

QUESTIONNAIRE MANAGEMENT AND DDI: THE QDDS CASE



By Oliver Hopt with Alerk Amin, Arofan
Gregory, Jannik Jensen, Dan Kristiansen, and
Mary Vardigan

02/23/2010

DDI Working Paper Series -- Use Cases, No. 5

This paper is part of a series that focuses on DDI usage and how the metadata specification should be applied in a variety of settings by a variety of organizations and individuals. Support for this working paper series was provided by the authors' home institutions; by GESIS - Leibniz Institute for the Social Sciences; by Schloss Dagstuhl - Leibniz Center for Informatics; and by the DDI Alliance.

Questionnaire Management and DDI: The QDDS Case

BY OLIVER HOPT WITH ALERK AMIN, AROFAN GREGORY, JANNIK
JENSEN, DAN KRISTIANSEN, AND MARY VARDIGAN

ABSTRACT

A questionnaire is usually revised several times during construction, and documenting those revisions in detail is important for understanding the lineage of the questions asked and for further questionnaire development in future studies. This documentation is also indispensable for teaching survey research methods.

Nonetheless, this documentation is usually done only partially or not at all, due to a lack of useful documentation tools. The aim of the Questionnaire Development Documentation project was the development of a system to permit permanent electronic documentation of questionnaire development and the final state of the instrument.

QDDS is a useful and effective system but may be improved by moving to DDI 3 to take advantage of enhanced functionality, especially in terms of versioning. DDI 3 also has the potential to provide the framework for a question bank.

BACKGROUND

The first version of QDDS was implemented at the University of Konstanz but is no longer accessible. This was a system to display all versions of a questionnaire as jpg snippets of Word document screenshots put together by a large ASCII property file parsed by a Perl engine. One could move through different versions of the questions and get the history of questions through the process. This was a proof of the concept only, proposed by Rainer Schnell and Frauke Kreuter.¹

It was decided as a next phase to integrate the questionnaire development process, documentation, and the questionnaire itself. This has to be done before archiving by the researcher him- or herself. Software needs to be able to create a questionnaire and to document what changed in subsequent versions and why. One needs a large base of questionnaires to get any advantage from this, and thus, it was necessary to create an application independent of the questionnaire platform – a sort of abstract application and data storage. The developer had experience with XML and thus chose DDI 2 initially to align with standards.

¹ For details see Schnell, R., Kreuter, F.: Neue Software-Werkzeuge zur Dokumentation der Fragebogenentwicklung (2001); in: ZA-Informationen, 48, S. 56-70.

At first QDDS focused on print questionnaires and then moved to CAI systems. Now there is an export to Blaise and other systems including OpenDocument Format, and imports from Voxco, NIPO, and Blaise in the pipeline. Web surveys may be next. There have also been requests to import Word documents.

For single researchers (but not for panel studies), it is sufficient to have a single user client application to create a questionnaire.

Currently, QDDS uses DDI 2.1 plus structured elements, surrounding DDI by structures not found in the standard. It was necessary to extend the standard in specific areas such as study-specific data, formatting, and presets.

Storing everything in one XML structure proved too slow and inefficient, so a decision was made to break up the structure. One must deal with several files to create the history. The change documentation is currently outside of the system. There are several categories of changes, and this element uses a controlled vocabulary. It may be possible to use Version Rationale in DDI 3 to cover the various reasons for change. There is currently a new DDI instance each time a question changes, and the system does not capture who makes a change.

One can request to retrieve what has changed since an earlier version to the present and get a summary, which can be exported as HTML so one can get the same table structure to reuse for documentation purposes and for reports to funders.

Using the versioning and grouping features built into DDI 3 as well as its extension methodology if necessary would restructure QDDS to be more flexible and to implement question bank functionality.

How DDI 3 handles complicated question blocks is one area that must be explored further. It will also be necessary to explore how best to handle resorting of items in DDI 3. This is currently done manually.

There are two ways that the QDDS structure might be handled with DDI 3 Groups. One might have one large group with everything common at the top and deviations/changes indicated in subgroups using the add, replace, delete mechanism in DDI 3 -- that is, a Group that is a stack of Study Units and all but the first is change tracking. Alternatively, each question scheme could be versioned when it changes.

The current architecture of QDDS would make it easy to move to DDI 3.

USE CASE / REQUIREMENTS

There are usually as many arguments for adopting a standard as there are not to adopt it. In our case there were even more on the "pro" side for the usage of DDI in general. The major reason was that the Information Processes in the Social Sciences (IPS) department, which is part of GESIS, cannot advocate the adoption of standards to fulfill its tasks (for example, Dublin Core and OAI for exchange of bibliographic data) and at the same time ignore an existing standard in the field for new software it is developing. Both demanding and adopting standards are driven by the growing need for interoperation and exchange of data with partners, for which standards must be the basic agreement.

When QDDS started with GESIS, DDI 2.1 was available and therefore the format of choice. On that basis the decision was made to build an XML structure to meet the needs for developing a questionnaire and documenting this process. One might wonder why QDDS developers would consider changing QDDS to a new format. The decision to move to DDI 3 focuses mainly on two reasons:

1. The structure of DDI 3 covers a lot more if not all the needs of the QDDS functions, while DDI 2 does not. This means not only aligning with a standard but fulfilling it. Efficiency will improve by using a format in which the content of QDDS is covered.
2. Using DDI 2 had some limitations (mainly the artificial equality between variables and questions) that prohibited a straight reuse of question items and blocks of them and also the straight reuse of the produced documents themselves. So with DDI 2, building up a question bank would require extra operations to prepare elements for importing. On the other hand QDDS documents have to be transformed to make them fit into DDI 2 archival structures.

In addition, the plans of several data archives to switch to DDI 3 make it necessary to deliver in that standard.

POSSIBLE SOLUTIONS and DESIGN CHOICES

Versioning

The main approach for versioning within QDDS is to create new versions of the entire questionnaire for the purpose of documenting distinct events. These events could be pretests, meetings, consultations, and so on. Then, within a version, several questionnaire elements may be added, changed, moved, or deleted, and these elements are then flagged and the flag is displayed in the overview tree. Every change made should then have comments in a specific screen accessible for each element individually. This keeps all decisions and mistake corrections preserved for further usage.

In DDI 2 the storing of versions is not supported directly. Therefore the QDDS structure had to be extended in two ways. The first was to surround the DDI with all information on the version in general, which means number, reason, creation date, and the documentation of structural changes. The second was storing comments with the elements. This was solved by putting some new XML elements inside the notes of the variable structure. And when a new version is created, the entire DDI structure is duplicated and cleaned up along with change flags and comments.

DDI 3 provides this necessary structure in a satisfying way. Several substructures are versionable by themselves so that the main thing is to align the DDI 3 versioning structure with versioning as it is meant in QDDS. Therefore the DDI 3 grouping mechanism can be used, with just one big `g:Group` containing a `dc:QuestionScheme` for each version. In common use cases, only the first of these `dc:QuestionSchemes` will contain the complete questionnaire with subsequent versions referencing the unchanged questions. When there is a change on an element in a given version, the data for that element will be copied and then changed. Possible types of change (called actions) are "Add," "Update," and "Delete". This means that QDDS will not have to duplicate the questionnaire for each version.

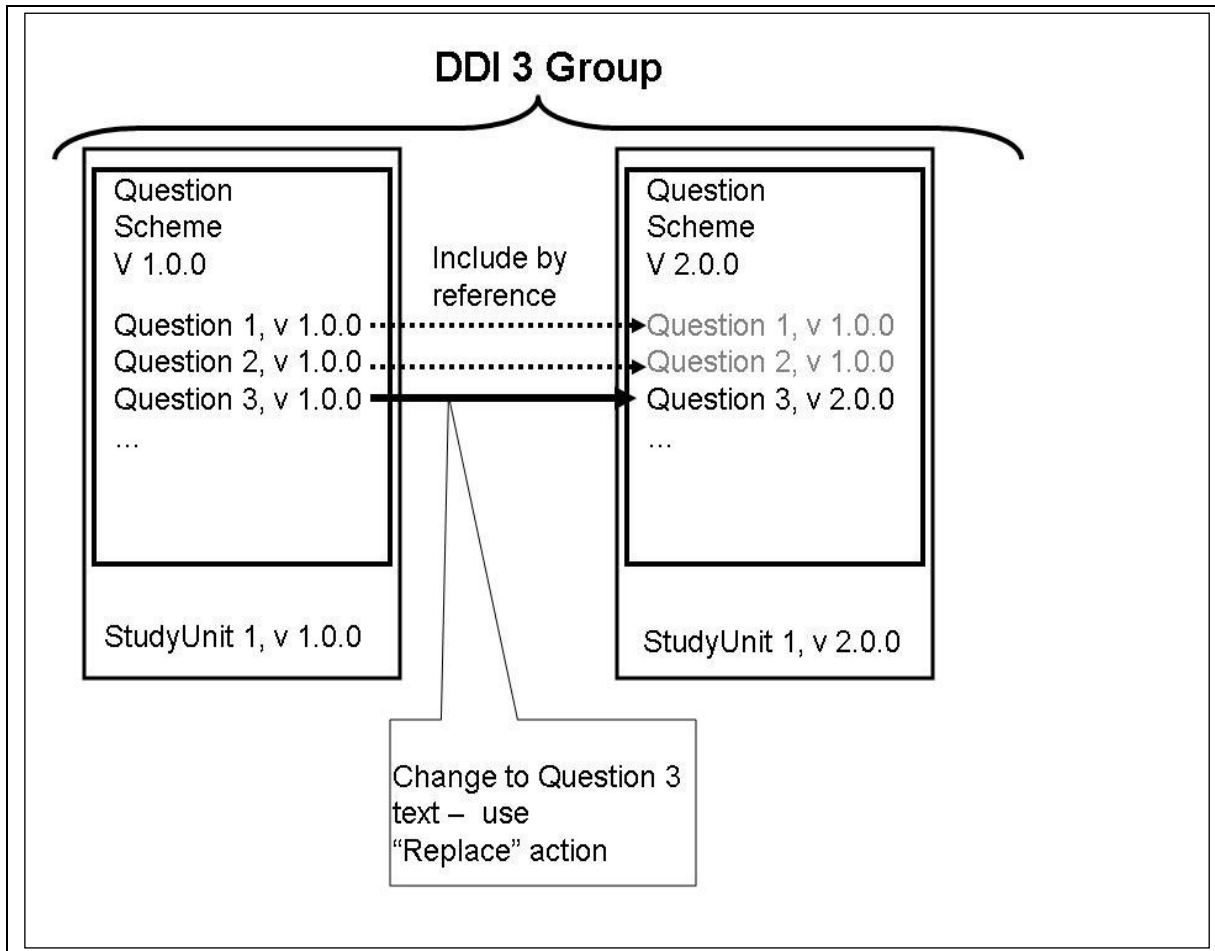


Figure 1: A DDI 3 versioning example

However, within the list of change types an important one is missing, which is "position changed". The solution for this is to separate the `dc:QuestionItems` from their order of appearance in the questionnaire.

One of the core concepts of QDDS is to document not only what has changed but also why the change happened. Therefore a sophisticated mechanism of commenting on changes is needed. This includes the ability first to select general categories of and reasons for the change and second to describe the motivation with some text. All this information can later be accessed through the version history of the given element or the change summary of the entire questionnaire. This information can be stored within DDI 3 as `r:VersionRationale`. Because `r:VersionRationale` currently has no structure to specify any type or other textual identifier for the mentioned general category or reason, a first approach to its textual content will be a key value pair consisting of type identifier, delimiter and comment.

Structure

Currently QDDS supports the arrangement of questions within blocks, which are mainly meant to depend on content. So, for example, a block may be defined to hold all questions on demographic aspects.

The way QDDS renders this in DDI 2 is to create a variable group that references the questions and blocks it contains. But all variables and variable groups exist on the same hierarchical level. As a result of that, the tree

structure of questionnaires, as displayed in QDDS, has to be calculated by evaluating all variable groups for the elements they have to contain. The order of appearance is determined just by the appearance of variables and variable groups within the DDI structure of the questionnaire document.

In DDI 3 there is a mechanism to address this: dc:Instrument in combination with dc:Sequence defines the flow through a questionnaire and thereby the order and appearance of given questions within the questionnaire. This allows one to directly read the questionnaire structure from the XML file to process it. This is first a point of performance and second a point of robustness of the software. But on the other hand this brings up the question of where to store information on the block level like intro or instructions. One way would be to use an instance of dc:QuestionItem that is not intended to be an independent question within the flow and numbering of the questionnaire, but just the information container. This should be declared by a special type of id notation.

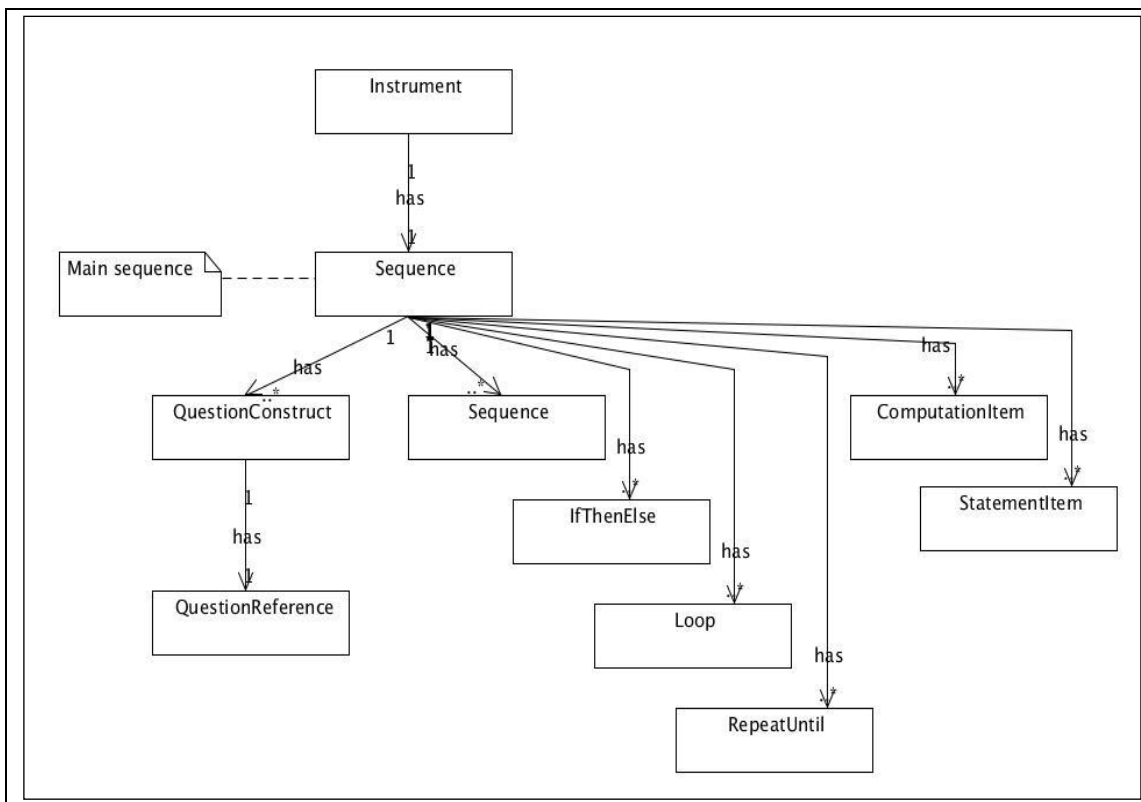


Figure 2: Instrument structure in DDI 3

Reuse

Currently, reusing questionnaire elements in QDDS is managed on the basis of the entire questionnaire file, but portions can be reused. Therefore an existing QDDS file is selected in an import dialog. Within this dialog an overview tree is displayed, similar to the tree in the editor itself. One of the given questionnaire elements then can be selected and imported to the new questionnaire, which is edited at the actual position in the file. If the selected element is a block of questions, the block will be imported with all included subblocks and questions. If questions are imported that will steer a filter in the source questionnaire, sometimes called a “go-to” statement, the destinations will be erased during the import.

To benefit from the questionnaires entered into QDDS files within the project, QDDS offers an information system containing all questions from these questionnaires. Within this repository, one can search by free keywords and by structural and environmental, or contextual, properties of questions. For example, one may search for an open-ended question containing “education” in the text followed by a closed-ended question with two answer defaults. The search result can be displayed as an overview list or in a detailed view. The latter then contains a link where the corresponding QDDS file may be downloaded to be used for import and other purposes.

DDI 3 can clear the way to a more robust and more direct import of items. While question items may now exist separate from variables within the DDI structure, in DDI 3 they can be stored inside a question bank and be referenced to the questionnaire or even questionnaires they appear in. The same is true for sequences / blocks of questions and answer scales. Entire structures can be stored locally by researchers as a favorite set of reused elements. By supporting DDI 3, QDDS will be prepared to implement question banks even before the full functionality is available.

ISSUES / RESTRICTIONS

QDDS was and is still being developed with a limited amount of programming resources. This is obvious when one sees that QDDS is in version 0.7.x after a project period of nearly four years. As a result, some features are still to be improved and others are missing. However, the set of functions in QDDS is already sufficient to seriously support questionnaire development.

Moving QDDS to DDI 3 will require time for implementing a new set of classes on the level of data manipulators. Additionally, some extra features will become possible and are worth putting into practice. Because the QDDS developers suggested moving to DDI 3 support in various presentations and discussions, it will may be possible to bring out this general functionality in coming months. The rest will then depend on future decisions that may impact QDDS.

OUTLOOK / CONCLUSION

The MISSY editor developed for the German microdata census covers variables already and could be extended to questions to make it relevant for QDDS. This would connect parts of the life cycle within DDI 3.

APPENDIX A

The paper is one of several papers which are the outcome of a workshop held at Schloss Dagstuhl - Leibniz Center for Informatics in Wadern, Germany, November 2-6, 2009.

Workshop title:

Workshop on Implementation of DDI 3 - Advanced Topics

Organizers:

Arofan Gregory (Open Data Foundation, Tucson, Arizona, USA)

Wendy Thomas (Minnesota Population Center, University of Minnesota, USA)

Mary Vardigan (Inter-university Consortium for Political and Social Research [ICPSR], University of Michigan, USA)

Joachim Wackerow (GESIS, Leibniz Institute for the Social Sciences, Germany)

Link: <http://www.dagstuhl.de/09452>

This series was edited by Michelle Edwards, Larry Hoyle and Mary Vardigan.

The authors of the paper would like to acknowledge others who participated in this workshop.

Alerk Amin, CentERdata, Tilburg University, the Netherlands

Michelle Edwards, University of Guelph, Canada

Bryan Fitzpatrick, Rapanea Consulting, United Kingdom

Oliver Hopt, GESIS, Leibniz Institute for the Social Sciences, Bonn, Germany

Larry Hoyle, Institute for Policy and Social Research, University of Kansas, USA

Sanda Ionescu, Inter-university Consortium for Political and Social Research (ICPSR), University of Michigan, USA

Jannik Jensen, Dansk Data Archive (DDA), Denmark

Uwe Jensen, GESIS, Leibniz Institute for the Social Sciences, Köln, Germany

Mari Kleemola, Finnish Social Science Data Archive (FSD), University of Tampere, Finland

Dan Kristiansen, Dansk Data Archive (DDA), Denmark

Agostina Martinez, University of Cambridge, United Kingdom

Martin Mechtel, Institute for Educational Progress, Humboldt-Universität zu Berlin, Germany

Olof Olsson, Swedish National Data Service (SND), Sweden

Ørnulf Risnes, Norwegian Social Science Data Services (NSD), Norway

Wolfgang Zenk-Möltgen, GESIS, Leibniz Institute for the Social Sciences, Köln, Germany

APPENDIX B

Copyright © DDI Alliance 2009, *All Rights Reserved*

<http://www.ddialliance.org/>

Content of this document is licensed under a Creative Commons License:
Attribution-Noncommercial-Share Alike 3.0 United States

This is a human-readable summary of the Legal Code (the full license).

<http://creativecommons.org/licenses/by-nc-sa/3.0/us/>

You are free:

- to Share - to copy, distribute, display, and perform the work
- to Remix - to make derivative works

Under the following conditions:

- Attribution. You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).
- Noncommercial. You may not use this work for commercial purposes.
- Share Alike. If you alter, transform, or build upon this work, you may distribute the resulting work only under the same or similar license to this one. For any reuse or distribution, you must make clear to others the license terms of this work. The best way to do this is with a link to this Web page.
- Any of the above conditions can be waived if you get permission from the copyright holder.
- Apart from the remix rights granted under this license, nothing in this license impairs or restricts the author's moral rights.

Disclaimer

The Commons Deed is not a license. It is simply a handy reference for understanding the Legal Code (the full license) — it is a human-readable expression of some of its key terms. Think of it as the user-friendly interface to the Legal Code beneath. This Deed itself has no legal value, and its contents do not appear in the actual license.

Creative Commons is not a law firm and does not provide legal services. Distributing of, displaying of, or linking to this Commons Deed does not create an attorney-client relationship. Your fair use and other rights are in no way affected by the above.

Legal Code:

<http://creativecommons.org/licenses/by-nc-sa/3.0/us/legalcode>