

Decolonizing Metadata: Lessons from Stolen Relations' Controlled Vocabulary Development

Mairelys Lemus-Rojas¹  , Patrick Rashleigh²  , and Khanh Vo²  

¹ University of Central Florida, ² Brown University

Abstract

Metadata should be understood as an interpretive practice and not just as a technical framework for describing digital objects. It holds power and facilitates community engagement. Within the digital humanities arena, metadata plays a pivotal role in narrating and recovering stories that have remained obscured or misrepresented in historical records. This raises a fundamental question: how can we more humanly describe Indigenous communities whose identities and relationships to kinship and culture have been misrepresented in colonial records? This paper examines the role of controlled vocabularies in shaping the representation of Indigenous histories in *Stolen Relations'* digital humanities project, positioning metadata as a form of archival intervention. It demonstrates how iterative feedback informs the refinement, implementation, or creation of controlled vocabularies and positions metadata as a space where descriptive practices are examined and reshaped.

1. Introduction

The project *Stolen Relations: Recovering Stories of Indigenous Enslavement in the Americas* is a collaborative digital humanities endeavor undertaken by Brown University in partnership with Indigenous communities in the region commonly known as New England (Dawnland), an area of land that is part of what we know today as the United States [1]. While *Stolen Relations* aims for a hemispheric documentation of the history of Indigenous enslavement, the initial start of the project primarily focused in this area. It started in 2015 with the vision of recovering the stories of individuals from tribal communities whose lives were impacted by enslavement [2], [3]. By doing so, the project brings visibility to a part of American history that is too often either underrepresented or misinterpreted. It recognizes that by uncovering and sharing these stories, one-by-one, we can begin to understand the struggle that Indigenous communities endured and overcame. Metadata, through reparative and controlled vocabularies, continues the efforts of the project to decolonize and reconcile this history.

2. Related Work

Questions of power, representation, and historical accountability have long shaped scholarship on metadata and knowledge organization in libraries, archives, and information systems. Since at least the late twentieth century, researchers in library and archival studies have argued that metadata and knowledge organization systems are not neutral structures, but rather reflect the cultural values, historical contexts, and power relations in which they are created [4], [5]. More recently, librarians such as Emily Drabinski have advanced interdisciplinary approaches that draw on queer theory to critically examine the “categories and systems of naming that characterize library knowledge organization schemes” [6]. Institutions including Northeastern University

Full Paper

Available Aug 01, 2026

Correspondence to

Mairelys Lemus-Rojas
mairelys.lemus-rojas@ucf.edu



Copyright © 2026 Lemus-Rojas *et al.*. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license, which enables reusers to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator.

and the University of Utah have similarly developed tools and assessment frameworks for inclusive and reparative metadata that seek to address harmful or exclusionary language in archival description [7]. Other scholars have called for a rethinking of archival practices that prioritize reciprocal relationships with Indigenous communities, adherence to Indigenous protocols, and accountability to the people and histories represented in archival records [8], [9].

These concerns have also informed the development of large scale digital projects focused on slavery and enslavement, where decisions about metadata shape how historical experiences are documented, interpreted, and made discoverable. *Peoples of the Historical Slave Trade* project, commonly known as *Enslaved.org* [10], further explained in the next section, has been influential in advancing critical approaches to metadata design and has provided a model from which other digital projects, including *Stolen Relations*, have drawn their metadata structure. *Freedom on the Move* [11], a project that collects and makes available all surviving North American runaway slave advertisements from 1706 to 1865, relies on crowdsourced data collection and cleanup, thus adopts relatively few standardized controlled vocabularies in its descriptive metadata to accommodate a broad range of contributors. Other initiatives, such as the *Digital Library on American Slavery* [12], emphasize comprehensive documentation through robust descriptive metadata and schemas informed by existing works of historians including Loren Schweninger and Lathan Windley. Building on these efforts, *Stolen Relations* adopts an iterative approach to metadata development that centers care, respect, and accountability in the representation of individuals and communities documented through archival materials.

Beyond libraries, archives, and digital projects, related questions about representation, documentation, and the values embedded in data have also emerged in computational disciplines. From a machine learning perspective, Miceli et al. critique the concept of data “bias,” arguing that it fails to acknowledge the power relations embedded within data as well as the labor involved in data collection and documentation practices [13]. Their observations are particularly relevant to projects such as *Stolen Relations*, which rely extensively on colonial records that reflect the perspectives and priorities of the institutions that produced them. Similarly, Scheuerman et al. examine the values embedded within computer vision datasets, contrasting efficiency, universality, impartiality, and model work with often overlooked values such as care, contextuality, positionality, and data work [14]. Heger et al. further explore the frameworks through which machine learning practitioners document and communicate their modeling decisions [15].

This body of scholarship demonstrates that metadata is not just a technical structure, but a means through which historical narratives and relationships to communities are represented. These concerns are particularly relevant to *Stolen Relations*, where records documenting Indigenous enslavement must acknowledge both the limitations of colonial sources and the perspectives of the communities represented within them.

The project's approach to descriptive metadata emerges from this broader conversation, seeking to support discovery while also advancing more accurate and respectful representation.

3. Descriptive Metadata

Descriptive metadata in the context of cultural heritage information systems not only facilitates discovery, but it can also be a powerful tool for reflecting historical accuracy as it shapes how records are interpreted. Metadata frameworks define core elements to be used in describing cultural heritage resources while also recommending the use of controlled vocabularies to improve consistency and interoperability. Controlled vocabularies are used within descriptive metadata fields to provide a standardization of the concepts represented across a dataset. In other words, they constrain the use of values that are allowed within a specific field. This standardization ensures consistency and facilitates disambiguation, allows interoperability across systems, and enables the computational use of datasets. However, there are many reasons why existing vocabularies may not adequately meet the needs of a project, particularly when they reinforce colonial language present in historical records. The *Stolen Relations* project worked in collaboration with tribal representatives from the region commonly known as New England (Dawnland) to develop controlled vocabularies to more accurately and humanly represent the individuals found in the historical records.

The project, itself, is built on a Wordpress site with a search interface that is powered by a SQL (Structured Query Language) database. The dataset is structured around "sources" (bibliographic data) and "records" (biographical data) although these two categories constitute records within the database at a general level. Through an entry form, contributors are able to create bibliographic metadata to describe historical sources containing information on individuals affected by enslavement and those who facilitated enslaving systems. Information from the sources is then used to create entries for individuals whereby individuals are grouped by "records". These entries contain biographical data both unstructured and structured.

An important consideration of the project was to avoid reproducing dehumanizing terminology while maintaining historical transparency. Harmful and derogatory terminologies, while historically accurate, are preserved in transcriptions, but this information is then reinterpreted and contextualized before it is mapped to controlled vocabularies. The vocabulary development process included the adoption of terms from *Enslaved.org*, a domain-specific project which developed its own specialized vocabulary [16]. This adoption ensured interoperability with other datasets documenting enslavement. In instances where these vocabularies were found insufficient, they were extended or refined to provide more specificity. In addition, new vocabularies were crafted to more accurately represent the complexities found in the historical records documenting Indigenous enslavement.

Before the project launched in May 2025, user experience (UX) testing was conducted with Tribal community members, scholars in Indigenous Studies, and librarians. Feedback consistently highlighted the increased accessibility of search functions enabled by the implementation of controlled vocabularies. Users also emphasized that the project's defined metadata terms added a layer of decolonizing context that helped humanize and center stories of Indigenous people rather than provide a reductionist list of names. By adopting and extending controlled vocabularies through this collaborative approach, the project works to support more accurate, nuanced, and respectful representation of Indigenous communities.

3.1. Vocabulary Development Process

The vocabulary development workflow started with a “change” document. This document was used as a space to present the issue, outline factors to consider, explore possible solutions, and draft an implementation plan. This brainstorming approach allows for discussions and collaboration. Team members can either support the proposed solution or suggest other alternatives. The work focused on four descriptive fields, namely Age Category, Gender, Racial Descriptor, and Status. A “change” document was created for each where the discussions and recommendations were recorded. This included the data modeling strategy, whether that meant implementing an existing vocabulary, extending it to meet project needs, or creating a new one; data mapping workflow where controlled vocabulary values are mapped to unstructured values; definitions for newly created values to specify what each value means; and field names to ensure field labels clearly reflect the data they represent. In most cases, the project made use of the *Enslaved.org* vocabularies as a starting point while also ensuring new terms created provided the necessary balance and representation. These vocabularies are referred to throughout this article as the “referenced vocabulary.”

Defining these fields and assigning values to them proved deeply complicated, demanding years' worth of intense discussion within the project team. A central question was—and remains—*which reality* should the data represent?

1. The reality as recorded in the colonial and racially biased source documents?
2. The (inferred) lived experience of the individuals described in those documents?
3. Or a constructed reality designed to make the public web application useful, legible, and less harmful to our intended audiences?

Understanding that whichever realities we choose can never fully represent the lost or stolen lives of the people within the database, the project never settled on a single definitive answer. We chose instead to address each field on a case-by-case basis. The case studies below include descriptions of some of the complications encountered along the way.

Table 1. List of Age Category terms, definitions, and examples of language found in the source documents. Asterisks indicate the new terms added to the vocabulary.

Age Category Terms	Definitions	Language from Sources
Infant Age Group	Age category covering period from nursing to walking, or from birth up until but not including 2 years of age.	suckling childe, papoose
Child Age Group	Age category covering period from walking to puberty, or from 2 up until but not including 14 years of age.	12 or 14, a little Indian boy, boy, child, children, girl, little, young girl
Youth / Young Adult Age Group *	Age category covering from 14 up until but not including 25 years of age.	24 or 25, about 14, likely young, young, young adult, young man, young woman
Adult Age Group	Age category covering from 25 up until but not including 40 years of age.	30s, about 25 years of age, adult, between 35 and 40, man, near 40 years, woman
Older Person Age Group	Age category covering period from 40 up until 55 years of age.	about 40, abt 50 years of age
Elder Age Group *	Age category from 55 onwards.	anchant, elderly, of full age, old, old Indian woman, old wench, remarkably old

3.1.1. Age Category

Information found in the source documents of individuals identified as meeting the criteria for inclusion in the project did not always conform with modern age categorizations. Often this information was expressed using diminishing or offensive language (see the Language from Sources column on Table 1). The Age Category field was extended beyond its original referenced vocabulary to more accurately represent individuals' age ranges. This vocabulary extension involved the creation of two new terms (Youth/Young Adult Age Group and Elder Age Group) with corresponding definitions. Language drawn from the source materials was mapped to these terms and included in the documentation to better illustrate the curatorial decisions driving this work. Establishing this vocabulary ensures that individuals can still be appropriately classified and the database searchable even when historical records lack specificity. For instance, when descriptors such as "papoose," "remarkably old," or "young" were used.

The use of categories represents a compromise between qualitative and quantitative approaches to representing age. A fully qualitative approach, using the full text of the source, preserves the richness and ambiguity of the source materials, but is unwieldy in the context of a search filter. A quantitative approach, representing age as a number, works well for range filtering and analysis, but is often reductionist and requires unreasonable editorial assumptions. For instance, what numeric age would one assign to someone described as "remarkably old" or "little"?

This approach is not without complications. Defining a category such as "child" or "elder" requires setting explicit age boundaries—deliberate curatorial choices based on editorial judgment. This difficulty is compounded by historical and cultural context: what constituted childhood or old age differed meaningfully between Indigenous peoples and European settler societies, and shifted between historical and modern times, or even across the periods represented in the project.

3.1.2. Gender

The Gender field was refined to reflect community-informed terminology. The equivalent field in the referenced vocabulary uses Sex as its label. While this terminology reflects the content found in historical documents, the team decided to update the label to Gender to better align with descriptive practices reflecting the cultural dimensions of gender identity. In addition, an existing term from the referenced vocabulary was excluded and a new term added and defined (See Table 2).

Table 2. List of Gender terms and definitions. Asterisk indicates the new term added to the vocabulary.

Gender Terms	Definitions
Female	To indicate that the human subject is female.
Male	To indicate that the human subject is male.
Two-Spirit*	To indicate a distinct gender in Indigenous cultures.

The gender field presented its own surprising complications. Many project members felt strongly that the schema should not reify a binary conception of gender and argued for the inclusion of at least one non-binary category. An argument against including a third value was that Colonial source documents would likely never identify a non-binary individual explicitly. This counter-argument was eventually proved wrong, as a document surfaced that described a self-emancipated individual named Louis that “in his general appearance and manner resembles a woman as much as a man, constituting a singular and unique cast, which seldom fails to create a surprise, and to suggest the idea of a hermaphrodite.” [17].

Another complication was the tradeoff between historical certainty and search utility in the work of data entry. When a source implies gender without stating it (e.g., through a pronoun or a traditionally-gendered name), strict fidelity to the document may result in a value of “not recorded” when no gender indicators are present, which excludes the entry from gender-based searches. Inferring gender from context improves search utility, but this solution is also inherently imperfect as it may risk misgendering. In most cases, the primary documents are from colonial sources that more consistently used binary gender identities. While this does not fully justify the approach, assigning gender values can support high-level research, such as illustrating overall gender distributions within the project. The team acknowledges the risks and ethical implications of assigning gender based solely on given names and, instead, implemented a methodology where gender is assigned using both names and contextual information when available such as pronouns and titles (e.g., Mr., Mrs.). In these instances, a gender based on given name, pronouns, or title can be assumed, though this work always requires editorial judgment, awareness, and additional research particularly given the variability of gendered language across historical and cultural contexts.

As of writing, approximately 22% of individuals in the database are listed as *Female* and 66% as *Male*, with 0.25% identified as *Two-Spirit* and approximately 12% having no value assigned. Due to rounding, percentages may not sum to 100%.

3.1.3. *Racial Descriptor*

The *Racial Descriptor* field represents a new controlled vocabulary developed in order to facilitate the findability of individuals represented in the database. It is not meant to define anyone's identity but rather help represent individuals. Racial information is recorded verbatim in the referenced vocabulary through the *Race or Color* field. The *Racial Descriptor* vocabulary provides a normalized set of descriptive terms that is applied across the dataset.

Table 3. List of Racial Descriptor terms, definitions, and examples of language found in the source documents.

Racial Descriptor Terms and Definitions	Language from Sources
Black This term usually refers to a person having origins in any of the sub-Saharan groups of Africa, including African descended people of the Americas. We recognize that Black has many different present-day political and social meanings that go beyond this project's usages. This term was imposed as a category on Indigenous and multi-racial people to justify enslavability and subjugation and has also contributed to the erasure of Indigeneity in the archives. Indigenous This term usually refers to a person or group of people having origins in any of the original peoples of North and South America (including Central America and the Caribbean). We recognize that the term "Indigenous" has global resonances that are linked to people from specific places with knowledge of and respect for original ways. Due to the diverse nature of these groups, there is no single designation for all Indigenous peoples. Multi-Racial In the records of this project, this term usually refers to those who descend from more than one racial background, usually white and Indigenous or Black, or Indigenous and Black. These designations were highly subjective and inconsistent, and were almost exclusively imposed by colonizers for their own benefit. The derogatory terms that were often used for mixed-race Indigenous people often had the effect of challenging "authentic" Indigenous identity. In cases of Black-Indigenous multiracial terms, foregrounding their association with Blackness justified the enslavement and marginalization of these people. The end result was often the archival erasure of Indigeneity. White A person having primary family or ethnic origins in Europe. Like other racial designations, whiteness was often subjective and although it was claimed by Euro-descended peoples as an exclusive status of power, control, and privilege, it held legal weight and could be claimed by multiracial people as a basis for freedom. "White passing," whether claimed or imposed, was/is another strategy of erasure. In the records of this project, the vast majority of enslavers were white.	<i>Negro, Black, East India Negro; also skin color descriptions (dark-skinned, etc.)</i> <i>Indian, Native American, Native, American Indian</i> <i>Creole, Criollo, Dark Mulatto, Half Indian, Half Negro, Indian Mulatto, Mestiza, Mestizo, Mixed, Mulato, Multi, Mustee</i> <i>Spanish, English, French, Dutch, Creole ("creole" could refer to a white European who was born in the Americas, as well as to a multi-racial person of color); in many early records, "Christian" also signifies white Europeans</i>

Like other controlled vocabularies introduced by the project, this approach involved a tradeoff: the richness and complexity of the source materials was reduced in order to enable effective filtering in the interface. Perhaps more importantly, it also reduced the potential harm of offensive source terminology by keeping it out of the filter options, while still leaving it accessible for study through source transcriptions.

As of writing, the representation of each racial descriptor in the database is *Indigenous* 50%, *White* 29%, *Black* 4%, *Multi-Racial* 4%, with 17% having no racial designation because racial descriptors were not recorded in the source document. These are not mutually exclusive categories, which is why the sum of percentages exceeds 100%.

3.1.4. Status

The *Status* field was the most complex component of the vocabulary development process as it indicates a person's freedom status or connection to slavery. If *Stolen Relations* has an overarching scholarly argument, it is that Indigenous freedom was stolen in ways that often differed markedly from the theft of African freedom [18]. As such, the *Status* field formed the heart of that argument.

The initial process included drawing terms from the referenced vocabulary for the *Person Status* field where terms are represented as a flat list of values. This basic classification does not explicitly represent the hierarchical relationships among different categorical terms. To address this limitation, the team worked on extending the *Person Status* vocabulary to create a hierarchical taxonomy (see Figure 1). This model which in the early development process represented a four-level hierarchical structure was refined into a three-level one facilitating the query of entries at different levels while maintaining the specificity assigned to individuals. This three-level hierarchical structure includes two broad categories, *unfree person* and *free person*, which allow for classifications to be assigned more broadly when specificity in the source is lacking. With the understanding that a person's freedom status or connection to slavery could have changed at different points in their life, the *Status* field allows a single person to be associated with multiple terms. In addition, in this model, individuals associated with subcategories *formerly unfree* and *prospective unfree*, which are both under the *free*

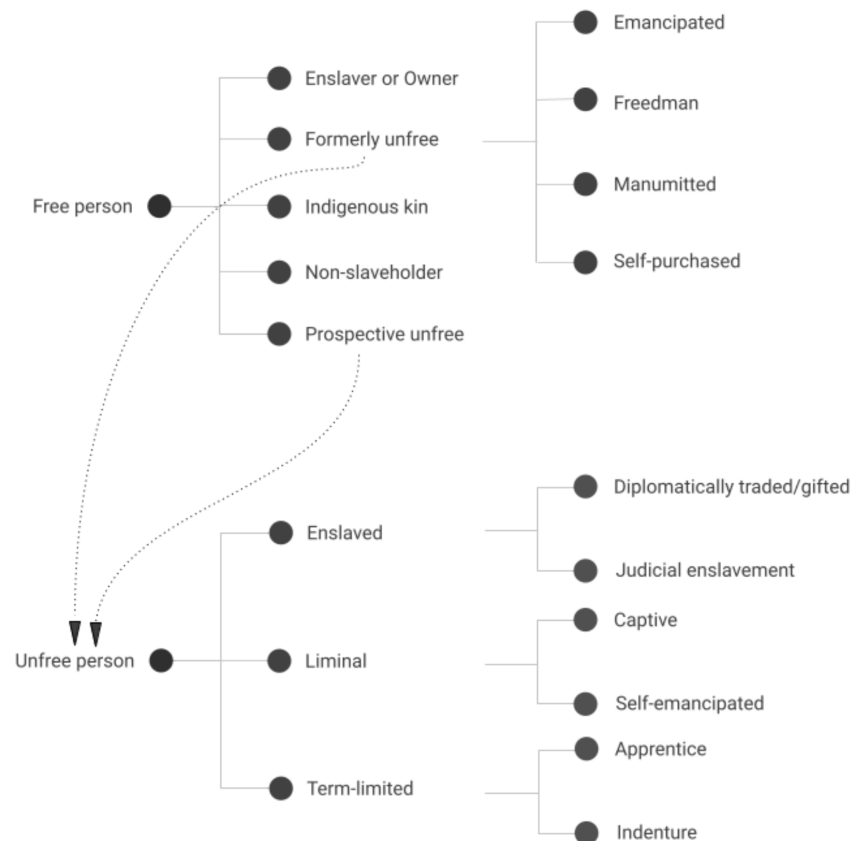


Figure 1. Status field hierarchical taxonomy.

person umbrella, also have a connection to *unfree person* (see Figure 1) and are labeled as such.

Contextual information was used to inform the development of the terms and their corresponding descriptions (see Appendix A). This was an iterative process requiring extensive internal discussion among the core project team as well as consultations with partners from Indigenous communities and scholars in the field. As of writing, the categorical term *unfree person* accounts for 54% of individuals in the database while *free person* represents 46%.

3.2. Representing Unnamed Individuals

The absence of personal names in historical records is another challenge that the team encountered. These omissions reflect common historical practices by colonizers and enslavers where individuals' names were excluded from the historical documents for reasons ranging from dismissal of Indigenous names to consolidations of enslaved people on lists to deliberate omission. A way in which the team addressed this issue was to introduce a controlled value to represent unnamed individuals. The adoption of this value aligns with contemporary museum practices aimed to restore identity and acknowledge archival loss [19]. A semantic change from *Unknown* to *Name once known* brings humanity into the database and affirms the presence of individuals even when attribution is lacking. This indicates that their names were not recorded which

reflects broader patterns of omission in colonial record keeping. But unlike an unknown name, “once known” reminds us that these individuals were known by their family and held ties to their communities. This practice has increasingly become the standard for curatorial exhibits within the last few decades, particularly within the field of Native American art and material culture [20]. As of writing, 21% of individuals in the *Stolen Relations* database are unnamed and therefore contain the value *Name once known*—their names are lost rather than unimportant and unrecoverable. Most importantly, it challenges us to think about memory, erasure, and the historical records; whose names are preserved and whose stories are left out; and how our current archival practices can leverage metadata as a tool to start that reparative work.

4. Conclusion

This paper underscores the role that controlled vocabularies can play in shaping how Indigenous histories are represented and interpreted. By framing metadata as a form of archival intervention, it highlights the responsibility that we have to challenge existing narratives. While the project keeps a verbatim transcription of the source document to ensure historical language is accessible for interpretation and continued scholarly research, working on extending, refining, or creating controlled vocabularies for these four fields has enabled the project team to more accurately represent the individuals found in historical documents. At the same time, adopting the designation *Name once known* represents a meaningful step toward affirming individuals' humanity. These efforts demonstrate the power of metadata to decolonize the archives and restore visibility to the stories of enslaved Indigenous people.

Appendix A. Appendix A

A.1. *Free person*

Person who possesses personal liberty, autonomy, and the ability to exercise their rights and make choices without coercion, constraint, or enslavement.

A.1.1. *Enslaver or Owner*

Person or organization who has some ownership or dominion over the services, body, labor, and reproductive rights of another person. Includes people who own slaves, servants, indentured servants, apprentices, as well as capturers, sellers/buyers, former enslavers, prospective enslavers, concessionaires, and overseers.

A.1.2. *Formerly unfree*

Person who was once unfree but now lives in legal freedom. All **formerly unfree** individuals should also have an additional status as an **unfree person**.

Emancipated

Person released from slavery through legislative or judicial action.

Freedman

Freedmen were individuals who had been enslaved by Native nations in Oklahoma (the so-called Five "Civilized" Tribes) and, after the Civil War, were given their freedom. The U.S. government required the Native nations who had enslaved them to give them citizenship in their nation. These formerly-enslaved people could be African-Americans or multi-racial, including Indigenous-Black individuals. The status of freedmen was contentious at the time, since many freedmen were not Indigenous at all, and has continued to be controversial up through the present day.

Manumitted

Person released from slavery through direct action of their enslaver..

Self-purchased

Person released from slavery by purchasing their own freedom.

A.1.3. Indigenous kin

A free person who is Indigenous, usually related or a member of the same community as an unfree Indigenous individual.

A.1.4. Non-slaveholder

A free person who does not hold slaves.

A.1.5. Prospective unfree

A free person who may experience unfreedom at a future point. All prospective unfree individuals should also have an additional status as an unfree person.

A.2. Unfree person

Person who is deprived of personal liberty and autonomy, is compelled to provide labor or service, or is subjected to varying degrees of control and exploitation by others.

A.2.1. Enslaved

Person whose body, labor, and reproductive rights are legally and indefinitely owned by another person or organization. Unless otherwise noted, enslaved persons are assumed to be enslaved for life. Enslaved individuals are treated as property, have no legal rights, and can be inherited, bought, and sold. The children of enslaved women are also presumed to be enslaved by law and custom.

Diplomatically traded/gifted

Person whose body, labor, and reproductive rights are traded between two parties as a means of diplomacy. Common in French and Iroquoian contexts.

Judicial enslavement

A person who has been convicted of a crime and sentenced to slavery or servitude through a judicial process.

A.2.2. *Liminal*

Person in an indeterminate, intermediate, or transitional status between servitude and freedom.

Captive

Person who was forcibly seized, captured, or taken prisoner by European colonizers during periods of colonial expansion and conflict. Captives were often subjected to various forms of exploitation, including forced labor, servitude, sale into slavery, or assimilation within the colonial settler communities, but may not yet be in a formal role of enslavement or servitude.

Self-emancipated

A person who has escaped bondage. Commonly known as runaways or fugitives, they neither live nor are legally recognized as free or freed. The term “self-emancipated” is a decolonial term that adds agency to the actions of the enslaved person.

A.2.3. *Term-limited*

Person whose labor and service are bound to a person or organization for a fixed term, sometimes under a contract, including as collateral against a debt. May be voluntary or involuntary.

Apprentice

Person whose term-limited unfreedom includes the acquisition of a skill or trade.

Indenture

Person who enters, usually but not always voluntarily, into an agreement to provide labor or service for a fixed period. Though sometimes voluntary, it still involves a limited form of unfreedom during the contractual period. A person may enter an indenture voluntarily or be forced into the arrangement through judicial processes or their parents (if they are a minor).

References

- [1] G. W. Miller, “Q & A with Linford D. Fisher.” [Online]. Available: <https://www.oceanstatestories.org/q-a-with-linford-d-fisher>
- [2] “Stolen Relations: Recovering Stories of Indigenous Enslavement in the Americas.” [Online]. Available: <https://stolenrelations.org/>
- [3] J. Kimball, “Collaborative project yields new digital archive on hidden history of Indigenous enslavement.” [Online]. Available: <https://www.brown.edu/news/2025-05-07/stolen-relations>
- [4] S. Berman, *Prejudices and Antipathies: A Tract on the LC Subject Heads Concerning People*. McFarland & Company, 1993.
- [5] W. M. Duff and V. Harris, “Stories and Names: Archival Description as Narrating Records and Constructing Meanings,” *Archival Science*, vol. 2, pp. 263–285, 2002, doi: [10.1007/BF02435625](https://doi.org/10.1007/BF02435625).
- [6] E. Drabinski, “Queering the Catalog: Queer Theory and the Politics of Correction,” *The Library Quarterly*, vol. 83, no. 2, pp. 94–111, 2013, doi: [10.1086/669547](https://doi.org/10.1086/669547).

- [7] U. o. U. J. Willard Marriott Library, "Marriott Reparative Metadata Assessment Tool (MaRMAT)." [Online]. Available: <https://www.marmatproject.org/>
- [8] D. Lee, "Indigenous Knowledge Organization: A Study of Concepts, Terminology, Structure and (Mostly) Indigenous Voices," *Partnership: The Canadian Journal of Library and Information Practice and Research*, vol. 6, no. 1, 2011, doi: [10.21083/partnership.v6i1.1427](https://doi.org/10.21083/partnership.v6i1.1427).
- [9] K. Christen and J. Anderson, "Toward Slow Archives," *Archival Science*, vol. 19, no. 2, pp. 87–116, 2019, doi: [10.1007/s10502-019-09307-x](https://doi.org/10.1007/s10502-019-09307-x).
- [10] Enslaved.org, "Peoples of the Historical Slave Trade." [Online]. Available: <http://enslaved.org/>
- [11] "Freedom on the Move." [Online]. Available: <https://freedomonthemove.org/>
- [12] "Digital Library on American Slavery." [Online]. Available: <https://dlas.uncg.edu/>
- [13] M. Miceli, J. Posada, and T. Yang, *Studying up machine learning data: Why talk about bias when we mean power?*. in Proceedings of the ACM on Human-Computer Interaction. 2022. doi: [10.48550/arXiv.2109.08131](https://doi.org/10.48550/arXiv.2109.08131).
- [14] M. K. Scheuerman, E. Denton, and A. Hanna, *Do datasets have politics? Disciplinary values in computer vision dataset development*, no. CSCW2. in Proceedings of the ACM on Human-Computer Interaction, no. 5. 2021. doi: [10.48550/arXiv.2108.04308](https://doi.org/10.48550/arXiv.2108.04308).
- [15] A. K. Heger, L. B. Marquis, M. Vorvoreanu, H. Wallach, and J. Wortman Vaughan, *Understanding machine learning practitioners' data documentation perceptions, needs, challenges, and desiderata*. in Proceedings of the ACM on Human-Computer Interaction. 2022. doi: [10.48550/arXiv.2206.02923](https://doi.org/10.48550/arXiv.2206.02923).
- [16] "Peoples of the Historical Slave Trade - Controlled Vocabularies," 2024. [Online]. Available: <https://docs.enslaved.org/controlledVocabulary/v4/enslavedcontrolledvocabularies-v4.pdf>
- [17] "One hundred dollars reward will be paid by Dr. Hayes, Bourbon street, New Orleans, for the apprehension of his two slaves, Willian and Louis, runaways, or fifty dollars for either," *The Daily Delta*, 1853. [Online]. Available: <https://www.newspapers.com/image/262146337/>
- [18] L. Fisher, *Stealing America: The Hidden Story of Indigenous Slavery in U.S. History*. Liveright, 2026.
- [19] M. Hartman, "Why do some labels say "Artist once known" at the Hood Museum of Art?." [Online]. Available: <https://sites.dartmouth.edu/meanwhileatthemuseum/2024/03/04/why-do-some-labels-say-artist-once-known-at-the-hood-museum-of-art/>
- [20] S. Chasse, K. Kramer, and L. Morgan, "Peabody Essex Museum's New Gallery of Native American and American Art." [Online]. Available: <https://www.incollect.com/articles/peabody-essex-museum-s-new-gallery-of-native-american-and-american-art>