A Discovery of Trees: A Comparison and Evaluation of Five Web-Scale Discovery Tool Services

By:
Dana M. Caudle
Midge Coates
Cecilia M. Schmitz
With assistance from Nadine Ellero
Web-Scale Discovery Tool Services (DTSs)

... promise a search experience that is:

• Fast: one box, one search
• Easy: faceting technology to narrow large search result sets
• Unified: all (or nearly all) a library system’s owned and licensed resources
• Effective: proprietary algorithms for greater relevance in search results
Research Question

How well would these tools perform for discovering a set of known items deemed authoritative by a subject expert?

We consulted a local expert for a list of materials relevant to a specific research topic ...
Effect of Tebuthiuron (herbicide) on Live Oaks (Quercus Virginiana)
Methodology: Monographs
Methodology: Journal Articles
Methodology: DTSs
Methodology: Institutions

- 2 institutions/implementations for each DTS
- Public access (no log-in) for search required
- Land grant institutions preferred
  - Similar collection materials
  - If no land grant institution available, then substitute ARL institution
Methodology: Searches

• OPAC known item searches = monograph titles, journal titles (not article titles)
• DTS known item searches = monograph titles, article titles
  – With/without quotation marks
• DTS keyword searches = pre-selected phrases
  – With/without quotation marks
  – With appropriate wild card character, as necessary
• For each search:
  – How many results were in the list?
  – Was our item on the list? (looking at first 2 pages only)
  – If yes, what position on the list?
DTS Results for Monographs

Known Item Searches
• Book title searches worked well for all DTSs.
• Quotation marks reduced the total number of results, but didn’t affect an item’s position on the list.

Keyword Searches
• Never more than 2 books found (out of 5)
• Not the same 2 books, either
• Non-relevant books at the top of the results list
• Hard to exclude these non-relevant books
Oak Forest Ecosystems

• OPAC—
  – known-item search: title
Oak Forest Ecosystems

- OPAC—
  - known-item search: title
- DTS—
  - known-item search: title
Oak Forest Ecosystems

- **OPAC**—
  - known-item search: title

- **DTS**—
  - known-item search: title
  - keyword search: oak*
Oak Forest Ecosystems

- OPAC—
  - known-item search: title
- DTS—
  - known-item search: title
  - keyword search: oak*
  - keyword search: live oak*
DTS Results for Journal Articles

Known Item Searches

• Title searches less successful for articles than for monographs ...
  – Some results hidden from “guest” searchers
  – Some journals not indexed by DTS?

Keyword Searches

• Only one article found frequently (a few others found sporadically)
• Narrowing with subject facets
  – Reduced the results pool and moved articles higher up the results list
  – Did not reveal any additional articles from the target list
  – Many irrelevant articles remained after limiting by facets
“Control of live oak (Quercus Virginiana) and under-story vegetation with soil-applied herbicides” Weed Science

• OPAC—
  – known item search: journal title
“Control of live oak (Quercus Virginiana) and under-story vegetation with soil-applied herbicides” Weed Science

- OPAC—
  - known item search: journal title

- DTS—
  - known item search: article title
“Control of live oak (Quercus Virginiana) and under-story vegetation with soil-applied herbicides” Weed Science

- **OPAC—**
  - known item search: journal title

- **DTS—**
  - known item search: article title
  - keyword search: oak*
“Control of live oak (Quercus Virginiana) and under-story vegetation with soil-applied herbicides” Weed Science

- OPAC—
  - known item search: journal title

- DTS—
  - known item search: article title
  - keyword search: oak*
  - keyword search: live oak*
“Control of live oak (Quercus Virginiana) and under-story vegetation with soil-applied herbicides” Weed Science

- **OPAC**—
  - known item search: journal title
- **DTS**—
  - known item search: article title
  - keyword search: oak*
  - keyword search: live oak*
  - keyword search: oak* + tebuthiuron
“Control of live oak (Quercus Virginiana) and under-story vegetation with soil-applied herbicides” Weed Science

- **OPAC**—
  - known item search: journal title
- **DTS**—
  - known item search: article title
  - keyword search: oak*
  - keyword search: live oak*
  - keyword search: oak* + tebuthiuron
  - keyword search: live oak* + tebuthiuron
“Control of live oak (Quercus Virginiana) and under-story vegetation with soil-applied herbicides” Weed Science

• OPAC—
  – known item search: journal title

• DTS—
  – known item search: article title
  – keyword search: oak*
  – keyword search: live oak*
  – keyword search: oak* + tebuthiuron
  – keyword search: live oak* + tebuthiuron
  – keyword search: oak* + tebuthiuron + quercus
“Control of live oak (Quercus Virginiana) and under-story vegetation with soil-applied herbicides” Weed Science

- OPAC—
  - known item search: journal title
- DTS—
  - known item search: article title
  - keyword search: oak*
  - keyword search: live oak*
  - keyword search: oak* + tebuthiuron
  - keyword search: live oak* + tebuthiuron
  - keyword search: oak* + tebuthiuron + quercus
  - keyword search: live oak* + tebuthiuron + quercus
“Vegetation management in the Cross Timbers: response of woody species to herbicides and burning” Weed Technology

- OPAC—
  - known item search: journal title
“Vegetation management in the Cross Timbers: response of woody species to herbicides and burning” Weed Technology

- OPAC—
  - known item search: journal title
- DTS—
  - known item search: article title
“Vegetation management in the Cross Timbers: response of woody species to herbicides and burning” Weed Technology

• OPAC—
  – known item search: journal title
• DTS—
  – known item search: article title
  – keyword search: oak*
“Vegetation management in the Cross Timbers: response of woody species to herbicides and burning” Weed Technology

- **OPAC**—
  - known item search: journal title

- **DTS**—
  - known item search: article title
  - keyword search: oak*
  - keyword search: live oak*
“Vegetation management in the Cross Timbers: response of woody species to herbicides and burning” Weed Technology

- **OPAC**—
  - known item search: journal title
- **DTS**—
  - known item search: article title
  - keyword search: oak*
  - keyword search: live oak*
  - keyword search: oak* + tebuthiuron
“Vegetation management in the Cross Timbers: response of woody species to herbicides and burning” Weed Technology

• OPAC—
  – known item search: journal title
• DTS—
  – known item search: article title
  – keyword search: oak*
  – keyword search: live oak*
  – keyword search: oak* + tebuthiuron
  – keyword search: live oak* + tebuthiuron
“Vegetation management in the Cross Timbers: response of woody species to herbicides and burning” Weed Technology

- OPAC—
  - known item search: journal title
- DTS—
  - known item search: article title
  - keyword search: oak*
  - keyword search: live oak*
  - keyword search: oak* + tebuthiuron
  - keyword search: live oak* + tebuthiuron
  - keyword search: oak* + tebuthiuron + quercus
“Vegetation management in the Cross Timbers: response of woody species to herbicides and burning” Weed Technology

- OPAC—
  - known item search: journal title

- DTS—
  - known item search: article title
  - keyword search: oak*
  - keyword search: live oak*
  - keyword search: oak* + tebuthiuron
  - keyword search: live oak* + tebuthiuron
  - keyword search: oak* + tebuthiuron + quercus
  - keyword search: live oak* + tebuthiuron + quercus
Discussion: Can We Improve DTS Searching?

Simple keyword searches don’t work very well for disciplinary resources ...
Non-relevant materials (poetry, history, etc.) get in the way.

• Better facet narrowing
  – more categories
  – easier to find

• Disciplinary profile – users indicate what category of materials they are interested in (science, art, literature, etc.), and DTS uses this to pre-limit search results. (already in the works?)
Overall DTS Search Experience ...

• Fast ?
  Yes.

• Easy ?
  Easy to search; difficult to use facets effectively.

• Unified ?
  Hmmm ... Not really, but that wasn’t obvious.

• Effective ?
  Yes, for known item searches; No, for keyword searches.
Overall DTS Search Experience ...

• Any overall winners or losers?
  No, they were fairly equal in performance.
  • Ease of use
  • Search results
“Taking relevancy to the next level in sophistication will have to go far beyond the ways that keywords match the query”
(Breeding 2012a, 29)

Contact Information:

Dana M. Caudle
caudlda@auburn.edu

Midge Coates
coatemi@auburn.edu

Cecilia M. Schmitz
schmice@auburn.edu