



NASIG News

NASIG Webinar: Tracking Down the Problem: The Development of a Web-Scale Discovery Troubleshooting Workflow Reported by Sofia Sluts kaya

Todd Enoch, head of Serials and Electronic Resources for the University of North Texas (UNT) Libraries in Denton, started the presentation by giving a definition of web-scale discovery and describing the UNT Denton discovery set-up. According to Enoch, web-scale discovery is a service that indexes materials from many different sources. When a library subscribes to a discovery service and a user performs a search, the discovery service reaches out to the institution's knowledgebase (Serials Solutions at UNT Denton). The knowledgebase returns results that are available to users in full-text. When users click on the link, the content is retrieved using an OpenURL link resolver (Serials Solutions 360 Link for UNT Denton).

The presenter noted that for many libraries, this discovery process often breaks down due to a variety of reasons, including:

- the discovery product has incorrect metadata or linking syntax problems;
- an institution fails to update the knowledgebase holdings, proxy configurations, and/or subscription information;
- user misunderstanding of their search results.

To preface a discussion about troubleshooting workflows, Enoch provided some background information on UNT Denton's web-scale discovery service (Summon) implementation. Summon was

introduced in 2012 and was promoted mostly as a full-text article search interface. A survey conducted shortly after the Summon implementation showed a 71% positive response to the new service. However, despite the lack of help tickets, there were many unofficial complaints about the failure of the new discovery service to produce good results. Acting on the anecdotal evidence, Enoch initiated a meeting with public service librarians that confirmed their dissatisfaction with Summon's performance.

It became clear to Enoch that the existing error-reporting mechanisms were not sufficient and did not enable users and public service librarians to easily capture enough information to effectively diagnose access issues.

The solution to the problem was to embed an error-reporting link on the Summon's search results pages. The form enables users to select a type of error and include additional optional comments. It also harvests metadata from Summon. In the initial implementation, the patron's name and contact information were optional and were included in the comments fields. Later, the patron's name and e-mail address were entered separately into their own fields for easier follow-up. Submitted error reports are routed to the e-resources' e-mail address.

Enoch noted that the greatest advantage of this approach is that the report contains a lot of information harvested directly from Summon: the full citation, the "problem" URL that the patron clicked on, and the search results page URL to help recreate the search context and to simplify troubleshooting. In the first

month of the form's existence, Enoch received 200 error reports. Since its implementation 4 years ago, 7,347 error reports were submitted by library patrons. The number of error reports was so large that it was impossible for just Enoch to manage them and it became necessary to develop a workflow and to train staff members and student workers to handle error reports.

In his presentation, Enoch outlined the workflow steps:

- error reports are received into "*Active Summon Errors*" folder;
- student assistants and staff members retrieve 5-10 error reports at a time and move them to their personal "*In progress*" folders;
- email is moved to the "*Completed*" folder and statistics are recorded once the problem has been identified and responsible parties notified.

Follow-up communications are handled on an as-needed basis. Each individual working with the issues maintains their own statistics.

Enoch noted that the most challenging part of the workflow is correctly diagnosing and troubleshooting the issues. This requires knowledge of e-resources and "detective" skills. Enoch spent the largest part of his presentation going over ways of identifying errors. The first step is usually to evaluate the error message. However, it should be noted that error messages are not always available, and some search results may appear as errors to end users even though they are not (for example, the OpenURL resolver does not take users to the specific article but rather to a database/journal landing page).

If there is no error message, staff members working on a ticket should still verify that the full-text is accessible. If it is not, they should check the status of the subscription and verify if the holdings are correct in the knowledgebase. All cancelations and additions should be reflected in the knowledgebase in a timely manner. Even if full-text is accessible, it is still important to verify

that it is the correct article and that all pages are legible and to notify the content provider if that is not the case.

In many cases, even after extensive investigation the staff is not able to diagnose the problem. Sometimes, technical issues are already resolved or the problem occurs on the patron side (cookies, firewall settings, etc.) and sometimes users have unrealistic expectations or cannot interpret the results.

After addressing the identifying issues workflow, Enoch explained how errors are categorized for statistical purposes and shared some statistical data collected over the last four years: 37% of reported errors required some follow-up action for them to be resolved and 10% of errors were "no action taken" problems. In 53% of reported cases, the staff was not able to identify or replicate the error.

The following types of problems were identified by Enoch as requiring follow-up action: citation errors, DOI was incorrect or not registered, duplicate entry, embargo not accurately reflected, holdings incorrect, knowledgebase is returning false positives, linking errors, missing articles on provider site, proxy not configured, and subscription problems.

The types of errors that require no follow-up action are browser problems on the patron's end, granularity issues (i.e. when the discovery layer and the content provider index materials on a different level), problems with Open Access articles that are not set-up to properly communicate with link resolvers, temporary technical difficulties, and user errors.

In addition to the types of errors, Enoch maintains statistics for the cause of the error. The highest percentage of errors (45%) occurred because of discovery service/knowledgebase issues. 37% of errors were caused by aggregators, 11% by publishers, 6% by the library, and only 1% were user errors.

Enoch sees multiple benefits of using the error reporting workflow. Allowing patrons to report errors alleviates some of their frustrations. It also brings staff

attention to issues such as incorrect holdings or proxy configurations that might not be discovered otherwise. Gathered data helps in educating public service librarians about web-scale discovery. He observed that giving users and librarians the ability to report problems has resulted in a change of attitude towards Summon. Reporting vendor, publishers, and knowledgebase issues also improves experiences for users at other institutions.

Enoch concluded the presentation by discussing recent changes in the error reporting workflow. A user information field was added to the form. Users are still not required to provide their contact information but are encouraged to do so if they want to access an article in question. Since the change was implemented, over half of the error reports included user information. This change allowed staff to better prioritize the error reports by first addressing the ones requiring a response.

Enoch answered many questions from the audience, including a question was about using a similar workflow for e-book troubleshooting. Enoch explained that the link currently only displays for full-text articles. He also believes that e-book URLs are more stable and create fewer issues. Another questioned if the error reporting form is embedded in other database interfaces. The presenter stated that it is currently only embedded in Summon because Summon enables harvesting of metadata.

A few questions were related to staffing and using student workers to support the troubleshooting workflow. Enoch answered that he tries to hire students with analytical ability and provides one-on-one training. He stated that there is currently only 1 student working 20 hours per week who handles all error reports with help from staff members as needed. Reporting errors to vendors is also handled by a student worker and is done through the error reporting mechanism provided by each vendor. All follow-up communications with vendors are managed by a student as well, except for more complex cases and issues related to budget and payments.

Enoch was asked about recording and using statistics. He said that statistics are recorded in an Excel spreadsheet. Each person records their own statistics. Personal spreadsheets are compiled by Enoch. The data is mostly used internally and for training public service staff.

One participant asked if the number of error reports decreased over time. The presenter did not see a significant decrease but noted that the number fluctuates depending on the time of the semester.

Finally, the presenter was also asked about scheduling and turnaround time. He stated that questions are answered during normal business hours. The average turnaround time is 24 hours, but it takes longer for questions received on Friday, Saturday and Sunday.