Critical Community-Centeredness: Ethical Considerations for Computational Archival Studies

Madelynn Dickerson
Digital Scholarship Services
University of California, Irvine Libraries
Irvine, USA
http://orcid.org/0000-0002-4541-4800

Audra Eagle Yun
Special Collections & Archives
University of California, Irvine Libraries
Irvine, USA
https://orcid.org/0000-0002-6284-0123

Abstract—**In this paper, we call for computational archival studies to prioritize social justice and community-centeredness. Our initial research findings, as well as the work of community archives, provide evidence of the need to elevate and truly center the voices of those depicted (or underrepresented) in large-scale digital archives, leveraging the power of computational thinking with the transformative experience of seeing oneself represented (or representing oneself) in digital collections.**

Keywords—community archives, digitization, digital archives

I. INTRODUCTION

As practitioners in an academic library environment, we are looking to bridge theory and praxis of community-centered archives in order to sustain social justice in digital archives. We are building a coalition that pushes critical community-centeredness as a foundational element of archival work. In this paper, we call for computational archival studies to center social justice to counter digital practices that may be essentializing and unintentionally harmful.

Occupying the unceded lands of the Acjachemen and Tongva people, the University of California, Irvine (UCI) is a nearly majority first-generation school that is federally recognized as an Asian American Native American Pacific Islander Serving Institution (AANAPISI) and a Hispanic Serving Institution (HSI). We also aspire to become an institution committed to Black thriving and to be more accountable to the Indigenous communities of the region. UCI is located in Orange County, the third most populous county in California and sixth most populous in the United States, home to an increasingly diverse population of 3.2 million [1]. Nearly half of Orange County households report that their primary language in the home is not English. The county is arguably at the center of polarizing debates about immigration, housing insecurity, and the conflation of ethnic studies and critical race

theory, to name a few important issues. These are relevant contexts for the work of community-centered archives practice.

Critical digital archives necessitates critical, computational archival studies. Acknowledgement of intentional archival curation and reparative metadata practices is paramount in ensuring that context matters in digital archival aggregations. As meanings and context change over time, computational archival studies must center and foster community-centeredness iteratively.

II. BACKGROUND

Computational approaches to archival access, discovery, and analysis offer myriad opportunities to understand humanity and synthesize this understanding on an increasingly large scale. The possibilities and potential benefits of digital archives are many - from the ability to make connections across fragmented collections, to digital aggregation, digital curation, and beyond. The egalitarian vision of a unified digital memory of the world is all at once inspiring and aspirational.

It is in this environment of potential that Devon Mordell [2] proposes a fifth archival paradigm, "archives as data," joining four concepts established previously by Terry Cook: evidence, memory, identity, and community. This framework follows on the idea of "collections as data," which encourages "computational use of digitized and born digital collections" [3]. Mordell articulates the "archives as data" paradigm in recognition of the ongoing expansion of computational approaches to archival processes, but she also uses this proposed paradigm to make a call for scholars and practitioners to consider the ethical and social justice implications for such a concept.

"As computational methods begin to constitute a larger part of archival work, the efforts of the archival profession over the past three decades to pluralize the archival endeavor and to

introduce a social justice orientation - incorporating critical race theory, feminist theory, queer theory, and other overtly politicized modes of inquiry - into archivy may be stifled or even undone. If an archives-as-data paradigm is genuinely taking shape, then the archival profession has a responsibility to ensure that a social justice critique is maintained within it" [2].

Archives are people, and people are more than the records (and data) that document them [3]. In this paper, we are primarily concerned with agency, rights, and care as they relate to the *subjects* of (digital) archives. This builds on Cook's concept of community in the archival paradigm, which suggests "empowering communities to look after their own records, especially their digital records, by partnering professional archival expertise and archival digital infrastructures with communities' deep sense of commitment and pride in their own heritage and identity" [2].

Omissions, misrepresentations, and erasures have always been present in archival and library descriptive standards, hindering self-discovery. From the catalog to the archival finding aid, our fields have a long legacy of bias that presages algorithmic bias in contemporary search engines [5]. The impact of literal absence or technical erasure through lack of sufficient descriptive context is equivalent from the standpoint of those unable to see themselves represented or representing in archives. Research conducted as part of the Institute of Museum and Library Services (IMLS)-funded "Transforming Knowledge Transforming Libraries" project (2017 – 2020) reveals patterns of harm and negative affective responses to not seeing oneself in archives. Among over 700 undergraduate college students surveyed at UCI, the majority (79%) expressed a negative sentiment to the experience of symbolic annihilation in archives, both digital and physical [6]. In the landscape of archival data, likewise, archival silences can be exacerbated by the same structures, systems, and tools designed to streamline access and discoverability.

Michelle Caswell mapped the concept of "symbolic annihilation" (the ways in which members of marginalized communities are absent, underrepresented, or misrepresented in mainstream media) to mainstream archives [7]. Caswell, Cifor, and Ramirez developed a framework for the potential of community archives to foster "representational belonging" and counteract symbolic annihilation through three strategic areas of impact: *ontological* impact (in which members of marginalized communities get confirmation: "I am here"); *epistemological* impact (in which members of marginalized communities get confirmation: "we were here"); and *social* impact (in which members of marginalized communities get confirmation: "we belong here") [8]. These ideas reinforce the fact that communities (and their archives) exist and have the power to represent themselves independently.

As computational projects continue to develop, "archives as data" is a useful lens through which to interrogate the transformation that takes place when the *people* at the center of archival collections are not only digitized, but "datafied." This datafication does not intrinsically support the needs of communities that have been marginalized in the historical record, as datafication requires structure. Structured data involve some element of loss, reduction, and/or essentialization of

context, and as Mordell further explains, "supporting community-oriented goals is not an inherent feature of datafication" [2].

Datafication takes the problem of harmful essentialization to a magnified scale, further separating what is personal and contextual from the material of memory. People, and their histories, are absent from archives in the aggregate, or misrepresented in the aggregate, whether they want to be or not. In his 2019 keynote speech at the ARL-CNI Fall Forum in Washington D.C., computer scientist and experimental artist Jaron Lanier described the library as a "keeper of context." In a world of increasing corporate surveillance culture, he explained, "the library has a role to play in facilitating personhood" [9]. If datafication is always reductive, we will always need archives that are rooted in personhood and ethical engagement with the people being represented in these records. As Bowker states, "What we need is a strongly humanistic approach to analyze the forms that data take" [10].

The aspiration of providing universal access to all human knowledge both necessarily and symbolically annihilates community archives as a function of the Western/colonial enterprise. A social justice imperative in archives – elevating the voices and expertise of those (mis)represented, maligned, or otherwise marginalized in the historical record – does not serve the purposes of objective, structured, all-knowing, complete data. In this way, the pluralized archival paradigm and the centering of representational belonging become relegated to the margins of archivy, irreconcilable with the essentialization required of structured, computational archives. We must always be willing to release the archival debt levied upon those of us who steward the material of memory [11]. Within this context, we have a responsibility to ensure that computational archival studies continuously values community-centeredness, and look to this field for ethical computational solutions that reflect the needs and autonomy of people represented in archives.

III. COMMUNITY-CENTERED ARCHIVES PRACTICE MODEL

In a multi-year initiative funded by the Andrew W. Mellon Foundation, the practitioners and stakeholders at the University of California, Irvine Libraries seek to cultivate commitment among higher education institutions to sustained, critical engagement with community archives. Institutions that are committed to community-centeredness must also support ethical and responsible, large-scale representation of marginalized histories in digital collections. An ethics of care and responsible stewardship centers the perspectives of those who have been harmed by institutions. Within the project, we are conducting an assessment to identify actionable strategies for large-scale digital collections and seek to elevate the expressed needs of community archives in relation to digital collections.

Community-centered archives practice (C-CAP) is a model of archival work based on collaboration with community stakeholders, critical interrogation of our own collections and knowledge gaps, shared authority with communities, and flexibility in how we work with diverse groups. It consists of teaching, archival practice, curatorial work, outreach, and everyday action that privileges memory keeping and making in the service of self-representation. Supporting creation, preservation, and access to community-centered archives is a

central part of our effort to do this work while acknowledging our positionality within an academic institution.

Community-centered archives come into being through collaborative partnerships between mainstream archival institutions and communities that are underrepresented in the historical record. The goal is to empower communities, including community-based organizations that would not define themselves as community archives, in the process of telling and preserving their own histories. In a community-centered archives partnership, the principles espoused by academic institutions include:

- Attentiveness to inequities reflected in archives: An institution should seek to understand how communities have been misrepresented, absent, or maligned in historical documentation.

- Responsiveness to the community's needs: An institution must be flexible, adaptable, and take an iterative and ethical approach to responding to how community memory and evidence is preserved, described, and made accessible. This means being willing to bend and stretch how archival work is defined to reflect what matters to the community.

- Collaboration through shared authority: In a community-centered approach, the institution focuses on shared authority, making decisions together and respecting the value, expertise, and perspective brought to the partnership by the community.

- Awareness of the divergent priorities of communities: Community-institution partnerships must vary depending on the needs of each community, from the level of involvement by specific contributors to decisions about what archival material to collect.

Our work is to cultivate understanding among higher education institutions on the value of community-centered archives approaches as a sustainable and mutually-beneficial pillar of library, archives, and museum work. In this way, the project seeks to simultaneously solidify the ability and role of academic libraries to critically engage and contribute to social justice focused scholarship, training, pedagogy, and partnerships. The transformation of systems that lock out marginalized communities in the US is inextricably linked to the ability of students, researchers, and the public to find trustworthy sources, seek out and build community archives, and comprehend a fuller, more inclusive history. We have seen, and our research illustrates, the transformative impact of ethnic studies, social justice, and community archives as part of critical engagement with the stuff of history [12]. Through coalition-building with a network of stakeholders we are working to establish extensible community-centered archives models supporting ethnic studies and social justice education in the United States – particularly in a digital collections environment.

We envision a future where community-centered principles are always incorporated into the values and principles of CAS. The Advanced Information Collaboratory's recent project in Asheville, North Carolina, "Measuring the Impact of Urban Renewal," provides an example of how digital mapping,

historical records, and community advocates can coalesce in an ethical way, in service of repairing historical harm and trauma. In California, community-based archives and community-based organizations have curated their own historical narratives while tapping open-source tools and implementing concrete reciprocal partnerships with institutional collaborators, when necessary. In Santa Ana, California, this looked like a research solidarity project using historical maps and aerial photographs, led by environmental activists and informed by health data, lead concentration data, and archival records from institutions in the Orange County region, documenting the historical and continued impact of toxic chemical exposure [13]. An ethical, community-centered approach means yielding to the knowledge of communities as well as prioritizing the intended uses of archival data that best serve the articulated desires of these communities: crowdsourcing that is focused on those whose lives and identities are impacted by archival computation. In this way, we recognize and appreciate the potential for CAS to bring reparative practices to the forefront of digital archival work.

IV. MORE INCLUSIVE DIGITAL AGGREGATIONS

A major component of the ongoing C-CAP initiative is a broad research assessment that seeks to understand the ways regional and national digital collection aggregators can work towards more representative and inclusive aggregation, as well as support the development of a more responsible and inclusive framework for digital exhibitions. This assessment was developed collaboratively between the University of California, Irvine Libraries and the California Digital Library, which stewards the Calisphere (https://calisphere.org/) digital aggregation platform.

The C-CAP research assessment complements and extends recent National Historical Publications and Records Commission (NHPRC)-funded research activities conducted by Shift Collective and LYRASIS to assess the needs of small and diverse cultural memory organizations -- as well as other initiatives aimed at creating more inclusive aggregations, such as the Institute of Museum and Library Services (IMLS)-funded "Building a National Finding Aid Network" project.

Formally beginning in 2023, the research assessment is being conducted by metadata scholar Sharon Mizota, whose activities include analyzing the composition of cultural heritage organizations represented in digital aggregations and identifying gaps in representation, with an initial focus on evaluating the network of contributor hubs to the Digital Public Library of America (DPLA) through an environmental scan and a combination of surveys and interviews with community-centered organizations and DPLA hubs.

As of September 2023, Mizota has completed a preliminary environmental scan that surfaces potential barriers for participation in digital aggregation by smaller community-centered organizations. Mizota conducted a digital collections policy scan, reviewing policy and contributor participation documentation created by hub organizations that collect and push digital content to the Digital Public Library of America. The early findings from this analysis suggest that participation in digital aggregation with the Digital Public Library of America requires a level of technical and metadata expertise that may exclude smaller, less-resourced institutions. For example, of the

DPLA hubs included in the study, 79% (26) require contributors to comply with a minimum standard for metadata, 73% (24) require geographical constraints on content or institution, and 67% (22) require contributors to be institutions or organizations, not families or individuals [14].
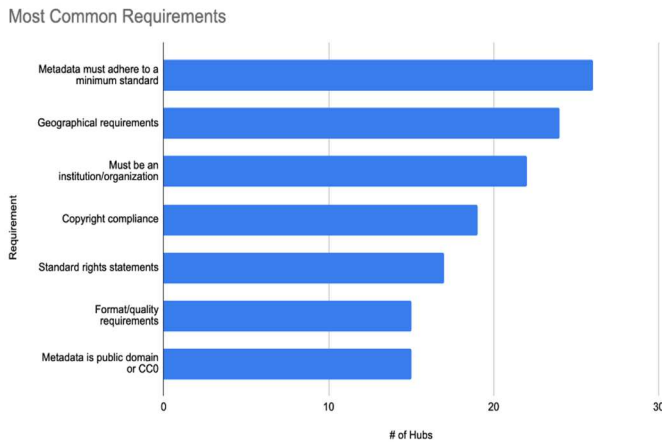

Fig. 1. Chart demonstrating participation requirements specified by at least 45% of DPLA Hubs

The research assessment will culminate with a public "Guide to Collaboration" aimed at an audience of both traditional institutions and community organizations recommending strategies and action plans that support the needs of community-centered archives and other organizations, while mitigating barriers to participating in aggregations. The "Guide to Collaboration" will also include a shareable toolkit that can be readily adapted by practitioners, outlining practical approaches and ethical considerations with developing digital exhibitions, and summarizing responsible practices for presenting research resources within the content of those exhibitions.

V. CONCLUSION

The need for critical, community-centered CAS cannot be overstated. Archivists steward historical records, representations of memory, and material culture that are by their very nature evidence of our humanity. The vastness of the digital record and, by association, archival data, beckons us to find a way not only to categorize, but to distill, the historical nuance of the human record in a package that can be read by machines. Due to the scale of archival data, digital records foster a false sense of documentary comprehensiveness. While on the precipice of significant expansion in the power of generative artificial intelligence and language learning models, we must resist the pull of essentialization, marginalization, and erasure and be willing to incorporate principles of community-centeredness into CAS. Our initial research findings, as well as the work of community archives, provide evidence of the need to elevate and continuously center the voices of those depicted (or underrepresented) in large-scale digital archives, leveraging the power of computational thinking with the transformative experience of seeing oneself represented (or representing oneself) in digital collections.

REFERENCES

[1] Orange County Business Council, First 5 Orange County, Orange County United Way, CalOptima and Orange County Community Foundation, 2020–21 Orange County Community Indicators, 2020, https://www.ocbc.org/wp-content/uploads/2020/09/2020-Community-Indicators-Report.pdf

[2] D. Mordell, "Critical questions for archives as (big) data," *Archivaria,* 87, pp. 140-161, 2019, https://www.muse.jhu.edu/article/724731.

[3] Always Already Computational - Collections as Data, "The Santa Barbara statement on collections as data - version 2," 2023, https://collectionsasdata.github.io/statement/.

[4] A. Eagle Yun, "Archives are people: love, hope, and courage for our future," Society of California Archivists Keynote, 2022, https://escholarship.org/uc/item/37z011gc.

[5] S. Nobel, Algorithms of Oppression: How Searchc Engines Reinforce Racism. New York, NY: University Press, 2018.

[6] A. Eagle Yun, J. Zavala,, K. Tribbett, T. Vo Dang. "Transforming Knowledge, Transforming Libraries - Researching the Intersections of Ethnic Studies and Community Archives: Final Report" 2020, https://escholarship.org/uc/item/47c2h0dd.

[7] M. Caswell, "Seeing yourself in history: community archives in the fight against symbolic annihilation," *The Public Historian*, 36:4, November pp. 26-37, 2014.

[8] M. Caswell, M. Cifor, M.H. Ramirez, "To suddenly discover yourself existing: uncovering the affective impact of community archives," *The American Archivist,* 79 (1), pp. 56–81, 2016, https://doi.org/10.17723/0360-9081.79.1.56.

[9] M. Dickerson, "ARL-CNI Fall Forum 2019 explores research libraries as catalytic leaders," 2020, https://www.arl.org/arl-cni-fall-forum-2019-explores-research-libraries-as-catalytic-leaders/.

[10] G. Bowker, "Data flakes: an afterword to 'raw data' is an oxymoron," in "Raw Data" is an Oxymoron, L. Gitelman, Ed. Cambridge, MA: The MIT Press, 2013, pp. 167–172.

[11] J. Cuellar, A. Eagle Yun, J. Meehan, J. Tai, "Defining archival debt: building new futures for archives," *Journal of Contemporary Archival Studies*, 10 (8), 2023, https://elischolar.library.yale.edu/jcas/vol10/iss1/8.

[12] A. Eagle Yun, K. Tribbett, T. Vo Dang, J. Zavala, "Fostering transformation: ethnic studies as critical intervention for primary source pedagogy, in Ethnic Studies in Academic and Research Libraries," R. Pun, M. Cardenas-Dow, K.S. Flash, Eds. Chicago, IL: Association of College and Research Libraries, 2021, pp. 17-26, https://escholarship.org/uc/item/1qs4m355.

[13] Santa Ana, General Plan, Orange County Environmental Justice, 2018, https://www.santa-ana.org/orange-county-environmental-justice/.

[14] S. Mizota, "C-CAP digital collections & digital exhibitions environmental scan," 2023, https://escholarship.org/uc/item/3xb9b3tr.