

End User Evaluation of the FAIR4Health Data Curation Tool

Mert GENCTURK^{a,b,1}, Alper TEOMAN^{a,b}, Celia ALVAREZ-ROMERO^c,
Alicia MARTINEZ-GARCIA^c, Carlos Luis PARRA-CALDERON^c,
Beatriz POBLADOR-PLOU^d, Matthias LÖBE^e and A. Anil SINACI^a

^aSRDC Software Research & Development and Consultancy Corp., Ankara, Turkey

^bDept. Of Computer Engineering, Middle East Technical University, Ankara, Turkey

^cGroup of Research and Innovation in Biomedical Informatics, Biomedical Engineering and Health Economy, Institute of Biomedicine of Seville (IBiS), Virgen del Rocío University Hospital, CSIC, University of Seville, Seville, Spain

^dEpiChron Research Group, Aragon Health Sciences Institute (IACS), IIS Aragón, Zaragoza, Spain

^eInstitute for Medical Informatics (IMISE), University of Leipzig, Leipzig, Germany

Abstract. The aim of this study is to build an evaluation framework for the user-centric testing of the Data Curation Tool. The tool was developed in the scope of the FAIR4Health project to make health data FAIR by transforming them from legacy formats into a Common Data Model based on HL7 FHIR. The end user evaluation framework was built by following a methodology inspired from the Delphi method. We applied a series of questionnaires to a group of experts not only in different roles and skills, but also from various parts of Europe. Overall, 26 questions were formulated for 16 participants. The results showed that the users are satisfied with the capabilities and performance of the tool. The feedbacks were considered as recommendations for technical improvement and fed back into the software development cycle of the Data Curation Tool.

Keywords. Software Evaluation, User-Computer Interaction, Software Ergonomics, FAIR Data, Data Transformation, Data Curation, HL7 FHIR

1. Introduction

Data sharing is a major challenge in healthcare and this challenge is fundamental for nearly all further research using health data such as data mining and machine learning. The FAIR (findability, accessibility, interoperability, and reusability) Principles [1] were developed to aid data sharing not only in healthcare but also in all domains where data need to be shared. Since FAIR Principles lay out general foundations, there are efforts [2] to provide domain specific, specialized implementation guides to achieve FAIRness in healthcare, specifically health research.

FAIR4Health [3], is a European Commission supported Horizon 2020 research project which aims to apply the FAIR Principles to health research data and then perform data mining algorithms on these FAIRified datasets of different health research

¹ Corresponding Author, Mert Gencturk, SRDC Software Research & Development and Consultancy Corp., ODTU Teknokent Silikon Bina K1-16 06800 Cankaya, Ankara, Turkey; E-mail: mert@srcd.com.tr.

organizations. Data Curation Tool is one of the software components described in the FAIRification Workflow [2] of the FAIR4Health project. The tool has been developed to help transform raw health data into a standard format, HL7 FHIR [4], utilizing FHIR profiling so that FAIRness can be achieved. It is available as open-source software [5].

Data Curation Tool is a highly specialized Extract-Transform-Load tool which can extract data from relational databases and spreadsheets, applies user-defined transformations and then loads the transformed resources into an HL7 FHIR repository. The tool communicates with an HL7 FHIR repository to process the available FHIR profiles and helps its users apply transformations accordingly, in the sense that FAIR Principles are achieved once data is loaded to the repository as described in [2].

This study aims to build an evaluation framework for the user-centric testing of the Data Curation Tool to receive early end user feedback and improve not only the tool itself, but also its role in the FAIRification workflow. We built the evaluation framework inspired by the Delphi method [6,7]. A series of questionnaires were prepared and sent out to a set of experts with different roles from different regions of Europe. The results were compiled and fed back into the software development cycle for the Data Curation Tool to improve it based on end user feedback.

2. Methods

In the scope of the FAIR4Health project, first, the requirements for the Data Curation Tool were defined and the tool was designed as a set of subcomponents satisfying those requirements. Based on the architecture of the tool, five subcomponents were implemented, and the requirements were categorized around them: Data Source Analyzer, Metadata Mapper, Terminology Linker, Data Transformer and Data Validator. “The system shall support connection with various data formats, such as Excel, CSV file, relational database (SQL)” is an example of an elicited requirement of the Data Source Analyzer, while “The system shall provide mechanisms to map source data elements to the target fields in HL7 FHIR profiles” is a requirement for the Metadata Mapper.

After the development of the Data Curation Tool, the end user evaluation framework was built by following a methodology inspired from the Delphi method. The objective of Delphi method is to obtain the most reliable consensus of a group of experts by subjecting them to a series of questionnaires [8]. In our framework, we conducted two rounds of questionnaires by following the four steps below: first-round questionnaire, identification of experts, second-round questionnaire, end-user evaluation.

Based on the analysis of requirements, the first step was to generate an initial list of questions aiming to understand the broad view of experts, hence the questions at this step were kept as generic as possible, such as “Were you able to install and run the tool on your computer?”. The questionnaire was conducted to only one IT expert who is not one of the developers of the Data Curation Tool, but has an extensive knowledge on the data curation process and the concepts such as SQL, healthcare standards, code systems etc.

As the Delphi method relies on a panel of experts, before conducting the second-round questionnaire, we identified 16 experts not only in different roles and skills, but also from various parts of Europe (7 countries, 10 organizations). Data scientists, health research scientists, IT professionals and clinicians were included in the expert panel from the following organizations and countries from the FAIR4Health project:

- Six pilot sites: (i) Andalusian Health Service, Spain; (ii) Aragon Health Sciences Institute, Spain; (iii) Geneva University Hospital, Switzerland;

(iv) University of Porto, Portugal; (v) Catholic University of the Sacred Heart, Italy; (vi) Institute for Pulmonary Diseases of Vojvodina, Serbia.

- Two universities: (i) Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Germany; (ii) Peter L. Reichertz Institute for Medical Informatics, University of Braunschweig, Germany
- Two private companies: (i) SRDC Software Research & Development and Consultancy Corporation, Turkey; (ii) ATOS Spain SA, Spain.

Based on the answers to the first-round questionnaire, we identified the incomplete or erroneous system functionalities, and modified, deleted, or added questions to be able to go deeper and clarify the issues. This constituted the final list of questions in two categories: (1) Yes/No questions, such as “Were you able to fetch the FHIR resource types and profiles successfully?”, (2) Scoring questions through Likert scale (0 to 5), such as “The mapping procedure was easy”. The questions were also grouped by the corresponding subcomponent, e.g., Metadata Mapper. Finally, the questionnaire was made available online through Microsoft Forms [9] and sent out to the identified experts.

During the end-user evaluation phase, the experts received an executable of the Data Curation Tool since it is a standalone desktop application, training materials including a user guide and video tutorial explaining step-by-step activities on the demonstrators, and a sample data to be used to get familiar with the tool. Responses to the questionnaires were monitored through Microsoft Forms and once they are completed the analyses were performed to see the end user evaluation results.

3. Results

26 questions were formulated in five categories associated with the five subcomponents of the Data Curation Tool. For each category, a number of Yes/No questions examining users’ success on performing a task, and scoring questions evaluating the easiness and understandability of tasks were asked to the experts. This section presents the detailed results including the responses of participants to “Yes”, “No” and “N/A” (for the functionalities that are Not Applicable for a specific participant), average results for the scoring questions as well as examples of formulated questions. Table 1 shows a summary of the response rates to the Yes/No questions and average scores for the scale questions.

Table 1. Responses and average scores grouped by subcomponents of the Data Curation Tool

Subcomponent	“Yes”	“No”	“N/A”	Score
Data Source Analyzer	97%	3%	0%	4.5
Metadata Mapper	88%	7%	5%	3.68
Terminology Linker	44%	37%	19%	3.67
Data Validator	100%	0%	0%	3.56
Data Transformer	88%	12%	0%	3.73

The **Data Source Analyzer** related questions received a high number of “Yes” responses (97%). The questions related to the connection with source data, such as “Were you able to connect your data source?” and “Were you able to upload your data?” were responded positively by all (100%) participants. The only “No” responses (3%) were retrieved on the question “Were you able to connect to the FHIR endpoint successfully?” indicating that although the process of data source connection was easy and worked fine in the majority, some improvements can still be made on FHIR endpoint connection part.

For **Metadata Mapper**, average responses were 88% for “Yes”, 7% for “No”, and 5% for “N/A”. The participants responded “Yes” to all the questions (e.g., “Were you able to fetch the FHIR resource types and profiles successfully?”, “Were you able to match the attributes in your source file with FHIR resource attributes successfully?”) except for “Were you able to assign default value(s) for the FHIR attributes that are required in the profile but do not appear in your source data?”. This feedback told us that although this feature was not needed for some of the users (no empty values in their data), it was needed for most of them, but half of them had a difficulty in default value assignment process. Apart from this, the mapping procedure was found easy and understandable by the majority with a high score (4.1).

The **Terminology Linker** related questions received the lowest “Yes” responses (44%). The questions in this category were related to linking source code systems to target code systems (i.e., SNOMED, ICD-10), that is, mapping the coded fields. The two lowest “Yes” question score was with “Were you able to find the source code system?” and “Were you able to annotate the data elements with code from terminologies?”. The users were able to connect to the Terminology Server, but they were not able to find the desired source code system, hence could not perform the terminology linking. When we asked the participants to provide further details on this issue, it appeared that they were using proprietary code systems in their source data, hence it must have been defined in the Terminology Server before the curation process. Otherwise, the participants gave good score (3.66) to “Terminology linking procedure was easy”.

88% of the participants gave “Yes” answers to the questions related to **Data Transformer**. It is a matter of course that some data may not be transformed during transformation process. In the scoring questions, the participants highlighted the insufficiency of information shown for the data which could not be transformed with a score of 3.0, but they gave the highest score (4.57) among the others for “Transformation of raw data to FHIR was fast enough”, which is a crucial positive feedback for the tool.

The **Data Validator** related questions, e.g., “Were you able to validate the mappings”, received the highest “Yes” responses (100%). The only feedback we received on Data Validator from the scoring questions was “The details of validation operation was not understandable all the time” (low score of 2.87) indicating that providing further details on the results of validation process would be useful for the users.

4. Discussion

We recruited 16 experts with relevant expertise covering the research and technology spectrum for FAIRification of health data from different organizations spread around Europe. Together with the satisfactory evaluation results coming from these experts, we can strongly conclude that the FAIR4Health Data Curation Tool addresses major challenges in terms of health data FAIRification and meets fundamental requirements.

We followed a sound and methodological approach to build our evaluation framework which helped us produce systematic and comparable results for different subcomponents and features of the Data Curation Tool. We applied two-round questionnaires; further rounds with additional and improved questions could help us to get more detailed feedback from the experts which can be listed as a future work for us.

The results showed that the Data Source Analysis and (Meta)Data Mapping processes are easy and understandable, and most of the users were able to perform these steps successfully. Furthermore, the speed of transformation of raw data to FHIR was

appreciated by the experts. No shortcomings were identified in Data Transformer and Data Validator except for the insufficient details in the GUI.

We observed that most of the “No” and low-scored responses to the questionnaire were due to lack of information in the GUI for the results of validation and transformation operations. The participants found the presented details imperfect. In addition, it was observed that terminology linking process should be improved to enable users to define and find proprietary code systems in an easier way. In this regard, instead of selecting source and target code systems from a list of URIs, an advanced text-based search mechanism utilizing the name of code systems can be implemented.

5. Conclusions

In this study, we explained how the user-centric testing of the FAIR4Health Data Curation Tool was performed by a set of experts with different roles from different regions of Europe. We presented our evaluation framework built on top of Delphi method, where a series of questionnaires were sent out to the experts.

The aim of this study was not only to evaluate the functionality of Data Curation Tool against its requirements, but also to explore how to improve it further based on the experts’ feedbacks. Overall, the evaluation was performed successfully with satisfactory results. The feedback was retrieved from the experts, and the results were compiled and fed back into the software development cycle of the Data Curation Tool.

6. Acknowledgements

This work was performed in the framework of FAIR4Health project. FAIR4Health has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement number 824666.

References

- [1] Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten JW, da Silva Santos LB, Bourne PE, Bouwman J. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*. 2016 Mar 15;3(1):1-9.
- [2] Sinaci AA, Núñez-Benjumea FJ, Gencturk M, Jauer ML, Deserno T, Chronaki C, Cangiolli G, Caverro-Barca C, Rodríguez-Pérez JM, Pérez-Pérez MM, Erturkmen GB. From Raw Data to FAIR Data: The FAIRification Workflow for Health Research. *Methods of Information in Medicine*. 2020 Jun;59(S 01):e21-32.
- [3] FAIR4Health Project Website. Available from: <https://www.fair4health.eu/>
- [4] HL7 FHIR. Available from: <http://hl7.org/fhir/>
- [5] Data Curation Tool. Available from: <https://github.com/fair4health/data-curation-tool>
- [6] Okoli C, Pawlowski SD. The Delphi method as a research tool: an example, design considerations and applications. *Information & management*. 2004 Dec 1;42(1):15-29.
- [7] Linstone HA, Turoff M, editors. *The delphi method*. Reading, MA: Addison-Wesley; 1975 Dec.
- [8] Dalkey N, Helmer O. An experimental application of the Delphi method to the use of experts. *Management science*. 1963 Apr;9(3):458-67.
- [9] FAIR4Health User Acceptance Questionnaire of Data Curation Tool. Available from: <https://forms.office.com/Pages/ResponsePage.aspx?id=dLDIxFSvR0y7jjPuiKzmjooon8KP6AWVDuL2ji8m4dltUOE9KRzVGQIFMUUzVDMwTFBCUFYwSDFXRCQIQCN0PWcu>