"A FUNDAMENTAL CLINICAL SKILL"
THE GENDER, SEX, AND SEXUAL ORIENTATION (GSSO) ONTOLOGY AND FACILITATING COMMUNICATION IN HEALTH CARE

Clair Kronk, PhD
Yale Center for Medical Informatics
About Me

◦ Postdoctoral fellow at Yale University.
◦ Thesis focused on the Gender, Sex, and Sexual Orientation (GSSO) ontology, a vocabulary of over 14,000 LGBTQIA+-related terms (part of the OBO Foundry, available here: http://www.ontobee.org/ontology/GSSO).
◦ Consultant for Canada Health Infoway, DICOM, HL7, Mayo Clinic, Merck, NAS, OutHistory, Queensland Health, and SNOMED CT.
◦ Board member at Homosaurus, the Trans Metadata Collective, and AMIA Inclusive Language Guidelines Task Force.
LGBTQIA+ Language and Health Care

- Linguistic stigmatization
- Language barriers
- Verbal abuse
- Perceived and actual language-based discrimination
- Slurs
- Former and current pathologizing and medicalizing language
- Contemporary versus non-contemporary terminologies

LGBTQIA+ Language and Health Care

[D]espite trying to find ways to improve my expertise, I just didn’t know where to go or who to talk to, or where to get the information [about trans people], and I felt really bad because some of my initial attempts to help these people—I sent them to people I wish I hadn’t sent them to.

—Anonymous Clinician (2012)

The biggest challenge is that the health care system is woefully unprepared to take appropriate care of LGBTQ people… It’s a dawning idea that needs to gain traction, that there’s also a whole field of sexual-gender minority health that providers need to have an understanding of.

—Kenneth Mayer, Fenway Institute (2018)


Introducing the GSSO

The Gender, Sex, and Sexual Orientation (GSSO) ontology is the first transdisciplinary ontology, including LGBTQ+ slang and medical terms side-by-side.

It includes thousands of unique and individually curated terms not present in any other vocabulary systems.

The GSSO is the first ontology to include a “drill-down” methodology for easy access to data sources.

The GSSO is more accurate, complete, concise, adaptable, clear, computationally efficient, and consistent than existing platforms.
The GSSO Construction Process

- Abstracting
- Marking
- Exploring
- Mapping
The GSSO Construction Process

- Identify Seed Terms
- Search terms in given databases
- Use resources to find new seed terms
The GSSO Construction Process

- Determine relevant databases
- Determine relevant search results
- Exploring
The GSSO Construction Process

Is it in another ontology?

Is its parent term in another ontology?

Is it or a parent in any hierarchy?

Can any new hierarchy be established?

Mapping
The GSSO Construction Process

Abstracting
Protégé

- Free
- Open-source
- Supports multiple ontology formats
- Supports W3C standards
- Hundreds of plugins
- Active user community
Releases

**Version 1.0 (June 2019)**
- Static website served in JavaScript with simple search interface only
- 6,250 classes and 0 instances
- 1,063 definitions
- 1,416 database cross-references
- 2.6 average number of annotations

**Version 2.0 (June 2020)**
- Dynamic model-view-controller (MVC) Angular application with complex search functions and additional features
- 7,121 classes and 2,939 instances
- 7,121 definitions
- 14,193 database cross-references
- 7.4 average number of annotations
<table>
<thead>
<tr>
<th>Ontology Name</th>
<th>Version 1.0</th>
<th>Version 2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>ChEBI</td>
<td>4</td>
<td>213</td>
</tr>
<tr>
<td>DO</td>
<td>62</td>
<td>193</td>
</tr>
<tr>
<td>FMA</td>
<td>43</td>
<td>327</td>
</tr>
<tr>
<td>GO</td>
<td>32</td>
<td>152</td>
</tr>
<tr>
<td>Homosaurus</td>
<td>NM</td>
<td>430</td>
</tr>
<tr>
<td>HPO</td>
<td>30</td>
<td>137</td>
</tr>
<tr>
<td>ICD-9-CM / ICD-10-CM</td>
<td>58</td>
<td>362</td>
</tr>
<tr>
<td>LCC / LCSH</td>
<td>NM</td>
<td>1,276</td>
</tr>
<tr>
<td>MedDRA</td>
<td>129</td>
<td>595</td>
</tr>
<tr>
<td>MeSH</td>
<td>261</td>
<td>904</td>
</tr>
<tr>
<td>NCIT</td>
<td>261</td>
<td>1,034</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td><strong>1,416</strong></td>
<td><strong>14,193</strong></td>
</tr>
</tbody>
</table>
Emphasizing Interoperability: Not Rebuilding the Wheel

- All properties are recruited from existing platforms, including RDFS, DCMI, Wikidata, and the Gene Ontology, among others.
- We focused on connections between the most consistently and heavily accessed vocabularies and ontologies in health care-related fields, using NCBO BioPortal, OntoBee, and the EMBL-EBI Ontology Lookup Service (OLS).