

DASISH

Data Service Initiative for Social science and the Humanities

7,5 M €, 3 years

CLARIN

Common Language Resources and Technology Infrastructure





Me
Birger Jerlehag
IT-coordinator at the Swedish National Data Service
snd.gu.se



Part of the Swedish research infrastructure
Governed by the Research Council
Hosted by the University of Gothenburg

FP7 Call

INFRA-2011-2.3.1: Implementation of common solutions for a cluster of ESFRI infrastructures in the field of "Social Sciences and Humanities".

A project under this topic should implement harmonised solutions for the ESFRI Infrastructures in the field of Social Science and Humanities on issues like, for example metadata frameworks, registries, single-sign-on systems and persistent identifiers.

Five PPPs

- EU funding for the PPPs of CESSDA, CLARIN, DARIAH, ESS and SHARE was provided under FP7
- From the results of these projects common challenges were apparent
- Common challenges call for common solutions

A Gateway to Social Science Data

is an umbrella organisation for social science data archives across Europe. Since the 1970s the members have worked together to improve access to data. CESSDA research and development projects and Expert Seminars enhance exchange of data and technologies among data organisations.

20 CESSDA member organisations serve some 30,000+ social science and humanities researchers and students each year, providing access to 25,000 data collections, delivering over 70,000 data collections per annum and acquiring a further 1,000 data collections each year.

CLARIN

Common Language Resources and Technology Infrastructure

CLARIN consists of 193 member institutions and 211 members sites (working units) in 33 countries, and is committed to establish an integrated and interoperable research infrastructure of language resources and its technology.

- aims at lifting the current fragmentation, offering a stable, persistent, accessible and extendable infrastructure and therefore enabling eHumanities
- is a large-scale pan-European collaborative effort to create, coordinate and make language resources and technology available and readily usable



The European Social Survey (ESS) is an academically-driven social survey designed to chart and explain the interaction between Europe's changing institutions and the attitudes, beliefs and behaviour patterns of its diverse populations. Currently in its fifth round and preparing for a sixth, the survey covers more than 30 nations and employs the most rigorous methodologies.



DARIAH

Digital Research Infrastructure
for the Arts and Humanities

14 Partners and 5 associate partners

The mission of DARIAH is to enhance and support digitally-enabled research across the humanities and arts. DARIAH aims to develop and maintain an infrastructure in support of ICT-based research practices.

DARIAH is working with communities of practice to:

- Explore and apply ICT-based methods and tools to enable new research questions to be asked and old questions to be posed in new ways
- Improve research opportunities and outcomes through linking distributed digital source materials of many kinds
- Exchange knowledge, expertise, methodologies and practices across domains and disciplines



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 more than 45,000 individuals aged 50 or over.

The survey's third wave of data collection, SHARELIFE, collects detailed retrospective life-histories in thirteen countries in 2008-09.

Common challenges

- How to achieve integration and interoperability beyond the borders of the individual projects given different data organizations?
- How can we manage to preserve our cultural and scientific memory and keep the records of science accessible?
- How to come from a down-load first scenario to a truly web-based usage scenario to optimally access and enrich the stored data to tackle the many big and small research challenges?
- How can we improve the quality of our data to enable advanced and cross-disciplinary access and enrichment operations?
- How to simplify access conditions for researchers?
- How to establish trust of SSH researchers in the infrastructure services?

Common solutions

- understand the different usage scenarios and architectural solutions to identify ways to come to common SSH solutions for data and service integration and interoperability (WP2)
- identify major quality issues and take serious efforts and measures to improve quality with foci on survey quality enhancements (WP3) and the quality of metadata and data access (WP5)
- establish criteria for long term persistency and curation of data and interact, preferably with the emerging data infrastructure, to push the quick deployment of first generally available services (WP4)
- work on all aspects that will foster shared data access and enrichment, starting with basic layers such as AAI based trust domain up to cross-disciplinary data enrichment frameworks (WP5)
- take care of legal & ethical issues that are of relevance for all SSH domains in a cross-disciplinary activity and work on simplified solutions (WP6)
- take a variety of measures in trust building and to engage the communities – in particular the young generation of researchers and perhaps even the public – in making use of advanced features of the infrastructures by education and training programs (WP7)
- disseminate the results according to proven channels (WP8)

Other initiatives

Each of the SSH initiatives has defined the kind of services that they need to implement to satisfy their communities and each established requirements for their centres to get the infrastructure and their services operational. On the other hand we expect collaboration in particular with GEANT/eduGain/TERENA to establish a European trust domain allowing single identity and single sign-on mechanisms

WP2:

☐ to get a deep understanding about the differences in the way data is managed and organized in the SSH with the goal of a stepwise harmonization towards a reference architecture, better integration and higher quality

WP3:

☐ to improve the quality of survey data to make it fit for the cross-country/cultural and lingual challenges

WP4:

☐ to offer robust deposit and long term preservation services for all SSH researchers preferably in collaboration with a European data infrastructure based on trust

WP5:

☐ to improve (meta) data quality and the infrastructure integration by sharing joint methods for basic services such as AAI, PID, metadata etc. as the basis for typical automated e-Research workflows

☐ to improve the visibility, accessibility and re-use of state-of-the-art tools and services that are useful in a cross-disciplinary context and provide a cross-disciplinary annotation platform

WP6:

☐ to bundle SSH forces to simplify access restrictions for researchers in particular due to the many IPR issues the SSH domain is faced with

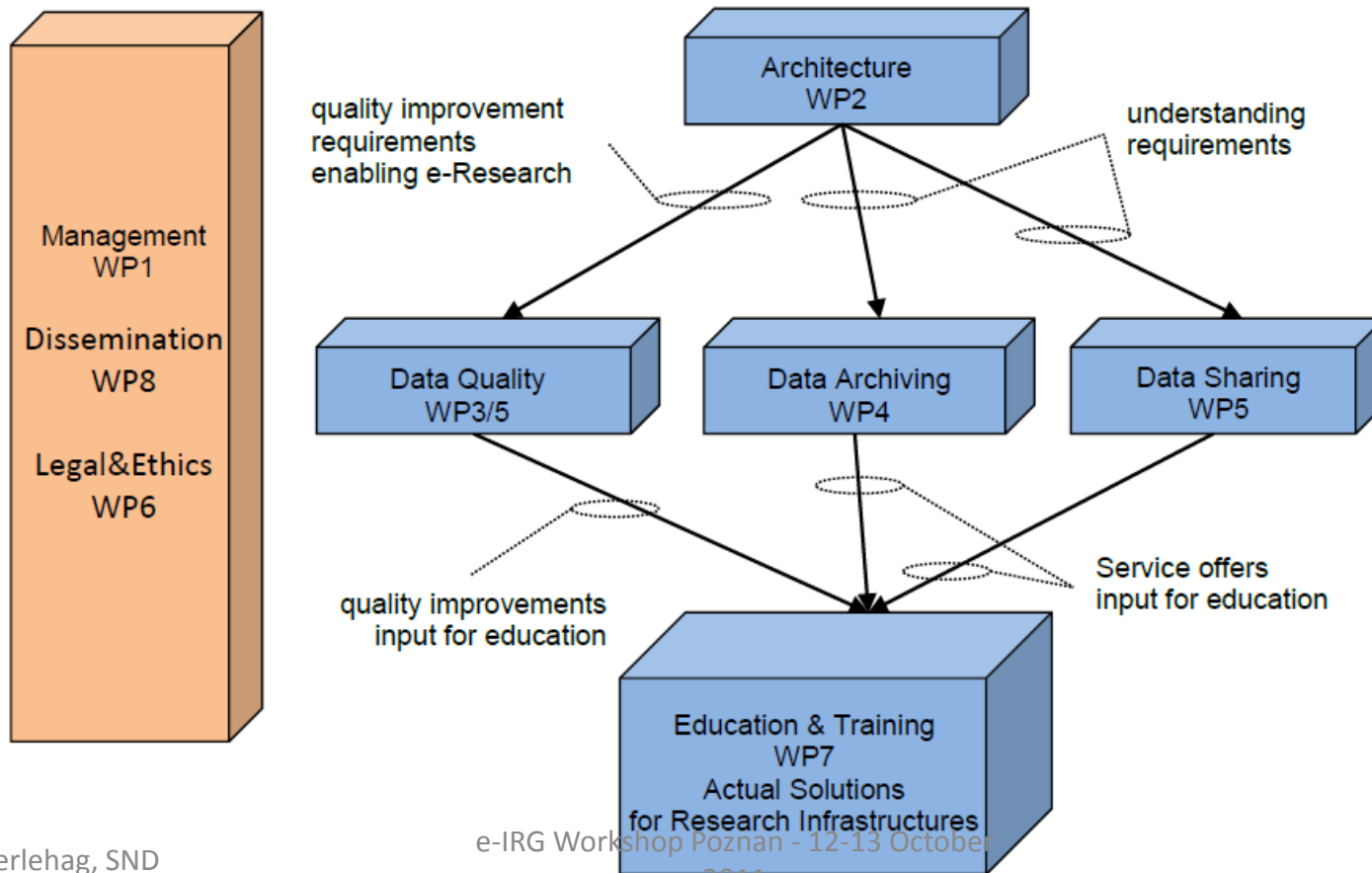
WP7:

☐ to take serious measures to educate and train SSH researchers about how to efficiently apply state-of-the-art methods and technology by taking profit from each other's knowledge, and to involve the public and pupils in these efforts

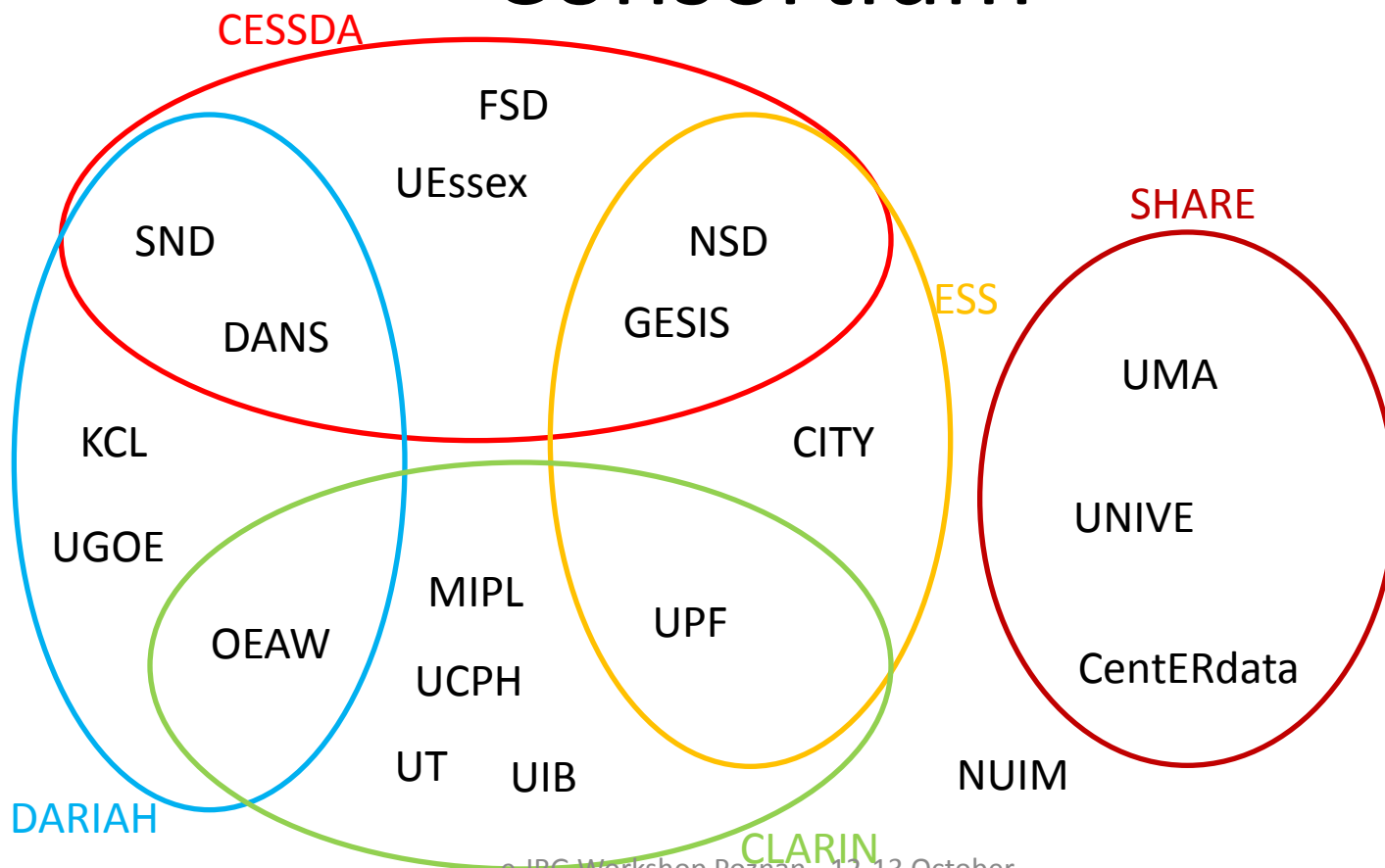
WP8:

☐ to reach out to the participating communities and beyond single sign-on mechanisms

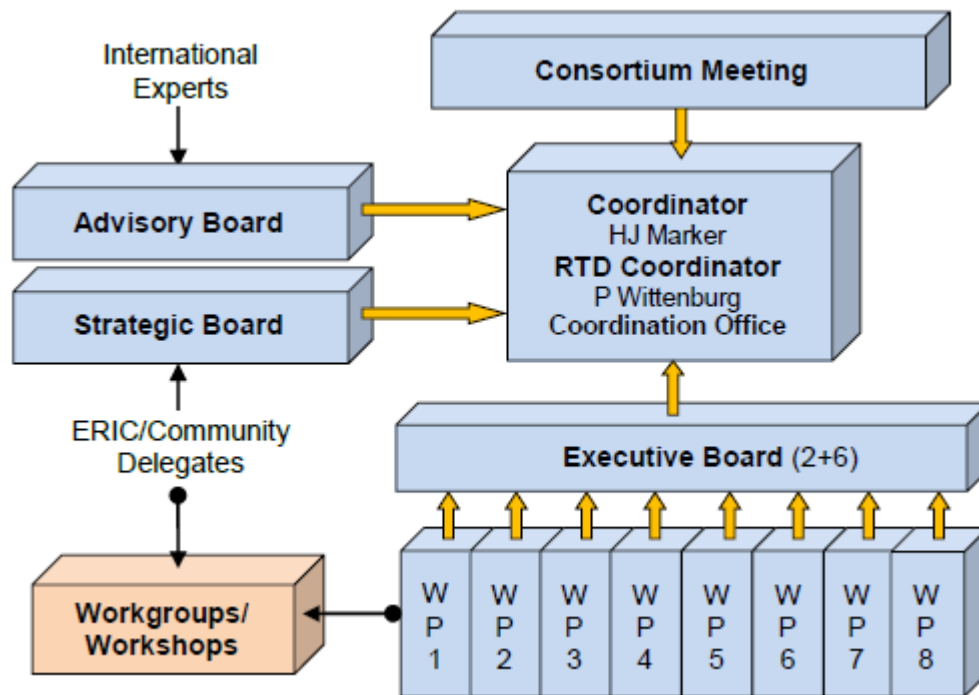
Work organization



Consortium



Management



CLARIN

Common Language Resources and Technology Infrastructure



Thank you