Aggre-culture: a view of what e-book aggregators offer libraries
Ray Lonsdale and Chris Armstrong

Ray Lonsdale is Reader in the Department of Information Studies at Aberystwyth University and Chris Armstrong is Managing Director of Information Automation Limited

During the last decade, e-books have grown to become a significant library resource; a 2007 international e-book survey showed that 88% of respondents “answered that they own or subscribe to e-books and nearly half of the respondents (45%) have access to more than 10,000 e-books” (ebrary, 2007). Although annual US wholesale e-book sales rose by 23.6% in 2007 (Rothman, 2008), take up has continued to be uneven in the UK. Higher and Further Education continue to dominate sales with a number of public and special libraries also committing to the format, but other sectors – such as schools – are only just becoming aware of the possibilities offered.

In the same way that some years ago aggregators developed platforms of e-journals, aggregators now have become an important means of access to e-books, providing large collections of titles from many publishers under a single contract and through a single interface. They now constitute a major source to which UK libraries turn for the provision of e-books, as evidenced in a recent JISC report: “Until 2005 roughly 50% [of HE institutions] had been buying e-books from aggregators, but this figure has risen to 63% in the last 18 months.” (Higher Education Consultancy Group, 2006)

Researching the article
Given the growth and importance of e-book aggregators, as well as the lack of information in the literature, it seemed an opportune time to explore developments in this sector. To undertake this task, a short study was conducted which focussed on major aggregators: Ebook Library (EBL), Dawsonera, MyiLibrary from Ingram Digital, Credo Reference (the re-badged xreferplus), Gale Virtual Reference Library (GVRL, from Gale/Cengage Learning), Knovel, KnowUK from ProQuest, Safari Books Online (available in the UK via ProQuest), NetLibrary, ebrary, and Questia. All offer information on their websites, and two – Knovel and Questa – allow searching from non-subscribers. It should be noted that the discussions which follow deal with the main aggregations for each company. Some companies, for example Ingram Digital which offers VitalSource for textbooks, may offer additional products. It is envisaged that the new Gale Directory Library will extend to include titles from other publishers and will thus become an aggregator product.

[Take in table of URLS]

<table>
<thead>
<tr>
<th>Credo Reference</th>
<th><a href="http://corp.credoreference.com/">http://corp.credoreference.com/</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Dawsonera</td>
<td><a href="http://www.dawsonera.com/">http://www.dawsonera.com/</a></td>
</tr>
<tr>
<td>Ebook Library (EBL)</td>
<td><a href="http://www.eblib.com/">http://www.eblib.com/</a></td>
</tr>
</tbody>
</table>
It was necessary to set some parameters for the study, and there are two areas which deserve a lengthier treatment than this article can provide: licensing and a comparison of software features.

Possibly one of the most difficult areas for librarians faced with collecting e-books is that of licensing and charging models. As the JISC report stated, “It is clear from our discussions with library staff that there is a considerable amount of ignorance and misunderstanding about what e-book deals are on offer and what terms can be obtained from aggregators and publishers” (Higher Education Consultancy Group, 2006). Licensing is a complex subject with probably as many approaches as there are aggregators or publishers. Some e-book aggregators offer a purchase model in which the library owns books that have been purchased – ownership, in this case, meaning a fixed number of accesses each year; this model may incorporate a small overhead for the platform and in some cases also offer a rent-to-buy approach. The other main model is the concurrent-use licence by way of an annual subscription to a collection or part of a collection. Some models allow users to download e-books to a laptop for a limited loan period; others allow library users to view titles in the aggregator’s master collection but to which the library has not yet subscribed. Finally, MyiLibrary has a pilot ‘inter-library loan’ system developed with the Canada Institute for Scientific and Technical Information (CISTI), which allows users to purchase one month’s access to an e-book by way of a credit card payment (http://cat.cisti-icist.nrc-cnrc.gc.ca/screens/myilibrary_e.html). Each aggregator will be able to give full details of their licensing/subscription model.

Secondly, the article does not provide a comprehensive listing and direct comparison of platform features, as the JISC is beginning to provide this through their Academic Database Assessment Tool (ADAT), which aims to “help libraries to make informed decisions about future subscriptions to bibliographic and full text databases” (http://www.jisc-adat.com/adat/home.pl). The most recently added area concerns e-book platforms, and five aggregators are compared there by means of over 70 criteria under 10 headings, which include: Access Control, Search, Indexing, Search Results, Restrictions and Usage Statistics. Here, we have explored some of the more novel features that aggregators offer users.

[Take in screenshot - © cleared]
JISC ADAT showing part of the comparison table
What follows is an overview of e-book aggregators and their services for those who are unfamiliar with e-books; while for those who are already developing e-book collections, the article highlights a number of trends in e-book aggregation and explores issues facing librarians and publishers. To facilitate this, we devised a set of questions covering five areas: content and coverage, features, pricing and licensing, working with publishers, and working with libraries. In seven cases, data were collected through face-to-face or telephone interviews while, in two others, aggregators chose to submit a written statement. The primary data was complemented by examinations of their websites; however in the case of Questia and KnowUK information was taken from their websites alone. Interview material was recorded and transcribed, all data were analysed using QSR NVivo7.

**Aggregator characteristics**

If, allowing for what has already been stated, aggregators can be typified, we would suggest that the provision of large collections of e-books from a range of publishers, all of which may be used on a single software interface, is the most important characteristic. Often, particularly but not only with reference collections, the full text of the entire collection can be searched at one go, and the book structure or model is visibly retained, often by an explicit and expandable table of contents in a separate pane on the screen. There will be a range of features such as the ability to search in-book, and to have password-specific book-marking, highlighting and note taking; for the most part, books are intended to be read online although a few aggregators may allow the e-book to be downloaded to a laptop. Because the majority of aggregators offer digitisations of printed books, there is little evidence of interactivity or multimedia elements in the collections. However, Credo Reference includes audio files (pronunciations, music and literature readings), animations and videos within selected encyclopaedia. Aggregators will all offer free trials, MARC records for individual e-books, and at least COUNTER-compliant or COUNTER-equivalent user statistics (John Cox (2008) compares user statistics from different e-book vendors); most will allow federated search access to the collection. The majority work with consortia as well as with individual institutions.

**Collection content**

With respect to subject coverage, EBL claims an STM emphasis and has the strongest Australian collection. Unusually, GVRL has focussed upon the humanities and on literature, in particular, although Gale anticipates that science and health are destined to develop. Dawsonera, MyiLibrary, NetLibrary, Questia and ebrary are truly multi-disciplinary (although MyiLibrary initially had strengths in the medical and health sciences area). Only NetLibrary includes fiction and children’s fiction, while both it and GVRL include collections suited to the schools market. Reference collections tend to be smaller and the four aggregators included here vary from around 120 titles (KnowUK) to some 2,000 (Knovel); the general aggregators vary from around 48,000 (Dawsonera) to 170,000 (ebrary). Typically each aggregation includes titles from 60 to several hundreds of publishers. Safari, which specialises in online IT and Business/management, has a smaller collection of around 6,000 titles from 26 publishers.
Four of our aggregations are clearly reference collections: Credo Reference, Knovel, KnowUK (a UK-related collection) and GVRL, with ranges of traditional reference sources including encyclopaedia, dictionaries, atlases, year books, and guides. Knovel focuses specifically on science and technology. The remaining aggregators do not focus on a particular form of e-book, although one – NetLibrary – has always offered a collection of free reference titles alongside of its much larger main collection. The more general aggregators all include a selection of e-book types including monographs and, to a limited extent, textbooks, but their collections also invariably offer more than books.

Safari also offers ‘Rough Cuts’, which are “books before they are published, so [enabling] authors to make their books available online – to get feedback from their audience at an early enough stage that it could feed into the final work” and ‘Short Cuts’ for authors who “have something to say, but it’s not enough to fill a book, 80 to 100 pages – somewhere between a super-article and a very short book … [it] probably won’t be printed – this is content specifically for online.”

The majority of aggregators will update collections as new editions become available. GVRL are moving towards dynamically updating the online content of certain of their titles, including *The New Catholic Encyclopaedia* and *Grzimek's Animal Life*, moving away from the need to wait for a new edition to be published, and analogous to the process of continuous revision.

There has been a marked move – possibly related to OCLC’s pronouncement (2004) that users are, or will become, “format-agnostic” – over the last couple of years for e-book aggregators to add non-book content to their collections, and to market them using the term ‘eContent’. As one interviewee expressed herself on adding journal material on an issue-by-issue basis, “I don’t think there’s any harm in doing that … if it makes sense for the patrons to find it there, or for the library to want to acquire it within our platform”. It brings together rich content in one place in response to a search query, certainly, which users find helpful. More than 80% of respondents to ebrary’s Global eBook Survey stated that integration of multiple online resources is “very important” (ebrary, 2007). Questia offers journal, magazine, newspaper and encyclopaedia articles; ebrary does not include newspapers, but adds sheet music, maps and reports (e.g. from Dun & Bradstreet and Datamonitor) to the mix. NetLibrary offers eAudioBooks, databases and e-journals; and Safari also has non-e-book documents and articles, while also adding some 600 hours of video instruction. MyiLibrary is moving in the same direction and has noted an increase in demand for certain of these formats, while Dawsonera speaks on its website of including “new media,” and audio was mentioned in the interview. GVRL, KnowUK, Credo Reference and Knovel – the four reference aggregators – vary: Knovel has also moved beyond the e-book to include “databases and other e-content”, while Credo Reference and KnowUK focus only on reference books.

*Take in Questia screen shot*

Search on Questia showing multiple econtent formats
While the facility for researchers to have a range of resource types available in one place sounds good and has evidently been welcomed by many libraries, it is necessary to sound a note of caution. Assessing bibliographic databases in the recent past, reviewers often commented on uneven coverage or on the lack of information on coverage – we have the same situation here. Other than GVRL, which links to InfoTrac for journals, it is not apparent how or why particular journal articles or issues have been selected for inclusion amongst the e-books, and one has to hope that library end-users will realize that they are being offered only a very small, and selective, window into a journal world.

Market trends
One of the principle objectives of our study was to explore recent trends in the markets aggregators are addressing for e-books. Historically, e-book publishers and aggregators focused upon the academic library sector, initially within higher education and latterly, through the initiative of the JISC e-Book Working Group, on post-16 education. Public libraries and special libraries too have, in recent years, benefited from a broadening of the range of e-book titles offered by aggregators. These sectors still constitute the principal market, as typified by ebrary which sees its users as “based in universities, further education colleges, corporations and public libraries”. However, our study revealed that in several instances, aggregators are now exploring the potential of the children’s and school library sectors. For several years, NetLibrary has expanded considerably its collection for North American school libraries so that it now constitutes a critical market. Whilst this remains a focus in the US, there is a suggestion that ultimately the collection will develop to encompass the needs of schools internationally. Similarly, GVRL has a significant portion of its collection which is used by US schools, and again, some of which will be internationally relevant.

Another aggregator that is currently dipping its toe into this pool is Credo Reference: “If a public library in the UK subscribes to Credo then we can actually create a second site for them. They’ll have access to two versions of Credo – our standard version and a customised student one. It’s a subset of books that can be targeted at a particular subject area.” Credo Reference are currently conducting book trials in a couple of libraries, and indicate that it is likely to focus on the secondary school student but, assuming that the market is viable, “there’s nothing to prevent us from targeting the more junior people”.

Over the past five years we have been conducting a series of workshops on e-resource provision for the school and children’s library sectors, and are cognisant that the time is ripe for a critical mass of e-book material to be made available. There remains, however, the problem of the fragmentary nature of this market with their often small devolved budgets, which aggregators have identified. This is gradually being acknowledged by school libraries and school library services, and also more recently by the JISC and Becta.

In respect of geographical orientation, several aggregators such as MyiLibrary are UK oriented, and are now looking to extend their international content. However, one of the early criticisms levelled by UK libraries at some of the general aggregators who were established in North America concerned the predominantly US bias of their collections.
Our study suggests that there is a discernible move towards offering a significant quantity of e-book titles from non-US publishers. NetLibrary, for example, indicates that over 60,000 titles from the collection of more than 160,000 emanate from non-US publishers; ebrary now offers more than 70,700 titles from over 74 non-US publishers and aggregators – including 35 major UK publishers – from its growing repository of more than 170,000 titles. Dawsonsera and Safari indicated that they were now becoming more Eurocentric, and the latter was drawing in content from Indian publishers; EBL has “a good mix of publishers in all regions – Australasia, UK and Europe”, and suggests that there is no discernible bias: “I think we have [none], because we’ve tried to be a little bit stronger with UK-based publishers to date”.

The move towards internationalism is also reflected in the development of multilingual interfaces and multi-language content. NetLibrary, Credo Reference, and MyiLibrary are amongst those that offer multilingual interfaces or are developing them. The range of interface languages is broad, extending beyond European to some of the Asiatic languages. The ebrary platform and interface is currently available in English, Spanish, Portuguese, French, and Swedish. This year, Credo Reference will be developing a new enhanced version of the system that will enable them to take multi-language works, and following market research into the needs of UK libraries they have identified Polish, Urdu, and Chinese as priorities. The new ebrary Reader, currently in beta, supports any language type, which means that any e-book, in any language, can be viewed and used.

Few other aggregators offer non-English language material, although some are addressing this dearth, and recognise the need to extend the volume of multi-language content. NetLibrary has been conscious of this issue for some years, and now offers approximately 7,500 non-English language e-book titles in over 20 other languages. Critically, there is a move towards serving the markets in China, Japan and India. ebrary’s selection also includes over 29,800 non-English titles. Credo Reference offers bilingual dictionaries and, given the predicted multilingual interface, anticipates developing collections of monolingual non-English reference titles. While the higher education market predominates for GVRL, the emphasis remains firmly on English-language titles, however they are conscious of the market potential and the collection already includes titles in Spanish.

**Developments in platforms**

Over and above the typical features previously mentioned, aggregator platforms offer some very sophisticated means of access to e-book content. Before discussing these it is interesting to note the comments of some students in the SuperBook survey of e-book usage in University College London libraries (Armstrong, Lonsdale and Nicholas, 2006; Armstrong & Lonsdale, 2007/8), who consistently used the browser or PDF reader find function rather than the much more elegant search feature provided within the software. When asked, they replied that they were unaware of, or had not noticed, the search box, let alone the advanced search feature. Allen McKiel noted in his analysis that the ebrary
survey, “provides indications that e-book collections and the research tools they provide are not well understood by a significant percentage of faculty and students.” (ebrary, 2007)

However, aggregators believe that they have a clear view of how their systems are used; as Safari noted: “we find it is used much more for practical learning where someone has a very definite problem to solve or a research topic – they spend on average about 30 minutes, so its not for immersive reading”, while MyiLibrary accepted that “it’s all about discoverability – if you’ve got stuff … we can improve discoverability.”

In addition to the conventional searching facilities, some aggregators offer more innovative approaches. Safari – because of its IT specialism – has developed the search facility to specify computer and programming code, and has cleared copyright on this content as well. This answers the recent debate on LIS-E-BOOKS on how back-of-book CD-ROMs are dealt with by some e-book publishers. Knovel has a range of features including searching for data by range, or data that is buried in a table or database: “Knovel will return just the relevant rows… Knovel is smart about units. If you search for a value in lb/ft³, Knovel will find the relevant results even if they are published in g/cm³.”

Credo Reference have looked at different learning styles – visual learners or individuals who are ‘exploring’ topics – and have developed tools specifically to help them. The Concept Map explores visually the terms associated with an initial search terms and provides a mind-map of subject nodes, which link to text in the various reference books: “I’ve seen people who are visual learners – their eyes just light up when they see that … you can choose a topic like global warming or perception or whatever, and just – it helps you get ideas about how to broaden that … you are looking at the inter-relationship and exploring a topic.” Further developments are underway to significantly enhance the Concept Map to give users more control of output and searching.

As we noted above, for the most part aggregator collections comprise digitisations of printed titles, but our study revealed a small degree of interactivity developed by some aggregators. While all aggregators are constantly improving their software, the reference aggregators – possibly because searching across the whole collection is so important – seem to have advanced further. Credo Reference has added user-definable timelines, interactive maps, and dynamic statistics tables. Similarly, “Knovel’s interactive tables can show or hide rows and columns, [and with] Knovel’s equation plotter, you can simply select the values of interest on the curve and export them to Excel for further manipulation, again with the source citation.”

Variations on book-marking, highlighting and note taking are common to most platforms: ebrary notes: “The current ebrary Reader™ allows end-users to take notes alongside individual pages and store them on their personal bookshelves for later retrieval. [The] new Java-based Reader, currently in beta, offers improved annotating features, with any combination of:

- Multiple highlights and notes per page
- Resizable and movable notes
- Highlights with or without notes attached
- Color coding of notes and highlights.”

Some aggregators offer referencing help by providing citations, sometimes in as many as half a dozen styles, and Knovel automatically exports a citation with any text or table. ebrary also offers automatic citations when text is copied and pasted and printed. ebrary’s citations, available in several different formats, also include a URL hyperlink back to the source. The interim results from SuperBook offer an important caveat in that, even where there is a choice of styles, these do not necessarily match the needs of students and academics. MyiLibrary is fully integrated, and Credo Reference and ebrary are working on integration, with citation software such as RefWorks and EndNotes. Those that do not offer this kind of help recognise it as desirable and are likely to make it available soon.

For the most part, e-book aggregations are marketed as independent collections. However, a different approach is offered by Credo Reference, the implication being that the collection is not necessarily totally self-contained but can and should be linked to the wider resources of a library, be they print or electronic: “we are going to make it possible to continue searching through other resources. Our search results and entry pages will provide a means for users to continue their search at an OPAC or additional library resources. These options will be library-defined and offer a gateway to further research. We really see reference works, particularly subject encyclopaedias, as being a foundation to your research, a gateway to the library.” Similarly, ebrary allows users to link out to additional electronic resources such as bibliographical databases, dictionaries, translations and other online resources in the library and on the web through ebrary InfoTools™ software. The InfoTools menu can be customized by the library to link to the online resources of their choice.

Other aggregators mention that their platforms do not prevent the possibility of facilitating hyperlinking to external sources if this has been undertaken by the publisher, and that they are willing to explore this further: “And I do have some ideas – you know I have these visions of ways that we might be able to inject the citation information with some ways to the OpenURL or whatever”. (EBL)

**Bibliographic issues**

A perennial problem that has been articulated for the past decade in research into e-books (Armstrong and Lonsdale, 1998; Armstrong, Edwards and Lonsdale, 2002; Higher Education Group, 2006) is the fraught issue of bibliographic control of the format. Given the lack of a single national bibliographic source for the foreseeable future, the identification of new titles can be a difficult and time consuming activity. One important development has been the integration of e-book titles within several international bibliographic sources such as Global Books in Print and WorldCat. Throughout the study some aggregators recognised the significance of this issue, and acknowledge the need to be more proactive in exploring the implications of bibliographic control. Several aggregators do put their titles in these bibliographies. For example, NetLibrary and Knovel have their titles within OCLC WorldCat, and GVRL within Global Books in
Print. Although they do not work through the major bibliographies, Safari note that, for their business library, all the titles are indexed in ABI-Inform, which again offers another form of discoverability. Ingram Digital has worked with Microsoft Live Search Books to make its e-book content discoverable through Live Search Books with a tie back to MyiLibrary.

Several aggregators commented that since they are integrated within a library supply system, and since libraries have traditionally used suppliers as the major source for selection and acquisition, they had felt there was no need to make e-book titles available in the international bibliographies. However, as a consequence of this study, one aggregator was interested in exploring the implications of having e-book titles listed in the international bibliographies.

Working with libraries
We were keen to explore the perceptions that aggregators held on working with libraries. There is an evident desire to offer an attractive deal for libraries while maintaining a balance in negotiations with publishers; as epitomised by the following approach: “you used to buy 500 of our 1000 titles for this much money – why don’t you buy 800 for the same money, because with the power of e- we can deliver that at no more cost to the publisher, and they can pay the authors what they always paid them. So everyone gets the same amount of money they used to get, but the institution gets more content.” (MyiLibrary)

One issue of particular concern related to the promotion and marketing of e-books within libraries. Over several years we have been running workshops about the management and promotion of e-book collections, and anecdotally it is evident that librarians from all sectors face the ongoing challenge of creating an awareness of new e-collections and developing skills strategies to exploit them. In the limited research that has been conducted into this sphere (Gold Leaf, 2003; Armstrong and Lonsdale, 2007/8) there is evidence to suggest that students and academic staff in university libraries appear ignorant of newly established e-book collections, and that it is incumbent upon librarians to do more to facilitate usage. Some aggregators share this view: “there is an issue with publicising the stuff and having a collection that is attractive ... I’m sometimes frustrated – a library will buy 12 e-books and then after a year they’ll go ‘well nobody used them’, and I’m thinking well no because they weren’t made available. I think libraries have to go through an educational process.” (MyiLibrary)

Whilst libraries are seeking new and innovatory ways of promoting the e-book collections – with varying degrees of success, several aggregators were of the opinion that libraries might make greater use of their promotional products. NetLibrary, for example, were concerned that there had been lack of take up of their marketing materials (e.g. their marketing kit is available at: http://library.netlibrary.com/marketingkits.aspx). Credo Reference has promotional materials “Promote it to your Library” and case studies of ways of using the reference collection, and GVRL have posters. Meanwhile, ebrary offers promotional materials as a part of its a “getting started kit” for new customers, as
well as videos explaining product features, quick start guides, and complete user guides; and recently launched a live web-based training program.

**Impact of new technologies**
All of the aggregators to whom we spoke are very conscious of new developments, particularly the increase in interest over e-book readers and mobile technology. While EBL noted that titles were “already are available on handheld readers – the books can be transferred to a PDA, and we are working some device companies”, and Credo Reference said that “part of our development for our next version will enable mobile technology, but actually Credo today looks pretty good on the Blackberry – you know it does work”, Safari acknowledged this would be delivered soon: “we are looking at 2 things – one is making Safari available on hand-held browsers – could be phones, could be a variety of things”. Likewise, NetLibrary acknowledge that they are exploring the possibilities, and Dawsonera also offer some PDA compatibility, “Yes. Some will work with the software, some wont” but were less convinced by e-book readers, which were not seen as “very important yet”.

**Conclusions**
Overall, aggregators represent a major, perhaps the major, source for e-books in libraries, bringing together vast collections of general-academic titles and quite large reference collections for easy access. While English-collections dominate, there is a discernable move towards developing multilingual access and content. Markets are focussed in academia and public libraries, but schools and special libraries are increasingly represented, and all aggregators acknowledge their needs. Aggregators provide access to, and the ability to interrogate, the digitised texts in innovative and helpful ways using an array of software features. While charging models are complex, aggregators believe that they represent a fair deal to both the publishers and the libraries, both of whom have to be accommodated in negotiations.

During the study we were listening to the voice of the aggregator: to their perceptions of how they may be able to enhance the library culture of e-book collection and use. Whilst many offered summaries of strengths that were unique to their collections and interfaces, several offered insights into the aggregators’ role as trusted intermediaries. These perceptions are reflected in the closing extract:

“a lot of our work is about bringing together the publisher and the library … [with] thousands of publishers and millions of consumers (libraries, or whatever it might be), someone needs to pull this together … And that requires a bunch of stuff – technology and transfer mechanisms and delivery mechanisms and all that stuff; but it also requires relationships with publishers and trusted relationships… [with] libraries and library consortia.” (MyiLibrary)

**References**


