OMEGA-PSIR: An ecosystem for Building University CRIS Network in Poland

Henryk Rybiński¹, Jakub Koperwas², Łukasz Skonieczny³, Wacław Struk⁴

¹H.Rybinski@ii.pw.edu.pl, ²J.Koperwas@ii.pw.edu.pl, ³L.Skonieczny@ii.pw.edu.pl, ⁴w.struk@elka.pw.edu.pl

Faculty of Electronics and Information Technology, Warsaw University of Technology
ul. Nowowiejska 15/19, 00-665 Warsaw, Poland

Keywords
knowledge base, Current Research Information System, Institutional Repository, Research Profiling System, research management, open science, open access,

1. ABSTRACT

OMEGA-PSIR, developed at the Warsaw University of Technology (WUT), is an institutional Current Research Information System (CRIS). At University it can be used for various purposes, such as presenting the University research, knowledge management, internal research assessment purposes, as well as reporting to the National Authorities in Poland. Since 2015 one can observe a growing interest among the Polish universities. It is now adopted by 15 universities in Poland. In the paper we will discuss how the bottom-up initiative of the Polish universities became a successful approach for building a low-cost network of the institutional CRIS systems, cooperating with the national CRIS system.

2. BACKGROUND

One of the outcomes of a dedicated project SYNAT (launched in Poland in 2010) was the system OMEGA-PSIR, designed and implemented by a team of Warsaw University of Technology (WUT). The project aimed at providing means for building an institutional knowledge base system (Rybinski et al., 2015). Its main idea was to integrate various university needs related to research.

Soon after installing the system at WUT, the system aroused interest among the Polish universities. Since 2015 the system has been adopted by 15 Polish universities, whereas few others advance the work towards installing the system. A special User Group has been launched for coordinating the system development.

In the paper we will discuss how a bottom-up initiative of the Polish universities became a successful approach for building a network of the institutional CRIS systems, cooperating with the national CRIS system. To this end we will discuss functionalities of the system, in particular, we will show to which extent combining CRIS, IR and RPS makes OMEGA-PSIR so attractive to the universities. We also discuss how the low costs of the system deployments at the universities can be preserved. In addition, we will present how the functionality of the system can essentially improve the data quality, and at the same time reduce the maintenance costs.

An important factor concerning interoperability with other systems will be discussed. We will discuss various aspects of interoperability with global systems, such as:

- OpenAire
- Sherpa-Romeo
- CrossRef
- ORCID,
- Commercial databases, like WoS and Scopus

We briefly present a kind of an ecosystem that has emerged around OMEGA-PSIR in the last few years. The main components of the ecosystem will be discussed; in particular we will discuss the following issues:

1. the software development and maintenance,
2. support options that are provided to the universities in Poland during the system deployment and exploitation;
3. exchange of experience between the partner universities;

Last but not least, it will be demonstrated how our approach can be applied/reused in other countries.

3. MOTIVATION, GOALS AND KEY ASSUMPTIONS

Observing contemporary institutional information systems, one can see two main different solutions: (1) on one side, an approach represented by systems like Fedora-Commons or D-space (see e.g. Berman (2008)), which focus mainly on the repository functions, such as storage and indexing of research-related documents, including also aspects of long term durability; (2) on the other side, an approach represented by systems, which are very well suited to the research-centered model of Current Research Information Systems (CRIS).

For building university research knowledge bases the institutional repositories were dominating for a long period. Their applications were especially driven by the idea of open access. The systems within this approach provided rather simple end-user functionality, mainly limited to browsing and querying the repositories, with documents organized in collections. They were bibliography oriented, usually document-centric ones, and did not provide-end users with sophisticated presentation capabilities. According to Russell and Day (2010) the main motivations for building institutional repositories were inter alia:

- to provide a showcase for scholarly output from an institution (e.g. facilitating increased visibility; generating indicators of academic quality);
- to improve dissemination of research output;
- to ensure the long-term preservation of resources; and
- to break down access barriers to content (i.e. reforming the scholarly communication system).

Since 2007 one could observe in the literature a criticism of the institutional repositories (Aschenbrenner et al. (2008), Davis and Connolly (2007), Salo (2008)). To this end, one can see an increasing interest of universities in using such institutional systems that describe and present various aspects of university research, not being just limited to bibliographic outcome. For this approach CRIS type system becomes a valid choice.

Nabavi et al. (2016) define CRIS as the one, which main goal is to provide a tool for supporting research decision makers (institutional or national, respectively) to carry out the following tasks:

1. defining research priorities;
2. supporting and automating the process of assigning budgets for research groups, teams, faculties (at the institutional level);
3. evaluating and ranking research institutes based on defined criteria.

This approach to institutional CRIS requires that the research be assessed, so that, as noted by Bittner and Müller (2011), CRIS systems are very formal and “bureaucratic”.

Since 2008 one can observe systems that somehow recall CRIS approach, but are focused on researchers, and their profiles. Examples are such global systems like ArnetMiner, Microsoft Academia, or networking systems, like ResearchGate. In the domain of institutional systems a good example is VIVO. All of them are researcher-centric, rather than research-centric. Nabavi et al (2016) distinguishes them from CRIS, and calls them Research Profiling Systems (RPS). What can be easily noticed, the RPS systems are especially appreciated by researchers. This explains why the combination of RPS with CRIS within the Elsevier’s system PURE turned out to be so attractive to many universities.

Bearing in mind all these tendencies, we have concluded that our solution should take into account the best features from all those three types of systems, i.e. IR, CRIS and RPS. We realized that integrating various functionalities within one system in not an easy task, as it means that we have to integrate various, sometimes conflicting, needs of different user groups. It should be beneficial for, and motivating to, quite different groups, including (but not limited to) researchers, students, university strategic management, administration, librarians. What is very important, it should
guarantee low maintenance cost, at the same time should be as much user friendly as possible. These were the key assumptions for developing the OMEGA-PSIR software.

4. OMEGA-PSIR ECOSYSTEM

Soon after installing OMEGA-PSIR at WUT, the system aroused interest among Polish universities. As the system has been developed with the public money, the license of the systems for Polish institutions was free. Since the beginning we were aware that OMEGA-PSIR, being a low-cost solution, is a real chance for the academic community in Poland with a modern approach to institutional information systems. Two years after installing OMEGA-PSIR at WUT, the first two institutions were licensed and started implementing the system. For the WUT team it was an experiment showing various kinds of problems and difficulties with installing the system in various environments. It was also a risk that with limited human resources our problems with supporting a growing number of institutions might arise.

On the other hand it was also clear for us that with a larger number of institutions using OMEGA PSIR, there are real chances for finding funds for the system maintenance and development, making the OMEGA-PSIR project sustainable.

For the WUT team the growing number of the installations of the system in Poland meant that:

1. the quality of the system documentation and instructions should be essentially improved;
2. there should be a way to make possible the system adaptation to specific needs without loosing compatibility of given installations with the standard version of the system;
3. there should be forums for the participating universities for exchanging experience and sharing ideas in solving various technical and organizational problems;
4. there should be procedures in updating common databases (e.g. journal database, conferences, funding organizations, etc.).

We have quickly realized that forums for exchanging experience and sharing ideas were of the highest priority. Therefore, after first two successful installations we have decided to organize an “interest group” (a form of club) for institutional users. The users group meets few times a year. During the meetings we discuss new functionalities of the software and plans for the development.

Independently of the working meetings once a year we organize Annual OMEGA-PSIR Seminar, which is open to a larger audience in Poland. It is devoted to the Open Science issues, in particular legal issues of Open Access, Open Data, as well as the developments of the national CRIS - POL-on. Also, the OMEGA-PSIR users present their problems and solutions, often in the context of integrating OMEGA-PSIR with other institutional systems, like e.g. e-learning, ERP or HR.
Yet another form of platform for information exchange is a dedicated forum site, open to all. At the OMEGA-PSIR forum various questions and problems can be freely discussed, so it is used mainly for:

1. self-assistance
2. sharing experience
3. populating new ideas and solutions.

Since the last few years Poland is on the way of implementing reforms in the Higher Education system. This fact causes that the WUT team is involved in implementing various changes in the system, resulting from changes at universities, including evaluation procedures of the universities. For this reason, we have decided to closely cooperate with the body which in Poland is responsible for the national-wide POL-on.

In addition, since 2018 we have started cooperating with a private company which is involved in organizing training and webinars, where various technical and organizational issues are presented.

As a result of our activities, the number of national licenses and installations is growing (see Fig. 1). We have distributed more than 24 licenses to the institutions, of which 18 are members of the OMEGA-PSIR Interest Group. Of those 18 universities, there are 15 institutions that have already the system installed. In 2019 we expect granting license to four other universities.

What is important, the activity of the group member institutions increases year by year. Recently, responsibility of the members for journals database has been determined in such a way that particular universities declare, which journals (or even journal domains) are monitored by them, so that the work on the journal database is shared. For this reason we plan implementing a central database for journals with mechanisms for automatic synchronization of this database with local instances. The same mechanism will be applied for periodical conferences, publishers and database of funding organizations.

5. FUNCTIONALITIES THAT MAKE OMEGA-PSIR ATTRACTIVE

The development of the system was driven by user needs and has been progressing iteratively. The first phase took the period of three years. The process of moving the system to the University level at WUT was simple, as the team had experience with organizational and training issues at the faculty level. Such approach caused that the system was confronted with its users from the very beginning, and the developers were confronted with real user needs, so that when the system was finally ready to be shared with other universities in the form of a complete Ω-ΨR package, it was already mature, well-tested and well-documented.

Already in 2013 the functionality of the system went beyond the typical functionality of Institutional Repository, turning towards the functionality of CRIS, additionally the functionalities of RPS have been already implemented.

Since the beginning the system has been subject of interest at other universities in Poland. Once the system became subject of distributing to other universities, the development of its functionality has been driven not only by the needs of WUT, but also the needs of other universities, also the smaller ones. For this reason, one of the important tasks was the reduction of the maintenance costs. Due to applied intelligent tools, such as acquisition tools, reporting functionalities (Koperwas et al, 2017), the maintenance efforts of Knowledge Base are essentially reduced, as compared to the typical solutions. In particular, with increasing the interoperability with other systems, such as WoS, Scopus, or CrossRef the acquisition costs have been essentially reduced. Fig. 2 illustrates the development phases. It also shows how the interoperability with other systems was changing.

5.1. Acceptance of end-user functionalities

The motivations for incorporating within OMEGA-PSIR various functionalities of IR, CRIS and RPS system have been discussed by Rybinski et al (2017). Here, we discuss how the main functionalities are assessed by various users groups.

---

1 http://forum.omegapsir.ii.pw.edu.pl/top?order=activity
In OMEGA-PSIR within the end-user functionalities one can distinguish the ones that are typical for IR systems (mainly accessing the documents), as well as the functions specific for CRIS and RPS. So, the typical functions available for end-users are:

1. Accessing the documents after browsing or searching in bibliographic data (IR);
2. Accessing the profiles of researchers (searching by names) and units (CRIS);
3. Viewing the cooperation networks (RPS);
4. Searching for experts in a domain (CRIS, RPS).

At the end of 2016, we have decided to verify how the system is used by end-users. We have performed two types of research:

1. a survey research among the active and potential users;
2. in addition, taking advantage of using Google Analytics since May 2013 we have analyzed the Google Analytics data.

**Survey research**

In total, 424 opinions were collected. When asked which group will find the data provided by the OMEGA-PSIR system most useful, the respondents indicated that it is mainly employees of Polish and foreign scientific units and universities.

A high level of relevance and usefulness of such data to Polish and foreign entrepreneurs was also indicated. It is worth noting that each group of surveyed beneficiaries declared that the resources would be very useful for them. The results of the survey can be summarized as follows:

1. The largest group of the system recipients are academic staff of Polish and foreign universities and scientific institutions. Open access to full-text scientific publications brings many benefits to both scientific entities and to the whole scientific community. It is important to mention the possibility of multi-criteria search and automatic generation of information about the achievements, areas of interest of individual scientists, units, and teams. All these criteria have a positive effect on intensifying the exchange of scientific ideas and establishing new cooperation between institutions.
2. Another group of recipients are entrepreneurs who, with help of the OMEGA-PSIR system, can cope with the existing barriers impeding the establishment and effective cooperation of science
and business. The biggest problem in this area is insufficient knowledge of the entrepreneurial sector related to the areas of expertise of Polish universities, which is also related to the lack of experience and good practices in effective presentation of the university academic achievements and research.

3. PhD students, undergraduate and graduate students, but also candidates to study, are also an important group. The system is particularly helpful in recognizing the character of the unit and the quality of its services.

4. The system is also helpful for media representatives who can have access to information about experts in various fields, which is necessary to get professional comments on current events.

5. And last but not least important group is the university management staff. Having the OMEGA-PSIR system with the CRIS elements, they have an important tool for assessing the work of individual scientists and teams, and also a tool for developing a scientific policy of the unit.

In addition to identifying the potential stakeholders, the survey also enabled identification of the key needs of potential recipient groups and identification of data types and analytical services in the systems with open access to scientific information. The most important issue for the respondents using these types of systems is free access to the data. They also pointed out that in this context, it is important for them to have open access to as many scientific publications as possible as well as to raw research data. In addition, we have asked for the opinions concerning especially those non-typical ones:

1. graphical visualization of cooperating teams;
2. graphical visualization of profiles of individual researchers or university units;
3. search for experts.

It turns out that researchers evaluate equally high the functionalities of visualizing cooperation between units, researchers, as well as, visualization of the research areas of units and individuals, which is provided in the form of terms clouds. An overall view of the users preferences is summarized in Table 1.

**Google Analytics statistics**

The GA data for the WUT Knowledge Base are available since 2013. They show a growing interest for Data in the WUT Knowledge Base. In 2016, some 1,000 sessions per month (10% of total traffic) were started by users out of Poland, in 2018 the number of sessions started out of Poland reaches 3,000. The geographic distribution shows that the traffic from USA and EU countries is significant (now more than 15%).

<table>
<thead>
<tr>
<th>Groups</th>
<th>Interactive graph of cooperation between faculties</th>
<th>Interactive graph of cooperation between researchers</th>
<th>Visualization of research areas of the university and its units</th>
<th>Visualization of research areas of researchers</th>
<th>Multicriterial search for experts and teams in a given domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Researchers</td>
<td>7,32</td>
<td>7,43</td>
<td>7,41</td>
<td>7,57</td>
<td>8,57</td>
</tr>
<tr>
<td>Business</td>
<td>5,15</td>
<td>4,8</td>
<td>6,05</td>
<td>6,2</td>
<td>8,15</td>
</tr>
<tr>
<td>Public admin.</td>
<td>4,00</td>
<td>3,73</td>
<td>3,95</td>
<td>4,00</td>
<td>5,05</td>
</tr>
<tr>
<td>PhD students</td>
<td>6,84</td>
<td>6,89</td>
<td>6,84</td>
<td>7,37</td>
<td>8,21</td>
</tr>
<tr>
<td>Students</td>
<td>6,93</td>
<td>6,67</td>
<td>6,82</td>
<td>6,82</td>
<td>8,09</td>
</tr>
</tbody>
</table>

With GA we can assess the statistics for the users’ behavior. In particular, we have analyzed the types of the starting screens, and the distribution of the functions used by the users. It turns out that quite a meaningful number of users start their session with the screens that are not available within typical repositories, such as viewing researcher profile (17 percent), the screen for searching...
experts by domain (10 %), or the screens with profiles of the university units, such as faculties, institutes, etc. (7 %).

As an example, Figure 3 shows behaviour of users from USA in 2013-2016. One can see that in addition to the standard repository functions, users often refer to the function of looking for experts.

Unfortunately, with the GA data we cannot decide, which groups of users generate requests for these functionalities, but the analysis confirms the results of the survey research concerning high interest in those functions that are not typical for institutional repositories, and result from the researcher-based approach.

Some functionalities were not covered by the research for two reasons:
1. they were not ready at that time;
2. they were not so explicitly exposed to the users.

For (1) it refers mainly to the following functions:
1. Options for researchers to download their bibliographic data to their profiles from Scopus;
2. Possibilities for researchers to automatically generate their CVs;
3. Possibilities available to all users to visualize cooperation of the university units with the other universities (see Fig. 4). Although this is quite new feature, we expect it will also be an important “must have” standard in the institutional knowledge base systems, not only for end-users, but also for the management.

Examples of functionalities that are not so explicitly exposed refer mainly to interoperability with other systems. For example, interoperability with CrossRef brings a lot of savings in data acquisition. The same refers to interoperability with Scopus and WoS. Yet another value of the system is achieved with connecting our journal database with Sherpa Romeo. We have implemented assistance to the researcher, who in time of depositing documents to the repository can see the openness policy of the journal.

Recently, a very important interoperability has been added, namely in cooperation with EuroCRIS we have implemented communication of OMEGA-PSIR with OpenAIRE by means of the CRIS exchange format CERIF.

5.2. Knowledge Management Functionalities

Reports
The system is equipped with a set of tools that make possible predefining various reports that are requested by the University authorities. The reporting tools are used mainly to build analytical views of data for the managerial staff of the university. They turned out to be especially important with the advent of reforms of Higher Education, currently taking place in Poland. The basic reports cover
1. Annual reporting of the University units
2. Bibliometric analyses for various disciplines and teams
3. Personal reports of annual achievements, CVs

The reports can be defined \textit{a priori} in a form of “generators”, so that the final used can define specific parameters. An example of a bibliometric report is provided in Fig. 5.

With the reporting tools however the main problem is that preparing any simple form requires an advanced expertise in Java scripting. To this end we have considered designing of a special tool that would be applicable for the administrative staff. As a result, the essential development of functionalities addressing knowledge management issues took place in the last few months. Namely, we have developed a module for generating pivot tables, which is addressed directly to the administrative staff and does not requires deep knowledge about java scripting or even data structure nuances.

\textbf{Pivot tables}

The main idea is that for each object type the most important attributes are defined. Some attribute are equipped with special parameters, for example for numeric attributes one can easily define ranges, for hierarchical attributes (like affiliation) one can define which levels have to be used in the resulting table, etc. Given a set of attributes, the end user can easily define columns, rows, and select the aggregating functions for the cells, and then having the table definition, run the pivot to get the table. As an example, selection of Unit for rows, Start year for columns, and number of project as an aggregating function (Fig. 6a), gives a simple table of numbers of projects per year (Fig. 6b).
The tests with our librarians and the staff responsible for research projects have been performed, showing that no special IT support is needed for running advanced statistics with the pivot tables. The users can modify their pivots by themselves, store the modifications for future use, and, if needed, share them within their units with other staff.

Among the functions for generating the pivot tables, one important feature refers to matrix-type tables, aiming at showing advanced measures of cooperation between units, and teams (including cooperation between disciplines).

Visualization of cooperation
A variety of visualization tools have been implemented in order to facilitate data analysis. Some of them, related to automatic building of the research interest cloud are presented in (Koperwas et al 2017; Rybinski et al, 2017). Recently, new visualization tools have been developed:
1. One of the examples is shown on Fig. 3, where cooperation of the university staff with other universities is shown geographically. In this visualization we use affiliation of external co-authors.

2. Another example is to visualize a cooperation of authors from various disciplines (see Figure 7).

The main idea of visualizing the cooperation of teams internally and externally is very important for university management. In the case of (1) it simplifies discovering the most active teams from the point of view of internationalization of research. For (2), the idea is to discover interdisciplinary research groups. For this case the disciplines of researchers are visualized by bottom color bars. Marking a cooperation edge between two researchers one can see their cooperation in terms of common publications, projects, patents, etc. This kind of visualization can be very useful to visualize the matrix of cooperation between disciplines, obtained by the pivot module.

![Figure 7 New visualization tools in OMEGA-PSIR](image)

6. CONCLUSIONS, FURTHER PLANS

As reported by Ribeiro et al (2016), building an in-house CRIS system is quite a popular decision at universities, and often it is seen as a remedy for expensive solutions based on a commercial offer. Our experience shows that with current technologies, often available in the form of open source, one may develop an advanced system with many interesting functionalities.

What is of crucial importance, the system should obtain a high acceptance of various groups of end-users (in our case researchers, but also managers). According to our observations confirmed by survey research, the incorporation of the RPS features into the OMEGA-PSIR system makes our solution more attractive for researchers. On the other hand, the reporting and statistical tools make the system attractive to the managers. They play a crucial role in strategic planning of the research activities, staff development, etc. The survey and analysis of GA data prove that our idea to combine various functionalities of IR CRIS and RPS into one system was a good decision.

Another good decision was to implement interoperability of OMEGA-PSIR with other systems. With the possibilities to refer to CrossRef and global databases (Scopus and WoS) it made our system cheaper in terms of the database maintenance costs, whereas interoperability with Sherpa Romeo made our system more supportive to Open Access initiative.

One thing is to create a good institutional system, on the other hand though, a real challenge is to maintain system sustainability after the end of the project. To this end, after the completion of the SYNAT project (the one that financed the development of our system in 2010-2013), we immediately started to create a kind of ecosystem for the efficient maintenance and further development of OMEGA-PSIR. In particular, we have initiated a grassroots initiative, building a user group around the
Our approach turned out to be successful, and resulted in building a network of the cooperating universities, sharing the same software, and the same methodologies for building institutional information systems, exchanging data with the national CRIS system.

Our short term plans refer to the changing communication with the national CRIS system (POL-on), mainly towards automating the exchange of data requested by the national system for the university evaluation.

Our longer term plans cover:

1. steps towards providing Linked Open Data interface;
2. interoperability with science oriented knowledge resources, such as ontologies and thesauri (Mesh, Agrovoc, etc.)
3. interoperability with Mendeley, ORCID and other global systems.

Last but not least we look for interested universities abroad, considering cooperation in developing the system.

7. REFERENCES


8. AUTHORS’ BIOGRAPHIES

Prof. Henryk Rybinski

Henryk Rybinski is full professor at Institute of Computer Science. In 2008-2016 he was Director of the Institute. Since 2010 he is responsible for the development of Knowledge Base software for WUT. His current research is focused on using text mining techniques for knowledge discovery from text data. He has published more than 130 scientific publications in the area of information systems. For some 35 years Prof. Rybinski has been conducting projects for building information systems for many international bodies (i.a. FAO, UNESCO, UNEP, IFRC, IUCN).

Dr Łukasz Skonieczny

Łukasz Skonieczny, Ph.D, assistant professor at Institute of Computer Sciences, is one of the main developers of the Ω-Ψ system. His research interest is in database systems, data-, text- and web-mining, graph theory and web development. He has in his record 10 scientific papers and, 3 edited books. He participated in a bunch of research projects, and cooperated with many institutions, inter alia France Telecom, Samsung, UNEP, FAO, IUCN.

Dr Jakub Koperwas

Jakub Koperwas, PhD, is an assistant professor at Institute of Computer Sciences, Warsaw University of Technology and lead consultant and partner in IT consulting company - Sages. His research interests are data mining of semi-structured data, especially for bioinformatics and distributed data mining. He has published 10 scientific publications in the area of information systems. He provides software development lectures for students of Warsaw University of Technology.

Wacław Struk, M.Sc

Wacław Struk, M.Sc., works at the Faculty of Electronics and Information Technology since 2010. He is responsible for the FEIT information systems infrastructure. He participated in the development of the Ω-Ψ. His professional interest is in information systems, especially text oriented databases. He participated in various research projects, and cooperated with many international institutions, inter alia Infoterm, TERMNET, UNIDO, FAO, IFAD, WFP.