

“Seek and Ye Shall Find” - Researching across Multi-Disciplinary Content

by Dianne Miles, Director of International Library Services, Ingenta PLC
presented by Richard Smart, Regional Manager Library Services

The title of this presentation is a quotation “Seek and Ye Shall Find” which seemed an appropriate one to me, with the sub-title of “Researching across Multi-Disciplinary Content”. This presentation is within the session covering the area Internet Information Supply and the Role of Search Engines in the Future. First of all I would like to briefly review the current information environment and then summarise some points about the supply of information via the Internet and the different access routes, go on to talk about the need for and benefits of searching across multi-disciplinary content and in this context how Ingenta has developed its online services to researchers searching for scholarly information, and finally, speculate on the development of such searching in the future.

The Information Environment

We are in an information environment where researchers and professionals require access to information at their workstation, wherever that may be and whenever they might need that information. These needs increasingly mean a desire for electronic online delivery and a particular need for remote access, especially to support distance learning.

As a corollary to this there is less physically centralised access to information than before. Indeed we are seeing a generation that has grown up with the Internet, and which seems to hold the belief that information is only of value if it can be accessed directly at the desktop, with an unwillingness to physically visit the library to find the information they need. For remote researchers this may not even be a realistic option. The library however can be seen increasingly as a facilitator in both organising electronic access and customising that access to the needs of the different constituencies.

Also users of information want searching to be a simple process involving few stages in order to access the information they need.

Internet Information Supply

There are a number of routes for researchers to access full text scholarly journal material via the Internet. This could be directly from the website of a publisher or indirectly via an aggregator. By its nature the publisher’s website will give access to that publisher’s material. We can loosely define what an aggregator does as:

- bringing together content from multiple publishers on one system
- providing a common interface to that content

Generally they allow cross searching of the bibliographic header and sometimes of the full text. They increasingly offer interlinking of content. Most importantly from the library's point of view they offer a single one-time authorisation layer. We can broadly divide aggregators into three types:

- traditional full text aggregators
- gateway aggregators
- hosting aggregators

Some examples may help to differentiate these "new" and traditional aggregators.

Good examples of traditional aggregators would be ProQuest, Ebsco Publishing and Gale Group. Core characteristics are that they

- hold the full text
- add value as a one stop shop to libraries
- offer publishers access to new markets albeit at a heavy discount

The aggregator bears the technology and marketing costs and there is little or no entry cost for the publisher. From the point of view of the library a "package", often focussed on a particular discipline, is sold complete and cannot be significantly unpicked. Since the bundled discounts are sometimes arguably viewed as a threat to subscription income publishers are increasingly embargoing the most recent material in order to avoid this.

The second category of aggregator I have identified here is "gateway aggregators". Examples would be the subscription agent gateways such as SwetsNet Navigator and Ebsco Journals Online. These have been joined more recently by newer gateways such as The Scientific World and TDNet.

The defining characteristic of most but not all of these gateway aggregators is that they point to content, by and large they do not host the material. Usually they hold bibliographic records, typically including the abstract, so that they can provide their users with a useful retrieval interface. The full text is held elsewhere on either a publisher's website or a hosting aggregators website and they pull the full text from a third party source. This type of gateway often forms part of a bundle of services that is offered to libraries as part of a major subscription package.

The third type of aggregator I have listed I have called the hosting aggregator. Examples would be Project Muse, CatchWord (prior to its integration with Ingenta), Allen Press and HighWire, although HighWire rejects this description of themselves.

They are differentiated from the gateway aggregators in that they exclusively host material. This type of aggregator can add a different mix of value, which includes a series of extensive and complex technology and distribution services which from the libraries point of view can lead to a wider choice of access routes to subscribed material and means that users can reach content regardless of their chosen access method.

Like other aggregators they offer libraries a one-stop shop. Typically they are paid by the publisher and so are in a sense “Digital Presses”, as HighWire describes itself and, like a conventional printing press will be paid for the services they provide. The technology platform used to support the aggregation service has seen further development to support quite independent subject or publisher focussed web sites and portals.

Additional services that are now typically available from hosting aggregators include the creation, maintenance and resolution of reference links, TOC alerting services and the ability to search for related articles. Some services also offer a rapid “pre-publication” service. Additionally access can be made available for non-subscribers on an e-commerce basis.

Distribution services for content can also provide a valuable service not only to the publisher who would like their content readily available but also to the library and researcher who is given a choice of routes to content. This means that users can reach their subscribed content regardless of their chosen access method and also allows for access to non-subscribers via pay-per-view or document delivery. An additional benefit to both publishers and librarians is that this allows for the collection and interpretation of statistics and provides valuable feedback on researcher behaviour and needs.

Ingenta is a hybrid in that it hosts the majority of its content but also points to a large proportion as well. Ingenta offers the researcher a comprehensive, multi-disciplinary collection of content which includes 5400 full text online journals from more than 180 publishers as well as 20000 further titles with text available for traditional document delivery via fax and Ariel. The user can freely search the bibliographic database for all these titles and access the electronic full-text without additional charge where they have verified subscription rights. Where there is no subscription typically pay-per-view or document delivery is available.

Ingenta’s proposition to the publisher is to offer content conversion and distribution in a model that both supports their subscription revenue and allows for e-commerce. This is particularly attractive to smaller publishers who cannot afford to do this on their own. Outside of the top eight STM publishers, there are something like a further 20,000 titles published by more than 5000 publishers.

Searching across Multi-Disciplinary Content

Some reasons for searching via a multi-disciplinary source of content as opposed to a discipline specific source might be:

Ease of access and breadth and richness of material.

Area of desired information is not clearly discipline specific or no obvious place to start the search?

Know where to start but want to access the most comprehensive material available?

Have some interesting material and want to access other articles referenced?

Want to check for related material from unlikely sources?and so on.

Traditionally a researcher might have started with a bibliographic or A& I service and then looked for the full text in paper in the library, via interlibrary loan or traditional document delivery. Now researchers use online search engines and customised alerting services to identify articles of interest which they can access online either through their subscription access rights or by paying for the individual articles.

Research commissioned by the Ingenta Institute last year¹ found that in the sample group studied articles were being accessed by both these methods in order to maintain awareness of current research. The study also found that researchers used the recommendations of colleagues and peers extensively to identify articles of interest and noted that a service that helps users provide references that make it easy to locate articles would add value to this process.

Ingenta and others are increasingly incorporating technologies which allow reference linking and the use of the OpenURL to allow the creation of direct links to an article. Ingenta is now a valid target for any OpenURL resolver such as SFX and several other similar products which are currently emerging.

Although not the usual context in which it is mentioned the well known study by OhioLINK² regarding use of its Electronic Journal Center which found that 52% of usage was from titles not previously held on subscription at the user's campus reinforces the benefits of carrying out searching across a broader spread of journals.

The Future

I believe in the future we will see a variety of developments of both the quantity and type of scholarly material available with both more journals becoming available online and also an increased use of the abilities of electronic media to display information in ways that are not just page facsimiles. An example of the latter would be for example gene mapping, which is not effectively displayed two dimensionally. There will also be an increased tendency to be able to customise services in ways that better meet the needs of individuals and institutions. There will be more use of reference linking and other linking initiatives such as the ability to link from an article to other relevant types of material, for example to patent databases. Specifically in the area of linking, Ingenta is starting to broaden our offering of reference linking capability. More work will be done on electronic archiving solutions and there are a number of initiatives already in process in this area. There will also be an increasing use of statistical information allowing the refining of services provided and used by researchers.

Summary

In an increasing electronic environment there are many benefits to multi-disciplinary searching particularly where the search can be carried out on a broad based group of scholarly journals using methodology that incorporates the ability to make further connections to relevant material.

References

- 1) "A behavioural study of end-users of full-text articles", David Worlock for The Ingenta Institute 2001 Proceedings
- 2) A method out of the madness: OhioLINK's collaborative response to the serials crisis, Sanville T J, *Serials* 14.2 pp163-77, 2001