



# Implementing a Discovery Layer

## A Rookie's Season

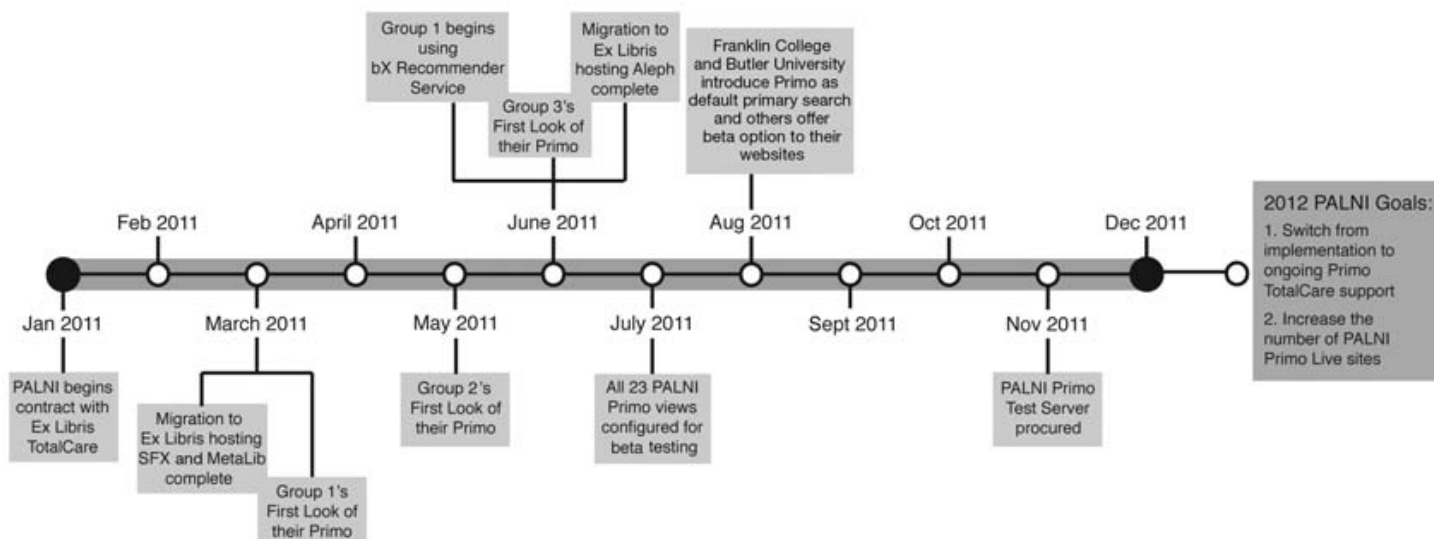
BY NOAH BRUBAKER, SUSAN LEACH-MURRAY, AND SHERRI PARKER

The year 2011 quickly moved us from rookies to more experienced racers in working with our discovery engine and all its supporting components.

It is now 2012, and it seems so long ago that we sponsored our 2010 Discovery Layer 500 race to choose a new search interface for the PALNI (Private Academic Library Network of Indiana) consortium. It would take another year to complete the task. In fact, 2011 was the consortium's "rookie season" for the implementation of Primo. In this article, we will report on our transition to the cloud within Ex Libris Ltd.'s Primo TotalCare environment: our preparation, the steps involved to move to live production, and our reflections on our consortium's first season.

»

## PALNI'S ROOKIE SEASON TIMELINE



Implementing the discovery layer for PALNI was a yearlong effort.

During the 2010 Discovery Layer 500 race, a PALNI Technology Advisory Group (PTAG) was in the driver's seat, evaluating all major discovery layers to make a final recommendation for the winner. This group looked equally at the following discovery layers: Ex Libris' Primo; Innovative Interfaces, Inc.'s Encore Synergy; OCLC's WorldCat Local; EBSCO Publishing's EBSCO Discovery Service; and Serials Solutions' Summon. There were many factors involved with our analysis of these systems: the impact on traditional federated searching, utilization of FRBR, whether or not the vendor was a content provider, weighting (relevancy ranking) of results, and how the discovery layers would work with our Aleph ILS and data. Yet, PTAG managed to identify major attributes that were important to the consortium and individual institutions. Some of these major attributes were resource availability and call number in results, content neutrality, peer-review filters, the presence of a mega-aggregate cloud, and full-text filters. The complete coverage of our experience in choosing Primo as our 2010 Discovery Layer 500 race winner for PALNI can be found in the article, "Shapes in the

Cloud: Finding the Right Discovery Layer," in the March/April 2011 issue of the journal *ONLINE: Exploring Technology & Resources for Information Professionals*.

### Gearing Up for Our Rookie Season

**Contract negotiations.** Before our season began, we had to negotiate our contract with Ex Libris, which began in October 2010. The PALNI consortium of 23 libraries decided to subscribe to Primo TotalCare, and because our existing support contract was nearing renewal for Ex Libris' SFX, Aleph, and MetaLib, we chose to migrate these services to the TotalCare support environment as well. Ex Libris TotalCare is a full-service subscription that provides hosting and support for Ex Libris customers. This service is an annual support contract, but some libraries sign for multiple years. Deciding to move our services to the cloud environment provided new opportunities for PALNI, but it also greatly impacted our Primo implementation timeline and introduced new challenges, considering that we were moving 23 different schools from one service provider to another. In addi-

tion, because we were migrating to a new vendor using a software-as-a-service model, data security and the ownership of responsibility in case of a data breach for Aleph patron data became the most time-consuming negotiating factor.

**Priming the Primo project.** As with all rookies, many of those in the consortium (directors, instruction librarians, and systems librarians) felt pressures due to our inexperience of working with any discovery layer; however, we were excited and ready to face our new challenges. To prepare for our 2011 Primo implementation season, a crew chief was designated for each PALNI-supported organization (usually the person chosen was responsible for Aleph systems administration). These individuals were invited to a TotalCare kickoff meeting in January, where Ex Libris demoed Primo once again and made time for questions, provided migration and implementation timelines, and outlined the kind of service and support we should expect after the migrations were complete. A plan was laid out for the 23 institutions to be divided into several implementation groups. Group 1 was created to include six institutions that would be diverse in size and terminal degrees offered (we



## WWW.PALNI.ORG

There are 23 libraries in the PALNI consortium.

needed to ensure seminary and music degrees were represented). The idea was that this group would “sign off” on implementation prior to subsequent group implementations; however, this plan did not work out due to the complexity of our consortial implementation.

During the kickoff meeting, in order to initiate Primo project communication, Ex Libris introduced a newly created listserv that included each crew chief and the Ex Libris Primo project team. The intention of the listserv was to ensure that everyone was informed about new developments and impending deadlines. It was also meant to help raise awareness about reported issues or product implementation choices among schools. Ex Libris stressed that communication should stay to the listserv as much as possible, so it could be easily seen by everyone at all steps of the process—in a consortium such as ours, almost every issue identified by one library could or would affect others. Unfortunately, the idea of this list encompassing all communication became an issue, and the use of it had to be tweaked during implementation.

**Project communication goes critical:** Managing and using effective

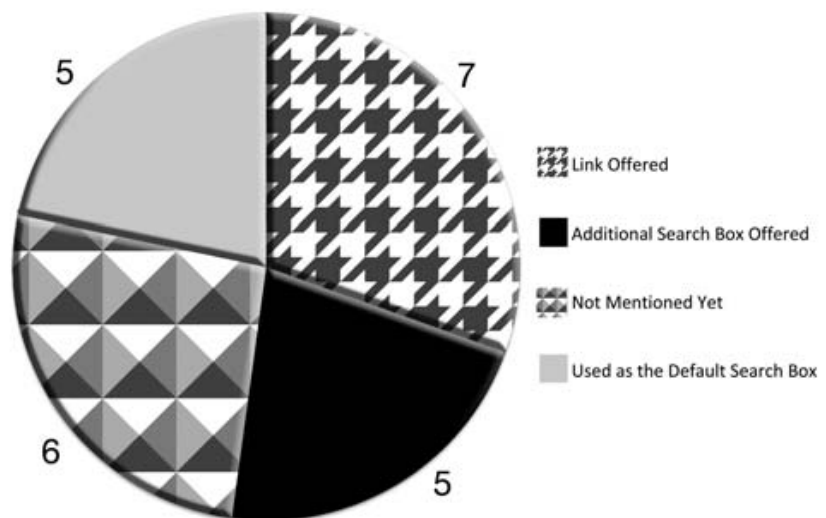
communication processes during this project quickly became essential for our success, but they also became points of contention. We attempted many forms and combinations of communication before we arrived at what we use today. While we tried to manage all communication via the Ex Libris listserv, we found that approach to be very cumbersome in addressing the broad information requests, specific questions, and problem reports originating from each institution. It did not take long for Ex Libris to realize that there were too many posts being sent to this single listserv for its project team to be efficient in progressing with its project responsibilities. In addition, we soon realized that there were posts that really only pertained to us as a consortium, those that really did not need the vendor’s attention. We remedied that by creating an internal listserv, separate from the vendor’s discussion list, and we limited the vendor’s list communication to be used primarily for vendor communication to us as a group. We then created a shared Google Docs spreadsheet, in order to post Primo problems to Ex Libris in lieu of the listserv. At that point, we could add our institution’s name to any

posting to indicate the degree to which an issue was global.

Eventually, Ex Libris made its support ticketing system available to us. In retrospect, we think access to this system should have been provided sooner. While the Google Docs spreadsheet did serve us well to coordinate all Primo questions and issues as well as project priorities, we think that it could have been eliminated if we had access to the support ticketing system earlier. Within the Ex Libris support ticketing system, we identified two types of support incidents: global and individual institution tickets. We also discovered a great way to enhance consortiumwide ticket communication by adding our consortium listserv email address to the additional contact email address. Then, all listserv recipients received ticket updates for that support incident. This discovery has and continues to serve us well.

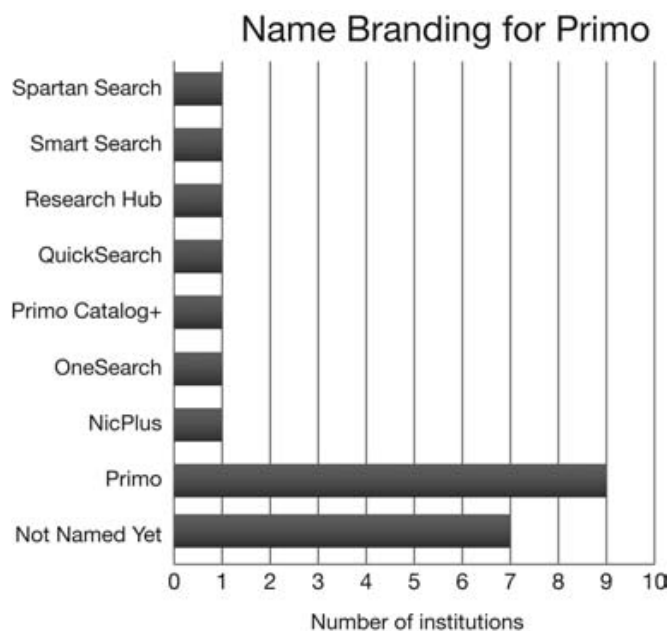
PALNI currently has in place a regular meeting schedule between Ex Libris, the PTAG steering group, and the chair of the cataloging advisory group. The PTAG steering group consists of three crew chiefs plus the consortium executive director. These meetings allow our vendor to communicate progress

How Primo Is Introduced on Library Websites for 23 PALNI Institutions  
Jan. 15, 2012



The Primo discovery layer is not yet deployed across all PALNI libraries. Here's how it was deployed in January 2012.





*This is how Primo is branded by the libraries in the network. The chart shows the number of libraries in the consortium that call Primo by each name.*

related to implementation and problem resolution, providing general question/answer sessions for our consortium and handling questions cropping up over the prior week. These meetings also allow us to prioritize our project progress and help influence how Ex Libris chooses to provide resources devoted to solving our particular problems.

One of the primary communication methods PALNI has made use of is a wiki. Learning the wiki interface did require some time and explanation. But at this point, most of us are actively engaged in consuming information from the wiki, and many are updating and contributing to it. With so many individuals at various geographic sites and with different skill sets, we think the wiki tool has proved very useful in creating a go-to place for project-related information.

In addition to the wiki, the internal consortium listserv is frequently used to get quick answers to questions not addressed via the wiki, to help resolve system problems or outages, and to quickly poll members for decisions. Some listserv topics end up as wiki pages. All institutions have access to the listserv and can post new topics or replies to questions, allowing for those with answers to

or follow-up questions about a topic to quickly get information exchanged.

A final aspect to our project communication revolved around disseminating information about what options are available during implementation and beyond. We disseminated information to our supported organizations by offering presentations that covered topics such as overviewing Group 1 display options, streamlining the SFX menu options, demoing direct linking via SFX, providing an overview of experiences with the discovery APIs, and explaining batch loading of MARC records to be indexed by Primo. We realize that the complexity of current integrated library systems requires continual education for librarians, in order to understand and envision all possibilities for system functionalities, so we will need to continue to provide learning opportunities.

### Setting Up and Tuning Our Four-Cylinder Primo Engine

PALNI contracted with Ex Libris to configure and set up Primo to include integration from a supported ILS (Aleph and Unicorn), OAI-compliant digital repositories (CONTENTdm and

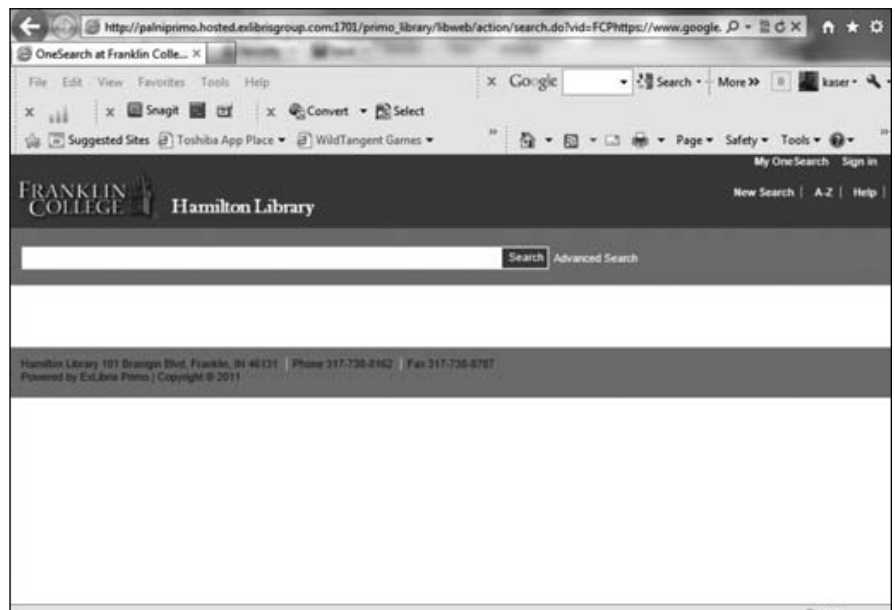
Digital Commons), library website pages, an existing link resolver (SFX), a federated search system (MetaLib), bX Recommender service, and Primo Central. While all but one of our supported organizations use Aleph, this integration has been by far the most time-consuming to configure due to how Aleph was set up for our consortium. Originally, we contracted for a single Primo installation, with multiple institutional views. However, about 10 months into the project, we realized we needed a second Primo installation that could serve as a test environment. With 23 institutions in different phases of implementation, we discovered that it was not in our best interest to implement changes that adversely impacted live Primo views. The lesson learned here was that we should have procured a test installation much earlier than we did. This would have allowed us to try things in small batches on one server without impacting the production server; then we could compare things, note specific problems, and view proposed changes.

**The ILS piston.** All but one of the PALNI-supported organizations share the Aleph ILS, and the other institution hosts its own instance of Unicorn. The PALNI consortium's Aleph system architecture was set up with both a consortium-shared bibliographic table and a holdings record table, but there are individual institutional item record tables. Our patron records are also shared. Our crew chiefs have learned more about Aleph than ever before with our Primo implementation, but we have definitely experienced some stalled engines and piston misfires! The biggest challenge for us has been that individual institutions really only want to see their own Aleph information within Primo, and they do not desire to see other members' holdings and call numbers. Ex Libris has spent an enormous amount of time working hard to separate our Aleph data within Primo, because the Aleph architecture was initially set up for a shared environment. In addition, we have discovered that numerous Aleph

configurations were not set up properly or that they required data cleanup that had been overlooked previously.

Other considerations for Aleph data have related to how Primo handles deduplication and FRBR. By default, Primo shows preference to online holdings. Again, because our holdings are shared, this became problematic when some institutions owned only a print version while others owned electronic copies of the same resource. It took us a while to figure out that the best plan of action was to turn off the deduplication feature across the board. By changing this configuration, print records were no longer combined with electronic records, alleviating this issue within our shared record environment. Also, after much thought and deliberation, we decided to also turn off FRBR-ization within Primo for Aleph data. There were just too many issues with the way records were being combined within the Aleph data, and we felt it was easier to deal with this later, when FRBR could be tested and evaluated format by format on the test server. Ex Libris still needs to enhance the default record within Primo that is displayed so that it represents the work, without call numbers, manifestations, etc. Currently, it is misleading to the user when a specific record's information and format are represented with "More Versions" available. We feel it is likely Ex Libris will address a solution to this globally in the near future.

One of the other challenges related to Aleph data has been in relation to refining the resource type facet options. What complicates this is the fact that Primo harvests metadata from multiple sources (ILS, digital repositories, etc.). By default, Primo combines audiovisual into a single resource type facet. For our school of music students, this does not suffice. Ex Libris is working with us to offer a more detailed list; however, we may lose sources when facets appear that only represent one data source. We are still working on our test server for this change and feel confident we will come up with a solution that will work for our users. Another item on our to-do



*A simple page for an elegant search solution—how one library implemented Primo*

list is to harvest the Aleph reserves database(s) for those institutions that use Aleph's reserve functionality.

Our consortium is currently trying to decide what to do about content that is not currently included in the ILS or the Primo Central index. For example, work is being done to load ebook titles from databases into our local ILS via a batch load so the content can be harvested into Primo. This allows those titles to be found. Before, they were not cataloged and not included in the Primo Central index. We have questioned how much time and effort we should put into batch loading content that is not included in the Primo Central index. How long is it OK to wait to see if it will be included? How else will our patrons be aware of this subscribed content if we do not add it to the ILS?

#### **The digital repository piston.**

Our harvests of metadata from CONTENTdm have gone fairly smoothly, but there were two slipped connecting rods. First, by default, OCLC's OAI harvester does not separate subject facets into single XML entries for fields that provide multiple metadata items separated by semicolons. For example, one Dublin Core subject field may include several subject terms separated by semicolons (e.g., Methodist church;

clergy; alumni). By default, Primo received this metadata as a single subject because OCLC does not split it up into multiple subject fields in its harvest. So, Ex Libris had to set up a rule in Primo to split this metadata out at the semicolons.

Our second issue has been related to OCLC providing multiple OAI harvester URLs. So far, when things have not been harvested completely, OCLC reports we need to use a different harvester URL address; we are up to three different URLs now. So, we are waiting to hear from OCLC, which is hosting CONTENTdm for PALNI, on which address is the appropriate one to count on more permanently.

#### **The Primo Central piston.**

Primo Central is the name of the central repository of content hosted by Ex Libris. Each institution has the ability to activate which sources it wants included as part of its Primo Central integration into Primo. We did have issues with relevancy ranking of results at the beginning of our implementations; however, Ex Libris made great strides in improving the algorithms.

Our biggest challenge regarding Primo Central has been in the management of librarians' expectations on what makes discovery systems such as



As our discovery layer project begins to move from implementation to production, we believe our efforts will have a major positive impact on our patrons.

Primo different from anything we have ever dealt with in the past. We had to stress not to lose sight of the fact that this new “race car,” which is craved by the undergraduate students, is a very different engine from anything we have had in the past. We had to emphasize that these mammoth discovery systems are really attempting to be giant indexes of every possible piece of scholarly content out there—aka our own library “Google.”

When we had conversations about relevancy ranking, we also had to keep sight of the magnitude of what was going to be included in the discovery systems and all the different types of searches that relevancy-ranking rules have to take into consideration—not to mention the complexity of what it takes for vendors to get it right. When users go to a database, such as EconLit, they know the subjects of economics, business, and such are covered well, so they cannot necessarily expect that they can compare an initial Primo hit list in the same light when it comes to relevancy ranking. And, when we search Shakespeare in a system such as Primo, we have to respect that there are scientists whose last names are Shakespeare and that scientists do seem to quote William Shakespeare more often than some of us may have realized. While in our dream world, librarians might expect that systems such as Primo would recognize the subject of a specific set of search terms and adjust the relevancy accordingly; these systems are in their infancy. We do have hope that over time, they will get “smarter.” Actually, Ex Libris has recently announced ScholarRank, a relevance-ranking algorithm project that will factor in new capabilities when providing search results, such as the user’s background

and information requirements at the point of need (e.g., college major), citation information, and the number of click-throughs to full text from the bX Recommender service ([www.exlibrisgroup.com/category/NewsletterJanuary2012](http://www.exlibrisgroup.com/category/NewsletterJanuary2012)).

In addition, just after we subscribed to Primo, EBSCO decided to throw a wrench in our crankshaft—it decided to no longer allow Ex Libris to harvest its metadata, instead requiring the use of its EBSCOhost API. This gets confusing for some in that, even though EBSCO pulled its citations out of Primo Central, the majority of the citations it offers were included because the citations either come direct via publishers or from competing vendors that do not use a “walled garden” approach to content. When librarians saw other vendor names such as Gale, part of Cengage Learning; JSTOR; Project MUSE; etc., and no longer saw EBSCOhost as a result facet, they thought the content covered by EBSCOhost was not included in Primo Central; however, the majority of citations were indeed included in the results. Conversations regarding this and relevancy-ranking expectations had to occur numerous times during implementation.

**The federated search piston.** Primo allows for the inclusion of external remote search options in addition to the vendor central repository (Primo Central) and locally harvested content. One option integrates federated search results via the discovery layer interface. The majority of the institutions chose not to pursue this option. The federated search was much slower in returning results than the central database, and the results were separate from other searches. One

caveat to this was that, in order to use the Primo-embedded A–Z database list within the discovery interface, at least one MetaLib-federated search scope must be enabled as a search option. A few schools have used a federated search scope option for WorldCat in order to also gain access to the Primo-embedded A–Z database list functionality. Some seminaries have also used the federated search option for ATLA, which is not yet included in the Primo Central index.

The second option for external remote searching was the development of an EBSCOhost API that would allow direct searching of EBSCOhost databases. Many of the institutions were eager for this development to become available, as citation metadata previously not shared by EBSCOhost would likely become less of a perceived issue with the ability to search those databases directly. Unfortunately, EBSCOhost restricted the developed API. It did not include any faceting information and had a recommended limit of 15 databases (as of August 2011), and if results were blended with the rest of the Primo content, then selecting any facet would remove all EBSCOhost content. For these reasons, most, if not all, schools have forgone this API approach.

**The link resolver spark plugs.** Some institutions have chosen to have their SFX holdings harvested by the discovery layer to offer an effective journal title search. Configuring and harvesting this information was a straightforward process that added one more search capability to our system. We do need to state, though, that bad parses occasionally occur when trying to parse from one vendor to another via our link resolver. The fact that the same citation may be cited numerous ways by various vendors is a concern when links fail as users attempt to access full text. Additionally, source vendors provide coverage to a journal run, but certain articles are omitted due to copyright, and there is currently no way for link resolvers to handle missing articles in thresholds.

## Reflections on Our Rookie Season

The year 2011 quickly moved us from rookies to more experienced racers in working with our discovery engine and all its supporting components. Because we were working to implement 23 institutions, finding the right mix of communication tools was essential. While we attempted to divide into implementation groups, there were too many factors that impacted all PALNI-supported organizations. What ended up happening is the PTAG steering committee members and the cataloging advisory group chair acted as the project managers for the entire consortium. They met with the vendor regularly, worked to recommend priorities, posted global tickets, surveyed crew chiefs, and coordinated feedback to the vendor. Due to the consortiumwide nature of decisions, the in-depth technical nature of the topics, and the large amount of communication, it was hard to keep the crew chiefs informed without overwhelming them. This project manager group attempted to do what it could to address these concerns. Regarding interface customizations, we had to remind ourselves time and again who Primo's audience was and that Primo is an interface for the patron and that it is not designed to cover every librarian's want and need. There is a difference between a librarian's search option expectations and what a typical undergraduate student might want to see. It has taken longer for a few of the reference and instruction librarians to embrace this radically different, patron-oriented search tool. We can foresee that there is a need for ongoing conversations regarding what makes Primo different from search tools we have encountered before.

At the end of our rookie season, the majority of the institutions have at least posted a Primo search box or link on their library webpages. Currently, only five of the 23 institutions are using Primo as their primary default search tool. We also still have a few institutions that are not ready to start

their discovery engines. It is easy to look back and think about how we could have handled things differently. Due to the amount of time it took to separate the display of our Aleph data, we could have instead discussed dividing our Aleph tables by institution prior to Primo implementation.

We have to state that while 2011 was a very taxing year on our crew chiefs, close bonds have been created among fellow consortium members. Support from Ex Libris was continuous through the project so far, and we believe it has gone above and beyond our expectations in some cases. We feel that everyone has been lifted to a higher level of system understanding, and this education cannot be understated. With the willingness to help each other, we found that in 2011, PALNI has been successful in finding solutions to our unique challenges at the consortium level as well as at the institutional level. So as our discovery layer project begins to move from implementation to production, we believe our efforts will have a major positive impact on our patrons.

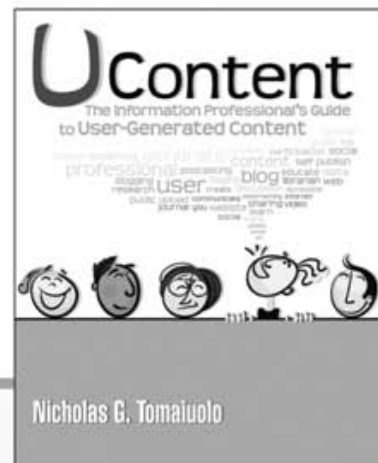


**Noah Brubaker, M.L.S.** (*nbrubaker@uindy.edu*), is the library technology systems manager at the University of Indianapolis. He coordinates and implements technology initiatives for the Krannert Memorial Library and is currently serving as chair of the PALNI Technology Advisory Group.

**Susan Leach-Murray, M.L.S.** (*SLeach-Murray@franklincollege.edu*), is the technology/technical services librarian at Franklin College. While her primary duties are cataloging and library systems administrator, she is also the Primo project manager for Hamilton Library.

**Sherri Parker, M.L.S.** (*Sherri.Parker@harrison.edu*), is now the dean of learning resource centers at Harrison College. Her previous work as coordinator of library technology at DePauw University Libraries included system administration for the Ex Libris suite of products. She also served as chair of the PALNI Technology Advisory Group until January 2012.

# An Essential Field Guide to Web 2.0 Technology



By Nicholas G. Tomaiuolo  
ISBN 978-1-57387-425-0 • \$49.50

"A thorough and thoughtful guide to Web 2.0, providing historical background on user-generated content and a field guide to Web 2.0 applications. Especially valuable for information professionals and all those who work with the web."

—Ron Gilmour,  
natural sciences librarian,  
Ithaca College

**UContent Is Available in Print and Ebook Editions. Visit Our Website for more Information and to Order.**

 Information Today, Inc.

[www.infotoday.com](http://www.infotoday.com)



Copyright of Computers in Libraries is the property of Information Today Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.