University of Illinois Library

Federated Search Pilot Project Implementation Team: Final Report

August 28, 2006

I. Introduction

The Federated Search Pilot Project Implementation Team (“the Feds”) was appointed in the fall of 2004 by the now defunct Access Strategy Team. When the Content Access Policy and Technology Group (CAPT) was created in 2005, the Feds came under the administrative oversight of this group.

The Feds were charged with: (1) overseeing and managing all aspects of the pilot implementation of an undergraduate information portal utilizing federated search techniques at the UIUC Library and (2) collaborating with the OpenURL Link Resolver Implementation Team (OLRIT) to insure that link resolution would be enabled for the information sources and services selected for inclusion in the pilot user interface.

Team members include: Susan Avery, Naun Chew, Jon Gorman, Lisa Hinchliffe, JoAnn Jacoby (replaced by Sue Searing), Kathleen Kern (replaced by Erik Kraft), Erik Kraft, Lori Mestre, Bill Mischo, Dawn Schmitz, Sue Searing, Peggy Steele (co-chair), David Ward (co-chair).

The present document serves as the Team’s final report to CAPT. It includes a set of recommendations to CAPT on future utilization of federated search technology in the Library and a narrative review of the Team’s activities and discoveries.

II. Recommendations

Based on the Feds experiences during the pilot project, we have the following recommendations for CAPT regarding the use of WebFeat, Search Assistant and other multi-database search technologies in the University Library.

1. We recommend against implementing the current version of WebFeat. We do recommend, however, that CAPT appoint a small working group to review the functionality of the new version of WebFeat when it becomes available and to assess its ability to meet the Library’s requirements for multi-database searching.
2. We recommend that the University Library continue to offer the Search Assistant undergraduate interface to our users, and that this system be migrated from pilot project status to a fully-supported production system.
3. We recommend that the availability of Search Assistant be publicized to the user community during the fall semester.
4. We recommend that CAPT, in collaboration with Library IT, promptly determine the level of technical support needed to maintain and further enhance the Search Assistant undergraduate interface as a core production
service of the University Library, and that Library IT immediately assign personnel to this task in order to migrate the system to production status as quickly as possible.

5. We recommend the creation of additional discipline or subject-specific Search Assistant portals associated with other subject-focused areas of the Library’s web space. Additional portals may be focused specifically on undergraduate research needs or may be designed to serve other segments of the academic community.

6. We recommend creation of a small Search Assistant working group comprising four public services/collections personnel and one or more IT personnel to plan and oversee ongoing development of the user interface(s) and the underlying search and retrieval software. Members of the advisory group should include one representative each from Undergrad, Reference, and two Library units outside CPS, plus the IT personnel assigned to support the system. We recommend appointing one public service representative and one IT representative to serve as co-chairs of the group. The working group should report to CAPT and regularly keep CAPT informed of its activities.

7. We recommend that the Search Assistant working group regularly review usage statistics and also perform periodic usability testing to inform future development of the system.

8. We recommend that CAPT sponsor the investigation of other technologies, beyond federated search as we know it today, for consolidating and simplifying user access to the Library’s digital and print collections and remote information resources.

III. Review of Team Activities

A. Initial Activities

The Feds first convened on December 7, 2004, and agreed to meet regularly twice a month for the duration of the project.

The Team initially expected to implement a WebFeat system for the undergrad information portal. WebFeat had been acquired by ILCSO (now merged with two other consortia to form CARLI) for use by its member libraries. A WebFeat customer support representative was assigned to work with both UIUC and the other ILCSO libraries which chose to implement the WebFeat product. Given the nature of the product, each ILCSO library worked independently with WebFeat on its own implementation, although ILCSO also established a coordinating committee to assist the individual libraries with initial training and common issues.

The Feds originally expected to roll out the initial WebFeat interface in time for Fall Semester 2005. This timeline proved to be unrealistic, primarily due to problems uncovered with the WebFeat translators and staff turnover within WebFeat customer support. (Translators are the small programs that negotiate the communication of search strings and search results between WebFeat and individual native databases.)
B. WebFeat

Over a period of months the Feds came to understand that the technology underlying the WebFeat translators was inflexible and that the translator programs could not be customized to meet the search needs of users at individual customer sites. This was not made clear up front; it took several rounds of communication between UIUC and WebFeat for the Feds to get a full understanding of the limitations of the translator technology. Additionally, despite repeated requests detailing a set of common, specific behaviors expected from each translator, Webfeat did not modify the translators to the committee’s satisfaction.

The translator problem had a significant negative impact on the quality and usability of the results retrieved through a WebFeat search. There was no consistency in the way strings of keywords were interpreted by the system across different databases. We asked WebFeat to design translators so that all keyword phrases enclosed in quotation marks would be treated as phrases, as occurs in Google and many other search interfaces. We also asked that strings of keywords not enclosed in quotation marks be treated as Boolean AND searches. WebFeat failed to create translators that could do these two things consistently. Thus, a term like Acquired Immune Deficiency Syndrome, whether enclosed in quotations or not, might be searched as a phrase, as Acquired AND immune AND deficiency AND syndrome, or as Acquired OR immune OR deficiency OR syndrome. As a consequence, databases with similar content would offer very different results -- tens of thousands of citations from some, fewer than a hundred from others. This was confusing and misleading.

In addition, Webfeat was unable to consistently “pass-through” the robust native system full-text links to their short-entry citations in the results displays. This feature was implemented over a few target systems, but not others, and even when it was properly implemented, the terminology used in the short entry display links was inconsistent and confusing.

Another frustrating aspect of testing the Webfeat system involved their inability to provide translators that would work with the target services we specified. The Feds team spent many hours attempting to test translators that failed to connect to specific native systems at all (and often displayed no error messages) and, in at least one case, connected to and searched the wrong system. Some translators would fail on an intermittent basis. This did not inspire great confidence in the viability of the translators.

The vendor assured us that a new, much more flexible version of the system was under development and would be in production by early Spring 2006. As of the writing of this report, however, the new version is still in development and a release date has not been announced.

UIUC’s attempt to implement WebFeat coincided with a period of staff turnover at the company which also occurred at the same time the vendor was taking on a significant number of new customers. The combination of these events made it difficult for WebFeat to work with us on a satisfactory basis.

The difficulties we experienced with WebFeat translators were also symptomatic of a more general problem. The company’s approach to the development and
implementation of the product often seemed not to be attuned to the needs of an academic institution. The low priority placed on consistent search logic is one example of this problem; another is the failure to provide satisfactory support for reference linking, an important application in an academic library setting. Such difficulties made it difficult to have confidence that problems with the product would be rectified in the long term.

C. Search Assistant

Fortunately for the Library, Grainger Library had developed an alternative in-house system for simultaneous searching of multiple information resources which was already in active use by the Library’s engineering and business patrons. This system is known as “Search Aid for Engineering and Business Information.” With the active cooperation of Grainger personnel, a version of Search Aid, customized to meet the Feds’ guidelines for undergrad use, was developed to fill the gap left by the unsuccessful attempt to implement WebFeat.

The undergrad version of Search Aid is known as “Search Assistant.” It is a straightforward search and discovery system and operates under a different model of returning results than WebFeat. On the initial results screen WebFeat returns individual citations from target databases, displaying a set number of citations which varies from database to database, with embedded links to full-text articles whenever they are available. Search Assistant, by contrast, gives the user a quick summary of the number of relevant articles in each database that was searched and enables the user to go directly into each native system to view the individual citations. Full text can be reached through options in the native interface, such as Discover links or native full text links. It is a clean and simple system, at least from the user’s perspective, and, most importantly, it has been designed to generate consistent search queries across databases—a claim that could not be made by WebFeat. Neither WebFeat nor Search Assistant merges results from multiple databases, and neither provides relevancy ranking of databases or citations. Search Assistant’s response time is noticeably faster.

Search Assistant made its unannounced public debut on June 12, 2006 via links on the Library Gateway and the Undergrad web page. The timing coincided with the beginning of Summer Session II.

The Feds offered Search Assistant training to Library personnel immediately prior to the June 12 system debut. Three sessions, each covering the same content, were scheduled in Undergrad 291 at various times of day during a two-week period, giving staff multiple opportunities to attend. Each session covered the history of the pilot implementation project, the decision to adopt Search Assistant, a brief introduction to Search Assistant, and information on how staff could deal with problems that might arise with the system.

At the outset of the project, the Feds expected to make a large number of commercial electronic information resources in many disciplines searchable via WebFeat, in addition to the UIUC and I-Share catalogs plus some of the Library’s local ContentDM databases. As our work progressed and we learned more about WebFeat, and federated search technology in general, we realized that selecting a more tightly focused collection of resources for our undergrad portal would give the
intended users a more successful search and discovery experience. Research done on federated search outcomes at other academic libraries confirmed this conclusion.

The Feds continued to work with this more exclusive (rather than inclusive) model for resource selection even after abandoning efforts to implement WebFeat and turning our attention to Search Assistant. As a result, the version of Search Assistant currently available to users offers the following options for general multidisciplinary searching: 1) a grouping of three multi-disciplinary journal article databases; 2) a grouping of four newspaper databases; 3) the UIUC and I-Share catalogs; 4) a grouping of reference databases, which has not yet been implemented.

In addition to these general resources, the current version of Search Assistant also offers options for searching three subject-specific database collections: Engineering, History, and Business. They serve as examples of how the Search Assistant could be used to create additional discipline or subject-specific portals associated with other subject-focused areas of the Library’s web space.

**D. Usability Testing**

Two sets of usability tests were conducted in the Spring and Early summer of 2006. The usability tests sought to determine whether students understood what the software was for, whether they could use it effectively, and what modifications needed to be made to the interface.

The first usability test compared Webfeat and Search Assistant head-to-head, asking students to perform the same tasks in each. Neither interface was the clear winner. One user disliked them both, but two users liked them both about equally and asked when they would be available. Overall, the Search Assistant interface was perceived as easier and faster to use. Options to manipulate results (e.g. sorting, emailing) were appreciated in WebFeat. WebFeat offered both pre-selected packages of databases to search and a list of individual databases, from which the user could customize a selection. Undergraduates had difficulty choosing appropriate databases.

The second usability test used only the Search Assistant Interface, which had been modified to accommodate suggestions from the first usability test. The results were generally positive, with the most difficulty coming in understanding how results were organized (that is, where the best results were for any search). Various suggestions for aesthetic and functional changes were adapted into the current version of Search Assistant.

**E. Use Statistics**

A transaction analysis system that records usage statistics has been developed and implemented. Daily and year-to-date statistics are available and a database recording user-entered search arguments and result resources “clicked-on” is being continuously built. Through August 2, the system has captured 5334 user search arguments. Since June 28, the 3312 performed searches have generated 4119 click-throughs to a retrieved search result.
F. Conclusion of the Project

With the conclusion of 2006 Summer Session II, the work of the Feds also comes to an end. We have implemented a viable new search and discovery tool for undergrads and tested it with summer session users for roughly two months. We feel that, despite a rocky start, the project has had a successful conclusion. The Library now has a working undergrad information portal, built on technology that can be enhanced and reused to build additional information portals for either undergrads or other segments of the UIUC academic community.

Much credit for this success goes to Bill Mischo and his colleagues at Grainger for their work in customizing their own locally-developed software to meet the specific needs of the pilot project the Feds were charged with managing. The Feds and the Library owe their thanks to the developers of Search Assistant.