

FROM AN INSTITUTIONAL REPOSITORY TO THE BASE OF KNOWLEDGE - CASE STUDY

Weronika Kubrak

W.Kubrak@bg.pw.edu.pl

Warsaw University of Technology - Main Library, POLAND

This paper is licensed under the Creative Commons licence: CC-BY-SA-4.0 (<http://creativecommons.org/licenses/by-sa/4.0/>).

Abstract

The Warsaw University of Technology Base of Knowledge it is not only a institutional repository but also good place to promote the scientific activities of the university staff. Unquestionable feature of this Base is that it allows present not only published papers but also collect f.e. patents and projects documentations, professional activity of our staff and students dissertations. The paper presents the advantages of system which combines functions of a repository and the Base of Knowledge functions.

Keywords

Repositories, Base of Knowledge, University Repository, Scientific Output

Introduction

The main objective of the Warsaw University Base of Knowledge (WUT Base of Knowledge), based on the OMEGA-PSIR¹ system, is to present and disseminate research achievements of the University faculty and students, both in Poland and all over the world. It provides access to academic and scientific publications, reports, dissertations and theses, as well as to information on the current research at the Technical University of Warsaw. Thus, the Base of Knowledge contributes to widening contacts and developing cooperation between the WUT academic staff and scientific and business communities in Poland and abroad.

One of the most important parts of the system is a repository which archives metadata and digitized full-text documents, such as monographs, journal articles, book chapters, reports and papers, such as dissertations and theses which are necessary for awarding university degrees and academic titles. The data in the repository support the needs of the university research, promote the university achievements and are also used for internal and external reporting.

Effects of implementation and usefulness of the system

During the implementation of the OMEGA-PSIR system and its later development, its creators and editors² responsible for data entry encountered a number of problems which turned out to be a barrier to build a joint University repository. The most important and most common problem was how to convince the authorities of some faculties to stop using their own databases on scientific achievements of the WUT researchers - some, faculties use their own systems, often incompatible with the new one. Therefore, the priority was to convince the authorities of each faculty to stop using their databases and to transfer all data to the Base of Knowledge. A decisive argument for such change was the Resolution of the WUT Senate No. XLVIII / 2012 of November 21, 2012 concerning the rules for creating a central system for recording and archiving of the writing, publishing and teaching achievements of the University staff, doctoral student and students, all University units and the WUT Repository. Then followed resolutions to specify the activities and the level of responsibilities of particular groups responsible for creating the WUT Base of Knowledge.

Thanks to the support of the University authorities, the Base of Knowledge became the main source of information to serve both recording and presenting the scientific achievements of the WUT academic staff and students. The Base is also used for reporting and for promoting science and research.

Capabilities of the OMEGA-PSIR system

The main purpose of the software authors was to create an integrating system, accessible to everyone, which would not only include scientific publications and other documents but

¹ OMEGA-PSIT it's software designed by a team from the Warsaw University of Technology Faculty of Electronics and Information Technology and is used by the University repository. The same team is involved in developing the research Base of Knowledge.

² Each University faculty has to appoint a faculty editor – a person responsible for entering the data on publications submitted by authors. Currently, there are over 120 editors. There are between several and over a dozen editors in a faculty, depending on the faculty size and the number of publications.

would also enable using the data for reporting purposes. The system functionality can therefore be divided into several types:

1. The repository functions related to recording achievements of the university scientists, archiving the achievements (in accordance with the copyright protection) and provision of the following:
 - Books and chapters
 - Journal papers
 - Engineer's and Bachelor's theses
 - PhD theses
 - Researchers projects
2. Presentation of documentation on the WUT projects and patents
3. Database on presented papers and published conference proceedings
4. Reporting on internal needs of the University, including generating reports for the academic staff assessment
5. Reporting for the needs of external units which evaluate the University activities (e.g.: Pol-on³, MNiSW⁴, PBN⁵) and communication with external systems (e.g. Google Scholar)
6. Presentation of the University scientific achievements and transfer of knowledge within the university and outside
7. Presentation of experts from various fields of science
8. Building a tag cloud presenting areas of research throughout the University, its individual units and each staff member.

Base Of Knowledge as a source of information about the university units and scientific researchers

The faculty "editors", appointed by the faculty authorities, are responsible for entering the data from their respective faculties into the repository, which is and integral part of the University Base of Knowledge. The University academic staff are obliged to submit the data to the Base of Knowledge for the purpose of updating it. It is also the only source of obtaining the data necessary for assessing the staff and the University unit, which is in turn necessary for the national system of assessing the functioning of the Polish universities (Pol-on, PBN). The lack of up-date data on the units means that it will not be taken into account in the reports.

³ POL-on is an integrated information system for higher education, supporting the work of other Polish systems

⁴ MNiSW - Ministry of Science and Higher Education

⁵ PBN - Polish Scholarly Bibliography is a portal of the Polish Ministry of Science and Higher Education, collecting information about publications of Polish scientists and Polish and foreign scholarly journals.

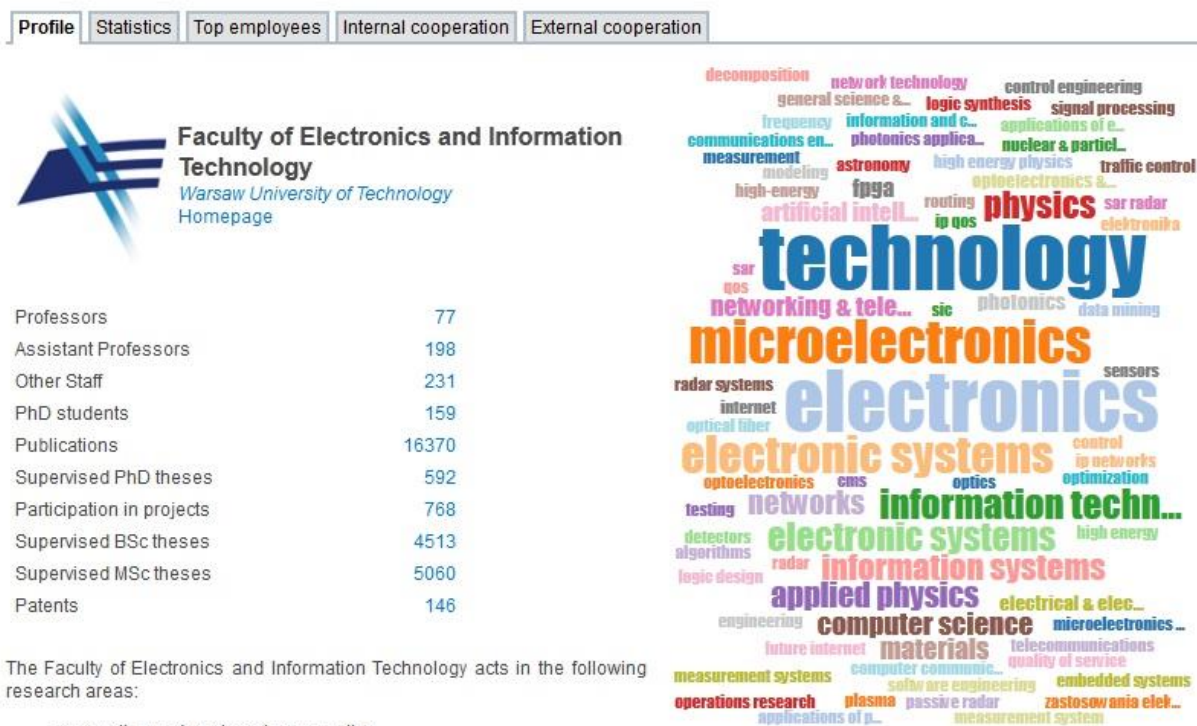


Figure 1 Profile presenting the characteristics of the Faculty

The figure shows an example of the faculty profile page where you can find information on not only the characteristics of the Faculty activities and the list of publications, but also:

- Statistics – graph demonstrating the increasing number of publications of the unit’s authors over the years and the points awarded for them in accordance with the Ministry’s regulations
- Top employees – graph presenting researchers with the highest number of publications in the unit
- Internal cooperation – schemata presenting all internal WUT units with which the unit cooperates
- External cooperation – refers to external units cooperating with the faculty
- BSc and MSc – list of BSc and MSc degree students thesis
- PhD – list of PhD students of the faculty
- Projects and Patents – list of projects implemented in the faculty and patents owned by WUT researchers.

The Cloud of Tags in profiles, generated at irregular frequency on the basis of key words added to the description of publications. Each keyword from the cloud of tags refers to publications linked to the given word.

The home page of the University researcher looks similar.

Profile
Publications
PhD
Projects
BSc and MSc
Activities
Citations
Statistics
Cooperation
Edit



Henryk Rybiński, PhD, DSc,
Professor
Professor
The Institute of Computer Science
Faculty of Electronics and Information Technology
Email: H.Rybinski@ii.pw.edu.pl
Phone: +48 22 234 7432, fax +48 22 234 6091
Room no: 204
Consultations: Monday 14.00-16.00



Researcher Report	
Publications	104
PhD theses	14
Participation in projects	41
Supervised BSc and MSc theses	35
Professional activity	10

h-index*: 12

MSc (1970), PhD (1974), DSc (1988), Tenured Professor (2001); Specialization: information systems; knowledge representation, data and text mining, databases, Professor, Director of the Institute (2008-), Head of Division of Information Systems (1994-2008), Co-ordinator of the Curriculum on Software Engineering and Information Systems (1994-2008), Co-ordinator of the Subject Class "Databases and Information Systems" (1995-2001), voting member of ACM and SIGMOD (1989-), Affiliate Member of IEEE (1990-1996); Member of several programme committees of international conferences and workshops, among others: IIS, ISMIS, IIPWM, AM, ISWC, RSFDGRC, RSKT, TKE, PKDD, PAKDD, MCD; member of CREST Working Group; expert and consultant of many UN agencies and European Commission; Member of Informatics Committee at Polish Academy of Science (2011-); Editorial Board Member of the Journal of Intelligent Information Systems (2012-); Editorial Board Member of the International Journal of Social Network Mining (2012-); Chair of the Rector's Committee for the strategy of developing ICT infrastructure for WUT in 2013-2020.

Figure 2 Researcher's profile

Similar to the faculty profile, also in this case, all information concerning the scientific activities of the researcher are divided into groups and provide the following information:

- Profiles – provides contact information and the characteristics of scientific activity
- Publications – list of researcher's publications
- PhD – list of doctoral theses promoted by the scientist
- Projects – list of metadata and attached documents, presenting scientific projects in which the researcher participated
- Activities - information on memberships, e.g. in organization/s, participation in editorial committees, committees organizing conferences or seminars
- Citation – statistics related to the author's most frequently cited publication. In addition, on the basis of this data, the Hirsch index is calculated - approximate calculation obtained in the Repository based on the scientist's publications (including auto citations) in the Repository and the Internet information analysis. The value is close to the value obtained with the Publish or Perish system. In general, it is higher than the value given by the Scopus or Web of Science sites.
- Statistics – the same as in the case of the faculty profile, the tab presents a graph showing increasing number of publications and the Ministry points awarded over the years

- Cooperation – graph shows scientific co-authors of the researcher's publications and researchers cooperating in projects implementation and patents development

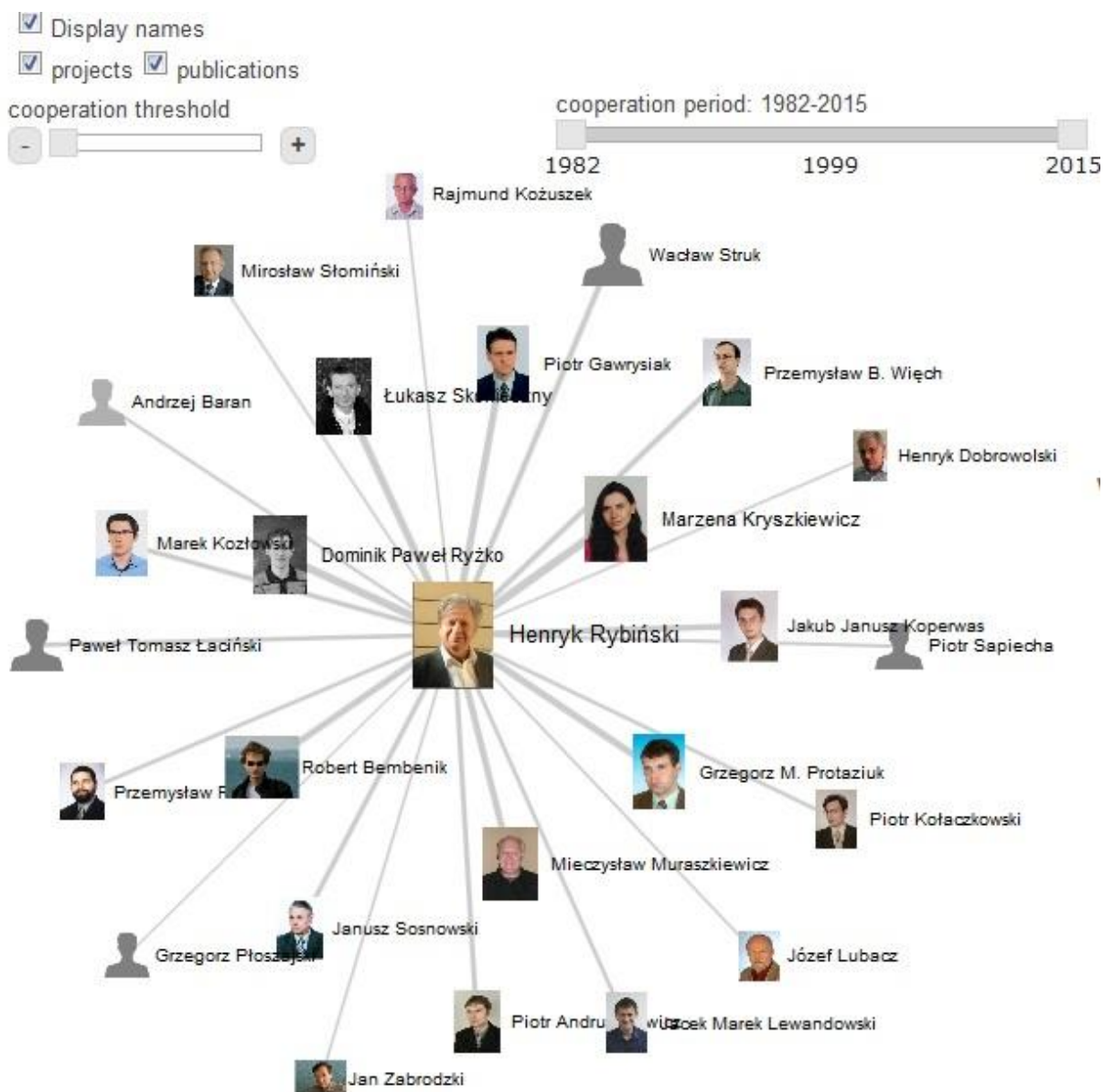


Figure 3 Cooperation graph in researcher's profile

Base of Knowledge used for reporting

The final function of the Base of Knowledge discussed in this paper is the capacity of generating various types of reports. The first and the simplest one is a report on the overall scientific activities of the researcher. Such report includes the University logo and can be attached to e.g. documents needed to assess the researcher while applying for professional

advancement. This document is generated as a PDF⁶ file and it can be also used as part of the researcher's resume. In addition, the repository can generate reports necessary for preparing the unit annual report or if needed by the faculty dean. Depending on the type of report, it can be generated as an Excel file, CSV⁷, Bibtext⁸ or as a bibliography presented on the Base of Knowledge website for internal use. The most comprehensive report is a survey of the unit and is prepared every four years. The report is generated in the form of an Excel file and its fields it can be, if necessary, edited and sent to the national system of assessing the Polish universities, assuring that the transmitted data are complete and do not require additional adjustments.

The Repository Statistics									
WUT - Warsaw University of Technology									
13-09-2015 / 14:36									
General statistics	Publications all	Publications 2009-2014	Publications 2014	Publications (Report) (Art+Roz+Ksi)	Publications (obce języki)	Publications (jęz. polski)	PhD, M.Sc, B.Sc. (BSc+MSc+PhD)	Patents (Zgl+Pat)	
Warsaw University of Technology	WUT	WUT	WUT	WUT	WUT	WUT	WUT	WUT	WUT
WUT Central Administration	5	3	1	0+1+0	0	1	0+0+0	0+0	
Main Library of Warsaw University of Technology	107	70	19	0+0+0	0	19	0+0+0	0+0	
WUT Information Technology Centre	2	0	0	0+0+0	0	0	0+0+0	0+0	
College of Economics and Social Sciences	195	186	66	31+17+4	15	51	11+0+0	0+0	

Figure 4 Repository statistics - several types of choice e.g. all publications in repository

Conclusion

Each OMEGA-PSIR function listed in this paper will surely facilitate reporting of the University faculties. The Base of Knowledge has been operating for not long, but we can already say that the implementation of this joint University system was a right decision. The process of preparing reports has been facilitated and standardized. The function of promoting knowledge and the scientific potential of the University and its staff is also a great advantage. All metadata entered into the Base are accessible through the Internet without any limitations. Some data are available together with full-text files that have been included in the Base in compliance with the Polish copyright law.

Undoubtedly, despite the advantages of our Base, there still exist several obstacles: the principles for assessing research units often change, new bases for assessment are being created, the authors are often reluctant to make their work public. Therefore, the creators of the Base of Knowledge and people who oversee its merits have to make sure that the Base is being advanced on ongoing basis, that it is compatible with such external systems as POL-on or PBN. They also have to broaden their knowledge on the changing legislature and be up-to-date with constantly developing Open Access in Poland and all over the world.

⁶ PDF - Portable Document Format

⁷ CSV - data storage format in text files.

⁸ Bibtext - tool for formatting bibliographies.

References

KOPERWAS, Jakub Janusz, RYBIŃSKI, Henryk and Łukasz Skonieczny. *Projekt i implementacja pilotowego systemu repozytorium dla prac dyplomowych (inżynierskich, magisterskich i doktorskich) oraz publikacji pracowników Politechniki Warszawskiej*. Warsaw: the Warsaw University of Technology, Faculty of Electronics and Information Technology, dec. 2010 [cit. 2015-12-07].

STEPNIAK, Jolanta. *Ewidencja dorobku i repozytorium Uczelniane* [online]. Warsaw: Biblioteka Główna Politechniki Warszawskiej, 2012 [cit. 2015-12-07]. Available from: <http://bcpw.bg.pw.edu.pl/dlibra/docmetadata?id=3869>.

NIEZGÓDKA, Marek. *Wdrożenie i promocja otwartego dostępu do treści naukowych i edukacyjnych: Praktyki światowe a specyfika polska. Przewidywane koszty, narzędzia, zalety i wady*. [online]. Warszawa: Ministerstwo Nauki i Szkolnictwa Wyższego, 2011, 288 p. [cit. 2015-12-07]. Available from: <http://depot.ceon.pl/handle/123456789/1545?show=full>.