

Enhancing Access and Research Possibilities through Critical Engagement with Historical Data

Award PW-285172-22

6/1/2022 - 4/2/2025

Jamie Rogers - ORCID: 0000-0002-4849-114X

Rhia Rae - ORCID: 0000-0002-9154-8033

Project Team

Jamie Rogers, Assistant Dean for Strategic Initiatives (Project Director)
Rhia Rae, Digital Assets Project Manager, Special Collections and University Archives (Co-Project Director)
Rebecca Bakker, Digital Collections Librarian (Key Personnel)
Molly Castro, Digital Humanities Librarian Consultant (Key Personnel)
Jill Krefft, Associate Dean, University of Tulsa Library (Key Personnel)
Dr. Diana Ter-Ghazaryan, Associate Professor of Spatial Sciences, University of Southern California (Key Personnel)
Luis Garcia Falcon, GIS Specialist
Sonia Santana Arroyo, Research and Data Repository Manager (Key Personnel)

Interns

Juan Lopez, Graduate Student, Department of History, Florida International University
Catherine Tibaquirá, Undergraduate Student, Department of Art + Art History, Florida International University
Rosemary Toval, Undergraduate Student, Department of History, Florida International University

Consultant

Dr. N. D. B. Connolly, Herbert Baxter Associate Professor of History, Johns Hopkins University

Advisory Board

Dr. Robin Bachin, Senior Associate Dean for Undergraduate Education, Director, Office of Civic and Community Engagement, Charlton W. Tebeau Associate Professor of History, University of Miami
Dr. Julio Capó, Jr., Associate Professor, Department of History, Deputy Director, Wolfsonian Public Humanities Lab, Florida International University
Dr. Marvin Dunn, Historian and Professor Emeritus, Florida International University
Dr. Sharony Green, Associate Professor, Department of History, University of Alabama
Dr. Kenneth Lipartito, Professor, Department of History, Florida International University
Dr. Chanelle N. Rose, Associate Professor, Department of History, Coordinator, Africana Studies Program, Rowan University
Dr. Daniel Royles, Assistant Professor, Department of History, Florida International University
Dr. Melanie Shell-Weiss, Chair and Associate Professor, IRIS: Integrative, Religious, and Intercultural Studies Department, Grand Valley State University
Althea Silvera, Head of Special Collections, Florida International University

Table of Contents

Summary	3
Project Background	4
Purpose.....	4
History.....	4
The Collection.....	5
Project Planning	7
Data Collection, Curation, & Access Methodologies	9
Unstructured Data.....	9
Structured Data.....	10
Geospatial Data.....	13
Project Outcomes and Impact	16
Deliverables.....	16
Developing a Community of Practice.....	18
Consultant Report.....	20
Future Work	22
Acknowledgements	23
Bibliography	24

Summary

While advances in big data and machine learning have generated considerable discourse on optimizing business processes and research capabilities, significantly less attention has been paid to ensuring that data practices are transparent, fair, and accountable. Libraries and archives, with their expertise in organizing, describing, providing access to, and stewarding information, are uniquely positioned to advance ethical data frameworks.¹

The initiative *Enhancing Access and Research Possibilities through Critical Engagement with Historical Data*, led by Florida International University and funded by the National Endowment for the Humanities, serves as a case study in developing an ethical, care-based community of practice around historical data. Based on the papers of Dana A. Dorsey—widely regarded as Miami's first Black millionaire—the project explores the interpersonal networks and investments within Miami's Black community during the pre-redlining era. The creation of openly accessible data resources aims to enhance understanding of the contributions of Black civic groups, trade groups, and community alliances to the economic and social infrastructure of the city.

By developing principles of intentional language, transparency, continuous stakeholder input, and iteration, the project seeks to address gaps in the historical record and model a replicable approach in data curation. In honoring the lives and stories embedded in archival documents, the initiative emphasizes the importance of data practices that are not only rigorous but also restorative.

Key deliverables include full-text transcripts, curated tabular datasets, georectified historical maps, and detailed property location data. These resources reveal patterns in early 20th-century legal and social norms, highlighting the roles of influential Black figures in shaping their communities. Enhanced access to location data allowed the project team to trace how urban development, particularly highway construction, led to the disruption and dismantling of established Black neighborhoods like Overtown. This work underscores the importance of ethical data practices and aims to provide broad access to these resources, capturing the full complexity of the region's history.

The [project website](#) offers comprehensive information about the initiative, including its objectives, team members, documentation, deliverables, media coverage, and a feedback portal. This digital presence ensures that the resources are not only widely accessible but also shaped by ongoing dialogue.

¹ Eun Seo Jo and Timnit Gebru, "Lessons from Archives: Strategies for Collecting Sociocultural Data in Machine Learning," in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, FAT* '20 (New York, NY, USA: Association for Computing Machinery, 2020), 306–16, <https://doi.org/10.1145/3351095.3372829>.

Project Background

Purpose

This report examines the development of an ethical, care-based community of practice for data collection and curation, through insights and outcomes of Florida International University's National Endowment for the Humanities-funded project, *Enhancing Access and Research Possibilities through Critical Engagement with Historical Data*. The resulting practice was also informed by the ongoing work of community archives, librarians, archivists, historians, and digital humanists committed to social justice.

Focusing on a collection of open data resources derived from the papers of Dana A. Dorsey (1868 – 1940), recognized as Miami's first Black Millionaire, the project includes warranty deeds, mortgages, legal documents, and correspondence, transcribed and analyzed to produce unstructured text, tabular data, and geospatial assets. By analyzing and interpreting the papers of Dana A. Dorsey, we foster a deeper understanding of interpersonal networks, community development, and investment within Miami's Black community, particularly during the pre-redlining era. This initiative will also help us to better understand the roles of civic groups, trade groups, and community alliances in the paths of success for Black people in this community. The set of open data resources produced through this work may also provide insight into south Florida's position within our nation's history of struggle for equity and racial justice.

The core values of this work include addressing the significant gaps in our historical record, critical engagement in data collection and curation processes that are rooted in humanity through the histories of individuals, and establishing a model for future human centered data work.

History

In the early twentieth century, Miami emerged as a boom town, with rapid real estate expansion creating opportunities in transportation, construction, and hospitality. For many Black migrants from the Caribbean and the American South, the city offered economic possibilities, including land and homeownership, unavailable in the rural South. Black real estate investments played a critical role in Miami's growth and led to the development of Colored Town (now Overtown), a center of Black and Afro-Caribbean cultural and economic life. Overtown became nationally recognized for its vibrant music scene, hotels, and restaurants, supported by Black investment in community institutions.

Black entrepreneurs and community leaders mobilized capital through building and loan associations, mutual aid societies, fraternal organizations, insurance companies, churches, and

funeral homes.² They invested in civic life, established business associations, and worked to strengthen their communities. However, navigating the structures of racism and segregation often required difficult compromises and left them vulnerable to exploitation by white-owned financial institutions.³ While some used their success to support civil rights efforts, others took a more cautious approach, focusing on protecting their families and communities within the constraints of Jim Crow.

By the mid-1930s, federally backed lending policies, including the Home Owners' Loan Corporation's redlining maps, further restricted Black access to homeownership. While exclusionary financial practices existed before, redlining institutionalized racial barriers to capital, exacerbating long-term disparities. Today, the racial wealth gap, particularly in homeownership, continues to shape Black economic mobility, resilience, and educational opportunities. This persistent inequity is rooted in a long history of structural racism, racial violence, and systemic barriers to Black success with lasting impact on generational wealth.⁴

The Collection

Dana A. Dorsey (1868 - 1940), who was regarded as a successful businessman of his time, accumulated a variety of properties and businesses in the region. Dorsey and his wife, Rebecca, also sold and leased property to many of Miami's early Black residents. The Dana A. Dorsey Collection, donated by Professor Emeritus Marvin Dunn to the Special Collections and University Archives at Florida International University (FIU), consists of warranty deeds, mortgages, legal documents, and correspondence, which include legal property descriptions of locations of what were the newly created Miami sub-divisions at the turn of the century through 1940. These documents hold significant research and educational value as they frame the growth of the Miami metropolitan region in highly racialized conditions.

The Dorsey papers consist of 291 records and over 620 pages, which have been digitized and are available online through the university's open access digital repository. The Dorsey papers

² Alexa Benson Henderson, *Atlanta Life Insurance Company: Guardian of Black Economic Dignity* (University of Alabama Press, 1990).; Walter B. Weare, *Black Business in the New South: A Social History of the NC Mutual Life Insurance Company* (Duke University Press, 1993), <https://doi.org/10.2307/j.ctv11cw3kd>.; *The History of Black Business in America: Capitalism, Race, Entrepreneurship* | Juliet E. K. Walker, accessed October 11, 2022.; Shennette Garrett-Scott, *Banking on Freedom: Black Women in U.S. Finance Before the New Deal* (Columbia University Press, 2019), <https://doi.org/10.7312/garr18390>.

³ N. D. B. Connolly, *A World More Concrete: Real Estate and the Remaking of Jim Crow South Florida*, Illustrated edition (University of Chicago Press, 2016).

⁴ William A. Darity and Samuel L. Myers, *Persistent Disparity: Race and Economic Inequality in the United States since 1945* (Cheltenham, UK ; Northampton, MA, USA: E. Elgar Pub, 1998). William Darity Jr. et al., "What We Get Wrong About Closing the Racial Wealth Gap," Samuel DuBois Cook Center on Social Equity (Insight Center for Community Economic Development, April 2018).; "Wealth Mobility and Volatility in Black and White," *Center for American Progress* (blog), July 29, 2008. Rakesh Kochhar and Richard Fry, "Wealth Inequality Has Widened along Racial, Ethnic Lines since End of Great Recession," *Pew Research Center* (blog), accessed October 11, 2022.

are a unique and valuable resource as they provide insight into the methods some Black Americans employed to become property owners, start businesses, buy homes, and invest in their communities after Reconstruction, despite profound racial violence and oppression.

The contents within this collection have garnered substantial online engagement, with over 168,841 views of the digital collection materials spanning from 2015 to 2025. However, although the content is openly accessible online, discovery was limited to basic metadata fields, including the titles of the documents, signatories, and legal lot and block description.

This project was designed to enhance access to the Dorsey papers through mapping of the properties, transcriptions of handwritten and typed text, along with structured tabular data. By including comprehensive details, such as all recorded names, gender, marital status, occupations, business affiliations, community organizations, and present-day property locations, we provide multiple entry points for exploration, enabling broader engagement and new avenues for analysis.

Project Planning

The project team remains critically aware of the risks associated with the datafication and potential commodification of human lives, particularly within the structures of academia. To mitigate these concerns, our implementation strategy prioritizes ethical considerations at every stage of our work. Rather than approaching these materials as mere data points, we emphasize the individuals, relationships, and narratives they represent. This commitment extends to the language we use, avoiding extractive terms such as mining and extracting when referring to personal histories. Similarly, we refrain from employing aggrandizing language when discussing figures like Dana A. Dorsey, ensuring a nuanced representation that acknowledges both individual achievements and the systemic barriers Black communities faced. We also take care not to conflate the success of individual entrepreneurs with broader notions of Black wealth, recognizing the persistent economic and structural inequalities that shaped their experiences.

Scholars and community archives have made significant strides in documenting the history of Colored Town (now Overtown), shedding light on the contributions of its residents as well as the zoning, redlining, and land use policies that constrained Black economic mobility. However, comprehensive analytical data on land transactions, business development, and capital distribution in the community from 1896 to 1925 has remained scarce or inaccessible behind newspaper paywalls, tax rolls, or county archives that require records requests. To address these gaps while respecting prior scholarship, we grounded our work in existing literature and consulted historians specializing in Black history in Miami.

Foundational texts such as Nathan Connolly's *A World More Concrete: Real Estate and the Remaking of Jim Crow South Florida*⁵, Marvin Dunn's *Black Miami in the Twentieth Century*⁶, and the Historic Virginia Key Beach Museum Park's *Miami Black History Documentary Series*⁷ have informed our research. Additionally, this project builds upon broader studies of race, housing, and displacement, including Dr. Robin F. Bachin's *Race, Housing, and Displacement in Miami*, *Mapping Inequality: Redlining in New Deal America*⁸, *Mapping Segregation in Washington, D.C.*⁹, and *Mapping Segregation in Washington DC*¹⁰—the latter of which closely aligns with our objective of using individual stories to illustrate systemic patterns of racial exclusion. Our ethical approach also draws from established best practices, including *The Santa*

⁵ N. D. B. Connolly, *A World More Concrete: Real Estate and the Remaking of Jim Crow South Florida*, Illustrated edition (University of Chicago Press, 2016).

⁶ Marvin Dunn, *Black Miami in the Twentieth Century*, Reprint edition (Gainesville: University Press of Florida, 2016).

⁷ "Miami Black History Documentary Series," *Historic Virginia Key Beach Park, Miami, Florida* (blog), accessed June 29, 2021, <https://hvkbmp.org/miami-black-history-documentary-series/>.

⁸ HIS 561 662 Spring 2020, "Race, Housing, and Displacement in Miami," ArcGIS StoryMaps, November 2, 2020, <https://storymaps.arcgis.com/stories/0d17f3d6e31e419c8fd9bbd557f0edae>.

⁹ "Mapping Inequality," accessed October 11, 2022, <https://dsl.richmond.edu/panorama/redlining/>.

¹⁰ "Mapping Segregation DC," accessed October 11, 2022, <http://mappingsegregationdc.org/>.

Barbara Statement on Collections as Data, developed by the *Always Already Computational: Collections as Data* project team.¹¹

Recognizing the importance of stakeholder engagement, we prioritized meaningful collaboration with those most connected to the history we are documenting. Our stakeholders include the communities represented in the collection, collection donor Dr. Marvin Dunn, and the Dorsey family, alongside scholars, librarians, and archivists. By fostering a diverse advisory board and project team—including experts in local and Black history, business history, archival practice, digital humanities, and GIS—we ensured that multiple perspectives informed our work. Our outreach efforts incorporated varied communication strategies to facilitate ongoing input, particularly from the libraries and archives community.

Transparency, feedback, and iterative refinement were central to our methodology. We incorporated a feedback form on our website, held regular meetings with our advisory board, and worked closely with our project consultant, historian Dr. Nathan Connolly. We also engaged in formal discussions with library and archival professionals and maintained an open dialogue with Dr. Marvin Dunn, the donor of the Dorsey collection. Throughout the project, our team remained flexible, reassessing our approach as needed and remaining receptive to new insights. By centering ethical engagement, historical context, and community collaboration, we have sought to create a research framework that is both responsible and impactful.

¹¹ “Collections as Data: Part to Whole Final Report,” accessed November 29, 2023, <https://doi.org/10.5281/zenodo.10161976>.

Data Collection, Curation, & Access Methodologies

By following the individuals associated with Dorsey's investments and property ownership, we have developed a comprehensive resource for examining the effects of discriminatory policies on our built environment, real estate, investment, and social networks striving for prosperity amidst racial oppression. The resources generated through this initiative include (1) unstructured data as transcripts of Dorsey's financial documents, correspondence, and other legal papers, (2) structured data as a dataset derived from his papers containing information about the individuals engaged in the Dorsey real estate transactions and, (3) a geospatial dataset pinpointing the properties involved in his financial transactions.

Data Resources

Data Type	Description	Data Format	Access Points
Unstructured Data	<ul style="list-style-type: none"> transcripts images 	<ul style="list-style-type: none"> TXT, PDF TIFF, JPEG 	<ul style="list-style-type: none"> digital library system data repository
Structured Data	<ul style="list-style-type: none"> tabular data 	<ul style="list-style-type: none"> CSV 	<ul style="list-style-type: none"> data repository
Geospatial Data	<ul style="list-style-type: none"> historical maps 	<ul style="list-style-type: none"> georeferenced TIFF GeoJSON, hosted feature service/layer 	<ul style="list-style-type: none"> ArcGIS data repository

In the following sections, we critically analyze and reflect on the methodologies used for data collection and curation, as well as the strategies employed to ensure the project's ethical implementation.

Unstructured Data

The project team produced full-text transcripts of the handwritten and typed materials in the Dorsey collection to create usable text data, enhance search and retrieval, and improve accessibility for individuals using screen readers. A total of 291 transcripts, encompassing over 620 pages, have been created from the Dorsey papers. This unstructured data represents a key deliverable in our efforts to make these historical documents more accessible and usable for a wide range of audiences.

The project began with an evaluation of several Optical Character Recognition (OCR) software options, including ABBYY, Tesseract, and Prime Recognition. After thorough testing, we found that the combination of paper quality, clarity of typed print, and the nature of the handwritten content made automated transcription unreliable. Consequently, we explored alternative

methods and determined that speech-to-text technology, supplemented with manual quality control and editing, provided a more accurate and feasible solution.

After testing various speech-to-text services, the team chose a low-tech yet effective strategy using Microsoft Word's speech recognition/dictation tool. This tool was selected for its accuracy and user-friendliness, particularly for our interns, who played a crucial role in the transcription process.

Throughout the project, several implementation challenges required careful consideration. One significant question was how to accurately transcribe documents that included annotations, edits, or other marginalia, as well as those containing complex tables and charts. We adhered to the Library of Congress "By the People" transcription guidelines, which provided a solid framework for handling most of these complexities. However, in cases where documents included lengthy tables that could not be effectively represented in text, we opted to have interns include a note directing users to the original document for full context. This approach allowed us to maintain a balance between accuracy and usability, ensuring that the unstructured data produced is both meaningful and accessible to users while preserving the integrity of the original documents.

Structured Data

Another key deliverable of the project was the creation of structured data from the content of the Dorsey papers. This data is intended to provide researchers with a comprehensive view of the individuals and businesses that played a role in real estate transactions and other community development activities in the Overtown area.

The process of selecting and recording data from the Dorsey collection involved a series of ethical and practical considerations to ensure our work was both responsible and aligned with the project's mission and values. Here is a summary of the questions we asked ourselves throughout this process:

- Is this information useful for gaining historical insights?
- Does it adhere to the mission and values of this project?
- Does this information violate personal privacy?
- Are we presenting this information in a way that conflicts with cultural values or norms of the groups represented?
- How can we ensure transparency in our data selection and recording practices?

This work has also compelled us to think more deeply about the ethics of collecting demographic data. Demographics can be useful in expanding historical representation as well as providing a deeper understanding of the complexities of life for this group of individuals. It also helps in identifying disparities and how intersectionality may impact individual experiences. At the same time, we recognize that the individuals in these legal documents and correspondence had little to no control over how they were identified, whether by race, gender, or other attributes, leaving us reliant on others' recordings of their identities.

Rather than collecting and exposing all possible data points, we adopted a data minimization approach, prioritizing relevance, ethical considerations, and historical integrity. We chose to not look into secondary sources like the census and to not project our own suppositions of race and gender. During our preliminary research, we used secondary sources like census and tax lists to confirm certain details, such as the spelling of names and other identifiers. However, we encountered inconsistencies—for example, one individual's gender was recorded differently in different census years compared to our documents. Given these challenges, we concluded that while race and gender information can be valuable for historical research, it quickly becomes problematic when such demographics are speculated upon, assigned, or reassigned without clear evidence. Therefore, in the interest of transparency, clarity, and ethics, we decided to record only the information directly available in the archival documents themselves. We intentionally avoided consulting secondary sources like the census or projecting our own assumptions about race and gender.

This methodology ensures that our data is both ethically responsible and faithful to the original documents, allowing researchers to draw their own conclusions without the influence of potentially speculative or inaccurate demographic classifications.

In addition to these ethical concerns, privacy considerations also played a crucial role in our decision-making process. The team took these issues seriously and consulted resources such as the National Archives and Records Administration (NARA) federal regulations regarding the release of personally identifiable information (PII). These guidelines specify that NARA withholds records not previously made public that are less than 75 years old. However, many of the records in the Dana A. Dorsey collection are public and exceed this threshold. Given this, we determined that making available the names of individuals along with limited demographic details, such as gender, marital status, occupations, and business affiliations, would not pose significant privacy risks.

The structured data for this project was derived through a meticulous manual process. Interns were primarily responsible for populating a Google spreadsheet with the data curated from the documents, which included information about over 880 individuals and more than 270 businesses, municipalities, and organizations such as law offices, insurance companies, and fraternal associations. To ensure accuracy and consistency, the interns' work underwent rigorous quality control checks, followed by additional review and follow-up by the project directors. This hands-on process allowed us to capture detailed and reliable data. This process not only enhances the utility of the structured data for historical research but also ensures that the data is presented in a manner that is ethical, transparent, and aligned with the project's goals. The structured data is available in the data repository as downloadable CSV files, ensuring accessibility for researchers and other users.

Data Documentation

A crucial component of establishing transparent data practices in this project has been the creation of consistent and publicly accessible data documentation. To achieve this, the team

adopted the Datasheets for Datasets methodology.¹² While this approach was originally developed to improve transparency in big data and machine learning, its application to small cultural heritage datasets is equally valuable for historians and digital humanists. The project team found the Datasheets for Datasets framework particularly useful due to its structured guidance, which includes broad documentation topics supplemented by detailed, specific questions. This documentation addresses key areas such as the motivation behind data collection, the dataset's composition, collection and normalization processes, anticipated uses, and strategies for mitigating potential misuse.

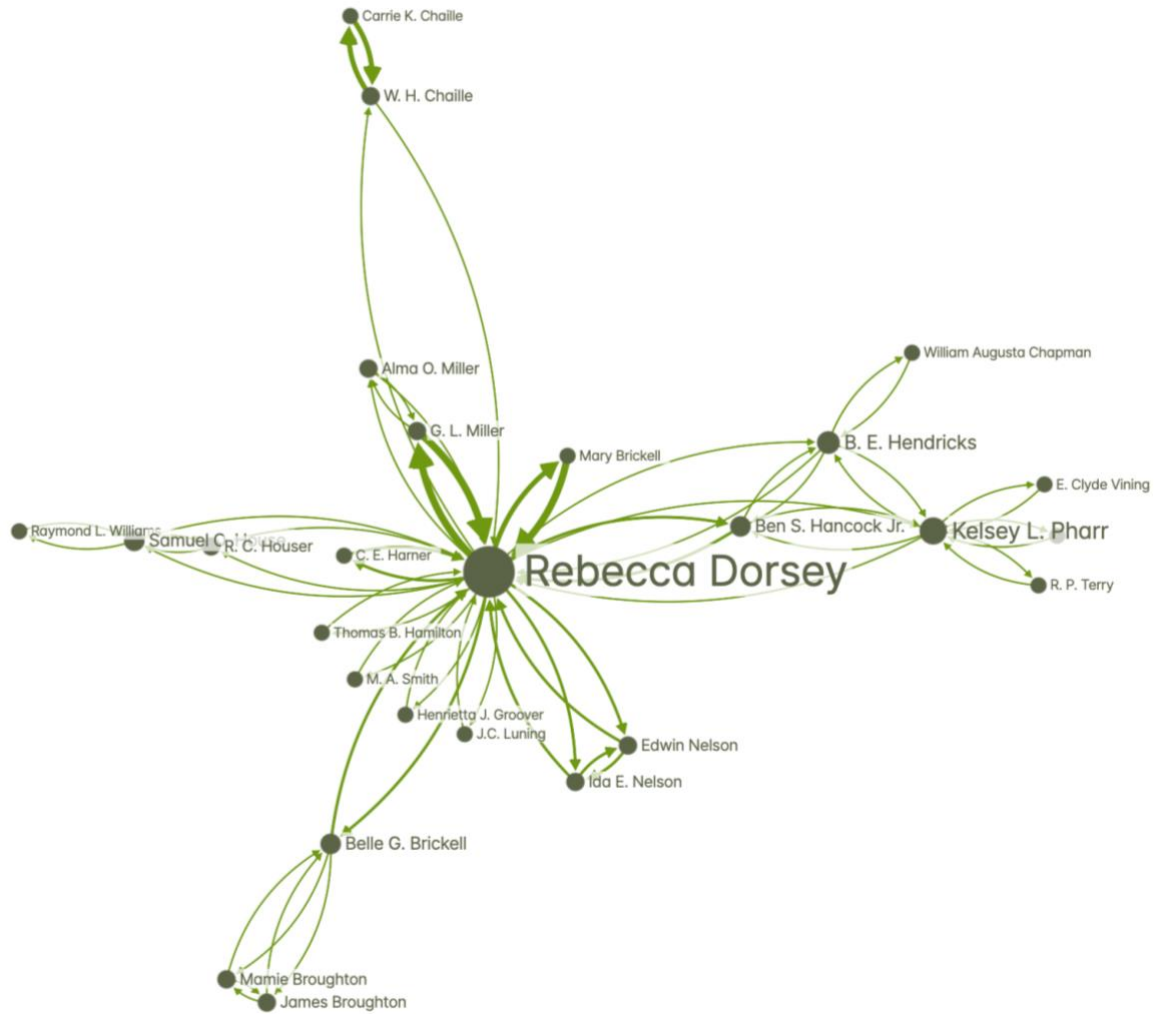
Suggested mitigation practices for the project dataset include:

- Researchers should not assume that the dataset fully represents the broader Black community. Supplementing findings with data from other archival sources can provide better historical context.
- When possible, users should cross-check data with original documents and acknowledge potential error rates in publications.
- Race and gender should be treated as historically specific constructs, avoiding the application of modern labels or assumptions. Transparency about categorization criteria is essential.
- To prevent misinterpretation, researchers should not imply any present-day property ownership claims based on historical records. Publications should include disclaimers to reinforce this distinction.
- Users should approach the dataset with sensitivity, ensuring that research does not lead to reputational harm or undue scrutiny of individuals. Where appropriate, consulting community stakeholders can help guide responsible usage.

Network Analysis

The project team utilized network graphing tools such Graph Commons to visualize relationships between individuals, groups, and organizations mentioned in the documents. By visualizing connections between entities whose names co-occur on documents, one may discern insights into community networks. The below sample network graph reveals that aside from Dana Dorsey himself, his spouse, Rebecca Dorsey, played a significant role in his real estate ventures—emerging as a central figure in the networks the team graphed. Other pivotal connectors include Kelsey L. Pharr and Belle Brickell, who extend from Rebecca Dorsey's "node." Utilizing network graphing tools, the team has been able to create customized visualizations that underscore particular relationships or patterns, quantify node centrality, and pinpoint clusters of interconnected nodes. This sample network exploration will hopefully provide historians with the necessary tools and data to visualize additional links between businesses, occupations, locations, and more, further enriching our understanding of these networks.

¹² Timnit Gebru et al., "Datasheets for Datasets," *Communications of the ACM* 64, no. 12 (December 2021): 86–92, <https://doi.org/10.1145/3458723>.



Geospatial Data

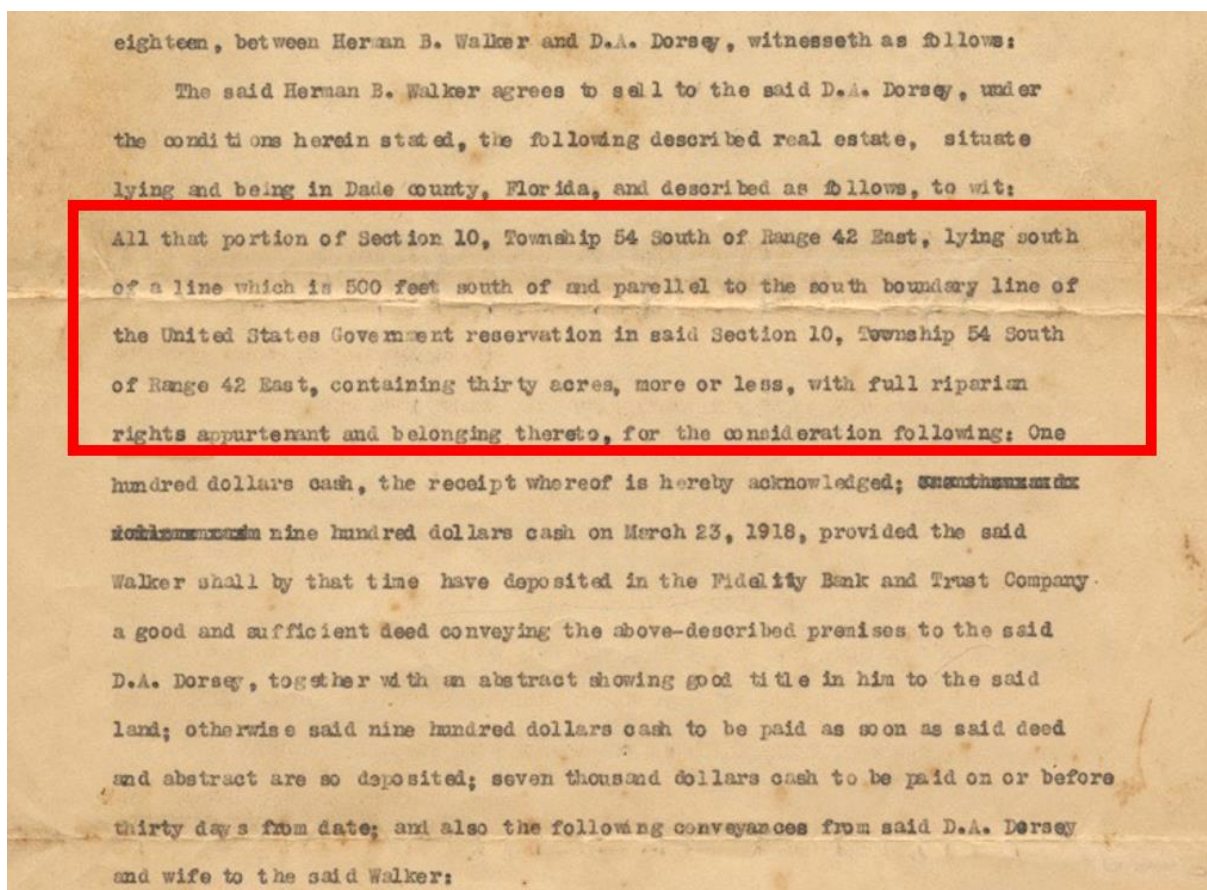
The geospatial component of our project incorporates both georectification and geolocation techniques. Georectification involves placing a historical map onto a spatial grid in relation to contemporary maps or satellite imagery, aligning geographic attributes. This process allows researchers to compare historical map data with present-day geographic features and locations. In order to compare the location of properties listed in the Dorsey papers, digitized maps from a 1925 Hopkins' plat book of Miami¹³ were stitched together and georectified¹⁴ so they aligned with current satellite imagery. Additionally, polygons were used to indicate the geographic area of the properties listed in the Dorsey documents, which can now be explored via an interactive

¹³ <http://dpanther.fiu.edu/dpService/dpPurlService/purl/FISC000211/00001>

¹⁴ <https://fiugis.maps.arcgis.com/apps/mapviewer/index.html?webmap=8669acbed3f44fdca5bbff9f71b72e6d>

online ArcGIS interface. Furthermore, this geospatial analysis produced a comprehensive, step-by-step methodology for georeferencing historical locations, serving as a valuable resource for future historical geospatial research.¹⁵

To demonstrate how this process works, please see the 1918 property transfer between Herman B. Walker and Dana A. Dorsey pictured below.¹⁶ The agreement provided a precise location (outlined in red) for the property: Section 10, Township 54 South, Range 42 East. This method for describing properties is known as the Public Land Survey System (PLSS), which is the surveying method developed and used in the United States to plat or divide real property for sale and settling.

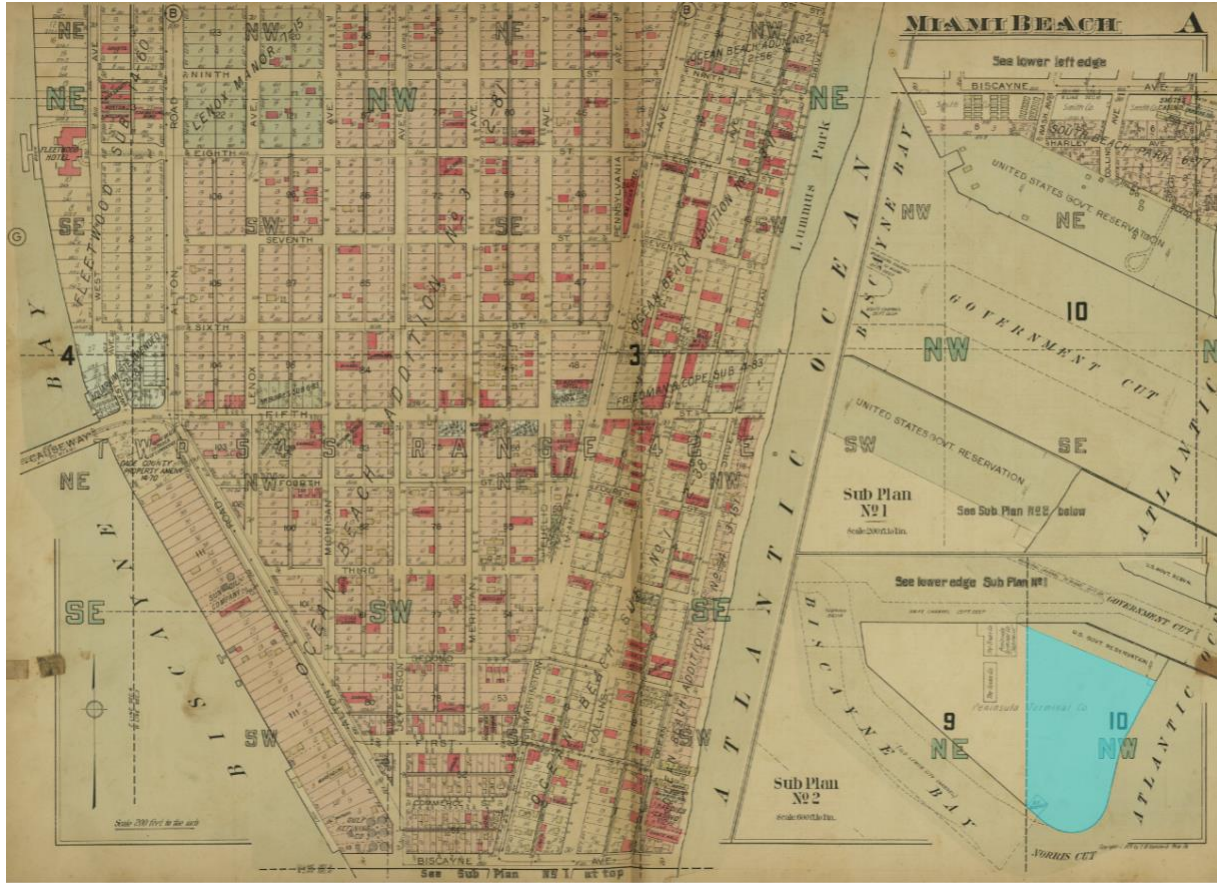


Box: 1, Folder: 2260. Dana A. Dorsey Collection, SPC5070. FIU Special Collections.

To illustrate property boundaries in the city of Miami, the G. M. Hopkins' Company created an atlas of cadastral maps that combine the PLSS and Lot and Block Survey System. By manually plotting the PLSS description from the property transfer between Walker and Dorsey, we were able to visualize the location of the property as a polygon (see below in blue) on the map sheet.

¹⁵ <https://storymaps.arcgis.com/stories/cfc11fa257cb411cba7e3db300c18139>

¹⁶ <http://dpanther.fiu.edu/dpService/dpPurlService/purl/FI05100405/00001>



When the properties' boundaries were overlaid on a modern satellite image, through the process of georectification, it clearly delineates the Dorsey properties within our modern landscape. The maps below show the geolocated property from the archival document, now visualized on the 1925 plat and a modern satellite image of Miami.



Project Outcomes and Impact

Deliverables

Through the course of this project, the team produced a total of 291 transcripts, identified 882 individuals and 272 entities mentioned within the archive papers, generated approximately 50 georectified maps, and created polygons for 145 individual lots and blocks, shedding light on the complex socio-cultural and legal landscapes of early 20th-century Miami. While transcribing the Dorsey papers and mapping property locations, the project team uncovered significant historical insights, including language reflective of the period's legal and societal norms, such as prohibition clauses, racial covenants, and gender-specific insurance exclusions.

An example from 1919, predating federal Prohibition, demonstrates local legal restrictions on alcohol: "Provided nevertheless that no alcoholic, malt, spirituous or intoxicating liquor be manufactured or sold thereon."¹⁷ This clause from a property deed illustrates local efforts to regulate alcohol, signaling community attitudes toward prohibition before it became national law. Racial covenants from the 1920s and 1930s, such as "None of the lots hereby conveyed shall be sold, leased, or rented in any form or manner by any title either legal or equitable to any person or persons other than of the Caucasian race...",¹⁸ underscore the legally enforced segregation of the time. Additional covenants explicitly restricted occupancy to "the negro race"¹⁹ in areas like Liberty City, reflecting the rigid racial boundaries embedded in property laws. Furthermore, a 1923 insurance document includes exclusions related to gender and morality, stating, "Benefits will not be paid in case of sickness or death from intemperance or immorality nor from disease contracted before joining the association nor for injuries received in a fight or while violating the law nor will sick benefits be paid for rheumatism, venereal diseases, disease peculiar to women or cancer of the mammae or of the uterus,"²⁰ highlighting societal biases in access to financial protections.

The transcription process also revealed valuable connections to prominent historical figures who significantly contributed to the infrastructure of Miami's Black community. The papers reference William B. Sawyer²¹, a pioneering African American doctor who established vital medical and hospitality institutions in Miami's historically Black neighborhoods. Kelsey L. Pharr²², the first Black embalmer in Miami and one of the few Black cemetery owners in the South (Lincoln Memorial Park Cemetery), is frequently mentioned. The documents also include references to Robert Freeman Burdine²³, a lawyer and treasurer of the Burdines chain of department stores,

¹⁷ <http://dpanther.fiu.edu/dpService/dpPurlService/purl/FI05091602/00001>

¹⁸ <http://dpanther.fiu.edu/dpService/dpPurlService/purl/FI05091611/00001>

¹⁹ <http://dpanther.fiu.edu/dpService/dpPurlService/purl/FI05101201/00001>

²⁰ <http://dpanther.fiu.edu/dpService/dpPurlService/purl/FI05090601/00001>

²¹ <http://dpanther.fiu.edu/dpService/dpPurlService/purl/FI05082311/00001>

²² <http://dpanther.fiu.edu/dpService/dpPurlService/purl/FI05082918/00001>

²³ <http://dpanther.fiu.edu/dpService/dpPurlService/purl/FI05100404/00001>

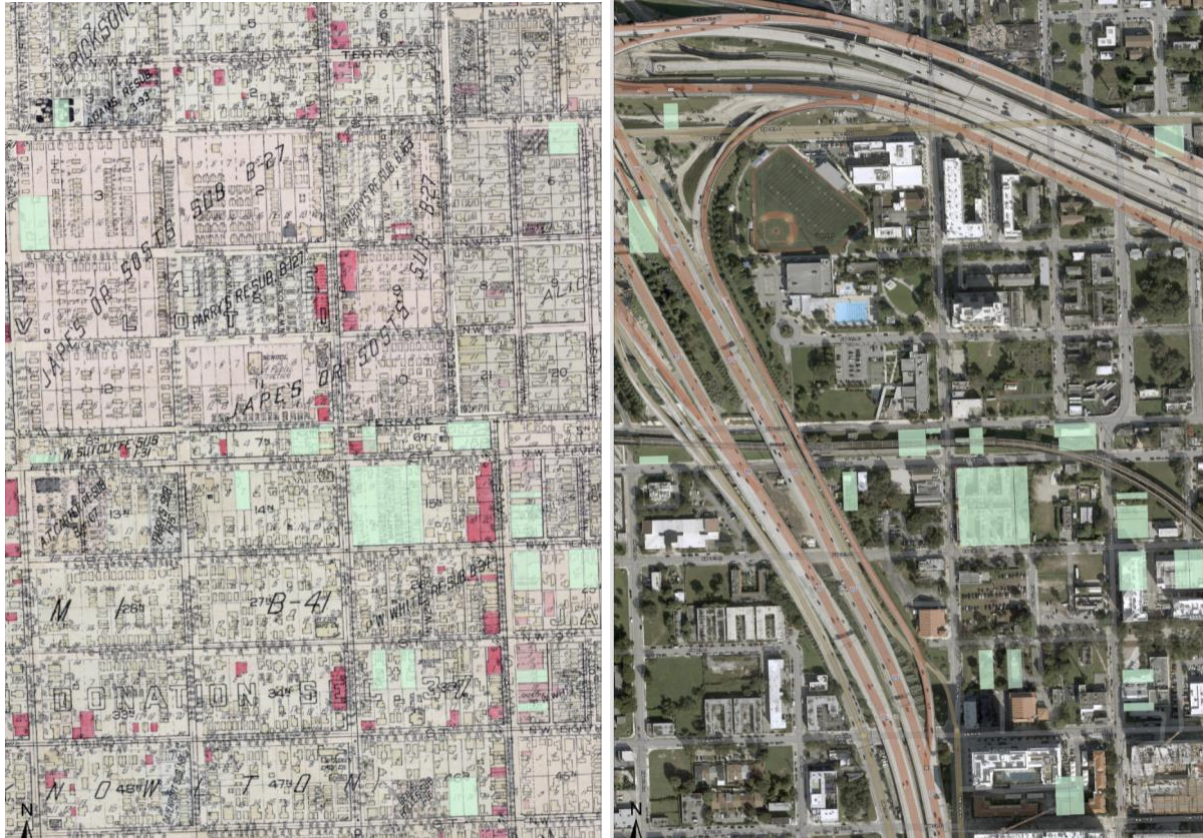
and Richard E. S. Toomey,²⁴ Miami's first Black attorney known as "The Soldier Poet," whose legal witness signatures appear in the collection. The Dorsey papers further document contributions to educational infrastructure, such as land donations for schools serving Black students²⁵ and promissory notes²⁶ supporting the building fund for Florida Normal and Industrial Institute (now Florida Memorial University). These connections not only emphasize the agency within the Black community but also provide a deeper understanding of how these individuals navigated and influenced a segregated landscape, fostering community amid systemic barriers.

During the mapping process, the team identified additional landmarks that reflect the community structures built under segregation, such as Dorsey Park, which was designated as a "Colored" city park in the 1925 Hopkins' plat book of Miami. The mapped residential properties, when overlaid with modern maps, reveal the significant transformations in the urban landscape over the past century. The routes of the I-95 and I-395 expressways and the Metrorail line now cut through areas that were once Black neighborhoods, including Overtown. This overlay not only confirmed the locations of former Black residences that were claimed through eminent domain for highway construction in the 1960s but also highlighted the extensive and often destructive impact of urban development. This expansion of infrastructure systematically dismantled established communities, displacing residents and eroding the social and economic fabric of Overtown.

²⁴ <http://dpanther.fiu.edu/dpService/dpPurlService/purl/FI05092703/00001>

²⁵ <http://dpanther.fiu.edu/dpService/dpPurlService/purl/FI05101805/00001>

²⁶ <http://dpanther.fiu.edu/dpService/dpPurlService/purl/FI05101812/00001>



This project underscores the significance of preserving and sharing the stories of this community, highlighting not only their historical importance but also the critical need for thoughtful curation and ethical access to the data. Our aim is not to dictate a singular narrative but to provide broad access to these resources, allowing for a fuller exploration of the diverse and complex history of this region. By making these materials widely available, we seek to honor the multifaceted experiences and contributions of the individuals and communities represented.

Developing a Community of Practice

Through this initiative, an ethics-driven community of practice has emerged, dedicated to the responsible handling of historical data from marginalized communities. The project team aims to grow this practice, guide the development of new resources, foster partnerships, and shape the next steps in enhancing access to historical data. Our approach prioritizes transparency, representation, iteration, and respect for the individuals whose lives are documented, setting a standard for future endeavors.

Critical Language Choices

Language plays a crucial role in implementing ethical data practices, guiding our project team to prioritize non-extractive approaches throughout our work. Language is not neutral, particularly in historical research and data curation involving marginalized communities. Throughout this

project, we intentionally avoided extractive or exploitative terminology that reduces individuals and communities to data points. Instead of "mining," "extracting," or "discovering" information, we describe our work as "documenting," "interpreting," or "curating" to acknowledge the agency and lived experiences of those represented in the materials.

Additionally, we took care in how demographic attributes were recorded, resisting deterministic assignments of race and gender. Instead, we preserved identifications as they appeared in historical documents while acknowledging the power structures that shaped these records.

Transparency

Transparency is foundational to ethical data work, ensuring that research processes and decision-making frameworks are accessible and accountable. We established transparency through:

- **Public Documentation:** Detailed documentation of our methodologies is available, allowing others to understand how data was selected, structured, and interpreted.
- **Clear Attribution:** Each dataset, image, and visualization includes citations to source materials; and public facing project outcomes recognize the individuals and institutions that contributed.
- **Explicit Decision-Making Processes:** When faced with ethical dilemmas, such as whether to include inferred demographic data, we documented our choices and rationale to encourage dialogue and critique.
- **Community Input:** We created avenues for stakeholder feedback, including advisory board discussions, project consultant input, and an open feedback forum on our website.
- **Data Provenance Tracking:** Every data point is linked back to its original document, ensuring users can verify authenticity and context.

Representation

The project emphasizes broad accessibility and representation, ensuring that data practices reflect the diverse communities represented in the Dorsey papers. The goal is not just to document these histories but to design systems that affirm the presence, contributions, and agency of historically marginalized groups.

The project team included historians, librarians, GIS specialists, and digital humanists, ensuring that multiple perspectives shaped the methodologies and interpretations, creating resources that better serve scholars, students, and community members seeking to engage with Miami's Black history. All project outputs, including structured data, transcriptions, and geospatial maps, are available through open access platforms to remove barriers to use. The structured data includes fields that highlight affiliations with Black businesses, mutual aid societies, and civic organizations, providing alternative entry points for exploring community life beyond property transactions.

Iteration

Ethical data curation is not a static process but an ongoing practice of refinement, reassessment, and adaptation. Throughout the project, we employed an iterative methodology that allowed for continuous feedback and improvement. Regular dialog with our advisory board and project consultant ensured that we could adjust our approaches in response to new insights. As our understanding of the collection deepened, we reassessed how to structure data fields to better reflect relationships and social dynamics. When OCR proved unreliable for transcription, we shifted to a human-assisted speech-to-text workflow, prioritizing accuracy over automation. Iteration allows us to remain responsive to ethical concerns, scholarly advancements, and evolving understandings of best practices in data ethics.

The innovative data methodologies established through the Dana A. Dorsey project provide a valuable framework for creating more inclusive and comprehensive resources on Black history throughout Florida. By emphasizing ethical considerations and community engagement, this project offers a model for others seeking to construct data resources that not only preserve history but also empower and serve the communities they represent. As we continue to expand and share these data resources, we hope to inspire broader efforts that embrace and promote access to historical knowledge.

Consultant Report

[NEH Consultants Report \(2024\).pdf](#)

The project team engaged historian Dr. Nathan Connolly as a consultant to provide expert guidance and insight for the Dana A. Dorsey Collection project. Dr. Connolly, a renowned historian and author of *A World More Concrete: Real Estate and the Remaking of Jim Crow South Florida*, is recognized for his scholarly work on racism, capitalism, politics, cities, and migration in the late-nineteenth and twentieth centuries. As a consultant, Dr. Connolly played a crucial role in advising the team at various stages of the project, helping to shape its direction, and ensuring its historical integrity. He also authored a final report that includes recommendations for enhancing the project's impact and accessibility.

The report highlights several key aspects of the project, including the historical significance of the Dorsey papers, which documents the experiences of African Americans in Miami between the 1910s and 1940s, revealing the complex relationships between black business people and their white counterparts and providing insights into racial, social, and economic interactions during this period. The collection also provides foundational data for understanding broader historical patterns, including economic practices like redlining and the role of black business people in the Civil Rights Movement.

Dr. Connolly also addresses the project website as a platform designed to be accessible to both novice and expert researchers, providing tools for document analysis and educational use. The report highlights the project's emphasis on a transparent and iterative process that aims to democratize access to the archives and support the present-day needs of marginalized

communities. The Dorsey papers are valuable for researchers studying African American history, urban development, and social dynamics in early Miami.

Lastly, the report suggests that additional investments could make the collection more user-friendly, such as improving search capabilities, adding a glossary for complex real estate terms, and creating a dedicated website. It also recommends further digitization efforts, including enhanced keyword search functions and expanded digital mapping.

Future Work

One of the most exciting aspects of this project has been envisioning the various avenues of archival, digital humanities, and data ethics exploration we can follow. As the project progressed, we gained deeper insights into the community and individuals involved in Dorsey's correspondence and real estate transactions. This expanded understanding opens the door for future phases of the project to more directly engage with the community, including descendants of those connected to Dorsey and the businesses he interacted with. Building these connections could offer valuable perspectives that inform how the information we locate can be utilized in ways that benefit the community, beyond its scholarly significance. Furthermore, this engagement could lead us to develop new and more effective methods of communicating the project's findings that resonate with the represented community and other audiences, ultimately fostering a broader impact and potential for enacting change.

The project team is also considering enhancements to the current data dissemination, possibilities of expanding the regional scope and interpersonal networks, collaborating with historians to further develop narratives from these resources, as well as developing ways to incorporate the resources in teaching and learning.

Some of the ways we may be able to enhance the current dissemination are to develop more advanced search functions that allow users to navigate the collection more effectively, such as by moving between properties, people, and associations, and linking these data points to a broader geographic context. For instance, we are exploring the possibility of providing a website table lookup feature that would allow users to match names with their sources directly.

To expand the scope, we may also explore avenues for partnerships. Additional data resources and/or archival collections may be incorporated as well as organizations and individuals who may be potential collaborators.

There are also many pedagogical possibilities. The team is open to working with educators to create lesson plans, research guides, and other educational materials to support the use of the Dorsey papers in K-12 and higher education settings, emphasizing primary source analysis and historical inquiry. This is just one avenue for encouraging greater involvement from local communities, scholars, and educators in using the Dorsey Papers. The team may also continue to host workshops, webinars, and collaborative projects to promote the use of the collection in academic and community research.

Acknowledgements

On behalf of Jamie Rogers, Associate Dean for Strategic Initiatives and Project Director, and Rhia Rae, Digital Assets Project Manager, Special Collections and University Archives, and Co-Project Director, we extend our sincere appreciation to the dedicated project team at Florida International University (FIU) Libraries. Their collaboration and expertise were essential to the success of this initiative. Special recognition goes to Rebecca Bakker, Digital Collections Librarian; Molly Castro, Digital Humanities Librarian Consultant; Jill Krefft, Dean, University of Tulsa Library; Dr. Diana Ter-Ghazaryan, Associate Professor of Spatial Sciences, University of Southern California; Luis Garcia Falcon, GIS Specialist; and Sonia Santana Arroyo, Research and Data Repository Manager. Their commitment and tireless efforts were crucial in advancing the goals of this project.

We are grateful for the ongoing collaborations with our faculty colleagues in FIU Department of History, Dr. Kenneth Lipartito, Dr. Julio Capó, and Dr. Daniel Royles, who generously contributed their time to work with the Libraries in planning and writing of the grant proposal. Furthermore, we would like to thank Dr. Lipartito for his valuable contribution to the proposal, offering his expertise and helping to articulate the historical context and broader relevance of the Dana A. Dorsey collection.

We also acknowledge the invaluable contributions of our project interns, Juan Lopez, Catherine Tibaquirá, and Rosemary Toval, for their diligent work in transcription and data curation. Their enthusiasm and dedication were key to the project's success.

The project team extends our deepest gratitude to the collection donor, Dr. Marvin Dunn, and our project consultant, Dr. Nathan Connolly, from Johns Hopkins University, for their expert guidance, insightful recommendations, and support throughout the project's various stages.

We also thank the members of our advisory board: Dr. Robin Bachin, Dr. Julio Capó, Jr., Dr. Marvin Dunn, Dr. Sharony Green, Dr. Kenneth Lipartito, Dr. Chanelle N. Rose, Dr. Daniel Royles, Dr. Melanie Shell-Weiss, and Althea Silvera. Their expertise, feedback, and guidance have greatly enhanced the scope and impact of this project.

In the fall of 2020, Dr. Levente Juhasz from our university's GIS Center devised a way to use scanned plat map sheets from the Hopkins 1925 Plat Map of Greater Miami to georeference the properties in the Dorsey papers. Working together, Dr. Juhasz, Dr. Diana Ter-Ghazaryan, and GIS intern Adeola Kolapo-Oluwo, geolocated 30 properties from the Dorsey documents.²⁷ It was the success of this work that led to this larger effort.

Finally, we extend our deepest appreciation to the National Endowment for the Humanities for their sponsorship and support, which made this project possible.

²⁷ <https://storymaps.arcgis.com/stories/cfc11fa257cb411cba7e3db300c18139>

Bibliography

- Center for American Progress. "Wealth Mobility and Volatility in Black and White," July 29, 2008. <https://www.americanprogress.org/article/wealth-mobility-and-volatility-in-black-and-white/>.
- "Collections as Data: Part to Whole Final Report." Accessed November 29, 2023. <https://doi.org/10.5281/zenodo.10161976>.
- Connolly, N. D. B. *A World More Concrete: Real Estate and the Remaking of Jim Crow South Florida*. Illustrated edition. University of Chicago Press, 2016.
- Darity, William A., and Samuel L. Myers. *Persistent Disparity: Race and Economic Inequality in the United States since 1945*. Cheltenham, UK ; Northampton, MA, USA: E. Elgar Pub, 1998.
- Dunn, Marvin. *Black Miami in the Twentieth Century*. Reprint edition. Gainesville: University Press of Florida, 2016.
- Garrett-Scott, Shennette. *Banking on Freedom: Black Women in U.S. Finance Before the New Deal*. Columbia University Press, 2019. <https://doi.org/10.7312/garr18390>.
- Gebru, Timnit, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé Iii, and Kate Crawford. "Datasheets for Datasets." *Communications of the ACM* 64, no. 12 (December 2021): 86–92. <https://doi.org/10.1145/3458723>.
- Henderson, Alexa Benson. *Atlanta Life Insurance Company: Guardian of Black Economic Dignity*. University of Alabama Press, 1990.
- Historic Virginia Key Beach Park, Miami, Florida. "Miami Black History Documentary Series." Accessed June 29, 2021. <https://hvkbmp.org/miami-black-history-documentary-series/>.
- Jo, Eun Seo, and Timnit Gebru. "Lessons from Archives: Strategies for Collecting Sociocultural Data in Machine Learning." In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 306–16. FAT* '20. New York, NY, USA: Association for Computing Machinery, 2020. <https://doi.org/10.1145/3351095.3372829>.
- Kochhar, Rakesh, and Richard Fry. "Wealth Inequality Has Widened along Racial, Ethnic Lines since End of Great Recession." *Pew Research Center* (blog). Accessed October 11, 2022. <https://www.pewresearch.org/fact-tank/2014/12/12/racial-wealth-gaps-great-recession/>.
- "Mapping Inequality." Accessed October 11, 2022. <https://dsl.richmond.edu/panorama/redlining/>.
- "Mapping Segregation DC." Accessed October 11, 2022. <http://mappingsegregationdc.org/>.

Spring 2020, HIS 561 662. "Race, Housing, and Displacement in Miami." ArcGIS StoryMaps, November 2, 2020.

<https://storymaps.arcgis.com/stories/0d17f3d6e31e419c8fdffbd557f0edae>.

The History of Black Business in America: Capitalism, Race, Entrepreneurship | Juliet E. K. Walker. Accessed October 11, 2022. <https://uncpress.org/book/9780807859100/the-history-of-black-business-in-america-capitalism-race-entrepreneurship/>.

Weare, Walter B. *Black Business in the New South: A Social History of the NC Mutual Life Insurance Company*. Duke University Press, 1993.

<https://doi.org/10.2307/j.ctv11cw3kd>.

William Darity Jr., Darrick Hamilton, Mark Paul, Alan Aja, Anne Price, Antonio Moore, and Caterina Chiopris. "What We Get Wrong About Closing the Racial Wealth Gap." Samuel DuBois Cook Center on Social Equity. Insight Center for Community Economic Development, April 2018.