

Back to [StaffWeb Home](#)

# Enriching the Catalog with Table of Contents Data

## A report for the Cornell University Library

Prepared by David Banush, with assistance from Jean Pajerek

March 7, 2002

### Contents

#### [I. Executive Summary](#)

#### [II. Introduction](#)

#### [III. Overview of TOC Services Available](#)

#### [IV. Considerations for TOC enrichment](#)

#### [V. Vendor information](#)

#### [VI. Costs and timing of implementation](#)

#### [VII. Complementary services](#)

#### [VIII. Recommendations](#)

## I. Executive Summary

The addition of table-of-contents (TOC) enhanced bibliographic records to the Cornell University Library (CUL) catalog has been under consideration for several years. In December 1997, Marty Crowe prepared a report for CUL senior management called "Table-of-Contents Enhancement of the Catalog." In it, she identifies a number of reasons for TOC enrichment. Probably the most important of these is giving users significantly enhanced intellectual access to library materials at the point of searching, regardless of where the search takes place. The addition of TOC data is a user service that adds value to our records and saves users' time in identifying relevant items.

The potential value of TOC enrichment has not diminished since the original report was issued. Recognizing the benefits of enhanced access to printed monographs, even as we increasingly direct our efforts toward the provision of digital resources, the CUL Digital Futures Plan explicitly calls for adding tables of contents to bibliographic records for monographs in the online catalog (section II.A2). TOC-enhanced bibliographic records would enable CUL to provide a new dimension of service in anticipation of the heightened expectations of library users.

This report updates the 1997 document by taking a fresh look at the TOC enrichment marketplace. It provides a general overview of the services provided by four principal vendors, the costs of these services, and the added costs of handling the processing internally. Timing, alternative and complementary services, and other implementation factors are also considered.

### Considerations for TOC enrichment

*Prospective or retrospective enrichment:* Costs will be significantly higher to retrospectively enrich records than to enrich records only in a prospective fashion. However, the added benefits to library users may justify the expense, and that cost would be largely a one-time expenditure.

*MARC tagging and display:* Lengthy contents notes are common for titles in certain disciplines, mostly in the sciences. The Voyager OPAC display of these records may be difficult to read and interpret for many users. Stakeholders will have to determine if the drawbacks of the Voyager display can be minimized, or if they can be justified by the additional benefit of increased access to materials.

*Keyword indexing and searching:* Adding TOC data, especially if both retrospective and prospective enrichment are selected, will significantly increase the size of the keyword index. This increase will affect searchers' results, with an increase in recall and a decrease in precision. TOC data will most likely be accessed through keyword searching, though the Library could choose to create a new keyword search that would index only TOC data to improve precision.

*Cataloging and database maintenance:* The largely automated fashion in which TOC data would be loaded would have no effect on the day-to-day workflow of cataloging staff. Enrichment would require additional programming effort from staff in LTD and careful coordination with staff in the Database Quality and Enhancement unit (DBQ & E) to ensure that the TOC enrichment process works harmoniously with other automated maintenance and enhancement functions. It will also be necessary to ensure that enriched data not be exported to OCLC WorldCat or RLIN.

## TOC vendors

Four potential vendors of TOC data were identified. Each offers a variety of options, with varying degrees of flexibility as to how TOC data may be taken. Coverage for each vendor is limited to English-language monographs; no serial or non-book material is included.

The vendors are: Blackwell's, Syndetic Solutions, MARCIVE, and the OCLC MARS service. Highlights of their services, with advantages and disadvantages, are given below:

### TOC Services

Vendor	Advantages	Disadvantages
<b>Blackwell's</b>	<ul style="list-style-type: none"> <li>▪ Most experienced vendor</li> <li>▪ Largest files -- data from 1992 to present</li> <li>▪ Choice of 505, enhanced 505, or 9XX</li> <li>▪ Re-enrichment at no charge</li> <li>▪ Higher hit rates</li> <li>▪ 1 day turnaround time (prospective enhancement)</li> </ul>	<ul style="list-style-type: none"> <li>▪ Highest costs (\$1.05 per record enriched)</li> </ul>
<b>Syndetic Solutions</b>	<ul style="list-style-type: none"> <li>▪ Lowest costs (\$.50 per record enriched)</li> <li>▪ Choice of 505, enhanced 505, or 9XX</li> <li>▪ Re-enrichment at no charge</li> </ul>	<ul style="list-style-type: none"> <li>▪ Relatively new vendor with no long-term track record</li> <li>▪ Smaller data files, with lower hit rates</li> <li>▪ 2-3 day turnaround time</li> </ul>
<b>MARCIVE</b>	<ul style="list-style-type: none"> <li>▪ Lowest costs (pricing identical to Syndetic Solutions)</li> <li>▪ Choice of 505, enhanced 505, or 9XX</li> <li>▪ Established library service provider</li> </ul>	<ul style="list-style-type: none"> <li>▪ Smaller data files, with lower hit rates</li> <li>▪ Reliant on Syndetic Solutions for data</li> <li>▪ No re-enrichment without additional charge</li> </ul>

**OCLC MARS**

- |  |  |  |
|--|--|--|
|  | <ul style="list-style-type: none"> <li>▪ Choice of data from Blackwell's, Syndetic Solutions, or both</li> <li>▪ Pricing identical to use of Blackwell's and/or Syndetic Solutions</li> <li>▪ Choice of 505, enhanced 505, or 9XX</li> <li>▪ Re-enrichment at no charge</li> <li>▪ Established library service provider</li> </ul> | <ul style="list-style-type: none"> <li>▪ 3-5 day turnaround time</li> <li>▪ Reliant on both Blackwell's and Syndetic Solutions for data</li> </ul> |
|--|--|--|

**Costs**

Cost estimates are as follows:

For prospective enrichment only, the estimated annual cost range runs from \$2800 for 5600 enriched records per year to ca. \$18,100 for ca. 20,000 records per year, depending on the vendor(s) selected.

For retrospective enrichment (covering materials with an imprint date of 1992-2001), the estimated cost range runs from \$14,000 for 28,000 records to \$169,260 for just over 175,000 records, depending on the vendor(s) selected. Retrospective enrichment would be a one-time cost.

The great variation in the number of records and the costs for both prospective and retrospective enrichment reflects the significant differences in the services available and the size of the vendors' backfiles of TOC data.

For full information about costs, please see section VI of the report.

**Timing of implementation**

Because of ongoing projects, including retrospective conversion, and the loading of records for Maeda and Wiley, it is estimated that TOC enrichment could not begin until sometime during summer 2002, barring the addition of new high-priority project, unforeseen problems with ongoing projects, or with the TOC enrichment process itself.

**Recommendations**

Based on the suitability of the services offered by the vendors discussed here, the benefits to users, and the costs of enrichment, the following options are offered for consideration. They are

listed in order of potential benefit to library users.

### **Option 1.**

#### ***Contract with OCLC MARS to get data from both Blackwell's and Syndetic Solutions for prospective and retrospective enrichment***

##### ***Advantages:***

Maximum benefits to users: Electing both retrospective and prospective enhancement would allow us to add TOC data for a ten-year period, greatly enhancing access to thousands of recent English-language titles in many subject areas.

Flexibility: Choosing the OCLC MARS service would allow the library to retain all of the flexibility offered by each vendor. Retrospective enrichment could be done over the course of one or several years while prospective enrichment moves forward. Retrospective enrichment could be done over a number of years. The library could also choose to take retrospective data from one supplier while also opting to do prospective enrichment from both suppliers.

Value: By taking all of the Syndetic Solutions records, and only the unique Blackwell's records, we could reduce costs while still getting the best services from both suppliers.

##### ***Disadvantages:***

Complexity: While prospective enrichment would present fewer technical problems for DBQ &E and LTD, retrospective enrichment would require additional programming and staff time. Implementation may therefore take longer than prospective enrichment alone.

Costs: Choosing both retrospective and prospective enrichment would entail the highest costs, both in staff time and in vendor charges. It should be noted however, that most of the cost would be a one-time expenditure, unless the retrospective enrichment were spread out over time.

### **Option 2.**

#### ***Contract with OCLC MARS to get data from both Blackwell's and Syndetic Solutions for prospective enrichment only***

##### ***Advantages:***

Flexibility: Choosing the OCLC MARS service would allow the library to retain all of the flexibility offered by each vendor.

Value: By taking all of the Syndetic Solutions records, and only the unique Blackwell's records, we could contain costs while still getting the best services from both suppliers.

Relative ease of implementation: Programming work to extract records for export to OCLC WorldCat is already complete and would require only minor tweaking to be used for prospective enhancement. Implementation would be simpler and may move forward more quickly.

Lower costs: Without adding TOC data from the backfiles from each vendor, we would reduce costs of enrichment significantly.

***Disadvantages:***

Reduced benefits for users: With fewer enriched records, users would not realize the same level of benefit as they would from both retrospective and prospective enrichment.

***Option 3.***

***Contract with Blackwell's exclusively for prospective and/or retrospective enrichment***

***Advantages:***

Simplicity: Choosing the Blackwell's service alone would allow the library to retain all of the benefits the vendor offers (flexibility, higher hit rate), without having to worry about the relationship between Blackwell's and third parties like OCLC.

Turnaround time: Turnaround time would be faster with Blackwell's than using the MARS service; Blackwell's turnaround time for current customer (using prospective enrichment) is 24 hours.

Potential partnership with Syndetic Solutions: The possibility that Blackwell's may itself partner with Syndetic Solutions means we may enjoy the same benefits by working with Blackwell's directly as we would by working with OCLC. There may even be additional benefits like faster turnaround. However, the exact nature of this partnership and its potential advantages are at present uncertain and may never happen.

***Disadvantages:***

Reduced benefits for users: Without the supplemental records from Syndetic Solutions, fewer records would be enriched, and users would not realize the same level of benefit as they would if we elected to use the data from both vendors.

Cost: By taking all data exclusively from Blackwell's, rather than only its unique records, the Library would incur Blackwell's higher costs for every record enriched.

**II. Introduction**

The addition of table-of-contents (TOC) enhanced bibliographic records to the Cornell University Library (CUL) catalog has been under consideration for several years. In December 1997, Marty Crowe prepared a report for CUL senior management called "Table-of-Contents Enhancement of the Catalog." In it, the author identifies a number of reasons for considering the enhancement of the catalog with TOC information. Searchable tables of contents allow the user a deeper level of access to the Library's collection of printed monographs, regardless of whether the user has physical access to the material or not. For users who connect to the catalog from locations outside the library, and for those interested in titles that are housed in remote storage, the ability to search and examine tables of contents via the online catalog is unquestionably a great convenience.

Since the original TOC report was issued, there have been important changes in the marketplace for TOC enhancement services as well as within CUL, most notably the implementation of the Voyager library management system. The potential added value of TOC enrichment, however, has not diminished. Recognizing the benefits of enhanced access to printed monographs, even as we increasingly direct our efforts toward the provision of digital resources, the CUL Digital Futures Plan explicitly calls for adding tables of contents to bibliographic records for monographs in the online catalog (section II.A2). TOC-enhanced bibliographic records would enable CUL to provide a new dimension of service in anticipation of the heightened expectations of library users.

Like its predecessor, this report is intended to provide the library's senior management with the relevant information necessary to decide whether and how to enhance the Cornell OPAC with TOC data. It provides a general overview of the services provided by four principal vendors, the costs of these services, and the added costs of handling the processing internally. Timing, alternative and complementary services, and other implementation factors are also considered.

The Information about the vendors' services comes from conversations with their representatives as well as documentation they have provided, including their Web sites. Input has come from numerous colleagues throughout CUL. Colleagues at other institutions have also provided valuable feedback, in particular Jennifer Bowen of the University of Rochester and Steven Miller of the University of Wisconsin-Milwaukee.

### **III. Overview of TOC Services Available**

Since the last TOC report was written in late 1997, there have been some changes in the bibliographic record enhancement market. However, it remains similar in terms of available services and providers. Blackwell's Book Services, now as then, is the largest and most important player in the field, with the most extensive backfile of TOC data. Blackwell's also offers its customers a number of options for taking the data, but customers pay for these benefits in higher costs. A new company, Syndetic Solutions -- made up of many former Blackwell's employees -- has emerged in the past two years to serve a slightly different segment of the market with similar services at very attractive prices. MARCIVE has partnered with Syndetic Solutions to offer TOC enhancement, and OCLC 's MARC Record Service (MARS) is offering



customers the choice of either Blackwell's or Syndetic Solutions' data (or both) as part of its suite of bibliographic record processing services. Meanwhile, other suppliers of TOC data have withdrawn from the market. After tentatively entering the field in late 1997, Yankee Book Peddler has dropped plans to offer TOC services. As of December 2000, RLG has terminated its agreement with Blackwell's to supply enhanced records via RLIN.

For this report, four potential vendors of TOC data were identified. Each offers a variety of options, with varying degrees of flexibility as to how and TOC data may be taken. Coverage for each vendor is limited to English-language monographs; no serial or non-book material is included.

The vendors are: Blackwell's, Syndetic Solutions, MARCIVE, and the OCLC MARS service. Their services are briefly summarized below. More detailed information about each is provided in Section V below.

### ***TOC Services***

<b><i>Blackwell's</i></b>	<ul style="list-style-type: none"> <li>▪ Largest files, going back to 1992 for North American imprints</li> <li>▪ Most experienced vendor</li> <li>▪ Most flexible data options</li> <li>▪ Fast turnaround time</li> </ul>
<b><i>Syndetic Solutions</i></b>	<ul style="list-style-type: none"> <li>▪ Modest backfiles, but current files growing as quickly as Blackwell's</li> <li>▪ Covers materials outside of Blackwell's scope</li> <li>▪ Offers some flexibility in data options</li> </ul>
<b><i>MARCIVE</i></b>	<ul style="list-style-type: none"> <li>▪ Lowest costs</li> <li>▪ Uses Syndetic Solutions' data</li> <li>▪ Longtime library service provider</li> <li>▪ Fast turnaround time</li> </ul>
<b><i>OCLC MARS</i></b>	<ul style="list-style-type: none"> <li>▪ Offers use of both Syndetic Solutions' and Blackwell's files</li> <li>▪ Most flexible vendor</li> <li>▪ Longtime library service provider</li> </ul>

## **IV. Considerations for TOC enrichment**

### ***Prospective or retrospective enrichment***

One question that will need to be addressed is whether the library wishes to enrich bibliographic records only prospectively or would like to enrich older records as well. All of the vendors under consideration will handle either prospective or prospective and retrospective enrichment. A



primary consideration is cost: a retrospective enrichment would involve more staff time and significantly higher vendor charges than prospective enrichment alone. However, the expenditure for retrospective enrichment would be largely a one-time outlay, and the benefits of having nearly 10 years of enriched materials may be worth the added expense. Costs of both retrospective and prospective enrichment are covered in more detail in the cost analysis section of this report (Section VI.)

### ***MARC tagging and display***

Whether we use a standard 505 contents field or the enhanced 505, in which title and author information are subfield delimited, the public display in Voyager will be the same, i.e., in the format governed by AACR2, using ISBD punctuation to separate titles and statements of responsibility. For some users, the display may be confusing; differentiating between titles and authors formatted in this fashion is not always easy. This difficulty already exists in Voyager for records that may already have a contents note, but the number of those records is relatively small; the addition of many more such records could be more problematic.

The size of the 505, particularly for titles in the sciences, can be considerable. The Voyager Long View of such records may be overwhelming to some users, since much of the screen is taken up with the contents note. In such cases, scrolling is a necessity, especially on smaller monitors, and holdings and circulation information in the long view appear after the TOC data, at the bottom of a very lengthy page.

Some library management systems (e.g., Innovative Interfaces) can display TOC information taken from 9XX fields in a formatted fashion. Each chapter-level title and statement of responsibility is loaded into a separate 9XX field. Such formatting mimics the layout of a printed title page and presents a much more intuitive display to the end user. Voyager does not currently offer this functionality. We are therefore restricted to the standard contents note display and the use of either standard or the enhanced 505 in the MARC records. If in the future Endeavor updates the Voyager OPAC functionality to include the formatted display, we could re-enrich our records, moving the 505s to a locally defined (9XX) field. Three of the four vendors considered here will perform this re-enrichment without additional charge.

### ***Keyword indexing and searching***

Adding TOC data, particularly if done both retrospectively and prospectively, will significantly increase the size of the keyword index. This does not represent the same technical problem that it did in the NOTIS environment; keyword regeneration is already a standard routine run each week by systems staff. However, the increase in the size of the index will affect end users' results when searching. Recall will of course be greater, but almost certainly at the expense of precision.

505 fields are not included in the left-anchored indexes for Voyager. The only way to access data in these fields is through keyword searching. In order to facilitate searching of TOC data, the Library could create a new keyword search that would index only the 505 contents note. The

University of Rochester has created such a search in its Voyager OPAC. A left-anchored search could also be created, but would require customized programming from Endeavor at a significant fee. Moreover, the utility of a left-anchored search may be very limited, depending on our choice of vendors. Because the data in the 505 fields would include initial articles, and authors' names would appear in direct rather than inverted order, a left-anchored title search would only function if searchers included articles in their queries. That search method runs counter to all other searching and would be confusing for staff and users alike.

If the Library were to choose Blackwell's data exclusively, we would have the option of omitting initial articles from the titles in the 505 field, thus increasing the utility of a left-anchored title search. Authors' names, however, would remain in direct order, and data from Syndetic Solutions, if it were taken either alone or in combination with Blackwell's data, would contain the initial articles. Thus paying to have a new left-anchored search created for TOC data, while an appealing idea, may be very impractical.

### ***Cataloging and database maintenance considerations***

Because TOC enrichment would be handled in a largely automated fashion, there would be no effect on the workflow for cataloging staff who create or edit bibliographic records in Voyager. However, enrichment would require additional work on the part of LTD and staff in the Database Quality and Enhancement (DBQ & E) unit to prepare the files for output to the vendor(s) and to load the enriched records back into Voyager.

Several considerations would need to be addressed. The first would be producing an extract of data to send to the vendors. Since we are already preparing such files for output to the utilities, extracting the data from Voyager is unlikely to require much, if any, additional work, at least for prospective conversion. Retrospective conversion would require additional programming to extract older records. A second factor is the restriction, imposed by all vendors, that the enriched TOC data not be exported to the utilities. The algorithms that send our records to RLG and OCLC would thus need modification to ensure that vendor data in the 505 is not loaded into RLIN or WorldCat. Another consideration is the Marcadia record overlay process. We will need to ensure that Marcadia records do not overlay our enriched records.

### **V. Vendor information**

The advantages and disadvantages of each vendor's services are listed in the table below. Following the table is a narrative profile of the four vendors.

Vendor	Advantages	Disadvantages
--------	------------	---------------

<b>Blackwell's</b>	<ul style="list-style-type: none"> <li>▪ Most experienced vendor</li> <li>▪ Largest files -- data from 1992 to present</li> <li>▪ Choice of 505, enhanced 505, or 9XX</li> <li>▪ Re-enrichment at no charge</li> <li>▪ Higher hit rates</li> <li>▪ 1 day turnaround time (prospective enhancement)</li> </ul>	<ul style="list-style-type: none"> <li>▪ Highest costs (\$1.05 per record enriched)</li> </ul>
<b>Syndetic Solutions</b>	<ul style="list-style-type: none"> <li>▪ Lowest costs (\$.50 per record enriched)</li> <li>▪ Choice of 505, enhanced 505, or 9XX</li> <li>▪ Re-enrichment at no charge</li> </ul>	<ul style="list-style-type: none"> <li>▪ Relatively new vendor with no long-term track record</li> <li>▪ Smaller data files, with lower hit rates</li> <li>▪ 2-3 day turnaround time</li> </ul>
<b>MARCIVE</b>	<ul style="list-style-type: none"> <li>▪ Lowest costs (pricing identical to Syndetic Solutions)</li> <li>▪ Choice of 505, enhanced 505, or 9XX</li> <li>▪ Established library service provider</li> </ul>	<ul style="list-style-type: none"> <li>▪ Smaller data files, with lower hit rates</li> <li>▪ Reliant on Syndetic Solutions for data</li> <li>▪ No re-enrichment without additional charge</li> </ul>
<b>OCLC MARS</b>	<ul style="list-style-type: none"> <li>▪ Choice of data from Blackwell's, Syndetic Solutions, or both</li> <li>▪ Pricing identical to use of Blackwell's and/or Syndetic Solutions</li> <li>▪ Choice of 505, enhanced 505, or 9XX</li> <li>▪ Re-enrichment at no charge</li> <li>▪ Established library service provider</li> </ul>	<ul style="list-style-type: none"> <li>▪ 3-5 day turnaround time</li> <li>▪ Reliant on both Blackwell's and Syndetic Solutions for data</li> </ul>

## Detailed vendor profiles

### *Blackwell's Book Services*

Blackwell's Book Services is the largest and most-experienced TOC vendor. TOC data date back to 1992 for U.S./Canadian editions of English language titles, and 1995 for U.K./European editions of English language titles. Blackwell's captures TOC data for materials in its approval and new titles services. The focus is on high distribution monographic titles published by university, scientific, technical, trade, and specialty publishers, where the titles are of interest to academic institutions. Conference proceedings, medical titles on the Brandon-Hill list, as well as the top 5,000 best-selling popular titles are also included. (A list of North American publishers covered by Blackwell's approval plans is available online at <http://www.blackwell.com/shelf/tools/core.htm>. The list of European publishers is on the Web at <http://www.blackwell.com/shelf/tools/coreEUR.htm>.) TOCs that do not contribute to an understanding of works are excluded, e.g., tables of contents from dictionaries, novels, and travel guides. Approximately 800 TOC records are added per week; the total number added per year, ca. 40,000.

Blackwell's offers libraries the option of taking TOC data in the 505 field or in locally defined (9XX) fields. 505s may be basic or enhanced. The price structure remains the same regardless of which option is chosen. When using the 505 field for TOC data, Blackwell's will omit initial articles from titles, or leave them as they appear in the item. If the library chooses to take the data in 9XX fields, Blackwell's offers the option of providing authors' names in both direct and inverted form; they also will provide the authorized form of an author's name in inverted form in a 9XX at no additional charge. A library may also elect to take the data into the 505 initially; Blackwell's will then re-enrich the records later and put the TOC data into a 9XX without additional charge.

With Blackwell's, customers may choose either prospective or retrospective enhancement, or both, with no difference in unit price. Costs are based solely on the number of matches, regardless of the date range of the material being enhanced.

Blackwell's also provides its TOC data to the OCLC MARS service (see below.) They are also in negotiations with Syndetic Solutions to provide a one-stop option for TOC enhancement. The details of this potential partnership are not yet publicly available.

On February 25, a file of 3,590 records, representing our cataloging for the weeks of March 4, 2001, and May 13, 2001, was sent to Blackwell's for an enrichment test. 446, or 12.4% of the records, were enriched with TOC data. Given that CUL creates approximately 100,000 bibliographic records per year, and assuming a variation in the hit rate of +/- 20%, we could reasonably expect to enrich between ca. 10,00 and 15,000 records per year using Blackwell's service.

The strength of Blackwell's lies in its experience, the size and relevance of its files for academic libraries, and the flexibility it offers customers. The promise of rapid turnaround time is also a plus. The primary disadvantage is cost. Blackwell's currently charges \$1.05 per record enhanced, more than double what Syndetic Solutions or MARCIVE charges. These costs have risen considerably over the past few years; the current charges represent an increase of 40%

over less than five years for essentially the same services. It would be prudent to assume that costs could continue to rise over time, even as we recognize the difficulty of predicting the rate of increase.

## ***Syndetic Solutions***

Syndetic Solutions (<http://www.syndetics.com/>) has been providing TOC data since May 1999. Many of the staff at Syndetic are former employees of Blackwell's. Syndetic Solutions reports adding approximately 60,000 TOCs annually for new English-language non-fiction titles published and/or distributed within the US and Canada, and claims a current database of about 110,000 records. Coverage begins with 1997. TOCs that do not contribute to an understanding of the work (e.g., tables of contents from dictionaries, novels, and travel guides) are excluded. Syndetic relies on two primary sources for its data: Ingram Book Company and Booknews. Unlike Blackwell's, Syndetic does not supply a list of publishers covered by its services.

In general, Syndetic Solutions' services are similar to Blackwell's, but there are some important differences. TOC data may be taken in either the 505 field (basic or enhanced), or in locally defined (9XX) fields. If a 505 is chosen, the library may re-enrich the records and put the TOC data into a 9XX later without additional charge. The cost structure is based solely on matches, with no difference vis-à-vis prospective or retrospective enhancement, or the 505/9XX field option. However, Syndetic is not quite as flexible as Blackwell's. The company does not allow libraries to omit initial articles from 505s, and authors' names in both direct and inverted form cannot be taken in a 9XX. In addition, Syndetic does not provide the authorized form of an author's name in inverted form in a 9XX. Finally, the coverage of the two services is not identical. While there is some overlap between the two, the scope of Syndetic Solutions' data is different. The database is smaller and is more heavily weighted toward trade publications.

Like Blackwell's, Syndetic Solutions provides its TOC data to the OCLC MARS service (see below.) As noted above, Syndetic is currently negotiating with Blackwell's to provide a one-stop option for TOC enhancement. Such a partnership would presumably compete directly with the MARS service, though details are not publicly available at this time.

On February 25, the same test file that went to Blackwell's was also sent to Syndetic Solutions. 249, or 7%, of the records were enriched with TOC data from Syndetic's file. Given that CUL creates approximately 100,000 bibliographic records per year, and assuming a variation in the hit rate of +/- 20%, we could reasonably expect to enrich between ca. 5600 and 8400 records per year using Syndetic's service -- considerably fewer than Blackwell's.

With its narrower range of options and smaller files, Syndetic Solutions is in many ways a less desirable choice than Blackwell's. In addition, it is a new firm, without the long-term record of an established vendor like Blackwell's. The company's principal advantage is its lower pricing. However, the same caveat noted above regarding Blackwell's pricing applies here as well.

## ***MARCIVE***



Founded in 1981, MARCIVE (<http://www.marcive.com/HOMEPAGE/WEB1.HTM>) offers a variety of services for its customers, including customized MARC records, retrospective conversion, authorities processing, and bibliographic record enrichment. The MARC enrichment service, which includes enhancing MARC records with tables of contents, added entries for fiction and biography, and summaries, was initiated in the spring of 2000.

MARCIVE's TOC service draws its data from Syndetic Solutions. Service options and costs for enrichment are identical to those offered or incurred when dealing with Syndetic directly. The primary incentive for contracting with MARCIVE lies in combining TOC enrichment with one or more of the vendor's other services, like authorities processing. MARCIVE also promises slightly faster turnaround time for record enhancement (1 working day) than Syndetic Solutions, but only to customers that also use one of its other services. Since CUL is no longer contracting these services out, there is no real advantage to selecting MARCIVE as a supplier of TOC data. Moreover, MARCIVE is completely dependent on Syndetic Solutions for its data. Were Syndetic to withdraw from its agreement with MARCIVE, or go out of business, the service would no longer be available.

## **OCLC MARS**

OCLC's WLN MARC Record Service (MARS) (<http://www.oclc.org/western/products/mars/index.htm>) became part of the OCLC Authority Control Suite with the January 1999 merger of WLN and OCLC. MARS provides a wide range of database preparation and authority control services, including TOC enrichment. A library may use the MARS TOC enrichment service alone, or it may combine that service with other MARS services, like authorities processing.

The MARS service draws its data from both Blackwell's and Syndetic Solutions. Libraries send their files to OCLC, which then coordinates the data transfers from one vendor to another. Typically, the files are run against Blackwell's first and then Syndetic Solutions, since Blackwell's files go back further and are larger, but the customer may specify that the order be reversed. Only records that are not enriched are run against the second vendor's file, so that the library would be charged only for unique hits from each file. OCLC also sorts the library's file of records first to weed out records that usually would not contain TOC data, such as sound recordings and other non-book formats. Both the shuffling between the vendors and the record sorting are handled by OCLC at no additional charge. Turnaround time for handling is 3-5 business days. Charges for enhancement through MARS are the same as when using Blackwell's and Syndetic Solutions' services directly.

A great strength of the MARS service is that it permits customers to use both of the TOC vendors without having to contract with each separately or manage the transfer of data themselves. By allowing records to run first against Syndetic Solutions' database before Blackwell's, MARS allows libraries to cut costs while still getting the fullest possible enrichment. The principal disadvantage is the reliance of the service on the outside suppliers. It is possible that if a partnership between Blackwell's and Syndetic Solutions does emerge, restrictions

imposed because of that deal could force MARS to alter or drop the service.

## VI. Costs and timing of implementation

### Costs

#### *Vendor charges: Prospective enrichment*

Cost estimates vary significantly, depending on the vendor chosen. Assuming a variation of +/- 20% from the test file hit rate, we could expect annual vendor charges for single-source prospective enrichment to fall within the following ranges:

Table 1. Cost estimates for single-source supplier

Vendor/Data supplier	Annual Costs
Blackwell's (directly or through OCLC MARS)	Low range: \$10,500 (10,000 records enriched/year @ \$1.05/record)  High range: \$15,750 (15,000 records enriched/year @ \$1.05/record)
Syndetic Solutions (directly or through OCLC MARS or MARCIVE)	Low range: \$2800 (5600 records enriched/year @ \$.50/record)  High range: \$4200 (8400 records enriched/year @ \$.50/record)

The cost of using the MARS service to get records from both Blackwell's and Syndetic Solutions is more complicated to estimate. Neither OCLC nor any current customers taking data from both suppliers were able to provide information about the percentage of unique records in each vendor's file. Jeff Calcagno of Syndetic Solutions supplied data indicating a roughly 40% overlap between the TOC files from the two vendors. Thus we can estimate that 40% of Syndetic's hits could be subtracted from the sum of the two vendors' combined matches to give a rough total of unique records.

The following example illustrates the formula used to estimate the overall number of records enriched and the costs when using both vendors' data. (The hit rate is only illustrative and not based on our test file results):

Example - 3000 record monthly file

Blackwell's (Standalone) - 1000 total enriched

Syndetic Solutions (Standalone) - 250 total enriched



$1000 + 250 - 100$  (40% of 250) = 1150 unique enriched records.

Since the price difference between Blackwell's and Syndetic Solutions is \$.55/record, the cost would be as follows:

\$1050 (1000@ \$1.05/record (Blackwell's))  
 + 125 (250@ \$.50/record (Syndetic))  
 - 55 (100 non-unique records from Syndetic @ \$.55 less than Blackwell's)=  
 \$1120

Based on these estimates, we might expect annual charges for using both vendors' data via MARS to fall within the ranges in Table 2a below. The hit rates below are based on the results of the test enrichment, again with a variation of +/- 20%. Tables 2b and 2c break down the calculations used for the number of records and the costs in Table 2a.

Table 2a. Cost estimates for dual-source supplier (OCLC MARS)

Vendor	Annual Costs
OCLC MARS (using both Blackwell's/Syndetic Solutions)	Low range: \$12,068 for 13,360 records enriched/year
OCLC MARS (using both Blackwell's/Syndetic Solutions)	High range: \$18,046 for 19,920 records enriched/year

Table 2b. Calculation of records enriched and estimated costs: Low range

Number of records	Vendor	Costs
10,000	Blackwell's (10,000 records/year)	\$10,500 (10,000 @ \$1.05/record)
+ 5600	Syndetic Solutions (5600 records/year)	\$ 2800 (5600 @ \$.50/record)
- 2240	Non-unique records, Syndetic Solutions (40% of 5600)	(\$ 1232) (2240 @ \$.55/record less than Blackwell's)
13,360 unique records		\$12,068

Table 2c. Calculation of records enriched and estimated costs: High range

Number of records	Vendor	Costs
15,000	Blackwell's (15,000 records/year)	\$15,750 (15,000 @ \$1.05/record)
+ 8400	Syndetic Solutions (8400 records/year)	\$ 4200 (8400 @ \$.50/record)
- 3360	Non-unique records, Syndetic Solutions (40% of 8400)	(\$ 1848) (3360 @ \$.55/record less than Blackwell's)
20,040 unique records		\$18,102

### *Vendor charges: Retrospective enrichment*

The cost estimates for single-source retrospective enrichment are given below. It should be noted that because the backfiles for the two suppliers of TOC data vary considerably in size and depth, a true direct comparison between the two is not possible. The figures below assume that the hit rate for past years will be similar to the test file results and again use a variation of +/- 20% in the hit rate from those results.

Table 3. Cost estimates for retrospective enrichment, single-source supplier

Vendor/Data supplier	Costs
Blackwell's (directly or through OCLC MARS)	<p>Low range: \$105,000 for 100,000 records enriched (10,000 records enriched/year @ \$1.05/record X 10 years (1992-2001))</p> <p>High range: \$157,500 for 150,000 records enriched (15,000 records enriched/year @ \$1.05/record X 10 years (1992-2001))</p>

Syndetic Solutions (directly or through OCLC MARS or MARCIVE)	<p>Low range: \$14,000 for 28,000 records enriched (5600 records enriched/year @ \$.50/record X 4 years (1997-2001))</p> <p>High range: \$20,500 for 41,000 records enriched (8200 records/year @ \$.50/record X 5 years (1997-2001))</p>
---	---

Estimating the cost of using the MARS service to get records from both Blackwell's and Syndetic Solutions for retrospective enrichment is complicated. The following table provides an estimate based on the assumption that we would receive only Blackwell's records for 1992-1996, and records from both suppliers for 1997-2001. For those years where the two vendors' files overlap, the formula from Table 2 above is used to calculate total unique hits and costs. Tables 4b and 4c break down the calculations used for the number of records and the costs in Table 4a.

Table 4a. Estimate of costs, retrospective enrichment, dual-source supplier (OCLC MARS)

Vendor	Cost
OCLC MARS (using both Blackwell's/Syndetic Solutions)	Low range: \$112,840 for 116,800 unique records enriched, 1992-2001
OCLC MARS (using both Blackwell's/Syndetic Solutions)	High range: \$169,260 for 175,200 unique records enriched, 1992-2001

Table 4b. Calculation of records enriched and estimated costs: Low range

Number of records	Vendor and Date Range	Costs
100,000	Blackwell's, 1992-2001 (10,000 records/year for 10 years)	\$105,000 (100,000 @ \$1.05/record)
28,000	Syndetic Solutions, 1997-2001 (5600 records/year for 5 years)	\$ 14,000 (28,000 @ \$.50/record)
-11,200	Non-unique records, Syndetic Solutions, 1997-2001 (40% of 28,000)	(\$ 6160) (11,200 @ \$.55/record less than Blackwell's)
116,800 unique records		\$112,840

Table 4c. Calculation of records enriched and estimated costs: High range

Number of records	Vendor and Date Range	Costs
150,000	Blackwell's, 1992-2001 (15,000 records/year for 10 years)	\$157,500 (150,000 @ \$1.05/record)
+ 42,000	Syndetic Solutions, 1997- 2001 (8400 records/year for 5 years)	\$ 21,000 (42,000 @ \$.50/ record)
-16,800	Non-unique records, Syndetic Solutions, 1997- 2001 (40% of 42,000)	(\$ 9240) (16,800 @ \$.55/ record less than Blackwell's)
175,200 unique records		\$169,260

**Staff time**

Regardless of the vendor chosen, staff time costs would be roughly the same. The more significant consideration is prospective or retrospective enrichment. Retrospective enrichment would entail more LTD and DBQ & E staff time than prospective enrichment.

With either option, most of the staff time cost would be upfront. Tasks would include data extract, loading enriched records, and that TOC enrichment process works harmoniously with other automated processes (e.g., Marcadia), and that the contractual obligation not to export vendor-supplied TOC data be met.

Jim LeBlanc estimates that prospective enrichment alone would require approximately 48 FTE academic and programming staff hours (6 days); retrospective enrichment would add another 16 hours (2 days). Once the process is underway, LTD and DBQ & E would seek to automate as much of it as possible, and the ongoing staff commitment would be minimal.

**Other projects/priorities**

Currently, there are seven high-priority projects in the queue for LTD and DBQ & E staff, including retrospective conversion, Maeda records, and Wiley updates. Jim LeBlanc estimates that TOC enrichment could begin sometime over the summer, barring the addition of new high-priority project, unforeseen problems with ongoing projects, or with the TOC enrichment process itself.

**VII. Complementary services**

Several services offering TOC data are currently in place throughout CUL, enhancing access to both monographic and serial items. Though none of the services duplicates what the vendors

covered by this report are offering, all offer enhanced access to materials in one form or another.

### ***Mann Library New Book Shelf***

Mann Library has been scanning the TOC for items on its New Book Shelf for over eight years. The service is accessible through the Mann Library Web site at <http://www.mannlib.cornell.edu/cgi-bin/newbooks.cgi>.

Users may click on the highlighted titles in the citation to access the scanned image of the book's table of contents. The New Book Shelf is both physical and virtual: the list of approximately 20 selected titles is updated every two weeks as the actual books are added to or taken from the shelf as part of the regular workflow. The New Book Shelf is a joint product of Mann Library's Technical Services Division and the Information Technology Section (ITS).

Staff scan the TOCs and then FTPs them to a holding file. ITS staff matches the bibliographic ID number to the scanned image. Access is then available to the new TOCs through the Mann Library New Book Shelf. As the materials are removed from the New Book Shelf, links are created in the catalog records to provide ongoing access to the scanned TOC images via the OPAC.

### ***IndexMaster (Law Library)***

A project currently underway in the Law Library provides catalog users with table of contents information by hyperlinking from within bibliographic records to scanned TOCs supplied by IndexMaster (<http://indexmaster.com/>). The URLs that link to the scanned TOC data are being inserted into the bibliographic records manually, using a master list of titles given to the Law Library by IndexMaster. Clicking through the hyperlink displayed in the OPAC, the user is led (via an intermediate screen) to the scanned TOC in PDF.

The hyperlinked TOCs in .PDF are not searchable via the online catalog. Retrieval is therefore not enhanced via the indexes. While PDFs are available as links through the catalog, the TOC information itself does not actually reside within the bib records or on a locally controlled server. Thus the Law Library is dependent upon the continued existence of IndexMaster and its server for access to the TOCs. Nonetheless, the Law Library decided to take advantage of IndexMaster's TOC service because it covers many more legal titles than services such as Blackwell's. There is very little overlap between the titles provided by IndexMaster and those coming from Blackwell's.

### ***MyContents***

MyContents will become an official part of the MyLibrary suite of services on July 1, 2002. The service is currently in a pilot stage. It is an extension of the service offered by Mann Library for the last several years, in which users sign up to automatically receive TOCs for selected journals

via email. More information about MyContents is available online at <http://mylibrary.cornell.edu:8080/mycontents/enter.do>.

MyContents will allow users to select journal titles for which they would like to automatically receive Tables of Contents (TOC). Data are provided from two sources, Ingenta and ISI, though the system is designed to add new vendors easily. The journal TOC will be delivered to the e-mail address specified by the user. The most current TOC for the selected journal(s) will also be available from the user's MyContents personalized space, if he or she has created one.

In the current pilot project, there is TOC data from ca. 500 journals, primarily in the sciences, though about 20% of the titles are in the social sciences. The MyContents team anticipates that more titles and subject areas will be represented by the July 1 launch.

MyContents will cover only serial titles, not monographs. MyContents will not enhance access to individual titles through the catalog, though it will provide a valuable service to those interested in the contents of particular journal issues.

## **VIII. Recommendations**

Based on the suitability of the services offered by the vendors discussed here, the benefits to users, and the costs of enrichment, the following options are offered for consideration. They are listed in order of potential benefit to library users.

### ***Option 1.***

#### ***Contract with OCLC MARS to get data from both Blackwell's and Syndetic Solutions for prospective and retrospective enrichment***

##### ***Advantages:***

Maximum benefits to users: Electing both retrospective and prospective enhancement would allow us to add TOC data for a ten-year period, greatly enhancing access to thousands of recent English-language titles in many subject areas.

Flexibility: Choosing the OCLC MARS service would allow the library to retain all of the flexibility offered by each vendor. Retrospective enrichment could be done over the course of one or several years while prospective enrichment moves forward. Retrospective enrichment could be done over a number of years. The library could also choose to take retrospective data from one supplier while also opting to do prospective enrichment from both suppliers.

Value: By taking all of the Syndetic Solutions records, and only the unique Blackwell's records, we could reduce costs while still getting the best services from both suppliers.

##### ***Disadvantages:***



**Complexity:** While prospective enrichment would present fewer technical problems for DBQ &E and LTD, retrospective enrichment would require additional programming and staff time. Implementation may therefore take longer than prospective enrichment alone.

**Costs:** Choosing both retrospective and prospective enrichment would entail the highest costs, both in staff time and in vendor charges. It should be noted however, that most of the cost would be a one-time expenditure, unless the retrospective enrichment were spread out over time.

## ***Option 2.***

### ***Contract with OCLC MARS to get data from both Blackwell's and Syndetic Solutions for prospective enrichment only***

#### ***Advantages:***

**Flexibility:** Choosing the OCLC MARS service would allow the library to retain all of the flexibility offered by each vendor.

**Value:** By taking all of the Syndetic Solutions records, and only the unique Blackwell's records, we could contain costs while still getting the best services from both suppliers.

**Relative ease of implementation:** Programming work to extract records for export to OCLC WorldCat is already complete and would require only minor tweaking to be used for prospective enhancement. Implementation would be simpler and may move forward more quickly.

**Lower costs:** Without adding TOC data from the backfiles from each vendor, we would reduce costs of enrichment significantly.

#### ***Disadvantages:***

**Reduced benefits for users:** With fewer enriched records, users would not realize the same level of benefit as they would from both retrospective and prospective enrichment.

## ***Option 3.***

### ***Contract with Blackwell's exclusively for prospective and/or retrospective enrichment***

#### ***Advantages:***

**Simplicity:** Choosing the Blackwell's service alone would allow the library to retain all of the benefits the vendor offers (flexibility, higher hit rate), without having to worry about the relationship between Blackwell's and third parties like OCLC.



Turnaround time: Turnaround time would be faster with Blackwell's than using the MARS service; Blackwell's turnaround time for current customer (using prospective enrichment) is 24 hours.

Potential partnership with Syndetic Solutions: The possibility that Blackwell's may itself partner with Syndetic Solutions means we may enjoy the same benefits by working with Blackwell's directly as we would by working with OCLC. There may even be additional benefits like faster turnaround. However, the exact nature of this partnership and its potential advantages are at present uncertain and may never even happen.

***Disadvantages:***

Reduced benefits for users: Without the supplemental records from Syndetic Solutions, fewer records would be enriched, and users would not realize the same level of benefit as they would if we elected to use the data from both vendors.

Cost: By taking all data exclusively from Blackwell's, rather than only its unique records, the Library would incur Blackwell's higher costs for every record enriched.

3/21/02--db

---

[StaffWeb Home](#)

[Library Gateway](#)

<http://www.library.cornell.edu/staffweb/TOC.html>