



Volume 4, General Issue (2016)

Measuring Altruistic Impact: A Model for Understanding the Social Justice of Open Access

Margaret Heller, Franny Gaede

Heller, M. & Gaede, F. (2016). Measuring Altruistic Impact: A Model for Understanding the Social Justice of Open Access. *Journal of Librarianship and Scholarly Communication*, 4(General Issue), eP2132. <http://dx.doi.org/10.7710/2162-3309.2132>

External Data or Supplements:

Heller, M. & Gaede, F. (2016). Appendix A: Social Justice Term Analysis & Appendix B: Social Justice Overlap Template & Geographic Usage. <http://dx.doi.org/10.7910/DVN/XU5IBN>, Harvard Dataverse.



© 2016 Heller & Gaede. This open access article is distributed under a Creative Commons Attribution 4.0 License (<https://creativecommons.org/licenses/by/4.0/>)

RESEARCH

Measuring Altruistic Impact: A Model for Understanding the Social Justice of Open Access

Margaret Heller

Digital Services Librarian, Loyola University Chicago

Franny Gaede

Scholarly Communication Librarian, Butler University

INTRODUCTION Traditional assessment of ways in which open access initiatives and institutional repositories have provided a return on investment normally use pragmatic measures such as download counts and citation benefits. This pragmatic approach misses out on the powerful altruistic impact of improving access to international and/or marginalized communities. Using a frame of social justice, this article considers the importance of developing altruistic measures of repositories, particularly for institutions with missions specifically related to social justice and related themes. **METHODS** Using web analytics data for search keywords from eight institutions and geographic usage data from nine institutions, the authors were able to determine how well social justice related content is accessed by search engines and how much overall content is accessed internationally, particularly by lower-resourced countries. A social justice term list was developed to permit corpus overlap analysis with each institution's search keywords, while the World Bank country income lists were used to determine international access by low and low-middle income countries. **RESULTS** Universities with mission statements explicitly mentioning social justice or Catholic social teaching had greater overlap with the social justice corpus. Low and low-middle income countries as defined by the World Bank were among the most engaged users. All institutions had at least one social justice search term in their top ten; Marquette University had five. Collection development in social science and environmental sustainability at Loyola University Chicago successfully increased this term overlap year-over-year and increased user engagement as measured by session length. **DISCUSSION** The results of this exploratory study indicate that it is possible to use repository data to evaluate the success of an institution's open access and social justice initiatives. The year-over-year improvement of Loyola's numbers suggest in addition that it is possible to increase social justice impact through collection development. Performing an analysis of social justice impact can be used as an overall strategy for repository success and outreach on campus, particularly for institutions where social justice is an important part of the campus identity. For repositories in need of further resources, the ability to quantify impact for university administrators and decision-makers may be of use. **CONCLUSION** For institutions with a social justice mission, improving social justice content may improve repository ranking in social justice related search results. Collection development strategies should focus on departments and/or individuals who are working in social justice-related areas, which defined broadly could encompass much of an institution. For institutions that emphasize social justice, it may be easier to approach faculty who might not otherwise have an interest in open access issues.

Received: 01/26/2016 Accepted: 07/05/2016

Correspondence: Margaret Heller, Cudahy Library, Loyola University Chicago, 1032 W. Sheridan Road, Chicago, IL 60660, mheller1@luc.edu



© 2016 Heller & Gaede. This open access article is distributed under a Creative Commons Attribution 4.0 License (<https://creativecommons.org/licenses/by/4.0/>)

IMPLICATIONS FOR PRACTICE

1. The model presented can be used by repository managers to quantitatively assess the success of their repositories in providing content with a social justice bent and how well all content is reaching countries with limited access.
2. Altruistic impact assessment can support pragmatic arguments for maintaining a repository and help repository managers justify their work to university administrators.
3. The authors provide strategies based on their own experiences for repository managers to improve the social justice impact of their repositories through collection development.

INTRODUCTION

As systems of scholarly communication have changed radically over the last few decades, librarians have been on the front lines of trying to navigate a more complex and potentially less fair system. Barbara Fister (2010) introduced the concept of “liberation bibliography” in her *Library Journal* column, challenging libraries to allocate fiscal and human resources to help combat injustice in our practices. These injustices range from forcing librarians into the role of gatekeeper for increasingly expensive scholarly journals to seeing alumni deprived of the scholarly resources they have been trained to use. The need to demonstrate value in an era of shrinking higher education and library budgets only compounds these issues, giving rise to rhetoric trending towards quantification and validating return on investment.

This paper argues that an assessment of an open access repository must exist on a continuum between purely altruistic and purely pragmatic considerations. The number of institutional repositories has increased over the last ten years (OpenDOAR, 2015; Morrison, 2015), but paying to make something available for free could be seen as economically unsound, whether that comes in the form of supporting a repository or paying an article processing fee. Assessments of open access institutional repositories generally have focused on citation advantage, since these are quantifiable and may have a direct institutional benefit in terms of faculty prestige. Studies of this phenomenon indicate that making an article open access tends to improve its chances of citation elsewhere (SPARC Europe, 2015). These quantitative analyses miss some important aspects of open access, however. This article will establish that providing access to people who would otherwise lack it is a crucial role for libraries in sustaining the public sphere. Libraries fail to make the argument for why they are a social good if they ignore the altruistic impacts of repositories. This is an important consideration for institutions funding repositories based on their institutional character or mission of social justice or social responsibility.

This paper argues that libraries must move beyond pragmatic justifications for institutional repositories. Repository managers must understand their work in the context of social justice, lest they become complicit in unjust scholarly communication systems. To that end, this paper presents a model for assessing the social justice impact of open access institutional repositories in higher education and some demonstrations of the model.

To create a “social justice impact” metric, we have measured two separate but related types of institutional repository usage—access via search engine of social justice-related content in the repository and international usage of all content in the repository, particularly by lower-resourced countries. These two methods, taken together, begin to paint a picture of the social justice impact of the repository.

LITERATURE REVIEW

Open Access and the Public Sphere

In this section, we will ground the idea of open access in the theory of the public sphere and demonstrate that open access is a social and public good. The notion of a public sphere comes from German philosopher Jürgen Habermas’s *The Structural Transformation of the Public Sphere* (1989). This work examines the structural transformation of society from the rational-critical debate of 18th century salon culture to a manipulated and inauthentic consumer culture constructed primarily by mass media and other private interests. These interests have monopolized public opinion and political thought because they are no longer challenged by an external, autonomous public authority engaged in rational discussion about politics. Habermas suggests that state intervention in the economy in the 19th and early 20th century (linking economic issues with political ones) and the emergence of uncritical consumption from new media disrupted the fragile socioeconomic balance that permitted the public sphere to be formed in the first place. Habermas remains optimistic that it could be recreated if we, as a society, vigorously interrogate the interests mentioned above and reclaim our intellectual autonomy. Though Habermas draws on historical example, it is possible to identify the Internet as a modern-day arena of the public sphere. Just like 18th century salon culture, not everyone is part of the conversation, but those who do have the privilege are able to participate.

Scholar of critical pedagogy and higher education Henry Giroux suggests that neoliberal politics threaten the academy, and the humanities in particular, through increased corporatization. This leads to “the slow death of the university as a center of critique, vital source of civic education, and crucial public good” (Giroux, 2014, p. 16). As the university becomes increasingly irrelevant as a public sphere, faculty are demoralized. Rather than

fighting for change, many become cynical and “retreat into a sterile form of professionalism” (p. 17). Academic work becomes cut off from any relevance to social or civic problems, and universities and intellectualism become suspect as “part of an ongoing attempt to destroy higher education as a democratic public sphere that enables intellectuals to stand firm, take risks, imagine the otherwise, and push against the grain” (p. 19).

John Buschman examines the role of libraries in the public sphere in his 2003 book *Dismantling the Public Sphere: Situating and Sustaining Librarianship in the Age of the New Public Philosophy* and follow-up article in *Library Philosophy and Practice* (Buschman, 2005). He suggests that the continued defunding of libraries and schools can be traced back to the long-term crumbling of the public sphere under neoliberal economics; much of this argument is drawn from Giroux’s work. Buschman identifies a profound shift from “access, freedom of information, and information as a public good in a democracy [...] toward the commercial information industry” (2003, p. 28). In response to this cultural change, libraries have been remaking themselves in a corporate image, modeling physical spaces after bookstores and integrating business rhetoric into methods and practices to “justify our existence in this new environment” (Buschman, 2005, p. 5). This new vision for libraries moves them further away from a mission to support a right of access to information (Buschman, 2003, p. 48). In Buschman’s view, library collections are vital to “further democratic inclusion by curating collections and resources to reflect historical and current intellectual diversity” (2003, p. 47). Similarly, librarians “enact [...] the] principle of unfettered information and transparency” as they “commit [...] to preserving a wide variety of balanced collections over time” (Buschman, 2003, p. 46).

We argue that participating in open access is a critical part of preserving the public sphere, and that we, as librarians, must build and assess our open access initiatives with the understanding that they are a vital public and social good. They undercut a hegemonic model of scholarly publishing, make academic discourse more accessible to those outside the academy, and allow more diverse voices to participate.

Open Access and Social Justice

We believe social justice impact is a natural focus for evaluating an institutional repository for altruistic impact for two reasons. First, many institutions have a mission that explicitly mentions social justice or closely related concepts and requires all campus units to demonstrate alignment with said mission (Reason, Ryder & Kee, 2013, p. 13; Atkinson, p. 380-1; Torres-Harding, p. 90). University libraries can contribute to the fulfillment of university mission statements related to Catholic social teaching, education for personal and social responsibility, and community engagement and development by making research

products and social justice-related content openly available. Catholic social teaching exhorts the wealthy to aid the less fortunate proactively (Leo XIII, 1891). Education for personal and social responsibility focuses on civic engagement, volunteerism, and serving the public good (Reason, Ryder & Kee, 2013, p. 13). Embeddedness in the community “establish[es...] intellectual and institutional resources [that] can make genuine contributions to improving the quality of life” (Checkoway, 2001, p.139). This paper will return to the idea of university missions in the discussion of results.

Second, social justice aligns closely with the ideals of open access defined by the Budapest Open Access Initiative Declaration (2002):

Removing access barriers to [...] literature will accelerate research, enrich education, share the learning of the rich with the poor and the poor with the rich, make this literature as useful as it can be, and lay the foundation for uniting humanity in a common intellectual conversation and quest for knowledge (para. 1).

Refining our understanding of social justice further, we will consider John Rawls’s *A Theory of Justice* (1971), which underpins most modern discussions of social justice and Catholic social teaching. Rawls developed his theory as a rejection of utilitarianism, where social structures exist to maximize the sum total of happiness, and to provide a systematic theory that would consider justice when envisioning a well-ordered society. His idea of society is a “cooperative venture for mutual advantage” though “marked by a conflict [...] of interests since persons are not indifferent as to how the greater benefits produced by their collaboration are distributed” (Rawls, p. 4). This conflict is to be resolved by a set of social justice principles that “provide a way of assigning rights and duties in the basic institutions of society and [...] define the appropriate distribution of the benefits and burdens of social cooperation” (Rawls, p. 4). The basic structures of society include a system of government: laws, organization of economy, and cultural conditions. But which kinds of these basic structures best exemplify justice? Ultimately, Rawls believes “all social values—liberty and opportunity, income and wealth, and the bases of self-respect—are to be distributed equally unless an unequal distribution of any, or all, of these values is to everyone’s advantage” (p. 54). This last clause is particularly important; unequal distribution is acceptable if and only if it “improves the expectations of the least advantaged members of society” (Rawls, p. 65). Access to information is vital to success in our information economy. The commodification of information in the last few decades has continued the unequal distribution of resources among classes. Those without access, whether to research databases, high-speed internet, uncensored social media, or electricity for computers, are at a distinct disadvantage.

The Justice of International Access

Major publishers have created programs for developing nations that allow some access to scholarly research databases, but use of those resources is dependent on the conditions under which research is conducted in those countries. Scholars in developing nations face challenges to using online material, such as minimal computer access and lack of time (Harle, 2010, p. 4). There is some indication in the literature of a preference for open access literature in under-resourced African universities. A study of Nigerian researchers found that nearly all participants used and cited open access articles, despite power outages and limited computer time that affected their access (Ivewighrehweta & Onoriode, 2012). Another study found African researchers in corrosion chemistry were more than twice as likely to cite open access articles than those in non-African countries (Taha & Kraus, 2013).

Stratification of institutions by research funding and the ability to secure long-term access to monetary and scholarly resources leaves some researchers ahead and others left out of the international scholarly community. Maha Bali describes the personal perspective of a researcher from the Global South who feels at a disadvantage, even though her own institution is relatively privileged, and calls for more participation in open access in the developing world (Bali, 2015). At least one open access publisher, BioMed Central, uses a list of low and low-middle income countries produced by the World Bank (The World Bank, 2016) to determine whether an author should receive a fee waiver for open access publishing (BioMed Central, 2016). Even if a publisher like BioMed Central removes the financial hurdle of an author fee for open access publication, this does not mean researchers will have the necessary resources to participate in publishing.

Mauritius, a small island nation in the Indian Ocean, provides a good illustration of the issues facing such researchers. It is geographically isolated but has increased its information technology sector over the past two decades and so has relatively good internet connectivity (Oolun, Krishna, Ramgolam & Dorasami, 2012, p. 161-168). The University of Mauritius, like many institutions, has identified strategic goals emphasizing the production of impactful research, but does not have the means to provide its researchers with all the tools to do so. The university provides access to ScienceDirect, but usage by faculty was not as high as anticipated. “UoM FoS [University of Mauritius Faculty of Science] scholars say that they use academic databases most often (74%) for finding econtent. This is followed by searching through aggregated journals (47%), Google Scholar (43%) and pre-print repositories (40%). This is a common pattern of usage in institutions that do not subscribe to large numbers of journals, but rely on package subscriptions with a few big publishing firms” (Trotter, Henry, Kell, Willmers, Gray, Beeharry & King, 2014, p. 91). Searching Google Scholar is not as reliable, since there may not be an accessible copy of an article. While

interlibrary loan services are available, due to the length of time to delivery, some scholars rely on international research partners or colleagues for access to materials through semi-illicit or illicit sharing. This is not the ideal scenario, in the words of a researcher interviewed for this study:

I think it is very important that we are not seen from the North as beggars. I'm sorry to say the term. I think we need to show people that they can partner with us because they will gain from our science. The challenge is not to send one of our students to the USA or Europe or South Africa or Japan for a postdoc, the challenge is to get those people coming to see us... knowing they come here not because Mauritius has got sandy beaches and so on, but because they know that we are doing good science (Trotter et al., 2014, p. 94).

Making articles and datasets available open access is an important way for privileged institutions to improve research results and enable partnerships without putting their less-resourced colleagues in an uncomfortable position. Jonathan Harle describes the difficulty researchers from sub-Saharan Africa countries in the British Commonwealth face from journal editors: “Revisions often stall or take too long when academics are not able to locate the material they need to address the deficiencies highlighted by editors” (2010, p. 56). A good illustration of the condescension some of the academic community has for these researchers is a 2013 “sting” by John Bohannon, which created a poor quality paper from a fake African researcher that was submitted to open access journals to prove certain journals were low quality (Bohannon, 2013).

The Justice of Search Engine Traffic

Universities with a social justice or responsibility mission need to ensure their research is part of the public conversation, and assessing how effectively users access the repository content through search engines can be a demonstration of whether that is happening. Search engines form an integral layer through which information is accessed, in particular by novice searchers who are unfamiliar with the non-indexed web and are unlikely to find material not in search engines (Halavais, 2009, p. 41). Unique library content must be included in search engines to be part of the modern knowledge landscape and to improve the quality of results, particularly for people with limited search skills (which are correlated with income, race, and education levels) (Halavais, 2009, p. 88). Sociologist Alexander Halavais urges cultural institutions to improve their openness and participation in the web. He suggests that the shift to open access is part of the web's cultural imperative, and that it facilitates the creation of more diverse ideas in the academy and the world. While this may be discomfiting to scholars used to traditional modes of publication, it is crucial for universities and libraries

to be in the space where knowledge creation is happening. Individual institutions have limited power to improve their rank in search engine results against corporate search engine optimization, but improved metadata and more references to external content can help break the mainstream American corporate hegemony that threatens the promise of the web and the public sphere (Halavais, 2009, p. 108). Researchers often do not go into a search with a particular vision of what they want to find; rather, what they find tends to shape what they are looking for (Halavais, 2009, p. 87). Therefore, a repository needs to do well in search results to contribute to public discourse.

METHODS

To create a “social justice impact” metric requires quantifying these various types of justice. First, how well-accessed by search engines is the social justice-related content that exemplifies the social justice mission of the university? Second, how much of the content is accessed internationally, particularly by lower-resourced countries? While we consider each factor independently, it is necessary to consider cases where a repository may be weaker in social justice content but much higher international usage from low income countries, or vice versa, to equalize the types of access across institution type (for instance between a small liberal arts college and a land grant university). The methods below describe two different types of assessment, but they are meant in the end to be considered together.

Sample

The sample for this study consists of ten institutions, two of which are the authors’ home institutions. One author requested assistance from the Digital Commons email list in October 2014 for additional institutions to contribute their own data. The sample included a variety of types and sizes of institutions, all of which had repositories functioning since at least the beginning of 2012. Each institution’s mission statement was acquired from its website and an extract containing mention of social justice, Catholic social teachings, and related concepts was recorded.

Time Period

The analysis considered data from nine institutions for geographic analysis and eight institutions for keyword analysis, covering the period of October 1, 2013–October 1, 2014.

Data Collection

Google Analytics do not give an exact measure of everything that happens across the repository

(for instance, hits on an article from Google Analytics might be far below server logs of downloads of articles), but they do indicate browsing patterns.

After some experimentation and an initial small conference presentation which elicited helpful suggestions, Margaret Heller determined that the most useful Google Analytics reports at the time dealt with geographic usage (specifically, the Google Analytics report for Audience > Geo > Location (by country)) and search engine keywords (specifically, Acquisition > Keywords > Organic).

Geographic analysis

The report from Google Analytics contains columns with the country/territory name, the number of sessions, percent of new sessions, number of new users, the bounce rate (i.e. the percent of users who left the site), the number of pages per session, the average session duration, and several columns related to advertising that contain no data in this case. Creating the reports required removing extra rows such as dates and totals, followed by sorting the data in several ways to answer different questions. In addition to the total volume of international usage in a year, it was important to know which countries had particularly engaged users, even if their total volume was quite low. This was determined by sorting by number of sessions (i.e. which countries used the repository the most) and copying that data into a new worksheet, followed by the same procedure for the number of pages per session (i.e. which countries used the most repository items in a sitting) and the average session length (i.e. which countries had users who spent a long time on the repository due to engagement or poor internet connectivity).

A comparison between the World Bank country income lists and the country reports created an estimate of the possibility that users were less privileged (The World Bank, 2016). As described in the literature review, this is the list used by BioMed Central for its open access fee publishing waiver. The World Bank country list was copied into each country list and the cells highlighted that matched low and middle-income countries in order to determine the percentage overlap, as well as specific countries that tended to appear on each list. This gave institutions a “score” for engaged international users from low and low-middle-income countries, and a list of which countries were the most engaged.

Social justice term analysis

The search engine keywords report contains columns for the keywords, number of sessions, percentage of new sessions, numbers of new users, the bounce rate, pages per session, and average session length. Similar to the geographic analysis, the goal was to examine both the

highly used keywords and an aggregate of how many of them covered social justice research topics, which required a form of corpus overlap analysis. Using search engine keywords as an analog for social justice research topics requires a series of complex and subjective steps, which are illustrated below for the sake of transparency.

Creating the Social Justice Corpus

The corpus is available for others to use (Appendix A) without following the steps below. Ten websites were selected from social or governmental organizations promoting social justice research, based in the United States, Canada, the United Kingdom, Australia, and South Africa. This provided a broad pool of candidate terms from across the English-speaking world. Non-English languages were outside the scope of the available data. These websites were processed through Voyant Tools, a graphical text-mining tool (Sinclair & Rockwell, 2016), which created a word list from all the websites consisting of 9,382 terms, excluding a built-in stopword list. These were downloaded as a plain text file and opened in OpenRefine (OpenRefine, 2016) for initial processing.

Using the clustering and filtering functions in OpenRefine, the following types of terms were removed: Any term with only one appearance in the corpus, terms with only 2 letters, URLs, digits, numbers (written out), roman numerals, possessives, file extensions (pdf, html), contractions, abbreviations, country names from the data sets (Australia, Canada, USA, UK, New Zealand), and directional names. For additional cleanup, natural language processing was run on the list using WebNLP (Tirlea & Pörsch, 2014), which is a graphical UI for Python Natural Language Toolkit (NLTK). First was the lemmatize process, which removes stems such as -ly and -ing from terms, which resulted in 4,612 terms, though this was not complete. Second was the parts of speech tagging process. This process creates a comma-separated list of terms with the parts of speech for the word identified.

After loading this last list into Excel, parts of speech that were not nouns or verb stems were removed. After a final deduplication this resulted in a list of 3,125 candidate terms, which we used for initial testing of the corpus for false positives. Due to many false positives, the majority of terms were removed after comparing to a list of the 2,000 most common English words (Hakuta & Wientjes, 2010). This left a final corpus of 722 terms. This process did not remove all potential biases and subjective views of social justice-related research, but our later testing revealed that a large corpus introduced too many false positives. The corpus is available for anyone to use and improve upon or adapt to a particular situation.

Overlap Analysis

Unlike geographic reports, keyword reports contain thousands or tens of thousands of rows, which makes individual analysis or coding extremely time-consuming. Normalization of search terms allowed grouping related terms and removing outlier searchers. OpenRefine provides clustering algorithms to remove differences created by spelling errors and slight differences in punctuation or word order, which made it the preferred tool. In cases where two synonyms were used for the same concept, one was chosen and the other converted to match. For example, one repository had hundreds of variations on the term “ethical marketing”, so that was the term chosen to represent all of the searches for that concept. For additional outliers not caught by the normalization process, a column was created in OpenRefine called Multiples that indicated whether a particular term showed up more than once or had multiple sessions. Any term that did not meet these criteria was deleted.

After the cleanup and reduction of terms, it was possible to perform some manual coding to get a sense of the general categories into which the terms fell and to remove potential false positives before the overlap analysis. Remaining categories of terms that could be identified included institution or department names, the repository name, the library name, a specific document URL, a standard identifier such as a DOI, a specific journal or citation, yearbooks, specific non-historical people, specific historical people, or Google Scholar referrals.

The last step after reviewing the data was to determine how many terms were potentially related to social justice. Many sophisticated methods exist for determining corpus overlap, but other repository managers could replicate an Excel-based method without special training. The coded column of search terms was copied into an Excel workbook on the first worksheet, and the 722-word corpus into the next worksheet. The following Excel formula was pasted in the first cell in the column next to the terms, and dragged down the column to apply to all rows. It indicates whether any word in the coded terms column matches a word in the social justice corpus:

```
=IF(ISNUMBER(LOOKUP(2^15,SEARCH(Sheet2!A$1:A$722,$A1))),”Match”).
```

A template is available for others to use (Appendix B). The new column was filtered to the cells with “Match” and the percentage of “social justice terms” was calculated. Franny Gaede conducted a manual review of the matches for Butler University, and determined that there could be as high as 30% error rate with false positives, and as high as 10% missed matches, so an error rate of around 20% is assumed. While this

is high, such a subjective measure will be hard to reduce without manual review. This rough estimate can, however, help to spot trends and show changes year to year.

RESULTS

Institutions with explicit mentions of social justice, Catholic social teaching, or a related concept in their mission statement tended to have more overlap between search terms that reached the repository and social justice–related terms (Table 1).

Institution	Mission Statement Extract	Social Justice Term Overlap
Marquette University	Catholic social teaching	35%
DePaul University	special concern for the deprived members of society	31%
Illinois Wesleyan University	social justice	25%
Loyola University Chicago	Service that promotes justice	24%
Butler University	community service	20%
University of Montana	ethical, and engaged citizens	17%
SUNY Brockport	civic engagement in a culturally diverse society	17%
Washington University-St. Louis	enhance the lives and livelihoods	13%

Table 1. Social justice content of mission statement and overlap analysis

Comparing these numbers to the geographic analysis, there are a few parallels, as well as some important distinctions. No institution had any countries from the low-income list in most number of sessions, but all of them had at least some lower-middle-income countries. For the measures for engaged users of most pages and longest sessions, the numbers varied, but all institutions had at least some countries from both the low and lower-middle-income lists (Table 2).

Institution	Most Sessions			Most Pages			Longest Sessions		
	Low	Lower-middle	Total	Low	Lower-middle	Total	Low	Lower-middle	Total
Marquette University	0%	16%	16%	20%	36%	56%	16%	28%	44%
DePaul University	0%	20%	20%	20%	8%	28%	32%	12%	44%
Illinois Wesleyan University	0%	20%	20%	8%	12%	20%	8%	16%	24%
Loyola University Chicago	0%	20%	20%	8%	24%	32%	12%	20%	32%
Butler University	0%	16%	16%	8%	16%	24%	16%	12%	28%
University of Montana	0%	12%	12%	16%	16%	32%	16%	16%	32%
SUNY Brockport	0%	16%	16%	12%	12%	24%	16%	24%	40%
Washington University-St. Louis	0%	12%	12%	20%	12%	32%	24%	12%	36%
University of Nebraska	0%	32%	32%	32%	8%	40%	40%	28%	68%

Table 2. Summary of geographic usage by World Bank country income level

Overall, the only lower-middle-income country that appeared on every repository country list was the Philippines, followed by Pakistan and India on nine lists, and Indonesia on eight lists. Some of this may be related to English being widely spoken in these countries. Mauritius, while not a low-income country, appeared in 5 of the lists for most sessions and 9 for most pages and longest sessions, which matches expectations from the literature. Table 3 illustrates the spread of specific lower and lower-middle-income countries represented in the data (listed in numerical order).

Multiple Most Pages/Session Report Appearances	Multiple Longest Session Report Appearances
Cambodia	Cambodia
Democratic Republic of the Congo	Democratic Republic of the Congo
Eritrea	Eritrea
Ethiopia	Ethiopia
Gambia	Gambia
Madagascar	Malawi
Malawi	Mozambique
Mozambique	Nepal
Rwanda	Rwanda
Sierra Leone	Sierra Leone
Somalia	Somalia

Table 3. Countries appearing on multiple reports for high engagement and on World Bank low income list (numerical order)

A complete record of the data is available as online spreadsheets in Appendix B.

DISCUSSION

These results show a correlation between the explicitness of social justice concepts mentioned in institutional mission statements and the number of searches matched with social justice-related terms. Additionally, these institutional repositories tend to have engaged users from low and low middle-income countries. Several exceptions to this trend indicate that it is worthwhile to consider keywords and geographic usage in tandem, as we have in this study. Washington University in St. Louis has a focus on scientific research; while the number of searches for social justice-related terms in its repository is low, lower-income users are using the material. We do not have terms for the University of Nebraska, but its repository focuses on agricultural research, which may be of particular interest to users outside the academy.

Some institutions with relatively few low or lower-middle countries in the “most sessions” category nevertheless had more engaged users from those countries. For example, only 16% of the top 25 countries represented in Marquette University’s statistics for most sessions were from the low or lower-middle income countries list. However, 56% and 44% of the countries represented in the statistics for most pages and longest sessions, respectively, were from that list. With 35% matches for social justice terms, this institution ranks at the top for social justice impact. The term analysis showed that all institutions had at least one social justice-related term in the top 10 coded keywords. While a search for a yearbook or a specific person (usually a faculty member) might appear as the top result, a social justice-related term would appear somewhere else in the top 10. Half of the top ten terms in Marquette University’s list were social justice-related.

One of the goals of this project was to assess whether the content in repositories reflected the institutional character. Margaret Heller conducted an initial version of this research in late 2014 and found Loyola University Chicago’s results showed no keywords related to several important social justice-related initiatives on campus, indicating a major gap between Loyola’s mission and repository content. She engaged in collection development in areas related to social sciences and environmental sustainability, leading to a measureable improvement. Social justice term overlap went from 24% to 30% and engaged users went up for longest sessions from 32% to 48%.

The results of this exploratory study indicate that it is possible to use repository data to evaluate the success of an institution’s open access and social justice initiatives. The year-over-year improvement of Loyola’s numbers suggest in addition that it is possible to increase social justice impact through collection development. Performing an analysis of social justice impact can be used as a strategy for repository success and outreach on campus, particularly for institutions where social justice is an important part of the campus identity. For repositories in need of further resources, the ability to quantify impact for university administrators and decision-makers may be of use.

Limitations

As mentioned in “Creating the Social Justice Corpus,” we developed the social justice-related keyword corpus based on the text of ten government and social justice organization websites. This list is neither bias-free nor comprehensive, but we present it as a starting place for more nuanced ways of understanding repository use and impact.

Just as we do not presume all users from a low-income country are underprivileged, we do not assume all social justice-related search topics are part of social justice research projects.

Rather, our data shows that an open access repository was being used in the public sphere of the open web on a social justice topic.

Future Research

Further refinements of this method should test the amount of manual cleanup of Google Analytics reports necessary to achieve an accurate result and verify that changes to Google Analytics have not changed the method of creating the correct report. However, we recommend that repository managers seeking to apply this method engage in a large amount of manual analysis at least once to review terms and understand the types of incoming traffic to the repository and whether this aligns with institutional goals.

In terms of scope, we focus on international usage in this study, but a great many potential readers of academic research in the United States have similarly limited access to scholarly literature. For state universities, it may be just as important to measure rural use in their own states as use in other countries.

In developing this model, the authors hope other institutional repository managers will be able to assess how well their repositories are providing content with a social justice research bent and how well all their content is reaching countries with limited access. Individual repository managers can use the information above to understand their repository's social justice impact in the context of the institutions studied here and/or with peer institutions. Additional research is needed to confirm the assumptions made in this article, but it is possible to draft some implications for practice.

CONCLUSION

For institutions with a social justice mission, improving the depth of social justice content collected may improve repository ranking in social justice-related search results. Collection development strategies should focus on departments and/or individuals who are working in social justice-related areas, which defined broadly could encompass much of an institution. For institutions that emphasize social justice, it may be easier to approach faculty who might not otherwise have an interest in open access issues.

By balancing our pragmatic and altruistic motivations, we can demonstrate far-reaching institutional impact. Open access to the scholarly and creative output of our institutions contributes a vital academic good insofar as prestige and reputation are concerned, but the social good is something extraordinary and should excite us more. In reclaiming our role as facilitators of democratic discourse, we demonstrate the change we believe in and live out our bibliography.

REFERENCES

- Atkinson, T. N. (2008). Textual mapping of imitation and intertextuality in college and university mission statements: A new institutional perspective. *Semiotica*, 2008(172), 361-387. <http://dx.doi.org/10.1515/SEMI.2008.104>
- Bali, M. (2015). Open access & copyright: A view from the South [Blog]. Retrieved from <http://chronicle.com/blogs/profhacker/open-access-copyright-a-view-from-the-south/61308>.
- BioMed Central. (2016). BioMed Central open access in the developing world: For authors. Retrieved from <http://gateways.biomedcentral.com/oaddw/authors/>
- Bohannon, J. (2013). Who's afraid of peer review? *Science*, 342(6154), 60–65. <http://doi.org/10.1126/science.342.6154.60>
- Buschman, J. E. (2003). *Dismantling the public sphere: Situating and sustaining librarianship in the age of the new public philosophy*. Westport, CT: Greenwood Press. Retrieved from <http://www.ebrary.com>
- Buschman, J. (2005). On Libraries and the Public Sphere. *Library Philosophy and Practice*, 7(2). Retrieved from <http://digitalcommons.unl.edu/libphilprac/11>
- Chan, L., Cuplinskas, D., Eisen, M., Friend, F., Genova, Y., Guédon, J.-C., ... Velterop, J. (2002, February 14). Budapest Open Access Initiative Declaration. Retrieved from <http://www.budapestopenaccessinitiative.org/read>
- Checkoway, B.. (2001). Renewing the civic mission of the American research university. *The Journal of Higher Education*, 72(2), 125-147. <http://doi.org/10.2307/2649319>
- Fister, B. (2010, April 1). Liberation bibliography. *Library Journal*, 135(6), 38–40.
- Giroux, H. *Neoliberalism's War on Higher Education*. Toronto, ON: Between the Lines, 2014.
- Habermas, J. (1989). *The structural transformation of the public sphere: An inquiry into a category of bourgeois society*. Cambridge, MA: MIT Press.
- Hakuta, K., & Wientjes, G. (2010). WordSift. Retrieved from <http://www.wordsift.com/wordlists>

Halavais, A. M. C. (2009). *Search engine society*. Malden, MA: Polity Press.

Harle, J. (2010). *Growing knowledge: Access to research in east and southern African universities*. London: The Association of Commonwealth Universities. Retrieved from <https://www.acu.ac.uk/focus-areas/arcadia-growing-knowledge>

Ivewhreghweta, O., & Onoriode, O. K. (2012). Open access and scholarly publishing: Opportunities and challenges to Nigerian researchers. *Chinese Librarianship*, (33). Retrieved from <http://www.white-clouds.com/iclc/cliej/cl33IO.pdf>

Leo XIII. (1891). *Rerum Novarum: Encyclical of Pope Leo XIII on capital and labor*. Vatican City: Libreria Editrice Vaticana. Retrieved from http://w2.vatican.va/content/leo-xiii/en/encyclicals/documents/hf_l-xiii_enc_15051891_rerum-novarum.html

Morrison, H. (2014, October 1). Dramatic growth of open access September 30, 2014: Some useful numbers for Open Access Week [Web log post]. Retrieved from <http://poeticeconomics.blogspot.com/2014/10/dramatic-growth-of-open-access.html>

Oolun, K., Ramgolam, S., & Dorasami, V. (2012). The making of a digital nation: Toward i-Mauritius. In *The Global Information Technology Report 2012*. World Economic Forum. Retrieved from <http://www.weforum.org/reports/global-information-technology-report-2012>

OpenDOAR. (2015, October 20). Worldwide Growth of the OpenDOAR Database. Retrieved from <http://www.opendoar.org/>

OpenRefine. (2016). Retrieved from <http://openrefine.org/>

Rawls, J. (1999). *A theory of justice* (Revised ed.). Cambridge: Belknap Press.

Reason, R.D., Ryder, A J., & Kee, C. (2013). Higher education's role in educating for personal and social responsibility: A review of existing literature. *New Directions for Higher Education*, 2013(164), 13-22. <http://dx.doi.org/10.1002/he.20072>

Sinclair, S., & Rockwell, G. (2016). Voyant Tools: Reveal your texts. Retrieved from <http://voyant-tools.org/>

Taha, M., & Kraus, J. R. (2013). The citation of open access resources by African researchers in corrosion chemistry. *Library Philosophy & Practice*. Retrieved from <http://digitalcommons.unl.edu/libphilprac/889/>

Torres-Harding, S. R., Diaz, E., Schamberger, A., & Carollo, O. Psychological sense of community and university mission as predictors of student social justice engagement. *Journal of Higher Education Outreach and Engagement*, 19(3), 89-112. <http://openjournals.libs.uga.edu/index.php/jheoe/article/view/1485>

The World Bank. (n.d.). Country and Lending Groups. Retrieved from <http://data.worldbank.org/about/country-and-lending-groups>

Tirlea, B., & Pörsch, J. (2014). WebNLP. Retrieved from <http://dh.mi.ur.de/>

Trotter, H., Kell, C., Willmers, M., Gray, E., Beeharry, G. K., & King, T. (2014). *Scholarly Communication at the University of Mauritius Case Study Report. Scholarly Communication in Africa Programme*. Retrieved from <http://openuct.uct.ac.za/article/scholarly-communication-university-mauritius-case-study-report>

APPENDICES

Appendix A

Heller, M. & Gaede, F. (2016). Appendix A: Social Justice Term Analysis. <http://dx.doi.org/10.7910/DVN/XU5IBN>, Harvard Dataverse.

Appendix B

Heller, M. & Gaede, F. (2016). Appendix A: Social Justice Term Analysis. <http://dx.doi.org/10.7910/DVN/XU5IBN>, Harvard Dataverse.

Heller, M. & Gaede, F. (2016). Appendix B: Social Justice Overlap Template & Geographic Usage. <http://dx.doi.org/10.7910/DVN/XU5IBN>, Harvard Dataverse.