

Link-Checking the E-Journal Portal: Trust, But Verify

Alexis Ackel, Electronic Resources Librarian & Clayton Crenshaw, Web Services Librarian | Gibson D. Lewis Health Science Library



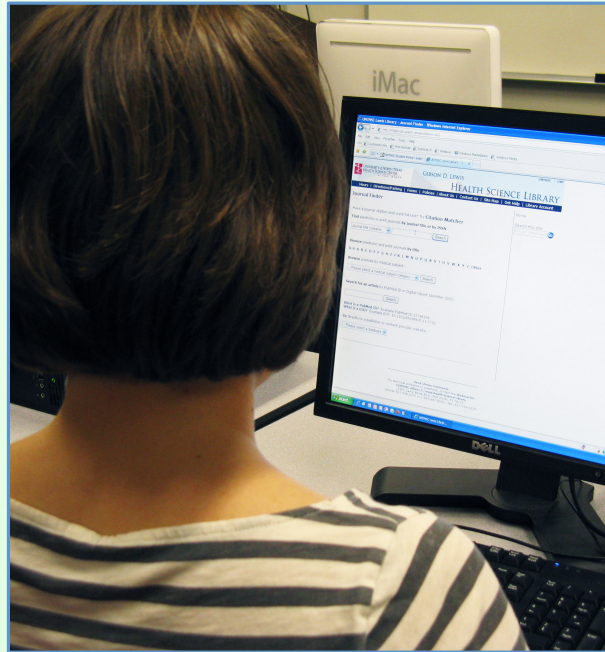
Question

Can an academic library proactively check the validity of links provided by a third-party e-journal portal?

Setting

The UNT Health Science Center serves more than 1,500 students in biomedical sciences, public health, allied health, and medicine.

Lewis Library provides access to more than 20,000 journals online.



Methods

Using Excel spreadsheets, text files, and the free Xenu Link Sleuth software, we developed a procedure for testing the links provided through our e-journal portal:

- Download an Excel file containing the links from the e-journal portal vendor.
- Remove all data except the URLs and, using Find and Replace, strip out the proxy strings within URLs. Including proxy strings caused the link-checking process to exceed the proxy software's connection limits.
- Divide the URL list into smaller groups of approximately 5,000. If the files were too large, Link Sleuth would run but would fail to generate a report.
- Save the resulting URL lists as text (.txt) files. This was necessary in order to use Link Sleuth's "Check URL list" function; it would not process data within Excel files.
- To prevent checking beyond the URLs provided, set the Maximum Level in the Link Sleuth Options to zero. Then run Link Sleuth against each text file and generate a report.
- Using the Link Sleuth report and the original Excel file, match up broken links with titles and ISSNs.
- Verify the broken links and resolve them. We handled these in one of several ways:
 - Directly replace the URL used in our e-journal portal with a functioning one.
 - Provide a functional URL to the e-journal portal vendor and ask them to correct their database.
 - Simply deactivate the link in the e-journal portal. This is typically used for journals that have ceased publication or that have moved from a free site to a commercial site.

Results

Each time URLs were checked, broken ones were found; most of them were associated with freely available journals. In three checks of URL files generated several months apart, a total of 57 journals were found to have broken links. The journal error rates for the three tests were .06%, .03%, and .06%.

In addition to those for e-journals, we found that broken links were being generated for a significant portion of the titles within a single newspaper database. Further investigation revealed problems with both the links and the title list. By working with both the database publisher and the e-journal portal provider, these problems were largely corrected.

Report from Link Sleuth

Broken links, ordered by link:

<http://bjoc.bellstein-journals.org/articles/browse.asp>
error code: 404 (not found), linked from page(s):

<http://bjp.rcpsych.org/supplements.shtml>
error code: 404 (not found), linked from page(s):

<http://brain.oxfordjournals.org/>
error code: 12002 (timeout), linked from page(s):

<http://brain.oxfordjournals.org/archive>
error code: 12002 (timeout), linked from page(s):

<http://brain.oxfordjournals.org/archive/>
error code: 12002 (timeout), linked from page(s):

<http://cebp.aacrjournals.org/search>
error code: 12002 (timeout), linked from page(s):

[http://onlinelibrary.wiley.com/journal/10.1111/\(ISSN\)1268-7731](http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1268-7731)
error code: 404 (not found), linked from page(s):

[http://onlinelibrary.wiley.com/journal/10.1111/\(ISSN\)1467-8721](http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1467-8721)
error code: 404 (not found), linked from page(s):

<http://www.apfmj-archive.com/archive.htm>
error code: 12007 (no such host), linked from page(s):

<http://www.clinicalcardiology.org/productcart/pc/previous.html>
error code: 403 (Forbidden request), linked from page(s):

<http://www.ehealthinternational.org/articles/browse.asp>
error code: 404 (not found), linked from page(s):

Conclusions

We found that periodically checking our e-journal portal links is a feasible project that can reduce the number of broken links being delivered to our patrons, thereby improving their online experience with the Library. We also demonstrated the utility of this process for identifying linking problems at the database level. Though the broken links represent a very small percentage of the total, we know from experience that our patrons do find and report them. Therefore, we plan to continue the checks on a quarterly basis.

Contacts

Alexis Ackel
Alexis.Ackel@unthsc.edu

Clayton Crenshaw
Clayton.Crenshaw@unthsc.edu

Gibson D. Lewis Library
University of North Texas – Health Science Center
3500 Camp Bowie Boulevard
Fort Worth, TX 76107
(817) 735-2070

	URLs Checked	Number Broken	Percentage
March	36,384	23	.06%
July	38,630	104 (12 journals, 92 newspapers)	.27% (journals = .03%)
Sept	39,189	22	.06%

Journals with Broken Links

American Indian and Alaska Native Mental Health Research
Annals of General Hospital Psychiatry
Archives of Gastroenterohepatology
Asia Pacific Family Medicine
Beilstein Journal of Organic Chemistry
British Journal of Psychiatry, Supplement
Canada Communicable Disease Report : CCDR weekly
Clinical Cardiology
Cuadernos de Historia de La Salud Pública
Current Allergy & Clinical Immunology
Current Directions in Psychological Science
Ehealth International
Foot and Ankle Surgery
Genes to Cells : Devoted to Molecular & Cellular Mechanisms
Herpes
Humanidades Médicas
Indian Journal of Nuclear Medicine
Iranian Journal of Pharmacology & Therapeutics
Journal of Indian Academy of Oral Medicine and Radiology
Journal of Indian Prosthodontic Society
Journal of Oral and Maxillofacial Surgery
Journal of the Anatomical Society of India
Journal of the Pakistan Medical Association
Nursing Outlook
Open Access Rheumatology : Research and Reviews
Open Access Surgery
Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontics
Otolaryngology--Head and Neck Surgery
Perspectives in Health Information Management
Planning for Higher Education
PMC Biophysics
PMC Physics. A
PMC Physics. B
Reumatologia
Revista Archivo Médico de Camaguey
Revista Cubana de Enfermería
Revista Cubana de Estomatología
Revista Cubana de Hematología, Inmunología y Hemoterapia
Revista Cubana de Higiene y Epidemiología
Revista Cubana de Investigaciones Biomédicas
Revista Cubana de Medicina
Revista Cubana de Medicina General Integral
Revista Cubana de Medicina Militar
Revista Cubana de Medicina Tropical
Revista Cubana de Obstetricia y Ginecología
Revista Cubana de Oftalmología
Revista Cubana de Ortopedia y Traumatología
Revista Cubana de Pediatría
Revista Cubana de Plantas Medicinales
Revista de Salud Animal
Revista Habanera de Ciencias Médicas
Revista Mexicana de Reumatología
Texas Medicaid Bulletin
Vaccinmonitor
Viewbox
World Psychiatry
Yonsei Medical Journal