

Specification and immediate visualization of a questionnaire, a metadata approach for a mixed-mode data collection infrastructure

Benoît Werquin (Insee, France)

benoit.werquin@insee.fr

Abstract

Insee has been developing during the last years a web based survey platform for the data collection with a meta-driven approach. The first step is the possibility of responding online via a single web portal: the Coltrane Project (acronym for “COLlecte TRANsversale d’Enquêtes” in French, Cross-Disciplinary Survey Data Collection). Coltrane is a data collection platform, currently used for web establishment surveys. It is composed of several modules including Eno, a metadata-driven questionnaire generator.

This tool takes as input a formal description of the questionnaire complying with the DDI standard and executes a completely automated chain of transformations to produce the current survey questionnaire (web and paper questionnaires). The questionnaires can be personalized and, depending on the output format, a variable amount of flow logic can be implemented. Thanks to the automation of the process, Eno referring to DDI standard allows traceability of the changes in the questionnaires and reinforces the soundness of the data collection process.

Currently, nearly 15 survey questionnaires were produced by Eno over the last two years, and this figure is expected to grow to more than 40 survey questionnaires by the year 2020. Future work will allow these numbers to grow significantly with Eno providing interviewer questionnaire.

An additional module to this generator is a questionnaire design user interface, codenamed Pogues, that connects with the generation process. More specifically, Pogues produces the DDI description of the questionnaire which is then submitted to an embedded instance of Eno. Using this tool, a survey manager or questionnaire designer can specify his web questionnaire in a friendly way and visualize the generated result in one click.

The current version supports the main functionalities needed for business surveys, and the roadmap foresees the development of more complex logic flows that can for example be found in household surveys. Other future developments will enhance the possibilities to specify questionnaire controls that could be used during the data collection process as well as during the data editing and imputation processes.

Pogues is open source and natively internationalized, and in consequence can be used directly by other statistical agencies. Other national statistical institutes have already shown their interest in this software.

The next step is to build a mixed-mode data collection infrastructure with the same metadata approach: the Metallica Project (acronym for “METadonnées Actives et Logiciels Libres et Infrastructure pour une Collecte Assistée” in French, Actives METadata and Free Software for a Data Collection Infrastructure). The main principle for questionnaire generation is to start from a single description of this questionnaire, regardless of the mode. The same questionnaire description must be able to be implemented in different modes: a web questionnaire, a paper questionnaire, an interviewer questionnaire or another format if necessary.

As the Coltrane project began a few years ago, the main challenges were:

- to extend, modernise, industrialize and standardise mixed-mode data collection for all households surveys (using the same tools that business surveys) ;
- to set up a metadata-driven approach. It reduces development costs, permits to have the documentation of the survey always in sync with the questionnaire in production. And beyond building collection tool, this principle could be implemented on other sub-processes in the GSBPM ;
- to provide a single access point for the household and the respondent easier ;

A full meta-driven approach allow to have a standardized approach to statistics production process while increasing data quality and reducing workload.

References

- [1] T. Dubois, “Specification and immediate visualization of a questionnaire – a metadata-driven approach”.
- [2] A. Hussein-Skalitz, “Modernisation and industrialisation of the online data collection”, NTTS 2017
- [3] T. Dubois, “Pogues & Eno generator : Specify and view a web questionnaire generated from its formal description”, International Conference on Establishment Surveys (2016).

Keywords: survey, data collection, questionnaires generator, mixed-mode, metadata-driven approach, innovation, DDI