Proceedings of the Second ICSU/UNESCO International Conference on
Electronic Publishing in Science
held in association with CODATA, IFLA and ICSTI at UNESCO House,
Paris 20–23 February 2001

The SciELO model for electronic publishing and measuring of usage and impact of Latin American and Caribbean scientific journals

Abel L. Packer

SciELO, FAPESP/BIREME/PAHO/WHO

Abstract. After three years of continued operation, the Scientific Electronic Library Online (SciELO) Model (*http://www.scielo.org/*) is progressively consolidating as an adequate, advanced and long-term solution to move the Latin American and Caribbean (LA&C) scientific communication to the era of Internet electronic publishing. By the end of 2000, SciELO operates online the full text of 85 journal titles, organized by means of national collections of selected journal titles from Brazil, Chile and Costa Rica and an international collection of the best public health related journals from Ibero-America.

1. Introduction

The Scientific Electronic Library Online (SciELO) is emerging as a regional model for the electronic publishing of scientific journals, intended to cover primarily publications from Latin America, Caribbean and Spain. In the long term, the library is devised as an instance for the measurement of usage and impact of the scientific journals.

It operates currently through the www.scielo.org Internet Web portal, that references the decentralized national collections of selected scientific journals organized by country of publication and regional collections by organized by subject areas.

SciELO was launched in 1997 as cooperative project among the State of São Paulo Science Foundation (FAPESP), the Latin American and Caribbean Center on Health Sciences Information (BIREME), a center of the Pan American Health Organization (PAHO/WHO) and a selected group of Brazilian scientific editors. FAPESP www.fapesp.br is a very dynamic agency that promotes scientific research in São Paulo and has a program that supports scientific journals. Its fundamental motivation toward SciELO is the production of indicators of usage and impact of Brazilian scientific journals. BIREME (*http://www.bireme.br/*) is a specialized center on the promotion of technical cooperation in health sciences information through out Latin America and the Caribbean (LA&C). BIREME's main objective is to strengthen the flow of health sciences information and its current operational strategy is the construction of the Virtual Health Library (VHL) as a network of information sources in the Internet. A specific product of the VHL is the LILACS (LA&C health sciences literature) bibliographic database that indexes the regional literature in health sciences. Its fundamental motivation on SciELO is to move the best journals indexed by LILACS to electronic format. The partnership between FAPESP and BIREME involves all subject areas.

By the end of 2000, SciELO operates regularly collections of the best journal titles published in Chile and Brazil in all scientific areas, and it is being implemented in Cuba and Costa Rica as well but covering health sciences related journals. A separated collection of the best public health journals from Latin America covers titles from Mexico, Brazil, Spain and the Bulletin of the Pan American Health Organization. In total, in the beginning of 2001, SciELO operates online 85 titles and about 12.000 articles. The number of online access to all these collections is increasing constantly.

The strength of the SciELO model resides in the fact that

Paper number EPS02016.

The online version of the paper was published 20 June 2001.
DOI: 10.1244/02.eps2/016.2.00

in addition to promote the transition of paper based journals towards the Internet electronic media, it intends to address and contribute to overcome the traditional problems inherent the publications running out of the mainstream scientific communication expressed by the titles indexed in the major international bibliographic databases, such as those produced and commercialized by the Institute of Scientific Information Inc. data bases, the MEDLINE produced by the National Library of Medicine (NLM) of the United States National Institute of Health (NIH), the United States IEEE data bases, the American Phycological Association database, etc.

2. The Traditional Vicious Circle That Affects Developing Countries Scientific Journals

The problems that traditionally affect the non-mainstream journals, specially the publications from developing countries, are concerned with the quality and more specifically with the perception of the quality of the scientific communication they accomplish.

There is no local systematic quality control of scientific communication carried out by developing countries based on time series of bibliometric and informetric indicators. In consequence, the measurement of developing countries scientific production relies most of the times on indicators produced by developed countries information products and services, specially the ISI Journal Citation Reports (ISI JCR). This situation represents a tremendous limitation due to the fact that most of the scientific journals published by developing counties are not indexed and not measured systematically in terms of usage and impact by the traditional developed countries indexes. For example, ISI JCR 1999 Social Science Edition indexes 8 titles from LA&C and ISI JCR 1999 Science Edition indexes 43 titles. NLM MEDLINE indexes 43 titles from a total of more then 600 titles indexed by LILACS.

The limited number of journals indexed in the major developed countries indexes has historically limited the visibility and accessibility of a great volume of developing countries publications. This phenomena has been a constant on the analysis and discussions of the problems faced by LA&C scientific journals and was identified by *Gibbs* [1995] as "Lost science in the Third World".

Without a systematic quality control, developing countries journals have been considered in many circles as second class in terms of quality. Worst than that, even the titles from developing countries indexed as main stream journals are frequently not perceived as such due to the fact that they come from developing countries. This perception permeates the academic community and the agencies that support science research in both developed and developing countries. In consequence, the stigma or the myth of second class often is assigned to developing countries journals, which, in the view of many, can only be minimized when the title is indexed in the main stream indexes. The practical consequence is to reinforce authors and institutions preference to publish on journals from developed countries because they will better rewarded. So, under this stigma, local journals are relegated or viewed as simply recipients of manuscripts not accepted outside or with a high probability to be rejected due a variety of reasons, including the authors perceived value of the manuscript, the subject focussed on local problems or interests, language, etc.

This is the context that feeds the circle vicious that has affected the evolution of the developing countries journals. And there is no perspective to break this vicious circle if developing countries journals continue to be judged by the unique criteria of being or not being indexed by developed countries bibliographic databases, in particular, by the indicators supplied by the ISI JCR.

The SciELO Model was conceived and it is being developed to break this vicious circle by promoting a context where the publishing of non-mainstream journals can have a positive feedback with their potential or target community of readers and authors and be progressively judged by its real dimension and value. In this line, the SciELO Model intends to provide their journal collections with advanced mechanisms of indexing and measurement that complements ISI JCR and other developed countries information products and services.

3. The Conception and Evolution of the SciELO Model

The SciELO Model is based on the assumption that the universe of scientific journals published by developing countries varies largely in terms of quality, including the fact that there are high quality titles not indexed by developed countries bibliographic databases due to a variety of reasons. In addition, the contents communicated by developing countries journals are in a significant extent related to local problems and therefore the promotion of their credibility and wide dissemination is crucial for the usage of scientific information in the social and economic development process. Therefore, a strategic objective of SciELO is to conform and strengthen the regional or peripheral scientific communication envisaging and promoting its integration in the global information flow Internet is creating. An expected result of this strategy is to foster the development of science in and for developing countries as a consequence of the wide dissemination of the local scientific research results.

In practical terms, the SciELO Model promotes the operation of a network of decentralized collections of selected titles of journals that publishes original articles that communicates results of original scientific research and other original works, case reports, technical reports, reviews, and communications related to scientific research. SciELO sites are organized to increase the visibility, accessibility and usage of individual titles. In addition, SciELO encompass a series of measurements of usage and impact of journal titles envisaging to value and monitor the quality of the journals titles and the contents they communicate.

For the purpose of its development and dissemination, the SciELO Model evolved according three components: SciELO Methodology, SciELO Site and SciELO Network.

The SciELO Methodology comprises standards, guidelines and software for the electronic publishing of integrated collections of scientific journals. It makes intensive usage of Internet related technologies to provide users with the capability to browse and search texts in the scope of a collection of titles as well as of individual titles, look up at table of contents of individual issues, display and print abstracts and full texts, which are presented in HTML format and optionally PDF format. The methodology comprises also a set of selection criteria and guidelines for including and maintaining titles in a collection as well as an integrated set of tools to measure usage and impact of the journal titles texts.

The texts are marked up using SGML based international standards so each text is accompanied by its metadata, which can be used to exchange records with different bibliographic data bases as well as to establish hyperlinks inside a SciELO collection or with external information sources. A practical example, is the interchange of records and the automatic linking between LILACS and MEDLINE bibliographic records and SciELO articles.

The methodology requires constant improvements in order to increase its efficiency, answer new needs and follow the advancements in electronic publishing. A key characteristic of SciELO Methodology is combining the compliance with international standards and practices and the adaptability to the prevalent conditions of developing countries in terms of infrastructure of information technology, economic and human resources. The first version of the methodology was developed during the pilot phase of the project, from March 1997 to May 1998, by operating a selected group of 10 Brazilian titles. A second version was launched in 1999 and the third version fully based in XML is expected to be released by the end of 2002.

The second component of the model is the SciELO Site, which comprises all issues related to the actual production and operation of a collection of electronic journals in Internet, according to the SciELO Methodology. A SciELO Site requires an established organization to deal with its daily management and operation, including text conversion, editing, markup, storage, publishing, exchanging metadata and linking with external databases, producing bibliographic indicators and reports of usage and impact.

SciELO Sites are planned to be operational at national and regional levels. The first SciELO Site was **SciELO Brazil** (*http://www.scielo.br*), which started its regular operation in June 1998 with 10 titles, reaching 54 titles by the end of 2000 covering all scientific areas. SciELO Brazil is expected to reach around 80 titles by the end of 2002. The SciELO Brazil, which was created and developed by the partnership between BIREME, FAPESP and a group of Brazilian scientific editors was the origin of the SciELO Model. SciELO Brazil is financed by FAPESP and operated by BIREME, which also coordinates the dissemination of the model.

The success of SciELO Brazil was followed by **SciELO Chile** (*http://www.scielo.cl*), which operates regularly since the end of 1999 under the coordination of the Chilean National Council of Science and Technology (CONICyT). It started with 5 journal titles and reached 20 titles by the end of 2000. SciELO Chile is expected to reach around 35 titles by end of 2002 covering all scientific areas.

SciELO Costa Rica (*http://www.scielo.sa.cr*) is operating in a pilot way under the coordination of the Library of the Social Security Institute of Costa Rica (BINASS). It is expected to start its regular operation by April 2001with 5 titles related to health sciences.

SciELO Cuba (*http://www.scielo.sld.cu*) is also expected to start its regular operation by April 2001 with 5 titles related to health sciences. SciELO Cuba is coordinated by Infomed unit of the National Center of Medical Sciences Information of the Ministry of Health of Cuba (CNCM).

In next two years, SciELO Model is expected to be adopted by at least three news countries from LA&C.

BIREME and the Instituto de Salud Carlos III from Spain are working on a cooperative project to implement SciELO Spain covering health sciences journals published in Spain. The participation of Spain in SciELO will increase the availability of scientific information in Spanish language for LA&C countries.

At regional level, **SciELO Public Health** (*http://www.scielosp.org*) includes 5 titles from Brazil, Mexico, Spain and the Bulletin of the Pan American Health Organization. The Bulletin of WHO is expected to be added into the collection by April 2001.

SciELO Sites, by putting together selected collection of journals titles, they maximize the development of individual titles in several dimensions. First, it makes feasible the electronic publishing of journals in an advanced and compatible way, which would be impossible or would take long time and would cost much more if published individually. Most of the editors and publishers of LA&C countries journal do not have economic and technological conditions to move their journal to electronic format in a sustainable way. Second, it provides the collection and each individual title higher visibility and accessibility when compared to paper based distribution as well as to electronic publication in scattered sites in Internet. The collection per se stimulates the browsing, navigation and cross-links among the different title texts. SciELO also increases incoming links from external sources. The library adds value to the time of the users by maximizing the relationship between recall and precision in the searching process when compared to searching over Internet non-compatible sites. Third, SciELO succeeds in quality control it will be progressively recognized as a context that privileges quality and therefore become a reference for authors, readers, editors, publishers, research-related agencies, etc. which will increase the credibility of SciELO journals.

In order to guarantee the continued search for excellence, the operation of a SciELO Site is recommended to be assisted by an advisory committee responsible for the application of the selection criteria to include news titles into the collection and to monitor the performance of individual titles regarding the minimum requirements to remain in the collection.

The SciELO Site produces automatically a set of unique numeric indicators of usage and impact, which provide the advisory committee, editors, publishers and the agencies that support scientific communication empirical data to mon-

itor the performance and to identify the weakness and flaws that affects the collection as a whole and each individual title. Examples of indicators are: Web pages visited, articles visited by journal, issues visited, citing and cited journals, etc. As these indicators acquires critical mass, countries using SciELO will have systematic and updated series of bibliometric and informetrics indicators to evaluate and improve their scientific communication. In other words, the SciELO Site is expected to create an environment that induces the constant monitoring and enhancement of the quality of journals.

Finally, the systematic evaluation of SciELO site helps developing countries to build and improve local capacity on scientific writing, editing and publishing.

The third component of the SciELO Model is the development and operation of the network of SciELO sites. Currently the network is at its initial stage. Its current public expression is the Web Portal www.scielo.org that links to the different decentralized SciELO Sites described before.

The dissemination and development of the SciELO network is promoted primarily by BIREME as part of its major program to strength the flow of health scientific and technical information among the countries of LA&C through the as Virtual Health Library (VHL) technical cooperation framework (*http://www.bireme.br*). The VHL is conceived as a virtual space in Internet to be co-operating by information producers, intermediaries and user envisaging the equity access to health information. Inside the VHL, SciELO texts are linked with other information sources such as the LILACS (Latin American and Caribbean Health Sciences) and MEDLINE databases, directories of authors, institutions, etc. These links increases the number of access points and therefore the public exposition of the texts. Inside the VHL, users can search for literature and have access to the original documents be they on paper or electronic media. For example, searching LILACS database (*http://www.bireme.br/bvs/I/ibd.htm*) for "chagas disease Cochabamba," the VHL returns references which points to electronic texts when available and to the inter-library loan services for location of the document in paper. Selecting the link to SciELO in the first reference, the full text article is displayed with links to authors resumes and from the bibliographic references at the bottom of the article.

The dissemination of the model is carried out through the promotion of partnerships with national institutions supporting scientific research and communication. SciELO Model is designed to cover all scientific subject areas and, whenever possible, the implementation of SciELO at national level is coordinated by the national agency or agencies responsible for supporting scientific research and communication. When this is not possible in the short term, priority is then moved to the publishing of health science journals.

Once a national institution assumes the responsibility to operate the SciELO Site, BIREME transfers free of charge all the SciELO Methodology to operate a collection of journals. In addition, the national institutions involved in the operation of SciELO are also expected to cooperate on the development and further dissemination of the Model at national and regional level. It takes about 1 year to have a SciELO site operating in a regular basis.

The SciELO Network founds its operation on decentralized SciELO sites that are operated at national level. The decentralization offers long term advantages because it promotes the development of national capacity to run local information flow as part of the regional and global information flow. This approach requires that local needs and idiosyncrasies be taken into account, including language priorities, national agendas for scientific research and communication, financing policies, etc. The decentralization may take more time than a centralized approach but in the long term it assures more sustainability.

The decentralization also poses challenges on the harmonization of policies and criteria as well as on the interoperability of the collections.

The quality control for both inclusion and permanence of titles in SciELO collections will demand constant discussion and evaluation in order to balance the search for excellence and the promotion of local scientific communication. It will be necessary the continued development of a common set of indicators that take into consideration local and regional conditions.

The inter-operability of the collections will demand constant technological improvement, including the provision of adequate connectivity, searching and mining databases of texts in different languages, set up of dynamic exchange of records and hyperlinks with external databases, etc. In addition to facilitate user access to information, an efficient inter-operability among SciELO Sites will stimulate the cross citing among local and regional authors.

Finally, as SciELO become fully operational in several countries in the coming years, it will be possible to work on the elaboration of regional and sub-regional bibliometric indicators, which will complement the statistics currently available for the mainstream journals.

References

W. Wayt Gibbs, Lost science in the Third World. *Science, 2(273),* 1995, pp. 76–83.

Abel L. Packer, SciELO, Rua Botucatu 862 São Paulo, Brazil, e-mail: abel@brm.bireme.br

(Received 25 June 2001)