	<p style="text-align: right;">Date : 14/08/2007</p> <p><b>Between the devil and the deep blue sea: how the publishing department of an international organization tries to address the challenges of information delivery in a changing environment.</b></p> <p>Valentina Kalk Rights Manager The World Bank Washington, DC USA</p>
<p><b>Meeting:</b></p>	<p><b>119 Government Information and Official Publications</b></p>
<p><b>Simultaneous Interpretation:</b></p>	<p>Yes</p>
<p style="text-align: center;"><b>WORLD LIBRARY AND INFORMATION CONGRESS: 73RD IFLA GENERAL CONFERENCE AND COUNCIL</b> 19-23 August 2007, Durban, South Africa <a href="http://www.ifla.org/iv/ifla73/index.htm">http://www.ifla.org/iv/ifla73/index.htm</a></p>	

Abstract

*The World Bank, one of the largest development agencies in the world, publishes about 150 publications in print every year. Its publishing department (“The Office of the Publisher”) launched between 2001 and 2007 four online, subscription-based services called [World Bank Online Resources](#), consisting of two statistical databases, one portal of analysis of global economic conditions, and a collection of about 4,000 e-books and working papers.*

*The experience with online collections has been overall very positive so far. However, a number of questions and challenges present themselves as technology evolves and demand grows at a time of technology renaissance and a second internet boom. This paper discusses the experience of the Bank as regards the following: changes needed in internal workflows and organizational habits; the tension between mission and cost-recovery as regards free electronic access to the new online services; the challenges posed by technology developments, which create the need to upgrade and invest in the systems fairly frequently and extensively; the rise of new actors such as Google; the increasing demand for online content and tools to make content more user-friendly on the side of the users’ community – librarians, intermediaries and end-users; which practices work and which do not in developing vs. developed countries, and how to ensure that the end-users in developing countries can get access to a wealth of information they need.*

It all started with Netlibrary. In the year 2000, The World Bank's Office of the Publisher (which I will call World Bank Publications from now on, and will refer to as "we" every now and then) signed a contract with Netlibrary which granted Netlibrary the right to create electronic versions of our books and sell them to the library market worldwide. Over 1000 books in print were sent to a warehouse somewhere in the Midwest to be scanned and digitized and more (including full cataloguing, assigning Dewey and Library of Congress numbers, and creating Marc records). It should be noted that, at that time, World Bank Publications did not have an e-book or digital conversion workflow: production was primarily aimed at print products, and a few CD-ROMs. Between the year 2000 and 2001, new e-content aggregators started activities in the USA (for example ebrary and Questia), bringing new opportunities for making World Bank content broadly available digitally in libraries, as well as a new revenue stream. At this point, I should say a couple of words regarding the business model of World Bank publications: our total budget is a combination of Bank funds and revenues from sales of our publications. Although we are not expected to be fully self-sustainable, at least at this time, we do have to sell in order to cover all of our costs. As our primary goal is dissemination, we have established a pricing structure that allows us to sell publications in low and lower income countries at 75% off the list price, and at 35% off in upper middle income countries. In addition, we have approximately 200 depository libraries worldwide that receive free copies of our publications and / or free access to our Online Resources. But let's go back to the period 2000-2001.

It was at that time that we started asking ourselves a number of questions: should we try to work with as many aggregators as possible? Should we consider developing an electronic collection ourselves? The content aggregator boom had shown that e-books and electronic collections were indeed in demand; that although this was still very much a US-Western Europe phenomenon, it gave us high hopes for increased access to World Bank knowledge in academic environments in developing countries, at least those with decent to good internet connectivity, by solving the well-known problems of distribution and physical infrastructure in those countries where buying a book published abroad often takes a very long time, with high mark-ups by local distributors, and outrageously high transportation costs. It was then, I guess, that we decided to develop World Bank Publications' own, subscription-based e-book collection. We started in a rather casual way by buying back in bulk the Bank's files that had been digitized by Netlibrary. Importantly, though, the "what to do with 1000 xhtml files" was a question we were still not addressing. In the meantime, things were not so rosy for Netlibrary, which filed for bankruptcy in November 2001.

Around the same time, we launched the online versions of two World Bank statistical databases: World Development Indicators and Global Development Finance. These two databases are compiled by the World Bank's Development Data Group and were previously available only in book and CD-ROM. The success of these databases proved that an online collection of text-based publications would be welcome by the library community and their users.

Fast forward in time, in 2002 Ingenta was selected as technology partner for the development of the World Bank e-Library, and a contract was signed in January 2003. In

view of a relatively ambitious launch plan (May 2003), we had to put ourselves at work to make this happen.

Three dramas soon started:

- 1) The metadata drama
- 2) The file drama
- 3) The internal turf drama

The metadata drama was the sudden realization that we had 1000+ files from Netlibrary, but apart from the title names, we missed (or did not keep in one place) rather obvious information such as other basic bibliographic data, abstracts, keywords, and other classification needed to power the e-Library's search engine. This metadata construction work was immense, but worthwhile: not only the information we collected and created was pivotal for the e-Library, but "The" metadata sheet we have scrupulously maintained, normalized, updated and expanded since has become the major reference point and one of the first starting points for all subsequent e-publishing initiatives we have been involved in.

The file drama was generated by an actually very simple question: how to ensure that for all new and forthcoming publications, an e-book file would be created. As paradoxical as it may sound, this was not an easy task. First, it was not obvious to decide who should create an e-book file. Our production is decentralized: we use different editors, typesetters and printers for different books. The first idea was that the printer would also provide an e-book file. It became clear from the beginning that this would cause inconsistency in file provision. We also needed to agree on file standards, and ensure that one file would work for our own e-Library and for all content aggregators. It took a while to complete this very simple step, because of our lack of experience but primarily because our production team's work was all geared to print book publishing. This brings me to the internal turf drama. E- and print production started to be perceived as conflicting priorities: the e-Library team, for every new book published, wanted the files by... yesterday. The production team wanted the physical book to be published first, and all the rest was an accessory. A culture change was necessary and it took a while for everybody to see that all that was actually needed was simply an additional step in our workflow, as well as an expert vendor who could do file conversion for us in a quick, consistent and efficient way.

The results were slow to come, but the effects have been remarkable: e-book files started to flow easily from the production department to the e-Library team, who in turn made sure that they would be associated with the relevant metadata and uploaded on the e-Library as well as distributed to all other content aggregators we do business with. In this process, we also became the main providers of e-book files to our authors and, more importantly, to the Bank's archives. This leads me to describe another major issue we faced, and are still facing: the tension between mission and money. The World Bank prides itself to be a "knowledge bank", and the predominant view within our institution is that the knowledge we create should be public good. While it is perceived as somewhat justified to sell books (although there are different views on this, too), because of the

costs of the physical production and distribution, the dominant view is that online distribution costs zero - despite the fact that is very easy to demonstrate that this is not entirely true - and for this reason we should not charge for online services. In 2004, in an attempt to compromise among different views, my office drafted an electronic access policy which states that it is the authors' discretion to decide whether their e-book should only become available through our commercial services, or also through the Bank's corporate web site (intranet and extranet), free of charge. The vast majority of the authors like the idea of having both. As a matter of fact, this reflects the fact that increasingly, World Bank staff members use and share with others electronic resources and publications from the Bank's web site, as the graph below show. A few ask to first make the electronic version of their work available from the (commercial) e-Library, and after a few months, free of charge, through the Bank's corporate web site. And very, very few (not more than 2 or 3 per year; we publish as many as 150 books/year) decide not to make their work available free of charge.

So we could say that my department is trying to sell what is universally available free of charge. Why, then, do hundreds of institutions still subscribe to our priced services?

One answer is that content seems to no longer be key – as long as it's good, of course! Our purpose is not to sell access to the repository of all Bank's disclosed content. We want to operate some form of editorial selection (in our case, the best of the research carried out at the Bank: all formal publications, the best working papers, and the top databases). Decision-makers as the librarians, as well as their user community, clearly want a number of services associated with the e-content: from cataloguing information (MARC records and the like) to reference linking services; from permanent identifiers such as DOI to automatic alerting systems; from a well-conceived, simple search engine that is helpful to readers rather than confusing, and delivers relevant results, to citation export tools. Last but not least, it is also important to offer good, competent and responsive customer service.

But there is another, perhaps more interesting answer to the question why institutions still subscribe to our priced services. In a 2007 survey of the "Doing Business" database, currently a free resource ([www.doingbusiness.org](http://www.doingbusiness.org)):

- 45% of the librarians interviewed said they were aware of the database
- 61% rated the database unique, 60% relevant
- 3 out of 4 indicated an interest to subscribe if it becomes a subscription product
- But only 11% of those aware of Doing Business currently provide links for their patrons

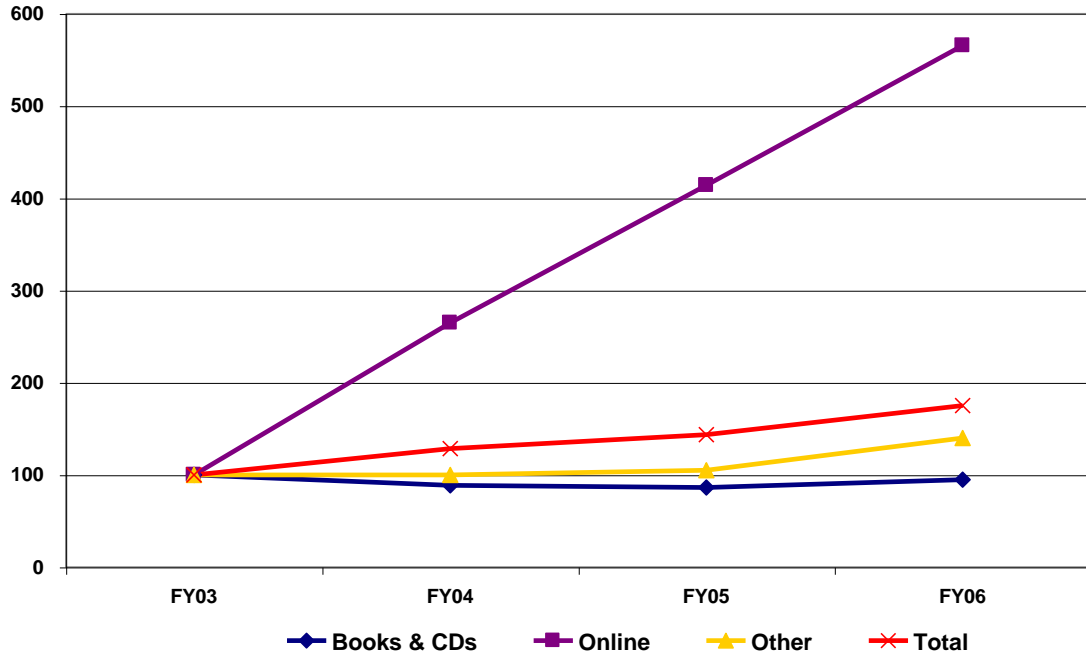
This, in turn, creates another challenge. How to keep platforms, systems and files up to speed with technology developments so that the increasingly sophisticated demand from clients is met? File-wise, it is a success to achieve an optimal workflow and file consistency, but which one is the right format? PDF-based e-books created from application files are in certain respects *passé*, but transitioning to an xml-based workflow creates issues not unlike the ones I described above. Converting a backlist of 1500+ e-books into proper, tagged xml is extremely expensive and time-consuming (e.g. quality

check with vendors); having “hybrid” platforms with a combination of PDF for the backlist and xml for the front list is less than ideal. The other dilemma, especially for a book publisher, is reference linking. Most CrossRef members are journal publishers. While STM books often cite journals, the opposite is more rare. This means that most books are cited in other books. To create a fertile web environment where cross-referencing among books really happens, most of the STM and social science book publishers should be CrossRef members. But we are far from this. We were hoping that partners such as Google (Google Book Program) or Microsoft (Live Book Search) could be instrumental in helping with the cross-referencing process. This has not happened so far; my personal impression is that their referencing technologies (e.g. Google Scholar) are still far from being reliable, so they are waiting. But in my opinion, this is going to change very soon. There is also a big potential for Google and Microsoft to partner with content aggregators and technology providers for publishers, but this is still very much at the beginning.

In sum, there are a lot of open questions about how to evolve our systems and platforms in order to be as user-friendly and up-to-date as possible. An important one is how to find the right compromise between advanced technologies and accessibility in those developing countries where connectivity might not be as good or as reliable as elsewhere. This includes choices such as avoiding plug-ins, proprietary formats, or “heavy” files. On the business side, as mentioned before, we do offer discounts in developing countries; still, organizations like INASP (International Network for the Availability of Scientific Publications) are instrumental in reaching local libraries, institutions and consortia, helping them sorting out the complexities of license agreements, and negotiating with publishers nationwide agreements. We currently have over 40% of authorized users coming from developing countries, which in our mind shows that a combination of accessible technology and affordable prices seem to work.

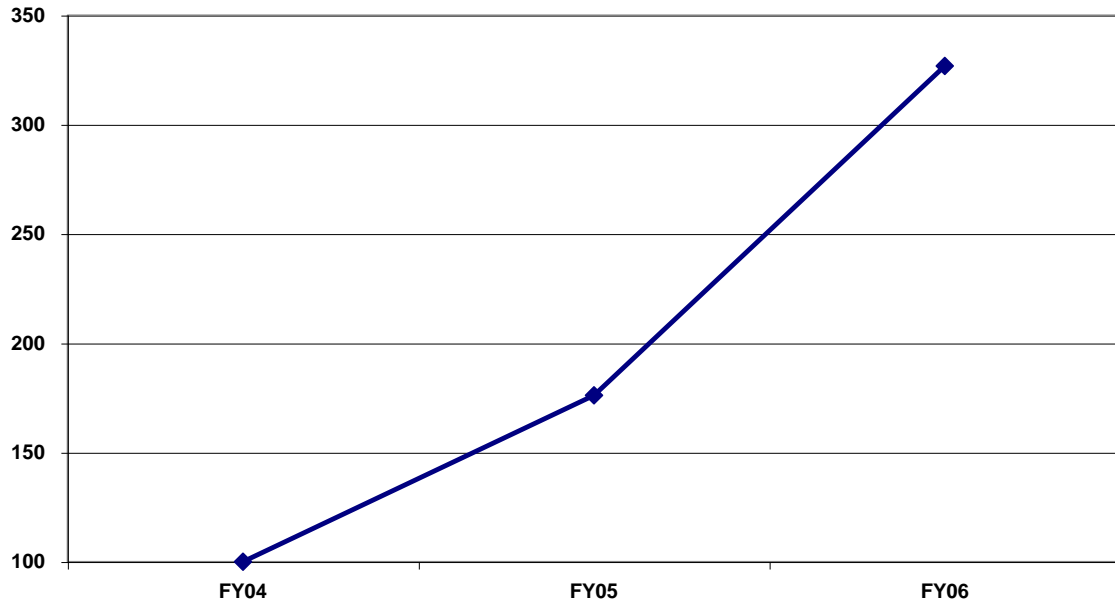
# Online business has taken off

Revenues, by Source  
FY03 = 100

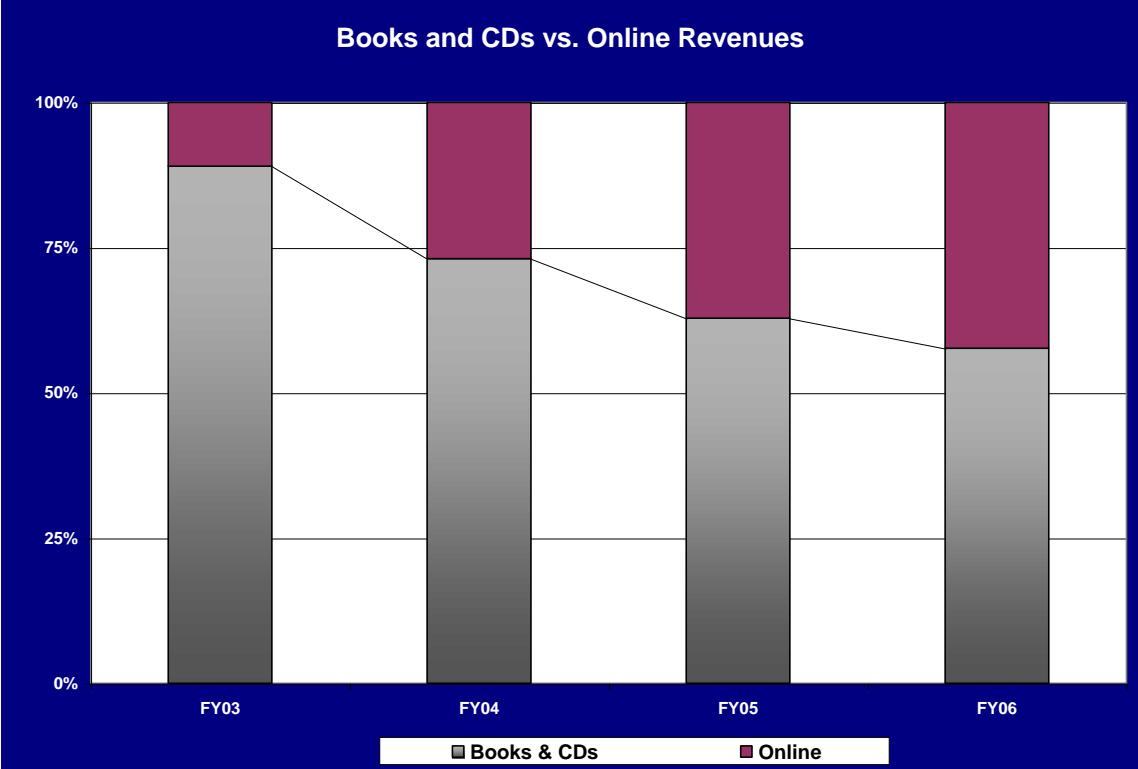


# And the Number of FTEs Subscribed to Online Resources was about 9 million at the End of FY06

World Bank Online Resources  
Full-Time Equivalent Authorized Users  
FY04 = 100



**Revenues from Books and CDs Still Exceeded Revenues from Online until FY06; now they are approx. 50%-50%**





## 98% of Staff Used Bank Publications Most staff use both print and online formats

In the past year, have you personally used a Bank publication, either in print or online?

