

Denison University

Denison Digital Commons

Faculty Publications

2021

Why Digital Humanists Should Emphasize Situated Data over Capta

Matthew Lavin
Denison University

Follow this and additional works at: <https://digitalcommons.denison.edu/facultypubs>

Recommended Citation

Lavin M. Why digital humanists should emphasize situated data over capta. *Digital Humanities Quarterly*. 2021;15(2). <https://denison.idm.oclc.org/login?url=https://www.proquest.com/scholarly-journals/why-digital-humanists-should-emphasize-situated/docview/2565416574/se-2>.

This Article is brought to you for free and open access by Denison Digital Commons. It has been accepted for inclusion in Faculty Publications by an authorized administrator of Denison Digital Commons.

Why Digital Humanists Should Emphasize Situated Data over Capta

Lavin, Matthew . Digital Humanities Quarterly ; Providence Vol. 15, Iss. 2, (2021).

[ProQuest document link](#)

ABSTRACT (ENGLISH)

This essay looks back on Johanna Drucker's "Humanities Approaches to Graphical Display" (2011) ten years after its initial publication in *Digital Humanities Quarterly*, in particular Drucker's call to "reconceive all data as capta." Drucker makes several crucial points about humanistic inquiry, but this essay argues against her embrace of capta as a replacement term for data in two ways: (1) furnishing a revised and expanded etymology for the terms data and (2) exploring the benefits of embracing concepts such as situated data rather than capta.

FULL TEXT

As we pass the ten-year anniversary of Johanna Drucker's "Humanities Approaches to Graphical Display" (2011), I believe it is an opportune time to revisit Drucker's call to "reconceive all data as capta" [Drucker 2011].[1] Drucker's argument for such a reimagining is based on "the etymological roots" of data and capta – data comes from the Latin "given," and capta from the Latin "taken" – and raises a larger point about the differences between scientific realism and constructivist critiques of realism.[2] Realist approaches to visualization, Drucker argues, assume "transparency and equivalence, as if the phenomenal world were self-evident and the apprehension of it a mere mechanical task" [Drucker 2011]. Drucker goes further in stating, "Nothing in intellectual life is self-evident or self-identical, nothing in cultural life is mere fact, and nothing in the phenomenal world gives rise to a record or representation except through constructed expressions" [Drucker 2011]. Drucker lays out a lexical argument between data and capta that more or less parallels the realist and constructivist modalities of thought as she describes them (i.e., data suggests realism and capta suggests constructivism).[3]

In this piece, I want to argue in favor of embracing some of Drucker's points about humanistic inquiry while simultaneously arguing against capta as a term to be used in place of data. As a humanist, I do see value in emphasizing that data are taken and not given, but I believe there is a richer etymological narrative, and a richer history of the word *data* in English, to be described. There is also the more complicated question of whether all forms of empiricism require proceeding "as if the phenomenal world were self-evident," and the degree to which social constructions mediate our experiences and understanding of the world. I will gesture at these larger questions in this essay, but my primary concern is with how scholars in digital humanities should approach conversations about data.

The bulk of my essay is a revised and expanded etymology for the terms *data* and *capta*, which considers their Latin roots, as well as the close ties of the word *data* to the publishing history of Euclid's *Data* (1625) and subsequent translations and editions of that work. Euclid's own use of the term "given" was a subject for discussion and debate, as was the degree to which his use of "given" resembled other Greeks' use of the term. A fully fleshed out history of the term *data* in English includes its use in two related but distinct senses: geometrical and empirical, the latter of these emerging from the prior. To appreciate the difference between these senses, we must understand the degree to which early modern thinkers synthesized geometrical and numerical thought into what is now regarded as the singular discipline of mathematics. Further, these thinkers adapted and extended mathematical thinking to translate observations of or measurements from the natural world into generalizations and explanatory systems.

In the closing section of this essay, I will turn to the benefits of embracing concepts such as situated data over *capta*. Such an approach allows humanists to contest oversimplifications of poorly executed data-driven inquiry and simultaneously to create more opportunities for conversation with other disciplines. The tone of these conversations could be positively affected, as well, with digital humanities speaking and listening in equal proportion. Lastly, I will discuss the new possibilities that this rhetorical shift could create for how we teach data analysis, which has the potential to advance digital humanities as a discipline.

Drucker's Conceptual and Terminological Intervention

My response to Drucker begins with a certain degree of uncertainty about how "Humanities Approaches to Graphical Display" has been received in digital humanities. Generally, I take it to have made a strong, positive impression. Drucker was not the first scholar to point out that *capta* might be a more appropriate term than *data*, but her work is arguably the most influential to make this point, especially in digital humanities.[4] This essay, or the corresponding material in Drucker's book *Graphesis* (2014), is widely referenced in digital humanities, and I have seen it listed on many DH course syllabi.[5] There appear to have been numerous conference presentations, blog posts, and discussions of *capta* on social media. In various contexts, public displays of one's agreement with Drucker might take the form of statements like, "There is no such thing as 'data.' There are only 'capta.'"[6] Broadly speaking, Drucker has set the terms of the debate on the subject of data and *capta*, and this influence has been both descriptive and normative. Simultaneously, I have not seen signs of a large-scale movement to purge documents of the word *data*, nor have the word *capta* or its cognates become especially prevalent in digital humanities scholarship.[7]

One might argue that Drucker makes a conceptual intervention and not a terminological one. I would concede that the essay is most concerned with reconceptualizing data as *capta*, and the importance of such a shift. With regard to data visualization, Drucker argues that "the rendering of statistical information into graphical form gives it a simplicity and legibility that hides every aspect of the original interpretative framework on which the statistical data were constructed" [Drucker 2011]. Data do not "pre-exist their parameterization" because they are "constructed as an interpretation of the phenomenal world, not inherent in it" [Drucker 2011]. Drucker does not explicitly argue that humanities scholars should never say or write the word *data*, but she does claim that "all data is *capta*" and that because of the etymological roots of the terms, *capta* "is 'taken' actively while *data* is assumed to be a 'given' able to be recorded and observed" [Drucker 2011]. As a result, I reject the idea that one can so easily separate Drucker's conceptual intervention from the terminological one supporting it.

Another way of looking at Drucker's influence would be to say that there is widespread agreement with her views on data and *capta*, but practical considerations outweigh other concerns. Such practical considerations include (1) maximizing clarity for particular audiences, (2) limits on space that would make it difficult to explain the choice to use *capta*, and (3) the question of how to invoke concepts like big data, data visualization, databases, datasets, metadata, and open data. One might attempt to stake out a rhetorical middle position by mentioning that data are really *capta* and to proceed with the word *data* thereafter, as Rob Kitchin does in *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences* (2014). Kitchin adds, "since the term data has been so thoroughly ingrained to mean *capta*, rather than confuse the matter further it makes sense to continue to use the term data where *capta* would be more appropriate" [Kitchin 2014]. This seems like it would be an appealing option for many people, and I suspect it is widespread, although I have not investigated this question in any serious way. Regardless, I am aware of no published work that disputes Drucker's core points. That is:

DETAILS

Subject: Big Data; Open data; Etymology; Realism; Digital humanities; Visualization

Business indexing term: Subject: Big Data

Publication title:	Digital Humanities Quarterly; Providence
Volume:	15
Issue:	2
Publication year:	2021
Publication date:	2021
Section:	Editorials
Place of publication:	Providence
Country of publication:	Providence
Publication subject:	Humanities: Comprehensive Works--Computer Applications
e-ISSN:	19384122
Source type:	Scholarly Journal
Language of publication:	English
Document type:	Editorial
ProQuest document ID:	2565416574
Document URL:	https://denison.idm.oclc.org/login?url=https://www.proquest.com/scholarly-journals/why-digital-humanists-should-emphasize-situated/docview/2565416574/se-2
Copyright:	© 2021. This work is published under http://creativecommons.org/licenses/by-nd/4.0/ (the "License"). Notwithstanding the ProQuest Terms and Conditions, you may use this content in accordance with the terms of the License.
Last updated:	2021-08-28
Database:	Publicly Available Content Database

LINKS

[Check for Full Text](#), [Find full text](#), [Check Full Text Finder for Full Text](#)

Database copyright © 2022 ProQuest LLC. All rights reserved.

[Terms and Conditions](#) [Contact ProQuest](#)