**Linked Modernisms NLP Research Project**

**Objective**: To devise a method for using NLP on a body of unstructured data (encyclopedia entries) in .TXT format to infer relationships among entities in the data. The inferences must be determined by the Casaubon information schema (to be provided) and output in .TTL format for ingestion into an RDF triple-store.

**Resources**:

* .TXT files (encyclopedia entries) for processing
* the Casaubon information schema
* sample hand-annotated training sets
* sample .TTL outputs
* access to the project’s existing codebase on Github for relevant Python scripts that have already been developed
* limited guidance from the project’s PI and from one of the original developers
* existing documentation for the Linked Modernisms project

**Current Iteration**:

The site is live at linkedmods.uvic.ca

The git repository is at https://github.com/LinkedModernismProject/web\_code

**Relevant Papers**:

* Buckner, Cameron, Matthias Niepert, Colin Allen. 2010. “From encyclopedia to ontology: toward dynamic representation of the discipline of philosophy” <https://inpho.cogs.indiana.edu/papers/2010BucknerNiepertAllen.pdf>
* Long, Hoyt and Richard So. “Network Analysis and the Sociology of Modernism” pre-print article available from Professor Ross ([saross@uvic.ca)](mailto:saross@uvic.ca))
* Pattuelli, Cristina. 2012. “FOAF in the Archive”
* Pattuelli, M. C. and Miller, M. (2015). Semantic network edges: A human-machine approach to  represent typed relations in social networks. Journal of Knowledge Management, (19)1, 71-81. [[Abstract](http://www.emeraldinsight.com/doi/abs/10.1108/JKM-11-2014-0453)] [[Preprint](https://cpattuelli.files.wordpress.com/2010/04/preprint_pattuelli_miller_jkm.pdf)]
* Wang, Yushi, Jonathan Berant, Percy Liang. 2015. “Building A Semantic Parser Overnight” <http://nlp.stanford.edu/pubs/wang-berant-liang-acl2015.pdf>

Students may also wish to familiarize themselves with the Dbpedia Spotlight tool (<https://github.com/dbpedia-spotlight/dbpedia-spotlight/wiki>), which may form part of the project’s design, depending on how students wish to proceed.

**Online Resources**:

Comparable projects:

* The Indiana Philosophy Ontology (InPhO): <https://inpho.cogs.indiana.edu>
* LinkedJazz: <https://linkedjazz.org>
* Read the Web Research Project at Carnegie Mellon University: <http://rtw.ml.cmu.edu/rtw/resources>

An index of online resources pertaining to the following subjects will be provided:

* JavaScript
* CSS (within JavaScript Directory)
* Polymer Desgin (Within JavaScript Directory)
* D3
* Huviz
* VSP
* RelFinder
* Virtuoso
* Sparql
* XML
* Natural Language
* Bash