

The Role of Metadata in the Discovery, Selection and Acquisition of e-Books

Case study: Bangor University, 3rd October 2011

Mieko Yamaguchi

Head of Collection Management

Bangor University Library

Phone: +44 (0)1248 382970

m.yamaguchi@bangor.ac.uk

What was/is the problem which PDA was implemented to address (i.e. what does PDA solve?)

In Bangor most book selection is done by academics rather than by library staff. Different schools have different policies; some buy core texts while others do not. Book budget is devolved to academic schools and the library has little influence on what is added to its collection. We sometimes see reading lists with books which we don't even have in the library but all we can do is to alert the schools concerned.

We wanted to give end users the power of selection and have been looking at PDA for a while. We had built up a modest collection of e-books over the years but there wasn't a critical mass of e-books for users to look specifically for e-books when they search the library catalogue. Although some e-books have been heavily used, a number of them have never been "borrowed" since we purchased them. Overall our policy/preference is for 'e' rather than print so PDA for e-books fitted into that well. We hoped PDA would help to build up a critical mass of e-books that our users actually want to use.

Like many libraries we are short of space. We have replaced many print journals with e-journals but it would be impossible to replace much of our print book collection with electronic equivalent. Some academics are anti-ebooks but we try to convince them that e-books will help to safeguard the future of little used print material which they are adamant we should retain.

At the end of the 2009/10 financial year we had some unspent money in our central budget which we would have lost if we did not spend it. We sought permission to make a prepayment of £60k for a PDA pilot during the 2010/11 academic year.

What benefit does it have (or might potentially have) over other procurement solutions

The benefits to the end user

PDA expands the choice of what is available and makes material available instantly when the user wants it even if it's out of hours. Many of our students live in rural areas and a substantial number of them are mature students who find it difficult to come to the library to access print material needed for their study. Students can access e-books when it suits them and can read around the subject through PDA. Students should be encouraged to decide what to read from a wider selection of material than a library can possibly hold rather than just from a prescribed list of what they need to read to pass exams or get degrees.

The benefits to the library/institution

We are better able to meet student demand and hope that PDA helps to enhance student experience. Demand can be patchy and we might find 30 copies of a recommended book sitting on the shelf most of the time yet there aren't enough copies to meet peak demand. PDA and e-books in general help us to meet that kind of demand better. We have already found that students prefer to use e-books although print copies are also available in the library. We hope to find more evidence like this and inform the academics and to convince them and the university that e-books and course reading material in electronic format are the way to go.

PDA makes a better use of our investment on our Encore discovery tool. What's the point in having a discovery tool if all that students need to find is books on reading lists or titles pre-selected by teaching staff? We hope that through PDA our collection will be a better-balanced one meeting the need of our users.

The role of metadata in the discovery, selection and acquisition of e-books

We use EBL exclusively for PDA and load their MARC records to our catalogue. On the whole the quality of records is good. Some have non-authorized subject headings which are rejected but subject headings are less important in keyword searches and searches via discovery tools. We no longer offer subject searches (using LCSH) in our catalogue in any case as most users don't understand what they are. We were going to upgrade MARC records for purchased e-books with OCLC records but on evaluation decided this might not be necessary. Judging by the level of usage, our students seem to have little trouble finding e-books in the catalogue and the discovery service.

The default search of our catalogue is a Quick Search which searches the Encore discovery service. We have deliberately decided not to offer Encore as a separate search interface as many libraries do. An Encore search is a keyword search with the result sorted by relevancy by default and with a preference to electronic format.

Use of metadata for profiling

Selecting or de-selecting by subject would have been very difficult and we might have excluded some useful material. So profiling the PDA e-book collection was by publisher (we selected the main academic publishers) then language (English and Welsh) then price – the limit was \$150 last year but we raised it to \$200 this year to ensure quality material is included on advice from EBL.

We excluded e-book titles we already owned, e.g. NetLibrary and Springer subject collections. This was achieved by giving EBL the ISBNs of e-books already in our catalogue so they could exclude them from the MARC files supplied to us. On advice from EBL we deleted all PDA MARC records from 2010/11 except for those which have been purchased, and loaded 80,000+ new MARC records. This ensured duplicate records were not loaded inadvertently. We have been more careful with duplicate records this year (e.g. where the library holds the title in print). It is possible to identify them because MARC records from EBL include print ISBNs.

Impact of Discovery services

As already explained Encore search is a keyword search. There is a pull down menu on our main search screen for other types of searches but the main focus is on keyword—most users stick with the Encore 'quick search'. Through a 'quick search' users can search words in author, title, subject, chapter heading, content or descriptive note, which are included in most EBL MARC records, and pre-

limit their search to e-books only or filter the initial result list to e-books only. Encore gives preference to electronic format but this does not always work due to the quality of metadata.

The accuracy of metadata is important when similar titles, different editions of the same work or the same work in print and electronic format are held by the library. We expect different versions of the same work to appear together in the catalogue so users can decide which edition or format to select. However, if a subtitle is missing or diacritics are not used correctly, users may miss that there is an e-book version of the work in addition to print copies they see in the catalogue. The role of metadata is suddenly very important in a discovery service where resources in different format are searched together.

How does the PDA scheme work?

We loaded 40,000 records from EBL last year (October 2010). This year we've just loaded a little over 80,000 records. The collection was profiled (see above). We did not tell our users about PDA, so anything they found was found through normal catalogue searches. For the first 3 STLs (short term loans) the library pays a percentage of the purchase price. The 4th access to any 'unowned' title triggers an auto-purchase.

We tried to put in place as few barriers as possible. There was no mediation by library staff last year but we have introduced a limited amount of mediation this year based on the STL charge. We do limit the number of STLs a user can have in a single day to 3 per person

Last year we started in October and ran out of money we had set aside (£60k) by February. In order to control demand we suppressed the records from the catalogue for those e-books that we had not already purchased. This caused a very sudden drop in demand so it's clear that the primary discovery route is the library catalogue. Users could have gone to the platform (a link to the EBL platform is on the library's e-resource page) and seen 'suppressed' records and initiated a PDA. A number of users did. If a user complained that they had seen something before which they could no longer see in the catalogue we made the title visible.

This year we are collecting information about the users. The first time they try to access any EBL e-book, users are asked for their school (e.g. social science) and whether they are undergraduate, postgraduate or staff. This information is associated with the patron's Shibboleth ID so we can track what group of users are using e-books. We have also added a few new features to try and make the money (£105k for 2011/12) last longer. The first 5 minutes of online reading is free, then a message appears asking 'do you wish to continue?' and informing that the library will be charged if they do. If the user confirms it becomes a short term loan and the library pays. Unfortunately this seems to upset some users but we hope most people will understand why we display this information. We also started showing the price of the book which we don't already own. There was a suggestion at a London e-book focus group that users actually prefer to see this information.

Role of evaluative metadata

We are uncertain whether students use the catalogue metadata to evaluate any e-book before viewing it. The library is using metadata for evaluation, as already explained, to pre-select the material available to the end users through PDA. The library catalogue is also linked to Google Preview so there is the option to 'look inside' and of course the first 5 minutes of online reading is free and the user is asked if they wish to continue. The library catalogue/Encore offers tagging, review and rating options – but it is still early days and these features are little used at present.

The outcome/lessons

We find it very encouraging that some lecturers are beginning to put e-books more consciously on their reading lists. We will engage more with academic staff to promote e-books in general and monitor usage so that core material may be purchased without waiting for 3 STLs plus a 4th access. This will save us money. If we are in any doubt whether certain recommended titles will be used or not, we can always decide not to purchase them upfront though it would cost a little more should we end up purchasing them through PDA in the end. We will also try and encourage academics to select any e-books not already in our PDA profile from EBL's public catalogue. In terms of user satisfaction we will see what students have to say in LibQUAL next year. We may also run focus groups.

We haven't seen any evidence of 'frivolous' purchases though we see some short term loans lasting 0 minute in usage reports. We would like to find out why and hope that telling the users the library will be charged if they continue might discourage 'frivolous' access.

Some comparison of e-books bought through PDA and e-books bought in other ways:

- We had bought 350 EBL e-books requested by academics before we first started the PDA pilot. These e-books were used 368 times in 4.5 months between October 2010 and February 2011.
- 450 e-books purchased via PDA during 4.5 months between October 2010 and February 2011 were used 3,012 times during this period.

- Average number of access of auto-purchased titles was 16.3 (including 3 STLs prior to purchase) between October 2010 and August 2011.
- Average number of access of already owned titles was 4.8 between October 2010 and August 2011.

- Average cost spent on e-books bought through PDA was £70.07 including 3 STLs prior to purchase. (£52.17 + 3 x £5.97)
- Average cost of EBL e-books purchased before PDA was £118.18.

- Average cost of PDA access per title was £8.94.
- Average cost of PDA access per use was £3.89.

We are still working through usage statistics to get more detailed information.

Have you achieved a sustainable model for PDA? What factors are important in terms of sustainability?

We had hoped demand might level off but that has not happened yet and looking at usage it may not happen for a long time. As explained already we are putting in additional measures to control expenditure but it seems unlikely we will find equilibrium in the near future.

When we sent our initial report to the Library Task Group (a group of senior academics and library staff to discuss strategic matters) with usage of PDA e-books for the first 4.5 months, their support was unanimous and we were asked to submit a bid to the University to continue the initiative. However, we have found the library would be able to fund it for another year from its own resources and withdrawn our bid as there were other projects for which we were seeking financial support.

We hope we will be able to collect enough evidence that students prefer e-books to print copies of course reading material. We also hope to demonstrate students' use of library resources is enhanced by PDA and that material obtained through PDA is not just to meet the demand of a few users but will continue to be used by many more readers. It is hoped that we will be able to convince the university that PDA enhances student experience to secure future funding and at the same time get the right balance of resource to demand through better profiling and parameterisation of PDA.

We heard some horror stories of libraries which publicised PDA and quickly ran out of money. We did not tell our users about PDA in the first year and imposed very few barriers. We have just started our second year of PDA and are still reviewing parameters. We are offering more titles this year by increasing the number of publishers as well as raising the maximum price. We have found that STL charges by some publishers are 15-20% of the list price rather than 10% which we believed was the norm. So we have decided to mediate any STL request costing more than \$20. There is no point in including any title if we are going to deny STL access. However, we may ask the requester if the title has been recommended for a module or if there may be others who are also interested in the book. Or we may decide it's a title we should have in the library anyway and decide to purchase it upfront. EBL is believed to be developing a way to allow libraries to flag certain titles for immediate purchase as soon as a request is made.

We have also started explaining to our users how PDA works (at least partially). We might even explain more fully about PDA. We were concerned that the system might be abused but I think our users understand resources are not unlimited and that the service should not be abused.

I hope we will develop a sustainable model by further fine-tuning of parameters and by managing user expectation through user education and more information. Of a small sample of recent auto-purchased titles (3 STLs plus 1 further access), 45% are titles we already hold in the library in print (some are older editions) and 25% are recommended readings. We hoped PDA would enhance our book collection by adding new titles but somewhat unexpectedly a large proportion of e-books acquired through PDA has turned out to be titles we already hold in the library in print – and in multiple copies and/or on reading lists. This is actually good news for the library and the university. Instead of buying multiple copies of recommended books in future we could buy just one or two copies in print plus an electronic version. The latter should satisfy the need of a majority of students who are already voting with their feet, or their laptops, that their preferred format is e-books. We hope the money saved may be used to continue PDA even on a reduced scale.