

LINCS Ontologies Adoption & Development Policy

Last updated: 04 July 2021

Date of last review by LINCS Ontology Working Group: 17 December 2020

Authors: Erin Canning (University of Guelph), Susan Brown (University of Guelph)

Contributors: Kim Martin (University of Guelph), Alliyya Mo (University of Guelph), Sarah Roger (University of Guelph)



Linked Infrastructure for
Networked Cultural Scholarship

Table of Contents

Purpose of this document	2
Guiding principles	2
Values	2
Practicalities	3
Design approaches	4
Policies on using external resources	6
Ontologies	6
Incorporating external ontologies	6
Vocabularies	6
Using URIs to uniquely identify entities and concepts	6
Referencing vocabularies and authorities	7
Implications of referencing an external URI independent of its ontology	7
Structuring relationships with external resources	7
References and resources	8

Purpose of this document

This document serves as a statement of the shared values regarding LINCS ontologies. It will be used as a reference document to guide the selection, adoption, and development of ontologies.

This document should reflect the values of the LINCS community, as articulated by the members of the Ontology Working Group (OWG) and as arising from broader consultation with LINCS researchers. It should be considered as a living document that will be updated when necessary as LINCS work progresses to include new needs of the community as they surface.

Guiding principles

Values

A real challenge of linked data for humanities research is incorporating nuance and epistemological difference, including the accommodation of [boundary objects](#) that have different meanings in different fields. LINCS includes researchers whose datasets embed alternative understandings of history, and of how knowledge works. LINCS will work to ensure that its ontologies can represent non-hegemonic epistemologies and push alternative knowledge representations into the Semantic Web. In its information architecture, LINCS will attend to difference and diversity, including the ways of knowing of marginalized groups.

In order to achieve these goals, LINCS ontologies will be selected, adopted, and developed with an attention to intersectionality, multiplicity, and difference. Where these ontologies deal with description of people, it will seek to convey an understanding of identity categories and social classification as being contextually embedded: culturally produced, intersecting, and discursive. LINCS will aim to present these kinds of categories as not fixed classifications, but instead as grounds for investigation, starting points for understanding. Furthermore, LINCS will seek to foreground the contextual data of all of the data involved in this project, not just that dealing directly with the description of people, places, or groups, as well as the relationship between the data and its sources.

LINCS's ontology adoption and development process will proceed with a recognition of the risk of coding sexism, racism, and other forms of discrimination into our digital infrastructures. In order to

mitigate these risks, LINCS commits to a collaborative process of ontology adoption and development. Furthermore, LINCS commits to the rejection of ontologies that seek to encode their domains in ways that are at odds with the values stated in this document.

Table 1. How LINCS will seek to address systemic biases in its ontology work

Bias	Method
Racism	Center the perspectives of people of colour in discussions of race.
Colonialism	Adopt or develop ontologies that promote equal views of epistemological differences and/or promote Indigenous ways of knowing.
Cissexism	Center trans perspectives in discussions of the gender binary; use trans-inclusive language.
Heteronormativity	Resist assumptions about family structures, sexual and romantic preferences and identities, and gender roles.
Ableism	Develop tools with accessibility considerations included from the point of inception.
Classism	Develop solutions that can be used by partners/others without requiring additional expensive technologies; provide training and education resources.
Discrimination	Promote understanding of different worldviews along with histories, cultures, and contexts; use justice-oriented language.

Table adapted from *Data Feminism*, pp. 218–19

Practicalities

LINCS will seek to manage the risk of scope creep in its ontology work by maintaining focus on the biggest problems for humanities researchers related to linked data, as determined by consultation and discussion during the inception and development of the project, which point to three main requirements:

1. Link to each other;
2. Link to external resources and other data communities;
3. Support data storage and retrieval, and the use of the tools required to do so.

LINCS will determine specific, actionable requirements for the ontology work and products through a process of explicitly gathering and documenting ontology competencies and linking them to the appropriate ontological structures, tools, and processes.

LINCS accepts that there will always be a tension between completeness, accuracy, precision, and usability. After all, any model is an abstraction and simplification. In ontology adoption and development, LINCS will maintain focus on the main project goals by focusing on providing ontology solutions that highlight connections between LINCS datasets and between these datasets and external data sources, convey the complexity and local meanings of the data being represented, and show the full content of the ingested datasets.

Design approaches

LINCS ontology design—which refers here to the process of selecting and adopting as well as developing extensions or new ontologies—will follow a dynamic, lean approach. This means that the ontology development will be a continual process, starting with the main priorities and subject to ongoing community evaluation. This is an iterative, participatory process. This is an important approach for LINCS to commit to as it emphasizes that ontologies are linked to the communities whose data they seek to represent, and thus their adoption or development should involve those voices at all stages of the process.

The LINCS ontologies must be a part of, connect to, or connect with other ontologies that are being used by the community or otherwise considered important. This list will be revised as the project progresses, but preliminarily includes:

- [Web Annotation Model](#)
- [CIDOC-CRM](#)
- [CWRC](#)
- [Wikidata](#)

The LINCS ontologies should be relatively lean and general—as simple as possible, but no simpler—to ensure flexibility and breadth of use. This means:

- Vocabularies/thesauri exist external to the ontologies
- Use existing ontologies where possible while incorporating as few ontologies as possible

- No new classes of entities without the need for new properties pertaining solely to those classes
- Modular or suite/network of ontologies: core properties to function as a cohesive base ontology, with modular extensions for domain-specific requirements
 - Describe shared domain use cases or competency questions consistently while allowing local extensions for local problems
- Attempt to adhere (within reason) to Gruber's five criteria of ontology design:
 - Clarity: All elements of the ontology should be documented in plain language that is easy for the reader to understand, and include examples as well as logical axioms if possible.
 - Coherence: An ontology should support inferences that are consistent with the definitions of the classes, properties, and overall ontology. If a term is defined using logical axioms and a statement can be inferred from the axiom that contradicts the plain-language definition provided, then the ontology is considered incoherent.
 - Extendibility: An ontology should offer a foundational structure for a range of expected uses (as documented in use cases or competency questions), and be designed so that a user is able to define new terms for special uses based on the existing vocabulary in a way that does not require revising the existing ontology.
 - Minimal encoding bias: An ontology should exist and be understood as a shared conceptualization at a knowledge level, not purely on the level of encoding: an ontology is not its OWL file, it is the concepts and worldview that is then encoded using a language such as OWL. An encoding bias occurs when choices are made for the convenience of notation or implementation, as opposed to what is best for the conceptualization.
 - Minimal ontological commitment: An ontology should require the minimal ontological commitment sufficient to support the intended knowledge sharing activities. This will help to support extendibility, as users should be able to specialize and instantiate the ontology as required; this is also another reason for vocabularies to be stored externally from the ontology itself.

Policies on using external resources

Ontologies

Incorporating external ontologies

LINCS will link to external ontologies on a case-by-case basis. In order to support connection to external resources and other data communities, LINCS will also explore mapping to commonly used ontologies, such as [FOAF](#), that can then be used as a crosswalk to access and ensure LINCS materials are accessible to a broader linked data world. In incorporating external ontologies, LINCS will be incorporating the structure of the ontology—meaning classes and properties from the ontology under consideration - into the LINCS ontological environment.

Vocabularies

Using URIs to uniquely identify entities and concepts

LINCS will reuse existing URIs and build on existing linked data work where possible. In taking this approach, LINCS also seeks to highlight and incorporate the work that domain experts have already undertaken in the many scholarly areas that this project will touch on. URIs already exist for many of the entities that LINCS will be interested in referencing, and minting our own LINCS-specific URIs in these cases where URIs already exist will only reduce compatibility with the wider linked data world instead of expanding it. Use of these URIs require the data holder to agree that they are referencing the same entity or concept as defined by the authority: this means that they do not have to agree with all of the properties that are related to the entity by the external authority, but that there needs to be a shared understanding that the entity is the same for both cases. For example, while two data sources might disagree on the exact birth date of a person, they can agree that they are both talking about the same person—this is the level of agreement that is required when looking to use an existing URI. Adopting an external URI does not imply agreement with the description, properties or relationships, or even the names or labels associated by the external authority with that entity.

LINCS will provide guidance on sources for these URIs for projects to look to, as well as open documentation on, cases where the data holder disagrees with the definition of an existing URI to the extent that they do not feel it is compatible for their project.

Referencing vocabularies and authorities

Linking to established authorities (for example, [Getty Vocabularies](#), [VIAF](#), [GeoNames](#), [Wikidata](#), [ORCID](#), [Homosaurus](#), etc.) is essential for LINCS's participation in the larger linked data community. These references can be used to uniquely identify named entities as well concepts. LINCS will support references to, without necessarily endorsing, these external authorities. Referencing vocabularies will allow LINCS to leverage the existence of URIs for entities and concepts that have already been minted: this will support shared reference by projects within LINCS, while also preparing LINCS data for participation in the larger linked data community by establishing references that will be shared by other data sources as well.

LINCS is developing a list of preferred vocabularies in consultation with researchers. Projects may wish to rank those vocabularies according to their own order of preference.

Implications of referencing an external URI independent of its ontology

When using an URI as an authority reference, LINCS will not import the ontology that may be associated with the URI. For example, Wikidata URIs can be used to uniquely identify entities, but unless LINCS decides to incorporate the Wikidata ontology, the Wikidata ontological structure and related properties will not be also included through this reference unless a data user specifically generates a [federated query](#) to independently explore other ontologies and datasets that also reference that same URI. Any decision to import an authority's ontological structure along with its terms (URIs) will be made by the Ontology Working Group in consultation with the researchers whose domain or areas of expertise are involved.

Structuring relationships with external resources

LINCS will explore options for structuring relationships with external vocabularies and ontologies, including leveraging SKOS- or OWL-based mapping properties such as `skos:closeMatch`, `skos:exactMatch`, `owl:sameAs`, `owl:equivalentClass`, and `owl:equivalentProperty` to define these connections. LINCS will also consider predicates such as `cwrc:functionalRelation` which indicate that two terms may be treated as related for functions such as querying and retrieval but do not assert a semantic relationship between the terms or ontologies in which they are based.

By reusing existing ontologies where possible, searching out domain-specific vocabularies to incorporate, and positioning the project as a resource willing and able to support other projects that

wish to embark on more extensive domain ontology development, LINCS seeks to balance domain-specific ontology requirements with the general needs of the project as a whole.

LINCS acknowledges that many of the relationships that the project seeks to forge with the broader linked data world are not purely utilitarian in nature, and that these connections will likely need to be contextualized so as to clarify the nature of the relationship held with the term or ontology being referenced. In light of this, LINCS will approach the question of what to link to and how with an intent to create space to expose tensions or significant differences, where they exist, and propose solutions that are devised in collaboration with project participants. Linking to a particular ontology or vocabulary, in whole or in part, may be eschewed if the LINCS community concludes that to do so would run counter to the values articulated in this policy.

References and resources

- Brown, Susan, Abigel Lemak, Colin Faulkner, Kim Martin, Rob Warren. 2017. "Cultural (Re-)formations: Structuring a Linked Data Ontology for Intersectional Identities." *Digital Humanities 2017*. <https://dh2017.adho.org/abstracts/580/580.pdf>
- Canadian Writing Research Collaboratory. 2019. *CWRC Ontology Preamble*. <http://sparql.cwrc.ca/ontology/cwrc-preamble-EN.html>
- Cifor, Marika, Patricia Garcia, TL Cowan, Jasmine Rault, Tonia Sutherland, Anita Say Chan, Jennifer Rode, Anna Lauren Hoffmann, Niloufar Salehi, Lisa Nakamura. 2019. *Feminist Data Manifest-No.* <https://www.manifestno.com/>.
- Cummings, Joel, and Deborah Stacey. 2018. "Lean Ontology Development: An Ontology Development Paradigm Based on Continuous Innovation." *Proceedings of the 10th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management—Volume 2: KEOD*, 367–74. <http://doi.org/10.5220/0006963003670374>.
- Duarte, Marisa Elena, and Miranda Belarde-Lewis. 2015. "Imagining: Creating Spaces for Indigenous Ontologies." *Cataloging & Classification Quarterly*, 53(5–6): 677–702. <http://doi.org/10.1080/01639374.2015.1018396>.
- D'Ignazio, Catherine, and Lauren F. Klein. 2020. *Data Feminism*. Cambridge, MA: The MIT Press.

Flanders, Julia. 2019. "Building Otherwise." In *Bodies of Information: Intersectional Feminism and Digital Humanities*. Edited by Jacqueline Wernimont and Elizabeth Losh. Minneapolis, MN: University of Minnesota Press.

<https://dhdebates.gc.cuny.edu/read/d02c3ed5-0c55-4de9-88de-5f543fec130/section/f627035f-5fd0-4bd6-ad74-361374ed9a2a>

Gruber, Thomas. 1993. "Toward Principles for the Design of Ontologies Used for Knowledge Sharing." *International Journal Human-Computer Studies* 43, 907–28.

<https://tomgruber.org/writing/onto-design.pdf>

Linked Infrastructure for Networked Cultural Heritage. 2018. "LINCS Application Narrative."

Loukissas, Yanni Alexander. 2019. *All Data Are Local*. Cambridge, MA: The MIT Press.