The IDRC Digital Library: an open access institutional repository disseminating the research results of developing world researchers

Barbara Porrett

Research Information Management Services Division
International Development Research Centre (IDRC CRDI)
P.O. Box 8500
Ottawa, ON
K1G 3H9
Canada
e-mail: bporrett@idrc.ca
web: http://www.idrc.ca/

Abstract

The International Development Research Centre (IDRC) has recently launched the OAI-PMH compliant IDRC Digital Library (IDL), a DSpace institutional repository. The digital library has been developed to enhance the dissemination of research outputs created as a result of Centre-funded research. The repository has a number of unique qualities. It is the public bibliographic database of a Canadian research funding organization, its subject focus is international development and the content is retrospective, dating back to the early 1970s. Intellectual property issues have been a major factor in the development of the repository. Copyright ownership of a majority of IDL content is held by developing world institutions and researchers. The digitization of content and its placement in the open access IDL has involved obtaining permissions from hundreds of copyright holders located in Africa, Asia and Latin America. IDRC has determined that obtaining permissions and populating the repository with developing world researchers' outputs will help to improve scholarly communication mechanisms for Southern researchers. The expectation is that the IDL will make a contribution to bridging the South to South and South to North knowledge gap. The IDRC Digital Library will serve as a dissemination channel that will improve the visibility, accessibility and research impact of southern research.

Keywords: developing world research; institutional repository; open access; DSpace; IDRC Digital Library; International Development Research Centre

Introduction

The subject of this presentation is an institutional repository called the IDRC Digital Library [1]. The repository has several unique qualities that distinguish it from other DSpace institutional repositories now accessible on the Internet. It is the repository of a research funding organization, it serves as the organization's public bibliographic database for the dissemination of funded research outputs and public corporate documents, its content is retrospective, dating back to the early 1970s and as a result, its development and management has presented some significant intellectual property issues. Notwithstanding these and other challenges, the IDRC Digital Library is developing into a significant resource, sharing the research results of developing world researchers with the international research community.

IDRC stands for the International Development Research Centre, a Canadian Crown corporation that works in close collaboration with researchers from the developing world in their search for the means to build healthier, more equitable, and more prosperous societies. The creation of IDRC was Canada's response to a climate of disillusionment and distrust that surrounded foreign aid programs during the late 1960s. Maurice Strong, a distinguished Canadian who through his contributions in both private industry, at the United Nations and other NGOs has become one of the world's most influential political and environmental activists played a key role in the founding of IDRC. He and others urged the then Canadian Prime Minister Lester B. Pearson to establish a "new instrument" to provide forward-thinking approaches to international challenges that could not be addressed by way of more conventional programs. This led to the establishment of the world's first organization devoted to supporting research activities as defined by developing countries. IDRC's

objectives, as stated in the *International Development Research Centre Act* of 1970, are "… to initiate, encourage, support, and conduct research into the problems of the developing regions of the world and into the means of applying and adapting scientific, technical, and other knowledge to the economic and social advancement of those regions."**.** IDRC is guided by a 21-member, international Board of Governors and reports to the Canadian Parliament through the Minister of Foreign Affairs. In 2007/08, IDRC received CA$135.3 million in funding from the Parliament of Canada.

IDRC and the Dissemination of Funded Research Results

IDRC has, from the onset, placed a great deal of importance on the sharing of the research outputs that are created as a result of Centre-funded research. Although the copyright ownership of the outputs has always remained with funding recipients, it has been a condition of funding that IDRC maintains the ability to disseminate the research outputs supported by Centre funding. An archive of these outputs has been maintained since 1970, originally on paper but now increasingly in digital format. Bibliographic management of this archive has been done through a library catalogue and more recently through an online public access catalogue that has been accessible to the research community on the IDRC Internet web site.

In an effort to enhance the dissemination of these research outputs and to provide an improved scholarly communication mechanism for Centre-funded researchers, it was decided in the fall of 2005 to explore the possibility of building an Open Access Initiative (OAI) compliant institutional repository. Under the guidance of a Steering and a Stakeholders Committee, and a policies and governance document [2], a project team of two librarians and a systems analyst undertook the initiative. In April 2007 a DSpace open access institutional repository, called the IDRC Digital Library or the IDL was launched.

Content of the IDRC Digital Library

The IDL provides access to information about the IDRC research output archive dating back to the Centre's beginnings. The database holds 34,000 Dublin core metadata records, approximately 30% of which provide links to digital full text. The subject coverage reflects the international development focus of IDRC research funding, with strong representation from the sciences and social sciences. The subject areas of research that have been supported by the Centre have changed over time. Research funding currently focuses on the following five themes: Environment and Natural Resource Management; Information and Communication Technologies for development; Innovation, Policy and Science; Social and Economic Policy; and Global Health Research. An average of slightly over 500 IDRC-funded research projects are active at any point in time and approximately 750 research outputs are added to the archive each year.

Audience of the IDRC Digital Library

The primary audience of the repository is the international research community. This includes researchers, applicants for IDRC funding, donor agencies, policy makers, and development practitioners. The repository's purpose is to share the research results of developing world researchers, to facilitate the discovery of research literature in the fields of international development and to identify researchers, research institutions and civil society organizations that have undertaken research in the fields of international development. The IDL not only enhances the public accountability and transparency of IDRC-funded research, but also demonstrates the Centre's commitment to contribute to the global "public good" contribution of the research it supports. It also ensures that the research results will be freely accessible in order to contribute to the public debate on development issues for public benefit. The research literature in the IDL can be accessed and used for non-commercial purposes in accordance with a definition of open access based on the Budapest Open Access Initiative. The expectation is that the IDL will make a contribution to bridging the South to South and South to North knowledge gap. These channels of scholarly communication and scholarly publishing are less heavily traveled than the North to North and North to South. [3] The IDL will serve as a dissemination channel that will improve the visibility, accessibility and research impact of southern research.

Focus of the Presentation

This presentation will focus on three aspects of the IDRC Digital Library: IDL content and how it will continue to develop, copyright permission challenges presented by the repository's retrospective content, and IDL services. Evidence of use of the repository will also be discussed.

Development of IDRC Digital Library Content

The bulk of the content disseminated by the IDL is in the form of final technical reports. The reports present the research results produced by Centre-funded researchers. They are submitted by funding recipients as a requirement of research funding. The IDL also disseminates books published by the Centre, documents and other writings by staff and IDRC governors for and about IDRC, as well as other substantial works related to the Centre's programs, projects and activities. This second category or collection of content represents about 25% of the digital library's content.

These two collections have, historically, been housed in the Centre's Library and managed in the library catalogue. The Library's catalogue records are the source of the majority of the digital library's metadata. These were mapped and migrated from the Library open public access catalogue or OPAC into MIT's DSpace. This kind of undertaking can be perilous, even under the best of circumstances. The library OPAC software, called MINISIS was home-grown, originally designed to be used by developing world libraries. The non-standard/non-MARC record structure of the MINISIS bibliographic records presented significant challenges. Further, the record content and database structure had changed over time. Migrating this content into DSpace was much like opening a Pandora's box. Countless unanticipated challenges had to be overcome but after a great deal of problem solving, some metadata field customization and two migration attempts, an IDL database with an acceptable level of integrity has emerged. The IDL now serves as the Centre's public bibliographic database.

Submission Process of the IDRC Digital Library

The submission and metadata creation process for the IDL is centralized in the IDRC Library. The Centre's research subject specialists, called program officers, review final technical reports received from research project recipients. Although the review is not a true peer-review, it can lead to redrafting of the reports by the funding recipients to ensure that they meet Centre funding requirements. Once finalized, the program officer determines if the report is eligible for public dissemination in the IDL. For example, reports containing politically sensitive or patentable information are not added to IDL holdings. Further if a researcher requests that their research results not be placed in the IDL because, for example, they have published or wish to publish with a publisher that does not permit dissemination on an OA platform, similarly, the report will not be added to the IDL. A Centre-funded researchers' ability to choose not to have their final technical reports published in the IDL results from the fact that the contractual agreement for research funding has only just been modified to make provisions for OA and digital dissemination of final technical reports. IDRC-funded outputs produced by research projects approved after January 2008 will submit outputs to their program officers in digital format and the researchers will have granted the Centre permission to disseminate their funded research results in the OA IDL. A soon to be released IDRC publishing policy reiterates these new conditions of IDRC funding. The impact of this contractual and policy change is that submission of final technical reports to the IDL will be mandatory. However, the implications of this will not be seen by the IDRC Library for a number of months, because the outputs are received only after the research has been completed.

If a report is destined for the digital library, the program officer determines where it will be placed within the IDL's browsing structure. That is, in which collection of his/her DSpace community. Incidentally, the community and collection structure of the IDL has been created in collaboration with the Centre's programming staff. This approach to the development of the IDL browsing structure is an example of how the Library has attempted to share ownership of the IDL with the Centre's program branch. The researchers and the program officers are asked to provide uncontrolled vocabulary or keywords for the report's IDL metadata. This is being done with the hope that keywords in the metadata that have been recommended by the subject specialists and/or the researchers will help to enhance the retrievability of the digital library's content.

Four pieces of information, an indication that the report is destined for the IDL, the appropriate collection name, plus keywords and the report are emailed by the program officer to a records management staff member. These are then placed in the Centre's digital records management system by records staff. This information is transferred manually into the IDL by a library cataloguing technician who completes the metadata creation and submission process. Additional subject

description is added to metadata records using the OECD Macrothesaurus. Automating this process to enable migration of this information from the records management system to the digital library is planned.

IP Issues and the IDL

Seventy percent of the outputs described by IDL metadata are in paper format and as mentioned earlier, the copyright ownership of funding recipient created research outputs is owned by the researchers. In order to comply with Canadian copyright law, permission must be obtained from the copyright holders before the format of the outputs can be changed from paper to digital and made accessible through the open access digital library. This then leads to the subject of copyright permissions and digitization. Developing world researchers have encountered and continue to face barriers to the publication of their work. To ensure that IDRC research funding does not further impede efforts to publish, the contract between the Centre and its researchers places copyright ownership of final technical reports with the funding recipient. Obtaining permission to digitize and to place final technical reports in the IDL has been a full time occupation of a library staff member since the fall of 2006. To date, approximately 450 copyright holders have been contacted and asked to complete and sign a license granting IDRC permission to digitize and place their research results in the IDL. Many of the copyright holders are developing world institutions that hold the copyright of numerous works.

The success rate in obtaining permissions is in the 65% range. It has not been difficult to obtain permission to digitize and place Centre-funded outputs in the IDL if it was possible to contact the copyright holder. For the most part, the copyright ownership of the outputs has not been transferred to publishers and, with just a few exceptions, copyright holders were willing to grant permission. Copyright holders have not make requests for further information about open access. How this should be interpreted is not clear, however, the correspondence requesting permissions has been carefully drafted in an effort to ensure that its intent is not misunderstood. Recipients have asked to be notified when their outputs were accessible in the IDL, in one case because they planned to place their digitized research results on their own web site. The impediments to obtaining permissions can be summarized as follows: the copyright holder is deceased, the research institution that received the research funding no longer exits, it was not possible to identify the copyright holder or a reply to correspondence requesting permission just has not been received.

The Library has the capacity to continue to request permissions from copyright holders and to undertake in-house digitization with the objective of expanding the digital content in the IDL. However, not surprisingly, it has been difficult to locate the copyright holders of many older outputs. Regrettably, it is unlikely that the IDL will be able to disseminate the digital full text of all the final technical reports that its metadata describes.

IDRC Digital Library Services

This then means that IDL metadata will continue to describe final technical reports that are not delivered digitally by the IDL. In an effort to enable users to access these outputs, the IDRC Library does its best to offer a document delivery service. Users are invited to enquire about options for accessing the research results.

IDL users can, of course, visit the IDRC Library in Ottawa but this is not a practical choice for many researchers. The Centre's contractual agreement with recipients funded after February 2004 enables the Library to digitize an output and make it available on the IDRC web site but not in the OA IDL. In cases where IDRC can not obtain permission from copyright holders to disseminate their outputs and the project contract predates February 2004, the Centre may rely, to a limited extent, on the so–called 'fair dealing' exception under Canada's *Copyright Act.* This exception provides only a very narrow exclusion to allow a library to copy and distribute a portion of a work without it infringing copyright in that work. The library must be satisfied that the use of the work will be for research or private study. The law does not set clear limits on what portion of a work may be copied under the fair dealing exception. But, what is clear is that copying an entire work would not be permissible under the fair dealing exception. This is not an ideal situation but the document delivery staff attempt to do their best to meet the information needs of requesters.

Another service being explored by the IDL is the hosting of works authored by developing world researchers who are not IDRC-funding recipients. A Centre-funded project has developed a research methodology that is being applied by non-Centre funded researchers in the developing world. The project's lead researcher recognized the value of managing and disseminating the results of this

disparate group of researchers. A partnership was established to create a DSpace community that makes the research results OA accessible through the IDL. A service agreement addressing issues such as content review, intellectual property, metadata creation and termination of the collaboration was developed to formalize the partnership. This led to the creation of the Social Analysis Systems2 (SAS2) Community [4] in the IDL. This community not only disseminates developing world researchers' work, but also facilitates the aggregation of a body of knowledge.

Integrating the IDRC Digital Library into Other Centre Systems

The IDL has been designed to integrate with a suite of other repositories of information created by IDRC. For example, as described earlier, final technical reports are filed in the records management system. The documents and their skeletal metadata are reused in the digital library. This eliminates the need for submission to both the records system and the IDL.

Further, IDL content can be accessed through the Centre's project database, called IDRIS+ [5] and its persistent URLs are widely used in the IDRC web content management system. Integration enables reuse of the IDL's content through other IDRC systems and will help to ensure long term funding and survival of the repository.

Use of the IDRC Digital Library as a Research Resource

Preliminary data indicates that the IDL is on its way to accomplishing the objective of making a contribution toward bridging the South to South and South to North knowledge gap.

The context of this data is as follows. The absence of links to digital full text in 70% of the IDL's metadata records has made it possible to gather some information about who is using the IDL as a research resource. Users are contacting the IDRC library to enquire about receiving the full text of outputs that are described but not delivered digitally. The majority of the requests are received by email. Although it is not always possible to confirm that the requester originates from the developing world (many developing world researchers are studying and working in developed world institutions and organizations), these requests for full text do reveal the following. The total number received between July 2007 and mid April 2008 was 96. The mailing address and/or signature indicate that 53 were writing from the South, approximately 55% of the total. The majority of these came from Africa and Latin America. A smaller number were received from India, Vietnam and Cambodia The origin of the remaining 45% was, in order of frequency, Canada, the U.S., the UK, and France.

The IDL's server log files are not available for analysis at this time. However the DSpace application provides a statistical summary that reveals some interesting information about the system's use. Data collection began in November 2007. The IDL has been searched an average of 16,000 times per month, an average of 81,000 items have been viewed and an average of 35,500 bit streams or digital files have been accessed each month. The words searched by IDL users are also noteworthy. The French, Spanish and English languages are equally well represented among the terms being used. Although the presence of French and English is not surprising, the numerous terms in Spanish may indicate that the IDL has caught the attention of Latin American researchers. Search terms such as reformas, gouvernance, poverty, tecnológicas, rurale, policy as well as developing world geographic locations are high on the list of frequently searched words. All of these terms reflect the research areas funded by IDRC. The terms also suggest that there is a strong potential that searchers' information needs are being met by the IDL.

Conclusion

By way of conclusion I would like to note that IDRC is the first Canadian research funding organization to build an OAI-PMH compliant institutional repository to disseminate its funded research results. It was the vision of Marjorie Whalen, the IDRC Library director that led to the creation of an IDRC institutional repository to enhance the dissemination of southern researchers' research results. The experience of the project shows that challenges, some expected, others not, were inevitable but not insurmountable. The collaborative nature of this undertaking has been enriching for all of us at the Centre. But content development of the IDL is far from complete. Obtaining consent from copyright holders to distribute their works through the IDL will remain a high priority for some time. This is consistent with the belief at IDRC that open access will lead to the maximization of the societal benefits of investment in research.

To close, I would like to share the following which were sent to us by a researcher.

Thank for your email request for permission to include my works in the IDRC Digital Institutional Repository.

I am a firm believer on the universal right of the people of the World to have free access to knowledge. Especially when the knowledge created is a result of communal effort as is the case for all IDRC projects.

References

1. 1. The International Development Research Centre Digital Library/ La Bibliothèque numérique du Centre de recherches pour le développement international. URL : http://idl-bnc.idrc.ca
2. WHALEN, M. *IDRC Digital Library Policies and Governance*. 2007. URL: http://idl-bnc.idrc.ca/dspace/handle/123456789/35334
3. CHAN, L.; KIRSOP, B.; ARUNACHALAM, S. Open Access Archiving: the fast track to building research capacity in developing countries,' *SciDevNet*, February 11, 2005. URL: http://www.scidev.net/en/features/open-access-archiving-the-fast-track-to-building-r.html
4. Social Analysis Systems2 (SAS2)/ Les Systèmes d'analyse sociale2 (SAS2). URL: http://idl-bnc.idrc.ca/dspace/handle/123456789/34882
5. IDRIS+. URL: http://idris.idrc.ca/app/Search