Digital Professionalism in Health and Care: Developing the Workforce, Building the Future
P. Scott et al. (Eds.)

doi:10.3233/SHTI220900

Digital Health Multilingual Ontology to Index Teaching Resources

Stéfan DARMONI [a,b,1], Arriel BENIS [c,d], Emeline LEJEUNE [a,b], Flavien DISSON [a,b], Badisse DAHAMNA [a,b], Patrick WEBER [e], Pascal STACCINI [f], Julien GROSJEAN [a,b]

[a] *Department of Biomedical Informatics, Rouen University Hospital, France*
[b] *LIMICS Laboratory of Medical Informatics and Knowledge Engineering in e-Health, INSERM U1142, Sorbonne Université, France*
[c] *Faculty of Industrial Engineering and Technology Management, Holon Institute of Technology, Israel*
[d] *Faculty of Digital Technologies in Medicine, Holon Institute of Technology, Israel*
[e] *EFMI, Le Mont-sur-Lausanne, Switzerland*
[f] *RESINES, Université Côte D'Azur, Nice, France*

Abstract. The aim of this paper is to present the use of Medical Informatics Multilingual Ontology (MIMO) to index digital health resources that are (and will be) included in SaNuRN (project to teach digital health). MIMO currently contains 1,379 concepts and is integrated into HeTOP, which is a cross-lingual multi-terminogy server. Existing teaching resources have been reindexed with MIMO concepts and integrated into a dedicated website. A total of 345 resources have been indexed with MIMO concepts and are freely available at https://doccismef.chu-rouen.fr/dc/#env=sanurn. The development of a multilingual MIMO for enhancing the quality and the efficiency of international projects is challenging. A specific semantic search engine has been deployed to give access to digital health teaching resources.

Keywords. Digital Health; Language; Teaching; Vocabulary, Controlled

1. Introduction

Digital Health and Health Informatics are at the crossroads of medicine and health sciences, computer science and engineering, information and communication sciences, mathematics, statistics, technology, and innovation management.

The European Federation for Medical Informatics (EFMI) is built around 30 member-countries affiliated societies [1]. One of the objectives of EFMI is to encourage academic and professional collaboration and cooperation between all its members (both from academia and industry), to enhance health and medical informatics research, education, knowledge, and industrial transfer.

In the 1990s, some glossaries and thesauri, comprising a few hundred to 2,000, were developed in Medical Informatics. Two teams led these projects, one from Missouri,

[1] Corresponding Author: Stefan Darmoni, stefan.darmoni@chu-rouen.fr.

USA [2, 3] and the other from Liverpool, UK [4]. These paper-printed vocabularies were monolingual, in English only. Moreover, according to our knowledge, there is no similar prior work dealing with multilingual issues.

Under the umbrella of EFMI, a multilingual ontology on health informatics (and beyond, digital health, including health information sciences), called MIMO (Medical Informatics Multilingual Ontology) was developed since 2019 [5, 6].

In 2022, the French government launched an open call to teach digital health for every health student [7], which includes medicine, odontology, pharmacy, midwives, but also nurses and Physiotherapists. 71 million euros were devoted to this specific call. The Rouen & Cote d'Azur (Nice) Universities have successfully answered to this call, with the SaNuRN project, which will begin September 1st, 2022 for a five years duration.

The objective of this paper is to present the use of MIMO to index digital health resources that are (and will be) included in SaNuRN. This work has been partially granted by the European project (granted by Horizon 2020) HOspital SMART development based on Artificial Intelligence (AI; HosmartAI) [8]. HosmartAI is a European project involving 24 partners in 12 countries. HosmartAI promotes an effective and efficient healthcare system transformation through technological developments in AI and robotics. HosmartAI is creating a common open integration platform with the necessary tools to facilitate and measure the benefits of integrating digital technologies (robotics and AI) in the healthcare system. This work will be integrated into this platform to improve the interoperability among the various tools of this project.First paragraph.

2. Methods

The SaNuRN project are based on existing teaching resources that were developed by the two Department of Digital Health (DDH) from Rouen & Nice. In Rouen, since 1999 with the creation of the French Medical Virtual University [9], all the teaching resources are freely available in "open data" at the following URL: https://www.cismef.org/cismef/d2im/cours/.

Since 1993, the Rouen DDH has developed CISMeF, a catalogue of French-speaking health resources currently containing 129,979 resources, including 10,032 teaching resources [10]. Since 2000, a semantic search engine (Doc'CISMeF) was developed using primarily the MeSH thesaurus. A multiterminology approach was used since 2010, allowing any resource included in CISMeF to be indexed by more than one health terminology, while the MesH thesaurus remains the pivotal terminology.

The Medical Informatics and Digital Health Multilingual Ontology (MIMO) currently contains 1,379 concepts. MIMO has been developed with the following process and integrated into the SaNuRN project. The MIMO team has analyzed the main controlled vocabularies in health and being a part of UMLS (Unified Medical Language System) [11]: MeSH (Medical Subject Headings), SNOMED CT, NCIT (National Cancer Institute Thesaurus), MedDRA or ICD (International Classification of Diseases; in their ninth, 10th, and 11th revision). Based on this manual review by two medical informaticians (AB & SJD), the MeSH thesaurus was assessed as having the most comprehensive number of medical informatics concepts. Indeed, the MeSH thesaurus is the unique reference terminology, where a tree describing medical informatics and, more broadly, information science exists. Accordingly, to complete this thesaurus as exhaustively as possible, the MIMO team has decided to expand the concepts to be included not only extracted from the MeSH hierarchy "medical informatics" but to be

extracted from the broader concept “information science.” In other words, there is an “IS_A” relation between “medical informatics” and “information science,” where “medical informatics” is the narrower concept. No such tree exists in SNOMED CT or in NCIT, where only a tree of “software” exists in these two reference terminologies, and both are currently integrated into MIMO. Therefore, the MIMO concepts were primarily extracted from existing reference health terminologies: out of 1,379 MIMO concept, 460 are MeSH Descriptors, 220 MeSH Concepts, 230 NCIT concepts, 170 SNOMED CT concepts, and more than 300 newly created concepts. MIMO is continuously updated to comprise as recent as possible concepts and their translations in more than 30 languages. Moreover, the MIMO’s development team constantly improves MIMO content and supporting information. Since January 2021, MIMO development has been partially granted by the Horizon 2020 European project “HOspital SMART development based on Artificial Intelligence” (AI; HosmartAI) [8].

To disseminate MIMO, the MIMO team has chosen HeTOP, which is a crosslingual multiterminogy server (URL: https://www.hetop.eu/hetop/en) [12]. HeTOP contains more than 2 million concepts in English, 100 health terminologies, and ontologies, 55 languages, including non-European languages (e.g., Arabic, Hebrew, Japanese, Korean or Mandarin). Moreover, the generic model of HeTOP is compatible with the ISO 25964 standard and allows interoperability with other vocabularies platforms [13,14]. Termino-ontologies are regularly updated and are accessible via a website and a dedicated web service. HeTOP has been designed as a reference multi-terminology and multilingual server to help librarians, translators, students, and medical professionals retrieve resources and knowledge across various complex medical fields. In France, HeTOP is also largely used by the French Agency of Digital Health to populate and update the national terminology server to disseminate the main reference terminologies. The proof-of-concept of the MIMO was integrated, in mid-2019, into the HeTOP terminology server and was available to the consortium using an identification. Since the beginning of 2022, MIMO is freely available via HeTOP.

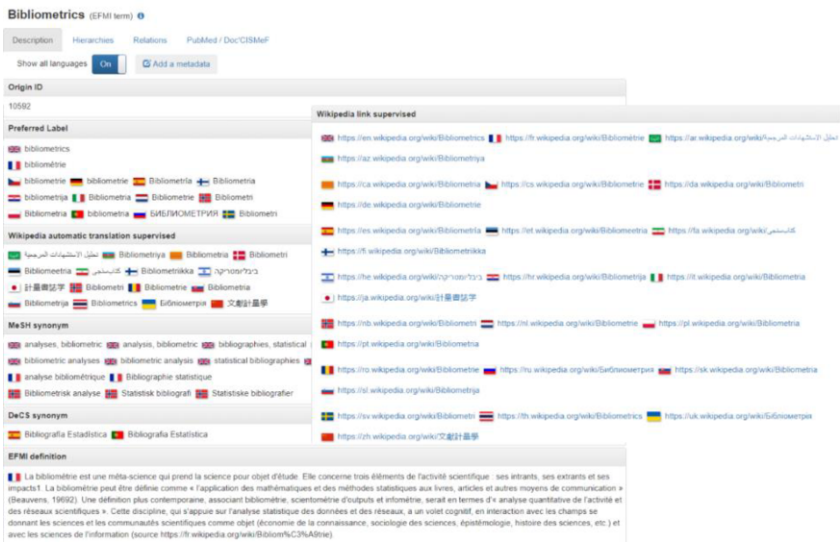

Figure 1. HeTOP page with the MeSH descriptor “Bibliometrics”

(also available at: https://www.hetop.eu/hetop/en/#rr=MSH_D_015706&q=bibliometrics)

3. Results

The Medical Informatics and Digital Health Multilingual Ontology is freely available through the HeTOP interface at the following URL: https://www.hetop.eu/hetop/rep/en/EFMIMIMO/. Figure 1 displays an example of MIMO, using the "bibliometrics" concept. It provides the translation in various languages, mainly European in this example as "bibliometrics" is a MeSH descriptor. A copy of these translations was integrated into MIMO. An added list of translations was added from the WikiMeSH algorithm [15] that we have recently developed to enhance the MIMO translations in European languages and above. A specific module was devoted to the manual curation of these translations because this mapping could generate false positives in case of polysemy (multiple related meanings) or ambiguous acronyms. In May 2022, MIMO was also integrated into the information system of the Rouen DBI, to allow the manual indexing of SaNuRN digital health teaching resources by medical librarians. Currently, 345 out of 10,000 teaching resources are in the scope of digital health. These teaching resources are reindexed, using MIMO to be as precise as possible. Figure 2 displays an example (in French) of a digital health teaching resource indexed using MIMO. The following MIMO concepts were used: data annotation, data warehouse, NoSQL, search engine, semantics, SQL.

A specific semantic search engine (CISMeF SaNuRN) has been deployed that only focuses on digital health teaching resources. It is available at: https://doccismef.chu-rouen.fr/dc/#env=sanurn

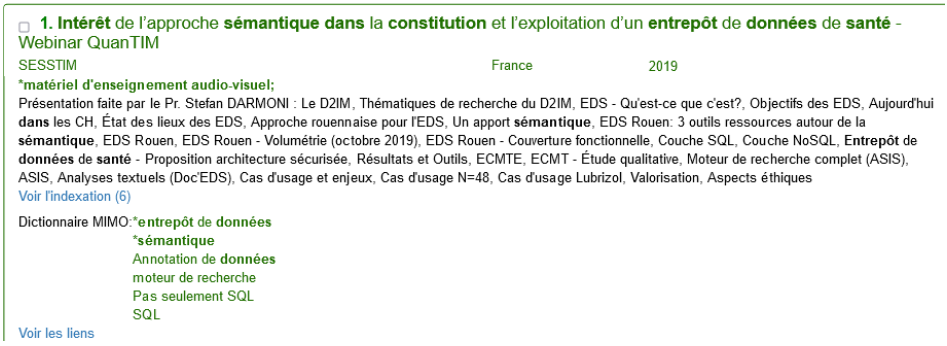

Figure 2. An example of a digital health teaching resources indexed with MIMO

4. Discussion and conclusion

The development of a multilingual ontology related to Medical Informatics and Digital Health for enhancing the quality and the efficiency of international projects is challenging. The MIMO developers aim to support all the actors interacting in multilingual projects at all the levels of professional health (industry) communication (e.g., education, research, engineering, knowledge dissemination, marketing) such as the one required in the European context. In this work, the MIMO team has presented the usefulness of MIMO to index digital teaching resources.

As the current corpus is composed of 345 resources, we plan with the development of digital health teaching in each French medical faculty in September 2022 (n=31) that

this corpus will double in the next year. In the future, a challenge should be to extend this kind of approach to other countries and languages.

Acknowledgement

This work was partially granted by the HosmartAI project (#101016834).

References

[1] Benis A, Crisan-Vida M, Stoicu-Tivadar L. The EFMI Working Group "Healthcare Informatics for Interregional Cooperation": An Evolving Strategy for Building Cooperation Bridges. Stud Health Technol Inform. 2019 Aug 21;264:1907-1908. doi: 10.3233/SHTI190707. PMID: 31438401.
[2] Ogg NJ, Sievert M, Li ZR, Mitchell JA. Construction of a medical informatics thesaurus. Proc Annu Symp Comput Appl Med Care. 1994:900-4. PMID: 7950054; PMCID: PMC2247813.
[3] Sievert MC, Moxley DE, Ogg NJ. Thesaurus of Health Informatics, IOS Press, Amsterdam, 1998.
[4] Rada R, Ghaoui C, Russell J, Taylor M. Approaches to the construction of a medical informatics glossary and thesaurus. Med Inform (Lond). 1993 Jan-Mar;18(1):69-78. doi: 10.3109/14639239309034470. PMID: 8366694.
[5] Benis A, Crisan-Vida M, Stoicu-Tivadar L, Darmoni S. A Multi-Lingual Dictionary for Health Informatics as an International Cooperation Pillar. Stud Health Technol Inform. 2019 Jul 4;262:31-34. doi: 10.3233/SHTI190009. PMID: 31349258.
[6] Benis A, Grosjean J, Billey K, Montanha Meireles Martins JG, Dornauer V, Crisan-Vida M, Hackl WO, Stoicu-Tivadar L, Darmoni SJ. Medical Informatics and Digital Health Multilingual Ontology (MIMO): a tool to improve international collaborations. Submitted to International Journal of Medical Informatics.
[7] Appel à manifestation d'intérêt (AMI) « Compétences et métiers d'avenir ». Available at https://travail-emploi.gouv.fr/actualites/l-actualite-du-ministere/article/appel-a-manifestation-d-interet-ami-competences-et-metiers-d-avenir (Last accessed: 2022 June, 12).
[8] HOSMARTAI – AI for the smart hospital of the future. Available at https://www.hosmartai.eu/ (accessed December 30, 2021).
[9] Beux P, Duff F, Fresnel A, Berland Y, Beuscart R, Burgun A, Brunetaud JM, Chatellier G, Darmoni S, Duvauferrier R, Fieschi M, Gillois P, Guillé F, Kohler F, Pagonis D, Pouliquen B, Soula G, Weber J. The French Virtual Medical University. Stud Health Technol Inform. 2000;77:554-62. PMID: 11187614.
[10] Darmoni SJ, Leroy JP, Baudic F, Douyère M, Piot J, Thirion B. CISMeF: a structured health resource guide. Methods Inf Med. 2000 Mar;39(1):30-5. PMID: 10786067.
[11] Lindberg DA, Humphreys BL, McCray AT. The Unified Medical Language System. Methods Inf Med. 1993 Aug;32(4):281-91. doi: 10.1055/s-0038-1634945. PMID: 8412823; PMCID: PMC6693515.
[12] Grosjean J, Merabti T, Dahamna B, Kergourlay I, Thirion B, Soualmia LF, Darmoni SJ. Health multi-terminology portal: a semantic added-value for patient safety. Stud Health Technol Inform. 2011;166:129-38. PMID: 21685618.
[13] International Organization for Standardization, ISO 25964-1:2011, Int. Organ. Stand. (2011). https://www.iso.org/standard/53657.html.
[14] International Organization for Standardization, ISO 25964-2:2013, Int. Organ. Stand. (2013). https://www.iso.org/standard/53658.html.
[15] Dusenne M, Billey K, Desgrippes F, Benis A, Darmoni SJ, Grosjean J. WikiMeSH: Multi Lingual MeSH Translations via Wikipedia. Stud Health Technol Inform. 2022 May 25;294:403-404. doi: 10.3233/SHTI220483. PMID: 35612105.