



BREAKING **EPISTEMIC CULTURES** **OF DIGITAL HUMANITIES PROJECT**

The background is a vibrant green with large, faint blue circles. Two stylized white documents with blue horizontal lines are positioned diagonally. In the center-right, there are molecular structures with blue, orange, and pink spheres connected by purple lines, and two white, cloud-like shapes.

Author

Ayom Mratita Purbandani

Editor

Hafiz Noer

Design & Layouter

Muhammad Fanani Arifzqi



⇒ Introduction

Studies in the 'classic' humanities have traditionally privileged textual aspects as the starting point of analysis. Despite the humanities being an academic field with various methods and traditions, textual sources are argued to be essential for understanding closely how meaning is constructed and conveyed in the epistemic traditions of humanities research. At this point, the development of digital humanities (DH) in recent decades occupies a unique status in the development of humanities studies. The insertion of the word 'digital' in digital humanities is not just a prefix. Instead of being merely a set of machines, networks, and databases, 'the digital' also brings forth connections between various things that are not related to each other. Alexander Galloway calls this the wonder of the digital concept itself.¹

The simultaneous work between digital and data is predicted to produce an epistemic tradition that is no longer understood by humans, where human capability is no longer the primary criterion for epistemic success.² This highlights the digital as a deliberative and democratized space. The digital enables other freedoms that need to be considered: It opens up broad access for anyone to act as both knowledge producers and knowers.

Thus, digital technology is assumed to have noble aspirations towards knowledge. As a new arena in knowledge acquisition, digital technology brings about new ways of knowing and remembering various epistemic sources from the digital society obtained through the digital space itself. As a part of knowledge attainment, Lehigh University uses digital technology to host a repository of digital humanities projects, including databases, maps, and timelines.³

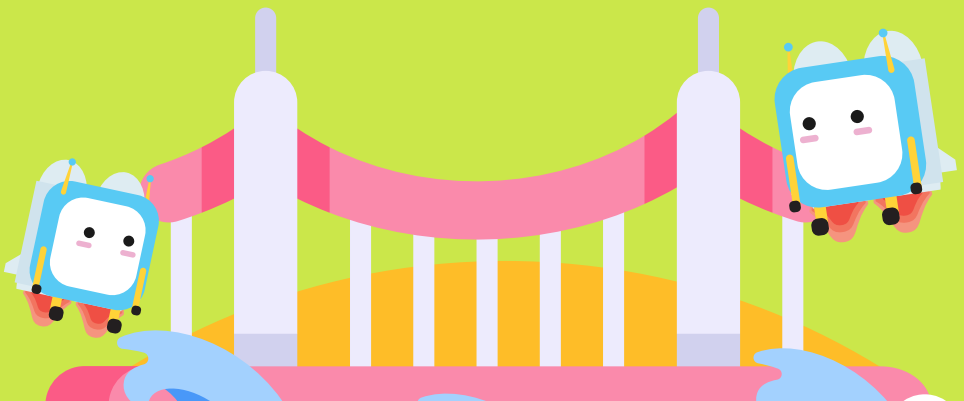
Since evolving far from contextual computing in the humanities, digital humanities has expanded its reach but remains connected to the goals that have animated it from the beginning: using information technology to clarify human records and bringing understanding of human records to bear on the development and use of information technology.

Several arguments about research practices in digital humanities, with a special focus on the role of data and methodology, are raised in this case study. In this writing, epistemic tradition refers to the framework and assumptions underlying how we understand knowledge, truth, and ways to obtain it.⁴ Reflecting on the development of studies in the digital humanities related to knowledge production, this case study does not intend to offer a 'conceptual' novelty, but to re-examine epistemological issues and clarify the complex epistemological assumptions of the epistemic tradition and research practices in digital humanities. The result is expected to provide a robust framework and epistemic basis for investigations in digital humanities.



➔ Bridging “two cultures”

At least until a decade ago, tensions between 'classic' humanities and digital humanities remained strong.⁵ Thinkers of 'classic' humanities still found it difficult not to be skeptical of what the digital offers in digital humanities. "The Two Cultures," a famous cliché metaphor coined by Charles Percy Snow in 1959 to describe the dichotomy between science and humanities, still represents the structure of intellectual debate in recent decades.⁶ Snow explained that the dichotomy between science and the humanities is a historically-rooted formulation stemming from the developments in knowledge post-World War II.⁷ The mention of The Two Cultures in science and the humanities is situated within an epistemological context that perceives the differences in modes of thinking between them, which has implications not only in the academic world but also in addressing humanitarian issues. Despite some critics' reservations about its terminology, the argument that science and humanities are two distinct and entirely unrelated scientific activities has become widespread. Although digital humanities seems to appear as an epistemological mediator that consolidates science and humanities, Ingvarsson instead refers to it as "The Two Cultures 2.0," because the view that digital humanities is a separate activity from 'classic' humanities remains prominent.⁸





The separation between them encompasses differences in focus between those prioritizing adopting computational methods, massive data analysis, and algorithmic approaches and those prioritizing qualitative research, critical theory, and humanistic inquiry.⁹ The former position seeks to transform humanities studies by analyzing large datasets and uncovering patterns. Meanwhile, the latter position observes digital tools as complementary to humanistic approaches. Both may be true to some extent. However, such simplifications tend to be negligent and instead evade proper thought.

Amidst these tensions, Anne Balsamo suggests that fundamentally we need to stop perceiving digital as a channel for disseminating data-driven humanities studies.¹⁰ Digital tools are not just mere instruments, but they also contribute to restructuring how we manage and distribute knowledge and offer new opportunities for pedagogy. On the one hand, there is a demand to explore how technology ensembles are involved in reconfiguring knowledge production across human cultural domains. On the other hand, there is a call for a new approach to digital innovation that considers culture in all forms of digital development.¹¹

The anxiety of the two "cultures" demands digital humanities to adopt a stance: "counter-culture" or "correspondence." Counter-culture tends to sympathize with the Gutenbergian view that digital humanities scholars should actively resist and challenge problematic trends inherent in digital culture.¹² These trends may include data surveillance, algorithmic bias, information overload, and privacy erosion. Scholars adopting a counter-cultural stance can advocate for critical engagement with digital technology, promote awareness of its social-political implications, and advocate for ethical practices in digital research and pedagogy.

Instead, the "correspondence" stance argues for embracing digital culture as a valuable resource and mindset in academic work and education.¹³ Its proponents advocate for leveraging the capabilities of digital technology to enhance scientific research, teaching, and public engagement. Katherine Hayles sees the need to effectively communicate humanistic capabilities in a digital environment that caters to the language and media of younger generations.¹⁴ The correspondence stance emphasizes the potential of digital humanities to democratize access to knowledge, facilitate interdisciplinary collaboration, and promote new modes of scientific communication.

The choice between a counter-cultural or correspondence stance in digital humanities is not always binary. Adopting elements of both becomes essential for the foundational design of digital humanities. Interdisciplinarity is at the heart of the ethos and methodology of digital humanities, making their foundational design significant from an epistemological perspective.



➔ Considering the particular in the universal

The expression "data-driven" has spread across various experimental and media studies, as if the increased available data is the next scientific breakthrough. Generally, data are considered objective evidence when they are used in this expression. At this point, digital humanities needs to be cautious with this expression. Despite the fundamental elements of digital humanities involving data in its methods, the widely accepted term "data-driven" raises epistemic concerns that view the simultaneous nature of data and digital as a naive positivistic perspective. Some thinkers argue for the importance of large-scale data sets, advocating computational methods such as text mining and network analysis to uncover patterns and trends in extensive text collections or cultural artifacts. Others emphasize the importance of curated and contextualized data sets, reasoning that the value of small-scale and deeply annotated collections prioritizes richness and interpretive specificity over mere volume.¹⁷

Digital humanities continues to grapple with various challenges, including questions regarding the nature of data and data selection, reliability, and representation. Nevertheless, data is based upon human constructs, conditioned by time, place, -



language, and the ideologies of those concerned with its collection. Skepticism toward the positivistic nature of digital is not a concern that can be easily brushed aside. The tendency of digital-related studies still draws its spirit from representational philosophy.¹⁸ Representational approach, with its slogan "data speaks for itself," naively assumes data production equates to capturing a ready-made world picture that can be studied systematically. Thus, data is regarded as an object with fixed and unchanging content whose value, as a representation of reality, is determined solely by its suitability for its intended use.¹⁹

In the philosophy of science, the statement "data speaks for itself" has long been debated. There are many layers ranging from differences in data types to differences in interpretation mechanisms. Representational view holds that there are right and wrong ways to interpret data, and it is the analyst's responsibility to discover them. In representational views, data are seen as objects with fixed, context-independent meanings, contradicting the observation that massive data is consistently seen as moving objects. As the objective foundation for objectivity, data that facilitates knowledge accumulation inductively conflicts with the principle that big data can be reused in countless ways. This is indeed an impasse.

An alternative to this impasse can be unravelled by embracing findings and abandoning the idea that data constantly represents reality. A relational view provides opportunities in this regard, where data is treated as potential or actual evidence for scientific claims in ways that, at least in principle, can be examined and accounted for. The meaning attributed to data depends on its origin, physical features, and what is considered to represent those features, as well as the motivations and instruments used to visualize and maintain certain interpretations.²⁰ Data reliability thus depends on the credibility and precision of the processes used to generate and analyze it. Data presentation—how data is identified, selected, and included (or even excluded) in databases—and the information provided to users for recontextualization are fundamental.²¹ For digital humanities researchers like Tom Boellstorff, data is dense, interpretative, and contextual. Thus, digital humanities consider "thick data" more acceptable.²² Thick data refers to qualitative, detailed, and context-rich information. As opposed to big data, which has a high volume and cannot be processed or analyzed with traditional database tools and software, thick data emphasizes understanding the "why" behind the patterns and trends found. By focusing on qualitative elements, we are able to uncover hidden insights that are not visible when analyzing quantitative data.

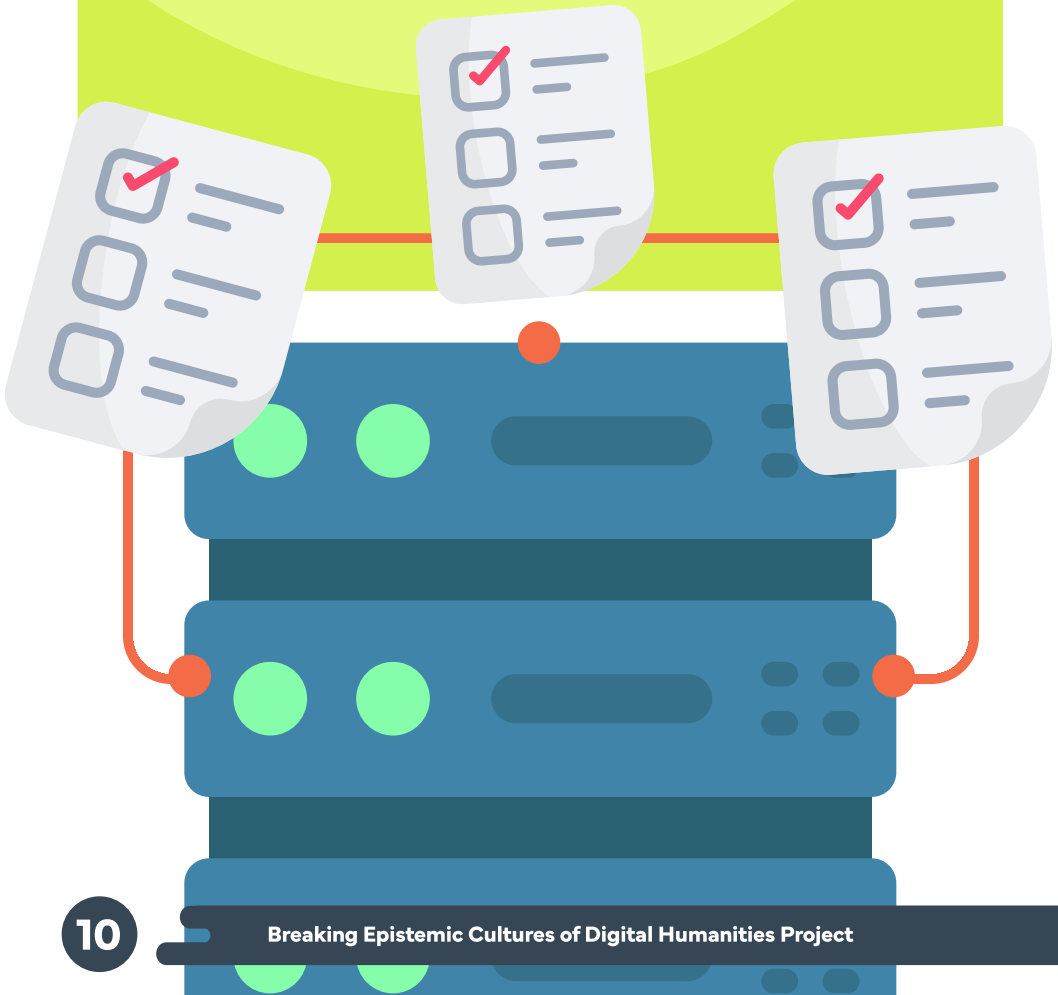
Further, the technosocial ensemble used to create computationally inferred knowledge carries a great deal of power since its epistemological assumptions impact both the results and the presentation of them. Bhattacharyya, in "Words in a World of Scaling-up," explains the invisibilization that occurs in digital space.²³ Invisibilization refers to the

tendency of text analysis and visualization to overlook the emergence of materials from non-English speaking countries, mainly in the Global South.²⁴ As a result, epistemologically heterogeneous data can become unreadable in representation schemes based on standardization. The normative assumptions of these schemes tend to prioritize data that fall into standard categories, while non-standard data may appear if the variations are made available.

The logic of standpoint theory, which emerged from the social turn in knowledge theory and is rooted in feminist epistemology, can provide an effective framework for theorizing about some of the epistemic issues discussed above. Standpoint becomes significant epistemically in technosocial knowledge apparatuses through the presentation and visualization of data in digital humanities projects that can perpetuate or challenge existing power dynamics and perspectives. Standpoint theory primarily considers whose perspectives are privileged or marginalized in information management. Considering data in the digital humanities in the context of standpoint theory highlights the importance of critically examining the social locations and perspectives that shape data production and interpretation.



Here is an example of a socio-ecological digital humanities project that considers data collected based on social location and demonstrates the particularity within the universal. *Feral Atlas* is a digital humanities project that explores interactions between humans and non-humans in the context of global environmental change.²⁶ 79 Anthropocene field findings were curated in this project, which illustrates the fallacy of imposing a single and systematic perspective on the environment, even as a fundamental fallacy. These recordings offer an interdisciplinary approach that shows the material processes through which the environment changes profoundly and visualizes complex patterns of interaction between humans and the environment. In *Feral Atlas*, many processes are depicted through maps and timelines.





➞ Materiality of knowledge, epistemic dilemma

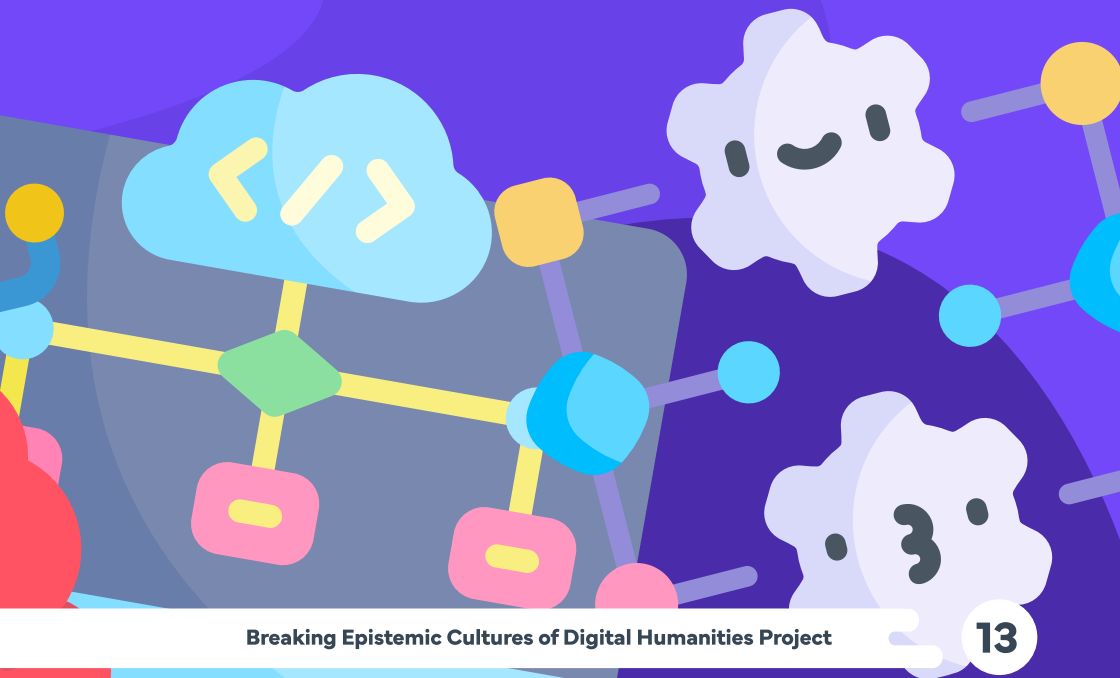
It is true that knowledge as a whole is a non-material abstraction, but the processes and systems that combine small amounts of data together to produce larger ensembles of data reside in the material realm. Unlike purely informational systems, systems that produce knowledge through procedural operations have material effects.²⁷ In digital humanities, I outlined the epistemological dilemma that could affect the readability of cultural production. Large-scale knowledge and networked knowledge can produce different forms and expressions at the same time. It is nevertheless true that both are designed in a large-scale manner and detached from the situationality of the small scale.

In "Small Data, Data Infrastructures, and Big Data," Rob Kitchin and Tracey Lauriault observe that knowledge production, especially academic knowledge, has evolved over the past few centuries using what can be termed as "small data" research, which is investigations based on highly selectively generated data samples to answer specific questions.²⁸ In the humanities field, we may consider this constitutive of the paradigmatic activity of meticulously reading texts and switching to data on a large scale, offering alternative epistemologies. Kitchin and Lauriault indicate that one consequence of the emergence of techniques for analyzing big data is that "small data" is made more like big data through the development of new data infrastructures that "collect, scale, and link small data" into larger data sets, opening it up for investigation using "big data" analysis.²⁹

The materiality of knowledge in digital humanities involves understanding how information technology processes and systems represent and process non-material information in material form. In digital humanities, information is often stored in digital formats without physical form, such as text, images, or sound. However, to be used and understood, this information must be transformed into a material form that computers, such as bits, can process. Therefore, there is a blend of the non-materiality of information itself with the materiality of the processes and systems used to access, store, and manipulate that information.

Understanding how the operational operations of information technology systems impact knowledge production in digital humanities requires understanding their influence. Information technology systems such as databases, search algorithms, and digital platforms play a crucial role in facilitating access to information, organizing and analyzing data, and facilitating collaboration among researchers.³⁰ However, using these systems can also shape frameworks of understanding, limit access to certain information, and direct attention to specific aspects of research. The impact of procedural operations in information technology systems in shaping knowledge is therefore crucial in the context of digital humanities.

These tools, adapted from Jacques Derrida, exteriorize human capacity.³¹ Referenced by Stiegler, Derrida deconstructs these technologies as therapeutics and pain-causing.³² It is possible to solve the problem of invisibility caused by dominant epistemic applications using 'crowdsourced' techniques based on tertiary retention. A fundamental aspect of understanding the digital humanities epistemological dilemmas is understanding how non-material abstractions become material during knowledge production.



Information is often abstracted from its original context in digital humanities and represented in digital format. In the Lehigh University repository database I mentioned earlier, several digital humanities projects illustrate this trend, including the following: *New York Public Library's Menu Project*, which digitizes collections of historical restaurant menus and allows readers to observe culinary trends of the era; *What Middletown Read Project*, which collects documentation and research about the experiences, contributions, and roles of women in Muncie, Indiana, as well as Middletown studies as a whole; and featured in *Kindred Britain* is a map that shows nearly 30,000 individuals, including many iconic figures in British culture, are related by blood, marriage, or other connections.³³ This abstraction process allows information to be efficiently processed and analyzed by computers, but it can also reduce the complexity and nuance of its original context. Therefore, it is important to critically reflect on how this abstraction affects our interpretation and understanding of cultural, historical, and human phenomena in general. Thus, reflection on epistemological dilemmas in digital humanities raises questions about how we understand and respond to transforming non-material abstractions into material in the context of knowledge production.

⇒ Conclusion

A series of explanations above attempt to address epistemic problems relating to understanding, studying, and formulating theories about digital infrastructure in knowledge production. In this regard, it becomes crucial to shift from treating infrastructure as a substrate to treating infrastructure as a substance. As digital humanities resources and technologies expand, digital does not simply mean an extension of humanities. Borrowing from Basalmo's assertion, a digital humanities project must consider digital not only as a technological application but also as a complex socio-techno-cultural ensemble replicating

→ Bibilografi

¹ See Galloway, A. R. *Against the Digital*. (Minneapolis, MN: University of Minnesota Press, 2014), 63.

² T. Nickles, "Alien Reasoning: Is a Major Change in Scientific Research Underway?" *Topoi* 39, no. 4 (2018): 901–914, <https://doi.org/10.1007/s11245-018-9557-1>; K. K.

³ Lehigh University. (2023, July 6). Projects - Digital Humanities. Library Guides @ Lehigh. Retrieved March 23, 2024, from <https://libraryguides.lehigh.edu/digitalhumanities/projects>

⁴ Cetina, *Epistemic Cultures: How the Sciences Make Knowledge* (Harvard University Press, 1999), <https://doi.org/10.2307/j.ctvxw3q7f>

⁵ J. Ingvarsson, *Towards a Digital Epistemology* (Springer Nature, 2021)

⁶ C. P. Snow, "The Two Cultures," *Leonardo* 23, no. 2/3 (1990): 169–173, <https://doi.org/10.2307/1578601>.

⁷ C. P. Snow. *Two Cultures*. The REDE Lecture. (1959). <https://apps.weber.edu/wsuiimages/michaelwutz/6510Trio/Rede-lecture-2-cultures.pdf>

⁸ J. Ingvarsson. *Towards a Digital Epistemology*. (Springer Nature: 2021)

⁹ H. Porsdam. (2014). *Working at the Intersection of the Humanities, Law and Technology: Digital Humanities and the "Two Cultures"*. *Pólemos*, 8(1), 75. <https://doi.org/10.1515/pol-2014-0005>

¹⁰ See Balsamo, A. *Designing Culture: The Technological Imagination at Work*. (Durham, NC: Duke University Press;), 137.

¹¹ Idem, p. 151.

¹² N. K. Hayles, *Electronic Literature: New Horizons for the Literary* (Notre Dame, IN: University of Notre Dame, 2008).

¹³ N. K. Hayles, *How We Think: Digital Media and Contemporary Technogenesis* (Chicago, IL: Chicago University Press, 2012).

¹⁴ Idem, p. 57

¹⁵ Rojas Castro, A. (2017). *Big Data in the Digital Humanities*. *New Conversations in the Global Academic Context: Digital Humanities, Library Science, Information Science, Museums--Study and Teaching*. Retrieved from <http://dx.doi.org/10.17613/M6434X>

¹⁶ See Bod, R. *A New History of the Humanities*. (Oxford University Press: 2013), 146.

¹⁷ Idem, p. 156

- ¹⁸ Sabina Leonelli, "What Difference Does Quantity Make? On the Epistemology of Big Data in Biology," *Big Data & Society* 1, no. 1 (2014): 205395171453439, <https://doi.org/10.1177/205395171453439>.
- ¹⁹ Bas C. Van Fraassen, *Scientific Representation: Paradoxes of Perspective* (Oxford: Oxford University Press, 2008), doi:10.1093/acprof:oso/9780199278220.001.0001.
- ²⁰ Jim Bogen, "Theory and Observation in Science," in *The Stanford Encyclopedia of Philosophy* (Spring 2013 Edition), ed. Edward N. Zalta, URL: <https://plato.stanford.edu/archives/spr2013/entries/science-theory-observation/>.
- ²¹ Idem
- ²² Tom Boellstorff, "Making Big Data, in Theory," *First Monday* 18, no. 10 (2013)
- ²³ Sayan Bhattacharyya, "Words in a world of scaling-up," *Sanglap: Journal of Literary and Cultural Inquiry* 4, no. 1 (2017).
- ²⁴ Idem, p. 35
- ²⁵ Sandra Harding, *Whose science? Whose knowledge?: Thinking from women's lives* (Ithaca, NY: Cornell University Press, 1991).
- ²⁶ Anna Lowenhaupt Tsing, Jennifer Deger, Alder Saxena, & Feifei Zhou, *Feral Atlas: The More-Than-Human Anthropocene* (October 22, 2020), <https://doi.org/10.21627/2020fa>.
- ²⁷ S. J. Shep, "Digital Materiality," in *A New Companion to Digital Humanities*, 322–330 (November 27, 2015). <https://doi.org/10.1002/9781118680605.ch22>
- ²⁸ R. Kitchin and Tracey P. Lauriault, "Small Data, Data Infrastructures and Big Data," *SSRN Electronic Journal*, <https://doi.org/10.2139/ssrn.2376148>. p. 470
- ²⁹ Idem, p. 471
- ³⁰ S. J. Shep, "Digital Materiality," in *A New Companion to Digital Humanities*, 322–330 (November 27, 2015). <https://doi.org/10.1002/9781118680605.ch22>
- ³¹ J. Derrida, *Of Grammatology*, trans. Gayatri Chakravorty Spivak (Baltimore, MD: Johns Hopkins University Press, 1976).
- ³² T. Colony, "Epimetheus Bound: Stiegler on Derrida, Life, and the Technological Condition," *Research in Phenomenology* 41, no. 1 (2011): 72–89



Center for Digital Society

Faculty of Social and Political Sciences
Universitas Gadjah Mada
Room BC 201-203, BC Building 2nd Floor,
Jalan Socio Yustisia 1
Bulaksumur, Yogyakarta, 55281, Indonesia

Phone: (0274) 563362, Ext. 116
Email: cfds.fisipol@ugm.ac.id
Website: cfds.fisipol.ugm.ac.id



facebook.com/cfdsugm



[Center for Digital Society \(CfDS\)](https://www.linkedin.com/company/center-for-digital-society-cfds)



[cfds_ugm](https://www.instagram.com/cfds_ugm)



[@cfds_ugm](https://twitter.com/cfds_ugm)



[@cfds_ugm](https://twitter.com/cfds_ugm)



[CfDS UGM](https://www.youtube.com/channel/UCfDSUGM)