

# Multi-terminology cross-lingual modeling in the Health Terminology/Ontology Portal

**Julien Grosjean, MSc, Tayeb Merabti, PhD, Nicolas Griffon, MD,  
Badisse Dahamna, MSc, Lina Soualmia, PhD, Stéfan J. Darmoni, PhD**  
TIBS-LITIS EA 4108 & CISMef, Rouen University and Hospital, France

## Abstract

*The Health Terminology/Ontology Portal (HeTOP) is a repository dedicated to health professionals and students. It provides access to 32 health terminologies (including MeSH, ICD-10, etc.) available in 23 different languages (English, French, German, Russian, etc.). Several methods and technologies have been developed to create this multi-terminology server, dedicated to both users and computers. HeTOP is a valuable tool to help in indexing, as well as for teaching and performing audits in terminology management.*

## Methodology

Semantic interoperability through HeTOP relies on a Unifying Model of Vocabulary (UMV) described in [1]. Different sources of data for each available terminology were merged (UMLS [2], official national sources of ICD-10, etc.). UMV is a common model for all terms, whatever is the terminology they belong to. It can be viewed as a meta-model or an upper ontology designed to support broad semantic interoperability between terminologies that fulfill it. UMV provides ‘mapping’ properties such as those defined in the SKOS language. The meta-model is basically cross-lingual because preferred terms, synonyms or other textual attributes can be defined by a language code (en for English, fr for French, etc.). Each terminology T (such as MeSH, ICD-10, FMA, SNOMEDCT, ...) included in HeTOP is built as an enrichment the meta-model of UMV (UMV-T). Each enrichment defines its own specializations of Descriptor. For example, UMV-FMA defines the descriptor type FMAentity and the associations FMAinnervation, FMA drainageveineux (venous drainage). The terminologies included in the HeTOP are implemented as light OWL ontologies. Moving from an ontological representation to a terminological one is based on a *reification* process. By this way, formal ontologies are “degraded” to fit to this multi-terminology model. This method allows to keep the original data of an ontology and to enrich the server with health lexicons.

## Results

A total of 32 terminologies are included in HeTOP, with 1,280,000 concepts, 2,400,000 synonyms, 222,800 definitions and 4,100,000 relations. The bilingual version of HeTOP has been used on a daily basis since January 2010 by the CISMef librarians to index in multi-terminology mode. It is also used by 500 unique machines. To consult the HeTOP (<http://www.hetop.eu/index.html?lang=en>, click on “Log in”; login=fmauser, password=fmapass).

## Conclusion

Currently, the HeTOP is a necessary basic tool to index any document in a multi-terminology multilingual mode. Other portals propose searching and navigating T/O such as NCBO Bioportal [3] and the EBI Ontology Lookup Service [4]. These tools are also very friendly but they do not allow users to navigate through terms or search among synonyms in different languages; neither are they adapted to indexing on a daily basis. To our knowledge, this kind of multi-terminology and cross-lingual portal is the first in the health sector. Moreover, much hard work has been carried out to achieve interoperability between terminologies and for their enrichment (addition of synonyms, translations, etc.). Non-European languages have also been integrated recently: MeSH in Japanese, in Chinese or ICD-10 in Arabic. The conceptual approach used in the model allows integration of any language while maintaining correct relations between concepts.

## References.

1. Grosjean J, Merabti T, Dahamna B, ... and Darmoni SJ. Health Multi-Terminology Portal: a semantic added-value for patient safety. *Stud Health Technol Inform.* 166:129-38 (2011).
2. Bodenreider O. The Unified Medical Language System (UMLS): Integrating biomedical terminology. *Nucleic Acids Res.* 32:267–270 (2004).
3. Noy N.F, Shah N.H, Whetzel P.L, ... and M.A Musen. Bioportal: ontologies and integrated data resources at the click of a mouse. *Nucleic Acids Research*, 37:170–173 (2009).
4. Cote RG, Jones P, Apweiler R and H Hermjakob. The ontology lookup service, a lightweight cross-platform tool for controlled vocabulary queries. *BMC Bioinformatics*, 7(1):97 (2006).