Homosaurus and Digital *Transgender Archive*[1]
http://homosaurus.org/ and https://www.digitaltransgenderarchive.net/

Reviewed by Brian Watson, Archivist-Historian and Researcher, Kinsey Institute Library

Homosaurus is a linked data vocabulary used as a controlled vocabulary on the *Digital Transgender Archive.* Despite this modern form, it has deep historical roots: it is based on the internal thesaurus of Netherland's International Homo/Lesbian Information Centre & Archives (IHLIA). Two separate institutions—the Homodok research library at the University of Amsterdam and the Anna Blaman Huis of Friesland—pooled resources for LGBTQ+ history to form IHLIA, creating one of the most extensive queer-specific library and archives in the world.[2] Upon their union, the newly-formed IHLIA (now called IHLIA LGBT Heritage) discovered a need to describe their combined collection, but found that there were little to no applicable subject terms. The resulting project, *Queer Thesaurus: An International Thesaurus of Gay and Lesbian Index Terms* (1997) was edited by Ko van Staalduinen, Henny Brandhorst, and Anja Jansma. From 2013 to 2015 Jack van der Wel and Ellen Greenblatt revised, edited, and transformed the vocabulary into linked data.[3] All versions are published in both Dutch and English. This year, the Homosaurus editorial board will release an abridged version "with only LGBTQ-specific terms, that archives and libraries can use as a supplement to LCSH."[4]

Currently, the most thorough application of Homosaurus is the *Digital Transgender Archive* (*DTA*), and thus the selection for this extended review. The *DTA* was formed in 2008 as a result of a conversation between K. J. Rawson, associate professor of English at the College of the Holy Cross, and Nick Matte, historian at the University of Toronto.[5] Their discussion centered on how difficult it was to research trans* history among varied and disparate collections around the world, how difficult it was to sort out the different terms and descriptors used in those archives, and the

---

[1] The author would like to thank K. J. Rawson of the *DTA* and Jack van der Wel of IHLIA GLBT Heritage for their eagerness to discuss their project at depth. Additional thanks go to Marika Cifor and Shannon Devlin for their comments on an earlier version of this review. Final thanks to the insightful edits and critiques of Bethany Anderson and Gloria Gonzalez of the SAA Reviews Portal.
[2] Staalduinen, Ko van, and Henny Brandhorst. 1997. *A queer thesaurus: an international thesaurus of gay and lesbian index terms*. Amsterdam: Homodok.
[3] Personal communication about Homosaurus between Jack van der Wel and author on April 12, 2019.
[4] Personal communication, April 12, 2019.
[5] "Overview," *Digital Transgender Archive*, accessed March 3, 2019, https://www.digitaltransgenderarchive.net/about/overview.

fact that there were little-to-no primary sources available online.[6] The rough sketch became, by 2019, a collaboration of more than fifty different institutions, private collections, and archives around the world, currently directed by Rawson. The College of the Holy Cross hosts and maintains the *DTA*, with additional support from the American Council of Learned Societies.[7] Although the Homosaurus and the *DTA* are separate projects, they are intimately linked: the [Homosaurus homepage](#) refers immediately to the *DTA*; the *DTA*'s motto is "Trans History, Linked," which, if not a direct reference, is at least a hat-tip. For these reasons, and because *DTA* is the best case study for the Homosaurus, the two projects are reviewed here in tandem.

Although historians in *The Journal of American History* and the *American Historical Association* reviewed the *DTA*, neither review included archival perspectives or paid much attention to the underlying technologies like the Homosaurus.[8] Indeed, the Homosaurus, alongside the *DTA*, makes unique technological contributions to the archival profession as a whole that are worth considering *in toto*. For example, the use of metadata crosswalks and the Homosaurus allow the *DTA* to be the largest archives of transgender materials in the world, and the leading exemplar for other digital archives. Finally, although there have been a spate of research projects into linked data vocabularies in galleries, libraries, archives, and museums (GLAM) contexts, there has not been an extensive examination of the interplay between the

---

[6] "Overview."

[7] "Overview."

[8] Finn Enke, "Digital Transgender Archive," *Journal of American History* 104 (June 2017): 315–16, [academic.oup.com, doi:10.1093/jahist/jax165](#); Sadie Bergen, "Trans-Ing History on the Web: The Digital Transgender Archive," *AHA: Perspectives on History*, accessed February 25, 2019, https://www.historians.org/publications-and-directories/perspectives-on-history/summer-2016/trans-ing-history-on-the-web-the-digital-transgender-archive.

two.[9] This review seeks to bridge the gap in current knowledge. Finally, Rawson's discussion of the *DTA* will inform our conversation below.[10]

Homosaurus

To first turn to the Homosaurus: although it is a linked data vocabulary, it occupies the latest manifestation of a long history of controlled thesauri. The first thesauri were publicly outlined in 1947 and then developed and operationalized by DuPont, the US Department of Defense, and the American Institute of Chemical Engineers in 1959 and 1960 to provide better access to company- and subject-specific information for experts and employees.[11] Specialized thesauri support internal information organization, but they can also, as Donna J. Drucker points out, "be powerful tools for challenging and remaking information hierarchies and the social hierarchies embedded within them."[12] Drucker's argument could extend into the current moment: continued collaborations between users and GLAM professionals are necessary for confronting and helping to remedy injustices (racial, sexual, gender embedded in broader informational, social, and political systems—which is something the *DTA* and Homosaurus do, as described below).[13]

The Homosaurus contains nearly 2,000 different terms, all of which are interlinked. Clicking on a term like 'polyamory' brings the user to a term page (i.e., http://homosaurus.org/terms/polyamory, also known as a URI). Each term page contains a variety of boxes that elucidate the term in varying depths. For example, the polyamory page notes that the "Preferred Label" is polyamory, includes a

---

[9] See Gary Munnelly, Harshvardhan J. Pandit, and Séamus Lawless, "Exploring Linked Data for the Automatic Enrichment of Historical Archives," *The Semantic Web: ESWC 2018 Satellite Events*, ed. Aldo Gangemi et al. (Springer International Publishing, 2018), 423–33; Jinfang Niu. "Linked Data for Archives," *Archivaria* 82 (2016): 83–110; Sai Deng, "A Step Forward: Adding Linked Data Vocabularies to Digital Repositories," *Faculty Scholarship and Creative Works*, Accessed May 24, 2018, https://stars.library.ucf.edu/ucfscholar/637; Antoine Isaac and Thomas Baker, "Linked Data Practice at Different Levels of Semantic Precision: The Perspective of Libraries, Archives and Museums: Linked Data Practice at Different Levels of Semantic Precision: The Perspective of Libraries, Archives and Museums," *Bulletin of the American Society for Information Science and Technology* 41 (April 2015): 34–39, https://doi.org/10.1002/bult.2015.1720410411; Jacob Jett, Timothy W. Cole, Myung-Ja K. Han, and Caroline Szylowicz, "Linked Open Data (LOD) for Library Special Collections," *2017 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, 1–2. Toronto, ON, Canada: IEEE, 2017, https://doi.org/10.1109/JCDL.2017.7991604; Cliff Landis, "Linked Open Data in Libraries," *AUC Robert W. Woodruff Library Staff Publications*, Accessed January 1, 2019, http://digitalcommons.auctr.edu/libpubs/30; Erik T. Mitchell, "The Current State of Linked Data in Libraries, Archives, and Museums," *Library Technology Reports* 52 (December 17, 2015): 5–13; and María Hallo, Sergio Luján-Mora, Alejandro Maté, and Juan Trujillo, "Current State of Linked Data in Digital Libraries," *Journal of Information Science* 42 (April 2016): 117–27, https://doi.org/10.1177/0165551515594729.

[10] See K. J. Rawson, "The Rhetorical Power of Archival Description: Classifying Images of Gender Transgression," *Rhetoric Society Quarterly* 48 (August 8, 2018): 327–51, https://doi.org/10/gfkhvc.

[11] Donna J. Drucker, "How Subjects Matter: The Kinsey Institute's Sexual Nomenclature: A Thesaurus (1976)," *Information & Culture: A Journal of History* 52, no. 2 (2017): 207–28, https://doi.org/10/gfrkv8.

[12] Drucker, "How Subjects Matter."

[13] Drucker, "How Subjects Matter."

definition ("practice of romantic relationships with more than one person, simultaneously, with the knowledge and consent of all parties"), additional metadata such as creation date, and then links out to "Broader," "Related," or "Narrower" terms as needed. In the example case, the broader term is "sexual relationships" and a related term is "polygamy."

The Homosaurus is not comprehensive—which would be impossible for any vocabulary. The most noticeable lacunae are in sexuality and romance orientation: there is little coverage of asexual community beyond the term "asexuality" such as "aromantic" or "ace." "Polyamory" is included, but terms like "compersion" or "fluid-bonding" are not; another omission is apparent in BDSM/kink terms such as, say, kink. These are items that remain for further revision. Personal communication with members of the Homosaurus' board has revealed that there are plans to allow the submission of new terms via a website form. Despite these arguably minor oversights, the Homosaurus is useful in many academic or GLAM institutions, notably those using modern platforms like Omeka S or Scalar, which support linked data. It is especially worth considering in contexts where information about gender and sexual minorities might be limited or missing. As the case of the *DTA* below illustrates, in the right contexts, linked data vocabularies can create powerful resources. With this understanding of the vocabulary, it is possible to move on to understanding its role in the *DTA*.

The Digital Transgender Archive
As a digital archive, the *Digital Transgender Archive* is profoundly well designed, showing that a great deal of thought has gone into how to organize and present materials in the most efficient and accessible way possible. The *DTA* landing page features an uncluttered header and logo, visible links to "Browse," "Learn," "About," and "Contact" pages. Below the header is a search bar and links to different ways to explore the archives: users can search by an interactive map, or by "Institution," "Collection," "Topic," "Genre," and "Latest" (added). Additionally, the website is screen reader and colorblind-friendly, a requirement for any digital archival project.[14]

The *DTA* itself is organized into nearly a hundred collections representing the holdings of fifty-six different institutions around the world. The total number of items is ever-expanding, but as of January 2019, it includes nearly 7,000 different items along with 250 finding aids that point to other collections, institutions, and items. The *DTA* is free, does not require institutional affiliation, and is accessible to anyone in the world with an internet connection. Because of this, the *DTA* almost-automatically becomes one of the most useful classroom teaching and reference

---

[14] A number of useful resources are available online: Pablo Stanley, "Designing for Accessibility Is Not That Hard," *UX Collective*, June 27, 2018, https://uxdesign.cc/designing-for-accessibility-is-not-that-hard-c04cc4779d94; and "What Is Accessibility?" The Interaction Design Foundation, accessed April 4, 2019, https://www.interaction-design.org/literature/topics/accessibility.

sources for researchers or students of sexuality, gender, and queer history, to name just a few categories.

Owing to the nature of the wide-flung and disparate sources that the *DTA* brings together—including some material which would be difficult to access elsewhere—the *DTA* includes a broad, yet thoughtful definition for transgender:

> The *DTA* uses the term transgender to refer to a broad and inclusive range of non-normative gender practices. We treat transgender as a practice rather than an identity category to bring together a trans-historical and trans-cultural collection of materials related to trans-ing gender. We collect materials from anywhere in the world with a focus on materials created before the year 2000.[15]

Rawson has been open and transparent about the historical practices that guide the construction of the *DTA*. Writing elsewhere, he notes that "Being all too aware of the consequences of exclusion, we enact a firmly queer commitment to err on the side of inclusion as we try to bring in any materials that seem to relate to trans-ing gender, irrespective of the identities of the individuals involved."[16] Additionally, he has noted flaws in *DTA*'s representations of and around nonwhite races, nonbinary genders, and nonwestern identities that originate from the archive's dependency on northeastern American universities. Rawson and the *DTA* are taking steps to correct this trend by "seek[ing] out partnerships with collections around the world. . . [and making] extra efforts to collect "zines and oral histories and other genres and formats that help to correct some of the biases." [17] The same criticisms of the *DTA* outlined above can also apply to the Homosaurus: it was created for and used to describe a collection of mostly European material and people. Regardless, Rawson and others involved with the project recognize the real power of linked data vocabularies and are making conscious attempts at improvement.

Indeed, the language and links on the archive are where the *DTA* goes from being merely curious to exceptional. Each item in the *DTA* occupies a page, which splits into two sections. The first section contains a title, an image of the resource (usually a photograph or screenshot), a download link, social media sharing buttons, a citation generator, and a map widget identifying the physical location or institution of record for the digitized item. The second half of the page includes all of the item-specific metadata—which is perhaps the most engaging part for archivists and librarians. *DTA* item metadata is straightforward, self-explanatory linked data: this is where the Homosaurus really shines.

---

[15] "Policies," *Digital Transgender Archive*, accessed February 24, 2019, https://www.digitaltransgenderarchive.net/about/policies, and https://www.digitaltransgenderarchive.net/about/overview.
[16] See Rawson, 333.
[17] Rawson, qtd in Bergen.

In the example (figure 1) above, an early-twentieth-century work of sensation fiction called *The female-impersonators. . . (1922)*, the item's basic descriptors (author, publisher, publication date, language) are pulled from the Rubenstein Rare Book and Manuscript Library at Duke University. Duke, however, classifies the item under the incredibly dated and offensive subject heading "paraphilia," which is a fantastic argument in favor of the *DTA*/Homosaurus' redescription and reappropriation. *DTA*-specific terms are all linked and, as a result, incredibly useful. For example, clicking on the Places field (New York) brings the user to a list of 800 items tied to New York City, which means that an early-twentieth-century sensational text is placed alongside and contextualized amongst Spanish-language news clippings about the Stonewall Riots and images of posters from 1965 drag show. To take another example (figure 2), a user clicking through the metadata fields for *A Girl At Last* (1975), a work of "crossdressing erotica," juxtaposes erotic literature with powerful historical images and events.

It is these moments of surprise and astonishment that make the *Digital Transgender Archive* a true boon: through *DTA*'s linked vocabulary, trans* history becomes, well, linked to broader cultural narratives and histories. A user can "suddenly discover [them]selves existing."[18] By embracing new technologies and concepts, like linked data vocabularies, the *DTA* has placed itself at the cutting edge of archival, information, and library science and demonstrates potential roads to empowerment for other minority communities. Digital archives can become spaces of reclamation, redescription, reappropriation, and renaming.

Both academics and members of the broader public are well-served through Homosaurus and the *DTA*. The most notable shortcoming of item-level pages is the fact that there is no link back to the original item in its parent institution. While this might be a part of the broader reclamation that the *Digital Transgender Archive* aims for, the current author was unable to find reasoning for this choice. As a result, it was difficult to track down original context and descriptors to compare with the *DTA*'s. While this choice may be intentional—it potentially obscures or eliminates vocabularies that could be inaccurate or offensive—a statement of intentional choice by the *DTA* could help to strengthen its case. Another noticeable deficit originates from those parent institutions: many images and most physical publications only print or display in black and white. Finally, some users might have minor quibbles with the fact that citations are only available in Chicago format and not MLA or APA.

These are minor details indeed. The Homosaurus is an impressive and extraordinary demonstration of the power of linked data vocabularies. The *Digital Transgender Archive* is a testimony to the capabilities of modern tools and concepts. Additionally, the *DTA*'s thoughtful, accessible, and clean design should serve as an

---

[18] Michelle Caswell, et al., "'To Suddenly Discover Yourself Existing': Uncovering the Impact of Community Archives," *American Archivist* 79 (June 2016): 56–81, https://doi.org/10.17723/0360-9081.79.1.56.

aspirational model for any GLAM-associated digital platform to take note of and emulate.