

The electronic jungle: Managing electronic resources

Alan Hopkinson
Head of Library Systems
Middlesex University

1. Introduction

What a jungle is the world of electronic journals! Very much a tropical jungle with fast growing trees which suddenly disappear only to be replaced by others which quickly grow to the same height. One stops growing and another suddenly appears by its side and outgrows the parent.

2. A case study

What is this strange world? I can best illustrate it with an example. While planning this seminar I thought it would be useful to show my colleagues the article on COBISS published in *Program* (*Program* 36(2) 2002 89-98). So I looked in our library catalogue for the journal *Program* (PPT 2) and from the catalogue record I found a link to the computer of EBSCO which I followed. This company is one of the companies that provide us with our hard copy journal subscriptions and who have committed themselves to providing access to electronic journals where publishers make their electronic journal available free of charge to those libraries that purchase the hard copy. So I pursued the link which I had found in the catalogue:

<http://www-uk.ebsco.com/online/direct.asp?JournalID=104929>

This took me to the EBSCO Host page for *Program* (which actually has a different URL, but their system knew to transfer me automatically):

<http://ejournals.ebsco.com/Journal.asp?JournalID=104929> (PPT 3)

This told me that this journal was offered by EBSCOhost as a remote title and linked to the publisher's site. This is in fact slightly out of date as they do not publish their own journal any more and the EBSCO website made no mention of this. I followed the link. Aslib's *Program* page (PPT 4) allowed me to see Vol.30 (1996) to vol.35 (2001). I could also click on current issue but that gave me vol.35 issue no.4. I actually had the library's hard copy of vol 36 no.2 in front of me, so I knew that vol.35, no.4 was not the current issue!

I knew that Aslib had ceased to publish their journals themselves and had sold out to Emerald who now publish the journal in the context of what they call Library Link. I could (as I had the journal in front of me) have read this in the journal and found the URL there. Instead I searched for a journal which I knew was also published by Emerald, Library Hitech News. I found this not in our catalogue but from our list of journals available in electronic format (PPT 5) and looked for the page of journals beginning with L (PPT 6). I then accessed the Emerald page for Library Hitech News (PPT 7). Having reached that page, as this was not the journal I really wanted, I went to the Emerald home page (PPT8), clicked on Emerald Journals and searched for *Program* (PPT9). One record was retrieved (PPT 10). I clicked on this and discovered that 23 issues of *Program* are available from Vol 31 (1) to Vol. 36 (3) (PPT 11)

I was at least able to find the article I wanted.

Another approach would have been to access *Program* in our web list. Let's do that now (12). You will see it is in our list but requires a password. You will see that the word Password conceals a clickable link. On clicking on it a list of passwords is revealed, sited on our intranet so the passwords

are revealed only to those accessing from a computer situated in Middlesex or by the direct dialup which is run for us by Easynet. In fact we have a page of passwords for different journals so I'm not going to display it now or else you'll all be able to go and consult these journals free of charge from the internet here!

So, let's make a note of the password and access the journal by clicking on its title. This brings me a different front page which tells me that current subscribers may access the articles contained in the 1979-1999 volumes and inviting one to apply for a password through an on-line form which actually turns out when you click on it to be no longer available (PPT 13). Not to worry. If I move that page down you can see we can access previous issues (PPT 14), and here we find again we are back at the page we saw before which allows us access to vol 30 (1996) to vol 35 (2001). (PPT 15)

So what can I do about it as the person responsible for enabling access to electronic journal files?

We have two sources of information, Aslib's own website which gives us access from 30-35 after which no further journals have been added, and Emerald's which starts at 31 and goes to the current. Because neither gives fuller coverage than the other, I decide that we have to have both. Therefore we must put both in the catalogue and both on the journals list. It seems to me that Aslib have made the journal freely available on the web as I did not have to enter the username and password at any stage. So I print out Aslib's page with the list with links to vols 30-35, write on it that this should be clarified in the catalogue as being only for those volumes, and then print out the Emerald page with the list and the URL and add a note suggesting that this should be added to the catalogue entry. I send this to our journals cataloguer. I also inform my assistant who deals with web access and supports me in enabling access to electronic resources and ask him to remove the link to the password and the password for *Program* from the special password page itself. I remind him that the existing link gives 30-35, this new one gives 31-36 and will presumably continue.

By the time I give this paper I hope these amendments will have been made and I can add them to this powerpoint demonstration.

How did this situation arise? Other users at Middlesex University could have found *Program* by going to the Emerald website. How did we get conflicting entries and not even have the latest access? This is an interesting situation but by no means unusual. In this case, Aslib journals were taken over by Emerald. In Middlesex University there is no subject librarian responsible for librarianship so there is no one who would need to be aware that the website had changed. In this case a professional association has sold its journal to a publisher and the old site is still there but it appears not to be updated. Interestingly it seems to be freely available from the association's website. We have a blanket subscription for Emerald management titles and although we have added many of these titles to our catalogue, we have not done it systematically. If we had checked each one we might have ignored this. We should have added it as an extra source of the journal since both sources provide one year which the other does not. Our database librarian tells me that even the contract with Emerald is unclear as it is not clear if it applies to new journals that they launch on their website.

Another way we find URLs that need updating is when we find that a URL does not work, often reported by a library user to a member of staff at the enquiry desk. In this case I myself found it but it was not that clear cut because it appeared that nothing had been published since last year. So it was not obvious to anyone that we were not getting the full run, except to anyone who had the journal to hand, or at least a later issue of the journal than the last in the digital library.

3. Where are electronic journals located?

Where are the journals?

It is important to have some idea of where the electronic journals are located. At least if you know when things go wrong you have a better idea as to what has happened and who to contact to get the matter put right.

1. Some are on publishers' own websites.
2. Others are on the websites of intermediaries. Some time ago there were two such intermediaries, Ingenta and Catchword. On Ingenta's website it says: 'Since its launch in May 1998, Ingenta has developed and grown to be the leading Web infomediary (not a word I knew before) empowering the exchange of academic and professional content online.' They offer to publishers the opportunity to allow free access to subscribers who may be a subscriber directly to the electronic version or may get it by virtue of subscribing to a hard copy. Or they offer a pay-as-you-use service to the publishers' material aimed at non-subscribers.
3. Additionally the intermediaries can provide links from their computer direct to the correct address in a publisher's own website.

Incidentally, because the intermediaries have thousands of articles on their databases they are beginning to offer searches across these journals, either through author and title or even through full text.

4. What has to be done?

4.1 Pointing the user to the electronic resource

Internet resources have been with us a number of years now, and MARC has had a field 856 in the catalogue for a number of years which can be used to link to an external resource. So the problems of links to these resources are well known. Thank goodness it is my brief in this paper only to talk about electronic journals, though e-journals are probably the reason why everyone has to use links now especially as some journals are not available in hard copy! So, now no library can avoid implementing links.

There is more than one way to allow users access to electronic journals and indeed we have seen in the case study that Middlesex uses two methods, access through the catalogue, and access through a special set of webpages. The second is fiddly and involves duplicate work. The cataloguers enter the URLs in the catalogue records and the library web master enters them in a list. We are considering a third possibility, a database which allows access to electronic resources. We would set this database up by copying records from the catalogue with a field 856 and importing them into another database. Alternatively we should consider thinking of having an extra index in our library catalogue based on materials with a field 856. One way or another, we should not have dual data entry to the extent we have at the moment.

The case study covered only one scenario, a URL changed when the publisher changed. In this case a new URL was set up but the old one remained. But there are many other scenarios. Publishers may move URLs because they reorganise the way they put the journals on their computers. They may change their host computers since they do not always host their own data. Journal suppliers may also do the same.

As an aside, sometimes journals have easy-to-follow URLs and sometimes they have very unmemorable strings. Of course the way the cataloguers enter these strings is by cutting and pasting from the webpage which is the target of the link. They copy the URL from the webpage itself into the catalogue record. Heaven help those users of any library system where the operating system does not allow cutting and pasting! Sometimes these URLs are generated by a database. So journal publishers may switch from a direct address to one generated by a database.

If journal suppliers change the URL they should have a link from the old address which allows automatic redirection to the new. But they cannot keep the two addresses for ever. Catchword recently taken over by Ingenta had servers to cover different parts of the world and the UK linked through Manchester. Now this has changed and libraries need to update their catalogues.

We have seen in the case study a journal which had more than one means of access. This can happen for other reasons: a journal may be available on more than one host. The hosts need to know who qualifies and there are various ways in which they do this. Electronic Journals may be obtained by subscription or they may be obtained by virtue of having a print subscription. The agent who supplies your journals should inform the publisher or database host when this is available (but they do not always do so).

4.2 Enabling access

A common situation in our library is that a reader accesses a journal and is told they have to pay for an article. Since the links we have been talking about usually go to a page devoted to that journal title, there is nothing to prevent a library from linking to every journal it can find on the web whether or not there is a subscription to the electronic journal. indeed there is some value in that because journal 'home' pages are often accessible to anyone and have links to contents pages and abstracts. And abstracts that are accessible to anyone can be quite useful. So if your library does not subscribe to a journal the user can get to a page where they can pay for an article themselves. Ironically, if I try to access Program through Ingenta I get this, because we never registered our Emerald subscriptions through Ingenta. You might think this does not matter because users at Middlesex can access Program through Aslib or Emerald. But if someone has found Program by finding an article of interest in Ingenta's article database, they would not necessarily know that Middlesex takes Program electronic version through another route. We can search for VTLS and Tomsk (PPT 18), this finds an article in *Program* (PPT 19) Click on the article availability and we are asked to pay £16.26 (PPT 20). This is a message that annoys students who retrieve an article that we are not entitled to access. They in turn get annoyed at Library staff because they expect a free service. It is even worse if they get this message because an intermediary does not know we subscribe and will not allow us access even when we are actually entitled to it.

4.3 How is access enabled

How is access enabled when everything is as it should be?

4.3.1 Access by IP address

Many publishers who host their own systems, and intermediaries such as Ingenta who host journals for others, validate requests coming from external sources by IP address. Every computer searching over the internet has an IP address. It may keep this permanently or it may be allocated it just for the session. Universities have ranges of addresses for their computers and these can be registered with the publishers or intermediaries so they know to accept requests from those computers. The publishers' computers also have to be told what each university has access to. Publishers know which libraries are subscribing to which of their journal (or they should do). Sometimes a journal which is taken is not given access and so library staff have to inform the controller of the host computer either directly or via a publisher. There are different methods of doing this and the hosts have set up self-registration procedures which usually end up with an email to the help desk since the self-registration often does not result in the database hosts getting the correct information. Here is ingenta's registration page. I search for the title *Program* (PPT 22). I enter a subscription number and the information is checked by Ingenta staff who will then add it to their access list so it becomes available or more likely when they check it they will not find we have access to it and we will have to correspond by email usually involving them speaking directly to our book supply agent. Unfortunately as issues are added with time, we do not always get access to the latest issues, so we have to keep undertaking this self registration activity and going through all the procedures I have just mentioned, and usually an email

is necessary to explain the situation. Publishers who own their own databases often have a variation of self registration. Elsevier Science work differently in that you contact their help desk via email and they add your entitlements.

A further complexity has recently been added both to Elsevier and Blackwells Science and Humanities. Before Blackwells Science required that you subscribed to Blackwells Science journals as well, nothing was free even if you were a subscriber to the hard copy. Now this has changed. With Elsevier's Science Direct it was different. You had to subscribe to Science Direct even if you had hard copy and this was expensive. Now they have changed their rules and have inaugurated Science Direct Web Editions (we don't have to pay to access this) to which they have now added a new feature which means we can see the last year of a journal we subscribe to without having to pay more. But this is at their website Web Editions where there is a list of the journals covered by this entitlement (PPT 26).

4.3.2 Access by password

Some providers of data require a password to be entered before accessing the data. The communication of this can cause problems to potential users. We place the key to the password on a website on our intranet so that it can be seen only by users in the university. Of course they may be told this over the phone if they do not have access to the intranet.

4.3.3 Access by authentication

In Britain there is a national authentication system run under the auspices of the Higher Education Funding Council. Details may be found at <http://www.athensams.net> (PPT 27). The system allows universities and colleges to allocate passwords to their library users. Athens administration keep a database of the resources to which each university is entitled for its users. Publishers may or may not join the scheme. Some allow IP authentication as well as this method, so that external users will access by this means. Some publishers do not trust the system which of course is open to abuse as a username and password can be communicated to anyone, though this is of course illegal as it constitutes fraud.

5. Who is involved in this?

As you can see the situation with electronic resources really is like a jungle with many different activities in different areas which do not relate well to each other.

On the provider side we have the publisher, database hosts, consolidators who are akin to the subscription agents of printed journals and often are the same. We forget about the authors in this context!

Between the provider and the end user we have the library service. Here we have people who initiate orders for journals (usually in electronic format). This will usually be the subject librarians. Next we have staff who do the actual ordering. Other staff register journal issues (usually in print format). We have staff who catalogue journals and ensure that the electronic journal is also catalogued. Other staff recommend journals to readers in their subject areas (subject librarians). We also have staff who check in journals. Often electronic access to a journal requires a PID which can be found on the wrapper of the printed copy. Then we have to speak to the local campus serial librarian to find that. Additionally we need involvement in ensuring that these journals are accessible in electronic format. As we have seen, this requires special efforts, in fact relating to identifying the journals and testing that they work. Identifying is the job of the subject librarian though other people may contribute to this activity. In the case of *Program* the hard copy continued to arrive and access to the digitised resource continued to remain available though not added to and another additional one became available. It would have been very difficult for anyone to identify this in the library service. The publisher or consolidators should have done the work on their systems by making it available to customers, but did not and in practice do not. (There is an issue here of subscriber *versus* membership.

Consolidators are much better at knowing about subscriptions than journals obtained through membership of organisations).

We also have on one campus a database librarian whose job it is to look after electronic materials.

Added to all these groups mentioned before, in Middlesex University the Systems Team is also involved. I mentioned earlier the problems of access to electronic resources. Athens the authentication system has a deputy administrator in the systems team. registration of IP addresses is felt to be technical, erroneously so perhaps, and is the responsibility of the systems team. The systems team provide the links to electronic resources which are not in the catalogue and now we add URLs to the catalogue a selection of those which are so we get involved. The systems team also provide tools for checking URLs. The systems team are also registered as administrators of those electronic resources which require an administrator. Some of these resources have facilities for listing the hard copy holdings so that users searching a database can see if the library holds journal articles found there in hard copy We have not set this up in many cases. This is usually done by the system team. However, the database librarian also does this for the materials on her own campus.

So you can see that there is a huge division of responsibilities.

Added to these the subject librarians should be reporting journals which are added or subtracted from the acquisitions.

6. What tools are available

Many of the tools I have been talking about are actually provided by the database providers or the consolidators. They offer mechanisms for the administrator to be able to see what journals the institution has registered as having subscribed to and what is the usage.

There are other tools which we can use to check the URLs in our library catalogue. We can produce from our system a list of records with a link using MARC 856 field to external resources. We can then feed this into a website (1) which will check each link and report back on them. The URL page it produces indicates whether links are non-existent, diverted, or OK.

7. Conclusion

It is difficult to see how the situation will improve but you can see that the situation regarding electronic resources is already dire. The scenario is caused by publishers needing to ensure they make money from their publications. Otherwise only half of the activities which I have been outlined would have been needed. Some journals are available free of charge on the internet, for example the British Medical Journal. Because this is used by doctors as a magazine, they still subscribe to it despite the availability on the internet and the publisher has reckoned that putting it on the web will not deplete their subscriptions. Issues relating to copyright could make the situation worse rather than better. Even the procedures in place at the moment as described above are not foolproof. Publishers may devise some methodologies which might be even more difficult to implement to ensure that fewer people who are not entitled to the resources can gain illegal access. So the future is not a rosy one when it comes to hacking our way through the electronic jungle.

References

1. W3C *HTML validation service*. <http://validator.w3.org> [PPT22]