národní
úložiště
šedé
literatury

Role akademické knihovny v šíření šedé literatury – případová studie: A. Mickiewicz University Repository

Rychlik, Małgorzata
2015

Dostupný z http://www.nusl.cz/ntk/nusl-200843

THE ROLE OF THE ACADEMIC LIBRARY IN DISSEMINATING GREY LITERATURE – ADAM MICKIEWICZ UNIVERSITY REPOSITORY AS A CASE STUDY

Małgorzata Rychlik

rychlik@amu.edu.pl

Poznań Uniersity Library, Poland

Abstract

One of the main and crucial strategic projects that any university, or an institution of higher learning, has been recently involved in is the creation and development of an institutional repository. The AMUR repository is an open web-based archive of scholarly material that originated in 2010 and has aggregated a collection of more than 12,000 diverse digital objects. This paper presents the types of grey literature that have been archived in the repository. The following key elements concerning grey digital objects are presented: policy of the repository, legal issues, standards for metadata and the language coverage. In addition, usage data are provided.

Keywords

Adam Mickiewicz University Repository, AMUR, Grey Literature, Open Access, Usage Statistics, Scientific Communication

Introduction

The libraries of universities and research institutions all over the world contribute to development of scientific communication by launching OA repositories, among others. As a result, the intellectual output of an institution can be widely visible in the world. OA repositories collect, preserve and disseminate different types of digital objects. These objects include those distributed by traditional channels, such as articles and books published by traditional publishing houses. However, part of digital content archived in repositories can be described as the so-called underground literature (Boukacem-Zeghmouri, Schöpfel, 2005). Until recently, that type of content has not been fully indexed and could be encountered usually in the "Invisible Web" area (Derfert-Wolf, 2007). Thanks to OA repositories that comply with the Open Access Initiative Protocol for Metadata Harvesting and provide persistent URLs, it is now possible to search and retrieve materials defined as grey literature. For the purpose of this article I will use currently most frequently used New York definition (Myška, Šavelka, 2013) that characterizes grey literature as follows: "[Grey literature is] that which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers i.e., where publishing is not the primary activity of the producing body" (Schöpfel, 2010).

An analysis of the content types in OpenDOAR[1] shows that grey literature is by no means a marginal part of the content. Among five most frequent content types of OA repositories, three of them belong to a group defined as grey literature. At present, OpenDOAR indexes 2,973 OA repositories. Electronic Theses and Dissertations (abbrev. ETD) are collected by 1,643 OA archives, unpublished reports by 1,078 and conference and workshop papers by 1,063 OA repositories (Figure 1).
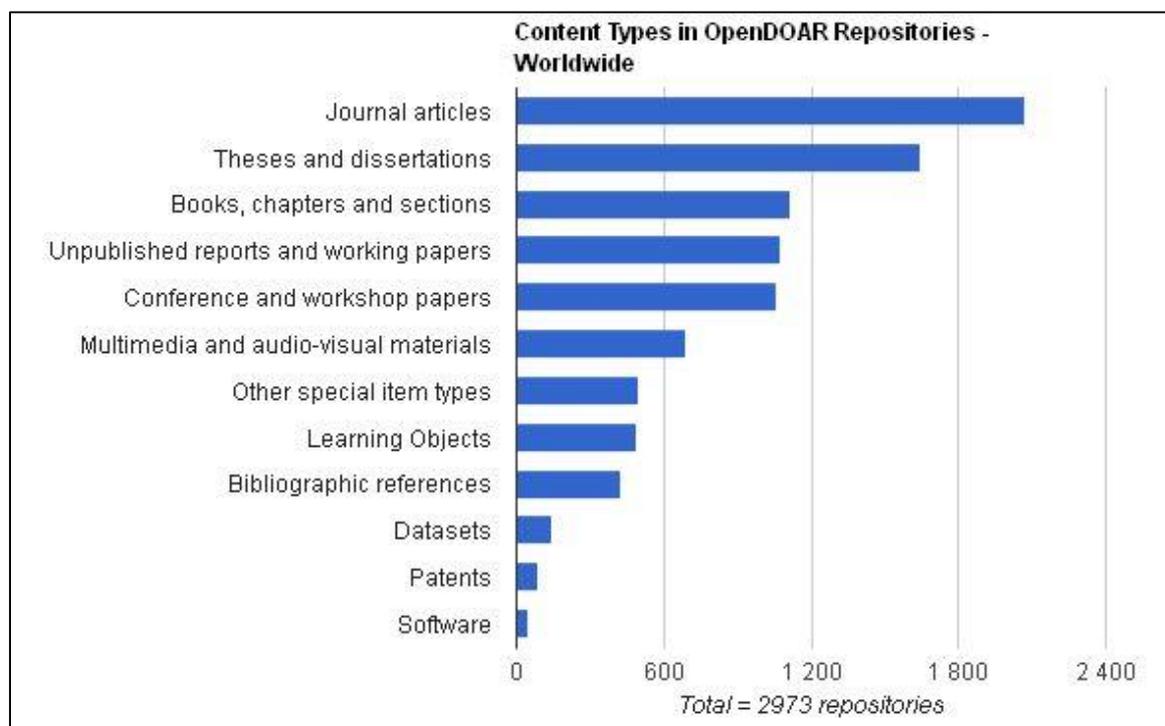

Figure 1 Content types in 2,973 OA repositories

---

[1] **http://www.opendoar.org/index.html**

Adam Mickiewicz University Repository (AMUR)[2]

The first institutional repository in Poland was established at the Adam Mickiewicz University in Poznań in 2010. The University is one of the finest Polish universities. More than 40,000 students and 1,300 PhD students attend the University. There are over 3,000 scholars employed. The AMUR repository is a big institutional archive with more than 12,000 digital objects. It ranks 155th among 2,275 indexed repositories in the Ranking Web of Repositories (July 2015 edition). The ranking places the AMUR platform in the first place among Polish repositories. This repository is based on DSpace software.

Types of grey literature in the AMUR repository

Communities and collection in the AMUR repository have a hierarchical structure. The basic framework of the repository is composed of 15 communities of the university faculties plus the community of journals published at the university and the community of ETD. About 89% of digital objects collected in the repository accounts for white literature (articles, books and book chapters). The remaining 11% belongs to grey literature. ETD outnumbers the remaining collections that archive several dozen digital items each (Figure 2). The total number of grey items in the AMUR repository is 1,411.
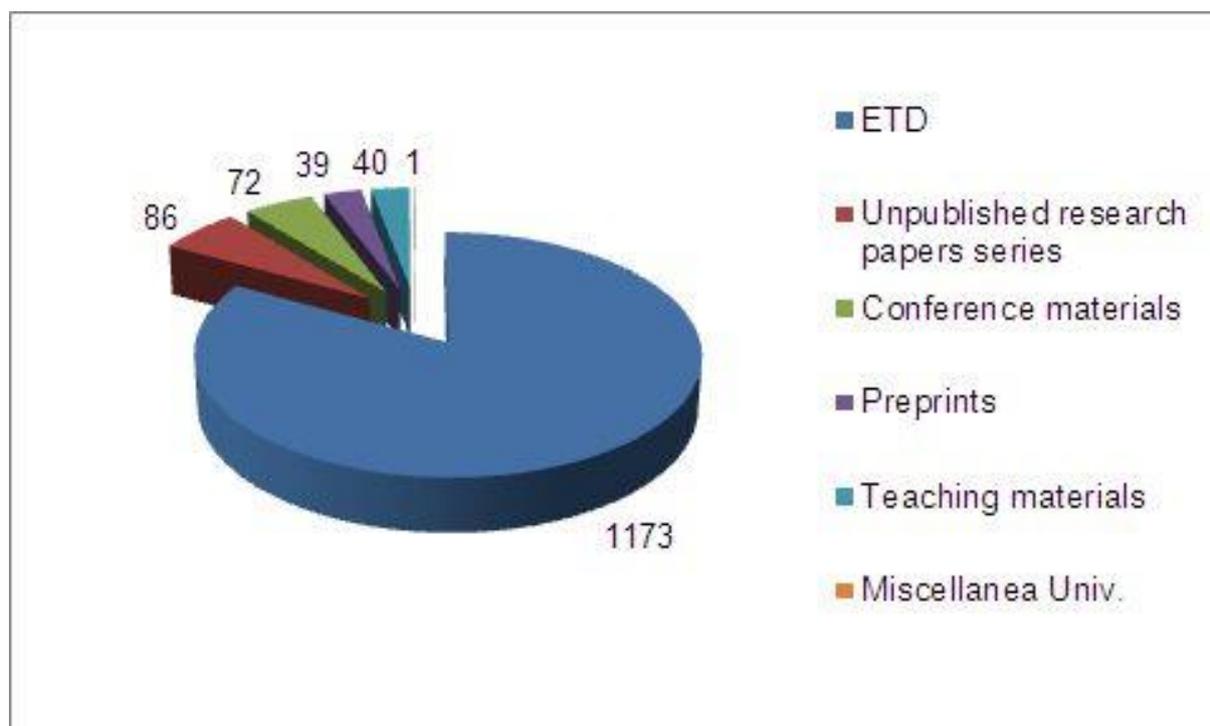

Figure 2 The share of different types of grey literature in the AMUR repository

---

[2] http://repozytorium.amu.edu.pl

Policy

One of the main aims of OA repositories is to disseminate the scientific and scholarly output of an institution. And one of the primary ways to improve content acquisition is to implement the open access mandate. This requires researchers to make their works open access by self-archiving. So far, more than 500[3] research organizations around the world have obliged scholars to deposit their works in OA repositories. It is essential to notice that these OA mandates are not always fully open and are subject to internal regulations of an institution. In the AMUR repository, there is a regulation implemented by the Rector of Adam Mickiewicz University which says that all the PhD theses defended at the University since 2010 have to be deposited in the repository. Each doctoral candidate signs a non-exclusive license and grants either the right to use PhD thesis in open access or limits the use of PhD thesis to just on-campus-only access. Hence, about 7% of the content of the repository is not OA.

In the AMUR repository all white content is open access, but as regards grey literature the percentage of PhD theses with restricted access is quite high. As many as 820 of ETD are not open due to on-campus restrictions. It gives 59% of grey literature as non-OA, whereas 41% remains open access.

Metadata

Many repositories implement the Dublin Core scheme composed of 15 classic metadata set. The scheme provides the ability to expand by adding qualifiers (subattributes to specify the main attribute). This allows the order of the attributes to be changed. Selected data can be modified, which may be useful for particular local applications. However, it is essential to realize that such local solutions often bring information noise. In the AMUR repository, we use a set of 15 elements (title, creator, subject, description, publisher, contributor, date, type, format, identifier, source, language, relation, coverage and rights). One specific element is created as regards grey literature - dc: thesis supervisor. This element consists of the name of a PhD thesis supervisor.

Legal issues

Generally, digital content of repositories is protected by copyright. In the Open Science area there is a distinction between two kinds of OA: gratis OA and libre OA. The one is defined as free, perpetual, open access which grants the right to use the work in the frame of fair use (i.e. for personal use and didactic purposes) to all users (Siewicz, 2012). Gratis OA occurs widely in OA repositories, although better and more efficient for the users would be the other type, i.e. libre OA that removes not only price barriers (as gratis OA does), but also at least some permission barriers (Suber, 2012).

For users, it would be much easier to have access to a work which is offered under a free license. Hence, it is really essential to make scholars aware of the necessity of possessing corresponding rights enabling them the use of adequate free licenses. As it happens, an author who does not possess such rights, offers his/her work under the CC license (Polčák, 2010). Many authors while self-archiving their works in OA repositories are in the possession

---

[3] **http://roarmap.eprints.org/**

of publisher's agreement for depositing the work only in the repository (gratis OA). Therefore, gratis OA is much more common than libre OA in OA repositories.

In the AMUR repository, approximately 10% of the whole content is offered under CC licenses. In the case of grey literature only 41 digital objects are under the CC licenses.

In our repository, each PhD candidate signs a non-exclusive license (see Policy). This results from the lack of regulations concerning PhD theses both in the Act of 4 February 1994 on Copyright and Related Rights as well as in the Act of 27 July 2005 on Law of Higher Education.

Usage data [4]

Usage data provide information about the use and access to works deposited in repositories. The data enable the use of repositories content to be measured and help in promoting the content. Usage statistics are gathered at the level of individual objects as well as they are available for the whole content within its different aspects.

While analysing usage statistics of grey literature in the AMUR repository we can see that both ETD and teaching materials are downloaded at similar levels - their average number of downloads is around 1,500 per item. Conference materials and preprints have got much lower number of downloads per item. Both, Miscellanea Universitatis and the collection of unpublished research papers series represent the extreme results - the highest and the lowest (Figure 3).
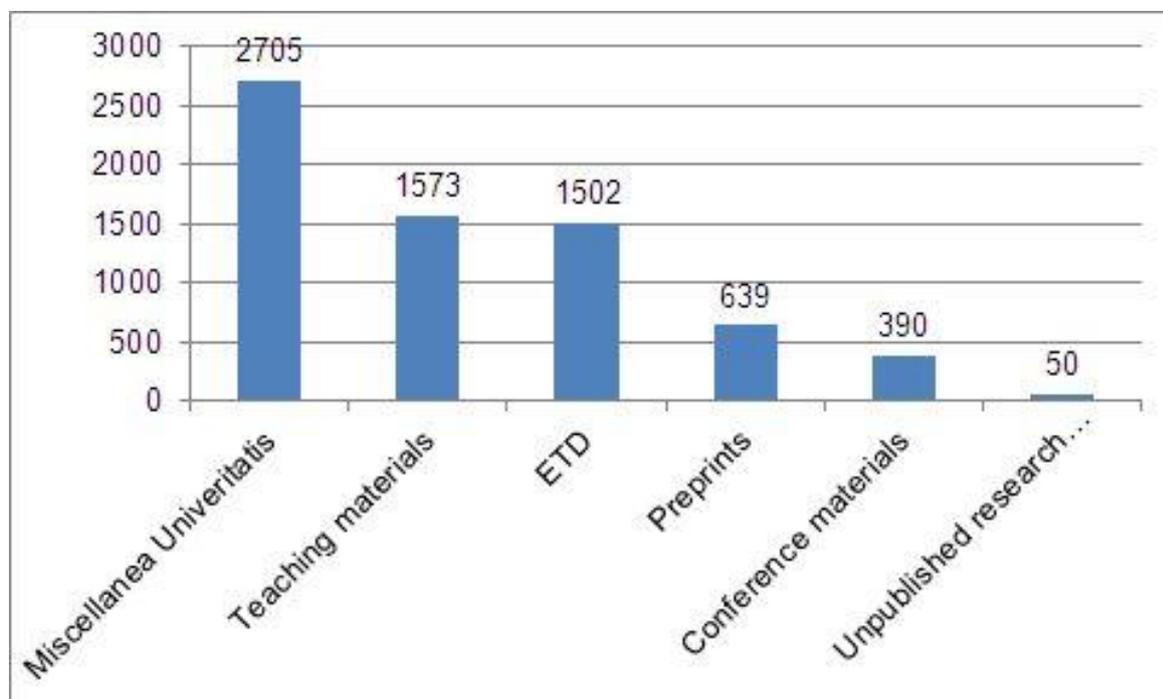

Figure 3 Average number of downloads per item of grey literature in the AMUR repository

[4] Usage data are provided only for grey OA items in AMUR (820 non-OA PhD theses are excluded)

While comparing usage statistics of white and grey literature we can observe that grey items are downloaded more frequently than white, which is well visible (Figure 4). For a comparison, I used three most frequently downloaded types of the white content (articles, books and book chapters), and three most frequently downloaded types of the grey content (teaching materials, ETD and preprints). I excluded Miscellanea Universitatis as this collection contains so far only one item and therefore is incomparable. The average number of downloads for grey literature is 1,238, whereas for white it is 579,7.
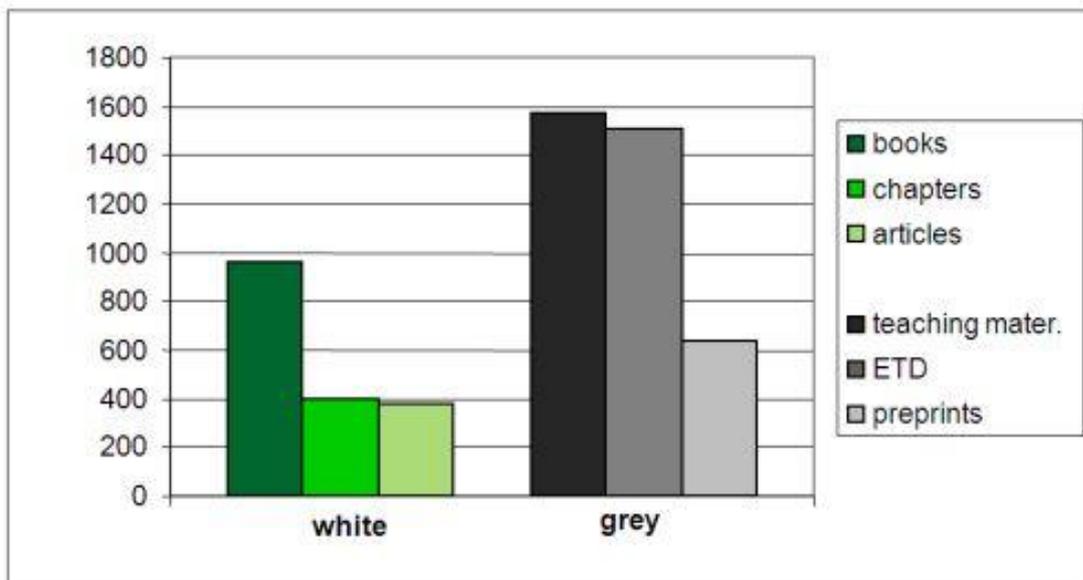

Figure 4 Comparison of average number of downloads between grey and white literature

Our usage statistics can help determine our users. We can trace their IP addresses and, based on download statistics of grey literature, are able to say that 50% of our "grey" users are from Poland, 17% from the States and the remaining 33% come from other countries (all continents are represented). Download origins are related to language coverage. The digital objects in the primary language of the repository (and most often the country) are more frequently downloaded from IP addresses located in that country than from other addresses. As to the language coverage in our repository, grey items written in Polish (70%) outnumber those written in English or other languages (Figure 5).
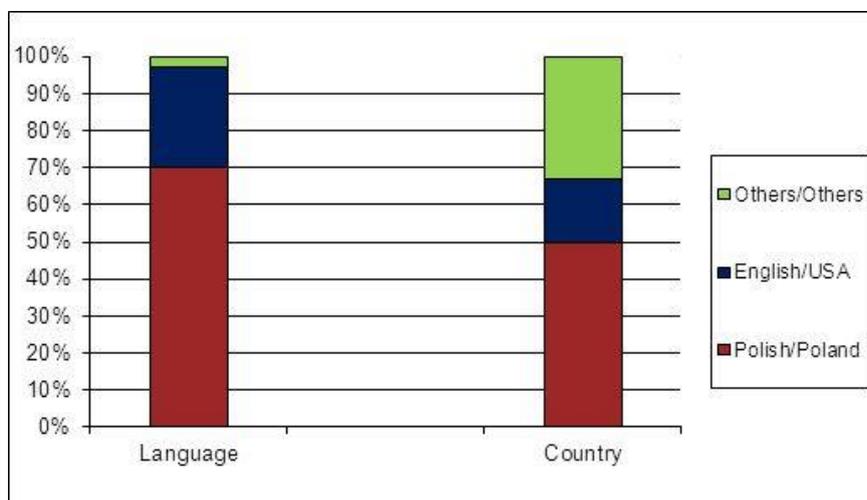

Figure 5 Language coverage of grey literature in the AMUR repository and download origins (by countries)

Conclusions

A determination of the core issues concerning grey content in an OA repository is a vital step towards better understanding of the role of grey literature.

- In the case of AMUR, the ratio of the total amount of digital items to grey literature ratio is almost 10 to 1. As we see in Figure 4, the average number of downloads of grey literature is much higher than that representing white literature, therefore we should implement a promotion strategy which would be helpful in gaining more grey literature into the repository.
- Undoubtedly, legal issues are the ones of most frequent barriers relating to self-archiving. The awareness of scholars in the context of granting free licenses needs to be raised. The lack of access to 820 PhD theses is an immense disadvantage for our users. Unfortunately, this situation is not going not change considerably for some time until a new law concerning theses and dissertations has been enacted in Poland.
- There are some decisive elements concerning metadata scheme that we should work on. Unquestionably, new elements regarding conference materials should be created.

A robust and efficient repository would be a project based on a cooperation between librarians and researchers. It is important to strengthen a synergy between these two groups. Librarians should remain responsive to new developments in the Open Science area. They should implement new tools and standards which can be useful in disseminating knowledge. While the task of scholars is to actively build a repository content by self-archiving their works (both white and grey). A growing number of scholars believes that such an activity can bring tangible benefits to both researchers and their institutions in terms of greater visibility and citation advantage.

References

BOUKACEM-ZEGHMOURI, Chérifa and Joachim SCHÖPFEL. *Access and document supply: a comparative study of grey literature*. 7th International Conference on Grey Literature, Dec 2005, Nancy [online]. [cit. 2014-09-03]. Available from: **http://archivesic.ccsd.cnrs.fr/sic_00091843/document**.

DERFERT-WOLF, Lidia. Odkrywanie niewidzialnych zasobów sieci. In: *II seminarium z cyklu Infobroker: Wytwarzanie i przetwarzanie cyfrowych informacji, Warszawa 17 kwietnia 2007*. [online]. [cit. 2014-09-14]. Available from: **http://eprints.rclis.org/8862/1/derfert_CPI.pdf**.

MYŠKA, Matěj, ŠAVELKA, Jaromír. *A model framework for publishing grey literature in Open Access* [online]. 2013, vol. 4, no. 2, p. 104-115 [cit. 2011-09-05]. ISSN: 2190-3387. Available from: **http://www.jipitec.eu/issues/jipitec-4-2-2013/3744/myska-savelka.pdf**.

POLČÁK, Radim. Legal Aspects of Grey Literature. In: P. Pejšová ed.. *Grey literature repositories* [online]. Zlín: VeRBuM, 2010 [cit. 2014-09-05] Available from: **http://eprints.rclis.org/8862/1/derfert_CPI.pdf**.

8th Conference on Grey Literature and Repositories: proceedings [online]. Prague: National Library of Technology, 2015 [cit. 2015-12-15]. Available from: **http://nrgl.techlib.cz/index.php/Proceedings**. ISSN 2336-5021.

SCHÖPFEL, Joachim. *Towards a Prague Definition of Grey Literature Twelfth International Conference on Grey Literature: Transparency in Grey Literature,. Prague, 6-7 December 2010, Czech Republic* [online]. 2010 [cit 2014-09-03]. Available from: **http://archivesic.ccsd.cnrs.fr/sic_00581570/document**.

SIEWICZ, Krzysztof. Otwarty dostęp do publikacji naukowych. Kwestie prawne [online]. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego, 2012 [cit. 2014-09-03]. Available from: **http://repozytorium.ceon.pl/bitstream/handle/123456789/335/K_Siewicz_Otwarty_doste p_do_publikacji_naukowych.pdf?sequence=4**.

SUBER, Peter. *Open Acces* [online]. Cambridge: The MIT Press, 2012 [cit. 2014-09-14]. The MIT Press Essential Knowledge Series. ISBN 978-0-262-51763-8. Available from: **https://mitpress.mit.edu/sites/default/files/9780262517638_Open_Access_PDF_Version .pdf**.