Bringing the Neighborhood In: Exploring the Inequalities Within and Outside of Big Data

JOSEPH GIBBONS, DEPARTMENT OF SOCIOLOGY, SAN DIEGO STATE UNIVERSITY

SPECIALIST MEETING ON HUMAN DYNAMICS IN THE MOBILE AGE, AUGUST 11-12

2015

How do neighborhood effects impact the creation of big data?

ANSWERING THIS ALLOWS US TO UNDERSTAND HOW BIASES IN PHYSICAL SPACE TRANSLATE INTO CYBERSPACE

Biases in Neighborhood and Connectivity

- Community connection is often stratified along racial and class lines across neighborhoods
  - Lack of communication between new media users in white and nonwhite communities?
- The ever present digital divide which disproportionately impacts poor and nonwhite neighborhoods
  - Leaving out voices from big data

Key Questions

- How is the information collected by big data is being disseminated across different neighborhoods?

- How do we study big data in ways which meaningfully account for these physical divisions?

- *Looking at twitter followers is a good start, but we can do more!*

Approach 1: Closer Examination of New Media Users

- ## Often gaps in what we know about the users themselves
  - Where do they live? Who outside of new media do they tell about tweets? Do they communicated across racial ethnic lines?

- ## Traditional Social Science as a Supplement
  - Secondary Data sources
  - Random Sample Survey of known users
  - In-Person interviews

Approach 2: New Questions

- Deeper and more substantive questions are needed which look at the content of the data and how it relates to location

- Example of race and twitter usage:
  - How do racist tweets vary by location?
  - Are they connected to the local segregated character of an area?
  - How does the nature of the discourse vary?

Conclusion

- We need to bring location more into the discussion of big data
  - How are neighborhood divisions by race and income impacting the dissemination
- This will require better conceptual questions and empirical frameworks
- Doing so will offer better meaning to these data as it reflects the people who generate it