This is an Accepted Manuscript of an article published by Taylor & Francis in Public Services Quarterly on March, 11 2019, available online: https://www.tandfonline.com/doi/full/10.1080/15228959.2018.1522989.

FUTURE VOICES IN PUBLIC SERVICES

Miriam L. Matteson

Column Editor

Column description. The Future Voices in Public Services column is a forum for graduate students in library and information science programs to discuss key issues in academic library public services, to envision what they feel librarians in public service have to offer to academia, to tell us their visions for the profession, or to share research being conducted in library schools. We hope to provide fresh perspectives from those entering our field, in both the United States and other countries. Interested students in library and information science programs are invited to contact Miriam L. Matteson, mmattes1@kent.edu.

Research Article Institutional Repositories and Usage: Thoughts on Realized and Potential Value

http://orcid.org/0000-0002-3603-0071 William Jones, Miriam Matteson. Department of Library and Information Studies, University of Buffalo, Buffalo, New York, 14260-2200, USA Risman Plaza, Kent State University School of Information, Kent, Ohio, USA. Comments and suggestions should be sent to the Column Editor: Miriam L. Matteson, Risman Plaza, Kent State University School of Information, Kent, Ohio, 44242-0001, USA. Email: mmattes1@kent.edu

W. Reed Jones is a Master’s student in the Department of Library and Information Studies at the University of Buffalo. In this essay, he analyzes three issues relating to the adoption and effective use of institutional repositories: potential impact of the material, interest in submitting materials to the repository, and the efficacy with which materials are available and accessible in the repository. He argues that institutional repositories should continue to focus on collecting both published and unpublished materials, and can also play a significant role in data set storage and management within institutions.

In spite of the existence of Institutional Repositories (IRs) for over a decade there is still some ambiguity surrounding IRs and their role in scholarly communication. The idea of a central repository for all digital material created by those affiliated with an institution is an exciting model which fosters greater access to scholarship without the barriers many traditional models of publication carry. The institutional repository is often linked to the larger Open Access movement that has existed for decades and is gaining more support from both within and outside of academia. Two federal acts have accelerated interest and support of this movement: FASTR (2013) and FRPAA (2012) mandate that federally funded scientific research must be made freely accessible after 12 months. This has altered the structure of grant applications and increased the relevancy of IRs for researchers. IRs provide permanent storage of data sets mandated by the grant funders by creating permanent URL links. Though this column does not address the state of OA, it assumes that interest in reducing barriers to access is an important driver of future dissemination techniques and, therefore, influences decision making for suppliers and consumers of intellectual products.

The institutional repository is a dedicated digital space for acquisition, preservation, and dissemination of intellectual products of those directly affiliated with an institution. This includes papers, data sets, and reports that are of interest to researchers, but it also contains institutional data and other material that must be contained behind access barriers for confidentiality purposes. This paper focuses on the institutional repository’s (IR) role in increasing barrier-free public access to material. To do so, this analysis covers three topics concerning the efficacy of IRs in barrier-free dissemination: impact, interest, and efficacy.

IRs have a distinct ability (potential or realized) to increase the availability and dissemination of scholarly work. However, current conditions show that IRs have not realized their potential. One reason for this is copyright law, which restricts IRs from providing free, open access to material under copyright by another party. A recent study of the copyright policies of 100 IRs found that the average IR can only provide about 65% of its content as open access (Dawson & Yang, 2016). For IRs to increase access to materials the libraries must actively work with researchers to negotiate permission rights with publishers. Russell et al. (2016, p. 63) argues that working with groups like...
ScienceDirect and Elsevier to facilitate access to works created by institutional researchers maximizes “visibility, impact, and dissemination” of the work. The Dean of Libraries of the University of Florida, Judith Russell, advocates linking published works into the IR to appease researchers who prefer to have the published versions, instead of preprints, available in the IR (Russell et al., 2016). In accordance with this line of thinking, for IRs to most effectively provide access, there must be continued dialogue with traditional publishers to create more favorable terms regarding the copyright permissions on the scholarship. More importantly, library professionals must take an active role in these discussions instead of putting the burden on the authors.

Perhaps, in focusing on generating access to materials deemed traditional sources of scholarly communication (academic publications and journal articles), the benefits of the institutional repository are lost. This led some research to conclude IRs “are not consistently collecting materials of the highest impact, suggesting that the value of these repositories is relatively low” (Wacha and Wisner, 2011, p. 386). Instead, according to this argument, institutional repositories tend to focus on their library goals, often collecting student works, theses and dissertations. This argument misses some key factors about IRs. To get a better assessment, it is more important to look at the constraints of the IR in its current state. It is accepted that scholars still look to traditional methods of publishing as their primary method of disseminating scholarly works. At best, the IR will harbor duplicate copies of this material. When the initial material is already available openly and without cost, this creates a redundancy. Self-archiving is not a common practice; a 2011 survey found a significant number of faculty were unaware of IRs (40.1%) (Kim, 2011). Factors including altruism, copyright concerns, accessibility, and trust impact the decision making of researchers when considering depositing material in an IR. Kim argues that preservation and easy access are the major benefits of the IR and should be promoted when advertising the use of IRs by content creators (Kim, 2011). There must be short and long term initiatives and in its current state, to strengthen the position of IRs now efforts should focus on making thesis, dissertations, and other university products discoverable. The long term goal should be to foster greater use and submission to IRs.

This is not to say that pursuing the acquisition of published articles in IRs should not be pursued. Instead it acknowledges the current state of scholarly communication and argues that focus should remain on areas that create the greatest impact, while generating interest and outreach based on how faculty interact with and perceive IRs. Koler-Povh et al. (2014) reviewed the University of Ljubljana IR and found that statistics spanning two years indicate more than 2,800 downloads of Ph.D. theses from within the collection, representing an increase from previous years. Other data collected in this survey show interesting download statistics including materials were used in more than 90 countries (Koler-Povh et al., 2014). This broad geographic finding demonstrates how an IR can be an excellent platform to increase exposure to specific materials instantly.

However simply deciding what materials are publicly and digitally accessible does not suffice. Accessibility, as noted in Kim (2011), carries a significant amount of influence over those looking to deposit material. IRs can only succeed if the material they store is discoverable by the end users. The work of Linde et al. (2011) examines a specific type of IR material: conference papers. Linde exposes issues arising from self-archiving material that “has first been published in a scientific journal or proceedings” (Linde et al., 2011, p. 261). Without the same policy tools of journal articles (SHERPA/RoMEO), conference proceedings complicate the process of self-archiving materials. Using 1,787 papers spanning from 2005 to 2009 across six Swedish institutions, the research concludes 15% of materials could not be found in the designated search channels, something that indicates a discoverability problem (Linde et al., 2011). Lee et al. (2015) uses the IR at Florida State University as a case study to examine true accessibility to material. They found that IRs inflate their accessibility numbers by equating availability with accessibility, with some available material limiting access to it (Lee et al., 2015). Arlitsch and O’Brien (2012) analyze discoverability through Google Scholar by detailing a comprehensive and multi-phased approach to increasing discoverability at the University at Utah through careful testing and evidence-based enhancement. Most researchers use Google (65% market share of general searches) as the primary search tool to begin research (only 2% researchers surveyed used library sites as the primary tool in 2005 and 0% in 2010), this paper details many methods for improving results for IR discoverability on the most used tool (Arlitsch and O’Brien, 2012). With an average indexing ratio in Google Scholar of 30%, it shows that metadata must be crafted to interact with Google Scholar’s crawlers to make material accessible (Arlitsch and O’Brien, 2012).

Institutional repositories can make an important impact on the dissemination of scholarly work. In its current state, the actual dissemination of material is hindered by three major issues: researchers do not know they exist, current copyright practices dissuade individuals from storing material in IRs, and discoverability is hindered by metadata
practices that are not standardized to maximize use with the search engines used by researchers. To address these issues libraries must take an active role in fostering development in each area.

Before libraries dedicate resources to this endeavor, the issue of what types of documents should be stored must be discussed. As stated by Arlitsch and O’Brien (2012, p.70):

“Since use of the site operator in GS reveals only the primary versions of the articles, the average indexing ratio of 30 percent indicates that most IRs do not contain very many primary articles. This raises some interesting questions about the purpose of IRs. Specifically, how much value is really derived from having pre-prints in the IR, given the amount of labor required to put them there, particularly if the primary publisher is open access as well?”

It is clear that duplicate copies held in the IR do not have the same impact as other materials. The materials that have the greatest impact are ones that are not digitally available otherwise, as shown by Koler-Povh et al. (2014). Before implementing any changes to IR practices, the institutions need to refine the missions of their IR to maximize the impact of the work of those who run them. Mission statements should go beyond the ambiguous definition that an IR is simply a container of documents, and explain what types of material should be prioritized and where professionals should focus their development.

The potential value of IRs is exciting to consider. In their simplest forms, these digital repositories allow for all users to access material through any of their internet capable devices. Instead of travel, and time costs associated with visiting a physical location, these virtual structures give anyone with access to internet a portal to instantly view documents and data. At the heart of the IR should be the user. So when Wacha and Wisner (2011) expressed concern about the library goals of IRs detracting from the open access movement, what library professionals should take away from that concern is that library goals should always place the user at the center of the equation. In this case, working to make previously inaccessible material available and discoverable digitally is paramount to the IR’s success.

Finally, when looking towards the future the IR has great potential to foster novel methods of research. As the need for permanent data links grows for grant funding, the amount of easily accessible data will continue to grow. Where IRs could play a major role is connecting this data across institutions. The concepts of accessibility and discoverability could extend beyond the search engines to consider how to make their data easily mined for large meta-analyses. Advancements in artificial intelligence and machine learning have placed a new focus on large data, something that IRs could extend beyond the search engines to consider how to make their data easily mined for large meta-analyses.

The Serials Librarian


AUTHOR QUERIES

Query: AQ1: Please review the table of contributors below and confirm that the first and last names are structured correctly and that the authors are listed in the correct order of contribution. This check is to ensure that your names will appear correctly online and when the article is indexed.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Prefix</th>
<th>Given name(s)</th>
<th>Surname</th>
<th>Suffix</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>Miriam L.</td>
<td>Matteson</td>
<td></td>
</tr>
</tbody>
</table>

Response: 1 | W. Reed | Jones2 | Miriam L. | Matteson

Query: AQ2: At the top of the manuscript the word “Article” appears; this is actually a column.

Response: Resolved

Query: AQ3: Please insert zip code in correspondence paragraph.

Response: Resolved

Query: AQ4: Should “Column description.” continue to be included at the beginning of the column’s abstract? And, should the part of the abstract beginning with “Reed Jones” be separated, to appear as the actual article’s abstract, distinct from the column description abstract? Response: Resolved

Query: AQ5: The spelling of “Dawson & Young, 2016” has been changed “Dawson & Yang, 2016” to match the entry in the references list. Please provide revisions if this is incorrect. Response: Resolved

Query: AQ6: The reference “Hahn & Wyatt, 2014” is listed in the references list but is not cited in the text. Please either cite the reference or remove it from the references list. Response: Can be removed. Unable to alter text on references

Query: AQ7: The reference “Waugh et al., 2015” is listed in the references list but is not cited in the text. Please either cite the reference or remove it from the references list. Response: Can be removed. Unable to alter text on references

Query: AQ8: The reference “Zheng Ye & Yu, 2015” is listed in the references list but is not cited in the text. Please either cite the reference or remove it from the references list. Response: Can be removed. Unable to alter text on references

Query: AQ9: The ORCID details of the authors have been validated against ORCID registry. Please check the ORCID ID details of the authors. Response: Resolved

Query: AQ10: As per style ‘doi’ is mandatory for journal type references. Please provide doi for references wherever applicable. Response: doi added. Lee et al. and Russell et al. do not have doi.

Query: AQ11: Please check and resupply corresponding author email address if the given is inaccurate. Response: Resolved