theories and methodologies

Categorically Provisional

SUSAN BROWN

FEMINIST LITERARY HISTORY BALANCES COMMITMENT TO A DIFFERENT FUTURE, ONE BETTER THAN THE PRESENT WITH RESPECT TO GENDER, with an orientation toward the past, whose ways of knowing it seeks to supersede even as it engages with them. The revision of our cultural past through the lens of gender has, by drawing on past categorizations of authors as female, necessarily invoked problematic paradigms in the service of critique and epistemological change. The relation of the digital humanities (DH) to category work is similarly fraught. I offer here my take on the power and peril of classification based on category making in the pursuit of digital feminist literary history through the Orlando Project, an ongoing experiment in using semantic markup for online scholarship.[1] Orlando is known for its online textbase, published with Cambridge University Press, but the team has produced a number of exploratory interfaces and translations of the material into other forms. Over the course of a quarter century of grappling with "the *digital as difference*" (Wernimont and Flanders 430) alongside other feminist projects, I have changed my understanding of classification as my collaborators and I have tried to represent the difference that gender analysis makes when undertaken in a computational environment. I here argue that category work, always vexed, always provisional, is crucial to realizing the potential of DH for representing, analyzing, and fostering difference.

Other terms closely related to category work are relevant here, including *knowledge representation*, *classification*, *knowledge organization*, *epistemology*, and *ontology*. The array of such terms in fields ranging from philosophy to library and information science to computer science to sociology underscores the central role this kind of work plays. Reflecting on categorization in DH highlights a particular set of challenges for our times. We have moved, argues Ronald E. Day, from older documentary systems of classification and indexing to data-based ones that assign "personal and documentary identity as a *function* and a *product* of sociotechnical systems" (35), which is to say that social identities do not precede but are instead

SUSAN BROWN holds a Canada research chair in collaborative digital scholarship at the University of Guelph. She is a founding editor of the Orlando Project and leads two infrastructure projects: the Canadian Writing Research Collaboratory and the Linked Infrastructure for Networked Cultural Scholarship.

© 2020 SUSAN BROWN
PMLA 135.1 (2020), published by the Modern Language Association of America

https://doi.org/10.1632/pmla.2020.135.1.165 Published online by Cambridge University Press

created by and through systems ranging from citation indexes to health-care management systems to social media, which direct current needs and actions as well as shape potential futures. This reflects an epistemic shift in which the "dominance of digital mediation for the overall social positioning and expressions of the subject gives rise to a regime of digital governance" that operates "through promoting and trimming cultural forms and social norms . . . and therefore controlling the affordances for actions" (133, 135). As a result of this shift, classification is immanent in the increasingly digitally mediated world within which identities are continuously performed and thereby constituted.

Scholarship engages pervasively in classification work: "classification in the broad sense involves the identification of the properties of objects of study, and can hardly be dispensed with in any coherent discourse" (Sperberg-McQueen 377). Outlining an area of inquiry, framing an argument, and defining a set of materials for investigation all involve decisions about how to group things and what to include and exclude. However, humanities scholarship can also eschew categorization. Eve Kosofsky Sedgwick noted that *Epistemology of the Closet* is most queer in its refusal to use the word *queer*, arguing that "the specificity, materiality, and variety of sexual practices . . . can be done better justice in a context where the impoverished abstractions that claim to define sexuality can be treated as not authoritative"; at stake is less the "conceptual incoherence" of categories than their political effects (xvi, xvii). The effects of categorization have been charted in numerous fields. "To classify is human," argue Geoffrey C. Bowker and Susan Leigh Star (1), who demonstrate the ubiquity and material force of how categories operate in a wide range of contexts including apartheid, censuses, the field of nursing, and the World Health Organization's International Classification of Diseases, as well as information infrastructures. Bowker and Star stress the ethical importance of studying the often-invisible systems of classification in our lives: "Each standard and each category valorizes some point of view and silences another. This is not inherently a bad thing—indeed it is inescapable. But it *is* an ethical choice, and as such it is dangerous—not bad, but dangerous" (5–6).

Sedgwick's "refusal to pretend to make sense within a bifurcated discourse that did not make any sense of me" (xvii) makes perfect sense as a discursive strategy for unsettling categories responsible for pain and violence. Yet much digital humanities work simply does not work without categorization: computers rely on formal categories to "read"—that is, to store, organize, access, process, and render materials. All computing requires classing or categorizing things, and any computational work that goes beyond the rendering of relatively static materials, such as randomly ordered images, to the organization of bodies of content for search and analysis—whether building archives, text mining, text encoding, or building databases—relies on classification as a condition of possibility.

Jacques Derrida observes of classification, "literary genre but also gender, genus, and taxonomy more generally . . . brings with it the question of law, since it implies an institutionalized classification" (221). DH category work, like any classificatory act, either draws on or makes a bid for authority. In the case of DH, that authority is reinforced by scientism (Drucker), which accords disproportionate ideological weight to methods associated with computation or quantification. Classification work also participates in informational infrastructure building and the ideology of postindustrial knowledge work that relies on "the separation of content from material instantiation or formal presentation" (Liu 216). Even on a modest scale, the potential for digital content to be copied, replicated, relocated, decontextualized, and divorced from its source makes classification

https://doi.org/10.1632/pmla.2020.135.1.165 Published online by Cambridge University Press

work on contested cultural categories such as gender, race, or sexuality more perilous than work destined for print or digital print surrogates. As they structure data, classificatory actions become not just discursively but operationally performative: they provide a means by which content can be sorted, queried, retrieved, transformed, suppressed, or exposed in environments distant from those in which it was originally produced, concurrently silencing and obscuring some aspects of the data associated with that content. Classifications in a DH context instantiate categories that wield power in assigning people and things to categories that shape the infrastructure of our information environments with far-reaching consequences, such as whether one can even search for a particular identity category within a given system.

The interlocking relations of classification to scientism and informational infrastructures is precisely why DH must not relinquish it. Much of the relevance of humanities work in computational contexts depends on engagement with machine-readable classification schemes in ways that allow scholars to critique them and to try to intervene with the alternative classifications that advance humanities-driven inquiries and the machine-readable representations of the world that increasingly shape human lives. Feminist literary history is but one example that shows the feasibility and the urgency of intervening critically in digital category work despite the risks and limitations.

Feminist analysis begins with a critique of past and present gender constructions and ideologies. However, critique alone is insufficient if investigation of those constructions and ideologies is to proceed computationally. To reject classification by gender is to render the effects of gender invisible. The challenging task is to make women count, differently (Mandell). If authors are not categorized by sex or gender, one simply cannot analyze an archive or a corpus of texts for certain kinds of gender bias. Such classification is essential to developing representative corpora, which are in turn necessary for advancing stylistic and literary-historical inquiry based on large-scale text analytics.

The binary model of sex as m/f is of course inadequate in relation to contemporary theories of gender, according to which the distinction arises not from fixed or essential biological difference but from a complex intersection of social and signifying practices. As Miriam Posner argues, "most of the data and data models we've inherited deal with structures of power, like gender and race, with a crudeness that would never pass muster in a peer-reviewed humanities publication." Modeling, a process of making explicit the "shaping choices we make in representing or analyzing the materials we study" (Flanders and Jannidis 3), has been theorized extensively as a core DH activity (McCarty) that is not only representational and context-dependent (Ciula and Eide) but able to foreground the constructedness and contingency of knowledge (Piper 9). Modeling fosters "the proliferation of many worlds," argues Roopika Risam, by placing "the liminal, abject, or subaltern . . . at the center of inquiry in the digital cultural record" in such a way as to challenge the hegemonies and epistemologies of the "real" world (33, 34, 33). However, it is a challenge, given technological constraints, to create a model of gender that reflects the complexity of understanding we have reached through decades of debate in women's and gender studies. How can we represent gender as a spectrum or a historical process while also reifying it enough to analyze its effects and affects across different datasets and domains in which the meanings will be, finally, incommensurable? Jacqueline Wernimont and Julia Flanders have discussed the process of semantically encoding texts with Extensible Markup Language (XML). XML embeds metadata in a digital text, indexing it through the insertion of human-

https://doi.org/10.1632/pmla.2020.135.1.165 Published online by Cambridge University Press

readable and machine-processable labels that categorize features of the text. For example, XML might be used to tag a personal name—"<persName>Virginia Woolf</persName>"—thereby marking it for recognition and processing by a computer programmed to use the tags for indexing, searching, retrieving, or counting. Wernimont and Flanders argue that modeling digital scholarly editions of early women's writings, for instance by formalizing the types of information such as variants or notes that they contain, enables productive category work precisely by confounding "the processes of categorization, explication, and description central to digital text markup" (428). They align the Women Writers Project with Donna Haraway's "Cyborg Manifesto" in the "different, more intimate experience of boundaries and their constructedness" provoked by digital as compared with traditional editing, for instance, in realizing the futility of attempting to sequester the feminized genre of romance from considerations of empire and state (429). Problematizing category boundaries is critical not only for the feminist project of bridging from legacy vocabularies to new ones but also for basic digital literacy, given that categories and the values they embed govern information systems.

The Orlando Project sought a new means of inscribing social identity, also using XML. It modeled the complexity of cultural formation by embedding tags for identity categories within prose narrative, rather than providing decontextualized descriptors in, for instance, a table or list, and by devising tags to indicate the discursive contexts such as religion, language, ethnicity, social class, sexuality, race, or color in which identity descriptors occur, while leaving the text in the tag unconstrained by a fixed vocabulary. It combined a degree of processability (one can gather all descriptors related to a specific context, such as sexuality) with a refusal of standard(ized) terminology such as a predetermined list of terms that might, for instance, permit use of the term *lesbian* but not the older *sapphist*. The aim was to inculcate awareness of the historical variability and contingency of categories even though classification, in the sense of making "explicit a particular view concerning the nature of the objects being classified" (Sperberg-McQueen 377), was endemic to Orlando's approach to literary history.

Such reliance on unstructured or natural language and nonce vocabularies can work for independent, stand-alone resources, though it may make it harder for people to find things. To make digital resources accessible through common interfaces so that their knowledge can be combined, however, requires a greater degree of regularization. Hence authoritative vocabularies such as the Library of Congress Subject Headings (LCSH), a system created in 1898. The limitations of vocabularies mirroring past epistemes are legion—consider the imperialist implications of the LCSH terms "Indians of North America" and "Eskimo," and the heterosexism of "Butch and Femme (Lesbian culture)"—underscoring the importance of category work that reflects current epistemologies and debates. Notwithstanding our allegedly new knowledge infrastructures, such vocabularies are still used extensively in the catalogs of thousands of libraries globally, and the LCSH vocabulary is increasingly being leveraged to enable the search, discovery, and use of millions of digital objects online. This kind of legacy classification that, in the words of Hope A. Olson, "marginalizes and excludes *Others*" and ignores "concepts outside of a white, male, eurocentric, christocentric, heterosexual, able-bodied, bourgeois mainstream" (142) will prevail well into the future unless we work toward more permeable, capacious, and dynamic alternatives (227).

The tension between nonce vocabularies and standardization made itself felt as Orlando prepared to migrate to the Canadian Writing Research Collaboratory (CWRC), a platform for producing and publishing di-

https://doi.org/10.1632/pmla.2020.135.1.165 Published online by Cambridge University Press

verse research projects in literary and cultural scholarship within a common interface. CWRC aims to push content into the Semantic Web, an initiative to make online resources more discoverable, shareable, and interoperable by making them machine-readable, in part through employing unique identifiers not only for named entities such as people but also for concepts such as relations, categories, and personal properties, so that dispersed online resources can be interlinked to form a vast, structured body of information akin to a database.[2] This means that to disseminate Orlando content on the Semantic Web requires transforming Orlando terminology, including those cultural form tags embedded in natural language as described above, to produce a formalized, which is to say a less flexible and more regulated, classification strategy in the shape of what is called an ontology.

An ontology in a computational context means not the abstract study of being, as it does in philosophy, but a model of a domain of knowledge in machine-processable form. Semantic Web ontologies require the definition of semantic structures, which is to say the definition of concepts, categories, and the logical relations among them. Semantic Web assertions take the form of subject-predicate-object structures called triples. Relations in this context are the predicates, also called properties, that specify in precise ways that computers can process how one thing, such as a concept or a person, is connected to another. So to return to that <persName>Virginia Woolf</persName> tag, an ontology could establish that names can be properties of persons but not of books, which would mean that the assertion "Virginia Woolf [using an identifier for Woolf that was classed as a person] hasName 'Virginia Woolf'" would be valid but the assertion "Virginia Woolf [the book by Hermione Lee] hasName 'Virginia Woolf'" would not be. For the book you would use a relation or predicate like "hasTitle." So the different classes distinguish two different entities with identical labels and also govern the types of properties those entities have. Semantic relations thus organize and process data, which is to say digital content. Ontologies are meant to provide the basis for a "shared conceptualization" (Gruber 909), but formalizing them for machines often seems to take precedence over attempts to communicate that conceptualization to human beings. In contrast to the strategies used by many ontologies and vocabularies, the CWRC strategy has been to provide full documentation of the relations among the ontology's most significant terms, such as those for social identities, and to provide definitions for categories and key terms that cannot be sourced from external ontologies. The International Committee for Documentation's Conceptual Reference Model (CIDOC-CRM), an established ontology for describing cultural heritage objects, also adopts this strategy of ample documentation, and it consults closely with its community as it evolves (Oldman et al. 261). This approach is more involved than assuming that a term's meaning is self-evident, but it has the advantage not only of promoting a common approach and shared conceptualization but also of allowing the model to change in response to the evolving understanding of the community that uses it. Because humanities disciplines recognize that knowledge production is far from neutral, and because they employ a range of theoretical frameworks in different areas of specialization, this kind of documentation is crucial, forming a basis for evaluation and debate of the intellectual work done by ontological classification.

Because Semantic Web and linked open data technologies do not work optimally with natural language or so-called strings, the CWRC ontology will translate a good portion of Orlando's descriptors into terms formally defined in the ontology. To avoid blurring the distinctions among contested terms occurring in multiple cultural formation tags, the ontology retains these distinctions through a

https://doi.org/10.1632/pmla.2020.135.1.165 Published online by Cambridge University Press

layered strategy of classification and labeling. String descriptors such as *Jewish* are translated in the ontology into properties (that is, semantic relations) that are categorized according to the tag that contained them, so that the categorization reflects the way the terms were contextualized in the narrative text by tags, such as ethnicity and religion. This strategy respects Semantic Web requirements for defined descriptors without reducing complex identities to a unique term with a singular meaning. Jewish cultural identity is thus represented in the ontology through multiple terms that retain their discursive context by means of classification. The ontology translates the range of terminology associated with Jewishness within Orlando into four terms belonging to different classes to indicate that those terms are different ways of framing identity (see fig. 1). There are thus multiple meanings for a term that in other datasets—such as a census—might be treated as singular and fixed. Classing the same term in different ways clearly indicates instability and slippage, as does allowing for individuals to be associated with more than one version of a term, as shown by Grace Aguilar and Eva Figes in figure 1.

To bring together terms in different contexts, the ontology groups granular descriptors under labels that are explicitly representations. So the label *Jewish* assembles different discursive constructions of Jewishness, foregrounding their *différance* yet allowing anyone searching for Jewish writers to find them. The terms representing Jewish identity in the ontology are not comprehensive; rather, they are based on the text of Orlando, which reflects both its subject matter and the theoretical approaches of its creators. Indeed, as indicated in the diagram, the model is designed to retain the links between Semantic Web assertions and the sources from which they are derived, since documenting sources is crucial to humanist scholarship. The CWRC vocabulary for cultural forms could be extended on the basis of other content to accommodate further terms and classes, such as the distinction between Sephardic and Ashkenazi Jews, or different denominations of Judaism.

In effect, the discursive framing and contextualization of Jewishness in the ontology enables the data extracted from Orlando to indicate the social and discursive complexity of the terms and how they relate to individuals. The diverse classifications, when combined, make an embedded but discernible argument about the historical shifts that Jewish cultural identity has undergone with re-

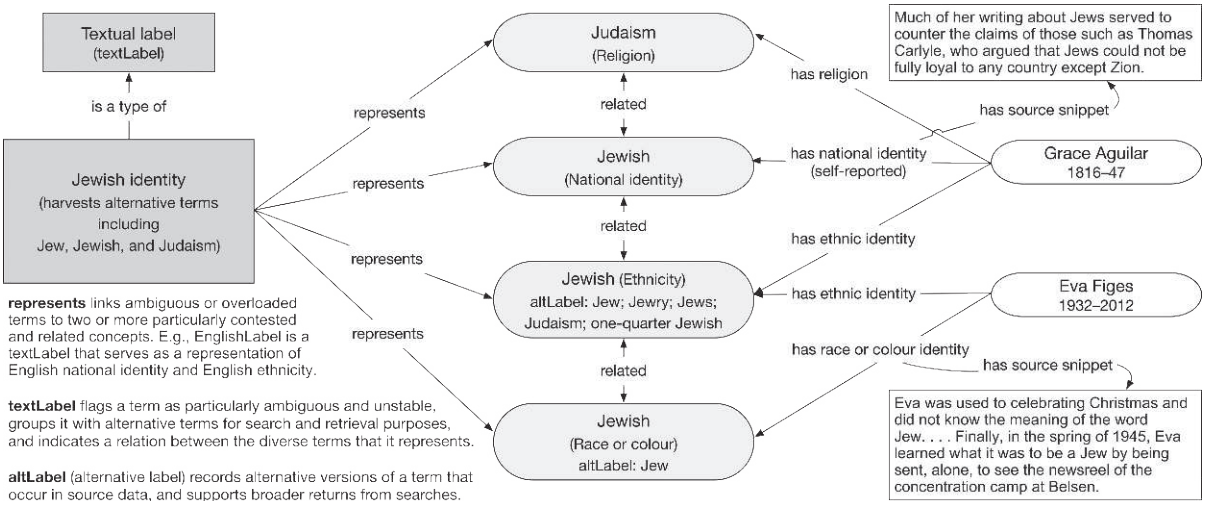


FIG. 1
Simplified diagram of CWRC ontology terms and relations that structure descriptions of the cultural identities of Grace Aguilar and Eva Figes. Assertions between elements in boxes are read in the direction of the arrows.

https://doi.org/10.1632/pmla.2020.135.1.165 Published online by Cambridge University Press

spect to British women writers, revealing that while religion was invoked most frequently in relation to earlier writers, ethnicity has been more frequently invoked in recent years. Discursive contexts can reflect the complex tension between identity as externally imposed and identity as internally embraced, such as when Jewishness is framed not only as the national identity of the Victorian writer Aguilar but also in opposition to Thomas Carlyle's rigid notions of singular, exclusive nationality: she also defines herself as an English national. Such contextualizing can also point to historical watersheds. The earliest framing, measured by the birth year of the author, of the term *Jew* as a race or color in Orlando occurs in the description of Figes, who, as a German whose family had emigrated to England, "learned what it was to be a Jew by being sent, alone, to see the newsreel of the concentration camp at Belsen" (Brown et al.; see fig. 1). Jewishness here is evoked alongside Jewish ethnicity, German national heritage, English nationality, and the Christian religion. The ontology is designed to support distinctions between related social identities, making space for nuance, complexity, and apparent contradiction in the representation of how social conditions inform literary history.

The distinction between labels and categories emerges from the fact that technologies for creating and processing classification systems were not designed to deal with ambiguous and overlapping social identities. The difficulties in a Semantic Web context stem from the limited relations or properties available for asserting similarity in the Web Ontology Language (OWL), which governs how ontologies are written. OWL's "sameAs" property asserts complete identity between the two things that are the subject and object of the "owl:sameAs" predicate, and there is no alternative for things that are close but not exactly the same. This deficiency has led to great imprecision in how sameness is expressed computationally (Halpin et al.). Humanities scholarship demands a better means of dealing with this problem than conflating sameness and near sameness, precisely because scholars engage regularly with the conjunction of relatedness and incommensurability, for instance when they draw on sources that employ binary categories for sex while employing intersectional gender analysis that eschews binaries. Being able to hold related but distinct terms together without conflating them is crucial to critical inquiry and to the articulation of different worlds, as invoked by Risam, that can break free of past epistemes. Experimenting and comparing alternative systems of classification from a humanities perspective thus stands to advance the provisional world making of both modeling and classification technologies.

CWRC faces the challenge of addressing the tension between sex and gender in relation to the widely adopted categories of the International Standards Organization's "codes for the representation of human sexes," ISO 5218:2004, according to which numerals signify values: 0 is "not known," 1 is "male," 2 is "female," and 9 is "not applicable" ("ISO/IEC"). The OWL "sameAs" designation is not an appropriate way to relate these terms to gender terms, since gender and sex are not the same. However, asserting some relation is required in order to link gender data computationally to that of legacy systems that employ sex classification. Many Web resources, such as collections of digitized texts containing metadata about their authors, employ these codes; it is desirable for certain purposes, such as matching authors or enhancing information about them, both to draw on and interlink with this content without endorsing the conceptualization of sex embedded in the codes. CWRC therefore devised a formal means to allow terms to be related without being equated to external vocabulary terms, including those with which the project vehemently disagrees. The predicate "functionally related" indicates a term that is semantically

https://doi.org/10.1632/pmla.2020.135.1.165 Published online by Cambridge University Press

incommensurate with the linked CWRC term but that occupies a position that may be considered functionally parallel to that term in other datasets.

As it did for Jewishness, the ontology groups an array of terms for gender by means of labels, in this case labels such as *queer*, *man*, and *woman*. These broad labels can represent related instances of gender; the latter, for example, represents *cis woman* and *trans woman*. When it comes to linking to the International Standards Organization classifications, the terms *woman* and *man* are "functionally related" to the ISO 5218:2004 numerals for sex, 2 and 1 respectively, a property that links them without asserting their equivalence (fig. 2). Indicating a functional relation between terms lays the ground for using machines to bridge categories while emphasizing that they are semantically distinct.

As will be evident by now, the classification strategies I am advocating establish categories provisionally, working from specific materials or data in a porous structure that allows for revision and expansion over time. To be widely adopted, an ontology for gender would need to accommodate other terms that emerge from literary-historical work quite different from Orlando's. For instance, Laura Mandell is working to derive from eighteenth-century fiction gender terms that are alternatives to the ones that became hegemonic, while the text of the project Lesbian and Gay Liberation in Canada, led by Constance Crompton and Michelle Schwartz and building on the work of the archivist Don McLeod, contains recent terms such as *butch* and *femme* that do not appear in Orlando sexual identity tags. The CWRC ontology could expand to accommodate both. But one ontology cannot do it all. To ensure that projects such as The People and the Text, an Indigenous literary history project led by Deanna Reder, have access to terms that reflect Indigenous ways of knowing, domain-specific ontologies such as CWRC's can coordinate with broader initiatives, such as partnerships between Indigenous communities and libraries, to decolonize description. Such initiatives aim through reciprocity with Indigenous communities to wrest archival and library vocabularies from the legacies of colonialism and align them with Indigenous ways of knowing ("Surfacing Knowledge"). As machine analysis based on large bodies of digital content becomes more prevalent, it will also be important to figure out how human-defined classifications can interact effectively with those emerging from algorithmic processes such as text mining, topic modeling, and machine learning (Brown and Mandell). Dynamic and flexible models for relating diverse terms and classificatory systems will be crucial.

These are all serious challenges that require engagement, experimentation, and collaboration across multiple communities. Yet engaging with and in category work can be satisfying and productive precisely because

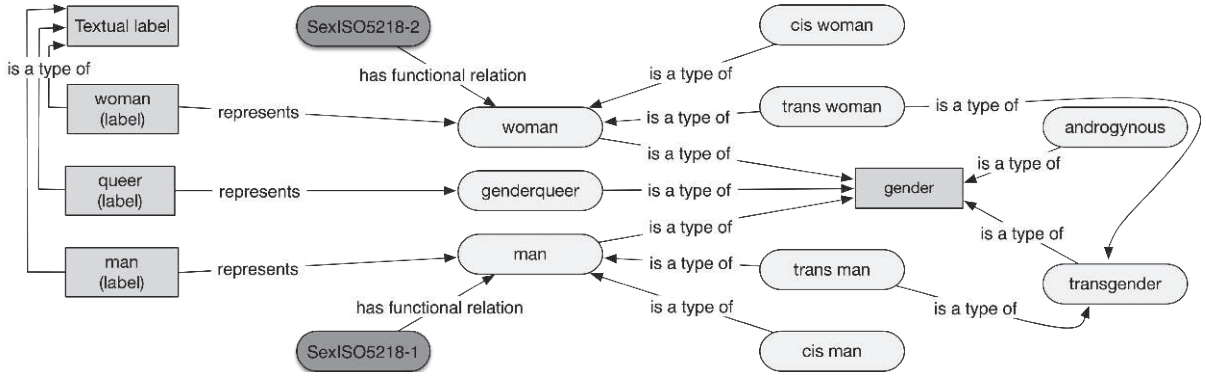


FIG. 2
Gender term relations, including their functional relation to the ISO values for sex.

https://doi.org/10.1632/pmla.2020.135.1.165 Published online by Cambridge University Press

of the stakes and magnitude of the task and because of the ways experimental classification strategies can be mobilized by machines. *Orlando*, the Web resource published by Cambridge University Press, has been credited with encouraging readers "to think in new ways" (Nixon 379) and with providing a model that supplants encyclopedic print resources (Harner). One can experience and critically assess what the project's category work does representationally and analytically in the complex interactions among its backend classifications, its content, and its interface. The work is imperfect and provisional, and by design open to critique and revision in ways that enable feminist literary history to move forward with a difference.

DH work offers opportunities to intervene, with experimental models, in the knowledge structures of our time. Such interventions require the leveraging of past classification schemes as well as the forging of alternatives. Like feminist literary history, DH takes a double stance in that it engages the models of the past while understanding that such models have often been superseded intellectually.

Cultural critics like Alan Liu have described infrastructure as a defining feature of our moment. DH category work has the potential to be persistent, pervasive, and powerful as a part of a global online knowledge infrastructure, and as such offers hope for countering in a small way the weight of commercial and hegemonic interests. It has the potential to make interventions that disseminate more nuanced category work to trouble oversimplifications of identity and difference. The problems are significant, and there are no easy solutions. Continual engagement that allows for flexible reworkings will only ever yield partial and inadequate models for the unevenness and complexities of human culture and activities, which incessantly change and can never be fully modeled or classified. Yet that is no more reason for abandoning such work than is the impossibility of writing the last word on any subject. At the conclusion of his meditation on the complicated, vexed, and also productive interactions among labels for race, gender, ethnicity, class, and the diverse lived experiences of identities, Kwame Anthony Appiah notes, "We are denizens of an age in which our actions, in the realm of ideology as in the realm of technology, increasingly have global effects" (218–19). Ideology and technology meet nowhere more intimately than in the classification of identities with computers. For Kant, the categorical imperative provided an objective, unconditional basis for action. I have argued instead that acts of categorization representing contested and contingent cultural concepts in machine-processable form are necessary precisely because they are provisional. Provisional categorization enables us to work with machines both to critique the past and to articulate possible futures with a difference.

Notes

My profound thanks to my collaborators in the Orlando Project and the Canadian Writing Research Collaboratory, and in particular those with whom I have had the most sustained debates over markup, taxonomies, ontologies, and vocabularies: Jeffery Antoniuk, Katharine Binhammer, Patricia Clements, Joel Cummings, Jasmine Drudge-Willson, Sharon Farnel, Susan Fisher, Isobel Grundy, Susan Hockey, Mihaela Ilovan, Abigel Lemak, Kimberley Martin, Alliyya Mo, Mariana Paredes-Olea, and Deb Stacey.

1. The project, begun in 1995, produces the Web resource *Orlando: Women's Writing in the British Isles from the Beginnings to the Present* (Brown et al.) and explores the potential that semantic encoding offers for literary history. See Holland and Brown.

2. For an introduction to the Semantic Web and the linked open data technologies it uses, see Berners-Lee et al.; Oldman et al.; and Powell and Hopkins. For the more technically experienced, see the World Wide Web Consortium page on linked data ("Linked Data").

Works Cited

Appiah, Kwame Anthony. *The Lies That Bind: Rethinking Identity*. Liveright / W. W. Norton, 2018.

https://doi.org/10.1632/pmla.2020.135.1.165 Published online by Cambridge University Press

Berners-Lee, Tim, et al. "The Semantic Web." *Scientific American*, vol. 284, no. 5, 2001, pp. 34–43.

Bowker, Geoffrey C., and Susan Leigh Star. *Sorting Things Out: Classification and Its Consequences*. MIT Press, 1999.

Brown, Susan, et al. *Orlando: Women's Writing in the British Isles from the Beginnings to the Present*. Cambridge UP, 2006–19, orlando.cambridge.org/.

Brown, Susan, and Laura Mandell. "The Identity Issue." *Journal of Cultural Analytics*, 2018, doi:10.22148/16.020.

Ciula, Arianna, and Øyvind Eide. "Modelling in Digital Humanities: Signs in Context." *Digital Scholarship in the Humanities*, vol. 32, no. 1, 2016, pp. 33–46.

Day, Ronald E. *Indexing It All: The Subject in the Age of Documentation, Information, and Data*. MIT Press, 2014.

Derrida, Jacques. *Acts of Literature*. Edited by Derek Attridge, Routledge, 1992.

Drucker, Johanna. *SpecLab: Digital Aesthetics and Projects in Speculative Computing*. U of Chicago P, 2009.

Flanders, Julia, and Fotis Jannidis. "Data Modeling in a Digital Humanities Context." *The Shape of Data in Digital Humanities: Modeling Texts and Text-Based Resources*, edited by Flanders and Jannidis, Routledge, 2019, pp. 3–25.

Gruber, Thomas R. "Toward Principles for the Design of Ontologies Used for Knowledge Sharing." *International Journal of Human-Computer Studies*, vol. 43, nos. 5–6, 1995, pp. 907–28.

Halpin, Harry, et al. "When Owl:sameAs Isn't the Same: An Analysis of Identity in Linked Data." *The Semantic Web—ISWC 2010*, edited by Peter F. Patel-Schneider et al., Springer, 2010, pp. 305–20.

Harner, James L. *Literary Research Guide*. 5th ed., Modern Language Association, 2008.

Holland, Kathryn, and Susan Brown. "Project, Process, Product: Feminist Digital Subjectivity in a Shifting Scholarly Field." *Bodies of Information: Intersectional Feminism and the Digital Humanities*, edited by Elizabeth Losh and Jacqueline Wernimont, U of Minnesota P, 2018, pp. 409–33.

"ISO/IEC 5218:2004." *International Standards Organization*, 2015, www.iso.org/standard/36266.html. PDF download.

Lesbian and Gay Liberation in Canada Project. Directed by Constance Crompton and Michelle Schwartz, lglc.ca/.

"Linked Data." *W3C Consortium*, www.w3.org/wiki/LinkedData.

Liu, Alan. *Local Transcendence: Essays on Postmodern Historicism and the Database*. U of Chicago P, 2008.

Mandell, Laura. "Gender and Cultural Analytics: Finding or Making Stereotypes." *Debates in the Digital Humanities 2019*, edited by Matthew K. Gold and Lauren F. Klein, U of Minnesota P, 2019, pp. 3–26.

McCarty, Willard. *Humanities Computing*. 2005. Palgrave Macmillan, 2014.

Nixon, Cheryl. "A Sampling of Results." Review of *The Orlando Project*. *Eighteenth-Century Fiction*, vol. 22, no. 2, 2009, pp. 375–79.

Oldman, Dominic, et al. "Zen and the Art of Linked Data: New Strategies for a Semantic Web of Humanist Knowledge." *A New Companion to Digital Humanities*, edited by Susan Schreibman et al., Blackwell Publishing, 2016, pp. 251–73.

Olson, Hope A. *The Power to Name: Locating the Limits of Subject Representation in Libraries*. Springer Science and Business Media, 2013.

Piper, Andrew. *Enumerations: Data and Literary Study*. U of Chicago P, 2018.

Posner, Miriam. "What's Next: The Radical, Unrealized Potential of Digital Humanities." *Miriam Posner's Blog*, 27 July 2015, miriamposner.com/blog/whats-next-the-radical-unrealized-potential-of-digital-humanities/.

Powell, James, and Matthew Hopkins. *A Librarian's Guide to Graphs, Data and the Semantic Web*. Chandos/Elsevier, 2015.

Risam, Roopika. *New Digital Worlds: Postcolonial Digital Humanities in Theory, Praxis, and Pedagogy*. Northwestern UP, 2019.

Sedgwick, Eve Kosofsky. *Epistemology of the Closet*. U of California P, 1990.

Sperberg-McQueen, C. M. "Classification and Its Structures." *A New Companion to Digital Humanities*, edited by Susan Schreibman et al., Blackwell Publishing, 2016, pp. 377–93.

"Surfacing Knowledge, Building Relationships: Indigenous Communities, ARL and Canadian Libraries." *York University Libraries*, scbrwiki.library.yorku.ca/.

Wernimont, Jacqueline, and Julia Flanders. "Feminism in the Age of Digital Archives: The Women Writers Project." *Tulsa Studies in Women's Literature*, vol. 29, no. 2, 2010, pp. 425–35.

https://doi.org/10.1632/pmla.2020.135.1.165 Published online by Cambridge University Press