

Technology Challenges and Change at NREL¹ Library

Rob Finger, Carol M. Fitzgerald, and Iris Martinez²

Employees of the Midwest Research Institute under Contract No. DE AC36 99GO1 0337 with the U. S. Dept. of Energy have authored this work. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non exclusive, paid up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for United States Government purposes.

The National Renewable Energy Laboratory (NREL) is the nation's primary laboratory for renewable energy and energy efficiency research and development. The Laboratory's mission is to develop renewable energy and energy efficiency technologies and practices, advance related science and engineering, and transfer knowledge and innovations to address the nation's energy and environmental goals.³ The NREL Library plays an important role in supporting this mission by providing library services and information resources to scientific, technical, and administrative staff at the Laboratory.

In 2001 and again in 2002, the NREL Library unsuccessfully attempted to secure funding to replace its integrated library system (ILS), then verging on obsolescence. We spent another year testing and rejecting our ILS vendor's new SQL product, a frustrating exercise that left us determined to find a more creative solution. By late 2004, we had come to the conclusion that we needed a new strategy, both financially and technologically, to ensure that we could continue to provide a high level of service to our clients.

Realizing that without the aid of supplemental funding we would have to work within our annual budget to achieve our goals, we assembled a team and began to explore our options. Our initial plan was to seek out vendors willing to provide us with some form of creative financing, allowing us to spread out payments for a new system over several years instead of the usual lump sum payment upon delivery. This proved to be more challenging than we had anticipated, and the list of potential vendors from which we could choose our new system was quickly reduced to a short list of three. Based on the responsiveness of the three vendors' sales representatives, we eliminated one more vendor and thus narrowed our options to just two companies, Dynix and Sirsi. This was not an unhappy prospect since they ranked first and second largest ILS vendors, respectively, at the time.⁴

While both Sirsi and Dynix said they would be willing to work with us on a payment plan, they also offered another alternative that appealed to us from a financial perspective: the application service provider or ASP model. The ASP model is essentially the outsourcing of all hardware

and software requirements normally associated with managing an in-house ILS. The vendor purchases and maintains the hardware, provides all software and upgrades, and keeps the system up and running so that library staff can focus on managing the content and providing service to customers. Access to data is via an Internet connection to the vendor's servers which can be geographically distant from library users. This model can dramatically reduce up-front expenses without a significant increase in yearly maintenance costs.

Potential trade-offs with the ASP model are a loss of control over (or access to) data, and the potential for downtime due to a loss of Internet connectivity. Both caused some initial reluctance on the part of Library staff. Our greatest concern was the potential for barriers to interoperability with third party applications. We foresaw the possibility that using an ASP would make us far too dependent on our vendor if we wanted to incorporate additional features at some point in the future, such as consolidated searching or a link resolver. Would we be able to purchase and use third party applications with our ASP database or be forced to go with the vendor's products? Another question that arose was the capability to customize the online catalog and whether ASP vendors would provide Web development server space for testing such changes (Dynix did not and Sirsi only would for a price we could not afford). With all of these issues in mind, and having done preliminary research into Sirsi Unicorn and Dynix Horizon that included contacting library customers of both, in addition to comparing many online library catalogs, we scheduled demonstrations with sales representatives from both ILS vendors.

We invited the entire Library staff to both demonstrations, as well as two members of the NREL Systems and Network Infrastructure group to provide their technical expertise on the different solutions being offered. Prepared with a five-page Microsoft Excel spreadsheet detailing everything we were looking for in a new ILS, we entered the demonstrations and began to take notes.

Both vendor representatives addressed questions from our attendees and presented their ASP systems “live,” then discussed their previously submitted quotes in depth. We followed up the demos with a trip to view a working ASP system at a Denver area library where we were reassured concerning the dependability of their ASP’s Internet connectivity and real-world usability. That trip, combined with our Systems and Network advisors’ endorsement of the ASP option, turned out to be all the convincing we needed. While it was not what we had originally set out to purchase, when confronted with the imminent obsolescence of our ILS and increasingly tighter budgets, we discovered that current ASP vendors might well be what Marshall Breeding described as “better able to operate an industrial-strength data center and to maintain server hardware, operating systems, and application software.”⁵ Choosing the ASP alternative allowed us to combine our need for a technologically advanced system with our requirement to keep costs within a particularly tight budget.

The ILS team met with the entire Library staff one final time to review everything we had seen in the demonstrations and to make sure we had consensus on a final decision between vendors. Each member of the staff was asked to participate in the discussion. With the decision made, we notified the sales representatives that we had chosen the Dynix Horizon ASP option. Within days, Sirsi and Dynix announced that they had merged. The news probably had less impact for us than for longtime users of either vendor’s system. The general expectation here has been that we now have the support of a doubly viable ILS supplier.

Once the contract details were completed and the agreement signed, we faced the task of migrating our data. ETL (extraction, transformation, and loading) of data is a challenge with any migration from one ILS to another, but ours presented a few more challenges than most. Several factors contributed to our taking on the extraction and transformation of our system data ourselves: no exit provision had been written in the contract with our old ILS vendor, we had cancelled our maintenance contract the previous year, the company did not respond to our request for an extraction services quote, and most importantly, outsourcing the process would have put us beyond our budget. Both of our proposed vendors had informed us early in the process that we would need to deliver our data in MARC format but they would work with tab delimited files for a portion of the non-bibliographic data. We eventually sent our core bibliographic data to SirsiDynix in MARC format, including authority cross-references. Other data, such as circulation, borrower, and serials inventory records were ftp’ed as tab delimited files.

Our old ILS, which was not in MARC format, did have an export utility for mapping fields to MARC tags. However, the utility only provided the capability to export data into MARC format with one subfield per tag and no filing indicators. With close to 34,000 bibliographic records which extracted into a 27 megabyte MARC file, we needed a robust file editor that could handle batch changes and saves. For example, many of our subject headings contained general (\$x), geographical (\$z) or genre (\$v) subdivisions and practically every MARC tag of every record needed multiple subdivisions identified and properly tagged. An open source product, MarcEdit (version 4.6.63; last modified December 22, 2004)^{6,7} proved to be an efficient and flexible software solution for us. The software comprises a range of programs, including MarcEditor Batch Editing Tools. Not only does the MarcEditor allow global search and replace insertion of subfield tags, it also has the flexibility to prepend data within a subfield. The global search and replace tool also allowed us to change location or collection codes to match our new system’s design. Although we could not edit record by record due to our very short migration and training schedule, we were able to quickly and efficiently insert many subfield tags and edit other data portions to create an acceptable core database for migration.

Despite not having MARC in our previous system, we had assiduously maintained AACR2 standards of punctuation, spacing, and format within our catalog records. Thus our search and replace operations could be easily based on that standard with confidence. For example, to insert the “statement of responsibility” subfield codes in main titles, we were able to search the 245 tag for *[space] / [space]* and replace that with *[space] / [space] \$c*:

```
=245 10 $aPractical ideas for the design, operation,  
and maintenance of plant energy systems / $cThomas  
C. Elliott, editor.
```

Because all of our subject subheadings were separated by double dashes, we were able to replace those with subfield “x.” We then re-coded subfield “v” and subfield “z” (our most common variation) subdivisions by searching for *\$x [term]* and replacing both subfield code and term:

```
=650 \0 $aWind power $z California $v Atlases.
```

We continue to use MarcEdit even now after the migration is completed. We periodically use the MarcEdit software to search the extraction file sent to SirsiDynix while fine-tuning data configurations in our new system. With all of its flexibility, however, MarcEdit could not help us tag all subfields correctly. We continue to work with our bib records within the Horizon system to bring them to full MARC format.

One unforeseen drawback of migrating a non-MARC database has been the necessity of editing all of our subject records to correct errors in the 000 (leader) and 008 tags. Because we had nothing equivalent to these tags in our previous database and no experience with their coding, we were unaware that they loaded into a template not properly coded for those records. We did not catch the problem in the testing phase or before we went live. It was not until we were creating new records in Horizon, selecting and inserting existing subject terms from our new authority files, that the erroneously coded terms caused noticeable problems. We have not been able to devise a way to use MarcEdit for this and have therefore begun a project to correct the codes with Horizon's editing functions.

Utility programs supplied with MarcEdit include a Delimited Text Translator which is designed to help users import delimited data into MarcEdit's readable, or "mnemonic" as it is called in MarcEdit, MARC file format with a .mrk file extension. This feature allowed us to create MARC formatted files of our authority cross-references even though we could not extract that data in MARC format from our ILS. The authority cross-references (*see* and *see also* records for subjects, authors, corporate authors, and series) could only be exported to flat files, which we first edited in Microsoft Excel, saved as tab-delimited files and, finally, pulled into MarcEdit where we applied the MARC tags. A variety of conversion tools have also been written for MarcEdit, including MarkMaker, which converts the mnemonic file format to MARC format file with a .mrc extension, or the machine-readable format.

Our vendor accepted tab delimited files for other data we could not extract as MARC: circulation, borrower, and serials inventory records. Because our ILS was not a relational database we were limited to reports that could pull data from across modules into a single file. By comparing many reports, we were able to select the most appropriate for the data we needed, edit formats where possible to add fields, and produce flat files for both circulation and serials inventory. The easiest way to pull borrower data, however, was to request that NREL's online phonebook of employee contact information be sent to us as a tab delimited file. This provided us with the most updated list of employees and included information such as email addresses and full telephone numbers that we had not been keeping in our previous borrower records, all data that we would have spent many hours adding to any file we pulled from our old system and which we now wanted to include in our new system.

We then massaged the data in MS Excel, performing many search and replace operations to change codes

to our new system's profiles. We had decided early on to make changes in our collection codes, sublocations of materials, and other organizational designations for our collections and borrowers. This required that we translate those changes in all files prior to transmitting them to our new vendor. We then saved the files in tab delimited format.

The migration process of extraction, transformation, and loading covered a three-month time period, including the loading of a test database, during which our library staff attended many training sessions on the Horizon system they would soon be using. NREL Library went live with its new Horizon ASP on November 18, 2005.

In a Washington Post article⁸ published November 30, 2005, Steven Pearlstein noted that Nicholas Carr, former executive editor of the Harvard Business Review, is proposing the next model for IT: utility-scale computing services. According to Pearlstein, Carr's model reflects early twentieth century manufacturers dismantling their own electric power generating stations and turning instead to centralized electrical power providers as an evolution of "infrastructural technologies."⁹ Carr describes the similarities between early adopters of centralized electrical utility services and today's computing outsourcers:

"When businesses began to turn to utilities for their electricity supply, smaller organizations led the way. Lacking the cash to build their own power plants, they had little choice but to buy power from outside suppliers. The most aggressive early adopters of utility computing also have tended to be capital-constrained organizations: small and medium-sized businesses, government agencies and nonprofits."¹⁰

The National Renewable Energy Laboratory fits neatly into the above description of early adopters as a not-for-profit contractor to the U. S. Department of Energy. As with most other libraries, we often face funding constraints, yet our mission is to support leading edge research. The ASP option, what Carr refers to as "utility computing," allowed us to migrate to a standards-based system with long-term viability developed by a company with a strong track record in the industry. The further commitment of SirsiDynix to continue development of the Corinthian product, designed for academic and research libraries, is also welcome news to us. We plan to use this next generation system when it becomes available to ASP customers to help us maintain a high standard of technological advancement well into the future.

NOTES:

¹NREL is a national laboratory of the U. S. Department of Energy, operated by Midwest Research Institute and Battelle,

and located in Golden, Colorado.

²Rob Finger is a Sr. Information Specialist in Publications, Information Resources, at NREL and an MLIS student at the University of Denver. Carol M. Fitzgerald is an Information Scientist and Iris Martinez is a Sr. Information Specialist, both in the Library, Information Resources, NREL.

³National Renewable Energy Laboratory. "NREL Overview." <<http://www.nrel.gov/overview/>> (accessed 19 January 2006).

⁴"ILS Vendors Sirsi and Dynix to Merge," *LibraryJournal.com* (June 21, 2005) <<http://www.libraryjournal.com/article/CA609704.html>> (accessed 19 January 2006).

⁵Marshall Breeding, "The Trend Toward Outsourcing the ILS: Recognizing the Benefits of Shared Systems," *Computers in Libraries* May (2004): 36-38.

⁶Terry Reese, "Information Professionals Stay Free in the MARC Edit Metadata Suite," *Computers in Libraries*, September (2004): 24-28.

⁷Terry Reese, "MarcEdit Homepage: Your Complete Free MARC Software." Oregon State University. <<http://oregonstate.edu/~reese/marcedit/html/index.html>> (accessed 29 December 2005).

⁸Steven Pearlstein, "Economy of Scale Might Inspire Companies to Ditch IT Departments," *Washington Post*, 30 November 2005. <<http://www.washingtonpost.com/wp-dyn/content/article/2005/11/29/AR2005112901096.html>> (accessed 4 January 2006).

⁹Nicholas G. Carr, *Does IT Matter? Information Technology and the Corrosion of Competitive Advantage*. (Boston: Harvard Business School Press, 2004), p.20-28.

¹⁰Nicholas G. Carr, "The End of Corporate Computing," *MIT Sloan Management Review* Spring (2005): 67-73.

Rob Finger, Carol M. Fitzgerald, and Iris Martinez are all on the staff of the NREL Library (see footnote 2 above).

CAL