

From Tailored Databases to Wikis: Using Emerging Technologies to Work Together More Efficiently

***Amanda Regolini, Frédéric Berger,
Emmanuelle Jannès-Ober, and Luuk Dorren
Cemagref, France***

**amanda.regolini@cemagref.fr; frederic.berger@cemagref.fr
emmanuelle.jannes-ober@cemagref.fr; luuk.dorren@cemagref.fr**

Abstract

This article describes the emergence of collaborative work between researchers and information professionals through the presentation of two tools developed within Cemagref. The first tool is a traditional database and the second a Wiki; they were both implemented by information professionals with the contribution of information technology (IT) experts.

The role of Cemagref information professionals is to provide researchers with expertise on information technology. The first task of the ten Cemagref documentation departments is to supply our researchers with suitable data.

This paper develops this experience through the description of two existing systems. Not only can researchers now retrieve information, but they are also able to feed the databases with their own information.

Keywords: scientific and technical information, scientific publishing, monitoring, database, wiki.

Introduction

Cemagref is a French public research institute that targets results directly useable in land and water management. Cemagref aims at producing new knowledge and technical innovation for managers, decision-makers, and companies by responding to specific questions raised by society concerning resource management, land use, and development. In Grenoble, France, three teams devote most of their research to mountainous regions.

Ten documentation departments manage scientific and technical information. The documentary organisation is based on the body's geographical distribution; each site has a documentation centre. Twenty-two staff handle documents, with the primary aim of meeting the needs of the establishment's researchers in terms of information.

Material published as part of this publication, either on-line or in print, is copyrighted by the Informing Science Institute. Permission to make digital or paper copy of part or all of these works for personal or classroom use is granted without fee provided that the copies are not made or distributed for profit or commercial advantage AND that copies 1) bear this notice in full and 2) give the full citation on the first page. It is permissible to abstract these works so long as credit is given. To copy in all other cases or to republish or to post on a server or to redistribute to lists requires specific permission and payment of a fee. Contact Publisher@InformingScience.org to request redistribution permission.

The documentary network is coordinated by the Scientific and Technical Information (STI) Department, at General Management level (Jannes-Ober, 2005). The STI Department acts as the driving force in encouraging scientific research-related exchanges and supporting our strategy within the institution as well as with outside partners. To achieve such goals, the STI Depart-

ment offers specific tools and methods allowing users to access useful information, assist the production and dissemination of relevant knowledge and expertise, develop scientific production, and further the activities of external and internal specialist networks.

This work was carried out at Cemagref Grenoble documentation centre, together with the Mountain Ecosystems research team, i.e. 27 researchers and technical engineers and six PhD students. The research unit focused on the dynamics of mountain ecosystems. It aimed at providing management tools for such environments. Its research contributed to improving what is known about the ecosystem (the functioning of plant communities, process modelling, the validation of ecological concepts, and theories on mountain systems as well as spatial modelling) and to developing operational methods and tools for ecological management and the engineering of mountain ecosystems.

The research within the Mountain Ecosystems team is multidisciplinary, experimental, and technical. This is one of Cemagref's assets. Scientists contribute to enhancing the knowledge that must be operational for ecological management and the engineering of mountain ecosystems. They also contribute to the scientific knowledge of mountain ecosystems.

Scientific knowledge is complex. Information professionals' first task is to meet the needs of individual researchers. However, can they leave the user alone to decide on their needs? Information professionals have to support scientific activity (Turner, 1996). This is a real challenge for information professionals who have to cope with the information researchers require. Information must be multidisciplinary, scientific, and technical.

The Cemagref documentation centre's main objective is to help researchers save time when seeking scientific and technical data. Information professionals offer their expertise to solve information and communication issues.

According to W.A. Turner (1996), before taking action there may be cooperation; yet, before deciding to collaborate, managing on-going social exchanges successfully is required. He argues that databases should be considered as tools for this type of management. The author draws the conclusion that modern information professionals should work in this direction.

At Cemagref, strategic plans from 1999 to date (Cemagref, 1999, 2004) have expressed the need to strengthen the links between researchers and information professionals. In the first plan, information professionals were asked to assist researchers with their work. In the current strategic plan, they act as actual support of scientific and technical management.

Over the last five years, different systems to help researchers have been explored at our centre: from a "traditional database" tailored to meet Cemagref's needs to a new open source tool from Web 2.0. The contexts, objectives, methodologies, and results are developed below for those two examples. This paper aims at describing how information management problems have been solved through cooperative approaches.

A Database to Help People Publish

Scientific journals are subject to ranking, particularly when scientists are evaluated. Scientific articles count in terms of quality and quantity. The quality of an article is associated with the impact of the journal used as support (Guédon, 2001). Choosing the journal properly is capital. To maximise impact, researchers must provide high quality research and publish in the highest impact journals. Cemagref has always been well known at European and international levels for such scientific subjects as water resources, uses, and risks. However, for some others, like mountain ecosystems, a scientific network had to be created. In fact, in the late 1990s the Mountain Ecosystems unit needed to make an effort to publish. Little publishing in scientific journals could be explained by the particular evolution of the institution. In fact, Cemagref was born from the

reunion of two technical institutions. The engineering and technical employees of Mountain Ecosystems unit had to share a publishing culture. That was even more difficult for PhD students since they only had 3 years to define their field of research. PhD students under contract usually spend three years at Cemagref. They have to publish at least two articles before defending their thesis. If publishing was difficult for researchers who worked in one area, it was even more so for Cemagref researchers dealing with multidisciplinary sciences.

In the four-year Cemagref – State contract (Ministère de l’Agriculture et de la Pêche, Ministère de la recherche, Cemagref, 2001), this aspect was pointed out and the increase in the number of papers published in good quality journals was one of the targets for the end of 2004.

Information professionals usually identify the core papers in each specialty to meet the needs of the local research scientists. They are often asked about the best journals in one particular field. Furthermore, when looking for a journal to publish in, researchers often ask some of the following questions: What is the impact factor? How many times a year is the journal published? Who is on the editorial board?

Other information centres had faced the same questions before Cemagref. In France, the Agricultural Research Centre for International Development (CIRAD) answered this request in 1997 by publishing a selective guide to scientific journals in the fields of agricultural and rural development (Boussou-Pélissier, 1997).

In 2003, a working group was created within the research unit to help PhDs, young researchers, and engineers to publish. It consisted of scientists, engineers, and PhD students.

For main journals, the process of peer reviewing can be very long; it may take more than one year before an article is eventually accepted for publication after all requested corrections have been made (Bach, 2007). Researchers wanted to share their experiences on how to be published successfully. Moreover, they aimed to collaborate with PhD students and young researchers to help them find the best scientific and technical journals to publish in.

Twice a month, they would organise meetings in which their projects were presented and reviewed. Those presentations were very appreciated at the beginning, but were then replaced by personal experiences and success stories in academic publishing.

At the same time, researchers wanted to set up a list of relevant journals in their field. Yet, instead of focusing on a core of very specialised sources, they wanted this list to include a larger number of sources. A multidisciplinary working team was therefore created for this specific project. It included researchers of the Mountain Ecosystems team, IT experts from the headquarters, and information professionals from the Grenoble Documentation Centre.

Promoting this Cooperation through an Organisational Formula: The Network Group

The experts came from different departments of the organisation; it was a network-type management with hierarchical participation, as the unit Research Director participated in the meetings. The group consisted of:

- ❖ Scientists and PhD students with personal knowledge of academic publishing, coming from the research unit;
- ❖ Information professionals from the documentation centre who acted as experts in scientific journals, IT experts in databases coming from STI headquarters.
- ❖ The Research Unit Director doing research as well.

The relationship within the group combined good relational communication, agility in decision, and reciprocity between members, which can be described as network management. However, because of the reliance on the Research Unit Director and on the experts in publishing leadership, the organisational formula could be described as spiral management (Obata & Shiizuka, 2003).

The scientists in this group were willing to share information, so a direct method was used to question researchers during personal interviews.

From an individual point of view, many models were studied (Wilson, 1999) to analyse information behaviour. Polity (2001) describes researchers' needs for information and draws the conclusion that researchers show no specific behaviour when seeking for information. Like other users, researchers use the least time-consuming approach when dealing with their activities. They choose familiar and easy access sources. Such behaviour confirms Mooers' law: "An information retrieval system will tend not to be used whenever it is more painful and troublesome for a customer to have information than for him not to have it" (Mooers, 1959)

Researchers in this group were willing to share information if they could find this familiar and easy access tool to archive and retrieve information.

Before the database was created, the answers to users' questions about publishing were disseminated throughout several websites. Navigation was complicated and did not respect the three-click rule (Porter, 2003): "If users can't find what they're looking for within three clicks, they're likely to get frustrated and leave the site." We were faced with the problem of linking up all this information coming from various information systems. We needed to integrate it into one single system.

Finally, there was rich internal knowledge kept by researchers after publishing, including knowledge of who had published in a journal, who read that journal regularly, who monitored its content, or the time taken from submission to publishing. This knowledge was not easily shared among researchers as no tool facilitated the flow of information.

That resulted in identifying a real need for the creation of a new database meeting both the need for consolidating researchers' expertise on publishing and for gathering the answers to the researchers' most frequently asked questions before publishing.

The Database

The researchers' needs and questions were identified and translated into database specifications with the help of IT people.

At the same time, a database was developed and new software Cadic from Cadic Corporation was purchased by the Scientific and Technical Information Department. The user-friendly interface of this software was taken as a true advantage for using the database as well as its appropriation by researchers.

- Two web-based interfaces; one for searching and another for entering data;
- Entering data was facilitated by pull-down menus;
- The workflow allowed information professionals to index and complete the references (links, IF, key-words, etc.).

At the time of creation, the database was filled with 150 items. They were selected by researchers who reviewed two main lists of articles. The first one included articles from INIST (French National Institute for Scientific and Technical Information) while the second one was recorded from Thomson ISI WoS (Web of Science is a database from the Institute for Scientific Information®). After a selection process, the list was reduced to 10 - 15 journals by each researcher.

To increase the quality of the paper and, therefore, its chances of being accepted, an internal review team was created. Any researcher was able to request feedback from his/her peers before submission.

Now containing more than 350 references, the database is considered as a true added-value to the Documentation Centre in the Institute. Not only does this concept of database provide a tool to gather information but it also allows quick access to it. Comparing journals is also made possible as they are described with the same common items. The fields to be filled in are the same for all journals.

After using the database for three years, its main objective has been reached. The number of articles in scientific journals has more than doubled in three years. While seven articles were published in scientific journals in 2003 (out of a total 51), in 2006, the number increased to eighteen (out of the 64 published that year).

Cemagref Grenoble's Wikindx: A Wiki as the Main Tool for Documentary Monitoring/Intelligence

Scientific and technological intelligence aims to identify new research areas and experts in the domains concerned and to watch a specific technology and its applications.

Cemagref's network of information professionals offers advice and tools for researchers. For example, information officers monitor databases, websites, and journals that focus on Cemagref's different research fields. They provide help to find the best query strategy for each database. The results are given to individual researchers to enhance their own monitoring. This system allows the researcher to be informed in real-time on the latest articles related to his/her domain.

Researchers have always needed to spread and share their bibliography with other research teams. Developing a variety of tools to manage and share a bibliography has been made possible by electronic access to information. Other solutions have recently appeared with Web 2.0.

In its 2004-2008 strategic plan, Cemagref placed the scientific and Technical Information Department as a key support to scientific and technical work.

An STI team supporting scientific intelligence was created in 2006. The objective was to provide methodological and logistic support to leaders, decision-makers, and researchers through different solutions. At Cemagref Grenoble, the solution came from an open source tool which would help researchers and information officers work together more efficiently: Wikindx.

In April 2007, the documentation centre decided to implement a Wiki database to enhance the monitoring of research units. This new tool was used to input relevant references and documents. References could be retrieved in different export formats so as to be saved and then reused in each researcher's bibliography.

A Wiki is a CMS (Content Management System), which is a collaborative website where users can create and edit pages. Wikis are decentralised: there is no editorial oversight or approval (Choate, 2006). A Wiki is "the simplest online database that could possibly work" explained Cunningham (Leuf & Cunningham, 2001). (Ward Cunningham and Bo Leuf designed WikiWikiWeb to be an open, collaborative community Website where anyone can contribute). A Wiki, from the Hawaiian term for quick, is essentially a small piece of server software that allows users to freely create and edit Web content using any Web browser and no other special tools. No HTML (Hypertext Markup Language) or programming knowledge is needed to contribute to a Wiki.

“WIKINDX is a single or multi-user bibliography environment storing searchable references, notes and citations and integrated with a WYSIWYG (an acronym for What You See Is What You Get, used in computing to describe a system in which content during editing appears very similar to the final product) word processor for the authoring of publication-ready articles automatically formatted to chosen citation styles” (WIKINDX, n.d. Wikindx is an open source platform developed under the GNU GPL (General Public License) licensed by SourceForge.net. It may run on all major operating systems (Linux, Apple OSX and Windows). This Wiki makes it possible to deposit bibliographic references with full text attachments via the Web. Actually, Wikindx allows multiple attachments for each bibliographic resource. This tool allows multiple users to access shared bibliographies. The users can comment upon the references and export them in different formats including RIS (The RIS file format is a tagged format for expressing bibliographic citations), BibTex (a tool for formatting lists of references), Endnotes (commercial reference management software), XML (Extensible Markup Language), RTF (Rich Text Format), PubMed (a free search engine offering access to the MEDLINE database of citations and abstracts of biomedical research articles), and Zotero (a free, open source extension from Mozilla for the Firefox browser). Users can export bibliography directly in various bibliographic styles (American Psychological Association (APA), Institute of Electrical and Electronics Engineers (IEEE), for example). Furthermore, it is possible to write an article entirely within Wikindx’s WYSIWYG word processor. Contributors can access their bibliography from any point on the net and share it with other members of their research team.

To promote this service, the documentation centre decided to prepare a newsletter with all references entered last into the Wiki. After consulting with the research teams, it was decided to publish four monitoring newsletters every two months. The documentation centre sent the monitoring letter via a mailing list. To avoid sending the information to everybody, four mailing lists were created. Researchers were invited to subscribe to one or more mailing lists. As shown in Figure 1, the Wiki is fed “almost” automatically, from different sources:

- ❖ the traditional documentation database where books, reports, proceedings in full text are saved;
- ❖ references from a commercial database;
- ❖ references from the press;

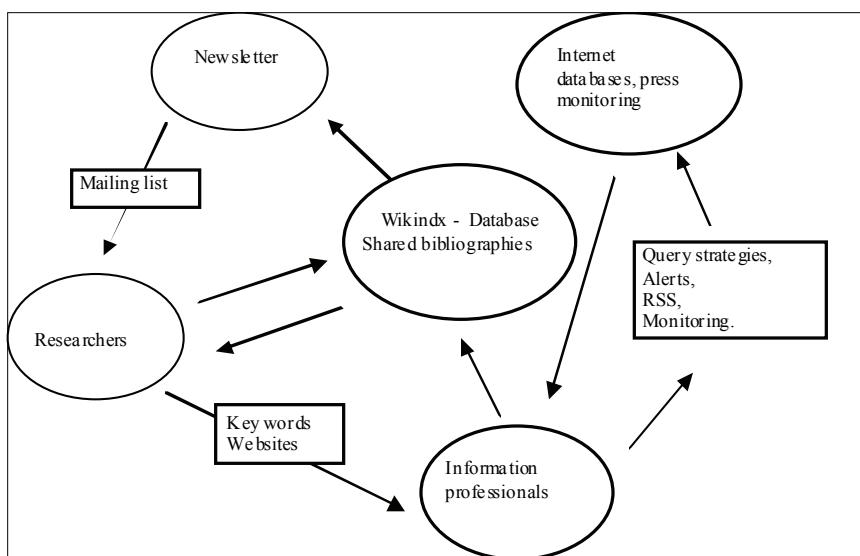


Figure 1: Information Flow

- ❖ from websites;
- ❖ references from individual researchers.

A validation stage is necessary to target the researchers' subjects. Close contact is maintained with researchers to determine the needs of the monitoring team. The data collected and validated is referenced and easily found at a later stage.

The number of subscribers to the mailing list has increased. Nevertheless, to target the right subjects this monitoring service has to be updated constantly with the keywords of new comers.

Scientific and technical monitoring remains the researcher's task. To promote a constant flow of information, collaborative behaviors' are needed. Tools coming from Web 2.0 may contribute to this process.

Wikindx is the reference manager software. It is free and ready to be used. As reference managers are used to sharing bibliographic data, it can be compared to EndNote, the other bibliographic software used in our institution, as shown in Table 1.

Table 1: EndNote compared to Wikindx		
Software	EndNote	Wikindx
Developer	Thomson Corporation	Mark Grimshaw
Open source	no	yes
License	proprietary	GPL
Runs Web server	no (in our version)	yes
Operating system support	Windows Mac OSX	Windows Mac OSX Linux BSD Unix
Export file formats	BibTex, Endnote/Refer/BibIX, RIS, Xml	BibTex, Endnote/Refer/BibIX, RIS
Cost	public price	free software
Document attachments	yes	yes multiple attachments

Both are able to exchange data in a great variety of formats.

Some interesting points about Wikindx are:

- This software gives the possibility to share one's bibliography with other members of the team or research unit. It runs on a desktop computer and on a web server so people can access their bibliography from any point on the web.
- Multiple files can be attached to each bibliographic resource. This makes it a useful tool to manage information.
- People can store quotes, paraphrases and thoughts for each resource so that they can share their opinions or tag any reference.

References may be managed at three levels:

- General level: users can enter and publish references.
- User group level: some members share bibliographies.
- Individual level: for personal bibliographies only.

Information for this monitoring Wiki comes from different internal and external resources; the data is already saved in other databases, so long conservation data warranty is not needed here. What really counts is fresh information.

The Main Lessons Learned from the Customised Database and the Wiki

It is interesting to compare the two databases and report their differences in order to choose the appropriated tool when necessary.

Cadic, a commercial document retrieval software program, had just been bought when it was decided to create the journal database. We chose this software for this particular application because we could have technical assistance and data conservation; transfer to new systems was also guaranteed. This software from Cadic Corporation, France was bought after evaluating the needs and surveying the document retrieval software programmes on offer: i.e. for choosing this kind of solution, there must be a Scientific and Technical Information department within the organisation. Many items have to be defined: meta-data, tables, index, mandatory fields to fill in, online user help, etc. Information professionals translate the users' needs into specifications. Then IT experts implement the specifications in a tailored database.

Today researchers are users more than they are contributors. They still give preference to titles and express their feelings for journals they wish to publish in. Yet, information professionals are responsible for the annual updates (IF, link surveys,). Each year, ten days' work is necessary to update the data. Besides, once a year, one full day is dedicated to a meeting with IT experts at headquarters to implement the database improvements. The database requires hard work to be an actually reliable tool.

The Wiki for our specific application was ready to be used; IT experts installed the tool in very little time. Although the database workflow was difficult to install, two levels of users were suggested. Three levels of group users were easy to set up in the Wiki as the administrator's interface is friendly. This tool is very useful for the documentation centre. The newsletter is easily done with the information retrieved from the Wiki. We hope the tool will be adopted by researchers. We have been working towards this objective. This Wiki is presented in documentation trainings, internal meetings and seminars. A link has been added directly to the intranet portal. The number of researchers to subscribe to the mailing-lists is 66 (55%). Nine users have entered references into the Wiki so far.

Table 2 compares the two tools.

Table 2: Cadic -Wikindx

Software	Customised database Cadic	Wikindx
Developer	Cadic Integrale	Mark Grimshaw
Cost	public price	free software
Open source	no	yes
License	proprietary	GPL
Operating system support software	Windows, Linux, Sun Solaris, HP-UX	Windows, Mac OS X, Linux, BSD, Unix
Export file formats	yes (word, excel)	yes
Import file formats	no	yes

Password protection	yes	yes
Workflow	yes, difficult to create	yes, easy to install
RSS alerts	yes (not in our version)	yes
Navigation by clouds	no	yes
Navigation by key-words	yes	yes

A researcher's point of view

This system allows us to save time in our daily work. More specifically, the selection of key-words by information professionals and the availability of a searchable database as well as an alert system (i.e. e-mail updates of the latest relevant database results based on a topic) allow us to optimise literature search and monitoring. On a daily basis, this entails an update of our knowledge and more productive writing. This system can be exploited to its full potential only if the user agrees to be trained on how to use the tools. In addition, he/she should contribute to the literature monitoring system, e.g. by informing others on the appearance of interesting articles.

Constant dialogue between researchers and information experts can only enrich the activities of both. To researchers, it seems essential to involve the information professional during the publication process.

Conclusion

How are such tools going to develop? It is a little too early to say anything about Wikindx. For the database however, it has been planned to promote the tool in the 10 Cemagref documentation centres and among researchers. It has also been requested to provide information via open access journals.

This project was made possible thanks to daily cooperation between researchers, IT engineers, and information professionals.

The two systems presented here are examples of tools developed to help researchers with their everyday requirements and needs for information.

Information technology is in constant evolution. Information professionals should remain aware of such developments and implement them in their documentation platform whenever necessary. Not only do they have to implement those new tools, but they also have to train researchers on how to use them efficiently and ensure constant access to a high level of information within the research institute. Nevertheless, such tools cannot be developed by information professionals only. They help people to work efficiently because they are the fruit of a close and daily multidisciplinary collaboration between information professionals, information technology experts, and researchers.

The particularly interesting aspect of Wikis reflects in the personal use of collaborative tools by Cemagref scientists and their outside partners. In this respect, the STI Department has assisted scientists in taking over such tools and persuaded scientific information experts that they are as useful, for instance, as the introduction of shared scanning tools for scientific research-related exchanges and knowledge production. Moreover, the progressive use of Web 2.0 tools within professional spaces (intranet or extranet) is also well worth mentioning.

References

- Bach, J. F. (2007). Problems raised by peer review. *France Institute meeting. Evolution of scientific publications: The point of view of scientists.* Paris, France, 14-15 mai 2007
- Boussou-Pélissier, C. (1997). *Où publier ? Catalogue d'information sur les revues internationales en agriculture tropicale.* [Where to publish? A list of international journals in tropical agriculture.] Montpellier, CIRAD.
- Cemagref (1999). *Cemagref strategic plan 1999-2003.* Antony, Cemagref
- Cemagref (2004). *Cemagref strategic plan 2004-2008.* Retrieved January 8, 2008, from
<http://www.cemagref.fr/English/PS04/ps04intro.htm>
- Choate, M. (2006). What makes an enterprise wiki? *CM Watch.* Retrieved February 12, 2008, from
<http://www.cmswatch.com/Feature/145-Wikis-in-the-enterprise>
- Guédon, J. C. (2001). In Oldenburg's long shadow: Librarians, research scientists, publishers, and the control of scientific publishing. *Proceedings of the Association of Research Libraries 138th Annual Meeting: Creating the Digital Future,* Toronto, Ontario, Canada. Retrieved January 8, 2008, from
http://eprints.relis.org/archive/00003951/01/ARL_Proceedings_138_In_Oldenburg's_Long_Shadow,_by_Guedon.htm
- Jannès-Ober, E. (2005). Le Cemagref la fonction IST. [A new plan for the Scientific and Technical Information Department of Cemagref]. *Rencontre des professionnels de l'IST 2005, Nancy, France, 20-22 juin 2005*
- Leuf, B., & Cunningham, W. (2001). *The Wiki Way: Quick collaboration on the web.* Boston: Addison Wesley.
- Mooers, C. N. (1959). Mooers' Law; Or why some retrieval systems are used and others are not. *American Documentation*, 11(3), i.
- Ministère de l'Agriculture et de la Pêche, Ministère de la recherche, Cemagref (2001). *Contrat entre l'Etat et le Cemagref 2001-2004*, p. 20. Retrieved January 8, 2008, from
<http://www.cemagref.fr/informations/Presentation/contrat2001/Contrat-quadr.pdf>
- Obata, K., & Shiizuka, H. (2003). Spiral management for e-business. *Journal of the Asian Design International Conference*, (G-01), 1-7. Retrieved February 14, 2008, from
http://www.idemployee.id.tue.nl/g.w.m.rauterberg/conferences/CD_doNotOpen/ADC/final_paper/499.pdf
- Polity, Y. (2001). Le comportement des chercheurs dans leur activité. [Teachers' attitude in the context of their investigations (especially towards information material)]. *Perspectives Documentaires en Education*, 52, 81-87. Retrieved February 12, 2008 from <http://www.iut2.upmf-grenoble.fr/R13/Adbu.htm>
- Porter, J. (2003). Testing the three-click rule. *User Interface Engineering.* Retrieved January 8, 2008, from
http://www.uie.com/articles/three_click_rule/
- Turner, W.A. (1996). The modern information professional: Towards a new partnership. *L'information scientifique et technique: Nouveaux enjeux documentaires et éditoriaux.* Colloque INRA, 21-23 Octobre 1996, Tours
- Wilson, T .D. (1999). Models in information behaviour research. *Journal of Documentation*, 55 (3), 249-270. Retrieved February 12, 2008, from <http://informationr.net/tdw/publ/papers/1999JDoc.html>
- WIKINDX. (n.d.), Retrieved from <http://sourceforge.net/projects/wikindx/>

Biographies

Library director, **Amanda Regolini** is a research engineer at Cemagref, where she has worked since 1990. She has been responsible for the bibliometric analysis at the institute since 2006.



Frédéric Berger is Forester and research engineer at Cemagref in Grenoble. He leads the Protection Forest Research Group and is specialised in rockfall and avalanche protection forests.



Emmanuelle Jannès-Ober is geographer and has almost 20 years of experience in managing scientific and technical information. She has worked for different international organizations in Africa and 10 years for the Institut Pasteur in Paris. Since 3 years, she is the head of the Scientific and Technical Information Department of Cemagref.



Luuk Dorren is Physical Geographer and works currently as senior researcher at Cemagref in Grenoble. He is specialised in the interaction between natural hazards and mountain forests in the Alps.