



HAL
open science

L'IST au prisme de l'Europe

Armelle Jézéquel

► **To cite this version:**

Armelle Jézéquel. L'IST au prisme de l'Europe: Actes des FRéDoc 2011. FRéDoc 2011: L'IST au prisme de l'Europe, Oct 2011, Bordeaux, France. pp.200. halshs-00868912

HAL Id: halshs-00868912

<https://shs.hal.science/halshs-00868912>

Submitted on 2 Oct 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

BORDEAUX

10 au 13 octobre 2011

L'IST au prisme de l'Europe



FRéDOC 2011

Crédit photo : Thomas Samson | mairie de Bordeaux

L'IST au prisme de l'Europe

Actes des FréDoc 2011

Collectif, Armelle Jézéquel (coord.)

Remerciements

ARPIST (Réseau des professionnels de l'information scientifique et technique de la région Aquitaine-Limousin) porteur de cet événement tient à remercier la MRCT (Mission ressources et compétences technologiques), le Conseil régional d'Aquitaine, le PRES de Bordeaux, Sciences Po Bordeaux, la ville de Bordeaux, Le CIVB (Conseil interprofessionnel des vins de Bordeaux), la Librairie Mollat, EBSCO, la MAIF, la Banque populaire, CASDEN, l'INSHS (Institut des sciences humaines et sociales) et la délégation Aquitaine du CNRS qui ont fortement soutenu ces journées organisées par le réseau national des professionnels de l'information scientifique et technique du CNRS : Renatis.



Avant-propos

Présenter l'information scientifique au prisme de l'Europe, telle était la mission que le comité d'organisation des FréDoc (réseau Rénatis-CNRS) s'était fixée dans le cadre de ses rencontres à Bordeaux du 10 au 13 octobre 2011¹.

Les interventions de ces journées ne prétendaient pas être exhaustives quant à la multitude de projets en cours ou déjà réalisés, mais avaient pour ambition d'amener les professionnels de l'IST à réfléchir sur l'Europe autour de quatre axes : le cadre institutionnel de l'IST ; les grands projets européens et grandes initiatives ; la mise en commun des ressources et données de la recherche ; les pratiques et projets IST ; et enfin l'organisation professionnelle.

La sélection de contributions réunies dans cet ouvrage aborde dans un premier temps le cadre institutionnel de l'IST en Europe à travers quelques grands projets et grandes initiatives européens. Un représentant de la Commission européenne, **Jarkko Siren** expose la stratégie de l'Europe dans le cadre de l'IST et de la recherche, par la mise en place de programmes et d'infrastructures structurantes liées à ces domaines. Comment ces initiatives doivent-elles financer, coordonner et/ou promouvoir le libre accès, conserver et diffuser les données de la recherche, au sein des États membres ? **Thérèse Hameau** dresse, elle, un panorama des projets européens en faveur du libre accès mis en place grâce à diverses impulsions de la Commission européenne, sans omettre de soulever les multiples problématiques induites par la diversité des politiques nationales des pays membres. L'aspect économique prenant de plus en plus d'importance, la diffusion des ressources est associée à un développement des contenus numériques : payant ou gratuit. La volonté européenne d'offrir à tout citoyen un accès libre et gratuit aux résultats issus de la recherche et financés par des fonds publics oblige à reconsidérer les modes de fonctionnement.

La deuxième partie traite de la mutualisation des ressources et données de la recherche. Dans un premier temps, **Alain Monteil** et **Fouil Bretel** témoignent de l'implication de l'INRIA au sein du projet PEER (*Publishing and the Ecology of European Research*). Ainsi, permettre une diffusion de la production scientifique en ligne sous-entend de trouver des solutions techniquement adaptées, mais aussi de tenir compte du circuit de diffusion commercial des revues et de la place des éditeurs au sein du mouvement des archives ouvertes. Côté Mathématiques, **Thierry Bouche** présente le projet EuDML. Il s'agit d'une bibliothèque virtuelle globale disposant d'un point d'accès au contenu distribué, spécialisée recherche en mathématiques et ce grâce à des interfaces faciles à utiliser. L'ambition est présente dès le départ car il s'agit de constituer « une bibliothèque numérique mondiale et exhaustive », où le libre accès est total et les résultats mis à la disposition de tous dans le temps, en accord avec les éditeurs. Pour ce faire, des moyens techniques et financiers sont affectés afin de permettre la numérisation de revues de mathématiques, mais aussi des résultats beaucoup plus anciens en tenant compte des particularités du domaine, telles que

¹ Les communications (diapos et vidéos) ainsi que les posters des FréDoc2011 sont consultables en ligne <<http://renatis.cnrs.fr/spip.php?article199>>, <<http://renatis.cnrs.fr/spip.php?article200>>, <<http://www.arpist.cnrs.fr/>>.

les formules mathématiques. **Ivana Roche** aborde, elle, le concept de l'évaluation par les pairs et de la production d'indicateurs. Nos décideurs, ainsi que la Commission, ont besoin, lors de leurs examens, d'éléments d'analyse pour mieux évaluer la qualité des productions scientifiques. Comment des outils statistiques automatisés peuvent-ils les aider à saisir l'innovation ? à comparer les résultats ? les aider lors de la prise de décision... Vient ensuite la présentation de l'Observatoire virtuel astronomique par **Thomas Boch**. Il s'agit dans ce cas, de la mise en place d'un prototype au sein d'un ambitieux projet international, basé sur un système d'interopérabilité, de partage et de standards, qui constitue un outil au service des immenses quantités de données produites par la recherche astronomique. Enfin **Jean-Luc Pinol** aborde l'exemple du TGE² Adonis et de ses multiples collaborations au sein de projets européens. Il montre comment l'activité d'un réseau peut innover en SHS³ dans le domaine du numérique et des infrastructures qui lui sont inhérentes, dans le but de développer des outils à destination des chercheurs.

La troisième partie s'articule autour d'exemples de projets déjà à l'œuvre ou en cours grâce aux témoignages de divers acteurs en Europe. Ainsi **Barbara Szczepanska**, nous présente les initiatives polonaises ayant permis de mettre en place et de développer des bibliothèques numériques ainsi que le libre accès au sein des bibliothèques universitaires. Puis comment en s'unissant, dans le cadre du Consortium polonais des bibliothèques numériques et de la Fédération des bibliothèques numériques, l'éventail des contenus s'est élargi afin de ne plus simplement satisfaire la science, mais de s'ouvrir aux champs de l'éducation et de la culture, avec le soutien financier de l'État et parfois même de fonds européens. À travers plusieurs exemples, elle nous présente les choix et les structurations adoptés tout en nous éclairant sur les difficultés rencontrées de par la loi nationale relative au droit d'auteur. Sur un autre plan, **Teresa Abejón** dresse un panorama des systèmes d'accès à l'IST en Espagne. Bien qu'au départ fondées sur des catalogues de bibliothèques spécialisées et sur des bases de données bibliographiques, les ressources se transforment rapidement en de multiples portails en accès libre. Développer ces nouveaux outils devient un objectif stratégique dont s'emparent les organismes institutionnels de recherche qui représentent aujourd'hui la majorité des dépôts (soit 69 % pour les universités et centres de recherche). Cette organisation autour de la diffusion des ressources grâce au numérique tend aujourd'hui à une centralisation institutionnelle des accès, mais il perdure cependant le problème de l'articulation de ces divers systèmes entre eux. S'ajoute la contribution de **Marian Koren**, qui s'attache à nous exposer la manière dont FOBID (Forum des bibliothèques néerlandaises) s'implique activement dans la politique d'information scientifique nationale afin de toujours concilier innovation et infrastructure au profit des bibliothèques numériques aux Pays-Bas. Au moment où les économies budgétaires sont de rigueur, comment intensifier les collaborations par des coopérations nationales et européennes tout en fédérant autour du développement de nouveaux usages ? Enfin, **Deborah Shorley**, nous expose la résolution de problèmes liés au manque de place au sein des bibliothèques universitaires britanniques grâce à la participation au réseau UKKR (*United Kingdom Research Reserve*).

La dernière partie de ce volume a trait à l'organisation professionnelle. **Sylvie Dalbin** consacre sa réflexion aux métiers et à l'organisation professionnelle évoluant dans le cadre

² Très grand équipement.

³ Sciences humaines et sociales.

de perspectives européennes. Comment l'IST et surtout ses acteurs doivent-ils s'unir et s'adapter en permanence à la multiplicité des formats, des supports, des données, produits par la recherche ? Quelles sont les possibilités d'apprentissages qui sont mis à leur disposition au cours de leur cheminement professionnel ? Comment l'Europe marque-t-elle par ses décisions les métiers et les organisations de l'IST ? ... Enfin, IdRef nous est présenté par **Yann Nicolas**. Il s'agit de référentiels Sudoc mis à disposition des professionnels de l'IST afin de résoudre des problèmes d'homonymie et de dispersion des références au sein des bases de données bibliographiques universitaires françaises.

Dans un paysage en évolution, devant l'abondance de données et de publications qui se confondent de plus en plus, leur organisation, leur mise à disposition, dépendent de projets communs. De nouveaux enjeux apparaissent pour les réseaux. Ces journées ont permis d'éclairer un certain nombre de questions soulevées par l'engagement dans les projets européens de la communauté de l'IST et de lui donner des clés afin d'affronter les nouveaux enjeux.

Armelle Jézéquel

Coordination de l'ouvrage

Sommaire

Remerciements	I
Avant-propos Armelle Jézéquel	III
Sommaire	V
Cadre institutionnel de l'IST en Europe	
Grands projets européens et grandes initiatives	
La DIST au CNRS (Direction de l'information scientifique et technique) <i>Francis André</i>	3
Strategies for Scientific and Technological Information in Europe <i>Jarkko Siren</i>	5
Panorama des projets européens en faveur du libre accès à l'IST <i>Thérèse Hameau</i>	13
Removing the Barriers! Promoting Open Access to Research through Advocacy and Collaboration <i>Lars Björnshaugen</i>	23
Mise en commun des ressources et données de la recherche	
Facilitating Open Access through International Networks: OpenAIRE and COAR <i>Birgitt Schmidt</i>	27
PEER, <i>Publishing and the Ecology of European Research</i> , retour d'expérience d'INRIA sur un projet associant éditeurs et archives ouvertes <i>Alain Monteil et Foudil Bretel</i>	29
EuropeanaLocal et la bibliothèque numérique européenne. La Gironde, département français pionnier <i>Nathalie Gascoin</i>	45
Projet EuDML, une bibliothèque numérique européenne de mathématiques <i>Thierry Bouche</i>	47
A Concept for Inferring 'Frontier Research' in Research Project Proposal <i>Ivana Roche et al.</i>	57
L'Observatoire virtuel astronomique : interopérabilité et standards au service de la recherche astronomique <i>Thomas Boch</i>	73
Le TGE Adonis et les projets européens <i>Jean-Luc Pinol</i>	83

Pratiques et projets IST en Europe

AllianzInitiative, Germany <i>Ralf Schimmer</i>	93
Pratiques et projets IST en Pologne <i>Barbara Szczepanska</i>	95
Ressources pour l'accès à l'information scientifique publique en Espagne : le scénario actuel, les projets et les tendances dans la diffusion de l'IST <i>Teresa Abejón</i>	109
Clouds and Crowds in the Netherlands Library Landscape <i>Marian Koren</i>	119
UKKR, today's solution to yesterday's problems <i>Deborah Shorley</i>	141
LIBER and EU Projects <i>Wouter Schallier</i>	149
IFLA (<i>International Federation of Library Associations and Institutions</i> , Fédération internationale des associations de bibliothécaires et d'institutions) <i>Pascal Sanz</i>	151
Organisation professionnelle	
Métiers et organisation professionnelle, perspectives européennes <i>Sylvie Dalbin</i>	155
Outils et la qualité des services universitaires de documentation : Libqual et Qualibib <i>Olivier Chourrot</i>	173
IdRef. Les référentiels Sudoc comme services aux réseaux documentaires <i>Yann Nicolas</i>	175
Conclusion <i>Herbert Grüttemeier</i>	183
Références	185
Glossaire	197
Comité d'organisation FréDoc 2011	199

Partie 1
Cadre institutionnel

La DIST au CNRS (Direction de l'information scientifique et technique)

Francis André

*Directeur adjoint de la DIST, CNRS
3 rue Michel-Ange, 75016 Paris, France*

Résumé

Les missions de la DIST concourent à la valorisation et la diffusion de la production scientifique : collecte, traitement, archivage et diffusion de la production scientifique nécessaire aux différentes communautés de recherche. La DIST est la tutelle du CCSD et de l'INIST rénové, elle collabore également avec le CLEO, ADONIS. Parallèlement, la DIST mène, en collaboration avec les autres établissements publics de recherche et d'enseignement supérieur, une réflexion sur la politique d'acquisition des ressources numériques dans un cadre mis en place par le MESR, en particulier dans le cadre de la Bibliothèque scientifique numérique (BSN).

La DIST agit dans plusieurs réseaux et collaborations européens, OpenAire, Eurohorcs, *Knowledge Exchange*, MedOANet...

Abstract

The DIST at the CNRS

The DIST's mission is to contribute to the promotion and dissemination of scientific production. This covers the collection, processing, archiving and dissemination of the scientific production required by the different research communities. The DIST is the authority responsible for the CCSD (*Centre for Direct Scientific Communication*) and the reformed INIST (*Institute for Scientific and Technical Information*) and also works in collaboration with the CLEO and ADONIS amongst other partners. In parallel to these activities, DIST is leading a study of French acquisitions policy regarding digital resources in collaboration with other public sector research and higher education institutions within a framework defined by the French Ministry for Higher Education and Research (MESR) and with particular focus on the French National Digital Library Infrastructure (BSN).

The DIST also participates in several networks and partnerships at the European level, including OpenAire, Eurohorcs, Knowledge Exchange and MedOANet.

Strategies for Scientific and Technological Information in Europe

Jarkko Sirén¹

European Commission DG INFSO, 25 Avenue de Beaulieu
Bruxelles, Belgique

Abstract

The Scientific and Technological Information (STI) strategy of the European Commission addresses policies for scientific information and related ICT infrastructures. Significant progress has been made since the start of FP7 while ensuring a good communication with stakeholders and coordination with European Member States. Leading by example has been a major concern and therefore the European Commission launched concrete initiatives such as the FP7 Open Access Pilot addressing the issues of access to the EC funded research results.

This paper summarises the history of EC STI policy since 2007 when a Communication² on Scientific Information in the Digital Age was first published. It highlights the current emphasis on open access, sustainable preservation, and on building reliable and participatory e-infrastructures that give rise to an 'online European Research Area'. More coherence and better coordinated national policies and Union wide infrastructures are at the core of the EC strategy to achieve the policy goals in a cost effective way.

1. Introduction

Although several other divisions are possible, for reasons of clarity the Scientific and Technological Information (STI) strategy of the European Commission can be divided in (1) Scientific Information policies, seeking to define priorities for STI in Europe as well as to harmonise and complement national policies and (2) e-Infrastructures for scientific information, seeking to create cross border infrastructures that enable creation, access, (re)use and preservation of scientific information.

These two aspects are two sides of the same coin; a good example is the access to the results of EU-funded research which stem from EC Open Access policies and rely on a distributed and participatory e-Infrastructure. This combined approach helps bridging gaps

¹ Disclaimer: The views expressed in this presentation are those of the author and do not necessarily reflect the views of the European Commission.

² A Communication is a formal mechanism for the EC to interact with the European Parliament, The Council of EU Member States and the Committee of the Regions.

between Member States and the digital divide between the more and less developed countries.

The EC's STI strategy is traced back to the 2007 Communication on Scientific Information [1]. This was followed by a Communication on ICT infrastructures for e-science in 2009, including a strategy to develop Scientific Data Infrastructures in Europe [3]. As a result of the subsequent discussions with EU Member States the European Commission launched the Open Access Pilot in FP7 (the pilot was adopted in August 2008 and the supporting OpenAIRE³ infrastructure became operational in 2010). The pilot aims to make results of EC funded research in seven scientific domains available for free to researchers and interested EU citizens. The Pilot will run until the end of the 7th Framework Programme (2013).⁴

The Commission started in 2011 the revision of the associated policies that make up the recently adopted umbrella Communication on "Open data: An engine for innovation, growth and transparent governance". This revised package includes a specific Communication and a Recommendation on Scientific Information (foreseen in the first half 2012 [2]⁵), as well as a proposal for an update for the PSI directive which was published in 2011 together with the umbrella Communication on Open Data⁶ [2, 4].

The EC has already expressed its intention to extend the current OA policies to its next Framework programme, Horizon 2020 (H2020). Proposals have been made to include all scientific domains (compared to 7 domains in FP7; see also chapter 4 below) as well as research data.⁷ Relevant information on EC future strategies for STI can be found also in the Digital Agenda for Europe [5]⁸ and Innovation Union Communications [6].⁹ These initiatives are part of the Europe 2020 strategy that aims to increase employment, productivity and social cohesion in Europe by 2020.

2. Policy on Scientific Information

Three fundamental traits are present throughout the EC policy on STI. The first of them, digital technologies' reshaping of research information, has a decisive influence on the other two. These are the objectives of (1) better access to scientific information (to boost collaboration in and quality of research) and (2) preservation of research results for future generations (to capture the added value of research and explore economies of scale). There

³ <<http://www.openaire.eu/>> consulted on March 1st 2012.

⁴ Relevant documents and additional information on FP7 OA Pilot can be found in <<http://ec.europa.eu/research/science-society/index.cfm?fuseaction=public.topic&id=1300&lang=1>> consulted on March 1st 2012.

⁵ <http://ec.europa.eu/information_society/policy/psi/docs/pdfs/opendata2012/open_data_communication/opendata_EN.pdf>, p. 12, consulted on March 1st 2012.

⁶ <http://ec.europa.eu/information_society/policy/psi/index_en.htm> consulted on March 1st 2012.

⁷ A good synthesis of OA proposals for Horizon 2020 can be found in <http://ec.europa.eu/research/science-society/document_library/pdf_06/open-access-in-horizon-2020_en.pdf> consulted on March 1st 2012.

⁸ P. 23: "...publicly funded research should be widely disseminated through Open Access publication of scientific data and papers".

⁹ Particularly p. 19.

are other important and related objectives of the EC STI policy. The most significant one from the 2007 communication all the way to the Europe 2020 strategy is the idea that these policies should increase the quality and innovative potential of European research [1-6]. The former research policy objectives (better access and preservation of knowledge) thus supports the second socioeconomic objective based on the idea that the former will help Europe reach its competitiveness (Lisbon strategy) and growth (Europe 2020 strategy) targets [1, 6].

The Open Access publishing model for scientific journals is present in many of the issues and solutions discussed in the 2007 Communication. The main motivation behind this is the so called serials crisis [1]. In addition to calling for wider access to scientific journals, the Commission aligned itself with the 2004 OECD Ministerial Declaration on Access to Research Data from Public Funding suggesting that publicly funded research data should be available to all [1, 7].

The 2012 Communication on Scientific Information will be accompanied by a Recommendation to the Member States. The preliminary steps for this Communication and Recommendation included a public hearing in May 2011 and a public consultation from July to September 2011 as well as several workshops organised by the EC.¹⁰ The Recommendation, a soft law coordination instrument directed to EU Member States, will focus on access to and preservation of scientific information. These developments take place in a strategic moment when Europe is preparing the next Framework Programme for research and innovation, Horizon 2020.

During this process the Commission produced and published a provisional version of the Open Data Communication (December 2011). This Communication discusses the benefits of open government data and emphasizes that data originating from publicly funded research should also be made widely available to strengthen Europe's innovative potential and economic growth. It proposes a strategy for all types of public data consisting of three key elements [2]:

- 1- Adapting the framework for data re-use, including legal, soft law and policy measures;
- 2- Mobilising financing instruments by prioritising open data in R&D&I and infrastructure programmes;
- 3- Facilitating coordination and experience sharing across the Member States.¹¹

As part of the Europe 2020 strategy also the Innovation Union Communication seeks to promote free movement of knowledge and open access to publications and data. Both are seen as means to strengthen the European Research Area, a 'single market' for research and innovation [6].

¹⁰ <<http://ec.europa.eu/research/science-society/index.cfm?fuseaction=public.topic&id=1301&lang=1>> consulted on March 1st 2012.

¹¹ P. 8.

3. ICT Infrastructures for scientific information

In almost all fields of science huge datasets are being created with data from observations and simulations. In these fields, new ways of doing research have created the concept of “data driven science”. Consequently ICT infrastructures for scientific data are increasingly emerging world-wide. However, many barriers still exist across countries and disciplines making interoperability and sustainability difficult to achieve. Therefore commonly agreed governance, legal frameworks and funding models are needed. To cope with the extremely large or complex datasets generated and used in research, it is essential to take a global approach to promote interoperability and discoverability of scientific information resources.

In its capacity building role, the EC promotes and funds ICT infrastructures for research, enabling a rapid transition to the so called e-science. The 2009 Commission Communication on ICT infrastructures for e-science highlights the strategic role of IT in contemporary scientific discovery process and seeks to promote increasing adoption of ICT in all phases of this process [3].

In the area of scientific data the 2009 Communication highlights the need to tackle the accelerated and uncontrolled proliferation of data. It also focuses on the urgency to develop a coherent strategy to overcome the fragmentation in data infrastructures and to enable research communities to better manage, use, share and preserve data.¹² In its Council Conclusion of December 2009, the Council of the European Union invited Member States and the Commission to broaden access to scientific data and open repositories and ensure coherent approach to data access and curation [9].

More recently, in October 2010 the High Level Expert Group on Scientific Data submitted its final report to the Commission. Although not an official Commission document, the conclusions of this report are likely to be considered when deciding the Commission’s future priorities for research infrastructures [10]. The main conclusion of the report is that there is a need for a “collaborative data infrastructure” for science in Europe and globally, as well as for the development of a broad conceptual framework on how different stakeholders would interact with such a system. The vision this system would enable is described in the following terms:

“Our vision is a scientific e-infrastructure that supports seamless access, use, re-use, and trust of data. In a sense, the physical and technical infrastructure becomes invisible and the data themselves become the infrastructure – a valuable asset, on which science, technology, the economy and society can advance.”¹³

These conclusions have already made their way to the objectives of the recently launched projects OpenAIREplus and EUDAT.¹⁴ Both projects will contribute to the efforts to establish an open forum on standards, APIs, policy rules, etc. for a collaborative data infrastructure in Europe and globally.

¹² P. 6-7.

¹³ P. 4.

¹⁴ <<http://www.eudat.eu/>> and <<http://www.openaire.eu/>> consulted on March 1st 2012.

4. Access to the results of EU-funded research

The 2007 Communication was published soon after the beginning of FP7. It was followed by Council Conclusions that, while inviting the Member states to reinforce their access and preservation policies and enhance their mutual coordination, also invited the EC to experiment with Open Access to scientific data and publications resulting from projects funded by the EU [8]. Consequently the Commission launched its FP7 Open Access Pilot in August 2008. The pilot covers seven scientific areas that represent approximately 20% of the FP7 budget. It includes a mandate to self-archive peer-reviewed research articles in open access as 'best effort'. Embargo periods of 6 or 12 months are foreseen when necessary.¹⁵

The FP7 Open Access Pilot is supported and monitored through the OpenAIRE project. It provides a technological infrastructure that enables deposition of peer-reviewed articles and harvesting of metadata from articles deposited in other repositories. Its human infrastructure supports the EU-funded researchers through helpdesk, training and dissemination activities. Its key objective is to enable fast and reliable access to EU-funded research results, in particular peer-reviewed research articles.¹⁶

In addition to FP7 related developments and coherently with its promotion of open access, the Innovation Union Communication also foresees that the Commission "will aim to make open access to publications the general principle for projects funded by the EU research Framework Programmes".¹⁷ [6]

The development of the so called "online ERA" is an important policy aspect in e-Infrastructures area in general and in the data infrastructures domain in particular. As part of the Digital Agenda flagship initiative the development of e-Infrastructures is envisioned as means to connect researchers, instruments, data and computation resources throughout Europe. These efforts are expected to create a seamless knowledge territory or "online ERA". [11]

5. Conclusion

This paper provides a summary of the main components of the EC strategy for STI. More coherence and better coordinated national policies and Union-wide infrastructures are at the core of the EC strategy. Furthermore, as a research funding body, the Commission is increasing its efforts towards wider access to the results of the projects that have received its support.

¹⁵ <<http://ec.europa.eu/research/science-society/index.cfm?fuseaction=public.topic&id=1300>> consulted on March 1st 2012.

¹⁶ <<http://www.openaire.eu/>> consulted on March 1st 2012.

¹⁷ P. 19.

Acknowledgements

I would like to thank Kostas Glinos (Head of Unit DG INFSO e-Infrastructures) and Carlos Morais Pires (Head of Sector DG INFSO Data Infrastructures) for their contributions and support.

Bibliography

- [1] European Commission, Commission of the European Parliament, the Council and the European Economic and Social Committee, *Communication on Scientific Information in the Digital Age: Access, dissemination and Preservation*, Brussels, 14 February 2007, **COM(2007) 56 final**. <<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2007:0056:FIN:EN:PDF>> consulted on March 1st 2012.
- [2] European Commission, Commission of the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, *Communication on open data: An Engine for Innovation, Growth and Transparent Governance*, Brussels, provisional version **COM(2011)XXX draft**. <http://ec.europa.eu/information_society/policy/psi/docs/pdfs/opendata2012/open_data_communication/opendata_EN.pdf> consulted on March 1st 2012.
- [3] European Commission, Commission of the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, *Communication on ICT infrastructures for e-Science*, Brussels, 5 March 2009, **COM(2009) 108 final**. <<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2009:0108:FIN:EN:PDF>> consulted on March 1st 2012.
- [4] European Commission, Commission of the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, *Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL Amending Directive 2003/98/EC on re-use of public sector information*, Brussels, provisional version **COM(2011)XXX draft**. <http://ec.europa.eu/information_society/policy/psi/docs/pdfs/opendata2012/revision_of_PSI_Directive/proposal_directive_EN.pdf> consulted on March 1st 2012.
- [5] European Commission, Commission of the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, *Communication A Digital Agenda for Europe*, **COM(2010) 245 final/2**, Brussels, 26 August 2010. <<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:0245:FIN:EN:PDF>> consulted on March 1st 2012.
- [6] European Commission, Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, *Communication Europe 2020 Flagship Initiative. Innovation Union*, Brussels, 6 October 2010, **COM(2010) 546 final**. <<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:0546:FIN:EN:PDF>> consulted on March 1st 2012.
- [7] OECD, *Principles and Guidelines for Access to Research Data from Public Funding*, **OECD, 2007**. <<http://www.oecd.org/dataoecd/9/61/38500813.pdf>> consulted on March 1st 2012.
- [8] Council of the European Union, *Council Conclusions on scientific information in the digital age: access, dissemination and preservation. 2832nd COMPETITIVENESS (Internal market, Industry and Research) Council meeting Brussels, 22 and 23 November 2007*. <http://www.consilium.europa.eu/ueDocs/cms_Data/docs/pressData/en/intm/97236.pdf> consulted on March 1st 2012.
- [9] Council of the European Union, *Conclusions on the future of information and communication*

technologies research, innovation and infrastructures. 2982nd COMPETITIVENESS (Internal market, Industry and Research) Council meeting, Brussels, 3 December 2009.

<http://www.consilium.europa.eu/uedocs/cms_Data/docs/pressdata/en/intm/111719.pdf>

consulted on March 1st 2012.

- [10] European Union, *Riding the Wave: How Europe can gain from the rising tide of scientific data. Final report of the High Level Expert Group on Scientific Data. A submission to the European Commission*, Brussels, 2010. <<http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf>> consulted on March 1st 2012.
- [11] European Commission, Commission of the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, *Communication on On the Response to the Report of the Expert Group on the Interim Evaluation of the Seventh Framework Programme for Research, Technological Development and Demonstration AZactivities and to the Report of the Expert Group on the Interim Evaluation of the Risk-Sharing Finance Facility, COM(2011)52 final*, Brussels, 9th February 2011.
<http://ec.europa.eu/research/evaluations/pdf/archive/other_reports_studies_and_documents/commission_response_fp7_ie_report_2011.pdf> consulted on March 1st 2012.

Panorama des projets européens en faveur du libre accès à l'IST

Thérèse Hameau

INIST-CNRS
2 allée du parc de Brabois CS 10310 54514
Vandœuvre-lès-Nancy Cedex, France

Comprendre le développement du libre accès en Europe, nécessite d'en rappeler quelques notions avant de broser le panorama des projets et d'en voir les développements en Europe. Enfin, il convient de se pencher sur la question des problématiques, qui bien qu'exprimées dès le départ du mouvement, sont toujours d'actualité.

1. Rappel sur le libre accès

On peut diviser le mouvement du libre accès en trois grandes périodes.

Les années 1990 durant lesquelles des personnes comme Stevan Harnad et Paul Ginsparg ont été des précurseurs : le premier avec le lancement des revues en libre accès ainsi que d'une archive ouverte, le second avec la création, en 1991, d'une archive en physique connue maintenant sous le nom d'arXiv.

Puis, avec le début des années 2000, le mouvement du libre accès acquiert plus de visibilité avec la lettre ouverte¹ de *Public Library of Science* (2001), qui encourage la création d'une « bibliothèque publique en ligne qui fournirait le contenu intégral des résultats publiés de la recherche et des textes scientifiques dans le domaine de la médecine et des sciences du vivant sous une forme en libre accès, interrogeable dans son intégralité et interconnectée ». Pendant cette période une prise de conscience plus générale se manifeste avec un grand nombre de déclarations et de prises de position à ce sujet. Se succèdent ainsi des déclarations, dont celles des 3 B : Budapest² (2002), Bethesda³ et Berlin⁴ (2003). Avec la déclaration de Berlin, le mouvement prend un tournant, car contrairement aux deux autres, ce sont des institutions qui sont signataires et les données brutes sont alors prises en compte. Celles-ci seront reprises dans la déclaration de l'OCDE⁵ en 2004.

Enfin, la période après 2006 est une phase de consolidation par la mise en œuvre de politiques sous forme de mandats et en corollaire la création de diverses archives ouvertes. La majorité des mandats incitent les chercheurs à déposer leurs articles, peu les appellent à publier dans des revues en libre accès. La publication de ces mandats est effectuée par des

¹ <<http://openaccess.inist.fr/spip.php?article6>> consulté le 2 mars 2012.

² <<http://openaccess.inist.fr/spip.php?article9>> consulté le 2 mars 2012.

³ <<http://openaccess.inist.fr/spip.php?article20>> consulté le 2 mars 2012.

⁴ <<http://openaccess.inist.fr/spip.php?article38>> consulté le 2 mars 2012.

⁵ <<http://openaccess.inist.fr/spip.php?article28>> consulté le 2 mars 2012.

institutions ayant des rôles différents au sein du processus de recherche. Soit il s'agit d'agences de financement soit d'organismes de recherche dont de nombreuses universités. Leur rapport avec le chercheur est différent, et la condition pour suivre la politique est, pour les premières basée sur le financement de la recherche, et pour les seconds, sur l'appartenance du chercheur à l'organisme.

Si au départ, ce sont les articles de revues scientifiques qui sont visés, on assiste rapidement à un élargissement à d'autres types de documents : les ouvrages, les données de la recherche, ainsi que tout ce qui a trait à l'enseignement, comme les cours et les thèses. Cette évolution est liée au fait que ce sont des archives institutionnelles qui se développent en lien avec les mandats des institutions. En effet, si les premières archives sont thématiques, de plus en plus d'institutions créent leur archive qui n'est pas dédiée uniquement aux articles des chercheurs appartenant à l'institution, mais destinée à recueillir l'ensemble des documents créés par l'institution, dont des documents pédagogiques, des cours et des thèses.

Par ailleurs, elles demandent le dépôt du texte intégral, mais beaucoup de chercheurs ne font que signaler leurs publications lors du dépôt des métadonnées.

Certains mandats demandent également le dépôt des données en lien avec les articles publiés.

Deux moyens ont été préconisés depuis le début du mouvement et notamment dans l'*Initiative de Budapest pour l'Accès Ouvert* :

- l'auto-archivage des articles dans une archive ouverte, appelée plus tard voie verte ;
- la publication dans des revues en libre accès, ou voie dorée.

Pour les archives ouvertes, on peut remarquer que ce terme est délaissé au profit de réservoir, de dépôt et même de bibliothèque.

La publication dans des revues répond quant à elle à différents modèles : l'accès gratuit pour les lecteurs, mais la publication est gratuite ou payante pour les auteurs ou leurs institutions, le modèle hybride proposé par de plus en plus d'éditeurs. Il associe, pour une même revue, le libre accès et l'accès par abonnement. Si l'auteur choisit de payer pour être publié, le lecteur accède gratuitement à son article ; les autres articles sont accessibles si le lecteur ou son institution a souscrit à un abonnement.

Dans les différentes déclarations et politiques, il s'agit d'articles évalués par les pairs. Mais, le dépôt concerne les différentes versions de l'article en relation avec les politiques des éditeurs⁶ : de la prépublication « manuscrit envoyé par le chercheur à l'éditeur » à la version de l'éditeur.

Les archives sont un des composants de l'infrastructure de la recherche. Pour éviter d'avoir des réservoirs isolés, ceux-ci doivent être capables de communiquer et de s'interconnecter : on parle d'interopérabilité. OAI-PMH⁷ est un protocole permettant ce

⁶ <Voir SHERPA / ROMEO : <<http://www.sherpa.ac.uk/romeo/>> consulté le 2 mars 2012.

⁷ *Open Archives Initiative's Protocol for Metadata Harvesting.*

dialogue. Il date de la convention de Santa Fé en 2000 qui elle-même succède à un séminaire de l'*Open Archives Initiative* en 1999⁸.

Les disciplines scientifiques se sont coordonnées plus ou moins rapidement : la physique et les mathématiques font partie des sciences pionnières ; les domaines SHS se singularisent par le fait que les chercheurs publient sous forme de monographies. Certains acteurs ont été moteurs, comme les *National Institutes of Health* aux États-Unis qui ont mis en place PubMed Central qui rassemble la production mondiale dans les domaines du vivant.

2. La Commission européenne et l'Europe

La Commission européenne a émis un certain nombre de documents ayant trait au libre accès qui peuvent être déclinés en quatre étapes.

La première est l'étude de 2006 portant sur le marché de l'édition scientifique en Europe *Study on the Economic and Technical Evolution of the Scientific Publication Markets in Europe*⁹. Une de ses recommandations est que « les résultats de recherches financées par des fonds publics devraient être accessibles rapidement après leur publication ». Une autre implique « d'expérimenter des modèles économiques, d'apporter un soutien pour recherches sur copyright, formes alternatives de distribution et les développements technologiques ».

Une autre instance, le Conseil européen pour la recherche, déclare en 2006¹⁰ que : « These considerations lead the ERC Scientific Council, like other research funding bodies, to stress the attractiveness of policies mandating the public availability of research results in open access repositories reasonably soon (ideally, 6 months, and in any case no later than 12 months) after publication. »

Puis, se succèdent des communications et des consultations, dont celles de 2007 qui vont déboucher sur un projet-pilote. Dans la communication sur l'information scientifique à l'ère du numérique¹¹, la Commission prend position en indiquant que : « Des initiatives qui conduisent à un accès plus large et à une meilleure diffusion de l'information scientifique sont nécessaires, en particulier en ce qui concerne les publications et les données brutes générées par les activités de recherche soutenues par un financement public. ». Elle poursuit en annonçant que « Dans le cadre du 7^{ème} PCRD, la Commission prendra des mesures pour promouvoir un meilleur accès aux publications qui découlent de la recherche soutenue par ses programmes. Dans ce contexte, les coûts des projets associés à la publication des résultats, y compris l'édition en libre accès, seront éligibles pour la contribution financière de la Communauté. »

Et en 2008, démarre un projet pilote¹² dans le cadre du 7^e PCRD dans lequel « Les bénéficiaires de subventions seront tenus de déposer dans une bibliothèque en ligne leurs

⁸ *The Santa Fe Convention of the Open Archives Initiative*, DOI : 10.1045/february2000-vandesompel-oai.

⁹ <http://ec.europa.eu/research/science-society/pdf/scientific-publication-study_en.pdf> consulté le 2 mars 2012.

¹⁰ <http://erc.europa.eu/sites/default/files/press_release/files/erc_scc_statement_2006_open_access_0.pdf> consulté le 2 mars 2012.

¹¹ <<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2007:0056:FIN:FR:PDF>> consulté le 2 mars 2012.

¹² <<http://ec.europa.eu/research/science-society/index.cfm?fuseaction=public.topic&id=1300&lang=1>> consulté le 2 mars 2012.

articles scientifiques ou manuscrits définitifs évalués par des pairs et résultant de projets menés dans le cadre du 7^e PC. Ils devront s'efforcer d'assurer un libre accès à ces articles, dans un délai de six ou douze mois – fixé en fonction du domaine de recherche – après leur publication. Cet embargo doit permettre aux éditeurs de publications scientifiques de rentabiliser leur investissement. » Sept domaines de recherche sont ainsi concernés, dont la santé et l'énergie, soit 20 % du budget du 7^e PCRD qui atteint 50 milliards d'euros.

La dernière consiste en une consultation publique¹³ qui s'est terminée en septembre et qui doit déboucher sur une communication officielle prévue en fin d'année¹⁴. Les parties intéressées étaient invitées à exprimer leur point de vue sur les grandes questions en matière de politique scientifique : l'accès aux articles et plus largement aux résultats de la recherche, la réutilisation de ces résultats et leur préservation à long terme.

De multiples projets européens ont vu le jour autour de la problématique du libre accès.

Concernant les archives, on peut citer deux projets : DRIVER et OpenAIRE.

DRIVER¹⁵ (*Digital Repository Infrastructure Vision for European Research*), dans sa première phase, a dressé un inventaire des archives ouvertes au sein de l'Union européenne. Des guides techniques ont été publiés afin d'aider à créer une infrastructure paneuropéenne dans ce domaine. La phase 2 consiste à conforter les résultats de la phase 1 en termes d'infrastructure, par l'extension du réseau des archives (confédération européenne de réservoirs numériques), et la mise en place de fonctionnalités pour les chercheurs et le grand public.

Le projet OpenAIRE (*Open Access Infrastructure for Research in Europe*) fait l'objet d'une communication dans cet ouvrage par Birgit Schmidt. En résumé, OpenAIRE¹⁶, a comme objectif la construction d'une archive destinée à recevoir les publications concernées par le projet pilote du 7^e PCRD.

Ces deux projets peuvent être associés à COAR¹⁷ (*Confederation of Open Access Repositories*) dont font partie un grand nombre de pays européens. Il s'agit d'une association lancée en octobre 2009 réunissant 59 institutions de 23 pays d'Europe, d'Amérique latine, Asie et Amérique du Nord. Sa mission est d'accroître la visibilité des résultats de la recherche en s'appuyant sur le réseau des archives ouvertes.

Les projets OAPEN, PEER et SOAP sont eux plus orientés vers les publications.

OAPEN¹⁸ (*Open Access Publishing in European Networks*) porte sur la publication d'ouvrages en libre accès en sciences humaines et sociales. L'objectif technique est de réaliser, à partir de l'existant, une plate-forme évolutive d'édition qui sera disponible en *open source*. C'est maintenant une fondation.

¹³ <http://ec.europa.eu/research/consultations/scientific_information/consultation_en.htm> consulté le 2 mars 2012.

¹⁴ La communication n'est pas encore publiée début janvier 2012.

¹⁵ <<http://www.driver-repository.eu/>> consulté le 2 mars 2012.

¹⁶ <<http://www.openaire.eu/>> consulté le 2 mars 2012.

¹⁷ <<http://www.coar-repositories.org/>> consulté le 2 mars 2012.

¹⁸ <<http://project.oapen.org/>> consulté le 2 mars 2012.

PEER¹⁹ (*Publishing and the Ecology of European Research*) étudie les effets du dépôt des articles sur l'accès par les lecteurs, la visibilité du chercheur et la viabilité de la revue (cf. dans cet ouvrage la contribution d'Alain Monteil et de Foudil Bretrel).

SOAP²⁰ (*Study of Open Access Publishing by Key Stakeholders*) analyse les pratiques des éditeurs et des chercheurs vis-à-vis des publications en libre accès. Il a pris fin en février 2011.

Le projet MEDOANET est né lors d'un séminaire réunissant six pays du sud de l'Europe autour du thème du libre accès, à l'issue duquel les participants ont publié la *Déclaration de l'Alhambra sur le Libre Accès*²¹ (mai 2010). Ce projet a comme objectif de coordonner les stratégies et les politiques à un niveau régional et national.

Après avoir rappelé quelques notions et abordé l'action de la Commission européenne, voici une illustration chiffrée des mandats, des archives et des revues en libre accès en Europe.

- JULIET²², service créé par SHERPA, dénombre 61 mandats d'agences de financement, dont 40 en Europe (tableau 1). ROARMAP²³, développé par l'université de Southampton, répertorie 300 mandats d'agences et d'institutions, dont 213 en Europe (tableau 2). Il faut noter que la notion d'agence est plus ou moins bien respectée selon le répertoire et les pays.

Le Royaume-Uni arrive encore en tête pour le nombre d'archives. Cette situation peut s'expliquer par le fait qu'un mandat a plus ou moins d'effets selon l'institution qui le porte : agences de financement ou institution isolée. Les sept conseils scientifiques britanniques et le *Wellcome Trust*, principaux bailleurs de fonds de la recherche, ont émis un mandat, ce qui a conduit à ce grand nombre d'archives recueillant une grande partie de la science produite dans ce pays. Une autre particularité est la création d'*OpenDepot.org* qui est le dépôt pour les chercheurs britanniques dont l'institution à laquelle ils sont rattachés ne possède pas d'archive institutionnelle.

La France occupe la sixième position (tableaux 1-2, figure 1).

Le contenu des archives est représenté majoritairement par des articles, suivis par des thèses, puis des ouvrages.

¹⁹ <<http://www.peerproject.eu/>> consulté le 2 mars 2012.

²⁰ <<http://project-soap.eu/>> consulté le 2 mars 2012.

²¹ <<http://openaccess.inist.fr/spip.php?article321>> consulté le 2 mars 2012.

²² <<http://www.sherpa.ac.uk/juliet/>> consulté le 2 mars 2012.

²³ <<http://roarmap.eprints.org/>> consulté le 2 mars 2012.

Localisation	Nombre d'agences de financement avec mandat
Royaume-Uni	18
Irlande	4
Espagne	3
Allemagne	2
Europe	2
France	2
Italie	2
Suisse	2
Autriche	1
Belgique	1
Hongrie	1
Norvège	1
Suède	1

JULIET 25-Sep-2011

Tableau 1

Localisation	Nombre d'agences et d'institutions avec mandat
Royaume-Uni	50
Finlande	28
Italie	28
Espagne	13
Allemagne	11
France	11
Portugal	10
Suède	10
Europe	7
Belgique	6
Irlande	6
Norvège	6

ROARPAP 25-Sep-2011

Tableau 2

Proportion of Repositories by Country Europe

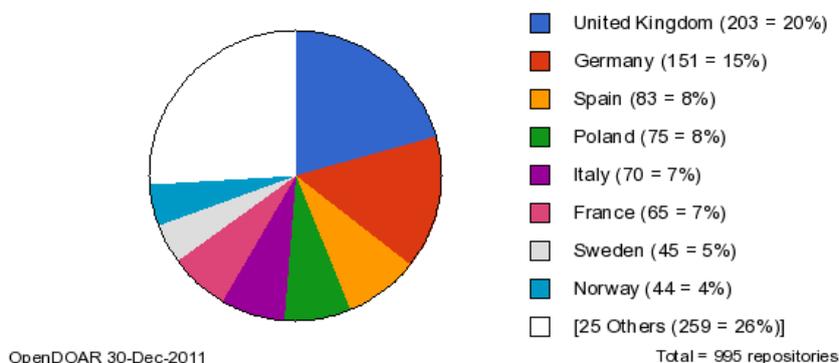


Figure 1

Content Types in OpenDOAR Repositories Europe

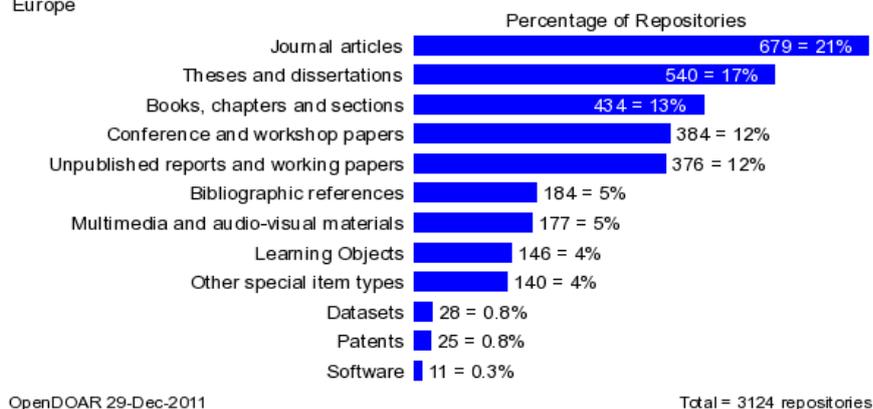


Figure 2

Le DOAJ²⁴ (*Directory of Open Access Journals*) ne répertorie que les revues qui offrent un libre accès sans paiement ni pour le lecteur ni pour l'auteur. Il contient plus de 7 000 titres. Le tableau 3 montre les dix pays européens qui possèdent le plus de titres.

Ces différents chiffres confirment que le Royaume-Uni est le pays le plus avancé sur le plan des politiques et de celui des réalisations.

²⁴ <<http://www.doaj.org>> consulté le 2 mars 2012.

Pays	Nombre d'agences et d'institutions avec mandat
Royaume-Uni	520
Espagne	383
Allemagne	235
Roumanie	207
Italie	187
France	134
Pologne	128
Suisse	101
Croatie	81
Serbie	77

ROARPAP 25-Sep-2011

Tableau 3

3. Quelques problématiques

Même si le mouvement du libre accès a pris de l'ampleur tant par le nombre de mandats, d'archives et de revues, il reste des problèmes à la fois politique, juridique et économique, mais aussi sur le plan de l'évaluation des chercheurs induits par ces nouveaux médias.

Les politiques ne sont pas homogènes tant au niveau des archives que des revues.

La plupart concernent les articles déjà publiés, mais certaines englobent les données issues de la recherche. La période d'embargo est très variable, mais la majorité se situe généralement au-delà de 6 mois. Cependant la question centrale est celle du choix entre incitation ou obligation. La majorité des institutions incitent les chercheurs à déposer leurs articles dans leur archive institutionnelle, car l'obligation pose des difficultés notamment d'ordre juridique.

Les coûts de publication demandés par les éditeurs sont parfois pris en charge par l'institution.

Le droit d'auteur désigne l'ensemble des droits dont jouissent les auteurs sur les productions issues de leurs recherches. Il régit ainsi l'accès, la diffusion, l'usage, la traduction, la réutilisation et la préservation des articles et des données. Généralement, les éditeurs demandent aux chercheurs de leur céder leurs droits patrimoniaux, condition nécessaire pour être publié.

Trois leviers peuvent être actionnés pour supprimer ces contraintes juridiques, l'un n'excluant pas les deux autres.

Le premier est de demander aux chercheurs de ne pas céder tous leurs droits d'exploitation aux éditeurs de manière à pouvoir au moins déposer leur production dans l'archive de leur institution ou dans une archive thématique.

Le deuxième se situe à un niveau national ou fédéral à l'exemple des États-Unis et de l'Espagne. Ainsi, à la suite d'une loi votée en 2009²⁵, les chercheurs des *National Institutes of Health*, dont les recherches sont financées par des fonds publics, doivent déposer leurs articles, publiés dans une revue avec comité de lecture, dans l'archive PubMed Central. L'extension de cette obligation aux autres agences fédérales fait toujours l'objet de vives discussions outre-Atlantique. Une loi similaire²⁶ votée en Espagne en juin 2011, oblige les chercheurs financés par des fonds publics à déposer leurs articles dans une archive.

Le troisième levier est à l'échelon de l'Union européenne (UE). Un pas a déjà été franchi par la Commission européenne avec le projet pilote pour le 7^e PCRD. La Commission se rapproche ainsi des États-Unis en obligeant le dépôt de tout article issu de recherches financées par l'UE. Mais on peut aussi agir sur les droits d'auteur en harmonisant notamment les exceptions pour la recherche et l'enseignement supérieur qui sont différentes selon les pays de l'UE. On pourrait aller plus loin en créant une spécificité pour la science et ne pas traiter un article scientifique à une œuvre musicale. Pourquoi ne pas admettre l'accès à la Science comme un droit fondamental pour tous les citoyens ?

L'aspect économique joue ici un rôle fondamental. Deux mouvements coexistent.

L'un vient des instances nationales, fédérales ou supranationales qui veulent imposer le dépôt de tout résultat de recherches financées par des fonds publics.

L'autre, plus ancien émane des éditeurs. BioMed Central et PLoS ont été les initiateurs de la formule auteur-payeur. Les auteurs ou les institutions leur paient une certaine somme pour que les articles soient accessibles gratuitement à l'ensemble des lecteurs. Cette formule a été reprise et adaptée par de nombreux éditeurs. Ce modèle baptisé hybride mêle l'auteur-payeur et le lecteur-payeur qui correspond au modèle classique de l'abonnement. Les coûts de publications sont hétérogènes d'un éditeur à l'autre.

Le chercheur reste maître dans le choix de la revue dans laquelle il souhaite publier ses résultats. Le facteur d'impact des revues joue un rôle primordial dans ce choix, car il est jugé sur ce critère. Il est évalué au prorata du nombre d'articles parus dans des revues avec un haut facteur d'impact. Les usages évoluent avec le dépôt de textes dans les archives ouvertes, mais également celui de données brutes. D'autres systèmes de mesure se mettent en place comme le *Journal Usage Factor* de COUNTER.

Avoir accès à la littérature scientifique est une condition *sine qua non* pour les chercheurs. Les éditeurs demandent une rémunération en contrepartie du service rendu à la communauté. La publication ayant un coût, un accord doit être passé entre les différentes parties de manière à ce que chacun puisse continuer à exercer son métier et en tirer des bénéfices. Un équilibre doit être également trouvé entre les détenteurs des droits d'auteur et les lecteurs de manière à ce qu'aucun acteur ne soit pénalisé.

²⁵ <<http://grants.nih.gov/grants/guide/notice-files/NOT-OD-09-071.html>> consulté le 2 mars 2012.

²⁶ <http://www.boe.es/diario_boe/txt.php?id=BOE-A-2011-9617> consulté le 2 mars 2012.

Removing the barriers ! Promoting open access to research through advocacy and collaboration

Lars Björnshaugen

Member of SPARC Europe¹ Board

Abstract

The presentation will introduce SPARC and SPARC Europe. SPARC established in 1997 (by ARL) and SPARC Europe (2002) has been active since the early days of the open access discussions. Promoting new scholarly communication models that expand the dissemination of scholarly research, supporting the development of new business models and services, educating and training and not least advocacy and lobbying are the important activities in the agenda. With the increasing take up of the open access concept by researchers, universities, research funders and even governments not to mention publishers it might be the time for closer collaboration between the various (new) organizations engaged in open access.

Résumé

Lever les barrières !

Militer et collaborer pour promouvoir le libre accès à la recherche

Cet exposé présentera SPARC et SPARC Europe. SPARC, créé en 1997 (par l'ARL), et SPARC Europe (en 2002) agissent depuis qu'est apparu le débat sur le Libre accès. Les activités importantes de leur ligne d'action sont : promouvoir de nouveaux modèles dans la communication savante, soutenir le développement de nouveaux modèles économiques et services, sensibiliser et former, et, sans que ce soit la moindre des choses, militer et faire du lobbying. Avec l'émergence croissante de l'idée du libre accès chez les chercheurs, les universités, les financeurs de la recherche et même au sein des gouvernements, pour ne pas citer les éditeurs, le moment pourrait être venu pour les (nouveaux) organismes engagés dans le libre accès de collaborer plus étroitement.

¹ *Scholarly Publishing and Academic Resources Coalition, Coalition des publications savantes et des ressources académiques), une initiative de l'Association of Research Libraries*

Partie 2
Mise en commun des ressources
et données de la recherche

Facilitating open access through international networks: OpenAIRE¹ and COAR²

Birgit Schmidt

Scientific manager OpenAIRE
Göttingen State University Library, Germany

Abstract

All research depends on former work, and depends on researchers' being able to access and share scientific information. These aims are facilitated through international networks promoting open access policies and practices, repository networks and support structures.

In Europe, the European Commission conducts an open access pilot project requiring researchers funded in seven selected areas of the Seventh Framework Programme to deposit their publications in a repository and to make them available in open access. Similarly, the European Research Council stipulates the deposition of publications in repositories through their open access guidelines. OpenAIRE (Open Access Infrastructure for Research in Europe), a three-year project which started in December 2009, supports the implementation of these policies by developing an infrastructure for repository networks, a European-wide helpdesk system and exploring disciplinary practices of research data management. On the European country level, the implementation is supported by national partners (the Couperin consortium in France), which also helps to create synergies between European and local open access initiatives. National partners have a strong role in activating researchers and project coordinators, librarians, research managers, policy makers as well as publishers and journal editors. Further developments are envisioned in the forthcoming OpenAIREplus project by linking publications to scientific datasets, enriching the information space and establishing connections with other infrastructures.

This links to the international approach of the Confederation of Open Access Repositories (COAR), which promotes the interoperability of open access repositories to enhance the visibility of research outputs and to strengthen the international open access implementation. Currently, COAR coordinates the collaborative efforts of more than 60 members and partners world-wide. With its actions COAR aims to support regional and national repository initiatives, to disseminate best practices and to facilitate the discussion on interoperability. COAR's three working groups address the acquisition of content for repositories, interoperability, and training and support actions. In particular, WG1 explores best practices and expertise in the area of populating repositories, WG2 will soon release a briefing paper on

¹ Open Access Infrastructure for Research in Europe.

² Confederation of Open Access Repositories.

the topic with the aim to develop interoperability guidelines. Moreover, WG3 develops a mentoring service and workshops addressing the needs of members and partners.

Further information: <<http://www.openaire.eu>> and <<http://www.coar-repositories.org>>.

Abstract

Comment faciliter le libre accès grâce aux réseaux internationaux : les projets OpenAIRE et COAR

Toute recherche dépend des travaux précédents et de la capacité des chercheurs à accéder à l'information scientifique et à la partager. Ces objectifs sont facilités grâce à des réseaux internationaux qui encouragent les politiques et les pratiques de libre accès, les réseaux d'archives et les structures de soutien.

En Europe, la Commission européenne est à la tête d'un projet pilote qui requiert que les chercheurs dans sept des domaines sélectionnés par le 7^e programme-cadre déposent leurs publications dans une archive pour les mettre à disposition en libre accès. De même, le Conseil européen de la recherche préconise dans ses directives sur le libre accès le dépôt de publications dans des archives. OpenAIRE (Open Access Infrastructure for Research in Europe, ou infrastructure en libre accès pour la recherche en Europe), un projet de trois ans lancé en décembre 2009, porte la mise en œuvre de cette politique en réalisant une infrastructure composée de réseaux d'archives, d'un service d'assistance à l'échelle européenne ainsi que l'analyse des pratiques par discipline en matière de gestion des données de la recherche. Au niveau de chaque pays européen, cette mise en œuvre est soutenue par des partenaires nationaux (le consortium Couperin en France), ce qui permet aussi de créer des synergies entre des projets européens et locaux dans le domaine du libre accès. Les partenaires nationaux jouent un rôle de poids pour mobiliser chercheurs, coordinateurs de projets, professionnels des bibliothèques, administrateurs de recherches, décideurs politiques, tout comme des éditeurs et des rédacteurs-en-chef de revues. D'autres développements sont envisagés dans le prochain projet OpenAIREplus où seront reliés publications et ensembles de données scientifiques, en enrichissant l'espace d'information et en établissant des liens avec d'autres infrastructures.

Ceci est à rapprocher de la démarche de la Confédération des archives en libre accès (Confederation of Open Access Repositories - COAR), qui encourage l'interopérabilité des archives en libre accès pour améliorer la visibilité des résultats de la recherche et pour renforcer la mise en œuvre du libre accès sur le plan international. Actuellement, COAR coordonne les travaux collaboratifs d'une soixantaine de membres et partenaires dans le monde. Par son action, COAR a pour objectif de soutenir des projets d'archives nationales et régionales, pour diffuser des bonnes pratiques et pour faciliter la discussion sur l'interopérabilité. Les trois groupes de travail de COAR traitent des questions d'acquisition de contenu pour les archives, d'interopérabilité et des actions de soutien. Plus particulièrement, le premier groupe identifie et analyse les bonnes pratiques et l'expertise dans le domaine de l'alimentation des archives, le deuxième groupe va bientôt publier un document d'information sur ce sujet dans le but de définir des préconisations en matière d'interopérabilité. Enfin, le troisième groupe prépare un service de tutorat et des ateliers destinés à répondre aux besoins des membres et des partenaires.

Pour en savoir plus : <<http://www.openaire.eu>> et <<http://www.coar-repositories.org>>.

PEER, *Publishing and the Ecology of European Research*, retour d'expérience d'INRIA sur un projet européen associant éditeurs et archives ouvertes

Foudil Brétel¹ et Alain Monteil²

¹ Ingénieur DSI

INRIA Lyon la Doua, Bat. CEI-1, BP 52132, 66 bd. Neils Bohr
69100 Villeurbanne Cedex, France

² Ingénieur IST

INRIA Sophia Antipolis Méditerranée, BP 93, 2004 route des Lucioles
06902 Sophia Antipolis cedex, France

Résumé

L'INRIA (Institut national de recherche en informatique et automatique) participe au projet européen PEER, *Publishing and the Ecology of European Research*, qui regroupe éditeurs et représentants de la communauté scientifique. Nous présenterons ce projet qui a pour objectifs d'étudier l'appropriation et l'utilisation des Archives Ouvertes par les scientifiques, et de poser les fondements d'un nouveau système économique respectant à la fois les règles du marché de l'édition et le principe du libre accès à la connaissance. Puis nous détaillerons les avancées techniques développées autour de l'application servant d'interface entre les éditeurs et les archives ouvertes : le *PEER Depot*. Enfin nous évoquerons le bilan pour l'INRIA d'une participation à ce type de projet.

1. Introduction

Le monde éditorial suit de près les évolutions du paysage de la diffusion de la production scientifique, et notamment celui de la communication scientifique directe par les chercheurs vers les chercheurs. Plusieurs rapports et études ont été menés sur les serveurs de publications et la mise en ligne des documents, mais plus du point de vue des institutions ou communautés scientifiques que de celui de l'édition traditionnelle.

STM (*International Association of Scientific, Technical & Medical Publishers*) qui est l'association des plus grands éditeurs scientifiques a souhaité mener ce type d'étude afin d'avoir une vision claire de la situation. Le programme *eContent+* de l'Union européenne a permis de donner un cadre à une étude de cette nature.

2. Le projet PEER

2.1 L'avant-projet

Lors de la phase préliminaire au projet, STM s'est rapproché de la MPDL (*Max Planck Digital Library*) comme interlocuteur car la Max Planck est à la fois un institut de recherche pluridisciplinaire et un acteur majeur dans le libre accès. Laurent Romary qui était à l'époque directeur de la MPDL a aussitôt mentionné le cas français et souligné l'opportunité de faire participer une archive comme HAL (Hyper article en ligne, Centre pour la communication scientifique directe) puisque très bien positionnée dans le paysage scientifique hexagonale, et plus particulièrement l'INRIA comme partenaire technique.

STM a donc regroupé autour du projet une agence de financement, ESF (*European Science Foundation*), trois instituts de recherche UGOE (*Max Planck Digital Library*), MPG (*Max-Planck-Gesellschaft*) et INRIA afin de répondre à l'appel à projet *eContent+* dans le cadre du programme cadre FP6 de l'Union européenne. Ce programme a pour objectif de favoriser et aider des initiatives de collaboration entre le monde du privé et les instituts de recherches publiques et mettre à disposition des citoyens du contenu scientifique. STM a fait appel à une consultante indépendante, durant six mois, pour l'animation et la récolte des informations nécessaires à la rédaction de la réponse de l'appel à projet.

Les compétences de cette assistance ont permis de compenser l'inexpérience des personnels IST en matière de projet européen. Nous pensons que c'est un des facteurs de réussite qui a permis au projet PEER de proposer un dossier solide et de qualité, et d'être retenu et financé par l'Union européenne.

2.2 Principes généraux

Le projet PEER propose la mise en place d'un observatoire, détaillé ci-après, à disposition des groupes de recherche afin d'étudier l'impact du libre accès dans l'écologie de l'édition scientifique européenne. Cet observatoire, après discussions, a été conçu comme un espace d'étude qui ne devait pas être « pollué » par des éléments extérieurs afin de refléter la réalité au plus près. C'est le premier principe de fonctionnement. Pour l'illustrer il a été décidé qu'aucune action de type incitation ou publicité ne serait faite par les archives participantes ou des documentalistes auprès des auteurs afin qu'ils déposent leurs articles. Pour cette raison, mais aussi pour des besoins techniques, chaque archive a bâti une archive parallèle et indépendante, telle l'instance PEER de HAL. Cette indépendance garantit des influences possibles vis-à-vis de l'archive institutionnelle.

Second principe général : celui de la définition des articles éligibles au projet PEER. STM a établi une liste de 240 titres de journaux représentant un échantillon relativement complet des différentes disciplines scientifiques, intégrant la notoriété par leur facteur d'impact et leur ancienneté. À partir de ce corpus, tous les articles dont l'auteur principal est résident d'un pays européen est éligible. Puis le volume est divisé en deux, une moitié sera diffusée par les éditeurs, l'autre par les auteurs.

Le dernier point fondamental du projet PEER concerne la définition du document devant être déposé. Il ne s'agissait pas de proposer les PDF des éditeurs, mais la version soumise et acceptée par le comité éditorial de la revue avant la mise en forme éditoriale. Ces

documents sont appelés « *Stage 2* » dans le cadre du projet PEER. Le « *Stage 1* » est la version auteur avant soumission et pour le « *Stage 3* » il s'agit de la version finale publiée par les éditeurs.

2.3 L'observatoire

L'observatoire est donc l'ensemble des processus et applications permettant au corpus d'articles scientifiques, rédigés par des scientifiques européens, d'être diffusé au travers des archives ouvertes. Le *PEER Depot*, réalisation principale de l'INRIA dans le projet, est au cœur du système et fait l'objet des avancées techniques les plus importantes. Elles sont détaillées dans la deuxième partie de cette contribution. Comme indiqué figure 1: l'alimentation devait se partager à égalité entre les éditeurs et les auteurs. Cette double alimentation doit être mesurable et observable au travers de logs afin d'étudier les usages concernant le dépôt dans cette archive ouverte.

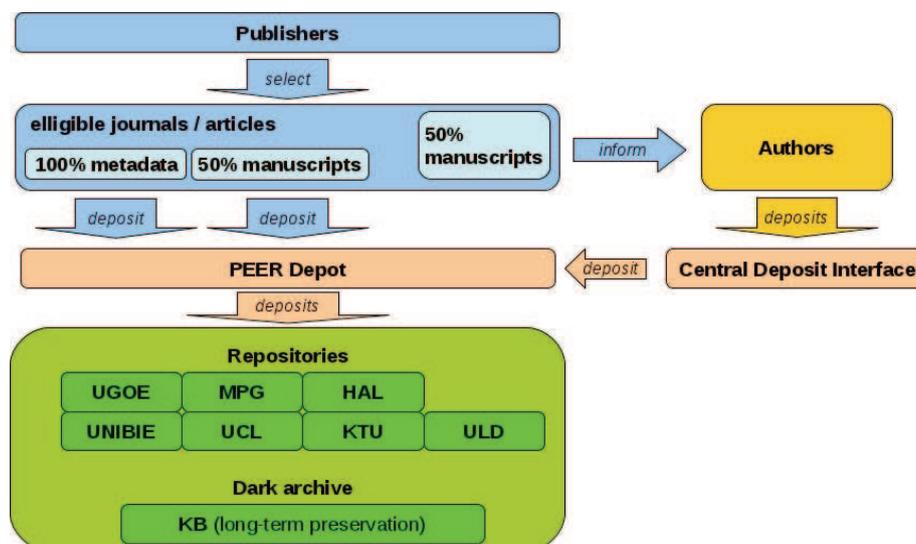


Figure 1 : Schéma des flux

Afin d'avoir matière à étudier, une masse critique de 10 000 références a été fixée par l'observatoire PEER. Au-delà de la volumétrie, chaque archive a également eu à mettre en place les éléments de recueil statistique, sans avoir de réel cahier des charges car les utilisateurs de ces données étaient inconnus.

L'atteinte de la masse critique n'a pas été aussi simple qu'il n'y paraît. Le temps de mise en place a été plus long que prévu, ce qui a entraîné un certain retard du projet. Pour ouvrir l'observatoire aux groupes de recherche, la masse critique devait être atteinte lors de la deuxième année. Cumulé au retard lié à la mise en place de l'observatoire, le dépôt direct par les auteurs n'a malheureusement pas été à la hauteur de nos espérances. Plus de 6 000 invitations aux dépôts ont été envoyées aux auteurs et seulement 147 ont répondu

favorablement par un dépôt d'article sur l'interface dédiée au projet PEER, gérée par la *Max Planck Digital Library*.

Face à ce constat le projet PEER adopta des mesures pour atteindre l'objectif fixé. Première mesure : l'extension de 40 titres de journaux passant par le canal de dépôt éditeurs, les embargos ont été réduits accélérant la diffusion des articles, enfin ceux plus anciens ont été intégrés aux projets par certains éditeurs. Grâce à ces mesures le seuil a été atteint en mars 2011.

Concomitamment une demande de prolongation du projet a été demandée à l'Union européenne afin de donner suffisamment de temps aux groupes de recherche pour mener à bien leurs études. Prolongation accordée pour neuf mois supplémentaires portant à juin 2012 la fin du projet.

2.4 Objectifs

L'observatoire PEER doit permettre de déterminer les impacts du dépôt à grande échelle d'articles en libre accès. Son fonctionnement s'appuie sur trois groupes de recherche. Un premier prend en charge la thématique de l'usage des articles scientifiques tant côté plateforme éditeur que côté archives participantes par l'étude des logs et l'appréciation de la diffusion et l'accès *a priori* plus grands des résultats scientifiques. Le second groupe de recherche ; a pour objectif de suivre les tendances, expliquer les modèles éditoriaux des auteurs et le comportement des utilisateurs dans le contexte du libre accès, ainsi qu'e les impacts possibles sur la conduite et l'exercice de la recherche elle-même. Enfin, le dernier groupe étudie les aspects économiques avec pour objectif d'analyser les effets sur l'économie de la communication savante du dépôt de grande ampleur, d'enquêter sur le coût de celui à grande échelle, y compris du point de vue de l'efficacité économique ou sur le coût du processus de dépôt.

L'observatoire PEER permet également : d'étudier les facteurs influents sur le dépôt en libre accès des documents scientifiques, les facteurs favorisant mais aussi les freins ; d'étudier et produire des modèles éditoriaux illustrant la possible coexistence entre l'édition traditionnelle et l'auto-archivage en libre accès.

Le projet PEER et la mise en place de l'observatoire participent de la reprise d'une relation de confiance et de compréhension mutuelle entre les éditeurs et le monde académique soutenant les archives ouvertes.

Aujourd'hui plus de 16 000 articles sont disponibles grâce au projet PEER, et sont diffusés par les archives participantes, cependant seulement 0,2 % proviennent du dépôt des auteurs après invitation des éditeurs.

3. Brève description technique du *PEER Depot*

Le *PEER Depot* est l'intégration, sur un serveur *Linux RHEL 5*, d'une collection de *librairies* et *scripts Perl*, associés à un serveur FTP¹ *Pure-FTPd*. Les statistiques en vue de l'administration de la plateforme, ou à destination de l'équipe de recherche sur les usages, sont recueillies dans une base de données *PostgreSQL*. Le traitement est déclenché directement par le serveur FTP à la réception de nouveaux articles. Les dernières tâches (complétion des métadonnées, gestion de l'embargo, distribution vers les archives) sont automatisées par *cron*.

La plateforme conçue et développée à l'INRIA est constituée de logiciels libres. Les spécifications de départ indiquaient une moyenne de 50 articles à traiter par jour. Mais le serveur a régulièrement été amené à en traiter plusieurs milliers. La conception a suivi le principe KISS (*keep it simple and stupid*). Grâce à cette simplicité, on peut imaginer d'agrèger plusieurs instances du *PEER Depot* en grappe, afin de supporter une plus forte montée en charge (*scalability*).

Une précision quant à la fonction du *PEER Depot* : nommé « *depot* » historiquement, il doit pourtant être considéré comme un *hub*. C'est-à-dire un relai de distribution. Les fonctions de stockage sont minimales. L'accent a été principalement mis sur le traitement et la distribution des articles.

4. Élaborer et gérer les flux : difficultés et solutions

4.1 Avancées organisationnelles

Un des grands accomplissements du projet, qui a significativement contribué à son succès, est l'établissement de conventions. Celles-ci ont nécessité un effort considérable de la part des participants, les monopolisant fortement durant les six premiers mois du projet, en générant de nombreuses réunions (partenaires PEER et éditeurs notamment). Ces conventions ont porté essentiellement sur le *workflow* de dépôt des éditeurs, et la fourniture des métadonnées [WP3.1]. Ce travail essentiel pour le *workpackage 7* est dédié à l'élaboration de modèles pour la cohabitation de l'édition traditionnelle avec l'auto-archivage.

Lors de l'élaboration de ces conventions, l'équipe de recherche en charge des études des usages n'était pas encore constituée. Les partenaires impliqués dans le *workpackage 2* ont donc dû anticiper les futurs besoins (aussi bien pour les données de statistique que pour la fourniture des logs par les archives). Par la suite, la nécessité d'une collaboration plus rapprochée entre l'équipe de recherche et celle du *PEER Depot* ne s'est pas fait ressentir. On peut donc penser que les besoins ont été suffisamment bien estimés et couverts.

¹ *File Transfer Protocol*.

4.1.1 Le dépôt éditeur

Si l'on considère le nombre d'organisations participantes (12 éditeurs, 6 archives), et leur diversité, en termes de contenu et de fonctionnement, l'échelle du projet était déjà une gageure, pas seulement pour l'établissement de conventions, mais aussi pour le traitement des spécificités de chacune. Ainsi au départ, il a été envisagé une période d'essai pour le dépôt des éditeurs à l'issue de laquelle ils devaient tous entrer simultanément en phase de production. Le délai nécessaire à l'adaptation technique (dénommée « validation » au sein du projet), aussi bien du *PEER Depot* que des éditeurs, n'a pas permis de respecter cette date commune. Aussi, le projet a dû se résoudre à traiter les cas un-à-un, avec une entrée de chaque éditeur séquentiellement.

DublinCore-like name	Comment
Title*	Article Title
Creator*	Corresponding Author's name: Last Name, First Name
AuthorEmail	Corresponding Author's email address
Description	Abstract
Date*	Date of Publication
Identifier*	DOI or PublisherArticleId ²
Coverage	Geographic location of the Contributing Author: ISO 3166
Journal	Journal Title
Affiliation	Multi-tier organisation list: Country, Organization, Laboratory
ISSN	These elements are not mandatory to electronic publication, and can be derived from CrossRef after DOI is provided, and may therefore not be provided by publishers. Possible use of CrossRef for DOI resolution
Volume	
Issue	
Page	
Type*	Default value = article. Mapped to <i>info:eu-repo/semantics/article</i> , <i>info:eu-repo/semantics/acceptedVersion</i>
Subject	Subject headings; Scientific classification (defaults to what is provided in the general STM Journal table)
Language	ISO 639-3 (defaults to "eng")
Embargo	Embargo Period (defaults to what is provided in the general STM Journal table)

Tableau 1 : Métadonnées requises

² Identifiant article propre à l'éditeur.

4.1.2 Les métadonnées

La liste des métadonnées nécessaires, qui pouvaient être effectivement fournies par les éditeurs ou d'autres sources, a été établie par recensement auprès des archives participantes. Globalement, le nombre de métadonnées nécessaires était peu élevé. Le projet aurait pu se limiter aux recommandations de DRIVER³, c'est-à-dire à 4 métadonnées issues du *Dublin Core*. Cependant, HAL et la *Koninklijke Bibliotheek*⁴ en requéraient une douzaine. Il a donc été convenu d'en distinguer trois catégories : obligatoires, requises, optionnelles. Dans la pratique, tous les éditeurs à l'exception d'un ont été en mesure de fournir les métadonnées obligatoires et requises, les optionnelles quant à elles pouvaient être obtenues par d'autres sources (CrossRef ou fichiers).

Le tableau 1 est extrait du rapport D3.1[1] détaille les métadonnées utilisées par le projet.

4.1.3 Les standards

Se conformant aux recommandations de DRIVER, le projet a utilisé au maximum les standards (normes ISO notamment). Le DOI⁵ y tient aussi une place particulière car les éditeurs avaient à cœur qu'il soit utilisé. En toute rigueur, cet identifiant renvoie au document publié, même si le projet traite du *stage 2*. Le DOI a cependant été retenu comme composante de l'identifiant des articles, sous la forme : PEER_stage2_[URL_encoded_DOI]. Le DOI est encodé pour des raisons pratiques de noms de fichiers qui ne peuvent pas contenir le caractère « / » (barre oblique) sous Linux. Ce caractère a posé problème à au moins un serveur Web d'archive pour lequel une option de sécurité précisait que les requêtes HTTP⁶ ne peuvent pas contenir de barre oblique. L'archive a décidé de renommer les fichiers en transformant « / » en « _ ». Ce qui n'a pas particulièrement perturbé l'équipe de recherche en charge des études des usages, qui fonde ses recherches sur les URL enregistrées dans les fichiers logs des serveurs Web des archives.

4.1.4 Le filtrage

Hormis la distribution des articles proprement dite, le rôle du *PEER Depot* devait recouvrir également des fonctions de filtrage selon plusieurs critères : période d'embargo, type d'article, contenu européen, éventuellement thème⁷.

La gestion de l'embargo était une préoccupation importante pour les éditeurs. Le calcul de celui-ci est identique dans les deux cas de flux : dépôt auteur et éditeur. Dans la mesure où les éditeurs fournissaient une date de publication (papier ou en ligne), et où le *PEER Depot* disposait des durées d'embargo par journal, leur gestion n'a pas posé problème. Voir ci-dessous *Interface de dépôt auteur* pour plus de détails.

Ne pouvant les prendre à leur charge, beaucoup d'éditeurs ont souhaité que des filtrages de contenu soient effectués sur le *PEER Depot* : par type d'article ou par journal. Ces informations font partie des métadonnées requises par PEER. La principale difficulté a été

³ *Digital Repository Infrastructure Vision for European Research*.

⁴ Bibliothèque royale des Pays-Bas.

⁵ *Digital Object Identifier*.

⁶ *Hypertext Transfer Protocol*.

⁷ Cas de l'archive thématique *Social Science Open Access Repository* (SSOAR).

d'adapter le filtrage, tout au long du projet, aux spécificités variables des éditeurs (qui n'utilisent pas toujours de manière orthodoxe les champs dédiés).

Pour le filtrage du contenu européen, il a été convenu qu'il serait basé sur le pays de l'auteur correspondant. Cette donnée était issue des affiliations (métadonnées requises). Les noms de pays et codes divers étant transformés en ISO-3611-A2. La possibilité d'inclure le pays de l'auteur principal dans la sélection ainsi que celui des autres auteurs a été évoquée en cours de projet. Elle n'a pas été retenue car l'origine de chacun n'est pas systématiquement fournie par les éditeurs.

Le filtrage par thème a la particularité d'être effectué en sortie du *PEER Depot*, tandis que tous les autres filtrages s'effectuent à l'entrée. En effet, il ne concerne pas toutes les archives. Le critère est dans ce cas l'ISSN⁸. Le filtrage en sortie étant assez modulaire de conception, d'autres solutions et critères auraient pu être envisagés aisément.

4.1.5 La distribution vers les archives

4.1.5.1 Le format d'échange

Concernant l'interfaçage entre le *PEER Depot* et les archives, la principale question a été celle du format de livraison des articles, et particulièrement le choix du format XML⁹. Les formats en lice étaient NLM¹⁰, format adopté par 6 des 12 éditeurs, et TEI¹¹, standard parfaitement maîtrisé par l'équipe INRIA, et largement utilisé dans l'édition scientifique et le monde académique. Le choix s'est arrêté sur la TEI. Un schéma *RelaxNG* est en cours d'élaboration afin de décrire formellement le format TEI utilisé dans PEER.

Ce choix de la TEI comme format pivot unique s'inscrit dans une démarche de standardisation. En effet, il s'agit d'une initiative internationale de référence pour les sciences humaines numériques, utilisée par les principaux acteurs institutionnels en édition électronique française (Persée, Revues.org, Presses universitaires de Caen). Elle propose une plateforme unifiée de description de documents textuels (manuscrits, livres, dictionnaires... et articles), ce qui permet d'éviter l'effet de silo et la fragmentation des formats créés par la NLM. Le *framework* est donc modulaire, extensible et propose plus de 500 éléments, y compris des extensions telles que SVG¹², MathML CALS, etc. Elle s'accompagne d'outils précieux, tels qu'un langage de spécification (ODD¹³) permettant d'adapter facilement un schéma à ses besoins et de le générer automatiquement (DTD¹⁴, XML ou RelaxNG¹⁵). Enfin, elle est gérée de manière consensuelle et ouverte, par une communauté très active¹⁶.

⁸ *International Standard Serial Number.*

⁹ *Extensible Markup Language.*

¹⁰ *National Library of Medicine.*

¹¹ *Text Encoding Initiative.*

¹² *Scalable Vector Graphics.*

¹³ *One Document Does it all.*

¹⁴ *Document Type Definition.*

¹⁵ *REgular LAnguage for XML Next Generation.*

¹⁶ Pour une réflexion sur les formats XML, voir [2].

4.1.5.2 Le protocole de dépôt

Le projet a choisi d'utiliser le protocole SWORD¹⁷ pour la distribution des articles du *PEER Depot* vers les archives. Toutes ont eu à implémenter le support de SWORD. Dspace¹⁸ incluant nativement ce support, les archives qui utilisaient ce logiciel n'ont eu qu'à adapter la transformation (XSLT¹⁹) de TEI vers leur propre format de métadonnées. L'INRIA ayant fourni un modèle générique. SWORD avait déjà été adopté par quelques rares archives (ArXiv par exemple), mais l'utilisation d'un protocole standard unique était une nouveauté. D'ailleurs SWORD étant assez souple (il continue d'évoluer), cependant l'implémentation n'a pas été comprise de la même manière par tous. L'implémentation retenue dans PEER ne convenait pas exactement aux archives pour lesquelles le dépôt s'effectuait en plusieurs étapes (*via* une validation humaine par exemple). L'avantage principal de SWORD pour le projet est le mécanisme d'acquittement qu'il inclut : l'archive transmet directement dans sa réponse l'identifiant du document déposé.

4.1.6 Interface de dépôt auteur

La spécification initiale du dépôt auteur n'avait pas prévu que, hormis HAL, les archives étant des archives institutionnelles, elles n'autoriseraient le dépôt qu'aux utilisateurs de leur propre institution. Il a donc été nécessaire de concevoir une interface de dépôt auteur générique. Hébergée à la *Max Planck Digital Library* (MPDL), elle transmet ces dépôts au *PEER Depot*, qui les distribue aux archives à la fin de la période d'embargo (de 0 à 36 mois). Cette date de distribution est calculée selon la formule : date de publication + période d'embargo = date de distribution. La date de publication est issue des métadonnées fournies par les éditeurs. La durée d'embargo définie par journal est issue de la liste des journaux intégrée au *PEER Depot*. Les éditeurs fournissent les métadonnées pour les journaux à destination du flux « éditeur », mais également celles à destination du flux « auteur ». Le *PEER Depot* se charge d'établir la corrélation, à partir d'un petit ensemble de métadonnées fournies par l'auteur. Dans la pratique, l'appariement s'effectue sur le titre du journal et celui de l'article uniquement²⁰.

4.2 Avancées techniques

4.2.1 Le développement

La conception et les spécifications du *workflow* de dépôt des éditeurs ayant requis plus de temps que prévu, le développement lui s'est réalisé dans des délais très courts, et a représenté environ 3 mois.homme pour le premier prototype. Ce qui comprend : l'écriture des XSLT, du code *Perl*, l'intégration de la plateforme, et la configuration système. Les besoins

¹⁷ "a JISC-funded initiative to define and develop a standard mechanism for depositing into repositories". SWORD est un profile du protocole *Atom Publishing Protocol* (APP). Cf. <<http://swordapp.org/>> consulté le 22 mai 2012.

¹⁸ Logiciel libre utile à la construction d'archives électroniques ouvertes.

¹⁹ *Extensible Stylesheet Language Transformations*.

²⁰ Voir le projet similaire *Open Access-Repository Junction* : <<http://oarepojunction.wordpress.com/about-the-broker/>> consulté le 22 mai 2012.

étant très spécifiques, pratiquement aucun code existant n'a pu être réutilisé, hormis quelques *librairies* fondamentales telles que XML::LibXML. Une tentative d'associer d'autres partenaires du projet au développement n'a malheureusement pas abouti, principalement pour des raisons d'usages techniques différents.

4.2.2 Fusion XML

Pour chaque article les métadonnées pouvaient être fournies en deux passes. Le seul vrai défi technique a été la fusion XML des métadonnées. Elle survient à la seconde passe. Par convention, l'opération se limite à la complétion des métadonnées, ce qui simplifie le problème. C'est-à-dire qu'il n'y a pas d'écrasement : seules les données dont on ne dispose pas déjà sont prises en compte. Cependant, comment comparer des nœuds pour lesquels l'information discriminante se situe plusieurs niveaux en dessous ? L'exemple ci-dessous illustre ce problème :

```
<author>
  <persName>
    <forename type="first">Foudil</forename>
    <surname>Brétel</surname>
  </persName>
  <email>foudil.bretel@INRIA.fr</email>
  <affiliation>
    <orgName type="department">Direction des Systèmes d'Information</orgName>
    <orgName type="institution">INRIA</orgName>
    <address>
      <postCode>75001</postCode>
      <settlement>Paris</settlement>
      <country key="FR">France</country>
    </address>
  </affiliation>
</author>
```

Dans cet exemple, lorsqu'on considère le nœud *author*, il faut prendre en compte ses descendants afin de pouvoir l'identifier (en vue de le comparer et de l'ajouter éventuellement). Sans entrer dans les détails, nous avons adopté un algorithme original, qui compare les descendants sur au plus deux niveaux. Celui-ci a donné des résultats satisfaisants.

4.2.3 Tests XML

Nous n'avons pas pu implémenter suffisamment de tests pour couvrir toute l'application – notamment en raison de la difficulté à automatiser les tests d'un *workflow* complet avec un espace de stockage de test spécifique. En revanche, il est devenu urgent de se prémunir d'effets de bords indésirables lors de modifications des transformations XSLT. Puisqu'il n'existait pas vraiment d'outils simples dédiés, nous avons élaboré le nôtre. Nous disposons d'un jeu d'environ 150 tests. Notre réalisation a montré des performances supérieures aux dispositifs similaires.

4.2.4 Conversion des fichiers source

En première analyse, et sur demande des éditeurs, notre équipe avait proposé la possibilité de convertir, sur le *PEER Depot*, lorsque le fichier PDF n'était pas disponible, les fichiers sources des articles, du format MS Word ou LaTeX vers PDF. En dernier recours, cette technologie aurait pu être empruntée au CCSD²¹. Mais les échantillons ont montré que cela n'était pas réalisable du fait de la diversité des sources : le fichier MS Word ne contenait pas toujours les figures, qui pouvaient être placées dans des sous-répertoires variés. La conversion n'était donc pas automatisable, en tout cas pas sans une convention préalable. Le *PEER Depot* a bien reçu des fichiers sources sans PDF, et dans ce cas, les articles ont été considérés comme perdus pour le projet.

4.2.5 Confidentialité

La question de la confidentialité des fichiers de log fournis par les archives à l'équipe de recherche sur les usages s'est posée tôt, et de manière plus insistante pour HAL et la MPDL, qui obéissent à des législations plus restrictives dans ce domaine (comparées par exemple à celle du Royaume-Uni). La solution a été laissée à la discrétion de chaque archive, tant que les origines des requêtes HTTP pouvaient être identifiées. Ainsi pour HAL, les logs sont traités avant livraison afin de transformer l'adresse *IP* et le champ *Referrer* en des identifiants uniques (SHA1).

4.2.6 Reporting

Le *reporting* sur l'activité du *PEER Depot* a aussi été envisagé très tôt. Hormis l'équipe de recherche sur les usages, d'autres instances du projet devaient pouvoir accéder aux données enregistrées par le *PEER Depot*. Le choix de l'outil de stockage s'est donc naturellement porté sur une base de donnée relationnelle (*PostgreSQL*), dont l'accès a été ouvert aux personnes intéressées. Il n'a pas été nécessaire de développer une interface spécifique pour la consultation de l'activité : la structure de la base de données restant simple, les outils de consultation de bases SQL (tel *pgAdmin*) ont suffi. En revanche, l'équipe en charge du monitoring a élaboré un mécanisme de génération automatique de rapports exhaustifs.

4.2.7 GROBIB

Le projet a été une nouvelle occasion d'expérimenter GROBID, l'automate d'extraction de données bibliographiques à partir de documents PDF, développé en collaboration avec

²¹ Centre pour la communication scientifique directe.

l'INRIA [3]. Le besoin est apparu avec un éditeur qui ne pouvait pas facilement fournir les affiliations. La solution adoptée a été d'extraire l'ensemble des métadonnées des articles en PDF, et de les compléter éventuellement par d'autres sources ; celles fournies par l'éditeur étant ignorées. Cette solution a nécessité des ajustements, en termes de format (se conformer à la TEI telle qu'utilisée dans PEER) et de contenu.

Les résultats sont encourageants (environ 2 200 articles traités, dont 950 retenus), et la solution, dans le cadre de PEER, pourrait être encore améliorée : surtout en automatisant le traitement, qui souffre d'être actuellement semi manuel, avec un *workflow* spécifique parallèle, sur un mode « essai-erreur », et donc enclin aux erreurs.

4.3 Contingences

Des difficultés sont survenues en cours de projet. En partie parce qu'elles n'avaient pas été abordées dès la conception, ou pas suffisamment développées, en partie parce que l'état de l'art ne permettait pas de les pallier.

4.3.1 Constat sur qualité des données

Une première surprise a concerné la qualité des données. Celles dont disposent les éditeurs proviennent souvent directement des auteurs, qui n'ont pas les moyens de les consolider, comme en les extrayant de référentiels. On pense notamment aux affiliations, mais on pourrait aussi évoquer les codes pays... Une des conséquences est qu'une partie non négligeable des articles restent incomplets en termes de métadonnées, et donc en attente sur le *PEER Depot*.

En termes techniques, les archives, qui étaient associées à l'élaboration du *workflow*, ont été assez autonomes. Cependant, il est intéressant de constater que toutes ont déployé une archive distincte dédiée à PEER. Les raisons sont d'ordre éditorial pour les archives de type institutionnel (PEER délivre des articles sans distinction d'affiliation d'auteur, or les archives souhaitent recevoir les articles de leur organisation), et d'ordre technique pour la seule archive de type national (les contraintes sur les métadonnées ne pouvant pas être satisfaites). On peut déplorer que la communauté de l'IST n'ait pas encore réussi à s'organiser afin de proposer un référentiel unique des affiliations.

4.3.2 Visibilité des articles déposés

Les équipes de recherche du projet (usages et comportements principalement) étaient préoccupées par la visibilité des articles déposés dans ces instances d'archives ouvertes dédiées à PEER. Cette visibilité était souhaitée identique à celle des instances principales. Les équipes de recherche prévoyaient notamment que les chercheurs-lecteurs parviendraient à ces instances par l'intermédiaire des moteurs de recherche (tel *GoogleScholar* notamment). Or la visibilité dans ces moteurs n'est pas maîtrisable : on ne peut que déclarer son site au moyen d'un *sitemap*²², puis attendre que Google indexe les notices. Il est à noter que *GoogleScholar* ne supporte plus OAI-PMH pour l'indexation, mais *OpenURL*. Cela a exclu une archive candidate²³ du projet : par manque de ressources pour l'implémentation.

²² <<http://www.sitemaps.org/>> consulté le 22 mai 2012.

²³ *Kaunas University of Technology*.

Globalement, une certaine visibilité a été atteinte, mais avec des résultats disparates d'une archive à l'autre.

4.3.3 Quantité de contenu déposé

À un certain stade du projet, l'équipe de recherche sur les usages a considéré que la masse critique d'articles déposés nécessaire à la recherche ne serait pas atteinte en temps voulu. Le projet a donc été prolongé de 9 mois. Et la quantité de contenu déposé dans les archives est devenue une priorité. Le projet a pu atteindre son nouvel objectif (10 000 articles déposés fin février 2011, soit un an après les premiers dépôts), en prenant des mesures dont certaines modifiaient sensiblement les conditions initiales, dont :

- passage de certains journaux du flux auteur vers le flux éditeurs ;
- inclusion de certains types d'articles ;
- réduction d'embargos pour certains journaux ;
- inclusion de collections plus anciennes.

4.3.4 Retirer des articles de PEER

La possibilité d'un retrait d'article à la demande d'un auteur avait été envisagée assez tôt, et les archives s'étaient engagées à permettre une forme de retrait. Le détail de ce processus sur l'ensemble du *workflow* n'a pas été spécifié et n'a été implémenté que lorsque le besoin s'est déclaré. Or celui-ci est venu des éditeurs, et sur des collections entières (ce qui a représenté parfois des centaines d'articles). Les principales raisons ont été : la présence de commentaires des relecteurs ou des auteurs, et du contenu manquant. La difficulté était de garder une cohérence sur l'ensemble du *workflow*. La modalité retenue a été la suivante : demande de retrait d'un auteur ou d'un éditeur auprès des interlocuteurs désignés sous forme d'une liste de DOI ; le retrait est effectué sur le *PEER Depot* qui en garde la mémoire (et donc de toutes les informations associées dont les dépôts correspondants vers les archives) ; le *PEER Depot* informe les archives du retrait au moyen d'une liste de DOI avec les identifiants propres à chaque archive ; les interlocuteurs contrôlent le retrait effectif des archives.

5. Bilan et perspectives

5.1 En termes métier

Pour le réseau IST et pour INRIA la participation au projet PEER fut une première dont les principaux enseignements peuvent être tirés.

En interne, en termes de support administratif nous avons rencontré des difficultés d'accompagnement. En effet la légitimité pour les services administratifs d'accompagner une ligne métier et non des scientifiques pour un projet européen n'était pas évidente. Le manque d'effectifs et le nombre important de projets européens obtenus par INRIA nous a parfois fait passer à un niveau inférieur de priorité. Nous avons cependant pu bénéficier de ce soutien indispensable pour la bonne marche d'un tel projet.

L'expérience du réseau IST en matière de gestion de projets européens était très faible, et principalement réduite à celles du conseiller à l'IST Laurent Romary, du délégué à l'IST Jacques Millet, et Alain Monteil chef de projet HAL-INRIA relativement novice ! La participation a été variable dans le temps, assez forte au début du projet puis de plus en plus faible, sauf pour J. Millet membre de l'*executive board*²⁴. Cette expérience nous a permis d'accroître notre connaissance du paysage des archives ouvertes en Europe, ainsi que de la vision que porte le monde éditorial sur le libre accès.

La confrontation avec nos partenaires d'autres structures et d'autres pays a permis de partager des expériences, de mieux se connaître permettant d'initier de possibles partenariats. Nous avons effectué une visite auprès de l'équipe de la MPDL à Munich afin de partager nos pratiques en matière d'archives ouvertes. Autres points positifs à signaler, le dialogue et contact direct avec les éditeurs autour des échanges d'information et le traitement des documents pour une bonne gestion et signalement dans les archives ouvertes. L'expertise INRIA en matière IST est à présent mieux connue et nous pouvons nous mettre en ordre de marche pour d'autres projets européens comme partenaire potentiel. Il est à signaler que la France est peu représentée dans ce type d'action, nous devrions être plus pro-actif en la matière à l'avenir.

5.2 En termes technique

La réalisation et l'exploitation du *PEER Depot* est un des éléments forts de réussite du projet PEER et de visibilité des compétences d'INRIA. Le traitement des données comme la mise au format TEI ou l'extraction de métadonnées ont permis à ses ingénieurs d'acquérir une haute expertise reconnue. Les avancées devront être diffusées et réutilisables par tous, notamment l'extraction de métadonnées à partir des fichiers sources qui pourra être intégrée comme une offre de service d'assistance au dépôt dans les archives ouvertes.

Côté éditeurs, les solutions techniques permettent de diffuser *via* des archives ouvertes institutionnelles la documentation scientifique, par ailleurs elles ouvrent des opportunités pour le modèle auteur-payeur. L'exemple de BioMedCentral et la diffusion *via* HAL-INSERM en utilisant le protocole SWORD, également utilisé dans le cadre du projet PEER, est emblématique de ce que pourrait être la communication scientifique ouverte.

En matière de perspectives, le projet PEER apportera à la fin un corpus de plusieurs milliers d'articles scientifiques librement accessibles, conservés de manière pérenne. Chaque archive participante pourra si elle le souhaite intégrer ce corpus dans sa base.

Participer au projet PEER a apporté une sensibilité nouvelle au réseau IST, visible par l'intérêt pour d'autres projets comme OpenAIRE et MedOANet. Cette expérience du réseau IST INRIA pourra être capitalisée en vue de la participation à d'autres projets de ce type.

Bibliographie

- [1] The eContent*plus* Programme
<http://ec.europa.eu/information_society/activities/econtentplus/closedcalls/econtentplus/programme/index_en.htm> consulté le 22 mai 2012.

²⁴ Organe de pilotage du projet réunissant principalement les représentants des partenaires.

- [2] Les rapports du projet PEER : <<http://www.peerproject.eu/reports/>> consulté le 22 mai 2012.
- [3] Publications et présentations autour du projet PEER : <<http://www.peerproject.eu/publications-presentations/>> consulté le 22 mai 2012.
- [4] M. Bijsterbosch, F. Brétel, N. Bulatovic, D. Peters, M. Vanderfeesten, J. Wallace, "Guidelines for publishers and repository managers on deposit, assisted deposit and self-archiving", May 2009. <<http://www.peerproject.eu/reports/>> consulté le 22 mai 2012.
- [5] J. McDonough, "Structural Metadata and the Social Limitation of Interoperability: A Sociotechnical View of XML and Digital Library Standards Development," communication à The Balisage: The Markup Conference, Montréal, Canada (2008).
- [6] P. Lopez, "GROBID: Combining Automatic Bibliographic Data Recognition and Term Extraction for Scholarship Publications," dans M. Agosti, J. Borbinha, S. Kapidakis, C. Papatheodorou, G. Tsakonas, Research and Advanced Technology for Digital Libraries. **Éd. Berlin, Heidelberg, 5714**, p. 473-474 (2009).
- [7] L. Romary, "TEI and Scholarly publishing – experience from the PEER project", <http://dho.ie/sites/default/files/events/teipublishing/TEIScholPub_Dublin.pdf>. consulté le 22 mai 2012.
- [8] Schematic of PEER Depot Workflow after PEER Annual Report D9.8. ECP-2007-DILI-537003. PEER Annual Report – Year 2. (2010) <http://www.peerproject.eu/fileadmin/media/reports/D9_8_annual_public_report_20100930.pdf>. consulté le 22 mai 2012.
- [9] F. Bretel, P. Lopez, M. Medves, A. Monteil, L. Romary, "Back to meaning – information structuring in the PEER project". Author manuscript, published in TEI Conference (2010) <http://www.peerproject.eu/fileadmin/media/ppt_about_peer/PEERBreakingNews.pdf> consulté le 22 mai 2012.
- [10] P. Lopez, "GROBID: Combining Automatic Bibliographic Data Recognition and Term Extraction for Scholarship Publications". In Proceedings of ECDL, 13th European Conference on Digital Library, Corfu, Greece (2009).
- [11] J. Fry, C. Oppenheim, S. Probeta, C. Creaser, H. Greenwood, V. Spezi, S. White, "D4.1 PEER Behavioural Research: Authors and Users vis-à-vis Journals and Repositories – Baseline report" <http://www.peerproject.eu/fileadmin/media/reports/Final_revision_-_behavioural_baseline_report_-_20_01_10.pdf> consulté le 22 mai 2012.
- [12] P. Dubini, "Complementary Article Dissemination via Journals and Repositories: Economic Evidence from the PEER Project", presented at APE 2011 - Academic Publishing in Europe, Berlin Brandenburg Academy of Sciences, Berlin, 11-12 January 2011. <http://www.peerproject.eu/fileadmin/media/ppt_about_peer/APE_presentation.pdf> consulté le 22 mai 2012.

EuropeanaLocal et la bibliothèque numérique européenne La Gironde, département français pionnier

Nathalie Gascoin

*Archives départementales de la Gironde
Chargée du projet EuropeanaLocal
Bordeaux, France*

Résumé

Le Conseil général de la Gironde est le représentant français du projet européen EuropeanaLocal.

Les objectifs sont de valoriser, pour la France, les collections des Archives départementales et du réseau Estuaire, et de diffuser la mise en place de bonnes pratiques en direction des institutions et associations détentrices de ressources culturelles sur le territoire girondin.

Les Archives départementales <<http://archives.gironde.fr>> ont piloté ce projet qui s'achève en 2011 par la mise à disposition d'1.2 million de documents numérisés sur le portail européen <<http://www.europeana.eu>>.

Abstract

Europeana, The European Library and the Europeana-Local project

The General Council of Gironde is the French representative in the European project EuropeanaLocal.

The aims are to promote access to the collections of the departmental archives service and the "Estuary" network and to disseminate the development of good management practices within institutions and associations owning cultural resources in the Gironde area.

The Departmental Archives <<http://archives.gironde.fr>> led this project which ends in 2011 and will have made 1.2 million digitized resources available on the European portal <<http://www.europeana.eu>>.

Le projet EuDML, une bibliothèque numérique européenne de mathématiques

Thierry Bouche

Université de Grenoble I, Institut Fourier (UMR 5582) & Cellule Mathdoc (UMS 5638)
BP 74, 38402 St-Martin-d'Hères cedex, France
thierry.bouche@ujf-grenoble.fr
<http://www-fourier.ujf-grenoble.fr/~bouche/>

1. Les enjeux spécifiques de la documentation en mathématiques

Les résultats mathématiques constituent l'outil primaire de nombreuses disciplines scientifiques. De fait, la formalisation et le savoir mathématiques jouent un rôle fondamental dans presque tous les domaines scientifiques et technologiques. Une remarque importante est qu'il n'y a pas de corrélation évidente entre la date à laquelle un résultat mathématique est publié et celle où il trouvera une application. En effet une même théorie peut trouver des applications dans des champs et à des époques très éloignés les uns des autres. Comme le dit Jean-Pierre Serre : « Les mathématiciens se contentent de mettre leur production à la disposition de tous, comme sur des étagères où l'on peut venir se servir. »¹

On constate donc que la bibliothèque est non seulement la principale production et l'outil de travail quotidien du mathématicien, mais aussi une ressource essentielle pour les autres sciences. Les articles anciens jouent pour cette discipline le rôle des données de la recherche, ailleurs : ce sont les briques à partir desquelles la recherche s'élabore. Aussi, les mathématiciens sont les héritiers d'une longue tradition d'édition de qualité. La relecture par les pairs avant la publication d'un article peut durer plusieurs années, si cela s'avère nécessaire pour valider les arguments. Ainsi, l'ensemble des résultats publiés forme un corpus relativement fiable et de dimensions réduites. Un résultat nouveau ne remplace pas les résultats valides précédemment connus, mais au contraire s'appuie sur eux en se contentant d'explicitier les références nécessaires.

La question de la conservation et de l'accès sur des durées très longues se pose de façon aiguë pour le corpus mathématique.

L'équipement principal des laboratoires de maths à travers le monde est une bibliothèque spécialisée. On peut considérer que le réseau mondial des bibliothèques universitaires et

¹ Michel Broué, « Jean-Pierre Serre et le métier de mathématicien », *Images des mathématiques*, CNRS, 2006. P. 28–29, <<http://images.math.cnrs.fr/Jean-Pierre-Serre-et-le-metier-de.html>>, consulté le 2 juin 2012.

spécialisées formait une bibliothèque (virtuelle) de référence à l'époque où le papier était l'unique support de diffusion et d'archivage. Depuis une vingtaine d'années, la localisation d'un document à l'aide des catalogues fusionnés et la pratique du prêt entre bibliothèques (PEB), voire les services de fourniture de documents, permettent au chercheur de satisfaire la grande majorité de ses besoins documentaires, l'identification des besoins étant elle-même considérablement facilitée par la conversion en bases de données des traditionnelles revues recensant la production mathématique mondiale, selon une classification par sujet *ad hoc* (MSC, *mathematical subject classification*). Pour l'essentiel, les questions cruciales de la conservation, de la fouille, et de l'accès étaient résolues.

Il n'en est pas de même dans le paradigme numérique actuel. La conversion rapide de l'édition classique au numérique, au moins pour les revues, est *a priori* un atout formidable pour la facilité et la rapidité d'accès aux textes. L'indexation plein texte et la navigation hypertexte ont démultiplié la visibilité et l'accessibilité des collections. Cependant, un certain nombre de questions fondamentales n'ont pas obtenu de réponses satisfaisantes.

- **Conservation.** Les bibliothèques ne possèdent plus les collections physiques et n'assurent donc plus cette fonction. Si elles reçoivent un exemplaire papier, il est de plus en plus souvent issu d'une impression laser impropre à une préservation à long terme. Dans le cas d'une faillite ou du rachat d'un éditeur, il n'y a aucune garantie que les collections resteront disponibles.
- **Accès.** La propriété de l'exemplaire papier mis à la disposition des lecteurs, y compris hors les murs grâce au PEB, n'a pas d'analogue dans le monde numérique, où les éditeurs conservent l'intégralité de leur production sur leur serveur, et contrôlent tous les accès. Les quelques systèmes d'archives plus ou moins noires mis en place soit au niveau national, soit par des sociétés comme Portico, soumettent toute communication d'un article au contrôle de son éditeur sur des périodes d'embargo extrêmement longues. Pour la première fois de l'histoire, un chercheur peut se trouver dans la situation où il n'a pas de moyen légal de se procurer un texte nécessaire à son travail. L'auteur qui se réjouissait de communiquer aux quatre coins du monde ses tirés à part a généralement signé un contrat qui lui interdit d'expédier le PDF de l'éditeur par courrier électronique.

Je ne traite pas ici des aspects économiques et contractuels qui accompagnent ce passage au tout électronique et à la commercialisation des revues par bouquets, lesquels aboutissent à vider de leur sens d'autres fonctions essentielles d'une bibliothèque (notamment la capacité de sélectionner les acquisitions sur des critères de qualité et de pertinence).

2. Une bibliothèque numérique mondiale et exhaustive ?

Spectateurs de ces changements qui ont un impact considérable sur leur environnement de travail quotidien, mais aussi acteurs (notamment à travers les sociétés savantes qui sont impliquées dans l'édition mais aussi dans la production de livres, de revues, mais aussi de bases de données recensant leur production : MathSciNet est un service de l'*American*

Mathematical Society, ZBMath est supervisé par la Société mathématique européenne²), les mathématiciens de la fin du xx^e siècle se sont mis à rêver d'une bibliothèque numérique mondiale et exhaustive. Cette bibliothèque numérique (dite DML : *Digital mathematical library*³), conçue à l'origine, et de façon un peu naïve comme la transposition électronique de la bibliothèque que chaque mathématicien fréquente dans son laboratoire, comporterait dans ses rayonnages virtuels la totalité de ce qui s'est publié depuis les *Éléments* d'Euclide. Elle serait étroitement liée avec les bases de données, et surtout aussi navigable et gratuite que possible.

L'exhaustivité n'est pas un rêve : c'est seulement une question d'argent et de volonté. La totalité du corpus ne couvre pas plus de 100 millions de pages, selon des estimations crédibles : c'est moins que ce qui s'est publié dans le seul domaine biomédical depuis l'an 2000, cela représente une masse d'information modeste qui tiendrait sur un seul disque dur !

La « navigabilité » est une des motivations fortes pour la communauté, du fait du caractère cumulatif de la littérature mathématique. Quelle facilité si la machine fournit le lien *ad hoc* chaque fois que l'on souhaite suivre une référence ! À comparer avec le nombre de sites qu'il faut connaître, de stratégies différentes qu'il faut y développer pour trouver quelque chose aujourd'hui.

Quant à la gratuité, elle intéresse bien peu les acteurs économiques. Il y aurait de grands bénéfices à en tirer, notamment parce qu'il serait bon que cette ressource basique soit disponible bien au-delà de l'environnement professionnel du chercheur en mathématiques. Une autre raison, plus pragmatique, est que la navigation est infructueuse si tous les liens pointent sur des portes closes. Or, du fait de la désorganisation des accès électroniques, aucun laboratoire dans le monde ne possède toutes les licences d'accès nécessaires.

En tant que projet *global*, la DML a peu avancé au cours des dix dernières années, malgré le soutien de l'Union mathématique internationale (IMU), de la Société mathématique européenne (EMS) et la tenue de nombreuses rencontres internationales sur le sujet.

Dans l'intervalle, si nous sommes toujours très loin d'avoir satisfait les ambitions du projet DML, le contenu mathématique existant au format numérique est devenu tout à fait considérable : il représente plus du tiers de la totalité du corpus, tous supports confondus. Le tiers des articles référencés dans les bases de données dispose d'un lien vers son texte intégral, bien que de très nombreuses ressources leur échappent, en général parce qu'elles sont produites et gérées par des non-professionnels. C'est au niveau des barrières d'accès que la situation est la plus décevante : conçue comme un service public dans un monde qui perd cette notion, la DML peine à trouver un modèle économique viable.

L'IMU a proposé depuis longtemps que les articles deviennent librement accessibles à l'issue d'un certain délai d'embargo pendant lequel l'éditeur récupère ses coûts de production

² MathSciNet, *Mathematical Reviews on the Web* : <<http://www.ams.org/mathscinet/>> ; ZBMATH, the *Zentralblatt MATH* database : <<http://www.zentralblatt-math.org/zblmath/en/>>, consultés le 2 juin 2012.

³ Voir <<http://www.mathunion.org/WDML/>>, consulté le 2 juin 2012.

(valeur généralement suggérée : 5 ans). Si cette recommandation a été entendue de nombreux acteurs académiques, elle n'a pas ébranlé les éditeurs commerciaux. Conscients de l'intérêt que les mathématiciens portent à leur patrimoine scientifique, les éditeurs commerciaux ont entrepris de numériser leur propre fonds qu'ils vendent désormais comme un nouveau produit. Les projets américains sans but lucratifs, mais sans financement public récurrent, comme JSTOR ou Project Euclid⁴, assurent leur pérennité en commercialisant l'accès à leurs services par les bibliothèques à un tarif qui peut nous paraître prohibitif. Au final, la quantité des projets apportant une partie des services attendus de la DML s'est considérablement accrue, et avec elle le nombre de conditions et de barrières d'accès !

Je décris dans ce qui suit deux projets de nature assez différente dans lesquels j'ai été impliqué au sein de la cellule Mathdoc : le projet de numérisation NUMDAM et son extension par des chaînes d'acquisition nativement numérique ; le projet EuDML qui a pour objet de porter au niveau européen une vision à la fois ambitieuse et pragmatique de la DML.

Je signale que l'histoire ne s'arrête pas là et qu'une rencontre organisée par l'IMU a eu lieu à Washington en juin 2012 pour relancer l'effort mondial.

3. Le programme NUMDAM

NUMDAM a été conçu comme un projet de numérisation de revues françaises de mathématiques (d'où le développement officiel de cet acronyme hybride en « Numérisation de documents anciens mathématiques ») dans une logique de soutien à la visibilité et à la diffusion la plus large possible des archives des revues académiques nationales. Le financement initial, de l'ordre de 450 000 €, a été débloqué par le ministère de la Recherche à une époque un peu particulière (la fameuse « cagnotte de Jospin »). Si les motivations étaient assez claires et convaincantes pour emporter cette décision, les modalités de mise en œuvre n'étaient pas finalisées du tout : les précédents directeurs de Mathdoc s'étaient inspirés du peu d'informations disponibles sur les rares projets de numérisation déjà actifs (JSTOR aux États-Unis, ERAM en Allemagne, Gallica en France, le projet européen DIEPER...) et avaient pour principal objectif de viser la meilleure qualité de numérisation. En ce qui concerne l'exploitation des données, rien n'était précisé : la diffusion des archives numérisées pourrait avoir lieu sur le site de chaque revue, sur la plateforme d'un sous-traitant, l'accès pourrait se faire moyennant une redevance modeste pour compenser ces coûts...

Les travaux préparatoires avaient commencé avant que l'argent ne soit versé, courant 2000, ils s'achevèrent par la rédaction d'un cahier des charges technique pour un appel d'offre de numérisation publié en juin 2001. Dans l'intervalle, des accords avaient été passés avec les responsables des 6 séries concernées par la première phase. Un recensement exhaustif du contenu et des particularités physiques des 6 collections avait été réalisé par des vacataires. Deux postes d'ingénieurs avaient été affectés par le CNRS, l'un sur un emploi permanent, l'autre en CDD, pérennisé un an plus tard.

⁴ Voir <<http://www.jstor.org/>> et <<http://projecteuclid.org/>>, consultés le 2 juin 2012.

La numérisation proprement dite débute en janvier 2002, chez le sous-traitant lauréat du marché. Une chaîne de contrôle qualité (manuel sur échantillons, et exhaustif pour tout ce qui pouvait être automatisé) est développée en interne ; il faut encore attendre six mois pour que la qualité du prestataire corresponde aux standards définis par le cahier des charges, et que la production prenne enfin sa vitesse de croisière. En parallèle, nous commençons à étudier les options pour la mise en ligne des articles et nous contactons les auteurs pour obtenir leur autorisation — l'un des éditeurs profitera de cette occasion pour faire signer des contrats d'édition à tous ses auteurs. Nous mettons en place des post-traitements pour les fichiers utilisateurs : ajout d'une page de garde, linéarisation et protection des PDF contre les modifications ; ajout de liens vers les bases de données pour les articles et ceux qu'ils citent en bibliographie... Un ingénieur de Mathdoc adapte un logiciel qu'il a développé dans un autre contexte (interface de recherche et de consultation du ZBMath) de façon à ce que nous disposions d'une maquette simple d'accès aux collections et commençons à réfléchir aux modalités de la mise en ligne. Cette maquette s'avère tellement fonctionnelle que nous décidons de l'adopter après quelques ajustements. Les deux premières revues apparaissent en ligne le 20 décembre 2002, les six revues de la première phase sont en ligne un an plus tard, soit 7 000 articles couvrant 200 000 pages dont la quasi-totalité sont en libre accès.

NUMDAM, en tant que projet pilote dont la première ambition était de numériser l'intégralité des archives de cinq ou six revues académiques de mathématiques publiées en France des origines jusqu'à l'an 2000, est alors dans une situation intéressante : moins de la moitié de l'argent alloué à cette opération a été consommée bien qu'elle ait duré 6 mois de plus que prévu. Les objectifs un peu vagues formulés lors du financement du projet sont largement atteints, mais le produit final ne ressemble pas du tout à ce qui avait pu être imaginé à l'époque, aussi bien sur le plan du modèle d'affaires (les articles sont en libre accès sur une plateforme entièrement gérée par la cellule Mathdoc) que de l'exploitation des données capturées. En fait, ce service sera le modèle explicite pour les projets qui naîtront en Europe à ce moment-là (DML-E en Espagne, DML-CZ en République Tchèque, BDIM en Italie), et ailleurs (*Canadian Journal of Mathematics...*).

Grâce au soutien constant de nos tutelles, la disponibilité du budget alloué en l'an 2000 ne sera pas remise en question pendant un peu plus de six ans : un deuxième marché pourra ainsi être financé et tripler la taille des collections. La troisième phase du programme NUMDAM s'est achevée en 2011, portant le nombre de séries numérisées à 66, dont 2 séries de mémoires et une collection de thèses. La couverture disciplinaire est désormais très large : de l'informatique théorique à la théorie des catégories, de la physique aux statistiques, de l'arithmétique aux sciences humaines. Des revues italiennes et hollandaises nous ont rejoints. Les collections débutent avec le premier article des *Annales de Gergonne*, première revue exclusivement consacrée aux mathématiques publiée dès 1810 à Nîmes, et comportent plus de 600 articles publiés depuis 2010, sur un total de 52 000.

Au-delà du seul programme de numérisation, le serveur NUMDAM abrite en effet la production des éditeurs qui veulent bien la lui confier. Cela concerne une vingtaine de titres publiés par une dizaine de structures, ce qui représente de l'ordre de 10 % des contenus

diffusés. Nous avons à cette fin développé des chaînes d'acquisition numérique qui nous permettent de convertir les métadonnées des éditeurs à notre format interne.

Au final, le projet pilote de l'an 2000, prévu pour une durée de 3 ans, s'est discrètement transmué en un service pérenne de bibliothèque de référence archivant *a posteriori* une partie substantielle du corpus mathématique produit en France et alentour. Le coût consolidé sur ses dix premières années avoisine le million d'euros, soit à peu près un euro par page numérisée.

4. Le projet EuDML

4.1 Présentation

Depuis l'année 2003, la SME n'a cessé d'appuyer la constitution d'une bibliothèque numérique de mathématiques, volet européen de la DML. Elle a à cette fin présenté plusieurs propositions à l'occasion de divers appels à projets (5^e PCRD, eContent+, ESF, 7^e PCRD...). C'est pour finir en 2009 que la première proposition fut retenue, dans le cadre du programme cadre pour la compétitivité et l'innovation, volet d'appui stratégique en matière de technologies de l'information et de la communication, thème : « bibliothèques numériques », objectif : « libre accès à l'information scientifique ».

Voici la description des projets attendus dans le cadre de cet appel⁵ :

« Des actions pilotes B seront financées pour mener des expériences concluantes de libre accès à des bibliothèques numériques de contenu scientifique/savant, y compris des expériences explorant de nouveaux paradigmes pour l'évaluation par les pairs, l'affichage, l'interrogation et la constitution de liens pour des contenus scientifiques/savants et (optionnellement) les ensembles de données sous-jacentes associés. »

Et voici l'impact espéré : « Accès libre à plus de contenu scientifique/savant et/ou le développement de nouvelles façons d'évaluer, afficher, interroger et lier du contenu scientifique/savant. »

Pour répondre à cet appel et à « l'amicale » pression de la SME, j'ai formé un consortium composé de 14 partenaires européens⁶. Une dizaine d'entre eux sont des institutions publiques qui apportent les principales collections de mathématiques numériques (en majorité numérisées) en Europe⁷. Un seul éditeur commercial (EDP Sciences) fait partie du consortium. Il participe avec cinq journaux français dont la série ESAIM, éditée sous les auspices de la Société de mathématiques appliquées et industrielles (SMAI) et les

⁵ *Competitiveness and Innovation Framework Programme (CIP), ICT Policy Support Programme, ICT PSP Work Programme 2009*, appel publié le 29 janvier 2009. [La traduction est de moi.]

⁶ La liste détaillée se trouve sur le site Web du projet : <<http://project.eudml.org/>>. La vision institutionnelle du projet par la Commission européenne est en ligne à l'adresse <http://ec.europa.eu/information_society/apps/projects/factsheet/index.cfm?project_ref=250503>, consulté le 2 juin 2012.

⁷ Projets « nationaux » déjà opérationnels : les tchèque DML-CZ, français NUMDAM, espagnol DML-E, polonais DML-PL, portugais PtDML, grec HDML, et des projets émergents comme le bulgare BulDML.

descendants de la revue *RAIRO* qui ont été numérisés par NUMDAM. De nombreux autres éditeurs sont associés au projet à travers leurs archives numérisées par un partenaire ou plus directement parce qu'ils utilisent pour leur édition électronique l'une des plates-formes partenaires du projet⁸. L'ensemble des partenaires représente une diversité notable de compétences techniques, depuis les bibliothèques numériques et les services d'édition électronique jusqu'au traitement automatisé du savoir mathématique.

Le consortium est renforcé par l'apport de ses deux partenaires associés, qui ne perçoivent aucun financement européen : la Société mathématique européenne (SME) est associée au projet comme autorité morale fixant les objectifs et évaluant l'utilité des résultats du projet. Elle préside un comité scientifique consultatif. La bibliothèque universitaire de Göttingen apporte les journaux numérisés par les projets ERAM et RusDML, ainsi que la plus grande collection de livres numérisés de mathématiques.

Nous avons rédigé un argumentaire (insistant notamment sur le fait qu'en mathématiques le libre accès à *terme* est un objectif pertinent puisqu'il est primordial de disposer d'un corpus de référence validé et accessible sans barrières) et un plan de travail sur 3 ans découpé en 11 *work packages*, jalonné d'objectifs précis. La coordination du projet a été dévolue à un collègue portugais, seul membre de notre consortium doté de compétences managériales.

Le projet a été accepté sur la base d'une durée de trois ans (à partir du 1^{er} février 2010), d'un volume de travail global de 487 personnes-mois, pour un financement maximal de 1,6 M€.

4.2 Objectifs

Par référence aux objectifs de la DML, le projet EuDML met l'accent sur quelques-uns des chantiers, espérant bénéficier d'une double masse critique : l'importance des contributions européennes en mathématiques, particulièrement au cours de la période très représentée des XIX^e et XX^e siècles, et la quantité de textes fondamentaux rendus disponibles numériquement par l'un de ses partenaires.

Un portail global. Le premier aspect, degré zéro de l'intégration, porte sur la mise en place d'un point d'accès unifié à l'ensemble des collections à travers un portail Web personnalisable permettant de feuilleter, fouiller, et naviguer dans les collections. Ce système central dispose de services pour transformer les références à ces articles en liens hypertextes, de façon automatisée.

De la sorte, on espère faciliter la fouille dans le corpus mathématique présent sur le portail, donner plus de visibilité à ce corpus précédemment éclaté, et mettre toutes les références hébergées à un clic de l'utilisateur.

⁸ Ce qui signifie qu'environ 90 revues fournissent directement leurs articles récents : elles sont diffusées par EDP Sciences, le projet CEDRAM de la cellule Mathdoc, ou le service ELiB du FIZ Karlsruhe. Leurs éditeurs sont pour la plupart des petites structures, allant d'une équipe de volontaires à des sociétés savantes en passant par des structures académiques ou de petites sociétés privées. <<http://publications.edpsciences.org>>, <<http://www.cedram.org>>, <<http://www.emis.de/journals/>>, consultés le 2 juin 2012.

Le défi est ici de définir un format d'échange suffisamment riche pour décrire de façon fine et précise la totalité des contenus (articles de revues, communications dans des séminaires ou des congrès, monographies, œuvres en plusieurs volumes...).

Une archive répartie. Une idée force pour la réalisation de la DML est non plus de partir d'un projet centralisé et pyramidal, mais de mettre en réseau les réservoirs de textes existants, et d'atteindre ainsi une masse critique. Le schéma proposé est que chaque éditeur identifie le partenaire du réseau le plus approprié pour y déposer ses archives (y compris récentes), et décide avec lui des modalités de cet archivage, à la façon d'un dépôt légal volontaire.

C'est ainsi que pourrait se constituer une archive pérenne de référence, indépendante des aléas économiques de l'édition, prise en charge par des institutions scientifiquement fiables et durables.

Cet aspect du projet n'a cependant pas convaincu grand monde au-delà des partenaires pour l'instant. D'autres bibliothèques numériques au profil similaire à celui des partenaires actuels se sont associées en cours de route : citons un projet italien (BDIM) et un russe (math-net.ru).

Libre accès à terme. La modalité de fonctionnement suggérée pour les éditeurs partenaires est la suivante :

- lorsqu'un éditeur a complété sa production, une copie archivable (métadonnées et textes intégraux) est fournie à l'institution *ad hoc* ;
- ces contenus sont validés et enregistrés ;
- les nouveaux textes sont indexés et apparaissent dans les résultats de recherches ;
- l'accès aux textes intégraux se fait sur le site de l'éditeur sous son contrôle ;
- à l'issue du créneau mobile, la copie locale devient librement accessible.

De la sorte, les textes intégraux sont archivés par des tiers pérennes, la quantité de textes de référence en libre accès augmente, La mathématique établie, mère de toutes les sciences, est enfin à la portée de tous !

Innovation. Un aspect particulièrement intéressant et ambitieux de ce projet est d'aller bien au-delà de la simple intégration dans un « guichet unique », et d'essayer pour la première fois de déployer des techniques avancées de gestion des savoirs mathématiques, de telle sorte que les spécificités des textes mathématiques (haute formalisation, emploi de formules...) ne soient pas un handicap, mais un ingrédient essentiel pour la fouille et la navigation dans le corpus.

Pour ce faire, nous privilégions les métadonnées en MathML, que nous cherchons à produire à partir de toutes les sources disponibles : OCR pour les textes scannés (avec le concours du logiciel *Infty Reader*), conversions à partir de sources LaTeX, extraction et reconstruction à partir de PDF nativement numériques.

À partir de là, le projet vise à permettre la recherche de formules, la mise en relation des textes à partir de calculs de similarité de leurs contenus mathématiques. Un pan entier du

travail est aussi dédié à rendre le corpus plus accessible aux utilisateurs handicapés par la dyslexie ou des problèmes de vue.

4.3 Résultats

Les collections prises en charge sont actuellement de 225 périodiques et séries, 235 000 textes, 2 600 000 pages.

Le serveur actuel contient en principe tous les types de documents prévus.

Sources	Nombre de textes
Allemagne ERAM/JFM, GDZ, ELibM	85 000
France Gallica-Math, NUMDAM, CEDRAM, EDPS	57 000
Rép. Tchèque DML-CZ	32 000
Russie RusDML	17 000
Pologne DML-PL	14 000
Grèce HDML	4 000
Espagne DML-E	6 400
Portugal SPM/BNP	2 000
Bulgarie BulDML	600
Italie BDIM	2 000

Au cours de ce projet, qui s'achèvera début 2013, un certain nombre de tâches ont été entreprises avec succès :

1. l'analyse des contenus disponibles et la définition d'un format d'échange. Ce format est basé sur le schéma XML "*Journal Archiving and Interchange Tag Set*" développé initialement pour PubMed Central, adopté par JSTOR ou Portico, et devenu récemment un standard NISO ;
2. une conférence a été organisée à Prague en octobre 2010 afin de discuter avec des partenaires potentiels (Springer, éditions de la *London Mathematical Society*...) de nos choix politiques et de leur capacité à attirer ou faire fuir de nouveaux partenaires, notamment parmi les éditeurs (libre accès à terme – créneau mobile, nouveautés fournies par les éditeurs et indexées rapidement, archivage partagé dans un réseau de bibliothèques numériques de référence). Nous n'avons pas entendu de *non* ferme et définitif, mais pas non plus de *oui* enthousiaste de la part des éditeurs recherchant le profit...
3. la constitution d'une base de données avec 235 000 références soit environ 1/5 du contenu numérique existant (1/3 de l'existant « DML ») ;
4. une série d'outils destinés à améliorer l'indexation, la navigation, l'accès et la visibilité pour ce corpus spécifiquement mathématique ;
5. un prototype désormais fonctionnel est en ligne et opérationnel depuis l'été 2012 sur <<http://eudml.org/>>. Le contenu disponible en ligne dépasse en quantité les prévisions initiales, quelques collections ayant continué à croître, et un projet extérieur s'étant associé (d'autres sont en cours de rapprochement). La diversité

des types de contenus pris en charge (articles, livres, communications dans des conférences) couvre l'essentiel du corpus visé ;

6. la recherche fonctionne assez bien, on peut trier et raffiner les résultats par facettes. On dispose d'une recherche plein texte sur la moitié des textes et d'une recherche expérimentale sur les formules, dont la pertinence reste à évaluer. Un feuilletage est disponible par journal, classification mathématique ;
7. toutes les métadonnées disponibles ont été converties en MathML, si bien que les informations sur un texte sont parfois mieux présentées sur le site <euDML.org> que sur le serveur d'origine ;
8. la création systématique de liens est opérationnelle (liens vers les bases de données MathSciNet et ZBMath depuis les textes EuDML ou leurs références bibliographiques, mais aussi de nouveaux liens internes entre les textes EuDML), certaines fonctions Web 2.0 sont actives ;
9. les interfaces de service qui permettent des appels directs depuis des machines distantes de la base de données EuDML viennent d'être rendues publiques. Ceci ouvre notamment la possibilité de transformer de façon automatique toute référence à un texte en lien vers EuDML depuis toute page Web, donc réalise une des fonctions principales d'une bibliothèque numérique de référence : mettre ses collections à la disposition de ses utilisateurs, en un clic, où qu'ils soient. Il reste à faire adopter l'outil de la façon la plus large.

Au final, le projet EuDML, qui avait affiché de très hautes ambitions aura atteint une partie substantielle de ses objectifs. La cause de la DML a progressé grâce à lui, par la mise en œuvre de la première réalisation transfrontalière. Il reste un assez grand nombre de questions ouvertes et de défis à relever dans la nouvelle période qui s'ouvrira à la fin du financement européen. Il n'est en effet pas simple de concevoir et développer une infrastructure pérenne avec des moyens dépendant d'appels à projets.

Les directions à suivre pour faire évoluer le prototype vers une infrastructure mondiale et complète sont évidentes : il faut

- étendre la couverture en associant des partenaires de tous les continents et en intégrant des contenus de provenances variées ;
- consolider l'architecture technique pour disposer d'un service souple et facilement configurable pouvant s'adapter à de nouvelles collections et intégrer des services évolués.

L'avenir nous dira si nous aurons les moyens de poursuivre les voies qui ont été ouvertes par ce projet.

A concept for Inferring ‘Frontier Research’ in Research Project Proposals

Dirk Holste¹, Ivana Roche², Marianne Hörlesberger¹, Dominique Besagni²,
Thomas Scherngell¹, Claire Francois², Pascal Cuxac², and Edgar Schiebel¹

¹ Austrian Institute of Technology, Donau-City Strasse 1,
1220 Vienna (Austria)
dirk.holste@ait.ac.at

² INIST-CNRS, 2 Allée du Parc de Brabois, CS 10310,
54519 Vandoeuvre-les-Nancy, France

Abstract

This paper discusses a concept for inferring attributes of ‘frontier research’ in peer-reviewed research project proposals under the scheme of the European Research Council (ERC). The concept serves two purposes: 1) to conceptualize and define, automatically extract, and comparatively assess attributes of frontier research in proposals; and 2) to build and compare outcomes of a statistical model with the review decision in order to obtain further insight and reflect upon the influence of frontier research in the peer-review process. To this end, indicators (including scientific ‘novelty’, ‘risk’, or ‘interdisciplinarity’) across scientific disciplines and in accord with the strategic definition of frontier research by the Council are elaborated, exploiting textual proposal information and other bibliometric data of grant applicants. Subsequently, a concept is discussed to measure ex-post the influence of indicators on the decision probability (or, alternatively, the odds) of a proposal to be accepted. The final analysis of the classification and decision probabilities compares and contrasts review decisions in order to, e.g., statistically explain congruence between frontier research and review decision or reveal differential representation of attributes. Ultimately, the concept is aiming at a methodology that monitors the effectiveness and efficiency of peer-review processes.

Communication aux 13th International Conference on Scientometrics and Informetrics, 4-8 July 2011, Durban, Afrique du Sud.

1. Background and Objectives

Scientific disciplines use peer review as an essential mechanism for resource allocation and quality control [1]. From the point of the performed research, it serves either to determine what concluded research deserves to be published (ex-post review) or what proposed research deserves to be funded through national or regionally operating agencies (ex-ante review). Reviewers face the challenge to sort out the “best” scientific research in accord with a journal’s or agency’s strategy. Typically, both journals and grant schemes are sufficiently different such that there are no “best practices” in either category. Peer-review systems are rooted in critical rationalism, widely accepted and actively supported quality control instrument at hand. Still they are not free from criticism on a number of issues, including poor reliability (congruent opinions); fairness (biased opinions of non-scientific merit); predictive validity

(decisions in accord with merit to the scientific community); or inefficiency (e.g., delay, resources spent, opportunity cost). Several issues are 'systemic' in that biases and quality filtering are not necessarily incommensurable. *E.g.* making scientific progress is fundamentally a conservative act by building incrementally upon previous work and requiring evidence resp. balance of risk/reward before publishing (funding) radical papers (ideas); or social system dynamics of subjective human interaction giving rise to lower reliability but are offset by higher predictive validity [1; 2]. Because of its central role for not only the scientific community but also publishers/editors and funding agencies, monitoring peer review processes is essential to shed light on to what extent set goals are actually accomplished by review decisions. The need for monitoring effects and the implicit reorientation of peer review processes is subject to current research activities [3; 4; 5; 6].

Where the scope and specialization of science has become too complex and sophisticated to be coped with by any one reviewer, scientometric methods offer a 'helping hand' to either support the decision-making process or evaluate its outcome. In fact, scientometric evaluation has been witnessing a significant attention in the rising need to get a grip on science output and efficiency [7; 8].

On the one hand, these methods carry strengths in that they are precisely defined and reliable, objective, efficient, and need no intervention; on the other hand, their weakness comes mainly in terms of limits of interpretation, applicability, confounding factors, and predictive validity – all of which are more or less debatable. Table 1 lists selected examples of scientometric indicators that are performance-centred and use easy-to-measure (in principle) volume data. Other indicators incorporate data extracted from networks of co-authorships, citations or bibliometric coupling; curriculum vitae, *e.g.*, age, gender, institution (teaching, research); referee scores; or socioeconomic relevance (previously attracted, requested funds).

Naturally a numerical model cannot be expected to substitute for expert peer review/process and the merit of the scientific community, each of which being hard to quantify. More faithfully models of this kind may serve to verify decisions, deliver support data efficiently, or hint at biases of the review process [9].

The nature and objectives in journal reviews (co-authorship, selection, improving quality of published research) is different from reviews of proposed research projects (individual investigator, allocation of resources to inherently risky, speculative projects). This is particularly evident in different expectations on predictive validity and, consequently, the choice of indicators tailored to the underlying strategy, mission and policy of funding bodies to establish interpretable and useful cause-effect relationships. Meanwhile discrepancies between the selection decision on the set of "best" proposals derived from review systems and scientometrically predictive identifications are expected. Because measures, numbers, and comparisons among peers can deliver a powerful message and impose normative behavior [10], a number of studies have shed more light on underlying reasons [*e.g.*, 7; 11; 9].

Indicator	Measurement	Interpretation options
Authorship	Number of publications or co-publications	Research output, productivity
Citation	Number of citations	Research influence, international impact
Publication-citation	<i>h</i> -index	Research productivity and impact
Journal impact factor	Total number of journal citations in year <i>n</i> of papers published in years <i>n</i> -1 and <i>n</i> -2 over the number of articles published in <i>n</i> -1 and <i>n</i> -2	Reputation of scientific journals
Online access	Number of times a paper is accessed online in some time period <i>T</i>	Global spread, attention in scientific community and beyond
Network properties	Social network parameters	Beyond stand-alone, reflect system properties and influence in network interconnectedness and speed of information exchange

Table 1 : Selected classical scientometric indicators for measuring scientific performance

This paper looks at bibliometric evaluation of research project proposals:

- From a grant point of view, it focuses on proposals submitted to the prestigious European Research Council (ERC) in the scientific domains "Physics & Engineering" (PE) and "Life Sciences" (LS)¹. Scientists from all over the world, who are intending to work with a host institution based in a EU Member State or associated country, can compete for two different types of grants: Starting Grants (SGs) for investigators with 2-12 years of experience after their PhD at the stage of starting or consolidating their independent research team; and Advanced (AGs) for already established investigators with at least 10 years of experience and significant research achievements [12]. Grants are to support pioneering, far-reaching research endeavours, combine high risk/high impact potential, break established disciplinary boundaries, or explore

¹ PE (LS) holds ten (nine) main and ~170 (100) subcategories. The third domain "Social Sciences & Humanities" is not considered as it is expected to differ in terms of publishing, citation behaviour, and other features from those observed in PE and LS (e.g., national/regional orientation, less publications in form of articles, different theoretical 'development rate', number of authors, non-scholarly publications), which make it less assessable for approaches developed for the natural and life sciences [19; 9].

new productive lines of scientific enquiry, methodology or techniques. Each project can receive up to 2 m€ (SGs) or 3.5 m€ (AGs) for a maximum of 5 years (*cf.* Table 2).

- From a methodological point of view, it complements the standard approach to scientific excellence, which is classically based on quantity of papers, citations, etc. [11], and takes into account textual features related to the content and quality of 'frontier research' [13] detectable in individual research proposals – being the sole criterion for awarding ERC grants to young or senior investigators [14].

Due to the single evaluation criteria (scientific excellence), ERC grants provide a suitable test-bed for content analysis/text-mining and modelling in the field of bibliometric evaluation [15]. The primary interest is the extent to which research proposal comply with attributes of frontier research and the influence of these attributes on the selection of awarded grants.

ERC grant (year)	Total budget (m €)	Number of proposals submitted	Total number of grants awarded	Number of grants awarded (in PE and LS)
SG (2007)	335	9,167	299	242
AG (2008)	553	2,167	282	198
SG (2009)	325	2,503	244	187
AG (2009)	515	1,584	244	202

Source: [16]; since 2009, SG and AG grants are awarded annually

Table 2 : Number of proposals submitted and grants awarded by the ERC in 2007–2009

The remainder of the paper is structured as follows. After introducing the ERC peer review system and the definition of frontier research, the outline a concept of the review process is presented. Subsequently scientometric and text-analytic methods [17; 18] capturing desired attributes of frontier research are laid out. Then, a discrete choice model is adopted to approximate the selection function and how indicators influence the decision probability for a proposal to be accepted. Finally, a discussion of the concept closes the paper by elaborating a comparison of peer review process and model outputs and the assessment of the "influencing power" of the individual indicators.

2. A concept based on joint scientometric and content analysis

2.1 Peer review process set up by the ERC

The first European research funding body targets research at the highest level of excellence in any scientific discipline. It supports investigator-driven projects aiming at broadening the scientific and technological knowledge without regard for established

disciplinary boundaries (*frontier research*²) through open and direct competition. The selection of proposals for is based strictly on peer review. The ERC has established a process which is to identify scientific excellence of frontier research as the sole evaluation criterion for funding decisions (ERC 2010). Internationally renowned scientists and scholars constitute two sets of review panels (for the SGs and AGs), each of which subdivided into 25 individual panels that cover the entire range of disciplines and fall into in the domains PE, LS, and SH³. Each panel is composed of 10-12 members and headed by a chair. If further expertise is required, external reviewers may be consulted by providing assessments on a proposal-by-proposal basis.

Indicator	Definition (Frontier research...)
NOVELTY	"... stands at the forefront of creating new knowledge and developing new understanding. Those involved are responsible for fundamental discoveries and advances in theoretical and empirical understanding, and even achieving the occasional revolutionary breakthrough that completely changes our knowledge of the world."
RISK	"... is an intrinsically risky endeavour. In the new and most exciting research areas, the approach or trajectory that may prove most fruitful for developing the field is often not clear. Researchers must be bold and take risks. Indeed, only researchers are generally in a position to identify the opportunities of greatest promise. The task of funding agencies is confined to supporting the best researchers with the most exciting ideas, rather than trying to identify priorities."
PASTEURESQUENESS	"... may well be concerned with both new knowledge about the world and with generating potentially useful knowledge at the same time. Therefore, there is a much closer and more intimate connection between the resulting science and technology, with few of the barriers that arise when basic research and applied research are carried out separately."
INTERDISCIPLINARITY	"... pursues questions irrespective of established disciplinary boundaries. It may well involve multi-, inter- or trans-disciplinary research that brings together researchers from different disciplinary backgrounds, with different theoretical and conceptual approaches, techniques, methodologies and instrumentation, perhaps even different goals and motivations."

Source: definition: [13]; indicator: own data

Table 3 : Relation between the definition of frontier research and correspondence of indicators

- ² The clear distinction between basic and applied science has become blurred because new areas of science and technology often display substantial elements of both. Frontier research reflects that reality and calls for concentrated efforts on a more receptive attitude to unconventional thinking [20].
- ³ In addition, a fourth domain 'Interdisciplinary' has been defined for cross-panel and cross-domain proposals.

Proposal reviews are carried out in two steps. 1) Individual panel member assess both the qualification of PI (Criteria-1) and originality of the proposal (Criteria-2), discuss jointly in panel meetings, score⁴ PI and proposal by at least three reviewers based on a scoring system⁵, and preliminary rank and eliminate low-ranking proposals. 2) Both panel members and remote referees individually assess Criteria-1,2 and the research environment (Criteria-3), and interview SG applicants; subsequently proposals are further discussed, scored and ranked, upon which chair meeting consolidate results across different panels and compose a final rank order. The ultimate outcome is a list of top ranks inside the year's budget, next-in-rank (~50% of budget) pending and lower ranks above threshold with no chance of funding.

2.2 Defining frontier research and corresponding indicator

In response to the insight that the classical distinction between 'basic' and 'applied' research has ceased to exist, and that emerging areas of science and technology often embrace elements of both, a High Level Expert Group used the term "frontier research" to denote research that reaches beyond horizons of existing knowledge by being intrinsically risky endeavours without regard for established disciplinary boundaries. Table 3 cites key attributes of frontier research according to the Group's report [13].

⁴ SG, intellectual capacity to reach beyond the state of the art and creativity (achievements, publications); grant as contribution to establish independence or its consolidation; commitment (time in-residence, on project). AG, ground breaking nature of research, potential impact; interdisciplinarity; novel concepts; highly novel and/or unconventional methodologies justified by high gain/high reward balance); feasibility; recourses claimed justified added-value of team members.

⁵ Categories and marks used: "non-competitive" (1), "very good" (2), "excellent" (3), or "outstanding" (4).

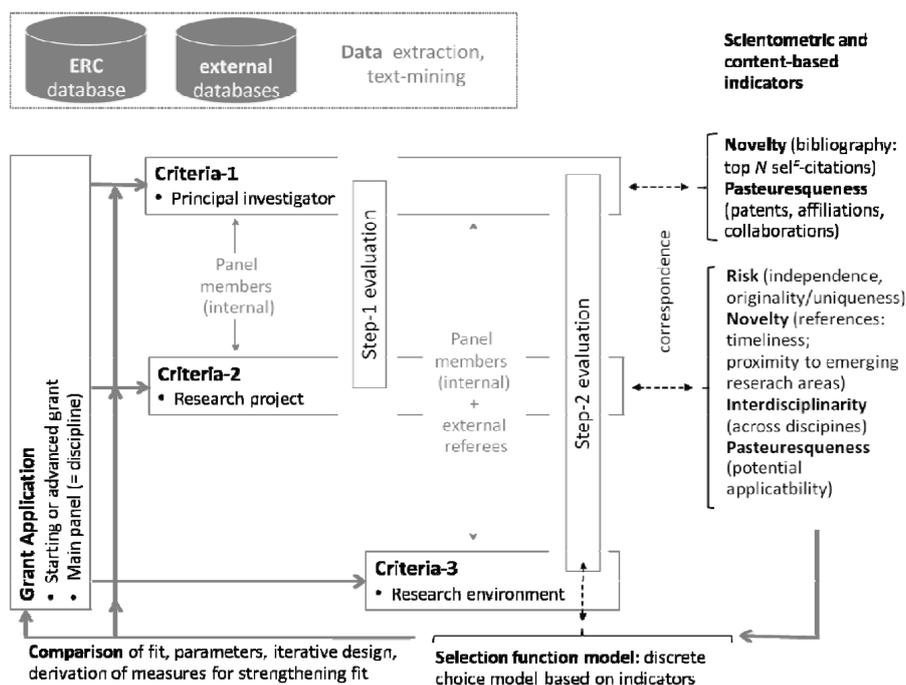


Figure 1 : Core review process and corresponding indicators and the selection function

2.3 Overall concept

The identification of frontier research through a combination of scientometric and text-mining methods is ambitious. The concept introduced here builds upon the understanding of the High Level Expert Group's notion of frontier research. The relation between a sought quantitative model and the above definition is made transparent through the correspondence between each identified key attribute and its indicator as well as a selection function (Figure 1).

Although each indicator has a clear measurement function and subjective interpretation, it is insufficient to represent frontier research. A faithful representation is achieved in combination, which is implemented in form of a selection function. Clearly the notion of 'revolutionary breakthrough' (*cf.* Table 2) is practically inaccessible by scientometric and textual methods alone. Here two indicators capture different albeit related aspects of the research activity in question: the "timeliness" (one aspect of novelty) of the knowledge-base explicitly used by the author and the "proximity to emerging research topics" (another aspect of novelty) of the proposed research project inferred through the dynamic change of the scientific research landscape pertinent to this discipline.

In computing indicators, an initial step identifies from a corpus of grant application relevant scientometric (*e.g.*, publications, citations, patents) and content data (*e.g.* text-strings, keywords) bearing relevance to frontier research, extracts and subjects them to data mining. This is essentially a filtering step to pre-process raw data of high quantity into

input data of lower quantity but higher quality. In a subsequent step, actual indicators are automatically and robustly computed and subjected to a selection function for comparison between empirical and model parameter. Finally, cross-validation and iterative variation of thresholds, classification criteria, and metrics as well as expert feedback from panel members and chairs, refine in a last step the performance of the model to sufficiently high usability. The following sections describe indicators and selection function in more detail. For various reasons the comparison of proposals is only meaningful within one scientific discipline. Thus we obtain indicator values and apply the selection function for each discipline (panel) individually.

2.4 Indicators of frontier research

The indicator TIMELINESS illustrated in Figure 2 is used as a proxy to infer the “innovative degree” of the proposal through bibliographic references; somewhat similarly, it can be applied to the top-N listed self-citations of the author. The underlying hypothesis is that the more recent references are, the more likely the work is at the frontier of science. Therefore, the bibliometric concept focuses is on references and the time elapsed since the publication listed to validating a proposal, and TIMELINESS is obtained from the difference between the times of proposal submission referenced publications.

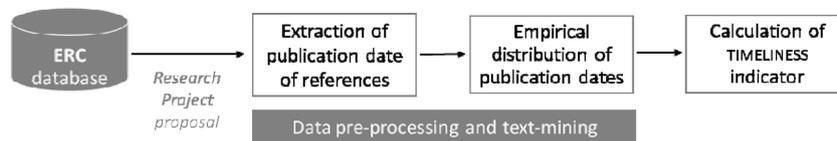


Figure 2 : The core bibliometric concept of the TIMELINESS indicator

While there are alternative data sources, references of the proposal are considered appropriate because they constitute published documentation of highest velocity and directly relate to the research project. After identifying references and extracting publication dates in actual texts, the indicator can be calculated from the set of values: The simplest way to obtain TIMELINESS is to use the arithmetic mean. In addition, because it may be influenced by statistical outliers, other statistics (*e.g.* mode or median of the distribution) are considered to quantify the distribution and re-sampling methods like bootstrapping are applied. Finally, it is tested whether TIMELINESS follows a known theoretical distribution (*e.g.* Poisson or negative binomial distribution).

The indicator PROXIMITY illustrated in Figure 3 is used to infer the “innovativeness” of a proposal. The more sophisticated bibliometric concept employs a “publication landscape” (represented by a cluster map of scientific resp. technological information and characterized by the level of change over time), in which it positions and assigns to each proposal a PROXIMITY value in dependence to the distance and ranking of nearest clusters. The underlying hypothesis is that the closer a proposal is to clusters of positive dynamic change, the more innovative it is.

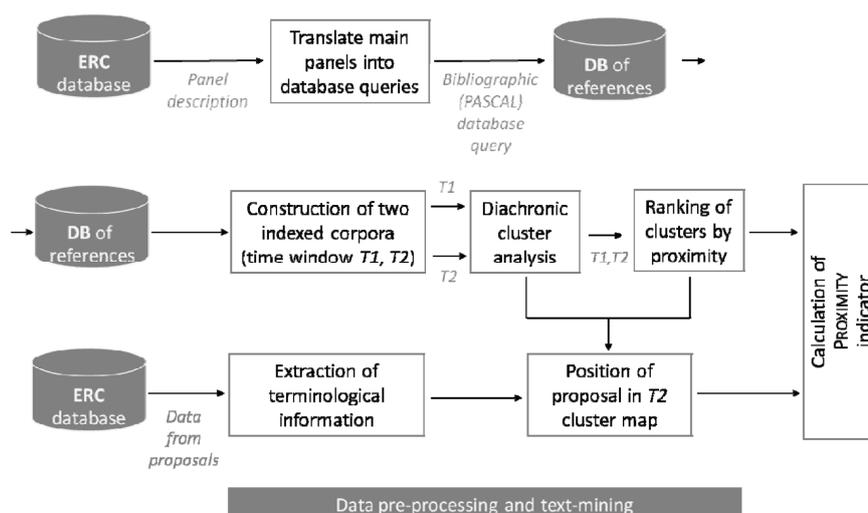


Figure 3 : The core bibliometric concept of the PROXIMITY indicator

To this end, raw data are obtained from external bibliographic reference databases (for international scientific and technological literature) and ERC panel descriptions to identify and extract discriminating terminological information. A clustering algorithm is employed to obtain a cluster map that groups similar references (through related keywords) and represents the publication landscape corresponding to the considered panel. A diachronic analysis [17] is used to study the evolution of the landscape across two time windows, by considering the content of each cluster and its relative location in the network of clusters.

Structural alterations of the network of clusters between the two time periods are identified and described: splitting or disappearing of clusters as well as persistence and emergence of clusters are investigated; as well cluster status changes (e.g. the evolution of a cluster the periphery toward a central position). In addition, techniques of association rule extraction are applied to complete the cluster evolution analysis [21]. The result is a ranking of clusters of the more recent map with respect to its dynamics and for each proposal the distance of a proposal to nearest clusters.

The indicator PASTEURSQUENESS illustrated in Figure 5 is used to infer the applicability of expected results of each proposal, by considering evidence for immediate or intended application. Input data are obtained from bibliographic and patent databases and proposals.

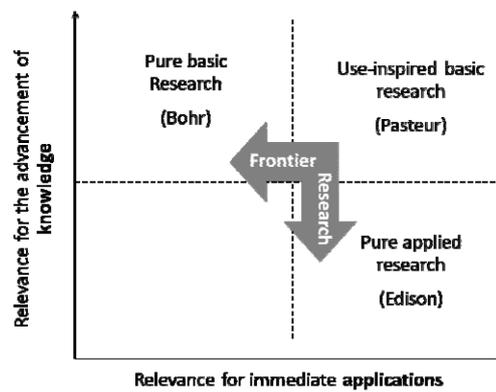


Figure 4 : Illustration of the Pasteur Quadrant

The term PASTEURESQUENESS originates from the definition of Pasteur's Quadrant [22], which describes scientific research/methods that seek both fundamental understanding and at the same time social benefit (*cf.* Figure 4).

PASTEURESQUENESS consists of in all three indicators: i) patents granted to the PI; ii) information related to business involvement; and iii) self-citations published in journals with prescribed 'applicability' (scoring function).

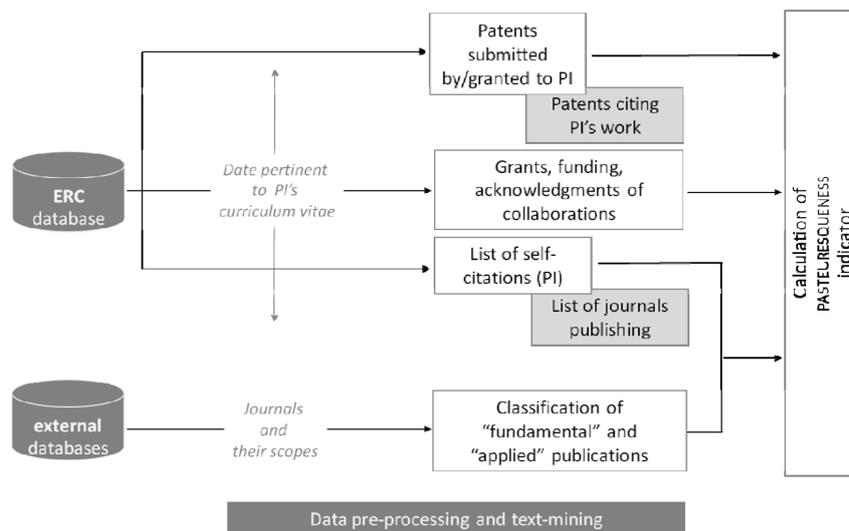


Figure 5 : The core bibliometric concept of the PASTEURESQUENESS indicator

The indicator RISK illustrated in Figure 6 is used to infer the "personal risk" of the PI in executing the proposed research under the assumption that references (of proposals, papers) serve as one informative source for constructing a 'bibliometric research profile'. In comparing the profile of a proposal to both past research by the PI and relevant research originating

inside the same field, observed overlaps are used to categorize anticipated constant resp. aligned, modified or changed research directions. On the one hand, one hypothesis is that the lower the overlap between two reference profiles (past vs. proposed research), the more risk-affine is the proposal for the PI, because it is indicative of a change from previous directions. On the other hand, another hypothesis is that the lower the overlap to all other '(field) profiles', the more unique is the proposed research. The computation of RISK can be refined by "exotic terms" (rare keywords) in textual information of the proposal; exotic in the sense that the terms are less frequently used than established terms in the same research area [17; 18].

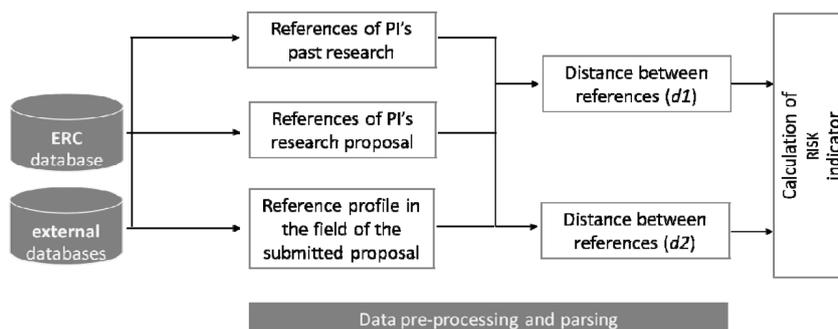


Figure 6 : The core bibliometric concept of the RISK indicator

The indicator INTERDISCIPLINARITY illustrated in Figure 7 is used to infer self-consistently the presence and proportions of characteristic terminology or textual traits, which "belong" to different ERC '(home) panels', and thereby reveal the intra or inter-panel/disciplinary character of each proposal. It is build upon the assumption and previously successfully tested concept [18] that the frequency of occurrence and distribution of discipline specific keywords in scientific documents can be used to classify and characterize disciplines.

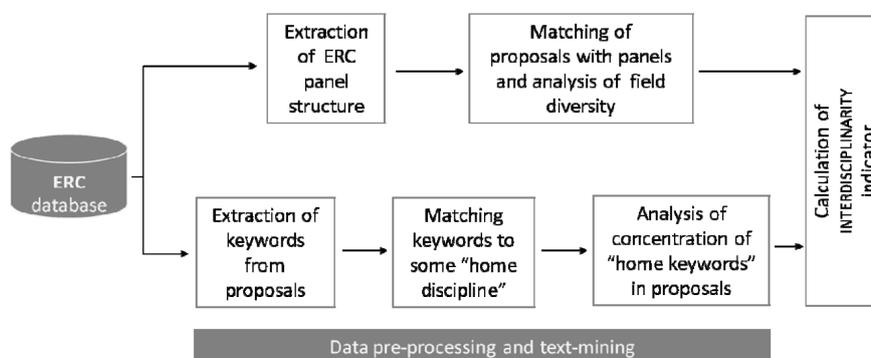


Figure 7 : The core bibliometric concept of the INTERDISCIPLINARITY indicator

To this end, each keyword is labeled according to its statistical frequency of occurrence across all PE or LS panels, filter are applied to distinguish relevant from irrelevant (*i.e.* panel unspecific) keywords, and the concentration of keywords from different home panels is assessed to classify proposals as very high, medium-high, low-medium, or very low in concentration. In addition to free keywords extracted from textual information of proposals, the analysis is supplemented by using fixed keywords from the description of ERC panels.

2.5 Modelling the selection function

The indicators described above are expected to have a positive effect on the decision probability of a grant application, which has in principle three possible outcomes: Type-A) above threshold and funded, Type-B) above threshold and not funded, and Type-C) below threshold. Yet as there are a number of other possible factors influencing the peer review process, a statistical analysis (discrete choice model) determining the actual association between indicators and the funding decisions is carried out separately for SG and AG applications.

A discrete choice model is used to estimate how various exogenous factors (captured by the indicators, *cf.* Table 3) influence the probability for a project proposal to get accepted. Discrete choice models have come into wide-spread use in econometrics to represent discrete alternatives rather than a continuous measure of activity [23]. In a first attempt, the selection model assumes a binary choice (types A/B vs. C) as they are the two central outcomes of the dependent variable: the rejection or acceptance of a project proposal. To this end, the dependent variable takes the form (empirical data on the $i = 1, 2, \dots, n$ proposals):

$$Y_i = \begin{cases} 1 & \text{proposal is accepted} \\ 0 & \text{otherwise.} \end{cases}$$

The decision probability of a proposal to be accepted, $\Pr(Y_i = 1)$, depends on a set of independent, exogenous factors (indicators) summarized in a matrix \mathbf{X} such that

$$\begin{aligned} \Pr(Y_i = 1) &= F(\mathbf{X}_i, \boldsymbol{\beta}) \\ \Pr(Y_i = 0) &= 1 - F(\mathbf{X}_i, \boldsymbol{\beta}) \end{aligned} \quad (2)$$

where \mathbf{X}_i is a set of k observed factors for proposal i , $\boldsymbol{\beta}$ is the estimated k -by-1 parameter vector reflecting the impact of changes in \mathbf{X}_i on the probability $\Pr(Y_i = 1)$. It is common practice to define F by the logistic distribution function, leading to binary Logit model:

$$\Pr(Y_i = 1) = \frac{\exp \mathbf{X}_i \boldsymbol{\beta}}{1 + \exp \mathbf{X}_i \boldsymbol{\beta}} \quad (3)$$

Technically, the parameter estimation is based on Maximum-Likelihood [*cf.* 23], and the interpretation in the context of the research question can be comfortably framed in terms of odds-ratios, which are given by $\exp(\boldsymbol{\beta})$ [*cf.* 24].

2. Discussion

The above concept aims at developing quantitative methods for determining and examining the relationship between peer review and decisions on grant allocation in terms of attributes of frontier research:

- Can attributes of frontier research be faithfully represented and validly quantified to evaluate the grant allocation decision by bibliometric approaches?

The detailed development has focused on the ERC grant scheme (with data in 2007-2009), but the concept might be applicable more generally, depending on mission, review process and guidelines, attributes and correspondence of indicators for grant schemes. The implemented concept is intended to yield a bibliometric model in which indicators are expected to have a positive effect on the decision probability for ERC grant applications. Thus a follow up and specific question is:

- How well do bibliometric indicators and the decision probability discriminate between grant applicants accepted resp. rejected for funding?

In particular, instead of relying on performance measures the concept builds on a combination of content analysis and data-mining in combination with scientometric methods. The concept presents a new approach in that it explicitly utilizes information present in research proposals submitted to a grant agency and relates it to the bulk of information drawn from activities of the larger research community in a specific field. To this end, textual information (e.g. keywords and longer strings extracted from proposals) as well as references (e.g. papers or patents) are combined using, e.g., human expert systems, clustering techniques, relational mapping, networks, bibliometric functions, various indices and data filtering techniques. Note that this approach makes consistent use of one initial core source of information, namely data by the ERC (e.g., on panels, proposals). Subsequently, additional sources are employed to enrich the core: 1) the first supplementary data basis to add to the ERC panel data through the construction of a scientific landscape (e.g., PASCAL, MEDLINE, or INSPEC); and 2) the second supplementary data basis to add to ERC proposal data (e.g., Web of Science, or patent databases, or SCOPUS).

Described indicators and selection function are currently implemented. First ex post analyses between model and review process can be expected a mixture of similarity between peer reviewer selection (i.e. Types-A/B) and dissimilarity (i.e. type C); cf. section Modelling the selection function. Depending on the found correlation between the discrimination of types A/B vs. C obtained from the bibliometric model and selection by peer review, numerical algorithms for the computation of indicators might need refinement, resulting in a modification of the model, or improvements of the future operation of the peer review process can be envisioned. In any case, it requires careful investigation (statistical independence, positive vs. negative correlation, outliers), explicit differentiation between measurement concept and interpretation, and careful feedback on parameters on a discipline-specific basis [25]. In order to serve its purpose, the development and refinement will be hand-in-hand with experts involved in the review process to determine what the metrics could be used for and how it is affecting the review process.

Ultimately the concept shall result in a methodology that allows the grant agency to monitor the operation of the peer review process from a bibliometric perspective and thereby provide a basis for the further refinement of the peer review process, including the *ex ante*

bibliometric evaluation of future grant applications to support reviewers with orientation knowledge for the review assessment.

Acknowledgments

The authors acknowledge the support that this work was partially funded by the Ideas specific programme of the EU's FP7 Framework Programme for Research and Technological Development (project reference no. 240765). The authors thank Helga Nowotny and Jens Hemmelskamp of the ERC for stimulating discussions, and Manuela Kienegger for time and energy in drafting the manuscript.

References

- [1] L. Bornmann, Scientific peer review. *Annual Review of Information Science and Technology*. **45**, p. 199-245 (2011).
- [2] K. Powell, Making the Cut. *Nature*. **467**, p. 383-385 (2010).
- [3] M. Hojat, J.S. Gonnella, A.S. Caellegh, Impartial judgment by the "gatekeepers" of science: Fallibility and accountability in the peer review process. *Advances in Health Sciences Education: Theory and Practice*. **8(1)**, p. 75-96 (2003).
- [4] Sweizer, Collen (1994).
- [5] L. Bornmann, Daniel (2008).
- [6] H.W. Marsh, U.W. Jayasinghe, N.W. Bond, Improving the peer-review process for grant applications: Reliability, validity, bias, and generalizability. *American Psychologist*. **63(3)**, p. 160-168 (2008).
- [7] P. van den Besselaar, L. Leydesdorff, Past performance, peer review, and project selection: A case study in the social and behavioral sciences. *Research Evaluation*. **18(4)**, p. 273-288 (2009).
- [8] R.v. Norden, A profusion of measures. *Nature*. **465**, p. 864-865 (2010).
- [9] P. Juznic, S. Peclin, M. Zaucer, T. Mandelj, M. Pusnik, F. Demsar, Scientometric indicators: peer-review, bibliometric methods and conflict of interest. *Scientometrics*. **85(2)**, p. 429-441 (2010).
- [10] D. Ariely, You are what you measure. *Harvard Business Review*. **6**, p. 38 (2010).
- [11] L. Bornmann, L. Leydesdorff, P. van den Besselaar, A Meta-evaluation of Scientific Research Proposals: Different Ways of Comparing Rejected to Awarded Applications, *Journal of Informetrics*. **4(3)**, p. 211-220 (2010).
- [12] M. Antonoyiannakis, J. Hemmelskamp, F.C. Kafatos, The European Research Council Takes Flight. *Cell*. **136(5)**, p. 805-809 (2009).
- [13] EC – European Commission. Frontier research: The European Challenge. High Level Expert Group Report, EUR 21619. (2005).
- [14] ERC – European Research Council. ERC Work Programme 2009. <http://rp7.ffg.at/upload/medialibrary/l_wp_200901_en.pdf> (2008) consulted on June 18, 2012.
- [15] B. Yoon, S. Lee, G. Lee, Development and application of a keyword-based knowledge map for effective R&D planning. *Scientometrics*. **85(3)**, p. 803-820 (2010).
- [16] ERC – European Research Council. Facts & Figures. Retrieved January 7, 2011 from: <<http://erc.europa.eu/index.cfm?fuseaction=page.display&topicID=165>> (2011).

- [17] I. Roche, D. Besagni, C. Francois, M. Hörlesberger, E. Schiebel, Identification and characterisation of technological topics in the field of Molecular Biology. *Scientometrics*. **82(3)**, p. 663-676 (2010).
- [18] E. Schiebel, M. Hörlesberger, I. Roche, C. Francois, D. Besagni, An advanced diffusion model to identify emergent research issues: The case of optoelectronic devices. *Scientometrics*. **83(3)**, p. 765-781 (2010).
- [19] A.J. Nederhof, Bibliometric monitoring of research performance in the Social Sciences and the Humanities: A review. *Scientometrics*. **66(1)**, p. 81–100 (2006).
- [20] ERC – European Research Council, ERC Grant Schemes Guide for Peer Reviewers Applicable to the ERC. Starting Grants and Advanced Grants (Work-Programme 2011). Updated September 2010. Retrieved January 7, 2011 from http://erc.europa.eu/pdf/Guide_ERC_Peer_Reviewers_2011.pdf (2010).
- [21] H. Mahgoub, D. Rösner, N. Ismail, F. Torkey, A text mining technique using association rules extraction. *International Journal of Computational Intelligence*. **4(1)**, p. 21-28 (2008).
- [22] D.E. Stokes, *Pasteur's Quadrant*. Basic Science and Technological Innovation. **Washington (DC), The Brookings Institution Press** (1997).
- [23] W.H. Greene, *Econometric analysis*. 5th ed. **Prentice Hall, Upper Saddle River (NJ)** (2003).
- [24] J. Johnston, J. Dinardo, *Econometric methods*. 4th ed. **New York/London, McGraw-Hill/Irwin** (1997).
- [25] M Haindl, Scientometry leading us astray. *ERICIM News*. **82**, p. 8 (2010).

L'Observatoire virtuel astronomique : interopérabilité et standards au service de la recherche en astronomie

Thomas Boch

*Centre de Données astronomiques de Strasbourg,
Observatoire astronomique, UMR 7550
11 rue de l'Université, 67000 Strasbourg*

Résumé

L'Observatoire virtuel (OV) astronomique est un projet international dont le but est de développer un nouvel outil de recherche permettant un accès transparent et homogène aux énormes ensembles de données réparties à travers le monde. Le développement de l'OV est piloté par l'*International Virtual Observatory Alliance* (IVOA), union des différents projets OV nationaux, chargé de faciliter la coordination et les collaborations internationales nécessaires à la production des standards d'interopérabilité, piliers de l'OV. Après un résumé de l'organisation de l'IVOA, je présenterai l'architecture de l'OV et les standards sous-jacents. Je détaillerai ensuite les développements effectués dans le cadre du projet européen Euro-VO.

Abstract

The astronomical Virtual Observatory (VO) is an international project aiming at supporting astronomical research through the development of a new tool enabling a transparent and homogeneous access to the large amount of datasets distributed across the world. VO development is driven by the IVOA (*International Virtual Observatory Alliance*), an alliance of the different national VO projects. IVOA mission is to facilitate the international coordination and collaboration necessary for the creation of interoperability standards, which are the pillars of the VO. After summarizing how the IVOA is organized, I will first describe the VO architecture and the standards it relies on. Then, I will give an overview of the developments performed in the frame of the European-funded project Euro-VO.

1. Introduction : le contexte de la recherche en astronomie/astrophysique

1.1 Les étapes de la recherche en astrophysique

De façon schématique, le processus de recherche en astrophysique basé sur la collecte des données passe par les étapes suivantes :

- 1- **Acquisition** par des télescopes et instruments dédiés ;
- 2- **Analyse et réduction** ;
- 3- **Mise à disposition des données réduites** dans des centres de données ;
- 4- Ces centres effectuent un travail de **data curation** : ajout et homogénéisation des métadonnées, sélection et agrégation de contenus, etc. ;
- 5- Travail de recherche à partir des données réduites et **publication**.

1.2 Avalanche de données

Les progrès dans l'instrumentation (utilisation des capteurs CCD¹ depuis les années 1980, augmentation du nombre de détecteurs et de leur résolution) ont conduit l'astrophysique à connaître une véritable avalanche de données.

À titre d'exemple, le relevé SDSS (*Sloan Digital Sky Survey*), démarré en 2000, a produit plus de 15 téra-octets de données (environ 3 000 DVD). Dans le futur proche, le projet LSST (*Large Synoptic Sky Survey*) produira 5 péta-octets par année, c'est-à-dire 14 téra-octets par nuit. Le réseau d'antennes SKA (*Square Kilometer Array*) générera lui 20 péta-octets de données par année de fonctionnement.

De tels volumes ne sont pas aisément transférables via le réseau : en supposant un débit moyen de 10 méga-octets par seconde, 18 jours sont nécessaires pour télécharger 15 téra-octets.

Le phénomène d'avalanche de données est également amplifié par la multiplication du nombre d'observatoires, au sol ou spatiaux (satellites). La figure ci-dessous donne un aperçu non exhaustif des observatoires actuellement en opération.



Figure 1 : Observatoires au sol et spatiaux

¹ Charge-Coupled Device.

1.3 Approche multi-longueur d'ondes

L'astrophysique moderne se base sur des observations effectuées dans plusieurs domaines de longueur d'onde (voir figure 2). Cette approche est indispensable à la compréhension globale des mécanismes physiques sous-jacents.

2. Concept d'Observatoire virtuel (OV)

Des spécificités évoquées dans la section précédente a émergé le concept d'*Observatoire virtuel astronomique* (OV).

Il s'agit de rendre accessibles, depuis un simple ordinateur, l'ensemble des archives et bases de données astronomiques. L'Observatoire virtuel n'est pas une base de données centralisée ni un système *peer-to-peer*. Il faut le considérer comme un *nouvel outil pour la science*², les données d'archive constituant un ciel virtuel auquel les astronomes accèdent de façon transparente. Cet outil est basé sur un ensemble de centres de données interopérables.

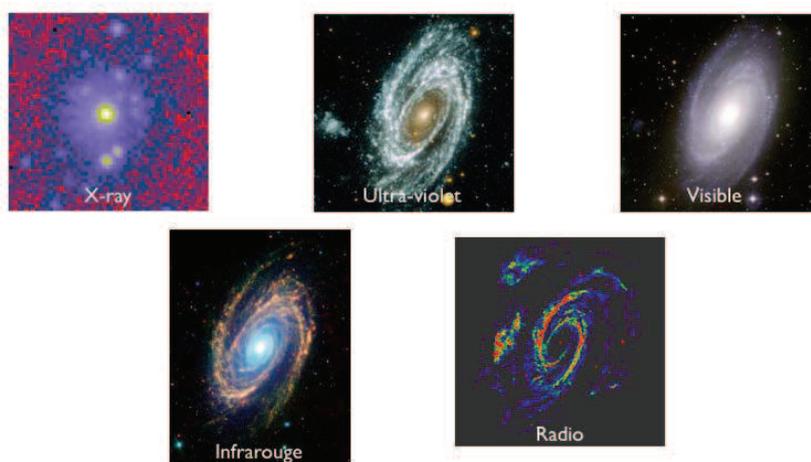


Figure 2 : La galaxie spirale Messier 81, observée dans différentes longueurs d'ondes

2.1 L'astronomie, un terrain favorable

La mise en place de l'Observatoire virtuel astronomique bénéficie d'un terrain favorable.

En effet, les données astronomiques sont le plus souvent librement accessibles après une période d'embargo relativement courte (généralement de 6 à 12 mois). Par ailleurs, des standards de données et d'échanges sont établis de longue date (le standard FITS – *Flexible Image Transport System* – existe depuis plus de 30 ans), et les centres de données ont une

² La page <<http://www.euro-vo.org/pub/fc/papers.html>> liste les publications qui ont utilisé l'OV, consultée le 21 mai 2012.

longue tradition de collaboration, notamment pour la mise en place de liens entre données et références bibliographiques.

3. IVOA : *International Virtual Observatory Alliance*

L'*International Virtual Observatory Alliance*³ est une alliance des projets d'OV nationaux. Créé en 2002, il a pour mission

« de faciliter les coordinations et collaborations nécessaires au développement et déploiement des outils et systèmes qui permettront une utilisation internationale des données d'archives au travers d'un observatoire virtuel intégré et interopérable ».

À ce jour, l'IVOA comporte 19 membres (16 projets nationaux, 2 agences internationales, et le projet européen Euro-VO).



Figure 3 : Les projets membres de l'IVOA

3.1 Fonctionnement de l'IVOA

L'IVOA est organisé en différents groupes de travail (*working groups*) consacrés chacun à un aspect particulier de l'interopérabilité.

Des réunions de travail ont lieu tous les 6 mois, et entre ces réunions, les discussions et échanges se font par le biais de listes de diffusion dédiées à chaque groupe de travail.

Le mode de fonctionnement est calqué sur celui du W3C⁴. La procédure de standardisation passe par les étapes suivantes : *note*, *draft*, *proposed recommendation*, RFC, etc.

³ <<http://ivoa.net/>> consulté le 21 mai 2012.

⁴ World Wide Web Consortium <<http://www.w3.org/>> consulté le 21 mai 2012.

3.2 Architecture

La figure ci-après résume l'architecture générale de l'Observatoire virtuel, au niveau le plus basique. Les utilisateurs (en haut de la figure) accèdent de façon homogène et transparente à des ressources (en bas) en utilisant les standards de l'OV.

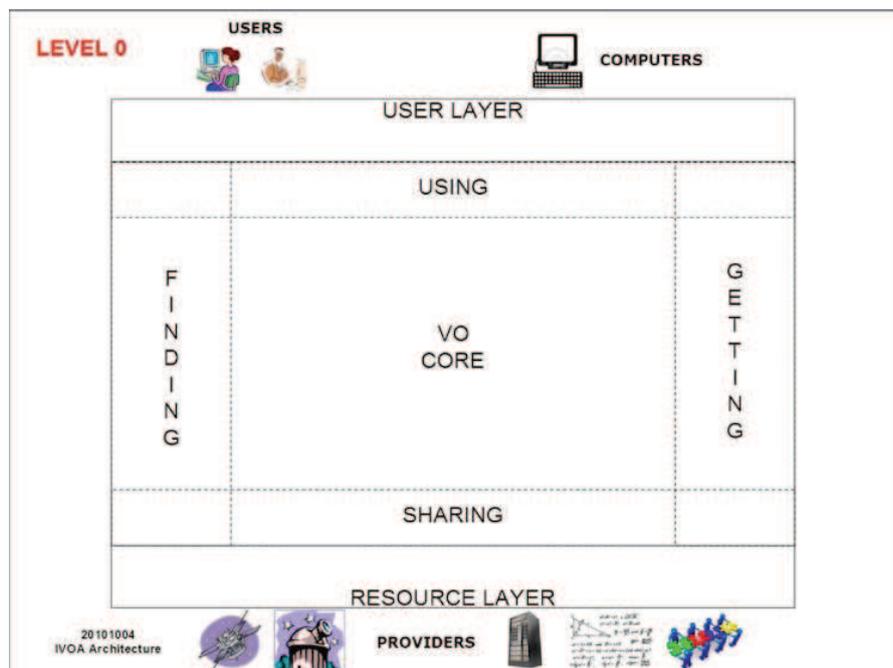


Figure 4 : Une vision schématique de l'architecture de l'OV

3.3 Groupes de travail

Dans cette section, je vais détailler le travail réalisé dans certains groupes de travail.

3.3.1 Registry

Les *registries* constituent l'annuaire, le système de pages jaunes de l'Observatoire virtuel. Leur rôle est de lister les services disponibles au sein de l'OV.

Un *registry* permet :

- de localiser les ressources (services, données) ;
- d'obtenir des informations sur ces services (métadonnées, types de données, domaine de longueurs d'ondes, etc) ;
- de savoir comment interroger ces services.

Le *registry* IVOA se caractérise par une architecture distribuée, composée de différents *registries* (cf. figure 5) qui se moissonnent les uns les autres selon le protocole OAI-PMH⁵

⁵ Open Archives Initiative's Protocol for Metadata Harvesting.

popularisé par l'*Open Archive Initiative*. Les ressources sont décrites selon un schéma reprenant les éléments du *Dublin Core* et complétées par des extensions spécifiques au domaine de l'astronomie.



Figure 5 : Carte des registries en fonctionnement

3.3.2 Data Access Layer

Le groupe Data Access Layer définit les standards d'accès aux données. Ces standards reposent généralement sur le protocole HTTP⁶ pour la couche de transport.

- le standard **Conc Search** décrit la façon d'accéder aux données tabulaires sur la base d'une interrogation par position ;
- les standards **SIAP** (*Simple Image Access Protocol*) et **SSAP** (*Simple Spectrum Access Protocol*) reposent sur des requêtes en deux temps :
 - on récupère d'abord une liste de données correspondant aux contraintes posées dans l'interrogation (e.g. : *liste des images autour de Messier 81*) ;
 - on sélectionne ensuite une ou plusieurs données d'intérêts et on les récupère ;
- le standard **TAP** (*Table Access Protocol*) définit un protocole complexe d'accès aux tables. Il se base sur un langage de requêtes dérivé de SQL⁷.

3.3.3 Data Model

Le groupe de travail *Data Model* est chargé de modéliser les métadonnées permettant la description des données observées ou simulées.

Ces modèles de données sont ensuite repris et utilisés par les autres groupes de travail afin d'être intégrés aux standards qu'ils développent.

À titre d'exemple, on peut citer le modèle STC (*Space Time Coordinates*) qui fournit un cadre pour décrire les coordonnées spatiales et temporelles.

⁶ *Hypertext Transfer Protocol*.

⁷ *Structured Query Language*.

3.3.4 Sémantique

Le groupe *Sémantique* a pour charge de standardiser les vocabulaires utilisés dans le domaine de l'astronomie, et de mener un travail exploratoire sur l'utilisation des ontologies afin de décrire les données astronomiques et leurs relations entre elles.

Il travaille notamment sur la construction d'un vocabulaire contrôlé décrivant les quantités physiques que l'on trouve dans une table astronomique, sur un document standardisant la description des unités ou encore sur une ontologie des types d'objets astronomiques. Une ontologie est une représentation formelle des connaissances dans un domaine donné. La figure 6 montre la branche *Infrared source* de l'ontologie des types d'objets.

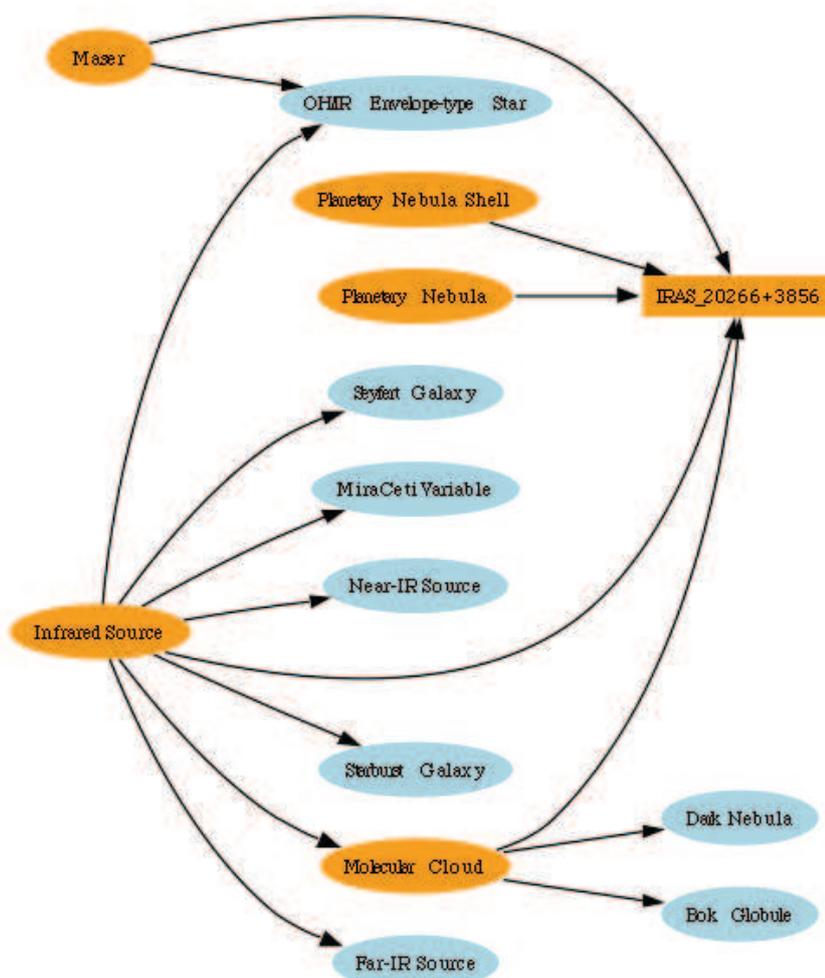


Figure 6 : Extrait de l'ontologie des types d'objets

4. Projet Euro-VO

Le consortium Euro-VO est constitué des membres suivants :

- 2 agences européennes :
- ESA (European Spatial Agency) ;
- ESO (European Southern Observatory) ;
- 6 projets nationaux :
- CNRS, France (CDS, OV France) ;
- INAF, Italie (Trieste, VObs.it) ;
- INTA, Espagne (SVO) ;
- Université de Groningue, Pays-Bas (NOVA) ;
- Université d'Édimbourg, UK (AstroGrid) ;
- Université d'Heidelberg, Allemagne (ARI, GAVO).

4.1 Projet AVO

La phase A d'Euro-VO a démarré par le projet AVO (Astrophysical Virtual Observatory). Ce projet d'une durée de 3 ans et d'un budget de 5 millions d'euros, financé à 50 % par l'Union européenne (FP5) et à 50 % par les partenaires, avait pour objectif d'établir une *preuve de concept* de l'Observatoire virtuel.

Le projet AVO était ponctué par une démonstration annuelle et a abouti au premier article dans une revue à comité de lecture utilisant l'OV : P. Padovani, M.G. Allen, P. Rosati, N.A. Walton, "Discovery of optically faint obscured quasars with Virtual Observatory tools", *Astronomy and Astrophysics*, 424(2), 2004, p. 545-549.

4.2 VOTech, Euro-VO DCA...

Dans le programme cadre FP6, les propositions de projets ont été découpées en plusieurs sous-projets afin de saisir différentes opportunités. Cette stratégie a abouti au financement du *Design Study VOTech* et de l'action de coordination *Euro-VO DCA*.

4.2.1 VOTech

Le projet VOTech s'est organisé en différents "*Design Study*" :

- 1 Architecture ;
- 2 Nouveaux outils ;
- 3 Découverte intelligente de ressources ;
- 4 *Data mining* et visualisation.

Des *Technology forums*, points d'avancement et de rencontre des différents acteurs, ont ponctué le déroulement du projet.

4.2.2 Euro-VO AIDA

Le projet AIDA, continuation de VOTech, a pour objectif d'assurer la transition de l'OV européen du prototypage vers la mise en production. Il doit permettre le déploiement à grande échelle des standards de l'OV par les centres de données. Il doit également créer une communauté de scientifiques utilisateurs. Enfin, il comporte un volet d'actions techniques (évolution des standards d'interopérabilité, veille technologique) et un volet d'actions de dissémination (*public outreach*) pour le grand public et les étudiants.

Support aux utilisateurs – Le support aux utilisateurs est assuré *via* les actions suivantes :

- des appels d'offre *Science Initiatives* permettent aux projets sélectionnés de bénéficier d'une aide technique sur les aspects Observatoire virtuel ;
- l'organisation de tutoriaux à destination des étudiants en thèse et des post-doctorants ;
- l'organisation de workshops à destination des centres de données « Comment publier vos données dans l'OV » ;
- la mise à disposition de tutoriaux consultables en ligne.

5. OV France

OV France est une action spécifique CNRS/INSU. Pilotée par le Centre de données astronomiques de Strasbourg (CDS), elle a pour buts de coordonner la participation française, de diffuser les techniques et méthodes de l'OV, et constitue le point de contact national pour les projets européens.

6. Le CDS dans l'OV

La participation du CDS à l'OV est multiple. Il est fortement impliqué dans OV-France, et participe à l'ensemble des projets Euro-VO. Au niveau de l'IVOA, il contribue à l'élaboration de standards, et plusieurs groupes de travail sont dirigés ou codirigés par des membres de l'équipe CDS.

Le CDS organise également des tutoriaux OV pour la communauté française. Enfin, il implémente les standards de l'Observatoire virtuel dans ses différents services, notamment les interfaces d'accès aux données.

Ainsi Aladin, visualiseur du ciel interactif développé au CDS, est un des outils phares de l'OV (figure 7).

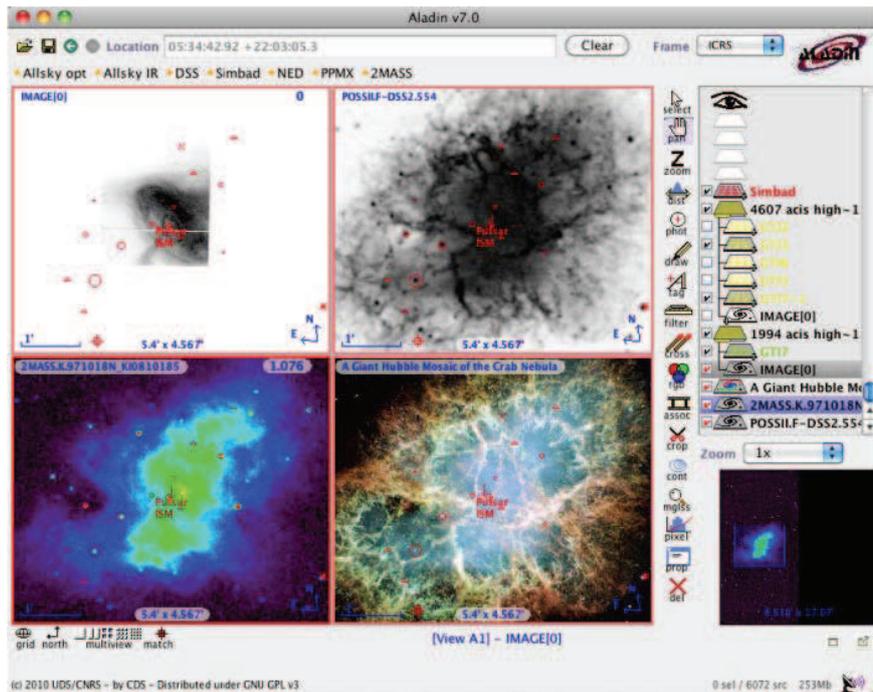


Figure 7 : Copie d'écran du logiciel Aladin

7. Conclusion et perspectives

En s'appuyant sur des collaborations existantes, une communauté réceptive aux problématiques d'interopérabilité et d'accès aux données, et en réutilisant des solutions techniques existantes (e.g. OAI-PMH, Dublin Core, SKOS⁸) lorsque cela était possible, l'Observatoire virtuel astronomique est passé en 10 ans d'un statut de prototype à un outil désormais en production.

Les standards étant désormais suffisamment complets et robustes, le prochain défi est de faire des outils de l'Observatoire virtuel des outils incontournables, connus de tous et utilisés quotidiennement par les chercheurs en astrophysique. Plusieurs actions sont menées dans ce sens :

- d'une part, l'IVOA est en contact étroit avec les grands projets observationnels afin qu'ils rendent leurs données disponibles à travers des standards de l'OV ;
- d'autre part, des actions de diffusion et de formation sont menées à destination des étudiants en thèse et des doctorants, afin de les sensibiliser à l'OV, et aux solutions qu'il propose pour appréhender et gérer le tsunami de données que connaît l'astrophysique moderne.

⁸ Simple Knowledge Organization System.

Le TGE Adonis et les projets européens

Jean-Luc Pinol

*Professeur à l'ENS de Lyon, directeur du TGE Adonis
CNRS UPS 2916 SupADONIS, 4 rue Lhomond, 75005 Paris, France*

Les projets européens auxquels participent le TGE¹ Adonis, en particulier DARIAH² s'appuient sur les réalisations des différentes infrastructures concernées. C'est pourquoi je commencerai par présenter la situation actuelle du TGE aujourd'hui avant de préciser sa participation à différentes initiatives européennes.

1. Le TGE Adonis, au cœur d'un réseau

Le TGE fonctionne en réseau avec de nombreux acteurs du numérique, dans le domaine des SHS, mais aussi avec des infrastructures numériques qui œuvrent pour les autres sciences. Dans le domaine des sciences humaines et sociales (SHS), il agit aussi avec de nombreux réseaux de métiers comme celui des documentalistes et des bibliothécaires. Il opère aussi avec des porteurs de projets de recherche, c'est-à-dire l'ensemble des chercheurs et des enseignants-chercheurs qui peuvent avoir besoin de ses services. C'est la condition même de son bon fonctionnement que d'aller au-devant des utilisateurs pour avoir des retours d'expériences qui permettent d'améliorer les services rendus. Par exemple, nous allons dans les maisons des sciences de l'homme afin de connaître les besoins et les attentes de tel ou tel laboratoire, de tel ou tel réseau disciplinaire. Nous entretenons des liens avec des producteurs de ressources comme le CLEO³ (centre pour l'édition électronique ouverte) ou Persée. Sans ces liens étroits avec les producteurs, la plateforme de recherche ISIDORE⁴ ne pourrait pas jouer son rôle.

Mais le TGE Adonis, s'il est au cœur des SHS, collabore aussi avec de nombreux acteurs du numérique qui sont étrangers aux SHS ou dont le domaine ne se limite pas à elles. Qu'il s'agisse de l'Agence bibliographique pour l'enseignement supérieur (ABES) et du système universitaire de documentation⁵, ou du Centre pour la communication scientifique directe (CCSD)⁶ que le monde des SHS connaît surtout pour HAL-SHS, ce sont là des partenaires quotidiens du TGE. Le TGE ne pourrait pas être ce qu'il est sans des organismes comme le

¹ Très grand équipement, <<http://www.tge-adonis.fr>>.

² *Digital Research Infrastructure for the Arts and Humanities*.

³ <<http://cleo.cnrs.fr/>>.

⁴ <<http://www.rechercheisidore.fr/>>.

⁵ Voir <<http://www.abes.fr/>> et les systèmes universitaires de Documentation, le SUDOC <<http://www.sudoc.abes.fr/>> consultés le 5 mars 2012.

⁶ <<http://www.ccsd.cnrs.fr/>> consulté le 5 mars 2012.

CINES⁷ ou le centre de calcul de l'Institut de physique nucléaire et de physique des particules (IN2P3)⁸, un des plus vastes d'Europe, où sont installés ses serveurs. Ainsi des liens avec des organismes internationaux et un environnement impliquant des institutions nationales qui sont des acteurs majeurs du numérique sont des conditions requises pour que se développent les recherches des SHS.

Le ministère de l'Enseignement supérieur et de la Recherche et le CNRS ont fixé pour mission au TGE Adonis de constituer un espace de navigation unifié pour les principaux documents numériques multimédias en usage dans la recherche en SHS. Il importe surtout pour les chercheurs que les projets de recherche, tout en poursuivant leurs objectifs scientifiques propres, mettent en œuvre les bonnes pratiques et respectent leurs standards d'interopérabilité qui seuls autoriseront la réutilisation des données et leur interopérabilité. Ce dernier point est fondamental puisqu'il est l'un des principes de base de la plateforme ISIDORE, créé par le TGE (*cf. infra*). Si les formats ne correspondent pas aux standards internationaux, les ressources numériques constituées peuvent être d'un grand intérêt, mais elles ne pourront pas être facilement communiquées, voire réutilisées – dans le respect de la propriété intellectuelle, ce qui est une autre question. Le respect des bonnes pratiques en matière de numérique est aussi un gage de visibilité pour les ressources élaborées. Il est important de pouvoir passer aisément des ressources constituées pour une recherche à l'exposition des résultats, voire des ressources, sur un site Web. Si les bonnes pratiques n'ont pas été respectées, cette étape devient complexe, et une nouvelle saisie des données sera même parfois nécessaire. Ce sont ces écueils auxquels il faut prendre garde. Avant de se lancer dans la mise en œuvre d'un projet, il est nécessaire de faire un état de l'art sur des questions similaires afin d'éviter les impasses où l'on risquerait de s'engager.

Pour ce faire, le TGE organise des formations qui permettent justement de faire connaître l'état de l'art des méthodes de numérisation des données. Des formations *ad hoc* sont aussi dispensées. C'est le cas par exemple pour acquérir les principes de base de la TEI (*Text Encoding Initiative*), une norme pour l'encodage des textes. Lou Burnard, de l'Université d'Oxford, a été missionné afin de développer une initiative de « Mutualisation d'expériences d'encodage textuel » (MEET). On trouve sur son site les offres de formation et les recherches qui utilisent cette norme⁹.

Le temps du chercheur isolé est révolu ; il est indispensable de s'intégrer dans des réseaux et de partager ses expériences avec d'autres tant sur le plan de la recherche que sur celui de l'ingénierie. Il est aussi nécessaire de disposer de services qui permettent d'assurer la sauvegarde et la pérennisation des recherches réalisées. C'est le rôle de la grille de services du TGE¹⁰.

⁷ Voir le site du Centre informatique national de l'enseignement supérieur <<http://www.cines.fr/>> consulté le 5 mars 2012.

⁸ <<http://cc.in2p3.fr/>> consulté le 5 mars 2012.

⁹ Voir <<http://meet.tge-adonis.fr/>> consulté le 5 mars 2012.

¹⁰ Les services du TGE ne sont pas réservés aux unités CNRS, ils sont ouverts à l'ensemble des équipes de recherche. Les exemples que j'ai pris sont développés dans les universités. La conférence des Présidents d'université est par ailleurs représentée au comité de pilotage du TGE Adonis.

2. Une grille

Une fois que les données des équipes de recherche sont constituées, classées et numérisées, ces dernières peuvent utiliser la grille de service du TGE. Une grille (de l'anglais *grid*) est une infrastructure qui agrège plusieurs machines, ce qui permet d'obtenir une puissance de calcul et des espaces de stockage de données considérables¹¹. Les serveurs du TGE sont installés au centre de calcul de l'IN2P3 du CNRS à Villeurbanne.

Cette grille offre aux équipes les possibilités suivantes :

- 1- un hébergement Web selon différents systèmes (PHP, MySQL, TOMCAT, Python) pour les projets de recherche scientifiques. Cet hébergement propose une large gamme de solutions privilégiant les technologies libres et ouvertes. Il s'adresse aux projets de recherche soutenus par une institution publique de recherche ;
- 2- un stockage intermédiaire sécurisé afin de mettre en sécurité, hors de leurs murs, des données produites dans le cadre de projets de recherche ;
- 3- un service de calcul scientifique sur des masses de données utilisant la ferme de calcul du centre de calcul de l'IN2P3 (modélisation, calcul scientifique pour les statistiques, rendu d'images 3D, etc.) ;
- 4- un accompagnement dans l'archivage à long terme des données de la recherche, et ce sur le modèle international de l'*Open Archival Information System* (OAIS). L'archivage est opéré par le CINES. Dans cette opération, le CINES et le TGE travaillent avec le SIAF, Service interministériel des archives de France, qui vérifie périodiquement la qualité technique de l'archivage.

Une partie des fonctionnalités de la grille et, en particulier, tout ce qui concerne la sauvegarde et le stockage hors les murs de bases de données en cours de constitution, a été présentée dans le numéro 10 de la lettre de l'INSHS¹² dans le cadre de la tribune d'Adonis, et plus spécifiquement dans la rubrique une question/une réponse qui répond par exemple à la question « Mes données sont-elles bien conservées ? ».

Le calcul scientifique est, par exemple, utilisé par des géographes qui veulent recalculer toute une série de photos satellites ou par des utilisateurs de reconstitutions en 3D, qui nécessitent des moyens de calculs considérables. Et nous avons vu que des historiens modernistes ou contemporanéistes pouvaient utiliser ce type de reconstitutions.

Pour ce qui est de l'archivage à long terme, le but est de préserver des fichiers et de maintenir les logiciels permettant de les utiliser. Les collègues qui font de l'histoire orale peuvent avoir des besoins importants dans ce domaine. Néanmoins, parmi les utilisateurs réguliers de la grille de service, les historiennes et les historiens des périodes moderne et contemporaine ne sont pas les plus nombreux, si l'on compare leur implication à celles des médiévistes ou des archéologues.

¹¹ Cf. <<http://www.france-grilles.fr/spip.php?rubrique14>> consulté le 5 mars 2012. Le TGE participe à l'Institut des grilles et à France Grilles.

¹² <http://www.cnrs.fr/inshs/Lettres-information-INSHS/lettre_infoINSHS_10.pdf> consulté le 5 mars 2012.

3. La plate-forme de recherche ISIDORE

L'objectif de la plate-forme ISIDORE est de permettre un accès unifié aux entrepôts de ressources numériques dévolus à la recherche. La plateforme a été lancée le 4 avril 2011. La situation avant que n'existe un accès unifié peut être schématisée de la manière suivante :

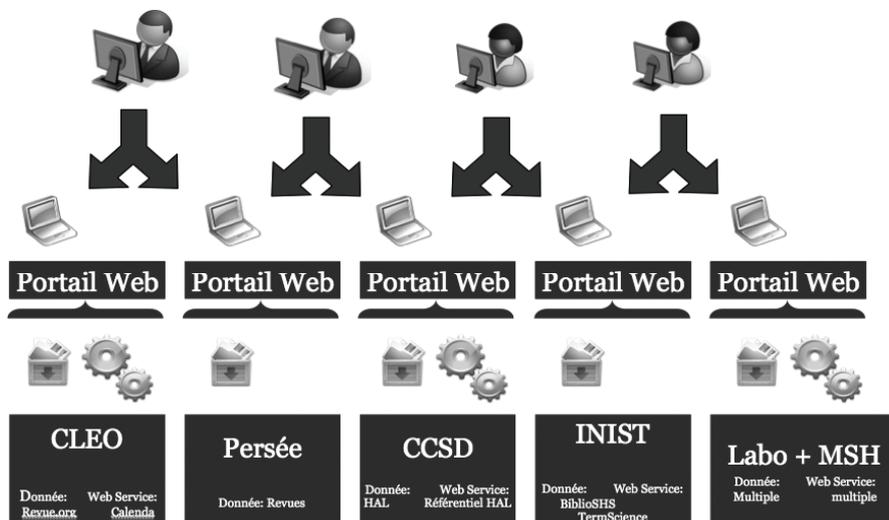
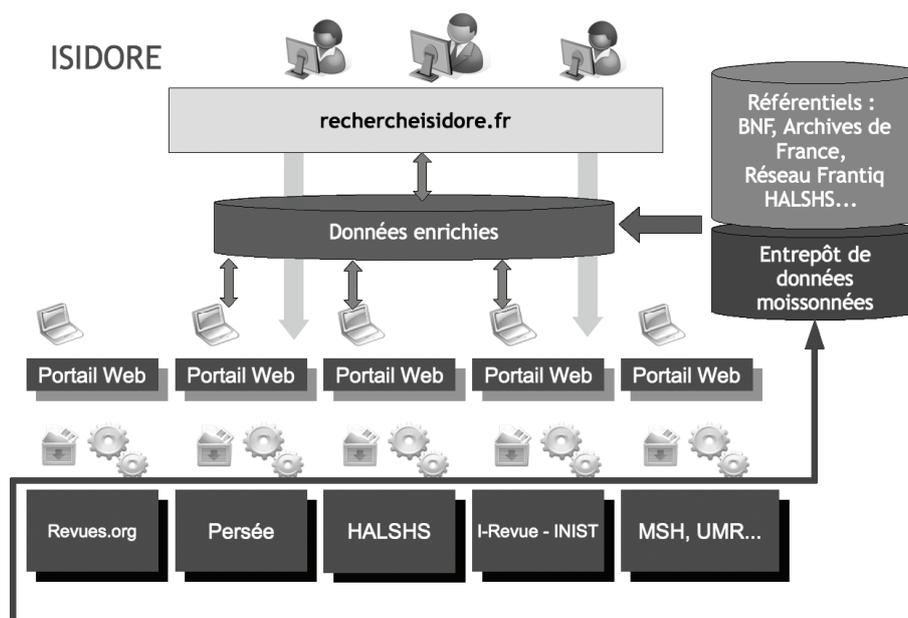


Schéma n°1 : La situation des entrepôts de ressources numériques avant l'accès unifié proposé par ISIDORE

Chaque entrepôt correspond à une ou plusieurs bases de données, qui peuvent être successivement interrogées par les chercheurs. Avec ISIDORE, on pose la question une seule fois. Cela paraît simple, mais le processus implique de nombreux préalables qui supposent le respect, par ces entrepôts et ces bases de données, des bonnes pratiques dans la structuration des données et des métadonnées par exemple.



Le schéma n° 2 : Les principes du fonctionnement d'ISIDORE

Les données collectées, ou moissonnées (*harvested* en anglais) dans les différents entrepôts sont regroupées en un seul, fondé sur un standard ouvert (RDF, *Resource Description Framework*). Elles vont être enrichies, c'est-à-dire indexées avec des termes issus de plusieurs référentiels. Cela va permettre de déduire de nouveaux concepts à partir des mots clefs, des mots du titre, du texte intégral, des résumés qui sont déjà présents dans les données, comme le montre le schéma n° 3.

Après avoir collecté plusieurs centaines d'entrepôts en différents formats (OAI-PMH, RDFa, des flux d'actualités RSS et Atom) les différentes ressources sont indexées et enrichies par l'utilisation de référentiels comme Rameau (BNF), les thésaurus Pactols (Réseau Frantiq) ou Geonames (Référentiel de localisation international). Cela signifie que de nouveaux mots clefs ou concepts vont être ajoutés dans la description des métadonnées. Prenons l'exemple d'une photo, celle de la mosquée de Hama qui figure dans l'archive ouverte MédiHAL¹³. Les mots clefs qui figurent dans la ressource originale sont : Syrie – Syria – Hama – architecture – mosquée – culte – islam ; lorsque l'on retrouve le même document dans ISIDORE, la liste des mots clefs est devenue mosquée, Hama, Syrie, sciences, Sciences, minaret, trésor, Architecture, architecture, Culte, pratique religieuse, Islam, islam¹⁴. Les termes « trésor », « minaret » et « pratique religieuse » ont été ajoutés. Ils résultent de l'indexation et de l'enrichissement auxquels a procédé la chaîne de traitement d'ISIDORE. Dans le cas d'un article ou d'un ouvrage, l'accès au texte de la ressource permet aussi de compléter les notices initiales.

¹³ Voir <<http://medihal.archives-ouvertes.fr/>> consulté le 5 mars 2012.

¹⁴ Les doublons s'expliquent par le fait que l'un des référentiels commence par des lettres minuscules et qu'un autre débute lettres majuscules, dans un premier temps cela nous permet d'identifier les référentiels.

L'accès à ISIDORE se fait avant tout par le Web, sur le site <<http://rechercheisidore.fr>>. Il donne actuellement accès à plus d'un million trois cents mille ressources numériques, issues du moissonnage de plus de mille sources de données différentes. En complément, les producteurs qui fournissent des données à ISIDORE ont accès aux données enrichies d'ISIDORE via une interface spécifique SPARQL endpoint, qui permet d'interroger, selon les principes du Web de données¹⁵, l'ensemble des informations structurées et normalisées par ISIDORE. Les producteurs ont, de cette façon, la possibilité d'améliorer la qualité de leurs données. Ainsi, entre ISIDORE et les producteurs, il y a des relations permanentes, dont Stéphane Pouyllau, en charge du développement d'ISIDORE au sein du TGE, s'occupe directement.

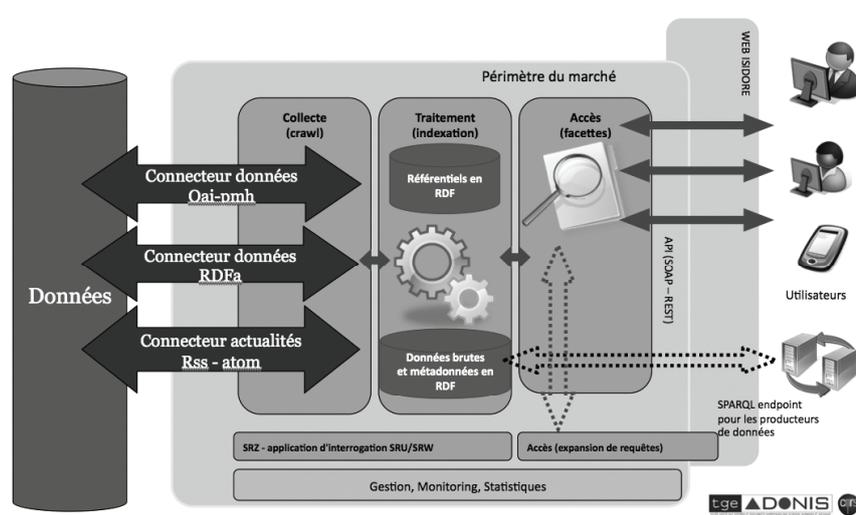


Schéma n° 3 : Le fonctionnement d'ISIDORE

La puissance d'ISIDORE permet de capitaliser rapidement les ressources qui sont numérisées par différents portails de revues comme Revue.org, CAIRN ou Persée, mais aussi des entrepôts de taille plus modeste qui peuvent être très utiles sur les sujets plus pointus. Par exemple, Criminocorpus, portail sur l'histoire de la justice, des crimes et des peines¹⁶ ou les bibliothèques virtuelles humanistes¹⁷ qui permettent d'accéder à des ouvrages rares, propriété de bibliothèques étrangères, tels *Les Essais* de Montaigne dans une édition de 1652 annotée par Jean-Jacques Rousseau mise en ligne en juillet 2011¹⁸...

Les ressources qui sont proposées sont de différentes natures (images, articles, ouvrages, mémoires et thèses, ressources bibliographiques, annonces de séminaires...) et les facettes qui existent sur le côté droit permettent de les sélectionner.

¹⁵ Le Web de données est en fait le Web actuel augmenté de données structurées et d'information scientifique. Il est en cours de réalisation au niveau international et de nombreux partenaires français y participent : la BNF, l'ABES, le CNRS, les universités, etc.

¹⁶ <<http://www.criminocorpus.cnrs.fr/>> consulté le 5 mars 2012.

¹⁷ <<http://www.bvh.univ-tours.fr/>> consulté le 5 mars 2012.

¹⁸ <<http://www.bvh.univ-tours.fr/Consult/index.asp?numfiche=766>> consulté le 5 mars 2012.

4. Les projets européens

Fort de ses réalisations, le TGE est engagé dans plusieurs projets européens et, tout d'abord, dans DARIAH.

L'objectif est de développer une infrastructure numérique distribuée qui prenne en charge l'ensemble du processus de recherches en SHS. La méthode de construction est *bottom-up*¹⁹ en ce qu'elle part des réalisations nationales. Ces dernières sont rendues accessibles et/ou (re)développées de manière collaborative [entre les partenaires européens].

Quatre centres virtuels de compétences (*Virtual Competency Centers* ou VCC) doivent prendre en charge la construction de l'infrastructure :

VCC1 *Infrastructure*;

VCC2 *Research and Education Liaison*;

VCC3 *Scholarly Content Management*;

VCC4 *Advocacy and Outreach*.

Chacun de ces centres de compétences est multinational, sous la responsabilité de deux pays. La France est en charge avec les Pays-Bas du VCC3. En liaison avec certains domaines du TGE, ses orientations privilégiées sont les données et leurs formats, le respect des normes et standards, les questions de stockage et d'archivage ainsi que le respect des bonnes pratiques. Un certain nombre de ses objectifs sont communs avec le VCC1 en charge de l'infrastructure générale, comme la structuration informatique de référentiels (listes d'autorités, thesaurus et ontologies) pour la « curation »²⁰ des données.

Jusqu'au mois de février 2011, la phase préparatoire du projet a permis l'élaboration des statuts, les premières simulations budgétaires et les premières évaluations par les ministères des différents pays. Aujourd'hui, les pays impliqués dans la phase de transition qui doit aboutir à la création d'un ERIC (*European Research Infrastructure Consortium*), sont l'Allemagne, l'Autriche, le Danemark, l'Irlande, les Pays-Bas et, naturellement, la France. Deux institutions suisses et italiennes interviendront à titre de *Cooperating Partners*. D'autres pays continuent de réfléchir à la forme que prendra leur implication dans le projet. On peut citer : la Lituanie, la Grèce, la Croatie, la Serbie, la Slovénie, le Royaume-Uni. D'autres contacts ont été pris avec l'Albanie, le Luxembourg et la Norvège.

La France doit déposer le projet d'ERIC et accueillir le siège social de l'organisation internationale lorsqu'il sera effectif. Par ailleurs, elle participe au DARIAH *Coordination Office* avec l'Allemagne et les Pays-Bas. L'ensemble de ce projet devrait être effectif dans le courant de l'année 2012.

Pour autant, DARIAH n'est pas le seul projet européen au sein duquel le TGE Adonis est impliqué. Sophie David, qui suit ce dossier pour le TGE, participe également au projet Nedimah (*Network for Digital Methods in the Arts and Humanities*). Il s'agit d'un projet de type RNP (*Research Network Program*), financé par la Fondation européenne pour la science (ESF). Il rassemble des institutions émanant des pays suivants : Allemagne, Bulgarie,

¹⁹ Le *bottom* renvoie ici aux réalisations nationales et le *up* aux réalisations européennes.

²⁰ Ensemble des activités et opérations nécessaires à une gestion active des données de recherche numériques, tout au long de leur cycle de vie.

Croatie, Danemark, Finlande, France, Irlande, Norvège, Pays-Bas, Portugal, Royaume-Uni, Suède et Suisse

Le travail est consacré à l'identification de l'ensemble des méthodes actuellement utilisées telles que : la représentation de l'espace et du temps, la visualisation en 2D et en 3D, la construction d'ontologies et le *Linked Data*, la constitution de grandes collections, l'exploitation de grandes collections, l'édition numérique scientifique. Deux groupes de travail transversaux s'attachent quant à eux à la publication numérique et à la construction d'une ontologie sur les méthodes. À l'issue du projet seront disponibles une carte représentant l'écosystème, soit l'ensemble des techniques, des méthodes et des modes de collaboration, et une ontologie sur les méthodes. Le rôle du TGE est ici d'encourager les membres de la communauté française des humanités numériques à participer aux travaux des groupes de travail.

Le TGE est aussi membre du *TEI Board (Text Encoding Initiative)*.

Enfin, il est partie prenante du projet Webdatanet (*Web-based data Collection – Methodological Challenges, Solutions and Implementations*²¹). Ce projet, de type COST (*European Cooperation in Science and Technology*), rassemble une trentaine d'institutions issues de 21 pays : Allemagne, Belgique, Chypre, Croatie, Espagne, Finlande, France, Grande-Bretagne, Grèce, Islande, Israël, Italie, Luxembourg, Malte, Norvège, Pays-Bas, Pologne, Roumanie, Slovaquie, Slovénie, Suisse. Laurent Lesnard (CDSP²²) est le représentant officiel de la France dans le projet et Sophie David, en charge des projets internationaux au TGE, est la représentante suppléante.

L'objectif du projet est de construire un réseau multidisciplinaire, plus particulièrement issu des sciences sociales, intéressé par les questions méthodologiques posées par l'utilisation du Web dans le recueil et le traitement des données.

Ainsi le TGE, fort de ses réalisations dans le domaine des Humanités numériques à destination des communautés scientifiques françaises en SHS, s'implique dans les projets européens dans le droit fil des orientations d'ESFRI (*European Strategy Forum on Research Infrastructures*).

²¹ <http://www.cost.eu/domains_actions/isch/Actions/IS1004> consulté le 5 mars 2012.

²² Centre de données sociaux-politiques.

Partie 3
Pratiques et projets en IST

AllianzInitiative, Germany

Ralf Schimmer

*Head of Department for Scientific Information Provision,
Max Planck Virtual Library
Germany*

Abstract

In their 5-year Priority Initiative "Digital Information" (2008-2012) the German Science Organisations have agreed to coordinate their activities and to expand on the ideal of the innovative information environment on behalf of their researchers. Equipping scientists and scholars with the information infrastructure best suited to meeting their research needs is the guiding principle of this priority initiative. Key areas of activities are: National Licensing; Open Access, National Hosting Strategy, Research Data, Virtual Research Environments; Legal Frameworks. In this presentation an overview over these activities and future plans will be given. From the experience of the Max Planck Digital Library, it will also be reported how the achievements in the various working groups influence the daily work in an institutional library.

Résumé

Allianzinitiative, Allemagne

Dans leur initiative « Information numérique », programmé sur cinq ans (2008-2012), les organismes allemands de recherche se sont entendus pour coordonner leurs activités et développer pour leurs chercheurs un environnement informationnel innovant. Ce projet se fonde sur la volonté d'offrir aux scientifiques et aux universitaires l'infrastructure informationnelle la mieux adaptée à leurs besoins.

Les axes de travail privilégiés sont les suivants :

- les licences nationales ;
- le libre accès ;
- la définition d'une stratégie nationale d'archivage numérique ;
- les données de la recherche ;
- l'environnement de recherche virtuel ;
- les cadres juridiques.

L'exposé présentera une vue d'ensemble de ces activités ainsi que des projets à venir. Il montrera également, à partir de l'expérience de la Bibliothèque numérique Max Planck, comment les réalisations des différents groupes de travail peuvent influencer sur le travail quotidien d'une bibliothèque institutionnelle.

Pratiques et projets IST en Pologne

Barbara Szczepanska

*Kierownik Biblioteki i Zasobów Informacyjnych / Library and Information Services
Manager Hogan Lovells (Warszawa) LLP (Spółka partnerska)
Oddział w Polsce
ul. Nowogrodzka 50
00-695 Warszawa, Polonia*

Résumé

Longtemps cantonnées à un niveau local, les initiatives relatives à la bibliothèque numérique et à l'Open access pour le soutien à la recherche et à l'éducation se structurent au niveau national à partir de 2008.

En matière de numérisation et de contenu, cette date charnière voit la prise de conscience croissante et l'implication grandissante de l'État au travers du ministère de la Culture. Des programmes ministériels se développent, particulièrement ces trois dernières années, avec la formulation de stratégies, l'intégration de démarches, la mise en place de modalités opérationnelles univoques pour le développement de contenus numériques via la Fédération des bibliothèques numériques (Federacja Bibliotek Cyfrowych) qui rassemble 70 bibliothèques et la bibliothèque numérique Polona. Tous ces efforts ont ainsi permis à la Pologne de diffuser ses ressources culturelles organisées au sein de la CBN Polona (Bibliothèque nationale polonaise) et à participer au projet Europeana.

La diffusion de ressources concerne aussi le patrimoine intellectuel de la communauté scientifique par leur mise à disposition dans des archives ouvertes, notamment l'AMUR (Adam Mickiewicz University Repository), unique archive institutionnelle de dépôts en libre accès. La diffusion se fait aussi par un portail proposant des contenus éducatifs sous licence Creative Commons (Open AGH). La conservation et la pérennisation des contenus numériques et des données sont assurées par l'ICM UW (Centre interdisciplinaire de modélisation mathématique et informatique de l'Université de Varsovie/Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego Uniwersytetu Warszawskiego) qui stocke des contenus hétérogènes, de la publication à la base de données en libre accès, et y donne accès via la plateforme YADDA.

Enfin, le droit reste un frein au développement de l'accès libre aux ressources : malgré l'adaptation de licences, telles que Creative Commons, le droit polonais, très strict en matière de droit d'auteur, empêche les auteurs de disposer librement de leurs œuvres.

Abstract

Practices and projects in Poland

Initiatives linked to the digital library and open access in support of research and education have been structured at the national level since 2008, with only local level initiatives prior to this date.

2008 was thus a turning point which brought growing awareness of the issue of digitalisation and contents and increased State involvement in this field through the Ministry of Culture. Ministerial programmes have been developed, particularly over the last three years, involving the formulation of strategies, integrating working methods and the implementation of clearly defined operating modes through the Federation of Digital Libraries (Federacja Bibliotek Cyfrowych) with its 70 member libraries and the Polona Digital Library. This work has enabled Poland to disseminate Polish cultural resources through the CBN Polona (Polish National Library) and participate in the Europeana project.

The dissemination of resources also concerns the intellectual heritage of the scientific community by making the resources available in open archives notably AMUR (Adam Mickiewicz University Repository), the only institutional open access archive of repositories. Resources are also disseminated via a portal giving access to educational content under a Creative Commons (Open AGH) licence. Hosting and long term preservation of digital content and data is managed by the ICM UW (Interdisciplinary Centre for Mathematical and Computational Modelling of the University of Warsaw/Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego Uniwersytetu Warszawskiego) which stores heterogeneous contents from the publication to the database and provides access to these via the YADDA platform.

Finally, rights issues still hinder the development of open access to resources. Despite the adaptation of licences such as Creative Commons, Polish law remains very strict on copyright which prevents authors from doing whatever they wish with their work.

1. Bibliothèques numériques et libre accès des bibliothèques universitaires

Pendant très longtemps, la Pologne n'avait pas de politique concernant la création des contenus numériques. Bien que certains éléments de ces régulations soient signalés dans quelques documents gouvernementaux, ils ne constituaient pas un système cohérent et n'étaient pas mis en pratique, *a fortiori* sur celui du financement de la numérisation. Toutes les initiatives de la période 2001-2004 relatives au développement des collections numériques avaient un caractère local et étaient financées par les fonds propres des institutions ou plus rarement par des Fonds structurels de l'Union européenne. L'objectif de ces initiatives était de soutenir la recherche et l'éducation ainsi que de réduire l'écart entre les bibliothèques polonaises et celles des pays européens technologiquement les plus avancés.

La période 2004-2008 a été capitale pour la numérisation au sein des bibliothèques polonaises : elles ont acquis un savoir-faire dans le domaine de la création de collections numériques. Le Centre d'informatique et des réseaux de l'Académie polonaise des sciences à Poznań (PCSS) a élaboré puis développé des outils informatiques peu onéreux afin de

créer des contenus numériques (le moteur dLibra ainsi que des logiciels libres). De surcroît, des possibilités de financement pour la numérisation se sont présentées par le biais de fonds communautaires aussi bien que nationaux. Avec le temps, l'attitude du ministre de la Culture et du Patrimoine national de la République de Pologne a également évolué. Nous avons pu noter la consolidation des modèles de fonctionnement de la bibliothèque numérique. C'est le cas notamment de la Bibliothèque numérique de la Grande Pologne (*Wielkopolska*), de la Bibliothèque de la région de Poméranie (*Kujawsko-Pomorska*) de Wrocław ou bien, de la Bibliothèque numérique de Silésie : *Śląska Biblioteka Cyfrowa*.

L'année 2008 a apporté des activités considérables et intensives en vue de la construction de bibliothèques numériques universitaires. Celles des grandes bibliothèques académiques qui avaient individuellement lancé la construction des contenus numériques ont travaillé sur la consolidation de leurs activités. Plusieurs centres scientifiques majeurs (15 exactement) dont deux bibliothèques publiques, ont décidé de lancer une coopération dans le cadre du Consortium polonais des bibliothèques numériques. Son objectif stratégique est d'assurer à toute personne intéressée un accès libre et commun à des contenus numériques cruciaux pour la science, l'éducation et la culture, grâce au développement du réseau des bibliothèques numériques et des répertoires, constituant une structure solide et viable.

On y souligne notamment les objectifs suivants :

- soutenir la création, la collecte et le traitement des données ;
- rendre accessibles les contenus numériques et les archiver ;
- recommander et diffuser les standards et les bonnes pratiques internationales relatifs à la création, le traitement, la mise en ligne et l'archivage des contenus numériques ;
- soutenir les initiatives assurant la préservation du continu des collections numériques.

La fondation de la Fédération des bibliothèques numérisées (*Federacja Bibliotek Cyfrowych*, FBC) représente une étape majeure dans la structuration des réseaux dispersés des bibliothèques numériques et des répertoires en Pologne. Son nom correspond à sa nature : il s'agit d'un ensemble de services de réseaux avancés reposant sur des contenus numériques mis à disposition dans des bibliothèques numériques polonaises, ainsi que des répertoires dans le cadre du réseau PIONIER, résultat de la coopération de nombreuses institutions scientifiques et publiques : universités, bibliothèques, archives, musées ou centres de recherches. Le service FBC est entretenu par le Centre d'informatique et des réseaux, tandis que son développement est assuré par le Groupe des bibliothèques numériques (*Zespół Bibliotek Cyfrowych*, PCSS).

La mission du service FBC consiste à :

- faciliter l'exploitation des contenus des bibliothèques numériques et des répertoires polonais ;

- accroître la visibilité et la diffusion sur Internet des contenus numériques des bibliothèques polonaises et des dépôts ;
- mettre à disposition des utilisateurs d'Internet et des créateurs des bibliothèques numériques des services nouveaux et avancés du réseau fondés sur des contenus de bibliothèques numériques et de répertoires.

En ce qui concerne le cadre formel et organisationnel, la coopération des bibliothèques numériques et des répertoires avec le service FBC est très simple. Elle ne requiert aucun paiement, aucun travail de la part des administrateurs du répertoire, ni la transmission des publications numériques au FBC. Il suffit de fournir à ses administrateurs une adresse URL qui permet de télécharger des données du répertoire selon le protocole OAI-PMH¹, qui est tout simplement l'adresse de l'interface OAI-PMH dudit répertoire. Le schéma de description de base est le *Dublin Core*.

La Fédération des bibliothèques numériques rassemble actuellement 70 bibliothèques régionales et institutionnelles, soit environ 200 institutions. D'autres bibliothèques envisagent également de rejoindre l'initiative ou bien, sont déjà en phase d'intégration. Elles mettent à disposition leurs contenus numériques *via* le site de la Fédération ou par ceux de leurs universités.

L'année 2008 marque la stabilisation du nombre de bibliothèques numériques ainsi que la création de nouvelles bibliothèques. Les bibliothécaires disposent déjà d'un certain savoir technique, technologique et juridique.

Certaines bibliothèques numériques polonaises donnent accès à presque toutes les catégories de leurs ressources. Depuis peu, les écoles polytechniques et les universités de médecine cherchent aussi à mettre à disposition leurs derniers travaux, ce qui justifie donc la prise en en considération et l'ajout de leurs ressources didactiques. L'analyse du contenu des bibliothèques numériques en Pologne démontre que le matériel le plus souvent numérisé concerne la presse régionale du XIX^e siècle et du début du XX^e siècle (environ 60 %). Certaines bibliothèques (entre autres la Bibliothèque numérique de la Petite Pologne, *Małopolska Biblioteka Cyfrowa*, la Bibliothèque numérique de Jelenia Góra, *Jeleniogórska Biblioteka Cyfrowa*) – conformément aux accords passés avec les éditeurs – présentent également la presse contemporaine. À la deuxième place, arrivent des documents d'un feuillet à savoir des photos, des graphiques, des cartes postales, etc. La numérisation de la presse, très souvent sous forme de microfilms, est relativement simple et peu onéreuse. Les bibliothèques numérisent également des ouvrages, des manuscrits, d'anciens imprimés, des partitions musicales et des objets cartographiques. Cependant en Pologne, le contenu numérisé contient peu de matériel audiovisuel. Les bibliothèques numériques mettent également dans leurs collections de nouvelles revues et des travaux de recherche venant des universités polonaises. Au regard de la diversité de leurs collections, les contenus des bibliothèques numériques peuvent constituer un important support didactique et scientifique, surtout dans le domaine des sciences humaines. Le développement des bibliothèques

¹ *Open Archive Initiative Protocol for Metadata Harvesting.*

numériques dans les universités techniques et médicales permet d'avancer l'hypothèse que prochainement, la didactique dans ces domaines sera fortement enrichie par des ressources numériques.

La demande de contenus numériques en Pologne est immense. Les collections numériques sont visiblement appréciées par les lecteurs. Bien que les attentes des utilisateurs ne soient pas encore analysées, le suivi des visites des sites (Web/ou portails) des bibliothèques permet de constater que le nombre moyen des visites dépasse 50 000, ce qui est un résultat considérable. Les données statistiques sont régulièrement fournies par la Fédération des bibliothèques numériques. Il en résulte par exemple que son seul site accueille près des 28 000 visites de différents types d'utilisateurs. Dans chaque bibliothèque, des utilisateurs entrent également indépendamment par les sites individuels WWW.

Malgré certaines différences dans les caractéristiques de base et les principes de fonctionnement, les bibliothèques numériques et les répertoires en libre accès (*Open Access*) reposent sur le modèle ouvert et réalisent les objectifs du libre accès, leurs contenus étant complètement ouverts dans la plupart des cas. Seule une partie est soumise à un contrôle, rendant la disponibilité limitée à des membres d'institutions particulières. Les bases de données reposent surtout sur le protocole d'échange de données OAI PMH, ce qui permet d'introduire les données dans la circulation à l'échelle mondiale et les rendre ainsi accessibles grâce aux plus importants moteurs de recherche scientifiques comme OASTER, BASE ou *Google Scholar*.

À partir du 12 décembre 2009, le Centre informatique et des réseaux de l'Académie polonaise des sciences à Poznań sont aussi fournisseurs de données pour Europeana, ce qui a permis de considérablement augmenter les contenus polonais. Pour Europeana, ces derniers sont agrégés à partir de deux sources : la Fédération des bibliothèques numériques (qui compte actuellement 551 175 objets associés) et la Bibliothèque numérique Polona organisée au sein de la Bibliothèque nationale (*CBN Polona*) contient 29 800 publications et a été visitée par 31 722 802 utilisateurs depuis sa création en 2008.

Comme Europeana est orientée sur le patrimoine culturel, la FBC a coopéré avec DART-Europe² et à partir de janvier 2010 elle a aussi donné accès à ses données sur cette plateforme. Actuellement, il y a 2 105 publications scientifiques transmises à la base DART via FBC.

L'œuvre scientifique polonaise est présente dans les répertoires les plus importants. Le *Directory of Open Access Journals* contient les données sur 99 revues polonaises disponibles en *Open Access*, dont 13 avec licence. La majorité de ces revues est postérieure à 2000. Les contenus avant 1970 sont mis à la disposition par *Acta Ichthyologica* et *Piscatoria*. Les revues OA sont éditées aussi bien par les universités, les instituts de recherche que des entreprises privées.

² Portail de thèses européennes.

Open DOAR enregistre 48 répertoires polonais, dont seulement deux sont des répertoires classiques. Les autres appartiennent à des bibliothèques numériques avec une fonction d'archivage des travaux de recherche, tandis que le portail Driver note 671 publications scientifiques polonaises et deux répertoires, dont un est en réalité celui de la Fédération des bibliothèques numériques.

1.1 AMUR : Adam Mickiewicz University Repository

L'AMUR, créé en 2008 à l'Université Adam Mickiewicz de Poznań, est le premier – et à vrai dire, l'unique – répertoire institutionnel classique en Pologne basé sur le logiciel professionnel DSpace. C'est un répertoire des collections numériques et du patrimoine intellectuel de la communauté scientifique de l'université. Conformément à la réglementation, son contenu institutionnel est défini par l'institution. À caractère scientifique, cumulatif et permanent, il est accessible en mode libre accès OA. Il se caractérise par l'interopérabilité (conformité avec l'OAI-PMH, qui permet d'effectuer la recherche dans les contenus des répertoires à l'aide des applications consacrées telles que *Google Scholar*, Scirus et autres moteurs de recherche).

Les travaux sur le répertoire AMUR ont duré presque trois ans. La première étape de sa création a consisté à promouvoir l'idée du libre accès au savoir et à faire connaître aux universitaires le fonctionnement des répertoires institutionnels ainsi que les processus d'archivage. Le but était de présenter les informations sur des répertoires lors de la délibération du conseil des facultés et d'en créer une section sur le site de la bibliothèque universitaire. Une enquête effectuée au sein du milieu universitaire, par les créatrices du projet, a permis de connaître les attentes pour cette initiative de créer le répertoire institutionnel. Son résultat s'est montré remarquablement favorable.

Le répertoire a une structure hiérarchisée. Il se compose de groupes et de collections. Les sections englobent des facultés ainsi que d'autres entités universitaires. Parmi les collections les plus importantes, il faut mentionner celles des revues scientifiques (y compris certaines de leurs archives parues à l'UAM), des thèses de doctorat, des articles scientifiques, des actes de colloques ainsi que du matériel didactique. L'archivage des documents peut être réalisé selon une démarche facultative ou obligatoire. Dans le cadre d'AMUR, il est obligatoire d'y déposer les thèses de doctorat (conformément à l'ordonnance émise par le recteur), tandis que le reste du dépôt de l'œuvre intellectuelle de l'université est facultatif.

1.2. Open AGH

*Open AGH*³ est un nouveau portail dédié aux contenus éducatifs. Son objectif est de promouvoir l'ouverture et le développement de l'éducation et de la recherche. Les données accessibles à tous grâce à *Open AGH* ont été élaborées par le personnel et les étudiants de l'Académie Górniczo-Hutniczej à Cracovie. Conçu sous licence *Creative Commons*

³ Académie Górniczo-Hutniczej.

« Paternité – Pas d'utilisation commerciale – Partage à l'identique des conditions initiales 2.5 Pologne » ce portail permet une utilisation dans un but non commercial et éducatif, ainsi qu'une adaptation au matériel déjà existant, selon les exigences individuelles.

1.3 Programmes ministériels

Pendant les trois dernières années nous avons pu remarquer les efforts du ministère de la Culture et du Patrimoine qui, à travers le Département du patrimoine national, cherchait à stimuler la création de contenus numériques en Pologne. Le 24 avril 2006, le ministre de la Culture et du Patrimoine national a créé une section de numérisation qui est un service auxiliaire affilié à la Bibliothèque nationale.

Son objectif est de :

- 1 – formuler une stratégie uniforme de la numérisation du patrimoine culturel et de l'œuvre scientifique en Pologne, quel que soit le statut juridique ou le lieu du dépôt des documents ;
- 2 – établir les modalités uniques de la numérisation, de l'accessibilité et du stockage d'une ressource numérique pour les bibliothèques, archives et musées ;
- 3 – intégrer les démarches entreprises par les bibliothèques, musées et archives en privilégiant la numérisation du patrimoine national.

Le groupe plus spécifiquement responsable de la numérisation (*Zespół ds. Digitalizacji*) a préparé entre autres le Programme de numérisation des biens culturels et de collecte, stockage et accessibilité des ressources numériques en Pologne de la période 2009-2020. Ce programme est susceptible d'englober toutes les institutions de mémoire.

Il prévoit la numérisation dans les domaines suivants :

- **bibliothèques** : 9-10 mln⁴ entités (état actuel : environ 600 000) ;
- **archives** : environ 37 mln entités d'inventaire (état actuel : environ 2 mln) ;
- **musées** : 13 mln objets (état actuel : environ 300 000).

Le programme a été conçu afin de développer les contenus numériques polonais (un exemplaire obligatoire des documents numériques de la bibliothèque, archivage des documents numériques natifs, archivage des ressources et archivage d'Internet) ; protéger les ressources numérisées (élaborer des standards uniformes, créer les entrepôts de données, des centres de compétence et de formation), ainsi qu'assurer :

- l'accès aux contenus numériques, y compris par la mise à disposition des ressources polonaises *via* des portails comme APENET et Europeana ;
- la mise à disposition en ligne des ressources du domaine public et la promotion de la politique du libre accès aux contenus numériques ;

⁴ Mètre linéaire.

- l'organisation du système de l'octroi de la licence pour la publication numérique d'ouvrages soumis au droit d'auteur.

En février 2011, le ministère de la Culture et du Patrimoine national a annoncé l'ouverture du programme *Kultura+* (Culture plus), dont l'objectif est d'améliorer l'accès à la culture et d'en augmenter le taux de participation dans des communautés locales, particulièrement au sein des communes rurales et semi-rurales. Parmi ses priorités figure la numérisation qui devra être réalisée grâce à :

- l'organisation du réseau d'ateliers de numérisation dans les centres de compétence et autres grands centres qui devraient accélérer la numérisation des œuvres culturelles ;
- la création du réseau de répertoires numériques professionnels dans les centres de compétence et autres centres suffisamment grands afin de stocker lui-même ses ressources numérisées ;
- la mise à disposition en ligne des contenus mis en forme des musées, bibliothèques, archives numériques ainsi que des ressources audiovisuelles ;
- l'augmentation de l'accessibilité aux ressources des musées, bibliothèques, archives ainsi qu'aux ressources audiovisuelles ;
- l'augmentation du nombre des utilisateurs des services des bibliothèques et des maisons de la culture.

On doit souligner que, selon le règlement, les structures qui mettent gracieusement à disposition par le biais du domaine public 85 % de ressources numérisées, bénéficieront de points supplémentaires au cours de leur évaluation future.

Pour la réalisation des objectifs définis dans le cadre du programme *Kultura+* (Culture plus) dans les années 2011-2015, la somme de 516 millions de zlotys (129 millions d'€) a été affectée.

2. Le Centre interdisciplinaire de modélisation mathématique et informatique de l'Université de Varsovie ***Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego Uniwersytetu Warszawskiego (ICM UW)***

Depuis plusieurs années, l'ICM UW œuvre pour le développement des archives ouvertes en Pologne. Ses objectifs se concentrent surtout sur la construction de son infrastructure. Le centre concourt également à des projets d'acquisition de contenus scientifiques, et à la mise en place de bases de données de différents domaines scientifiques. Elle offre aussi des outils juridiques en appui à l'accès à la science.

« La Bibliothèque scientifique virtuelle » *Biblioteka Wirtualna Nauki* (BWN), qui fonctionne depuis 1995, donne accès à des publications scientifiques sous licence à des centaines

d'institutions de recherche. À partir de 2007, l'ICM diffuse et met en œuvre le modèle du libre accès au savoir.

L'ICM UW propose aussi un répertoire, la plateforme YADDA.

Il s'agit d'une plateforme destinée à la création de répertoires numériques. Elle sert à stocker, conserver et mettre à disposition des contenus (articles ou ouvrages). L'idée directrice de la création de YADDA a été de pouvoir y stocker de très grandes collections de données sur de longues périodes de manière pérenne et efficace. Elle permet d'intégrer des contenus hétérogènes autant du point de vue de leur taille que de leurs caractéristiques, à savoir, à partir d'une seule publication jusqu'à des bases de données contenant des millions d'enregistrements. Grâce à elle, il est possible de créer un ensemble de contenus scientifiques numérisés accessible en ligne. Cet ensemble est homogène, intégré et actif dans toute la Pologne.

La plateforme YADDA, actuellement utilisée au sein de l'ICM UW, est composée des trois éléments suivants connectés par un système de services en réseau :

- un répertoire de données composé des archives des contenus, leur catalogue, ainsi que des index qui permettent de visionner la collection et d'y effectuer une recherche ;
- le système YADDAWEB, assurant aux utilisateurs la possibilité de visionner le contenu et d'y effectuer une recherche à travers la Toile (WWW) ;
- du logiciel DeskLight : qui permet l'introduction, l'édition et la possibilité d'accéder aux données.

Le répertoire rassemble des ressources scientifiques variées : collections Elsevier Science, BazTech – la base de données bibliographiques et de résumés d'articles parus dans des revues scientifiques polonaises des domaines techniques, ainsi que certaines revues de sciences exactes et sur la protection de l'environnement, BazHum⁵, AGRO⁶, une Bibliothèque virtuelle des mathématiques⁷. L'accès à ces bases s'effectue au moyen d'une interface opérationnelle, qu'elles aient été acquises aussi bien dans un cadre marchand que dans le modèle ouvert.

C'est sur la plateforme YADDA que fonctionne également la Bibliothèque scientifique virtuelle. Durant de nombreuses années, l'ICM UW jouait le rôle d'intermédiaire dans l'accès aux bases scientifiques de données à caractère commercial. Les contrats de licence étaient conclus entre des bibliothèques et des consortiums. En 2010, le ministère de l'Éducation en coopération avec l'ICM a lancé le projet de la Bibliothèque scientifique virtuelle, entièrement financée par le budget de l'État. Pour accéder à la Bibliothèque il faut être enregistré. Le coût

⁵ Base bibliographique de données du contenu des revues polonaises du champ des sciences humaines et sociales.

⁶ Base bibliographique répertorient des articles parus dans des revues polonaises en sciences naturelles, agricoles et techniques.

⁷ Base de données de texte intégral comprenant les articles archivés et des articles récents venant des revues de mathématiques polonaises.

des licences pour l'acquisition des bases de données scientifiques en 2010 s'élève à 150 millions de zlotys. L'accès au contenu de la Bibliothèque scientifique virtuelle est possible aussi bien dans les universités publiques que privées.

Un autre projet de l'ICM UW s'intitule « Ouvre un livre ». Son objectif est de créer la collection numérique d'ouvrages scientifiques contemporains mis à disposition par leurs auteurs. La collection est composée surtout d'ouvrages dits « épuisés » dont l'acquisition en version papier n'est plus possible, bien qu'ils aient toujours une valeur scientifique et éducative. Le projet est réalisé sous le patronage honorifique de l'Académie polonaise de la science, en coopération avec la Bibliothèque nationale (responsable de la numérisation des ouvrages) ainsi que du Musée d'Histoire de la Pologne. L'initiative a reçu le soutien de pays tels que : l'Islande, le Liechtenstein et la Norvège, ainsi que des subventions de l'Espace économique européen et du norvégiennes dans le cadre du Fonds des bourses et de formation.

Les auteurs des ouvrages sont titulaires des droits d'auteur. Dans le cas, ils mettent les œuvres à la disposition du public sous licence CC. Pourtant, dans certains cas, l'auteur dispose du droit d'exploitation d'une œuvre sur Internet seulement. Les droits liés à cet ouvrage restent alors réservés, mais il est disponible dans le modèle ouvert et il peut être exploité selon le principe de l'exploitation autorisée. La collection comprend actuellement 88 ouvrages couvrant les domaines de l'histoire, la culture, la sociologie et la gestion. Chaque œuvre est accessible sous forme d'un document scanné (qu'il est possible d'agrandir et de visionner pour la recherche), en version texte et en format PDF téléchargeable.

L'ICM UW travaille également sur un nouveau projet SYNAT, dont le but est de créer une plateforme d'hébergement et de communications, qui soit universel et ouvert à des contenus du réseau relatifs aux sciences, à l'éducation toujours dans la perspective d'une société de la connaissance ouverte. Outre les travaux relatifs à la plateforme informatique, ses créateurs mettent l'accent sur l'élaboration d'un recueil de propositions de modèles juridiques visant au développement de nouveaux modèles de communication ouverts. Ils envisagent des solutions susceptibles de permettre la future commercialisation du projet.

L'ICM prend également part à des projets pour l'accès libre. Il héberge par exemple *Creative Commons Polska*, il est le partenaire technique du projet OpenAIRE et un membre-fondateur de COAR (*Confederation of Open Access Repositories*) – organisation qui défend le modèle du libre accès à l'échelle mondiale.

3. La Coalition pour l'éducation ouverte ***Koalicja Otwartej Edukacji***

La Coalition pour l'éducation ouverte *Koalicja Otwartej Edukacji* a été créée le 27 novembre 2008 par quatre organisations : l'Association des bibliothécaires polonais *Stowarzyszenie Bibliotekarzy Polskich*, ICM UW/CC-Polska, la Fondation La Pologne moderne *Fundacja Nowoczesna Polska* et l'Association Wikimedia Polska. Actuellement, elle

associe onze organisations qui œuvrent pour le développement et l'exploitation des contenus éducatifs ouverts.

Les membres de cette organisation réalisent entre autres les projets suivants :

- **Manuels libres** – projet de la Fondation La Pologne moderne. Il s'agit pour les enseignants de préparer des ouvrages en accès libre. L'action suit l'idée du mouvement de la libre culture selon laquelle le savoir de l'humanité, notre patrimoine commun, devrait être accessible et ouvert à tous. Il crée une communauté qui œuvre dans le cadre de la vitrine d'une plateforme collaborative. Plusieurs auteurs peuvent simultanément écrire, commenter et corriger les modules d'un manuel ;
- **Lectures libres** – est un projet réalisé par la Fondation La Pologne moderne. Lancé en 2007, son but est de donner accès à des lectures scolaires obligatoires recommandées par le ministère de l'Éducation nationale. Il concerne des ouvrages qui sont déjà dans le domaine public, dont l'exploitation n'est pas régie par le droit d'auteur. Ces ouvrages accompagnés d'un commentaire sont disponibles en plusieurs formats (html, odt, txt et pdf). Il est possible de les consulter, télécharger et partager gratuitement. Leur accessibilité se fait sous licence CC Paternité ;
- **L'Association Wikimedia Polska** est chargée de Wikipedia, du Wikidictionnaire *Wikisłownik* (Wiktionary), du recueil des citations Wikicitations *Wikycytaty* (Wikiquote), du service manuels Wikibooks, du service de textes sources Wikisource *Wikiźródła* (Wikisource), du service d'information (*Wikinews*), du répertoire des médias (*Wikimedia Commons*) ainsi que du catalogue des espèces biologiques ;
- **Le Bulletin EBIB** – revue en libre accès publiée à partir de 1999 par l'Association EBIB. Il s'agit d'une revue en ligne destinée aux bibliothécaires, spécialistes en information et à toute personne intéressée par les problématiques de la profession. Conformément à l'idée du libre accès, la revue donne accès sous licence CC Paternité à tous les articles sans aucune restriction ni aucun paiement. Elle traite des problèmes relatifs aux bibliothèques, y compris numériques, de la numérisation, ainsi que des thématiques de l'information, d'Internet, et de la théorie de la communication, mais aussi de l'intérêt pour la lecture et les technologies d'information, etc.
- La Coalition de l'éducation ouverte *Koalicja Otwartej Edukacji* est un groupe responsable de la formation et de la promotion de l'idée d'éducation ouverte en Pologne. Elle organise différents événements tels que des conférences, séminaires et ateliers, ayant pour but de renforcer la coopération entre les organisations engagées dans des projets similaires de création de contenus éducatifs ouverts. Il est important de souligner que son activité s'étend jusqu'au niveau gouvernemental où elle s'investit dans de nombreuses initiatives afin de sensibiliser les autorités à l'enjeu de la réalisation d'une éducation ouverte et des contenus ouverts, ainsi que pour les convaincre de l'intérêt commun de ces initiatives. Le groupe a élaboré des « Recommandations au libre accès pour la science » *Rekomendacje open access dla świata nauk* ainsi qu'un « Guide des contenus éducatifs ouverts » *Przewodnik po Otwartych Zasobach Edukacyjnych*.

Depuis 2009, en coopération avec la Bibliothèque nationale, elle organise la « Journée du domaine public » où il est possible d'assister à des présentations d'auteurs, dont des ouvrages tombent dans le domaine public. Les informations proviennent des fichiers d'entrées des bibliothèques.

4. Aspects juridiques

Le développement des archives ouvertes en Pologne est souvent freiné par des questions liées au droit d'auteur. La mise à disposition en ligne des contenus numériques est régie par la Loi sur le droit d'auteur et droits dérivés du 4 février 1994⁸, selon laquelle la diffusion publique sans limite ne concerne que des ouvrages dont les droits patrimoniaux de l'auteur ont expirés. Dans les faits, il est parfois très difficile – voire impossible – de l'établir comme par exemple dans le cas des ouvrages orphelins, ou bien, celui de ceux qui sont épuisés.

L'art. 28 de la Loi sur le droit d'auteur admet « une exploitation gratuite autorisée » par les bibliothèques, archives et écoles pour les exemplaires et les ouvrages protégés par ce droit, mais uniquement sur place. Conformément à cette loi, la diffusion de la reproduction numérique en ligne effectuée par une bibliothèque, des archives ou une école ne peut pas être considérée comme exploitation autorisée. Ce qui limite considérablement les possibilités de mise à disposition des contenus numériques lorsqu'il s'agit d'ouvrages ou autres documents publiés aux XX^e et XXI^e siècles, qui sont les plus consultés. Le développement de la coopération du milieu scientifique et des bibliothèques, soutenu par l'application des outils juridiques disponibles pourrait néanmoins contribuer à la création de collections de bonne qualité. Une solution pourrait consister en l'utilisation de licences universitaires (entre autre par la Bibliothèque numérique de Poméranie *Kujawsko-Pomorska Bibliotekę Cyfrową*) ou bien, de licences CC. Ces deux types sont fréquemment utilisés par les bibliothèques numériques. Malheureusement, et ce malgré l'adaptation des licences CC au système juridique polonais, certains experts en Droit mettent en question leur conformité à la loi polonaise dont les dispositions juridiques très strictes empêchent les auteurs de disposer librement de leur propre œuvre. La portée des licences universitaires est pourtant très souvent plus étroite, et surtout, elles ne garantissent aucun droit aux utilisateurs d'un ouvrage mis à disposition par une bibliothèque, mis à part celui de l'auto-exploitation. Ces licences diffèrent également quant à la qualité juridique, ce qui s'explique par leur efficacité et leur portée. Un autre problème est le manque de standardisation des conditions de mise à disposition des ouvrages par les bibliothèques numériques. S'y ajoute également la crainte qu'ils puissent entraver un futur échange de contenus, ainsi que la fusion des bibliothèques ou bien, la mise en œuvre des évolutions technologiques.

Grâce à l'accord des auteurs et des éditeurs, certaines bibliothèques numériques réussissent à mettre à disposition certains textes scientifiques, notamment des articles :

⁸ J.O. 2000, no 80, position 904 /Dz. U. 2000 r. Nr 80 poz. 904/.

preprints, postprints et des productions universitaires. Pourtant, au sein des bibliothèques polonaises prévalent les ouvrages dont les droits d'auteur ont déjà expiré et qui se trouvent donc dans le domaine dit public.

Remerciements

Merci à Joana Janick pour son aide lors de la traduction de ce texte.

Ressources pour l'accès à l'information scientifique publiée en Espagne : le scénario actuel, les projets et les tendances dans la diffusion de l'IST

Teresa Abejón

CSIC-CCHS, 26-28 calle Albasanz
Madrid, Espagne

Résumé

Le but de cette communication est de présenter un panorama actuel des systèmes d'accès à l'IST en Espagne. Les bases de données bibliographiques et les catalogues de bibliothèques spécialisées ont été longtemps les seuls outils qui permettaient aux chercheurs de localiser l'information scientifique, mais ces dernières années en Espagne, les portails d'accès à l'édition électronique et les dépôts de publications numériques se sont développés.

Cette contribution analyse les tendances actuelles dans la diffusion institutionnelle des contenus scientifiques et les problèmes que cela pourrait entraîner dans un accès à l'IST centralisé. La nouvelle Loi sur la science, la technologie et l'innovation envisage déjà le risque de fragmentation de l'information disponible sur les plate-formes diverses et prévoit que le ministère de la Recherche facilite un accès centralisé aux ressources. Il reste encore beaucoup à faire pour aboutir à une véritable articulation des systèmes d'information axés sur la diffusion des dépôts institutionnels.

1. Introduction

Bien que notre titre soit « Ressources pour l'accès à l'information scientifique publiée en Espagne », il aborde particulièrement les ressources d'IST qui se rattachent à l'accès libre aux contenus.

À ce sujet, tout d'abord, on exposera un portail qui contient une information assez complète sur l'accès libre en Espagne et on montrera quelques données sur le développement des dépôts de publications numériques en accès libre et trois exemples d'entrepôts basés sur le protocole OAI-PMH.

L'adhésion au mouvement pour le libre accès des organismes de recherche et des établissements d'enseignement supérieur (avec la signature de la déclaration de Berlin) et le développement consécutif des ressources d'information en accès libre ont marqué clairement les tendances dans l'organisation des ressources d'IST en Espagne.

D'autre part, il y a quelques limites aux dépôts d'archives ouvertes, aux entrepôts OAI et aux portails de journaux scientifiques en accès libre comme ressources d'information scientifique. On exposera quelques problèmes associés à ces limites.

Finalement, on envisagera ici le scénario possible qui pourrait s'ouvrir avec la nouvelle Loi de la science.

On pourrait dire que dans le panorama actuel des ressources produites en Espagne, si nous le comparons avec un écosystème, subsistent quelques « animaux préhistoriques » (tels les bases de données bibliographiques ou les catalogues des bibliothèques spécialisées) qui se sont adaptés à un nouveau contexte (Internet, existence de publications numériques et numérisées, la possibilité d'intégration avec d'autres ressources, etc.) mais ce sont les nouvelles espèces qui ont le rôle principal. Notamment, en Espagne, les sites Web ou portails de journaux scientifiques en accès libre, les dépôts institutionnels en accès libre ou encore quelques entrepôts basés sous le protocole OAI.

Ce travail aborde surtout des dépôts institutionnels en accès libre parce qu'ils sont un des enjeux qui mobilisent la politique de financement et d'organisation des ressources d'information des établissements d'enseignement supérieur et de recherche.

2. L'accès libre en Espagne

2.1 Le portail « *AccesoAbierto.net* »

Le portail « *AccesoAbierto.net* » est une ressource d'information sur l'accès libre en Espagne créé par un groupe de recherche qui appartient au CSIC¹. Ce groupe est dirigé par une chercheuse qui est très attachée au mouvement pour le libre accès, comme scientifique et comme éditrice d'une revue savante.

Ce portail offre des informations complémentaires aux sites d'information comme OpenDOAR ou ROAR (*Registry of Open Access Repositories*), mais essentiellement sur les dépôts d'archives ouvertes espagnols. Le portail contient :

- le répertoire des dépôts de ces archives « *BuscaRepositorios* » (*RechercheDépôts*), qui est semblable à OpenDOAR mais avec des champs d'information légèrement différents ;
- la base de données « *Dulcinea* », très semblable à ROMEO, qui recense les politiques des éditeurs des journaux espagnols en matière d'archives ouvertes et contient les titres et liens URLs aux journaux scientifiques espagnols en libre accès ;
- « *Melibea* », très semblable à JULIET, qui identifie et analyse les politiques mises en place par les financeurs de la recherche en matière d'archives ouvertes, mais pas seulement au niveau de l'Espagne, mais aussi au niveau international.

¹ Le Conseil supérieur de la recherche scientifique, ou CSIC (*Consejo Superior de Investigaciones Científicas*), est le principal organisme public de recherche en Espagne. Il est une institution jumelle au Centre national de la recherche scientifique.

2.2 Développement des dépôts de publications numériques en accès libre en Espagne (2001-2011)

Selon le répertoire OpenDOAR, il y a 76 dépôts de publications numériques en accès libre en Espagne, selon ROAR il y en a 93 et selon « BuscaRepositorios », le répertoire espagnol, il y a 75 dépôts. Le chiffre le plus fiable est 76, parce que le *Registry of Open Access Repositories* contient des sites qui sont en réalité le résultat du moissonnage des données d'autres archives ouvertes et de quelques portails qui ne sont pas vraiment mis à jour.

De toute façon, si on compare ce chiffre de 76 dépôts en Espagne avec les chiffres pour la France, 62 selon OpenDOAR et 59 selon ROAR, l'extraordinaire développement de ces ressources d'information en Espagne est flagrant.

La figure 1 résume le développement des dépôts de publications numériques en accès libre en Espagne (2001-2011). L'année présentée ici correspond au moment où le dépôt a commencé à être opérationnel.

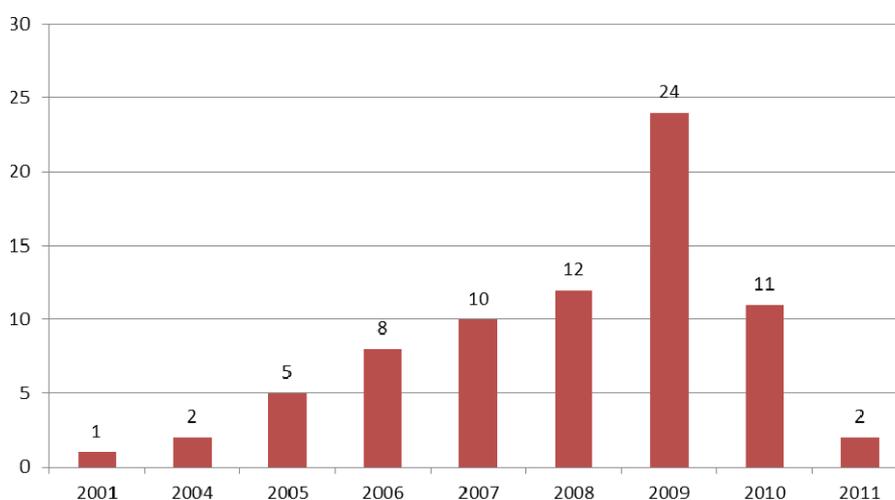


Figure 1 : Développement des dépôts de publications numériques en accès libre en Espagne

Le dépôt le plus ancien, en 2001, a été créé par le consortium des Bibliothèques universitaires de la Catalogne. Il s'appelle *Tesis Doctorals en Xarxa* (thèses de doctorat en réseau). En 2004 existait déjà le premier dépôt institutionnel en libre accès : celui de l'Université Complutense de Madrid, qui s'appelle « E-Prints ». La même année Dialnet démarre : un projet de l'Université de la Rioja qui se maintient avec la collaboration d'un grand réseau de Bibliothèques. En 2008 débute le dépôt « DigitalCSIC ». En 2009 on peut voir l'éclatement de ces ressources, avec des projets qui ne viennent pas seulement des universités et des centres de recherche. On peut trouver des dépôts des départements de la Culture des Communautés autonomes. En 2010 la plupart des nouveaux dépôts appartiennent aux institutions de l'enseignement supérieur. Et finalement ceux de 2011 appartiennent aussi aux universités

2.3 Prédominance des dépôts institutionnels des organismes de recherche et des établissements d'enseignement supérieur

La figure 2 présente le type d'institution selon le répertoire « BuscaRepositorios ». On constate une nette prédominance des dépôts institutionnels des organismes de recherche et des établissements d'enseignement supérieur et surtout des universités publiques, mais aussi des dépôts de trois universités privées.

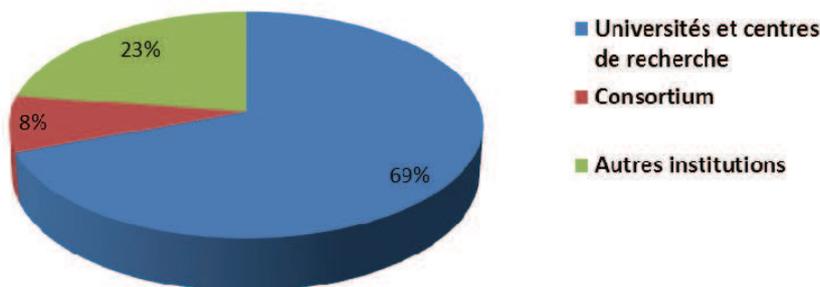


Figure 2 : Type d'institutions selon le répertoire « BuscaRepositorios »

Parmi les dépôts appartenant aux consortiums on peut en souligner quatre qui appartiennent au consortium des Bibliothèques universitaires de Catalogne : un qui recueille des thèses, un autre des documents pour l'enseignement, un autre de collections numérisées de photos, cartes, dessins, etc., relatifs à la Catalogne, et le dernier : « RACO » contient des revues catalanes en libre accès. Il est un des dépôts en accès libre le plus important d'Espagne en nombre de documents contenus.

Parmi les dépôts signalés comme appartenant à d'autres types d'institutions, on note celui du ministère de la Culture, ainsi que ceux de plusieurs départements de la culture et des bibliothèques régionales des Communautés autonomes, de fondations, d'académies ou d'associations professionnelles.

2.4 Types de contenus des dépôts en accès libre

La figure 3 présente des données concernant le type de contenus en libre accès. On peut remarquer que la plupart des dépôts qui contiennent des articles de revues, presque 50 dépôts, poursuivent les efforts des éditeurs de publications scientifiques des institutions afin de numériser les revues les plus anciennes. Souvent, les nouveaux titres sont électroniques dès la création de la revue. On remarque l'adhésion de beaucoup de ces éditeurs au mouvement pour le libre accès.



Figure 3 : Type des contenus des dépôts en accès libre.

Il est intéressant d'observer que l'accès libre s'attache aussi à ce que nous appelions auparavant la « littérature grise », avec de nombreux dépôts de thèses et de congrès. Des documents inédits de différente nature appartenant aux archives institutionnelles et des documents d'enseignement (manuels, cours, travaux dirigés, etc.) sont présents en quantité plus limitée.

D'autre part, selon les répertoires consultés, il y a un seul dépôt qui contient des données de la recherche : « Digital CSIC ». Il s'agit d'un ensemble de données sur le climat alimenté par un centre de recherche sur l'agriculture. Récemment, un institut universitaire de recherche linguistique y dépose ses données de la recherche. Il s'agit de données incluses dans les annexes des thèses, ainsi que des résultats des projets de recherche. Si d'autres organismes de recherche s'engageaient dans ces initiatives pionnières en Espagne, les dépôts institutionnels deviendraient des infrastructures scientifiques bien plus importantes.

Finalement, on peut souligner qu'il n'y a pas de logiciels en libre accès dans ces dépôts.

2.5 Les entrepôts OAI. Trois exemples de succès de l'interopérabilité des entrepôts via l'OAI-PMH

Le moissonnage de données des dépôts en accès libre permet de créer, d'alimenter et de tenir à jour automatiquement des entrepôts où les métadonnées de sources diverses peuvent être interrogées simultanément.

Au niveau national, l'entrepôt le plus important comme source d'information scientifique est le portail « Recolecta », que l'on peut traduire par « ramassage ». *Recolecta* est une initiative conjointe de REBIUN (Réseau des bibliothèques universitaires et de la recherche espagnoles) et de la Fondation espagnole pour la science et la technologie (FECYT), qui appartient au ministère de la Science et de l'Innovation. Les ressources sélectionnées sont principalement des journaux scientifiques publiés en accès libre par les universités et des dépôts institutionnels universitaires. Il s'agit d'une ressource vraiment remarquable par le

volume d'information et l'intérêt des contenus. Nous verrons ci-dessous ses limites en tant que ressource d'information.

« Hispana » est une ressource nationale plus orientée vers le patrimoine et la culture espagnole en général, mais elle peut être très utile pour d'autres domaines scientifiques. Cet entrepôt rassemble les collections numériques des archives, bibliothèques et musées dans le cadre de l'*Open Archives Initiative* et il a des fonctions de dépôt numérique similaires à ceux d'Europeana. Parmi les ressources sélectionnées, on trouve les dépôts institutionnels des universités espagnoles et les bibliothèques numériques des Communautés autonomes qui offrent l'accès à une masse croissante de nombreux types de documents du patrimoine bibliographique espagnol.

« *Portal de portales Latindex* » permet d'accéder aux contenus des revues en texte intégral disponibles dans les bibliothèques numériques d'Amérique latine, des Caraïbes, d'Espagne et du Portugal qui sont rattachées au mouvement du libre accès. L'objectif de ce portail est de diffuser des connaissances scientifiques publiées dans la région latino-américaine grâce au protocole OAI Harvester2 développé par PKP et basé sur le protocole OAI-PMH.

Ce portail documentaire est un projet établi en relation avec « Latindex », qui est un réseau de travail ayant fixé les critères de qualité pour les journaux scientifiques ou académiques d'Amérique latine, d'Espagne et du Portugal, et qui alimente un répertoire de revues. Au sein du *Centro de Ciencias Humanas y Sociales* du CSIC, nous coopérons au répertoire et aussi nous utilisons les critères de qualité Latindex pour sélectionner les revues analysées dans nos bases de données.

Portal de portales n'est pas une initiative espagnole, mais la source qui fournit la plus grande quantité de contenus : Dialnet l'est elle. *Portal de portales* a été créé au Mexique par l'Universidad Autonoma. En plus de Dialnet, le portail recueille les métadonnées de bibliothèques numériques d'Amérique Latine très importantes comme Scielo ou Redalyc, qui sont des ressources clés pour l'information scientifique produite en espagnol.

3. Tendances dans l'organisation des ressources d'information scientifique

Actuellement les institutions scientifiques et d'enseignement supérieur commencent à réorganiser leurs dépôts institutionnels et quelques-unes commencent à regrouper l'ensemble de leur production de recherche, leurs publications, leur documentation administrative, etc. par l'intermédiaire de l'intranet ou de l'accès libre. Dans cette réorganisation les ressources en accès libre jouent un rôle important pour la visibilité de l'institution.

Le dépôt GREDOS, de l'Université de Salamanque, est un exemple récent du regroupement de plusieurs ressources, il est ainsi devenu le dépôt institutionnel en accès libre au plus grand contenu en Espagne. Il se divise en quatre collections :

1. l'archive institutionnelle ;
2. la bibliothèque numérique, contenant des documents historiques et des collections numérisées ;

3. le dépôt dit scientifique ;
4. le dépôt pour l'enseignement.

Le dépôt scientifique regroupe plusieurs types de documents : les documents produits par les chercheurs de l'université (c'est-à-dire, la production scientifique), les revues éditées par l'université et les congrès organisés par l'université.

Suivant cette tendance de mettre ensemble toutes les ressources d'information de l'institution, le Conseil supérieur de la recherche scientifique vient d'approuver au mois de mai 2011 un document qui établit une nouvelle organisation des ressources d'information scientifique : le Système d'information scientifique. Ce projet veut constituer une nouvelle structure afin de réunir au sein d'une même plate-forme toutes les ressources du réseau des bibliothèques, le dépôt institutionnel en accès libre, les deux portails de publications édités par l'institution, la base de données de la production des chercheurs *ConCiencia* et le dépôt multimédia.

Dans le document on dit que le projet poursuit un modèle d'intégration pour manager et « communiquer », c'est-à-dire pour diffuser, la production scientifique de l'institution.

4. Quelques limites aux dépôts d'archives ouvertes et aux entrepôts OAI

Les dépôts en accès libre et les entrepôts OAI présentent quelques limites en tant que ressources d'information scientifique :

- risque de manque de précision dans les recherches ;
- risque de couverture peu cohérente ou de perte d'exhaustivité des documents recensés ;
- risque de ressources peu actualisées ;
- risque de manque de coordination interne des institutions.

Au sujet de la récupération d'information, il y a des limites, comme le risque de manque de précision dans les recherches, qui sont surtout dues à la qualité des métadonnées et quelques fois aux implémentations réalisées. Il y a des exemples de mauvais fonctionnement d'un entrepôt avec une interface correcte mais avec des métadonnées de qualité limitée.

Existe, d'autre part, un risque de couverture peu équilibrée, puisque beaucoup de ces dépôts se développent selon les dépôts volontaires des chercheurs ou que les professionnels de l'information voués à ces projets réalisent cette tâche. Par ailleurs pour certains domaines, on peut dire que ces ressources, même les grands entrepôts, sont incomplètes parce qu'un haut pourcentage de publications spécialisées, essentielles au domaine, sont publiées par des éditeurs privés et ne sont pas en accès libre. C'est le cas, par exemple, des revues sur le Droit en Espagne.

Une autre limite, pour les dépôts, est le risque de ressources peu actualisées puisque l'archivage dépend directement des politiques d'accès aux publications, de la volonté et de la sensibilisation des chercheurs, ainsi que de l'implication des professionnels en IST. Quelques

revues, par exemple, lèvent l'embargo d'accès aux contenus un an après la date de publication.

Parmi les problèmes associés aux limites des ressources, on peut trouver des difficultés lors de l'intégration de celles-ci au niveau institutionnel. Pour le développement de plates-formes comme GREDOS ou le Système d'information scientifique du CSIC, qui ont un accès centralisé aux ressources d'information propres, les différentes caractéristiques techniques des ressources et le besoin de coordination interne de plusieurs départements des institutions (des bibliothèques, des services de publications, etc.), peut être un handicap. Donc, il faut faire un effort de coordination important.

D'autre part, on parle beaucoup en Espagne de l'utilisation des dépôts institutionnels en accès libre à des fins d'évaluation de la recherche. Toutefois, il faut rappeler la faible fiabilité de ces ressources et quelques problèmes comme l'utilisation exclusive du nombre de téléchargements comme indicateurs d'usage.

Par ailleurs il faut tenir compte du fait que l'utilisateur peut oublier ou ne pas connaître les limites de ces ressources et plus apprécier la gratuité, la vitesse, la rapidité et la facilité d'accès que la pertinence, la précision, l'exhaustivité ou encore la fiabilité des ressources. Il faut donc faire un important travail de formation en information documentaire dès les premières années d'enseignement à l'université.

Enfin, la question du recouvrement de ressources dans un contexte de réduction des budgets publics s'impose, car les institutions, en voulant améliorer la visibilité de leur recherche et suivre la production scientifique des chercheurs et enseignants-chercheurs, donnent maintenant la priorité à ces projets de dépôts en libre accès et aux entrepôts. Le recouvrement de ces ressources pourrait remettre en question le besoin de valeurs ajoutées telles que l'alimentation, la maintenance ou l'acquisition toujours plus onéreuses.

Quant aux limites de la recherche interoperable, voici l'exemple d'un mélange de métadonnées de différente nature avec *Recolecta*, interface qui offre une recherche avancée par vedettes-matière, mais où le renseignement n'existe pas. La référence contient en réalité les mots clés fournis par les auteurs, puisque *Recolecta* moissonne essentiellement les métadonnées de revues en accès libre et les documents provenant de l'auto-archivage. En fait les résultats obtenus correspondent à une recherche par titre, résumé ou mots clés d'auteur.

Il serait possible de résoudre certains problèmes grâce à un bon développement des vocabulaires ou des ontologies intégrées dans la recherche, mais actuellement ce n'est pas le cas.

5. Le scénario futur : la nouvelle loi de la science

La nouvelle « Loi sur la science, la technologie et l'innovation »², adoptée en juin de cette année, envisage déjà le risque de fragmentation de l'information disponible sur les multiples plates-formes et prévoit que le ministère de la Recherche facilite un accès centralisé aux ressources, mais, à mon avis, elle ne prévoit pas de solutions pour les problèmes et les limites précédemment exposés.

L'article 37 relatif à la diffusion en accès libre veut promouvoir le développement des dépôts pour les publications du personnel de recherche, cependant la mise à jour des ressources reste un problème, car la loi établit que « Le personnel de recherche rendra publique une version numérique de ses travaux dès que possible mais pas plus tard que douze mois après la date officielle de publication ».

D'autre part l'article dit que « La version électronique sera publiée en archives en libre accès reconnues dans le domaine de la connaissance dans laquelle la recherche s'est développée, ou dans des dépôts institutionnels en libre accès » tandis qu'en même temps il y a des exceptions en fonction des droits d'auteur et de la protection des données. Cela complique le fait d'avoir l'exhaustivité la plus fiable, la plus complète et la plus rapide possible dans un domaine de connaissance donnée.

6. Conclusion

Les institutions, en voulant améliorer la visibilité de leur recherche et en appliquant la nouvelle loi de la Science et suivre, donnent maintenant la priorité aux projets de dépôts en libre accès et aux entrepôts.

En conclusion, on peut dire que malgré l'extraordinaire développement de ressources d'information scientifique en accès libre, il reste encore beaucoup de chemin à parcourir afin d'aboutir à une véritable articulation des systèmes d'information axés sur la diffusion des dépôts institutionnels. Il faut soutenir les projets collaboratifs au niveau national, européen et international. Et, d'autre part, pour les services que nous alimentons de ressources d'information dites « traditionnelles », le défi est peut-être d'explorer de nouvelles alliances pour disposer de ressources fiables, correctement soutenues et convenablement financées par les institutions.

Remerciements

Je voudrais remercier l'organisation des journées FRéDoc 2011 pour l'invitation à participer dans la session vouée aux pratiques et projets IST en Europe. Aussi je souhaite adresser mes remerciements les plus sincères à Geneviève Peyres, pour son aide et son travail de révision et de correction des diapositives et du texte de présentation. Je la remercie pour sa patience mais aussi de m'avoir encouragée à faire cette présentation en français.

² Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la Innovación.

Bibliographie

- [1] E. Abadal, F. Abad, R. Melero, J.M. Rodríguez-Gairín, Acceso abierto a la ciencia. Informe 2009 y prospectiva. Anuario ThinkEPI. **4**, 285-292 (2010).
- [2] P. Castro, « Digital.CSIC: el desarrollo de un repositorio institucional científico en el CSIC » dans *Interinformación: XI Jornadas Españolas de Documentación*. Ed. FESABID, Madrid. p. 249-256, 2011.
- [3] J.-C. Guédon, It's a Repository, it's a Depository, it's an Archive.: Open Access, Digital Collections and Value. Arbor. **185(737)**, 581-595 (2009).
- [4] R. Melero, E. Abadal, F. Abad, J.M. Rodríguez-Gairín, The situation of open access institutional repositories in Spain: 2009 report. Information Research. **14(4)**, paper 415 (2009).
- [5] R. Melero, El paisaje de los repositorios institucionales open access en España. BiD: textos universitaris de biblioteconomia i documentació. **20**, (2008).
- [6] E. Wulf Barreiro, Acceso libre a las publicaciones y jerarquía de la credibilidad. Revista General de Información y Documentación. **20**, 179-188 (2010).
- [7] Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la Innovación, Boletín Oficial del Estado, **131**, 54387-54455 (2011)
- [8] <<http://www.accesoabierto.net/>> consulté en septembre 2011>.
- [9] <<http://www.opendoar.org/>> consulté en septembre 2011.
- [10] <<http://roar.eprints.org/>> consulté en septembre 2011.
- [11] <<http://www.recolecta.net/buscador/>> consulté en septembre 2011>.
- [12] <<http://roai.mcu.es/es/inicio/inicio.cmd>> consulté en septembre 2011.
- [13] <<http://www.latindex.ppl.unam.mx/>> consulté en septembre 2011.
- [14] <<http://gredos.usal.es/jspui/>> consulté en septembre 2011.

Clouds and crowds in the Netherlands Library Landscape

Marian Koren

Head of Research and International Affairs
Netherlands Public Library Association
Grote Marktstraat 43, 2511 BH The Hague, Postbox 16146,
2500 BC The Hague, Netherlands
koren@fobid.nl

Abstract

Focussing on the political and scientific information policies in the Netherlands, on innovation, information infrastructure and digital library.nl, what consequences do they have for Dutch libraries and how are they coping?

Clouds and crowds, one may answer. The clouds refer to the dark clouds of budget cuts, which are affecting research at university level, education in schools and public libraries at the local level. Or it may refer to the increasing innovation of working, cataloguing, delivering "in the cloud". The crowds are the large variety of audiences to be served, or also: the many smaller and larger bodies of cooperation in the library field, among them FOBID¹ Netherlands Library Forum. As an umbrella organisation, it focuses on national policies and cross overs between different types of libraries. Expert committees are active in the fields of legal matters, library technical matters (cataloguing/metadata), professional education and international affairs, related to EBLIDA² and IFLA³.

FOBID serves as the international Office for professional exchanges and information on visits, but it is not a project bureau. FOBID as connector and advocate for libraries creates opportunities through a variety of relationships and networks.

Résumé

Des nuages et des foules dans le paysage des bibliothèques néerlandaises

Pratiques et projets scientifique et techniques en Europe⁴, FOBID (Forum des bibliothèques néerlandaises), Pays-Bas

Quelles sont les conséquences de la politique de l'information scientifique aux Pays-Bas sur les bibliothèques néerlandaises en matière d'innovation, d'infrastructure de l'information ainsi que sur bibliothèques numériques. Comment y font-elles face ?

¹ <<http://www.fobid.nl>> consulted on 15 May 2012.

² <<http://www.eblida.org>> consulted on 15 May 2012.

³ <<http://www.ifla.org>> consulted on 15 May 2012.

⁴ Note du traducteur (Richard Dickinson, INIST) : imiter le titre du roman de Steinbeck *Des souris et des hommes* m'a paru la solution la plus adaptée pour rendre ce jeu de notion sur *clouds* et *crowds* passant des acceptions informatiques à des métaphores sur les situations budgétaires et associatives du secteur dans ce résumé.

Nuages et foudres sont les mots qui viendraient à l'esprit pour répondre à cette question. Les uns pour qualifier un ciel chargé en coupes sombres qui frappent les budgets de la recherche universitaire, mais aussi l'enseignement scolaire et, localement, les bibliothèques publiques. Quant aux « foudres », il s'agit de ces nombreux groupes de coopération, plus ou moins développés, qui interviennent dans le domaine des bibliothèques. Parmi eux, FOBID. En tant qu'organisation « parapluie », elle se concentre sur les politiques nationales et les échanges entre les différents types de bibliothèques. Ses comités d'experts traitent des questions juridiques, techniques en bibliothéconomie (catalogage et métadonnées), de la formation professionnelle et des problématiques internationales, en lien avec l'EBLIDA et l'IFLA. FOBID sert d'agence internationale pour l'information et les échanges professionnels lors de visites, sans toutefois être une agence de projets. En tant que vecteur et représentante des bibliothèques, FOBID crée des opportunités en s'appuyant sur un grand nombre de relations et de réseaux.

As all societies are on the move and looking for directions, the Netherlands are no exception. The last decade has been one of rapid changes, due to information technology, globalisation and changes in user expectations and demands. They all affect the different types of libraries, in the performance of their mission, their services and relation to users and communities. How to describe such a complex development which has not come to an end, and probably never will. Are we living in a liquid society, as Zygmunt Baumann [1, 2] describes, then we will never encounter a fixed situation – something many of us, professionally and personally are already now experiencing.

The main actors in this moving Dutch library landscape are presented here in a developmental context. As the Dutch cherish their freedom and independence, one may be astonished about the number of organisations, institutions and groupings working in the information and library field, causing an abundance of abbreviations, for which I can only warn the international reader.

In the Dutch climate for libraries a number of positive aspects should be mentioned. The main one is the stress on innovation. In 2000, due to a major state-initiated innovation programme, all public libraries were connected to the Internet. Librarians were trained, in order to train citizens, groups of seniors, jobless people, immigrants, women and young people, to use the internet and new media for their benefit, needs for information, culture, curiosity, learning and personal development. Those first innovative steps caused a range of new orientations in all types of libraries. Students were trained to find online sources, to build up digital portfolio's etc. Within a decade, all working in libraries are convinced that they have to invent the library again or it will die. This sense of urgency to find new models of service, and to train staff in developing new skills, has been on the forefront of the Dutch library and innovation debate. Increasingly it was understood, that libraries could no longer just be a service on their own: collection – connection was the new orientation: put your library sources, your collection in context, otherwise it will be forgotten. Looking at social and sector developments, the government stated in the Library Charter a new vision:

- Social and technological developments are radically changing the way citizens use information and media. More and more information is being made available – not just in print but also digitally. In fact, a growing amount of information is now only

available digitally (born digital). Citizens have a growing need for digital information, and this is also affecting the position of the public libraries;

- Besides the physical public library ("a building with books, newspapers and magazines"), the digital public library is now also on the rise. The library of the future consists of two mainstays: the physical library and the digital library;
- The public library is providing new digital services and can also provide existing services (for example, media lending) in digital form;
- The digital library is under development. On a national level, the digital library is being constructed on the basis of content-related, system-related and efficiency considerations. In addition, part of the management and maintenance is also taking place on a national level;
- Digital products and services developed and agreed by parties on a national level will then be introduced in the local/regional library organisations;
- During this period, the development of the digital library will be the main area of innovation. Innovation will also take place in the collection policy, HRM and marketing.

The overall and leading idea is the need for a common and integrated Information infrastructure and digital library. The ideas were developed in the communities of public libraries, in university libraries and by the national library, *Koninklijke Bibliotheek*.⁵ Gradually these ideas found their way across these communities into a balanced concept of a single information infrastructure and architecture, a single national library catalogue, a digital library.nl. Recent budget cuts of which all, but especially the local public libraries (since 2010) are suffering, have not diminished the need for a common and coherent innovative library and information policy and service. How it will come about, will be the challenge in the years 2012 – 2015 to face. In order to understand the complex situation better, the main types of libraries, their role, organisation and development are described here.

1. Public libraries

Public libraries have been introduced around 1900, on initiative of well to do liberals, church groups and the labour organisations. This diversity in the roots of the public library service disappeared in the 1950ies when all libraries became truly public without denomination, are professionally run libraries and members of the Netherlands Public Library Association⁶ (VOB). The public library network in the Netherlands comprises approximately 170 public library organisations and 11 provincial library service organisations with a total of 900 branches and 550 service points. One characteristic did not disappear: most public service organisations are not a department of the public administration but kept their original legal body as association or nowadays foundation. This library foundation has a board of private persons; the foundation receives the public subsidy for the running of the public library. The library still has to cover around 15-19% of own income. Therefore, almost all public libraries work on subscription fees from the users.

⁵ Koninklijke Bibliotheek, National Library <<http://www.kb.nl>> consulted on 15 May 2012

⁶ <<http://www.debibliotheken.nl>> consulted on 15 May 2012

Regarding the public library system and administration: currently, there is no extensive library legislation. The Public Library Act 1975 introduced a three-layered library system and fee-free services for children up to 18 years of age. Decentralisation legislation in the 1980ies have brought the libraries under municipal responsibility. The three government levels: national, provincial and municipal are obliged to cooperate on library infrastructure and services. The provinces subsidise provincial service organisations which support municipal libraries in the field of e.g. management and training, finances, IT-networks and innovation, collection and service development, outreach services etc. Provincial subsidies may also include additional academic literature service in some larger public libraries (*Plusbibliotheken*).

At the national level, the government has assigned a number of national tasks (with funding) to the sector institute for public libraries (SIOB). This Netherlands Institute for Public Libraries⁷ has its origins in the Netherlands Public Library Association. Initiated by the legislature this was split in 2010 into three organisations. Besides the Institute (SIOB), these are the professional association (the Netherlands Public Library Association (VOB) and the project organisation: *Bibliotheek.nl*⁸ (Digital library products and services organisation). The reason for dividing the original association – which performed advocacy, developmental and governmental tasks – was that responsibility for the network as a whole does not always run along parallel lines with the representative functions of the professional organisation.

This change at the national level was the outcome of a process of reorientation in which the central government decided after a period (until 2000) of mere absence – due to decentralisation – to encourage innovation in public libraries. A process of up scaling was started to create larger and stronger library organisations. So the number of legal bodies for libraries, non profit foundations, has decreased in the period 2000-2010 from 542 to 166.

The sector institute was set up to create a strong network of public libraries. As digital developments have given customers new expectations of libraries, their needs are different from those of previous generations. The library sector has to respond to these needs if they wish to continue attracting visitors. A national strategy is necessary for coherence and to avoid that every library is choosing its own method for change. Commissioned by the Ministry of Education, Culture and Science⁹, the Netherlands Institute for Public Libraries (SIOB) coordinates all plans aimed at renewing and strengthening the library sector. It performs the following network tasks: (Inter)national promotion and representation; Information, reflection, debate and education; harmonization and coordination and includes also services for adapted reading for the blind and visually impaired. The Institute distinguishes seven programmes in carrying out these tasks.

Representation and cohesion includes the main tasks of maintaining (inter)national contacts, library advocacy, contributing to cooperation within the library network and with related sectors, as well as promoting innovative networking. In addition the Institute advises the Ministry on the subject of new library legislation and investigates the effects of this legislation on the sector and the way it functions.

⁷ <<http://www.siob.nl/>> consulted on 15 May 2012

⁸ <<http://www.debibliotheken.nl/>> consulted on 15 May 2012

⁹ <http://www.rijksoverheid.nl/ministeries/ocw?ns_campaign=M-OCW&ro_adgrp=OCW_algemeen&ns_mchannel=sea&ns_source=google&ns_linkname=ministerie%20ocw&ns_fee=0.00> consulted on 15 May 2012

The Institute also runs a research and knowledge programme, mapping the position of the library sector in total, with its strengths and weaknesses, its threats and opportunities. This analysis forms the basis on which the other programmes are formulated.

Digital innovation is vital for all libraries. The programme works for relevant content which should be accessible in the simplest possible way for as large an audience as possible. For that reason the BNL Foundation (*Bibliotheek.nl*) works on creating a national digital infrastructure for all libraries which will also include a National Library Catalogue. The Institute supervises and monitors *Bibliotheek.nl*.

The role of libraries in society is a programme working on relevant and intense relationships of libraries in the municipalities, making use of opportunities in the local communities. Cooperation with schools is particularly promising. Libraries need to stake a claim in contributing towards reading promotion, computer literacy and lifelong learning. An extensive development and service package supports libraries, especially in the field of reading promotion. This is done in cooperation with Foundation Reading (*Stichting Lezen*).

Library education needs much attention, as in the coming years an estimated 40% of the existing staff will leave due to retirement. An important challenge for the sector is to find the necessary new staff, and also to reflect on what competences are required of new employees. SIOB developed a competence index and looks for accredited education approaches, on both graduate and undergraduate level, and for library technicians.

The certification programme aims at guaranteeing the quality of library services. Standards have been laid down by the Netherlands Public Library Association (VOB) and the Association of Dutch Municipalities (VNG). The separate Netherlands Institute for Public Library Certification is responsible for the assessments. SIOB supports the Certification Institute by carrying out the actual work.

The public libraries are committed to a certification scheme. The first round of certification took place for the period 2006-2009. A separate foundation, *Stichting Certificering Openbare Bibliotheken* has been set up to organise and coordinate the audits and certification process. As many as 350 public libraries are assessed on the standards agreed upon by the public library sector and the government.

The public libraries have also agreed to use the INK Management model as the umbrella instrument for quality care. This means that libraries are supposed to implement quality care according to this model. In the same way, libraries have committed themselves to execute a position check and a customer satisfaction survey. The position check reveals strong aspects as well as point for improvement. The aim is to work on these improvement points according to an improvement plan.

Although quality care and certification are regarded as separate subjects they are closely related. Quality care could be considered a means to achieve the aim of certification.

In the year 2012, not only new standards for public libraries are adopted, but also work will start on the possible certification of the Provincial Service Organisations.

1.1 VOB

The Netherlands Public Library Association was established in 1908 and all public libraries are still members. It is an association which merged in 2010 with the Employers Association for Public Libraries. It means that the focus of the association – which has no individual professionals as members – is on employment issues, leadership and entrepreneurship for the benefit of libraries as organisation. Professional issues – if not of

importance for advocacy – are mainly left to the individual libraries, provincial service organisations and the sector institute. VOB has an office of around 11 fte.

The Association has only income from membership fees. The annual membership fees (around 46 € per 1.000 inhabitants of the service area/municipality) and additional contributions are used for advocacy and library promotion campaigns, the library portal *bibliotheek.nl* and a basic digital content package. A number of committees are active for strategy and advocacy, for interests in building up digital services, corporate marketing and branding of the library (*De bibliotheek*) and consortium tasks for buying digital content. E-book service is still under development. An agreement with the publishers association has been made, that publishers will be approached individually and asked to make their e-books available for services through libraries. NBD/Biblion is the central library supplier for all public libraries and provides all buying, binding and supply services necessary for libraries. It was set up in 1970 by the associations of publishers, booksellers and public libraries, to provide an efficient and cost effective library service. It delivers 3 million media annually, based on in house developed machinery and innovation of production processes to make media “library ready”.

1.2 *BibliotheekNL*

For setting up and maintaining the digital library services a special and separate development foundation was established in 2010: *Foundation BibliotheekNL*. Its main tasks are to work for and contribute to the national information infrastructure (GII), which also includes the National Library (KB) and the university libraries, which form a cooperation for licensing and access to scientific research sources (UKB)¹⁰. The KB has set the tone in its main strategy: Digitisation of all Dutch sources; one common infrastructure and general access for all users.

Instead of developing local websites one national White Label website is developed. It offers local libraries the opportunity to have a professional national website with all nationally organised services, which enlarges the local library offers. It also makes it possible to develop websites and digital services with other partners. The Digital Content Repository is the central place for gathered content, which is then used for websites, apps and widgets. Content connections have been made with the *Central Discotheek* in Rotterdam (music service), the National Library (KB) and the Digital Library for Dutch Literature. *Via de website*¹¹ users can shift with their local library and reverse. They only need one single log in. The portal gives access to the National Library Catalogue, and to Digital dossiers. Widgets for these KB-dossiers, and for Literature Plaza and Reading Plaza are available. Content is enriched with other sources, for example from public broadcasting.

In cooperation with the National Library, the foundation develops the National Library Catalogue, to combine the holdings of (eventually all) libraries in one national catalogue. The ambition of *Foundation BibliotheekNL* is to connect all public library organisations to the new infrastructure and services by 12-12-12. Contracts have been signed with OCLC to connect the Public Library collections – which are in a separate database – to the general catalogue system, as the basis for the National Library Catalogue including collections from all publicly

¹⁰ <<http://www.ukb.nl>> consulted on 15 May 2012

¹¹ <<http://www.bibliotheek.nl>> consulted on 15 May 2012

funded libraries. Other elements of the new digital library bibliotheek.nl are a widget store, digital show cases, a mobile website for demonstrations etc. The long tradition of public libraries to provide services to schools is now supported by the digital school library and projects with the national broadcasting service (NOS).

Back office services developed by *BibliotheekNL* include a data warehouse, marketing tools and a pilot for new ways of distribution of library materials.

1.3 Finances

The three levels of government are all involved in subsidizing the public libraries. The organisations themselves get around 15% of income through membership fees, borrowing fees and too late fees.

The general structure is as follows. The public library sector is decentralised, since 1987. It means that the municipalities have the main responsibility for the public library service. Municipalities pay around 457 mil € (2010), provincial/regional subsidies are around 53 million € and national subsidy is around 24 million €. The provincial subsidy consist of 40 million for the provincial service organisations (PSO) and 13 million € for the Plus libraries, these are municipal public libraries with a special task to deliver services including the level of higher education and first years of university studies. They form a go between the general level of public libraries and the university libraries.

A survey ranging from 1999 to 2009 shows an increase of municipal and provincial subsidies of around 50%. But of course, difference in bruto and netto subsidies are somewhat blurring the picture.

What is the national contribution? The national government has started to give a strong impulse for innovation in the public library sector. In 2012 around 6,1 million € is for the Sector Institute, recently established. It fulfils task for the whole of the public library system: and fulfils the tasks of cooperation and coordination; education, information and reflection, and (inter) national representation and promotion of the sector.

The sector institute has developed programmes around representation and coherence; research and knowledge sharing, digital innovation, societal relation, education and certification.

Separate tasks are SIOB's task the maintain the service for people with a print or reading disability: a subsidy of 11,3 million €, reduced to 10,5 million in 2013.

The national government continues it impulse for innovation. For 2012 17,4 million € is for a special task: digital innovation. (In 2013 this will be 17,1 million €). A special organisation: *Foundation Bibliotheek NL*, received until 2012 this innovation subsidy directly, now it has to work through the sector institute (SIOB).

Municipalities and some provinces announced budget cuts in 2010; some have been realised in 2011. It is unclear how the years 2012 and 2013 will show the effects of the financial crises affecting libraries. (Estimates indicate cuts ranging from 10-60%, whereby around 300 library establishments would disappear; the number of library service points in schools, institutes and community places would increase). In fact, a severe general cutting policy of around 20% for the whole art and culture sector is affecting all cultural institutions. The Sector institute, hardly started, faces a budget cut of 50%, leaving a subsidy of 3 million € for 2013.

The national government, the Ministry for Education, Culture and Science has committed itself to work for digital innovation, and to be responsible for the infrastructure.

The libraries themselves are buying digital content together. They pay an extra fee to the VOB, Netherlands Public Library Association, to buy content: In 2011 and 2012 this extra payment is 0,20 per inhabitant, meaning a sum of 3,2 million €. In order to make the organisation of the digital library more efficient, the national government has proposed (in agreement with the Association of Dutch Municipalities, VNG) to “take” a certain sum (15-25) million € out of the Municipal Fund in order to safeguard digital infrastructure and content. Apart from funds, lacking legislation and the publishers’ hesitation are the main obstacles for public libraries to buy e-books for lending services. As the associations of the public libraries and of the publishers could not come to an agreement on e-book lending (licensing), it was “agreed” that libraries could approach the publishers individually. The intention, also of the national government, is to have the digital library, including e-book service ready at the end of 2012. The Ministry will also look into appropriate possibilities in legislation.

1.4 New innovative approaches: application of retail concepts

A major discussion among Dutch public libraries has taken place (and continues) regarding the best way to ensure access to information, regardless of the format in which information is presented, and to maintain the principles of public service (non discrimination, democracy, transparency) and professional quality in services, collections and management, while likewise to provide evidence to (political) stakeholders of relevant service to citizens of all ages and walks of life.

Studies in other branches: bookshops, restaurants, banks, retail in general, have led to the development of a new library service concept based on retail applications. The national model includes applications of the new common public library logo in a wide variety of objects and materials, but has more to offer: a thorough study of current and potential users in the catchments area is fundamental to the type of services offered, the style of services and staff activities. The first library to present the new logo and retail concept was in a new branch of Zwolle; which proved to be successful from the beginning: enthusiastic comment from the increasing group of visitors and members; better display of titles, increasing loans, larger variety of requests. A separate order organisation (Biblionet) has been set up, supported by some provincial library service organisations (Groningen, Overijssel), to roll out the concept and sell it to all libraries.

Most libraries are positive to the concept but some keep their own house style on some communication forms, as they are part of the municipal administration or a cooperative multifunctional unit, e.g. a cultural centre or *kulturhus*.

Many libraries have no change to have a new building, but welcome the retail concept, as it proves to be a very attractive alternative for a makeover and innovation. One of the examples of a new library building and services (but without applying the common public library logo) is Almere, a new town project which started some decades ago, especially to offer housing for a lot of people from Amsterdam. Created in the land made province of Flevoland (a former sea Zuiderzee, now IJsselmeer), Almere is supposed to grow to 180.000 inhabitants. The new library (open March 2010) is a lively part in the inner city, a meeting place for multicultural Almere. Books and media are ranged according to themes, and have frontal presentations; easily accessible for reading on the spot, at home, for group work etc. A news cafe, a game circle for playing on the Wii or play station are other facilities. Rooms for teaching, lectures and theatre are available. Films are shown three times a week, alternating with theatre, debate or comedy. The example of Almere shows how successful a

new building and service concept can be: loans are up with 25%, visitors are up with 50%; 15 months after the opening the 1 million visitor could be welcomed!

As most libraries are facing budget cuts, announced since 2010, there is an extra challenge to come with innovative concepts. In spite of the economic crisis, some libraries manage to grow, not necessarily only in loans or members, but also in networking, service cooperation and local position. The Library Journal analysed best practices and their critical factors. (Bibliotheekblad 3, 2012). It shows that focus on network, digital innovation, presence of theatre and reading café, new (multifunctional) accommodation, retail concept and a mix of small and large scale development or even changes in the lending rules can help to grow as a library.

1.5 National strategy of libraries

A major strategic instrument in the ministerial innovation policy is the Library Charter (2009): It covers the period 2010-2012 and describes the distribution of roles between the parties, e.g. the local, provincial and national level of government, and describes agreements about the innovation of (public) libraries. The Charter is intended to gear the services of the public libraries more effectively to social and technological developments and to the needs of citizens. It includes around 20 million € a year to be spent effectively on library innovation, infrastructure and services.

Since the decentralisation of primary responsibility for public libraries to the municipal level, libraries have developed in various directions. Now, that changes in society, technology and economy are challenging libraries, it becomes important again to work according to a shared mission and common concept, using a common strategic agenda. Ultimately one may formulate this as reinventing the library, integrating the digital revolution, demonstrating the societal (added) value of libraries. During 2010-2011 several attempts and studies passed before the urgency was felt in the whole library community. Simply stated: this new agenda needs one recognisable library, one trade mark, one library pass, one catalogue, one infrastructure and one centre of expertise. It will require the commitment of all stakeholders to contribute to its realisation in the Netherlands. One of them are the academic libraries, the following section.

2. Academic libraries, UKB

Dutch universities have a long academic tradition. The first Dutch university was founded in Leiden in 1575 by prince William of Orange. Today the Netherlands has 14 research universities, all of which offer high-standard education and research. There are three universities of technology, in Delft, Eindhoven and Twente, and one university which focuses on agriculture and life sciences: Wageningen. The remaining ten research universities are broad-based. There are also eight university medical centres. The universities serve both education and research for around 230.000 students enrolled for an academic bachelor (430) or master study (890). The number of students has doubled since 1980. Every year around 50.000 new students enter the universities. In contrast to many other European countries, there is still a demographic wave in the Netherlands, causing an estimate of a growing influx until 2020-2021 of 36%. The 13 universities plus the Open University fear the coming decade regarding financial support: government investment and annual subsidy per student has

decreased from € 18.000 (until 2001) to 14.000 (2009). Academic staff has not increased at the same pace, save for PhD staff, and support staff has decreased. (total staff: 40.000 FTE).

University libraries are financed in different ways. The traditional situation was a strong link between faculty and its faculty library, where the faculty library resorted under faculty management, received its budget from the faculty and reported to faculty management. Central libraries had the difficult task to maintain some overall library service, receiving a budget from university management or "collecting" its budget from faculties in annual negotiations. In recent years in many universities libraries have been centralised, and budgets are a mix: for facilities provided centrally and for resources (acquisition, databases, media) provided by faculties.

Innovation started early by the proposal of the Tilburg University to build a new high tech library, acknowledging the importance of electronic information and IT technology for delivering services, opening its doors in 1992. The cooperation of library and computer centre resulted in Utrecht in the Electronic Library project (1995-1998). Increasing cooperation among the university libraries in this field led to a steering group for innovative projects, later integrated in SURF¹².

The 13 university libraries are working together in an informal consortium, called UKB, which also includes the *Koninklijke Bibliotheek* (KB, the National Library) in The Hague. Associate members are Open University Netherlands, The Royal Netherlands Academy of Arts and Sciences (KNAW) and the Working Group of Special Academic Libraries (WSWB). 19 research institutes are connected to the KNAW. The UKB consortium is not formalised. Although negotiations take place on behalf of the consortium, no one can commit or sign on behalf of the entire consortium. Apart from more traditional subjects (cataloguing, interlibrary loan, collection management, other items such as licensing policies, e-learning and innovation in the Digital Library are discussed in committees and working groups.

The UKB launched its first ambitious plan 2007-2010 with a strong emphasis on improvement of quality regarding delivery of scientific information, a national information infrastructure and innovation of services. Professionalisation of work process (e.g. shared acquisition) and cooperation are the focus, especially in the field of delivery of content, a publication infrastructure for the scientific output of the institutions, and the promotion of open access. It also worked for an efficient, integrated infrastructure for discovery and access, including both open and restricted information resources, and it wishes to develop facilities for e-learning and for e-science. The major results from this period include steps towards implementation of open access.

In general the Dutch library and research community has been active in promoting open access and the Dutch libraries have been on the forefront of signing the Berlin Declaration. Specific open access initiatives include the Dutch higher education sector declaring 2009 to be "Open Access Year", with the aim of boosting the concept; the development of repositories, such as the scientific portal, NARCIS¹³ and a knowledge bank with teachers' contributions for higher education. A recent debate with the minister resulted in agreement on the principle that all research financed with public funds, should be publicly available for everyone. But a national policy on open access has not yet been materialised.

¹² <<http://www.surf.nl>> consulted on 15 May 2012.

¹³ <<http://www.narcis.info/?wicket:interface=1>>.

UKB supports open access and managed to make agreements with major publishers: a new licensing model with Springer, and a pilot with Elsevier for delayed open access: medical articles of Dutch researchers will be in open access after 12 months. Nevertheless, UKB has also shown its teeth in 2011 when Elsevier changed its policy. It is the view of UKB that an author should in principle have the right to deposit his own article, preferably in the version produced by the publisher but in any case in the final author's version, a right which should not become dependent on (subsequent) agreements with publishers. UKB is particularly concerned about the fact that publishers may overrule agreements made between authors and funding bodies by means of this policy.

In 2009 all universities have agreed on open access *via* the green road: including an obligation to archiving, deposit for every institute; enlarging the content accessibility of repositories, storage and accessibility of research data, cooperation in publishing of open access journals.

The UKB actively seeks co-operation with other stakeholders in the scientific information network (such as the SURF foundation, university staff and students, other organizations in higher education, other library sectors), as well as information and systems vendors. UKB also actively takes part in international developments. One of them is the cooperation in Knowledge Exchange, which resulted in testing a national license for three databases, making them available to all library users.

Another successful project was UKB's High Potentials programme which aimed at extending management- and leadership skills of library workers. Sixteen young potentials participated, and its positive evaluation has resulted in executing the programme in international setting through LIBER¹⁴, the cooperation of European academic libraries.

Important steps have to be taken towards an outstanding national infrastructure for Dutch education and research. At the same time the university libraries are confronted with the demands of the parent organisation, their university and faculties. In its newest policy plan 2011-2015 UKB emphasises that cooperation remains crucial for libraries, but also that globalisation of availability of services requires up scaling and participation of libraries in an international information infrastructure. Modernisation and connection to the international context are important.

Being aware of budget cuts, UKB focuses on efficiency and reduction of costs. It works for a national repository infrastructure for electronic materials in open access; digitisation of collections is coordinated, and will possibly take place with some commercial partners. Together with the national library, UKB works for acquisition of e-content, including e-books, and use of the e-deposit in the National Library.

The embedded library is an answer to the need for data-intensive research and education, and for optimal integration of the workflow for researchers, teachers and students. Students can benefit from the UKB "central service for academic skills courses", and an academic writing workshop. UKB works with SURF on projects regarding library services for virtual work- and learning environments.

UKB has developed a practical benchmark, as a strong management tool. It will be extended with a model for impact measurement.

A main focus of the university libraries is working for Digital Collections, which requires negotiations with publishers on the prices for the increasingly E-format of sources. The battle

¹⁴ <<http://www.libereurope.eu/>>.

on prices for scholarly journals resulted in Big Deals, but means at the same time that a large part of library budgets is “fixed” in these deals, leaving less to negotiate and innovate. In 2008 UKB took the initiative to form a consortium with larger public libraries with some academic collections and function, under the name of Common Information Infrastructure (GII). It aims at cooperation of the back office of the digital library: loans, interlibrary loan, document delivery, acquisition and licensing, including identity management. The ideal situation for the near future is joint use of systems in the cloud, based on international standards and with international connections. An integrated, shared infrastructure will allow libraries to deliver tailor-made local information services which enable the universities to excel in research and education. Vast resources are needed to implement such an infrastructure and these can no longer be generated by individual institutions.

Cooperation in the field of digitisation of collections and sustainable access to collections are other concerns of UKB. *Libratory* is a mega digitisation plan designed by the KB and university libraries, UKB. The vision: to digitise all special collections from Middle Age manuscripts to printed materials until 1840: 44 million pages, 220.000 titles. The total costs are estimated 75 million €. Unfortunately, no government budget has been made available, but public money is needed: what is in the public domain should remain free and accessible, always and everywhere, is the vision. *Early Dutch Books Online* was a preparation to Libratory and still received public funds. Pragmatism may lead to public-private partnerships. Following the report by the European *Comité des Sages* The New Renaissance: by the end of 2016 all European member states must have delivered their major works to Europeana.

Data management of primary research data is another issue for UKB, as a number of libraries is offering services in this field.

UKB's working groups on innovation, Licensing, Collection Management, Open Access and other cooperation around the policy issues mentioned above, are fruitful, but UKB has a number of main concerns. Among them is the continuing pressure from budget cuts in higher education: more students, at least until 2020, but less staff and less budget for innovation facilities, which also affects the libraries. Increasingly, decentred library locations have been closed, whereas “central” university library buildings with modern equipment are becoming ever more popular with students. Good IT facilities, good physical facilities such as chairs and a restaurant are valued highly. Since the introduction of the bachelor/masters system there is also an increasing need for group work places. Recently, new university library buildings have been constructed in Utrecht and Wageningen, and drastic renovations have taken place in Amsterdam, Maastricht and Nijmegen. The average number of opening hours is about 80 a week. There is an increasing call for the extension of facilities as well as opening hours in most places. Customer relations become more important. Increasingly mobile services are developed for information about opening hours, free learning seats, etc.

Regarding reducing costs: licensing requires permanent monitoring of prices, and the UKB has therefore attracted specialists (SURF Office) in working for “big deals”. UKB participates in FOBID Netherlands Library Forum, especially in the Legal Committee, working as a watchdog on copyright and other legal matters, from the point of view of libraries and their users.

UKB reflects on the role of libraries, which wish to profile themselves stronger as supporters of education and research, and therefore contribute in improvement of academic skills and help through library learning centres. Increasingly libraries have to operate in an international world, and have to maintain their networks and knowledge sharing in this field.

Also the implementation of metadata and cataloguing according to new international standards requires frequent updates of knowledge and contacts, for which UKB groups provide the framework.

2.1 NVB

Another player in the library field is the Dutch Association for Library, Information and Knowledge Professionals (NVB) founded in 1912, which has approximately 1.600 members.

It is an association for all types of librarians, although mainly from academic and special libraries.

The association aims at: promotion of the profession in its many forms. It stimulates activities aimed at improvement of the level of professionalism of its members, and works for the promotion and lobby for the interests of the profession.

Among its activities promoting contacts and networking among its members is the primary means of achieving these aims. To this end workshops, seminars, conferences, training courses and study trips are organised. In cooperation with Essentials Media the NVB organizes an annual Conference and is one of the four members of the Federation for the Organisations in the field of Library, Information and Documentation Services (FOBID). The NVB promotes professional publications, both within the association itself and commercially and publishes a bi-weekly ezine (electronic newsletter).

Formal contact is maintained with the various organizations and institutions which provide training for information professionals, e.g. library schools, universities and *GO Education*, a private foundation for the training of librarians, documentalists, library managers and archivists.

The NVB is organized in chapters some reflecting the type of library (Research Libraries, College School Libraries, Special Libraries, Public Libraries) others reflecting a particular interest (Juridical Information, Biomedical Information, Serials Group, Users of Online Information, Literature Researchers). Each chapter has its own elected chapter committee and operates with a high degree of autonomy. The chapters have their own sections and working committees to meet their own policy requirements and activities.

The chapters are organized in four clusters: Prissma; Research, College and School Libraries; Public Libraries; Biomedical Information, Juridical Information and Special Libraries.

The Association has an Executive Committee, which is elected by the Association's members at the General Assembly. This Executive Committee meets 4 times a year. Each of the four clusters is represented through a chair in the Executive Committee.

At the semi-annual General Assembly meetings the Association's members decide on policy and budget issues. The General Assembly is the highest organ of the Association. The Association has a small secretariat (three part time employees) and a managing director. All other functions are executed on a voluntary basis. The Association is self-supporting and receives no governmental subsidy.

The Executive Committee is supported by permanent or ad hoc advisory committees, for example the Advisory Committee for Education and Training. External advocacy and lobbying on behalf of the profession is largely channelled through umbrella organizations such as FOBID, EBLIDA and IFLA.

The NVB has individual and institutional members as well as donators. The membership fee depends on the service package. The basic fee for individuals is € 55.00, membership per

chapter is € 7.50. For institutions the fee depends on the number of staff employed in its information section.

NVB members are entitled to a 50% reduction on subscriptions to the professional journal *InformatieProfessional*. Members receive an ezine (electronic newsletter), decide on policy at the General Assembly meetings and take part in workshops, seminars, conferences and training courses organised by the association and its chapters free of charge or at much reduced rates. NVB is known for its major annual congress, which attracts at least 1.000 participants and a large number of library and information suppliers.

3. National Library: *Koninklijke Bibliotheek*

The *Koninklijke Bibliotheek*, the National Library of the Netherlands was founded in 1798, starting with the decision of the parliament of the Republic (!) to turn the collection of the fled Governor, Stadhouders Willem V, into a national library. As the wall decoration in the KB tells: the royal sound has its roots in the revolution! The first 5.500 items have grown into 6 million, the title National Library was only legally attributed in 1982, when the KB moved to its current premises next to the Central Railway Station in The Hague. Another peculiar feature is that the systematic acquisition of all Dutch titles only started in 1974, and that there is no legal deposit, but publishers deliver these items on a voluntary basis. Now, the KB is the central deposit library for all printed and some electronic publications produced in the Netherlands, and also serves as an international deposit library for e-journals. It is responsible for the coordination of national preservation and digitisation projects. The library furthermore compiles a variety of national special collections in the humanities and social sciences, and research collections related to culture, history and language.

Ever since the introduction of new information technology, the KB has been at the forefront of acquiring funds to innovate services and to implement it for safeguarding cultural heritage through digitisation. In a good Dutch tradition, international cooperation was sought and initiated through various EU-funded projects, in which the KB often took leadership. Various earlier versions of what now is known as the European Library and Europeana were discussed and elaborated with the KB as a solid partner. The results of all these efforts can still be seen today, as many of these international collaboration initiatives and networks have their offices in the KB.

The necessary makeover for the KB has become visible with the Strategic Plan 2010-2013. It offers a clear and bold vision: "KB will work intensively in the coming four years to realise a digital library that will offer everyone access to all digital and printed publications that appear in the Netherlands. As the national library of the Netherlands the KB its task is also to foster the establishment of a new (digital) information infrastructure. Close cooperation between the KB, scientific and public libraries is essential to grant everyone in the Netherlands access to scientific information. In order to realise this vision the KB has established five strategic priorities. They are formulated in the modern management style of concrete promises to stakeholders:

- 1- We offer everyone access to everything published in and about the Netherlands;
- 2- We improve the national information infrastructure;
- 3- We guarantee long-term storage of digital information;
- 4- We maintain, present and strengthen our collections;

5- We develop the KB into a challenging organisation and an attractive employer.

So, the main activities of the KB are:

- Preservation, management, documentation and accessibility of the national cultural heritage;
- Deposit library for Dutch printed and electronic publications and the national bibliography;
- Research library for the history, language and culture of the Netherlands;
- Stimulating and coordinating a common information infrastructure for Dutch libraries;
- Centre of expertise for digitisation, preservation and digital preservation.

The implementation of the plan meant a restructuring of the organisation and services with a strong focus on the Digital Library. It also means, as in many other library organisations, to build capacity and train staff for new tasks requiring new skills, for the KB: with a redeployment of staff of 20%.

The new plan acknowledges the enormous growth of digital information, stronger than that of print, the expectations and the impatience of the newer generation of students and researchers; in short: the user-focus leads to priorities and investments in the digital library. Around 90 projects are coordinated in the programme Digital Library. A strategic collection development plan 2010-2013 includes the acquisitions policies for both print and digital collections, in three shifts: firstly, from print to digital: the digital collection will grow fast in relation to the printed collection. Secondly, the focus on Dutch history, culture and society, already initiated in the previous strategic period, will be strengthened. Thirdly, the inclusion of doubles in the collection will be reduced as much as possible: of each publication only one copy will be acquired, preferably the digital version.

The Annual Report 2010 gives the first results of this strategy. The KB continues to acquire precious prints and rare books, and special collections such as the *Biblioteca Philosophica Hermetica*. KB has also started to make use of "flipbooks", opening up important and valuable books to the general public. Since 1995 born digital publications (publications which are only published in digital form, such as websites, digital periodicals, e-books, etc.) have been collected. The e-deposit has reached the limit of 15 million articles. It has also started web archiving: a selection of Dutch websites will be important for future research and will therefore be harvested for sustainable preservation.

Major digitisation projects have been prepared by signing contracts with Google and Proquest, and started in 2011: Google scans out of copy right books in the KB collection dating 1700-1870, around 160.000 books; Proquest digitises old prints until 1.700, around 28.500 books.

Newspapers is another focus of the KB's ambitions. Agreements with the collective societies Lira (authors) and Pictoright (illustrators) have been signed to assure permission to give internet access to text and images of freelancers in newspapers. The agreements save time which otherwise had to be spent to ask each individual creator for permission. Access to historical newspapers is safeguarded through a new website which includes 1 million pages of 400 years of newspapers.¹⁵ The addition of a large collection of newspapers published

¹⁵ Historical Newspapers website: <<http://kranten.kb.nl/>> consulted on 15 May 2012.

during the war 1940-1945, including “wrong” newspapers to the website roused a debate around the freedom of access to information, which the KB defended with success. Together with the National Archive of Suriname digitisation of 34 Suriname newspapers will be undertaken to add to the website.

An important 6 years digitisation project has been finalised: 200 years of Parliamentary Records were collected at the site: Staten-Generaal Digitaal: all documents 1814-1995 are easily searchable for further research and interest, thanks to cooperation with the Dutch Second Chamber.

Metamorfoze is the title of a larger programme, which has now started with a selection of periodicals to be digitised. Early Dutch Book Online is a collaboration project with the university libraries of Leiden and Amsterdam on books from the end of the eighteen century.

Apart from digitisation of printed material, the KB undertakes collaboration for multiple collections with other institutes: the Rijksmuseum and the *Nederlands Instituut voor Beeld en Geluid* (Netherlands Institute for Image and Sound) to explore how the Netherlands paper and digital library, museum and audiovisual collections can be made searchable in their entirety. A comparable initiative is being undertaken between the National Archives (NA) and the KB, adding to the existing cooperation in the programme, website and exhibition: the Memory of the Netherlands.

The view of the KB on its encouraging role in achieving collaboration finds its background in its conviction that the national information infrastructure needs improvement and that fragmentation of services is no longer explainable to modern users and stakeholders. Therefore the KB ‘wishes to foster the establishment of a joint, national information infrastructure that exploits the possibilities of the digital world optimally.

A true customer focus requires quick and relevant service, difficult to achieve within a fragmented system. In order to realize a one stop service which makes the collections of all libraries in the Netherlands searchable and available to all customers a close cooperation between the KB, SURF (the service organisation for science and higher education), university libraries and public libraries is necessary.

In 2010 the formation of the Shared Information Infrastructure Consortium (*Consortium Gemeenschappelijke Informatie- Infrastructuur* (GII) consisting of the university libraries, the regional libraries (*Plusbiblioteken*), the Netherlands Public Library Association and the KB made a new step towards the establishment of a central point of direction regarding digital information provision. A vision document was drafted and circulated, on which most library organisations responded. The main elements are to make the total of library offers accessible for consultation *via* one desk; and also to provide regulation for service from discovery to delivery: a central regulation to organise access to the materials found, an integrated nationwide request- and transaction system, a single sign-on for library users, based on a.o. a national library card. The idea is to make information available for all and especially target groups in the general audience with services for information which is difficult to get access to. The vision includes also a central database for metadata for digital content. The latter was already proposed and explored in the Panorama-research project focussing on a search system for certified information, in order to provide users with the best, cheapest and easiest way to get full text documents. (Sieverts, 2009)

In the follow-up, it showed that a pragmatic approach prevails. Building an Open Index will make use of existing instruments such as the Joint Cataloguing system (GGC); In order to

meet the requirement of open standards, and avoid vendor lock in, library organisations discuss with OCLC how to arrange for this openness, so other systems can participate.

With a view to the national information infrastructure, the national library plays an important role in bringing various types of libraries and networks together. Especially the relationship to the public library network has been strengthened through various pilot projects. The aim is to make all acquired and licensed sources accessible and available to a larger audience. One of the pilots gives public library user of the regional library of Middelburg access to KB-licensed material. Possibly this will soon be followed for other provinces. Another pilot makes quality articles from Elsevier publisher available at low costs for non-scholars, via their own local library. The largest project includes the cooperation of the KB in building a national catalogue with holdings of all libraries (NBC) which will overcome the partly separate systems of the public libraries on one hand, and the research and university libraries and KB on the other hand. Another catalogue will be built for all digital information. KB and Foundation *Bibliotheek.nl* (developer of public library digital services and infrastructure) have granted a research assignment about the way in which to deal with metadata. In this way, the national information infrastructure will help to realize the aim: all users should have optimal access to all publications in all Dutch libraries.

Early 2012 the KB launched its platform: e-book Library, making 180.000 international scientific e-books accessible to its members. The e-books are downloadable for 1 week, have full text search and can be annotated and added to a personal reading list. The 10.000 electronic journals are presented *via* the Digital Library and EBSCO service (full text search).

The electronic sources are an increasing part of the KB collections (6 million items). With a budget of around 50 million €, a staff of 270 FTE, and 50 opening hours a week, apart from online services, the KB is well into the requirements of the 21st century.

3.1 Intensify existing cooperation: merging with National Archive

The latest news is that *Koninklijke Bibliotheek* (KB) and the National Archives (NA) will be integrated into one organization by July 2013, on the decision of the government, in line with its policy on compactness. Both institutions already work together in a number of areas, such as preservation and conservation. For instance, the joint program "Metamorfoze" for the preservation of paper heritage, and conservation research. In 2005 the Prince of Orange opened the exhibition space "The Legacy of the Netherlands" where both institutions share exhibitions of their treasures. The merging of the neighbouring institutions in interconnected buildings in The Hague offers advantages of scale and help to face the challenges of the digital era.

The administrative integration into one autonomous administrative authority (ZBO) will mean a change for the NA, as this is a ministerial agency. An amendment of the Public Records Act and the Higher Education and Research Act (WHW) are required.

3.2 FOBID

FOBID Netherlands Library Forum is the national umbrella organisation and forum for co-operation between the national library organisations:

- the Netherlands Public Library Association;

- NVB, the Netherlands Association for Library, Information and Knowledge Professionals;
- KB, National Library of the Netherlands;
- UKB, the co-operative Association of University Libraries.

The National Archive may become the next partner.

FOBID was founded in 1974 to improve the co-operation between different types of libraries and institutions for information and documentation services. Its aim is to defend the interests of these institutions and their professionals, and to enhance professional development. The main activities focus on advocacy of the library field's legal interests, on international networking and co-operation of professionals in the field of bibliographic matters and education. A division of tasks has been made into four Committees: legal aspects, bibliographic issues, professional education and international/IFLA-issues.

FOBID activities are mainly developed through its Committees. These Committees of Experts consist of 8-12 persons from a variety of library types. They include even experts from the archive sector, and experts from Flanders. The members are co-opted in the Committee, but appointed by the FOBID Board, which also approves the bi-annual plans.

FOBID monitors developments in society, legislation, the professional and international field, and acts where joint action of the library sector (and increasingly the archive or cultural heritage sector) is requested. FOBID maintains a collection of Relation Network Data to perform its work. The main activities focus necessarily on advocacy of the library field's legal interests. These are found in balancing copyright, being a watchdog for users' privacy, comments on proposed legislation etc. It is concerned and works for the visibility of libraries in society, and convinced that libraries must prove more than ever that they offer value for more, that they are added value to societal aims, be it first among the ranks of knowledge societies, entrepreneurial power or otherwise. The necessary change of the library services must be reflected in library professional education and includes reflection on a new type of profession. FOBID encourages cross-overs between different types of libraries, knowledge sharing and learning from each other; and increasingly, learning in the international context. This is necessary because libraries have always worked with international standards, and they need to be revised, for example in the field of metadata. Furthermore, librarianship is truly an international profession, and cooperation in projects is necessary in a globalising world.

3.3 Legal aspects

This FOBID Committee focuses mainly on the implementation of the EU Directives on Copyright and related matters and Dutch law. The Legal Committee co-operates with a specialist, hired by FOBID, to achieve the best possible results for libraries. It maintains close contacts with the Ministries of Culture, Education and Science, and of Justice. The Committee also advises on licenses, database protection legislation, Orphan Works, digitisation, home-consultation, e-Books, digital rights management and other related issues. It identifies the legal obstacles for libraries and looks for solutions, either through adapted legislation or negotiation for a good practice by suggesting pilots and monitoring practices. In this way a joint committee with Collective Societies was established to mediate for new practices in digitisation: Committee Digi@e. FOBID's Committee also plays a role in making Guidelines for Dutch professionals regarding negotiations for digitisation of library collections.

The Committee co-operates with EBLIDA (Expert Group on Information Law) and IFLA (Copyright and other Legal Matters, CLM) Committee in the field of intellectual property (WIPO) and World Trade (WTO) issues, for example on library exceptions, exceptions for the visually impaired etc. In this way local and international policies and advocacy are closely connected.

3.4 Bibliographic issues

Setting up and maintaining standards is one of the main aspects of this Committee. Different forms of bibliographic control are analysed with a view to international co-operation. The Committee advises on necessary decision-making in the bibliographic field, also related to international standards. It has worked on an advice regarding introduction of RDA, and consulted the library community on its implications and implementation.

3.5 Professional education

The Committee presented a report on future development in an aging professional field. Measures to be taken include pilot projects to recruit new staff and to improve professional curricula.

Participation in European educational projects is another means to attract a flexible professional workforce and increase knowledgeable staff. The Committee works on mapping the competences, function descriptions, and develops a vision for professional education. Those are the themes of recent expert meetings.

3.6 International issues

FOBID supports an international approach to library issues and advocacy. It serves as an International Office, for professional requests from abroad for visits and questions and vice versa. It also promotes Dutch library concepts, solutions, architecture in the international field.

The International Office also performs the coordination of Dutch representation in EBLIDA and IFLA.

FOBID works for increasing professional involvement on a European and international scale. About 25 colleagues from the Netherlands are involved in IFLA Standing Committees and bodies; a number of 40-50 persons visits the IFLA conferences annually. The Dutch library community supports both EBLIDA and IFLA in various ways; they are located in the National Library in The Hague, as does the FOBID office.

FOBID has a website¹⁶ with some information in English¹⁷ and encourages Dutch library institutions to be accessible in other languages.

3.7 The Hague – World Library Capital

¹⁶ <<http://www.fobid.nl>> consulted on 15 May 2012.

¹⁷ <<http://www.libraries.nl>> consulted on 15 May 2012.

FOBID has initiated The Hague – World Library Capital, a network of organizations in the international field of libraries, information and knowledge, based in The Hague. It is a mutual strengthening of the Dutch Municipality: The Hague, which presents itself as International City of Peace and Law and the Library Community working for freedom of expression, the right to information, equal access to a diversity of sources and aim at contributing to knowledge, culture and development in a democratic and well-informed society. The Hague is the only true World library capital. Nowhere else is there a city home to so many major international library organisations, such as IFLA, EBLIDA, LIBER, Europeana and supportive libraries such as the National Library, the libraries of the Peace Palace, of the Parliament, of the International Institute for Social Studies, and the Hague Public Library etc. The municipality and the library community may benefit both through local and international visibility. Members of the network, meeting a few times a year, sharing experiences, find an easy way of knowledge exchange and can plan events in a joint calendar. They can support each other in future projects and with strategic themes.

The City of The Hague has provided start funding, and an international Committee of Recommendation has been established.

The formal launch of The Hague – World Library Capital was on 14 April 2011, on the occasion of the IFLA Presidential Conference 'Libraries Driving Access to Knowledge', at the reception hosted by the municipality in the Public Library.

3.8 International involvement

Although FOBID has a small office, it has succeeded in international involvement through cooperation with others. For a number of years the Netherlands Public Library Association and FOBID organised international events in The Netherlands, under the title of Madurodam Conferences. Madurodam, being a museum of The Netherlands in miniature, was the best location to discuss: big issues in a small world. Those were themes like marketing, international standards, libraries on the agenda, green libraries, etc.

FOBID has been supportive in establishing the Dutch Caribbean Library Association 2012, and now serves as the main link and mediator for exchange and connections. It also supports IFLA's Programme on Building Strong Library Associations.

4. Concluding note on advocating for libraries/IST, an international approach

Through personal experience and professional work, it has become clear for me that much more is needed for library/IST advocacy, not only in the Netherlands, but also in France, in Europe, worldwide. As EU Directives affect libraries and IST increasingly, Advocacy is what we need in Europe, building a stronger network of associations as partners. This requires developing leadership skills, partnerships, team work and better communication.

The library/IST field has many forms of cooperation. They include:

- 1- associations of librarians/ ISTprofessionals;

- 2- associations of libraries, IST institutes;
- 3- network of directors;
- 4- agreements, consortia;
- 5- network of specialists.

All these forms have differences, strengths and weaknesses. In Europe the main organisations for advocacy are: EBLIDA and IFLA; for professional interest, one turns to LIBER. All these organisations are working with rather small offices. There is a general trend of developing professional interests for better advocacy. But this is still far from what is necessary: real focus on the main issues at stake and 1 voice for advocacy. Who is monitoring the European agenda from a common library/IST point of view? How is envisaging the future needs of libraries/IST? How do we identify the main theme relevant to the EU and to us? And how do we find the right persons in the European Commission and in the European Parliament, who can favour our issues? It can only be done when libraries/IST cooperate within a country and across borders.

5. What improvement in the library/IST situation is most urgent?

This final question of my presentation addresses you: What is most urgently needed in your IST situation? What is needed? How do you know? What strategy will you follow? Are there organisations who can help?

Will knowledge/reference to libraries/IST in similar situations in other countries help your case? If you think no longer only of your library island and show IST/Library spirit and energy, you can do the first steps, by making use of existing contacts and enlarge by those interested. It will work! It will work even better if you keep in mind and take to heart some advice, already realised by others:

- Speak another language beyond your own language;
- Encourage students and young professionals for exploration, reports, early research;
- Use European mobility programmes.

It will bring you beyond your limits: personally, professionally and for your institution. So take that challenge! We hope to see you in the Netherlands. FOBID will welcome you!

References

- [1] Z. Bauman, *Culture in a Liquid Modern World*. Polity. Cambridge (UK)/Malden (MA), 2011.
- [2] Z. Bauman, *Liquid Times: Living in an Age of Uncertainty*. Polity Press. Cambridge. 2006.

CBS: Statistics

- [3] <<http://statline.cbs.nl/StatWeb/publication/?DM=SLNL&PA=70763ned&D1=a&D2=a&HDR=T&STB=G1&VW=T>> consulted on 15 May 2012.
- [4] J. Kasperkowitz, M.H.L. van Tits and S. Fuhr, *De Leerfunctie van bibliotheken in beeld*. Sector Instituut Openbare Bibliotheken. The Hague. 2010.

- [5] *The Educational Function of Libraries in Focus*, Report 2009. Carried out on behalf of the Netherlands Public Library Association by: Kasperkovitz beleidsonderzoek en advies and IVA beleidsonderzoek en advies. Published by the Netherlands Institute for Public Libraries, **Sector Instituut Openbare Bibliotheken, The Hague**. 2010.
- [6] M. Koren, *New Library Buildings in the Netherlands, nr. 5, Vereniging van Openbare Bibliotheken/Bibliion, Leidschendam*. 2010.
- [7] M. Koren, "Public Libraries", in I. Abdullahi (ed.), *Global Library and Information Science: A Handbook for Students and Educators*. **Saur de Gruyter, München**. Series "IFLA Publications 136-137", p. 311-328. 2009.
- [8] Raad voor Cultuur, *Advies over de bestuurlijke organisatie van het openbare bibliotheekwerk. Raad voor Cultuur. The Hague*. 1998.
- [9] Research voor Beleid, *Aanpak van laaggeletterdheid door bibliotheken; kansen en mogelijkheden voor de toekomst*. **Sector Instituut Openbare Bibliotheken, The Hague**. 2011.
- [10] Research voor Beleid, *Approach policy to low-literacy by libraries: Chances and possibilities for the future*, Published by the Netherlands Institute for Public Libraries. **Sector Instituut Openbare Bibliotheken**. 2011.
- [11] B. Savenije, "Libraries in the information society: cooperation and identity", in by D. Baker and W. Evans (eds), *Libraries and Society: Role, Responsibility and Future in an Age of Change*. **Chandos Publishing, Oxford**. Series "Chandos information professional". p. 203-218, 2011.
- [12] B. Savenije, "Digital Library Economics: The Dutch Perspective", in D. Baker and W. Evans and C. Stevens (eds.), *Digital Library Economics: An Academic Perspective*. **Chandos Publishing, Oxford**. Series "Chandos information professional". p. 145-159, 2009.
- [13] E. Sieverts, *et al.*, *Panorama, FOBID Netherlands Library Forum, The Hague*. 2009.
- [14] E. Simons, "The New Learning Environment: Impact on Staff and Students". *Liber Quarterly* **20(2)**, p. 258-269, October 2010.
<http://www.shb-online.nl/dmdocuments/xplora_ellen_simons_liber_quarterly.pdf> consulted on 15 May 2012.
- [15] SHB, "Libraries to Universities of Applied Sciences"¹⁸, Samenwerkingsverband Hogeschoolbibliotheken:
<http://www.shb-online.nl/dmdocuments/xplora_ellen_simons_liber_quarterly.pdf> consulted on 15 May 2012.
- [16] N. Vallet, "Openbare bibliotheken: strategische partners in stadsontwikkeling", *Vlaams Tijdschrift voor Overheidsmanagement*, **16(2)**, p. 4-13, 2011.

Websites¹⁹

Reading Foundation: <<http://www.lezen.nl>>.

NAPLE: <<http://www.naple.info>>.

CDR music library services: <<http://www.muzeekweb.nl>>.

¹⁸ <<http://www.shb-online.nl/>> consulted on 15 May 2012.

¹⁹ Consulted on 15 May 2012.

United Kingdom Research Reserve (UKRR)¹

Today's solution to yesterday's problems

Deborah Shorley

Head, UKRR, Central Library, Imperial College London,
South Kensington, London, SW7 2AZ
d.shorley@imperial.ac.uk

Abstract:

By the end of the twentieth century UK university libraries were bursting at the seams. Indeed, paradoxically, many universities continued to regard the size of their library collections as a potent status symbols. But by the turn of the century many were beginning to admit quietly that something had to be done. Not only were we adding kilometers of shelving to our libraries every year, but much of the material kept on these shelves was also held by many other institutions across the country. And this, even though ever more legacy journal material was becoming available in durable digital form. We were also struggling to cater for hugely increased numbers of students, who expected bigger, brighter spaces to study in. Something had to give.

UKRR is a child of its time – conceived to protect the UK's research information infrastructure by replacing the just in case model with just in time access to material. But what exactly is it? At its simplest:

The UK Research Reserve (UKRR) is a collaborative, co-ordinated and sustainable approach to securing the long-term retention, storage and access to low-use printed research journals.

It aims to:

- *safeguard the long-term future of printed research journals;*
- *enable quick and easy access to research material;*
- *ensure efficient use of resources.*

UKRR Phase One was a pilot project, funded by the Higher Education Funding Council for England (HEFCE) for 18 months from early 2007, and led by Imperial College London, achieved all it set out to do, and far more. Not only did it effect the disposal of about 18 kilometers of low use research journals in the six partner institutions, but far more importantly it proved that this collaborative model for national collection development could work in the real world.

All in all, UKRR Phase One worked. It showed that the British Library had reassuringly few gaps in its collection, but that duplication across the sector remained very high.

¹ Christopher Brown, UKRR Data Coordinator (christopher.brown@imperial.ac.uk), helped prepare this article and is responsible for Fig 1).

It also showed conclusively that we needed a large scale national scheme to tackle the problem right across the sector. And this would demand serious money.

So UKRR Phase Two was launched in February 2009, funded by the Higher Education Funding Council for England (HEFCE) to the tune of 11.5m €.

Twenty-nine university libraries are participating in UKRR phase 2, which by June 2011 had processed 33 810 m of material. On average around 85% of the material offered to UKRR is cleared for disposal, with 13-14% retained for the research reserve. The remaining 1-2% will be found to be either outside the scope of the collection policy, or withdrawn from the process by the library. Some of the varied uses for space freed up through participation in UKRR include new study spaces, group study areas, learning cafes, and IT training facilities.

This is a groundbreaking initiative so we have of course encountered a few problems on the way. Issues we have faced so far include resistance within some disciplines to reduce their print collections, despite electronic access; and at operational level the variable quality of bibliographic and holdings data can make it difficult to match holdings across institutions and coordinate the disposal process as speedily as we would like.

All in all we are well pleased with the progress UKRR has made so far. Five years ago it simply couldn't have happened – and five years hence, we hope it will have freed up around 100kms of shelving across UKRR member libraries. Thereafter, who knows? Other journals, monographs, international collaborations? It is all to play for.

Meanwhile, one thing is certain: UKRR is already a key plank in the UK's research infrastructure. Our researchers, long recognised by many to be among the most effective in the world, need to find the information they need, when and where they need it. UKRR is a twentyfirst century way of making sure this will continue to happen.

Résumé :

United Kingdom Research Reserve (UKRR) Une solution d'aujourd'hui pour les problèmes d'hier

À la fin du xx^e siècle, les bibliothèques universitaires du Royaume-Uni menaçaient d'éclater. En effet, de manière paradoxale, beaucoup d'universités continuaient de considérer l'importance de leurs collections comme le symbole fort de leur prestige. Mais au tournant du siècle, beaucoup commençaient à admettre tout doucement qu'il fallait faire quelque chose. Non seulement nous ajoutions tous les ans des kilomètres de rayonnages à nos bibliothèques mais beaucoup des documents conservés en rayons se étaient aussi détenus par beaucoup d'autres institutions partout dans le pays. Et cela, malgré l'apparition croissante de revues traditionnelles disponibles durablement sous forme numérisée. Nous nous efforcions également d'accueillir des effectifs d'étudiants qui croissaient énormément, et qui désiraient étudier dans des espaces plus spacieux et plus lumineux. Il fallait agir.

L'UKRR est un enfant de son temps. Conçu pour protéger l'infrastructure de l'information scientifique du Royaume-Uni en remplaçant le modèle d'accès aux documents « au cas où ? » par le modèle « juste à temps ». Dans sa plus simple expression : UK Research Reserve (Réserve britannique pour la recherche -

UKRR) est une approche collaborative, coordonnée et durable visant à garantir la conservation à long terme, le stockage et l'accès à des revues scientifiques eu utilisées.

Elle a pour objectifs de :

- sauvegarder l'avenir à long terme des revues scientifiques imprimées ;
- permettre un accès libre et rapide aux documents de la recherche ;
- assurer une utilisation efficace des ressources.

La première phase de UKRR, un projet pilote, soutenu par le Conseil de financement de l'enseignement supérieur en Angleterre (HEFCE) pendant 18 mois à partir du début 2007 et conduit par l'Imperial College de Londres, a réalisé tous ses objectifs et au-delà. Non seulement elle a pu rationaliser la gestion de 18 kilomètres de revues scientifiques faiblement utilisées dans les six institutions participantes, mais, bien plus important encore, elle a prouvé que ce modèle collaboratif pour le développement de collections nationales était viable.

En somme, UKRR dans sa phase 1 a bien fonctionné. Chose rassurante, il a montré que le fonds de la British Library présentait peu de lacunes, mais que les doublons dans ce domaine restaient élevés. Il a aussi montré très clairement que nous avons besoin d'un plan national à grande échelle pour régler efficacement le problème, ce qui exigerait un soutien financier conséquent.

C'est ainsi que la phase 2 de UKRR a été lancée en février 2009, avec un financement du Conseil de financement de l'enseignement supérieur d'Angleterre (HEFCE) à hauteur de 11,5 millions d'euros.

Vingt-neuf bibliothèques universitaires participent à cette phase 2 de l'UKRR, qui avait déjà traité 33 810 m de rayonnages, à la date de juin 2011. En moyenne environ 85 % des documents proposés à UKRR ont été réformés et 13-14 % ont été retenus pour y être versés. Les 1-2 % restants ne sont pas dans le périmètre de la politique des collections, ou ont été retirés de la chaîne par la bibliothèque. L'espace libéré a permis de créer de nouveaux espaces de travail, des espaces de travail collectif, des cafés-formation et des espaces de formation.

Cette initiative est révolutionnaire, donc nous avons bien sûr rencontré quelques problèmes en route. Les questions auxquelles nous avons fait face jusqu'ici incluent la résistance dans quelques disciplines pour réduire leurs collections papier, malgré l'accès électronique ; et au niveau opérationnel la qualité variable de données bibliographiques peut rendre difficile la correspondance des données entre institutions et la coordination des procédures de réforme aussi rapidement que nous voudrions.

Globalement, nous sommes très satisfaits de la progression d'UKRR jusqu'à présent. Il y a cinq ans, cela ne se serait tout simplement pas fait - et dans cinq ans, nous espérons avoir libéré environ 100 km de rayonnage dans les bibliothèques membres de l'UKRR. Ensuite, qui sait ? D'autres revues, des monographies, des collaborations internationales ? La voie est libre.

En attendant, une chose est certaine : UKRR est d'ores et déjà un élément-clé dans l'infrastructure de recherche du Royaume-Uni. Nos chercheurs, longtemps reconnus pour être parmi les plus efficaces dans le monde, doivent trouver les informations dont ils ont besoin, quand et où ils en ont besoin. UKRR est un moyen en ce XXI^e siècle que cela continue à se produire.

1. Introduction

I am very pleased to have this opportunity to talk to you at the FRÉDoc conference about the United Kingdom Research Reserve (UKRR); an initiative we have undertaken in the United Kingdom to deal with some of our legacy collections of low use print research journals. We don't, of course, claim that what we are doing is a model for everybody all over the world. But we do believe that it is an entirely pragmatic solution to a problem which we in the UK could no longer afford to ignore. So far at least, UKRR has served us well.

2. UKRR initiative

In this presentation I shall talk about: what we have done, why we have done it, how we have done it, and our achievements so far. I shall also talk about our plans for the future. Before I begin discussing UKRR itself, it is important to remind ourselves of our primary role as librarians. Our primary responsibility is not to put material on shelves, nor to lend people material from our libraries. Rather, it is to make sure that we continue to take full responsibility for the stewardship of the information resources needed by our users. This sounds easier than it is. But it is very important, especially for those of us who work in libraries in large research institutions, to remember that we exist above all, as a service. And as a service we need to ensure that all our researchers have access to the information they need, in the form that they need it, when and where they need it. This is what Ranganathan said all those years ago, and it remains essentially true. Personally I don't believe that libraries are indispensable; but that they are, so far, the best mechanism anyone has found to provide the right information to users in the right way. This said we need to make sure that old practices are replaced by new, more appropriate, technologies, and that old ways of storing & sorting material and information are replaced by more effective methods wherever possible. The digital age has changed so many things, but it is fair to say above all for us as librarians, that it has changed the way we deal with our users' information needs more than any of us could have foreseen.

So now let me explain what UKRR is, what it does, how it does it, and why it is needed.

3. UKRR initiative purpose

UKRR is a partnership between the Higher Education (HE) sector and the British Library (BL) that facilitates, amongst other things, the releasing of space in HE libraries through coordinated deduplication. By the end of the twentieth century UK university libraries were bursting at the seams. Indeed, paradoxically, many universities continued to regard the size of their library collections as a potent status symbols. When all information was recorded in physical format and communications were slow this made good sense. But as we entered the second millennium many were beginning to admit quietly that something had to be done. Not only were we adding kilometres of shelving to our libraries every year, but much of the material kept on these shelves was also held by many other institutions across the country. And this, even though ever more legacy journal material was becoming available in durable digital form. We were also struggling to cater for hugely increased numbers of students, who expected bigger, brighter spaces to study in. Something had to give.

UKRR provides one means of dealing with these changes, and in that respect it is a child of its time – conceived to protect the UK's research information infrastructure by replacing the *just in case* collection model with *just in time* access to content. So what exactly is UKRR? At its simplest:

The UK Research Reserve (UKRR) is a collaborative, co-ordinated and sustainable approach to securing the long-term retention, storage and access to low-use printed research journals.

It aims to:

- safeguard the long-term future of printed low use research journals;
- enable quick and easy access to research material;
- ensure efficient use of resources;
- change the existing culture within libraries which sees collections defined by institutional boundaries.

UKRR's Phase One pilot project, funded by the Higher Education Funding Council for England (HEFCE) for 18 months from early 2007, and led by Imperial College London, achieved all it set out to do. Not only did it effect the coordinated disposal of over 10.5 kilometers of low use research journals in the eight partner institutions, but, far more importantly, it proved that this collaborative model for national collection development could work in the real world. It also gave an indication of some of the challenges that such a collaborative scheme would present.

It showed too that the British Library (BL), whose collections are checked to ensure they hold a copy of every issue of every journal offered to UKRR, had reassuringly few gaps in its collection. This was important as the completed holdings of the BL provide access to the content within UKRR through their existing document supply channels. Duplication of holdings among the collections of HE sector libraries was found to be high however, which demonstrated the need for a large-scale scheme to tackle the problem right.

So UKRR Phase Two was launched in February 2009, funded by the Higher Education Funding Council for England (HEFCE) to the tune of 11.5m €, and again managed at Imperial College.

4. How does UKRR work?

To start the UKRR process participating libraries identify print journals which they would like to remove from their collections; typically those which no longer support the teaching and research of the institution, or for which they have secure electronic access. The details of the journals and holdings are entered into a template Excel spreadsheet, which is uploaded by staff at the participating library into UKRR's IT system LARCH². The data is then passed to the British Library (BL), who shelf-check their loanable holdings to identify any items on offer not in their collections. Simultaneously, automated searches are performed on the SUNCAT³.

² Linked Automated Register of Collaborative Holdings.

³ Serials Union Catalogue for the UK research community <<http://www.sunca.ac.uk>> consulted on 15 June 2012.

union catalogue by staff at the national data centre EDINA⁴, to identify any holdings of the offered titles within other participating UKRR libraries. The results of these searches are then sent to the BL *via* UKRR, where staff manually identify correlations between the issues in the holding being offered, and those in other members' collections. The data resulting from these processes is then uploaded to LARCH, where it is used to identify material which can be disposed of, should be retained for the research reserve, or sent to the BL to complete their holdings. The list is then finalised and the decisions about what to do with the holdings are made available to the offering library as a report, which they can download in a variety of formats. Only material which is found to be held in two other members' collections is cleared for disposal, ensuring that in addition to the access copy held at the BL, two preservation copies remain within the HE sector (see *Fig 1*).

The role of LARCH is crucial as it creates a central record of the dispose/retain decisions made against all holdings offered to UKRR, so that they can be taken into account when making decisions against holdings offered in future. As removing holdings records from their catalogues and updating SUNCAT is typically one of the last steps libraries perform when disposing of material, LARCH ensures that the data relating to holdings elsewhere in the UKRR membership is adjusted to take into account previous disposals.

Twenty-nine university libraries participate in UKRR phase 2, which by November 2011 had processed 44 Km of material. On average around 85% of the material offered to UKRR is cleared for disposal, with 13-14% retained for the research reserve. The remaining 1-2% will be found to be either outside the scope of the collection policy, or withdrawn from the process by the library. Some of the varied uses for space freed up through participation in UKRR include new study spaces, group study areas, learning cafes, and IT training facilities. Most importantly we expect our members to dispose of duplicate material systematically and responsibly. To encourage them to do this we pay them approx €30 for every linear meter of low use research journal holdings they offer.

UKRR is a groundbreaking initiative, so we have of course encountered a few problems on our way. Researchers in some disciplines are reluctant to see their own library's print holdings reduced even when the electronic versions are instantly and ubiquitously available to them. We have undertaken a wide-ranging and sustained advocacy campaign to address these concerns, and have been mostly, if not entirely, successful in reassuring such researchers that their research will not be compromised by UKRR.

UKRR's progress has also been hampered by the variable, and at times unreliable, metadata which has made it difficult to compare holdings of different institutions without a good deal of manual checking.

⁴ <<http://edina.ac.uk/>> consulted on 15 June 2012.

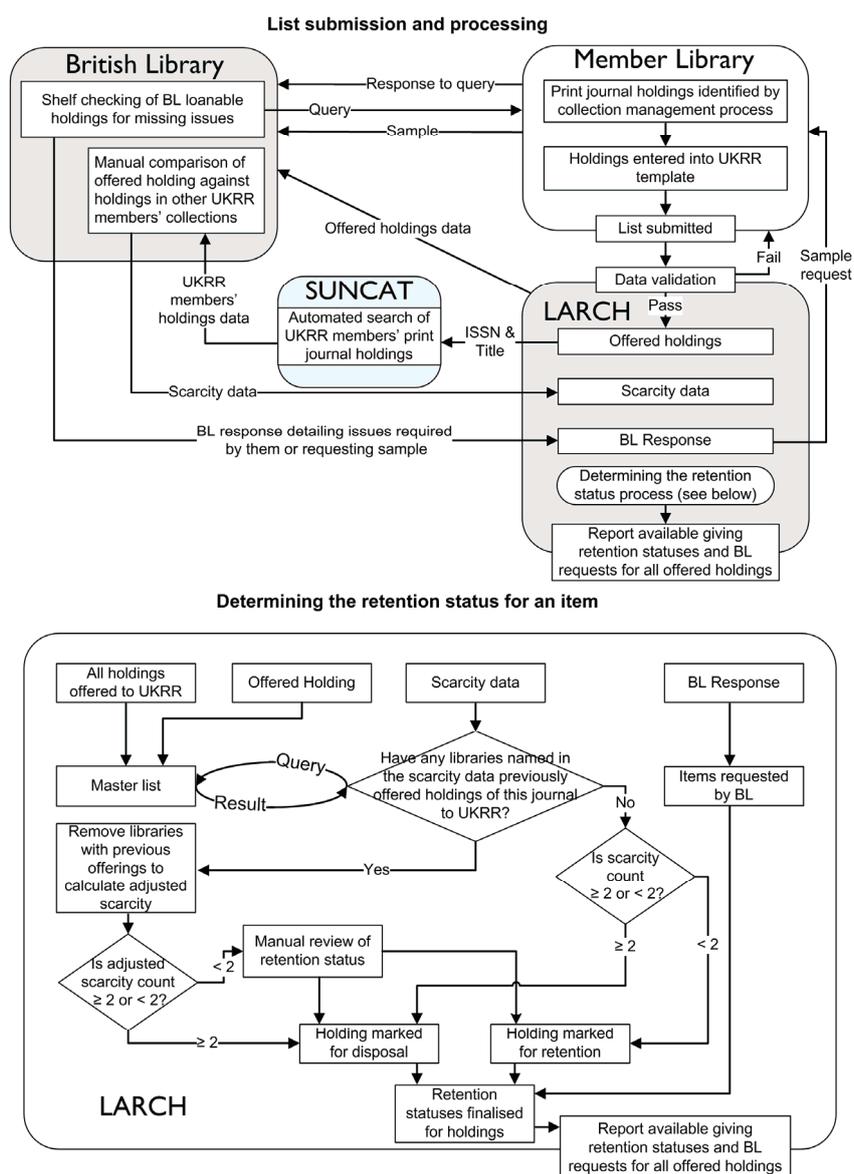


Figure 1

All in all we are well pleased with the progress UKRR has made so far. Five years ago it simply couldn't have happened – and five years hence, we hope it will have freed up around 100kms of shelving across UKRR member libraries. Thereafter, who knows? Other journals, monographs, international collaborations? It is all to play for.

This programme is not cheap, and its primary purpose is not to effect the disposal of surplus holdings of low use research journals, although this in itself is both useful and important. Rather UKRR's primary purpose is to encourage culture change among librarians and researchers. More specifically it aims to demonstrate that in our digital age it is no longer necessary for all universities to hold their own collections of low use print journals if they are safely preserved elsewhere.

Once our funding runs out in 2014 we shall need to make UKRR self financing. We don't yet know how we shall do this, but we are already exploring a number of possibilities.

5. Conclusion

Meanwhile, one thing is certain: UKRR is already a key plank in the UK's research infrastructure. Our researchers, long recognised by many to be among the most effective in the world, need to find the information they need, when and where they need it. UKRR is a twenty-first century way of making sure this will continue to happen.

LIBER and EU projects

Wouter Schallier

*Executive director of LIBER
Brussels, Belgium*

Abstract

LIBER (Ligue des bibliothèques européennes de recherche – Association of European Research Libraries) is the main research libraries network in Europe. It has a membership of 425 research libraries from more than 40 countries. Since its foundation in 1971, French research libraries have always been very well represented in LIBER.

LIBER is currently involved in several European projects. LIBER Executive Director Wouter Schallier will present these initiatives covering areas such as data exchange, digital preservation, digitisation, content aggregation and Open Access. He will explain how these projects contribute to a Digital Europe. Finally, Mr Schallier will highlight opportunities for French research libraries in future EU projects.

Résumé

LIBER et les projets européens

LIBER (Ligue des bibliothèques européennes de recherche – Association of European Research Libraries) est le principal réseau de bibliothèques de recherche en Europe. Il compte 425 bibliothèques représentant plus de 40 pays. Depuis sa fondation en 1971, les bibliothèques de recherche françaises y ont toujours été bien représentées.

LIBER est actuellement engagé dans plusieurs projets européens. Dans la présentation des initiatives LIBER faite par son Directeur exécutif Wouter Schallier, les questions d'échange de données, de conservation numérique, de numérisation, d'agrégation de contenus et du libre accès seront abordées. Il montrera comment ces projets contribuent à une Europe numérique. Enfin, elle mettra en lumière les possibilités de participation aux futurs projets de l'Union européenne pour les bibliothèques de recherche françaises.

IFLA (*International Federation of Library Associations and Institutions*)

Pascal Sanz

*Bibliothèque nationale de France,
président d'IFLA-France
Paris, France*

Résumé

L'*International Federation of Library Associations and Institutions* (Fédération internationale des associations de bibliothécaires et des bibliothèques), ou IFLA, consacre son action à tous les domaines de la bibliothéconomie, de la documentation et des sciences de l'information. Cette action est orientée et déterminée par :

- le plan stratégique de la Fédération ;
- les thèmes des présidents de l'IFLA ;
- les « activités fondamentales » (*Core Activities*) ;
- les évolutions scientifiques, techniques et professionnelles intervenant dans les domaines couverts par chacune des structures professionnelles de l'IFLA (43 sections, 11 groupes d'intérêt particulier).

Les tendances actuelles des actions de l'IFLA s'organisent, notamment, autour du développement des axes suivants :

- le développement des ressources électroniques (numérisation de collections physiques, acquisition de ressources électroniques natives) ;
- la question de la préservation à long terme des collections physiques et numériques des bibliothèques ;
- les problématiques techniques, juridiques et politiques de l'accès à ces ressources ;
- le rôle des bibliothèques, facteurs d'accès au savoir et de changement dans la société de l'information ;

la coopération entre pays développés, pays émergents et pays en développement.

Abstract

IFLA (International Federation of Library Associations and Institutions)

The International Federation of Library Associations and Institutions, or IFLA, focuses its activity on all fields of library, documentation and information science.

This activity is oriented and determined by:

- The strategic plan of the federation;
- The themes put forward by the presidents of IFLA;
- Core activities.

The scientific, technical and professional developments in the fields covered by each of the professional divisions of IFLA (43 sections, 11 special working groups).

IFLA's current activities concentrated around developments in the following areas:

- The evolution of electronic resources (digitalization of paper collections, purchase of native electronic resources);
- The question of long-term archiving of paper and electronic collections owned by the libraries;
- The technical, legal and political problems linked to access to these resources;
- The role of libraries, key factors of access to knowledge and change in the information society;

The cooperation between developed countries, emerging countries and developing countries.

Partie 4
Organisation professionnelle

Métiers et organisation professionnelle Perspectives européennes

Sylvie Dalbin

Consultante en organisation et ingénierie documentaires
fr.linkedin.com/in/sylviedalbin/

Résumé

La recherche et l'innovation sont fortement intégrées aux actions de l'UE depuis le début de sa création pour atteindre aujourd'hui un Espace européen de la recherche. Plus largement c'est le « triangle de la connaissance formé par l'éducation, la recherche et l'innovation » qui est au cœur de l'action européenne.

Comment les professionnels de l'information-documentation qui agissent traditionnellement soit au niveau local (France) soit au niveau international, se structurent et interviennent pour soutenir les activités des chercheurs au niveau européen ?

Des exemples issus du fonctionnement de l'UE, des programmes de recherche et de normalisation européenne nous permettront de pointer la pertinence de ce niveau d'intervention, et d'identifier les moyens pour les professionnels de l'information-documentation pour y participer activement au bénéfice de la recherche et de l'IST.

Abstract

National/international structuring work and Eurocertification

Research and innovation have been strongly integrated into the European Union's actions since its creation with the aim to have today a European Research Area. More broadly, the knowledge triangle composed of education, research and innovation is considered to be at the very heart of European strategy and actions.

How do information-documentation professionals operating either locally (France) or internationally structure themselves and intervene in order to support the activities of scientists at the European level?

Examples drawn from the way the EU functions, from research programs and European normalization will enable us to appreciate the relevance of this level of intervention, and identify the means required to enable information-documentation professionals to participate actively for the benefit of research and STI (scientific and technical information).

1. Introduction

Lorsque l'on aborde les métiers et les organisations professionnelles sous le prisme de l'Europe, il convient d'adopter une démarche à caractère historique, ceux-ci préexistant à la concrétisation du projet européen. Tout comme la recherche scientifique, la production d'information – qui n'avait pas besoin d'être qualifiée de scientifique – s'organisait alors localement mais surtout à travers des organisations nationales et internationales.

Une courte introduction sur les périmètres possibles de l'IST nous fournira un cadre général afin d'étudier dans une deuxième section comment les différentes politiques européennes – l'éducation et les métiers, la recherche et l'innovation mais aussi la société numérique, construisent une représentation particulière de l'IST au niveau européen.

Les conséquences de ces orientations européennes dans le secteur de l'information-documentation portent aussi bien sur le périmètre des informations prises en charge, les fonctions et les techniques professionnelles que sur l'organisation professionnelle qui pourrait être mise en place au niveau européen afin de couvrir ce large périmètre.

2. Quel périmètre pour l'IST ?

L'IST dans les politiques publiques françaises est fréquemment circonscrite au secteur de la « publication en IST »¹. Les orientations les plus récentes au niveau du CNRS témoignent d'une certaine ouverture [1], mais se projeter sur l'Europe nous oblige à prendre en compte un périmètre beaucoup plus étendu intégrant tous types de données ou informations quels que soient leurs formats et supports, produits ou utilisés dans le cadre d'activités de recherche scientifique et technique.

L'IST s'étend ainsi sur trois plans qui se recoupent :

- la nature de l'information prise en charge en s'ouvrant sur une production de données de nature organisationnelle, technique voire technologique ;
- le type de support en ne restreignant pas aux livres, thèses ou revues, le patrimoine à prendre en charge pour couvrir l'étendue des ressources soutenant la recherche scientifique telle qu'elle se pratique aujourd'hui ;
- enfin des techniques documentaires à développer pour capter, traiter et communiquer cette diversité de ressources.

La cartographie du secteur de l'IST – acteurs, politiques, méthodes, compétences et outils de traitement et d'organisation de l'IST, prend alors un tout autre visage.

En particulier la présence des sociétés savantes ou associations scientifiques et techniques auxquelles s'ajoutent les associations professionnelles ou la société civile dans les domaines comme la médecine, l'agriculture, l'informatique ou la sociologie au sein desquelles l'imbrication société/recherche semble incontournable, nous conduit à un

¹ Voir le rapport de 2008 du Comité IST du ministère de l'Enseignement supérieur et de la Recherche sur les « évolutions du secteur de l'IST » qui porte dans les faits strictement sur « l'édition scientifique (sous sa forme commerciale ou sa forme de libre accès) » (p. 3), excluant même les bases de données du périmètre d'étude. Consulté en septembre 2011, <<http://www.ladocumentationfrancaise.fr/rapports-publics/084000458/index.shtml>>.

recentrage vers les acteurs principaux du système de l'IST : la communauté de chercheurs, ses partenaires et plus largement la société civile au centre des préoccupations européennes.

3. Contexte créé par le développement de l'Union européenne

3.1. Place de la recherche dans les politiques de l'UE

La recherche scientifique – qui se distingue d'un Espace européen de la recherche² [2] – constitue un élément clé de la construction européenne : les traités CECA³ (1952) et Euratom⁴ (1957) dans les domaines respectifs du charbon/acier et de l'énergie nucléaire comportaient des dispositions en matière de recherche scientifique. Le 3^e traité instituant la Communauté économique européenne (CEE) en 1957 comportait quant à lui un article de portée générale qui permit le lancement dans les années 60 et 70 d'un certain nombre de programmes de recherche dans des domaines jugés prioritaires : l'énergie, l'environnement, puis la biotechnologie. La notion de programme-cadre structurant le développement de la recherche européenne apparaît alors au début des années 1980. Les orientations actuelles consolident la stratégie des programmes-cadre et la prolongent en proposant la construction d'un « espace européen de la recherche ». Cet espace mis en œuvre sous le pilotage de la Commission mais pas directement par elle, ne conduit pas nécessairement vers une « politique commune » [2] laissant les communautés dans leur domaine ou secteur, construire cet espace commun.

La recherche scientifique était donc initialement intégrée à des projets technico-industriels (CECA, Euratom). Mais si cette articulation entre recherche et usage des résultats de la recherche a toujours été une priorité européenne, aujourd'hui, elle intègre ou s'articule à d'autres politiques cherchant à rééquilibrer ce risque (fort) de dépendance à l'industrie ou aux entreprises. En premier lieu, une politique « **d'Éducation et formation tout au long de la vie** »⁵ [3-4] est associée à la recherche et innovation, formant le « triangle de la connaissance », cadre général des politiques de la recherche scientifique et technique de cette décennie. Ensuite les politiques développées autour de ce triangle de la connaissance (recherche-éducation-innovation) sont nécessairement articulées aux volets transversaux des politiques européennes depuis 20 ans, en particulier la « **société (sciences et société)** et le **numérique (technoscience et TIC)** ».

Essayons de pointer quelques caractéristiques de ces différentes politiques vis-à-vis de la politique de la recherche et de l'IST.

² Espace européen de la recherche, <http://ec.europa.eu/research/era/index_fr.htm> ; Service français d'accès à l'information sur la recherche en Europe, « L'actualité de la Recherche en Europe », <<http://www.eurosfaire.prd.fr/7pc>>, consultés le 20 juin 2012.

³ Communauté européenne du charbon et de l'acier.

⁴ Communauté européenne de l'énergie atomique.

⁵ Éducation tout au long de la vie, <http://ec.europa.eu/education/lifelong-learning-policy/doc30_fr.htm>, consulté le 20 juin 2012.

3.1.1. Apprentissage tout au long de la vie

L'éducation est une constante dans les actions de l'UE, mais elle a fait de l'apprentissage tout au long de la vie l'un des axes essentiels de sa politique pour la période 2007-2013, axes renforcés jusqu'en 2020 [3]⁶.

Différents instruments ont été élaborés dans le but de déployer cette politique formant un **réel écosystème autour de la formation et des métiers**⁷ [4] :

- le **Cadre européen de certification (CEC)**⁸ officialisé en 2008, constitue un cadre de référence commun pour l'éducation, la formation tout au long de la vie et la mobilité. Le CEC est l'outil utilisé sur l'ensemble du territoire européen pour transposer les différents systèmes et niveaux de certification⁹ tant pour l'enseignement général et supérieur (LMD)¹⁰ que pour l'enseignement et la formation professionnels. Le fonctionnement repose sur un point de coordination national : la Commission nationale de la Certification professionnelle. En 2011 tous les titres et diplômes nationaux ont été alignés à ce cadre de référence ;
- articulés au CEC, le Système de transfert et d'accumulation de crédits (**ECTS**)¹¹ et les Crédits d'apprentissage européens pour la formation et l'enseignement

⁶ La signification de la notion de « formation » est ici très étendue : elle inclut toute forme d'apprentissage formel, informel ou non-formel. Voir définition dans : La qualité dans l'enseignement et la formation – Glossaire, <<http://www.cedefop.europa.eu/EN/publications/17663.aspx>>, ou *Terminology of European education and training policy* (les 100 mots clés) <<http://www.cedefop.europa.eu/EN/publications/13125.aspx>>, consultés le 20 juin 2012.

⁷ Portail européen Formation, <http://ec.europa.eu/education/index_fr.htm> ; Agence « Europe-éducation-formation » en charge des programmes européens sur le territoire français, <<http://www.europe-education-formation.fr/agence.php>> ; Observatoire européen des normes des technologies éducatives, <<http://www.cen-ltso.net/>>, consultés le 20 juin 2012.

⁸ Attention à la perception que l'on peut avoir des 8 grades du CEC – « Les “descripteurs” (terme employé dans le CEC) rendent compte à la fois des spécialisations et des généralisations. Ainsi, atteindre des niveaux supérieurs n'implique pas nécessairement que les compétences et les savoirs soient plus spécialisés, même si c'est le cas dans de nombreux milieux universitaires ou de recherche. Dans certains contextes professionnels ou d'études, passer d'un niveau inférieur à un niveau supérieur peut marquer une généralisation ». Site officiel du CEC, « Le cadre européen des certifications (CEC) », <http://ec.europa.eu/education/lifelong-learning-policy/doc44_fr.htm>. Relié au Cadre des certifications de l'espace européen de l'enseignement supérieur, “Qualifications Frameworks in the EHEA”, <<http://www.ehea.info/article-details.aspx?ArticleId=65>>, consultés le 20 juin 2012.

⁹ Le terme de « certification » regroupe à la fois les « Diplômes » délivrés par l'Éducation nationale, les « Titres professionnels » délivrés via un ministère ainsi que les « certificats de qualification » qui sont des attestations dans une spécialité et d'une durée limitée.

¹⁰ Organisation licence-master-doctorat, <<http://www.enseignementsup-recherche.gouv.fr/cid20190/organisation-licence-master-doctorat-l.m.d.html>>, consulté le 20 juin 2012.

¹¹ *European Credits Transfer System*, <http://ec.europa.eu/education/programmes/socrates/ects/index_fr.html>, consulté le 20 juin 2012.

professionnels (**ECVET**¹²) permettent de décrire une certification (titres, diplômes) en termes d'unités d'acquis d'apprentissage (savoirs, aptitudes et compétences) transférables et capitalisables auxquelles sont associés des points de crédit, et ceci quels que soient la période ou le lieu sur l'espace européen. Ce principe applicable aussi bien aux cursus longs et diplômants qu'à des actions de durée plus courtes favorise la formation continue. Tout un chacun peut ainsi poursuivre le développement de ses compétences en fonction de ses acquis et besoins. Ce dispositif suppose qu'une visibilité sur les acquis et compétences (voir point suivant) et se met en place en articulation avec les principes généraux de la VAE – validation des acquis de l'expérience¹³ ;

- **l'Europass**¹⁴ est un outil type « portfolio de compétences » assurant la visibilité des compétences et des qualifications de son possesseur (CV Europass, passeport de langues, Europass mobilité, supplément au diplôme ou au certificat). Un travail particulier est fait au sein de ces programmes pour faciliter la mobilité des chercheurs au sein de l'espace européens¹⁵ ;
- le développement de référentiels de **certification, compétences et aptitudes professionnels** est un axe déjà ancien qui perdure au sein de différents programmes. Les réalisations montrent une grande diversité dans les orientations choisies pour ces référentiels : une fonction précise (« Assistant de managers »¹⁶), un métier (Multimédia audiovisuel – NAME¹⁷ ou Euroréférentiel des professionnels de l'information-documentation) ou un secteur industriel (l'informatique)¹⁸.

Le projet européen du *Référentiel pour les métiers de l'information-documentation* développé à partir de 1999 a été rendu possible par l'existence préalable d'un réseau professionnel actif au niveau européen, l'ECIA (*European Council of Information Associations*)¹⁹. Grâce à cette instance qu'ont été publiées en 1996²⁰ une *Déclaration sur le droit de la propriété* et en 1999 un *Code déontologique* qui fut à la base de bon nombre de codes déontologiques en France²¹. Le travail s'est donc poursuivi au sein de cette structure

¹² *European Credit system for Vocational Education and Training*, <http://ec.europa.eu/education/ecvt/index_fr.html> ; Programme-cadre pour la compétitivité et l'innovation (CIP), <http://ec.europa.eu/cip/index_fr.htm>, consultés le 20 juin 2012.

¹³ <<http://www.vae.gouv.fr/>>, consulté le 20 juin 2012.

¹⁴ <<http://europass.cedefop.europa.eu/europass/home/homav/Introduction.csp>>, consulté le 20 juin 2012.

¹⁵ Portail mobilité des chercheurs, <<http://ec.europa.eu/euraxess/>>, consulté le 20 juin 2012.

¹⁶ Assistant de managers, <http://www.euroast.eu/IMG/pdf/Le_guide_EURO-AST_aout_2010.pdf>, consulté le 20 juin 2012.

¹⁷ Nomenclature Analytique du Multimédia Européen <<http://www.namemultimedia.com/fr/presentation.html>>, consulté le 20 juin 2012.

¹⁸ Secteur informatique, <<http://www.ecompetences.eu/>>. Le Cigref en France a aligné son référentiel national à ce cadre européen, <http://www.cigref.fr/cigref_publications/2011/08/2011-metiers-systemes-information-grandes-entreprises-nomenclature-rh-cigref.html>, consultés le 20 juin 2012.

¹⁹ Voir section 4.2.

²⁰ ECIA 1996, <<http://www.adbs.fr/declaration-de-l-ecia-sur-la-propriete-intellectuelle-et-le-droit-de-copie-13612.htm?RH=REVUE>>, consulté le 20 juin 2012.

²¹ ECIA 1999, <<http://www.adbs.fr/code-deontologique-de-l-ecia-1980.htm>>, consulté le 20 juin 2012.

autour de deux projets européens : DECIDoc²² – avec l'EuroRéférentiel des compétences en 1999 et CERTIDoc²³ sur la certification en 2004 [7]. Ces deux actions ont eu un impact plus large auprès de pays européens n'ayant pas participé aux travaux²⁴ ou auprès de communautés spécialisées²⁵.

À ces outils spécifiques pour l'éducation, la formation ou le développement des compétences, s'ajoutent deux outils transversaux : des programmes de recherche dédiés comme Leonardo ou Grundtvig [4], et un environnement de normalisation puissant, le CEN²⁶ [3] plus particulièrement orienté par l'intégration des TIC à la formation ouverte et à distance (*e-learning*). Cette instance de normalisation est très peu connue des professionnels de l'information-documentation, mais très active dans différents secteurs spécialisés – espace, environnement, éducation, médecine, nanotechnologies...²⁷.

Cet ensemble de politiques et d'instruments, élaboré grâce à la participation d'institutions publiques, d'organisations professionnelles ou civiles, publiques ou privées vise à faciliter la mobilité certes géographique mais aussi professionnelle ou sociale des citoyens pour leur permettre d'aborder les évolutions des métiers, et de façon plus globale de la vie quotidienne, de façon positive.

3.1.2. La société numérique et les TIC

La dissémination des technologies numériques dans toutes les couches sociales dans la plupart des pays transforme les relations entre les individus et leurs relations à la vie quotidienne et au travail. Cela impacte nécessairement la recherche scientifique de plus en plus interrogée par la société.

La stratégie numérique pour la société proposée par l'Union européenne²⁸ a été initiée formellement il y a 20 ans. Elle vise à combattre la fracture numérique²⁹ entre pays (et on l'espère entre citoyens au sein d'un même pays) à travers différentes actions qui toutes visent à faire communiquer les citoyens entre eux. Sur ce terrain, des professionnels de l'information-documentation sont impliqués dans le développement de contenus numériques

²² Développer les EuroCompétences pour l'information et documentation.

²³ Certification des professionnels de l'information-documentation.

²⁴ Cartographie, <<http://referentieleninfodoc.wordpress.com/tag/euroreferentiel>>, consulté le 20 juin 2012.

²⁵ L'European Association for Health Information and Libraries avec un numéro de la revue sur la certification, *Journal EAHIL*, 6(4 *Library, Education Programmes and Certification*), November 2010, <http://www.eahil.net/journal/journal_2010_vol6_n4.pdf>, consulté le 20 juin 2012.

²⁶ Comité européen de normalisation.

²⁷ Secteurs normatifs du CEN, <<http://www.cen.eu/cen/Sectors/Sectors/Pages/default.aspx>>, consulté le 20 juin 2012.

²⁸ TIC : programme d'appui stratégique en matière de technologies de l'information et de la communication (TIC), <http://ec.europa.eu/cip/ict-ppp/index_fr.htm>, consulté le 20 juin 2012.

²⁹ Société numérique, <<http://ec.europa.eu/digital-agenda/>>; Contre la fracture numérique, "e-Inclusion", <http://ec.europa.eu/information_society/activities/einclusion/index_en.htm>; IST World, « Technologies for the growing information society », <<http://www.ist-world.org/default.aspx>>, consultés le 20 juin 2012.

(bibliothèques numériques, contenus en ligne), sur deux domaines particuliers : la culture avec de très nombreux projets³⁰ et l'information administrative.

Ces actions sont complétées par de nombreux programmes de recherche dans le domaine des technologies de l'information³¹.

La difficulté de toutes ces actions réside dans leur ancrage ou articulation avec les dispositifs nationaux de valorisation et développement des compétences (formation initiale et continue, formelle ou non formelle) et des métiers.

3.2. Impacts d'une recherche scientifique articulée à la société et à l'innovation

Les relations entre la science et l'industrie ou l'entreprise, que l'on retrouve actuellement sous la mention de recherche – innovation³², ont fait l'objet d'innombrables analyses (« finalisation » de la science ; science « entrepreneuriale » ; recherche finalisée...).

Dans les faits, la recherche orientée vers la résolution de problèmes techniques au sein de la recherche dite académique est très ancienne comme le montre la création dans les années 1880 des instituts techniques dans les universités technologiques ou encore l'association permanente par le savant Pasteur d'une recherche médicale fondamentale toujours articulée à une recherche clinique. Ce phénomène s'est poursuivi dans la deuxième partie du XX^e siècle en faisant « reconnaître des disciplines nouvelles, dédiées aux applications industrielles (toutes les variantes du génie, l'automatique, l'informatique, les biotechnologies, etc.) » [8] conduisant aux technosciences. Au-delà des deux modèles extrêmes de la recherche académique et de celui de l'innovation, plusieurs modes de coopération et de fertilisation croisée entre la recherche universitaire et l'entreprise se conjuguent aujourd'hui et se développent à travers les programmes européens de recherche et/ou d'autres programmes d'action comme ceux liés aux technologies de l'information³³ [3.6] ou à la société numérique³⁴.

Parallèlement à cette articulation de la recherche (et de ses résultats) aux milieux industriels et des entreprises, la recherche évolue dans ses pratiques et s'ouvre à une collaboration avec la société [10] par le biais de méthodes participatives³⁵.

Ce nouvel environnement de la recherche fait évoluer les besoins et diversifier les pratiques [11] d'une partie importante des chercheurs tout au long de leur carrière. Si cette évolution n'est pas née avec l'Union européenne, elle a été renforcée et dynamisée par

³⁰ Par exemple : le Réseau européen des Universités pour enfants <<http://eucu.net/>>, <<http://www.escity.org/>>, consultés le 20 juin 2012.

³¹ <http://ec.europa.eu/cip/ict-ppsp/index_fr.htm>, consulté le 20 juin 2012.

³² <http://ec.europa.eu/cip/index_fr.htm>, consulté le 20 juin 2012.

³³ <http://ec.europa.eu/cip/ict-ppsp/index_fr.htm>, consulté le 20 juin 2012.

³⁴ <http://ec.europa.eu/information_society/digital-agenda/links/index_fr.htm> ; <http://ec.europa.eu/information_society/activities/einclusion/index_en.htm> ; <<http://www.ist-world.org/default.aspx>>, consultés le 20 juin 2012.

³⁵ À un niveau national, nous pouvons citer les Instituts thématiques multi-organismes (ITMO) évoqués par l'Inserm, <<http://www.aviesan.fr/>>, consulté le 20 juin 2012.

l'organisation conjointe des politiques (directives en santé, environnement...) et les programmes de recherche proposés par l'UE. Ces changements portent sur les rôles que peuvent jouer chercheurs et scientifiques sollicités pour des expertises scientifiques [12] à des fins d'actions politiques ou pour des débats citoyens. Comme experts ou comme concepteurs de programmes et de projets, ils peuvent être amenés à dépasser leur propre spécialité, en développant des connaissances sur des domaines connexes à leurs spécialités. Ces nouvelles activités impactent également les pratiques de recherche elle-mêmes ainsi que l'organisation du travail coopératif aux différents niveaux (locaux, nationaux, internationaux puis européens), la variété des résultats et bien sûr des acteurs avec lesquels les chercheurs doivent interagir : chercheurs de cultures et pratiques différentes, industriels et professionnels d'entreprises, politiques, société civile.

Il apparaît alors que ces environnements de travail non seulement reposent la question du périmètre de l'IST et de l'information spécialisée, mais aussi des méthodes et techniques à déployer, ou encore de la nature du soutien à apporter demain aux chercheurs. Citons ici un seul cas : la recherche dite participative impose des rythmes de travail appropriés, plus longs et/ou différents de ceux habituellement requis pour une recherche non participative. De plus ses résultats peuvent porter sur des sujets hors de son propre champ disciplinaire. Dans ce contexte, la durée et le format de la thèse telle qu'ils existent aujourd'hui ne permettent pas de capitaliser toutes les connaissances réellement produites dans ces nouveaux contextes et de valoriser les activités des postulants. Une autre conséquence de ces nouvelles pratiques de recherche porte sur les « livrables » liés à tous ces projets de recherche depuis plus de 20 ans : Comment les tracer et les traiter ? Comment sont-ils valorisés dans le champ disciplinaire, au sein de l'institution et pour l'histoire des sciences ?³⁶

3.3. Place et périmètre de l'IST dans les politiques de l'UE

Si l'Union européenne intervient concrètement au niveau de la recherche et de son implication sociale (débat démocratique, culture et vulgarisation scientifique), elle n'a jamais agi directement au niveau de l'IST. L'UE intervient sur un plan plus global – la recherche, l'éducation et la culture, les métiers, les TIC, les contenus numériques, le droit... en vue de créer des *cadres d'action communs* permettant à chaque « communauté » – chercheurs mais aussi institutions publiques, société civile ou professionnelle, de développer des politiques et des règles adaptées à leur contexte. Le projet VOA3R³⁷ d'archives ouvertes dans le secteur de l'agriculture s'inscrit dans une action qui vise certes les sciences mais qui se trouve en réalité dans le « Programme d'appui stratégique en matière de technologies de l'information et de la communication (TIC) »³⁸.

³⁶ C'est dans ce contexte que le standard CERIF (*Common European Research Information Format*) a été mis sur pied à partir des années 1980 [7]. CERIF [EU7] est une norme pour l'échange et la gestion des données de recherche en cours – information sur les chercheurs, les organisations, les projets, les livrables et les financements tout au long des processus de recherche. CERIF est maintenant une Recommandation de l'Union européenne. Site officiel du format de données, <<http://www.eurocris.org/Index.php?page=CERIFreleases&t=1>>, consulté le 20 juin 2012.

³⁷ *Virtual Open Access Agriculture & Aquaculture Repository: Sharing Scientific and Scholarly Research related to Agriculture, Food, and Environment*, Journées Frédoc 2011, <<http://renatis.cnrs.fr/spip.php?article199>>, consulté le 20 juin 2012.

³⁸ <http://ec.europa.eu/cip/index_fr.htm>, consulté le 20 juin 2012.

Au niveau de la recherche scientifique, les actions de l'UE accordent une place prépondérante à l'intégration des technologies de l'information dans les sciences et aux « e-plateformes » [5] dans le contexte de la « société numérique ». Ces plateformes ne sont pas perçues comme des entrepôts de données ou de documents scientifiques mais plutôt comme des « laboratoires », des espaces de travail pour la recherche et la communauté de chercheurs³⁹. Ce sont des

« Laboratoire(s) scientifique(s) virtuel(s) permettant aux [collectif de] chercheurs, grâce à l'utilisation des nouvelles technologies tels que la réalité virtuelle et Internet, d'effectuer leurs travaux de recherche en collaboration, à distance et en temps réel, sans être soumis à des contraintes d'ordre géographique »⁴⁰.

Aujourd'hui le terme de plateforme de recherche et/ou d'infrastructure est plus courant, mais l'on retrouve la même idée sous-jacente de mutualisation et de travail de recherche... S'y trouve donc bien sûr les données scientifiques, utilisées et produites, mais aussi les instruments numériques, les modèles normatifs, les référentiels terminologiques, les projets et actions associées, les documents de gestion et d'évaluation, les profils de la communauté étendue... Les thèses ou revues, toujours centrales dans la production de la recherche, doivent ainsi partager le terrain des ressources scientifiques et techniques à manager⁴¹.

Au plan des acteurs, les intervenants sont nombreux : des institutions et organismes publics, mais aussi des entreprises et industrielles, des associations scientifiques et professionnelles ou des associations issues de la société civile.

On le voit, le périmètre de l'information scientifique et technique s'est largement étendu par rapport à celui couvert par les publications scientifiques plus traditionnelles.

Quels sont dès lors les impacts professionnels ?

³⁹ « Les grands équipements en sciences de la vie : quelle politique publique ? », V. Mangematin et A. Peerbaye, *Revue française d'administration publique*, 4(112-Administration et politique de la recherche), 2004, p. 705-718. <www.cairn.info/revue-francaise-d-administration-publique-2004-4-page-705.htm>, consulté le 20 juin 2012.

⁴⁰ Grand dictionnaire terminologique, <http://www.granddictionnaire.com/btml/fra/r_motclef/index800_1.asp>, consulté le 20 juin 2012.

⁴¹ Une revue numérique comme IPOL.im (*Image Processing On Line*) est plus facile à envisager dans ce type d'environnement de travail. Cette revue expose de nouveaux algorithmes dans l'analyse et le traitement des images numériques développés par les chercheurs. Ce produit éditorial propose à d'autres chercheurs d'utiliser ces algorithmes avec leurs propres données, et éventuellement d'ajouter les résultats obtenus à la revue comme autant de nouvelles extensions à l'article initial. Site de la revue : <<http://www.ipol.im/>>, consulté le 20 juin 2012.

4. Impact sur les métiers et les organisations professionnelles de l'information-documentation

Comme le dit Michel Serres⁴², nous travaillons tous avec ces drôles de machines que sont les ordinateurs (et aujourd'hui des tablettes) : comment dès lors distinguer les métiers entre eux ? Les citoyens, y compris dans des univers les plus manuels manipulent de l'« information », en produisent, la gèrent, la communiquent.

Comment se positionnent les professionnels de l'information-documentation ? Quelle place pour une Europe des métiers de l'information-documentation ?

4.1. Fonctions et métiers

Revenons, non pas aux structures professionnelles de l'information-documentation traditionnelles (archives, bibliothèques, documentation, musées, centres culturels scientifiques), mais plutôt aux activités de recherche dans le contexte étendu tel que nous venons de l'esquisser et tel qu'il apparaît dans le cadre des e-sciences. Il s'agit non seulement de capitaliser pour les générations futures, mais de soutenir le travail des communautés de la recherche (chercheurs et autres collaborateurs : partenaires, techniciens)⁴³ par des méthodes, des techniques et outils et des services adaptés.

D'une vision unifiée par rapport aux activités à réaliser (acquérir, décrire, indexer...), il apparaît que pour soutenir efficacement le travail des chercheurs et remplir les missions patrimoniales avec le même niveau d'engagement et de qualité pour ces deux grands missions, il convient de distinguer plus fortement :

4.1.1. Les spécificités des domaines scientifiques, les pratiques des acteurs, la variété des ressources à mutualiser sur les infrastructures de travail, leurs usages

Partant du principe qu'aujourd'hui – plus encore qu'il y a 30 ans au démarrage des techniques infométriques, une personne même au sein d'un collectif ne peut s'approprier par lecture directe toutes les connaissances qui lui sont nécessaires, deux types d'actions sont possibles pour assister les communautés de chercheurs : la production de ressources secondaires ou de bases de connaissances spécialisées telles que celles proposées à l'Inserm⁴⁴ et/ou le développement d'outillages automatiques capables de traiter des masses d'information et de données (infométrie). Ces services constituent une suite logique de l'offre traditionnelle qui va des bibliographies, aux synthèses et dossiers documentaires. Ces offres devraient pouvoir être proposées et pour partie fabriquées avec les équipes de recherche,

⁴² *Les nouvelles technologies : révolution culturelle et cognitive*, M. Serres, 11 décembre 2007, <http://interstices.info/jcms/c_33030/les-nouvelles-technologies-revolution-culturelle-et-cognitive>, consulté le 20 juin 2012.

⁴³ Which librarians for tomorrow?, M.-P. Baligan, A.-L. Achard, Cemagref Lyon, <http://archivesic.ccsd.cnrs.fr/sic_00001554>, au Congrès Euraslic "Open Water – Open source", Split, Croatia, 4-6th May 2005.

⁴⁴ L'IST à l'Inserm, N. Duchange et D. Pélaprat, 2011, ISIDORA. Bases de connaissances proposées par l'Inserm, <<http://extranet.inserm.fr/information-scientifique-et-technique/bases-de-connaissances>>, consulté le 20 juin 2012.

comme l'ont été en leur temps les bibliographies et outils de gestion bibliographique. En fonction des compétences des équipes et de la volonté des communautés de chercheurs, ces activités peuvent aller jusqu'à un appui plus soutenu autour des méthodes et outils développés au sein de chaque communauté spécialisée – par exemple le standard DDI (*Data Documentation Initiative*) de documentation des enquêtes et sondages en sciences humaines et sociales⁴⁵, la TEI⁴⁶ comme langage d'encodage de corpus textuels, jusqu'au développement de projets d'infométrie plus intégrés aux travaux courants des chercheurs ou un appui à la mise en place du format CERTIF pour le management des projets.

4.1.2. Les niveaux d'intervention – au plus près des chercheurs ou dans des actions patrimoniales hors de leur collectif

Actions au plus près des équipes de recherche - Les expériences d'un certain nombre de professionnels de l'information en appui opérationnel (recherche d'information mais aussi gestion de projets de recherche, gestion d'information et de dossiers d'activités de recherche), de conseil ou de formation au plus près des équipes, sur leur propre terrain (*embedded librarians*⁴⁷) montrent l'importance des besoins, l'intérêt pour les chercheurs, l'institution et les professionnels de l'information-documentation de ce type d'activités mais aussi la complexité de ces interventions. Dans ce contexte, associés aux techniciens des laboratoires, des professionnels de l'information-documentaire assurent des activités de soutien dans la gestion et le suivi de l'information scientifique ou managériale des projets de recherche⁴⁸.

Actions patrimoniales – À l'échelle de l'institution, le rôle patrimonial se renforce et les changements que nous venons d'évoquer au niveau des sources, des volumes et des flux de données et d'information, ou sur le plan de l'intégration des technologies de l'information impactent directement les méthodes, outils et compétences des professionnels de l'information. Citons le cas des publications produites dans le cadre des projets de recherche européens. Les rapports et en particulier les rapports organisationnels, méthodologiques, financiers, de recherche, techniques ou stratégiques liés à un programme ou un projet ne sont que très rarement collectés⁴⁹ et/ou difficiles à repérer en raison de pratiques de

⁴⁵ Le réseau Quetelet a participé au développement de ce format en France en faisant un important travail de traduction, <http://www.reseau-quetelet.cnrs.fr/spip/article.php3?id_article=140&var_recherche=DDI&lang=fr>, consulté le 20 juin 2012.

⁴⁶ *Text Encoding Initiative*.

⁴⁷ *S'intégrer à l'environnement de travail des utilisateurs d'information (embedded librarian)*, S. Dalbin, 2010, <<http://referentieleninfodoc.wordpress.com/2010/08/16/pid-integre-embedded-librarian/>>, consulté le 20 juin 2012.

⁴⁸ Par exemple par la participation au développement des outils de management communautaires comme le CERIF, <http://fr.wikipedia.org/wiki/Common_European_Research_Information_Format>, consulté le 20 juin 2012.

⁴⁹ Certes ils se trouvent dans les répertoires du service Communautaire d'information sur la recherche et le développement (Cordis), mais sans traitement approprié, il faut connaître leur existence pour les retrouver, <http://cordis.europa.eu/home_fr.html>, consulté le 20 juin 2012.

description ne permettant pas de valoriser ces matériaux⁵⁰. Une réflexion à la fois archivistique⁵¹, documentaire et bibliographique reste à conduire sur les genres documentaires liés à la gestion de projet (et de projets européens) et plus globalement sur les politiques d'acquisition et de traitement des dossiers de ces projets et de leurs contenus.

Si le xx^e siècle visait la rationalisation des techniques documentaires par une unification des pratiques, le xx^e siècle impose de déployer des services différenciés à la fois au plus près des équipes et/ou au niveau des disciplines et secteurs scientifiques dans des lieux intermédiaires patrimoniaux et dans des dispositifs communs d'accueil. Prendre acte de la distinction entre travail sur le patrimoine et soutien aux équipes dans leurs espaces de travail, c'est aussi accepter de mettre en avant et d'utiliser des outils, méthodes et pratiques issus des équipes ou domaines scientifiques.

4.2. Quelle organisation professionnelle de l'information-documentation pour l'Europe ?

4.2.1. Nouveau terrain d'intervention

Le périmètre de l'IST s'étend, les relations avec les équipes de recherche et plus globalement la société civile évoluent, les activités et modalités d'intervention se diversifient : quelle organisation professionnelle pourrait alors représenter ces métiers ?

Revenons à l'Europe. Le développement, ces dernières années, des projets de recherche sur ce territoire suscite un intérêt grandissant pour ce périmètre d'action en complément d'actions locales ou internationales. Les professionnels y participant, enthousiastes malgré la charge de travail⁵², comprennent les bénéfices apportés par ces démarches, bénéfices personnels et pour l'institution bien sûr, mais également pour l'IST et pour l'ensemble des métiers. Se pose alors la question d'une représentation européenne des professionnels de l'information-documentation. Pour qui et pour quoi faire ?

4.2.2. Revenons quelques décennies en arrière

Au xix^e siècle, des structures initiées en Europe mais résolument internationales se sont organisées autour de l'étude et du développement de techniques documentaires novatrices.

⁵⁰ Exemple issu du Sudoc. Comment trouver des rapports produits dans le cadre de projets de recherche de l'Union européenne ? Une recherche approfondie reste à faire mais quelques tests font ressortir certaines productions issues de projets en information-documentation. Les traitements bibliographiques ne fournissent que très peu de données sur le contexte, <<http://www.sudoc.abes.fr/DB=2.1/SRCH?IKT=12&TRM=095320156>>, consulté le 20 juin 2012. Si la mention d'auteur collectif et d'éditeur scientifique (Projet européen MINERVA) est bien mentionné, la politique d'indexation du sujet n'est pas clairement établie rendant difficile l'accès à ces ressources, les rapports (technique, de recherche ou d'organisation) issus de ces travaux ne sont pas distingués des « Livres », le type de coopération reste « international », le niveau européen n'étant pas mentionné.

⁵¹ Le dossier du projet de recherche dépasse le cadre de l'institution – comment préserver son intégrité ? Comment partager et mettre en cohérence les systèmes d'information touchant différentes institutions et/ou organismes ?

⁵² Il s'agit bien ici de projet de recherche en information et non d'un soutien d'un professionnel à des projets de recherche des chercheurs de l'institution.

Elles regroupaient des bibliothèques intégrant des services documentaires mais surtout des sociétés savantes ainsi que des organismes documentaires spécialisés d'unités de recherche ou de grands industriels (chimie, énergie, santé...). Ces organismes scientifiques ou industriels étaient, à cette époque, déjà structurés suivant trois logiques : des spécialités scientifiques ou sectorielles (énergie avec la houille, Société de géographie...), un niveau international et national par pays.

L'IIB (Institut international de bibliographie) est ainsi créée par Paul Otlet et Henri La Fontaine en 1895⁵³. Les études conduites par cet institut ne se limitent plus à la bibliographie et s'étendent au domaine de la Documentation. L'IIB change de nom en 1931 et devient l'IID (Institut international de documentation), puis en 1937, la FID (Fédération internationale de documentation). Ce mouvement visait à la fois à rationaliser des pratiques préexistantes, parfois anciennes de traitement de l'information par des méthodes et outils appropriés et à mettre en commun (mutualisation dirions-nous aujourd'hui) le résultat de ces travaux auprès des organisations de la recherche et de l'administration.

Pendant ce temps, les bibliothèques mettaient en place l'IFLA en 1929⁵⁴.

Après la Seconde Guerre mondiale, la configuration du paysage de l'information spécialisée change.

Les États se reconstruisent en établissant des politiques nationales de l'IST [13].

Dans le même temps, une place de choix fut donnée à une structure comme l'ONU⁵⁵. Si l'Unesco⁵⁶ a fait la promotion de structures de gestion nationales d'IST (programmes Unisist/PGI⁵⁷) et poursuit toujours ses actions dans le contexte d'une société des savoirs (éthique, éducation tout au long de la vie...), l'ONU par le biais de ses structures par grands domaines, a donné la priorité à la construction de systèmes documentaires internationaux spécialisés. Agrovoc/Agris⁵⁸ développé par la FAO⁵⁹ ou l'OCDE⁶⁰ sont emblématiques de cette démarche d'après-guerre.

4.2.3. Quelques pistes

Parallèlement, les sociétés savantes et les organisations scientifiques et professionnelles se déploient, couvrant des territoires géographiques de plus en plus importants avec l'arrivée plus récente des pays émergents. Elles se sont renforcées tout au long de ce XX^e siècle à la fois sur leur dimension internationale, continentale et locale, l'Internet et le Web ayant

⁵³ Sur le site du Mundaneum, <<http://www.mundaneum.be/index.asp?ID=424>> consulté le 20 juin 2012.

⁵⁴ Histoire de l'*International Federation of Library Associations*, <<http://www.ifla.org/en/history>> consulté le 20 juin 2012.

⁵⁵ Organisation des Nations unies.

⁵⁶ *United Nations Educational, Scientific and Cultural Organisation*.

⁵⁷ *United Nations International Scientific Information System/Programme général d'information*.

⁵⁸ Agrovoc est un thesaurus pour l'*International Information System for the Agricultural Sciences and Technology*.

⁵⁹ *Food and Agriculture Organization of the United Nations*, <<http://aims.fao.org/fr>> consulté le 20 juin 2012.

⁶⁰ Organisation de coopération et de développement économiques.

grandement facilité ces extensions. Chaque discipline, secteur ou domaine s'est doté de méthodes et d'outillages en propre permettant de produire et de capitaliser à la source, de structurer, d'organiser et d'annoter ses « connaissances ». Ces pratiques intègrent des techniques documentaires au fur et à mesure de leur apparition, celles développées en 1900 (nomenclatures, classification, bibliographie, index), dans les années 1950 (bases documentaires, recherche combinée et thésaurus), puis celles de la seconde partie du XX^e siècle comme l'infométrie ou pour les plus récentes, celles qui associent schémas de métadonnées/ontologies informatiques et « terminologies » spécialisées en vue d'annotation des ressources selon des structures d'analyse fortement enrichies par rapport aux outils bibliographiques et éditoriaux traditionnels.

Si ces organismes scientifiques ont été au XVIII^e et XIX^e siècles à l'origine de l'édition scientifique (les revues ainsi que les bibliographies scientifiques émanent historiquement de ces instances avant de passer dans d'autres mains, Éditeurs professionnels et Bibliothèques), elles ont toujours été le fer de lance des évolutions des principes, méthodes et outils de production, capitalisation, publication et communication de l'IST⁶¹. Nous pouvons citer le cas bien connu maintenant du développement des archives ouvertes par la communauté des physiciens ou l'Observatoire virtuel (OV) astronomique⁶².

Concernant la régionalisation de l'IST, les deux axes en présence au début du XX^e siècle – rationalisation du traitement documentaire d'un côté, et bibliothèques de l'autre, ont développé chacun des structures européennes soutenant leurs actions.

Pour la documentation, dans la lignée de l'IIB ou de la FID, une structure informelle s'était constituée dans les années 1970 entre 6 Pays⁶³ : le WERTID⁶⁴. Cette structure visait la « promotion des contributions des études et de la recherche, publiques ou privées ». L'ECIA prit la suite du WERTID en 1992 avec une structure plus formelle, et en étendant son périmètre d'intervention à l'information-documentation (voir section 2) avec des actions sur les métiers, le droit et les questions de déontologie.

Liber, la Ligue des bibliothèques européennes de recherche, s'est créée sous les auspices du Conseil de l'Europe en 1971 afin de « représenter et (de) promouvoir les intérêts des bibliothèques de recherches d'Europe »⁶⁵. Depuis 2002, Liber s'est doté d'un ambitieux programme de développement autour de 4 axes : la communication scientifique ; les services de bibliothèques ; le management des bibliothèques ; les services et opportunités offerts à ses membres. La réalisation de nombreux projets européens sous sa bannière permet de remplir ses objectifs en particulier les objectifs patrimoniaux⁶⁶.

⁶¹ VOA3R, *Journées Frédoc 2011*, <<http://renatis.cnrs.fr/spip.php?article199>>, consulté le 20 juin 2012.

⁶² Observatoire Virtuel (OV) astronomique, *Journées Frédoc 2011*, <<http://renatis.cnrs.fr/spip.php?article209>>, consulté le 20 juin 2012.

⁶³ Allemagne, Espagne, Belgique, France, Portugal, Royaume-Uni.

⁶⁴ *West European Round Table on Information and Documentation*.

⁶⁵ Article 2 des statuts de Liber, <<http://www.libereurope.eu/node/781>>, consulté le 20 juin 2012.

⁶⁶ *Report on integration of data and publications*, October 17th, 2011, dans le cadre du projet européen ODE, <<http://www.libereurope.eu/news/report-on-integration-of-data-publications>>, consulté le 20 juin 2012.

4.3. Une organisation européenne

Mais chacun des regroupements existants occulte une partie des métiers ou activités autour de l'information spécialisée : soit l'angle chercheurs-producteurs (les associations scientifiques et techniques sont peu présentes), soit l'éducation (avec une association type EBLIDA⁶⁷), soit d'autres acteurs patrimoniaux comme les musées ou les archives avec lesquels des relations plus fortes restent à construire, soit des catégories de traitements avancés (infométrie, veille).

Une structure de même nature que la structure française l'IADB (Interassociation archives bibliothèques documentation)⁶⁸ étendue à d'autres contextes comme l'intelligence économique, pourrait jouer au plan européen un rôle structurant plus large que les seules associations existantes ne peuvent couvrir. Bien sûr cela supposerait de cadrer son action à des projets-cadre communs sur les métiers, l'éthique, les usages des citoyens, les structures de normalisations voire certaines méthodes ou outils partagés. Un pilotage, qui s'il ne se limite pas à l'information dite scientifique, permettrait de répondre aux enjeux européens actuels : sciences et société, sciences et entreprises ou industries (innovation), sciences et administration.

Le territoire européen articulé aux territoires internationaux, nationaux et locaux, peut être un formidable vecteur de dynamisme et de réflexions autour des questions clés de l'information dans une société des savoirs.

Bibliographie

- [1] M. André, « Présentation de la DIST (Direction de l'information scientifique et technique) », ARPIST (Bordeaux 1^{er} février 2011), FREDOC 2011 (Bordeaux octobre 2011) et de la MISTRD par Françoise (ISIDORA 2001).
- [2] M. André, « Le 7^e programme-cadre dans l'histoire de la recherche européenne » (entretien), 2007, <http://ec.europa.eu/research/rtdinfo/special_fp7/fp7/01/article_fp709_fr.html>, consulté le 20 juin 2012.
- [3] HORIZON 2020 The Framework Programme for Research and Innovation, <http://ec.europa.eu/research/horizon2020/index_en.cfm?pg=home#>, consulté le 20 juin 2012.
- [4] Commission européenne, « Principaux développements et résultats politiques dans le domaine de l'éducation et la formation depuis l'an 2000 », <http://ec.europa.eu/education/lifelong-learning-policy/doc1120_fr.htm>, consulté le 20 juin 2012.
- [5] E-sciences, « Infrastructures TIC pour la science en ligne », Communication de la Commission au Parlement européen, au Conseil, Comité économique et social européen et au Comité des régions intitulées « Infrastructures TIC pour la science en ligne », **COM(2009) 108 final**, <http://europa.eu/legislation_summaries/information_society/internet/si0006_fr.htm>, consulté le 20 juin 2012.
- [6] R. Russell, *An introduction to CERIF*, UKOLN (United Kingdom Office for Library and Information Networking), June 2011, <www.ukoln.ac.uk/rim/documents/Introduction_to_CERIF_1.0.pdf>, consulté le 20 juin 2012.

⁶⁷ *European Bureau of Library, Information and Documentation Associations*, <www.eblida.org>, consulté le 20 juin 2012.

⁶⁸ <<http://www.iabd.fr>>, consulté le 20 juin 2012.

- [7] J. Meyriat, « CERTIDoc : un discret essai national transformé en un succès continental », *AIDaInformazioni*, **23(4)**, octobre-décembre 2005. <[http://eprints.rclis.org/bitstream/10760/10506/1/meyriat23\(4\).pdf](http://eprints.rclis.org/bitstream/10760/10506/1/meyriat23(4).pdf)>, consulté le 20 juin 2012.
- [8] M. Grossetti, « L'académisation des savoirs techniques. La lente progression des sciences d'ingénieurs dans les universités et la recherche académique », Communication pour le Groupe de Travail « Sciences, innovation technologiques et Sociétés », *AISLF*, Tours, juillet 2004 <<http://halshs.archives-ouvertes.fr/halshs-00476803/en>>, consulté le 20 juin 2012.
- [9] M. Grossetti et B. Milard, « Les évolutions du champ scientifique en France à travers les publications et les contrats de recherche », *Actes de la recherche en sciences sociales*, **3(148-Entreprises académiques)**, 2003, p. 47-56. <www.cairn.info/revue-actes-de-la-recherche-en-sciences-sociales-2003-3-page-47.htm>.
- [10] D. Spoel, *La démocratie participative, complément de la démocratie représentative*, 1^{er} novembre 2006, <<http://www.forum-civil-society.org/spip.php?article143>>, consulté le 20 juin 2012.
- [11] Y. Polity, « Le comportement des chercheurs dans leur activité (y compris face à la documentation) », *Actes de la journée d'études du XIX^e Congrès national de l'ADBU* (Association des directeurs des bibliothèques universitaires): Recherche et documentation, 17 septembre 1999, Dunkerque, <<http://www.iut2.upmf-grenoble.fr/RI3/Adbu.htm>>, consulté le 20 juin 2012.
- [12] B. Chevassus, *Repenser les relations sciences-sociétés : une contrainte de plus ou un atout pour relever les défis de l'agriculture et de l'alimentation mondiale ?*, Conférence à l'Inspection générale de l'agriculture, Paris, 24 novembre 2008, <<http://www.epbrs.org/PDF/chevassus.pdf>>, consulté le 20 juin 2012.
- [13] J.-M. Salaün, « Les politiques publiques en information scientifique et technique », in M. Poulain (dir.), *Histoire des bibliothèques françaises : les bibliothèques au xx^e siècle, 1914-1990* Paris, **Electre**, p. 685-701, 2009.

Outils pour la qualité des services universitaires de documentation : Libqual et Qualibib

Olivier Chourrot

*Pôle de contractualisation,
ministère de l'Enseignement supérieur et de la Recherche,
Paris, France*

Résumé

La LRU et le passage aux responsabilités et compétences élargies conduisent les universités à entreprendre des démarches qualité répondant à une triple exigence : s'autoévaluer afin de se préparer aux évaluations externes de l'AERES, améliorer la qualité de leurs services dans un contexte concurrentiel, adopter des standards de qualité communs aux autres établissements afin de développer la coopération. Les services de documentation ont, d'une certaine manière, anticipé ces évolutions en se dotant d'outils de mesure (Libqual+) et d'amélioration de la qualité (Qualibib), dont l'usage contribue à la performance des universités. Ces outils, qui peuvent intéresser tous les bibliothécaires et documentalistes, seront présentés dans leur contexte de mise en œuvre.

Abstract:

Providing tools for the quality improvement of university documentation: Libqual and Qualibib

Tools for the quality improvement of university documentation services: Libqual and Qualibib.

The LRU (*Loi relative aux libertés et responsabilités des universités*), a French law on the enhancement of power and responsibilities within universities, stipulates that those undertaking quality initiatives should meet the following three requirements:

- self-assessment to be prepared for evaluation by an external agency (AERES);
- quality improvement of their services in a competitive environment;
- implementation of common quality standards to develop cooperation with other institutions.

To achieve these criteria, university documentation services have anticipated these developments by using metrics (LibQUAL+) and quality improvement (Qualibib), both of which contribute to the universities' performance. These tools may be of interest to all libraries and information services and will be presented in their context of implementation.

IdRef

Les référentiels Sudoc¹ comme services aux réseaux documentaires

Yann Nicolas et Claude Oulion

ABES, 227 avenue Professeur Jean-Louis Viala
BP 84308, 34193 Montpellier Cedex 5, France

1. IdRef². Pourquoi ?

1.1 Des référentiels ? Pour résoudre quels problèmes ?

Un référentiel comme IdRef se propose de résoudre deux types de problèmes bien connus des gestionnaires de l'information en général et de l'information bibliographique en particulier :

- l'homonymie ;
- la dispersion des références.

Les difficultés posées par l'homonymie sont bien connues. Il suffit d'interroger *Google Scholar*, par exemple, avec un nom aussi banal que « John Smith » pour mesurer l'ampleur du bruit rencontré, faute de pouvoir distinguer entre les différents auteurs qui se cachent derrière cette appellation ambiguë. Ce bruit nuit à l'efficacité de la recherche documentaire, mais il brouille également l'identité numérique scientifique du chercheur : il est impossible et surtout dangereux de lui construire des indicateurs de réputation ou de citation sur un périmètre aussi confus.

La dispersion des références pose le problème inverse. Au lieu de rencontrer trop de références derrière la mention d'une personne, on en trouve trop peu. Il est difficile d'agréger toutes les références bibliographiques d'un auteur, tant les canaux de publication sont variés et parallèles : édition classique, archivage institutionnel, auto-archivage, etc.

Par ailleurs, le périmètre documentaire rattaché à un chercheur ne concerne pas seulement ses publications. Un chercheur, c'est également un directeur de thèse (soutenue ou en cours), un rapporteur, un auteur de rapport (littérature grise), un expert sollicité pour des jurys d'appels à projet, un auteur de ressources pédagogiques, etc. Cette collection de ressources documentaires dessine une identité numérique scientifique, mais elle est disparate, sans qu'aucune force centripète ne permette de faire converger efficacement les composantes de cette diversité.

¹ Système universitaire de documentation.

² IdRef pour Identifiants et Référentiels.

Pour illustrer l'impasse de cette identité centrifuge, il suffit de sélectionner un chercheur et de réciter la litanie des bases de données qui donne à voir une portion de son activité scientifique ! Prenons l'exemple de Claude Gnos :

- la liste de ses articles déposés sur HAL³ ;
- la liste des publications auxquelles il est lié dans le Sudoc, comme auteur, directeur de thèse ou éditeur scientifique⁴ ;
- un de ses articles numérisés dans Persée⁵ ;
- sa page sur le site de son laboratoire, affichant une sélection de sa bibliographie⁶ ;
- sa page dans <theses.fr>, qui liste notamment les thèses en préparation qu'il dirige⁷.

Rien ne relie ces différentes sources entre elles, si ce n'est un identifiant Sudoc qui est partagé par le Sudoc et <theses.fr>.

1.2 Les autorités comme solution : le cas du Sudoc

Les problèmes de l'homonymie et de la dispersion des références se posent au sein même d'un catalogue de bibliothèques. De nombreux catalogues, comme le Sudoc dès 2000, ont réussi à surmonter ces difficultés en ayant recours à la stratégie du contrôle d'autorité. Toutes les notices bibliographiques qui décrivent un document en rapport avec une personne P sont connectées à une nouvelle notice qui la décrit. Cette notice est repérée par un identifiant, et non par le nom de la personne décrite, ce qui neutralise le problème de l'homonymie. Cette notice d'autorité permet à la fois de rassembler logiquement et physiquement toutes les références liées à une personne et de ne plus confondre les homonymes.

³ Hyper article en ligne. <<http://hal.archives-ouvertes.fr/aut/Claude+Gnos/>> consulté le 26 juin 2012.

⁴ <<http://www.sudoc.abes.fr/DB=2.1/SET=2/TTL=1/REL?PPN=029950961>> consulté le 26 juin 2012.

⁵ <http://www.persee.fr/web/revues/home/prescript/article/ecoru_0013-0559_1992_num_207_1_4435_t1_0046_0000_3> consulté le 26 juin 2012.

⁶ <<http://leg.u-bourgogne.fr/pages-persos/gnos.htm>> consulté le 26 juin 2012.

⁷ <<http://www.theses.fr/029950961>> consulté le 26 juin 2012.

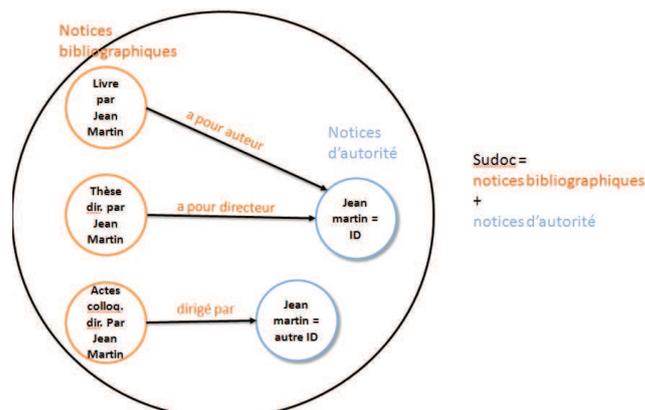


Figure 1 : Les autorités du Sudoc 2001-2010

1.3 Les autorités comme solution : le cas du Sudoc

À partir de 2006, au fur et à mesure où l'ABES⁸ ouvrait d'autres bases de données, la question des autorités s'est posée à nouveau. À la création de Calames⁹, fallait-il concevoir en son sein de nouvelles autorités ? La même question s'est posée à la création de STAR¹⁰, outil de dépôt et de signalement des thèses numériques. La réponse était évidente : Calames et STAR doivent réutiliser le référentiel d'autorités du Sudoc.

Dès 2006, l'ABES a donc imaginé et implémenté des solutions pour que les catalogueurs des deux outils puissent interroger les autorités Sudoc et y lier les manuscrits ou les thèses qu'ils décrivent, sans quitter leur application. Cette solution, basée sur la norme Z 39.50¹¹, avait des défauts (performances, précision des index), mais surtout une lacune fonctionnelle essentielle : il était impossible de créer une nouvelle autorité à partir de Calames ou STAR. Il fallait sortir de son application et utiliser l'interface de catalogue du Sudoc.

Afin de pallier ces défaillances et cette lacune, l'ABES a développé IdRef, une nouvelle application Web dédiée aux autorités Sudoc, mais indépendante de l'architecture technique de ce dernier. Les fonctionnalités de cette application sont décrites dans la seconde partie de ce document.

IdRef est conçue pour servir de référentiel partagé par différentes applications documentaires Web. Dire qu'il s'agit d'un référentiel d'autorité *partagé* a trois significations complémentaires :

- les données du référentiel sont partagées : on ne crée pas un nouveau référentiel pour chaque nouvelle base. En servant de pivot, le référentiel interconnecte entre elles les données qui lui sont liées. C'est la logique du Web de données ;

⁸ Agence bibliographique de l'enseignement supérieur.

⁹ CAtalogue en ligne des archives et des manuscrits de l'enseignement supérieur.

¹⁰ Dépôt national des thèses électroniques : Signalement des thèses électroniques, archivage et recherche.

¹¹ Protocole de communication informatique client-serveur.

- la responsabilité des données IdRef est partagée : il ne suffit pas de consommer un référentiel. Il faut l'alimenter, le maintenir, et ce collectivement. IdRef rend possible cette cogestion décentralisée, en offrant les moyens de créer de nouvelles autorités, ou d'en modifier ;
- IdRef est une solution partagée et donc centralisée. Mais en étant conçue pour se brancher sur d'autres applications Web, en fonctionnant comme un pop-up pour des applications diverses, IdRef cherche à trouver un équilibre entre centralisation et décentralisation.

IdRef : au-delà du Sudoc

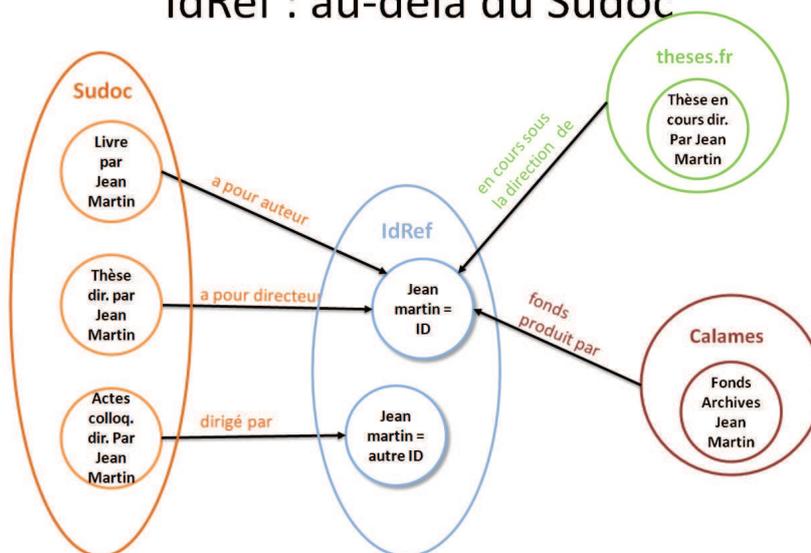


Figure 2 : IdRef : au-delà du Sudoc

1.4 IdRef : les autorités Sudoc, au-delà de l'ABES

Tout ce qui vient d'être dit à propos de Calames et STAR, « clients » d'IdRef, l'est aussi en général à l'échelle des applications documentaires de l'enseignement supérieur et de la recherche. Les archives institutionnelles, les archives ouvertes, les plateformes pédagogiques, les bibliothèques numériques, les systèmes d'information de la recherche (CRIS¹²), etc. ont besoin de s'appuyer sur des référentiels et IdRef souhaite répondre à ces besoins.

¹² National Research Information System.

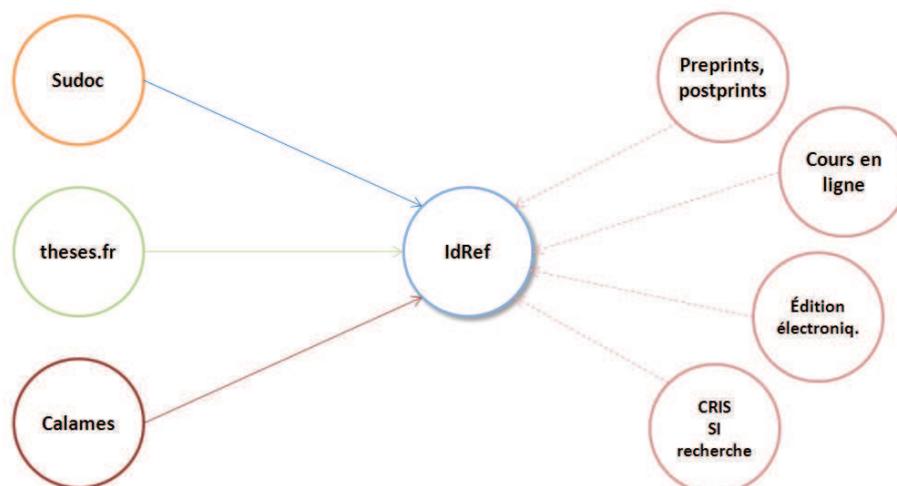


Figure 3 : IdRef

Contrairement à ce que peut laisser penser le schéma ci-dessus, IdRef n'a pas l'illusion d'être le seul référentiel de chercheurs sur le marché.

D'un côté, il existe des référentiels à vocation locale, à l'échelle d'un établissement et de son annuaire, par exemple. Mais ils ne peuvent suffire car on a besoin d'identifier un chercheur tout au long de sa carrière, parfois nomade, et à travers tous ses canaux d'expression et de publication, qui ne sont pas confinés à l'enceinte de son LDAP¹³.

De l'autre côté, il existe aujourd'hui des initiatives globales pour associer un identifiant à tout chercheur (ORCID¹⁴), voire à tout créateur ou producteur culturel (ISNI¹⁵, VIAF¹⁶). Mais comment organiser l'alimentation de ces bases mondiales ? Il nous semble plus fiable et plus prudent d'organiser la collecte et la maintenance des données à une échelle régionale, dans le cadre d'un réseau national par exemple.

Un référentiel à vocation régionale comme IdRef est intermédiaire entre les référentiels locaux et globaux. Ce qui compte, ce n'est pas de chercher le bon référentiel, mais de favoriser la mise en correspondances des identifiants de tous ces niveaux de référentiels. Ainsi, depuis décembre 2011, IdRef est intégré à VIAF, lui-même socle d'ISNI. Dès lors, travailler dans IdRef, c'est aussi contribuer à une visibilité internationale de nos chercheurs et de leurs publications. C'est aussi contribuer à construire le Web de données.

¹³ *Lightweight Directory Access Protocol.*

¹⁴ *Open Researcher and Contributor ID.*

¹⁵ *Code international normalisé des noms.*

¹⁶ *Virtual International Authority File.*

2. IdRef. Comment ?

2.1 Les autorités Sudoc : qu'elles aquò ?

La base d'autorité Sudoc est alimentée à la fois

- par les professionnels du réseau Sudoc, grâce au logiciel WinIBW, en cohérence avec les recommandations du *Guide méthodologique* mis à jour régulièrement par l'ABES ;
- par les professionnels des réseaux STAR, et Calames grâce à l'application IdRef.

Des chargements bimensuels sont effectués par la BnF (mises à jour du référentiel matière RAMEAU¹⁷ en base de production, autorités BnF en base dite d'appui¹⁸) et les mises à jour annuelles du référentiel FMesh par l'INSERM. Ces notices sont réutilisées par les catalogueurs du Sudoc, dans STAR et Calames.

2011

Entité	Janv.	Févr.	Mars	Avril	Mai	Juin	Juil.	Août	Sept.
Auteur / titre	12 924	13 001	13 065	13 124	13 210	13 258	13 321	13 372	13 453
Collectivité / Congrès	229 324	230 700	232 205	233 514	234 948	236 225	237 259	237 941	239 294
Descripteur chronologique	133	133	134	134	134	134	134	134	134
Descripteur iconographique	6	6	6	6	6	6	6	6	6
FMeSH	25 788	25 788	25 788	25 788	25 788	25 788	25 788	26 210	26 210
Nom commun	115 424	115 613	115 796	116 025	116 208	116 459	116 628	116 871	117 008
Nom de famille	2 861	2 877	2 902	2 916	2 931	2 941	2 962	2 965	2 978
Nom de marque	246	246	247	247	247	248	248	248	248
Nom géographique	53 022	53 245	53 464	53 668	53 858	54 049	54 188	54 287	55 066
Notice RCR	3 888	3 901	3 909	3 913	3 921	3 926	3 930	3 932	3 938
Personne physique	1 819 148	1 832 653	1 847 505	1 860 262	1 874 740	1 887 641	1 897 631	1 902 951	1 915 779
Titre uniforme	8 989	9 047	9 088	9 114	9 141	9 166	9 184	9 191	9 219
Total	2 271 753	2 287 210	2 304 109	2 318 711	2 335 132	2 349 841	2 361 279	2 368 108	2 383 333

Figure 3 : Voici une copie d'écran extraite de Webstats (Système de pilotage des applications de l'ABES)¹⁹

Ce tableau nous montre l'évolution en volume des notices d'autorités dans le Sudoc, de janvier à septembre 2011, *i.e.* une augmentation constante.

¹⁷ Répertoire d'autorité-matière encyclopédique et alphabétique unifié.

¹⁸ Cf. Guide méthodologique (site d'aide aux utilisateurs professionnels du réseau Sudoc) pour une description plus précise de la base d'appui : http://documentation.abes.fr/sudoc/manuels/interrogation/interrogation_professionnelle/appui.html > consulté le 26 juin 2012.

¹⁹ <http://webstats.abes.fr/webstats/> consulté le 26 juin 2012 ; plus d'informations sur le site Web de l'ABES : <http://www.abes.fr/Sudoc/Services-disponibles-autour-du-Sudoc/Webstats/%28langage%29/re-FR> > consulté le 26 juin 2012.

En l'espace de 9 mois, la base s'est enrichie d'un total de 111 580 notices, dont 96 631 notices personne physique et 9 970 notices nouvelles de type collectivité/congrès, ces deux types de notices étant les plus présentes dans la base, avec les autorités matière qui ne sont pas créées par nos réseaux.²⁰

2.2 Les services IdRef

2.2.1 *www.idref.fr* : le formulaire de consultation et d'édition

Ce sont les deux modes principaux d'utilisation d'IdRef : chercher la notice d'autorité à laquelle un professionnel de l'IST veut lier sans ambiguïté un document qu'il signale en le cataloguant, tout en laissant la possibilité à un partenaire de créer et modifier les notices (après login).

2.2.2 *Le mode consultation*



Figure 4 : IdRef consultation

Il permet d'effectuer des requêtes à partir des différents index (nom de personne, collectivité, matière, etc.), et d'afficher une liste de résultats ainsi que la notice recherchée lorsque la requête a été fructueuse. Des filtres sont possibles, ceux-ci correspondant au type de notice. Par exemple, ici, sur l'index « nom de personne », il est possible de filtrer la requête sur le critère « en vie ».

2.2.3 *Le mode édition*

Il exige une identification préalable, avec un login spécifique IdRef.

²⁰ Le référentiel rameau est néanmoins enrichi par les propositions de correspondants autorité du Sudoc. Ces propositions sont examinées par le Centre national Rameau qui décide ou non de la validité de la proposition d'enrichissement du référentiel. Si celle-ci est acceptée, la notice d'autorité écrasera la notice proposition dans le Sudoc. La création de notices « proposition rameau » est aujourd'hui possible avec le formulaire d'édition d'IdRef. Pour de plus amples informations sur Rameau, cf. : <<http://rameau.bnf.fr/index.htm>> consulté le 26 juin 2012.

Figure 5 : Notice IdRef

Il permet la création *ex nihilo* d'une notice, ou la correction de données dans une notice existante.

La logique d'IdRef est celle d'une application interopérable avec d'autres applications Web, mettant à disposition des données et des services réutilisables par des machines, dans la logique du Web de données.

2.3 Les Web services d'IdRef : des services ouverts

2.3.1 Web service de recherche

Il donne la possibilité de rechercher dans tous les index des autorités Sudoc (nom de personne, nom de collectivité, sujet, etc.). Ce service d'interrogation des autorités suit les spécifications Solr²¹.

Exemple : je veux toutes les notices d'autorité nom commun contenant le terme « France ». <http://www.idref.fr/Sru/Solr?q=subjectheading_t:france>²².

Il est possible de trier les résultats (ordre alphabétique, pertinence) et de les filtrer par facettes (types de notices ; langues ; vedette/subdivision ; en vie).

²¹ Disponibles à cette adresse : <<http://lucene.apache.org/solr/tutorial.html>> consulté le 26 juin 2012.

²² Consulté le 26 juin 2012.

2.3.2 Accès aux données d'une autorité par Web service, grâce l'identifiant PPN

Toute notice d'autorité Sudoc peut être récupérée en UNIMARC/MARCXML grâce à l'identifiant PPN²³, avec le libellé de la notice dans une zone 9XX non standard : <http://www.idref.fr/029950961.xml>²⁴.

La notice est également accessible en format RDF²⁵ : celle-ci contient à la fois les informations contenues dans la notice d'autorité, et la liste des notices bibliographiques liées avec un tri par rôle (auteur, directeur de thèse, traducteur, etc.) : <http://www.idref.fr/029950961.rdf>²⁶.

2.3.3 Micro Web service Biblio

Ce Micro Web service permet d'accéder directement aux références bibliographiques liées à la notice d'autorité, grâce encore au numéro PPN, identifiant Sudoc .

Exemple : www.idref.fr/services/biblio/029950961²⁷.

Cette URL renvoie de l'XML par défaut. Il est possible de demander grâce à la négociation de contenu à ce que cette URL générique renvoie du JSON²⁸.

2.3.4 Interaction entre les applications Web et les formulaires d'IdRef par Javascript

IdRef réceptionne les données que vous envoyez depuis votre application, et préremplit le formulaire de recherche. Il peut également préremplir le formulaire de création de notices si la requête a été infructueuse, et qu'une notice doit être créée.

IdRef donne ensuite la possibilité de lier des données Autorités Sudoc à votre application.

Un outil de test simulant cette opération interrogation/rapatriement de données est disponible à cette adresse : <http://www.idref.fr/autorites/aide/test.html>²⁹.

3. Rejoindre le réseau IdRef !

Pour toute question sur IdRef et son fonctionnement, savoir comment rejoindre son réseau : une adresse : le guichet d'assistance **ABES STP** <https://stp.abes.fr/>³⁰.

²³ *Pica Production Number.*

²⁴ Consulté le 26 juin 2012.

²⁵ *Resource Description Framework.*

²⁶ Consulté le 26 juin 2012.

²⁷ Consulté le 26 juin 2012.

²⁸ *JavaScript Object Notation.*

²⁹ Consulté le 26 juin 2012.

³⁰ Consulté le 26 juin 2012.



Figure 6 : ABES STP

Dans le domaine IdRef-, vous pourrez retrouver plusieurs FAQ³¹, et poser vos questions.

Vous pouvez en suivre l'actualité (entre autres !) sur le blog technique de l'ABES **Punktokomo** : <punktokomo.abes.fr>³².

Enfin, l'aide en ligne et la documentation pour les développeurs est disponible sur <www.idref.fr>³³ dans l'onglet « Aide ».

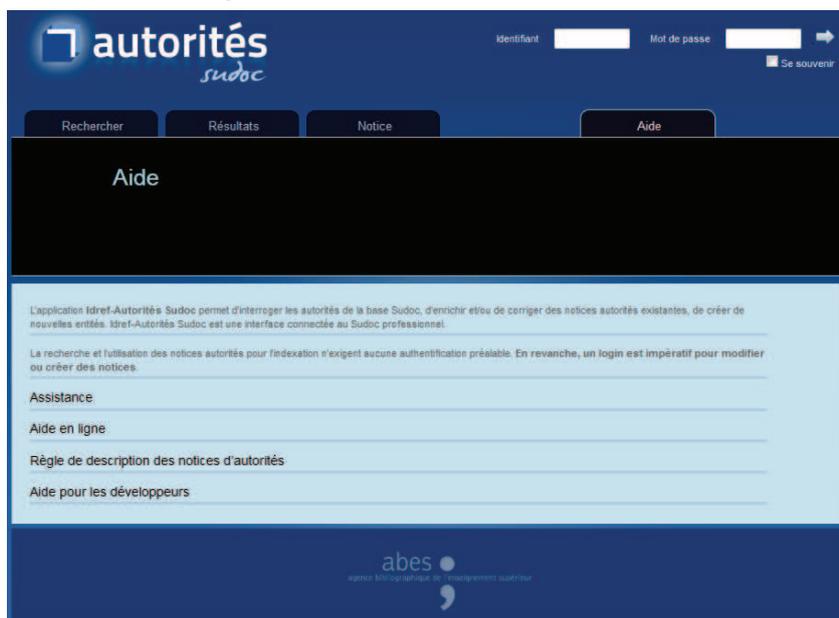


Figure 7 : Aide

³¹ Foire aux questions.

³² Consulté le 26 juin 2012.

³³ Consulté le 26 juin 2012.

Conclusion

Herbert Grüttemeier

Institut de l'information scientifique et technique (Inist), CNRS,
2, allée du parc de Brabois, CS 10310
Vandœuvre-lès-Nancy, France

Espèces menacées	<i>On est fauché!</i>	D'abord, il n'y a pas d'argent...
Dans ce scénario en IST, publications et données se confondent de plus en plus.	Pasteuresqueness	
Du « local » dans l'Europe ?	Sustainability of infrastructures	Avalanche de données
Interopérabilité	Futur challenge : réutilisation des données et Open Data	Clouds and crowds
		we have won the argument about open Access ! However difficult it seems, Library consortia must find ways to decompose the big deals.
		We incorporated green privileges. Linking to other information objects

D'abord, il n'y a pas d'argent...

On est fauché!

« Les données se confondent de plus en plus. »

« Il y a beaucoup de choses qui se ressemblent. »

Pasteurqueness

« Du « local » dans les infrastructures »

a different profession

« Interopérabilité »

Lobbying

« Futur challenge : réutilisation des données et Open Data »

Networking

« Clouds and crowds »

Libraries of the Future

« We have won the battle, but we have lost the war! »

« However difficult it seems, Library consortia must find ways to decompose big deals. »

« Nous aussi, on fait de gros deals. »

United we stand, divided we fall. (Aesop)

« On passé du temps à acheter des choses gratuites. »

nation objects

Partie 5

Références¹

Ouvrages IST au prisme de l'Europe

Bibliothèque publique d'information, *Regards croisés et perspectives, Bibliothèques publiques en Europe*, [actes du colloque organisé par la BPI, le British Council et le Goethe Institut, au British Council, Paris, 5-6 novembre 1998], Bibliothèque publique d'information (BPI) du Centre Pompidou, coll. « BPI en actes », 2000.

Blin Frédéric, *Enquête sur les activités de coopération internationale des bibliothèques des établissements d'enseignement supérieur*, Dossier de L'Association des Directeurs et des personnels de direction des Bibliothèques Universitaires et de la Documentation, 2006.

<https://www.sup.adc.education.fr/bib/Info/coop/Resultats_enquete.pdf> consulté le 7 octobre 2011.

Carbone Pierre, *Coûts, bénéfices et contraintes de la mutualisation des ressources électroniques : éléments de comparaison internationale et propositions*, Rapport à madame la ministre de l'Enseignement supérieur et de la Recherche, 2010-012, octobre 2010.

<http://media.enseignementsup-recherche.gouv.fr/file/2009/69/8/rapport_MRE_diffusion_159698.pdf> consulté le 7 octobre 2011.

Catanèse-Palanché Véronique, *Pour un tableau de bord commun aux grandes bibliothèques européennes : de la volonté de la BNF à une proposition de méthode*, Mémoire de recherche pour le Diplôme de conservateur de bibliothèque sous la dir. d'Aline Girard, Villeurbanne, Enssib, 2007.

<<http://enssibal.enssib.fr/bibliotheque/documents/dcb/catanese-dcb15.pdf>> consulté le 7 octobre 2011.

Chanier Thierry, *Archives ouvertes et publication scientifique - Comment mettre en place l'accès libre aux résultats de la recherche ?*, (1^{re} éd.), Paris, l'Harmattan, 2004.

Chauvigné Mathilde, *Le réseau de bibliothèques publiques de Mitte à Berlin : travail en coopération et corporate design*, Mémoire d'étude pour le Diplôme de conservateur de bibliothèque sous la dir. de Dominique Arot, Villeurbanne, Enssib, janvier 2006.

¹ Sélection réalisée à l'occasion des FréDoc 2011.

- Chevillotte Sylvie (dir.)**, *La formation à la maîtrise de l'information à l'heure européenne : problèmes et perspectives*, [actes des 3^{es} rencontres FORMIST, Villeurbanne, ENSSIB, 12-13 juin 2003], Villeurbanne, Presses de l'ENSSIB, 2004.
- Cornu Marie**, *Quel avenir pour les archives en Europe ? : enjeux juridiques et institutionnels*. Colloque international les 11 et 12 décembre 2008 à l'Université Paris-Sud 11, Faculté Jean Monnet à Sceaux / organisé par le Centre de recherche sur le droit du patrimoine culturel, CECOJI-CNRS-Université Paris-Sud 11 ; en collaboration avec la Direction des archives de France, ministère de la Culture et de la Communication [sous la direction de Marie Cornu et Jérôme Fromageau], Paris, l'Harmattan, 2010.
- De Castro Paola**, *Librarians of Babel: A Toolkit for Effective Communication*, Oxford/Cambridge, Chandos Publishing Ltd, coll. « Chandos information professional series », 2009.
- Demé Anne, Montagnon Sylvie, Russeil Gilles, Thompson Marie**, *Recensement critique des projets de numérisation en Europe : état des lieux de l'innovation et perspectives d'évolution*, Mémoire de recherche pour le Diplôme de conservateur de bibliothèque sous la dir. de Jean-Paul Metzger, Villeurbanne, Enssib, 2003.
<<http://www.enssib.fr/bibliotheque-numerique/document-978>> consulté le 7 octobre 2011.
- Dione Bernard, Savard Réjean (eds)**, *Managing Technologies and Automated Library Systems in Developing Countries: Open Source vs Commercial Options*, [Proceedings of the IFLA Pre-conference, August 15-16 2007], München, KG Saur Verlag, coll. "IFLA Publications Series 132", 2008.
- Erland Kolding Nielsen, Saur Klaus G., Ceynowa Klaus**, *Die Innovative Bibliothek : Elmar Mittler zum 65. Geburtstag [The Innovative Library To Elmar Mittler for his 65th Birthday]*, München, Saur, 2005.
- High Level Expert Group on Digital Libraries**, *Final Report. Digital Libraries: Recommendations and Challenges for the Future*, December 2009.
<http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/reports/hlg_final_report09.pdf> consulté le 7 octobre 2011.
- Joachim Schöpfel**, "Access to European Grey Literature", Petra Pejsova (ed.), *Grey Literature Repositories*, Radim Bačuvčík - VeRBuM, Zlín, 2010, p. 20-33.
<<http://nrgl.techlib.cz/images/Book.pdf>>
- Jougelet Suzanne, Vayssade Claire**, *Comparaison internationale de bibliothèques universitaires : étude de cas*, Rapport à madame la ministre de l'Enseignement supérieur et de la Recherche, n° 2009-0017, janvier 2010.
<http://media.enseignementsup-recherche.gouv.fr/file/2010/78/0/Rapport_etude_comparative_18_fevrier_2010_definitif_137780.pdf> consulté le 7 octobre 2011.

Jouguelet Suzanne, *Les Learning centres : un modèle international de bibliothèque intégrée à l'enseignement et à la recherche*, Rapport à madame la ministre de l'Enseignement supérieur et de la Recherche, n° 2009-022, décembre 2009.

<http://media.enseignementsup-recherche.gouv.fr/file/2009/33/6/Rapport_Learning_Centers_7-12_RV_131336.pdf> consulté le 7 octobre 2011.

Jouguelet Suzanne, *L'évaluation du numérique dans les bibliothèques françaises*, Rapport du groupe de travail *Évaluation du numérique en bibliothèque* au président de la Bibliothèque nationale de France, n°2009-000, 7 septembre 2009.

<<http://lesrapports.ladocumentationfrancaise.fr/BRP/104000149/0000.pdf>> consulté le 7 octobre 2011.

Jouguelet Suzanne, *L'évaluation de l'activité et des services documentaires français dans le cadre européen*, Rapport à madame la ministre de l'Enseignement supérieur et de la Recherche, n° 2008-003, juin 2008.

<http://media.enseignementsup-recherche.gouv.fr/file/Rapports/08/1/rapport_ind_10_juin_32081.pdf> consulté le 7 octobre 2011.

La coopération internationale dans le domaine des bibliothèques, Actes du séminaire, Sèvres – Centre international d'études pédagogiques, 8-9 mars 1995, organisé par la Fédération française de coopération entre bibliothèques en collaboration avec la Direction du Livre et de la Lecture, Paris, Fédération française de coopération entre bibliothèques, 1996.

Lemos de Annelise, *Une bibliothèque numérique européenne comme trait d'union*, Mémoire de recherche pour le Diplôme de conservateur de bibliothèque sous la dir. d'Yves Alix, Villeurbanne, Enssib, 2008.

Munde Gail, Marks Kenneth, *Making Good Libraries Better: An Introduction to Library Quality Improvement*, Oxford, Chandos Publishing Ltd, 2009.

Sanz Pascal, *Guide de la coopération entre bibliothèques*, Paris, Electre/Cercle de La Librairie, coll. « Bibliothèques », 2009.

Schöpfel Joachim (dir.), *Les Cahiers du numérique*, 5(2-Communication scientifique. Les nouveaux enjeux), 2009.

<<http://www.cairn.info/revue-les-cahiers-du-numerique-2009-2.htm>> consulté le 7 octobre 2011.

Tessier Marc, *Rapport sur la numérisation du patrimoine écrit*, au ministre de la Culture et de la Communication, 12 janvier 2010.

<<http://www.culture.gouv.fr/mcc/content/download/3520/23115/file/Rapport%20sur%20la%20numerisation%20du%20patrimoine%20ecrit.pdf>> consulté le 7 octobre 2011.

UNESCO, European Commission and the German Commission of UNESCO, *Open Access, Opportunities and Challenges: A Handbook*, Luxembourg, Publications Office, 2008.

http://ec.europa.eu/research/science-society/document_library/pdf_06/open-access-handbook_en.pdf consulté le 7 octobre 2011.

Articles IST au prisme de l'Europe

Aguiton Christophe, Cardon Dominique, « Web participatif et innovation collective », *HERMÈS*, 50(Communiquer – Innover : Réseaux, dispositifs, territoires), 2010, p. 77-82.

http://documents.irevues.inist.fr/bitstream/handle/2042/24155/HERMES_2008_50_77.pdf?sequence=1 consulté le 7 octobre 2011.

Arnaud Michel, « Propriété intellectuelle et accès public au savoir en ligne : Fractures dans la société de la connaissance », *HERMÈS*, 45, 2006, p. 61-68.

http://documents.irevues.inist.fr/bitstream/handle/2042/24035/HERMES_2006_45_61.pdf?sequence=1 consulté le 7 octobre 2011.

Banat-Berger Françoise, « La gestion de l'information et des archives électroniques en Europe », *Documentaliste*, 46(1), 2009, p. 12-14.

Barbe Lionel, « Wikipedia, un trouble-fête de l'édition scientifique », *HERMÈS*, 57(Sciences.com. Libre accès et science ouverte), 2010, p. 69-74.

Bargheer Margo, Schmidt Birgit, "Publishing services in an Open Access environment", *Information Services Use*, 28(28), p. 133–139, 2008, Göttingen, Göttingen University Press

Battisti Michèle, *Vers un fair use en Europe ? La France, village gaulois ou roseau ?*, compte-rendu, Colloque international, Association Française pour la Protection Internationale du Droit d'Auteur (AFPIDA), Les exceptions au droit d'auteur - État des lieux et perspectives dans l'Union européenne, Centre d'Études et de Recherche en Droit de l'Immatériel (Cerdi, Universités Paris 1 - Panthéon-Sorbonne & Paris-Sud 11), 21 mars 2011.

<http://www.adbs.fr/vers-un-fair-use-en-europe-la-france-village-gaulois-ou-roseau--100750.htm> consulté le 7 octobre 2011.

Battisti Michèle, *Exploiter les œuvres orphelines. Quelle position en Europe ?* compte-rendu, Séminaire Droit de l'information, EBLIDA, Paris, 15 avril 2008.

<http://adbs.test.oceanet.eu/exploiter-les-oeuvres-orphelines-quelle-position-en-europe--39623.htm> consulté le 7 octobre 2011.

Beaudry Guylaine, « La communication scientifique directe : un nouveau champ éditorial », *HERMÈS*, 57(Sciences.com. Libre accès et science ouverte), 2010, p. 51-57.

- Benabou Valérie-Laure**, « Les publications scientifiques : faut-il choisir entre libre accès et libre recherche ? », *HERMÈS*, 57(Sciences.com. Libre accès et science ouverte), 2010, p. 95-106.
- Bourcier Danièle**, « *Science Commons* : nouvelles règles, nouvelles pratiques », *HERMÈS*, 57(Sciences.com. Libre accès et science ouverte), 2010, p. 213-214.
- Boukacem-Zeghmouri Chérifa, Schöpfel Joachim**, « L'impact de l'édition scientifique en Europe. *Academic Publishing in Europe 2009: The Impact of Publishing* », *Bulletin des Bibliothèques en France*, 54(3), 2009, p. 81-82.
<http://archivesic.ccsd.cnrs.fr/sic_00333736/en/> consulté le 7 octobre 2011.
- Chartron Guylaine**, « Scénarios prospectifs pour l'édition scientifique », *HERMÈS*, 57(Sciences.com. Libre accès et science ouverte), 2010, p. 123-129.
- Chevillotte Sylvie**, « Formations en sciences de l'information et des bibliothèques : la coopération européenne : auberge espagnole ou châteaux en Espagne ? », *Bulletin des Bibliothèques de France*, 53(1), 2008, p. 34.
- Chourrot Olivier, Meller-Liron Élisabeth**, « Bibliothèques et insertion professionnelle », *BBF*, 2009, n° 2, p. 6-12.
<<http://bbf.enssib.fr/>> consulté le 06 octobre 2011.
- Cornu Marie**, « Création scientifique et statut d'auteur », *HERMÈS*, 57(Sciences.com. Libre accès et science ouverte), 2010, p. 85-93.
- Doray Pierre, Goldenberg Anne, Proulx Serge**, « Du laboratoire à la communauté : organiser l'espace pour innover », *HERMÈS*, 50(Communiquer – Innover : Réseaux, dispositifs, territoires), 2010, p. 131-138.
<http://documents.irevues.inist.fr/bitstream/handle/2042/24162/HERMES_2008_50_131.pdf?sequence=1> consulté le 7 octobre 2011.
- Doury-Bonnet Juliette**, « La coopération européenne des bibliothèques », *Bulletin des Bibliothèques de France*, 53(4), 2008, p. 74.
- European Commission Information Society and Media**, *i2010: Digital Libraries*, Luxembourg, Publications Office.
- Farchy Joëlle, Froissart Pascal**, « Le marché de l'édition scientifique, entre accès "propriétaire" et accès "libre" », *HERMÈS*, 57(Sciences.com. Libre accès et science ouverte), 2010, p. 137-150.
- Häkli Esko**, *Innovation through Co-operation: The History of LIBER 1971-2009*, Copenhagen, Museum Tusulanum Press, June 2011.
- Heaton Lorna, Millerand Florence, Proulx Serge**, « *Tela Botanica* : une fertilisation croisée des amateurs et des experts », *HERMÈS*, 57(Sciences.com. Libre accès et science ouverte), 2010, p. 61-68.

- Helie Thomas**, « Cultiver l'Europe. Éléments pour une approche localisée de l'européanisation des politiques culturelles », *Politique européenne*, 12, hiver 2004, p. 66-83.
<<http://www.cairn.info/revue-politique-europeenne-2004-1-page-66.htm>> consulté le 7 octobre 2011.
- Heurtematte Véronique**, « À la recherche de l'e-chercheur », *Livres-Hebdo*, n° 743, 29 août 2008.
- IFLA Journal*, SAGE Publications, dernier numéro paru vol. 37, n°2, 2011.
<<http://www.ifla.org/publications/ifla-journal>> consulté le 7 octobre 2011.
- « La cause des orphelines », *Livres Hebdo*, n° 795, 30 octobre 2009, p. 13-15.
- Lamouroux Mireille, Battisti Michèle**, « 73^e Congrès de l'IFLA. Bibliothèques du futur : progrès, développement et partenariats », *Documentaliste*, 44(4-5), 2007, p. 327-336.
<<http://www.cairn.info/revue-documentaliste-sciences-de-l-information-2007-4-page-327.htm>> consulté le 7 octobre 2011.
- Lang Bernard**, « Des cordonniers mal chaussés ou les informaticiens face au libre accès », *HERMÈS*, 57(Sciences.com. Libre accès et science ouverte), 2010, p. 81-84.
- Leblanc-Barriac Charline, Rasse Paul**, « Les enseignants-chercheurs face aux mutations de leur environnement documentaire », *HERMÈS*, 57(Sciences.com. Libre accès et science ouverte), 2010, p. 179-184.
- « Le projet Michael : plate-forme européenne de ressources culturelles », *Documentaliste*, 42(2), 2005, p. 95-96.
<<http://www.michael-culture.org/fr/about/european-service>> consulté le 7 octobre 2011.
- Meyer Morgan**, « Les courtiers du savoir, nouveaux intermédiaires de la science », *HERMÈS*, 57(Sciences.com. Libre accès et science ouverte), 2010, p. 165-171.
- Meyriat Jean**, « Coopérations européennes : pour la promotion des professionnels de l'I&D et de leurs compétences », *Documentaliste*, 44(2), 2007, p. 171-174.
- Mounier Pierre**, « Le libre accès : entre idéal et nécessité », *HERMÈS*, 5(Sciences.com. Libre accès et science ouverte), 2010, p. 23-30.
- Nowicki Joanna**, « Communication interculturelle et construction identitaire Européenne », *HERMÈS*, 41, 2006, p. 131-138.
<http://documents.irevues.inist.fr/bitstream/handle/2042/8963/HERMES_2005_41_131.pdf?sequence=1> consulté le 7 octobre 2011.
- « Numérisation mode d'emploi », *Livres Hebdo*, 1^{er} février 2008.
<<http://www.jouve.fr/section-actualites/presse/jouve-dans-les-medias/247-jouve-numerisation-fonds-editorial-editeurs.html>> conférence 2011.

ou

<http://www.salondulivreparis.com/site/FR/Programme/Conferences/_Animations/Zoom_Conference_Numerisation_mode_demploi_I2047_Zoom-ce70079222075bfb3d7bd445ff2eccac.htm> consulté le 7 octobre 2011.

Ott Florence, « Une logique européenne et un pragmatisme nord-américain », *Documentaliste*, 46(1), 2009, p. 40.

Perriault Jacques, « La norme comme instrument d'accès au savoir en ligne : fractures dans la société de la connaissance », *HERMÈS*, 45, 2006, p. 109-115.

<http://documents.irevues.inist.fr/bitstream/handle/2042/24037/HERMES_2006_45_77.pdf?sequence=1> consulté le 7 octobre 2011.

Reding Viviane, "The role of libraries in the information society", *CENL Conference*, Luxembourg, 29 September 2005.

Roche I., François C., Besagni D., « Détection de techniques prometteuses à partir de méthodes bibliométriques », *CIDE, 10^e Colloque International sur le Document Electronique*, Nancy, 4-7 juillet 2007, p. 193-197

Schöpfel Joachim, « L'édition scientifique en Europe. Academic Publishing in Europe 2010: Researchers, Librarians & Publishers and Their Roles in Scholarly Communication », *Bulletin des Bibliothèques en France*, 55(4), 2010, p. 77-78.

Shorley Deborah, "Past its Shelve by Date? United Kingdom Research Reserve (UKRR): A Twenty-First-Century Strategy to Protect Our Research Information for the Future", *New Review of Academic Librarianship*, 2008, 14(1), p. 115-120.

Shorley Deborah, "United Kingdom Research Reserve (UKRR) – tomorrow's solution to yesterday's problems", *LINK*, 6 April 2009, p. 3.

Simone Jérôme, « L'Europe face au défi de la conservation des documents numériques à long terme », *Documentaliste*, 45(1), 2008, p. 14-21.

<<http://www.cairn.info/revue-documentaliste-sciences-de-l-information-2008-1-page-14.htm>> consulté le 7 octobre 2011.

Stockinger Peter, « L'adaptation et la republication de ressources audiovisuelles numériques », *HERMÈS*, 56, 2010, p. 63-70.

Ulla Hakala, Ulla Nygrén, "Customer satisfaction and the strategic role of university libraries", *International Journal of Consumer Studies*, 34(2), 2010, p. 204-211.

Vinck Dominique, « Les transformations des sciences en régime numérique », *HERMÈS*, 57(Sciences.com. Libre accès et science ouverte), 2010, p. 35-41.

Zimmermann Christian, « La dissémination de la recherche en sciences économiques : les "cahiers de recherche" », *HERMÈS*, 57(Sciences.com. Libre accès et science ouverte), 2010, p. 43-49.

DRIVER

Dale Peters, Norbert Lossau, "DRIVER: building a sustainable infrastructure for global repositories", *The Electronic Library*, 29(2), 2011, p. 249-260.

eContentplus

B.F., « L'Union européenne lance le programme eContentplus », *Documentaliste*, 42(2), 2005, p. 99-100.

Sangwin C.J., "Have you seen this?... JEM - Joining Educational Mathematics: an eContentPlus Thematic Network", *MSOR CONNECTIONS*, 8(4), 2008, p. 39-41.

EuDML

Borbinha J., Bouche T., Nowinski A., Sojka P., "Project EuDML - A first year demonstration", *Calculus 2011 and 10th International Conference on Mathematical Knowledge Management MKM 2011 18th Symposium on the Integration of Symbolic Computation and Mechanized Reasoning (2011 07 18 - 2011 07 23 Bertinoro, ITA)*, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, v6824, 2011 08 11, p. 281-284.

Bouche Thierry, José Borbinha, Aleksander Nowiński, and Petr Sojka, "Project EuDML—A First Year Demonstration", James H. Davenport, William M. Farmer, Josef Urban, Florian Rabe (eds), *Intelligent Computer Mathematics proceedings of the CICM 2011 conferences*, Bertinoro, Italy, July 18-23, 2011 (*Lecture Notes in Computer Science*, Vol. 6824/2011, p. 281-284.

Bouche Thierry, « Démarrage du projet EuDML. La bibliothèque numérique de mathématiques », *SMF Gazette*, 124, avril 2010.

<<http://hal.archives-ouvertes.fr/hal-00559055/en/>> consulté le 7 octobre 2011.

Bouche Thierry, "Introducing EuDML—The European Digital Mathematics Library", *EMS Newsletter*, 76, 2010, p. 11-16.

<<http://hal.archives-ouvertes.fr/hal-00559060/en/>>.

Bouche Thierry, Sylwestrzak Wojtek, Borbinha José, Nowiński Aleksander, Sojka Petr, "Towards the European Mathematics Digital Library", *DML 2010, 3rd Workshop on Digital Mathematics Library*, Paris, 2010.

<<http://hal.archives-ouvertes.fr/hal-00559048/en/>> consulté le 7 octobre 2011.

Bouche Thierry, « Bibliothèques numériques de mathématiques en France et dans le monde », *Images des mathématiques*, Paris, CNRS, oct. 2009.

Bouche Thierry, "Digital Mathematics Libraries: The Good, the Bad, the Ugly", S. Autexier et al. (eds.), *AISC/Calculus/MKM 2008 proceedings*, (LNAI 5144, p. 331-332, Springer-Verlag Berlin Heidelberg 2008.

<<http://www.springerlink.com/content/v44230u011x722n7/>> consulté le 7 octobre 2011.

EUROPEANA

Alix Yves, « De la bibliothèque numérique européenne à Europeana », BULLETIN-BIBLIOTHEQUES DE FRANCE, 53(1), 2008, p. 78.

Bjørner Susanne, "Europeana and Digitization – The Collaboration is Only Beginning", Information Today, Inc. *Newsbreak*, September 10, 2009.

<<http://newsbreaks.infotoday.com/NewsBreaks/Europeana-and-Digitization-The-Collaboration-Is-Only-Beginning-56079.asp>> consulté le 7 octobre 2011.

Bjørner Susanne, "Thinking about Culture and Language: The European Library and Europeana", *Searcher*, 17(6), June 2009, p. 33-40.

Bouvier-Ajam Laurent, « Europeana : étude sur les usages et les attentes relatifs à l'interface de consultation de la future Bibliothèque numérique européenne » rapport final BNF, 2007.

<<http://www.enssib.fr/bibliotheque-numerique/document-1179>> consulté le 7 octobre 2011.

Europeana Foundation, *Charte du domaine public d'Europeana*, The Europeana Foundation, 2010.

<<http://www.enssib.fr/bibliotheque-numerique/document-48316>> consulté le 7 octobre 2011.

Reding Viviane, « EUROPEANA : un voyage à travers la culture européenne », Bruxelles, SPEECH/08/633, 20 novembre 2008.

<http://ec.europa.eu/information_society/newsroom/cf/itemdetail.cfm?item_id=4514> consulté le 7 octobre 2011.

Reding Viviane, « La bibliothèque numérique européenne : du rêve à la réalité », *Forum d'Avignon : Culture, facteur de croissance*, SPEECH/08/621, 17 novembre 2008.

<http://ec.europa.eu/information_society/newsroom/cf/itemdetail.cfm?item_id=4514> consulté le 7 octobre 2011.

Santantonios Laurence, « L'Europe de la culture en 1 clic », *Livres Hebdo*, n° 754, 14 novembre 2008.

LIBER

Commission européenne, *Role and Impact of Professional and Scientific Societies in ICT Research, Education and Innovation (SMART 2009/0061)*, rapport, 4 août 2011.

<http://ec.europa.eu/information_society/newsroom/infocentre/detail.cfm?id=7222&utm_campaign=isp&utm_medium=rss&utm_source=newsroom&utm_content=typ-e-news> consulté le 7 octobre 2011.

Liber Quarterly, The Journal of European Research Libraries, dernier numéro paru vol; 20, n°3/4, 2011.

Santantonios Laurence, « LIBER passe à l'offensive », *Livres Hebdo*, n° 785, 21 août 2009, p. 87.

Santantonios Laurence, « Il faut plus de Français à Liber », *Livres Hebdo*, 6 mars 2009.

LIBQUAL

Greenwood J.T., Watson A.P., Dennis M., "Ten Years of LibQual: A Study of Qualitative and Quantitative Survey Results at the University of Mississippi 2001-2010", *The Journal of Academic Librarianship*, 37(4), 2011, p. 312-318.

Morales Miguel, Ladhari Riadh, Reynoso Javier, Toro Rosario, Sepulveda Cesar, "An independent assessment of the unidimensionality, reliability, validity and factor structure of the LibQUAL+™ scale", *The Service Industries Journal*, p. 1-21.

Morales Miguel, Ladhari Riadh, Reynoso Javier, Toro Rosario, Sepulveda Cesar, "Factor structure and psychometric properties of a Spanish version of LibQUAL+™", *Performance Measurement and Metrics*, 12(1), 2011, p. 23-37.

Russell P., "Measuring up: the experience of LibQUAL+™ at ITT Dublin library", *SCONUL FOCUS*, 49, 2010, p. 47-51.

MINERVA

Guenther Rebecca MLS, Myrick Leslie PhD, "Archiving Web Sites for Preservation and Access: MODS, METS and MINERVA", *Journal of Archival Organization*, 4(1-2), April 2007, p. 141-166.

Publication gouvernementale internationale, *Coordinating digitisation in europe*, Projet européen MINERVA, Roma, Minerva eC Project, Ministero per i Beni e le Attività Culturali, 2008.

Taylor Robin, "MINERVA: Archiving Born-Digital Material at the Library of Congress", *Slavic & East European Information Resources*, 5(1-2), May 2004, p. 157-162.

FOBID

Besek June M., LeFurgy William, Muir Adrienne, Atkinson Benedict A., Carroll Emma F., Coates Jessica M., Fitzgerald Brian F., Mossink Wilma, *International Study on the Impact of Copyright Law on Digital Preservation*, report of The Library of Congress National Digital Information, Infrastructure and Preservation Program, The Joint Information Systems Committee, The Open Access to Knowledge (OAK) Law Project, The SURFfoundation, July 2008.

<<http://eprints.qut.edu.au/14035/1/14035.pdf>> consulté le 7 octobre 2011.

Koren Marian, "Working for a successful IFLA Presidential Meeting 2011: IFLA and FOBID, the international Dutch connection", *IFLA Journal*, 37(2), 2011, p. 165.

Koren Marian, "Highlights of IFLA Congress 2010 in Gothenburg", *New Library World*, 112(1-2), 2011, p. 94-96.

IVOA

Chilingarian Igor, Bonnarel Francois, Louys Mireille, Zolotukhin Ivan, Royer Frederic, Jegouzo Isabelle, Le Sidaner Pierre, Fernique Pierre, Boch Thomas, "3D Spectroscopy in the Virtual Observatory: Current Status", *arXiv:0711.0412v1*.

<http://arxiv.org/PS_cache/arxiv/pdf/0711/0711.0412v1.pdf> consulté le 7 octobre 2011.

Smillie Jonathan G., collaboration, for the MACHO, "IVOA Compliant Services for the MACHO Data Archive", 2009.

<http://arxiv.org/PS_cache/arxiv/pdf/0910/0910.5041v1.pdf> consulté le 7 octobre 2011.

Todd King, Jan Merka, Raymond Walker, Steven Joy, Tom Narock, "The architecture of a multi-tiered virtual observatory", *Earth Science Informatics*, 2008, 1(1), p. 21-28.

<<http://www.springerlink.com/gate3.inist.fr/content/t475202m07847536/fulltext.pdf>> consulté le 7 octobre 2011.

PEER

Laurent Romary, « Communication scientifique : pour le meilleur et le PEER », *HERMÈS*, 57(Sciences.com. Libre accès et science ouverte), 2010, p. 131-136.

SPARC

Johnson Gareth, "In the Kingdom of the Blind: Successfully Implementing Institutional Repositories in the United Kingdom and the SHERPA Partnership Experience", *New Review of Academic Librarianship*, 13(1-2), 2008, p. 13-33.

- Joseph Heather**, "Sparc: Creating Innovative Models and Environments for Scholarly Research and Communication", *Educational Futures Rethinking Theory and Practice*, 1(46-Digitisation Perspectives), Part V, 2011, p. 245-249.
- Jubb Michael**, "Supporting the Research Base: The Research Information Network and Scholarly Communications in the United Kingdom", *New Review of Academic Librarianship*, 13(1-2), 2008, p. 35-50.
- Peek R.**, "Focus on Publishing SPARC Europe Seal Clarifies Legal Rights on OA", *INFORMATION TODAY*, 25(6), 2008, p. 19.

Partie 6

Glossaire¹

(non exhaustif)

- AIBM** : Association internationale des bibliothèques, archives et centres de documentation musicaux (1951-...) ou **IAML** : *International Association of Music Libraries, Archives and Documentation Centres*
- APA** : *Alliance for Permanent Access* (Alliance pour un accès permanent)
- CENL** : *Conference of European National Librarians* (1987)
- CERL** : *Consortium of European Research Libraries* (Consortium européen des bibliothèques de recherche)
- COAR** : *Confederation of Open Access Repositories*
- DARIAH** : *Digital Research Infrastructure for the Arts and Humanities*
- DBF** : *Development and Verification of a Bibliometric model for the Identification of Frontier Research* (2009-2012)
- DigiCULT** : *Technology Challenge for Digital Culture*
- Driver** : *Digital Repository Infrastructure Vision for European Research* (Vision des infrastructures de dépôt numérique pour la recherche européenne)
- EBLIDA** : *European Bureau of Library, Information and Documentation Associations* (Bureau européen des associations de bibliothèques, de l'information et de la documentation)
- eContentPlus** : a pour but de rendre le contenu numérique plus accessible, plus utilisable et plus exploitable en facilitant la création et la diffusion d'informations – dans des domaines d'intérêt général – au niveau de la Communauté (2005-...)
- EDL** : *European Digital Library* (2007)
- EDLnet** : *European Digital Library network* (2007)
- ERC** : *European Research Council*
- Euclid** : *European Association for Library and Information*
- EuDML** : *European Digital Mathematics Library* (Bibliothèque numérique européenne des ressources mathématiques)
- EuroVO** : *European Virtual Observatory*
- FOBID** : *Federatie organisaties bibliotheek-, informatie-, dokumentatiewezen* (the Netherlands Library Forum)
- IAMIC** : *International Association of Music Information Centres*
- IASA** : *International Association of Sound Archives*

¹ Sélection réalisée à l'occasion des FréDoc 2011.

- IFLA** : *International Federation of Library Associations and Institutions* (Fédération internationale des associations de bibliothécaires et d'institutions)
- IMC** : *International Music Council*, organisation non-gouvernementale dépendant de l'Unesco
- IRO** : *International Repositories Organisation* (Organisation internationale de dépôt)
- IVOA** : *International Virtual Observatory Alliance* (Alliance des observatoires internationaux virtuels)
- JISQ** : *Joint Information Systems Committee*
- LIBER** : *Association of European Research Libraries* (Ligue des bibliothèques européennes de recherche)
- LibQUAL+** : *Charting Library Service Quality...* (Charte du service qualité des bibliothèques). Outil d'évaluation de la satisfaction des usagers.
- MICHAEL** : *Multilingual Inventory of Cultural Heritage in Europe*
- MINERVA** : *Ministerial Network for Valorising Digitisation Activities* (Réseau ministériel pour la valorisation des activités de numérisation)
- NEEO** : *Network of European Economists Online*
- OAI DAPHNE** : *Open Archives Initiative* (Données en Archéologie, Préhistoire et Histoire sur le Net)
- OAPEN** : *Open Access Publishing in European Networks* (Publications des archives ouvertes du réseau européen)
- OpenAIRE** : *Open Access Infrastructure for Research in Europe* (accès aux résultats de la recherche scientifique de l'UE)
- PEER** : *Publishing and the Ecology of European Research* (Édition et écologie européenne)
- PEERC** : *Production Engineering Education and Research Center*
- Prom-Tech** : *Identification and Assessment of Promising Emerging Technological Fields in Europe*, projet européen de détection des technologies émergentes par des techniques bibliométriques (2005-2007)
- QUALIBIB** : guide pratique pour l'amélioration de la qualité de l'accueil et des services dans les bibliothèques et centres de documentation (ISO 11620)
- SOAP** : *Study of Open Access Publishing*
- SPARC** : *Scholarly Publishing and Academic Resources Coalition* (Coalition des publications savantes et des ressources académiques)
- TEL** : *The European Library*
- Te-Learn** : *European Research on Technology-Enhanced Learning* (apprentissage assisté par les technologies)

Comité d'organisation des FréDoc 2011

Christine Cazenave	Les Afriques dans le monde UMR 5115 CNRS Sciences Po Bordeaux 11, allée Ausone 33067 Pessac Cedex
Élise Douat	Formation permanente CNRS - Délégation Aquitaine Esplanade des Arts et Métiers BP 105 33402 Talence Cedex
Isabelle Esqurial	PACEA - De la préhistoire à l'actuel : culture, environnement et anthropologie, UMR 5199 CNRS B18, Université de Bordeaux I, avenue des Facultés 33405 Talence Cedex
Françoise Girard	Laboratoire Hétéroéléments et Coordination UMR 7653 CNRS École Polytechnique - DCPH 91128 Palaiseau Cedex
Pascaline Hoël	INIST-CNRS 2, Allée du Parc de Brabois CS 10310 54519 Vandœuvre-lès-Nancy
Véronique Humbert	Archéologie des Sociétés Méditerranéennes – UMR 5140 CNRS 390 Avenue de Pérols 34970 Latte
Joana Janik	MI2S-UMS 3042 CNRS Maison Jean Kuntzmann 110, rue de la Chimie Domaine Universitaire – BP 53 38041 Grenoble Cedex 9
Armelle Jézéquel	Centre Émile Durkheim UMR 5116 CNRS Sciences Po Bordeaux 11 allée Ausone 33607 Pessac Cedex
Elizabeth Kneller	Bibliothèque mathématique Jacques Hadamard UMS 1786 CNRS Université Paris Sud, Bât. 425 Rue du Doyen Georges Poitou 91405 Orsay Cedex
Cyril Mauvillain	Bibliothèque de Recherche Mathématiques et Informatique *(BMI)* Université Bordeaux 1 351, cours de la Libération 33405 Talence Cedex
Geneviève Peyres	Maison de l'Orient et de la Méditerranée CNRS Université Lyon 2 5-7 rue Raulin 69365 Lyon Cedex 07

**Ainsi que les
membres du réseau
ARPIST**

Caroline Abéla, Bernadette Bergeret, Rodolphe Defiolle,
Gilles Dubus, Stéphane Toulon

**Et le service de la
formation permanente
Aquitaine**

Marie-Noëlle Besson, Philippe Cursol, Karima Benchabane,
Martine Rigal

ISBN : 978-2-7466-6304-6

EAN : 9782746663046