

---

Liaison Librarian

---

November 2007

# Implementing Endeca When You're HIP

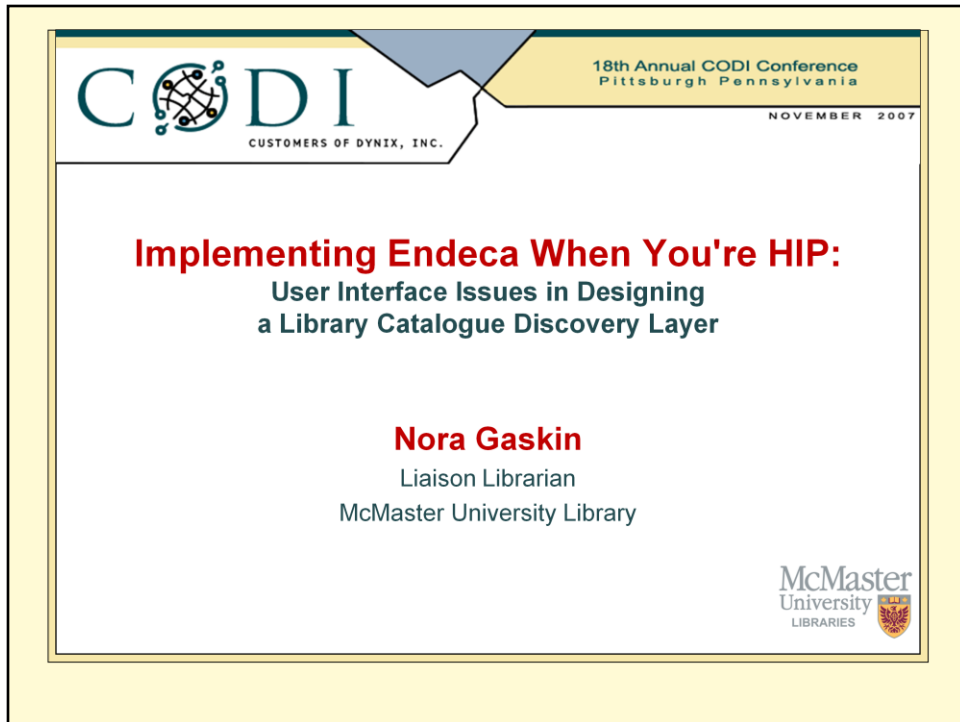
Contact  
Author

Start Your Own  
SelectedWorks

Notify Me  
of New Work



Available at: [http://works.bepress.com/nora\\_gaskin/1](http://works.bepress.com/nora_gaskin/1)



- Many thanks to Amanda Etches-Johnson, who generously allowed me to use her presentation:

[Forget the Lipstick, this Pig Needs a Complete Makeover](#) (presentation on McMaster's Endeca-powered OPAC), [Access 2007](#), Victoria, B.C., October.

<http://www.blogwithoutalibrary.net/?p=286>

...as the starting point for this presentation. Thanks, Amanda!

## Outline

- **why and how**
- before & after
- issues, choices, lessons
- what our users think
- future plans

## Why?

- LibQual results
  - improve e-resource discovery
- net generation
  - Google rules!
- Endeca
  - successful large academic library implementation @ [NC State](#); quick implementation!

- our LibQual results said our previous e-resource discovery tools were hard to use; we wanted to address this fast!
- most students search with Google syntax, i.e. no syntax! ...and expect relevance-ranked results
- Endeca could deliver a discovery layer quickly

## Endeca features we liked

- relevance ranking
- faceted browsing
- content spotlighting
- type-ahead
- stemming & thesaurus
- search within results

- Endeca did a demo for us in August 2006; we were impressed!
- faceted browsing allows you to refine your search by clicking on a variety of options
- content spotlighting provides a context-sensitive tool to highlight recommended e-resources, e.g. databases, encyclopedias, basic texts, new materials, etc.
- type-ahead suggests authors and subjects that match the user's search terms
- stemming file traps both singular and plural of searched-for terms
- ancillary stemming file traps both British and American spellings of searched-for terms
- we can add whatever terms we want to the thesaurus, e.g. movies = motion pictures

## About the implementation

- how long? 6 months from signing to launch
- teams:
  - Core Implementation Team (6)
  - User Interface Group (6-7)
  - ad hoc teams, e.g.
    - Content Spotlighting
    - Searchable Fields
    - LC/MeSH Mapping

- we signed at the beginning of October 2006
- launched our new interface in March 2007
- in addition to the core team and UI group, we had a number of ad hoc groups that we called together to make decisions as needed

### **About the implementation (cont'd)**

- cost in staff time: 2 FTE for 6 months
- Endeca did most of the development work

- cost in staff time is an estimate; we didn't keep strict track
- Endeca gave us 3 options: they could do all the development, we could work together to do the development, or we could do it all. Because we wanted to launch quickly, we went with option 1.

**Timeline:**  
**the 6 month whirlwind**

- October 17-18, 2006 -- Requirements Workshop
- December 2006 -- Requirements; Beta
- January 2007 – Staff testing
- February 2007 -- Initial deployment
- March 2007 -- Usability testing; tweaking
- March 18, 2007 -- Release to the public
- Since then -- tweak, tweak, tweak



## Horizon database to Endeca XML

- Horizon tables provide content
- MARCOUT utility creates MARC records from that content & puts them into a file, a binary representation of 1.6m MARC records
- That file goes into a pipeline where a custom-built utility preprocesses and streams it into the forge process
- The forge process produces a customized XML representation of the data to be used for search and display; Endeca properties and dimensions are created from MARC tag values and subfield values
- After the forge is the index process, dgidx, a d-graph indexer; this creates various Endeca search dimensions (indexes), and is the proprietary part of the process

- just in case anyone's wondering!

## Outline

- why and how
- **before & after**
- issues, choices, lessons
- what our users think
- future plans

## Before & after

- before: <http://morris.mcmaster.ca>
- after: <http://libcat.mcmaster.ca>

## Outline

- why and how
- before & after
- **choices, issues, lessons**
- what our users think
- future plans

**CODI**  
CUSTOMERS OF DYNIX, INC.

18th Annual CODI Conference  
Pittsburgh Pennsylvania  
NOVEMBER 2007

### UI Choices: search

- index dropdown location
- index labels
- search, clear, search within; oh my!

McMaster University LIBRARIES

- Where to put the index dropdown (after the search input box rather than before) was a fairly easy decision, since we wanted to stay consistent with our previous HIP interface and our major database interface, Scholar's Portal
- We looked at other catalogues for index labels we liked; still not sure we've got it right though – our users were used to having the browse indexes first & some are disoriented with a default to keyword searching; maybe adding Keyword to the index labels would help, e.g. Title Keyword, Keyword Anywhere
- When you haven't done a search yet, there's just a "Go" button to the right of the search box, but after you've done a search it gets a little complicated.
- We started with just a Go button, and a tick box underneath with the text "Search within results"; nobody saw search within, and people didn't know how to start over
- Added a Clear button to the right of "Go"
- Considered ways to make search within results more visible; went with a button, but then what to do with "Clear"?
- Finally figured out we could combine the functions of "Go" and "Clear" into "New Search" , so still only 2 buttons! This seems to have worked pretty well.

## UI Choices: defaults

- search type: keyword anywhere
- sort: relevance
- after a search, leave search terms in search box?
- after a search, if you just hit enter, new search, or search within?

Our decisions so far:

- after a search, leave search terms in search box? No
- after a search, if you just hit enter, new search, or search within? New search

## UI Choices: dimensions

- NC State's dimensions +
- Hierarchical or flat?
- Value sort within dimensions: alpha or most hits at the top?
- The look: Open? Closed?

- Started with NC State's dimensions & they kindly shared their coding choices with us!
- decided on 10 dimensions, removed them from the initial basic search screen for a clean look, and exposed them in the search results page and on the advanced and browse tabs

**CODI**  
CUSTOMERS OF DYNIX, INC.

18th Annual CODI Conference  
Pittsburgh Pennsylvania  
NOVEMBER 2007

### UI Choices: keyword vs. browse

- where to put alpha index search?
- popup dilemma

McMaster  
University  
LIBRARIES

- we initially separated the alpha browse search indexes from the keyword indexes by putting them in another tab – it didn't take long for us to change our minds, since no one could find them; we're learning most people just accept the default tab and the default index
- since Endeca doesn't do browsable indexes, we had to decide on how to deliver a HIP browse screen; decided on a featureless popup, because in testing, users maximized the HIP popup and never got back to the Endeca interface
- the featureless popup has problems of its own, though, when a user opens a link to an e-resource and wants to use navigation, print, etc.



## UI Choices: system messages

- null results page
- did you mean?
- corrected to...
- breadcrumbs

- null results screen provides suggestions for help
- did you mean and corrected to: we wanted to provide the user as much information as possible
- breadcrumb style: could have been horizontal or stacked; we thought stacked was easier to interpret

## UI Choices: results display

- results list = brief list
- more detailed record: AJAX layer
- HIP popup
  - complete MARC record; item & holdings
  - current status

- most of the info in the HIP popup is coming soon to the AJAX layer, with the exception of massive holdings records, which will be abbreviated with a link to HIP

### Issues: relevance ranking

- working well for subject searches
- not so well for some known-item searches, e.g.
  - Journal Title = nature
  - Author = smith a
- ongoing tweaking
- new relevance test environment

- we're trying to improve known-item searching by tweaking the relevance ranking, but there's only so much you can do! Sometimes a browse index is the only solution!

### **Issues: fuzzy logic**

- matching on any 2 or more terms is a tough one for librarians to wrap their heads around
- add stemming on plurals and spell correction and results can occasionally seem baffling

## Issues: data imperfections

- typos
- punctuation inconsistencies
- coding inconsistencies
  - \$v, \$x
- coding errors

- ...and incomplete catalogue records
- they show!

## Issues: you mean we can't have...?

- Boolean
- truncation
- browsable indexes

Coming Soon...

- current status
- summary of holdings & item records

- our advanced search interface does the equivalent of an “and” and “or” search using the “all the words” and “any of the words”
- the stemming file can handle plurals, but not other alternate endings
- current status and summary of holdings/item records are coming soon

### **Some interesting results from usability tests...**

- Only 1 of 5 subjects used the dimensions (oh no!)
- Some subjects relied on type-ahead to complete all queries

- Type-ahead: we've disabled the feature for now because it was truncating searches because of slow response time and it wasn't suggesting the matching authors and subjects with the most hits

## Issues: users -- hello!

- Dimensions?
- Advanced Search?
- Browse?
- Interactive search examples?

- we're not sure if any of these great features are getting used much
- have to do usability testing & examine search logs



## Outline

- why and how
- before & after
- choices, issues, lessons
- **what our users think**
- future plans

## Feedback

- newbies love it; returning students vary
- undergrads love it; grads & faculty vary
- biggest problem: known item searches
- biggest plus: relevance + stemming + match on any 2 or more terms = results

- undergrads tend to like it because it works well as a discovery tool;
- grad students & faculty, who tend to do more known-item searching, aren't always big fans, though some are

## Feedback

- people who used to Google-search the catalogue & get 0 hits love it
- people who used to search the catalogue efficiently and effectively no longer know what they're doing, have to learn new skills

## Outline

- why and how
- before & after
- choices, issues, lessons
- what our users think
- **future plans**

## Future plans

- current status
- “Available items only” search
- British and American spelling equivalents
- “My List”, emailing results, search history
- alpha indexes
- downloading records (Z39.50, etc.)
- analyzing search logs & more usability testing
- social layer, tagging

- current status is close; will be updated several times throughout the day
- search for available items only: once we’ve got live availability working, we’d like to add a search dimension that will allow our users to only search only for items that are in the library
- British and American spelling equivalents (ancillary stemming file)
- search history, the ability to add items to a “my list” feature and email those items to yourself or save/export them – our users are asking for these
- alpha indexes: find a better way to deliver them than the current featureless popup
- Z39.50: find a way to download citations into a citation manager using a z39.50 server



Check the video recording! Already forgotten ALL the questions (except the ones I couldn't answer, grrr.)

**Thanks!**

Nora Gaskin

McMaster University Library

email + MSN: [gaskinno@mcmaster.ca](mailto:gaskinno@mcmaster.ca)

<http://userfriendlylibrarian.wordpress.com/>