Library Faculty Publications and Presentations   University Library

Heard on the Net: Data, Data Everywhere

Jill Emery
Portland State University, jemery@pdx.edu

Citation Details

Heard on the Net

Data, Data Everywhere

By Jill Emery *(Collection Development Librarian, Portland State University)*   <jemery@pdx.edu>

There has been a tremendous amount of what I've taken to calling a "liberal freak out" made over the transition of power within the U.S. Government at the beginning of this year. websites have been drastically changed, removed, and/or replaced.[1] Federal data sets have been obscured, moved, or removed.[2] Librarians are rightfully concerned about these changes, and their reaction on social media sites like Facebook and Twitter have shown concern and distress over these changes. Government documents librarians have always been aware of the trade-off of administrative changes and administrative agendas in the federal government. Stuff changes with each new administration and each new administration has the agenda it wants to promote.

It can be said that the amount of change enacted as quickly and readily with this U.S. Presidential transition has been unprecedented. Furthermore, this is probably the first time in a vastly networked twenty-first century that we have been privileged to readily see these changes enacted. For most of the twentieth century, these changes happened via printed materials and thus happened slower and more consistently between each change of administration. It should also be noted that this points to how much more transparent the previous administration had become in providing information and documentation to the general population through the networked environment. Something as simple as providing a link to immediate Spanish translation of webpages was readily noted when that functionality disappeared from the main White House website.

Various organizations have stepped forward to help insure that the historical record of our governmental leadership is recorded. First among them is the Internet Archive,[3] a group of information professionals who have been instrumental in both creating archives of governmental websites[4] and also in creating blogs to help fact check the current administration.[5] In addition to the wonderful work being done by this organization, there have been librarians leading efforts to create data refuges[6] and performing what is being termed data rescue projects throughout the country.[7] These projects aim to capture and insure data sets that could be completely removed, considered controversial, or labeled "biased" by the current administration can be capture and warehouse for use by academics, scientists, and researchers throughout the world. An additional benefit to this coordinated effort is the application of data management standards and practices[8] to many of this information which had previously been provided in a more ad hoc way by different agencies and arms of the government.

While the rapid change of websites and governmental data is a relatively new phenomenon in regards to the United States, other countries have long term experience with this type of behavior. The issues and concerns regarding poor data and how this information may skew development efforts are exemplified in books like *Poor Numbers: How We Are Misled by African Development Statistics and What to Do about It*.[9] The lack of robust data is recognized as an issue with the ability to develop viable policies and support in many countries. The Latin American Network Information Center[10] run by the Teresa Lozano Long Institute of Latin American Studies at The University of Texas is another example of this work to capture and preserve data and websites during administrative changes in governments. Again, the need to do this is new in the United States but has been a regular practice in other countries for the past decade. Lastly, there are emerging global networks around data refuge and rescue.[11] There is quite a bit we can learn from our library colleagues working throughout the world and we should in some cases rely on their expertise and lead in these arenas as opposed to trying to recreate functionality and process that has been proven successful.

A final note on the "liberal freak out" phenomenon. While it is shocking for many of us in academia to be confronted with opinions that show propaganda, obvious bias, and a particular ethos, for many of our colleagues and friends, this situation has always been a part of their daily existence. There have been numerous articles and studies written that show that the information the federal government espoused and the data collected has always had inherent biases and problems.[12] One just has to read through governmental data from the nineteenth century regarding our Indigenous Nations to comprehend how "facts and figures" have been a mainstay for biased actions. There are groups attempting to correct some of these problems but whether these changes get enacted now is questionable.[13] Furthermore, we must learn to be both more contextual as well as more precise with accepting data, especially acknowledging that governmental data is flawed no matter who may be capturing it. We also should be vigilant as to how it is applied and utilized.[14] Another great book that gets at this problem and explains in a very direct way how data can be harmful is *Weapons of Math Destruction: How Big Data Increases Inequity and Threatens Democracy*.[15]

Librarians do have every reason to be concerned and react with dismay at what we've seen happen with the current governmental information systems. Governmental information is a mainstay resource of every library in this country and insuring we have reliable information to use and provide to our end-users is critical to all of our individual library missions. What we've learned in these early days is critical to our future success. We must recognize that this data has always had inherent flaws and that it has been applied unjustly. We must recognize that there is an investment to be made by libraries in a technological infrastructure that allows us to develop our own local networked archives to insure we can capture the information sources of the times as well as support those who can do it better than we can.

We also must recognize that this is a global issue and concern for librarians around the world and that the best efforts may be found outside of our national boundaries.

Author's References

1. <https://www.bostonglobe.com/metro/2017/01/20/why-are-whitehouse-gov-web-pages-disappearing/gd0HEAAU49hrL-ZMCiOQwuN/story.html>

2. <http://thehill.com/blogs/pundits-blog/technology/320511-why-is-federal-government-data-disappearing>

3. <https://archive.org/>

4. <https://blog.archive.org/2016/12/15/preserving-u-s-government-websites-and-data-as-the-obama-term-ends/>

5. <https://blog.archive.org/2017/03/01/how-to-use-the-trump-archive-to-find-tv-news-appearances-fact-checks-and-share-clips/>

6. <https://www.datarefuge.org/>

7. <http://www.ppehlab.org/>

8. <http://www.niso.org/apps/group_public/download.php/15375/PrimerRDM-2015-0727.pdf>

9. <http://www.cornellpress.cornell.edu/book/?GCOI=80140100939320>

10. <http://lanic.utexas.edu/project/>

11. <https://www.idare-portal.org/>

12. <https://www.census.gov/prod/2013pubs/p20-568.pdf>

13. <http://civilrightsdocs.info/pdf/reports/Census-Report-2014-WEB.pdf>

14. <https://www.washingtonpost.com/news/post-nation/wp/2016/07/11/arent-more-white-people-than-black-people-killed-by-police-yes-but-no/?utm_term=.bf9316432171>

15. <https://weaponsofmathdestructionbook.com/>