

Leveraging a powerful weapon in the fight against COVID-19—ontology

June 11 2020, by Bert Gambini



3D print of a spike protein of SARS-CoV-2, the virus that causes COVID-19—in front of a 3D print of a SARS-CoV-2 virus particle. The spike protein (foreground) enables the virus to enter and infect human cells. On the virus model, the virus surface (blue) is covered with spike proteins (red) that enable the virus to enter and infect human cells. Credit: NIH

The mass of growing and constantly changing data resulting from multiple disciplines represents one of the biggest challenges researchers and public health officials must confront while trying to manage the ongoing COVID-19 pandemic.

But several centers across the country, including the University at Buffalo's National Center for Ontological Research (NCOR), are working to develop ontologies to assist in the efforts to control the current outbreak, accelerate data discovery in future pandemics, and promote reproducible infectious [disease](#) research, according to Barry Smith, SUNY Distinguished Professor of Philosophy and director of NCOR.

Smith is among the co-authors of a new paper discussing how ontologies can assist in the fight against COVID-19.

To realize the scope of the challenge faced by scientists confronting COVID-19, consider the disciplines involved in the fight—everything from immunochemistry to behavioral population modeling.

All the data collected by biologists, pathologists, sociologists, geographers, physicians and epidemiologists require integration, but the [relevant information](#) is captured using discipline-specific terms and is often stored in ways that are accessible only to those working in the fields in which they originated.

"Ontology was designed to address that problem by creating common controlled vocabularies for data descriptions that everyone can use," says Smith, who was named one of the 50 most influential living philosophers in 2016 by TheBestSchools.org.

"It's nearly impossible, unless you're an expert in multiple separate disciplines, to join data deriving from multiple different sources. This

problem is especially acute in the face of a novel pathogen such as SARS-CoV-2, where no one can anticipate which combinations of factors will prove crucial in understanding how it affects its human hosts."

Accessing and integrating massive amounts of information from multiple data sources in the absence of ontologies is like trying to find information in library books using only old catalog cards as our guide, when the cards themselves have been dumped on the floor.

Ontologies are data sharing tools that provide for interoperability through a computerized lexicon with a taxonomy and a set of terms and relations with logically structured definitions.

Smith has been working for some 15 years with biologists and bioinformatics specialists to create a suite of ontologies to cover all the [life sciences](#).

The paper—with Shane Babcock (Niagara University), John Beverley (Northwestern University) and Lindsay G. Cowell (University of Texas Southwest Medical School)—has not yet been accepted for publication. However, in light of the urgency of the pandemic, it appears already on the preprint repository of the Open Science Foundation.

It presents, first, an Infectious Disease Ontology (IDO) Core, which contains terms relating to infectious diseases generally before describing how this IDO Core has been extended in a number of ontologies relating to specific infectious diseases, such as malaria, staph and flu. The paper concludes with a treatment of IDO ontologies for viral infectious diseases in general, for coronavirus [infectious diseases](#), and for COVID-19, specifically.

These ontologies help to fill the need for standardized terminology in describing coronavirus data and information, and because they are all

constructed in the same way, they make it easier to compare COVID-19 data with data pertaining to other coronavirus diseases, such as SARS, MERS— and the novel [coronavirus](#) diseases of the future.

"An infectious disease ontology can also contribute to solving the problem of reproducibility," says Smith.

Reproducing the results of experiments as part of the research process requires a precise description not merely of the results achieved but also of the protocols, statistics, equipment, samples and tests used.

"We believe that, when used in combination with other life science ontologies such as the Ontology for Biomedical Investigations, the IDO framework provides a promising strategy for the creation of comparable, integratable, and discoverable provenance metadata for the data generated in infectious disease research," Smith says.

More information: Shane Babcock et al. The Infectious Disease Ontology in the Age of COVID-19, (2020). [DOI: 10.31219/osf.io/az6u5](https://doi.org/10.31219/osf.io/az6u5)

Provided by University at Buffalo

Citation: Leveraging a powerful weapon in the fight against COVID-19—ontology (2020, June 11) retrieved 11 May 2024 from <https://medicalxpress.com/news/2020-06-leveraging-powerful-weapon-covid-19ontology.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--