The Application of Electronic Copyright Management Systems (ECMS) to Grey Literature

Dave Davis, Project Manager, Copyright Clearance Center

Abstract:

Electronic Copyright Management Systems, such as those provided by Copyright Clearance Center (CCC), enable maximum dissemination and impact for copyrighted content (including grey literature), while still returning aggregated use data to rightsholders (authors, publishers or others) as well as a range of royalty options, from 'zero' to "whatever the market will bear." New and unanticipated uses of grey literature, such as republication into a collective work (essays collected in books), or licensing for digital use on corporate intranets, or transactional licensing for distance education or other academic use on digital networks (such as professors making their class notes available for use by professors at other schools), are often most efficiently and effectively managed by a third party. A third party aggregator, such as CCC, matches potential users with a repertory of permissions (rights to use different works in various ways) in a web-enabled, customizable environment. Several examples are provided, with questions for further research included.

In 1788, referring to the proposed constitutional clause for copyright and patents in Federalist #43, James Madison wrote "The public good fully coincides in both cases with the claims of individuals." Whether this idyllic description fits any specific copyright regime is a matter beyond the scope of this paper, but that copyright law and systems based on that law generally are intended to benefit both the public and private concerns seems indisputable. The familiar twin clauses of Section 8 of Article I of the U.S. Constitution echo Madison's understanding of this rare 'full coincidence' of the public and private goods: "To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries."

> "To promote the Progress of Science and useful Arts,
> by securing for limited Times to Authors and Inventors
> the exclusive Right to their respective Writings and Discoveries."
> (U.S. Constitution, Article 1 Section 8)

In this paper, in working through the example of Grey Literature reuse (republication and redistribution to new publics), I will describe how Electronic Copyright Management Systems (ECMS) can be used to facilitate these noble and 'coincident' goals, both in 'promoting the progress of science' as well as 'securing' rights for private rightsholders.

Systems for managing rights and permissions, in the context of secondary publishing, have been around for many years. Both Charles Dickens and Mark Twain, for instance, successfully increased their fame by recasting material originally written for magazines into full-length books. More recently authors, literary agents and publishing organizations have had to develop routine procedures for answering requests for the re-use of materials that they control. In our new web-enabled, multimedia society the complexity of these requests, plus the potential for ever-increasing volume, implies an increasing need for more efficient systems to handle such requests. Rule-based automation facilitates this process, while honoring the constraints and requirements of both creators and licensees. ECMS were functionally described in extensive detail in a December, 1998 World Intellectual Property Organization (WIPO) paper, "Electronic Rights Management and Digital Identifier Systems" by Daniel Gervais (of CCC). Contemporary ECMS, such as those offered by Copyright Clearance Center (CCC) are able to handle the complex rights involved in licensing copyrighted content (text, audio, video, photographs, animation, streaming media, etc.). On the other hand, ECMS must also be able to meet rising user expectations that are now coming to include rapid, granular search & retrieval, bundled permissions, financial transactions, and content handling. Increasingly sophisticated, impatient users will require "Soup-to-Nuts" solutions, not handoffs and delays.

Although a technology-neutral, standards-based approach represents the path of wisdom to this goal, any complete market solution will tend to include at least these several technical elements, and perhaps others:

- Works level (Granular) Searching
- Standardized Metadata Tagging
- Persistent Resource Locator
- Maintenance of Rights Management Information
- Digital Content
- Real Time Access to Usage Data

Users will expect to search and locate at the article level or below. Downstream distributions (by users, as 'republishers' to new users) are possible, but not guaranteed. Standardized metadata tagging, as well as the use of persistent resource locators, enables potential users to find the material reliably, time after time, despite transfers of ownership or changes in distributors. Web-oriented users presume that the content they seek is begins its existence in digital form, and continues through its lifespan as bits, not as ink on paper or celluloid.

Meeting these presumptions is beneficial to both users and rightsholders. A secure container option for the material may be utilized to assure rightsholders that this is not the last sale or license for their material- this is a distribution, not an "abandonment of rights". This option is likely to be more appropriate for materials of high commercial value, or for documents of a confidential or sensitive nature. The last element, real time access to usage data, is also of primary value to rightsholders who need to know how well their materials are doing, both as a commercial property, and as an offering in the marketplace of ideas.

These technical elements are only part of the picture: human interaction, or just plain old great customer service, is also a nearly universal expectation of users and rightsholders in functioning copyright management systems. What CCC is finding, as a first-mover in this market space, is that technology *alone* is decidedly not enough - it is necessary but not sufficient, as my instructors in scholastic philosophy would say. Incorporation of this point, that high-touch customer service plus functional user-friendliness embedded within the technology plus recognition (based on experience) that the real world is not entirely susceptible of automation, is what distinguishes real world ECMS solutions from 'pretty on paper' business models that have not yet stood up to the acid tests of meeting a bottom line and meeting customer expectations, day after day, scalably and reliably.

As a context for the prospect of new uses for copyrighted material, in a recent report prepared for the U.S. Copyright Office, "Project Looking Forward, Sketching the Future of Copyright in a Networked World" (May 1998), Trotter Hardy wrote, "an appreciation for possible future market effects of a new use technology is essential to good copyright decision making in the present...." Similarly, aggregated information about the use and placement of a work may, in some cases, be more valuable than royalties! As the papers from previous GreyNet conferences demonstrate, a document within Grey Literature most commonly begins its life as a report or other publication intended to fulfill the requirements of a relatively small audience. While a significant portion of Grey Literature may be in the public domain, many other Grey Literature works are copyrighted and theoretically available for commercial exploitation. Or, if not offered for sale in hope of profit, then they are simply made available for redistribution to a wider audience, for low- or no- fees. An ECMS enables these redistributions to occur without unnecessary delays or loss of control.

While grey literature, almost by definition, often begins existence in a non-commercial context, multiple re-uses are possible and indeed often occur in the real world. Grey Literature papers are often republished into books, developed into "new media" (audio, video, CD-ROM) or even incorporated into a "community of interest", in the sense made popular by Larry Downes, Chunka Mui in their recent book Unleashing the Killer App : Digital Strategies for Market Dominance (Harvard Business School Press, 1998) by being excerpted or

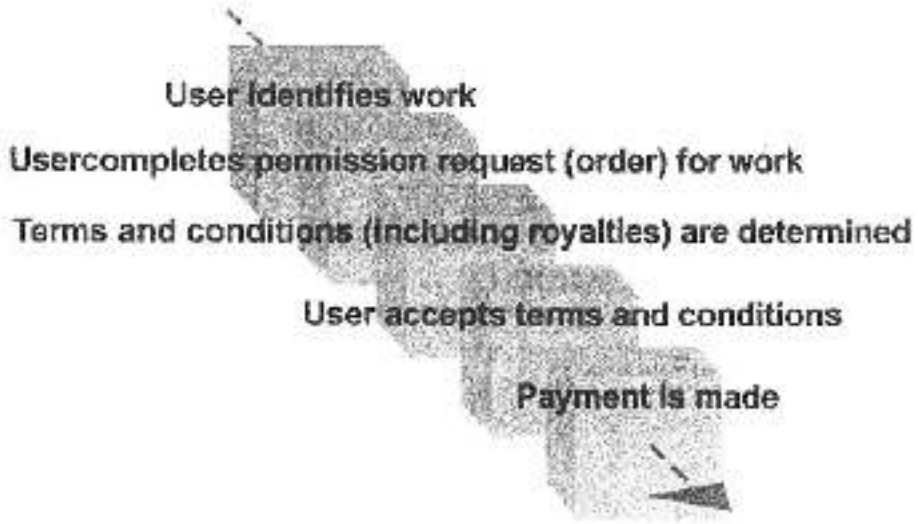
uploaded to a web-based discussion board. In that context, written works represent knowledge assets- they encapsulate findings, and, when of high quality, provide useful answers to real-world problems. They merit intellectual respect, and in the instance of Grey Literature, often deserve wider distribution than their initial publication achieves.

What are the prerequisites of a fully functioning ECMS? There are two principal aspects to ECMS- many ballyhooed commercial solutions seem to grab one or another of these aspects, but not both.

For the simpler aspect, the economic transaction, the requirements are:

- a Known, Discrete Work;
- a Rightsholder with 'Clear Title' ;
- a User with actual Intent-To-Use
- a Mutually Acceptable Agreement (and, possibly)
- a Payment Mechanism

Re-use as a Simple Economic Transaction



The second aspect, the communications cycle, entails a more complex system:

- Out of an earlier Work, a new Work is created by a creator for an audience; then
- The Work is made available (passively) or distributed (actively) to the audience; then

- The Work is received by its first audience- read, critiqued, acted on or disregarded, perhaps quoted or cited, or affirmed or rebutted; then
- The Work may be repackaged, translated, digitized or uploaded, excerpted or used in a new report, updated and revised, deeply annotated and discussed via hypertext; then, optimistically
- This now classic Work may become the basis of a new Work (Return to top of cycle and begin again)

Re-use as a Communications Cycle

This communications cycle is familiar to researchers and authors publishing in the sciences and social sciences, especially those utilizing peer review. From that perspective, any economic transactions that occur may appear of secondary importance, if they appear to bear any importance at all.

Copyright Clearance Center (CCC) is the largest, and the only not-for-profit, licenser of copyright reproduction and distribution licensing/permissions services in the U.S. CCC was formed in 1978 to facilitate compliance with U.S. copyright law. CCC provides licensing systems for the reproduction and distribution of copyrighted materials throughout the world. As part of our mission, CCC continually develops systems that utilize technology to support copyright compliance. Additionally, CCC's newly developed rights management technology allows us to respond quickly to opportunities on behalf of our rightsholders and users, and to provide a wide range of licensing programs tailored to various types of media.

Currently CCC provides all the elements of an ECMS in the sense discussed above through a suite of services:

- Direct Author Registration
- Various User Licenses: Transactional Reporting Service (TRS), Annual Authorizations Service (AAS), Academic Licensing Service (ALS)
- Republication Licensing Service (RLS)
- Plus others as well

Republication Licensing Service (RLS), in particular, fits well in this analysis, both from the perspective of *users* and from the perspective of *rightsholders*. The new RLS from Copyright Clearance Center is a fully automated, online service for licensing the rights to republish copyrighted works in both print and electronic media. In order to make efficient and effective use of previously published, copyrighted material, *users* have the ability to place their requests quickly (via web), to have the permission request either granted on the fly or, if necessary, automatically forwarded to the rightsholder. *Rightsholders*, on the other hand, gain the ability to have incoming requests made as standardized as possible, made for materials that they in fact control, and processed all the information necessary to making a permission decision. And, if royalties are part of the equation, they will be assessed accurately and according to the rightsholders' specifications. *Both* rightsholders and users benefit from knowing, in the aggregate, *what* material is used (a criterion of value), *who* is using it (a marketplace datum), and *where* (or to what new audience) it will be redistributed. CCC's existing services are capable of facilitating this complex process at a highly granular level (i.e., individual articles or photographs, etc.), and new service enhancements are planned to continue meeting these emergent user and rightsholder needs.

RLS also accelerates certain other user and rightsholder requirements for an ECMS. In the case of users, RLS obviates the need for time-consuming sleuthwork sunk into locating an obscure copyright holder. As an additional time-saver, in many cases the permissions will be granted instantaneously. Otherwise, RLS automatically contacts the rightsholder on the user's behalf, then reports the results to the requester online. Perhaps most importantly to users, use of the authorizations provided by RLS becomes a key component in their strategy of minimizing any risk of noncompliance. The additional requirements of rightsholders are: the option to decide which works to authorize for republication; what royalty fees to charge based on the type of use; whether to grant permissions on a pre-approved or case-by-case basis; and which formats (e.g., print, e-mail, Internet, Intranet, CD-ROM) to authorize republication within (based on the scope of rights held by the rightsholder). Finally, CCC distribution reporting aggregates the usage data and provides indications of the republished material's impact upon its new audiences.

Stepping away from the marketplace or service-level view, what are the ECMS functions described here?

- Aggregating users into a comprehensible marketplace
- Aggregating rightsholders and their works into a catalog of needed permissions
- Aggregating user data, works data, and rightsholder data into informative reports
- Processing financial transactions quickly and seamlessly
- Managing and maintaining rights knowledge accurately, and at the necessary degree of granularity

The complexity of administering this web of relationships is what renders rights management a field not for those faint of heart or new to the game. However, this overview demonstrates that CCC's existing services speed up and normalize *both* the transactional aspect – doing the deal, if you will – and the communications aspect. Additional services under development will do so with even greater efficiency and with wider scope.

Although CCC's services are online and operational *now* for the re-purposing of Grey Literature, several topics implicit in the discussion above remain as open areas for further research:

- To what degree is grey literature re-used in commercial publishing? What are the trends?
- What is the impact factor of grey literature, as measured by re-use? Does it differ from impact factor as measured by citation?
- What unexploited market opportunities exit for custom publishing of grey literature materials?
- Is standardized subject access (on a granular level) a prerequisite of increasing reuse?

Publishing and grey literature re-use are often analyzed transactionally. This works well from the perspective of economics, in Madison's sense of 'private good', and as e-commerce continues to advance we should expect to see more of this understanding applied. However, viewed as communication cycle, a different picture emerges- one of a feedback loop that continually increases the public good through its iterative turnings. It is the challenge of any ECMS, and of CCC's services in particular, to honor and facilitate both terms of this complex equation.

Finally, while we may not be able to agree with Mark Twain, who claimed that only one thing was impossible to God, to find sense in any copyright regime in existence, we can hold onto a vision where copyright systems enable and

facilitate both economic transactions **and** scholarly and professional communication, for Grey Literature and all others.

Additional Resources on ECMS

"Copyright Collectives and Libraries" (Interview) Copyright & New Media Law Newsletter (Spring, 1998).

"Copyright in a Digital Age: Practical Guidance for Information Professionals." Robert S. Weiner, CCC Online, May 1997.

"Copyright Protection for Electronic Publishing Over Computer Networks". A.K Choudhury, N.F. Maxemchuk, S. Paul, and H.G. Schulzrinne. IEEE Network, p 12-20, May-June 1995.

"The Economics of Print versus Electronic Publishing: An Illustration from Mathematics" Michael Rappa (draft) December 1998. URL: http://www.technika.org/copyright/model.html ; see also the accompanying graphic http://www.technika.org/copyright/model.pdf

"Electronic Rights Management and Digital Identifier Systems," by Daniel J. Gervais Journal of Electronic Publishing, University of Michigan Press (March, 1999) URL: http://www.press.umich.edu/jep/04-03/gervais.html .

"ERCOMS (1998) Market Survey of Existing Electronic Copyright Management Systems (ECMS) and Projects." URL: http://www.iiir.dmu.ac.uk/Projects/ERCOMS/ercomsm.html

Information Sources in Grey Literature . Peter Auger . Publisher: Bowker-Saur. 4th edition 1998.

"The invisible hand of peer review." Harnad, S. Nature [online] (Nov. 1998) URL: http://helix.nature.com/webmatters/invisible/invisible.html .

"A new generation of Grey Literature: the impact of advanced information technologies." G. A. Cotter, B. C. Carroll. Paper presented at GL '93 (December, 1993).

"PURLs: Persistent Uniform Resource Locators" Stuart Weibel, Erik Jul and Keith Shafer. URL: http://purl.oclc.org/oclc/purl/summary .

"Rise of the Phoenix: A review of new forms and exploitations of Grey Literature." Dominic J. Farace Publishing Research Quarterly. Vol. 13, #. 2, 1997.

"Survey on Electronic Rights Management" (1998 -) International Federation of Reproduction Rights Organizations (IFRRO), URL: http://www.ifrro.org/committees/surveyintro.html. See also IFFRO's Collective Management of Digital Rights (October, 1996) http://www.ifrro.org/papers/pp-digi.html

Copyright: Black and white or just making you see red?

Graham P Cornish,
Copyright Officer, British Library and
Director of the IFLA Programmed for Universal Availability of Publications

We live in an era in which rights are a matter of constant discussion and debate. But rights are far from uniform in their character and it is important to place the right to copy – copyright – in its proper context in order to understand some of the issues surrounding grey literature in its broadest sense.

We acquire rights in several different ways. There are what most people regard as the basic human rights – food and water, shelter, and the other basic necessities of life. We acquire these by virtue of our humanity and they require no action from us, nor should they need any legislation to enforce them. Other rights are bestowed upon us by the law of the land such as rights for compensation when we suffer inquiry or redundancy. These rights are ours whether we want them or not or whether we need them or not. We need to exercise them only when the circumstances are appropriate and we **may** choose not to do so. Then there are the rights that we acquire deliberately such as those that go with a piece of land or a trademark. Of course, occasionally we do buy something and acquire rights in the process we never knew existed. The case of the Birds' Eye Company buying a farm in Lincolnshire (England) and finding they also acquired the right to appoint the vicar of the local church is a quaint example!!

Copyright is a right which manifests several of these different characteristics. Whilst nobody would claim it was as fundamental as food and shelter, it still underpins the basic concept that what we create we ought to be able to control. This is because what we create is a projection of "self" in a very vulnerable and obvious way and it is usually considered a basic right that we can and should be able to protect ourselves. For those countries which have signed the Berne Convention[1] the rights conferred by copyright on the author are automatic. They are defined by the legislation of the country concerned and the author, as first owner of the rights, has no say over what rights are acquired in this way. Copyright is conferred by the state in such forms and within whatever legal context the legislators consider appropriate. Once a work has been created the author acquires certain rights (which may vary in content from one country to another) but which are absolute. Many authors never realize they have these rights, especially those who do not create works for gain, either financial or academic. Similarly, many users of copyright material never consider that any rights exist in them at all and those working in the field of copyright can quote example upon example of bizarre and strange ideas about copyright and the ways that one person wishes to use the intellectual property of another.

It is in this context of rights acquired whether they are wanted or not that I would like to discuss copyright and grey literature for I believe it is this passive acquisition of rights that can cause so many people to "see red" about copyright when discussing grey literature in its widest interpretation. There is no point in trying to define what grey literature really is. This is a matter of frequent debate and the general definition of something which has not been commercially

published for general distribution through the normal retail channels will suffice for the purpose of this paper.

Because copyright is automatically bestowed on the creator of an original work, that creator has no say in what rights are acquired or how s/he may wish to exercise them initially. Essentially copyright was invented to protect two distinct elements of the creative process: firstly, the economic value of whatever was created; second, the so-called "moral" rights of the creator, namely the paternity and integrity of the work created. In broad terms this means the right to have authorship acknowledged and the essential content of the work to be left in tact and not interfered with. The distinct nature of these two bundles of rights is a further complicating issue when looking at grey literature. This complication comes about because of the different ownership of these rights, depending on various legal and political circumstances, the complexities of which need not detain us here.

Authorship is, of course, central to the whole idea of copyright because, until something is created there can be no copyright. One of the distinctive aspects of copyright is that it subsists, rather than exists. In other words it is always necessary to create an object (in which other rights of ownership, for example, may reside) before copyright can subsist. [Perversely, copyright CAN subsist in something which no longer exists although it did once in the past. The classic example is a painting by Graham Sutherland of Sir Winston Churchill which Lady Churchill destroyed as soon as he died - but the copyright in the painting still subsists because there is a photographic record of it]. Most authors of grey literature have little or no interest in the economic exploitation of the publication they produce. In the first place, any rights in it are probably owned by their employer as discussed in a moment and secondly, they have no personal interest in the value of the actual document except in one crucial area. Although the so-called moral rights for authors may not be available to employees in some legislative regions, most authors of grey literature will have a strong interest in being acknowledged as the author of all or part of the document. Reputation for researchers in all fields, whether in the natural sciences, social sciences or the humanities is central to developing an acknowledged role in the specialized field in which a particular researcher works. Failure to be acknowledged as the author of a work, however esoteric or trivial it may seem as a publication, can result in serious loss of prestige in professional and academic circles with the resultant loss of status and possible career development. Therefore authors of grey literature actually have an interest in their work which is quite different from that which some employers or owners of economic copyright will have.
The other interest that authors have is the integrity of their text or other material. Statistical or chemical data can easily be manipulated to demonstrate quite different findings from those of the original researcher and the removal of one or two key phrases or even the word "not" will change the meaning of the text for ever. Therefore the integrity of the text is probably at least as important to researchers in all fields as acknowledgement of authorship itself.

So, now let us look at ownership, as this will determine attitudes to the economic elements of copyright. In most jurisdictions copyright in a work is owned by the employer if the work is created as part of employment, otherwise the copyright belongs to the person who created the work. So the copyright in a detailed report on a new drug prepared by a scientist working for a major pharmaceutical company will be owned by that company. But the company is far more

likely to be interested in the content of the research, which is not protected by copyright, than the copyright in the text itself which IS copyright. Once they have obtained all the information they need from the scientist and the report they are less likely to be interested in the text of the report itself. Having been cleared by the appropriate departments in the company from a confidential point of view, the company will be interested only in so far as no other company tries to copy the processes or research described in the report. The fate of the report itself will usually be of little or no interest. It may be released through a number of channels such as NTIS or the British Library or simply put on the company website where it can be freely downloaded, printed off, copied, stored or networked to nay number of interested parties. The company may well not mind any of this happening although it infringes several distinct rights enshrined in copyright including copying, distributing to the public and perhaps broadcasting[2]. This is in stark contrast to the commercial publisher who has a strong economic interest in the distribution of any work they publish. This may include hardback or paperback rights as well as geographical areas of distribution and the possibility of turning the book into a film or TV series. For academic publishers they will have an interest in mounting a scientific journal or report on CD-ROM to be sold for profit or at least putting it on a website to which access is limited to those who agree to pay. Very often the producer of grey literature will be grateful that someone copies or distributes the work as it actually saves the original company money in terms of staff time, consumable and warehousing. Where a work is made available electronically then, again, most producer of grey literature would be unwilling to invest large sums of money or technical expertise in constructing mechanisms to inhibit copying, downloading or distribution when these activities are actually the real reason the work was released in the first place.

However cavalier companies and institutions may be about the release and distribution of their reports, never the less they too may be sensitive about being acknowledged as the source of the information and may want their contribution to the increase of knowledge properly noted in any re-distribution of literature they produce.

We have therefore two sets of rights: moral rights which both authors and institutions will almost certainly wish to defend and economic rights which may be of little or no interest to those who actually own them. If they really saw any economic benefit in releasing material presumably they would have opted for the commercial publishing route anyway. This, in turn, leads to another approach to protecting copyright. Owners may not wish to go to the expense or bother of publishing material commercially but, equally, they may wish to prevent other from doing so as well. It is one thing to allow someone to download or copy a report for research or academic use: it is quite another to allow them to do the same actions for the purpose of commercial republication, either as a "stand-alone" document, or, more likely, to be integrated into a larger work either as a chapter for a book or part of an anthology on a specific topic.

But there are other concertinos in the field of copyright as well. Particularly important is the right of distribution. Owners of copyright have the right to control how their works are distributed or published. Although many owners of rights in grey literature may not have an economic interest in distribution, they may have another agenda to follow in terms of controlling *who* has ready access to the material and also *in what form* that access shall be. This may be because an institution wishes to discourage some other organizations or groups from gaining ready access to the material

or it may simply be that the owner wishes to know who wants to have access to give a picture of those working in a specific field of research. This links into the proposed [3] "Communication to the Public Right" which is likely to hit the statute book in many countries in the next five years. This right will give the owner the exclusive right to make material available electronically, such as on a website, whether or not that material is actually consulted. In the normal document supply world the algorithm is used that 80% of demand is met by 20% of the collection and if this is applied to the World Wide Web there are an awful lot of websites or bits of websites that are never visited and this is almost certainly true. However, just to put something on a Website will be an infringement of the owner's rights even though it may never be read by anyone. At the moment there are arguments about whether or not putting something a website actually constitutes publication[4] but this new right will take care of this legal nicety.

The creator and owner of grey literature can therefore demonstrate a whole range of rights which are acquired automatically upon creation of the work concerned. Neither the author nor the institution sponsoring the report has any control over what rights they acquire - they are automatic and given by law. Yet they may not wish to exercise any or all of them or may wish to do so only in very limited ways. The other side of this coin is that the user may understand that the owner has this bundle of rights and, equally, is unable to determine which of them the owner wishes to enforce, if any. Unlike physical property there is no ready way of signaling this. After all, if I own a plot of land I can put up a notice which says "Anyone may walk across this land" or, as a cricket field in south London in the UK has "You may walk your dog on this land but do not let it go on the grass." the variant meanings of the word "go" in this context not being explained. But for material which is copyright the user has no guidance. S/he may make an informed guess as to what the copyright owner might allow or consider acceptable but there is no guarantee. Frequently users consult specialists in copyright, only to be told what the law says but not what is the probable attitude of the owner. To be fair, who is prepared to guess on behalf of someone else in such a litigious area of law? They are then left to guess what might be possible without being taken to court. But nobody can be totally sure when guessing, otherwise it would not be a guess. Therefore the potential user is faced with the decision either to stick with the letter of the law in almost certainty that the owner will not take such a restricted view, or go ahead and do whatever they want to do with no assurance they will not bring down the full weight of the judiciary upon their heads.

Clearly this is an unsatisfactory situation when owners do not want to defend many of the rights given them by law and users usually want only to make use of a work in a sensible and realistic manner. What then is the solution?

It seems to me that some kind of internationally agreed code of practice on copyright for grey literature should be adopted. It would acknowledge the needs of users but pay due respect to the legal rights of owners as there will always be those who feel the need to enforce their full suite of rights. Essentially we end up with a table of needs and rights which looks something like the following. Naturally there is plenty of scope for refinement and additional situations to be included but this might provide the starting point for further discussion and implementation.

Action\Right	Code	Personal use	Internal distribution	Non-commercial publishing	Commercial publishing	Website use
Code		**1**	**2**	**3**	**4**	**5**
Always ask	**A**					
Acknowledge author	**B**					
Acknowledge source	**C**					
Paper copying	**D**					
Electronic copying	**E**					
Paper distribution	**F**					
Electronic distribution	**G**					
No restrictions	**H**					

By devising a set of agreed symbols using this sort of matrix would enable producers of grey literature to put a code on each document which could be understood internationally regardless of language, legal tradition or any other local circumstances. The code would act as a sort of proactive permission removing the need for users to seek unnecessary permissions and owners from having to handle requests in which they have no interest and which are nothing but a bureaucratic burden.

For example, using the highly simplified table above, an owner who did not wish to prevent electronic scanning of a work for personal use would code it E1,; a document coded B5 could be used on another person's website provided the author was always acknowledged and so on.

Users would need to be assured that such a code was accepted by owners and used only by those who owned the rights that they were effectively licensing to others. But such issues could be sorted out if there was an international will to do so. After all, the initial reaction to various codes for serial titles were greeted with scepticism but eventually found a permanent place in the world of library science, cataloguing and publishing.

Naturally a lot of work would need to be done on refining this matrix which is just presented as an idea and the different possibilities would be much larger and the restrictions placed would need to be refined but I believe this offers a basic way forward for the management of copyright in grey literature in the future. It is offered in the hope that copyright will no longer be seen as "black and white" ("yes you can" or "no you can't" depending on the attitude of the person consulted) nor will it make either owners or users see red any more.

[1] International Convention for the Protection of Literary and Artistic Works (Berne 1886 and subsequently revised) Article 5(2)

[2] The question of whether providing access to a work through a website is broadcasting hinges partly on whether a website is a broadcast or a cable programme service. This was a major unresolved issue in the *Shetland Times* and *Shetland News* case. Innes, John & McMillan, Fraser (1997) "The *Shetland Times*" Internet case. *Copyright World*, no 76, pp26-28

[3] Proposed by the WIPO (World Intellectual Property Organization) treaty in December 1996 and currently part of the proposed European Union reforms of copyright legislation.

[4] The case of Nottinghamshire County Council in the UK trying to stop a firm of consultants from putting a report on child abuse on the web because it constituted publication is just such a case. CORNISH, Graham P. Copyright and the Internet. In: *Theatre and the Internet: proceedings of a conference held at the University of Glasgow...May 1997.* pp.24-28.

Issues in Distribution of Grey Literature: Experience of the Japan Documentation Center (JDC)

Ichiko T. Morita
Japan Documentation Center

The focus of this paper is to present some issues in grey literature, its dissemination, its place in the collection, grey literature in electronic formats, and difficulty in dealing with languages of publication. Those issues are presented from the past five years of the author's experiences in establishing and developing the Japan Documentation Center. The Center collects, processes and disseminates Japanese policy related grey literature in the areas of politics, economics, business and industry, the environment, and national defense. The discussion, research results, and suggestions follow each.

The first issue is on grey literature and its users. Users of such information like Japanese policy related literature are not necessarily scholars and specialists in Japanese studies. They encompass lawyers, economists, legislators, scientists and many people of other disciplines who must deal with Japan in their own special field. Some are not even regular library users. To disseminate the information on the functions of the Center is a great challenge because one cannot pin-point target groups such as the Association for Asian Studies and the Council on East Asian Libraries.

The second issue is the place of grey literature in the collection development policy in relation to other Japanese materials in a collection. In the recent budgetary crunch in many libraries, the criterion for collection development are reviewed and tightened in general to its core materials of research value in the long-term perspectives of the library, whereas the value of grey literature is for its timeliness and urgent need for practical purposes but not necessarily for its historical value. How the library managers justify the efforts and costs needed for collecting grey literature is interesting.

The third is the development of electronic information distribution and coping with it. Besides the copyright issue which the author does not discuss here, establishing the good practice in handling increasing number of electronically distributed grey literature in relation to paper copies should be thought of.

The fourth is the language of the grey literature. This issue relates to all the above issues; users, collection policies, and electronic distribution of grey literature. In the case of the JDC in which the major portion of literature is in Japanese, how to reduce the difficulties of users dealing with Japanese is another serious matter for some users.

The author discusses each issue area further in detail, with analysis in some cases, and brings some findings or ideas.

Introduction

In conceptualization of the idealistic mapping of the dissemination of grey literature, one quickly realizes several issue areas which must be sorted out in order to seek out ultimate solutions. The focus of this paper is to discuss some serious issues in handling Japanese grey literature, based on the author's past 5 years' experiences in the Japan Documentation Center (JDC). The Center was created as a part of the Asian Division of the United States Library of Congress (LC) and began its function in March 1994. It is established with a special mission to collect the most current grey literature on Japan from primary sources in Japan, concentrating on the policy related information covering the areas of politics, economics, business and industry, society, the environment and national defense. Its functions are to process the collected information quickly, create bibliographical access, and disseminate the information in the most timely and effective manner. Thus, the issues the Center faces may not necessarily be identical with those of more clearly defined academic and research libraries. However, the author believes that some of those issues are of mutual concern and hopes that the JDC's experiences may offer insights to others who handle grey literature.

I believe it might be helpful to consider six specific issues that fall under three broad categories. Some issues are intrinsic to the nature of the grey literature. They would include:

1) Scope of users
2) The language of literature

Some are based on the situation where the grey literature collection is a part of a traditional library as a whole. They are as follows;

3) Shifting focus to address changing needs
4) Ephemerality of literature vs. historical value of research materials

Lastly, some relate to other issues which emanate from the issuing organizations' practice. In this category, the following is included:

5) Printed vs. electronically disseminated information
6) Need for archiving

There are numerous other major, as well as minor issues, a major one being copyright related matters. However, this paper will not deal with them, the underlying reason being that such matters are dealt with in a different arena and the discussion here will not be very productive. They will be touched upon as they pertain to my topics being discussed.

Scope of Users

Unless an agency is a long established one and well known for collection of grey literature, the process of reaching users and educating them on the nature and availability of the particular grey literature collection can be rather difficult. In case of JDC, which began from a zero base five years ago, developing a systmatic approach to identifying users and

disseminating information about the Center has been a serious issue throughout its early period. The fact that there has been much expressed need for Japanese policy information in the United States was a basis for establishing such a service. Moreover, providing for the urgent information needs of the U.S. Congress was the driving force behind establishing the Center at the Library of Congress. Once the JDC began its operations, however, it was extremely difficult to locate the diverse users who need the kinds of information which can be obtained only in form of grey literature.

Academic libraries and most research libraries can concentrate on serving their homogeneous constituents. In a Japanese studies library, for example, expected users are the Japanologists, including scholars, researchers, librarians and students. In case of the Library of Congress, users span a range of disciplines, and the JDC's services extends to everyone from Congress members to general public.

Characteristics of user groups, and issues

Among various categories of user groups, the academic researchers in Japanese studies were the only group the Center could target fairly effectively because they could be reached through already established networks. The JDC used two types of major networks: the Japan related Listserv's, each of which include Japan scholars, researchers, librarians, and students (though not limited to these groups); and professional associations, such as the Association for Asian Studies (AAS), the Council for East Asian Libraries (CEAL), and the Special Library Association. Messages are regularly posted on those Listserv's to inform their members of notable activities and new developments. The JDC staff make presentations and give demonstrations at those associations' meetings, and write articles on the Center's resources and services in various organizational publications. According to our informal studies, 39% of requests came from the academic community. The majority of these researchers focus their studies on Japan in such fields as political science, sociology, business and economics.

This means that the other 61% of users are non-academic and can be broadly considered under the general public category. We have found that these users are scattered in wide area, geographically covering the entire world and spanning numerous disciplines. In other words, there is no specific concentration. The population which needs and can make use of Japanese information, particularly in the Japanese language, has never been high, and there is no easy way to pinpoint who and where they are. The types of general public users of Japanese grey literature information include lawyers, technical experts, journalists, government administrators, exporters and business people, and other types of people who happen to deal with matters pertaining to Japan, among numerous other subjects. They would not likely know that such a specialized center like JDC exists, and for JDC, there is no systematic way to reach out to them. Some of them do not even think about libraries as an information source and some are not regular library users. Once the JDC home page was established in 1997 as part of the Library of Congress home page, it was much easier for users to learn of a specialized center like the JDC through their own web searching.

Another point which should be noted is that most of the users in the general public categories have been one-time users or extremely infrequent users, because they do not concentrate only on matters Japanese. Their needs are not constant. Therefore, it may seem that nurturing and encouraging such irregular users is wasteful, because they may never need the specific type of grey literature information again, and even if there is a next time, they might deal with a very different Japan related matter. A law firm is a good example. Once it has learned of the JDC's

services, it finds the Center very useful whenever it deals with a case with Japan. It may be arbitrary, however, as to when the firm takes on Japan related cases. As another example, an exporter appreciated the information we supplied very much because he needed to have information in order to begin exporting a certain item to Japan. Once that objective is accomplished, there is no further need for the JDC. Even in the Congress, the concerns with Japan are not continuous, though we regularly remind the Congressional members and their staffs of our services.

Outreach to users

Here we begin our endless outreach efforts. Even while the Center has built a strong grey literature collection, use of the information is a crucial measure of the Center's success. We have carried out the routine public relations steps, such as distributing brochures and newsletters and issuing announcements on Listserv's, and establishing a web presence. We routinely send printed literature to individual Congressional members and their staffs as well as to the Asia related committees. We maintain an international mailing list of think-tanks, academic and research institutions, organizations, agencies, and notable individuals, and include all that may have even a slight interest in contemporary Japan and public policy issues.

The JDC staff have presented papers on the Center, often accompanied by demonstrations, and have conducted separate workshops at local, regional, national, as well as international, conferences, meetings, and discussion sessions. To reach as many users and potential users as possible, we actively make professional linkages to organizations to be invited to present the JDC to their members at their meetings. In the past several years, these have included the U.S. Interagency Language Roundtable, Association of Japanese Business Studies, and the European Association of Japanese Resource Specialists.

The JDC has itself organized several very successful international symposia at the Library (1994, 1995, 1996), reaching out directly to a huge number of targeted and potential users through both printed and electronic measures. The symposia brought together experts to speak on a range of issues surrounding Japanese information and provided an excellent forum of exchange between providers and users of information, including grey literature. The JDC also hosted the Fifth JOHO Conference at the Library in 1997, in collaboration with the U.S. Department of Commerce's Office of Technology Policy. Publication of the symposia papers provides a further dissemination of information about the Center and on pertinent issues for those involved in research on Japan.

To summarize, it is clearly difficult to build a satisfactory number of so-called 'regular users' in a short time, in view of the reasons above. At the same time, those investing in and supporting the Center, who desire a quick and concrete outcome, look at certain measurements, such as numbers of users, numbers of documents delivered, and statistics on Congressional use. It has taken five years for the Center's existence to be recognized. We have found that constant, frequent, extremely patient and repetitious, reminders are necessary, and even so, usage patterns are not always consistent.

In addition to successfully making contact with users, there is a need to provide them orientation on how to effectively use the JDC collection and services. Many users are experienced and use the JDC resources either to locate specific information on well-focused topics or as a starting point to develop their research topics. We also direct them to other

resources as appropriate. On the other hand, we often receive questions beyond the scope of the JDC (e.g., historical, cultural), or inquiries from those who expect the Center staff to spoon-feed all the possible relevant information to them. The JDC home page describes the scope of the grey literature collection clearly explains how to request information and how to search the database. All the literature about the Center, printed and electronic, states that the reference and document delivery services are provided free of charge. We were therefore somewhat surprised that in a survey conducted in 1996 about the JDC, a number of people said the information is costly and that they could not afford to pay for the services. Such aspects of actual user behavior indicates to the JDC staff that continuous orientation is needed.

The Language of Literature

With the development of technology and subsequent capability of almost instant transmission of information in a global society, the language (e.g., Japanese, English) of the literature takes on dimensions. It is important to note that the nature of grey literature, which is often primarily aimed at a particular group of people, is that it is not necessarily issued in an international language, such as English. In the case of the types of documents the JDC handles, namely Japanese documents, these are mainly for Japan's policy makers and not particularly intended for a wider domestic or international distribution. More than 95% of the JDC's documents are in Japanese. As mentioned in the scope of users above, users of grey literature information at the Center are not limited to Japanologists who can read Japanese. In fact, over 60% do not read Japanese.

Thus the language abilities (more specifically, language limitations) of the users must be considered. Although the Center was established by the directive of the U.S. Congress which urgently needed Japanese policy information, we soon found out that there were very few Congressional members or staffs who can use documents in Japanese. In the Congressional Research Service of the Library of Congress where researchers investigate and prepare reports for the Congress on specific topics, there are several people in various disciplines who specialize in Japan. Even among them, not all can use JDC documents as they are. Most others need documents translated; some need them summarized; some need to wait for an English version, an availability of which is only arbitrary and certainly less timely.

Among JDC users, except for most of the academic researchers and a few federal government officers who specialize in Japan studies, the majority does not read any Japanese. However, it should be noted that some types of information are very useful for those who have some background on Japan but are not necessarily able to read Japanese. For example, the researcher who needs only statistics can successfully use them with only the translation of headings.

Minimizing language as an obstacle to access

As a partial solution, the Center provides an English abstract and assigns index terms for each document it acquires to enable users to easily identify potentially useful materials. The user, regardless of language abilities, can quickly grasp the content of a particular document and determine its usefulness from the bibliographic record before devoting time and effort to deal with the Japanese language document itself. This approach on the part of the Center requires special human resources.

The work of abstracting, namely going through a Japanese document, identifying the most

crucial aspects of its contents mentally, and writing an abstract in concise English requires quite an effort beyond routine abstracting work. Essentially, we do not have an option of finding experienced abstractors and to teach them the Japanese language. Instead, we have to search for people with Japanese and English language capabilities and teach abstracting. One soon realizes that it is much more difficult to create an abstract than to translate verbatim. It takes time and experience to become a good abstractor. Finding the appropriate personnel who can accomplish this work is not easily achieved. The work requires a native level of understanding more formal, written Japanese documents; training to extract the essence of information; and again native level of English to write clear abstracts. These steps must be done in a shortest possible time frame. For managing any grey literature collection, it should kept in mind that this approach, while effective for the user, is nevertheless costly in terms of human resources and time.

Based on its experience, the JDC has found that greater ease of access to the content of the document collection, particularly one in a non-English language, has a direct impact on usage of the materials.

Users' language abilities shape their needs

At the time of the review of the Center in 1996 to 1997, we found conflicting user opinions: pros and cons of providing abstracts, or the quantity versus quality issue. One opinion is that rather than spending so much effort in creating abstracts, the Center should devote those resources towards acquisition of more documents. The other is that the abstracts are extremely valuable in locating and verifying how useful the information might be before the user begins to cope with Japanese language documents. It is obvious that the former comes from the group of Japanese scholars and librarians with facility for reading Japanese. The latter from the group of non- Japanese readers and non-specialists as far as subjects on Japan are concerned. However, even among the Japanese specialists, there are voices that point out that the abstracts are useful for quick identification of documents, given the general or vague nature of many titles of public policy materials (e.g., "Towards an Affluent Society"). Hence, it was decided by the Review Committee that we continue providing abstracts as a value added feature of the bibliographic database.

It is clear in the Center's experience that provision of English abstracts to Japanese documents improves access to the source information and considerably reduces the seriousness of the issue of language as an obstacle. Without the abstracts, there would be many who would not even try to approach to the information. This was attested to by several users who stated that they would have otherwise had to compromise with the secondary sources in English coming from English speaking areas (such as Southeast Asia) about Japan's policies, even though they realize that the Japanese source documents would provide the precise information on policy matters. As a way to access this body of policy information, these users found the JDC database with abstracts for each item in the collection highly valuable.

Under the present stage of development of automated machine translation of Japanese into English, the output, especially for literature in the social science fields, is yet inadequate. The Center's approach seems to be the next-best solution. The author strongly believes that providing English abstracts of Japanese documents is a notable value of the Center's database, regardless of many problematical factors, such as scarcity of qualified abstractors, cost of abstracting, the time it takes for abstracting, interfacing of abstracts into the database, and maintenance of the database.

The above two issues, namely matters emanating from the types of users (and their language abilities) and from the language of the collection, are based on the intrinsic nature of the grey literature. In addition, the following issues arise from the fact that the grey literature collection is a part of the traditional library as a whole.

Effect of Political Changes

In general terms, the following issues concern the subject matter of a particular grey literature collection and how it relates to changes in need for that literature along with broader societal changes. In the case of the Japan Documentation Center, it is obvious that the need for policy information, especially the kinds that the Center collects, is affected by the political situation as well as by the economy and other factors. To mention a few examples, when the U.S. government negotiated for various aspects of Japan's deregulation, the information on the topic was in great demand; and when the political focus of the U.S. government shifted from Japan's internal policies to Japan's economic policy towards the Asia region, the requests for the Japanese documents on private business activities, global enviromental issues, economic assistance, and other related information increased. Avidly following its mission, the Center collected timely information focusing on the political, economic, and other topics in demand at a given point in time.

As a result, the Center's collection is not one of those that collects systematically on pre-determined topics in order to build a well-balanced collection for more historical and fundamental research. On the one hand, to maintain a relevant grey literature collection, the collection focus must shift as the needs shift. A possible short-coming of these circumstances, on the other hand, is the difficulty for the Center to pursue a certain topic and collect documents exhaustively. At the time of the Center's review, one reviewer pointed out that the Center lacks some reports in a series of report documents in a topic from the historical point of view. The Center tries as much as possible to collect all the related documents, but because of the nature of distribution of those documents, it is often impossible. For example, if policies on airport use come to be considered for revision, it is almost impossible to go back to the original policy documents and policy discussions, as well as subsequent revisions throughout the years.

The consequence of these two factors, collecting to meet changing needs and difficulties in exhaustive collecting due to the nature of how grey literature is distributed, must be understood as being both part of the strength and the limitations of a collection. It may be said that this is inevitable and stems from the nature and purpose of the collection, and therefore the grey literature collection must be considered within context the institution's whole collection.

Ephemerality of Literature

A common concept is to include the grey literature genre in so-called broader "Ephemeral literature" grouping because of its value being extremely time sensitive. Some information in grey literature is compiled and later incorporated into a more permanent form, such as annual volumes issued at a certain time of the year; some information becomes obsolete; while some others stay as they are. Providing timely information in a form of grey literature is highly

valued, but when evaluating it as a permanent collection of a library, quite different factors come into play. Under the current situation where many libraries are economically stringent, the retention and maintenance of collected grey literature simply come to be a matter of priority in terms of the library's broader collection development concerns. Grey literature must compete with other materials as to whether, or to what extent, efforts should be made to bibliographically control them as well as to expend costs for preservation measures.

Whether to absorb grey literature into the permanent library collection after fulfilling the role of providing timely information or to discard them is an issue that must be addressed. It is important to note that this is a library dependent decision. The policy of "once collected, never throw away" is no longer valid. Each library must make a wise decision, because the cost of retention is not nominal. In the Center's case, it has been determined that the document collection is to become a part of the Library of Congress' permanent collection.

Preservation of materials

Under this scenario, first the decision must be made as to how to preserve them. One of the characteristics of printed grey literature documents is that none are hard-bound, since generally speaking they are not meant to be kept and maintained. Many are issued in acidic papers and require some costly preservation measures. Furthermore, the spines of the original paper materials must be cut in order to be scanned. It costs much to render appropriate measure to each document and again, it has to compete with the rest of library materials which are of highly historical value for future research.

The Center fortunately has some funds of its own and is able to place the original printed documents in acid-free folders and in acid-free boxes and to organize them for permanent storage. However, the amount of funds is not large enough to de-acidify each document.

The possibility of microfilming can also be considered. In the case of the Center, the bulk of the documents has been digitized and the information can be easily printed out from the image database file. Therefore, it is justified to use the basic measures described above for retention of the paper documents. Furthermore, the image database itself can also serve as one of the preservation measures, although digitized images have not yet proven to be permanently secure.

Bibliographic control: the online catalog versus a special file

This brings up a second decision that must be made regarding access, i.e., bibliographical control. In the IDC's case, its bibliographic records are maintained in its own, separate database, and the structure of the record has been designed for the Center's particular needs. For example, the translated title appears first and is the one that is displayed in the listing of search results. Also, the Center has adopted the Legislative Indexing Vocabulary, developed by the Library's Congressional Research Service. If the materials are to become part of the permanent collection, the documents in printed format need to be processed, this time according to the Library's normal practice in order to be kept in a more permanent and regular book stacks. This means new catalog records must be created for each document and added to the Library's online catalog. For that purpose, ideally each record must be created following the current cataloging rules and contain various access points according to the pre-established headings. A degree of duplication of processing and cataloging may be unavoidable.

It should be noted that the reason why the Center painstakingly creates the bibliographic

records with index terms and English abstracts is that the access to collection is to be more orientated toward specific users of the type of information as well as towards a general duration of time in which it is needed. The approach is specifically tailored for the needs of the users, the majority of whom are not regular library catalog users. The Center's database which contains bibliographic records of those documents are now on the Internet as a part of the Library of Congress Web and is searchable. However, it may not be considered as a part of the permanent library collection. It is considered a special file.

Under the current economically stringent circumstances, it is practically untenable to catalog each document and therefore compromise would have to be made to catalog them on a collection level cataloging basis. This raises a new issue of whether it is worth spending that much effort when considering the effectiveness of limited access points. Each document contains specific information on specific topics and because of this, the document is valued. On the other hand, if many documents are grouped and given access by broad terms, it is difficult for users to have access to each document with its specific topic(s), unless the electronic database is maintained in some way on a permanent basis and provides cross references to the location within a group of documents.

In Japan, there is a center, an auxiliary to the Hosei University, which accepts various institutional collections of grey literature for control, maintenance and preservation. If there are several such centers organized in various parts of the world, this would be an ideal solution for the JDC and for other grey literature collections.

When the grey literature collection is part of a larger library, then how the collection is treated, given its ephemeral characteristics, must be determined by the institution. One issue discussed above touches on collection development, while another set of issues revolves around long-term treatment of the materials. Evaluation of cost versus benefit is an inevitable factor, particularly in context of the institution's planning and the realities of competing needs amongst its departments.

The second decision must be on the access, i.e., bibliographical control. The documents in printed format are to be processed according to the Library's normal practice in order to be kept in a more permanent and regular book stacks. This means that the documents must be cataloged and the bibliographic records must be added to the Library's online catalog. For that purpose, ideally each record must be created following the current cataloging rules, containing various access points according to the pre-established headings. Again, issues of human resources and costs must be seriously weighed against the benefits of having the materials and providing access to them.

Format of supplied information: printed versus electronic

From its inception, the Center's approach has been to collect printed documents. It is finding that an increasing number of documents in the grey literature category appear on web (in Japanese as well as in English). Starting in the mid 1990s, the Japanese government agencies have been making efforts that aim for wider distribution of their information, both for the Japanese public as well as the broader international public. Specifically, Japan's government documents on fairly specialized topics can be obtained in full text on the agencies' web. This is indeed a very welcome phenomenon. It would be extremely useful, if the following

information could be provided along with the reports and documents, even as a rough guideline:

1) Categories of the organization's information that are made available on web;
2) Timing of posting and how long the information will remain on the web, preferably in respect to the issuance of the printed format; and
3) Whether the information can be freely copied and disseminated for research purposes, or whether it is copyrighted

It would be helpful to the Center if we know which agencies are trying to put some information on web, and if possible, types of reports, whether full text or summary, presentation of graphs and charts, whether or not there are any restrictions, etc. Knowing this would eliminate redundant searches and the chase for printed versions on the part of the Center and its Tokyo Acquisitions Facility. At least the guidelines would give us a clue to decide whether or not to seek out the electronic version. If we know the timing of information being up on the web (in relation to the time of issuance of the printed document), that will give us good sense of locating information on the web in a timely way. This would also affect the decision of whether or not to seek printed materials, and yet enable us to be positioned to deliver the needed information to users within the time limits. Needless to say, it is also important to know how long specified information stays on the web.

This brings up a related issue. Most Japanese government agencies issue English translations of some documents. It should be noted that some are abridged versions or are short summaries. It helps to know the same information as above, such as which documents are translated, whether it is an abridged version, and how late these appear after the issuance of the original documents in Japanese. Those English translations are much sought after internationally and, therefore, Japanese government agencies must realize and respond to the needs accordingly. The author wishes to know other countries' practice.

Here is the Center's practice, which is as a related subject is certain to benefit users. When the Center provides a copy of a certain document, it sometimes finds an English summary prepared by the issuing agencies and listed on the web. We attach the print-out of those summaries to our original documents, and make a note in the bibliographic record of its availability as an attachment.

Japanese government documents are, with or without copyright notice, considered automatically copyrighted at the time of issuance. As for electronically disseminated information by the Japanese government agencies, an oral response to our inquiry was that it is open for everyone and can be freely copied. The Center's practice is that when it finds certain information on web that might be useful for research, it refer users to it but does not necessarily make a copy for the users.

Archiving

In relation to the above topic of electronically disseminated information, the author has been eagerly expressing the need for the Japanese government agencies to archive what they have put on web. While most individual organizations and companies do archive what they put on their web, there are fewer possibilities for the public to see their full texts; if the printed documents are reformatted to electronic form, the information is mainly for their organization's internal use and is generally not available for the public.

So far, the Center's finding is that the Japanese government agencies' full text documents on the web are in fact the electronic version of the printed documents, and it seems there has not yet been any documents that are issued in electronic format only. The author suspects that when such information is made available only in electronic form, it may be on Japan's special government local area network, such as the Kasumigaseki LAN and thus may not be available to the public, at least at the beginning. This is understandable because the information in those reports is, after all, issued mainly for the government's use to carry on the nation's business.

Again, the author's speculation is that under the current circumstances, provision of the electronic versions of various documents is an additional service to the public on the part of the agencies. The logic follows that since the printed version is the basis, the electronic version as a duplicate does not have to be archived. This assumption would have to change once the information is issued only in the electronic version, and it would be essential to provide archives and appropriate indexes. For the time being, whether to make a printed copy of electronically disseminated information for archival purposes for our own collection is something to be decided internally by the Center.

Conclusion

The above discussion covers a number of the current and important issues the Japan Documentation Center faces. Some are particular to the Center but others are of universal nature. In an ideal world, an information provider would be able to meet users needs by being able to obtain and provide that information in a timely way. The Center's users have proven to be diverse, with no major geographical concentration nor particular disciplines and having varying language abilities. Because of the collection being predominantly in Japanese, an approach was developed to maximize access to the collection through creating English language abstracts and indexing, and also by putting the database on the web. In order to reach users, and potential users, the Center continuously makes outreach efforts utilizing a variety of methods.

The Center collects the most recent public policy information in order to meet many researchers' need for timely information; the Center tries also to adapt its collecting focus to the political, economic and other changes in Japan in order to meet the changing research trends. Thus these characteristics of the grey literature collection bring up issues of how it would be treated, both bibliographically and physically, by the library institution of which this collection is a part, particularly on a long-term basis.

Finally, the Center has certain limitations stemming from practices of the issuing organizations. The challenge is to work with these outside factors so as to maximize those sources while minimizing redundant processing of the information.

There are yet other issues concerning grey literature, and the author wishes that they should be discussed vigorously in various forums so that better solutions could be found to benefit us all.

CONTROL AND ACCESS FOR GREY LITERATURE IN BRAZIL: A PROPOSAL.

ALMEIDA, Maria do Rosário Guimarães

ABSTRACT

Grey Literature in Brazil. Grey Literature is introduced in a constant transformation and incorporation of attributes, which makes it more and more different. There is a relation among its distinct classifications, definitions, characteristics and typology as of its importance. Grey Literature in Brazil is studied in accordance to its, visits, interviews and consultations of the database of the major Centres of Documentation and or Information in the Country. From this study, a proposal is made for the building of a Co-operative System of Grey Literature in Brazil, based on the System of Information on Grey Literature in Europe (SIGLE). System should be appropriate to our reality.

Key Words

Grey Literature;Technical Memory; Documental Memory; Institutional Memory; Technical Central File, Non-commercial literature; Non-conventional literature, Semi-published, Dark document, Invisible, Informal, Fugitive, Ephemeral.

1 INTRODUCION

The transmission of information is limited to a reduced number of people, and it is also necessary to have printed methods of comunication that may cover a large audition such as: journals, thesis, conferences records, written report, patents, databases, eletronic publications, CD-ROM, videodisc, including any message sent through the networks of communication

It is emphasized as the principal means of non convencional communication, the Grey Literature, accepted in almost all scientific areas. It is necessary to update the secondary services in order to obtain more efficient access to the documents produced. There is no doubt that Grey Literature covers a large and important amount of information, and for this reason it should be considered as the main source of information.

The term Grey Literature acquires relevant meaning and became common current use at the Seminary of York. That happened from 13 to 14 of December in York (United Kingdom). About this matter and, according to Bichteler, the expression was coined by the British librarians becoming of popular usage in 1980s.

Grey Litereture includes important informations to a considerable number of customers. It is a fact that the volume of scientific material covered, today, by this kind of literature escapes the editorials circuits, and represents an extraordinary source of valuable inedit informations. Therefore its growth is based on economic reasons that sets forth, today, a starting point towards information.

2 THEORETICAL REFERENTIAL

Grey Literature, definitely, is not a new kind of document, whose form we cannot clearly precise, nor define all its contents. Thus it is a documentation with potential existence.

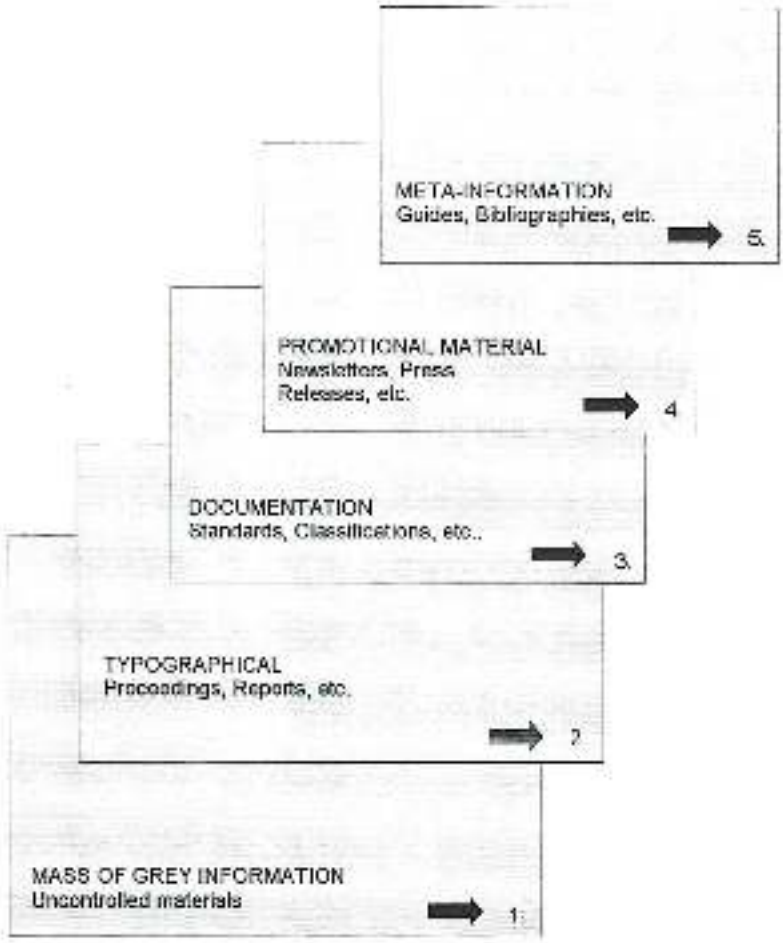
This kind of document is also known as non-conventional document, semi-published, dark document, invisible, informal, fugitive, ephemeral, and subterranean. It is characterized by a restrict circulation, and limited access and availability. This material is not submitted to any systematization process; it presents difficulties of bibliographic control and, consequently, difficult localization, reasons which make its acquisition economically expensive.

Grey Literature refers to documents of limited circulation which cannot be obtained through the conventional channels making its acquisition difficult. From this particularity, they are considered valuable documents and the most current examples are the technical and scientific report, thesis, private comunications, recommendations, official documents, etc that need a commercial diffusion.

We should mention the experiments of Leonardo da Vinci, who used to elaborate written reports about his experiments, but never worried about publishing them. So, this way, we may consider them as grey, in the sense that they have not been published. This does not mean that Grey Literature is an indifferent object, without name, without its own characteristics. It appears as something latent, not real, at the time of the greatest illustrator of the Sistina Chapel.

Grey Literature has been in use since 1920 and according to Schmidaier, mentioned by Auger, it was called *Kleinschrifttunm* - small literature. To our knowledge, the Germans had already had knowledge of this kind of literature through new publications, not available in libraries, but through the German National Bibliography since 1931.

Grey Literature is a subject that is in a constant transformation and incorporation of attributes, that makes it different from its own movements, by the inclusion of typical systems of meta-information or of abstracts. (see Figure 1)



Source: GREY Literature Network Service. NewsBriefNews.v. 5, n. 1, p. 2, 1996.
Figure.1 Evolution of the Grey Literature

According to WOOD[1] it is not convenient to define Grey Literature attending to this kind of material. Frequently, such documents are available to those interested in the time, lapse 12 and 18 months, including magazine articles as well as books.

The York's Seminary participants affirm that:

> "...the essential step of this literature class,.almost equivalent to so called non-conventional literature, is that it is not issued on the conventional circuits of commercial publications, and its access, is difficult in most cases. Furthermore, it's a literature that, in so many opportunities, has no intention of being published."[2]

[1] WOOD, N. David. Management of grey literature. In: DURANCE,C.J. (Comp.) **Management of record information**. Munich: Sauer, 1990. p.61-62.

[2] Id. Ibid. p.62

BRAUN considers that Grey Literature:

> "...is formed by scientific contribution, and also, by a considerable proportion of material whose character are informative and documental that, on frequently, has documental sources value. To BRAUN Grey Literature would be constituted by primary literature (written reports, thesis, and off prints), and also by other non-conventional materials such as: the international publications organisms, commercial chambers, banks, companies and official organisms"[3].

To WOOD, it would be:

> "...material which is not available through normal bookselling channels..."[4].

ALVAREZ-OSSORIO, defines it as "... the whole of documents, that have varying typology, and are not published through the usual channels of scientific transmission"[5].

Authors as CURRÁS and HENRICHS talk about documents:

> "... that offers a primordial information, in an non-conventional shape and in most cases, withhout specific classification because it does not have an International Stantard Serial Number (ISSN), International Standard Book Number (ISBN), Identification of Number of the Official Publications (NIPO), Legal Deposit (DL), and it cannot be acquired through the usual channels of diffusion"[6].

The term Grey Literature was defined once more in the Third International Conference on Grey Literature which took place in Luxemburg, November/1997, as follows:

> "that produced in a governmental, academical, business and industrial setting, in a print or elotronic form that is not controlled by the editorial channels"[7].

---

[3] BRAUN, H. Apud: MORENO-TORRES SANCHES, Rosario. El acceso a la literatura gris:actas de congreso y tesis doctorales. **Boletín de la ANABAD**, Madrid, v.36, n.4, p.673, oct./nov. 1986.

[4] WOOD, David N. Op. cit., p.61-62.

[5] ÁLVAREZ-OSSORIO, José R. Pérez. **Introducción a la información y documentación científica.** Madrid: Alhambra, 1988. p.29.

[6] CURRÁS, Emilia, HENRICHS, Norbert. **Estructuración de un programa de sistema experto e hipertexto para reconocimiento y localización de ciertos documentos de literatura gris.** Madrid, [1993]. Pretirada.

[7] FARACE, Dominic. Grey literature and publishing. **NewsBriefNews**, Amsterdam: TransAtlantic/ GreyNet, v.7, n.1, p.4, 1998.

The author defines Grey Literature, from all the definitions studied, as:
The set of documents notwithstanding its typology and support or eletronic print form, emitted by universities centers, research centers, enterprises, industries, public and private academical associations, without the intention of being published and which are of vital importance for the transferring of knowledge.

It constitutes common aspects of the many definitions, access difficulty, circulation and the non-comercialization of such documents.

2.1 Grey Documents

When we report to the typology and the support of the grey documents, it becomes difficult to typify them in relation to the information, or the access and known area. We conceptualize some kind of grey documents. (see Table 1)

TABLE 1 Classification of kind of the Grey Documents

CLASIFICATION OF GREY DOCUMENTS	DEFINITIONS
PROCEEDINGS	Publications, or serie of publications, that contains .the texts about expositions or communicated, orally facts transmitted in a conference, society or institution, proceeded of all the areas of knowledge.
NORMS	Documents that reflects in a consent of a community in a way to fix a model that should be practiced.
REPORTS	Documents that describes the development or the results of dominion of a tectonic or scientific research.
PATENTS	Documents that contain an invention which is garanteeded by the government by the government with a right also of product, use, sell, obtain and benefit itself for a determined number of years.
COMMERCIAL LITERATURE	Edited documents by manufactured or commercial enterprises with the finality of information or propaganda, that contain affable about products and services.
THESES (Ph.D.)	Original documents of research, presented to obtain the degree of doctor
BULLETIN	Documents have emanated from determined institutions or societies.
OFFICIAL PUBLICATIONS	Edited documents by a Government organism, with themes related by administrative,

2.2 The typology of Grey Literature

In ROSA DI CESARE's view point from the National Research Center (CNR) of the Central Library of Italy, Grey Literature is found distant from the conventional or white literature. The author uses the grey tones to classify such literature. (see Figure 2).

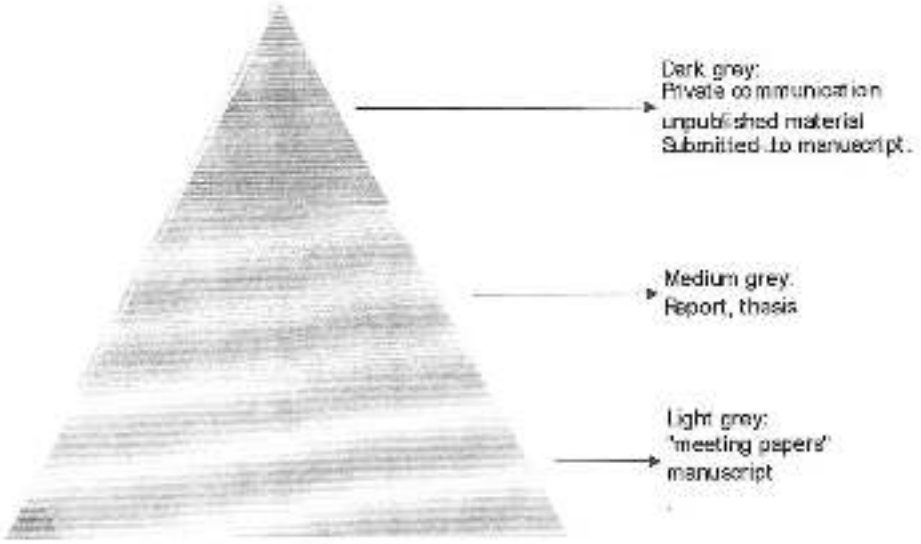


Fig.Classifications of Grey Literature by tons

2.3 Characteristics of Grey Literature

Grey literature presents specific characteristics which differentiates from the conventional ones. It has limited circulation, because it is constituded of few documents. It contains data to user which are quickly sent, and reache few readers.

The following peculiarities of this literature are pointed out:

a) variable rules of editing and production, produced through difference output mechanisms of reproduction, including the most simple;

b) ineficacious bibliographical control: the literature is not found in the banks of information agencies, that is, it is not submitted to ISSN, ISBN, LD and NIPO*. Therefore, it does not go through the commercial sources;

c) difficulty in handling: much of this information is supported on microfilms and microfiles or other kind of material, where handling is difficult and turned down for consultation;

d) conciseness: it has a less detailed character than the conventional literature, since it is not primarily for the purpose of publication. As mentioned before, in general some months later it is published by the commercial channals, in the form of magazine articles or books, because of the need to make the information's known is a necessity;

e) institutional production: as a general rule, Grey Literature is the responsibility of government institutions, universities, and research centers. Finally, it is almost always published by an institution.

What we can see, in an international scale, is a tendency in which the selection, the analysis and the treatment of the grey documents follow the standardized systematizazion of in/out registration, in the means of having a precise, national and international registration aiming at the availability and accessibility of the material without any problem on the course of going from the producer to the user. These works in a net system are made in collaboration with the researcher, the library science or information experts.

Therefore, for the treatment of the grey material, some agencies that follow the rules based on the norms of International Standard Organization (ISO), the Anglo American Cataloguing Rules (AACR2), the Classification of Dewey, of specific tesaurs in the areas of performance, descriptions of documents according to the rules of the Committee on Scientific and Technical Information (COSATI).

Another aspect to have in mind is the technological transference. There are many ways and requisites to develop. However, the more effective tranferences

* This characteristic has gone through changes since the 80's

occurs when fundamented in the interchange of personnel or in the interpersonnel transmission through the new world of micro computers and the information network, that offers new potentiality, as are the electronics publications, visualizing technics, virtual reality, among others, with the growth of the information industry.

2.4 The electronic grey document

Experts on this theme, worried in giving continuity to the cooperation among the organizations aiming at providing a better access to the world of interconnected information, that it is necessary to know well the systems of hypertext to obtain information from other sources.

The interconected information is not a linear extension of the printed information. The fast developement of an informative and interconnected world, and, specially, with the Internet has been having a giant impact on the access of information and how we deal with this new faceted, bringing great advantages to the Grey Literature, as well as one of its charateristics which is the absence of bibliographical control.

As to the volume and usage of the grey publication, these play and important role on the scientific information. Examples are mentioned at the International Conferences of Amsterdam in 1993 (GL´93), Washington, D.C. (G L'95) and Luxemburg (GL´97). The next Conference will be held on October 4 and 5, 1999 in the city of Washington, D.C., and will have the following theme: The new frontiers of the Grey Literature.

It is concluded that, with the development of Internet, new kinds of electronic grey documents are ocuring, such as: the electronic printed mail, the grey files and all the technical documentation retrived through WEB. This new kind of Grey Literature is characterized by the fact that it is a combination of documents that are not textually complete making it less understood. Through new hyperlinks, it is permitted that new kinds of Grey Literature become virtuality unlimited.

It is verified that the grey information constitutes an important part on the scientific and technical comunication. The grey documents remain unique and are rarely followed by a printed article. Part of the Grey Literature is published by well known institutions whose names garantee its quality.

With regards to the production increasingly volumous of Grey Literature, it is estimated to be 3 to 4 times greater than the conventional or commercial one, according to DI CESARE[8]

> "...Internet offers new means of scientific and technical communication, and new paradigms of communicaton are emerging.[...] the world information changing and electronic information sources, and particularly. Grey Literature are increasing. In this way, problems related to bibliographic control, will not be solved only by using new tecnologies. Therefore appropiate standards for successful identification and easy access to desidered information will be more necessary."[9]

2.5 The production of Grey Literature in Brazil

Through bibliographical search in database and Brazilian Libraries, we detected a strong tendency of typological literature that registers the production of Grey Literature such as: Technical Memory, Institutional Memory, Documental Memory or Technical Central File.

Along with this case, the necessity of establishing a terminology for the Grey Literature and the need for norms to standardize the treatment for this literature was detected.

A nationwide study of the main institutions was carried out. A list was produced per State and per Institutions within each State (see Table 2)

---

[8] GREY LITERATURE NETWORK SERVICE. NewsBriefNews. Amesterdam: TransAtlantic, v.6, .n.1, p.2,1997.

[9] DI CESARE, Rosa, SALA, Cesare. The use of grey literature in the agricultural economics field: a quantitative analysis. In:INTERNATIONAL CONFERENCE ON GREY LITERATURE, 2. , 1995, Washington. **Conference Proceedings**... Washington, D.C.: TransAtlantic/ GreyNet, 1995. p.168.

TABLE 2 Institutional sources/Brazil

INSTITUTIONS	SIGLE	CITY
National School of Public Administration Foundation	ENAP	Brasília
Agricultural Documentary Informational Coordination	CENAGRI	Brasilia
National Center of Environmental Information	CNIA	Brasilia
Science and Technology Information Brazilian Institute	IBICT	Brasilia
Documentation and Information of the Federal Senate Undersecretary	SDISF	Brasília
Latin America and Caribbean Health Science Information Center	BIREME	São Paulo
Technology Research Institution of São Paulo State	IPT	São Paulo
Technology Information Sector	CETEC	Belo Horizonte
Nuclear Information Center	CIN/CNEN	Rio de Janeiro
Information and Development Center of Petrobrás	CENPES	Rio de Janeiro

A research was performed within huge national information centers and/or documentation, pointing out the following areas of knowledge: Health, Public Administration, Law, Politics, Environment, Nuclear Energy, Agronomy, Politics and Sciences of Information, among others; the production and usage of Grey Literature in Brazil was quantified this way.

In the case of Brazil, isolated managemnet for Grey Literature were detected, but it was proved the lack of a co-ordinating organism that would unite and make known this kind of literature, in order to speed up the accessibility and availability of the grey documents.

It is worthy of comment that the only Brazilian experience on Grey Literature, is from the co-ordination of Professor Dinah Aguiar Población:

> "...since 1992, the National Association of Research and Post-graduation in Information and Library Science - ANCIB) concentrated its investigation activities in integrated projects that are composed by many subprojects supported by the National Council of Research and Development (CNPq). Two of these subprojects develop research in scientific production, considering Grey Literature as all kinds of documents created with non conventional characteristics. The thesis and events that constitutes the object of the subprojects were selected. That research had the objective identify the production and the producers characteristics of the communications that were

presented on the events taking place in Brazil..."[10]

In this sense, the problem to be studied will have an empirical object to establish a system to administrate the Grey Literature in Brazil, considering the importance of this material for the scientific and technological development. In spite of the co-operative systems, it was observed in Brazil the need for a system turned to the treatment and dissemination of Grey Literature.

Keeping this need in mind, the study intends to incentivate the Brazilian institutes to try to solve this matter, highlighting the manegment of policy based on the System for Information on Grey Literature in Europe (SIGLE).

For the establishment of such system, a proposal has been formulated and it will be detailed as follows.

3. PROPOSAL FOR CONTROL, ANALYSIS AND TREATAMENT, DISSEMINATION AND USE OF GREY LITERATURE IN BRAZIL

Grey Literature is the object of concern in the Centers of Documentation and Information, Libraries, Research Institutes and others, due to the difficulties of access found at the time of colleting the material in the institutions which study, analyse and produce this literature.

In view of the international experience of cooperation among information centers which work with the management and dissemination of Grey Literature, it is proposed that this system should follow the structure of the System for Information on Grey Literature in Europe (SIGLE); adjusting it to our reality.

[10] ASSOCIAÇÃO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO E BIBLIOTECONOMIA. **Eventos em arquivologia, Biblioteconomia, Ciência da Informação, Documentação e Museologia realizados no Brasil**: 1951-1994. São Paulo, 1994. p.1-2.

3.1 Objectives

3.1.1 General objective

Building up the structure of a co-operative system to manegment Grey Literature in Brazil is such a way as to make it possible to identify its writers; the establishment of rules and technical standards to control, the dissemination or spreading and transfering of grey information in a national and international setting.

3.1.2 Specific objectives

a) Identify the units which produce Grey Literature in Brazil.

In order to obtain such objectives, registering units wil have to be built for those who produce Grey Literature and integrate the Brazilian co-operative system, to make it possible to adminstrate. The system should incorporate some units of work located in different kinds of institutions such as:

Academies
Foundations
Laboratories
Observatories
Societies
Governement and no- government Agencies
Hospitals
Institutions
Museums
Educational Centres
Investigation Institutes
Universities

b)Establish rules and standards of bibliographical description and compensation for Grey Literature, compatible with other existing systems.

The rules and standards of bibliographical description and the compensation of Grey Literature can follow the guidline and manuals

adapted by SIGLE, internationally, well known through the descriptive cataloguing rules based on the International System of Nuclear Information (INIS).

To classify the themes, a modified version is adapted from the original classification of the Scientific and Technical Information Committee (COSATI) which belong to the United States Science and Technology Federal Council. And it is also known by the name of COSATI classification system (classification composed of letters and numbers). This classification is in use by SIGLE.

c)Establish rules for recovering, analysing, dissemination and use Grey Literature produced in Brazil.

The rules will be established to the integrated and non-integrated units of the system, detailing its recovering, treatment and dissemination of the literature in order to systematize the services related to the grey documents, so it can be sent to the coordinator unit because it is a decentralized co-operative system that acts as such decentralizing the tasks and coordinating them.

d)Create mechanisms to make possible the interface communication with similar systems of production, treatment and dissemination of Grey Literature.

The mechanisms which allows the communication's interface consist of having the electronic support and the knowledge on the system to be implanted. The recovery of the information to optimise the physical or the electronic access of the grey documents is expected from this procedure.

4 THE METHODOLOGY

4.1 Scope of the research

It is constituted of the material available in : Academies, Foundations, Laboratories, Observatories, Societies, Government and non Government Agencies, Hospitals, Institutions, Museums, Educational Centers, Institutes of investigation and Universities in Brazil

4.2 Procedure

a) contact with the several institutions which produce Grey Literature;
b) localization and retrieval of the available literature;
c) doing the collection of Grey literature;
d) identify the kind of material and support of Grey Literature;
e) develop a program of database elaboration;
f) find out the scientific production of researchers and
g) quantify, through bibliometric studies, the Grey Literature produced in Brazil.

5 SYSTEM MANAGMENT STRUCTURE

The manegment of the system can be reached by an Institute that has network experience. There is the Brazilian Institute of Scientific and Technical Information (IBICT) as an example. It is composed by other units and forms a new administrative body of the co-operative system. The folowing it is recommended:

® a National Coordinating Center, as in the IBICT case. In all co-operative centers, it is necessary to designate an organ or organs to take care of the different activities.

® a technical group is composed of representative(s) of Co-operative Centers from the National Network. This group should be in charge of the technical part of the national system, so it can gives a better eficacy of their services and products.

® an advising committee integrated by representatives from Education and Culture Ministry (MEC), National Association of Research and Post-Graduation in Information and Library Science (ANCIB), Getúlio Vargas Foundation (FGV). These are prestigious and capable institutes to guide the system in its totality.

® Co-operative Centers will be comprise all the units that produce and compose part of the National Coordinating Center.

® a GreyNet representative, being the most important network in dealing with grey material.

For the development of Grey Literature, it is necessary the promotion and dissemination of event calendars, seminaries, workshops, the incentive of electronic report editions, press reports, the creation of web pages and networks in the institutions which are working with Grey Literature, the production of secondary or meta-information sources.

Due to enournous economic crisis that is reflected upon the scientific production, the proliferation and the use of Grey Literature can be evidenced. It is constituted as a positive factor to the administration of such proposal in Brazil. That would greatly benefit students, as well as teachers, investigators and other professional fields which are growing with the advent of Internet.

Even though Grey Literature is an important source of information that can reach scientists and researchers in a quick and efficient way, the transmission of information is limited to a reduced number of people. Printing methods for the transmission of information, that can reach a great audience, such as: journals, thesis, conference records, reports, patents, rules, review on specific topics, summaries, indexes and database.

It seems necessary to discuss about the importance and effectiveness of this system. The materialization of this initiative, no doubt, will be

an action leading to what includes its recognition, speeding up its production and, consequently, the stablishment of Grey Literature.

By means, of what was stated above, we can see how opportune it is a system of such importance to Brazil, specially because of the few access to precise and concise information. The current proposal shall be implemented aiming at covering faults in the treatment and access to Grey Literature, giving concrete and significant elements to fortify the activities inherent to the researchers, librarian and library science and information professionals in Brazil.

BIBLIOGRAPHY

ALMEIDA. Maria do Rosário Guimarães. La Literatura Gris: sistemas y redes en el ámbito nacional e internacional. Una propuesta para Brasil. Madrid: Universidad Complutense de Madrid,1998. 573p. Tese (Doutora em Ciencia da Informação) – Departamento de Biblioteconomia e Documentação. Universidad Complutense de Madri. 1999. .547p.

ÁLVAREZ-OSSORIO, José R. Pérez. **Introducción a la información y documentación cientifica.** Madrid: Alhambra, 1988. 107p.

ASSOCIAÇÃO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO E BIBLIOTECONOMIA. **Eventos em arquivologia, Biblioteconomia, Ciência da Informação, Documentação e Museologia realizados no Brasil**: 1951-1994. São Paulo, 1994. 128p.

ASSOCIATION FRANÇAISE DE NORMALIZATION. **Vocabulaire de la Documentation.** Paris, 1983. [Projet de norme enregistrée].

BICHTELER, Julie, Geologist and gray literature access, use and problems. **Science & Tecnology Libraries,** v.11, p.39-50.

CARROL, Bonnie C., COTTER, Gladys A. A new generation of grey literature:the impact of advanced information technologies. In: INTERNATIONAL CO

NFERENCE ON GREY LITERATURE, 1., 1993, Amsterdam, **Conference Proceedings...** Amsterdam: TransAtlantic, 1994. 333p.

COMBEROUSSE, Martine. Les nouvelles tecnologies au service de la littérature grise. **Bulletin des bibliothèques de France,** Paris, v.40, n.2, p.51, 1995.

CURRÁS, Emilia, HENRICHS, Norbert. **Estructuración de un programa de sistema experto e hipertexto para reconocimiento y localización de ciertos documentos de literatura gris.** Madrid, [1993]. Preprint.

Di CESARE, Rosa, SALA, Cesare. The use of grey literature in the agricultural economics field: a quantitative analysis. In:INTERNATIONAL CONFERENCE ON GREY LITERATURE, 2. , Washington, 1995. **Conference Proceedings...** Washington, D.C. Transatlantic/ GreyNet, 1995. 333p.

DI CESARE, Rosa. The use of literature in the agricultural economics field: a quantitative analysis. In: INTERNATIONAL CONFERENCE ON GREY LITERATURE, 2., 1995. Washington, **Conference Proceedings...** Washington: TransAtlantic, 1995. 333p.

FARACE, Dominic. Grey Literature and publishing. **NewsBriefNews**, Amsterdam, TransAtlantic/ GreyNet, v.7, n.1, p.4, 1998.

SCHMIDMAIER, D. Ask no questions and you'll be told no lies: or how we can remove people's fear of grey literature. **Libri**, v. 36, n.2, p. 98-112, 1986. Apud: AUGER, Charles Peter **Information sources in grey literature**. London: Bowker, 1994. 170 p.,p.6

WOOD, David N., SMITH, A W. SIGLE: a model for international co-operation. **Interlending & Document Supply**, Boston Spa, v. 21, n.1, p.22, 1993.

INIST : Tracking grey literature in a changing environment

by Joachim Schopfel
Institut de l'Information Scientifique et Technique du Centre
National de Recherche Scientifique INIST/CNRS :

Ladies and gentlemen, Mr. Chairman,

I'm glad to be with you today and to represent INIST, the Institute for Scientific and Technical Information in Nancy, France. There are many French projects concerning non-conventional literature but to this day, there has been no clearly defined policies. In other words, we have to deal with " low policy, low economics ". INIST is an element in a greater national and international context. Our contribution has to be set within the rapidly changing context of documentation and information. But, even if our situation may be slightly different from other institutions, there are at least some common problems and issues.

(1) INIST

First, let me give you some basic information on our institute :

INIST, the French institute for scientific and technical information, is one of the major European scientific information centers. Our goals are to collect, process and disseminate the results of scientific research. Our current acquisition policy is to acquire all material likely to be used by the research, higher education and industry communities, regardless of subject and language.

Nevertheless, as a part of the CNRS, the French national center for scientific research, our most important mission is to work for and with the French scientific community i.e. the main public research organizations, universities, industrial R & D units and so on.

Our principal activities are database production (Pascal and Francis), document supply and scientometries research. Our databases contain nearly 15 million bibliographic records in science, technology, medicine, social sciences and humanities. Our library covers about 24 km of shelf space. The major part of INIST document resources are serial publications. About 9,000 serial titles are currently received by INIST, with another 14,000 titles in our holdings. 400,000 articles are digitized every year.

Copies from all these documents can be ordered from our online service platform (www.inist.fr). The documents which INIST does not possess are supplied via a network of more than 100 European referral libraries.

(2) Grey literature at INIST

Our " non-conventional " collections include about 100,000 doctoral dissertations, 60,000 conference proceedings and papers, 56,000 scientific reports. Currently, we receive more than 10,000 grey documents each year.

In the changing context of database production, e-journals and the evolution of the web, INIST is engaged in a collective reflection on its acquisition policy, services and products. I would like to outline some of the perspectives for non-conventional literature.

1. **Electronic documents:** Until now, we have limited our acquisitions to print documents, spending much time on traditional library activities like cataloging. Today, most non-conventional literature exists in electronic format. Therefore, we are re-orienting our grey acquisition policy to an " all-electronic " policy. Examples : some weeks ago, INIST started to encourage its partners to submit report or dissertation files by email ; we are working on a web portal in order to simplify submission of e-documents ; there are national projects for electronic dissertations and theses in which INIST could play a major role. And of course, we are engaged in " metadata discussions ".
2. **Access providing:** Until now, we have owned in our collections most of the documents needed for our databases and document supply. Today, this approach has reached its limits, as the example of NTIS shows. There are too many documents, on the web and elsewhere, for us to deal with, considering our limited financial and technical resources. In the era of the Internet, we do not need to have all the documents in our holdings. Instead, we have to know where to get them. This means

a systematic search on the web and a direct relationship with the main producers of grey literature. Here, I hope the GL compendium project will be helpful.

3. **Consulting and assistance :** Down the road, this also means assisting grey literature producers, in order to facilitate the search for this special type of documents. Two years ago, INIST became the French national ISRN agency. The International Standard Report Number contributes to the identification of report producers ; it facilitates document management and identification in databases and catalogs; it may also contribute to the creation of URNs. Last but not least, we are contributing to the standardization of report presentation.
4. **National public research :** I said that we did not need to have all the documents in our holdings. That is true. But there is one exception. INIST plays a specific role for French public research. We produce the science and technology part of the national thesis database, we collect scientific reports from public administrations and other research organizations, and proceedings of the major French conferences. La Documentation Française proposes an access to the administrative reports via an online database; currently, we are exploring the opportunity of a public STM report database on the web, together with other partners.
5. **CNRS database:** Today, we are also involved in a project of a CNRS database of the publications from the CNRS research units. Tomorrow, this may include pre-prints and other unpublished papers. Creating and providing a sort of web gate to these documents is a part of our role as keeper of our " documentary heritage ". In the future, such a database will also be a kind of guarantee against " dirty-grey documents ".
6. **Links:** It may be possible to link this documentary database to information on French experts and research units in science and technology.
7. **European level:** Facilitating access to French research production includes participation in the European EAGLE network. INIST is the French input center for the SIGLE database. Next time, I think we should discuss the future of this database in the context of the internet, the full text, and free access.
8. **Education and training:** The changing landscape of information and documentation calls for new skills and abilities. INIST is implementing an expensive program for the training and development of its workforce, together with ENSSIB, the major French

Library Science school at Lyon, and other institutions. INIST participates in the higher education of documentation specialists; each year we welcome trainees for research and other projects, we acquire French dissertations in information science, and we are creating at INIST a kind of " library-laboratory " linking higher education, user training, documentation, informatics and research. Assuming its role in the national research context, INIST will organize in 2000 the 2nd conference for the CNRS documentation specialists where grey literature will certainly play an important role.

To conclude, I would like to insist on the dual function of INIST in knowledge mediation :

- on the one hand, we provide and facilitate access to grey literature wherever it may be, on the web or elsewhere ;
- on the other hand, by archiving French grey literature, we act as keeper of a special document heritage,

We all know that the lifetime of non-conventional literature is rather short. In the era of electronic documents and in a culture of ephemerality, it may be necessary to save at least some of these documents from oblivion and extinction.

The Competitive Advantage for the New Millennium: Knowledge Management

Atefeh S. McCampbell, Ph.D.

Linda Moorhead Clare

Scott Howard Gitters

Florida Institute of Technology
Building 5442
Aberdeen Proving Ground
Aberdeen, Maryland 21005

ABSTRACT

The following paper defines the newly emerging concept of knowledge management. The topics presented include: principles and practices of knowledge management and the competitive advantage gained by industries employing this discipline. The technical applications and tools currently utilized within this discipline will also be discussed.

A survey was conducted to quantify the level of corporate awareness of knowledge management, the effectiveness of knowledge management practices on leveraging intellectual capital, the measurement of knowledge management initiatives and the technology employed to achieve those initiatives. The survey was distributed to the following industries: manufacturing, transportation, communication, financial services, retail trade, healthcare, government, and business and professional service organizations. The results of the survey indicated a majority of companies are employing knowledge management practices with positive effects on profitability, as well as, customer satisfaction and quality of service. According to the survey, technology solutions, inclusive of, smart documents and networks, are playing a pivotal role in the successful implementation of knowledge management strategies.

Knowledge Management Case Studies are also included on the following firms: Teltech, Ernst & Young, Microsoft, and Hewlett Packard. These case studies are analyzed to determine the competitive impact of knowledge management strategies.

Finally, conclusions are drawn regarding the strategic direction of this new discipline and its effect on competition, productivity and quality for the business of tomorrow.

INTRODUCTION

Knowledge Management-the Competitive Advantage

The study of knowledge management evolved from the need for companies to manage resources more effectively in a hyper-competitive, global economy. Ikujiro Nonaka stated in The Knowledge Creating Company, "In an economy where the only certainty is uncertainty, the one sure source of lasting competitive advantage is knowledge...Successful companies are those that consistently create new knowledge, disseminate it widely throughout the organization, and quickly embody it in new technologies and products" (Nonaka, 1991).

The American Productivity & Quality Center (APQC), a nonprofit education and research organization which fostered the creation of the Malcolm Baldridge National Quality Award , defines knowledge management as "the strategies and processes of identifying, capturing and leveraging knowledge" to enhance competitiveness (Manasco, 1996).

As the intellectual assets of companies become more complex, the ability to inspire employees to invest in knowledge will become more, not less, important. (Leonard-Barton et al., 1995).

According to Gartner Group findings for 1998, implementation of knowledge management systems has begun, and is currently being deployed by most large companies. One third of Fortune 1000 companies are now including knowledge management initiatives in their 1999 plans (Smalley-Bowen, et al. 1999).

According to a study of successful knowledge management projects , knowledge management was at least partially responsible for a major transformation of one large consulting firm. The transformation was extensive in terms of a marked improvement in financial results for the firm while engaged in knowledge management practices. Line consultants drew heavily from the firm's centralized knowledge centers, accessing previous presentations to other clients, process and system design specifications, work plans and other project-oriented collateral and artifacts. Senior managers described knowledge management as the core of the consulting strategy and the concept was pervasive in the company's internal and external documents (Davenport, et al , 1998).

According to Yogesh Malhotra, author of Deciphering the Knowledge Management Hype (1998), "Knowledge Management caters to the critical issues of organizational adaptation, survival and competence in face of increasingly discontinuous environmental

change...Essentially, it embodies organizational processes that seek synergistic combination of data and information processing capacity of information technologies, and the creative and innovative capacity of human beings" (Malhotra, 1998).

The new world of knowledge-based industries is distinguished by its emphasis on precognition and adaptation, in contrast to the traditional emphasis on optimization based on prediction. Arthur suggests that the new world of knowledge-based business is characterized by a continuous redefinition of organizational goals, purposes, and an organization's "way of doing things". This new business environment is characterized by radical and discontinuous change and demands anticipatory responses from organization members who need to carry out the mandate of a faster cycle of knowledge creation and action based on this new knowledge (Arthur, 1994).

With companies increasingly being viewed as mechanisms for coordinating knowledge, the mechanisms that companies utilize to integrate the specialist knowledge of individual members into goods and services become increasingly important (Grant, 1996).

A recently published 1999 Executive Briefing, eBusiness Imperative, by The Concours Group concludes that the ability to deploy knowledge is the principal differentiator in the Information Age. The emerging 21st century information age enterprise is anything but linear, rather growing outwardly. Businesses are competing on the basis of business models with brief half-lives. Products and services are differentiated by knowledge and intellectual capital. It is clear, in the move from the industrial age to the technological age, intellectual property or knowledge impacts profitability.

Principles of Knowledge Management

Knowledge Management is an emerging discipline which is a dynamic process. According to a study conducted by American Productivity & Quality Center's International Bench-marking Clearinghouse, of eleven organizations participating in the arena of knowledge management: *Arthur Andersen; Chevron Corporation; Dow Chemical Company; Hughes Space & Communications; Kaiser Permanente; Price Waterhouse; Sequent Computer Systems; Skandia AFS; Texas Instruments; USAA; and National Security Agency*, the following was recognized:

1. The practice is dynamic, but often starts by creating, finding and collecting internal knowledge and best practices.
2. Sharing and understanding those practices so they can be used.

3. Adapting and applying those practices to new situations

Also noted in this report was the direct reflection of top down management enabling, encouraging or undermining the process. Strong professional ethic and pride, supported by skills in cross- functional teaming, cultural support such as reward and recognition programs and CEO's as visible advocates for the sharing of knowledge and best practices, resulted in the promotion of success.

The most widespread strategy among the participants in the study was the transfer of knowledge and best practices in order to improve operations or to embed them in products and services. This includes systems and practices to obtain, organize, restructure, warehouse or memorize and distribute knowledge. Most of the firms also underscored the importance of teams, relationships, and networks.

These three elements are the basis for the effective transfer of knowledge. The study showed that the participating firms had taken part in a number of approaches to encourage collaborative knowledge transfer (Manasco, 1996).

Highly satisfied employees drives customer capital creation. Research has proven that skilled and motivated employees correlates with increased customer satisfaction and loyalty which further correlates to increases in market share and profits (Stewart, 1997). The team, relationship and network approach encourages employee satisfaction and motivation.

Conventional job descriptions mask the ways in which people actually work, but they also ignore the learning and innovation that are generated within informal communities of practice. The key to harnessing innovative energy lies in the effective design of organizations and the ways communities are linked (Seely-Brown et al, 1991).

The more common type of success in knowledge management involves operational improvements limited to a particular process or function. According to a study reported in the Sloan Management Review, the following factors lead to knowledge management project success:

- Link to economic performance or industry value
- Technical and organizational infrastructure
- Standard, flexible knowledge structure
- Knowledge-friendly culture
- Clear purpose and language
- Change in motivational practices
- Multiple channels for knowledge transfer
- Senior management support

Knowledge management is one of many components of good management. Sound planning, savvy marketing, high-quality products and services, attention to customers, efficient structuring of work and thoughtful management of an organization's resources are all critical to compete in today's marketplace. Knowledge management may help create the competitive edge in today's global environment (Davenport, et al, 1998).

Several recent studies report results that solidly show knowledge management initiatives growing in large corporations, up more that 100 percent from 1998, encompassing more than half of all major enterprises (Petch, 1999).

Methodology

A total of one hundred questionnaires consisting of seven questions were distributed to the following industries: manufacturing, transportation, communication, finanacial services, retail trade, healthcare, government and business and professional service organizations. The purpose of the survey was to quantify the level of corporate awareness of knowledge management, the effectiveness of knowledge management practices on leveraging intellectual capital, the measurement of knowledge management initiatives and the technology employed to achieve those initiatives. The survey also included demographic questions regarding industry type and respondents position within the organization. Analysis is conducted to determine existing and projected trends.

Summary of Findings

According to the survey, if knowledge management is defined as the strategies and processes to capture, leverage and disseminate tangible and intangible knowledge, 74% of the companies surveyed participate in the practice of knowledge management. The data indicated that 26% of the respondents did not participate in the practice of knowledge management.

According to the 74% of respondents who participated in the practice of knowledge management initiatives, 100% responded that knowledge management increased quality and productivity within their respective organizations. This is a clear indication of the highly successful outcome of knowledge management strategies across industries. Also, the improvement of quality and productivity within these organizations, fosters the competitive edge in the global business arena of today.

Effectively leveraging intellectual capital proposes a unique challenge. The survey indicated that knowledge management practices had varying effects on this issue. The data reports that 10% of respondents indicated that knowledge management practices had *no effect*, 28% of respondents indicated *some effect*, 39% of respondents indicated that their organizations knowledge management practices were *effective* at leveraging intellectual capital, while 18% of respondents indicated *very effective* and 5% indicated *extremely effective*. An analysis of these statistics results in the finding that over 60% of those surveyed, believe that their organizations knowledge management practices are, at the very least, effective at leveraging intellectual capital. This indicates a positive trend within the discipline to capture and leverage tacit knowledge. As knowledge management practices improve within an organization, one might predict an increase in this trend.

The survey raised the question of how organizations measure the success of knowledge management initiatives. The data reports that 13% of the respondents utilize ROI as the measurement tool. 36% of those surveyed, indicated Product/Service Innovation was the measurement indicator. Matching that percentage rate of 36% was Improved Profitability, while 29% indicated a variety of measurement tools inclusive of:

- Quality of Service
- Production Goals
- Customer Satisfaction
- Minimized Waste and Work Redundancy
- Employee Turnover
- Performance Improvement
- Cost per Access Line
- Process Cost
- Market share increase

3% of the respondents indicated no form of measurement was utilized. These findings indicate the trend to utilize not only measurement tools which reflect on the bottom line, but also, innovative concepts to measure the value of knowledge management implementation. In the opinion of the authors, further refinement of measurement techniques will result in a larger number of organizations participating in knowledge management practices.

The use of technology solutions to employ knowledge management practices is widespread across all industries. The survey indicated the following results regarding technology employed: Knowledge Process Mapping 15%, Smart Documents 15%, and

Networks 68%. 18% of respondents indicated the use of other forms of technology to implement knowledge management inclusive of:

- Technology-based Training Sessions
- Intra/Internet Knowledge Base for exchange of solutions by both client and employee

The use of networks, according to the survey results, is an integral facet of the successful capture and exchange of knowledge, not only internally, but externally. Respondents indicated the use of networks for research, strategy documents, training, flowcharts and organization-wide Smart document applications. As security is enhanced on the internet and firewalls improve, one might predict increased usage of internet based knowledge management data repositories world-wide.

Following are the demographic results of the survey by industry category and position within the organization:

Industry: Manufacturing 5%; Transportation 2%; Communication 2%; Financial Services 5%; Retail Trade 7%; Healthcare 39%; Business and Professional Services 23%; and Other (inclusive of Government, Construction, and Publishing) 13%.

Position: Chief Executive Officer 5%; Chief Information Officer 2%; Management 65% and Staff 26%.

Knowledge Management and Information Technology

As shown in the results of the survey, a key enabler for the implementation of knowledge management is information technology. IT's role is emerging as an integrator of communications technology, rather than solely a keeper of information. The critical role for IT lies in its ability to support communication, collaboration, and those searching for knowledge and information, not static repositories of "best practices", according to the American Productivity and Quality Center International Benchmarking Clearinghouse study (Manasco, 1996).

Karl Erik Sveiby, the author of The New Organizational Wealth: Managing and Measuring Knowledge-Based Assets, contends that the confusion between knowledge and information has caused managers to sink billions of dollars in information technology ventures that have yielded marginal results. Sveiby asserts that business managers need to realize that unlike information, knowledge is embedded in people, and knowledge creation occurs in the process of social interaction (Sveiby, 1997).

Knowledge can be managed as a strategic resource by focusing on the value of information technologies in exploiting organizational knowledge (Earl, 1994).

New design precepts are emerging as we accumulate experience. Participants interacting with the computer need to be able to interpret what's happening on the screen in terms of the real world. A well-designed learning computer lab leaves its participants with the skill to communicate without depending on a computer (Sense, 1994).

Innovative information technology is applied to support and facilitate organizational learning processes involved in the development and exchange of organizational members underlying opinions, assumptions and interpretations of the environment (Hine, M. et al, 1998).

There are four constructs linked to organizational learning: knowledge acquisition, information distribution, information interpretation and organizational memory (Huber, 1991).

The current business arena requires the need for a vast and complex interpretation of information outputs generated by computer systems. This variety is necessary to encompass the multiple global views of an unpredictable future. Brook Manville, Director of Knowledge Management at McKinsey & Company in Boston, views the implementation of these issues in terms of the shift from the traditional emphasis on transaction processing, integrated logistics, and workflows to systems that support competencies for communication building, people networks, and on the job learning (Malhotra,1998).

A lot of intellectual capital resides in the minds of IT workers. Companies such as Andersen Consulting, Ford, and Monsanto encourage employees to put "tacit" knowledge, the know-how in their heads, into "explicit" form, such as written reports or video presentations. This captured knowledge is then stored in repositories such as databases and intranet Web servers, all of which users can search. The ownership of intellectual capital has become a priority for many organizations, with several recent court cases having thrust this issue to the forefront. Wal-Mart Stores, Inc. filed a lawsuit last year against Amazon.com that alleges theft of trade secrets. Wal-Mart claims the defendants carefully recruited members of its staff who, as a group, have the knowledge to replicate information systems and business processes Wal-Mart has spent years and financial resources developing. A recent Supreme Court action affirmed that business methods linked with software can be patented. Savvy technology managers now use a combination of legal defenses and common sense to protect their companies' most innovative IT systems and to retain their talented information technology staff (Hibbard, 1999).

Increasing the number of new patents, accelerating to retirement obsolescence of patentable ideas, decreasing the time to market for new goods, eliminating waste by consistently applying best practices and investment in information systems, training and education are all practicess which add value and measurement of knowledge management (Edvinsson, et al, 1997).

Despite initial fragmentation, knowledge management technologies are quickly evolving and converging, spurred by requirements of top global organizations, attention by consultants and integrators and efforts by pioneering vendors (Mantelman, 1999).

Most knowledge applications have evolved from pre-existing types for managing documents, databases, workgroups and customers. Most explicit knowledge lives in documents, Web or PC files, paper or scanned images. Smart systems do more than track or store information. They help organizations manage content-in the context of what people know and need to know. These documents are now known as "Smart Documents".

Knowledge management requires collaborative group support. Software tools are available which let people build communities and take part in virtual teams; brainstorm, develop, present and deliver knowledge; share documents or applications; discuss and manage projects; and coordinate activities.

Teltech: The Business of Knowledge Management Case Study

Teltech was formed in 1984 and offers instructive lessons to companies wishing to better manage their knowledge and information assets. The company has built a successful business on helping companies gain access to external technical expertise and information.. Teltech's vision includes the view that purely technical approaches to information and knowledge provision will rarely add value such as the incorporation of the human element. The knowledge analyst assists in defining what information is desired, clarifies concepts and terms, interprets search results and knows when and where to seek further information.

A key premise of Teltech's business model is that people are not only guides to information , but important repositories of expertise. A key focus of the firm's product line development has been the development of software providing an integrated view of sources of information on a particular topic known as an integrated source map. It will present virtually all information that a customer may want on a particular topic in a matrixed environment (Davenport, 1997).

Knowledge Management at Ernst & Young

Ernst & Young is one of the "Big Six" professional services firms which traditionally offered audit, tax and management consulting. In 1993, the operational vision for the consulting practice involved five key processes: sales, service delivery, people and knowledge. By 1995, the knowledge strategy had been formalized into an approach entitled *Accelerated Solutions Environment* which involved the rapid application of knowledge, models and approaches to client situations.

The firm created a position of Chief Knowledge Officer with the following responsibilities: oversight of processes and technologies related to knowledge. Also, several high level committees were created to examine the direction of knowledge management within the firm.

Due to the scope and geographical distribution of the Ernst & Young knowledge base and its users, technology had to be used as an enabler for its knowledge workers. Lotus Notes was selected as the primary technological platform for capturing and disseminating internal knowledge. Thousands of databases were in use for network and focus groups, as well as, key documents. In 1996, exploration began for use of a Web based intranet. Infrastructure investment occurred which benefited knowledge management and technology (Davenport, 1997).

Microsoft Knowledge Management Case Study

Microsoft's IT group has invested time and resources into identifying and maintaining knowledge competencies. Microsoft hired a program manager to take on the issue of knowledge competencies. The manager's goal was to create an on-line competency profile for jobs and employees. The project was called "Skills Planning Development" or "SPUD". SPUD is not focused on entry level competencies, but rather competencies needed to stay on the leading edge of the workplace. The five major components of the project were:

- Development of a structure of competency types and levels.
- Defining the competencies required for particular jobs
- Rating the performance of individual employees in particular jobs based on competencies.
- Implementing the competencies in an online system.

- Linkage of the competency model to learning offerings.

SPUD involved building an on-line system that contained the competency structure, the job rating system and ratings database, and the competency levels for employees. The system had a Web front end for easy access through Microsoft's intranet.

Implementation of these knowledge management strategies proceeded on an international level involving a large number of employees and utilizing a cross-section of all job types. (Davenport, 1997).

Knowledge Management at Hewlett Packard

In mid-1995 there were several knowledge management initiatives at Hewlett Packard. A series of workshops on the topic were facilitated by the senior leadership of the firm. The goal of the workshops was to bring together a diverse group of people who were already sharing knowledge or were interested in the topic. The workshop objectives were:

- Sharing of knowledge through informal networking
- Establish a common language and management framework for knowledge management

As a result of the workshops, HP discovered twenty sites for knowledge sharing. One example was a training database ; an on-line knowledge database that contained training issues, topics and techniques. Other Hewlett Packard knowledge sharing operations included a guide to human knowledge within HP's laboratories. The goal of this database was to identify expert profiles, or guides to the backgrounds and expertise of individuals who were knowledgeable on particular topics. Hewlett Packard Product Processes organization developed three knowledge sharing products: competition information, research information, and marketing intelligence (Davenport, 1997).

CONCLUSIONS and RECOMMENDATIONS

The results of the survey indicate an integral part of enterprise activity is knowledge management. From massive intranet based repositories to corporate-wide smart document applications, a constantly changing environment seeks to optimize knowledge management practices and leverage intellectual capital.

In each of the case studies, it is clear that if knowledge is to be leveraged, it must first be categorized. The thesaurus-based, matrixed approach used by Teltech may be the "best practice" since knowledge is usually communicated and sought in words.

Substantial progress has been made in Ernst & Young's approaches to knowledge management, but several challenges remain, inclusive of: embedding knowledge in an ever changing technological environment; maintaining support with increased usage, and buy-in from front line staff.

Also, the task of measurement of return on investment is difficult at best. A 'dashboard" was created assessing such topics as value delivered, reusable content created, thought leadership, presence of subject matter expertise and state of networking environment. It was noted that "some level of faith" was required to fully justify knowledge management investments. The revenues in the US consulting practice in 1995 rose by 44% which surpassed all targets ((Davenport, 1997).

The survey reflected a small percentage of industries utilized ROI as a measurement tool. Creative measurement tools are being utilized such as access by line, employee turnover, process cost, performance improvement and product/service innovation.

Standards bodies, ranging from the International Accounting Standards Committee to The Conference Board are making important progress on performance measurement methodologies (Petch, 1999). Performance measurement will be a key issue in knowledge management initiatives since there is little precedent upon which to establish ROI. Balanced scorecard measures inclusive of financial, customer, internal processes, and innovation have been proposed for knowledge capital (Kaplan R., et al, 1996).

As an emerging and dynamic discipline, the creation of a standard measurement of knowledge management reflected on the balance sheets, is still in the formation stage. Once achieved, the result will be a rapid response from global business leaders, to implement knowledge management "best practices" in order to remain competitive.

In the case of Microsoft, it is evident that for a project to be successful it would require the involvement and teamwork by everyone in the organization. This is a model which is clearly valid in business practices, but essential to the discipline of knowledge management. The Microsoft case study reveals security issues pertaining to building a competency database which included access rights based on the nature of the data; "people data".

The survey results showed 68% of respondents utilized networks as a technological tool for implementation of knowledge management practices. The continual development of adequate security for knowledge-based network applications is critical to this emerging discipline. In order to effectively leverage intellectual capital through technology based solutions, the issue of security must primarily be addressed.

Hewlett Packard has recognized that there are several knowledge sharing initiatives, such as training techniques, marketing intelligence, etc. which can advance their position in a global marketplace.. It is clear that the level of resource commitment in gathering data and dissemination of data is high, but the gain is important to remain competitive in a global marketplace.

Analysis of these conclusions, implies that important aspects of this discipline are regarded in similar manner throughout the intellectual and business community. These aspects include a direct correlation between the implementation of knowledge management "best practices" and quality improvement, as well as, improved productivity.

Hedlund and Nonaka reinforce the important idea that not only the success, but also, the very survival of organizations will depend, in large part, on how well they create, transfer and exploit their knowledge resources (Hedlund, G. et al, 1993).

Knowledge management implores you to look at informal networks and protocols, any and all approaches to sharing experiences and know-how, as well as any and all cultural, technological and personal elements that spur creativity and innovation in response to changing stimuli. The effects of knowledge management on quality and productivity are evident by an internal and external awareness of collective strength and the ability to respond and instantly organize to meet market demands and opportunities. The competitive advantage for the new millennium is knowledge management.

Literature Review

Arthur, B. (1994), *Increasing Returns and Path Dependence in the Economy*, University of Michigan Press, Ann Arbor.

Davenport, T., DeLong, D. and Beers, M. (1998) " Successful knowledge management projects", *Sloan Management Review*, Vol 39 No 2, pp. 43-57.

Davenport, T. (1997) "Knowledge management case studies",Graduate School of Business, University of Texas at Austin URL: www.bus.utexas.edu/kman.html

Earl, M. (1994) "Knowledge as strategy: reflections on Skandia International and Shorko films", *Strategic Information Systems: A European Perspective*, John Wiley & Sons, Chichester.

Edvisson, L. and Malone, M. (1997), *Intellectual Capital*, Harper, New York.

Grant, R. (1996) "Toward a knowledge-based theory of the firm", *Strategic Management Journal* Vol 17, pp. 109-122.

Hedlung, G. and Nonaka, I. (1993), "Models of knowledge management in the West and Japan", *In Implementing Strategic Process: Change, Learning and Cooperation*, Blackwell, Oxford.

Hibbard, J. (1999) " Protecting innovation", *Information Week*, Issue 722, pp. 51-65.

Hine, M. and Goul, M. (1998) "The design, development, and validation of a knowledge-based organizational learning support system", *Journal of Management Information Systems*, Vol 15 No 2, pp. 119-152.

Huber, G. (1991) "Organizational learning: the contributing processes and the literatures, *Organization Science*, Vol 2 No 1, pp. 88-115.

Kaplan, R. and Norton, D. (1996) *The Balanced Scorecard*, Harvard Business School Press, Boston.

Leonard-Barton, D. (1995), *Wellsprings of Knowledge*, Harvard Business School Press, Boston.

Malhotra, Y. (1998) "Deciphering the knowledge management hype", *Journal for Quality & Participation*, Vol 21 No 4, pp. 58-60.

Manasco, B. (1996) "Leading firms develop knowledge strategies", *Knowledge Inc.* Vol 1 No 6, pp. 26-29.

Mantelman, L. (1999) " Technologies of knowledge management", *Knowledge Management*, Vol 2 No 2, p. 67.

Nonaka, I. (1991) "The knowledge-creating company", *Harvard Business Review*, Vol 6 No 8, pp. 96-104.

Petch, G. (1999) "Memo to executives: get out of swann's way", *Knowledge Management*, Vol 2 No 3, p.16.

Seely-Brown, J. and Duguid, P. (1991) "Organizational learning and communities of practice: toward a unified view of working, learning and innovation", *Organization Science*, Vol 2 No 1, pp. 40-57.

Senge, P. (1994), *The Fifth Discipline Fieldbook: Strategies and Tools for Building a Learning Organization*, Doubleday, New York.

Smalley Bowen, T. and Scannell, E. (1999) "Making knowledge mean business", *Info World*, Vol 21 No 8, pp. 69-70.

Stewart, T. (1997), *Intellectual Capital: The New Wealth of Organizations*, Doubleday, New York.

Sveiby, K. (1997), *The New Organizational Wealth: Managing and Measuring Knowledge-Based Assets*, Berrett Koehler, San Francisco.

Archiving Electronic Grey Literature of Korean Academic Societies in Science and Technology

Cho, Hyun-Yang ; Choi, Soon-Heui ; Choi, Sung-Pil ; Shim, Hyung-Sup
Korea Research and Development Information Center (KORDIC)

Abstract

Korean academic societies in the field of science and technology are the major sources for producing STI (scientific and technological information) in Korea. They produce various kinds of printed materials including journals, proceedings, preprints, seminar/workshop materials of the study group in a society, and so on. While journals and proceedings are published and distributed to members via conventional routes, however, other materials are not. The secretariat of societies can play important roles on collecting grey literature produced by supporting and controlling academic activities in each society. However, there are only few Korean academic societies to create databases by themselves on materials generated from academic activities.

KORDIC(Korea R&D Information Center) is responsible for archiving and distributing technical reports, generated by the result of government-supported researches. KORDIC also implements the project to support Korean academic societies to cultivate their own capability of creating databases and providing information services. 130 Korean academic societies in science and technology are participating at this project. The project consists of three different parts. The first part of the project is to create bibliographic and full-text databases on all the products of academic societies. The second part is to construct Internet homepage for each academic society. And the last is to assist to provide users with information service via Internet. KORDIC in lieu of attending academic societies collects all grey literature from members, constructs various types of databases on all the materials collected, and provides users with information services.

In this paper we will introduce the present status of archiving electronic grey literature generated by Korean academic societies and its services to users via Internet. We plan to create English version of electronic grey literature database produced from Korean academic societies. Although this project is initial stage, construction of databases on electronic grey literature and the study of efficient information services will be continued.

INTRODUCTION

Throughout all over the world, academic societies in the field of science and technology have been the important producer of current scientific and technological information that is produced by academic society members by national and international research activity and interchange.

Because of newest research developing related information of every field of researcher announced through research journals of relating society, proceeding, seminar, workshop as of first priority, we are in need of information service that user can reach out for the information that has been produced and circulated faster and easier.

In Korea, with exception of huge academic societies, most of academic societies are struggling to bringing it to life and barring it alive due to the lack of finance and uncertain establishment of gathering, manufacturing information, and service system, even though they realize the importance of information services.

As of today, in Korea, there are about 350 academic societies in science and technology field and due to division of studies and developing of new era, more of detailed new academic societies are appearing. At this time, about 10 Korean academic journals are being reported in SCI expended and also making a constant quality progress. This is the time that we are in need of constructing databases of all types of information that the societies have produced. To compromise this situation, KORDIC (Korea Research & Development Information Center) is starting society informatization project about academic societies in the field of science and technology in Korea. Society informatization project is to construct databases of information that each society produces, and to provide financial and technical supports for information service through an Internet homepage. In this study, we will discover the category and characteristics of societies' producing and circulating research information including grey literature and analyze the service KORDIC is proceeding and inquiring into more affective society's information service idea and system for later on.

WHAT DOES KORDIC DO?

KORDIC

KORDIC is a non-profit government research institute under the supervision of Prime minister of Korea that is built to suit the purpose of gathering and manufacturing the scientific and technological information and constructing it to databases and then provide service it through Internet. In Korea, non-profit sci-tech laboratories are established in

specialized fields, with the support from the Korean government, and as the laboratory that offers effective information service about all fields of science and technology KORDIC was established. Since 1993, after the establishment, with constant database construction we are managing 14 database service throughout several types of scientific and technological information and hold around 3 millions of data.

KORDIC's database structure forms mostly with bibliographic databases, facts and statistics databases, in depth and analytical information databases. Most of databases consist of Korea's own contents about information occurred in Korea.

KORDIC's bibliographic databases can be divide into two different types; one is the database that contains bibliographic information, such fields as title, author, descriptor, abstract, etc. The other is the database that includes both bibliographic and full-text information (image file or text file form). Informatization for Korean Academic Society is to meet the second example.

Informatization Project for Korean Academic Society

In Korea, the actual academic society informatization has started with manufacturing and distributing CD-ROM title. And lately, 3-4 years ago, since the Internet became an active method of academic information circulation we have launched informatization through Internet. Since the WWW which made Internet popular was develop among few physicists and biologists, also here in Korea, informatization has been tried in this field at the first time. At the present time, the some societies run their own information system. These are as follows; The Korean Chemical Society(http://www.kcsnet.or.kr)[1], The Korean physical Society (http://mulli.kps.or.kr)[2], The Biochemical Society of the Republic of Korea (http://www.biochem.or.kr)[3], Korean Mathematical Society (http://www.kms .or.kr)[4]

The Korean Chemical Society has been chosen by Korea Science Foundation as a trial society to support sci-tech society informatization running the KCSnet that serves through Internet. [5] The details of the informatization project are to build a homepage on Internet and construct various databases for information services to target users.

The academic society informatization has huge difference among societies. It differs from the knowledge and enthusiasm of the president in each society. And also, there has been a case that started with related foreign society's request.

Recently, 2-3 years ago, national organizations that support the academic society have set the informatization measure as the important evaluation measure, and relating foreign societies have asked the informatization measurement for International Corporation. These circumstances have come to maturity that the academic society has to be informationized. Most of academic societies in sci-tech field have been asked building a homepage on the Internet and creating a database on all types of information, produced by societies.

With these understanding, KORDIC 's society informatization area has been set, and main contents was set to build an Internet homepage that each society asks and to create bibliographic and full-text databases.

The Selection of Academic Societies for Informatization Project

The actual goal of society informatization project that KORDIC carries out is to build a information system, so that members of the all societies and information users can access the full-text database and use it in their own computer through the network. To do this possible we have decided to support the academic society with the respect to building the society academic information database, supporting for information service system, computing facilities, and cyberspace on the network. In the given space, each society builds its' homepage and manages it by itself. Data structure for the database of journal articles and current society's data runs after MARC, the standard format and supports basic computer and communication facilities. To make the goal come to life more effectively, we have set trial societies according to their capabilities of producing academic information. As of this time there are about 160 societies are participating.

Table 1 shows the number of participating societies and a representative society in different subject category.

<Table 1> Number of Participating Societies

Details / Subject Category	Number of Participating Societies	Representative Society
Basic Sciences	31	Korea Environmental Sciences Society
Life Sciences	9	The Ecological Society of Korea
Agriculture	23	Korean Agriculture Society
Medicine/Pharmacy	13	The Pharmaceutical Society of Korea
Engineering	33	The Korean Society for Composite Materials
Machinery	9	Korea Society of Automotive Engineers
Electric/Electronics	6	The Korean Institute of Electrical Engineers
Information/Communication	8	Korean Information Science Society
Management Sciences	4	Korean Operations Research and Management Society
Home Economics	10	Korea Society of Costume
Others	16	Linguistic Society of Korea
Total	162	

We developed criteria to select qualified societies for informatization project since KORDIC plans to supply financial and technical supports to societies. Criteria to be selected are years of existence, publications, academic activities, subject category in science and technology, number of members, and whether or not registering at Korea Science Foundation, etc. Besides, we considered the size of the society since we believe a huge society in a specialized field might lead to medium and small size of societies in informatization. Criteria will be shown at Table 2.

<Table 2> Selection Criteria of Academic Societies for Informatization Project

Details / Category	Minimum Range	Remarks
Years of Existence	5 years or more	
Publication	At least 2/year	Academic journals, etc.
Academic Activities	At least 1/year	Conference, Workshops, Seminars to be held
Subject Covered	Science & Technology	Engineering and related areas
Member Size	500 or more	
Reputation		

SYSTEM DEVELOPMENT FOR ACADEMIC SOCIETIES

Information Resources

We have made a survey of information resources, which will be used to create databases for pre-selected trial society. We have identified the information resource of society by analyzing data that have been collected from 143 societies as of now and could divide into two different types of information, academic and non-academic. Information resource from an academic society can be categorized by journals, technical reports, conference proceedings, and workshop/seminar materials, etc.. There are also non-academic resources, such as announcement, member information, and so on.

Resources for academic information are magazines, journals, proceedings that has been published in society and materials from seminars, lectures, and workshops which hold academic value. These are the important resources for creating databases of informatization project.

Member information includes not only the private information of each member, but also the academic achievement, major area, present research field, so when someone looks for the professional of the field, can be utilized. And furthermore, this information with citation index on published articles of every member can later on be used in academic database project such as SCI.

Information on academic and non-academic events and data, sponsored by each society, such as all kinds of table, application form, article of association, members of the board, rules of submission of articles, and ceremony rules is also available. Among these, materials from conferences, academic seminars, lectures can be grouped as grey information source.

Collecting Academic Information

All the data produced by the societies are gathered in society office and being sent at KORDIC, based on the agreement between KORDIC and each society. The data that have been sent are all of information we have described above. Among these, substantial amounts are grey literature data. The data, the journal, magazine, proceedings, seminar materials, workshop materials that have been published in their society are collected in paper or digital format.

Creating Databases for Academic Societies in Science and Technology

Among the collected data, journal articles with academic value were built up with bibliographic databases and full-text image databases. Title, author, co-author, affiliation, source, publisher, publication date, Korean/English abstract of the thesis are included in the bibliographic database..

Full text image database was built up with mainly with tiff image file(CCITT G4 Fax form, 300dpi) and some of data was built up with PDF format.

Structuring Homepages for Individual Society

Homepage of the society was built up with constructing a database. In order to design a model for homepage, we collect examples of domestic and foreign society's homepage. After then, we designed basic forms for each society, and asked them to check if designed forms are acceptable. Then uniqueness and characteristics of each society was applied. On the society homepage, society main page, introduction of each society,

members of the board, society news, publication status, proceedings and journals published, related site to be linked was included. When the menu of homepage was structured with BBS, it was made easy to put on the information for the manager and also to the users as well. So it made to grow the homepage uses. Part of managing and renewing of the society homepage is handle by KORDIC, but when society appoints the web-master, he/she takes care of it.

Information service system via Internet of society's bibliographic information and full-text image database was connected with society's homepage by using information retrieval system, KRISTAL II of KORDIC so user can search on WWW and also have access to the full-text. Each society's own information was unified to Academic Society Home Village (http://society. kordic.re.kr/). This homepage will be in English also sooner or later, so foreign users can have access to it.

Information Services for Grey Literatures

In addition to bibliographic and full-text database service for journal articles, another major information services of this project is to provide users with full-text image for proceedings, materials of seminar or workshops held by each academic society. KORDIC builds up these materials into the database having same format with journal paper. Included fields for workshop materials are date, place, researcher, affiliation, subject area, and full contents.

Information services for each society include not only journal article service, but about grey literature service as well which includes full-text image service on proceedings, materials of each seminar or workshops. These grey literature structures a database with a same format of journal articles. This database was provided users with full-text services separately or offers all the grey literature of one society service. And at the same time, we put together the same subjects of the society and give service hole in one.

And also, journal article services are running in association with societies using homepage BBS.

And further more, the system for management of society members and management of society's journal article has been developed and has set the system to provide information service for journal articles and grey literature from societies on publication.

FUTURE PLANS AND CONCLUSION

Society informatization project has its own goal in developing the system for the society itself to provide services for the information that each society produces. KORDIC's role in this is to support each society to build the information service system by itself. And with the corporation of societies KORDIC offers the unified service to users so they can access various kinds of information on societies in one place.

Future plans are to make more of trial academic societies. KORDIC will keep on studying and developing more of changing information technology so the system that users can access high quality information much easier and collect the academic information, produced by the society, in full text as much as possible, so we can offer full text service in multimedia.

References

[1] The Korean Chemical Society (http://www.kcsnet.or.kr)

[2] The Korean physical Society (http://mulli.kps.or.kr)

[3] The Biochemical Society of the Republic of Korea (http://www.biochem.or.kr)

[4] Korean Mathematical Society (http://www.kms.or.kr)

[5] Choi, Seon-Heui and Hyun-Yang Cho. (1988) "A Study on the Information System and Information Resources from Korean Academic societies in Science and Technology, *KOSTI '98 Proceedings*, pp.231-236

Mine Maps as Grey Literature

By Linda R. Musser,
Head of the Earth and Mineral Sciences Library,
The Pennsylvania State University, 105 Deike Building,
University Park PA 16802; 814-863-7073; Lrm4@psu.edu.

Maps are a tough format - frequently big in size, awkward to store, and of little or no interest to many librarians. Even among the map library community, mine maps are the ugly duckling of the map library world. A narrowly focused, unglamorous, rarely collected stepchild of the cartographic community. Yet mine maps are repositories of unique, difficult to replicate information illustrating location and extent of historic and current mining activities.

The majority of maps are governmentally produced, a trait that qualifies most maps for grey literature status. Fortunately, such are usually published in significant enough quantities that they can be collected and held in multiple repositories. Their content is frequently of broad enough interest that many map libraries collect, catalog, preserve and provide access to these materials. The same cannot be said for mine maps.

Firstly, these maps are not produced by the government in most cases. They are produced by the mine operator for internal use, although many governments require their creation and deposit with the state. Secondly, they are produced in single or very limited quantities, making them hard to collect. The subject matter is such that, to many librarians, it seems of limited interest or utility. This is far from the truth, of course.

Mine maps are of great utility to a variety of users. Homeowners (and realtors trying to sell homes) want to know if there is a mine underlying the property. Mining companies desire knowledge of the extent of previous mining operations in the event they choose to commence re-mining operations. Geologists utilize these maps to assist

in mineral reserve estimation, sample collection, exploration and mapping. Planners and engineers are interested from the point of view of project and community planning. Government inspectors require them for reasons relating to health and safety as well as environmental protection and redemption.

CHARACTERISTICS OF MINE MAPS

Mine maps, like all maps, come in a wide variety of sizes, colors and materials. Maps were frequently drawn on the most convenient and inexpensive material at hand. Thus mine maps are found drawn on newsprint, velum, linen, drafting paper, canvas, and mylar. Color may or may not be used depending on the mineral or mineral seams being represented in the map. For example, maps of anthracite mining operations commonly utilize color to indicate the extent of mining in a particular coal seam. Size is also tied to the type of material being mined. Again, using historic Pennsylvania coal mines as an example, anthracite mines tended to be limited in width but could be quite lengthy thus generating maps of reasonable width (137 cm.) but extreme length (366 cm.). Bituminous coal mines, on the other hand, were extensive in all dimensions requiring maps of equivalent proportions (e.g., 600 cm. x 900 cm.).

MINE MAP COLLECTIONS

As previously stated, mine maps are little collected by most libraries. Many of these maps are acquired by donation directly from the mining company as part of their archives. On occasion, mining engineers will also donate mine maps. It is quite rare, however, to find a mine map in a library outside a government agency. Given that many governments require that mine maps be deposited with the state, it is not surprising that the government is the most common collector of mine maps. Unfortunately, not all governments take good care of these unique records of industry. Nor are they uniformly concerned with providing access to these records. Conditions and access vary widely.

In the state of Pennsylvania, for example, mine maps are collected by no less than seven different agencies. Each agency has unique materials and separate inventory schemes. Additionally, the cooperation and communication among the agencies frequently leaves something to be desired. Also located in Pennsylvania is a national respotsitory for abandoned mine maps. Operated by the U.S. Office of Surface Mining Reclamation and Enforcement, the National Mine Map Repository (NMMR) collects and creates microfilm copies of abandoned mine maps loaned to the NMMR. Its holdings, while extensive, represent less than fifty percent of the mine maps in existence in the United States.

CONCLUSION

Similar repositories exist in other countries as well and they face similar challenges of collecting, cataloging, preserving and hopefully providing access to these unique materials. It is hoped that librarians will become more aware of the value of these not-so-pretty stepchildren of the map library world and endeavor to do their part in preserving these unique records of industry.

BIBLIOGRAPHY

Musser, Linda R. and Lisa A. Wishard (1999), "Mine Map Repositories in Pennsylvania and the United States", *Proceedings of the Geoscience Information Society*, v.29, (in press).

Office of Surface Mining Reclamation and Enforcement, U.S. Department of the Interior (1997), *Mine Map Respositories: A Source of Mine Data*, U.S.G.P.O., Washington, DC.

Automatic keywording of High Energy Physics

David Dallman, Jean-Yves Le Meur

CERN, Geneva, Switzerland
5 October 1999

Bibliographic databases were developed from the traditional library card catalogue in order to enable users to access library documents via various types of bibliographic information, such as title, author, series or conference date. In addition these catalogues sometimes contained some form of indexation by subject, such as the Universal (or Dewey) Decimal Classification used for books. With the introduction of the eprint archives, set up by the High Energy Physics (HEP) Community in the early 90s, huge collections of documents in several fields have been made available on the World Wide Web. These developments however have not yet been followed up from a keywording point of view.

We will see in this paper how important it is to attribute keywords to all documents in the area of HEP Grey Literature. As libraries are facing a future with less and less manpower available and more and more documents, we will explore the possibility of being helped by automatic classification software. We will specifically mention a project being carried out at CERN (European Laboratory for Particle Physics) for testing this automatic keywording.

SEARCHING DOCUMENTS BY SUBJECT

There are two main uses of a bibliographic database. The first one is to search for a specific item which one already knows about, and wants to find out if the library has it, and if so, to get access to the document. This is the so-called referral approach, a bit like looking up a piece of information in an encyclopaedia. You know it is there, you just need the answer.

The other main use is when one has a specific problem in mind, and wants to find documents which address that problem. It is only with this second type of use that we are concerned here. Basically, this means a subject-based approach to the library collection.

1) Subject connections via references

There is already a system of searching academic literature in a thematic way without any kind of intermediate database. This is via the references to other work which have been an accepted and important part of scholarly publication since the very beginning.
Starting from a core document, one can gradually widen the scope using the references and hopefully arrive at some fairly complete set of relevant documents. The electronic age has again enhanced such an approach without however changing it in principle. References in an electronic document can be links to the electronic versions of the documents referred to.

The main obvious drawback in this approach is that authors may not have referred to all the relevant material, either due to deliberate omission or just because they do not know about it.

Another possible disadvantage is that by definition one can only refer to what already exists at the time of writing a document ! From the core document, past documents will be reached but all new documents will be missed. However,

this could theoretically be solved through a database of such references, by forward searching from a document to retrieve all other documents which have referred to it later.

In practice, the connections via references is not an adequate approach for users needing for an exhaustive list of available documents related to a given topic. It takes too long and the full coverage is not guaranteed. The other solution - querying directly a bibliographic database - is much faster but it may still result in an incomplete result.

2) Recall and precision in searching

In evaluating the value of search results, two different concepts are of great importance: the first is the recall factor and the second is the precision.

The recall factor measures how many of the total number of documents which should be retrieved by the search, are in fact retrieved. Some relevant documents may be missed by a given search strategy. Taken on its own, this factor should obviously be as close to 100% as possible.

*recall = 100 * number of documents retrieved / total number of relevant documents*

The precision measures what part of the documents retrieved actually belong to the desired sample, the rest being undesired documents which have somehow also managed to satisfy the search criteria. Again, when considered alone, this should also be as close to 100% as possible.

*precision = 100 * number of relevant documents retrieved / number of documents retrieved*

The problem is that these two measures of search efficiency are not independent, in fact there is an anti-correlation between them. If you try to get the recall factor as high as possible by using a more complex search strategy, you will also tend to pick up more "background" documents which you do not want. Conversely, if you want all your retrieved documents to be relevant, you will have to pay the price of missing quite a lot of relevant documents too.

Database search engines may offer some features which are designed to improve the precision of the search. Words can be strung together as a phrase, and this phrase searched for. However searching for a phrase of more than three words is likely to result in a low recall factor, because of the flexibility of natural language (particularly English) in representing nuances of meaning by variations in word order.
In another approach, limits can be placed on the maximum number of intervening words which are allowed to occur between a pair of chosen words (proximity searches). The CERN Library database has such a functionality but it is rather cumbersome to use at present.

Of course, it is not just the search strategy which counts, but the result of the search strategy when applied to the data. Therefore, exactly which data for a given document are available for searching influences the search result.

3) Data from the document itself which are available for searching

Here we are only really concerned with data concerning the subject matter of the document, so things like author names do not play a role. Of course, searching for an author can result in retrieving a certain subject, but it is almost never the case that this author is involved with all the documents in that subject area.

Traditionally, the title is the item of bibliographic information which expresses the subject content. However, a title is usually far too short to contain a complete description of the subject area in a way which can be used efficiently by a search engine. A specialist reading the title may understand what the document is about, but he is using all sorts of prior knowledge into which context he plugs the new title. Therefore the recall factor of a title-based search is likely to be low. Furthermore, as the number of documents in the database steadily increases with time, the precision of title-based searches is likely to decrease as well.

An extension of this, which has become much easier to realise for electronic documents, is that more of the text than just the title can be used for searching. In particular, extending the search to the abstract is a very useful step. This has been done with the CERN HEP database since we started handling electronic preprints in 1994. However it has to be remembered that all words in the text are treated equally, so a mention in an abstract of a term by way of contrast and not because it is dealt with in the document, will still cause it to be indexed.

In principle, the full text of the document could be used for searching, but in practice this has not been done for documents in the HEP field. Considering the huge number of documents produced in this field, searching on full text would probably give a good recall factor but the precision would be far too low to be really useful. This is surely why we do not know of any project for indexing the whole text of HEP literature.

An alternative approach is to supply additional data concerning the subject material, and to use this for searching.

WHY DO AUTOMATIC KEYWORDING ?

Adding of subject material is called subject indexation or keyword enhancement. When we say "keyword" it could of course be a phrase of two or more words. There are two very different ways of doing this : to choose terms from a fixed thesaurus or to use free keywords which can be chosen by the indexer at will. The strategy of assigning keywords will obviously depend on which parts of the document itself (title, abstract, full document) are also available for searching.

1) Adding data to the documents

a/ free keywords

Allocating keywords on a free basis could also use terms which are not present in the document, but in practice this technique is mainly used for adding useful words or phrases taken from the text, such as section headings and other specific words which could help in improving the recall factor of the search. Free keywords can also be useful for indexing terms containing special characters which would not be completely recognised if they appeared in the title or abstract. For example, the CERN Library database normally breaks off indexing a word when it meets a non-alphanumeric character in a title or abstract, but it can be directed not to do this for a keyword field. Thus particles called W+ and W- would both be indexed as W in a title, but the full forms can be used as keywords and retrieved.

Free keywords can also be a useful way of adding synonyms of terms that appear in the text. But it would be better in general to handle synonyms at the search input end rather than adding them to each record when they occur.

b/ fixed thesaurus terms

The efficient allocation of keywords from a fixed thesaurus makes the most demands on the indexer, as the documents have to be well understood. The indexed terms may not appear in the same way in the text at all, which can give this method a big advantage over any strategy which just uses the text of the document.

Of course, such a method requires the existence of a complete, precise and up-to-date thesaurus, which is quite difficult to achieve in a rapidly-changing specialized research area like High Energy Physics.

2) Comparison between free keywords and fixed terms

In practice, these two forms of indexing are extreme cases. Real approaches have aspects of both, even though they may be closer to one than the other.

Thus, the drawback of having a fixed thesaurus is that the thesaurus itself has to be modified to keep up with developments in the field. This usually means that a new form of the thesaurus is issued at regular intervals, for example the DESY (Hamburg, Germany) HEP thesaurus has been updated every one or two years. Thus for searching

back over many years, each time period should in principle be combined with the relevant thesaurus terms for that period. In practice, this complex procedure is rarely undertaken by the searcher. It could be built in as a front end to the search, but this has not been done yet for any of the databases in our field which use fixed thesauri.

On the other hand, free keywording can be chosen to conform to a minimum set of rules, instead of being completely free and just taking the words as they appear. For example, it could be decided to choose singular forms instead of plurals. In fact, after a period of use, listing the terms which have been given as free keywords does give a sort of "thesaurus in practice", which can then be used to standardize the keywords which are subsequently assigned, in order to improve consistency.

3) Influence of the keywording on search quality

Including the abstract instead of just the title in searching (with no additional keywording) increases the recall factor but probably reduces the precision. Use of proximity searching could offset this loss in precision somewhat. It is straightforward to measure the change in precision (within a database which permits searching in title and abstract separately) but the absolute recall factor cannot be measured like this, as one has no way of knowing which relevant documents have not been retrieved at all !

The free keyword system is designed to be used in conjunction with the other data like title and abstract. If used with the title alone, it probably improves both recall and precision. But it does not give much improvement over using title plus abstract, as free keywords are most of the time words already present in either the title or the abstract.

The fixed thesaurus approach aims at describing each document by a series of thesaurus terms in such a way that both the precision and recall are 100%. This aim might not be achieved in practice if the expertise of the indexers leaves something to be desired. It is very important to realise that searches should only be made using the thesaurus terms assigned, all other text like title and abstract should be ignored. Some kind of measure of how much the thesaurus keywording improves search results can be obtained by searching for the particular thesaurus term in the title or in the title plus abstract.

The table below shows the numbers of documents found in some examples. For the chosen term in the DESY HEP Index (which covers published HEP literature), we look for the occurrence of this term as a keyword, then we look for its occurrence within titles in the same database. To compare with the abstract search, we use the CERN HEP database [1], whose coverage is similar.

We also look for the same term within the global scope of all eprints, published or not published. This gives an idea of the area which is not covered by the HEP Index.

Database:	DESY	DESY	CERN	DESY	CERN
Search performed Terms	By Keyword in *published* HEP	By Title in *published* HEP	By Title or Abstract in *published* HEP	By Title in *published and non published* HEP	By Title or Abstract in *published and non published* HEP
"Higgs Boson"	**8**	**703**	**510**	**1034**	**1305**
"Supergravity"	***5945***	***1791***	***1000***	***2714***	***2908***
"Duality"	6257	1073	779	1557	2591
"Interface"	409	95	272	422	1005
"Bifurcation"	57	41	80	80	287
"Dielectric"	293	111	108	240	450
"Graphics"	147	9	41	140	161

"Measure"	364	273	582	607	2556
"Pair"	228	1279	1000	2151	3804

Comparison of searches by thesaurus terms and searches by title/abstract in the DESY and CERN databases (28/09/1999).

The differences in the results are striking. The most reliable numbers in terms of precision and recall are the ones in the first column.
This means that when a user finds "Higgs Boson" in 703 titles of published HEP literature, only 8 of them only are really relevant to this topic. On the other hand, while a user may find the word "supergravity" 2714 times in the title, or 2908 times in the abstract of grey or published HEP literature, the number of documents actually relevant to this subject is more than double this. The other examples show the same kind of mismatch.

Moreover, it appears that the quantity of HEP literature without a classification (because it is not published) is quite large.

4) Conclusion

The added value of keywording based on a thesaurus is obvious, even when many other bibliographic fields are searchable. There is a direct relationship between the added value of keywording and the number of searchable documents: the more documents you keep, the more you need keywording.
A simple subject allocation cannot be satisfactory in the long term. Subjects need to be refined till they actually reach the precision of a thesaurus. The permanent increase of papers available in HEP will lead to a chaotic situation for Information Retrieval if a complete effective classification is not undertaken. In the next section, we will see what has been done so far in the High Energy Physics area.

Indexing by subject specialists is by the far the most precise method, but it is costly in terms of time and it requires highly-qualified people to do it. The question arises as to whether one could achieve a useful result by some automatic procedure based on the text of the title, the abstract or the full document.

TOWARDS AUTOMATION IN HEP

Before considering the automation itself, we give an overview of existing classifications in HEP. We describe the HEP specificity regarding the development of a keyword assigning expert system. Finally, we explain the tests that are currently being carried out at CERN.

1) Existing Classifications in High Energy Physics

Manual keywording has been carried out at DESY for more than 30 years. It covers all published articles in the various areas of HEP. The DESY HEP Index publication was the main output of this activity from 1963 to 1997. This publication itself then stopped but keywords are still allocated and they are searchable on the Web interfaces of the DESY [2] and SLAC [3] (Stanford, California) library catalogues.

A manual keywording activity used to be done at CERN as well. It started in 1983 with free keywording and was based on the HEP Index thesaurus (from 1989 to 1992). After this it was stopped due to lack of manpower.
Examples of fixed "commercial" thesauri are those used by INIS [4] (International Nuclear Information System, Vienna) and INSPEC [5] (Physics, Computing and Electrical Engineering Abstracts, UK). They are built manually and access is not free of charge. They are not sufficiently specialized in the HEP area and so are not really adequate for dealing with HEP literature.

Today, some articles do contain subject information supplied directly by the authors (usually only when the journal makes it a condition of publication!). So some journals have keywords, and quite a few journals have adopted the

Physics and Astronomy Classification Scheme (PACS) classification [6] supported by the American Physical Society. However, these approaches are far from being complete, so they are not useful for global searching. Also the PACS classification is still too broad for detailed searching in a narrow field such as particle physics.

On the contrary, in the case of books, where the Universal or Dewey Decimal Classifications are widely used, this approach can be very useful for retrieving all books dealing with a particular subject. A Web interface enabling searchers to browse a partial UDC index exists for CERN Library book catalogue [7], HEP preprints and published articles have no such world-wide recognised classification.

The CERN project, in collaboration with DESY and SLAC, is to use an expert system for automatically deriving keywords and then to map them onto the DESY HEP Index. In other fields such projects have already been carried out rather successfully. The Medical National Center (MNC) [8] and NASA [9] use machine-aided indexing for example, to speed up their classification.

2) Particularities of HEP literature

Natural language contains a huge vocabulary and the syntax of languages is very complex. In traditional literature, a text can be processed by considering words as individual items. A dictionary of single words can be used as the basis for creating a knowledge base. In scientific literature, we consider that the meaning is expressed mainly through multi-word terms ("noun phrases").
In HEP, documents contain many particle symbols or equations which may be among the most relevant noun-phrases in the document. Describing the syntax of the sentences present in HEP literature requires at least the definition of a new type of word: the particle symbol.
In addition, the knowledge base needs to be set up differently for experimental and theoretical documents. It is also planned to handle another specific dictionary for technology-related papers.

Another particularity is the size of its electronic grey literature. It amounts to more than 100 000 documents since 1994 and is growing at the rate of about 20 000 per year.

3) Sokrates Learning System

SOKRATES [10] stands for "Self-organizing Object-oriented Keyterm Recognition And Text-Editing System".
It derives from natural language key terms and keywords. Each new piece of information treated by the system is used to update a knowledge base. A learning system like Sokrates can be compared to a compiler: the input is a text written with a known syntax. The output is a condensed executable, like the set of key terms.

In terms of the earlier discussion, the Sokrates approach belongs to the free keywording type, where the free keywords must appear in the text and cannot be invented (except perhaps for synonyms if one chooses to build them in). So, for example, there is no way that this algorithm can return the term "Kac-Moody algebra" when the abstract says "graded Lie algebra", even though these are two names for the same thing.
The test of the software is divided into two parts: the derivation of the best key terms and their mapping onto the thesaurus.

a/ The Term Derivation

To run the extraction of the Key terms, three basic components are defined:

- ✓ A complete dictionary which is created and continuously updated. In this dictionary, individual words (any possible character type) are kept with the following two main attributes:
  - a code or type of word: "General", "Left", "Right", "Stop-word", "Particle", etc.
  - its frequency: number of times the word has been encountered in all documents processed so far.

- ✓ A knowledge base which stores all the key terms (single or multiple words) which have been selected together with their frequency.

- ✓ The rules for describing key terms (the "Text Description Language") which are expressed using the type of words. An example of a rule is:
  1. A P G R ... (Left, Stop Word, Particle, General, Right...), where L (R) enables one to specify that a word would only be significant when it appears to the left (right) of another relevant word.

An inference engine is able to match any rule to any text using the above rules and the dictionary of single words.

When dealing with a new document, Sokrates extracts all individual words and distinguishes old and new words. It counts repetitions and updates the dictionary. For new words (if any), it can ask an operator to provide the word characteristics (code).
In the next step, the term selector selects candidate terms. It uses the dictionary and the inference engine to extract all possible noun phrases. After this pass, a set of "valid" and "garbage" terms are available. The selector compares these derived terms to the established knowledge base, and keeps the ones for which a frequency threshold has been reached. The threshold can be defined differently according to the number of words in the noun phrases. For example, we could require that single words must have a frequency of 10, two-words terms a frequency of 5, etc.
If the number of derived keywords is too small, a third pass of the selector can be undertaken, with a reduced threshold definition.

b/ The Thesaurus Term Mapping

Different situations may occur:
- the key term exists in the thesaurus: the mapping is straightforward.
- the key term is similar to a thesaurus term: the correspondence can easily be established.
- the key term does not exist in the thesaurus: a subject indexer could associate one or more terms from the thesaurus to each key term found. This would only need to be done once for the whole dictionary; it would remain valid for all incoming documents. Only new key terms would need to be associated at later stage. Thus the savings over indexing on a document-by-document basis could be considerable.

4) CERN test: the status

The data for the test consists of about 1400 abstracts of preprints in each of the three fields of experimental high-energy physics, theoretical high-energy physics, and technological articles relevant for technology transfer. These have all been keyworded using the manual method.

The object of the test is to compare the automatic procedure with the manual method of keywording in use at DESY. We intend to find out whether the automatic procedure can be tuned to deliver keywords of a similar quality as DESY does with the manual method. Even if this turns out to be unrealistic, the expert system could be used as a Machine Aided Indexing (MAI) system to propose keywords to the indexer.

When the program has been tuned on these samples, it will be supplied with other samples of the same size. In comparison with the manually extracted keywords, the keywords extracted automatically should be as similar as possible in quantity (number per abstract) and in character.

70 000 words have been used as the "learning text" so far. Among the last 4000 words processed, only 200 words were unknown to the existing dictionary and required an input from an operator.

250 rules have been defined. Any sentence, parsed through these rules, will end up with one or many possible noun phrases.

The thresholds are still being defined: they need to be regularly adjusted as more documents are processed.

CONCLUSION

We can draw three main conclusions from the analysis and tests done so far:
1) The necessity of automatically keywording High Energy Physics Grey Literature is obvious.
2) We are optimistic that we will be able to build a valid knowledge base of noun-phrases, using Sokrates.
3) It is not yet clear how difficult it will be to successfully map this base onto the HEP Index thesaurus.

In all cases (whether the mapping is robust or not), the idea is that whenever a new document (with an abstract) is entered into the system, the expert system quickly delivers a set of key terms. This output can be added to the database straight away or it can be mapped to the thesaurus to try to deliver an "assigned term" and finally, it can be checked by experts before being loaded.

If the CERN test is successful, it will be run on a large scale in order to progressively cover all HEP preprints not yet classified. In addition to the traditional keyword searching option, a new utility will need to be developed to offer searchers a simple way of browsing through the thesaurus.

REFERENCES

[1] http://weblib.cern.ch/Preprints
[2] http://www.desy.de/library/homepage.html
[3] http://www.slac.stanford.edu/spires/hep/
[4] http://www.iaea.or.at/programmes/inis/inis.htm
[5] http://www.iee.org.uk/publish/inspec
[6] http://aip.org/pubservs/pacs.html
[7] http://weblib.cern.ch/share/udctree/udctree.html
[8] Wright, L.W., Grossetta Nardini H.K., Aronson, A.R., Rindflesch , T.C (1999) " Hierarchical Concept Indexing of Full-Text Documents in the Unified Medical Language System Information Sources Map", Journal of the American Society for Information Science, vol 50 No 6, pp. 514-523
[9] Silvester, J.P. (1997) "Computer Supported Indexing: a History and Evaluation of NASA's MAI System", Encyclopedia of Library and Information Science, vol 61 Jan 1998, pp. 76-90
[10] Sokrates Learning System, Dr. I. Steinacker, Intelligent Text Analysis and Management, Jenewein Burotechnik Innsbruck Austria.

Trends in Publishing Academic Grey Literature: Examples from Economics

Fethy Mili

Librarian, Département de sciences économiques
Université de Montréal
6128, Succ. Centre-ville Montréal , Canada H3C 3J7
Phone 1 (514) 343-6111 - Fax 1 (514) 343-7221
Email: fethy.mili@umontreal.ca

Abstract

What impact might electronic publishing have on grey literature? This work tries to give some answers based on changes in the publishing of economics research preprints or working-paper series.

Institutions and some members of the economics community have created archives to facilitate access to working paper series. In this document, we evaluate the prospects for electronic publishing of economics preprints using an inventory of material available from institutions' web sites. Also the paper contains an overview of some working-paper archives managed by economists.

1) Introduction

This paper deals with preprints in economics. Exchanging preprints or "working-paper series" as they are commonly known among economists had, been the traditional way for informal communication of recent research findings. Electronic dissemination of recent scholarly works has grown with the use of the Internet and the World Wide Web (WWW) facilities. Indeed, it is well documented now that computer and electronic dissemination of knowledge have changed the way of life within the information world. The incursion of the Internet as a gateway to the global information environment plays a significant role as a support in daily research activities.

Within universities, the Internet has become one of the major information vehicules particularly among the scientic community . It is quite easy to search information from the Internet for processing and manipulation. Indeed "the element of scholarly research most affected by the Internet is that of communication at all stages of the project, from study design and data gathering to project administration and dissemination of outcomes". P. Warren-Wenk and Vivienne Monty (1996)

For economists, as well as for many other academics, "Scholarly communication in the sciences holds, at its cornerstone, the tenets of publishing quickly the results of scientific inquiry or findings, and distributing the results of research expeditiously to the academic and industrial communities" as mentioned by J.Gelfand (1998). This situation has increased interests to find out quick channels to communicate research findings among colleagues and common academic public.

The delay associated with publishing in commercial journals, promotes the importance of working-papers series among economists. These documents were an assigned candidate for electronic publishing, distribution, and sharing via e-mail services and Websites. Robert P. Parks as pioneer of promoting the use of electronic infrastructure for economists describes with evident enthousiasm the success of E-Print Archives at Los Alamos National Laboratories which provide readers with easy access to preprints. This philosopy was the corner stone in creating the pattern used in managing EconWPA (http://econwpa.wustl.edu/wpawelcome.htm)

Since the beginning of the 1990s economists launched several on-line working-paper networks in order to intensify awarness among the economics community about recent findings. According to S. Karlsson and T. Krichel, the electronic working-paper series in economics appeared in 1993 and the production has grown as the use of the Internet features. Their paper pointed out the progress realised in electronic publishing of working-papers during the last few years. Indeed, many members of the economics community rely on the Internet to develop a self-publishing procedures. P. Warren-Wenk and V. Monty (1996) suggested that Web features provide faculty with a friendly environment in order to establish their own homepages and disseminate their writtings.

Others projects have been created in order to promote electronic publishing by building either central or shared archives with free access. These projects aim to enhance the speedy of access to publications. Economists were invited to participe by submitting their prepared papers to the archives in order to facilitate access to preprints. Wether they are centralised or decentralised, these archives aim to be a reference source to provide information about preprints in economics and to avoide the evident chaos (L.A. Davidson and Kimberlet Douglas, 1998) as well as "a growing disenchantment" within the Internet as described by A. Eisenberg (1997) . However according to W.L. Goffe and R.P. Parks (1997), "Authors who want their work to be read need to post their papers to a working-paper archive like EconWPA , or register the paper with an index like WoPEc".

The creation of such archives (NetEc, EconWPA,SSRN/ERN ...) has been associated with the arrival of local archives among academic and reseach institutions which have started publishing working-paper series in electronic format as "it is assumed that patrons would like the ability to have electronic access to serials either form their desktop or from a convenient location." as noted by Tenner, E. And al. (1998). For many academic institutions and research units, announcements of preprints in economics have become a part of their websites in order to inform Internet users about the reseach activites taken by faculty and other researchers. Also advertising research findings is a way to promote the dynamic work of research teams within the institution. Currently there are several hundred universities and research units with Web and FTP servers that provide online information about working-paper series.

In this paper we will try to sketch the changes which have occured in publishing and distributing paper series in economics. After several years of introducing electronic publishing, what does the world of grey literature in economics, espacially preprints looks like? In other words what has happened to the access to and availability of paper series in economics since institutions started to be interested in electronic publishing. Is it more easy to get a reference about working-papers? Do the announcements provided by the websites supply users with links to online papers? In short, has the incursion of the Internet into the reseach environment increases the possibilities for user to get working-papers and bibliographic information?

Meanwhile, the arrival of electronic full texts has not decreased the publication of hard copies. Indeed our daily experience shows that the production and the distribution of printed working-

paper series is still prosperous and only some academic institutions have thrown aside production in printed form while offering only electronic version of these papers. One might be surprised to realise that so many organisations still rely on the "conventional" method to advertise research results despite the promises of the electronic way of publishing.

In other words the use of the Internet to disseminate preprints in economics has multiple faces and we shall test the publishing services provided by the institutions in order to understand the trends in the availability of electronic full texts of working-paper series. This may help us as librarians to deal better with a changing world of information.

Like other grey products, investigating the situation of preprint literature in economics is a challenge because of the lack of a centralised project system for storing and distributing papers. As a result, it is difficult either to get or to define a clear view of what is effectively produced by different economic research units. In fact one can come across downloadable papers in homepages of authors as well as in other sites managed by academic departments and research centers. As this paper is concerned with the preprints prepared by the institutions (department of economics and research institutes ...) we don't attempt to evaluate the individual contributions widespreaded through the Internet. Also, we will not use the contents of particular projects like WoPEc. The reason is that this kind of archive relies basically on volunteer contributions and we believe that a reliable way to measure the changes is to see the institutions' participation to provide digital versions of working-paper series.

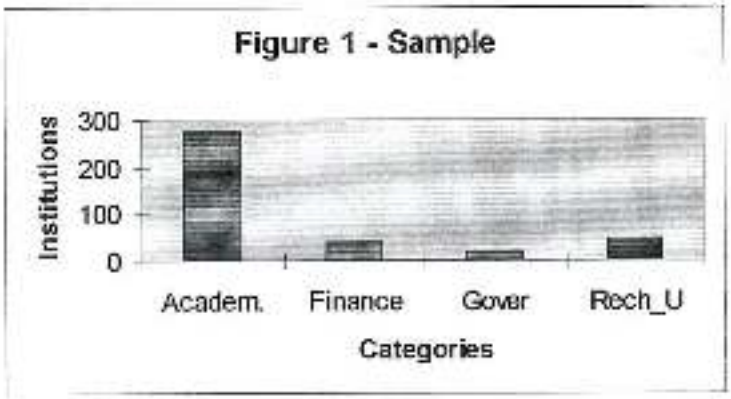
For this study we will use the series we receive on regular basis at our library. Our collection of working-papers counts more than 410 series. In 1993, these series were used to launch NetEc/BibEc with is a large bibliographic database of working-paper series. At present BibEc holds 50,000 bibliographic records of working-papers series. When created, the service inluded bibliographic records on some 35 000 economics working-papers (W.L. Goffe, 1994).

2) Data Description

The data we employed result from a first-hand compilation of elements related to working-paper management. However we are confident that the sample used would give a valuable appreciation of both past and future changes in electronic publishing of preprints in economics. Thus a sample of 383 Websites was used to gather information about the state of publishing

paper series via the Internet. We didn't contact providers and working-paper coodinators in order to suply details about the availabilty of their publications on the Internert. The reason was to simulate a searching process that could be made by a common user and to find out what kind of information he can get.

It is clear that we did not rely on any special scientific method of sampling and the sites we've examined were selected ramdomly. The series were drawn from our local collection as our library has a strong tradition in collecting and processing working-paper series in economics.

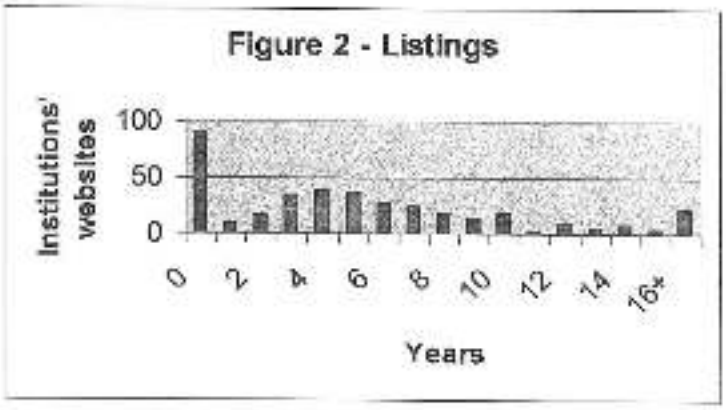


All the sites visited belong to departments of economics and research units. Figure 1 provides some details about the series sources. As shown, academic departments of economics present the major part of the sample. Indeed 276 websites checked belong to universities while the other providers are related to national and international financial institutions as well as to government agencies and to independent research units. As already estalished, academia is an important producer of knowledge; and, an important part of grey literature in economics is still produced by and for academics.

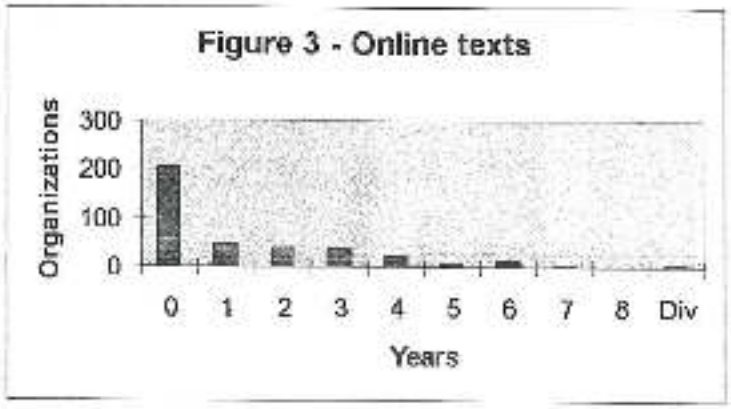
3)Result

Three different approaches have been observed while monitoring of the different websites:

3.1- Institutions that provide bibliographic lists of working-papers series produced by their faculty and searchers: These lists describe the contents of the series. An abstract is always given with each record at least for the recent years. These lists can be very helpful for people who need to be informed about the last preprints produced and in many cases the documents could be ordered either from the institution or from the author. In some cases a working-paper coordinator is associated to the Website, this makes it easy to get a paper when desired.We've also noticed that the great part of the sites visited have listings for 3 to 7 years. Only few institutions provide lists since more that 8 years, as Figure 2 shown



3.2- Institutions that make available access to the full text of papers: in that case, one can find information about how download the file as well as some details about the format in which the files are stored. Online files (downloadabale full text) are less available than one might expect. Our inventory shows that only 177 institutions (around 46% of the sample) are involved in offering online files of the paper series produced. The other indication is that 146 organizations (more than 80%) provide electronic version of the series for 1 to 4 years. See Figure 3.



3.3 Institutions without information about research preprints: these organizations are not persuaded that scholarly works could be well disseminated via the Internet. It is an important "indifferent group" (E. Shreeves, 1992) which is not committed or convinced about the usefulness of academics going digital in order to increase the communication of their findings.

Also the inventory permitted us to evaluate the updating frequency of the information provided by the institutions websites. In fact, one of the features that makes the information valuable is to keep it up-to-date. In reference to the listings available, 199 out of 293 contain data about publications produced in 1999 when 108 among 177 providers offering downlodable full text papers hold materiel produced in 1999. In addition, a link seems to exist between the starting of making the material available (lists & online files) via the Internet and its up-to-dating. Indeed institutions that keep the sites up-to-dated have mostly began an online services during the last 5 years.

In order to summarize the results of the inventory one can raise the following points:

1) Information about working-papers in economics is more and more available. It is the most positive point for enhancing awareness about economics' grey literature. Information professionals can be better informed about grey literature production. Indeed, besides its quality, the information we found while browsing the website illustrates the impact Internet is playing to insure better advertising to grey economic literature.
2) Despite the increase in information about preprints in economics, the effective participation of institutions in starting and maintaining full text Websites is still conservative. Even if electronic publication of working-papers seems to be accepted by economists, the results suggest that organizations are not making the necessary effort to guarantee the availabilty of the papers as is the case for producing printed working-paper series.
3) The electronic availability of paper series is quite recent. It is an emerging service and we may expect an improvement of the organizations' participation in the future.
4) Up-dating information is an important element for accuracy of the material. Many providers don't add their recent publications to the website. This would reduce users' interest to rely on institutions' websites to stay informed. The facts suggest that the information professionals dealing with grey literature should keep on gathering material according to the traditional methods if they need to provide customers with recent publications.

4) Discussion

Beyond the deficiency of the present context there are facts, and one should realize the great work done by organizations in order to give " grey literature in sciences gains a more prominent role in a range of information use, access and dissemination activities" (J. Gelfand 1997). In economics, electronic publishing of paper series is in progress and while browsing the institutions' websites we understand a real willingness among providers to disseminate information about the research preprints; however, one may require more committment among organizations in order to obtain an effective fully electronic access to working-papers. Scholarly grey literature in economics is subject to an important transformation as it is taking advantage of computer-based networks and all the attendant technological features. As observed by William L. Goffe and Robert P. Parks (1997) "while some changes to an electronic world might be long in coming, and while there may be some confusion along the way, we expect that the ultimate outcome will be exciting". Nevertheless at present, the eagerness to moving toward a digital future is not shared by all the community. Institutions seem to be unsure about how to proceed. As we've noticed, it some important organizations still stand aside from wave of change, mainly within a friendly environment created by software facilities for electronic publishing such as html and Postcript tools and pdf/Adobe Acrobat. Some common reasons could explain the reserve either to start or to improve electronic dissemination of research findings among some institutions.

- Technology;

Edward Shreeves (1992) observed that "The technology itself is in a state of constant flux with little probability for stability". We believe that the context is not better today and "this forces one to realize that technology is never at a standstill and that the Web remains at a frontier stage, more embryonic than developmental" (J. Gekfand 1998). Changing standardization and publishing technology can be a source of hostility against technological features as well as increasing doubt about the Internet's efficiency for communication.

- Commercial activity

Subscription to working-paper series is still frequently used if we need to get issues on a regular basis. Our experience confirms that we have always paid to receive paper series produced by some departments and by most major research institutions. Providing working-

papers series, through a subscription, has always prevailed among research institutions and the incursion of Internet and electronic publishing has not modified this traditional way of dissemination. In some cases subscription have become more expensive than before. Market laws don't incite these providers to change their behavior so they are not disposed to move towards digital publishing with free access.

- Cultural context

P. Warren-Wenk and Vivienne Monty (1996) have observed that "scholars seem to be using the Internet as a current awareness service". Indeed many faculty would appreciate using the Net to be entertained but they don't feel confortable to be involved within an effective electronic environment. As noted by many, electronic publishing creates some ethics concerns mainly for scholars seeking academic recognition so they must publish their findings in a formal print journal. Uncertainty influences the institutions' committment in to producing electronic versions of paper series on a regular basis. Peer reviews present an other great obstacle facing the acceptance of electronic publishing as a serious product. Many think that electronic publishing will be adopted by the scientific community only when it meets the needs of that community, solves problems related to print material and offers new opportunities that exceed the traditional way of disseminating knowledge. As long as these conditions are not satisfaied, electronic publishing will remain a marginal medium associated to "professional pop art" (Raney R. Keith, 1998).

Some more prosaic reasons can also explain why some institutions delay starting electronic publishing of paper series is explained by Sune Karlsson and Thomas Krichel (1999) while talking about economists hesitation to participate to a central archives of electronic working-papers "Our theory is that economists have a built-in distrust of monopolies. In their book of tales, there are a numerous accounts of the welfare losses caused by monopoly supply. They may also be afraid of power accumulated by a person who controls a hard disk where the complete output of the discipline is stored."

All these points can justify the levels of institutions' engagement to introduce, and to maintain and to improve electronic publishing of preprints. Nevertheless, our hypothesis is that more involvement could happen within institution when people feel ready enough to start a new way of doing things. We believe that all organizations, mainly universities in North America have

the know-how in order to go on disseminating the preprint in electronic format via the Internet, but obviously print material seems to be enjoyed by many users . Also institutions have to be convinced that using cyber space is not a fashion or an "electronic gadgetry" as pointed out by Edward Shreeves, but an irreversible stream for information management and that "the benefits more than compensate for the effort" (S. Karlsson and T. Krichel, 1999). In between, information providers (universities and research units) as well as information professionals and clientele will keep on building a meaningful framework for the forthcoming electronic days. In any case the present hybrid context of producing working-paper series can't be harmful to grey literature which will continue to benefit from visionaires experiments as well as from the wisdom of its traditional managers.

5)Community projects

This paper is interested in the institutional contribution to working-paper electronic publishing. However we can't have a complete view of preprints services if we ignore scholars' participation in building and managing electronic papers projects. In this section we present a summary of two major initiatives conducted by some visionary economists in order to facilitate access to full text working-papers series.

It would be pretentious to fix a date for the starting of electronic publishing of paper series; however, it is admitted that during the early 1990s resources for economists were introduced to Internet. In 1994 William L. Goffe published an exhaustive list of material available on the Internet. Among the projects described in that article, NetEc and Working-Paper Archive (EconWPA) were announced as services dedicated to promoting the usefulness of electronic publishing of working-paper series.

These two projects are still leading to make economic grey literature as available as possible to the Internet users.

EconWPA (URL: http://econwpa.wustl.edu/wpawelcome.htm) , is an electronic archive based at the Economics Department of Washington University. People can add their papers via a webinterface using any format. Working-paper are grouped in 26 subject areas, according to the classification scheme of the American Economic Association. Different methods of searching are available.

NetEc (URL: http://netec.mcc.ac.uk/NetEc.html) is an umbrella for a "group of projects to make the network more useful to the academic community in economics" (T. Krichel and T.

Wichmann, 1994) WoPEc is a service about electronic working-papers series. S.Karlsson and T. Krichel describe the service as a database focusing "on the collection of metadata about papers rather than the papers themselves". WoPEc uses a stansdard protocol called RePEc in order to exchange data between participants without extensive coordination. This decentralized archives allow providers to contribute to WoPEc through the acceptance of a defined template format to display the information. According to coordinators, the project has enough success to be considered now as "the primary source about recent research in economics". In fact, EconWPA and WoPEc are supplied by material from the institutions already producing electronic preprints and it would be difficult to evaluate the impact of these archives in order to convince more organizations without electronic projects to start new initiatives.

EconWPA and WoPEc have difference philosophies for managing paper series. The first site is claimed to be a central archive and its manager observes that "While economists are undertandably skeptical about monopoly providers, a single database for each type of information may well be preferable to multiple," (William L. Goffe and Robert P. Parks). On the other side, WoPEc stresses on the virtue of decentralized model. Currently, the two projects reflect the eagerness of a growing number of economists to participate effectively in the dissemination of research findings in economics. The projects maintain their activities and in both cases many efforts are used to keep the material available, accurate and easily accessible

6) Promises and Trends

6.1 Electronic vs. Priting

Over the past decade the face of grey literature in economics has changed profoundly. Preprints are much more available from different providers so students and faculty have virtually more opportunities to be informed about recent research. Print papers series are gradually being replaced by electronic publishing. Many organizations have shown a real willingness to move forward electronic publishing of working-paper series by cancelling the distribution of printed issues.

Nevertheless, one could ask whether the availability of the material via the Internet is a sufficient guarantee to ensure access to the publications. Our experience with working-paper

management advises us to be cautious. Indeed, the Internet free market cannot (at least presently) whiten the grey literature and make it more identifiable for the common user who still encounters important difficulties to get a valuable information through electronic media. Grey literature could be more spread through the network, however it can't be easily found. We still need human interfaces to gather the information because the Internet as a delivery medium creates a significant need among academics and other users to orchestrate the flows of information. It is quite important to recognize that publishing on the Internet "has created new models and opportunities, in an almostly magical way" (J. Gelfand, 1998), but one has to make the difference between the hope to see the Internet solve some of the library daily problems related to grey literature (locating, retriving ...), and the scholars who "are facing unprecedented information overload in their attempts to identify potentially relevant information sources" (G. McKiernan). As pointed out in W.L. Goffe and R.P. Parks (1997), "rational authors should realise quickly that readers will not find their work" unless it is available from a recognized provider. Since availability does not mean accessibility, one should continue gathering grey literature using the conventional pattern and using electronic material to complete users' needs. Many institutions are aware about the instability of electronic publishing on the Internet. So they keep on producing printed publications and we believe that this strategy will prevail for the long run nevertheless the pressure adopted by some " electronic-publication protagonists" (R. Heith Raney, 1998).

As shown in the inventory, electronic publishing of grey literature in economics is still a pilot venture and one has to wait in order to see whether incoming strategies which are being developed will influence the level of success or failure of publishing paper series. Sixteen years ago , J. Newcomb (1983) observed, in reference of electronic information future's that "the most obvious is that there is a lack of certainty as to precisely what the impact of the new technologies will be and how will they can be applied to publishing". As now, it obvious that technological tools increase the potential of electronic publishing but the chaotic reign within the Internet makes it difficult to get an effective use of the material available. So today's "most common complaints voiced by both experienced and new users of the WWW is its lack of structure and organisation" (David Revelt).

6.2 Standards

One of the major limiting factors to create a large acceptance of electronic paper series is the failure of agreement about a common way to advertise, to store and to display issues . While monitoring the different sites, we have encountred different presentation styles and storing formats. However some archiving projets such as WoPEc initiative, have established protocols that allow providers to exchange data. Currently, Postscript, PDF and HTML seem to be the most popular formats for publishing papers on the Internet because of their flexibility and the fact that many researchers can have access to the required technology to read the files. As observed by David Revelt (1998) "many sites have recognized that there is no common file format for electronic working-papers, and thus, that make no attemps to impose one". The Lack of standards could be a source of difficulties to access electronic paper series as users need to have the appropriate plug-in and viewers to display the files and many are convinced that "there is not, and most likely, will never be a standard format for working-papers or other Internet documents" (David Revelt,1998). We believe that the format war will not take place dispite of attemps by some archives managers to create a reference format to describe and to store electronic versions of working-papers. Standardization of electronic version for working-papers could not occur because paper series did not have a recognized model before introducing electronic publishing.

6.3 Providers

Our inventory related 3 major players within the world of electronic publishing of working-paper series: the institutions, the authors and some economics community members who have built special projects to coordinate the dissemination of paper series. We should note that each provider is motivated by a different interests in promoting the use of electronic material. The difference of opinions about going digital creates a additional obstacle to the acceptance of electronic publications by the scientific community. For organizations (Universities and research units), making available research findings on the Internet would be a signal for their commitment to innovate the procedures about advertising data about research related activities. Self-publishing is a result of the possibilities opend by the Web tools, so faculty can establish their own homepages in order to disseminate their works. Institutions and authors promote electronic publishing because of its optimal opportunity to enhance awareness about academic findings among the community and not to fight again traditional publishers. Indeed,

some economists involved in publishing projects consider electronic publishing on the Internet as an important step to breakdown the authority of the commecial publishers on academic knowledge. William L. Goffe and Robert P. Parks observed that "In academic, a primary goal is the growth, acquisition and dissemination of knowledge, which is aided by the freest possible access to information produced by and for academics" (1997). The electronic publishing and dissemination of reseach findings could help to keep off all the mediators.

6.4 Perspectives

The Internet state, the difficulties to use standards and providers different strategies confirm that electronic publishing still needs to define its contribution not as an alternative to printed material but as an effective tool to meet users' needs and to increase the effective dissemination of research findings. As argued by P. Warren-Wend and Vivienne Monty "electronic journal supplemented, but did not supplant print journal as a mechanism of monitoring developments in a field". In the same way, electronic working-paper series did not replace printed issues and our everyday experience attests that users are always seeking for hard publications for their research topics and we don't feel that some preprints on the Internet has changed neither the structure of our work, as librarians, nor the nature of the reference demands from the clientele. In any case, working-paper producers seem to continue publishing papers in printed form even if some organizations have initiated electronic projects. So the changes produced by electonic publishing on the conventional management of paper series are not so evident. As a librarian administring a large collection of working-papers, I am still discovering the ways to introduce Internet-based publications into ready reference services. In the short time, we believe that this part of grey literature in economics will continue to be produced in hybrid packages since the electronic world is a work-in-progress and many modifications in the actual system might take place.

Trends in electronic publishing of research in economics will be deeply influenced by many factors particularly:

(1) People: the willingness to participate in creating and maintaining updated electronic information. One can't rely on a website to find recent publications if one is not sure of finding a stable, reliable, standardized file.

(2) Institutions: organizations have usually sustained the publication of print working-papers. As now, the projects inventoried have always been built by a faculty member with little support from the institution. The initial eagerness of some committed faculty can't be

preserved if the institutions can't etablish a standard framework to produce electronic paper series on a regular basis and according to acceptable standards

(3) Technology: certain technological features have a real potential to increase scholarly electronic publications. The software used to create electronic files are commonly adopted so one can easier download files and display their contents. The challenge is to find the valuable information within the chaos of the Internet. As noted by David Revelt "finding information is often a difficult task, and while search engines and lists of links are the most obvious attempts to mitigate this problem, no practical solution has yet been found".

(4) Libraries: information units may play an important part in order to promote the use and acceptance of electronic paper series through the enhancement of electronic issues among the research community. Intergrating references to electronic material within libraries' catalogs could be a signal to adopt electronic publication. We believe that the committment of some economists to developing new strategies in order to disseminate preprints, should rely on the library expertise to get the necessarly credit for the new product.

Back to our basic inquiry about the trends in publishing paper series in economics: The survey presented in this paper shows that producing electronic publications is, at the same time, a source of hope and eagerness for some people and the origin of skepticism and uncertainty for others. We have noted more commitment among individuals (both researchers and librarians) than among institutions, with the consequence that individuals must work hard to convince institutions about the robustess of the cyber world as an emerging arena for the dissemination of knowledge.

7) Conlusion:

The preceding chapters have exposed some issues concerning the ongoing changes with the use of electronic publishing of working-paper series in economics. While institutions' contributions to promoting and maintaining access to online information of paper series is, in many cases, limited, there is an effective interest among the economics community to built central archives in order to enhance awareness about preprints. Despite the enthusiasm about the benefits of digital publications, many scholars are still uncertain about the technological answers to academic publishing difficulties.

Beyond the actual debate, one has to notice that electronic publishing is far from a redical change in communicating research findings, but a new method to create, and disseminate information among a special community as well as among the Internet users. Electronic

facilities provided by the Internet still have problems related to finding tools; standards, establishing source authority, URL's movements and access difficultics. Netherless, the electronomic environment "gives online users unprecedented power to follow unanticipated paths to unpremeditated payoffs",
(R. Keith Raney, 1998). Meanwhile, people will continue to enjoy both printed publications and online texts, and grey literature should take benefits form the two media.

References:

AGORA (http://hosted.ukoln.ac.uk/agora/scope/htm)
BibEc (http://netec.mcc.ac.uk/BibEc.html), Fethy Mili, Ed.
EconWPA, Economic Working Paper Archive,**(http://econwopa.wustl.edu),** Rober P. Parks, Coordinator
Eisenberg, A. (1997) "Dislinking the Internet", Scientific American, 6,
Feldman, S. (1998) "Conference Showcases the Move form Dream to Reality", Information Today, Vol. 15, No. 6
Gelfand, J. (1998) "Teaching and Exposing Grey literature: What the Information Profession Needs to Know - Examples from the Sciences." Collection Building , Vol. 17 No.4 19, pp. 159-166
Goffe, W.L. (1994) "Computer Network Resources for Economics", Journal of Economic Perspectives, Vol 8 No 3, pp. 97-119
Goffe, W.L. and Parks, R.P. (1997) "The Information Infrastructure in Economics". Journal of Economic Perspective, Vol. 11 No. 3, pp.75-94
P. Grenquist, "Why I Don't Read Electronic Journals: An Iconoclast Speaks Out," The Journal of Electronic Publishing, 3 (1), ISSN 1080-2711, 1997. (http://www.press.umi ch.edu/jep/03-01/Iconoclast.html)
Karlsson, S. and T. Krishel "RePEc and S-WoPEc: Internet access to electronic preprints in Economics. Unpulished
Krichel, T. and Wichmann, T. (1994) "Internet Primer for Economists, The Economic Journal, Vol 104, (November), pp. 1496-1523
Mackinzie Owen, J. (1999) "Role of the Library Sector", GreyNet Quarterly Newsletter, Vol. 8, No.1
NetEc (http://netec.mcc.ac.uk) , Thomas Krichel, Coordinator.
Newcombe, J. (1983) "Electronic Information Distribution. The Role of the Traditional Publisher and the Librarian", Special Libraries, April, pp. 150-155.
Tumer, E., Gyeszly, S.D., and Rholes, J.M. (1998) "Electronic and Traditional Sources for Newly Established Brach Library:: Product Availability and User Preferences", Collection Building, Vol. 17 No. 3 , pp. 123-128
Raney R. Keith (1998) "Into a Glass Darkly". The Journal of Electronic Publishing , Vol.4, Issue 2, December 1998. (http://www.press.umich.edu/jep/)
Revelt, D. "Electronic Working Paper Standards", (http://alfred.sims.berkeley.edu/working-paper-standards.html)
Shreeves, E. (1992) "Between Visionaries and the Luddites: Collection Development and Electronic Resources in the Humanities". Library Trends, Vol. 40(4), pp. 579-595
Warren-Wenk, P. and Vivienne Monty (1996) "Internet Use Patterns and the Scholarly Research Process" , in "The Challange Face of Reference", Edited by Lynne M. Stuart and Dena Holiman Hutto. JAI Press Inc

Grey Literature Based Surveys of Information Industry Development in Central and Eastern Europe

Krystyna SIWEK

Information Processing Centre, Al. Niepodleglosci 188b, PO Box 355, PL-00-950 Warsaw, Poland

Abstract

During transformation and accession period in Central and East European countries, specially after dissemination of Bangemann Report, various surveys, reports, studies dealing with computer and telecommunication industry development appeared in form of publications, on WWW pages and as internal documents. They are prepared by external bodies as EU, OECD or global telecom organizations. Some studies have been conducted by foreign and Polish experts affiliated to universities, research units or consultancy agencies.

There is still growing number of units dedicated to information society affiliated to government bodies, regional structures of administration and NGOs. Research organizations are involved in many European projects dealing with information technologies development in CEEC. One of such projects is ESIS - European Survey of Information Society.

The Information Processing Centre as a local partner of ESIS II is responsible for monitoring and reporting the information industry development in Poland utilizing the experience in grey literature oriented information services.

INTRODUCTION

Surveys, overviews, statements, reports and numerous statistic tables and analytic documents concerning Central and Eastern Europe countries in transition period are issued as books, articles, conference proceedings, internal reports, in printed and electronic form. All these documents are a grey literature type according to redefined description during GL'97 Conference in Luxembourg.

An example of information services on political, social and business changes in CEEC is CEEBIC - Central and Eastern Europe Business Information Center of US Department of Commerce opened in 1990. It issues reports on various aspects of business in form of monthly publications and articles on CEEBIC website. Business opportunities in industries of emerging markets are analysed. Among reports concerning computer and telecommunication industry in Poland are: "IT Investment by Sector" (September 1999), "Internet Services" (July 1999), "Industry Sector Analysis of the Personal Computer Market" (March 1999), 'Telecommunications Overview" (April 1999), "TPSA Privatization" (August 1999).

United Nation Organization agendas as UNESCO, UNDP, UNIDO provide support services established in special attempt to democracy and open market processes, aimed to support changes in post-communist countries, to develop their economies and industries and social

conditions as well. Needless to mention initiatives of the Open Society sponsored by Soros Foundation.

INFORMATION SOCIETY PROJECT OFFICE

After dissemination of Bangemann Report many European Information Society initiatives and projects are aimed to monitor and report information technologies development in European Union countries and in CEEC. Most important are: MISAC - Monitoring EU-CEEC Information Society Pilot Actions. GIP - G7 Global Inventory Project. TEN-34 - Trans European Research Networks. FEMIRC (A Fellow Member to Innovation Relay Centres) - a Inco-Copernicus/Innovation Programme project is one of awareness actions of Building Information Society Action Plan for CEEC. It has three action lines: information on RTD EU programmes, innovation and technology transfer, building Information Society.

The Information Society Activity Centre (ISAC) of the European Commission manages a multiannual programme to establish the Information Society in Europe. Its part is the Information Society Project Office (ISPO), which is an intermediary between EC and external relevant organizations. The other ISAC activities are Information Society Forum and international co-operation, including co-operation with Central and Eastern Europe countries. In accession countries harmonization with European Union standards and regulations is a crucial problem, so many of internal reports, statements and recommendations deal with necessity of structural changes of telecommunications. Supporting activities are based partly on PHARE programme.

Information Society Project Office develops two Information Society Inventories. One of them is mentioned above Global Inventory and the second - European Survey of Information Society (ESIS).

At the beginning of 1997, the Information Society Promotion Office of the European Commission had launched ESIS with the objective to build an inventory of projects and promotional activities. The 3rd EU/CEEC Information Society Forum that took place in Brussels in October 1997, and at the Euro-Med Net 98 Conference which took place in Cyprus in March 1998 stated to extend ESIS to Central and Eastern European and Mediterranean countries and territories. The project ESIS II started in March 1999 and will last until 2001.

METHODOLOGY OF EUROPEAN SURVEY OF INFORMATION SOCIETY

An inventory of Information Society projects - interactive projects with a societal dimension that use innovative technologies of information and communication and provide remote access. Each national contractor collects information about such projects in his country through bilingual questionnaires distributed to project promoters in printed and electronic form. The projects and action descriptions entries are sent to central database developed by management team of ESIS II. The central database is managed on the web by SQL server. An extranet web site is installed as a technical support and communication tool for management team and national ESIS partners. Every three months, the inventory of projects is updated.

Reports on regulatory developments and liberalization evolution, existing major network facilities in public utilities, statistical analysis of the IS data and indicators and overviews on key persons and organizations are prepared on national level, then separate syntheses for Mediterranean and Central and Eastern Europe countries are compiled. They are published on European Survey of Information Society Projects and Actions Web site.

ESIS II in Poland

The Information Processing Centre (OPI) is a local partner in ESIS II project for Poland. The leading partner is the Centre for Socio-Legal Studies (PCMLP) affiliated to the University of Oxford.

The Information Processing Centre, an information agency affiliated to State Committee for Scientific Research, is a publisher and databases vendor on research policy (RTD institutions, RTD projects and results etc). OPI has strong everyday cooperation with research and information units in all the universities and RTD institutes in Poland. These institutes are main information sources for OPI.

In the initial phase of ESIS II in Poland nationwide event for representatives of information technology and telecommunication institutions has been organized. Appropriate leaflets, e-mails and information on OPI server are disseminated. Training and individual consultancy is provided by OPI staff. Feedback for corrections and addenda of data between project promoters are provided as well.

There are differences in accessibility of telecommunication and information facilities. It depends on geographical region and differs from main research and industry centres as Warsaw, Cracow, Wroclaw, Poznan, Katowice and in remote regions. Specially North-Eastern part of Poland is less developed and there ESIS II survey can be conducted mostly by frequent personal visits and by standard mail.

SOURCES OF INFORMATION ON PROJECTS

Data for inventory of IS projects in Poland will be collected on Information Processing Centre server. Inventory will be conducted and developed be ESIS II in Poland team. Mostly electronic questionnaires will be used. Downloaded printed copies will be accepted as well. During two years about 100 projects and actions conducted in Poland or in co-operation with Polish partners are estimated to be included into ESIS II inventory. Estimation is based on evaluation of OPI databases on projects, experts and research institutions. Another reason of this estimation is number of projects in ESIS I database accessible on ISPO server. researchers, industry and administration. Projects under estimation started from early 1990s and were completed in 1997, 1998 or 1999. INCO-Copernicus and other European projects are the ones which started in the middle of 1990s mainly in 2 years period. KBN and university grants in average are scheduled for 2 years as well. Nation-wide projects founded by public administration or industry, depending of comprehensiveness and availability of financial support take time of 4-5 years, even longer. Increase will be stimulated by Information Society Programme in European Union Fifth Framework Programme for Research, Technology and Demonstration, in which partners from accession countries can participate on equal rules as participants from European Union, and growing interest among researchers, industry and administration.

The most frequent topics of current information society related projects identified by local ESIS II project team in the Information Processing Centre are Geographic Information Systems, CAD/CAM, technology convergence. Databases and specialised information services accessible by Internet are frequent mentioned by respondents. Projects are to be implemented in research and education, local administration, e-business, applications in various branches of industry and public services, media and entertainment, and telework. New are telemedicine, "intelligent house and office", internet services for disabled.

The estimated percentage of reports by sector is:

40% - Research and Education
20% - Media and telecommunications
30% - Business (including e-commerce)
10% - Administration

Potential respondents are identified through such sources as Information Processing Centre databases on Current and Completed Research SYNABA and POLISH RESEARCH DIRECTORY, through "INFOBAZA" inventory, Database of Grants funded by the State Committee for Scientific Research, proceedings and programmes of IS events, www services and articles in journals. CORDIS Database includes information of Polish projects conducted in framework of EU programmes.

INFORMATION SOURCES ON REPORTS

Data for reports are accessible from governmental and research documentation, Polish Press Agency, professional and scientific journals, WWW information services, specialized services as State Statistical Office, relevant databases and information services on national, regional or on European level.

Key organizations in the Information Society technologies and content sector in Poland include governmental and local administration, private business (Small and Media Enterprises), industry and services, research and technology institutions and associations, and the media sector.

Governmental information

Governmental statements on information society are those of the State Committee for European Integration, State Committee for Scientific Research, Ministry of Posts and Telecommunications and Ministry of Internal Affairs and Administration. The governmental document prepared in September 1998 by Ministry of Posts and Telecommunications is "A draft of telecommunications market development for 1998-2001" which includes the estimates of telecommunications development. Special departments or sections dedicated to Information Society development have been set up in almost all ministries. A lot of internal reports, surveys etc deal with EU integration. Special attempt is given to harmonization of the law before accession to EU, privatization of telecommunication structures, TV digitalization, electronic commerce development.

The Ministry of Internal Affairs and Administration is the national coordinator of Y2K Problem in Poland. The National Action Plan concerning Y2K was undertaken by the Council of Ministers on 24 march of 1999 and all the actions on governmental lever are planned and monitored. The Ministry's Governmental Registers, Telecommunications and Information Technologies Department conducts the information service on Y2K problem. It includes updated reports on the government's administrative preparations for Y2K, a timetable of deadlines for Y2K preparation in power supply, telecommunications, transport, banks, finances, health and social care sectors and other related issues. A lot of "grey documents" on Y2K are disseminated as governmental documents, surveys and guidelines. The State Committee for Scientific Research (KBN) as a governmental body responsible for science in Poland prepared a document "Year 2000 – The preparation of research computer infrastructure" based on inquiries among most important WANs, MANs and LANs providers concerning software and hardware quality.

Nongovernmental organizations

Relevant associations, foundations, chambers of commerce and industry, research and professional journals usually prepare yearly reports, surveys and rankings.

The representative examples of non-governmental organizations are the Polish Chamber of Information Technology and Telecommunications, Polish Chamber of Electronics and Telecommunications, the Association of Polish Software Market "PRO", the Polish Forum of Information Society, Cities in the Internet Association, the Polish Information Processing Society, Society for Open Systems Development, Scientific Society for Information, The Polish Commercial Internet Forum.

Promotional activities in Poland

Poland is an active member of many Information Society initiatives on international, European, national, regional and local level. The State Committee for Scientific Research (KBN) coordinates the growth of the Polish Internet and supports development of computer and Internet facilities for scientific community in Poland. On KBN home page, National Contact Point for Global Information Society service is conducted by Information Technology Systems department.

Main conferences and exhibitions, both international and nationwide, dedicated to computer industry are "Computer-EXPO", "COMNET", and "POLMAN", "Infosystem" conference, "Komtel" and "INFOBAZY'99".

Publications

Following the First Congress of the Polish ICT in December 1994, the Polish report was compiled and focused on the most important issues of IT development in Poland. The Second Congress, 30 - 2 December 1998 in "Treaty for building the Information Society in Poland" stresses the challenge for all the citizens. "Information Technology Development in Poland" report after The Second Congress is available in a printed version and on the Internet underlines obstacles of information society building in Poland.

Among computer and telecommunications magazines and journals in Poland are "Teleinfo", "Computerworld", "Internet", "Enter", "Networld". Many of them provides yearly rankings of companies. "Information Society Forum" ("Forum Spo lecze Istwa Informacyjnego") is an appendix to "Teleinfo".

The biggest computer and telecommunication companies active on the Polish market are mainly branches of giant foreign companies but the percentage of Polish businesses is still growing. In the new edition of "TELEINFO 500 – Polski Rynek Teleinformatyczny 1998. Raport", edited in June 1999, the 600 largest Polish computer companies, 200 computer companies operating in Poland, 150 largest Polish telecom companies and 25 top telecom companies operating in Poland in 1998 are listed.

Electronic sources

The expansion of web services in Poland is significant. The most popular to date are basic services such as e-mail and WWW pages – which are mainly targeted at institutional promotion. More advanced services such as electronic commerce are at the primary stage. The longest established are electronic book stores as "Ksi" Gki OnLine" (Books OnLine) – a part of Polska Online portal. A number of portals is still growing. ". The next is Onet.pl of Optimus S.A. and new portals of Internet Venture Poland (IVP) and Telekomunikacja Polska S.A.

For reports on regulations web pages of the Polish Parliament, The Ministry of Posts and Telecommunications, Ministry of Internal Affairs and Administration, The State Committee for Scientific Research and Committee for European Integration are frequently browsed.. Every telecom operator and Internet provider has established web service where up to date information on new technologies, telecommunication services, tariffs etc is presented. Databases are provided by research

institutions and libraries via internet. Information Processing Centre conducts on its own internet server bilingual databases as SYNABA, Dissertations and Polish Research Directory, accessible free of charge for remote users.

Conclusions

ESIS II survey can be a virtual grey literature collection for a growing number of noncommercial publications, internal documents and electronic multimedia resources concerning information and telecommunication market in Central and East Europe. Together with such a collection of documents and records gathered from EU countries during the first phase of ESIS project and with Mediterranean countries information resources it seems to be a very comprehensive inventory of European contribution to worldwide challenge in computer and telecommunication technologies. Much more effort should be given to disseminate these valuable knowledge resources broadly and provide access in "user friendly" way for different user groups of the global information society.

REFERENCES

Third EU/CEEC Information Society Forum. Brussels, 9-10 October 1997. Proceedings, European Parliament/European Commission, Brussels

Building the European Information Society for us all. Final policy report of the high-level expert group. (1997).European Commission. Brussels.

Coopers & Lybrand Technology Consultants (1998) *Impact of the enlargement of the European Union towards the associated central and eastern European countries on RTD-innovation and structural policies. Final report.*European Commission, Brussels

INFOBAZY'99. Bazy Danych dla Nauki. Gdansk [30 August - 1 Sept. 1999, Proceedings] Politechnika Gdanska, Gdansk

Polish Research Directory 1999. Information Processing Centre. Warsaw. Available http://www.opi.org.pl/en_baz.htm

Scientific Research in Poland.(1999) Information Processing Centre. Warsaw. Available http://www.opi.org.pl/en_baz/htm

Market Research. Poland. Central and Eastern Europe Business Information Centre. US Department of Commerce. Available http://www.mac.doc.gov/eebic/country/poland/market2.htm

Ministerstwo Lacznosci (1998) *Zarys strategii rozwoju rynku telekomunikacyjnegona lata 1998-2001".* Available http://www.ml.gov.pl

"Computerworld Online" Available http://www.computerworld.pl

"TeleInfo OnLine" Available http://www.teleinfo.com.pl

Use Pattern of Grey Literature in Rural Development An Indian case study

By T. Rama Devi & Anil Takalkar
Centre on Rural Documentation, National Institute of Rural Development
Rajendranagar, Hyderabad – 500 030. India

ABSTRACT

Rural Development Literature in India is growing as the expenditures meant for the sector are increasing under every Five Year Plan. There are number of institutions, NGOs and individuals undertaking research on various aspects of development. National Institute of Rural Development, Hyderabad, India is the apex research, training and consultancy organisation in India on rural development. It publishes several periodicals, monographs, proceedings of workshops and seminars, case studies etc. Journal of Rural Development (JRD) published by NIRD since 1982 covers subjects like agriculture, economics, political science, psychology, sociology, management and gender studies etc. This refereed and peer reviewed journal carries both theoretical and empirical research based articles pertaining to India and other developing countries. Besides, it publishes case studies, research notes, book reviews etc.

The present study examines the distribution, classification, and categorisation of articles published in JRD during 1982-1997 and the number of citations relating to grey literature. Based on the method of citation analysis, the use pattern, accessibility and availability of grey literature to the contributors of the articles were studied. Also an analysis of the formats of the grey literature pertaining to India covered in the two International Abstracting Services, Rural Development Abstracts by CABI, U.K. and Abstracts of Tropical Agriculture by Royal Tropical Institute, Netherlands was made and compared with that of JRD for the period 1982-97.

Study concludes that citation pattern is similar in theoretical and empirical articles. Increasing number of citations of grey references indicates the importance of and demand for grey literature. In order to help improve their access and awareness and consequent supply to enhance their utility, an action plan is suggested. Under this plan, a special programme of networking of socio-economic institutions in India especially the 25 State Institutes of Rural Development (SIRDs), spread across the vast country is to be initiated.

As a part of this activity, various government agencies, research and training institutions, NGOs, universities, banking institutions are being requested to supply their publications / list of publications which will form an input to the network. A separate database of Grey literature is being created at NIRD to make it available through Internet (www.nird.org).

INTRODUCTION

Grey Literature (G.L) plays an important role, as a means of information dissemination, needs no emphasis. The importance of such fugitive documents was brought to international attention by the report of the DEVSIS Study team (IDRC, 1975). To quote the report, " the G.L is made up of unpublished working papers, feasibility and pre-investment studies, theses, research reports and documents of governments and international organisations that are not widely

disseminated". The main characteristics attributed to such fugitive or grey literature are non-availability through normal trade channels, limited readership and limited distribution.

Development in general and rural development in particular assumes prime position on the agenda of activities of governments of various developing countries including India. Various inter-governmental organisations, non-governmental organisations, research and training institutions, development aid agencies etc. all over the world are engaged in rural development activities. Development initiatives of these organisations are manifest in a number of strategies and programmes. In the process, these organisations generate enormous literature that is mostly in the form of grey literature. According to an estimate, 5000 documents per year are produced in India alone on aspects directly relating to rural development by various institutions and organisations (Raju, K.A. 1994) .

In view of the increasing output of G.L in the field of rural development, it is often assumed that it is also being widely gathered, collated, documented or processed for use by the development community engaged in planning, policy making, training or implementing rural development programmes and projects.

As stated by the DEVSIS report major users of any development information system are those people who are themselves involved in the mission. It has been observed that the generators of literature in the development science are themselves the users of the system. Interestingly, this phenomenon is more conducive for reprocessing, repackaging for effective dissemination in a networked environment.

In India (as also in many developing countries), "development" is achieved through the planning processes which are commonly known as Five-Year Plans. The successive Five-Year

Plans initiated in 1951-52 in India are the major instruments for achieving the goals of development. Currently Rs.100 million are earmarked every year towards rural development which is slated to be doubled during the ongoing Five Year Plan (1997-2002). In formulating development plans, there is the need for a wide range of information covering many disciplines. Simultaneously, many institutions came into existence to undertake development research to provide new insights into various problems like growing poverty, unemployment, food security and for taking appropriate decisions. Generation of new knowledge through research in development is an essential pre-requisite and forms the basis of formulating and implementing development plans and programmes. National Institute of Rural Development at Hyderabad is one such leading institution that came into being in the beginning of 60's to undertake development research and training and help provide decision support to the development community.

National Institute of Rural Development (NIRD) is an apex organisation under the Ministry of Rural Development, Government of India. The mission of NIRD is to facilitate rural development in India to improve the quality of life of the rural poor through various programmes and projects. In order to achieve this mission, NIRD undertakes training, research and consultancy and information dissemination activities.

As a result of its varied activities over the years, NIRD has developed a vast reservoir of contemporary knowledge in rural development. As a part of its information dissemination function, NIRD brings out Journal of Rural Development, Newsletters, Research Highlights, Recommendations of Seminars and Workshops, Research Reports, Monographs, Occasional papers, etc.

JOURNAL OF RURAL DEVELOPMENT (JRD)

Journal of Rural Development is an official journal of NIRD and is seen as a representative journal in the field of rural development in India. NIRD was established in the beginning of 60's under the name of National Institute of Community Development (NICD) to reflect the then government's policy and activities on community development. As a vehicle to disseminate information about its contributions to the field, two journals were simultaneously being published. i) Behavioural Science and Community Development and ii) Panchayati Raj Digest. In the middle of 70's under the thrust of second development decade, a new policy paper on Rural Development was released by World Bank in the year 1975. Integrated rural development as a policy was initiated and India also adopted the same. In keeping with the developments, the name of NICD was changed to National Institute of Rural Development (NIRD) by the end of 70's. Consequently NIRD merged the two journals into one and started issuing the JRD from the year 1982 to 1997. Thus JRD in a way reflects the transition of development policies pursued by India. It acquired slowly a prestige and standing not enjoyed by any other contemporary Indian journal in the complex subject of rural development.

Rural development being interdisciplinary in nature, the coverage of the journal includes subjects like agriculture, economics, political science, local government (Panchayati Raj), psychology, sociology, management, gender, environment, energy, poverty, and industry. This refereed and peer reviewed journal carries both theoretical and empirical research based articles pertaining to India and other developing countries. Besides, it publishes case studies, research notes, book reviews etc. The journal aims at promoting study and research in rural development. It seeks to uncover links between the social sciences and rural development and to forge and strengthen them where necessary. It provides a forum for exchange of views between various social science disciplines and the policy makers, planners and the executives concerned with rural

development. Hence, for the purpose of the study, we have selected JRD since its inception 1982-97.

OBJECTIVES:

The objectives of the present study are

1. To examine the use pattern of grey literature with special reference to Rural Development.
2. To compare the coverage of grey literature on India in international abstracting services
3. To suggest an action plan for effective bibliographic control of grey literature.

METHODOLOGY

As the aim of our study is to understand the pattern of usage of grey literature by the social scientists with special reference to the faculty of NIRD, we have examined the data relating to the type of articles that are published in the journal during 1982 to 1997, their number, their focus (empirical or otherwise), the number of citations, the category of citations and their distribution together with many other details like classification of the subjects of the given citations, categorisation of the different author groups and the contribution of NIRD faculty.

As the citation analysis is one of the accepted methods to assess or determine the actual use of documents and resources, this method was chosen to investigate the use of grey literature. Analysis of such citation in a given scientific communication can reveal useful information like the relative use of different categories of documents such as conferences/ seminars / workshop proceedings, research reports, Committee and Commission reports, case studies, annual reports, dissertations etc. Additionally the data that was derived from JRD pertaining to grey literature was compared with that of the two international abstracting services 'Rural Development Abstracts' of

CABI, U.K. and 'Abstracts of Tropical Agriculture (TROPAG)' brought out by Royal Tropical Institute, Netherlands. We have also seen from this two international abstracting journals, the coverage of grey literature for the period 1982-97 pertaining to India. This was done in order to know the extent of coverage as well as the categorisation of materials to compare with the data obtained from JRD.

DATA ANALYSIS

JRD was published as a bimonthly during 1982-93 (12 years) and from 1994, the periodicity was changed to quarterly. Thus a total number of 88 (72+16) issues were checked and data collected. It is observed that a total of 596 articles were published during the 16 years. From these (a) articles based on empirical studies (b) articles contributed by single author and belong to NIRD or otherwise (c) articles with citations were differentiated. The following table gives the distribution as stated above :

Table I: Distribution of articles according to the given category

	Total No. of Articles : 596	
	Number	Percentage
Empirical based Studies	283	47.4%
Non-empirical based studies	313	52.6%
With citations	486	81.6%
Without citations	110	18.4%
Contributed by single author	333	55.8%
Contributed by more than one author	263	44.1%
Articles by NIRD Faculty	241	40.43%
Articles by Non-NIRD	355	59.5%

The above data gives an idea that there is more or less an equal distribution of the empirical and non-empirical based articles. The articles contributed by single author and those coming from outside NIRD are significant. We can infer that the journal's standing among the development community is increasing over the years attracting contributions from outside NIRD and most of the articles are having citations, that include a sizable grey literature.

CATEGORIES OF CONTRIBUTORS

We have examined the categories of persons who contributed the articles from NIRD and outside NIRD. The total number of contributors from NIRD is 362 and the rest are from different organisations. The following table gives the categorisation of the authors of the articles from NIRD and others. The number of contributors do not tally with the number of articles as some of the articles are written by more than one author.

Table II: NIRD and Non-NIRD Contributors

NIRD Contributors

Senior Professionals	Middle level	Junior Faculty	Total
151 (41.7%)	126 (34.8%)	85 (23.4%)	362

Non-NIRD Contributors

Academicians	Functionaries	Planners	Administrators	Others	Total
518 (86.9%)	48 (8%)	41(6.8%)	(2.68%)	7 (1.01%)	630

The category of senior professionals from NIRD consists of Directors, Deputy Directors of different centres. Faculty members like Assistant Directors, Senior Research Officers are categorised as middle level professionals and Research Associates are grouped into junior faculty.

As regards non-NIRD contributors, those from research and training institutions, and universities are grouped as academicians. Government functionaries like Block Development Officers, Project Directors of District Rural Development Afgencies (DRDAs), are categorised as functionaries, officials associated with planning and designing of development programmes, Chief Executives of different institutions have been categorised under planners. IAS officers who have contributed in a significant number have been categorised separately as administrators. Contributors who have not mentioned their designation or affiliation and could not be classified into any of the given groups, have been categorised as others.

The occupational category-wise distribution of the NIRD contributors shows that the senior faculty is contributing more often to JRD than the other categories. This can be attributed to the fact that senior faculty in NIRD are involved relatively more in research activities than the juniors. Also the projects that are designed in NIRD are directly related to either the rural development programmes of the Ministry of Rural Development, Government of India or of the thrust areas of the institution. In both the categories, the senior faculty are more involved. As a result, the outcome of the projects and the programmes heavily depended upon them.

As regards the distribution of outside contributors is concerned, the academic community occupy the first position followed by the others. Most of these academicians belong to the universities and other development oriented research and training institutions of the country. As stated earlier, the enormous development funds that are spent in the name of rural development by the Central and State governments increased the number of projects undertaken resulting in the production of different project reports and other attendant publications. These are mostly handled by the academicians which is reflected in their contribution. It is also seen from the data that authors from different backgrounds and institutional affiliations contributed to the output of literature in varying degrees.

Table III : Distribution of citations and percentage of grey references

	1982-86	1987-91	1992-97	Total
Total no. of cited references	572	2292	1903	4767
Total no. of grey references among the cited	148	597	581	1326
Percentage	25.8%	26%	30.5%	27.8%

The Table-3 shows distribution of articles in a five year interval along with the number of citations and the percentage of grey references among the cited. The Table-4 gives the idea about the grey references used by NIRD faculty. It is interesting to know that there is an increasing trend in the citation of grey literature progressively over the years. This can be attributed to two factors : the increasing production of grey literature and its relative availability. Similarly, the citations of grey literature by the faculty of NIRD reveal that the senior faculty perhaps has more awareness as well as access and availability of the grey literature than the others. Senior Faculty, by virtue of having information about the various programmes and projects that are ongoing in the country in their field of specialisation as well as due to their frequent visits to the headquarters of the several ministries in the state and central government, are relatively well off in their awareness of different kinds of material. Also 'invisible college' concept works equally well with Social Scientists.

Table IV : Type of Grey Literature used by NIRD Faculty

	Senior Professionals (Dir.,D.D.- 262)	Middle level Faculty (SRA, A.D-160)	Junior Professionals (R.A's-45)
Conferences/ Workshop	50 (19%)	19 (11.8%)	6 (13.3%)
Committee / Commission Reports	57 (21.7%)	24 (15%)	8 (17.7%)
Case Studies	7 (2.6%)	2 (1.3%)	1 (2.2%)
NIRD Reports	29 (11%)	30 (18.7%)	5 (11%)
Dissertations	18 (6.8%)	1(0.6%)	2 (4.4%)
Others *	101 (38.5%)	84 (52.5%)	23 (51.1%)
Total	**262**	**160**	**45**

* Annual Reports, Project Reports, Survey reports, Review papers, Evaluation reports, Progress report, guidelines etc.

SUBJECT COVERAGE OF GREY LITERATURE

We have also analysed the kind of subjects the cited grey literature cover as well as their type. The following table gives the various subjects covered in the articles contributed to JRD as well as the corresponding percentage of citation belonging to that topic of the contribution.

Table V : Subject Coverage of Grey Literature

Subject of Articles		Subject of Citations given in the articles	
Rural Development	116 (19.4%)	Rural Development	231 (17.4%)
Agriculture	90 (15%)	Agriculture	135 (10.1%)
Local Government/ Panchayati Raj	65 (10.9%)	Local Government	147 (11%)
Gender	38 (6.3%)	Gender	55 (4.1%)
Industry	31 (5.2%)	Industry	40 (3.0%)
Forestry	4 (0.6%)	Forestry	43 (3.2%)
Poverty	30 (5.0%)	Poverty	40 (3.0%)
Total	596		1326

Note : Remaining subjects like energy, employment, livestock, watershed, finances and many others covered are not given here for want of space.

JRD by virtue of its stated focus attracted articles more in number that relate to rural development followed by allied subjects like agriculture, local government, and gender studies. Citations also follow more or less the same pattern.

COMPARISON WITH OTHER SERVICES

In order to understand the kind of accessibility and availability of grey literature for the purposes of citation as well as bibliographic control, we examined two major international abstracting services namely Rural Development Abstracts and Abstracts of Tropical Agriculture. We have seen the coverage for the same period 1982-97 by using the CDs available under the name AgECONCD and TROPAGCD respectively. We have also restricted our search from the

CDs to the grey literature on India without any restrictions on place of publication. The following table gives the distribution of grey citations as against the total references in case of JRD and total entries in case of two CDs pertaining to India.

Table VI : Coverage of grey literature in different databases.

	JRD	AgECONCD	TROPAG
Total references for the period 1982-1997	4767	21887	14329
Grey references	1326 (28%)	3803 (17.3%)	1241 (8.6%)

It can be seen from the above table that the coverage of grey literature is insignificant, bringing home the point that there is a lack of bibliographical control of G.L. The findings are supported by a comparative study of three international abstracting services viz. Rural Development Abstracts, International Development Abstracts and Abstracts of Rural Development in the Tropics conducted earlier. The study showed that the coverage of grey literature in these abstracting services was inadequate compared to the other types of literature abstracted. It was also observed that the extent of duplication among the services was very insignificant (Raju, K.A. 1991)

Table VII : Types of materials covered under grey

	JRD	AgECON	TROPAG
Conference Proceedings	208 (15.6%)	535 (14%)	406 (32.7%)
Research Reports	44 (3.3%)	37 (0.9%)	3 (0.24%)
Thesis / Dissertations	102 (7.6%)	198 (5.2%)	10 (0.8%)
Case Studies	45 (3.4%)	872 (22.9%)	326 (26.2%)
Annual Reports	38 (2.8%)	671 (17.6%)	23 (1.8%)
Guidelines	10 (0.75%)	160 (4.2%)	51 (4.1%)
Manuals	16 (1.2%)	330 (8.6%)	104 (8.3%)
Project Reports	31 (2.3%)	5 (0.13%)	2 (0.16%)
Training programmes	6 (0.45%)	90 (2.3%)	36 (2.9%)
Others *	826 (62.2%)	905 (23.7%)	280 (22.5%)
Total	**1326**	**3803**	**1241**

* Working papers, Action plans, Evaluation Reports, Issue papers, Occasional papers, State of the art, Translation progress reports etc.

We have also seen the different types of grey literature (Table 7) such as conferences / seminars / workshop reports, research reports, case studies, thesis, dissertations, annual reports etc. covered in order to know the reach of these abstracting services in the bibliographical control of the grey literature emanating from India.

It is observed that the abstracting services are able to cover more the conferences/ seminars literature, case studies, dissertations. This may be because that the conferences and seminars are publicised more by the organisers nationally and internationally and the primary periodicals cover more case studies, list of dissertations. Also there are several other directories available which give the list of various organisations that operate in rural development. Information about the organisations facilitates getting annual reports relatively easier than any other documents. As a result, we find more coverage of such documents in the CDs and also their accessibility is reflected by their citation in the contributions to the JRD. However in the final analysis, there is still a need to have an indepth study of other abstracting services of national and international organisations in order to study the coverage of the grey literature, and their use pattern. Meanwhile, we are implementing an 'action plan' to increase the awareness and accessibility of grey literature taking advantage of the advances in information and communication technologies.

ACTION PLAN

1. A Computerised Library and Information Clearinghouse(CLIC) has been setup at CORD with a view to provide information on different aspects of rural development at a centralised location. The documents / information collected under CLIC falls mostly in the category of grey literature such as research reports, evaluation reports, case studies, success stories, conference / seminar proceedings, etc. Special efforts would be made to enrich the grey literature collection by approaching various government agencies, research and training institutions, NGOs, Universities, banking institutions and the like, who are the generators of Grey Literature. CLIC is presently accessible on internet (URLs: www. PanAsia.org.sg/nird and www.nird.org)

2. State Institutes of Rural Development (SIRDs) are the apex training and research institutions at state level. There are 25 such SIRDs. These are being brought into a network (Annexure–I) and would be assigned the task to collect the rural development literature generated at the state level and also in regional languages .The SIRDs would act as the depositories for such literature and send the information to CLIC for wider dissemination. The CLIC is programmed to work as a focal point for the grey literature on rural development at the national level on the lines of the European SIGLE.

3 . Internet would be searched for web sites relevant for the development activities. Material of interest from these sites, especially the grey literature would be identified and downloaded and made available under CLIC for the benefit of the development community.

4. International abstracting services in the discipline, which cover grey literature would be monitored on a continuous basis to identify the grey literature covered and to procure the same as far as possible.

CONCLUSIONS

It can be concluded from the above analysis that grey literature in various types, even in the present day context continues to be one of the important media of transfer of information and a source of recorded knowledge but eludes comprehensive bibliographical control. Increasing number of citations of grey references indicates the importance of and demand for grey literature. Though theoretical and empirical studies form distinct types of intellectual contributions, they follow similar pattern of citation of grey literature according to our study. It is observed that only those publications are usually cited most which are published in the preceding five years. The awareness about and accessibility to grey literature is responsible to a considerable extent as can be noted from the citations At the national level, inadequate bibliographical control of the grey literature is one of the major factors for this situation. Whereas at the global level as evidenced from the databases like RD Abstracts and Abstracts on Tropical Agriculture, the coverage of grey literature though appears to be numerically more but compared to other references in the databases, it is not very significant.

The implication of the study is that there is an urgent need at the national level, to coordinate the efforts of all major institutions/individuals in a given discipline which contribute significantly in the generation of grey literature. A centralized identified agency may be assigned with the responsibility of coordinating these activities on the lines of the European example SIGLE. This to a considerable extent may establish bibliographical control over grey literature and improve the awareness about, access to and consequent supply of grey literature among the concerned at the regional and national level. The initiatives of different countries in this regard may, in turn form the inputs to the agencies which are engaged in development of databases in respective disciplines with a world wide coverage to achieve bibliographic control at the global level.

With the advent of Internet, the concept of development as a global concern acquired new meaning. Many institutions are developing their own websites, homepages in order to publicise their activities, expertise, publications and products. These will definitely improve the awareness and accessibility of grey literature. Simultaneously several networks at regional or local level are emerging on the Internet, specialising in making available information generated by the participating institutions of different countries. One such is Pan Asia Network (PAN) sponsored by IDRC, Canada. Through PAN, the International Development Research Centre is aiming at promoting collaboration in research and development by connecting individuals and institutions for knowledge sharing across Asia. NIRD is also a member of PAN and it is deriving the benefit from the co-members of network in order to access information on grey literature and other useful network groups.

ACKNOWLEDGEMENT

We acknowledge with thanks Mr. R.C. Choudhury, Director General, NIRD for giving permission to submit the paper to the GreyNet 99 Conference. We also wish to acknowledge Mr. L.S. Ramaiah, Ex – Librarian of Central Institute of English and Foreign Languages, Prof.N.Laxamana Rao, Dept. of LIS, Osmania University, Hyderabad and Mr. K. A. Raju, Director (CORD), NIRD for their valuable comments and suggestions on the paper before submission to conference. The help rendered by Mr.P. Sudhakar in the tabulation of vast data before its analysis is gratefully acknowledged.

REFERENCES :

1. DEVSIS Study Team . "DEVSIS- Preliminary Design of an International Information System for the Development Sciences ". Ottawa, IDRC, p.18, 1976.
2. Raju, K.A. : Rural Development Literature : Preliminary results of a comparative study of three global abstracting services. IAALD Quarterly Bulletin, 36(4), 1991, 227-231p.
3. Raju, K.A. : Rural Development Literature : A Quick Estimate, Paper presented at the Review Meeting of an IDRC project on DEVINSA, New Delhi, 1994.
4. Commission of the European Communities, Brussels. "An up-to-date System for Information on Grey Literature produced in Europe (SIGLE)". A brochure, n.d.

ANNEXURE – 1

Figure 1 : Proposed Network of State Institutions of Rural Development (SIRDs) Connected

APARD, Hyderabad, (Andhra Pradesh)	SIRD, Shillong (Meghalaya)
SIRD, Itanagar (Arunachal Pradesh)	SIRD, Aizawal (Mizoram)
SIRD, Guwahati (Assam)	SIRD, Kohima (Nagaland)
BIRD, Patna (Bihar)	SIRD, Bhubaneswar (Orissa)
SIRD, Ahmedabad (Gujarat)	PUN-SIRD, Nabha (Punjab)
HIRD, Nilokheri (Haryana)	HCM-RIPA, Jaipir (Rajasthan)
HIPA, Shimla (Himachal Pradesh)	SIRD, Karfectar (Sikkim)
IMPARD –Srinagar (Jammu & Kashmir)	SIRD, Maraimatainagar (Tamil Nadu)
ANS-SIRD, Mysore (Karnataka)	SIRD, Agartala (Tripura)
SIRD, Kottarakara (Kerala)	DDU -SIRD, Lucknow (Uttar Pradesh)
MG-SIRD, Jabalpur (Madhya Pradesh)	SIPRD, Calcutta (West Bengal)
YASHADA, Pune (Maharashtra)	SIRD, Goa
SIRD, Imphal (Manipur)	

How do various Fugitive Literature Searching Methods Impact the Comprehensiveness of Literature Uncovered for Systematic Review?

Diane Helmer BFA MLIS
Isabelle Savoie MD MHA
Carolyn Green MSc

BC OFFICE OF HEALTH TECHNOLOGY ASSESSMENT

INTRODUCTION

The systematic review is a vital tool for health care providers, researchers, and policy makers, allowing efficient and effective review of research evidence on which appropriate policy or treatment decisions can be based. The systematic review uses explicit methods to evaluate and synthesize information in a manageable form so the current state of knowledge on the effectiveness of an intervention can be understood (Chalmer, 1995).

For researchers conducting systematic reviews and aiming to reduce publication bias, the nature of fugitive literature presents a number of problems not the least of which is to define its limits. Although considerably debated, the most authoritative definition of fugitive literature remains broad (GL'97, 1997). In addition, with the advent of advanced information technologies and the Internet, new forms of fugitive literature are being introduced and becoming more widely used (Caroll, 1997). As well, more organizations and writers have access to desktop publishing software. Overall, the fugitive literature is increasing at an exponential rate. In fact, according to one estimate, its rate of growth is three to four times that of conventional literature (Farace, 1997).

Some countries, particularly the US and the UK, have developed their fugitive literature resources and made them easier to access. This is not the case in many other parts of the world. Where no centralized agency exists to assist in its distribution, the acquisition of grey literature remains problematic (Alberani et al, 1990).

Although the accessibility of various kinds of fugitive literature and document delivery has been facilitated by the development of the Internet and resource-sharing initiatives between libraries, increased access to information through these means has only partly kept pace with the mass of information.

To establish standards for searching the literature a number of guidelines have been produced (Mulrow et al, 1997)(Goodman C, 1993).

Search protocols based on these guidelines will generally require researchers to:

- search the electronic literature databases;
- search the grey literature;
- hand-search key journals;
- scan reference lists;
- identify conference proceedings;
- consult other researchers in the topic area.

The assumption has been that this series of steps will reduce publication bias and improve the ultimate credibility and precision of the systematic review (NHS Centre for Reviews and Dissemination, 1996)(Stern et al, 1997). To date however, very little evidence appears to have been gathered to show the effectiveness of some of these steps.

In the light of these difficulties, this study was designed to determine if and to what extent various fugitive literature searching methods affect the comprehensiveness of literature uncovered for systematic reviews.

METHODS

The British Columbia Office of Health Technology Assessment (BCOHTA) at the University of British Columbia undertakes health technology assessment research, whose primary purpose is to examine scientific evidence on the effectiveness of current and developing health technologies. BCOHTA bases its appraisals on comprehensive evaluations of the scientific evidence, and consequently, the literature searches for each project can be intense.

In 1993, BCOHTA adopted the use of databases to manage the bibliographies associated with each project. Since 1993, the Office uses Dbtext Works by Inmagic Software Inc. to produce bibliographies.

In undertaking systematic review, a BCOHTA search is normally divided in two parts: the electronic literature (or Level 1 search) and the fugitive searches (New Zealand Health Technology Assessment, 1999).

The electronic literature search comprises detailed search strategies in databases such as Medline, Embase, HealthStar, Biosis, PsychLit or Current Contents. The selection of the databases is directly dependent on the topic of the systematic review. When the electronic literature search is complete, BCOHTA researchers select articles that meet the pre-defined inclusion and exclusion criteria, and this material is retrieved.

The second part of BCOHTA's search strategy is the fugitive search. In the fugitive search:

- subject specific and specialized databases are identified and searched;
- the Internet is searched by using meta-indexes such as OMNI. [2] Search terms similar to those in the electronic search strategies are applied to these meta-indexes. A select number of search engines are used to locate organizations and researchers;
- relevant non-indexed and indexed journals are identified. These journals are then handsearched to ensure comprehensiveness.
- The search for fugituve literature also makes use of directories to indicate organizations undertaking topic-specific research. From this point, relevant organizations and researchers can be identified and contacted.

For the purpose of this study, two systematic review projects were used as sources of data. The projects chosen were topics for which it was feasible both to apply a prospective methodology and to undertake a substantive fugitive search.

The systematic reviews aimed to gather as many randomized controlled trials as possible on their respective topics. A prospective analysis was applied to the literature items retrieved by both electronic searches and fugitive searches.

The fugitive search strategies for each project are detailed in **Tables** 1 & 2. For the electronic search, a strategy designed to be consistent across both projects was applied to Medline, Embase and Current Contents.

Items meeting the inclusion criteria for each project were retrieved and entered in a BCOHTA project-specific database using DbText. Each item was coded to describe how it was found . Items previously uncovered through the electronic literature search were downloaded directly from their respective database into DbText, and the unique identifiers assigned by Medline and

similar databases allowed such records to be differentiated from the fugitive records. A second code was assigned to describe the type of item uncovered. The item types were coded independently by one librarian and by one researcher. The coding results were then compared and any differences between the researcher and librarian's coding were resolved by discussion.

PRELIMINARY RESULTS AND DISCUSSION

In total, 1034 items were retrieved. Of these, 302 (29.2%) were identified through the fugitive search. Given the difficulties encountered in acquiring the fugitive literature, this may still represent an underestimation of the effectiveness of fugitive searches. For example, not all researchers have responded to our inquiry regarding trials and other relevant information.

Most of the fugitive literature uncovered takes the form of traditional reviews. The indexing and precision of randomized trials in the specialized databases may explain this finding. In addition, fugitive items are less likely to offer abstracts. Therefore, it is often harder to determine if the items meet the inclusion and exclusion criteria before ordering the item. The search of specialized databases was the most effective method of identification, followed by the review of reference lists.

As stated above, these results are preliminary. We are still in the process of following-up on a few letters that were sent. Also, some of the items selected have been requested through inter-library loan have not been received yet. We are waiting for a response from the Inter-Library Loans department to confirm whether or not these items are available.

CONCLUSIONS

The preliminary results show that the fugitive search can be an effective tool for uncovering material for the systematic review. The quality of the fugitive literature items uncovered as well as the likelihood of that literature affecting the results of the systematic review, is yet to be determined.

REFERENCES

Alberani V, de Castro Pietrangeli P, Rossi Mazza AM (1990) "The use of grey literature in health sciences: a preliminary survey", *Bulletin of the Medical Librarians Association*, Vol 78 No 4, p.358

Caroll BC, Cotter GA (1997) "A new generation of grey literature: the impact of advanced information technologies", *Publishing Research Quarterly*, Vol 13 No 2, pp. 5-14

Chalmer I, Altman DG, editors (1995) "Systematic reviews" BMJ Publishing Group, London, pp.1-2

Farace DJ (1997) "Rise of the phoenix: a review of new forms and exploitations of grey literature" *Publishing Research Quarterly*, Vol 13 No 2, pp. 69-76

GL'97 Luxembourg Convention (1997) "Perspectives on the Design and Transfer of Scientific and Technical Information" Third international conference on grey literature. Luxembourg, 13-14 November 1997

Goodman C (1993) "Literature searching and evidence interpretation for assessing health care practices" Swedish Council on Technology Assessment in Health Care, Stockholm, pp. 16-32

Mulrow CD, Oxman AD, editors (1997) "Cochrane Collaboration handbook" In: The Cochrane Library [database on disk and CDROM]. Oxford, Update Software, 1997, Issue 4

New Zealand Health Technology Assessment (1999). "NZHTA search protocol" Available from: URL: http:// nzhta.chmeds.ac.nz/nzhtainfo/protocol.htm

NHS Centre for Reviews and Dissemination (1996) "Undertaking systematic review of research on effectiveness : CRD guidelines for those carrying out or commissioning reviews. CRD report; no. 4. The Centre, York, pp. 2-3

Stern J, Simes RJ (1997) "Publication bias: evidence of delayed publication in a cohort study of clinical research projects" *British Medical Journal*, No 315, pp. 640-5

ENDNOTES

[1] <http://chid.nih.gov/subfile/subfile.htm>

[2] <http://www.omni.ac.uk>

TABLE 1. Fugitive Search – Lipid Lowering Drugs

Commercial databases	Web library catalogs (using Library of Congress or National Library of Medicine (MeSH) subject headings)	Internet peer-reviewed sites	Internet search engines
1. Cochrane Library 2. HSTAT (technology assessment gateway) 3. HSSEarl (MLM) 4. Dissertation Abstracts 5. Articles (OCLC) 6. Papers (OCLC) – conferences and paper abstracts 7. TRIP database (evidence-based medicine) 8. NTIS Database 9. CRISP (Computer Retrieval of Information on Scientific Projects) 10. HTA Database 11. CCACS	1. UBC Library Catalog 2. BC Ministry of Health Library Catalog 3. Canadian Institute of Scientific and Technical Information (CISTI) catalog 4. British Database (Buckinghamshire Health Authority Library) 5. HealthPromis (UK web catalog of health promotion) 6. National Health Information Center – Health Information Resource Database 7. CHIC Online (Consumer Health Information Database) 8. WorldCat 9. GAO Web Catalog 10. COPAC (Union catalog in the UK) 11. NLM Locator Plus	1. UK Academic Web Directory 2. UK Social Science Information Gateway 3. OMNI (Organizing Medical Networked Information) 4. Medical Matrix 5. Health Communications Network 6. Global Health 7. Health Index 8. Medweb Public Health 9. Medscape	1. Northern Lights 2. Alta Vista 3. Hotbot
In-house databases	**Directories**	**Organizations contacted**	**Journals hand-searched**
1. In-House Catalog 2. Analytic Database	1. SCRI: HealthCare Standards 2. UBC Technology Assessment Manual	1. US National Institutes of Health 2. University of Ottawa Heart Institute 3. Montreal Heart Institute 4. National Heart Lung and Blood Institute 5. Health Heart Program St. Paul's Hospital, University of British Columbia 6. World Health Organization 7. International Task Force on Coronary Heart Disease 8. International Society of Atherosclerosis 9. Institute of Clinical Evaluative Sciences 10. American Heart Association 11. National Cholesterol Education Program 12. College of Pharmacy and Faculty of Medicine, Dalhousie University 13. American College of Cardiology 14. Ontario Ministry of Health 15. Conseil d'Evaluation des Technologies de Santé 16. Australian Institute of Health and Welfare 17. University Hospital Consortium 18. Health Services Utilization Research Commission 19. Texas Institute for Health Services Research 20. Canadian Coordinating Office of Health Technology Assessment 21. International Society for Pharmacoeconomics and Outcomes Research 22. Therapeutics Initiative, UBC 23. National Pharmacy Continuation Council	1. Canadian Journal of Cardiology 2. American Journal of Cardiology 3. Circulation 4. Circulation 5. Evidence-Based Medicine 6. Cardiovascular Review Reports 7. JAMA 8. ACP Journal Club

TABLE 2. Fugitive Search – Acupuncture in Addiction Treatment

Commercial databases	Web library catalogs (using Library of Congress or National Library of Medicine [MeSH] subject headings)	Internet peer-reviewed sites	Internet search engines
1. Cochrane Library 2. HSTAT (technology assessment guidelines) 3. HSRProj (NLM) 4. Dissertation Abstracts 5. Articles (OCLC) 6. Periodicals (OCLC) – conferences and proceedings 7. TRIP database (evidence-based medicine) 8. Ovid Academic Search 9. Ovid Complete MAS 10. Alcohol and Alcohol Problems Science Database (ETOH) 11. ACUTASE (Pubmed) 12. CRISP (Computer Retrieval of Information on Scientific Projects)	1. UBC Library Catalog 2. BC Ministry of Health Library Catalog 3. Canadian Institute of Scientific and Technical Information (CISTI) catalog 4. BCHealth Database (Backing Canada Health Authority Library) 5. HealthPlanet (UK web catalog of health promotion) 6. National Health Information Center – Health Information Resources Database 7. CHID Online (Combined Health Information Database) 8. WorldCat 9. OAO Web Catalog 10. COPAC 11. NLM Locator Plus	1. UK Academic Web Directory 2. UK Social Science Information Gateway 3. OMNI (Organizing Medical Networked Information) 4. Medical Matrix 5. Health Communications Network 6. Global Health 7. HealGuide 8. Medweb Public Health 9. Metanoia	1. Northern Light 2. Altavista 3. Excite
In-house databases	**Directories**	**Organizations contacted**	**Journals hand-searched**
1. In-House Catalog 2. Analytics Database	1. TCCL Health Care Standards 2. CHC Technology Assessment Member	1. US National Institutes of Health, Office of Alternative Medicine (Also known as the National Center for Complementary Medicine) 2. Center for Complementary and Alternative Medicine 3. Center for Addiction and Alternative Medicine Research 4. The University of Texas Center for Alternative Medicine Research 5. Tai Chi Institute for Complementary and Alternative medicine (TIC) 6. National Acupuncture Detoxification Association 7. Richard and Hinda Rosenthal Center for Complementary & Alternative Medicine – Rosenthal Center Directory of Databases 8. Research Council for Complementary Medicine 9. Centre for Addiction and Mental Health 10. The National Institute on Alcohol Abuse and Alcoholism 11. Centre for Addiction Studies (UK) 12. NIDA Organizations 13. Canadian Centre on Substance Abuse 14. Foundation for Traditional Chinese Medicine (UK) 15. Hooper Detox Center 16. National Academy of Acupuncture 17. BC Women's Hospital 18. Vancouver Richmond Health Board (Detox Committee) 19. The Acupuncture Association of BC 20. Evergreen Treatment Services 21. Ministry for Children and Family (BC) 22. Okanagan County Medical Center 23. Metro West Center for Medical Research 24. Uranda Medical and Mental Health Centre 25. Yorus University of Traditional Chinese Medicine	1. Alternative Medicine Journal (peer-reviewed) 2. Alternative Therapies in Health and Medicine (peer-reviewed) 3. Journal of Substance Abuse Treatment 4. JAMA 5. American Journal of Acupuncture 6. Journal of Alternative and Complementary Medicine 7. Lancet 8. BMJ

Directory Database of Research and Development Activities (ReaD)

Hiroyuki Sato,
Kenshou Nakagawa, and
Yoshiyuki Maeda

JST, Japan Science and Technology Corporation

INTRODUCTION

Japan Science and Technology Corporation (JST) has developed a Directory Database of Research and Development Activities (ReaD) since fiscal 1997, and offered test service of ReaD since August 1st in 1998.

Now, the Japanese government aims to build up a nation creating science and technologies and invests actively in various basic researches on the basis of a science and technology master plan. These researches are aimed at creating new industries by the private business sector using the research results. In order to use the results, information such as the content of research has to be offered and opened to the public. JST conducted ReaD to spread the national research information by the Internet.

Generally, to grasp the research contents, we ask the responsible institute or researcher to confirm the research theme as a basic information. For that purpose, JST constructed a directory of research institutes, researchers, research themes and research resources as the nucleus of contents of ReaD for users as fundamental information (guidance).

When we developed the system, we formulated there principles of design namely, (i) collecting the open technologies as much as possible (standardization of design), (ii) open-ended design of the system and (iii) user friendliness. These principles are inherent in all ReaD functions.

In this text, the concept of development, outline of the system, functions and directory of ReaD are described.

DEVELOPMENT CONCEPTION OF READ

1. Abstract of ReaD

An overview of ReaD is shown in Fig. 1. The nucleus of ReaD consists of three research functions (Category search, keyword search and Multi-link function) and four kinds of directory database (Institute, Subject, Researcher and Resource DDB) with cross-references to other JST contents. ReaD is under way to make possible cross-reference with other related contents. The search function and four kinds of directory were developed in fiscal 1997. Other contents and the total system were upgraded in fiscal 1998. The Internet was selected as the insemination under consideration of foreign countries.

When an information disclosure system is designed, a computer system is important as a receptacle, but collection of information, that is information resources, is more important. ReaD has two collection methods: one is a questionnaire by JST and the other is information entered by research institutes and researchers.

Moreover, JST will construct a system which offers all the internal research information and ReaD play a window of national information in cooperation with the National Center for Science Information System, etc..

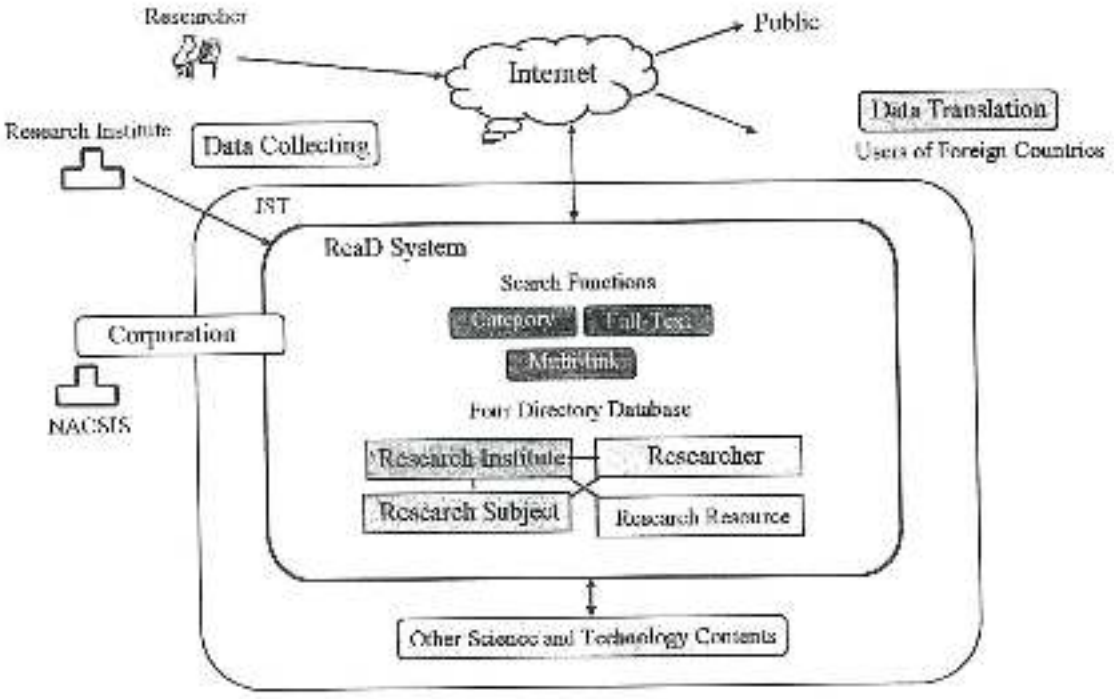


Figure 1 Concept of ReaD

2. Retrieval function

The most useful retrieval function is keyword (full-text) retrieval in the Internet. Although ReaD has keyword search, category search function is the main retrieval step, because user may retrieve names of research institutes and researchers and can search easily under a certain classification (category search) fixing the extent of the search breadth. Moreover, a part of them are added a squeezing function so as to offset a weak point of category search.

Keyword search displays various color of the appropriate character strings and changes the content of information at every directory and clarifies the retrieval results for user.

A multi-link function was developed as means of acquisition of all related information, because users might desire to obtain all information or research institutes and researchers without many retrievals.

3. Four kinds of directories

In the current printed medium, basic information are arranged under every research institute. The contents consist of research content (subject) of institute, researchers and offering service. JST distributes these data under various categories for use as one body by means of the Multi-link function. JST investigated the existing directory contents from the viewpoint of the kinds of information desired and which could be collected. Moreover, on the basis of know-how of JST and researches of the related organization, JST decided on the present four kinds of directories such as Institute, Subject, Researcher and Resource DDB.

4. Collection of information

Information is collected by a questionnaire on search institutes once a year at present. New entries and maintenance are carried out. Information on national organizations, such as National research institutes, collected by the Science and Technology Agency are used. Data on semigovernmental organizations and the public sector are collected by JST's questionnaire. These collection of data on data base is offered at any time by JST.

Information collection by the questionnaire is scheduled once a year, but updated every weak. Accordingly, a representation of data correction and addition is accepted and updated at any time.

Information on universities is not directly gathered, but it will offer in cooperation with the National Center for Science Information Systems of Ministry of Education.

Although information is currently collected by questionnaire, JST has a plan to construct a system updated directly by research institutes and researchers by means of the following function.

5. Research entry site of information

Entry of information by researchers is very important in the conception of ReaD.

Although the entry format of Web is not special, a good format for entry of science and technology information and updating of it by researchers using this format are very important.

Although the Internet, a mass media next to newspaper and television, has home pages, data entry by researcher is not sufficient and the data is not used enough. Under these conditions, ReaD can offer the chance of information input by researchers. We are modifying the present update system of ReaD to make better system.

ABSTRACT OF SYSTEM

1. History of development

We designed data items at first to select data items of each directory and then clarified the mutual relation between them. The subject directory contains two categories: one is ordinary research under the theme of one institute and the other is a project pattern by many institutes and universities (one search has a hierarchical structure consisting of a large theme, medium and small themes). This pattern is increasing.

An original database were created by transferring the existing data to the new format and by questionnaire using the new format.

The system – data addition or updating, verification of contents and updating – completes one cycle. Data addition or updating was kept on the basis of annal questionnaire by JST and correction of data by researchers and JST. The contents are verified by a quality control officer at JST for integrity assurance. Data is updated once 1 week (Fig. 2).

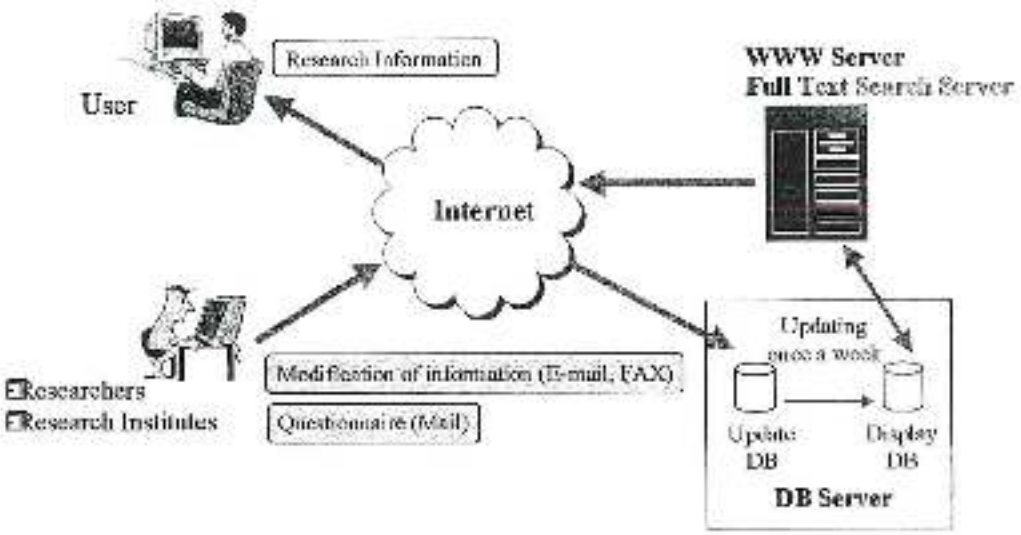


Figure 2 Image for ReaD system Operation

2. Hardware

The basic structure consists of two service systems, Service System 1 and 2, each made up of two servers such as WWW and full-text search processing server and a database server (Fig. 3). This structure can continue uninterrupted service during updating and in case of failures in either system. Moreover, an access load can be distributed to two systems for increasing further user convenience

All equipment are installed within the fire wall of JST network.

A database server in Service System 1 is WWW server for update processing via Web. Further, WWW of Service System 1 performs an electronic billboard function, statistics processing function, mailing list server function and output function of questionnaire results.

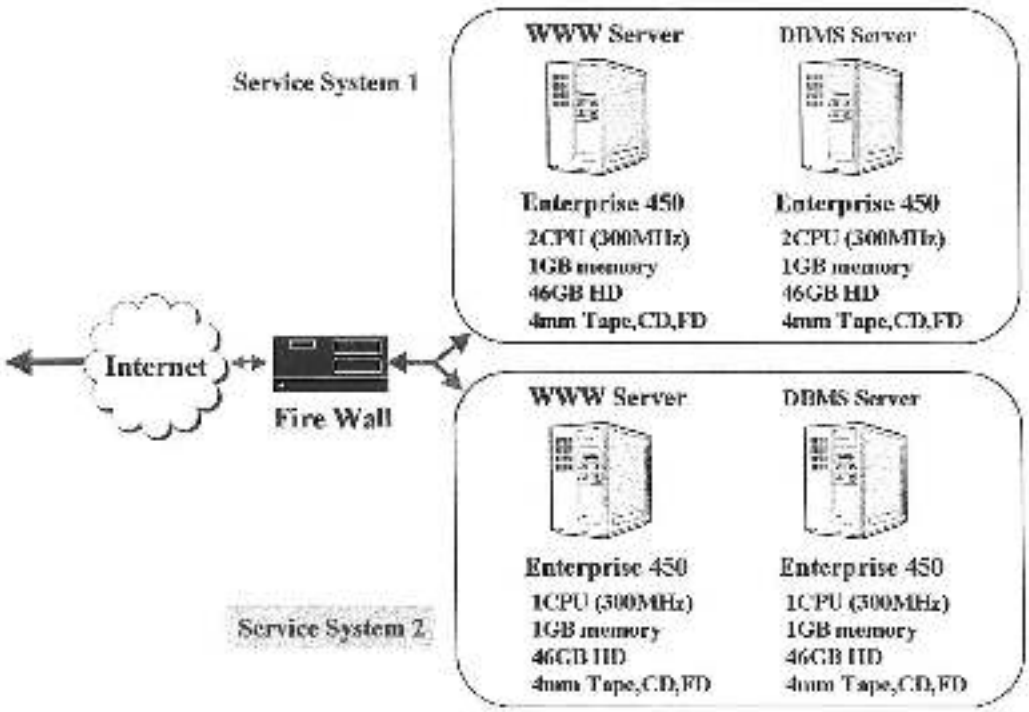


Figure 3 ReaD Hardware Structure

3. Software

The structure of software is shown in Fig. 4. Oracle Web Application Server 3.0 (WAS) is used as WWW Server. Oracle8 Enterprise Edition, Open Text and Livelink Spider are used as software of DataBase, Full-text Retrieval Engine and Robot Software, respectively.

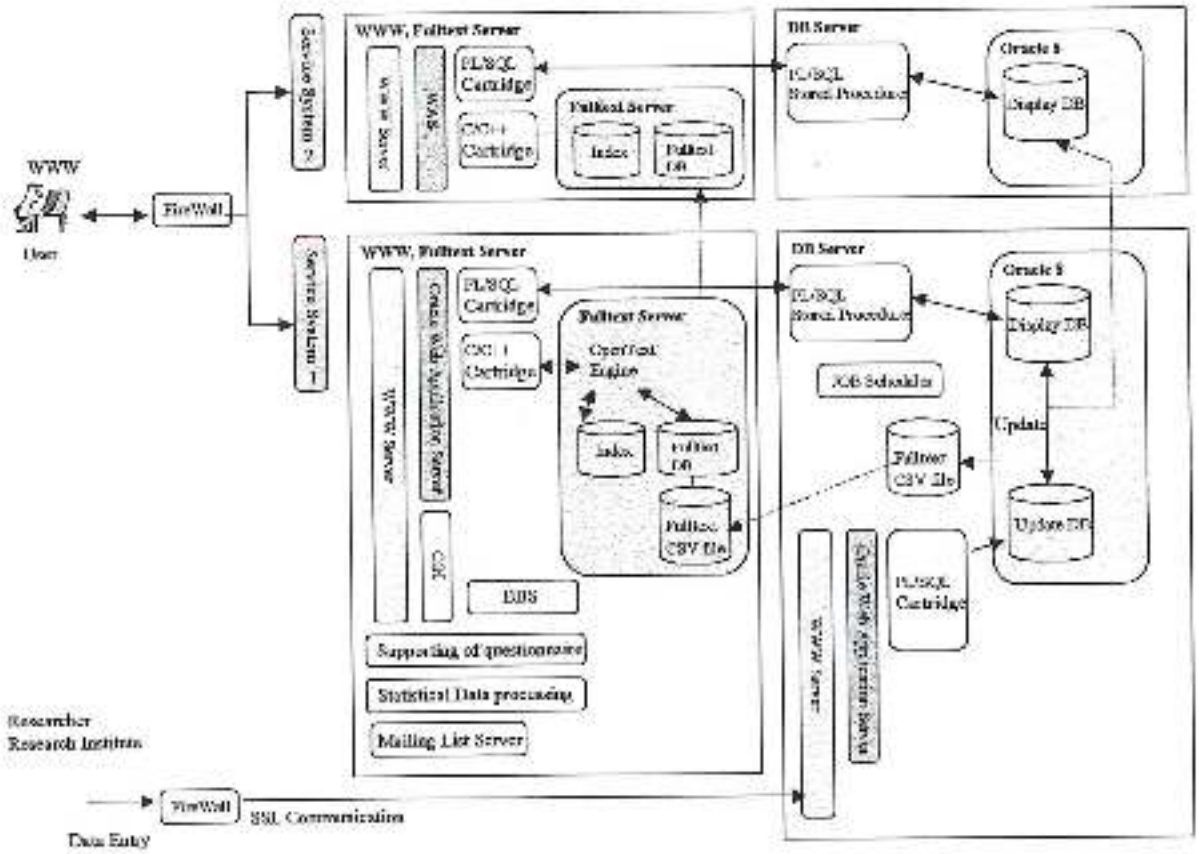


Figure 4 Software structure of ReaD

WAS consists of three parts: first is Web Listener as HTTP Server, second is Web Request Broken (WRB) equipment for HTTP Server API (Application Programming Interface) and third an external program named by cartridge (Fig. 5). A retrieval request entered by user's browser is accepted at once by Web Listener, then transferred to WRB for definite response to the cartridge and then the request can be processed by the selected cartridge. The retrieval results obtained by the above processes are returned to the user via the reverse route.

Cartridge automatically produces the number of process correspondences with the user's request while controlling load. The processing between cartridges is independent. Accordingly, error by one cartridge function does not affect other functions. Development of new Web applications means addition of new Cartridges. Use of cookies makes possible transaction management.

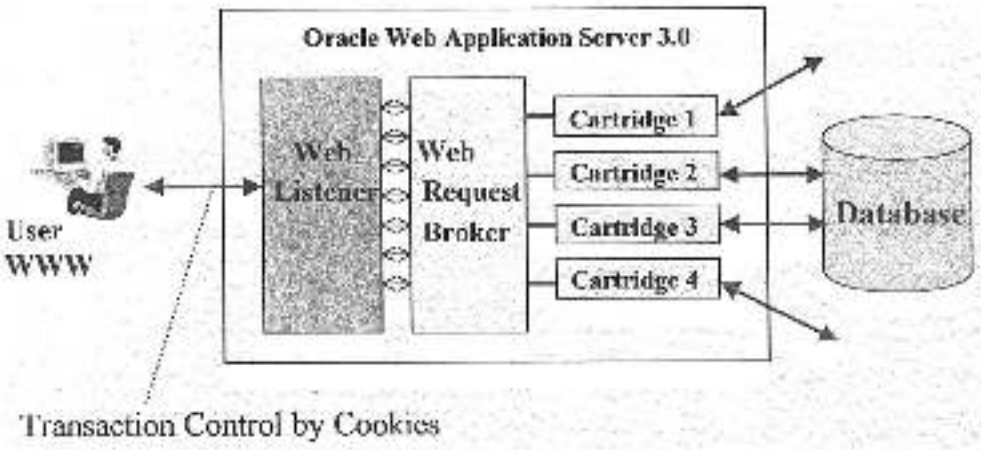


Figure 5 Link between WWW and DB

Next, a retrieval processing of ReaD is explained (Fig. 4). On the Category retrieval, a processing request from user's Web browser is accepted by WAS. When user requests "Announcement" or "What's New", HTML file is displayed directly to user. If user requires to access the directory database, the retrieval request is sent to database server via WAS, then the obtained retrieval results are displayed by HTML file after automatic reconfiguration. In the case of full-text search, the retrieval words and conditions entering into WWW Server are sent to full-text search engine via WAS, and then obtained results are displayed by HTML file. The retrieval requests of users are accepted by alternating acknowledge between Service 1 and Service 2 by means of the load distribution function of Fire Wall.

Updating processing can be done by WWW server on the Database Server of Service 1 as a special-purpose entry via Web. Reliability between user's browser and Server is increased by encryption using SSL (Secure Socket Layer) communication.

New data entered and update are placed on the update system database at once, and sent to update processing after verification of the contents by JST at week end. Subsequently, they are recorded on the database. Update processing of full-text system is carried out in the following order, the DB Server of Service 1 is updated and then CSV file (tag delimiter text file) for creation of index file of full-text search is swept. Then, DB System of Service 2 and full-text System of Service 2 are updated. The course of general update operation is terminated about half a day from the past records.

ABSTRACT OF FUNCTIONS

1. Retrieval function

There are two search methods of the directory type contents, one is a category search which the user follows the items of which are arranged by specific sorting and the other is a full-text search that the user can retrieve by use of any character string. ReaD offers these two search methods.

As category search, the user can obtain the objective information by sort items in each directory data (Fig. 6). ReaD offers "Squeezing search" function by which the user can squeeze the candidate on the way to retrieval (squeezing by the search system, research budget and research field in the research subject directory and squeezing by the research field in the researcher directory)

The buttons, showing link to DDB data and the related information, are set at the lowest part of the last display screen of each directory (Fig. 7). One of the important characteristics of ReaD is the relational links between 4DDBs named by a multi-link function. For example, user retrieves a research institute and then immediately can display its researcher names, themes and resources by means of it.

On the other hand, the full-text search function can be used by choosing DDB and input the retrieval words. Detailed retrieval using many retrieval words and the retrieval conditions such as AND and OR can be executed as well (Fig. 8).

With creating the index file of total Web server data in JST domain by using a robot software, the user can perform full-text search on them.

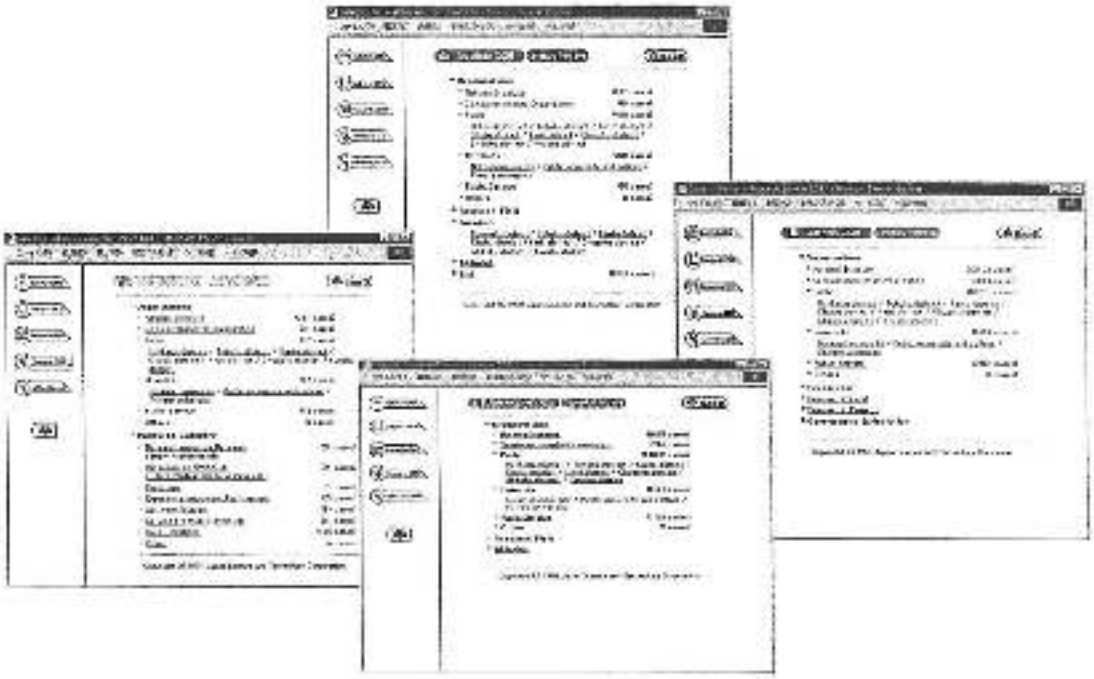


Figure 6 ReaD Category Search

2. Input data function

Input data to ReaD system is provided anywhere user can access the Internet and use a WWW browser. To use this service, users are requested to register user's ID and password issued by JST. Input function fixed the limit of data addition editing by creating three kinds of user categories, specifically, researcher, institute manager and JST manager. (Researcher can add and edit the researcher and research subject information, the manager of institute can do the above information and the research institute and research resource information and JST manager can carry out addition and editing of all data).

ReaD has also an output function of questionnaire results and a statistics processing function for managers. Both functions support the questionnaire work and analysis and control of information.

CONTENTS OF EACH DIRECTORY

The date is set automatically in all directories to express the date of input of the directory contents by the user. Moreover, it requires each research institute, researcher and JST data manager to pay attention to update.

1. Research institute directory

Table I shows the data items of research institute directory. The items explain about the institute such as its outline, history and purpose of activities. An address guide shows a map of the institute and the block plan of buildings. The items of research field, recent research results and department are used as an abstract of institute activities when user applies or utilizes it. The links are established between each directory page and home page of research institutes. It gives detailed information and can be used as link group.

Table I
Data Items of Research Institute DDB

1	Data Revision Date
2	Organizations
3	Address
4	Foundation Date
5	Telephone
6	FAX
7	History
8	Purpose of Activities
9	Main Activities
10	Representatives
11	Number of Staff
12	Number of Departments
13	Number of Attached Facilities
14	Area of Site
15	Area of buildings
16	Budget
17	Periodic publications
18	Attached facilities
19	Research Field
20	Departments
21	Recent Research Results
22	Number of Patents
23	Guidance of Address
24	Traffic

2. Research subject directory

Data items of the research subject directory are shown in Table II. The items consists of data on the name of researcher, accomplishment, completion/in progress and research field. The subject directory characterized to cope with a large research theme (main theme) with some smaller themes, for example, Exploratory Research for Advanced Technology (ERATO) of JST. In this case, the item of project, representative organization of project and representatives are used. It makes possible to retrieve research of a specific research system. Information about research results and evaluation of research supported by national fund are entered into it after due consideration that these research results are asked to be opened to the public. The research evaluation has no data open to the public, but we will develop a directory that you can look through the subject from beginning to the accomplishment, prize, etc., by entering data from the evaluation reports.

Table II
Data Items of Research Subject DDB

1	Data Revision Date
2	Research Subject
3	Research Organization
4	Departments
5	Researcher
6	Project
7	Representative Organization of Project
8	Representative
9	E-mail
10	Research Budget
11	Research Duration
12	Finish / Undergoing
13	Outline of Research
14	Research Field
15	Keywords for Research
16	Research Project
17	Government
18	Accomplishment
19	Prize etc.

3. Researcher directory

The data items of researcher directory are shown in Table III. There are many items about researcher personally, research contents and the past records which are not the required items but they are provided upon the researcher's discretion. In four directories, Researcher directory contains personal information, that is private information, so that only data input by researcher himself is offered and opened to the public including with data of the directory.

If the researcher desires to provide a lot of information, it is possible without any limitation on space. In this case, ReaD runs a set of operation that is repeated regularly in ten papers and books. When item of paper, book and patent stored many data, these items in Researcher directory become very useful information.

Home page URL of researcher, present research theme and free comments are offered for researcher to present himself. It is recommended to use them as much as possible.

Table III
Data Items of Researcher DDB

1	Data Revision Date
2	Name of Researcher
3	Organizations
4	Departments
5	Title
6	Birth Year
7	Graduate School
8	University
9	Degree
10	Mailing Address
11	Telephone
12	FAX
13	E-mail
14	Personal Homepage URL
15	Societies
16	Research Experiences
17	Career (Professional)
18	Research Field
19	Keywords for Research
20	Research Theme
21	Prize
22	Papers
23	Books
24	Patents
25	Availability (Books, Data, etc.)

4. Research resource directory

Finally, the items of research resource directory are shown in Table IV. Research resources in this article is a resource that is material and immaterial resources (hardware, software database, various kinds of systems and service, etc.) maintained by organizations and a outsider can use them by proceeding through the necessary formalities.

The items of the directory contain a list of organizations and contact persons who manage the resources, research subject category, resource information, condition for use and utilization.

Table IV

Data Items of Research Resource DDB

1	Data Revision Date
2	Research Resources
3	Mailing Adderss
3-1	Organizations
3-2	Departments
3-3	Contact Person
3-4	Address
3-5	Telephone
3-6	FAX
3-7	E-mail
4	Research Subject Category
5	Research Resource Information
6	Samples
7	Condition for Use
8	Utilization (Application)
9	Related Information
10	Formality for Use

CONCLUSION

At the end of introduction of ReaD, the objects of ReaD are explained.

(i) ReaD aims at one stop Web site of science and technology research information. JST is developing the necessary contents, for example, a symposium guidance directory and a science technology link group and user friendly retrieval functions.

(ii) ReaD performs an entry point to open research information of the nation, etc. Various kinds of research information for example, research results information without the present directories, are created and linked with ReaD and offered as an integrated database. ReaD will construct the system with research information of nation and universities in cooperation with the National Center for Science Information Systems.

(iii) ReaD aims at an entry site of research information by the research institutes and researcher. As indicated above, the entry site of information is very important for security of information entered by researcher himself and for collecting science and technology information in future. JST is going to develop the necessary functions which researcher himself can enter information and to spread use of the site.

Towards integration of information sources on grey literature. A case study

Rosa Di Cesare, Consiglio Nazionale delle Ricerche (CNR) - Biblioteca Centrale -, P.le Aldo Moro, 7 Rome - Italy; dicesare@bicc.cnr.rm.it
Gianni Lazzari, Biblioteca della Camera dei Deputati, Via del Seminario, 76 Rome - Italy; lazzari@camera.it

ABSTRACT

The institutional goals and the organisational contexts of GL producers, determine the level and the process of production and diffusion. Furthermore GL producers don't bring all the same "dowry": they have got neither the same resources, nor the same information policy. Inspired by such considerations the Italian reference centre of SIGLE (System of Information on Grey Literature in Europe), has based its activity on the Cupertino with GL producers.
To achieve this aim the Centre has evolved from a merely receptive role, consisting of pure and simple acquisition and bibliographic treatment of documents, to an active role, in order to control the very process of document production by exerting a consulting function on bibliographic aspects. In other words is the Centre that places itself at disposal of producers to solve all the problems connected with their participation in SIGLE database.
As an example of this policy the paper describes the interaction between the SIGLE Italian reference centre and the Library of Italian Chamber of Deputies which produces the database LEGO, which stands for Grey Literature Online, containing non commercial, official and administrative documents, produced and/or acquired by the Library since 1993. The consulting activity developed by the centre has allowed a fast and efficient acquisition of the document stored in LEGO.
Finally the paper describes the type of documents in LEGO, the selection criteria and the issues connected with information integration.

1. Introduction

The development of information technologies in the information and documentation field and the concomitant effects on working processes and products (electronic catalogues, interconnections, information sources integration) allow, albeit potentially, the effective distribution of the information sources and therefore, pave the way for change from the conception of documentary resources mainly based on "possession" to a conception based on "access" (8).
But if the improvement of the information and documentation technologies has implemented the capacity of producing and transferring information, on the other hand it has only partially solved the problems concerned with the management of data coming from different archives (or stored in different ways).

The problem is that the distinct information sources differ (inside) not only in data structure and in the treatment of information but also in organisational and technical models (2).
Integrating information is therefore a complex activity as it involves some basic aspects of work organisation. Different procedures and different "ways of working" and of treating information have to be integrated with technologies. An information system should solve these differences by identifying, both the basic requirements for integration (for example document types, management and bibliographic treatment), and the means of harmonisation (use of different systems for cataloguing and classification).
Information source integration is not new to the world of information and documentation. A first important way of integrating information sources has been in fact the realisation of collective paper catalogues, the evolution of which is testified by OPAC (6, 13).
An automatic transposition of the above experiences on the GL field is quite difficult, given the well known characteristics of GL itself: above all diversification of producers, fragmentation of information sources and the lack of bibliographical control. Moreover, even if the producers make a bibliographical control of these documents, they may do so in very different ways.
Inspired by such considerations, the SIGLE Italian reference centre, has based its activity on co-operation with GL producers, and with one in particular, namely, the Library of the Chamber of Deputies, which produces the LEGO (On line Grey Literature) database where official and administrative documents have been collected, produced and/or acquired by the Library since 1993.
In this work the introduction of the experience of the co-operation above is preceded by some remarks on GL collection and management. In section 2. some features relating to particular producers and those related to the activity of the SIGLE Italian reference centre are considered. In the last section some considerations of the experience are given.

2. Some remarks on GL collection and management in Italy

Generally speaking, in Italy the fragmentation of Libraries and Documentation Centre, apart from scientific and technical institutions (up to cultural foundations) has partially delayed the realisation of the co-operative information system. As regards GL the SBN (Italian National Library System) has considered other national and European GL databases, by introducing some bibliographic details peculiar to SIGLE standards. However, the input of GL documents to the SBN database is, up to now, extremely limited. It is probably due to the lack of specialist support, especially during the phase of selection and collection of documents.
Further isolated initiatives that are to be found are compartmentalised in nature and deal exclusively with specific document types such Doctorate theses which have only been dealt with in a specific publication in the Italian National Bibliography since 1995, or with specific disciplines. The key public research bodies handle scientific and technological GL while legal and administrative GL is handled by the Library of Chamber of Deputies (1).

2.1 GL producers: some characteristics of GL management

Turning to individual GL producers, the situation is varied. In 1996, the first survey carried out by the SIGLE Italian reference centre on individual producers' organisations revealed a

situation that was far from uniform. Italian GL producers adopt different approaches to the handling of documents and to the choice of appropriate technology.

The survey shows that acquisition of GL documents takes place via

- *a direct channel* between the "historical" producing bodies and the Italian reference centre (50% of the cases);
- *a direct channel* of legal deposit between the producing bodies and the Central Library of the National Research Council (CNR) to which the SIGLE Italian reference centre has selected access (25%);
- *a direct channel* between the "recent" producer and Italian reference centre, set up for the bottom up release of data for users of the Central Library and for librarians at annual training courses run by the Library itself (15%);
- *"informal" channels* relying on "informal" contacts mainly between academics, library users, librarians and CNR documentalists (3).

Further information on GL producers came to light on analysis of factors such as document layout and the organisation of bibliographic elements of cover page.

This information breaks down to cover three categories:

- Organisations that have a highly organised information network along with in house editing facilities and/or library, both internal and external GL documents are collected and subjected to bibliographical control;

- In the middle ground there are organisations that conserve in house GL making it the responsibility of a librarian or head of department. The latter nearly always applies personal criteria regarding the conservation and handling of documents, with scarce diffusion resulting. The SIGLE Italian reference centre has offered these organisations support facilities for bibliographic control;

- At the other extreme there are organisations that merely archive their GL in "bureaucratic" fashion, failing to carry out any bibliographic classification of the material. For this group the Italian reference centre prepares training schemes and provides information.

In the comparison between GL producing offices and organisations one common factor emerges. Generally speaking GL is not subject to efficient bibliographical control nor is utilisation monitored. It should be pointed out that some organisations possess advanced documentation systems in qualified hands, with access to quality data treatment systems in an environment with little interest in specific GL projects, especially when these require standardisation, extra personnel and funding. Having said this, there has been some recent interest in integrated acquisition of GL data.

2.2 The SIGLE Italian reference centre

Reviewing its activities, the Centre has had to embrace the wide variety of individual situations and general characteristics pertaining to GL information sources and management that has been outlined above. Key considerations taken into account are: 1) information sources and GL producers differ not only from an editorial and technological point of view but also in the bibliographic management of their GL documents; 2) the institutional goals and the organisational context of GL producers determine the level and the process of production and diffusion. Moreover, GL producers don't all bring the same "dowry", having neither the same resources nor identical information policy.
In the light of the above and in conjunction with its institutional functions, the centre has sought to concentrate on research and information sources as well as GL producers, trying to assure their participation in the SIGLE database. This work has itself given rise to new initiatives while at the same time consolidating and rethinking the traditional parameters of databases.
The centre's chief commitment has been to research into new producing organisations and to provide support for those already in existence, participating in the database, by promoting enhanced bibliographic management of GL documents in line with commonly agreed standards.
Pursuance of these objectives has been hampered by two elements: one to do with the inherent properties of GL, and the need to know the specific grey routes taken in the production and diffusion of GL documents, frequently limited to a specific production environment (a commission, a project etc.); the other due to organisational and functional shortcomings in libraries and heeldragging on the behalf of GL producers (both individual and organisations) who fail to appreciate the importance of documents produced in working groups and research projects and effectively block acquisition and circulation in the scientific community.
Given the situation, the Centre has emerged both in practice and in institutional terms as the national reference point for GL. In order to maximise producer involvement the Centre has evolved from a merely receptive role, consisting of pure and simple acquisition and bibliographic treatment of documents, to an active role, in order to analyse the very process of document production by exerting a consulting function when necessary and/or requested, on bibliographic aspects with a view to integration of the different information sources. The assistance and consultation services provided by the centre deal with both editorial presentation of documents (whether or not to bring out a new series, project differentiation according to content etc.) and with bibliographic description of document (attribution of catalogue numbers, series titles, standardisation of the producing organisation etc.).
In other words it is the centre that places itself at the disposal of producers in order to solve problems arising from their participation in the SIGLE database.
The above mentioned evolution in the centre's activity has been made possible by the strategies adopted by EAGLE (European Association of Grey Literature Exploitation) which produces and runs the SIGLE database. These strategies have permitted a more flexible handling of SIGLE standard both in terms of bibliographic format and bibliographic description of documents. (Implementation of new software that is directly operated by the various national reference centre of the countries adhering to the SIGLE system, simplification of the bibliographic level of description (monographic or serial) and

the capacity to acquire documents with the title in the original language so long as key terms are given in English.)

3. Toward integration between the SIGLE Italian reference centre and the Library of the Chamber of Deputies

The SIGLE Italian reference centre has opted for a form of Cupertino and integration of GL information resources with one selected producer, namely the Library of Chamber of Deputies. Four factors prompted this choice: 1) the library itself having set up a specific GL database; 2) its national role and institutional standing in the field; 3) the availability of GL documents produced by key organisations and their inclusion in the LEGO database; 4) the library having personnel skilled in the handling of such documents.

3.1 The LEGO (On Line Grey Literature) database

The LEGO database came into being in 1993 and currently holds some 10,000 bibliographic references with an annual growth of 1,700 -1,800 new inputs. It is a parliamentary grey literature database for material emanating from the Chamber of Deputies and the Senate or coming into their possession from governmental bodies, professional associations, universities, Regions or local authorities, research organisations etc. It provides day to day back up to parliament. The most striking characteristics of this unusual database is its dual function. The documentation that is acquired is produced both by parliament itself, and on the whole, does not appear in Parliamentary bills, as well as by data reaching parliament in the carrying out of its legislative and monitoring activity, above all in preliminary hearings of permanent commission.
Clearly such data is of great institutional and political importance. The availability to the public of this data is an Italian parliamentary commitment towards the divulgation of the legislative procedure, its inspection and monitoring activities, via open access to the texts at the basis of political decisions.
LEGO provides access to hidden channels of information that are often unknown and which vary in type and source. These are pooled in a single reference structure with a single reference system. This is to the advantage not only of citizens but also of the administration itself. The more sources are available, the greater the benefits.
Clearly the database seeks to meet the information needs of its primary users in the library and the Chamber. To be precise these users are Deputies, political groups, Parliamentary Commissions, service providers and offices. A co-ordinated and indexed collection of documents that are essential for the carrying out of specific functions is made available.
Documents are grouped in two categories: 1) internal; 2) external, and this reflects both means of acquisition and contents. Both categories are further subdivided according to the needs of the complex institutional activities of the Chamber. There is almost total bibliographic control of the first category and this permits a range of document products, normally organised in series. Nomenclature reflects the mainly grey channels followed by this literature that is produced in support of legislative activity at all levels. (Documentation and research, Decree dossiers, Actions dossiers).
The second category, those documents that are sent to the Chamber is extremely disparate reflecting the range of producers. The control carried out on these documents by the library is to date only partial and is beset with difficulties.

It should be noted that "accompanying slips" are attached to parliamentary documents enabling the successive steps of the legislation to be followed (Presentation, Discussion, Approval) and the elaboration of interesting statistics on parliamentary proceedings (12).

1.1. Cataloguing and structure of LEGO documents

Documents are recorded in the Library's DOBIS/LIBIS handling system in accordance with ISBD standards. The system has undergone various simplifications when it was felt that rigidity did not permit exhaustive bibliographic description of GL documents. Indexing is carried out by the TESEO Thesaurus, which is an instrument set up by the Senate for parliamentary documents based on Universal Decimal Classification (UDC).
Elements in the DOBIS index are inserted into the LEGO database. Some specific fields of selection and research in the bibliographical headings are of particular interest. Consider the *Tipdoc* field (document typology) and the *Freq* (frequency) field, which are used respectively to select documents according to the means of acquisition and the frequency of despatch of periodical reports. In the main body of bibliographic description we find the *Materia* (subject matter) field containing the TESEO thesaurus terms, the field *Fonte normativa* (source of legislation), where the standards of the legislation or the essential elements of Bills are recorded. The field *Iter* (course taken) records procedural information on documents named in the Chamber and assigned to a Commission.
The co-operation with the SIGLE Italian reference centre has benefited from the streamlined input of LEGO data into SIGLE thanks to the utilisation of TESEO. This is done by means of UDC subject categories in English, which has greatly facilitated the indexing of LEGO documents with SIGLE subject codes (10). Currently, SIGLE subject codes are added to TESEO markers on implementation of the LEGO database. This permits greater harmonisation between the documents that are inserted and the SIGLE database. The net result polarises into LEGO autonomy, and at the same time, LEGO compatibility with SIGLE. (The latter benefits by gaining enhanced value of its documentation with a minimum of extra document description work.)

3.3 LEGO documents in SIGLE

The following is an outline of the phases in the Cupertino. During its work use was made of an "agreement" defining techniques for the eventual overall integration of information sources (4).

- *Selection.* During this first phase LEGO selected only a part of the *Internal documents* that had already been subject to bibliographic control by the Library of the Chamber of Deputies. It was decided to acquire only those documents whose contents were of interest to a wider readership (Actions dossiers, Documentation and research etc..). The external documents selected were largely Government reports or those from the State Audit Court.

- *Bibliographic description.* Integration was made easier by the producer having adopted a flexible application of the ISBD standards. A case in point is the classification of LEGO registered documents according to document typology.

- *Format.* The SIGLE Italian reference centre provided a simplified conversion format for the FIBRE programme (5.) The format is based on a registration card organised according to document typology. Data travels from the Library of Chamber of Deputies to the centre by e-mail in standards SIGLE format.

- *Indexing.* During this first phase the Library of Chamber of Deputies uses SIGLE subject codes to index the documents. More detail on content is given by assigning UDC markers in the keywords field. An automatic conversion from UDC to SIGLE subject codes is being studied.

- *Titles.* The recent SIGLE decision to accept documents with titles in original language, as long as this is supported with key terms in English has enabled acquisition of LEGO documents.

- *Availability.* The agreement attributes an equal position to the two signatories that share responsibility for document delivery. This means that documents can be requested from either structure.

Up to now he has concentrated on improvements to procedure. We stress the benefits of having eliminated the transmission of documents by the producer to the centre. This has saved costs during registration and physical sorting of documents for all concerned, particularly for the centre. Producer harmonisation with SIGLE standards has freed resources and slashed SIGLE document acquisition costs.

4. Closing notes

Starting just after World War II, the CNR Central Library has collected GL from important national and international scientific, public and private institutions. This activity was extended and developed with adherence to EAGLE. Furthermore, under terms of legal deposit, the Library itself receives Italian technical/scientific publications. This legal right has habitually included GL documents.
The Centre's activity can be divided into two complementary phases. The first broadly coincides with the setting up of the SIGLE Italian reference centre at the Central Library of the CNR in 1985 (membership of EAGLE goes back to the same year), and continues until 1995.
During this first phase, the Centre's activity has concentrated on elaboration and management of Italian data (acquired from the Library as indicate in 2.1. paragraph). This data is periodically sent to the SIGLE database. Analysis of GL producer profiles was not carried out at this stage.
In the second phase the SIGLE Italian reference centre has implemented tailor made measures for individual producers in the drive for implementation of information resources. Producers' needs were analysed and care was taken to minimise interference with already elaborated technical-bibliographic procedure for GL documents.

This activity can be broken down as follows:

- *First level:* assistance to producers on grey information enhancement and the initiatives to be adopted. It should be borne in mind that rapid integration is out of question;

- *Second level:* aimed at producers who have already spontaneously adopted GL management measures. In these cases, the centre proposes an equal partnership aiming at the development of common strategy in GL management. This goes for theoretical considerations and/or the drawing up of guidelines. At this level, where the emphasis is no longer on urgency, more stable forms of collaboration can be established (common actions, "agreements"), that help towards integration of information sources.

Measures intended to involve producers must also:

Moderate the scale of participation: producer's involvement has to be calculated so as not to overload the SIGLE Italian centre with requests for assistance. Wider participation causes problems that are not only technical (different systems and instruments), but also problems of communication between the Centre and the various producers and between the producers themselves. In an integrated system, presumably producers must interact with one another and not only with the centre (e.g. the exchange of documents).

Monitor the producers and publicise the process of integration of resources: in the SIGLE Italian reference centre's current phase of activity, the monitoring is needed to have an idea of the progress being made (producer participation levels): data flow is still prevalently unilateral (from an individual producer to the centre).
As a result of this collaboration ground is being covered on the selection of individual producers to act as possible "hubs" in the widest possible network.
The aim is to publicise GL producers and their information resources through the Centre, and includes the benefits of elements of comparison arising from having to overcome common problems using analogous treatment procedure for grey information.
Monitoring can become a viewing glass for the harmonisation of procedure and for the transparency and identity of individual producers.
Further initiatives regarding integration could derive from "publicity" in appropriate channels (e.g. the Italian Libraries Association - Bulletin) on the resource integration initiatives that are underway.
It cannot be overlooked that in this potential increase in the multiplicity of communications the SIGLE Italian reference centre will have to find new partners and will need to reinforce its role of sorting and filtering grey information as well as that of putting producers, users and resources in touch. (When possible steering demand for grey information directly to where it is produced).

These new steps in the integration of information resources will involve the Centre in a *supervisory role* under contractual obligation with the different subjects involved, as in the case of the "agreement" with the Chamber of Deputies.
In order for integration policy to by successful, a communication culture is required. This should be understood in the etymological sense of the term i.e. pooling of objectives, results, criteria, procedure, and all else required for the development of information resources.

To sum up two issues arise. One is that the application of a standard model at best slows things, at worst making it impossible to proceed towards the objective. It is our experience that imposing a standard frightens off GL producers. The adoption of any standard requires technical, organisational and often cultural change. The other one is that the centralised organisation should be replaced since it is incompatible with the new scenario. The SIGLE Italian reference centre as a problem interpreter and shooter can, by means of integration, set up a beneficial circuit between producers, offering the end user an innovative "product" in terms of quality and quantity of GL information sources.

References

(1) Alberani,V.; De Castro Pietrangeli, P.; Minutoli, D. (1996) La letteratura grigia. 2. Convegno Nazionale, Rome, 20-21 Maggio 1996, Istituto Superiore di Sanita'.
(2) Colin J. (1998) Electronic technology and its impact on libraries, *Journal of librarianship and information science*, 30 (1) 7 -24.
(3(Di Cesare, R. (1996) Il Centro Nazionale di Riferimento SIGLE nella gestione della letteratura grigia prodotta in Italia: ruolo e prospettive. In: La letteratura grigia. 2. Convegno Nazionale, Rome, 20-21 Maggio 1996, Istituto Superiore di Sanita.
(4(Di Cesare, R.; Lazzari, T. (1999) La base di dati LEGO della Biblioteca della Camera dei Deputati in SIGLE, *AIB Notizie*, 11(5):1-13.
(5) FIBRE (Friendly Input of Bibliographic Records) produced by INIS/IAEA adapted to SIGLE standard
(6) Gatti, G. (1997) Macchine celibi? *Biblioteche Oggi*, 15(6) 6-21
Scritti in occasione del Centenario della FID (ed.) Augusta Maria Paci, Franco Angeli.
(7) Jamshid B. (1997) The evolving OPAC, *Cataloguing and Classification Quarterly*, 1997, 24 (1/2),163-85.
(8) Poppi, G. (1996) *Gestione di risorse integrate per il recupero dell'informazione e la fornitura del documento: esperienze italiane*, in La Documentazione in Italia. Scritti in occasione del Centenario della FID (ed.) Augusta Maria Paci, Franco Angeli.
(9) Rutstein, M.; Demiller, A.L.; Fuseler, E.A. (1993) Ownership versus access, *Advances in Librarianship*, v. 17.
(10) SIGLE Manual. Part 2: Subject category list, European Association for Grey Literature Exploitation, 1999.
(11) Timoty Sprehe, J. (1999) Government information: from inacessibility to your desktop and back again, *Journal of the American Society for Information Science*, 50(4), 340-345.
(12) Venturini, F. (1995) La letteratura grigia parlamentare. In margine a una nuova banca dati, *Bollettino AIB*, 35(3), 321-342.
(13)Weston, P.G. (1996) *Le tecnologie dell'informazione*, in La Documentazione in Italia. Scritti in occasione del Centenario della FID (ed.) Augusta Maria Paci, Franco Angeli.
(14) Zahiruddin, K. (1998) System migration: challenges for libraries in the Arabian Gulf region, *The Electronic Library* 16 (3), 171-174.

INTERNET' USE

Research of Grey Literature by teachers-researchers at the University of Constantine

SEMRA HALIMA
Institut de Bibliothéconomie
Université de Constantine

To have access to scientific and technical information, the university teachers-researchers always tend to use the documentary structures of the university of Constantine.

Evidently, they also go -or are orientated- to other information sources out of the university : public libraries, archive centers, documentation centers in companies..., according to their information needs.

The recent integration of Internet at the university of Constantine has incited us to lead a survery that will allow us to know whether this new mode of access to information will have a distinguishing place in the teachers-researchers' habits, and whether grey litterature is taken into account in electronic research.

1. INTERNET AT THE UNIVERSITY OF CONSTANTINE

The university of Constantine was born as a result of the reform of the Higher Education System which took place in 1971 and confered to this institution the statute of multi-disciplinary university in the whole eastern region of Algeria.

Its population does not stop increasing and has reached this year (1999) about 1500 university teachers and about 35000 students.

The documentary resources spread all over certain structures such as :

- The university library which represents the most important structure.
- The institutes'libraries (22).

There is a great interconnection between these units so that they may be able to answer their users'needs in terms of studies, teaching, activities and research.

The policy of establishing a national network of scientific and technical information has led the (CERIST)* to provide the structure of Higher Education and Research with the services offered by the « network of networks ».

The integration of Internet at the university of Constantine took place in november 7th 1998 within the university library. This service has got a specialist telephone line connected to the CERIST.

The exploitation of Internet by the teachers-researchers requires a pre-registration due to the limited available places. The exploitation is however free of charge and offers :

- The bibliographical research.
- The electronic mail.
- The file-transfer.

Now, 180 consultation sets have been provided at the disposal of users and spread over 5 university campuses.

* Centre de Recherche sur l'Information Scientifique et Technique (Alger)

2. THE SURVEY

a. Our choice has been focused on the university library, the first location where Internet has been integrated – November 7th, 1998 within a room (the Internet Room) that is equipped with 16 consultation sets.

b. How was the survey carried out?

A questionnaire was administered to the persons who were using Internet in two periods of time:

- At the moment of the integration of Internet, that is to say November-December 1998.
- Six months after its integration: April 1999.

With few questions that were simple and concise, the majority of the teachers were able to answer the questionnaire the day of its distribution. Few, however, handed it in later. 101 teachers answered the questionnaire.

c. The objectives of the survey

Our analysis intends to clarify certain questions to which we would like to find an answer:

- Who are the Internet users?
- What are their objectives in using Internet?
- What do they think of Internet? (advantages/disadvantages)
- How do they conceive the future of this technological tool?
- Does grey literature occupy a place in their electronic research?

Our hypothesis is based on the assumption that the use of Internet varies according to the subject.

3. THE RESULTS OF THE SURVEY

3.1. Characteristics of the population studied

Table 1: position of the population studied

Position	Answers
Professor	7
"Maître de Conférence"	12
"Chargé de Cours"	21
"Maître Assistant"	19
"Assistant"	38
Engineer	4
Total	101

The table reveals that the highest number of answers comes from the "Assistants". Nevertheless, when we observe the number of answers in terms of the global population – by position – we notice that the most important proportion of answers derives from both the extreme positions: Professor/Assistant.

This is illustrated by the following table:

Table 2: percentages of answers by position

Position	% of answers	Total of population by position
Professor	9.2	76
"Maître de Conférence"	8.4	142
"Chargé de Cours"	3.5	588
"Maître Assistant"	4.2	452
"Assistant"	16	235

3.2 The Internet use- objectives of uses

a. The great majority of teachers use Internet with research purposes (90 answers out of 142), most of the time during preparation of a thesis (Master or Ph.D.) – an answer given evidently by the "Chargé de Cours", the "Maître-Assistant" and the "Assistant".

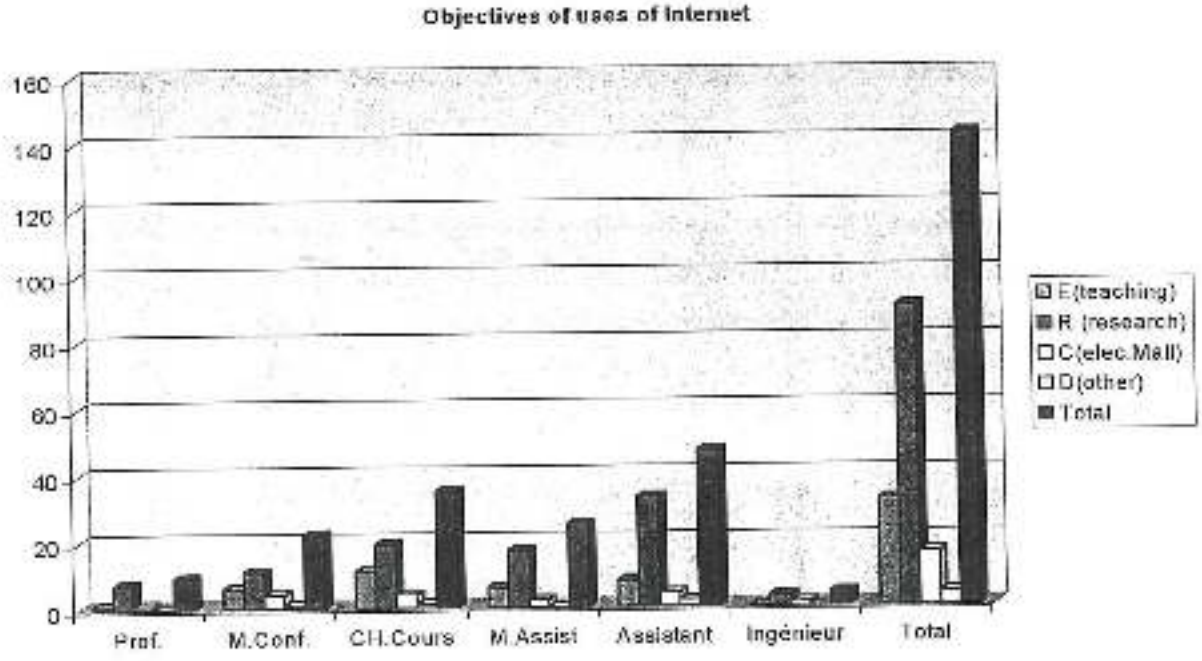
Some times – mainly – the Professors and the "Maître de Conférence" have recourse to tele loading of articles or preprints related to their field of research.

b. The second objective concerns teaching. Internet is used according to teachers, especially the "Chargés de Cours" and the "Maîtres –Assistants" for the preparation of their lectures.

c. Thirdly, the exploitatin of Internet is linked to the rapid exchange of information thanks to the Electronic Mail: an information means that cannot be neglected, but still for the time being less representative. We may say that we need time to be familiarized with such up- to – date tool.

d. From the other possibilities offered by Internet that attract the teachers' attention, we may cite: the press, entertainment- these cultural usage remain, however, of little use.

Thus, the exploitation of Internet concerns in priority the access to bibliographical references in the field of teaching and research.



3.3. The appreciation of Internet

a. Advantages

Most of the teachers agree t recognize that Internet has got advantages thanks to:

- The rapid access to information.
- The economy of time.
- The significant size of information available.
- The possibility to communicate with the whole world.

These parameters have led people to assume that this tool cannot be compared to anything else.

b. Disadvantages

Many teachers have thought about major disadvantage related to the time allocated to the consultation of Internet which seems to be not sufficient – a thing which is "frustrating" for certain questioners.

Other teachers criticize the slow process in getting the answers and the tele-loading which is not all the time possible.

Certain teachers underline the tremendous size of information – this raises the problem of saturation and over – consumption.

Few have dealt with tiresome as a result of Internet use.

We should mention that certain persons did not notice any disadvantage related to the use of this tool.

c. Suggestions

This tool is s appreciated that most of the teachers claim its valorization:

- By integrating Internet in all the institutes (institutes'libraries) – that is to say its extension to all university campuses.
- By providing more time for exploitation - at night if possible –
- By training a staff who may help the users.

This raises here the problem of redefining the role of the professional of information.

3.4. The place of grey literature in documentary research

The teachers - researchers seem to use various types of documents – but certain resources are largely more used than others.

a. Published articles

This type of document is the attraction of all: it comes in the first position with 95 persons out of 101. With a percentage of 41.30 %, it is then a priority.

b. Papers read in conferences

The recourse to such documents is important, especially for the teachers who occupy the highest academic positions : Professor, « Maître de Conférence », « Chargés de Cours ».

c. Books and theses

On the third position come these types of documents sought by means of Internet.

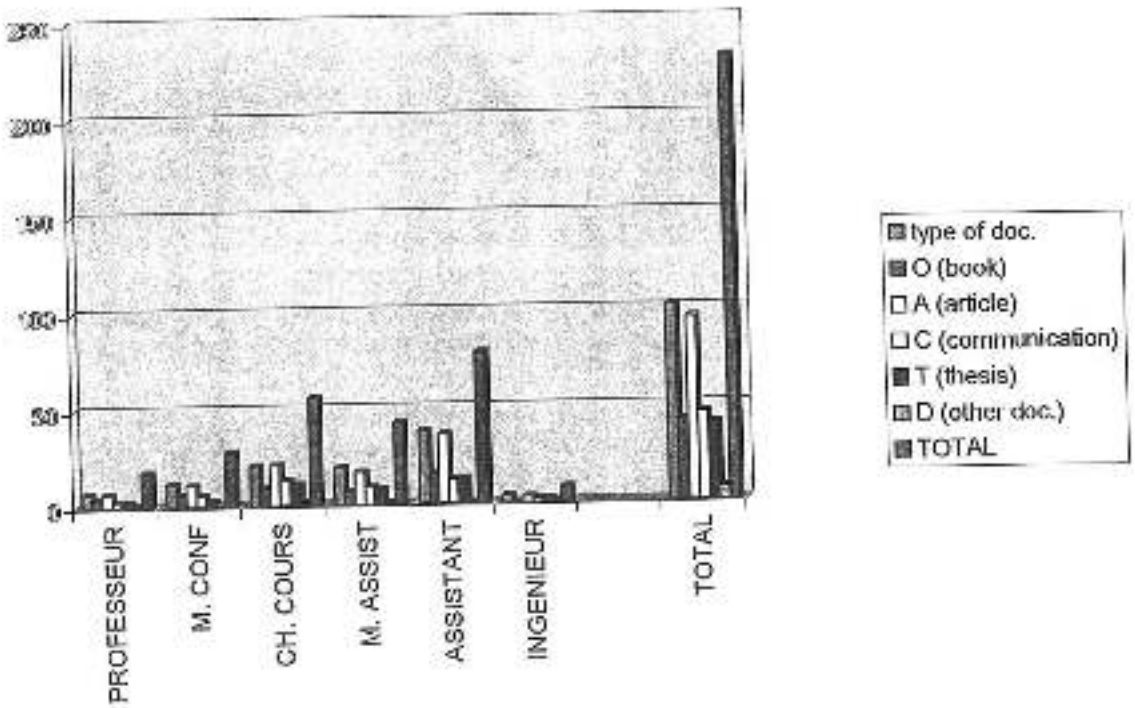
With a slight preference of one document to another, these documents seem to be used mostly by the "Chargés de Cours" and the "Maître-Assistant", whereas the books show more attraction again among the teachers who occupy the extreme position: Professor, "Maître de Conférence" and "Assistants".

d. Other documents

Among the documents that are mostly mentioned, we cite.

- Reports
- Softwares

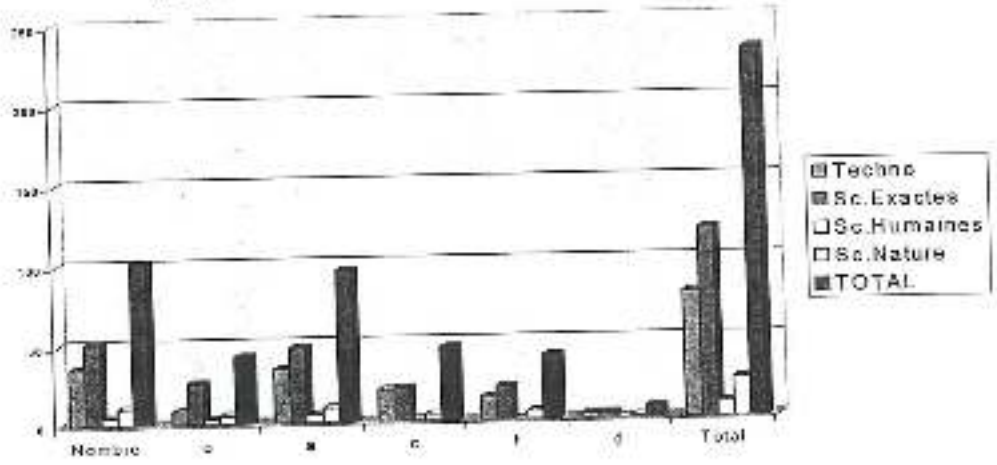
The place of grey literature in documentary research



We notice that the documents of grey literature that are mostly required are the " universitary documents": theses, papers, research reports (especially those related to the university research units).

Moreover, we notice that if the survey has shown that the technological and natural sciences are predominating in terms of the Internet use of grey literature, the human sciences, on the opposite, are practically absent in this field.

The use of Internet and grey literature in the different disciplines



4. COMMENTS ON THE RESULTS OF THE SURVEY

4.1. The appearance of Internet at the university seems to have enhanced an Infatuation for the electronic research among the teachers-researchers. We suppose that there are several reasons:

a. For many, Internet will solve all the problems related to the research and access to scientific and technical information. This attitude is a little bit naïve when we take into account certain assumptions that qualify this tool as "unique" and "without any disadvantage".
b. The use of Internet free of charge is another main reason of its exploitation.
c. For certain persons, it may be a fad to navigate and surf, a manner to show that they move with the movement of fashion.
d. For the majority of teachers, Internet is a chance to be integrated in the network of electronic communication. This electronic mail is tremendously mentioned and allows the questioners :
   - To meet other colleagues from other countries.
   - To be known and to make their research works known.
   - To publish their works in famous international reviews.

We have been convinced – during this survey- that a majority of teachers know what they look for because they know well this technological tool that they manipulate – in Algeria or abroad – for a relatively long time.

We should mention here that according to the person in charge of the service of the SERIST at the university of Constantine about 1000 teachers are homely connected to Internet.

4.2. This survey has also allowed us to know the place of grey literature in electronic research. It seems that the documents that are mostly required are those that are necessary for research and teaching activities – the article published in periodicals has got a priority – this document will certainly remain a pre-requisite research tool for the world scientific community.

The documents of grey literature, mainly universitary documents : theses, papers, reports are documents that – like any document of grey literature – show difficulties to have access to.

Does this type of resources satisfy teachers? This should be the object of another survey.

CONCLUSION

It may be too early to express one's opinion about the effective usage of the Internet by the teachers-researchers of the university of Constantine.

It is a recent means whose mastery is not totally acquired. However, through our modest inquiry, the premises are already significant.

Because it favors the research of technical and scientific information in a fast, efficient and free manner, the Internet will constitute, in the future, an incomparable means of scientific communication.

For the moment, the practices of the electronic research seems to be the exclusive fact of teachers stemmed from scientific or technical fields – as well for searching documents of grey literature.

What confirms our hypothesis, according to which there would be a predominance of scientific disciplines and an absolute absence of social sciences.

The most important is to give a following to this work: carry out inquires to follow the evolution of the Internet usage by the teachers-researchers of Constantine and focusing on important part which we did not mentiopn in our inquiry . The place of the information specialist as a mediator in these new practices of research.

List of Participating Organizations

AWI	Science and Innovation Administration	**BE**
BCOHTA	British Columbia Office of Health Technology Assessment	**CA**
BIOSIS	Publisher of Biological Abstracts and Zoological Record	**US**
BLDSC	British Library Document Supply Centre	**UK**
-	Camera dei Deputati, Biblioteca	**IT**
CASI	Center for AeroSpace Information, see NASA	**US**
CCC	Copyright Clearance Center	**US**
CERIST	Centre de Recherche sur l'Information Scientifique et Technique	**DZ**
CERN	European Centre for Higher Energy Physics, Library	**CH**
CIAO	Columbia International Affairs Online, Columbia University Press	**US**
CLRC	Central Laboratory of the Research Councils	**UK**
CNR	Consiglio Nazionale delle Ricerche	**IT**
CNRS	Centre National de Recherche Scientifique, see INIST	**FR**
CORD	Centre on Rural Documentation	**IN**
-	Cornell University	**US**
CPS	Centre for Property Studies	**CA**
-	CYBEK of New York	**US**
DOE	U.S. Department of Energy, see OSTI	**US**
DTIC	Defense Technical Information Center	**US**
EAGLE	European Association for Grey Literature Exploitation	**UK**
EMICH	Eastern Michigan University	**US**
FID	International Federation for Information and Documentation	**CA**
FIT	Florida Institute of Technology	**US**
GREYNET	Grey Literature Network Service	**NL**
GSFC	Goddard Space Flight Center, See NASA	**US**
GUKCC	Gallaudet University Kellogg Conference Center	**US**
-	Harvard University	**US**
HP	Hewlett Packard Company	**US**
HUT	Helsinki University of Technology	**FI**
ICSSR	Indian Council of Social Science Research	**IN**
IDA	Institute for Defense Analyses	**US**
IFLA	International Federation of Library Associations & Documentation	**UK**
-	Informed Strategies	**US**
INIST	Institut de l'Information Scientifique et Technique	**FR**
IRC	Information Resource Center	**US**

Author Index

A-B

C

D - G

H - J - K - L

M

N - O - P - R

S

T - V

GreyNet's Publication Order Form

INFORMATION RESOURCES		Nº OF COPIES x AMOUNT IN EUROS
GL'99 Conference Proceedings	ISBN 90-74854-24-9	_____ x € 55.00 = _____
GL'99 Program & Abstracts	ISBN 90-74854-22-2	_____ x € 15.00 = _____
GL'97 Conference Proceedings	ISBN 90-74854-17-6	_____ x € 55.00 = _____
GL'97 Program & Abstracts	ISBN 90-74854-18-4	_____ x € 15.00 = _____
GL'95 Conference Proceedings	ISBN 90-74854-08-7	_____ x € 55.00 = _____
GL'95 Program & Abstracts	ISBN 90-74854-09-5	_____ x € 15.00 = _____
GL'93 Conference Proceedings	ISBN 90-74854-03-6	_____ x € 55.00 = _____
GL'93 Program & Abstracts	ISBN 90-74854-04-4	_____ x € 15.00 = _____
GLOSSARY, a vocabulary of terms in GL	ISBN 90-74854-19-2	_____ x € 35.00 = _____
Annotated Bibliography on Grey Literature	ISBN 90-74854-20-6	_____ x € 30.00 = _____
Int. Guide to Persons and Organisations in GL	ISBN 90-74854-23-0	_____ x € 30.00 = _____
Notebook on Grey Literature	ISBN 90-74854-15-X	_____ x € 35.00 = _____
NewsBriefNews: Quarterly Newsletter on GL	ISSN 1389-1804	_____ x € 20.00 = _____

Postage & Handling (Holland)	€ 5.00 = _____
Postage & Handling (Europe)	€ 10.00 = _____
Postage & Handling (Other)	€ 15.00 = _____
SUBTOTAL	=€ _____
○ Reduction of 15 % for GreyNet Members	=€
TOTAL	=€ _____

Name: _____ Organisation: _____

Address / P.O. Box: _____ P-code/City/Country: _____

Please Check one of the boxes below for the method of payment

❑ **Bank check/draft enclosed and payable to GreyNet.**

❑ **Please remit to GreyNet | ABN◆AMRO Bank No. 52.93.61.841 in Amsterdam, Netherlands.**

❑ **MasterCard/Eurocard** ❑ **Visa/JCB Card** ❑ **American Express** ❑ **Diners Club Card**

No. _____ **Expiration Date:** _____

If the name on the credit card is not the same as mentioned above, print the name that appears on the card here

_____

Signature: _____ Date: _____

Please Mail or Fax this Signed Form to: GreyNet, Grey Literature Network Service,
Koninginneweg 201, 1075 CR Amsterdam, The Netherlands
TEL/FAX: 31-20-671.1818 - Email: GreyNet@inter.nl.net