ICMAN
Instituto de Ciencias Marinas de Andalucía

Covid-19 ontology resources for libraries

Enrique Wulff

Instituto de Ciencias Marinas de Andalucía (CSIC)

Workshop on COVID-19 Ontologies. Friday, October 23, 2020
8th Annual Workshop of the Clinical and Translational Science Ontology Group

Quality criteria

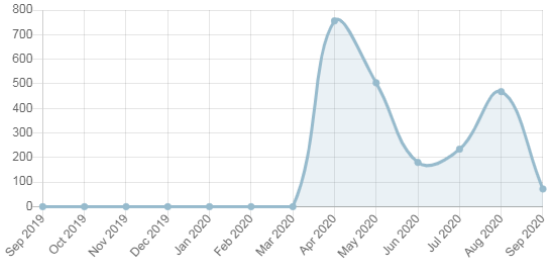
- **Library needs**
  - a framework for the comparison of different semantic data sources in the COVID-19 pandemic
  - an all-integrating ontology for coronavirus disease knowledge and data
- **Covid-19 ontologies an emerging technology in libraries**
- **Issues**
  - online services and/or applications based on ontologies or SKOS-based COVID-19 thesauri are still rare
  - in spite of using ontologies in the Semantic Web, meanings of concepts and relationships are still largely unrealized

Comparing the existing Covid-19 biomedical ontologies from the library

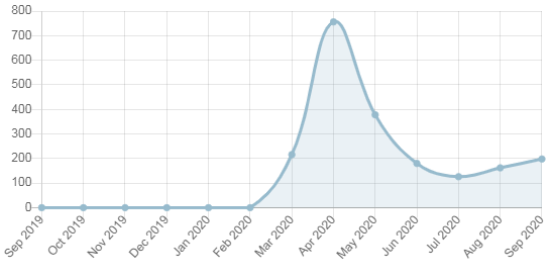
- **Ontologies building into a disease classification all the attributes**
  - WHO COVID-19 Rapid Version CRF semantic data model (COVIDCRFRAPID) and the COVID-19 Surveillance Ontology (COVID19) are organized according to these data retrieval rules
- **Ontologies separating attributes with metadata, evidence and provenance**
  - COVID-19 Infectious Disease Ontology (IDO-COVID-19) and the Community-driven Coronavirus Infectious Disease Ontology (CIDO) give out to new discoveries and concepts in their coverage

Comparing the existing Covid-19 biomedical ontologies from the library

- **COVIDCRFRAPID and the COVID-19 Surveillance Ontology manage Covid-19 data stored in electronic medical records**
  - Relevant issues
    - They afford with sources of clinical information in unprocessed textual formats
    - Insufficient described information hampers its interpretation
    - Difficulties for searching, summarization, decision support, and statistical analysis



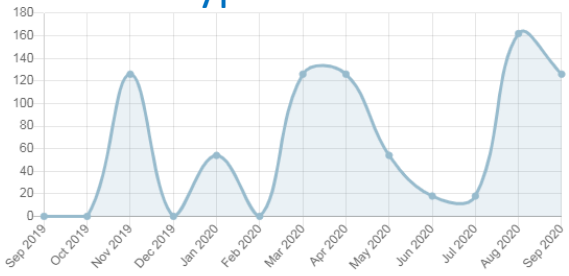
COVIDCRFRAPID Visits - NCBO BioPortal



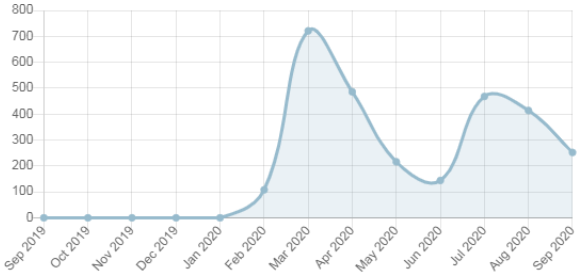
COVID-19 Surveillance Ontology Visits - NCBO BioPortal

Comparing the existing Covid-19 biomedical ontologies from the library

- **IDO-COVID-19 and CIDO are open to the capabilities of semantic annotators to filter ontologies by semantic types**
  - Main advantages
    - They can annotate free text and keep only certain broad types of annotations and textual formats
    - Their recognition of broad entity types (e.g. gene, drug, disease) is an utility of high interest for the BioNLP community
    - These major semantic groups are often used as coarse-grained groupings of the Unified Medical Language System (UMLS) Semantic Types



IDO-COVID-19 Visits - NCBO BioPortal



CIDO Visits - NCBO BioPortal

Libraries know the method used to develop the existing COVID-19 ontologies

- **Issue**
  - In the absence of canonical data structures it is impossible to contemplate COVID-19 successful biomedical research
- **Library help**
  - Research at the library is necessary to the effective practice of making biomedical ontologies work

Research Data Management for the Health Sciences: Metadata and Ontology Schemas

A way to have more research on ontologies available from Yale Library

Libraries know the method used to develop the existing COVID-19 ontologies

- **Issue**
  - Storing large-scale RDF has been an urgent need since the beginning of the COVID-19 pandemic

- **Library help**
  - Libraries have a great deal of experience in managing RDF data.
  - Expert Librarians are needed:
    - **for the creation of meaningful mappings**
    - **for finding associations between concepts of local data models and existing ontologies**

The research at the library is a call for the use of COVID-19 ontologies

Building semantics in the domain of COVID-19 data is feasible through an OBO Library approach

- **A new direction for research at the Library would be the assessment and ranking of COVID-19 ontologies**

- Library as an institution is interested by the formalization of shared terms
- Library should be involved in providing ontology's quality and suitability exercises
- Standardization efforts in the field of COVID-19 biomedical ontologies would be extremely valuable if done together with libraries

The work of Libraries concerning Covid-19 ontologies

- **outlines the core use cases**
- **characterizes the unique features and coverage of these ontologies**
- **explains the degree these ontologies are based on NLM resources**
- **describes the ontologies search features and limitations**
- **characterizes the importance and comparability of biomedical ontologies as an issue to share data and knowledge**
- **suggests a design ontological based approach to improve the semantic interoperability of coronavirus data**

Please ask me any question