MEDICAL RESOURCE INDEXING FOR AN EASY AND TARGETED ACCESS

A wide variety of medical resources are available on-line. Among the wealth of information to be found are classification systems such as ICD-10\(^1\), medical ontologies such as Snomed and UMLS\(^2\), and open data provided on the Semantic Web, such as Drugbank and DailyMed\(^3\) that are part of the « Linked Data Cloud »\(^4\).

How can a user find its way through all these platforms?
How does one grasp their content and get a hold of all the available resources?
What is the most efficient targeted access method?

The experts in natural language processing of the DETI group can help you understand, index and gain quick access to relevant information through available on-line resources, or within your own medical resources whether they are structured (database) or not (text corpus). They master indexing technologies such as Lucene\(^5\), and can adapt these technologies to meet your specific indexing needs for your own applications.

Here are 3 examples of free on-line medical resources explored and indexed in order to answer some typical questions. Navigation interfaces have been developed for each indexed resource.

**QUESTION:** What are the terms relating to the word *coeur*?

**INDEXED RESOURCES:** SNOMED, English and French\(^6\) versions.

**METHOD:** Pairing of French and English resources to allow for bilingual navigation of the terms and synonyms contained in the resources.

---

1. The new standard « International Classification of Disease » is available at the following site: [http://www.who.int/classifications/icd/en/](http://www.who.int/classifications/icd/en/)
2. Snomed (English) and UMLS are available under licences at [http://www.nlm.nih.gov/research/](http://www.nlm.nih.gov/research/)
4. Several Semantic Web resources are accessible at [http://linkeddata.org](http://linkeddata.org)
2 QUESTION: The patient experiences abdominal pains; to what can this be related?

INDEXED RESOURCE: UMLS (Unified Medical Language System)

METHOD: Exploration of the concepts and semantic relations contained in UMLS to find different information related to the searched term.

3 QUESTION: What is the generic medication for « Lazeta » and what symptoms does it treat?

INDEXED RESOURCES: Drugbank and DailyMed, 2 resources of Linked Data (Semantic Web).

ANSWER: Exploration of the information types contained in RDF triplets (standard Semantic Web representation) of the resources for the purpose of indexing.

RESEARCH AND DEVELOPMENT:
Caroline Barrière, Ph.D., researcher, DETI team, CRIM
Michel Gagnon, Ph.D., invited researcher at CRIM, École Polytechnique de Montréal

INFORMATION:
Sacha Leprêtre, Team Director, DETI, CRIM
Tel.: 514 840-1238 or 1 877 840-2746 - sachalepretre@crim.ca

www.crim.ca