Dance History and Digital Humanities Meet at the Archives: An Interim Project Report on Dunham’s Data

*This interim project report addresses the ongoing work of Dunham’s Data: Katherine Dunham and Digital Methods for Dance Historical Inquiry. The project centres choreographer Katherine Dunham’s transnational circulation, and takes a critical mixed methods approach informed by feminist and anti-racist discussions in the digital humanities in order to explore the questions and problems that make data analysis and visualization meaningful for dance history. Dunham’s Data sits on robust datasets that we have manually curated from currently undigitized sources— an iterative and evaluative process that approaches these archives and the histories that they contain from a granular perspective. This update contextualizes our particular conjunction of archival and digital methods within dance history’s precedents for curating data, and talks through our own datasets as tools for dance historical analysis in terms of Dunham’s global legacy.*

Keywords: Katherine Dunham, digital methods, dance history, data analysis, archives, manual curation, dance data

This interim project report addresses the ongoing work of *Dunham’s Data: Katherine Dunham and Digital Methods for Dance Historical Inquiry*.[[1]](#footnote-1) The project centres choreographer Katherine Dunham’s transnational circulation, and takes a critical mixed methods approach to explore the questions and problems that make data analysis and visualization meaningful for dance history. *Dunham’s Data* sits on robust datasets that we have compiled from her currently undigitized archival materials at Southern Illinois University, the Missouri Historical Society, and the Library of Congress, among other sources. Drawing almost entirely from materials that Dunham herself chose to save, these datasets document Dunham’s daily whereabouts, the works in her repertory, and the people she employed. At the same time as Dunham’s archives are substantial, every datapoint requires evaluation, with discrepancies and competing narratives offering further insight into Dunham’s history.[[2]](#footnote-2) This methodology is particularly informed by feminist and anti-racist discussions in the digital humanities.[[3]](#footnote-3) At present, our data represents over 5000 days from 1947-60 that Dunham spent in hundreds of cities, almost 200 performers who traveled with Dunham, and over 200 pieces they performed during that time.[[4]](#footnote-4) In this update, we contextualize our particular conjunction of archival and digital methods within dance history’s precedents for curating data, and talk through our own datasets as tools for dance historical analysis in terms of Dunham’s global legacy.

Even without recourse to digital methods, dance history has established precedents for curating data as part of its analytic practices, from performance timelines to appendices of works. As early as the 1960s, Christena Schlundt began chronicling the performances of Ruth St. Denis and Ted Shawn, Helen Tamiris, and Daniel Nagrin. These were based on her archival research into performance dates and locations as well as repertory performed.[[5]](#footnote-5) In recognition of the Ballets Russes’s centenary in 2009 and as part of a 2010 exhibition at the Victoria and Albert Museum, Jane Pritchard compiled an exhaustive list of performances given by the company from 1909-1929, which, like Schlundt, includes venues and repertory. Pritchard expanded archival work previously done by Roland John Wiley in 1971 and provided additional historicization.[[6]](#footnote-6) In 2004, Stephanie Jordan and Larraine Nicholas built a database ‘Stravinsky the Global Dancer’ that catalogued the premiere dates of 1252 dances choreographed by 703 choreographers to the composer’s music. Whereas Schlundt’s data was primarily a reference tool, Jordan developed scholarly arguments on the basis of examples and trends found in this database, including the ways Stravinsky might be reconceived historically beyond the Ballets Russes and Balanchine.[[7]](#footnote-7) *Dunham’s Data* builds on these analogue and early digital methodological precursors for tracking historical itineraries. Working in the context of a computational framework — where dance archives meet the newer technological practices and methods of digital humanities — both prompts precision at a different scale of granularity, and offers a means to manipulate large quantities of information toward complex analysis.

Such research datasets stand in abstracted form for aspects of archival materials at the same time as they form a secondary archive that combines data and metadata. Although Dunham was a rigorous self-archivist, any individual item in her archive is not necessarily reliable in itself, not only because of the fallibility of archival materials more broadly, but also because the precarity of Dunham’s financial circumstances, which resulted in regular last-minute changes to the company structure and touring itinerary.[[8]](#footnote-8) In manually curating data, we routinely cross-reference and reconcile information including from Dunham’s personal and professional correspondence, contracts, company documents from receipt books and payroll to costume lists, personal logs, programs, scrapbooks, lighting plots, and newspaper clippings. To these we have added supplemental data sources, such as immigration records, local newspapers, and historical transportation maps and schedules. Within *Dunham’s Data*, we affectionately refer to this process as ‘decadating’ because ‘triangulating’ is insufficiently rigorous; rather, we joke that we need ten or more documents to be confident about any single datapoint due to the high level of conflicting information and persistent changes of plan. We leave traces of this process in our datasets, including sources and annotations.

Key dates in Dunham’s historical record have already been corrected by scholars through the close analysis of archival materials, including Das who redates Dunham’s controversial *Southland* and also argues that she spent far less time that Caribbean on her anthropological travels than previously assumed.[[9]](#footnote-9) However, the systematic nature of datasets offer a means to methodically consider moments and events that may prove to be important but have not been thoroughly documented. For example, scholars have approached us regarding Dunham’s itinerary, for which we have been able to corroborate, add to, and correct dates of travel that have a bearing on those scholars’ interpretations, and ultimately contribute to a broader conversation regarding the transnational flows of Dunham’s company.[[10]](#footnote-10) Likewise, we are still in the process of building a dataset of Dunham’s repertory, but already note complexities in both naming and dating that can have ramifications for the ways scholars understand her oeuvre. These datasets are not thus only research tools, but are also research outputs in themselves that depend on a deep engagement with and interpretation of archival materials — to borrow from the HEFCE definition: ‘a process of investigation leading to new insights, effectively shared.’[[11]](#footnote-11)

The first dataset we began and continue to build out is Dunham’s daily itinerary--where she was every single day for approximately thirty years of her performing career, of which we currently have 1947-60 documented. Although scholars have tracked Dunham’s extensive travels on an ad hoc basis, our workrepresents the first attempt to build a comprehensive dataset of her whereabouts. We initially thought this Everyday Itinerary dataset would serve to help us understand the relationship between Dunham’s choreographic and ethnographic work, but soon realized that this dataset underpins the entire *Dunham’s Data* project. In addition to geographic location, we also track, albeit less comprehensively, the accommodation in which Dunham stayed each night; the theatres, nightclubs, television studios, and other places she and the company performed; the modes of transportation used when travel occurred; additional transit cities through which she passed; and whether or not Dunham was likely to be in rehearsals or giving public performances. Between 1950-53 alone, Dunham passed through at least 84 cities over 129 trips across five continents. Accounting for all of the many cities and venues through which Dunham travelled, including from a statistical perspective, rebalances the geography of Dunham scholarship, which emphasizes a few key sites at the expense of the many contexts that deeply informed both the repertory she choreographed, as well as the spread of her influence.

Whereas the daily itinerary accounts for Dunham’s own travels, the Check-In dataset accounts for the comings and goings of her dancers, drummers, and singers over time, helping to discern who among almost two hundred performers were in the studio and theatre together over the fourteen years from 1947 to 1960.[[12]](#footnote-12) As with the daily itinerary, data on check-ins come from scattered sources. We classify these as two types of check-ins: comprehensive and non-comprehensive, based on the thoroughness of data available. For example, a list of hotel accommodations in a particular city might give us an overview of the dancers, drummers, and singers with the company at that moment, whereas a note stating that a performer missed his entrance indicates that he was present that night, but says nothing of other artists. Whereas more conventional models of accounting for employment might involve start and end dates, Dunham’s almost-continual touring for years on end, together with the ongoing reconfiguring of her company for different performance contexts, meant that we had to develop a data structure more suited to marking the evidence we have for artists present at a given moment, and then using digital tools in order to draw inferences that bring these independent moments together, which we do in many ways, including visualized as an interactive flow diagram.[[13]](#footnote-13) While the Check-In dataset is developed in tandem with information from the everyday dataset, and vice versa, due to information available, there is a greater level of ambiguity in the check-in dataset and many dates are approximated in order to achieve chronology. With the Check-In dataset, we are able to trace potential lines of transmission of embodied knowledge within and beyond the Dunham company.

Our newest dataset concerns Dunham’s repertory, and catalogues the various titles and descriptions by which a piece might be known, the years in which it was performed, and all of the singers, dancers, and drummers who are ever listed as performing in it. We have also documented other aspects of each work such as composers of the music, the varying numbers of performers, and whether pieces were performed in concert venues, nightclubs, or both. We further examined program notes that indicate real or imagined connections between different pieces and different places or historical time periods. The process of building out the Repertory dataset drew our attention to the various scales at which Dunham repurposed choreographic elements, and therefore the alternative ways that works might connect individual performers. To track these fluid relationships among over 200 identified works in Dunham’s repertory, we began to categorize them in three ways. Primarily drawn from African American vernacular and social dance forms, “dances in dances” are often used as building blocks of larger pieces. Then there are named “pieces” that may stand alone, or be absorbed into one or more larger works. Finally, there are “containers” of different sizes and structural integrity, that draw together multiple pieces, whether for less than an act or an evening-length show.[[14]](#footnote-14) As with so much of *Dunham’s Data*, the manual curation of the repertory dataset changed the ways we understood the ontology of Dunham’s choreography and how it circulated within the company and throughout the world.

While we have so far described these datasets as independent entities, both the curation of these datasets and scholarly arguments based on them develop in tandem. Many of the problems that we are trying to solve involve reference to more than one dataset at a time. From a data curation standpoint, we often have undated programs in various languages but if we know when certain performers appeared, and also the dates in which the company was in particular countries that spoke those languages, then we can begin to fill in further information regarding performers, repertory, and itinerary. Conducting scholarly inquiry in conjunction with such datasets already scales to encompasses micro and macro perspectives — the granularity of data as well as the general trends when seen in their entirety from a distance.[[15]](#footnote-15) Joining datasets enables us to ask questions at even greater scales. For example, we know anecdotally that Dunham tended to travel around the world and pick up performers as she went, but by bringing together our datasets on daily Itinerary, check-ins, and performer attributes, we begin to provide evidence of just how systematic this was.[[16]](#footnote-16) By extension, this indexes the kinds of cultural knowledge that performers brought to Dunham’s diasporic practice.

Our first phase of research involved building out the 1950-53 itinerary as a proof of concept for working this way in Dunham’s archives, and also as a reference dataset for further work, for which we have 93% of the days. In the process, we developed a critical mixed methods approach in which we articulated the ‘everyday’ as an intersectional analytic through which to understand the global method that ties Dunham’s transnational mobility to her practices of making-do as an African American female artist in the mid-twentieth century.[[17]](#footnote-17) Among other visualizations we produced during this phase, the 1950-53 itinerary can be explored in an interactive online timeline.[[18]](#footnote-18) In the second phase, we expanded the itinerary to 1947-60 and also built out the Check-In, Personnel Attributes, and Repertory datasets for that time. Together with these datasets, we began to ask questions about the nature of embodied transmission within Dunham’s company as a dynamic movement community that traveled through the world; as new performers joined, they brought diasporic knowledge with them into the company, and touring Dunham’s repertory disseminated diasporic representations outward. In this phase, we further experimented with types of visualizations to not only develop inquiry through this data, but also to serve as visual arguments that themselves convey a sense of motion.

Looking ahead, our work takes many forms. First, is the further development of datasets and visualizations, and with them scholarly arguments, in particular related to the questions with which we began, regarding the relationships between choreography and ethnography in Dunham’s body of work. Second, we conceived of this project as always occurring in dialogue with those who hold deep knowledge of Dunham’s history, practices, and legacy. The team has already begun to conduct oral history interviews with surviving company members who appear in our data, including showing them our ongoing work. We will develop this polyvocality further by commissioning a series of essays from expert users to reflect on the work to date. Third is ensuring the accessibility and sustainability of this research; it is important to us that the datasets, static and interactive visualizations, code, and other outputs have multiple means through which they can be accessed and referenced in the future. We envision this as manifold, rather than lodged in a singular repository; for example because our datasets link together the resources we have consulted across multiple archives, some of the archivists have expressed interest in holding copies to benefit other Dunham researchers. Finally there is the onward dissemination of the knowledge developed by *Dunham’s Data*. In addition to digital and print publications, we are also working toward a teaching toolkit with a series of lesson plans so that dance history teachers can better represent Dunham in their courses and contribute to both archival and digital literacies among students.

1. Writing and datasets are equally co-authored by Kate Elswit and Harmony Bench; name order is alphabetical. *Dunham’s Data* is supported by a three-year research project grant from the UK Arts and Humanities Research Council (AH/R012989/1). Postdoctoral research assistants on the project are Takiyah Nur Amin, Antonio Jimenez-Mavillard, and Tia-Monique Uzor. [↑](#footnote-ref-1)
2. On the idea of data as story, see Harmony Bench and Kate Elswit, ‘Katherine Dunham’s Global Method and the Embodied Politics of Dance’s Everyday’, *Theatre Survey* vol. 61 no. 3 (forthcoming 2020). [↑](#footnote-ref-2)
3. Including [Catherine D’Ignazio and Lauren Klein, *Data Feminism* (Cambridge: MIT Press, 2020)](https://www.zotero.org/google-docs/?HiunfH); Jessica Marie Johnson, ‘Markup Bodies: Black [Life] Studies and Slavery [Death] Studies at the Digital Crossroads’, *Social Text* vol. 36, no. 4 (2018); Marisa Parham, ‘Digital Archives, Datum Storytelling, & the Future of Memory,’ https://www.youtube.com/watch?v=bxWv0pv1l40 (accessed 10 September 2018). [↑](#footnote-ref-3)
4. For video demos of preliminary visualizations, see Harmony Bench and Kate Elswit, ‘Interactive Visualizations: Work in Progress Videos,’ *Dunham’s Data Research Blog*, https://www.dunhamsdata.org/blog/work-in-progress-videos-1, 27 March 2020. [↑](#footnote-ref-4)
5. See Christena L. Schlundt, *The Professional Appearances of Ruth St. Denis & Ted Shawn; A Chronology and An Index of Dances, 1906-1932* (New York: New York Public Library, 1962); *The Professional Appearances of Ted Shawn and His Men Dancers: A Chronology and an Index of Dances, 1933-1940* (New York: New York Public Library, 1967); *Tamiris: A Chronicle of Her Dance Career, 1927-1955* (New York: New York Public Library, 1972); *Daniel Nagrin: A Chronicle of his Professional Career* (Berkeley: University of California Press, 1997). [↑](#footnote-ref-5)
6. Jane Pritchard, ‘Serge Diaghilev’s Ballets Russes – An Itinerary. Part 1: 1909–1921’, *Dance Research: The Journal of the Society for Dance Research* vol. 27 no. 1 (2009): 109–98; ‘Serge Diaghilev's Ballets Russes-An Itinerary. Part II (1922-9)’, *Dance Research: The Journal of the Society for Dance Research* vol. 27 no. 2 (2009): 254-360. [↑](#footnote-ref-6)
7. Stephanie Jordan and Larraine Nichols, ‘Stravinsky the Global Dancer: A Chronology of Choreography to the Music of Igor Stravinsky’, dataset and website, http://urweb.roehampton.ac.uk/stravinsky/, 2003; Stephanie Jordan, ‘The Demons in a Database: Interrogating “Stravinsky the Global Dancer”’, *Dance Research* vol. 22 no. 1 (2004): 57–83. [↑](#footnote-ref-7)
8. We elaborate this argument in Bench and Elswit (2020). [↑](#footnote-ref-8)
9. Joanna Dee Das, *Katherine Dunham: Dance and the African Diaspora* (New York: Oxford University Press, 2017). [↑](#footnote-ref-9)
10. See Harmony Bench and Kate Elswit, ‘Datasets are Research’, *Dunham’s Data Research Blog*, <https://www.dunhamsdata.org/blog/datasets-are-research>, 5 September 2019. [↑](#footnote-ref-10)
11. UK Research and Innovation, ‘REF 2021 Guidance on Submissions (2019/01)’, January 2019, 90. [↑](#footnote-ref-11)
12. The check-in dataset sits in relation to a second dataset in which we track Performer Attributes, from alternative names to geographical affinities. [↑](#footnote-ref-12)
13. Video demos are linked from footnote 4. [↑](#footnote-ref-13)
14. Video demos are linked from footnote 4. [↑](#footnote-ref-14)
15. On scale, see Harmony Bench and Kate Elswit, ‘Mapping Movement on the Move: Dance Touring and Digital Methods,’ *Theatre Journal* vol. 68 no. 4 (2016): 575-596. [↑](#footnote-ref-15)
16. See Harmony Bench and Kate Elswit ‘Visceral Data for Dance Histories: Katherine Dunham’s Pieces, Places, and Pieces, 1947-60’. Unpublished Manuscript. [↑](#footnote-ref-16)
17. Bench and Elswit 2020. [↑](#footnote-ref-17)
18. <https://dunhamsdata.org/1950-53timeline/> [↑](#footnote-ref-18)