Places on the Map and in the Cloud: Representations of Locality and Geography in Wikipedia

[Extended Abstract]

Randall M. Livingstone
University of Oregon
1275 University of Oregon
Eugene, OR 97403
1-541-246-8258

livingst@uoregon.edu

ABSTRACT

This poster will present preliminary results of a study that considers the efforts of WikiProject Countering Systemic Bias, a collective of editors dedicated to combating bias on the English-language Wikipedia. Through a content analysis comparing the project to a sample from the general population, the scope of this group's labor is gauged and discussed.

Categories and Subject Descriptors

H.5.3:n. [Information Interfaces]: Group and Organization Interfaces – Collaborative computing, Web-based interaction, Computer-supported cooperative work; K.4.3 [Computers and Society]: Organizational Impacts – Computer-supported collaborative work.

Keywords

Wikipedia, wiki, collaboration, users, bias, geography, content

1. INTRODUCTION

Nearly ten years ago, Manuel Castells [1] mapped the geography of the Internet for us in *The Internet Galaxy*. By exploring the infrastructure of our growing computer networks, Castells was able to demonstrate how offline areas of economic and technical wealth were influencing the development of our online worlds, defying the potential utopianism preached by some of the early digital advocates [2] [3]. Castells also brought into focus the social and political issues around "the digital divide," again implicating the influence of wealthy Western nations on the global diffusion of Internet technologies and highlighting the remarkable similarities between inequality offline and online. Finally, rather than the end of physically concentrated geographic power and a reduction in the defining characteristic of distance and space that networked information and communication technologies could facilitate, Castells predicted the rise of "sprawling metropolises" and quotes William J. Mitchell [4]: "the power of place will still prevail."

In more recent work, Castells [5] reflects on the development of information networks over the past decade, reasserting the relationship between physical and online geography. He points out: "There is an increasing contradiction between the space of flows and the space of places," citing a recent World Values Survey indicating that 47 percent of people feel a strong regional or local identity, while only 13 percent identify with the global culture that accompanies major information hubs. Despite the increasing awareness of our online activity and engagement (our online "lives"), we are still deeply rooted in the material realities of our favorite physical locales; we are still deeply rooted in the space of places.

In referring to the hardware of Internet technology, Castells reminds us: "the uses of the Internet are dependent not only on connectivity, but on the quality of the connection." The same dictum can be applied to the content flowing through our information and communication systems. *What is* making its way into the digital annals of our recorded online history? And more specifically, in light of our continuing need for geographically physical identity, how is the content of global locales being represented on the Internet? This study contributes to our knowledge of online geographic representation by exploring the work of a small community of Wikipedia editors working against information bias.

Wikipedia, the world's largest online encyclopedia, has faired well in studies comparing its accuracy to that of traditional encyclopedias like *Britannica*, including the now infamous 2005 *Nature* study. Wikipedia's open content system, however, provides the possibility of a much more robust representation of the world around us than a hardcopy source ever could. A few hundred invested editors have actively expressed their concern over the direction this content is headed by forming WikiProject: Countering Systemic Bias (WP:CSB), a loose collective that intends to combat biases in the English language Wikipedia by remedying omissions and contributing content on unrepresented and under-represented non-Western topics. As with other WikiProjects, though, these editors have few ties to each other beyond this desire to counterbalance bias, and coordination of their efforts is mostly passed along to other WikiProjects with regional interests (such as WikiProject: Algeria, WikiProject: Peru, etc.).

This study assesses the work of WPCSB (n = 314) through a quantitative content analysis of their edits, looking specifically at a sample of edits pertaining to global locations and physicalities. Their work is then compared to a comparable sample of editors (n = 314) randomly drawn from the general population of Wikipedia contributors. Finally, similar to some of Castells work, this data is plotted onto geographic maps of

the globe to further explore the extent of WPCSB's efforts compared to the control group. Descriptive data is drawn from editors' "User" pages to contextualize results and discuss findings.

2. METHODOLOGY

To explore the extent to which WPCSB addressed non-Western content, each editor's last 50 edits (from March 1, 2010 back) were coded for their primary geographical country of interest, as well as the conceptual category of the page itself ("P" for person/people, "L" for location, "I" for idea/concept, "T" for object/thing, or "NA" for indeterminate). Corresponding information on each country's associated continent, based on the United Nations Statistic Division listing, was added after the coding stage. The random sample of editors was chosen by utilizing the system's "Random article" function, then selecting the 10th most recent contributor to that article. Bots were excluded from the sample, and same coding procedure discussed above was used.

Only edits coded as "L" for location were considered for this analysis.

3. PRELIMINARY RESULTS

As coding for this research is in-progress, these preliminary results present random samples (n = 150) from both the experimental (WPCSB) and control (general population) groups. This sample yielded a total of 2650 location edits.

Table 1. Edit counts and Chi-square Goodness of Fit results

	WPCSB	Random Sample
North America	416***	646***
Latin America	30	44
Europe	324	358
Africa	68	78
Asia	185***	386***
Australasia	35	50
Antarctica	1	0
NA	8**	21**
Totals	1066	1584

Note: ** $p < .01$; *** $p < .001$

These results indicate that WPCSB is contributing nearly 33% less edits to geographical locations than the general population, instead devoting their time to content on people, things, and ideas. As would be expected, the random sample contributed significantly more to areas in North America and Asia (particularly Japan and South Korea), while no significant difference was found for any other areas of the globe.

As data analysis presented the opportunity, additional exploratory analyses were conducted on edit counts at the national level. Chi-square goodness of fit tests were calculated

to compare frequency counts of all individual nations (n = 126) appearing in at least one of the samples. Due to the large number of comparisons, a Bonferroni adjusted alpha level of .0003968 was used to gauge statistical significance. These nations are mapped in Figure 1 and shaded according to significance level and direction of significance. Nations in green are those more heavily edited by WPCSB sample, while nations in yellow are those more heavily edited by the random sample. These preliminary results show individual differences in nations that largely fit the systemic bias highlighted in the mission of WPCSB—heightened editing activity by the general population on areas of North America, Europe, and developed Asia—though WPCSB's editing impact seems minimal. With a full data set, these interpretations will be more fully explored in the final poster.

4. REFERENCES

[1] Castells, M. 2001. *The Internet Galaxy: Reflections on the Internet, Business, and Society*. Oxford UP, Oxford, UK.

[2] Negroponte, N. 1995. *Being Digitial*. Knopf, New York, NY.

[3] Rheingold, H. 1993. *The Virtual Community: Homesteading on the Electronic Frontier*. Addison-Wesley Pub Co., Reading, MA.

[4] Mitchell, W. 1999. *E-topia: "urban life, Jim—but not as we know it*. MIT Press, Cambridge, MA.

[5] Castells, M. 2010. Globalisation, networking, urbanization: Reflections on the spatial dynamics of the information age. *Urban Studies*, 47, 13 (Nov. 2010).

[6] Forte, A., Larco, V., & Bruckman, A. 2009. Decentralization in Wikipedia governance. *Journal of Management Information Systems*, 26, 1, (Summer 2009).

[7] Kane, G. C. 2009. It's a network, not an encyclopedia: A social network perspective on Wikipedia collaboration. *Proceedings from the 2009 Academy of Management Annual Meeting*, Chicago, IL.

[8] Ayers, P., Matthews, C., & Yates, B. 2008. *How Wikipedia Works: And How You Can Be a Part of It*. No Starch Press, San Francisco, CA.
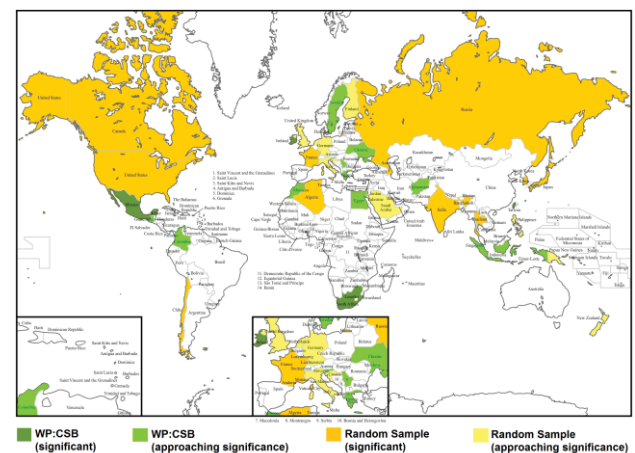
Figure 2. People Edits by Nation – Direction of Significance