DASISH

DATA SERVICE INFRASTRUCTURE FOR THE SOCIAL SCIENCES AND HUMANITIES

The DASISH Question Variable Data Base: Developments and Outlook

Håvard Venge Bakkmoen, Hilde Orten, Benjamin Beuster

NSD - Norwegian Social Science Data Services

EDDI 2014 - 6th Annual European DDI User Conference, December 2 – 3 2014

SEVENTH FRAMEWORK PROGRAMME

DASISH

DATA SERVICE INFRASTRUCTURE FOR THE SOCIAL SCIENCES AND HUMANITIES

Outline

- The Question Variable Data Base – purpose, aims and requirements
- QVDB – production pipeline, example use-cases
- Current structure and outlook for future developments

SEVENTH FRAMEWORK PROGRAMME

DASISH

DATA SERVICE INFRASTRUCTURE FOR THE SOCIAL SCIENCES AND HUMANITIES

The Question Variable Data Base (QVDB)

- Purpose:
  - Searchable database with broad public profile with user access to survey questions in original languages, concepts, variables etc.

- Primary aims:
  - To serve business processes of the ESS and other surveys
  - Browsing possibilities for researchers and students
  - Interoperability with other systems and tools
  - Reusable model and code

- Possible users (same as for the QDDT):
  - European Social Survey (ESS) is the usecase
  - Other DASISH survey projects
  - Projects outside DASISH
  - Researchers and students

SEVENTH FRAMEWORK PROGRAMME

QVDB (and QDDT), a selection of the requirements

- DDI - Lifecycle based storage structure for metadata elements of high level of granularity
  - multilinguality
  - study-independent components
- DDI - Lifecycle and DDI-Codebook export and import possibilities to/from tools/web-services
- Communication between the three tools should be possible; minimal human interaction
- Support boolean field level search
- Core module based on DDI components; DDI profile; Resource Packages
- Open source system
- User access rights

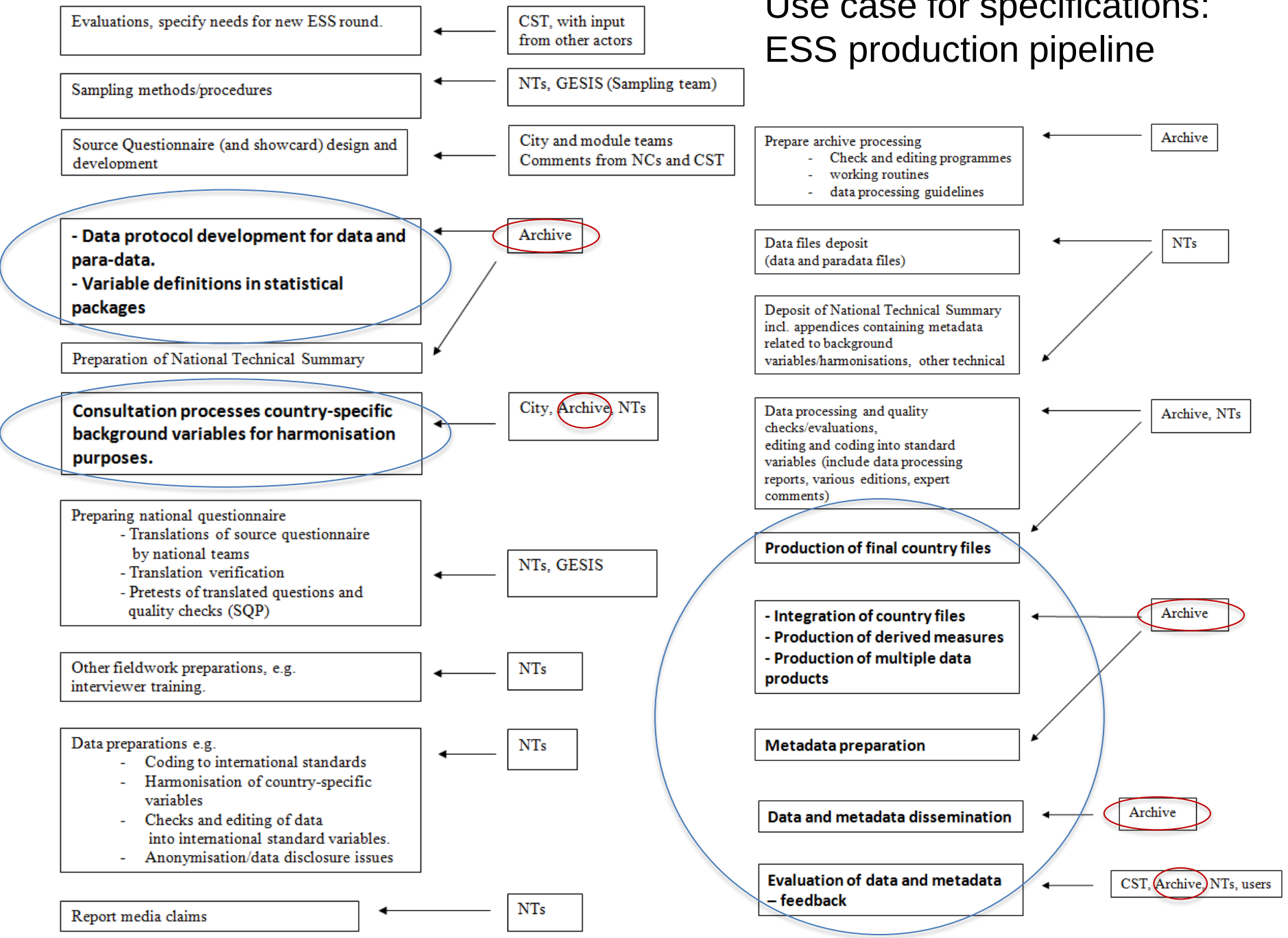
QVDB in the archive production pipeline

- The QVDB is under development for use in the archive business processes at different steps in the pipeline:
  - Creating new variables
  - Creating data protocols and variable specifications for statistical packages
  - Documenting harmonisations etc.

- The QVDB will support import of questions from the QDDT and translations from the TMT

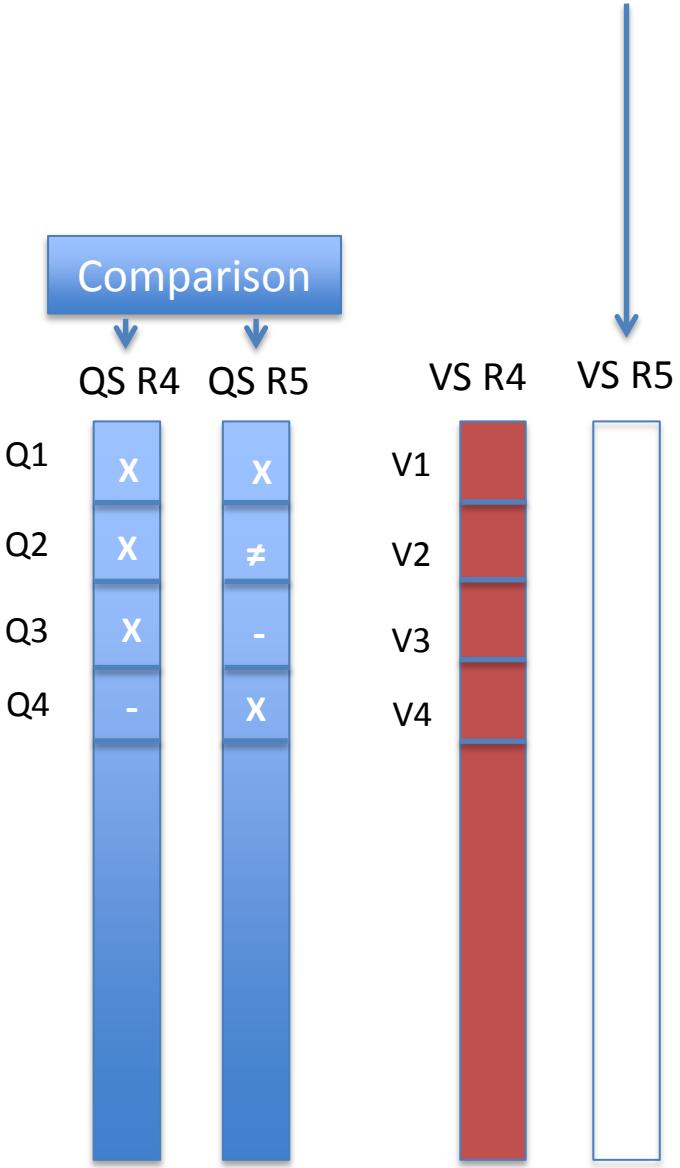
- This facilitates creating new variables as well as documenting them

Use case for specifications: ESS production pipeline



QVDB: possible usecase – archive processes:

Create new variables for a questionnaire module for a new ESS round



1) View source questions in the module of the current round (QS R5)
2) View comparison between question list for module in the current round and that of the former round (Comparison QS R4 and QS R5)

a) No changes in question list:

- View corresponding variable list and included elements
- Reuse variable list used in former round (VS R4) or update if necessary

b) Questions deleted from list:

- Delete and update corresponding variables from variable list used in former round (VS R4)

c) Questions added to list:

- Check if the question has been used earlier and if so if reusable corresponding variables exist in the database.
- Reuse, update or create new variable.

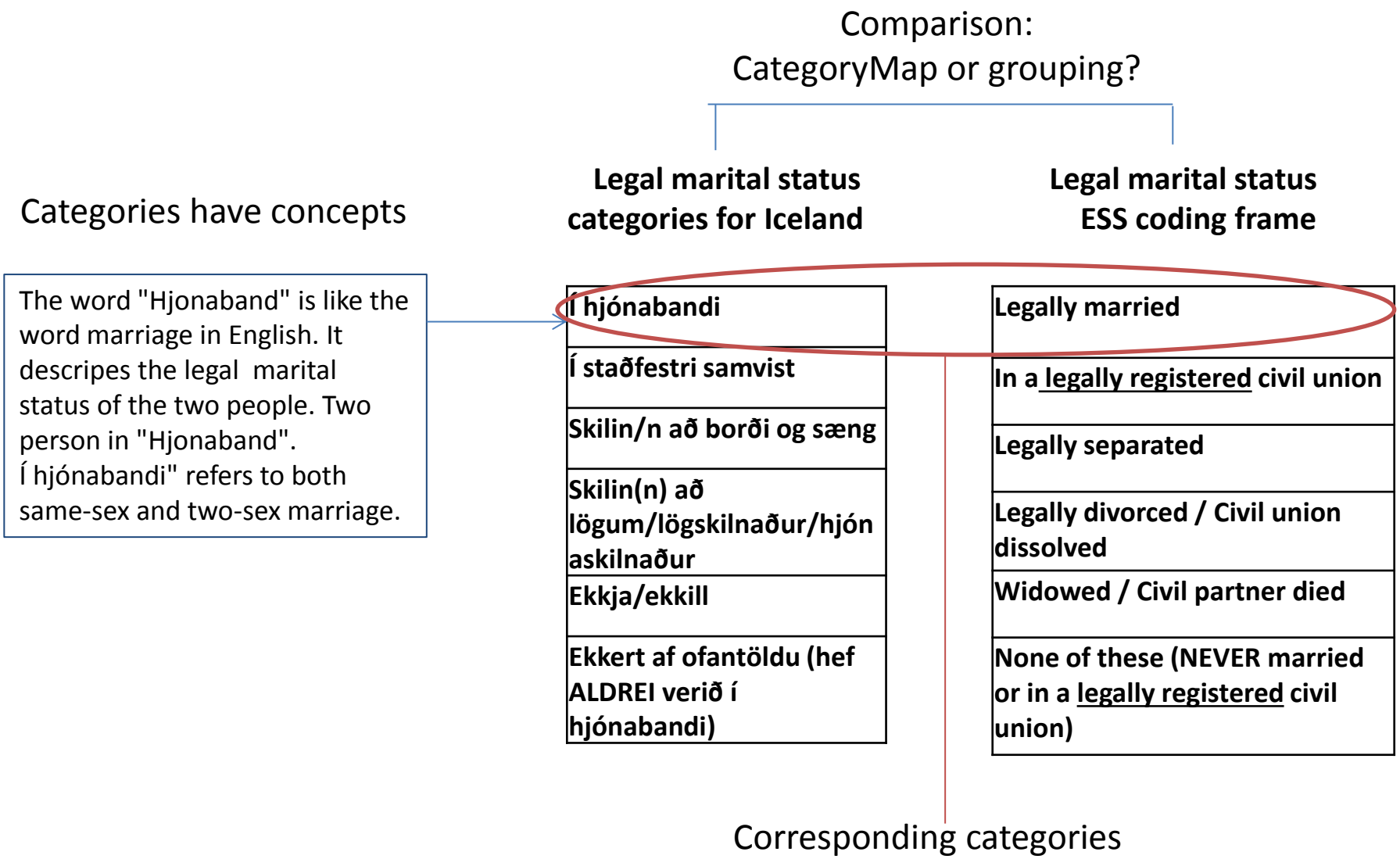
d) Questions changed:

- Check earlier versions of the question and whether corresponding variables exist that could be reused.
- Reuse, update or create new variable.

A new variable could have a based-on reference to a variable it is based on

QVDB: Aim - to support structuring of coding schemes and harmonisations

Example: ESS coding scheme for legal marital status



QVDB: possible usecase – archive processes (2):

Development and reuse of categories from the ESS coding scheme for legal marital status in response domains and variable representations

Legal marital status variable, Iceland

MARSTIS	1 Í hjónabandi
	2 Í staðfestri samvist
	3 Skilin/n að borði og sæng
	4 Skilin(n) að lögum/lögskilnaður/hjónaskilnaður
	5 Ekkja/ekkill
	6 Ekkert af ofantöldu (hef ALDREI verið í hjónabandi)
	66 Not applicable
	77 Refusal
	88 Don't know
	99 No answer

3) Categories reused in VariableRepresentations in QVDB

Legal marital status variable, ESS

MARSTS	1 Legally married
	2 In a legally registered civil union
	3 Legally separated
	4 Legally divorced/civil union dissolved
	5 Widowed/civil partner died
	6 None of these (NEVER married or in legally registered civil
	66 Not applicable
	77 Refusal
	88 Don't know
	99 No answer

Legal marital status categories for Iceland	**Legal marital status ESS coding frame**
Í hjónabandi	**Legally married**
Í staðfestri samvist	**In a legally registered civil union**
Skilin/n að borði og sæng	**Legally separated**
Skilin(n) að lögum/lögskilnaður/hjónaskilnaður	**Legally divorced / Civil union dissolved**
Ekkja/ekkill	**Widowed / Civil partner died**
Ekkert af ofantöldu (hef ALDREI verið í hjónabandi)	**None of these (NEVER married or in a legally registered civil union)**

1) Categories developed in the QVDB

Country-specifie question, Iceland

ÞEIR SPURÐIR SEM BÚA EKKI MEÐ EIGINMANNI / EIGINKONU / MAKA OG ERU Í SAMBÚÐ.

F11 **SPJALD 50** Þessi spurning er um lagalega hjúskaparstöðu þína en ekki með hverjum þú býrð eða býrð ekki. Hver er af eftirfarandi lýsingum á þessu spjaldi lýsir lagalegri hjúskaparstöðu þinni eins og staðan er nú?

SKRÁIÐ AÐEINS EITT SVAR: EFSTA MÖGULEGA SVARIÐ (til dæmis ef svarandi segist vera í hjónabandi (01) og fráskilinn (04) þá á að skrá 01)

Í hjónabandi	01
Í staðfestri samvist	02
Skilin(n) að borði og sæng	03
Er skilin að lögum / lögskilnaður / hjónaskilnaður	04
Ekkja / ekkill	05
Ekkert af ofantöldu (Hef ALDREI verið í hjónabandi)	06
(Veit ekki)	88

2) Categories copied to TMT and QDDT and reused in questionnaire response domains

Source question, ESS

ASK IF NOT LIVING WITH A HUSBAND / WIFE / PARTNER OR ARE COHABITING

F11 **CARD 50** This question is about your legal marital status not about who you may or may not be living with. Which one of the descriptions on this card describes your legal marital status now?

CODE ONE ONLY: PRIORITY CODE[101]

Legally married	01
In a legally registered civil union	02
Legally separated	03
Legally divorced / Civil union dissolved	04
Widowed / Civil partner died	05
None of these (NEVER married or in legally registered civil union)	06
(Don't know)	88

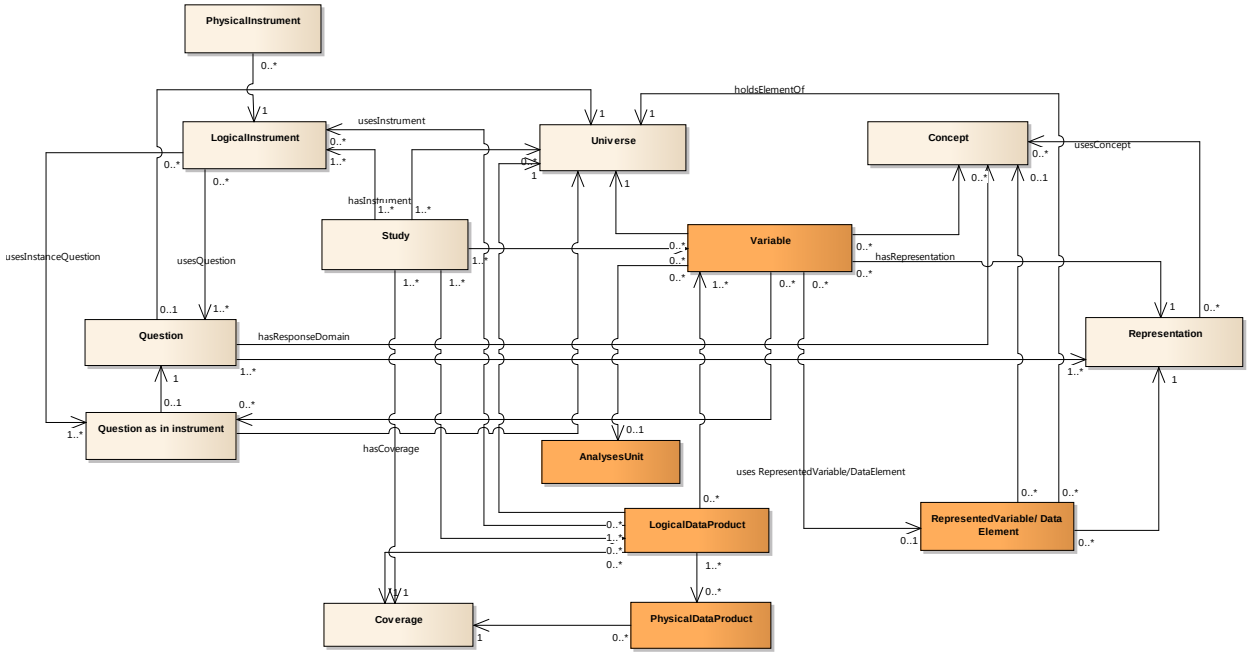
QVDB: Possible usecases - researchers and students

A researcher wishes to make an evaluation of the generated variables of a survey. He wishes to view all generated variables, as well as the syntaxes used to produce them	Find generated variables and the syntaxes they are based on Make PDF of search results
A student wishes to see which questions, variables and concepts have been used in a particular survey	Find questions, variables and concepts used in study, e.g. question lists, variable lists, concept lists, as well as information on single questions, variables and concepts and relations between them. Make PDF of search results
A methodologist wishes to see how repeated metadata components in a survey, like questions and variables have developed over time	Find all versions of the metadata element of interest (question; variable) Views any comparisons made between different versions of the metadata element, e.g. variables. Make PDF of search results
A researcher wishes to explore the coding standards used in a survey	Find and view the summary of DDI Resource Package Find Resource Package used in survey Create user defined enhanced report based on the content of the RP
A researcher wishes to import questions in multiple languages into a different system than the QVDB for reuse	Find questions in multiple languages Export DDI-XML instance or fragment from QVDB to different system
A researcher wishes to import questions in multiple languages into a questionnaire delivery tool	Questions in multiple languages structured in detailed DDI is exported from the QVDB to questionnaire delivery tool

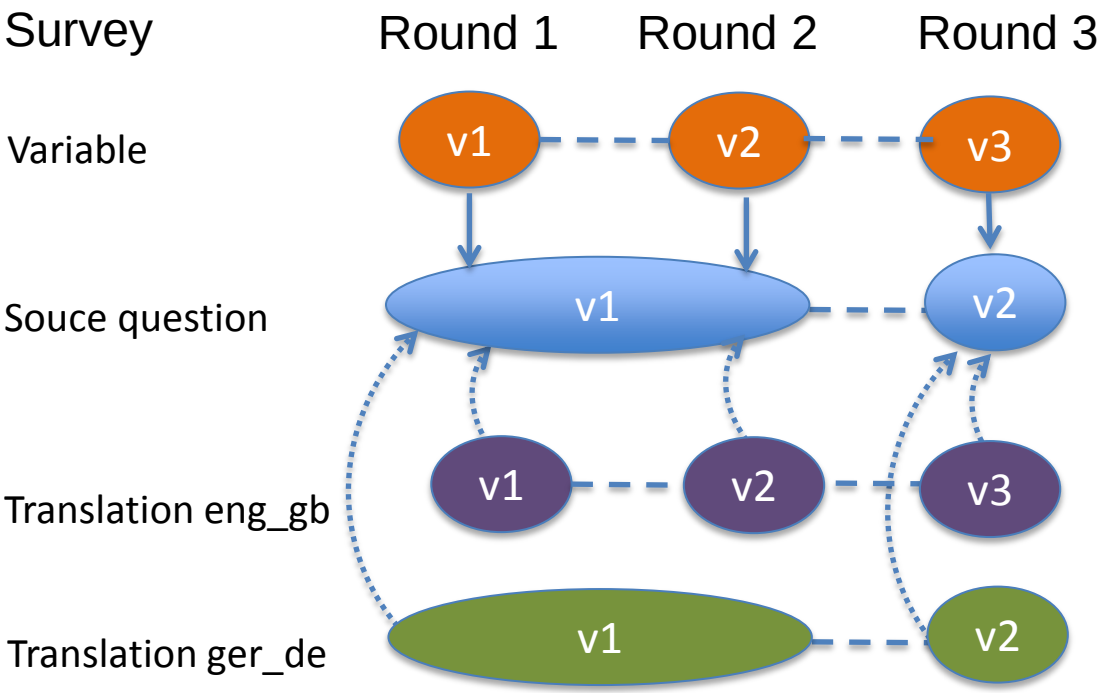
QVDB Structure

- The QVDB builds on the database structure of the QDDT, but the QVDB also includes variables

- The conceptual model is based on DDI-Lifecycle



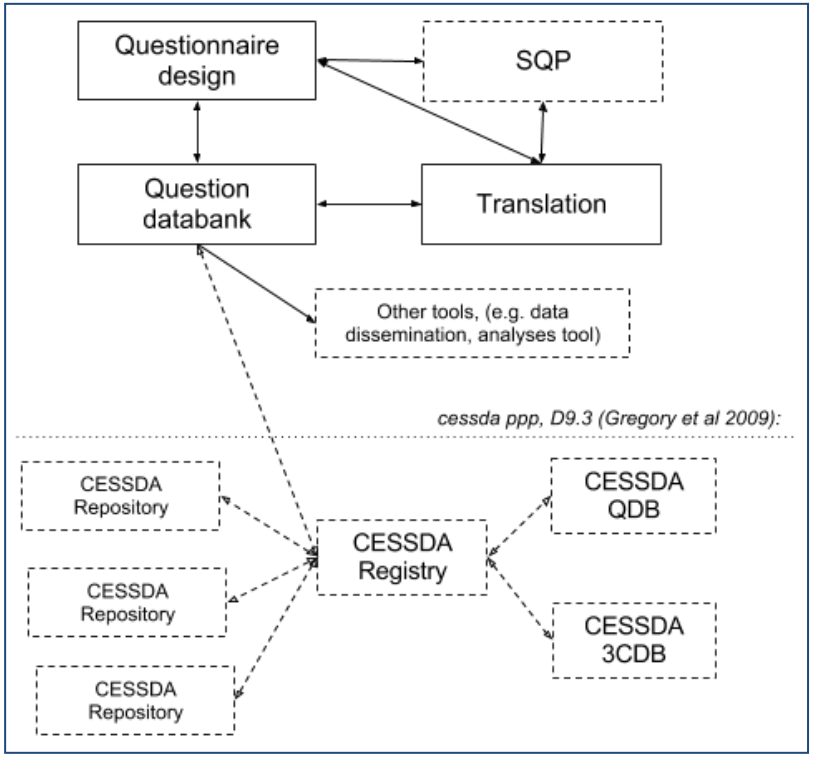
QVDB: Relationships between variables and questions



QVDB – future outlook

Same as for the QDDT:

- Further development
- Real life use by ESS and SHARE
- Real life use by other survey projects
- Interoperability with other DDI-L ifecycle based and compatible survey tools



The QVDB:

- Include functionality to support further archive business steps, most importantly data curation

- The database could serve as a repository in a possible CESSDA DDI-Lifecycle based network (CESSDA portal)

DASISH

DATA SERVICE INFRASTRUCTURE FOR THE SOCIAL SCIENCES AND HUMANITIES

Acknowledgements

Special thanks to Joachim Wackerow for consultancy

Thank you for your attention!

havard.bakkmoen@nsd.uib.no
hilde.orten@nsd.uib.no
benjamin.beuster@nsd.uib.no

SEVENTH FRAMEWORK PROGRAMME