

Towards Precise Descriptions of Medical Free/Libre and Open Source Software

Franziska JAHN^{a,1}, Michelle BINDEL^a, Konrad HÖFFNER^a, Maryam GHALANDARI^a, Birgit SCHNEIDER^a, Sebastian STÄUBERT^a, Verena DORNAUER^b, Thomas KAROPKA^c, Elske AMMENWERTH^b, Alfred WINTER^a
^a*Institute of Medical Informatics, Statistics and Epidemiology, Leipzig University*
^b*Institute of Medical Informatics, UMIT – Private University for Health Sciences, Medical Informatics and Technology, Hall in Tirol, Austria,*
^c*Medfloss.org, Germany*

Abstract. The web portal Medfloss.org lists over 360 medical free/libre and open source software (MEDFLOSS) projects. These projects are described with the help of a self-developed nomenclature. Due to inconsistencies, the nomenclature shall be replaced by HITO, the Health IT Ontology. HITO is developed iteratively based on different use cases. This paper aims to describe methods and results of the second HITO use case in which HITO is extended to improve the description, retrieval and comparisons of MEDFLOSS projects on Medfloss.org. We use a mixed-methods approach to add concepts and relationships to describe MEDFLOSS precisely. The resulting HITO version stresses functional descriptions based on features and supported enterprise functions, rather than just describing technical characteristics. However, describing a larger number of MEDFLOSS projects requires the commitment of the community.

Keywords. Health Information Systems, Software, Semantics

1. Introduction

Medical free/libre and open source software (MEDFLOSS) is widely used in academia and in low-resource settings, but also in healthcare institutions of industrial countries such as the US and UK [1]. The web portal Medfloss.org gives an overview of 363 MEDFLOSS projects, 111 professional service providers and 436 publications dealing with MEDFLOSS [2]. The portal is provided in cooperation with the FLOSS working groups of the European Federation of Medical Informatics, the International Medical Informatics Association and the International Society for Telemedicine & eHealth. About 3.350 visitors consult the website every month.

To describe software projects, Medfloss.org uses a self-developed nomenclature describing software by categories such as “application type”, supported “enterprise functions” and “license”. Since the establishment of Medfloss.org in 2010, the uncontrolled growth of the nomenclature has led to inconsistencies such as misassignments of descriptors to characteristics (e.g. “Virtual Reality” is listed as

¹ Franziska Jahn, Leipzig University, Medical Faculty, Institute of Medical Informatics, Statistics and Epidemiology, Härtelstr. 16-18, 04107 Leipzig, E-mail: franziska.jahn@imise.uni-leipzig.de

enterprise function) and the use of synonyms and homonyms for the descriptions of MEDFLOSS projects.

The Austrian-German research project HITO aims at the development of a **Health IT Ontology** for the systematic description of application systems and software products used in healthcare [3]. Whereas software products are tailor-made or purchased pieces of software that can be installed on a computer system, application systems are already installed software products that are adapted to a certain organizational environment [4].

HITO is developed iteratively by means of five use cases that currently lack precise application system and software product descriptions. Each use case is intended to add new concepts and relationships to HITO. In the first use case, we developed HITO V1 in order to improve the retrieval of evaluation studies examining application systems in healthcare settings [5]. HITO V1 systematically describes evaluated application systems by features, user groups and organizational units.

In the second use case, HITO V1 is extended by concepts and relationships for the precise description of MEDFLOSS (HITO V2) to improve retrieval and comparisons of MEDFLOSS projects on Medfloss.org.

This paper aims to describe methods and results of developing HITO V2.

2. Methods

We used a mixed-methods approach to identify important categories for the description of MEDFLOSS. The methods described in sections 2.1 to 2.4 were applied in parallel.

2.1. Surveying the Medfloss.org portal operators

In semi-structured interviews, two of the Medfloss.org web portal operators were asked about their experiences with the current terminology used on Medfloss.org and their ideas for improvement. On a 4-point scale ranging from “unimportant” to “very important”, both rated 8 out of 11 categories of the current terminology (enterprise function, application type, status, license, standard, language, client type, and platform) as “important” or “very important” for the description of the projects. The categories “popularity”, “database” and “programming language/toolkit” were each rated “less important” and “important” by one of the interviewees. They pointed out that the importance of these categories also depends on the perspective of the MEDFLOSS portal user. While health professionals are interested in functional software descriptions, software developers are more interested in technical categories.

However, from their point of view, the terminology currently lacks possibilities to describe the functionalities and application types more precisely. Therefore, the integration of existing terminologies such as WHO’s classification of digital health interventions [6] should be considered.

2.2. Comparison of three terminologies for describing MEDFLOSS functionalities

During the first HITO use case, we identified existing health IT taxonomies and ontologies [5]. To describe MEDFLOSS functionalities, we selected three of these taxonomies, which cover a broad range of health enterprise functions or application

system types systems, for a deeper analysis: (1) The reference model for the domain layer of hospital information systems (RDHIS) provides a hierarchy of enterprise functions (e.g. “medical admission”, “execution of nursing procedures”) [7]. Although it is originally developed for hospitals, many of these enterprise functions also fit in other healthcare settings. (2) The textbook “Health Information“ includes a list of application system types (e.g. radiology information system, patient administration system) used in healthcare and tries to avoid conceptual overlappings [4]. (3) Varshney et al. systematically developed a taxonomy of healthcare applications consisting of eight application system types [8].

For 356 projects listed on Medfloss.org, we counted how many of them could be described by the three taxonomies’ classes. The RDHIS gained best results in this comparison (Table 1). 71 % of the MEDFLOSS projects could be described by it. Only 13 of the 68 RDHIS’ enterprise functions could not be used for any description. With the help of application system list and the taxonomy of healthcare applications, only 42 % and 29 % of the projects could be described, respectively.

Table 1. Comparison of three taxonomies to describe 356 projects listed on Medfloss.org

	(1) Reference model for the domain layer of HIS [7]	(2) Application system list [4]	(3) Taxonomy of healthcare applications [8]
Number of classes	68	16	8
Number of Medfloss projects assignable to classes (n=356)	253 (71%)	151 (42%)	105 (29%)

2.3. Analyzing the Architecture of a MEDFLOSS-based Laboratory Environment

The MI-Lab at Leipzig University is a small computerized hospital information system architecture built up from several MEDFLOSS components [9]. Its analysis reveals that MEDFLOSS products can be provided as parts of other MEDFLOSS products. The Bahmni software used as electronic medical record (EMR) system in MI-Lab is composed of four other MEDFLOSS products providing a laboratory information system (LIS), an enterprise resource planning system (ERP), a patient management system and a picture archiving and communication system (PACS).

2.4. MEDFLOSS Literature Review

In order to find important characteristics for the description of MEDFLOSS products, we conducted a snowball search starting from a review on MEDFLOSS publications [1] which was complemented by a Pubmed and Google Scholar search on open source software characteristics and comparisons. Four out of seven articles providing structured descriptions of MEDFLOSS [10–13] list characteristics that are worth to be considered as HITO classes. However, the use of different self-developed lists of features or functionalities in the articles demonstrates the lack of a consented or at least organized terminology, but also unclear borders between features (e.g. “templates”) and supported enterprise functions (e.g. “billing”). Most of the categories used by Medfloss.org (2.1) can also be found in the articles: license, database, platform, programming language, language. Further categories describing privacy, security, usability, performance, customizability, user and developer support issues should be considered for inclusion into HITO.

3. Results

Based on our mixed-methods approach, we made the following design decisions for HITO V2 and explain them referring to the well-documented Bahmni software [14]:

- (A) All categories for the description of MEDFLOSS used on Medfloss.org are integrated into HITO V2 as classes because of their importance rating (2.1) and their use in MEDFLOSS literature (2.4). For example, Bahmni uses the license “AGPL 3.0” and supports the platform “CentOS”.
- (B) Functionalities of software products are described by combinations of supported enterprise functions and features (2.2, 2.4), e.g., Bahmni supports the enterprise function “Patient Registration” for which it provides the features “Register patients with their demographics”, “Identification contact”, etc. In a further step, these citations of software descriptions can be linked to entries of established catalogues for enterprise functions (e.g. [6,7]), and features (e.g. [6]), respectively. The enterprise function “Patient Registration” can be linked to the RDHIS function “Administrative Patient Admission”, for example. See Figure 1 for the corresponding HITO classes and relationships.

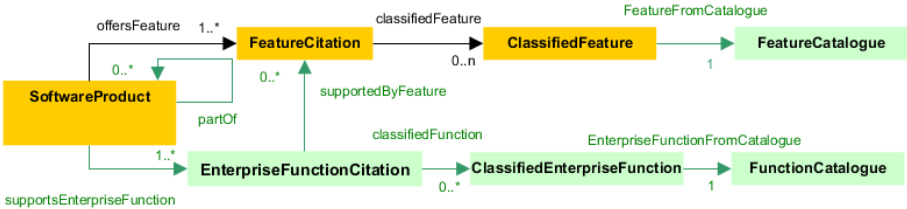


Figure 1. HITO V2: Relationships between functions and features as UML class diagram. Every citation of a MEDFLOSS description can be linked to a “classified” entry of a catalogue.

- (C) The RDHIS is transformed into a reference model for the domain layer of health information systems which includes enterprise functions for specific diagnostics and therapies, mobile health, research and public health (2.2).
 - (D) The class “Software product” is linked with itself by a “part of” relationship (2.3). Thus, the relationship “Open MRS is part of Bahmni” can be expressed.
- HITO V2 is available as Linked Open Data as RDF on <https://hitontology.eu> [3].

4. Discussion and Conclusion

Using a mixed-method approach, HITO could be extended by a variety of classes for the description of MEDFLOSS projects. Besides rather technical characteristics like platform or programming language, the precise description of software functionalities for medicine and health care plays an important role for the selection and comparison of MEDFLOSS. Thus, the domain-specific description abilities of HITO can be considered as one of its strengths and could, in the future, make it comparable to the EDAM ontology used in bioinformatics to describe operations and data [15]. However, due to the different terminologies for functionalities and application system types already in use, a unification of terms cannot be achieved in the short term. The HITO

project therefore integrates different existing terminologies for features and enterprise functions with the long-term goal of linking their terms.

Although a user searching for a specific software often relies on allegedly established terms for application system types (e.g. EHR, EMR, PACS, LIS), these terms are not clearly defined. For example, an “electronic health record (EHR) software” might be developed for the use in practices, hospitals or healthcare networks. It may be the electronic collection of patients’ documents or it may provide full medical process support. For this reason, HITO currently focuses on functional descriptions. After describing a large number of software products, clustering might help to improve definitions of application system types. However, coming to a huge amount of meaningful MEDFLOSS descriptions by HITO depends on good documentation of the software projects and the commitment of the community.

Acknowledgements

The HITO project is funded by the Austrian Science Fund FWF (I3726-N31) and the German DFG fund (WI 1605/11-1).

References

- [1] T. Karopka, H. Schmuhl, and H. Demski, Free/Libre open source software in health care: a review. *Healthc Inform Res* 20 (2014), 11–22.
- [2] H. Streidl, H. Demski, and T. Karopka, Medical Free/Libre and Open Source Software, <https://www.medfloss.org> [cited 2019 September 25].
- [3] HITO Team, HITO: A Health IT Ontology for systematically describing application systems and software products in health IT, <https://hitontology.eu> [cited 2019 September 25].
- [4] A. Winter, R. Haux, E. Ammenwerth, B. Brigl, N. Hellrung, and F. Jahn, *Health Information Systems: Architectures and Strategies*, Springer, London, 2011.
- [5] E. Ammenwerth, V. Dornauer, M. Ghalandari, F. Jahn, N. de Keizer, and A. Winter, An Ontology for Describing Health IT Interventions: Methodological Considerations. *Stud Health Technol Inform* 264 (2019), 1419–1420.
- [6] WHO, *Classification of digital health interventions: A shared language to describe the uses of digital technology for health*, Geneva, 2018.
- [7] G. Hübner-Bloder, E. Ammenwerth, B. Brigl, and A. Winter, Specification of a Reference Model for the Domain Layer of a Hospital Information System. In *Proceedings of MIE2005*, R. Engelbrecht, ed. IOS Press, Amsterdam, Washington, DC, 2005, pp. 497–502.
- [8] Varshney, U., Nickerson, RC and J. Muntermann, Taxonomy Development in Health-IT. In *AMCIS 2013 Proceedings*, 2013.
- [9] K. Brandt, M. Löbe, M. Schaaf, F. Jahn, A. Winter, and S. Stäubert, MI-Lab - A Laboratory Environment for Medical Informatics Students. *Stud Health Technol Inform* 228 (2016), 48–52.
- [10] A.E. Flores Zuniga, K.T. Win, and W. Susilo, Functionalities of free and open electronic health record systems. *Int J Technol Assess Health Care* 26 (2010), 382–389.
- [11] A.A. Zaidan, B.B. Zaidan, A. Al-Haiqi, M.L.M. Kiah, M. Hussain, and M. Abdunabi, Evaluation and selection of open-source EMR software packages based on integrated AHP and TOPSIS. *Journal of Biomedical Informatics* 53 (2015), 390–404.
- [12] S. Purkayastha, R. Allam, P. Maity, and J.W. Gichoya, Comparison of Open-Source Electronic Health Record Systems Based on Functional and User Performance Criteria. *Healthc Inform Res* 25 (2019), 89–98.
- [13] P.S. Millard, J. Bru, and C.A. Berger, Open-source point-of-care electronic medical records for use in resource-limited settings: systematic review and questionnaire surveys. *BMJ Open* 2 (2012).
- [14] Bahmni Coalition, *Hospital System for Low Resource Settings*, <https://www.bahmni.org> [cited 2019 September 27].

- [15] J. Ison, M. Kalas, I. Jonassen, D. Bolser, M. Uludag, H. McWilliam, J. Malone, R. Lopez, S. Pettifer, and P. Rice, EDAM: an ontology of bioinformatics operations, types of data and identifiers, topics and formats. *Bioinformatics* 29 (2013), 1325–1332.