

## USING BIG DATA TO VISUALIZE THE REPRESENTATION OF LATINE IDENTITIES

### **Project Justification**

#### **Introduction**

As Project Director (PD), Dr. LaTasha Velez, Assistant Professor in the University of North Carolina, Library and Information Science (LIS) Program, requests a \$282,201 Early Career Development Grant, with no cost-sharing, to investigate how Latine people – including but not limited to patrons, faculty, and library workers – are depicted in published, English-language LIS research. English is not the only language that produces valuable scholarship; however, this project is currently limited to English because English-language texts are most likely to be read by US academics and assigned in US LIS courses.

This exploratory research is grounded in critical Latinx Studies and critical discourse analysis (CDA). CDA relies on a thorough understanding of the societal context of the texts being analyzed, and Latinx Studies will be used to help provide that context. This exploratory research addresses the Laura Bush 21st Century Librarian Program Goal 2, Objective 2.3: Support the research of untenured tenure-track library and information science faculty, furthering the faculty member’s long-term research agenda, career trajectory, and professional development.

Requested funding will advance the research agenda of the Project Director by providing quantitative and qualitative data that examines the inclusivity of LIS research by responding to these key questions:

- R1. What are the conversations the LIS profession has about Hispanic/Latine people?
- R2. What patterns in LIS research will be revealed through bibliometric analysis?
- R3. What are some underlying biases or assumptions in the scholarship around Latine people?

The PD will lead a research team that includes two graduate assistants (GAs) and two project consultants (PCs). The GAs will be hired using grant funding. The PCs are Dr. Melissa Villa Nicholas, an Associate Professor at The University of Rhode Island who has done extensive research on Latine populations and Michelle Rosquillo, Research Manager at the Medicus Firm, who brings a much-needed practitioner perspective (see “consultants” section for more information). The study will result in the following outcomes: mapping of frequency trends in Latine research, bibliographic analysis, the development of an ontology around deficit vs asset-based conversations about Latine in LIS research, evidence-based recommendations to shift cultural assumptions in the field of LIS, and a methodological framework that will benefit future research and enable replication opportunities for researchers to conduct similar content analysis. The nature of this research project does not require a supplementary IRB process.

The term ‘*Latine*’ is used throughout the project description to describe people of Mexico, Central, Caribbean, and Latin America origins and heritage, using the ‘e’ to acknowledge that not all genders fall into the -o or -a category of male or female binaries. Many Latines do not identify with gendered terms but have not come to a consensus on one term in general. The emerging term *Latine* is a favorable non-gendered term that is more smoothly incorporated into Spanish speech and conjugation (Crochet, 2022). Sometimes, the word *Hispanic* will be used, as many other researchers use this term, and demographic information typically uses “Hispanic.” Some people make the distinction that *Hispanics* are from Spain or from Spanish-speaking countries in Latin America, while *Latine* are people from Latin America regardless of language. Other people use *Hispanic* and *Latine* interchangeably. Where other authors use the term *Hispanic* to denote those who are from Spain or Spanish-speaking countries, the PD will also use the term *Hispanic*.

#### **Relationship to Research Agenda**

Doctor Velez’s long-term research agenda explores equity, diversity, and inclusion (EDI) issues in LIS. Specifically, her work focuses on amplifying the needs and stories of Black, Indigenous, and People of Color (BIPOC), specifically those of Black and Latine people. Dr. Velez’s identity as an Afro-Latine, cisgender woman is integrally tied to her identity as a researcher and informs the topics and people that are researched. Being a member of this group is a strength that allows Dr. Velez to ask questions other researchers may not think to ask. Being an insider in a marginalized group offers insights to help understand the nuances of interpreted data.

Dr. Velez's research focuses on seeking to dismantle hegemonic narratives in LIS by usage of methodologies and theories that focus on words and language, such as collecting counter-narratives and using context-sensitive critical discourse historical analysis to identify the ways in which marginalized people tell their own stories and use words to show their agency in oppressive societies. Oliphant (2015) agrees that discourse analysis, including ontological analysis, is a useful tool for conducting social justice research. This project will benefit from and be significantly informed by Dr. Velez's previous experience with critical discourse analyses (CDA) and creating an ontology around [race and racism in LIS research](#).

Dr. Velez's prior research on Latine people has focused on the interruption of deficit narratives around Latine literacy, the purpose of which is to expand reader response theory by focusing on how racial and ethnic minorities relate to [diverse books](#). Specifically, Dr. Velez applied information gained from family interviews and auto-ethnographic information to explore how the Afro-Latine ethnicity and male gender of authors and characters related to how the readers in one Afro-Latine family interpreted a text. Her publications also include a CDA related to James E Shepard's discursive stratagems used to advocate for Black education in the Jim Crow [era South](#). Shepard was the president and founder of North Carolina Central University, currently the home of the EJ Josey papers and the only historically black college or university (HBCU) with a Master of Library Science (MLS) program. Dr. Velez also co-authored an article with Dr. Villa Nicholas, an identified consultant for this proposed project, which was clearly similar in concept and methodology by mapping the conversations about race and racism occurring in [LIS literature](#). This critical literature review benchmarked and identified gaps within the state of LIS discussions around race and racism.

The requested funding will further support, bolster, and broaden Dr. Velez's discursive work through an expanded analysis of the conversations in LIS research related to Latine people. It will also fuel the rigorous expansion of her cumulative methodological and theoretical knowledge by adding software-assisted bibliographic analysis and content mapping to her existing ontological methodology toolbox. This knowledge will further inform and strengthen future projects and enable the updating of previous work on discussions of race and racism in LIS research.

## **Statement of Need**

Librarianship's ethical codes, standards, and ideologies include EDI as goals and calls to action. The American Library Association (ALA) code of ethics explicitly states, "We work to recognize and dismantle systemic and individual biases; to confront inequity and oppression; to enhance diversity and inclusion; and to advance racial and social justice in our libraries, communities, profession, and associations through awareness, advocacy, education, collaboration, services, and allocation of resources and spaces" ([Committee on Professional Ethics, 2021](#)). Part of the mission of the ALA Office of Accreditation is to promote inclusion and diversity in LIS education ([Office for Accreditation, 2006](#)). Meanwhile, the Association for Library and Information Science Education (ALISE) ethical guidelines oppose discrimination ([Association for Library and Information Science Education, 2010](#)). Accordingly, some researchers have sought to include social justice in their work and explore the needs of marginalized communities ([Brook et al., 2015](#); [Colón-Aguirre & Ceja Alcalá, 2021](#); [Mehra et al., 2006](#); [Oliphant, 2015](#); [Velez, 2021](#); [Velez et al., 2022](#); [Velez & Villa-Nicholas, 2017](#)).

One such marginalized group is the Latine community. The Hispanic population is the second-largest ethnic group in the US and the fastest-growing minority ([Krogstad et al., n.d.](#)). The ALA released a report stating that Spanish is the number one language spoken by patrons of public libraries after English ([2008](#)). While Spanish is the most supported language in the library, corresponding services and programming for Spanish speakers are still of primary concern for many librarians ([ALA Office for Research and Statistics et al., 2008](#)). This is unsurprising because libraries must be valuable to their communities to continue to remain relevant. Libraries may be "marginalized" or even closed if they are not valuable to their communities ([Tyckoson, 2020](#)). As Tyckoson suggests, "Librarians must learn who comprises that community, what their information needs are, and how those needs are changing" ([2020, p. 4](#)).

More literature about and for Latines in LIS is a request spanning decades. Gilda Baeza wrote a history of library services available to Hispanics but was surprised to find a lack of materials that pre-dated 1970 ([1987](#)). Güereña and Erazo examined the timeline and history of services to Hispanics and found a dearth of literature on Hispanic librarians ([2000](#)). Guerena and Erazo explained that discussions of service to underrepresented groups has

been present in professional literature since the 1970s, but it is still an emerging concept that meets opposition in LIS (2000).

The professional organization REFORMA: The National Association to Promote Library & Information Services to Latinos and the Spanish Speaking 1997 report found that librarians of color ranked lower in the library hierarchy than their white counterparts. Like Guereña and Erazo the report noted, “There has also been much literature on library service to the Spanish speaking, but not much on Latina/o librarians themselves” (Gilton, 1997, p. 16). More recently, researchers have noted that the population growth of Latine people has necessitated additional research to better serve their Latine patrons (Colón-Aguirre & Ceja Alcalá, 2021; Velez et al., 2022). So, we know anecdotally that LIS literature around Latine peoples lacks a depth of scholarship; however, we do not have a full picture of the bibliometrics of the existing scholarship or what it has to say about Latine people. How conversations about Latine people are framed is important as authors also highlight the need to serve Latine people in empowering ways that do not produce barriers to access (Mitra et al., 2020). We do not completely understand if the existing literature that researchers, faculty, librarians, and students must draw from empowers Latines or perpetuates barriers.

This project builds on existing scholarship surrounding marginalized communities in LIS in general and Latine communities specifically. Understanding and serving this population is a recognized need in librarianship. Marginalized communities, such as Latine people, frequently face oppression (Card, 2019; Velez, 2021; Velez et al., 2022; Villa-Nicholas, 2020). LIS has long been acknowledged as a Euro-centric field that reinforces white supremacist values (Hathcock, 2015; Honma, 2005). Discursive ways of knowing through academic scholarship are one avenue in which the values of a field, be they empowering or oppressive, are reinforced (Hackney et al., 2018). LIS researchers engaged in social justice work seek to call out and end such oppression. To do so, we must interrogate the assumptions that underpin LIS. Before we can examine and interrogate the underlying discourses on LIS, they must first be made visible, which is what ontological research does. Hackney et al. argue that: “The literature of a discipline marks the key debates, ideas, concepts, and conceits of an academic field. As a sanctioned articulation of the field’s best work, peer-reviewed journal articles shape the discourse of LIS as an academic discipline” (Hackney et al., 2018, p. 11).

Unfortunately, a lack of comprehensive bibliometric data surrounding articles that engage with Latine people is available. Although some articles discuss conversations around marginalized populations in general, no definitive article in the LIS context specifically discusses the growing Latine population. Gaps in research about marginalized populations affect curricula as there are few articles for instructors to draw from and students to read. LIS students are completing their degrees without the necessary skills to work with diverse communities. LIS students also graduate without feeling prepared to work with members of different cultural communities (Patin et al., 2021). This project will help address this gap by deeply examining the articles that LIS produces. Providing such evidence can assist researchers by (a) benchmarking the current state of Latine literature in LIS, (b) bringing visibility to researchers in the past and currently who write about Latine people, (c) pointing to gaps in research, and (d) suggesting ways to reframe conversations from deficit-based conversations to asset-based conversations. As such, the target group for this project is other LIS researchers, LIS faculty, LIS students, and, ultimately, patrons. While we hope other researchers can and will build on this project to inform their own research, an integral result of this project is that LIS instructors will become more aware of Latine topics and that this subsequent research will provide information to disseminate with students through relevant curricula. The priority outcome is for instructors to gain additional, strengthened resources that will embed and infuse relevant research-based information about marginalized populations into LIS courses so that new professionals will seek to identify and address the needs of the array of underserved members of their own library communities.

## **Project Work Plan**

This project does not involve human subjects research and is therefore exempt from requiring IRB approval. This project seeks to understand the bibliometric and ontological aspects of LIS research about Latine people. It uses a grounded theory approach, which includes simultaneous data collection and analysis (Mansourian, 2006). This is mixed methods research using qualitative coding and quantitative bibliometric data to determine:

- R1. What are the conversations the LIS profession has about Hispanic/Latine people?

R2. What patterns in LIS research will be revealed through bibliometric analysis?

R3. What are some underlying biases or assumptions in the scholarship around Latine people?

This research will have three phases:

Phase 1: Corpus creation (Throughout year one)

Phase 2: Bibliometric analysis (Last 6 months of year one)

Phase 3: Ontological analysis (Throughout year two)

Throughout the duration of the project, the PD, PCs, and GAs will meet virtually twice a month to track progress, discuss results, and consult on articles and presentations. We also plan to meet in person once a year to facilitate to aid in meeting project timelines, enhance communication, and provide more of a mentoring and cohort experience for the GAs. The GAs will work 20 hours a week throughout the duration of the grant.

## **Methods**

### *Theoretical Framework*

The theoretical frameworks for this research are social justice (SJ) and Latinx studies. To borrow from Mehra, Albright, and Rioux's discussion of social justice, reasons for using SJ theories to ground this project include:

- Potential to contextualize LIS research within the lived experiences of marginalized people in ways that make a material difference in their socio-economic and socio-political experiences;
- Recognition of the potential for LIS to change people's lives; and
- Re-examination of LIS research within current social context ([Mehra et al., 2006, p.3](#))

Latin/o/x/e studies is also predicated on a tradition of social justice. Latinx studies is concerned with the lived experiences of Latine people. Latinx studies de-centralizes the experiences of the US-nation state and acknowledges existing hegemonic power structures within the United States ([Rúa & Ramos-Zayas, 2021, p. 6](#)).

Critical discourse analysis (CDA) is perfectly in line with an SJ theoretical framing. Echoing the words of Mehra et al., Fairclough states that CDA looks for a "better understanding of how societies work and produce both beneficial and detrimental effects, and of how the detrimental effects can be mitigated if not eliminated" ([2003, pp. 202–203](#)). Toward that end, this research adopts a constructivist approach to ontology. Constructivist ontology does not assume the existence of objective reality. Rather, constructivist ontology recognizes that meaning is subjective and constructed by the social realities in which meaning is interpreted and understood. This dovetails well with CDA. Discourse analysis looks at the role of language in identity formation, the construction of reality, and how language reproduces and perpetuates existing power structures ([Bourdieu & Thompson, 1991; Fairclough, 1989; Oliphant, 2015](#)). CDA analysis recognizes the power imbalances often inherent in language and uses methodologies that are meant to empower their objects of study, ensuring that recognition of social context is central to their analysis ([Reisigl & Wodak, 2005](#)). As such, CDA is clearly aligned with SJ, and Latinx Studies will provide the perfect context through which to interpret the language of LIS research.

I recognize that it is impossible to generalize the multiple identities and lived realities of Latine people. The intent of the ontology is not to generalize Latine people themselves but rather to uncover what assumptions are occurring in LIS research surrounding Latine identities. This ontology can also help uncover who is being discussed, i.e., Latine patrons, Latine librarians and staff, etc., and where there may be gaps in research coverage. To create the ontology, the PD will implement a modified CDA triangulation approach suggested by [Reisigl and Wodak \(2005\)](#):

- The intratextual language of LIS Latine research texts
- The intertextual language of the LIS Latine research corpus and the relationship between discourses
- The institutional frames of LIS as a profession, i.e., values, policies, ideologies, etc.
- The broader sociopolitical and historical context of Latine people in the US within which LIS discourses are situated ([p. 41](#)).

Particularly relevant to this project's social justice and CDA framing, the ontology uses the concept of asset vs deficit narratives to examine discussion around Latine people. The deficit model of research frames subjects as lacking some skill or knowledge as opposed to asset-based research that focuses on people's strengths or recognizing when deficits are based on structural inequality rather than a personal failing ([Glenn et al., 2018](#); [Martinez-Cola, 2020](#)). Leung and López-McKnight note that education researchers have explored race and racism's effect on educational inequities and have sought to "disrupt the dominant conversation the places pathologies and deficits on and within [marginalized] communities" ([2021, p. 11](#)). Moving into the realm of librarianship, they discuss the harm caused when librarians hold deficit perspectives toward Black, Indigenous, and people of color (BIPOC) youth ([Leung & López-McKnight, 2021, p. 20](#)). It is important to understand whether the field is causing harm by perpetuating deficit conversations around Latine people. As stated above, authors publishing about Latine research are concerned with ensuring that texts are empowering and do not reproduce deficit framings.

One approach to operationalizing definitions of asset vs deficit conversations is through the language of micro-aggressions (deficit) vs micro-affirmations (asset). Microaggressions are endemic humiliations, be they verbal, visual, or nonverbal, that communicate discrimination or insults to marginalized people. Sue et al. created a typology of microaggressions, including microassaults, microinvalidations, and microinsults ([Sue et al., 2007](#)). Microaffirmations disrupt oppressive and marginalizing narratives by offering empathy and community. Building on the work of others, Rolón-Dow and Davison offer a typology of microaffirmations that include microrecognitions, microvalidations, microtransformations, and microprotections ([Rolón-Dow & Davison, 2021](#)).

Researchers in fields including education, psychology, and LIS have operationalized microaggressions in their own research ([Alabi, 2015b, 2015a](#); [Rolón-Dow, 2019](#); [Solórzano & Pérez Huber, 2020](#); [Williams, 2020](#)). While microaffirmations are understudied, some education and psychology researchers have used microaffirmations as an analysis framework ([Koch et al., 2022](#); [Rolón-Dow, 2019](#); [Rolón-Dow & Davison, 2021](#); [Soriano, 2022](#)). Similar to the aims of this project, one such application of microaffirmation research, authored by Rolón-Dow and Davison, is also ontological research grounded in Latinx studies ([Rolón-Dow & Davison, 2021](#)).

### *Preliminary research*

Dr. Velez has prior experience writing about Latine populations, including a paper that won the 2022 Ethnic Multicultural Information Exchange Roundtable (EMIERT) of the American Library Association David Cohen Multicultural Award, which recognizes research that promotes multiculturalism in US libraries. She has also created similar ontological work mapping the conversations LIS had about race and racism. Her dissertation used CDA as its methodological framework.

This research is iterative as, prior to phase 1, Dr. Velez did preliminary research into English-language LIS articles about Latine people. Those preliminary results were presented at the ALISE conference. During that preliminary research, she focused on searching the author-supplied and controlled vocabulary subject terms of articles in the Library and Information Science Abstracts with full text database (LISTA). These preliminary results netted 106 out of 384,760 articles, including duplicate articles that had different subject terms. After removing duplicates, Dr. Velez ended up with 42 articles. Twenty of those 42 articles were in medical information journals. Removing gap words, nineteen articles included at least one term with negative connotations specific to Latine identities:

• Barrier	• Depression	• School dropout prevention
• Digital divide	• Disparities	• Sexually transmitted diseases
• Folk devils	• Hate crimes	• Problem-solving
• Illegal	• Low-income	• Moral panic

These observations do not imply that the discourse in the article as a whole leaned toward deficit narratives. However, the examination highlights potential biases in LIS research that warrant more in-depth study. Evaluating the preliminary research also revealed limitations Dr. Velez would correct in this project. Recognizing that the number of articles located was low, even for an under-studied topic, Dr. Velez found that some influential journals, such as *The Journal of Academic Librarianship* and *Library Management*, were not available through her



institution’s LISTA subscription but were included in other databases, so Dr. Velez plans to modify her data collection methodology, as shown below in “corpus creation”, to counteract this limitation.

### Consultants

The proposed project consultants, who demonstrate both lived and academic experience and expertise, will assist in assuring the quality, validity, and fidelity of this project’s overarching and specific components. Because meanings and discourses are contextual and socially constructed, it is imperative that this research is not completed in a vacuum, but instead informed by other professionals and stakeholders. The consultants will assist with visualizations, provide feedback on the themes identified through CDA, and provide informed and expert input on the resultant ontology. Bi-monthly online meetings and once-yearly in-person meetings will be conducted to discuss and resolve any challenges as they arise. The consultants will be included on the project website and on conference panels promoting the research. The team will be compensated for their time and can apply the ontological methodology and framework to advance their own research agendas and professional practice.

#### *Proposed consultants:*

Michelle Rosquillo is a research manager at The Medicus Firm. Ms. Rosquillo serves as the Diversity Chair for the Carolinas chapter of the Special Library Association, and won the 2022 Horizon Award for distinction in promoting the information profession. A naturalized American citizen of Filipino descent and member of the LGBTQIA+ community, Michelle has long experienced the dissonance that her ethnic heritage and sexual identity are framed as inferior to the American ideal and yet valuable enough to be traded for various academic and career opportunities. Ms Rosquillo aims to explore language neutrality and affective response, including asset- and deficit-based frameworks built for and around, but not by, underrepresented communities, particularly how these frameworks influence policy creation and implementation.

Dr. Melissa Villa-Nicholas is an associate professor in the Graduate School of Library and Information Studies at the University of Rhode Island. Dr. Villa-Nicholas’ work focuses on the information and technology histories and practices of Latines, immigrant information rights, and critical approaches to information science. Dr. Villa Nicholas has received numerous awards, including the Diversity and Inclusive Excellence Awards at the University of Rhode Island in 2021. Dr. Villa-Nicholas founded the Information Equity, Diversity Communities, and Critical Librarianship track for the Master of Library and Information Studies (MLIS) at the Graduate School of Library and Information Studies at the University of Rhode Island.

#### *Phase 1*

##### *Corpus creation, Year 1, Months 1-12*

Using the list of identity terms shown in the table below paired with the LIS terms listed later, the PD will conduct preliminary searches and review resultant articles to narrow down search strings that produce the most relevant results. The PD is proposing to hire two Graduate Assistants (GAs) with the requested grant funds. The GAs will each work 20 hours a week throughout the duration of this grant, including over the summer. During phase 1, the PD will provide the GAs with the narrowed down lists of search strings to search the UNCG Libraries’ unified catalog, which provides results through WorldCat for English language articles. Our preliminary list of identity terms includes:

Subject Heading	UF/BT/RT	Use
Afro-Hispanics		
Afro-Latinos	Afro-Latinas	
Argentinian Americans	Argentiniens	Argentinos
Belizean Americans	Belizeans	Belizeanos
Bolivian Americans	Bolivians	Bolivianos
Boricuas		Puerto Ricans
Brazilian Americans	Brazilians	Brasileños

Chicano Americans	Chicanos; Chicanas; Chicax	
Chilean Americans	Chileans	Chilenos
Colombian Americans	Colombians	Colombianos
Costa Rican Americans	Costa Ricans	Costarricences
Costarisences		Costarricences
Cuban Americans	Cubans	Cubanos
Dominican Americans	Dominicans	Dominicanos
Ecuadorian Americans	Ecuadorians	Ecuadorianos
Guatemalan Americans	Guatemalans	Guatemaltecos
Hispanic Americans	Hispanics	
Hispanos		Mexican Americans
Honduran Americans	Hondurans	Hondureños
Latine		
Latino Americans	Latin Americans; Latinos; Latinas; Latinx	
Mejicano	Mexicano	
Mestizaje		
Mexican Americans	Mexicans	
Nicaraguan Americans	Nicaraguans	Nicaragüenses
Panamanean Americans	Panamaneans	Panameños
Paraguayan Americans	Paraguayans	Paraguayos
Peruvian Americans	Peruvians	Peruanos
Puerto Rican Americans	Puerto Ricans	Puertorriqueños
Salvadoran Americans	Salvadorans	Salvadoreños
Spanish Americans	Spaniards	Españoles
Tejano	Texano	
Uruguayan Americans	Uruguayans	Uruguayos
Venezuelan Americans	Venezuelans	Venezolanos

The identity terms were derived from:

Resources for Thesauri/Vocabulary	Links
Chicano Thesaurus	<a href="https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Feslibrary.berkeley.edu%2Fsites%2Fdefault%2Ffiles%2Fthesauru.doc&amp;wdOrigin=BROWSELINK">https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Feslibrary.berkeley.edu%2Fsites%2Fdefault%2Ffiles%2Fthesauru.doc&amp;wdOrigin=BROWSELINK</a>
ANSSWeb	<a href="https://acrl.ala.org/anss/index.php/subject-headings-for-latin-hispanic-and-spanish-americans/">https://acrl.ala.org/anss/index.php/subject-headings-for-latin-hispanic-and-spanish-americans/</a>
Keywords	<a href="https://keywords.nyupress.org/latina-latino-studies/about-this-site/">https://keywords.nyupress.org/latina-latino-studies/about-this-site/</a>
DEI Controlled Vocabs Resource List	<a href="https://docs.google.com/spreadsheets/d/19solOX6tQTYvIF4lr_JNz2WlcsA76CcK3bxvYZ8cHzg/edit#gid=0">https://docs.google.com/spreadsheets/d/19solOX6tQTYvIF4lr_JNz2WlcsA76CcK3bxvYZ8cHzg/edit#gid=0</a>
Library of Congress Demographic Group Terms	<a href="https://www.loc.gov/aba/publications/FreeLCDGT/DEMOGRAPHIC.pdf">https://www.loc.gov/aba/publications/FreeLCDGT/DEMOGRAPHIC.pdf</a>

LIS terms were taken from the work of Hackney et al. (2018) with additions and edits from the PD. LIS terms include (but are not limited to):

• Librar*	• Archiv*	• Information science
• LIS	• Library study	• Information studies
• Library education	• Library science	

To ensure the internal validity of the dataset, each GA will be given the same list of search strings. The GAs will add each article that matches the search criteria to a private Zotero library. Zotero is a free citation management software that simplifies collecting, annotating, sharing, and citing research. Zotero also allows data to be exported in file formats, like csv and ris, that allow for data analysis and uploading to data analysis software. The PD will compare the resultant search results to ensure they are identical. Once the PD and GAs have determined that all available data has been collected and the data sets match, the PD will dedup the Zotero library. The dedupped data set in Zotero will be divided among the PD, PCs, and GAs. Each person will review their assigned articles' titles, subject terms, and abstracts to remove false positives. For example, we will look for and remove articles about the classical Latin language versus Latin as an identity marker. We will also remove articles that mention Latine only when speaking about demographic classifications of survey participants or area residents but have no other mention or discussion of Latine people. The cleaned Zotero library is the basis for the final dataset for the project. As there are likely to be many search strings and the number of articles could potentially be over 1000, the PD anticipates the data collection and cleaning will last for the first 6 months of this project. This initial phase of corpus creation will be sufficient to begin the bibliometric analysis and article writing occurring during phase 2. During the last two months of year one, the GAs will continue updating the corpus by adding full-text .pdf files of each article remaining in the Zotero library to a Google folder for use in phase 3. The GAs will also note articles where UNCG Libraries do not provide full-text access and submit interlibrary loan (ILL) requests for those articles. Once ILL requests are received, the GAs will also upload those articles to Google to complete the corpus creation.

#### *Phase 2: Bibliometric analysis, Year 1, Months 7-12*

The PD will experiment with importing the Zotero dataset as an ris file to an open-source bibliometric network analysis tool, [VOSviewer](#), to create visualizations to present bibliometric information. If the ris upload does not produce the necessary results, the PD will recreate the dataset through SCOPUS and upload a SCOPUS file to VOSviewer. VOSviewer accepts multiple formats, including ris, SCOPUS data, and Web of Science information. While this is occurring, the GAs will create frequency tables including but not limited to the journals that publish about Latines, the number of articles published per year, and the percentage of Latine articles published in high-impact journals relative to total articles published, etc. Analysis of the bibliometric and frequency information will form the basis for the first peer-reviewed article.

While the PD and PCs are engaged in the bibliometric analysis phase, the GAs, supervised by the PD, will identify and collect relevant articles for the literature review of the bibliometric article. During the final months of year 1, the PD will write the first planned article with input from the PCs and GAs. The first article will use the visualizations and tables created by the PD and GAs and the analysis of these visualizations to examine frequency and bibliographic trends in patterns of study, such as the years and frequency of study, journals that publish studies about all categories of Latine people (patrons, librarians, students, etc.), the number of articles, and citation ranges as well as bibliometric trends such as citation, bibliographic coupling, co-citation, or co-authorship relations.

After the GAs have aided in creating tables and locating articles for a literature review, they will return to completing the corpus by adding full-text .pdf files to a Google folder, as mentioned above in phase 1.

#### *Phase 3: Ontology, Year 2, Months 1-12*

Once the GAs have completed uploading the full text, they will spend the first months of year 2 collecting relevant articles for the ontological paper literature review. This grant funded research project expands on an existing Latine ontology created by Rolón-Dow and Davison (2021) and operationalizes those theories using the concepts of microaggressions vs microaffirmations. To that end, while the GAs are collecting literature review articles, the PCs and the PD will operationalize how they will apply the asset/deficit framework (see theoretical framework section above) to each article. They will also operationalize how to categorize articles by library type/subject. Articles will be divided using a similar framework to Velez and Villa-Nicholas' "[Mapping race and racism in US library history literature, 1997-2015.](#)" We will use categories such as Biographies, Institutions, Library Education, Library Associations, Historical Studies, and Print Culture. We will then subdivide categories (where appropriate). For example, Institutions may be subdivided by library type (academic, public, private,



special), or we may subdivide even further by the subject of the research article (i.e., Latine patron, Latine librarian, etc.). The PD will document the resultant methodology and store that document as a private Google document that the research team can consult during the critical discourse analysis.

The PD will check for internal validity by borrowing a methodology previously employed to study the themes present in LIS diversity research (Hackney et al., 2018). So, once the asset/deficit and categorization methods are agreed upon, the PD will choose 20 articles at random and ask each team member to code these articles to check for internal agreement. Once potential validity concerns are surfaced and addressed, The PD will divide the remaining full-text articles among herself, the GAs, and the PCs. We will use our methodological document to conduct the critical discourse analysis of each document to explore the social justice implications of the research. We will return to our methodological documentation to categorize articles by library type/subject. Each of these categorizations will be done in individual documents but shared in Google documents. This will result in a completed ontological map of asset versus deficit-based narratives in LIS research subdivided by into categories as noted above.

In consultation with the GAs and PCs, the PD will spend the final months of the grant writing the ontological article to uncover trends in LIS research conversations surrounding Latine identities, inclusive of all subjects we find (patrons, librarians, library staff, researchers, faculty). During the second year, the PD will also identify future article topics based on the collected data.

In both years, the PD, PCs, and GAs will submit to conferences such as the Reforma National Conference, iConference, and the ACRL, ALA, JCLC, and ALISE conferences. In later years, the PD hopes to use the methodology employed in this grant project to expand her research in multiple directions, including looking at other minority communities and applying our data approach to research conversations surrounding marginalized populations in other STEM fields.

#### *Dissemination of Findings:*

The PD will share findings and recommendations directly through peer-reviewed research articles, presentations, and a website. Presentation venues include the REFORMA, JCLC, ALISE, and ALA conferences. One of the grant supports offered by UNCG SOE is the creation and hosting of a website. This website will also be used to disseminate findings. We will advertise our website on various locations, such as the JESSE listserv, other listservs identified through the REFORMA resources and publications listserv page, the REFORMA newsletter, and other avenues the PD and PCs find. The website will include visualizations that come from the project and presentation .pdfs. It will also include a bibliography of the articles in our data set to amplify the voices of researchers doing work surrounding Latine communities. Two peer-reviewed articles are currently planned. One article will benchmark the bibliometric data that results from the study. The second will share the asset/deficit ontology.

### **Diversity Plan**

This project will strengthen librarianship's commitment to EDI by interrogating underlying assumptions that may be present in the research we publish, disseminate to LIS workers, and teach in classrooms. Understanding the patterns of how Latines are discussed may reveal biases such as racism, sexism, gender discrimination, and classism. Before we can change deficit-based conversations, we must know if they are occurring. This recognition will help the field change this structural discrimination by shifting the narrative around *who* Latines are in the library and beyond. Such honest evaluations of invisible biases can help LIS professionals work toward a climate of cultural humility, making the profession a more welcoming and inclusive space for marginalized populations such as Latine people. The PD and both consultants are members of marginalized groups and have done EDI or Latinx Studies research in the past, thus bringing valuable research and lived experience to this project. Finally, as the PD and consultants are members of marginalized groups, this project will advance the careers of underrepresented minorities in the LIS field.

## **Project Results**

The study will result in the following outcomes: the development of an ontology around asset versus deficit-based conversations about Latine people in LIS articles, bibliometric data regarding articles discussing Latine people, a bibliography of articles written about Latine people that others can draw from, evidence-based recommendations to shift cultural assumptions in the field of LIS, and a methodological framework that will benefit future research and enable replication opportunities for researchers to conduct similar content analysis.

This project will also result in the mentoring and professional development of the GAs. The PD has previously published articles with several students and maintains professional relationships with past GAs, including her former GA, Michelle Rosquillo, who is currently a consultant on this project. She is committed to creating lasting mentoring relationships with the GAs working on this grant and introducing them to other library professionals. The GAs will gain research experience, conference presentation experience, and author credit on resultant publications.

Results can also aid in potential future research avenues for the PD, PCs, and GAs, such as using the open-source software [MALLET](#) to create topic models from the initial corpus with the potential to further explore the themes in LIS research about Latine people beyond the asset/deficit ontology and bibliometrics. It will advance PD goals by enabling and supporting expanded research into marginalized communities. The PD will apply the framework from this research to study other marginalized populations and to continue to update ontologies surrounding those communities in the coming years. The requested funding will provide the PD with tools to progress and sustain her overarching research agenda for many years.

Project results will also be immediately applicable to the PD and Dr. Villa-Nicholas' teaching and will be integrated into our respective courses.

Ultimately, a better understanding of marginalized populations will spread to librarians through LIS curricula. This project will fill a gap in LIS research by providing bibliographic and ontological data about the Latine population. It will also uncover gaps in current research that others may wish to explore. Other researchers will be able to use the data gathered during this grant project to provide additional empirical support for their own projects.

This could have wide-ranging impacts beyond researchers and into LIS classrooms, offering faculty data on the growing population of Latine people that they can share with their students. These future librarians and information professionals will leave classrooms better prepared to serve the growing Latine population.

**Schedule of Completion**

<b>Tasks Year 1</b>	Aug 24	Sept 24	Oct 24	Nov 24	Dec 24	Jan 25	Feb 25	Mar 25	Apr 25	May 25	Jun 25	July 25
Project director (PD) will conduct preliminary searches to narrow search strings to those most likely to produce relevant results												
Graduate assistants (GAs) will collect data using search strings identified by the PD and add article citations and abstracts to Zotero												
PD, consultants, and GAs manually check Zotero data for relevancy, duplicates, false positives												
The PD will experiment with importing the Zotero dataset as an ris file to VOSviewer, to create visualizations to present bibliometric information. If the ris upload does not produce the necessary results, the PD will recreate the dataset through SCOPUS and upload a SCOPUS file to VOSviewer. to create bibliometric visualizations												
GAs create frequency tables												
PD and project consultants (PCs) analyze patterns in bibliometric visualizations												
PD and PCs conduct frequency analysis												
GAs identify and collect relevant articles for bibliometric article literature review												
PD writes bibliometric article with input from PCs												
GAs search for and add full texts of articles and add them Google folder to critical discourse analysis corpus												
GAs submit interlibrary loan requests for missing articles												

<b>Tasks Year 2</b>	Aug 25	Sept 25	Oct 24	Nov 25	Dec 25	Jan 26	Feb 26	Mar 26	Apr 26	May 26	Jun 26	July 26
GAs identify and collect relevant articles for ontological article literature review												
PD and PCs operationalize how articles will be labeled asset/deficit												
PD and PCs operationalize how to categorize articles by library type/subject												
PD chooses 20 articles at random for everyone in the team to catagorize, checking for internal validity												
PD, consultants, and GAs address potential validity concerns												
PD, PCs, and GAs use operationalized processes to label articles asset or deficit												
PD, PCs, and GAs use operationalized processes to categorize articles by library type/subject categorization												
PD writes ontological article with input from PCs												

### **Digital products plan**

This project will create visualizations, tables, methodological process documents, conference presentations, and publications. The visualizations and tables will show results related to critical discourse analysis, frequency analysis, and bibliographic analysis. These items and our data and methodology will be made available via a grant project website hosted by The University of North Carolina at Greensboro School of Education. Visualizations will be available as JPEGs, and PNGs. Tables will be available as PDFs. Publications and articles will PDFs. These will be final deliverables, not editable files. Visualizations will be made with open-source software such as VOSviewer, a tool for visualizing bibliometric networks.

The University of North Carolina at Greensboro School of Education supports grant projects by creating and hosting websites. Data will be made available through the website. Although every attempt will be made to publish articles in open-access journals, when this is not possible, raw visualizations, preprints, presentations, and other data will be available on the website.

The project director's goal is to make this research accessible. The methodology is open to use, with attribution, as a model for other projects of this nature. Visualizations, analysis, and results should be used with attribution.

We will preserve and maintain digital products on the cloud (Google and Zotero) but with a local backup to the PD's computer to prevent it from disappearing. Data will also be preserved on the SOE website. Data on the cloud and the website will be preserved indefinitely. While the cloud data will be private, the website data will be publicly accessible.



## Data Management Plan

*Using Big Data to Visualize the Representation of Latine Identities* will collect and share bibliographic data and full text of publications about Latine people within our research team. Data may include but is not limited to a spreadsheet of citation data, lists of search strings, and documentation explaining how operationalization of the asset/deficit framework for the PD and PCs to use when analyzing research articles.

Most of the raw data will be collected in the form of citations and full text of LIS article about Latine people. This data will be collected from web searches through the UNCG Jackson Libraries catalog and interlibrary loan. No sensitive data will be collected during this project. However, there are copyright issues with sharing full texts of articles, so full text will only be shared within the project team. To facilitate this sharing, the project director (PD) will use the Google suite and Zotero to share data between graduate assistants (GAs) and project consultants (PCs). Zotero is a free citation management software that simplifies collecting, annotating, sharing, and citing research. Zotero also allows data to be exported in file formats, like csv and ris, that allow for data analysis and uploading to data analysis software.

The PD will disseminate findings to the broader library community through a website created by the University of North Carolina at Greensboro (UNCG) School of Education (SOE). The PD, PCs, and GAs will also share findings and methodologies through presentations and research papers. The methodologies for all articles published will be thorough and detailed so that other researchers may replicate the methods in their own research.

In Year 1, the GAs will collect citation and abstract data through catalog searches. The list of search strings the GAs will use will be uploaded to a Google folder. This resultant data from catalog searches will be uploaded to a Zotero library shared with the GAs and PCs. Once the Zotero data is cleaned, the PD will publicize the collected citation information on the SOE website to amplify Latine research in LIS. This may be done via csv format to make it easily accessible to other researchers. At the end of year one the PD intends to write a bibliometric article. The bibliometric visualizations will be uploaded to a private Google folder created by the PD and shared with the PCs and GAs. Where acceptable by publishers, pre-prints will be made available on the website. Slide decks from project presentations, including analysis and visualizations, will also be available on the website.

In Year 2, critical discourse analysis will occur to create an ontology surrounding conversations in LIS research about Latine people. The graduates will save the .pdf full text files of articles to a Google folder, which is only accessible to the project team to avoid copyright infringement. Documentation about how to apply the asset/deficit framework and content analysis visualizations will be uploaded to a private Google folder. The PD will also write and publish pre-prints (where acceptable by publishers) of the ontological article on the website. Slide decks from project presentations, including ontological information and visualizations, will also be available on the website.