ADAPTION OF OPEN SCIENCE PARADIGM AT GDANSK UNIVERSITY OF TECHNOLOGY

TOMASZ BOIŃSKI

Department of Computer Architecture
Faculty of Electronics, Telecommunications and Informatics
Gdansk University of Technology
Narutowicza 11/12, 80-233 Gdansk, Poland

(received: 22 May 2015; revised: 26 June 2015;
accepted: 3 July 2015; published online: 1 October 2015)

Abstract: Open Access plays a growing role in sharing the results of today's research. The movement is very strong and many Universities, including the Gdansk University of Technology, are adapting this approach. This paper presents some general information about the Open Access movement and the approach adopted by the Gdansk University of Technology. The components of the Open Science platform as well as the procedures are described.

Keywords: open science, repositories

1. Introduction

Open Access plays a growing role in sharing the results of today's research. The movement is very strong and many Universities are adapting this approach. Some, the including Gdansk University of Technology (GUT), try to go further by providing not only access to their publications, but showing general openness in project implementation, contacts with other research institutions or finally cooperating with business. In this way Open Science is born, where not only publications, but information about projects, courses, research data and the software developed during this research is available.

Open Access movement is the foundation for Open Science. However, it is poorly understood by the scientific community and often misunderstood as something that forces researchers to give away everything that they have done for free and limits their possibility of publishing their work in renowned journals, as it is usually required that they should sign a copyright transfer form.

With the C^{2}NIWA[1] project (Centre of Competence for Novel Infrastructure of Workable Applications) the Gdansk University of Technology is finally able to embrace the Open Science paradigm. The aim of the C^{2}NIWA project is to provide comprehensive services in the area of modern platforms for creating applications (parallel, distributed, and mobile). The Centre will offer its users an advanced IT infrastructure, platforms for application development and wide consulting services. Open Science is a key element of the projects infrastructure.

The C^{2}NIWA project will provide all the necessary tools and systems allowing implementation of the Open Science paradigm. This will spring implementation of the procedures needed to fully implement the process.

This paper presents the basic principles standing behind Open Access and the tries do dispel the doubts behind them. Later the solution adopted by the Gdansk University of Technology is presented. The structure of the paper is as follows. Section 2 describes the principles behind Open Access and presents approaches adopted around the world. Section 3 presents the approach adopted by the Gdansk University of Technology. Section 4 describes an additional system that the University uses and that will be integrated into a complex Open Science platform. Section 5 proposes the procedures needed to publish a document in the repository. Finally some conclusions are given

2. The principles behind Open Science

The term Open Access [1, 2] has been defined in three public declarations: Budapest Open Access Initiative[2], Bethesda Statement of Open Access Publishing[3] and Berlin Declaration on Open Access[4]. In general Open Access is understood as wide, easy, public and free access to scientific publications. Furthermore the copyright owner should allow everybody to copy, use, redistribute, perform, modify and create derivative works [3][5].

2.1. The Green Way and The Golden Way

In Open Access two approaches to publishing can be distinguished: the golden way, where Open Access is realised by the publisher independently from the business model, and the green way, where Open Access is realised using repositories [4, 5]. With the green way, it is the self-archiving concept that is often used – the author publishes his or her papers to an open repository independently from the publisher (sometimes even despite the fact that the publisher does not do that).

1. http://niwa.gda.pl
2. http://www.budapestopenaccessinitiative.org/read
3. http://dash.harvard.edu/handle/1/4725199
4. http://openaccess.mpg.de/Berlin-Declaration
5. http://repozytorium.ceon.pl/handle/123456789/65

Both of these solutions complement each other. If an author does not have the possibility of publishing in one way, then there always exists another way. The researcher does not have to select prestige over openness, especially that many publishers allow publishing using the green way[6].

In a perfect world all reviewed and published papers would be available using the golden way. However, some of the publishers are not interested in providing free and open access to their papers as this would require a substantial change in their business model. Fortunately most of the publishers allow self publishing of at least a preprint form of the accepted paper. This however is a fact that is not too keenly advertised by publishers [1].

2.2. What should be published?

The basic principles standing behind Open Access are proposed with regard to articles and book chapters. Nonetheless, many solutions, including the one implemented at the GUT, propose an extension of the set of published materials [1]. In general it is advised to extend the openness of publications with at least the openness of the test data and the source of software that was developed during the research. Publication of such works can increase the trust in the results and allow easy verification and comparison of different approaches.

2.3. Bottom-up or top-down?

Open Access is generally a bottom-up initiative. In most of the western countries, with a long and established culture of knowledge exchange, the move towards an Open Access model is usually sparked by activities of downstream employees. It is in the researchers' best interest to increase the citation levels and availability of their work. However, in eastern countries openness is often not considered at all – all that matters is how prestigious the journal or conference is and how many points it is worth. The citation level has usually been a secondary factor. Many people still think that they will loose potential financial income when they "give away their work for free".

A further problem is the added bureaucratic strain put on researchers. Due to the amount of paperwork they often do not think about the openness of their research papers. Even if they would like to publish papers using Open Access, they usually do not want to spend more time looking for repositories or browsing the policies of publishers. EUA's report[7] shows that that encouragement policies have limited effect on the openness of publications, and only mandatory policies seem to boost self-archiving of research articles in the institutions' repositories[8,9].

6. The status can be checked using http://www.sherpa.ac.uk/romeo/

7. http://eua.be/Libraries/Publications_homepage_list/Open_access_report_v3.sflb.ashx

8. http://www.openscholarship.org/jcms/c_6215/effectiveness-of-open-access-policies

9. http://www.openscholarship.org/jcms/c_6226/en/open-access-policies-for-universities-and-research-institutions

Due to these reasons both local and European legislation try to impose Open Access for at least parts of research results that are conducted using public funds.

3. Open Science at GUT

3.1. Open Access Policy

The Gdansk University of Technology started preliminary work on implementing the Open Access policy several years ago. With the C^{2}NIWA project we now have the possibilities to implement the policy. The main points of the policy are:

1. the Gdansk University of Technology will adopt the policy based on the green way with encouragement to publish using the golden way;
2. The policy will be mandatory for scientific publications, chapters in monographs and complete monographs. The encouragement policy will be applied to different types of publications;
3. The repository should contain every scientific paper published by faculty members independently from its openness;
4. By default, the submitted work will be made public immediately. Any embargoes need to be justified by the author;
5. Whenever possible, final, reviewed versions should be published;
6. Exclusion from the policy requires a request in writing;
7. The papers should be published by the original authors.

3.2. Technical Aspects

The aim of the Gdansk University of Technology is to go beyond Open Access. The institutional repository is only a small part of the whole platform (Figure 1).

The Open Science platform comprises four main components:

1. The institutional repository that is built on top of widely used DSpace system[10] adapted to the needs of the University. The main task of the repository is storage and distribution of scientific papers, but it is also used as a catalogue of available platforms, courses, services, tutorials and solutions offered by the University;
2. A project support system based on Redmine[11] and GIT[12]/SVN[13] repositories. It allows joint project implementation and code integration. It also serves as a repository of projects done by the University and its employees and students;

10. http://www.dspace.org/
11. http://www.redmine.org/
12. https://git-scm.com/
13. https://subversion.apache.org/

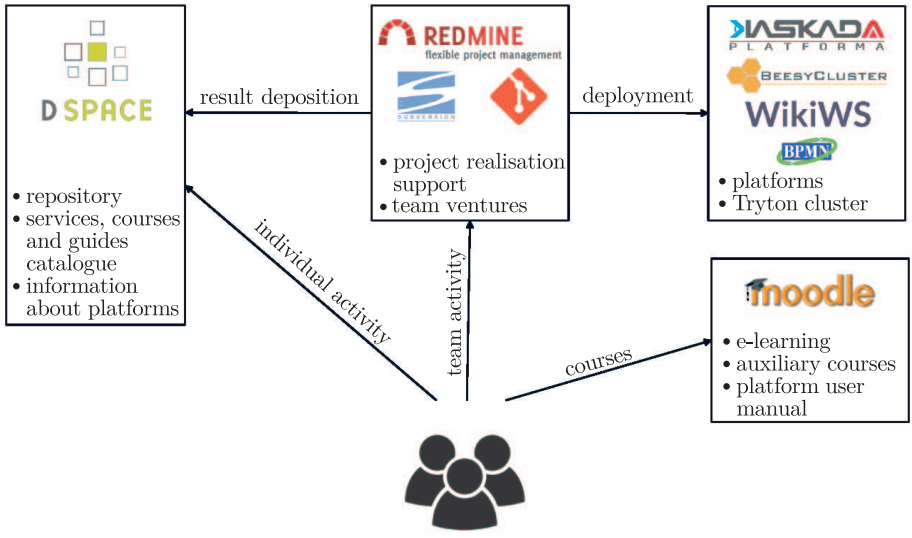

Figure 1. Open Science systems

3. An e-learning platform based on Moodle, used to publish courses strengthening soft skills of team members and teaching how to use rest of the systems in Open Science. All the courses are free;
4. Platforms (KASKADA [6], Wiki-WS [7], BeesyCluster [8], BPM) that support development and delivery of software and services, they give access to the resources of C^2NIWA, mainly the Tryton Cluster[14]. The Wiki-WS platform plays a special role in the Open Science as it is freely available and can be used for easy deployment of WebServices based solutions.

All those systems form a complete repository of activities conducted at the Gdansk University of Technology and thus allow easy communication and results sharing between research teams.

4. External systems

It is planned to connect Open Science systems with external systems already operating on the Gdansk University of Technology to further improve the quality of the provided solutions. In this way the whole system will be integrated with the current solutions implemented at the University and will fit in with the administrative structure (Figure 2).

First of all, proper implementation of the Open Access paradigm requires simplification of the deposition process. Without it, the bureaucratic burden would hinder the whole process. Integration with the MojaPG[15] system is needed for that purpose. During the period of their employment researchers employed at the GUT are required to register every publication written during the research

14. http://task.gda.pl/kdm/sprzet/tryton/
15. https://moja.pg.gda.pl/

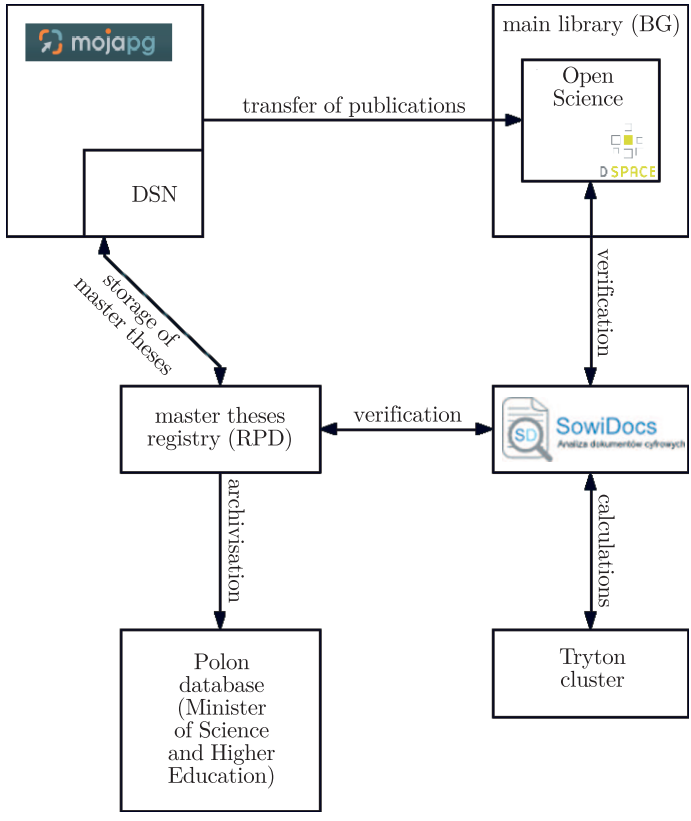

Figure 2. Open Science within University structure

process. Nonetheless, the internal database is simplified with regard to the needs of the Open Access repository. With some extensions, mainly an addition of some fields (Table 1), it can be used as an interface for the repository system allowing both registering the paper in the University database and the DSpace institutional repository with the verification process for both databases combined.

Another system that we are integrating is an anti-plagiarism SowiDocs system. By such integration we will manage to verify whether the paper or book published in the repository is not a plagiarism. Furthermore we will be able to verify also master theses and papers, if they do not plagiarize from previously published, publicly available documents. In this way the quality of the results generated by SowiDocs will also improve as its database will grow. This generates additional strain on the hardware – SowiDocs requires powerful hardware for calculating similarity measures between different documents. Due to that the calculations are performed on the Tryton Cluster. The possibility of connecting DSpace with the SowiDocs system will be also available as standalone plug-in allowing other universities to use DSpace to connect with our anti-plagiarism system.

Table 1. Fields from DSpace missing in MojaPG system

DSpace	MojaPG	Comment
Additional title	—	Not required
Day, month and year of publication	Year of publication	Only year is required
Publisher	—	May be required to fulfil copyrights
Cite reference	—	Can be generated
Sponsors	—	Can be stored in Additional information field
Description	—	Required
Attachments	—	Required
Embargo information	—	Required per attachment
CC and FLOSS licence	—	Required
Repository licence	—	Required

5. Proposed procedures

The deposition process requires specification of at least two main procedures:

- item deposition and verification;
- licence selection.

They include both the author and the University authority and will use other, external resources with regard to the repository.

5.1. Item deposition

The person responsible for publishing the paper should be the author or the representative of the authors. In case of older publications, to which the Gdansk University of Technology holds copyrights, the documents can be published by an employee of the Main Library. The deposition itself can be done in two ways:

- Through the MojaPG portal during the publication registration – the author, while registering his or her publication, enters all the data required to create records both in DSpace and MojaPG, this is a recommended way of depositing publications by the employees of the University;
- Directly using DSpace software where the document is registered using built-in wizard. This way should be used when depositing non-registered documents (like software, presentation, master theses, reports *etc.*) and when depositing older, already registered in MojaPG publications and when the deposition is done by people not working at the University.

Regardless of the method used to deposit the document it is being sent to SowiDocs for plagiarism verification. In parallel the document meta-data should be verified and corrected by an employee of the Department of Research (pol. *Dział Spraw Naukowych*, DSN). Currently the DSN employees perform verification

of records stored in MojaPG, this step should not add any additional burden, as all publications are currently verified in the same manner.

Only when the document is deposited directly via DSpace, it is then sent to SowiDocs, the DSN employees do not verify the meta-data as it does not need to meet the needs of DSN's classification.

The document is held back until the system sends the comparison results which are added to the meta-data of the attached items and DSN performed verification and correction of meta-data. Later, the item should be verified by employees of the Main Library (pol. *Biblioteka Główna*. BG). This verification step should take into account the score of the anti-plagiarism system and should verify the attachment of the proper version of the publication that is in compliance with the policy of the publisher of the paper. After the verification the library employee finally accepts the document and makes it visible in the repository.

5.2. Licence selection

The second procedure involves the licence selection for the deposited item. The author of the paper has 3 choices:

- Not to apply any licence – not recommended, in this case the work will be available based on the copyright law stating that the users of the repository will be able to read the paper or run the software but nothing more. There will be no permission to redistribute the work or create any derivative works;
- To apply one of the Creative Common licences – useful for any type of creation except software. It should be applied whenever possible and this automatically allows redistribution and usage of the work for creation of derivative works. By using the wizard built into the DSpace software the user can limit the rights of the repository users to some extent. It is recommended to provide permission at least to use the work for non-commercial purposes and allow creation of derivative works with the requirement to distribute them under the same condition as the original licence of the item. Nevertheless, the licence should comply with the rules of the original publisher;
- To apply one of the Free/Libre Open Source Licences (FLOSS) – recommended for assignment licence to the software. The user can choose between 2- and 3-clause BSD, MIT/X11, GPL and LGPL version 2 and 3. The type of the licence should be chosen dependent on the amount of rights that the authors would like to give the users. The exact algorithm of licence assignment is not trivial. Authors can however follow the flowchart in Figure 3 or consult one of the C^{2}NIWA employees.

6. Conclusions

Open Access plays a more and more important role in today's research process and is a key to a wider concept of Open Science. With growing costs of accessing publications alternative means of reaching for knowledge becomes

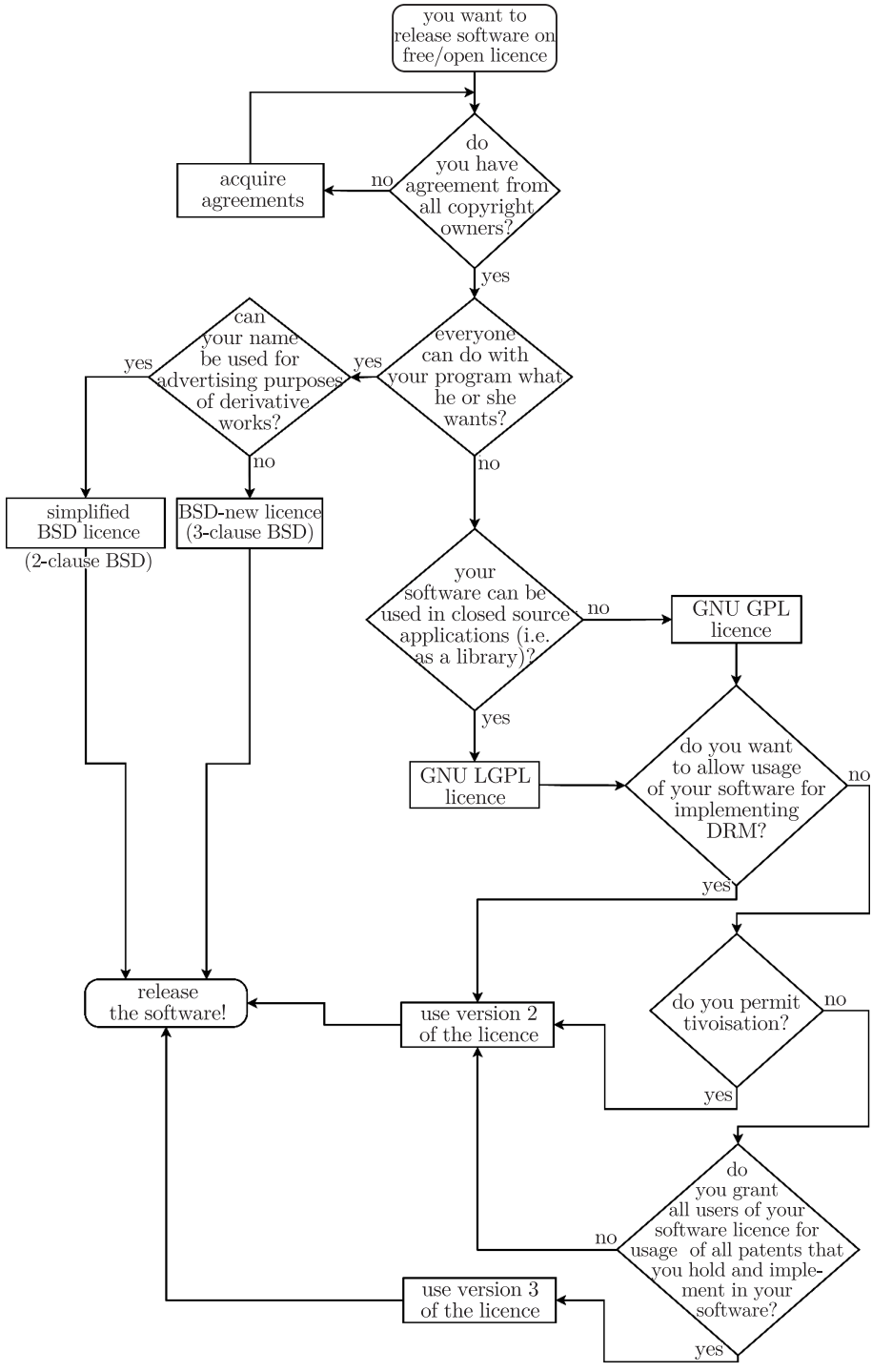

Figure 3. Free/Open licence recommendation

necessary. Modern technology, especially the Internet, gives us the opportunity to access the information and data in an easy fashion.

Traditional access to scientific papers has been guarded by many restrictions. The most important one is limiting access to articles by access fees. Fortunately with the Open Access movement this barrier is lowered and the results of scientific research are available to the extent that was never possible before. It is also troublesome to look for interesting papers when you have to browse different databases.

Most of the publishers already agree for self archiving and distribution of research results using the green way. More and more of them, supported by the growing awareness of researchers, are accepting the golden ways. With a better visibility and indexing possibility from search engines like Google Scholar the citation rate will get higher than ever.

Publishing articles using the Open Access model is not everything. With Open Access all the research results can be easily distributed, including software, test cases and test data. That in turn improves correctness and verifiability of scientific work.

With the C^2NIWA project the Gdansk University of Technology hopes to fully embrace the Open Science concept not only by giving access to the publications of its employees but also by providing easy access to its resources, the possibility to browse for solutions, take part in courses and finally create projects using modern platforms available on modern hardware.

Within the project we plan to establish fully fledged repositories where scientists will be able to deposit their work. Additionally we will enable licence verification and integration with systems that are currently used at the University. In the near future we plan to extend the policy to make deposition of publications mandatory while maintaining support for deposition of other types of work such as research data, software, *etc.* Furthermore, we would like to extend the availability of the repository to other Universities providing a general solution for the whole Pomeranian Province.

Acknowledgements

NIWA Centre of Competence for Novel Infrastructure of Workable Applications. Project co-funded by the European Union from the European Regional Development Fund within the Innovative Economy Programme, "Subsidy for innovation" Gdansk University of Technology, Narutowicza 11/12, 80-233 Gdansk, Poland.

References

[1] Suber P 2012 *Open Access*, The MIT Press
[2] Siewicz K 2012 *Open access to scientific publications: legal issues*, Wydawnictwa UW (in Polish)
[3] Hofmokl J, Tarkowski A, Bednarek-Michalska B, Siewicz K and Szprot J 2009 *Guide to open science*, Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego Uniwersytetu Warszawskiego (in Polish)

[4] Rychlik M and Karwasińska E 2007 *Institutional Repository as a factor supporting the development of science in the academic environment* (in Polish)
[5] Rybiński H, Skonieczny Ł, Koperwas J and Struk W 2014 *University Knowledge Base in the SYNAT project – the experience of Warsaw University of Technology*, INFOBAZY 2014 (in Polish)
[6] Krawczyk H and Proficz J 2012, Interactive Multimedia, InTech, Citeseer 289
[7] Krawczyk H and Downar M 2012, Intelligent Tools for Building a Scientific Information Platform, Springer 251
[8] Czarnul P 2013 *The Journal of Supercomputing*, Springer, **63** (1) 46