



Enhancement of data and metadata infrastructures for the CESSDA RI

Work Package 8 of the 2008-2009 CESSDA PPP under the 7th Framework Programme

(See <http://www.cessda.org/project/> for information on the CESSDA PPP.)

Data are the single most important component necessary for a science-based understanding of society. To enhance access to cross-national data is to promote research with complex data. Associated metadata are a key element not only for location and discovery but also interpretation and especially evaluation of the potential for relevance to data comparison, integration and harmonisation. A logical extension is to upgrade and build tools, which facilitate metadata capture and management for data producer along the life-cycle and provide user facilities for the virtual comparison and possible linking, integration or harmonisation of data across different locations.

The efforts to enhance data and metadata infrastructures must be persistent to keep up with the rising demands of the data users and to maximise the full potential of the new technologies to make the creation of metadata easier. It allows data producer, to make data access more direct and allowing data and metadata to flow from the producer to the user while retaining the linkage between them.

With these considerations in mind, WP8 sets out to work on three substantive key objectives.

Objective 1: To determine the metadata and data model requirements of the CESSDA Research Infrastructure (RI) for handling complex dataset types throughout the entire data life-cycle.

The discovery, identification and interpretation of the available data depend largely on the quality of the associated metadata. The CESSDA members have invested much effort in coming to terms with the increasing expectations of the data users concerning the efficiency of data and documentation access. The needs to enhance data and metadata models grow with the increasing complexity of the data collected. Metadata capture, management, publishing and the accompanying user demands to access such metadata require as a prerequisite adequate technical standards that allows designing and handling these complex metadata.

The starting point of this wide-ranging objective regards the need to evaluate the data documentation standard DDI XML version 3. Use cases will examine the new modular

structure and extended facilities to link and reuse metadata sources in relation to its capacity to support the life-cycle processes of complex datasets including preservation metadata.

Key issues regard the representation of metadata for comparative data and questionnaires and the management of datasets and their expression in the DDI schema. Particular aspects as the versioning and preservation theme are to examine in cooperation with further work packages. Based on the results potential development needs will then be discussed with the DDI Alliance.

A second focus will assess and define the requirements for data and metadata models in the field of preservation of complex data and documentation products.

Objective 2: To plan the strategic developments required for metadata, data models and software upgrades for data and metadata capture, management, processing, publishing and access within the CESSDA RI; to support more complex dataset types.

The first key task is to perform a conceptual analysis of technical developments required to upgrade existing software to handle complex data and its related documentation with metadata along the several key phases of the study life-cycle.

With special consideration of cross-national datasets and time series, the analysis has first to identify key phases and outcomes in data work. Further analysis of processes and particular process outcomes supports the generation of exemplary use cases describing systematically the complexities of different study and data types on the space and / or time dimensions. Conceptual analysis preparing for data models needed to support further complex data types like combined individual and aggregate data sets complete this work.

The structured overview of functional needs and specifications for metadata of cross-national surveys, time series and further complex data types is finally to transform them into strategic recommendations for software upgrades within the CESSDA RI considering different systems and data models serving metadata at particular phases of the life-cycle.

Objective 3: To recommend options for the organisation, sustainability and future funding of metadata, data model and software tool development.

Due to the changing technical landscape and the changing nature of user expectations, there is a general requirement to keep metadata development continuously under review as well as to enable sustainable developments of standards and tools.

Consequently, recommendations are of high demand, providing options and specified funding models for future resourcing and sustainability in developing and implementing metadata standards and software tools.

The second key issue under this third objective regard the development of possible models, which provides scenarios for the strategic organisation of technical developments in the future. Recommendations have to discuss in particular how distributed work of exiting teams can keep the balance with some activities at central locations to pool and exchange expertise, efforts and outcomes.

Both efforts – the provision of efficient and effective strategies for developing and implementing metadata standards and software tools - aim to provide complex metadata infrastructures to facilitate access to high quality social science research data to the scientific community.

Participation in Work Package 8 (contact persons at the national partner archives)

The project team of WP8 is comprised of staff from different CESSDA members.

Contact persons at each organisation:

Uwe Jensen,	GESIS Data Archive, WP leader)	uwe.jensen[at]gesis.org
Herve L'Hours,	UKDA	herve[at]essex.ac.uk
Mari Kleemola,	FSD	mari.kleemola[at]uta.fi
Atle Alvheim,	NSD	Atle.Alvheim[at]nsd.uib.no
Hans-Jorgen Marker,	DDA	hjm[at]dda.dk