Data Service Infrastructure for the Social Sciences and Humanities

Towards a common metadata understanding for the three DASISH WP3.2 tools

Report on work and agreements

Bergen
January 5th 2015

Authors:

Hilde Orten, Håvard V. Bakkmoen, Knut Kalgraff Skjåk (NSD)

Taina Jääskeläinen (FSD)

Brita Dorer (GESIS)

Joachim Wackerow (Consultant)

Contirbutors:

Yvette Prestage, Eric Harrison (ESSERIC HQ)

Eric Balster, Edwin De Vet, Maurice Martens (CentERData)

Anna del Castellarnau, Ricardo Fernandez (UPF)

Contents

Introduction

The three tools under development under the Data Service Infrastructure for the Social Sciences and Humanities (DASISH) project, Work Package 3.2 are a questionnaire design and documentation tool, a translation management tool and a question/variable database. The three tools are developed as individual tools that should be able to communicate with each other. To achieve this, a common metadata understanding between the three tools is needed. Work towards a common metadata understanding based on the metadata standard DDI (DDI Alliance 2013_1) is carried out in connection with the DASISH project. Requirements for a common metadata model have been specified. This regards which metadata elements will be transferred between the tools, the metadata flow between the tools, the mapping of metadata elements to the DDI, identification and versioning, how the exchange of metadata between the tools takes place, as well as administrative ownership of metadata. This report describes the work towards a common metadata model for the three tools carried out within DASISH, agreements made, as well as open issues to work forward with.

0. Towards a common metadata model for the three DASISH task 3.2 tools

0.1 The three tools of DASISH task 3.2, aims

The three tools under development under the Data Service Infrastructure for the Social Sciences and Humanities (DASISH) project, Work Package 3.2 are:

a) The Questionnaire Design and Development Tool (QDDT);
b) the Translation Management Tool (TMT);
c) and the Question Variable Data Base (QVDB).

The tools are planned to serve the new European Infrastructures like the European Social Survey (ESS) and Survey on Health, Ageing and Retirement in Europe (SHARE) in their survey processes.

These tools are currently under development as individual tools that should be able to communicate with each other.

The purposes of each the tools three tools are described below:

a) The Questionnaire Design and Development Tool (QDDT):

The ESS question module development process is iterative. It takes place over a period of 20 months from selection of the Question Design Teams (QDTs) in June of the first year, through to issuing the final source questionnaire in March of the third year.

To ensure transparency, each stage is documented, and reveals to data users the rationale behind the design decisions made en route. The new Questionnaire Design and Development Tool (QDDT) is designed to document this process via a database (rather than the current Word template). In addition to streamlining the process, this change will facilitate the production of usable outputs (such as printed questionnaires, and the status of development of individual items and concepts) for the administration team during the design process. It will also feed directly into the Translation management Tool (TMT) and the Question Variable Data Base (QVDB).

The tool will also document the results of pre-testing on items developed to measure the required concepts and will thus provide a much needed source of evidence for other scientists in the field who want to address similar topics in future.
We see the QDDT as a database with specific sections and fields that can either update automatically or be manually updated. The database should be accessible online to allow access by multiple users in multiple locations. It should be interactive. It needs to be able to handle text, numerical data and figures / graphics and provide usable outputs – preferably in both PDF and Word format. The information in the database may be made publicly available.
Developed originally with the ESS in mind, the QDDT will be designed to service projects internal as well as external to DASISH, to interoperate with other systems and tools, and to act as a reusable model for other questionnaire design and development tools. The system should be able to import and export metadata structured in DDI at a high level of granularity, which should be done in a manner where human interaction with the system is minimal.

ESSERIC HQ is responsible for specifying and testing the tool, with the questionnaire design procedures and needs of other cross-national survey programs in Europe in mind. FSD has also

provided input to the specification. NSD is responsible for ensuring its compatibility with the metadata standard DDI and for the programming (Prestage et al 2013).

b) The Translation Management Tool (TMT):

ESS and SHARE have led the development of improved translation procedures, many of which have now been adopted by other cross-cultural surveys. It is effectively a 'committee-based' system, aiming to achieve optimal translations via staged perusal and amendment by different actors, whose interaction plays a major role in ensuring cross-cultural equivalence. This task start with SHARE's Language Management Utility (LMU) and is developing into a comprehensive translation tool to facilitate and faithfully document the process. Its outputs will be usable during the translation design process and will then feed into the Question Variable Data Base. It will also be an open source 'plug in' available to other cross-cultural surveys. GESIS is responsible for specifying and testing the tool in collaboration with MPG (MEA) and FSD, CentERdata is responsible for programming whilst NSD advise on compatibility with DDI and the Questionnaire Design and Development tool and the Question Variable Data Base (Martens et al 2013).

c) The Question Variable Data Base (QVDB):

Access to survey questions in their original languages is often severely limited. The scope of the Question Variable Data Base is to create a searchable database with a broad public profile, providing user access to survey questions in their original languages, the concepts they are based on, additional material used in the data collection, the resulting variables and their generation, coding classifications and more. The primary aim of the tool under development aims to serve the ESS and other DASISH projects in their business processes, by facilitating reuse of metadata components at the different stages and sub-stages of these processes. The tool will also allow users from within the research community to browse the database (question, concept and variable mining) in order to design new research or to explore existing projects. Developed initially with the ESS in mind, the Question Variable Data Base is designed to also service projects external to DASISH. A goal for the QVDB is to interoperate with other systems and tools, and to act as a reusable model for other question databanks. NSD is responsible for specifying and testing the tool, in collaboration with ESSERIC HQ, UPF and FSD. NSD is responsible for the programming, as well as for the DDI compliance. (Skjåk et al 2013).

0.2 A common metadata model for the three tools:

A core aim of the task 3.2 tools is to serve surveys at different stages in their business processes.

At a very general level, one could say that the three tools cover work processes of questionnaire design and development, translation and archiving, and of providing metadata for reuse. The QDDT would typically serve questionnaire development related issues, while TMT would serve the translation process. The QVDB is planned as a multipurpose archiving tool, to serve the different stages in the archive business processes, and with the possibility to serve a variety of tools and web-services.

The system should work so that metadata developed at a certain stage of the survey lifecycle could be reused at the later stages. Interoperability between the three tools will allow for metadata

components developed at earlier stages of the work processes to be reused later, over the survey lifecycle.

The aim is to allow the three independent tools to communicate in an efficient way. A common metadata understanding between the three tools facilitates this. The DDI Lifecycle metadata specification (see chapter 3.1) has been chosen as the transport format between the tools, and also form basis for the QDDT and QVDB conceptual models. Possibilities for bridging between DDI metadata elements of the three DASISH tools and elements of the SQP (Survey Quality Predictor) has also started to be explored.

During the DASISH period, work towards a common metadata model for the three DASISH task 3.2 tools has been carried out in order to facilitate DDI-based transport of metadata elements between the tools. The work and decisions made is described in the following chapters.

0.3 Work towards a common metadata understanding /meetings

Early spring 2013 NSD started the work to define required metadata elements of relevance for the metadata transfer between the tools, based on design and archiving business processes of the ESS, in collaboration with the ESSERIC HQ. The work to map identified metadata elements to the DDI was also started. A consultant was engaged to advice on the development of the tools to be developed by NSD, as well as on the metadata exchange between the three DASISH tools (see Appendix 1).

Meetings; work towards a common metadata model

Three meetings have been held in order to progress the work on the common metadata understanding for the three DASISH tools.

The first meeting was held in Bergen 7th and 8th of May 2013. Challenges and solutions to a unified development/communication were discussed. An important outcome of this meeting was the identification of requirements for a common metadata model which are specified in chapter 0.4.

A second follow-up meeting was held in Bergen again the 4th and 5th of February 2014. This meeting was focused on each of the requirements for a common metadata model resulting from the May meeting 2013. Working groups were defined in order to work forward with particular topics and unresolved issues.

The last meeting was held in Tilburg the 7th and 8th of July 2014. At this meeting important decisions were made regarding the remaining open issues and the further work. For an overview of roles and working groups related to the current work, see Appendix 1; acknowledgement to participants of the three meetings, see Appendix 2).

Between the meetings discussions have been going on between the participants, ideas and policies have been worked out and metadata related tasks like identifying and mapping of metadata elements to DDI, development of DDI element hierarchies and example DDI instances (publication wrappers) have been exchanged between the partners for agreement.

0.4 Requirements for a common metadata model

The requirements for a common metadata model for the three DASISH task 3.2 tools are defined as follows:

1. Which metadata elements will be used in the transfer between the three tools?
2. The direction for the flow of metadata elements between the three tools, as well as the steps in the work process at which metadata components are exchanged.
3. Mapping between the metadata elements and DDI.
4. A common identification and versioning system.
5. How the exchange of DDI metadata takes place, which type of DDI instances or fragments that will be transported, and which type of web-service will be used.
6. Administrative ownership of metadata

In the following each of the above requirements, as well as agreements made, issues to be resolved and related work is described.

1. Which metadata elements will be used in the transfer between the three tools?

Metadata transfer is a core objective of the current work. The first requirement to arrive at a common metadata understanding between the three DASISH task 3.2 tools is to get an overview over which metadata elements will be transported between the three tools. The identification of metadata that will be transferred between the tools forms basis for a common metadata understanding, as well as the DDI profile for the metadata transfer between the tools.

In order to identify metadata elements for the transfer between the tools, work has been carried out to analyse the workflow of the business processes of the ESS data from the design stages to the archiving stages, as well as of relevant outputs at each stage. Particular focus has been put on the metadata elements for transfer between the tools QDDT and QVDB developed by NSD, and the TMT developed by CentERdata.

1.1 Identification of metadata elements to be transferred between the three DASISH tools

Metadata for transfer between the tools are typically questionnaire related metadata elements that also are relevant for translation. Examples are questionnaire modules as in the instrument, question items, instructions, categories and other response options, as well as complex question structures stemming from different types of question grids, for example.

A list of identified questionnaire related metadata items formed basis for an agreement on which metadata elements that are relevant for transfer between the tools. The metadata for transfer between the tools are rich metadata that require a structure of high level of granularity to allow for maximal flexibility in reuse in business processes, as well as in presentation, over the data lifecycle.

1.2 Identify between which tools the metadata elements will be transferred

Work has been carried out to identify which metadata elements are relevant to see, handle or store in each tool. This work formed basis for deciding between which tools transfer of identified metadata elements will take place.

While the transfer of some components seems rather straight forward, other would require some follow-up work on the side of the metadata structuring in the tools, or finding best practice solutions for the structuring of metadata in DDI. Alternatives for structuring and transport of the complex structures, like complex question structures, scale layout and comparisons are currently explored.

1.3 SQP workflow and possibilities for interaction with the system

In the current design process of the ESS, questions are tested in the Survey Quality Predictor programme (SQP). During the development of the source questionnaire questions are coded for quality checks by members of the Research and Expertise Centre for Survey methodology at Universitat Pomoeu Fabra, Barcelona. During the process National Coordinators code translated questions in SQP. In each case, the question text, responses and instructions are entered manually into the system and the full coding is manual as well. To make this process more efficient it could be useful if some metadata elements (question text, introduction text etc.) could be imported

automatically to SQP from the DASISH tools, to facilitate the entering of questions to the system and perhaps also in order to automate the coding of some of the characteristics. In return, it could be useful to be able to import quality reports resulting from the SQP automatically to the DASISH tools.

For SQP and the DASISH task 3.2 tools to communicate efficiently, a mapping between the DDI metadata elements and the SQP metadata elements needs to be developed, and possibilities for this has started to be explored. This regards the identification of SQP elements and the mapping of the internal SQP identifiers to the DDI identifiers of the DASISH model, which of the metadata elements that would be relevant for SQP that are included in the DASISH tools, the detailing of metadata elements, as well as how the metadata elements from the tools are structured in DDI for other purposes, compared to how they preferably should be structured to map to SQP. Exchange is possible but will need to be further explored.

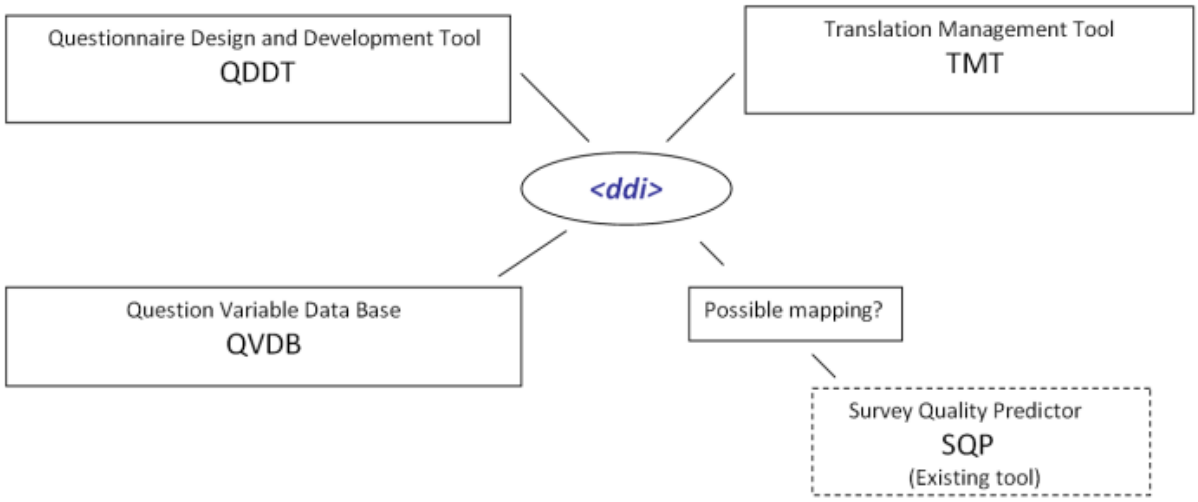

Figure 1. Possible communication between the three DASISH tools and SQP, through a mapping to DDI

Work towards the goal of a mapping between the DDI metadata elements of the DASISH tools and the SQP metadata elements has started, and some possibilities are explored. In some instances it will probably make sense to detail metadata from the DASISH tools in a way that facilitates import to SQP, and it should be explored further in which cases this could be relevant. It has for example been agreed to distinguish between 'Introduction text' and 'request for an answer text' in the DASISH tools.

To sort out the possibilities for inclusion and detailing of metadata elements in the DASISH tools, as well as to provide a full mapping between DDI and SQP would represent an interesting task for a follow-up project.

2. Metadata flow: The direction for the flow of metadata elements between the three tools, as well as the steps in the work process at which metadata components are exchanged

2.1 Which metadata should reside in which tool?

In order to achieve a common metadata understanding it is important to agree on where the different metadata elements are stored and where they are maintained. Maintenance in this context refers to any maintenance (storage and updating - changes, additions and deletions) of metadata elements.

The following has been agreed between the partners:

The QDDT stores the development history of questionnaire related metadata items like questions and concepts, as well as their final versions, and these metadata are maintained in the QDDT. The TMT stores translated metadata elements and the translation history, as well as their final versions, and these metadata are maintained in the TMT. The TMT could also store copies of metadata elements from the QDDT to aid the translation process, but these are maintained in the QDDT. The QVDB stores copies of final/publicly published metadata elements imported from the QDDT and the TMT, but these metadata elements are not maintained in the QVDB. Variables and classification schemes are stored and maintained in the QVDB.

In some more detail this means that the QDDT should contain the development history of the modules/questionnaires developed with the tool, i.e. all versions of modules (questionnaire and related elements incl. concept hierarchy) including final/publicly published versions. In the case of the ESS, this would mean that the modules of the source questionnaires are developed with this tool, that all versions are stored in this tool, and that management and maintenance of the source questionnaire modules takes place in the QDDT.

The TMT should contain the development history of module/questionnaires developed with the tool, i.e. all versions of module/questionnaire including final/publicly published versions. In the case of ESS, this means that all translations of the source questionnaire modules are developed in this tool and that all versions of the translated questionnaires are stored in this tool. Management and maintenance of translations takes place in the TMT. So do all versions of country module/questionnaires variants that are in the same language as the source questionnaire (For example Irish modules/questionnaires if the source questionnaire is English). Versions of the source questionnaire itself may also be imported and stored in this tool to aid the translation process, but management of the source questionnaire takes place in the QDDT. The source questionnaire may be commented on, and relevant comments could usefully be imported back to the QDDT.

Finally, the QVDB should store copies of the final/publicly published versions of source questionnaires and final/published versions of translated questionnaires (incl. detailed structured instrument and questionnaire related metadata elements as agreed). I.e. versions that are marked as publicly “published” in QDDT and TMT are imported and stored in QVDB. In additions to the imported metadata elements from the QDDT and the TMT, the QVDB contains variables, derived variables and coding schemes and other metadata elements. The management and maintenance of variables, coding schemes etc. should take place in the QVDB.

The table below gives an overview over where some important metadata elements are developed and maintained:

	Development and maintenance (storing, and updating - add, change, delete)	**Storing of copies**
Concept and concept hierarchies	QDDT	QVDB, TMT
Questions and questionnaire related metadata elements for source questionnaires/modules	QDDT	QVDB, TMT
Translations of questionnaire related metadata elements	TMT	QVDB
Represented (reusable) variable Variables	QVDB	
Classification or coding schemes/ classification code-lists	QVDB	QDDT, TMT

Table 1: Overview of where some important groups of metadata elements are develop, stored and maintained according to the proposed model.

2.2 Metadata flow between the tools

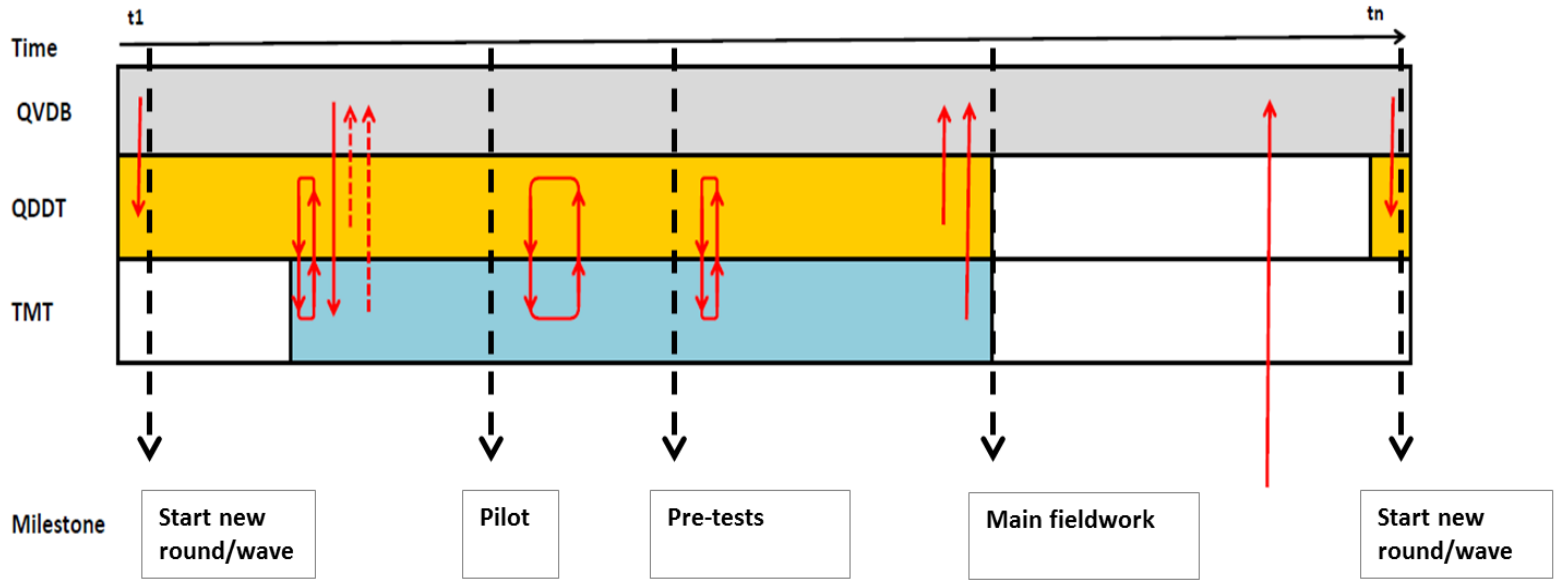

Figure 2: Possible metadata flow between tools during questionnaire design and translation – generic example

Figure 2 above gives a generic description of the possible metadata flow between the three DASISH task 3.2 tools. The arrows in the figure are explained below, using a selection of important use cases for possible metadata flow between these tools.

Figure 2: What the arrows mean:

Circles of arrows:

When a version of a questionnaire module or a single question is ready for translation, it should be possible to import copies of the module or question to the TMT from the QDDT. Export from QDDT, import to TMT. This transport will be of items that are not final. It should be possible to import copies of translation comments to the QDDT from the TMT. Export from the TMT, import to the QDDT.

Red dotted arrows:

Possible use-case: It should be possible to import copies of questionnaire modules and single questions that are not final (e.g. pilot questions) into the QVDB from the QDDT, and also import translations of those from the TMT. Export from the QDDT and TMT, import to QVDB.

Red arrows (drawn line):

- First red arrow:

  At the start of, as well as during the questionnaire design process, it should be possible to import copies of questionnaire modules or single questions from the QVDB to the QDDT for reuse. This would typically be questions or modules from other surveys than ESS and SHARE. Export from QVDB, import to QDDT.

  Another possible use-case for the first arrow: Import of copies of country-specific code lists from classification schemes developed and maintained in the QVDB for reuse in response alternative/ domain in QDDT.

- Second red arrow:

  It should be possible to import copies of translated questionnaire modules or translations of single questions to the TMT from the QVDB for reuse during the translation process. This would typically be questions or modules from other surveys than ESS and SHARE. Export from the QVDB, import to the TMT.

  Another possible use-case for the second arrow: Import of copies of country-specific code lists from classification schemes developed and maintained in the QVDB for reuse in the TMT.

- Third red arrow:

  It should be possible to import copies of final questionnaire modules or single questions into the QVDB from the QDDT. Export from QDDT, import to QVDB.

- Fourth red arrow:

  It should be possible to import copies of final translations of questionnaire modules and single questions from the TMT to the QVDB. Export from the TMT, import to QVDB

- Fifth red arrow:

Possible metadata transfer, to be explored:

It should be possible to enrich the QVDB with metadata elements from other sources or tools than the QDDT and the TMT, for example fieldwork monitoring metadata . Import to QVDB from external sources.

• Sixth red arrow:

See first red arrow.

Exchange between tools – file based and web-services:

Agreement has been made to develop web-services for each of the three tools, which allow machine access to metadata elements stored in the tools (for further description, see chapter 5.2). This will be in addition to the agreed DDI3.2 XML file based transport.

An open issue to be worked further with is how tools should know that metadata are available.

3. Mapping between the metadata elements and the Data Documentation Initiative (DDI specification)

The Data Documentation Initiative (DDI) is an acknowledged metadata specification describing data from the social, behavioral, and economic sciences, developed by a self-sustaining membership Alliance recently reconfigured with new Bylaws, and successfully adopted by national and international organizations and institutions throughout the world.

The DDI metadata specification has two branches: DDI Codebook and DDI Lifecycle. The DDI Lifecycle branch of DDI supports the entire research data life cycle, and its metadata specification accompanies and enables data conceptualization, collection, processing, distribution, discovery, analysis, repurposing, and archiving. Structuring of metadata in DDI-Lifecycle allows machine actionability and reuse of detailed metadata elements between surveys and within the same survey over the time and across modes of collection. Tools for questionnaire design, coding and data publishing can be built on the top of metadata structured in the DDI.
DDI Lifecycle, (current version 3.2) is a rich metadata specification well suited for structuring of metadata that requires a high level of granularity, like for complex surveys. Detailed metadata structuring in DDI facilitates use and reuse of metadata elements in the business processes of surveys from the design stage to the archiving, as well as for publication, and allows for saving time and money in the business processes in the long run.

The current DDI is expressed in Extensible Markup Language (XML) which is a simple, very flexible text format, designed to meet the challenges of large-scale electronic publishing, and also playing an increasingly important role in the exchange of a wide variety of data on the Web and elsewhere.

For more information about the DDI specification, see Wackerow 2013 and DDI Alliance 2013_1.

3.1 Export format: DDI XML

It has been agreed to use DDI 3.2 XML for the transport of metadata elements between the three DASISH task 3.2 tools. DDI 3.2 (DDI Alliance 2013_2) is the current version of DDI Lifecycle.

3.2 Which version of the DDI should be used for the mapping?

Accordingly, it was agreed to use DDI 3.2 in the mapping of metadata elements to DDI. The tools developed by NSD, the QDDT and the QVDB, are modeled for compatibility with DDI 3.2, as well as to allow for compatibility with other versions of the DDI. CentERdata also agreed to use DDI 3.2 for the mapping of metadata elements for the transport. Their data dissemination tool Questacy is already based on the current version of DDI Lifecycle, DDI 3.1.

Advantages of using DDI 3.2 for the mapping compared to 3.1 is that it contains a set of new features of relevance to the structuring of metadata elements from complex surveys e.g.:

- Question grid, that allows for structuring of complex questionnaire structures/grids.
- Scale domain, that provide interesting alternatives for display formats for scales.
    - Separate management of different types of response options/domains is possible (e.g. answer formats like scale, date-time and missing values)

- Transport of single maintainable/versionable metadata elements, e.g. questions, is possible using the fragment structure.
- DDI 3.2 also has a new identification system and versioning rules (https://github.com/DASISH/survey-tools-metadata)

3.3 Mapping of metadata elements to DDI - structuring of metadata elements

Work has been carried out to map the identified metadata elements for each tool to DDI, where the main focus has been put on metadata elements that will be transported between the tools.

DDI metadata element hierarchies which specify metadata elements and attributes relevant for usage in the QDDT and the QVDB, as well as for the transport between the tools (https://github.com/orgs/DASISH/teams/t3-2/repositories) are currently under development. DDI instances (publication wrappers) are exchanged between CentErData and NSD for agreement. Work towards a DDI profile for the transport of metadata elements between the three tools is in progress (see chapter 5.1).

Two issues have been of particular focus in the work related to the structuring of metadata elements have been:

1) Reusability: How to structure metadata in DDI in order to maximize the reusability of metadata elements in different contexts and by different surveys?
2) The handling of translations: Should source questions and translations be regarded as compound – or separate elements?

Structuring of metadata with reusability in mind

The tools should be designed to allow reuse of metadata elements, and it is important that metadata are structured in a way that facilitates their reuse.

Questionnaires typically contain more or less reusable components. It could make sense to define a hierarchy of reusability for different things, which could from basis for deciding which components could be reused in which contexts. For a question item the following components would typically be more or less reusable:

- **Most reusable**:
  Reusable within country-language variants of the same survey, across modes and over time, as well as between surveys.
  - Question text
  - Response domain
  - Question intent/concept (if used)

- **Reusable but less:**
  Reusable within country-language variants of the same survey over time.
  - Mode specific – reusable within modes or across some (for example PAPI or primitive web surveys) but not all modes.

- Administrative instructions (other than routing instructions). For example related to respondent or interviewer behavior.

  - Study specific - reusable within country-language variants of the same survey over time, across modes but not between surveys.
    - Question number/name

- **Less/not reusable:**
  Sometimes reused within, but not between surveys
    - Routing instructions
    - External aid reference (Showcard number reference)

Based on the above it has been agreed that the DDI element QuestionItem should contain only the most reusable content (question text and response domain). QuestionConstruct in DDI should contain the less reusable parts of a question, for example question number and administrative instructions.

To let QuestionItem contain the question text and response domain also makes sense from a researcher perspective, as changes in these components are likely to result in a change in the collected data. An instruction like 'Tick one box only' in PAPI self-completion mode, or 'Mark only one answer' in Web-survey is, however, expected to result in the same outcome.

It is also been agreed to keep variables reusable in the QVDB. The DDI distinction between element 'RepresentedVariable', that contains the core reusable parts of the variable, and 'Variable', which is the applied expression of a data item within a data set will be supported by the QVDB.

Work is currently carried out on how to handle complex question structures with reusability in mind, and agreements between the partners regarding this needs to be made. QuestionGrid in DDI will be supported by the QDDT and the QVDB in order to structure complex grid-like questionnaire setups where the individual 'questions' included are not suited for reusability as single question items. If the included questions of a question battery, for example, contains questions that could be reused as single questions, it would be better to structure them as a Sequence of QuestionItems and instructions.

Below are two examples of question batteries from the European Social Survey.The first example (Figure 3) could typically be structured as a QuestionGrid, while the second (Figure 4) could be structured as a Sequence of individual questions and instruction.

CARD 5 Using this card, please tell me on a score of 0-10 how much you personally trust each of the institutions I read out. 0 means you do not trust an institution at all, and 10 means you have complete trust. Firstly…**READ OUT…**

		No trust at all										*Complete trust*	*(Don't know)*
B2	…[country]'s parliament?	00	01	02	03	04	05	06	07	08	09	10	88
B3	…the legal system?	00	01	02	03	04	05	06	07	08	09	10	88
B4	…the police?	00	01	02	03	04	05	06	07	08	09	10	88
B5	…politicians?	00	01	02	03	04	05	06	07	08	09	10	88
B6	…political parties?	00	01	02	03	04	05	06	07	08	09	10	88
B7	…the European Parliament?	00	01	02	03	04	05	06	07	08	09	10	88
B8	…the United Nations?	00	01	02	03	04	05	06	07	08	09	10	88

Figure 3. Example of a question battery from the European Social Survey where the individual sub-questions are less reusable as individual question items.

CARD 11 Using this card, please say to what extent you agree or disagree with each of the following statements. **READ OUT EACH STATEMENT AND CODE IN GRID**

		Agree strongly	**Agree**	**Neither agree nor disagree**	**Disagree**	**Disagree strongly**	**(Don't know)**
B26	The government should take measures to reduce differences in income levels.	1	2	3	4	5	8
B27	Gay men and lesbians should be free to live their own life as they wish[16].	1	2	3	4	5	8

Figure 4. Example of a question battery from the European Social Survey where the individual sub-questions are more reusable as individual question items.

In order to maximize the reusability of metadata elements it has been agreed to include metadata elements in other elements by reference rather than inline in the metadata element if possible. This means for example that a response domain (for example a code list) is included by reference in a question, rather than 'belonging' in the question itself. This allows the response domain to be reused by different QuestionItems or QuestionGrids.

Handling of translations

It was discussed between the partners whether the source questions and translations be regarded as compound elements or separate elements, with the conclusion that translations of a source or master question should be handled as separate elements.

The rationale for this decision is that:

-In cross-national surveys, translated questions are mirror images of the source question. The source question is the real and true starting point. Giving a new version to the source question just because there has been a change in one translation (say, in one out of 15 languages) seems not really helpful but rather confusing.

-Researchers working with data from one country or a few countries need to see when the question has changed over time. From their point of view, only changes in the translated questions are relevant, not changes in the source. So they need to be able to pinpoint changes in these particular translated questions.

-Translators are interested in changes in both source questions and translated questions. Independent versioning enables them to detect and easily track down a change in any question.

An issue that is raised by handling of translations independently from that of the source questionnaire is related to multilinguality in variables. From the perspective of a variable (merged across countries), the definition and the categories could usefully have text and labels in multiple languages. Although not a requirement by the ESS, it would be desirable for the tools to support reuse of the categories from the question items. This is, however, not straightforward, and it is currently not clear how this could be done in DDI 3.2.

4. A common identification and versioning system

4.1 identification system

A common identification system is advantageous as it makes it easier to work with the same metadata in different tools.

The main purposes of the identification structure of DDI is to uniquely identify objects in a persistent manner to ensure accurate reference and retrieval of the object content , as well as to manage metadata change over time.

The' Canonical URN' of the DDI has a three part structure which is equivalent to a persistent identifier of an object. The structure consists in 1. an Agency identifier, which is a unique identifier for an agency registered in the DDI Alliance. The agency can have multiple sub-agencies; 2. a Unique id of the object within the context of the agency; and 3. a Version number of the object, to track versions over time.

Canonical URN in DDI has the form: "urn:ddi:agency[.sub-agency]:ID:Version"

It has been agreed between the partners to use Canonical URN that is recommended by the DDI Alliance.

The object id part of the Canonical URN was agreed to be a Universally Unique Identifier (UUID), which is scoped globally, but where the usage could be scoped to the agency or sub-agency. The advantage of using a globally unique identifier is that it is easy to generate and can travel frictionless through software systems (independently from the DDI agency id).

Local identifiers for other purposes can be stored as UserID in DDI. In this way the mapping between DDI and other identifier systems is possible. This applies for example to database identifiers or SQP identifiers.

4.2a Versioning system

The main purpose of a versioning system is to enable humans and machines to distinguish changes in metadata elements that are important from those that are not so important as well as to track provenance.

A versioning system relates to how versioning should work, how many versioning levels the tools and the DASISH metadata model should allow, how a change in an element that is included in another element affects the versioning of the parent element, if the system always should retrieve the latest version of an element (late bound references), and similar. A versioning system typically ticks in once metadata elements are regarded as final or are published.

It has been agreed between the partners to distinguish between the technical versioning and the business versioning

Technical versioning

The technical versioning regards how to keep track of revisions. It has been agreed that a system external to DDI could take care of the technical versioning, but it should be up to each tool host to decide regarding the technical versioning. It was also agreed that the identifier of an external versioning system could be stored as a UserID in DDI.

Business versioning

The purpose of the business versioning is to save any stages of the development which can be important for going back to a previous stage or for understanding the development process.

The business versioning is purely driven by the decision of the user. The user should assign a new version on the appropriate level according to the versioning policy, according to some version rationale decided by the versioning policy.

DDI is used for the business versioning. Each business version which should be made available to partners should be additionally indicated by the DDI attribute isPublished. “Published” means here not necessarily public, but available for usage.

How many versioning levels

It has been agreed to allow three meaningful versioning levels in the DASISH task 3.2 tool for versioning of published objects: ‘Major, Minor and Sub-minor version’. However, in DDI 3.2 an unlimited set of versioning levels is possible.

While three versioning levels will be possible in the tools, the DASISH metadata model will only use two levels ‘Major, Minor’ to keep versioning simple. Projects with a different versioning system will be free to use up to three versioning levels.

Versioning rationale and usage of RationaleCode

VersionRationale in DDI will be used to assign a new version at the appropriate level, as well as to document the rationale of the version changes. RationaleCode in VersionRationale supports a controlled vocabulary. It was decided to develop RationaleCode controlled vocabularies describing different types of changes to metadata objects. Below is an example of RationaleCodes developed for lists or Schemes in DDI, like category lists, question lists, variable lists etc.. Further examples of Rationale Codes are found at https://github.com/DASISH/survey-tools-metadata.

Rationale code value	Change	Comments
ListName.Added	Major	Adding a ListName.
ListName.Removed	Major	Removing a ListName
ListName.Changed.TypoOrNoMeaningChange	Minor	Minor changes such as changes in punctuation, spacing, capitalization or spelling, and other typographical and orthographical changes that do not change the meaning.
ListName.Changed.MeaningChange	Major	Any other change that could affect the meaning. If the amendment is considerable enough, review whether becomes a new element rather than a new version.
ListLabel.Added	Major	Adding a ListLabel.
ListLabel.Removed	Major	Removing a ListLabel.
ListLabel.Changed.TypoOrNoMeaningChange	Minor	Minor changes such as changes in punctuation, spacing, capitalization or spelling, and other typographical and orthographical changes that do not change the meaning.
ListLabel.Changed.MeaningChange	Major	Any other change that could affect the meaning.
ListDescription.Added	Major	Adding a ListDescription.
ListDescription.Removed	Major	Removing a ListDescription
ListDescription.Changed.TypoOrNoMeaningChange	Minor	Minor changes such as changes in punctuation, spacing, capitalization or spelling, and other typographical and orthographical changes that do not change the meaning.
ListDescription.Changed.MeaningChange	Major	Any other change that could affect the meaning.
ListOrder.Changed	Major	Any changes made to the order of items in the list.
Contentelement.Added	Major	Adding a Contentelement. Contentelement is an element included in-line or by reference in the scheme, for example, a question in a question scheme.
Contentelement.Removed	Major	Removing a Contentelement. Contentelement is an element included in-line or by reference in the scheme, for example, a question in a question scheme.
Contentelement.Changed.TypoNoMeaningChange	Minor	Minor changes such as changes in punctuation, spacing, capitalization or spelling, and other typographical and orthographical changes that do not change the meaning.
Contentelement.Changed.MeaningChange	Major	Any other change that could affect the meaning. If the amendment is considerable enough, review whether becomes a new element rather than a new version.
HierarchyLevel.Added	Major	Applies to schemes which allows hierarchical structures.
HierarchyLevel.Removed	Major	Applies to schemes which allows hierarchical structures.

Figure 3: RationaleCode controlled vocabulary describing types of changes to lists or schemes

The goal is that the system should work like this: The person changing a metadata element ticks or chooses a RationaleCode from the controlled vocabulary after having made the changes, thereby documenting what has been changed. The system automatically gives the element a new version number. Whether it is a minor or major change depends on the code chosen, and the user does not need to worry about the minor-major aspect.

It has been discussed whether the system should allow consecutive versioning (the person chooses a RationaleCode from the controlled vocabulary after each change) only, or if batch versioning (the person chooses two or more RationaleCodes from the controlled vocabulary after all the changes has been made to the element for a particular business version) should be supported.

It was agreed that the tool should support batch versioning, as this would result in less version numbers and would allow a quicker and more flexible workflow. A problem with this could be that the person doing the documentation could just choose one RationaleCode instead of two, recording of the reason behind one change only.

A solution here could be that the tool could provide a flag about which elements are changed since the last business version, and remind the person doing the work to choose appropriate

RationaleCodes for the different changes. Some extra checks could be included in the tool to detect instances where two or more changes are made to the same elements, for example one code being added to- and another removed from the same CodeList.

It was also agreed that the tools should support RationaleDescription where the person can describe the reason for the version change in own words. The usage of RationaleDescription should be optional for each project.

Usage of UserAttributePair and publication status controlled vocabularies

It has already been mentioned that each business version made available to partners or to the general public should be flagged with the DDI attribute 'isPublished'

In addition it has been agreed between the partners that it would be useful to be able to distinguish between versions that are not published, internal publications of material that is not final - for example for exchange between partners, and final content that will be published to the broader public. It would also be useful to indicate the business purpose of this publication. A typical example would be that a version is published internally for export from the QDDT and import to the TMT.

It has been decided to use UserAttributePairs in DDI for this purpose. UserAttributePair supports controlled vocabularies that can be used to 'flag' versions of metadata elements for a particular business purpose. When a metadata element is published for a particular business purpose it can be assigned a 'flag' which indicates the purpose of the publication. The tools could function so that the person who enters a version of a metadata element into the tool and wishes for example to publish this internally selects a key value from a drop-down list to denote for which purpose he or she is publishing. The same key value can be used at different stages of the translation process. When a source question item or a questionnaire is transferred from QDDT to TMT, it should arrive with version information since it will not be versioned in the TMT, but it would be useful to know if it is a development version or a finalized version.

Controlled vocabularies for publication status has been developed for the QDDT, the QVDB and the TMT (https://github.com/DASISH/survey-tools-metadata). Example key values for the QDDT vocabulary is included below.

Code value	Comment
Development.NotPublished	Development – Made available for the key members of the CST questionnaire design sub group, but not designed to be published internally.
InternalPublication.AvailableForDesignTeam	AvailableForQDT – Items and discussion made available internally for members of the relevant Questionnaire/Module Design Team to view and comment on.
InternalPublication.ExportToSQP	ExportToSQP – Items made available for 'export' to SQP for coding. The results from SQP will provide guidance to further development of the item, and may be used at multiple stages of the questionnaire process. Final items will also be exported to SQP.
InternalPublication.AvailableForNationalTeams	AvailableForNationalTeams – items and their development histories made available to National teams to view. They will be asked to comment, but these comments will only be added to the QDDT by the team at ESS HQ. There may be some aspects of the development history that we would like to 'hide' from national teams.
InternalPublication.ExportToTMT	ExportToTMT – Items made available for 'export' to TMT for translation. Items may be exported to the TMT at various times in the questionnaire design lifecycle, for example when questionnaires for pre-testing are produced (this may only be a selection of items) and when items are finalized.
ExternalPublication.ExportToQVDB	ExportToQVDB – once finalized, items will be exported to the QVDB to be made publically available.
ExternalPublication.ExportToPublic	ExportToPublic – in addition to the final question, the development history will be made available to the public. The published development history might however be an edited version of the development process that is stored in the database.

Figure 3: UserAttributePair controlled vocabulary for publication status for the QDDT, example ESS

The keys have been developed for the ESS and SHARE but could also possibly be used by other surveys. They are as a first step intended for usage for questionnaire modules, instruments, question schemes and questions. It will be considered later whether they would prove useful for other metadata elements as well.

'Upward versioning'

'Upward versioning', which is mandatory in DDI, has also been approved by the project group. 'Upward versioning' means that a version change in a versionable element residing in a parent maintainable packing element triggers a version change in the parent element. The version of the elements remains unchanged if the version of the parent element changes if this version change is caused by changes in some other metadata elements that the parent element contains, see chapter 3.1 'Versioning' in DDI Alliance 2013_3.

- A practical example of what this means: if an element included inline in a QuestionItem (e.g. the question text) changes, the version of the whole QuestionItem changes as well, as it is the parent element. However, elements that are included in the QuestionItem by reference (e.g. response options/domain or instructions) and versioned independently, will not change version because of this change.
- Another example: if a question belonging to a question list (QuertionScheme in DDI terms) is changed in a way that trigger a new version, then the question list (Scheme) also get a new version.

Late bound references

Late bound references allow you to access the latest version of an object. It was agreed to include possibilities for late bound in the tool, but that usage should be optional for each project.

4.2b Versioning policy

As different institutions and actors will be using the tool, a common versioning policy is needed.

Which metadata elements should be versioned?

It has been agreed that the fully developed tools should support versioning of all versionable DDI elements that are used. This includes for example concepts, instruments, questions, categories, code lists, variables as well as all types of lists or schemes in DDI.

When should the business versioning start?

A discussion regarding when versioning should start in the different tools has led to an agreement that business versioning should start when an object is available for search and archive. This could be from the start of the development, which is the most likely scenario for tools where it is important to track the full development history, like for the QDDT and the TMT, and at latest when an element is shared between partners, which is the most likely scenario for the QVDB.

How many versioning levels?

It is an aim to keep the versioning policy as easy as possible to understand as well as to work with (even if the issue in itself is quite complex).

It has been agreed to restrict the number of meaningful versioning levels to two levels for published items in order to keep versioning simple. Minor changes typically represent minor changes in punctuation, spacing, capitalization or spelling and other typographical and orthographical or minor wording changes that do not change the meaning, while all other changes will be of a major character.

Versions and new elements

To decide which change represents a version of an existing element and what is a ‘new’ element will probably be done on a case-to-case basis. When a ‘new’ metadata element is based on an existing metadata element, a BasedOn reference to the ‘old’ element could usefully be included in the ‘new’ element to track the relationship between them. A BasedOn reference is for example relevant to use in order to track the relationship between questions of the source questionnaire and their translations, as translations will be versioned separately from the source questionnaire elements.

Versioning of translations

Translation is the communication of the meaning of a source-language text by means of an equivalent target-language text. In case of shared languages, survey question items are sometimes translated identically and sometimes differently across countries sharing the language. This happens

because translations need to stay close to the everyday language in the country where the questions are fielded, and that some questions contain country-specific elements.

It has been agreed that translations will be versioned independently of questions from the source questionnaire (see chapter 3.3 on Handling of translations). Translated questions are based on the source question and will have a 'BasedOn' reference to the source. 'BasedOn' will also be used to show if a translation is based on a previous, similar but not equivalent translation of the same source questionnaire.

A change in the source question always triggers a change of version in the translated question, even in the very rare case that the translation does not change (e.g. source questions adopts a synonym of a previously used word and both synonyms are translated with the same word in the translated question). Source change triggers a new version of a translated question because this is a translation of a 'different' source question. In this rare case, Comparison Maps in DDI could successfully be used to document why equal items have different versions. The advantage of versioning translations independently is that source question version is changed only when there actually has been a change in the source, not when a translation has been changed. Otherwise upward versioning would trigger a change in the source as well, as this would be a parent element.

Regarding the versioning of translations there are two options: To version translations by language or by language and country. As several country-specific variants of a translated question item into a shared language exist, it would probably be too high level to version translated questions by language. Versioning by language and country is thus the approved alternative. It was agreed to use the xml:lang attribute as a 'locale' (combination of a language and a region or country), for example using IETF language tags: http://en.wikipedia.org/wiki/IETF_language_tag. This will apply to any language differences in questions. Other country-specific differences in questions should be documented by other means (not xml:lang). Questions with national political parties or educational programs as response categories would for example be versioned and maintained by each country as country-specific questions.

A similar versioning system could, however, be implemented in different ways, depending on translation goals and workflows. Decisions regarding the handling and versioning of translations are available at https://github.com/DASISH/survey-tools-metadata.

5. Metadata exchange between the tools

5.1 Metadata exchange

DDI profiles describe the subset of DDI objects used by an agency for a specified purpose. DDI profiles for the transport of metadata elements between the tools are currently under development. Work is currently carried out that specifies the content and the structure of the DDI publication wrappers for the planned instances of metadata transfer between the tools. These will form basis for the DDI profile for the metadata exchange between the tools. Some example instances nn and mm are available at xx.

5.2 Technical exchange and Web-services and file-based export

Agreement has been reached to use DDI 3.2 XML (see chapter 3.1) as the export format.

It has also been agreed to develop REST API web-services for the communication between the tools. A web-service is a software system designed to support machine to machine communication of XML based content over a network. Web-services could be used to provide any identified DDI item from one tool to another. All the tools will when final have web-services to provide easy access to the metadata maintained in the tools. This could be understood as a distributed storage system, where it is agreed which metadata elements are maintained in which tool (see chapter 2.1).

6. Administrative ownership of metadata

6.1 DDI agency

As mentioned in section 4.1, the persistent identifier of an object includes an "Agency identifier". The agency identifier is a unique identifier for an organization registered in the DDI Registry, and identifies the administrative owner of DDI elements. The administrative owner of an element is the organization creating and/or maintaining the metadata element.

The DDI Registry integrates with the Domain Name System (DNS) to provide global resolution for DDI-related services the organization runs. For example, the organization can allow queries for the URL of a Web Service-based repository that holds its metadata. This will facilitate for the meta data flow described in section 2. Another relevant example for the ESS would be to offer web services to international metadata portals or harvesters run by CESSDA and/or other international infrastructure organizations.

According to the DDI specifications, the agency name should be in the form [country code] dot [name], for example nz.statistics (Statistics New Zealand) and uk.census (UK censuses, Office for National Statistics). International organizations that up till now have registered as DDI agencies have used "int" as [country code], for example int.GSIM (Generic Statistical Information Model, under the umbrella of United Nations Economic Commission for Europe).

In principle the agency names are "meaningless" strings of characters, in the same way as the numbers in the book identifier ISBN. An organization might register in the DDI as be.xwpo4t, - the DDI registry will keep track of which organization the characters represent. However, as long as characters are used, organizations in the DDI Registry tend to represent themselves with "meaningful" names, mostly acronyms. The DDI Alliance must approve of the agency names.

The members of the Manager group in WP3.2 of DASISH (see Appendix 1) have proposed that the ESS and SHARE adapts the practice of other members in the DDI Registry. The DDI agency name will be **int.esseric** for the ESS. This abbreviation will distinguish the European Social Survey from other research related "ESS" activities like Eurostat's European Statistical System (ESS) and the European Spallation Source AB (ESS, listed in the first ESFRI roadmap). Equally, int.shareeric has been chosen for SHARE.

6.2 Sub agencies and version responsibility

The DDI registry allows for the use of sub agencies. It was discussed whether sub agencies could be useful in keeping track of the survey production pipeline, for example by identifying in which application or organization different meta data are created and/or should be maintained (see chapter 2.1).

In addition, the DDI opens up for documentation of who are responsible for the versioning of the different elements, so-called "Version responsibility".

As an outcome of the discussions it was decided not to use sub agencies and rather focus on version responsibilities within each project or agency.

Version responsibility should in a "generic" way represent responsibilities and tasks according to a project's business and work plans. One major problem is to find generic categories, since business plans, like people and organizations, change over time. This discussion has been very useful as it makes the projects focus on the need for good and detailed work plans where responsibilities in the operation phase of the applications are clearly and unambiguously allocated.

Further work

Follow up of the current work is planned in order to make agreements regarding structuring of complex question structures, as well as to complete the DDI profile for the exchange of metadata elements between the tools, and for each of the independent tools.

To further explore the communication between the three tools and SQP could be an interesting task for a follow-up project.

Summary of the agreements

The main agreements made under each requirement can be summarized as follows:

1) Which metadata elements will be used in the transfer between the three tools?
Metadata for transfer between the tools are typically questionnaire related elements that also are relevant for translation. Examples are questionnaire modules as in the instrument, question items, instructions, categories and other response options, as well as complex question structures. The identified metadata for transfer between the tools are rich metadata that require a structure of high level of granularity to allow for maximal flexibility in reuse in survey business processes.

2) The direction for the flow of metadata elements between the three tools, as well as the steps in the work process at which metadata components are exchanged.
Concepts and questionnaire related metadata elements like questions and instructions will be developed in the QDDT, while copies will be stored in the TMT and the QVDB. Translations are developed and maintained in the TMT, while copies will be stored in the QVDB. Variables and classification code lists are developed in the QVDB while copies of the classification code lists will be stored in the QDDT and the TMT.

Metadata will be exchanged between the tools at certain steps in the survey business processes from the start of a new round or wave, till the completion of the main fieldwork. Directions and steps in the metadata flow between the tools have been identified.

3) Mapping between the metadata elements and DDI – structuring of metadata elements.
DDI 3.2 XML is the agreed exchange format for the transport between the tools. Work has been carried out to map the identified metadata elements for each tool, as well as for the transport between the tools to DDI.

Reusability of metadata elements within and between surveys has been a guiding principle when decisions have been made regarding structuring of metadata elements in DDI:

- QuestionItems in DDI will for example contain only the most reusable parts of a question, like the question text and a response domain, while QuestionConstruct will include elements of questions as in the questionnaire, and will include for example question name and instructions.
- Variables will also be structured with reusability in mind. RepresentedVariable in DDI will contain the more reusable parts of a variable, while Variable will represent the applied expression of a data item within a data set.
- In order to maximize the reusability of metadata elements it has also been agreed to include metadata elements in other elements by reference rather than inline in the metadata element, if possible. This means for example that a response domain (for example a code list) is included by reference in a question, rather than 'belonging' in the question itself. This allows the response domain to be reused by different QuestionItems or QuestionGrids.
- As translations can change independently of changes in the source questionnaire, it has been decided to handle translations as separate reusable objects.

4) A common identification and versioning system

DDI 3.2 Canonical URN on the form urn:ddi:agency:ID:Version will be used for the identification of the metadata elements. The ID of the canonical URN will be a Universally Unique Identifier (UUID). Local identifiers, like for example database identifiers, can be stored as UserID in DDI.

Technical versioning, or keeping track of revisions, should be distinguished from the business versioning.

DDI is used for the business versioning. The user should assign a new version on the appropriate level according to the versioning policy, according to some versioning rationale decided by the versioning policy.

The main purpose of a versioning system is to enable humans and machines to distinguish changes that are important from those that are not so important as well as to track provenance. The agreed business versioning system has the following properties:

- The system should allow three versioning levels
- VersionRationale in DDI will be used to document version changes. RationaleCode controlled vocabularies for describing different types of changes to a metadata object are under development. A goal is that the user chooses a RationaleCode from a list to indicate what kind of change has been made, and the system automatically assigns either a major or minor change in the version number, according to the code chosen. A RationaleDescription will additionally be allowed in the tools, but its usage will be optional.
- Each business version which should be made available to partners or to the general public should be flagged by the DDI attribute 'isPublished'. A UserAttributePair controlled vocabulary in DDI should additionally be used to distinguish between elements that are not published, as well as between internal and external publications.
- 'Upward versioning' which is mandatory in DDI has been approved by the group. Upward versioning means that a version change in a versionable element residing in a parent element triggers a version change in the parent element.

- LateBound references, which allow users to access the latest version of an object, should be allowed in the tool. Its usage should, however, be optional.

The agreed versioning policy states that:

- All versionable DDI elements that are used should be versioned.
- Business versioning should start when an object is available for search and archive. For the QDDT and the TMT versioning is expected to take place from the start of the development process.
- Two versioning levels will be used: Major and Minor. Minor changes are typos or minor wording changes that do not change the meaning of the metadata element. All other changes are major changes.
- Batch versioning: Users can choose two or more RationaleCodes, if more than one type of change is made to the element at one go, to describe all changes to the metadata element since the last version.
- Translations are versioned by country and language and reuse of translated metadata elements should be allowed within and between countries.

5) How the exchange of DDI metadata takes place, which type of DDI instances or fragments that will be transported, and which type of web-service will be used.

DDI 3.2 is the export format for exchange of metadata elements between the three DASISH tools. An agreed goal is to develop REST API web services for the communication between the tools. A web-service is a software system designed to support machine-to-machine communication of XML based content over a network.

Full questionnaire modules including concepts and questionnaires, as well as single questions will be exchanged between the three tools. In addition it should be possible to export classification schemes from the QVDB to the QDDT and the TMT.

To further explore the communication between the three tools and the Survey Quality Predictor tool (SQP further could be an interesting task for a follow-up project.

6) Administrative ownership of metadata

The administrative owner of a metadata element is the organization or 'agency' creating and/or maintaining the metadata element. The DDI registry has a Domain Name System for agencies that develop and maintain metadata objects.

DDI domain names for the DASISH surveys are agreed as **int.esseric** for the ESS and **int.shareeric** for SHARE.

References

DDI Alliance. (2013_1). Data Documentation Initiative (DDI). [Online] Available http://www.ddialliance.org/[04 March 2014]

DDI Alliance. (2013_2) DDI Lifecycle 3.2. [Online] Available http://www.ddialliance.org/Specification/DDI-Lifecycle/3.2/ [04 March 2014]

DDI Alliance. (2013_3) DDI Lifecycle 3.2. Part 1: Technical document. Draft [Online] Available http://www.ddialliance.org/Specification/DDI-Lifecycle/3.2/XMLSchema/HighLevelDocumentation/DDI_Part_I_TechnicalDocument.pdf [04 March 2014]

Wackerow, J. (2013) DDI - A Metadata Standard for the Community. [Online] Available https://theuniversityofgothenburg.basecamphq.com/projects/8978593/file/178395805/DASISH_DDI_Bergen-May-2013.pptx [04 March 2014]

Prestage, Y. et al (2013) DASISH WP 3.2a: Questionnaire Design Documentation Tool https://theuniversityofgothenburg.basecamphq.com/projects/8978593/file/160981660/DASISH_WP3_Requirements_document_v3%208.7.13.docx[04 March 2014]

Martens, M. et al (2013) DASISH WP 3.2: Translation management tool [Online] Available https://theuniversityofgothenburg.basecamphq.com/projects/8978593/file/158756838/DASISH_WP32_1_Requirements_document_translationenvironment_06.docx [04 March 2014]

Skjåk, K. et al (2013) DASISH WP 3.2.c: Question Variable Data Base [Online] Available https://theuniversityofgothenburg.basecamphq.com/projects/8978593/file/178395655/DASISH_WP3.2c_QuestionVariableDataBase_Requirements_document_v2_2014-01-28.docx [04 March 2014]

Available at GitHub:

https://github.com/DASISH/survey-tools-metadata

DASISH 3.2 Versioning of translations in the TMT and the QVDB 2015_01_05.pdf

DASISH 3.2 RationaleCode controlled vocabularies for schemes, questions and variables v1 2015-01-05.xlsx

DASISH 3.2 DDI element hierarchies for questions and concepts, and their attributes v1. 2015_01_05.xlsx

DASISH3.2_PublicationStatusCV-QDDT_1.0_2015-01-05.htm

DASISH3.2_PublicationStatusCV-QDDT_1.0_Genericode_1.0_2015-01-05.xml

DASISH 3.2 PublicationStatusCV-QVDB_1.0_2015-01-05.htm

DASISH3.2_PublicationStatusCV-QVDB_1.0_Genericode_1.0_2015-01-05.xml

Conference papers and posters

http://www.eddi-conferences.eu/ocs/index.php/eddi/eddi13/paper/view/122

http://www.eddi-conferences.eu/ocs/index.php/eddi/eddi13/paper/view/135

http://www.eddi-conferences.eu/ocs/index.php/eddi/eddi14/paper/view/177/91

http://www.eddi-conferences.eu/ocs/index.php/eddi/eddi14/paper/view/166

Appendix 1

Work towards a common metadata model - roles and working groups:

Roles:

Work package leader DASISH WP 3:

Eric Harrison (City)

Consultant:

Joachim Wackerow

Working groups:

Technical group:

Håvard V. Bakkmoen (NSD)
Ricardo Fernandez (UPF)
Edwin de Vet (CentERData)
Maurice Martens (CentERData)
Joachim Wackerow (Consultant)

DDI group:

Håvard V. Bakkmoen (NSD)
Anna de Castellarnau (UPF)
Hilde Orten (NSD)
Yvette Prestage (City)
Edwin de Vet (CentERData)
Joachim Wackerow (Consultant)

Versioning policy group:

Eric Balster (CentERData)
Brita Dorer (GESIS)
Taina Jääskeläinen (FSD)
Hilde Orten (NSD)
Yvette Prestage (City)
Joachim Wackerow (Consultant)

Manager group:

Eric Harrison (City)
Eric Balster (CentERData)
Knut Kalgraff Skjåk (NSD)

Appendix 2

Acknowledgements to the participants that contributed in the discussions:

Participants May meeting 2013

Håvard V. Bakkmoen - NSD
Eric Balster - CentERdata
Brita Dorer - GESIS
Eric Harrison - City University
Taina Jääskeläinen - FSD
Kirstine Kolsrud - NSD
Maurice Martens - CentERdata
Hilde Orten - NSD
Yvette Prestage - City University
Diana Zavala Rojas - UPF
Knut Kalgraff Skjåk - NSD
Joachim Wackerow, Consultant

Participants February meeting 2014

Plenary
Håvard V. Bakkmoen - NSD
Eric Balster – CentERdata
Edwin DeVet - CentERdata
Anna de Castellarnau - UPF
Ricardo Fernandez - UPF
Eric Harrison - City University
Taina Jääskeläinen - FSD
Hilde Orten - NSD
Yvette Prestage - City University
Knut Kalgraff Skjåk – NSD
Fei Su – NSD
Sigbjørn Revheim - NSD
Joachim Wackerow, Consultant

Group 1 (technical)
Håvard V. Bakkmoen - NSD
Edwin DeVet – CentERdata
Ricardo Fernandez – UPF
Fei Su – NSD
Sigbjørn Revheim - NSD
Joachim Wackerow, Consultant

Group 2 (domain)
Eric Balster – CentERdata
Anna de Castellarnau - UPF
Eric Harrison - City University
Taina Jääskeläinen - FSD
Hilde Orten - NSD
Yvette Prestage - City University
Knut Kalgraff Skjåk – NSD
Joachim Wackerow, Consultant

Participants July meeting 2014

Håvard V. Bakkmoen - NSD
Eric Balster - CentERdata
Brita Dorer - GESIS
Eric Harrison - City University
Taina Jääskeläinen - FSD
Maurice Martens - CentERdata
Hilde Orten - NSD
Yvette Prestage - City University
Anna de Castellarnau - UPF
Edwin de Vet - CentERdata
Joachim Wackerow, Consultant