CRIS-Cross: Current Research Information Systems at a Crossroads

Eric H. Zimmerman (Keynote Speaker)
The Research Authority of Bar-Ilan University, Ramat Gan, Israel

Abstract

CRIS: Current Research Information Systems must take center stage as the mechanism of scientific information provision. In the ever-growing complex scientific world, we are faced with mounting challenges. Increasingly we are witness to pressures on universities and scientists to provide answers to the problems of society, while working in multi-disciplinary and cross-border teams, with increased funding competition. The need for quality scientific and technological information is apparent. Content management (with taxonomy) and sound information architecture is key. My thesis is that there is a strong need for intelligent CRISs, and that the Web search engines, powerful as they are, are not a replacement for a good CRIS. CRISs should be used for decision-making at all levels, for the management of research activities, and for the dissemination of results. CRIS, in this respect, is key for facilitating the processes of knowledge creation and management, and hence economic growth.

1 Introduction

With the ERA, the European Research Area, taking shape, and the FP6, the Sixth Framework Programme of the European Commission (significantly different from previous programmes), months away from its inauguration, the European – indeed, global - research community is at a crossroads. In recent years, science systems worldwide have been undergoing radical change. This change includes increased fiscal responsibility, limited research funds, greater numbers of students, increased competition, the changing nature of science in its relationship to society and global economies, and the growing internationalization of science. These changes have occurred just as new and emerging information and communication technologies have been impacting on organizations.

CRIS must take center stage as the mechanism of scientific information provision. In the ever-growing complex scientific world, we are faced with mounting challenges. Increasingly we are witness to pressures on universities and scientists to provide answers to the problems of society, while working in multi-disciplinary and cross-border teams, with increased funding competition. The need for quality scientific and technological information is apparent. Content management (with taxonomy) and sound information architecture is key.

This paper addresses the following points:

1. The nature of CRIS.
2. Why do we need CRIS systems in an age of super-powerful Web search engines?
3. What can CRISs do for us?

My contention is that there is a strong need for intelligent CRISs, and that the Web search engines, powerful as they are, are not a replacement for a good CRIS. CRISs should be used for decision-making at all levels, for the management of research activities, and for the dissemination of results. CRIS, in this respect, is key for facilitating the processes of knowledge creation and management, and hence economic growth.

The world of higher education, science and technology, research and innovation is facing increased complexity. The universities of late, as an example of the changes taking place, have become tremendous "knowledge factories" due to their overwhelming contribution to the production of new knowledge through research endeavors. This has not always been the case. As we increasingly recognize the importance of knowledge, it becomes apparent that the average knowledge worker will be in need of perpetual life-long education and re-training.

We are witnessing a paradigm shift on the global science scene. Science has been shifting from discipline-oriented to cross-disciplinary research. Gibbons (1994) called this Mode Two. This shift in turn is leading us from data through information to knowledge, and from knowledge to wisdom, and from wisdom to insight. This process is spurring increased quality of life and wealth creation, as industrial countries, worldwide, attribute ever-greater importance to R&D because of their acknowledged role as a stimulus for economic progress.

Until recently, much of research has been discipline-oriented, curiosity-led, and motivated and executed by an individual following observation, hypothesis, experiment, or proven method. However, the complex problems of today are such that they require teams with each member having a specialization contributing to the whole. These collaborative teams are often geographically dispersed and of differing disciplines. Researchers today must be proficient in their chosen discipline but also multidisciplinary in their approach to science. This is to enable them to see the „breadth" of problems, but have the requisite „depth" in order to solve the problems.

Increased knowledge, the paradigm shift, recognition of economic stimulus and collaborative interdisciplinary science lead inexorably to the need for systems to assist researchers, administrators, strategists, opinion-formers, entrepreneurs and innovators and also the general public. Systems are needed to provide both information for decision-making and support to the process of knowledge-creation. As Keith Jeffery writes, "The end-user requirement is for the relevant information (relevance, recall), at the right place (wherever worldwide), at the right time (when required), in the appropriate form (optimal presentation, integrated for further use in electronic information / office environments) (Jeffery, 1999, p.5)."

CRIS provide access to and dissemination of research information. This includes People, Projects, Organizations, Results (publications, patents and products), Facilities, and Equipment. The Common European Research Information Format (CERIF), developed by the European Commission[1], is a set of guidelines designed for everyone dealing with research information systems. It is intended to help in the development of new research information systems; to assist existing CRIS systems considering extensions; and to guide CRIS systems on how to structure and index their data.

The advantages of Using CERIF include usability; the common presentation of data providing homogenous view for people accessing data; and the defined format and database structure, simplifying CRIS development and information sharing. The *Added Value* is the guaranteed interoperability allowing for data exchange or homogenous access to heterogeneous information. Further advantages of using CERIF include: nurturing cross-border and cross-disciplinary research; a common model with defined meaning and structure of data, facilitating collaboration between researchers, industry, and policy-makers; common vocabularies making data searchable in a distributed environment; and XML and RDF solutions promoting distributed data retrieval based on leading-edge technologies.

Though the need for CRISs is apparent, the take-up has been slow, and questions have been raised as to their need, given the strength of current Internet search facilities. Commercial success stories are few. Failures are many.[2]

1 In 1991 and again in 1999.

2 INDARD, The Israel National Database of Academic Research and Development, is an example.

2 Best Practice: Developing a CRIS

(Research Information Systems, developed using a Common Information Format, according to Good Practice)

Current Research Information Systems (CRIS) and Common European Research Information Format (CERIF) efforts are not new concepts. As early as in the 1970s, serious efforts were being made among research information systems, in the field of international co-operation.

The problems of subject indexing were already tackled long before the Internet era. In the late 60s and the early 70s, both Smithsonian Science Information Exchange and UNESCO attempted to come towards an international standard taxonomy for fields of science and technology. A standard taxonomy is seen as an important tool in the management of scientific and technological affairs, especially in fields related to science policy and science statistics.

2.1 CRIS

Access to information on current research activities throughout Europe is an essential requirement for the success of EU innovation policy. The key asset in European R&D consists of ideas, technical reports, publications, patents, prototypes, products and know-how – leading to technology transfer and wealth creation, and to the generation of new R&D ideas. The key added value to be achieved is the pan-European approach to the generation of and exploitation of R&D. There is a need for information on currently relevant R&D information to be made available widely to encourage both innovation and new, improved R&D.

The innovators in industry and services, the academics pushing the frontiers of R&D, the decision makers in governments and R&D funding agencies all require easy-to-use desktop access to R&D information. The raw data sources are the R&D information held by funding agencies and other information providers in the EU. These are held for the particular purposes of the agencies and the particular clients of the information providers. They are heterogeneous and unconnected. The potential for European wide exchange is not being exploited enough.

There is a need for the consumer of research information to be able to access this data through a uniform familiar interface and to be able to integrate and compare the information between data sources. This “common interface” must not only address the content (what must be exchanged) but also the format of such information (how it should be presented). This information must be presented in a uniform way, at least at summary level. Classification should be consistent for all the research information sources. Subject indexing is required and a controlled terminology should have the same meaning in all languages. To this end, the Common European Research Information Format was proposed in 1991 and revisited in 1999.

The definition of this uniform information description platform requires:

- the definition a full CRIS data model which will cover the database structures of the majority of existing CRIS;
- the definition a set of data models which could provide examples for data exchange (since there are an infinity of possible exchange data models between CRIS;
- the definition of a metadata data model to provide a uniform summary-level view over heterogeneous information sources.

Easy access to information must address not only the availability of information with a common definition and format but also how the consumer could retrieve that information. The consumers need to be able to search, European-wide, for information on a particular research topic or theme. Subject indexing of the information is the obvious key in this respect. Classification should be consistent for all the research information sources; otherwise people will not get consistent results when they retrieve information. Since consumers also use different languages, the controlled indexing terminology proposed should have the same meaning in all languages.

2.2 CERIF

The European Union's Innovation policy aims at improving and strengthening the generation and exploitation of current and new Research and Development (R&D) projects as well as technology transfer. To this end, access to information on current research activities throughout Europe is an essential requirement. New R&D ideas can emerge thanks to a pan-European approach for information sharing and exchange. There is thus a need for a convenient tool to spread relevant R&D information widely to encourage innovation and improved R&D as well as wealth creation.CERIF: Common European Research Information Format provides a practical common standard for information contents and for subject indexing. Further, controlled value lists ease the collection and exchange of data. To provide easily searchable information implies adapting common rules. To this end, it is important to use standard controlled vocabularies. The use of standard controlled vocabularies and standardised data structures, as described further in the CERIF guidelines, should be as wide spread as possible, in order to make both data providers and end-users familiar with common CRIS characteristics.

CERIF is this common language that fosters the diffusion of information across Europe. The EU Commission's Green Paper on Public Sector Information[3] emphasizes the importance of access for European citizens to publicly funded information and equally the opportunities for economic growth and employment that it provides. All Member States are taking initiatives with regards to public R&D information – at an uneven pace. European Union policy should therefore aim to have all Member States arrive at the same point, and as quickly as possible. CERIF in addressing Public Sector R&D information is dealing with an area of economic activity with a high growth potential, and is crucial to the ongoing competitiveness of European industry.

2.3 The Code of Good Practice (CGP)[4]

There are many organizations, large and small, participating in research projects throughout both Europe and the rest of the world. The continuing evolution of the supporting technology (in particular the Internet and Computer-Mediated Communication tools) is making communication available to even the smallest organization. The Code of Good Practice was drafted to establish a framework for encouraging interoperation and harmonization between European and all research institutions worldwide. The intent was that this document be regarded as a set of good practices for existing research institutions.

The CGP was developed as a guide for both new and existing producers of CRISs. The intention was to focus clearly on the reasons for having a CRIS and on the main components of the system. The CGP is not a comprehensive guide to building and developing information systems.

As a stand-alone system each CRIS plays an important role for the host institution or organization. Together, a collection of CRISs is potentially a very powerful information tool, the true value of which can only be harnessed if interoperability can be achieved. Universal adoption of the CGP by CRIS producers for both new and existing CRISs will be a significant step towards realizing this goal and will provide CRIS users and data providers alike with a framework for knowledge transfer. Likewise, there will be greater scope for the CRIS institutions to exchange data for mutual benefit or commercial gain.

The cornerstone of the CGP is consistency. This is the single most important benefit that will result from the adoption of the CGP. Consistency will lead to:

- increased usability of data and value of CRISs to the users (often with varied demands and requirements);

3 Public Sector Information: a key resource for Europe", Green Paper on Public Sector Information in the Information Society, European Commission, COM (1998) 585 http://europa.eu.int/ISPO/docs/policy/docs/COM(98)585/ <accessed 2 May 2002>

4 Developed by euroCRIS <http://www.eurocris.org> and the European Commission (CORDIS), 1998.

- increased interoperability between CRISs;
- lower operating costs for CRISs;
- reduced effort for information exchange.

In order to realize the overall benefits offered by the CGP, it is first necessary for all relevant parties to adopt the CGP and then ensure its implementation within their working environment as a recommended standard or norm to be used. The increasing accessibility of information through developments in technology further emphasizes the need for consistency (with regard to information exchange) to ensure that the wealth (and potential diversity) of information available locally is accessible globally.

3 Why CRIS and Not a Web Search Engine?

The system should work harder so that the user need not. Rather than force the user to choose the best query term, the system should be able to perform intelligent searches in a multi-lingual environment, with the assistance of a controlled vocabulary (ontology / taxonomy). Below we discuss the problems associated with jargon and language through the collaboration of multinational research teams.

"Power" users, relying on CRIS for decision-making, must be able to find all relevant information in a homogenous fashion, with a high degree of precision, and with minimum noise. This is possible only with a dedicated CRIS. Internet search engines cannot provide users with all this – yet. CRIS, based on CERIF, offers the combination needed to aggregate information from distributed heterogeneous systems and allow the user to access the information via a homogenous interface.

As we have demonstrated above, developing research information systems according to agreed-upon conventions of good practice, and according to uniform data standards, offers a powerful tool, unmatched by the most powerful of today's Internet search facilities (engines and directories alike).

3.1 What can CRISs do for us?

The functions of CRIS are as varied as there are information systems. They can be used, among other things, to detect and map trends in science, identify discipline specialists, find specialized equipment or facilities, recognize innovations and results (to avoid duplication of effort), manage the grant process, produce statistics and reports, evaluate projects and assess science, promote science in society, and to locate funding sources. In these lines we examine but a few of the major applications of CRISs.

3.1.1 Information-Sharing (Knowledge-Sharing) and Communication Between Scientists

As some have pointed out, CRIS can be a facilitator of cross-discipline research. Jewitt and Görgens (2000, p. 410) point out that "multidisciplinary communication is one of the missing links of science." Koku et al writes (2000) that increased specialization has added urgency to the need for communication. CRIS can be used to power the communication networks of 21st century science.

An intelligently designed CRIS may provide context to science. With ever-increasing cross-disciplinary science being performed across borders, the importance of preserving context cannot be overstated. As Gibbons has stated (1994), Mode 2 knowledge is produced within a context of application. The involvement of different perspectives, languages (and discipline-oriented

jargon[5]) and cultural (national) settings, by teams of people with heterogeneous skills, mandates that a system to preserve context exists.

Gloria Mark (2000) writes that to build a culture of true information sharing on the Web, we must begin to think in terms of communal information repositories. No longer would we view the Internet as a set of individual knowledge warehouses.

CRIS can also be used as the engine of *Collaboratories* and collaborative white-boards. Access rights could be carefully monitored.

With the spread of electronic mail, the world witnessed a growth of invisible colleges. With the Web we are witness to the flourishing of "communities of practice". euroCRIS is a well-defined community of practice, and with many members in academia, we may be considered a college as well, as we promote the free sharing of ideas, information and insight.

3.1.2 Decision-Support and Statistical Analysis

CRIS can be used for Budgeting and Reporting at all levels. Also, strategic and operational decision-making at all organizations is based on good information. This information comes from multiple varied sources (i.e. experts), yet a lone individual often makes the decisions. CRIS data can fuel the support systems necessary to guide decision-makers.

R&D Unit and Researcher Assessment at the local (university) and national level will benefit from a system of comparable indicators. CRIS, taking into effect (and recording) global diversity, can develop a system of indicators for the international comparison of science.

It has been shown that science policy making worldwide has lacked the ability to compare data sets, especially in the social sciences.[6] CRIS could prove instrumental in integrating analogous data from different countries. Standardization is key to the success here, and the CERIF recommendation would play an instrumental role. This will prove to be of immense importance with respect to the ERA (see below).

3.1.3 Research Administration

CRIS can be used (and indeed is used) to support the management of institutional expertise, the administration of the proposal development process, the submission and peer-review process, the supervision of grants, and the publication of results. In this sense, when working as an integrated whole, CRIS is a full life-cycle research knowledge management system. In the United States this system is known as Electronic Research Administration (ERA – not to be confused with the European Research Area, also known by the same acronym). Hunt Williams (2000) termed this the "research process" – from the conceptualization of a research idea through the commercialization and exploitation of results. Jeffrey (1999) lists several of the life-cycle components, from a "products" angle. They are publications, patents, products, results, know-how and IPR, education and training, and publicity. Each stage or product must be managed.

Having the right information on funding opportunities is crucial for scientists, in today's increasingly competitive funding environment. Traditionally, argues Richard Tomlin (2000), funding opportunities information systems have been discipline oriented. With increasing interdisciplinary research and new research paradigms, Tomlin claims we might be doing harm by not allowing (helping) the individual scientist to see the bigger picture. By targeting information based on discipline-specific keywords, we might be reducing information overload, but we might also be inhibiting novelty, serendipity and innovation.

Barriers to grant-getting among junior faculty can be attributed to a lack of mentoring and to a need for faculty development programs (Boyer and Cockriel, 1998). CRIS can be used to fill this need by cataloguing writing guides, proven techniques, and the like.

---

5 See, for example, Jewitt and Görgens, p. 411.

6 http://www.oecd.org/dsti/sti/prod/intro-24.htm

3.1.4 Publishing and Access to Research Information

Gibbons (1990) writes that in Mode 2 science, results are disseminated among the producers in the first place. The diffusion of the results is part of the knowledge production process. Later, results are published via the normal academic channels. As stated above, it is imperative that the context – ever evolving, as new problems are tackled – not be lost.

CRIS can be used to publish research results in peer-reviewed electronic journals, making full use of multi-media and hypertext capabilities. Likewise, preprints can also be published via a CRIS-based system.

CRIS can be used as the backbone of digital repositories (digital libraries). Similarly, CRIS can be used to drive Web-based scientific sites.

3.1.5 Knowledge Management (KM)

KM has, it can be argued, two main functions: "community of practice" building and adding structure to a rich body of knowledge so that it might be used and reused. CRIS can serve both functions effectively.

The management of "communities of practice" necessitates the development of expertise databases (i.e. Community of Science) or "competency dictionaries".

Mode 2 Knowledge Production, as postulated by Gibbons (1994), is performed in transient group settings. Networks are developed on an ad-hoc basis. When a given problem is solved, these networks are oftentimes dissolved. It is vital that the experiences of these groups be captured. This knowledge management is crucial if new knowledge is to be transferred to solving a new problem.

CRIS can be the natural bridge between scientific information and innovation information; two different yet complementary types of information, necessary for the presentation of the "complete picture" of science. Whereas the general aim of science is to understand nature and society, innovation seeks to create new products, processes and services.[7] In recent years we are witness to the growing dependency of innovation on science. This rapidly expanding body of knowledge could be captured by CRIS.

With ever increasing amounts of information available it is becoming more important to be able to find the relevant information needed. Productivity cannot be jeopardized by the time needed to sift through the vast amount of data and information, of varying quality. Efficient retrieval can be accomplished by using CERIF to control the "language of science". The key here is to develop and maintain a global taxonomy of science. This will help all actors drive science forward, by filtering out the "noise" associated with high recall and low precision (relevance) of retrieval. Though some stages of the information coding can be automated, the intervention of human information specialists would be crucial in order to ensure the proper monitoring and control of language, concepts and rules.

Gray Literature is increasingly recognized as an important area of research information study, because "in a R&D environment [it] represents the cutting edge of this knowledge and so its management is of utmost importance (Jeffery, 1999, p. 1)". As Jeffery writes, an organization documents and stores its knowledge assets within gray literature.

3.2 European Added Value: CRIS as the Lynchpin of the European Research Area (ERA)

In its tender for the ERIS study[8], the European Commission stressed the importance of a European research information system as "… co-ordination of national and European research

7 The Management of Science Systems, OECD, 1999, p. 5.

8 Contract N° COPO-CT -2000-00002

programmes, including the mutual opening-up of national programmes, mapping of excellence, definition of a European approach to research infrastructure, better use of instruments of indirect support to research, benchmarking of national research policies…" The Commission cited that while initiatives exist (i.e. CORDIS, ERGO and euroCRIS), no comparative data of national research policies is readily available, under "one roof".

CRIS can be used as a Facilitator of Integration in that access could be given equally to less prominent researchers, less prestigious institutions, and less favored regions. This would have a democratizing effect on science.

It is known (Oxbrow and Abell, 2002, p. 26) that Europe suffers from a low entrepreneurial spirit (as compared to the United States for example) and less professional mobility than the US. Barriers to this no doubt include language and culture. CRIS can be instrumental in bridging these barriers. Europe's strength (as compared to the US) might be said to be in its networking capabilities. Here euroCRIS will prove instrumental in spearheading the drive for increased CRIS awareness.

The European Commission recently conducted its most recent "Eurobarometer" survey of attitudes towards science in Europe. As the survey shows, there is work to be done in the promotion of science in Europe. CRIS can play a role. The report demonstrates that scientists have a strong image in Europe, though the knowledge they possess affords them much power (European Commission, 2002, p. 3). The survey goes on to report that most Europeans do not believe science will help to "eradicate" poverty and famine (European Commission, 2002, p. 8), though a majority do believe a cure for AIDS and Cancer will come, and that science and technology will help pave the way for a better tomorrow. Basic science is also highly regarded. Finally, Europeans do not know much about European excellence in science (European Commission, 2002, p. 14), though they would like to see Europe more heavily involved. However, they think better organization is key here, not increased funding. This is where the role of CRIS can be pivotal.

4 Barriers to Wider Use of CRIS Systems on Global Scale

There are several issues that have yet to be settled, owing to the dynamic nature of the Internet and new forms of collaborative work. These are listed here, but left for a more critical review in a future study. The issues include information quality and timeliness, language and the ability to convey social cues and non-verbal communication, information presentation (maximizing the Web's capabilities), intellectual property rights, publication and copyright, liability and other legal issues, privacy, national security, and foreign access to sensitive data.

A major barrier to greater use of CRIS is strengthening awareness of CRIS among scientists, administrators and policy-makers. CRISs have generally been considered a starting point for locating information, not the repository of what is actually being sought.[9] Well-established researchers have long felt that they do not need the services of such systems to locate collaborators or information, as the formal and informal networks serve this purpose better. This may no longer be the case with the growing cross-disciplinary nature of science.

5 How are CRISs currently used around the globe?

During April-May of this year we are conducting a global survey of CRISs. The results will be presented at CRIS2002. It is expected that we shall learn from this study of the uses of CRIS, the problems facing CRIS developers, directions for future development, and how global CRISs can cooperate to enhance science. A catalog of CRISs is available on the Web, sponsored by NIWI,

9 Jostein Helland Hauge, Director of the Research Documentation Section, Bergen University Library, Norway, personal email, 8 April 2002.

the Netherlands Institute for Scientific Information Services. DRIS, the Directory of Research Information Systems, may be accessed at
http://www.niwi.knaw.nl/cgi-bin/nph-dris_search.pl?language=us.

6 Summation

CRIS can and must take center stage as the mechanism of information provision in the world of science. In the ever-growing complex scientific world, we are faced with mounting challenges. Increasingly we are witness to pressures on universities and scientists to provide answers to the problems of society, while working in multi-disciplinary and cross-border teams, with increased funding competition. The need for quality scientific and technological information, available to all, is apparent.

In our knowledge-based society, where so much is dependent on scientific discovery, we must be positioned to intelligently make use of information and apply it to address the problems and issues of the day.

7 Acknowledgements

The author wishes to thank his many colleagues at euroCRIS and Bar-Ilan University for years of stimulating conversations, both face-to-face and digitally – especially Keith Jeffery, Anne Asserson, Jostein Hauge, Bernd Niessen, Peter Finch, Richard Tomlin, Hunt Williams, Natalie Lapidot, and Andrei Lopatenko.

8 References

Boyer, Patricia and Cockriel, Irv. (1998): Factors Influencing Grant Writing: Perceptions of Tenured and Nontenured Faculty. In: *SRA Journal*, Spring, pp. 61-68, USA.

euroCRIS and the European Commission. (1999): *CERIF 2000 Final Report*, Luxembourg.

euroCRIS and the European Commission. (1998): *ERGO Pilot Project Final Report*, Luxembourg.

euroCRIS and the European Commission. (1998): *Code for Good Practice*, UK.

European Commission. (2002): Europeans, Science and Technology: Survey Findings. In: *RTDinfo*, Special Edition, March, Luxembourg.

Gibbons, Michael (1994): *The New Production of Knowledge*. London: Sage.

Jeffery, Keith G. (1998): The Future of CRIS. In: *CRIS98: The Nutcracker*. <Accessed 14 April 2002>.

Jeffery, Keith G. (1999): *An Architecture for Grey Literature in a R&D Context*, unpublished.

Jewitt, G.P.W and Gorgens, A.H.M. (2000): Facilitation of interdisciplinary collaboration in research: lessons from a Kruger National Park Rivers Research Programme project. In: *South African Journal of Science*, Volume 96, August.

Koku, Emmanuel, Nazer, Nancy and Wellman, Barry. (2000): Netting Scholars: Online and Offline. In: *American Behavioral Scientist*, Volume 43 (preprint of article), USA.

Mark, Gloria. (2000): Social Foundations for Collaboration in Virtual Environments. In: *Access to Knowledge: New Information Technologies and the Emergence of the Virtual University*, Tschang, Ted and Senta, Tarcisio Della, Pergamon Press, UK.

OECD. (1999): *The Management of Science Systems*. <Accessed 1 December 2001>.

Oxbrow, Nigel and Abell, Angela. (2002): Is There Life After Knowledge Management? In: *Information Outlook*, 6:4, April, Special Libraries Association, USA.

Tomlin, Richard. (2000): *CRIS and the Challenge of New Research Paradigms*. In: CRIS2000, June, Helsinki. <Accessed 14 April 2002>.

Williams, Huntington. (2000): Research Partnerships in the Internet Age. In: *Community of Science*, USA.

9 Contact Information

Eric H. Zimmerman
Bar-Ilan University
The Research Authority
The Begin Building
Ramat Gan 52900
Israel

e-mail: zimmee@mail.biu.ac.il