based on different instruments.

The four typologies of instruments are as follows.

**DIGITALISATION** Digitalisation of non-computerised research materials should be done to ensure both their documentation and preservation for posterity. This will also facilitate access, thereby overcoming the “death of distance” problem so that researchers from all over Europe and the world will have equal and easy online access to the data. This is an important element in the infrastructures for the humanities in particular, where great challenges exist to create virtual laboratories of image, sound, visual and textual research materials, each with associated resource discovery tools.

In the social sciences the cooperation between European data archives has already taken important first steps in this direction, for example with integrated and multilingual retrieval aids. In the humanities there is a need to create a digital archive of the cultural heritage of Europe, among other things in order to ensure that data sets and other information transferred to a digital form can be located, accessed, archived and preserved virtually.

Such a resource would improve the availability and exploitation of resources for researchers, especially at a time when these researchers are working in an increasingly interdisciplinary and collaborative environment. For individual humanities researchers, it would also offer opportunities to combine innovative developments in information technology with the riches of our traditional libraries, museums, and galleries; bring together at the user’s desk, different types of resources, irrespective of existing geographical or disciplinary boundaries, thereby facilitating efficiency as well as collaborative research within the user’s institution and between institutions; enable content to be exploited in new ways by, for example, collating content from different locations and facilitating the comparison of different data types; reflect developments in digital libraries (such as the National Digital Library at the Library of Congress and a similar development at the University of Virginia); and exploit interactive IT applications in assessing digital resources.

**SURVEY TECHNOLOGY** Survey technology is one of the most important instruments used in the creation of truly comparative data at the European level.

Survey technology can be divided into three major areas. In the first, sampling statistics, the precision, scope and level of the generated data are defined for each survey according to statistical sampling theory. The second area, observation technology, is often called “survey mode”. Here one has a choice of methods for collecting information from respondents. The third area is how to ask questions, for which numerous types of question formats are available.

Surveys can produce an enormous amount of data. For instance, a medium-sized survey with an average of 3,000 respondents in each the 33 states of the EU and neighbouring countries measuring 500 characteristics per respondent will produce just under 50 million data points. This may not be particularly large compared to some other scientific disciplines, but the multivariate analyses of such data can be extremely complex, especially given that some responses may be open-ended, and context and meaning can vary across linguistic and cultural boundaries. Equally, in order to achieve uniform quality for all these data points an extremely complex survey process engine has to be established, maintained and upgraded.

Moreover, these data points cannot be rendered interoperable without creating meticulous documentation, metadata and related ontology's. In the social sciences data archives exist to provide mechanisms and tools for easy access and exchange of data, but the generation of high-quality cross-national metadata is an arduous task. Similarly as with all instruments, surveys need calibration, otherwise no comparison is possible between different measures. Again, the social sciences have already developed schemes to guarantee comparison over time and across cultures/nations (the Data Documentation Initiative, for example). However, these schemes are still preliminary and often based not on calibrated evidence (measured), but on common sense.

There is therefore a need to ensure further development of this very important instrument in the infrastructure of the humanities and social sciences.

**EXPLORATORY INSTRUMENTS** Experimental instruments are needed for two main reasons: first for purely methodological reasons such as calibrating surveys, testing a specific hypothesis in an experimental or quasi-experimental setting, and testing new modes or question formats; second, for bringing traditional survey methods together with other kinds of measure-

ment of human behaviour and organisation, spatial context data and behavioural experiments, which not only face respondents with specific decisions but reward them depending on their decisions. All these kinds of new means of measurement bring researchers closer to the idea of experimental infrastructure, common to the natural scientists. The United States already funds such an environment for experimental surveys and this may serve as a useful instrument for EROHS.

The cultural and societal variety within the EU is an invaluable experimental setting per se. A European Experimental Survey would thus be highly attractive to scientists from within and beyond Europe.

**CORPORA – EXEMPLIFIED BY LANGUAGE RESOURCES** The term “language resources” denotes large amounts of speech or language data in machine-readable form, used e.g. for language studies, for constructing or evaluating algorithmic natural language and speech systems or as core resources for the software localisation and language services industries, electronic publishing, international transactions, subject-area specialists and end users.

Examples of linguistic resources are written and spoken corpora (i.e. collections of natural texts or natural speech), machine-readable dictionaries (often laboriously compiled by commercial publishers), terminology databases, and speech collection and processing. Basic software tools are also mandatory for the acquisition, preparation, collection, and use of these language resources.

The EU faces a radically new situation in terms of language policies, support, and service in May, 2004, as 10 new members entered the Union, speaking Slavonic (Polish, Czech, Slovak, Slovenian), Baltic (Latvian, Lithuanian), Finno-Ugric (Estonian, Hungarian), Semitic (Maltese), and Turkic (Turkish on Cyprus) languages that have neither been official languages of the European Union to date nor have been incorporated in the earlier corpus-oriented efforts. This enormous step alone calls for a substantial con-

certed EU investment in language-related infrastructure, especially in view of the ever-increasing general dominance of the English language in many spheres of language use.

Closely related to the question of sufficient and easily available corpora is the problem of readily available, free, open source language technology infrastructures for all EU languages (old and new). Free language technology tools (e.g. morphological and syntactic analysers of running text) are available off the web for English but not for most other EU languages.

It could be argued that the ready availability of free, open language technology tools strengthens the dominance of English, while the corresponding lack of free, open language technology infrastructure threatens the survival of many other languages. The EU has a responsibility to support all EU languages, so this problem should be addressed.

These four typologies of instruments will be supplemented by two horizontal instruments.

**STANDARDS, METHODOLOGY AND IT INFRASTRUCTURE** An infrastructure unlocking existing and future data sources and utilising them to their full potential requires international metadata standardisation, multilingual thesaurus support to breakdown language barriers to documentation processes and resource discovery, for use in advanced data access and dissemination systems, and calls for underpinning software technology to allow the integration of survey-based and aggregate-type data. This should be coupled with a rich arsenal of resource location techniques, such as multilingual searching, look-ups by topical classifications, geographical location and other search-related ontologies. For this to function, the infrastructure needs to be rich in content and allow easy access and instant browsing and analysis. The character of the instrument is to bridge diversity and to link together a distributed system.

Despite the recent development of advanced Internet discovery tools and better data sources, researchers in the humanities and social sciences still spend a disproportionate amount of time searching for relevant information, whether text or data, and then negotiating arcane and antiquated access arrangements in order to extract value from these resources. What is needed is an infrastructure and environment that aids researchers at every stage of the process – across resource discovery, access rights, browsing, analysis, publishing and dissemination of findings.

Use is the prime key to dynamic maintenance, renovation and renewal of large-scale databases. Access and availability fosters use and unlocks potential. From resource providers it requires active participation and some common tools. The most important tools are standards for quality assurance, standards that allow comparison across resources and standards that drive software. There are metadata standards available today that allow integration of heterogeneous data types, survey-type data, qualitative data, time series and administrative data. These may help content creators produce necessary contextual, scientific and explanatory information to ensure that the user has access to the best possible content. Yet these not only need to be applied, they also need to be further enhanced and developed, and EROHS should have an important role in this regard.

**COMMUNITY, COMMUNICATION AND TRAINING** Initiatives to improve the European research laboratories for the humanities and social sciences will make a major contribution to the creation of the kind of cooperation that is needed to produce European added value in the endeavour to attain the strategic goals for the future knowledge society.

The humanities and social sciences have a strong tradition in seeking cross-national collaboration in order to enhance the cumulative advantages of scientific work. European investment in strategic infrastructure facilities will add to this tradition and enhance communication and collaboration.

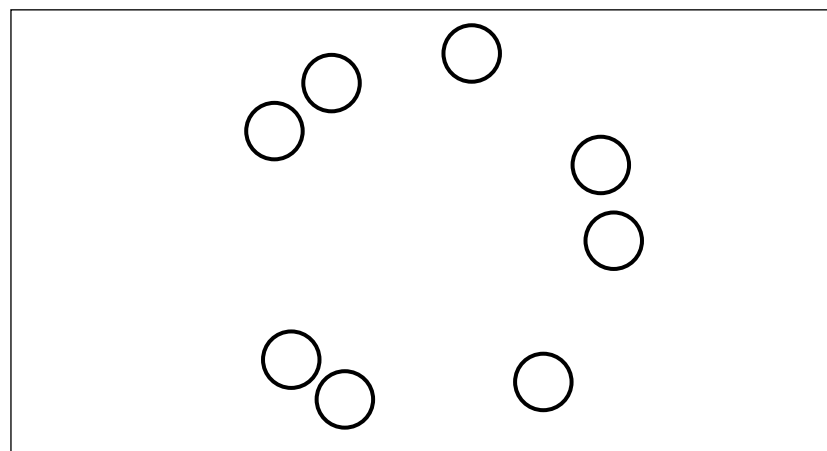
A specific and very important part of this will be the establishment and running of researchers training courses for the research communities to use the infrastructure components maintained and produced by the EROHS initiative.

**4.1.2 ORGANISATIONAL OUTLOOK** Taken together, the entire work package of EROHS represents a massive and ambitious enterprise. Designing an institutional framework that suits and supports this is equally challenging.

In general, one is confronted with the organisational dilemma of remaining sensitive to both the top-down and the bottom-up perspectives.

In order to maximise the impact of EROHS it will need to have a clear and recognizable focal point that ensures the connectivity, coordination and quality of the activities. This suggests a physical centre with sufficient leverage in scientific, administrative, technical and economic terms to set priorities, ensure the scientific quality of the schemes and exert the authority to back its decisions. In short, it needs to ensure that research infrastructure ideas are translated into effective action. Of course, accountability also comes into play here.

*Figure 1: The current state of affairs. The existing infrastructures at European level exist in isolation marring the humanities and social sciences with a raft of problems – lack of coherency and funding; accessibility; and standardisation and quality.*



While coherency and coordination is critical, it is equally clear that a single unit alone will not be able to solve all the tasks, as EROHS will have to base its work on the existing networks of expertise widespread across Europe. This flexibility will thus respect both the variety of approaches and data, characteristic of the humanities and social sciences and accommodate national sensitivities.

Against this background it is obvious that the actual institutional framework of EROHS cannot be a straightforward replica of the large scale single-sited facilities so successful within the natural sciences. Instead one needs to look to flagship models within the confines of the humanities and social sciences. The success and appreciation of CESSDA and ESS, for example, serve as ready inspiration, as do several national efforts within the composite disciplines.

EROHS will therefore be structured as a distributed model with a strong

physical hub working in close conjunction with a number of nodes across Europe – a virtual integrated and interoperable network of excellence. This is regarded as the optimal solution for a machine of this purpose, size, and complexity. Thus the requisite coherency and synergy will arise from a virtual community harnessing European expertise through a coordinated yet decentralised infrastructural network.

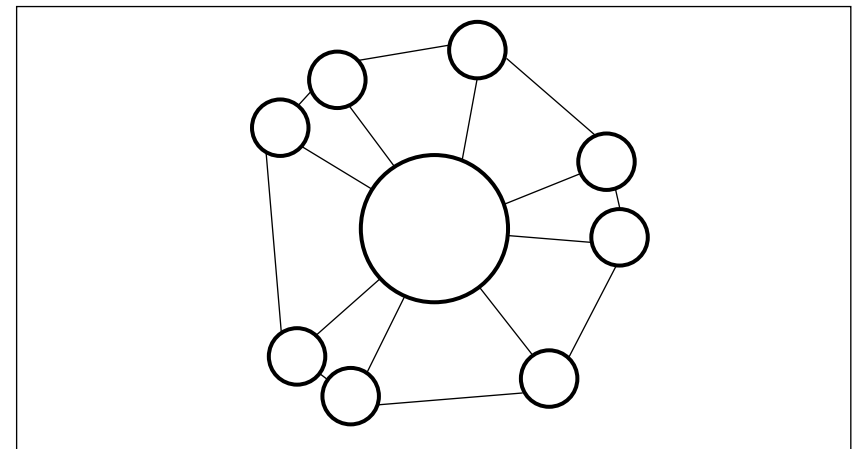
**THE HUB – AN OBSERVATORY** The hub will be based at one location and will obviously play a pivotal role for the success of EROHS. The physical presence of the hub will give EROHS the much wanted authority, solidity and visibility. The hub will also be acknowledged and recognized as the pivot of EROHS.

The hub will have responsibility for the following discrete activities:

- The key strategic role of coordinating the work of EROHS;
- Cooperation and negotiation with a view to guarantee access to data
- Communication;
- Monitoring scientific support and improving the scientific and methodological quality of data provision in Europe;
- Technical support with a view to improving distributed access, availability and use through the development and application of Grid technologies;
- Administrative support assisting the governing bodies and distributing funds to the nodes.

In turn these specific responsibilities demand the presence of both the strategic, administrative, scientific and technical competences and resources at the hub.

*Figure 2: After the first milestone. With the creation of the EROHS-hub the existing research infrastructures will be knit together into a system as nodes providing access to data and improving quality – the first step to form a coherent and integrated system.*



**THE NODES** The hub will have the competence to appoint a number of related nodes on the basis of excellence and through a process of open competition. Through the nodes a collection of “sub-observatories” will thus be spread across Europe, building on existing units of expertise. These may be discipline, subject or theme specific and will very much be where the main bulk of work under EROHS is to be undertaken. However, work within the nodes will not be performed in isolation from one another. Here especially, the role of the hub is of critical importance in ensuring that a semantic humanities and social science web is constructed, binding the various activities together, thus creating a single unified research infrastructure – a virtual conduit, facilitating both capacity building and knowledge transfer.

A few general remarks illustrating the necessary platforms should also be made in order to demonstrate the importance of the decentralised nature of the enterprise – or rather, why the centre cannot command the competence on all relevant issues:

- The nodes will represent different scientific fields – these may be organised along traditional disciplinary lines, or may be cross-disciplinary and organised according to subject, theme or approach. As quite disparate scientific cultures will be working together, there must be a system that brings these scientific communities together.
- The nodes may represent one or several competences, for example methods, sampling, data archiving, standards, accessibility, as this expertise is distributed across Europe.
- The nodes will reflect the rich diversity of national cultures and societies. All the data gathered are anchored in the context of national communities. The detailed knowledge of national rules, procedures and legalities are both respected and brought into play at European level.

**4.1.3 MANAGEMENT AND GOVERNANCE** EROHS must operate as an independent and autonomous body, drawing its expertise from the scientific communities. If EROHS is to secure the necessary legitimacy of the research communities within the humanities and social sciences at large it needs to be self-governed according to excellence alone.

The Observatory will take the lead in managing EROHS and it is envisaged that the co-ordinating hub will have the following main governing bodies:

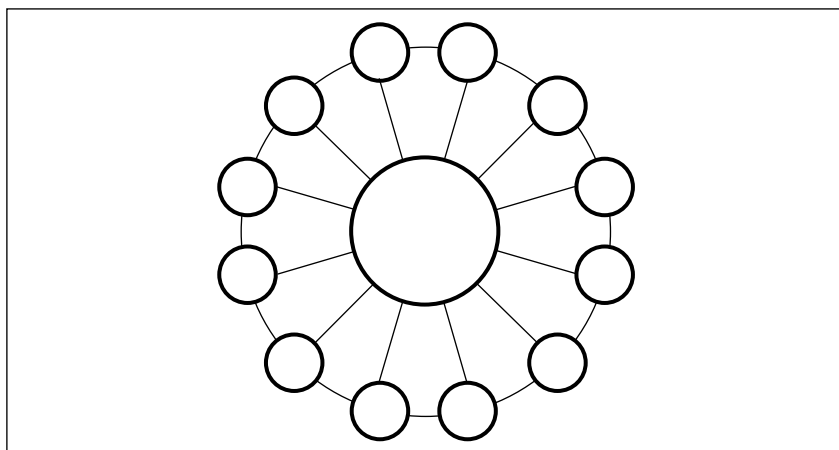
- The Council will have the responsibility for the strategic decision-making functions. It will decide on the overall strategic plans and priorities, and appoint the Director and the Board of Directors. The Council will ultimately be responsible for all important decisions; it will set out EROHS's policies in scientific, technical and administrative matters, approve the programme of activities, adopt budgets and review expenditure. It is suggested that alongside representation from the European Union, each member state should have two official delegates, one representing his or her government's administration, the other the national scientific interests.

- The Science Committee will be the forum ensuring the communication and contact between the research communities, policy-makers, data providers and other stakeholders, bringing together the key stakeholders at European and national level. The Committee will ensure the scientific value of the activities by making recommendations to the Board on the EROHS machinery. Its members will be scientists appointed by Council on the basis of scientific eminence without reference to nationality.
- The Director and Directorate will be the executive body responsible for the day-to-day management of EROHS, being accountable to the Board. Together they will be the drivers of EROHS. The Director must be an internationally highly respected scientific figure, supported by a Directorate-team of equal standing.

Further thought should be given to the governing bodies. Additionally the possibility of setting up a Finance Committee and an Audit Board to review EROHS at fixed intervals could be explored.

**4.1.4 ESTIMATED BUDGET** The funding necessary for the creation of EROHS should come from the EU. The perspective of the whole endeavour is truly European by nature and clearly exceeds the national competences. As such it offers the EU the opportunity to create the world's most efficient data infrastructure for the humanities and social sciences by reducing the duplications and inefficiency that currently desert the existing investments in research infrastructure and by developing new and truly comparative sources of data at European level, enhanced digitalisation, surveys and experimental data collections. Thus, the funding will very much be top-off funding, based on the national investments already made and hence maximising the synergy of those decentralised investments.

Figure 3: After the second milestone, EROHS will establish new research infrastructure facilities that generate new data and digitalise existing data eventually transforming the entire system into a coherent and integrated whole – “the dream machine of the humanities and social sciences”.



Later, additional funding should also come from other sources, but the core financing for this initiative must obviously come from the EU.

In order to have the wanted impact a volume of 25 million euro a year is proposed<sup>1</sup>. This will pay for the construction and operating costs of the hub. The costing will be expected to gradually rise as the activities of EROHS are extended and deepened.

**4.1.5 ROADMAP TOWARDS EROHS** Obviously EROHS cannot be created overnight. The implementation should be a gradual process with increased funding as the new organisation matures and demonstrates its capacity and

<sup>1</sup> This figure is arrived at by extrapolating from similar endeavours at national level and by taking the costing of CESSDA and ESS into consideration.

success. The following general milestones towards excellence could serve as a roadmap for the realisation of EROHS:

- The setting up of a consortium bringing together all the major stakeholders in Europe to facilitate the creation of EROHS
- The creation of the hub of EROHS with a focus on the first two objectives and the last objective of EROHS - sharing of existing data, raising quality and providing training programmes.
- The development of EROHS to also focus on the third objective – the generation of new and truly European data and the digitalisation of data.

The estimated timeframe from the setting up of the consortium to the full-blown existence of EROHS is five-seven years.

**4.2 THE EUROPEAN PROMISE. FACILITATING NEW KNOWLEDGE** Not before the last step is taken will EROHS have been fully established and the entire impact of the strategy realised.

To sum up, the creation of EROHS will facilitate new knowledge in the following ways:

- The existence, accessibility and comparability of European data will strengthen interdisciplinary and cross-border collaboration and the comparative research on a European dimension, and thus reinforce open scientific enquiry. Assumptions will be replaced with comparable and cumulative knowledge, thereby producing sounder observations and more robust empirical findings. The environment of improved and new data will also facilitate the generation of new research agendas and questions that cannot be addressed today.
- The existence, accessibility and comparability of European data will be the defining element of the Europeanization and internationalisation of the humanities and social sciences and in structuring of the disciplines at the European level. The present uneven distribution of facilities and capabili-

ties in Europe will not only be alleviated but will kick-start the building of research infrastructure capacity in the less well-resourced countries of today.

- The existence, accessibility and comparability of European data for the humanities and social sciences increases the opportunity to improve knowledge on social processes and holds great potential in terms of advising European and national policy-makers on how to accomplish the ambitions in the Lisbon declaration by managing societal change solving current and future problems in ways that ensures coherent societal development and European competitiveness.

## Appendix 1 Members of the working group

### MEMBERS OF RISSH

#### **Research Director Niels Ploug (chair)**

Vice-chair for The Danish Social Science Research Council  
The Danish National Institute for Social Research  
Denmark

#### **Director, Dr Maurice Bric**

Clinton Institute for American Studies  
University College Dublin  
Ireland

#### **Director Bjørn Henrichsen**

National Social Science Data Services  
Norway

#### **Professor Poul Holm**

Institute of History and Civilization  
University of Southern Denmark  
Denmark

#### **Nikolaos Kastrinos**

DG Research  
The European Commission

**Director, Dr, Jaap Kloosterman**

International Institute of Social History  
The Netherlands

**Professor Fred Karlsson Department of General Linguistics**

University of Helsinki  
Finland

**Assistant Professor Sylvia Kritzinger**

Department of Political Science  
Institute for Advanced Studies (IHS)  
Austria

**Research Professor Louisa D. Loukopoulou**

Research Centre for Greek and Roman Antiquity  
National Hellenic Research Foundation  
Greece

**Professor, Dr Peter Mohler**

Centre for Survey Research and Methodology  
German Social Science Infrastructure Services (ZUMA/GESIS)  
Germany

**President Gaston Schaber**

Centre for Population, Poverty and Public Policy Studies  
International Networks for Studies in Technology, Environment, Alternatives,  
Development (CEPS/INSTEAD)  
Luxembourg

**Director, Professor Kevin Schürer**

UK Data Archive  
University of Essex  
United Kingdom

**Professor Richard Topol**

Centre de Recherche en Epistémologie Appliquée (CREA)  
Ecole Polytechnique  
France

**Dr Bogdan Van doninck**

Belgian Federal Science Policy Office  
Belgium

**Director, Professor Gert G. Wagner**

The German Socio-Economic Panel Study  
German Institute of Economic Research (SOEP/DIW)  
Germany

**Head of Section Lars Christensen (secretary)**

The Danish Research Agency  
Denmark

Inputs have been received from many sources and the group is not only grateful for the attention and effort but also pleased by the widespread engagement in this matter.

The group has meet on three occasions: Copenhagen 1-2 December 2003, Berlin 2-3 March 2004 and Amsterdam 2-3 May 2004.

## Appendix 2 References to background material

Banda, Enric, *Declaration from the Steering Committee*, Conference Infrastructures: the Backbone of European Research organised by The European Commission, The European Science Foundation and The French Ministry for Research, 18-20 September 2000

Council of European Social Science Data Archive, [www.cessda.org](http://www.cessda.org)

European Commission, *Working Document on the Commission Services. A European Research Area for Infrastructures*, February 2001

European Commission, *Research Infrastructures 2002*, volume A, 2002

European Commission, *Research Infrastructures 2002*, volume B, 2002

European Commission, *Support for Research Infrastructures in the FP6 Specific Programme on "Structuring the European Research Area"*, *Workprogramme 2002-2003*

European Commission, *Europe and basic research*, January 2004

European Council, *Conclusions*, European Council, Lisbon 2002

European Research Advisory Board, *EURAB Recommendations on Research Infrastructure (RI)*, October 2003

European Research Advisory Board, *EURAB Recommendations on the Euro-*

*pean Research Area (ERA) and the Social Sciences and Humanities (SSH)*, January 2004

European Research Council Expert Group, *The European Research Council. A Cornerstone in the European Research Area. Report from an expert group*, December 2003

European Social Survey, [www.europeansocialsurvey.org](http://www.europeansocialsurvey.org)

National Science Board, *Science and Engineering Infrastructure for the 21st Century: The Role of the National Science Foundation*, February 2003

Nowotny, Helga, *Challenges for the social sciences and the humanities within the ERA in particular for the countries of Central and Eastern Europe*, proceeding from the Belgian EU Presidency research conference *Unity and diversity: the contribution of the social sciences and the humanities to the European Research Area*, 2001, Bruges, October 29-30, 2001

OECD, *Social Sciences for a Digital World. Building Infrastructures and Databases for the Future*, Proceedings of the OECD Workshop on Large-scale Infrastructures for Social Sciences (Ottawa, 6-8 October 1999)

Panel A, *Report on the Role of Infrastructures in Research*, Conference Infrastructures: the Backbone of European Research organised by The European Commission, The European Science Foundation and The French Ministry for Research, 18-20 September 2000

Panel F, *How to develop The European research infrastructure*, Conference Infrastructures: the Backbone of European Research organised by The European Commission, The European Science Foundation and The French Ministry for Research, 18-20 September 2000

Smith, John, *Implementation of the European Research Area in the Social and Human Sciences, Especially as Regards the Coordination and Opening-Up of National Programmes*, Discussion Paper prepared for the European Commission, June 2002

Theofilatou, Maria, *Report on the Workshop on "Support to Research Infrastructures in the 6th Framework Programme: New Opportunities for Social Sciences and Humanities"*, organised by The European Commission, 4 February 2004, Centre Albert Borschette, Brussels

Time-sharing Experiments for the Social Sciences, [www.experimental.org/tess](http://www.experimental.org/tess)