Marek Niezgódka

ICM (Interdisciplinary Centre for Mathematical
and Computational Modelling), University of Warsaw
Pawińskiego 5a, 02-106 Warsaw
marekn@icm.edu.pl

PROJECT REPORT: OPEN SCIENCE - AN OFFER FOR OPEN SOCIETY

Extended Abstract

The new age, age of information and communication, is the time for new models and paradigms in the world of research and education at any level. Scholarly publishing at all stages has been revolutionized by ICT breakthrough, hence contributing to a democratization (and symmetrization) of scientific creation.

New common trends result in opening: Open Access, Open Software have become landmarks of this world. What about open data, open technology or open knowledge: is this just a list of non-realistic ideas? How does this interplay with the generation change and with millenials starting to play more and more articulated role in society?

In the world of research the time has come for its overall reorientation towards open models. By the way, this is by no means any new discovery but rather a return to ancient roots. Public funding of research imposes clear obligations to its beneficiaries: they should provide public access to results achieved, any related documentation and publications. Public means here open and borderless.

With such openness, science should reach high transparency, this in particular applied to scientific evaluation, public debate and reduction of institutional/national borders.

During past decade, a number of initiatives arose contributing to fostering Open Science concept in the Polish academic world and its structures. Among those, dated back to mid-nineties, Virtual Library of Science (BWN) program (http://vls.icm.edu.pl) belongs to the most spectacular developments. In its new shape, the program has started to implement the open concepts:

- results of publicly funded research are to be openly accessible in the form of electronic publications,

- a legal framework should get accordingly adapted to new situation, in particular redefining the very concept of copyright,
- scientific data of publicly funded research should become open,
- the same for scientific software.

The underlying IPR concepts and their treatment have got revolutionized with the declaration of Creative Commons (and their follow-ups) legal solutions. Those new approaches are now becoming de facto standards, widely adopted by funding agencies and international bodies.

Initially, the BWN program was meant to set up a sort of virtual academic campus open to all Polish higher education and research institutions. As an underlying principle, the institutions were required to join virtual organizations granted unlimited access to specific leading-edge bibliographic databases, full-text content within selected disciplinary archives (starting from biomedical literature, within OVID system) and a range of scientific data resources (with Beilstein and Gmelin systems in the beginning).

Within following few years, the access gateway to the BWN system has reached an integrated form and a range of full-content collections were added to the offer. With complete ranges of ELSEVIER, Springer and Kluwer collections of journals, the BWN became full-dimensional system with well over 4K journals inside.

Among the underlying principles:
- full access symmetry, i.e., full unlimited access (from all IP domains) to entire collections for any of the member institutions,
- perpetual archiving on site,
- perpetual access rights (once gained, non-expiring),
- content unbundled from access software, hence freedom to use ICM's own access system and software, were introduced and got contractually implemented by ICM.

Underway, we have gained quite an experience with:
- successful implementation of the complete transition from print to digital content on country level,
- changing the attitude from journal-based to collection-based,
- monitoring a spectacular evolution of usage patterns (a complete documentation remains preserved, with resolution reaching single accesses),
- supporting emotional and intellectual efforts of libraries to re-position their roles.

To illustrate the scale of ICT operations performed, weekly load for just a single (even if the largest one) collection happened to approach nearly 100K single article downloads, with the annual mean value in the range of 70K+.

The list of collections has in-between been substantially enlarged, with nearly all research institutions joining the BWN campus. As for higher education institutions, certain stratification could be observed: all high-rank universities and academic schools in, whereas only few non-public colleges showing interest.

There is a spectacular record showing cost-efficiency of the system and both direct and, by far larger, indirect gain of individual institutions joining the BWN program. But what is even more fundamental, the multifold digital divide has been significantly reduced, geographically and institutionally, but also offering students new dimension of access to reference knowledge.

To facilitate an integrated concept of the BWN system, with single access gateway for the users to the entire resources, a development of an unifying IT platform has been initiated recently at ICM. The YADDA (Yet Another Digital Document Archive) platform (http://yadda.icm.edu.pl) is now fully functional, enabling access, authoring, indexing, browsing and multi-level text analysis. A derivative of the same platform has reached the status of the main technology in EU FP6 and FP7 initiatives DRIVER and, accordingly, DRIVER2, that aim at establishing a European Infrastructure for Open Scientific Repositories within next few years.

The latter implementation became possible due to preceding work on expanding YADDA into a unified platform for distributed multimodal system of federated knowledge archives and repositories, with heterogeneous access modes ranging from limited up to fully open.

The system is capable of handling bibliographic files, indexes and comprehensive metadata, along with full-text content, both structured (publishers' collections) and unstructured (single journals and, down to, single publications). Further, multimedia files and archives/repositories of scientific data can also not only be incorporated but also functionally connected to respective publications. A number of additional functional features are under running development, with an objective to foster novel future scholarly publishing models. In particular, a combination of scalable concepts attributed to Grid processing technologies with those considered underlying the convergence of communication media will successively get incorporated into YADDA platform.

Based on the resulting ICT infrastructures, scholarly communication and publishing will get revolutionized, far more than just separate events, they will become implementations of dynamical systems, with collaborative features and open structure of interactions. To what extent will this reshape the present organizational schemes of academic world, remains open. For sure, institutional barriers preventing free (open) creativity may get broken or, at least, lowered.

Aspects of the whole development that remain unclear are the questions on financing and business models as well as those on ultimate handling of IPR

issues for electronic media. Nevertheless, the future models will be substantially built atop of open paradigms.

The number of funding organizations and governmental bodies that already have introduced Open Mandates for scientific publications grows rapidly. This contributes to removing the contradiction between quality-assured peer-reviewed publishing and providing open access to the whole content on Internet. As for software, the position of free open licensing models has become widely accepted with high success impact, also on commercial usage. The component that remains somewhat behind is that of scientific data and their open accessibility. So, on one hand no open breakthrough is possible without new technological concepts, still regulatory support may prove critical as entering into more conservative dimensions.

ICM's development concept, so far based on public research funding, is aimed to enter also into other areas of relevance to the process of shaping open society.