

UCLA

UCLA Previously Published Works

Title

Editor's Introduction: Data as Media

Permalink

<https://escholarship.org/uc/item/3ng7f0rf>

Journal

Feminist Media Histories, 3(3)

ISSN

2373-7492

Authors

Posner, Miriam

Klein, Lauren

Publication Date

2017-04-01

Peer reviewed

Editor's Introduction

Data as Media

Accustomed though we are to films, files, moving-image apparatus, and books, media scholars often still find the concept of “data” jarring. As evidenced by any number of scenes of what might be called data montage—think of the reference library sleuthing in *Spotlight* (2015) or the now-iconic scrolling green type in *The Matrix* (1999)—we struggle to picture data, let alone begin to analyze it as a media object. Even in information studies and data science, fields in which data is a primary object of study, specific definitions of the term are contested and often unsatisfying. Does information become data when it is converted from undifferentiated content to structured form? When it is translated from text or image into machine-readable form? Or when marshalled as evidence or proof of results?

The information scholar Christine Borgman has argued for the contingent and time-based nature of data: information becomes data not by virtue of any inherent property, but in the process of being transformed for use as evidence in support of a claim.¹ The historian Daniel Rosenberg takes this line of argument further. Data, he contends, is “*specifically* rhetorical”—the necessary prerequisite to any analysis or claim.² These formulations of data are helpful, but to the media scholar, they omit some important features. We may select and arrange a series of sequences from a film or television show in support of an argument, but we almost never refer to this evidence as “data”—a fact that suggests that, for media scholars, the concept of data carries more ontological specificity than the primarily instrumental construct that Borgman and Rosenberg describe. So to Borgman and Rosenberg’s definitions, we add two features of “data” that help distinguish it from other forms of mediatic evidence. First, for a source to be data, it must be made computationally tractable.³ By tractable we mean, following Willard McCarty, that information can be read, manipulated, and programmatically transformed by a computer (or by computer-like operations).⁴ Computability is often accomplished by categorization, and these two qualities, categorization and computability, distinguish data sets proper from

the sources that media scholars more commonly analyze and interpret. Of course, almost everything we encounter is categorized in some way (as a sitcom, for example, or a biopic), but the presumption that these categories are themselves meaningful, amenable to aggregation and measurement, and possessing some salience across a domain is the hallmark of computable data.⁵ Thus, data is as much an orientation toward one's sources as it is a primary category of knowledge.⁶

Media scholars, oriented to their own sources, are well aware that the specificity of any particular medium becomes most apparent when its contents are shifted to another medium. Such was the case when analog film gave way to digital cinema and scholars renewed their focus on, for example, the material basis of celluloid film, or the indexicality of the photographic image. Such is also the case now, as we find our objects of study, and our own consuming habits, enumerated, analyzed, and segmented as never before. Netflix, powered by machine learning algorithms, offers not just drama and comedy, but “Cerebral Business Documentaries” and “Scary Cult Movies from the 1980s.”⁷ Social media platforms and search engines triangulate our personal preferences with our demographic information to recommend particular TV shows or films. Although we may not encounter them as spreadsheets or statistics, the objects of media studies—and especially film—have already acquired shadow lives as data, just as our viewing and consuming habits have. And yet, while a number of thought-provoking studies have emerged that make use of this (or related) quantified data about media, the field of media studies has yet to fully grapple with data as a medium in and of itself.⁸

This is where the contributions of media studies to the larger field of data studies begin to emerge. With its emphasis on the interrelation of form and content, on the specificity and material bases of media forms, and on how both individuals and social groups shape meaning, media studies offers a set of sensitivities and concerns that have yet to be sufficiently addressed by the adjacent fields of information studies, science and technology studies, and the history of science that have largely constituted the field of data studies to date.⁹ Indeed, media studies scholars have long attended to the medium-specific features of their objects of study, and how those features impact an object's function and significance. By the same token, media scholars have—for almost as long—considered how any particular media form might be required to define its features, as Erika Balsom helpfully summarizes, “in relation to the aggregate mixtures it enters into with other media.”¹⁰ These debates about medium specificity allow us to begin to identify certain features that inhere in

and around data, such as information segmentation or scales of measurement, that impact its significance and use.¹¹ In addition, an emphasis on the medium-specific properties of data point us toward the development of analogous media-specific methods, such as the critique of a data set or the close reading of code—still very much in their infancy—that can help surface some of the ways in which data structures our lives.¹²

Defining data through the lens of media studies also brings to bear a specific accounting of data's material properties. Materiality, as formulated through media studies, entails an attention to the specific conditions in which media forms emerged and are articulated, whether via photochemical film, VHS tape, or vinyl records. As a field, media studies has been alert to the ways in which the material conditions of a format (its sensitivity to heat and light, for example, or the speed with which it can convey information) shape its capacity and potential. Media scholars such as Lisa Gitelman and Jonathan Sterne have enriched and even transformed our understanding of contemporary digital file formats by showing that their apparent novelty is in fact rooted in much older media of communication.¹³ We are only beginning to understand how the specific affordances of data storage and retrieval technologies—the spreadsheet, the relational database, the graph database—might shape the means by which humans interact with information more intimately than they might think.¹⁴

Finally, a media studies approach to data entails an attention to the histories, cultures, and contexts that gave rise to it. As we see in many of the essays in this issue, data sets never arrive in the world fully formed, but are assembled from tangles of historical forces and ideological motivations, as well as practical concerns. For this reason, data must be analyzed through a sufficiently robust critical frame, one that allows for the actors and agents associated with the development of any particular data object to be fully acknowledged and explored. Media studies, which itself adapts traditional humanistic techniques such as close reading and historical synthesis to the medium-specific nature of its analyses, thus provides a model of how technical explanations might be buttressed by the cultural and critical analyses that are the hallmark of humanities scholarship.

In situating this special issue of *Feminist Media Histories*, we seek to model how one set of humanistic approaches, rooted in feminist theory, might be applied to enrich the field of data studies. Since the 1970s, feminist approaches have evolved to examine much more than representation: they seek to lay bare the ways in which often-unacknowledged forces structure our experiences of gender, knowledge, and the world.¹⁵ This has obvious applicability for the study

of data, a mode of information that depends so heavily on categorization, and the collection of which so often represents aggregations of power and capital; indeed, feminist theory is well positioned to challenge repressive systems of classification, and to expose how choices in what and how to categorize carry consequences greater than order alone.¹⁶ Feminist theory also emphasizes the importance of subject position in determining one's evaluation of truth, especially when examining purportedly neutral objects—another key consideration when encountering the various structures associated with data, such as objects or tables, that are often presumed to serve as mere containers.¹⁷ From feminist scholarship, too, comes a deep concern with labor, especially the modes of labor associated with women—such as transcription, layout, data collection, or tabulation—that have traditionally been overlooked or undervalued by the market.

Which is not to say that feminism in 2017 is uncomplicated, or uncontested. As scholars and thinkers from bell hooks to Angela Davis to Sara Ahmed have pointed out, multiple strains of feminism have enshrined middle-class white women's specific experiences as universal "womanhood."¹⁸ We do not all share the same disadvantages on the labor market, for example, nor do we experience the same assumptions about our character, or the same encounters with the apparatus of the state, or even the same biology. Thus a meaningful approach to feminism must be, at the least, intersectional.¹⁹ If it is to be useful to the field of data studies, a feminist approach must recognize that different people experience multiple, overlaying identities, advantages (or disadvantages), privileges, and outlooks. Feminist theory is also explicitly not just for people who identify as women; indeed, one of the most powerful facets of feminist theory is its ability to dismantle cut-and-dried divisions between "male" and "female" in favor of multiple, pluralistic categories (or the abandonment of categories altogether).

As the essays and projects in this issue make clear, there is also much to be gained from the study of data from the point of view of feminist media *history*. The media objects under consideration here range from online security questions to demographic data to border surveillance protocols. From context-specific excavations of data sets, we learn that our received wisdom about social phenomena is powerfully inflected by assumptions about race, class, and gender. In Shawn Shimpach's "Only in this way is social progress possible': Early Cinema, Gender, and the Social Survey Movement," the notion of "the audience" emerges as the configuration of a very particular set of assumptions about gender, demography, spectatorship, and data. Carole R. McCann, in "Figuring the Population Explosion: Demography in the Mid-Twentieth

Century,” shows us that the “population boom” at midcentury depended on highly gendered and raced ideas about fertility and appropriate behavior.

Once data exists in the world, it takes on remarkable power to shape identities and perceptions. We see in Juan Llamas-Rodriguez’s “The Datalogical Drug Mule” how border surveillance units have assembled a spectral body from predictive data about race, gender, and demography. Bonnie Ruberg’s “What Is Your Mother’s Maiden Name?: A Feminist History of Online Security Questions” offers a surprising, feminist history of the security question, that most mundane of digital challenges. Michael Eng shows how the artist Adrian Piper challenges some of our basic assumptions about data, including our ability to discern or define race, in “Lights! Race! Gender! Adrian Piper and the Pseudorationality of Data.” Natalie Wreyford and Shelley Cobb take on the question of how a feminist researcher should work responsibly with quantitative data about women’s lives in “Data and Responsibility: Toward a Feminist Methodology for Producing Historical Data on Women in the Contemporary UK Film Industry.”

Data’s underexplored textures, as well as our own assumptions about what data should be or do, are explored in several of the digital projects included in this issue. Lauren F. Klein and her coauthors, in “The Shape of History: Elizabeth Palmer Peabody’s Feminist Visualization Work,” probe our expectations about data visualization, introducing us to the largely unknown data visualization pioneer Elizabeth Palmer Peabody and asking whether a feminist work of data visualization is possible. In “Retrieving from My Digital Body: A Map of Abuse and Solidarity,” Marta Delatte experiments with a standpoint-specific “warm data” archive by documenting episodes of sexual assault. Rachel Devorah, in “Overmorrow,” sonifies data about gun violence, pushing us to examine our assumptions about neutrality and affect in data visualizations. Shelly Eversley and Laurie Hurson reflect on the process of building the history of sex and gender equality in “Equality Archive: Open Educational Resources as Feminist Praxis.” Finally, Gabriela Aceves Sepúlveda makes a powerful argument for remediating works that have tended to escape the archive in “[Re]Activating Mamá Pina’s Cookbook,” a work that plumbs and visualizes her own family’s feminist history. Taken separately and together, these works demonstrate what a field of feminist data studies could soon become. ■

LAUREN F. KLEIN is an assistant professor in the School of Literature, Media, and Communication at the Georgia Institute of Technology, where she also directs the Digital Humanities Lab.

MIRIAM POSNER is coordinator and core faculty for the Digital Humanities program at the University of California, Los Angeles.

NOTES

1. Christine L. Borgman, *Big Data, Little Data, No Data: Scholarship in the Networked World* (Cambridge, MA: MIT Press, 2015), 28–29.

2. Daniel Rosenberg, “Data before the Fact,” in “*Raw Data*” *Is an Oxymoron*, ed. Lisa Gitelman (Cambridge, MA: MIT Press, 2013), 15–40, 15.

3. This is not to say that evidence has to be literally entered into a computer to become data; information recorded on paper would do just as well, but it must be structured and arranged so that it can be programmatically transformed via regular operations. See Mary Poovey, *A History of the Modern Fact: Problems of Knowledge in the Sciences of Wealth and Society* (Chicago: University of Chicago Press, 1998), for examples of the transformation of “knowledge” to computable data during a period that far predates the computer.

4. Willard McCarty, *Humanities Computing* (New York: Palgrave Macmillan, 2005).

5. For a discussion of the many implications of classification, computation and otherwise, see Geoffrey C. Bowker and Susan Leigh Star, *Sorting Things Out: Classification and Its Consequences* (Cambridge, MA: MIT Press, 1999).

6. For additional definitions and histories of data, see Johanna Drucker, “Humanities Approaches to Graphical Display,” *Digital Humanities Quarterly* 5, no. 1 (2011): <http://digitalhumanities.org/dhq/vol/5/1/000091/000091.html>; Ed Folsom, “Database as Genre: The Epic Transformation of Archives,” *PMLA* 122, no. 5 (2007): 1571–79; Jonathan Furner, “Data: The Data,” in *Information Cultures in the Digital Age*, ed. Matthew Kelly and Jared Bielby (Wiesbaden, Germany: Springer, 2016), 287–306; Gitelman, “*Raw Data*” *Is an Oxymoron*; Lisa Gitelman, *Always Already New: Media, History and the Data of Culture* (Cambridge, MA: MIT Press, 2006); Ian Hacking, *The Taming of Chance* (Cambridge, England: Cambridge University Press, 1990); Orit Halpern, *Beautiful Data: A History of Vision and Reason since 1945* (Durham, NC: Duke University Press, 2015); Poovey, *A History of the Modern Fact*; Yanni Alexander Loukissas, “Taking Big Data Apart: Local Readings of Composite Media Collections,” *Information, Communication and Society* 20, no. 5 (May 4, 2017): 651–64; Theodore Porter, *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life* (Princeton, NJ: Princeton University Press, 1995).

7. Alexis C. Madrigal, “How Netflix Reverse Engineered Hollywood,” *Atlantic*, January 2, 2014, <http://www.theatlantic.com/technology/archive/2014/01/how-netflix-reverse-engineered-hollywood/282679/>.

8. Indeed, a number of media studies scholars have turned to digital humanities methods in order to better understand the data of media. See for instance Deb Verhoeven’s ongoing analysis of the social networks embedded in the Australian film industry, as documented on the Kinomatics site; Eric Hoyt et al.’s Project Arclight, a data mining and visualization tool (and accompanying publication) centered on film and media history; Christian Keathley and Jason Mittell’s concept of videographic criticism, as described in *The Videographic Essay: Criticism in Sound and Image* (Montreal: Caboose, 2016); and Maria Kanatova et al.’s computational analysis of trends in cinematic temporality, as described in “Broken Time, Continued Evolution: Anachronies in Contemporary Films,” *Stanford Literary Lab Pamphlet 14* (Palo Alto, CA: Stanford Literary Lab, 2017). For critical writing on media studies and digital

humanities, see Jentery Sayers, ed., *The Routledge Companion to Media Studies and Digital Humanities* (New York: Routledge, 2017); and Eric Hoyt and Charles R. Acland, *The Arclight Guide to Media Studies and Digital Humanities* (Sussex, England: REFRAME, 2016).

9. Thus far, the field of data studies has been largely populated by scholars trained in science and technology studies (STS), and its journals, methods, and academic programs reflect that disciplinary and methodological slant. See for example danah boyd and Kate Crawford, “Critical Questions for Big Data: Provocations for a Cultural, Technological and Scholarly Phenomenon,” *Information, Communication and Society* 15, no. 5 (2012): 662–79; Andrew Iliadis and Federica Russo, “Critical Data Studies: An Introduction,” *Big Data and Society* 3, no. 2 (2016): 1–7.

10. Erika Balsom, *Exhibiting Cinema in Contemporary Art* (Amsterdam: Amsterdam University Press, 2013), 73. See the section “Post-Medium Post-Mortem” for a summary of the debates around medium specificity in cinema studies and art.

11. On measurement and segmentation, see for example Norton Wise, ed., *The Values of Precision* (Princeton, NJ: Princeton University Press, 1995); and Jean-François Blanchette’s forthcoming *Running on Bare Metal: Materiality and Modularity in the Computing Stacks* (University of Chicago Press).

12. See for instance Mark Sample’s “data analysis” assignment, developed for the course Data Culture at Davidson College, <http://www.samplereality.com/2014/04/30/dig-210-data-culture/>; and the work of the Humanities and Critical Code Studies Lab at the University of Southern California, <http://hacslab.com/>.

13. Lisa Gitelman, *Paper Knowledge: Toward a Media History of Documents* (Durham, NC, and London: Duke University Press, 2014); Jonathan Sterne, *MP3: The Meaning of a Format* (Durham, NC: Duke University Press, 2012). In a similar vein, Lisa Nakamura offers a wonderful excavation of the materiality of the semiconductor in “Indigenous Circuits: Navajo Women and the Racialization of Early Electronic Manufacture,” *American Quarterly* 66, no. 4 (December 15, 2014): 919–41.

14. On the knowledge effects of specific data formats, see for example Paul Dourish, “No SQL: The Shifting Materialities of Database Technology,” *Computational Culture* 4 (November 2014), n.p.; Alexander R. Galloway, *Protocol: How Control Exists after Decentralization* (Cambridge, MA: MIT Press, 2004); David Golumbia, *The Cultural Logic of Computation* (Cambridge, MA: Harvard University Press, 2009).

15. A chronicle of the history of feminist theory lies beyond the scope of this introduction. For an overview of the field and a sense of its reach, see Carole R. McCann and Seung-kyung Kim, eds., *Feminist Theory Reader: Local and Global Perspectives*, 3rd ed. (New York: Routledge, 2013).

16. Key feminist works that have challenged the salience of received categories include Judith Butler, *Gender Trouble: Feminism and the Subversion of Identity* (New York: Routledge, 1990); Donna Haraway, *Simians, Cyborgs, and Women: The Reinvention of Nature* (New York, Routledge, 1991); Karen Barad, *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning* (Durham, NC: Duke University Press, 2007); Catharina Landström, “Queering Feminist Technology Studies,” *Feminist Theory* 8, no. 1 (April 2007): 7–26; Mel Y. Chen, *Animacies:*

Biopolitics, Racial Mattering, and Queer Affect (Durham, NC: Duke University Press, 2012).

17. On feminism's relationship to objectivity see for example Donna Haraway, "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective," *Feminist Studies* 14, no. 3 (Autumn 1988): 575–99; Sandra Harding, *The Science Question in Feminism* (Ithaca, NY: Cornell University Press, 1986); Susan Bordo, *The Flight to Objectivity: Essays on Cartesianism and Culture* (Albany: State University of New York Press, 1987); Patricia Hill Collins, *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment* (Boston: Unwin Hyman, 1990).

18. bell hooks, *Feminism Is for Everybody: Passionate Politics* (Boston: South Press, 2000); Angela Davis, *Women, Race, and Class* (New York: Vintage, 1983); Sara Ahmed, *Living a Feminist Life* (Durham, NC: Duke University Press, 2017).

19. For an overview of intersectionality from its coinage to its contemporary applications, see Patricia Hill Collins and Sirma Bilge, *Intersectionality: Key Concepts* (New York: Polity, 2016).