Open Access to Knowledge@ Novartis – new ways of knowledge sharing in the pharmaceutical industry

Alternative Title:
Open Access to Knowledge@ Novartis – the Rocky Way to Open Access

Author:

Dr. Christian Gumpenberger, MSc
Knowledge Center Manager
Novartis Knowledge Center Vienna
Novartis Institutes for BioMedical Research
Vienna, Austria

Introduction
The Open Access movement has gained momentum in recent years without question. This is true for academia, but how is the situation in industry?
Mostly perceived as an unwelcome free-rider the time has come for industry to also embrace the new trends in scientific communication and publishing.
This presentation is about the endeavours of a big pharmaceutical company to become more actively engaged in the Open Access movement.

Status quo in academia
The traditional publishing model has seen the arrival of Open Access publishing, and the number of newly launched or retrospectively converted OA titles is steadily growing. However, OA publishing is under a permanent debate and sustainability of the different business models is still unclear.
In contrary to the “Gold Road”, the term coined for OA publishing, another alternative has proven to be more successful to achieve Open Access on the short term: Open Access Self-Archiving (or the “Green Road). Already practiced by physicists, mathematicians and computer scientists since the early nineties, scientists of other disciplines discovered this way of sharing their intellectual output only recently.
Increasingly we now see traditional publishers allowing authors to self-archive copies of their articles in their Institutional Repositories. An estimated 90% of all scientific articles could be made Open Access this way globally.
In response, many universities and academic institutions worldwide have set up their own Institutional Repository permitting them to freely store copies of their own publications in one central place which are accessible for both their own communities but also to external visitors.

Status quo in industry
In contrast to the development in academia the Open Access movement has so far not really impacted the industry world, at least not visibly for the general public.
However, the whole world of scientific publishing and communication is changing, and industry can definitely not deny this global trend.
But how could that be achieved in the highly competitive corporate environment?

Would it be compatible with a corporate business model? And if yes, would it be a win-win situation for the company as well as for the general public?

Background

Transformational Change at Novartis

For the Swiss pharma company Novartis corporate citizenship is an integral part of how to operate and a key to success. The Singapore based Novartis Institute for Tropical Diseases (NITD) is one of the best examples to showcase this philosophy, as it is the institute's mission to improve access of medicines to the developing world. Thus the concept of Open Access to the company's scientific output fits perfectly well with the idea of sharing.

In 2005 the Novartis Institutes for BioMedical Research (NIBR) therefore started to support the Open Access movement through memberships to the big Open Access publishers BioMed Central and Public Library of Science (PLoS). Since then authors have the option to publish OA in order to increase the impact of their work and consecutively of the whole company. Nevertheless author uptake has not been overwhelming within the company so far similar to observations in academia (less than 1% of Novartis research has been published OA). This is because authors on the one hand still cling to traditional journal titles and their ISI impact factor. And on the other hand OA publishing can still not be praised as the ideal alternative due to the uncertain viability of the different business models and the extra cost on top of ongoing traditional journal subscriptions.

Project "Open Access to Knowledge @ Novartis (OAK)

Thus the next logical consequence for Novartis was to walk the "Green Road" as well.

The Novartis Knowledge Center (NKC - a global group of Information Specialists taking care of the external information acquisition and information delivery within Novartis) has been closely monitoring the new trends in scientific communication and publishing within the last years. It came forward with the idea of implementing a Novartis Institutional Repository in summer 2005 and should also take the future responsibility to run this pioneering project in industry.

Following benefits for the individual researcher AND the whole company were highlighted:

- Efficient, instantaneous access to scientific articles for Novartis scientists and collaborators worldwide
- Greater impact and visibility for Novartis-funded research
- Provide a showcase for the research carried out at Novartis to the global scientific community
- Marketing tool for the company, informing and attracting potential employees, collaborators and investors
- Enabling NIBR management to have a complete record of all the research published from the company; easy visibility for departments and business units to their research publications
- Enabling NIBR management to monitor and assess this research on an empirical basis

- Enabling text and data mining deposited in the repository
- Enabling Novartis scientists to compile reference lists for publications, personal publication records and other such documents simply and cost-effectively in terms of time
- Submission mandate would increase number of documents in repository

The Rocky Way to finally reach the Green Road

Senior Management recognized the benefits for the individual researcher and the company as a whole and decided to go for a Novartis Institutional Repository in principal.

However, in a corporate environment with strict information security policies an Open Access repository project could not be launched without overcoming many obstacles. So first of all NKC had to prove there was a real business case.

In order to prepare a solid project proposal for Senior Management a retrospective analysis of all Novartis publications from 2004 was done as a first step (one complete year serving as a role model).

The result showed that Novartis publishes close to 1000 original scientific articles annually across a wide range of medical and scientific journals. A majority of 70 – 80% could be self-archived and shared with the global scientific community easily, which was a reasonable number to proceed.

Until the end of 2005 the available literature on Institutional Repositories was studied. It was evident to go for an open source software solution (advantages of open source vs. proprietary software are evident and generally accepted), however, the question whether to build the repository ourselves or to outsource could not be answered immediately and needed further investigation.

A project proposal including major requirements was put together for Senior Management in Q1 2006. It covered basic information to implement a Novartis Institutional Repository on the DSpace or the EPrints software, looking at both either outsourcing or self-development. In the following 6 months NKC set up demo repositories for both software options and further investigated their functionalities.

In parallel the requirements specification was put together. It was clear that the standard version of both software packages would not meet all the needs of Novartis and further customization was necessary.

At this time of evaluation NKC became aware of a project of NIBR Novartis Research Publishing (NRP), a group within NIBR Communications. They define the process to manage and approve publication manuscripts from NIBR scientists prior to their submission to publishers. NRP also wants to keep track of all manuscripts that come from NIBR. The current process includes only one mandatory step and is so far paper based. The Publication Tracking System (PTS) should replace the paper based system, however, this system was planned to be built on propriety software for internal use only and did not address copyright compliant long-term archiving.

After a thorough analysis it was finally clear the PTS project could be linked to or even integrated with the Novartis repository.

In Q3 2006 NKC came up with comments and cost estimates for the different options and favoured an outsourced solution of a highly customized EPrints repository with a full integration of the so far planned processes for PTS. It should comprise of 2 sections:

1. OAK Internal Archive: containing all publications from “Draft” state to “Published” state; audience restricted to Novartis employees
2. OAK Open Access Internet Archive: containing all final drafts of already externally published peer-reviewed articles cleared for Open Access; to be accessed on the internet by anyone interested

The different sections should satisfy corporate Information Security requirements as well as the innovative Open Access aspirations.
Outsourcing was proposed because

- internal FTEs (and the associated technical knowledge) were not available to develop an institutional repository built on standard open source software and maintain it in the long term (if FTEs were only hired for development, but not for maintenance, the repository would not be sustainable)
- if outsourced the development and maintenance costs are very transparent and impact on the budget can be planned easily

EPrints as software package and EPrints Services as project partners were chosen because of

- their pioneering status in the repository history
- a commercially attractive offer
- the best fit with the elaborate project requirements specification
    - 2 separate archives (internal vs. external archive)
    - OAK to be built on the OAI-PMH (Open Archives Initiatives – Protocol for Metadata Harvesting) standard and is therefore interoperable with other open source repositories.
    - automatic authentication of internal users based on SSL (Secure Socket Layer) using PKI (Public Key Infrastructure) certificates. The PKI certificate identifies the user by means of a distinguished name (DN) which includes a serial number from a global Novartis directory database. The repository then extracts the serial numbers from the DN and maps it to the corresponding user accounts. Account data are regularly transferred to and stored in the repository.
    - external article linking (DOI, Entrez DB, SFX [for internal archive only])
    - metadata autocompletion (for internal archive only)
        - Autocompletion of journal title and ISSN → authority list derived from NLM Serfile
            - Once journal info is autocompleted, the publisher policy is automatically displayed based on the integrated RoMEO database
        - Autocompletion of internal author names → authority list derived from a combination of internal user account data and external author names already available in the repository
    - automatic Metadata import (for internal archive only)
        - Creation of pre-populated records by entering either PubMedID or DOI
    - implementation of Medical Subject Headings (instead of LoC)
    - Novartis organizational structure search- and browseable by Function, OrgUnit and SubUnit.
    - OAI interface to MetaLib (for internal archive only)
    - bulk import of records from internal proprietary archives
    - integration of an internal Review & Approval Process (for internal archive only)

- based on the design of the planned, but not launched PTS (Publication Tracking System)
    - 9 workflow states supported
    - different visibility and accessibility for items of each workflow state
    - internal reviewers and approvers check for scientific quality
    - information specialist ensures metadata quality, compliance with copyright and publisher policies and identifies items to be extracted into the external archive

Senior Management finally agreed on the outsourced EPrints repository option with fully integrated PTS processes. Thus project OAK was launched officially in Q4 2006. In parallel to the time consuming fine tuning of the requirements specification the project required several risk assessments due to global Novartis policies.

An initial High Level Risk Assessment (HLRA) addressed

- project purpose
- validation requirements (Does the system require any regulations?)
- Sarbanes Oxley requirements (Does the system support processes for financial management?)
- Information Security Requirements (What are the system's risks regarding confidentiality, integrity and availability?)
- any other regulatory or legal requirements

For the OAK Project the HLRA only revealed Information Security risks to be further investigated in a more detailed risk assessment. Moreover a vendor solution validation was carried out based on a document study (information provided by EPrints Services) and a telephone interview to touch on the vendor's IT security (logical security, logging & monitoring, virus protection, vulnerability management, patch management, change management, data backup / recovery, data security, network security, data center controls). The results were satisfying and written up in a report.

Project OAK status quo and outlook

After HLRA & vendor solution validation were done, the requirements were fine tuned, and all the necessary corporate project documentation has been set up, the development of OAK finally started in late August 2007.

Development of OAK v1 has been completed at the end of 2007 and Performance Qualification Testing in a core team of testers is currently work in progress.
User Acceptance Testing as well as several deployment preparations are on the agenda for the near future. OAK should come together with a sophisticated education programme not only about the repository application but also about Open Access in general. The challenge would be to get the buy-in within the company:

- Top-down support from Senior Management is required (endorse a mandate for self-archiving; integrate self-archiving in an overall publishing strategy)
- Bottom-up support from the repository users: compared to academia the pharmaceutical industry has a more competitive work environment. Thus buy

in for new systems is generally difficult in big companies due to the abundance of systems that permanently emerge & change.
Time spent on daily tasks is recorded in a time management application at Novartis ("Time is money!").Therefore a system supporting specific processes must be very intuitive and time saving to be really accepted by its users (i.e. using the system should result in spending less time on the process meant to be supported). Ideally already existent systems should be rather replaced by or linked to or integrated into a new system, rather than adding just a new one to the portfolio and increase the workload.
Apart from the time issue a complete change in the people's mindset is necessary. Implementing an OA repository built on Open Source software is a brand new way to deal with information in a corporate environment.

Unfortunately project OAK is currently delayed due to the internal complexity of project management and ongoing internal organizational changes. Thus the exact launch date of OAK is unknown right now.
OAK – despite delayed and not yet launched - is so far still the first institutional repository initiative in Pharmaceutical Industry. It is most desirable the project can be completed soon to demonstrate the transformational change also taking place in industry.

Lessons Learned & Recommendations

- Build a strong business case before addressing your management!
- Sustainability needs to be addressed realistically (internal vs. external development)
    - Advantages of outsourcing
        - No worries about employee turnover regarding technical development and maintenance
        - Impact on budget is easier to plan in the long term
    - Disadvantages of outsourcing
        - Potential risk of unauthorized data access as data are stored outside the company firewall
        - Much more bureaucracy to get clearance for an outsourced solution from Information Security (Risk Assessments)
        - Dependency on external partner(s)

  Note: Not everything can be oursourced! Plan for internal resources also to run the project and manage the future repository. Make sure your repository project leader has a deputy who can stand in easily if necessary.
- Focus on your individual institution's requirements and then choose the appropriate software to meet them (set up and test demo repositories for a given time). Rather then rushing a software decision and do the requirements specification as a second step only focused on the already selected software package.
- Find out about already existent systems and ongoing projects (either synergistic or hindering) BEFORE launching a repository project.
- Having an overall Open Access strategy in place BEFORE launching a repository project would be more straight forward than the other way round. This strategy would include all activities necessary to launch the repository successfully (policy, communication plan, education, ongoing support, planned associated services, other related Open Access activities complementing the repository).

- During development invest time in a proper project documentation.
- Communicate frequently! Get the buy in from all your stakeholders and keep them involved.