

# The Use of Big Data

By Tate Barend, Tan Pham & Kevin Lee  
Digital Humanities 101



# Our Proposition

Our understanding of Big Data:

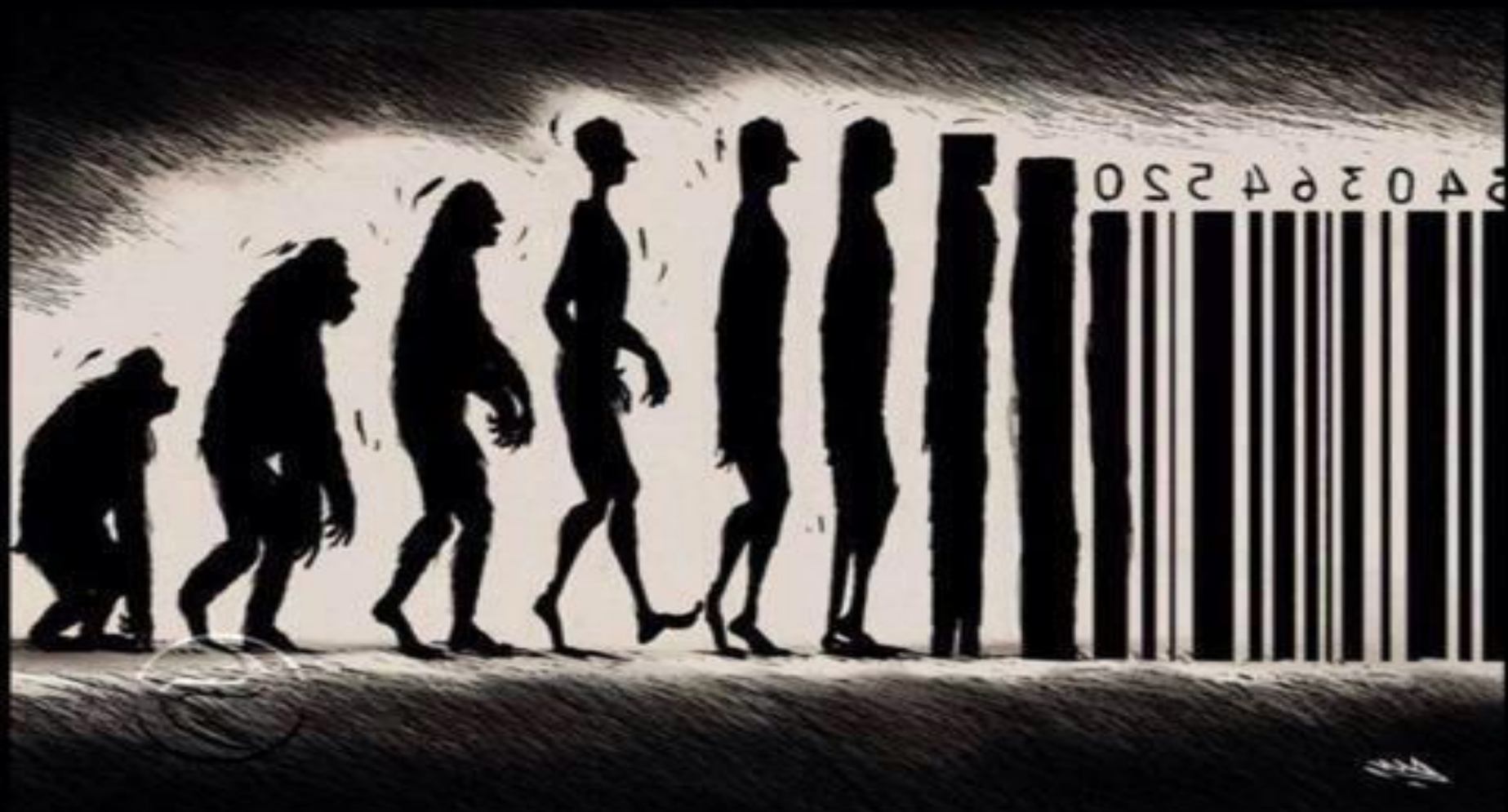
- Big Data is the massive amount of data that is being created every second, by everyone.
- While only a few are in the position to understand and analyze these Data sets.

Our presentation will focus on the accessibility of Data collection at an unprecedented scale and problems this poses as well as potential solutions within the Digital Humanities fields

# Introduction to Readings

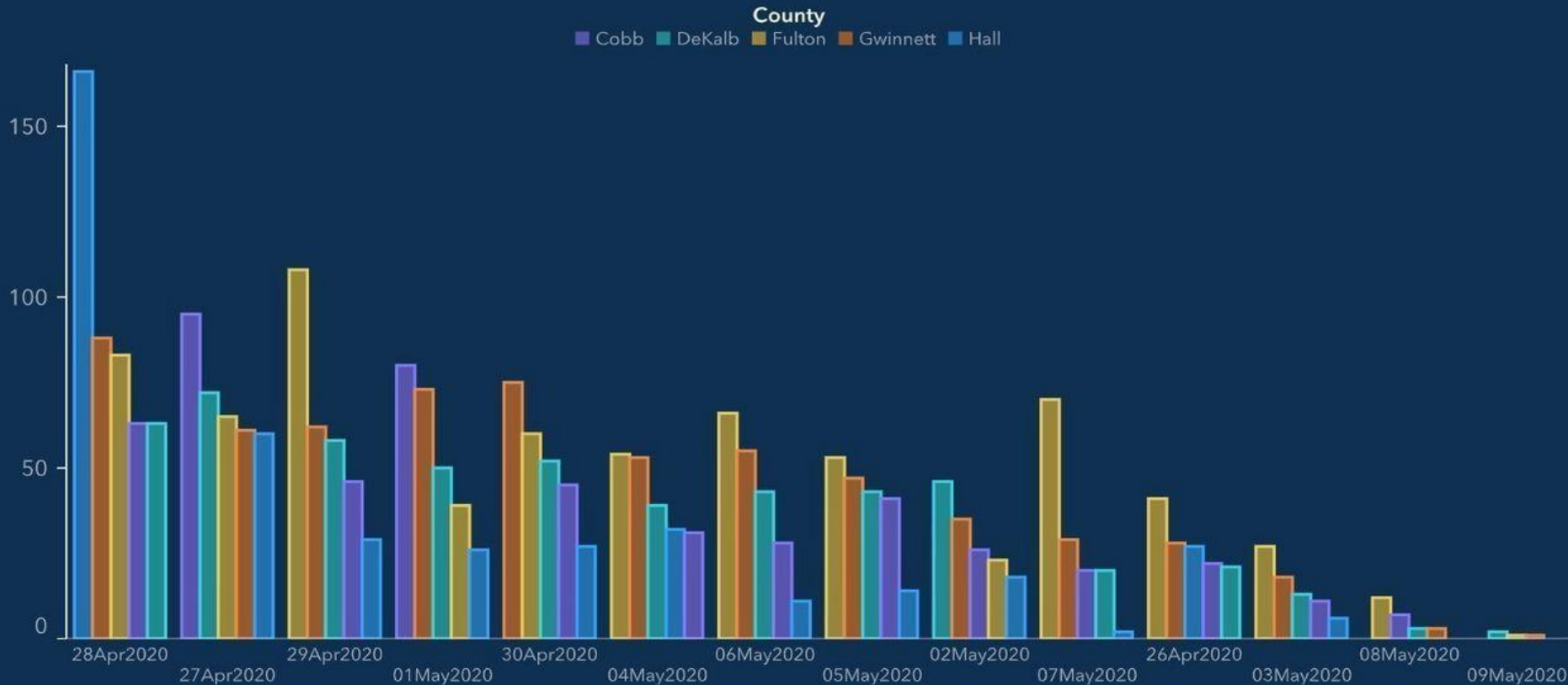
- “Critical Questions For Big Data” by Danah Boyd & Kate Crawford
  - A scholarly integration of the implications of Big Data
- “The differences between digital humanities and digital history” by Stephen Robertson
  - Insight into some of the entry barriers and disciplinary differences that impact big data
- “Against Cultural Singularity” by Alan Liu
  - specific practical applications to manage Big Data within the Digital Humanities themselves





# Top 5 Counties with the Greatest Number of Confirmed COVID-19 Cases

The chart below represents the most impacted counties over the past 15 days and the number of cases over time. The table below also represents the number of deaths and hospitalizations in each of those impacted counties.



# Why Big Data? (Robertson)

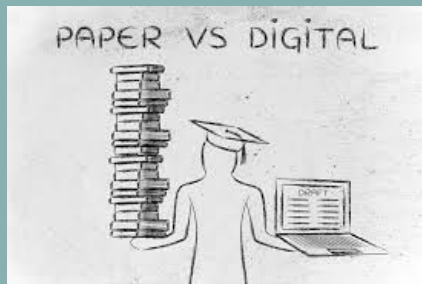
- Increased demand for digital expertise
- Accessibility
- Contextualization





# Drawbacks & Barriers (Robertson)

- Lack of Digitized Sources
- Costs of Accessibility
- Data Limitations



“Unless we can have machines scan, sort, and apply digital techniques to the full texts of documents, we can’t do sophisticated digital scholarship”

- Daniel J. Cohen

Yale & Princeton Alumni, American Historian, Founding Executive Director of the Digital Public Library of America





# Against the Cultural Singularity

**Liu:** cultural singularity refers to the point in time when all constituent parts of the present culture converge and are brought under the corporate framework and are capitalized.

- Culture as forms of knowledge that become essentially products of "information capital"

**The threat of “deep government and industrial drive toward gathering and extracting maximum value from data” without the representation of academia, and society at large. Also people like you and me.**

- Alan Liu in "Against the Cultural Singularity," May 2016.





# Against the Cultural Singularity

What remedies are there within the Digital Humanities sphere to help circumvent the imposition of power structures that hinder the flow of information in academia?

One possible solution is **Open-Source**

**Github** is the most popular platform for sharing open-source software





**teejee2008** commented on Nov 11, 2019

Owner

Author

...

I have archived this project. I will be releasing this project again under a paid licence at some point in the future (maybe next year if I get the time).

If someone wants to fork this project please go ahead. This is a open-source project that anybody can fork and develop.

For people who are still using this app there is no need to be disappointed. I'm going to take my time finishing this app before I launch it again. When it launches again, it will be a finished product with a regular development cycle. And as a paid app you will have some surety that the project is sustainable.



# Against the Cultural Singularity

Open-source as a code word for **'trial version'**

- Possibility of corporate developers exploiting open-access products from Github for free for their own motives and profit
- eventually Github users pay their developers for their own project

**Does this allow for large companies like Microsoft to take over "open-access"?**

Precludes Liu's expressed concerns and desires to prevent the mismanagement/take-over of Big Data within spaces like academia by corporate interests/big businesses as a clear example of "cultural singularity"



## Conclusion

**“The era of Big Data has only just begun, but it is already important that we start questioning the assumptions, values, and biases of this new wave of research.”**

**-Boyd**