FAST in Practice: a progress report from Cornell

The FAST Team

OCLC Research: Rick Bennett, Eric Childress, Kerre Kammerer, Ed O'Neil, Diane Vizine-Goetz
Cornell University: Gary Branch, Chew Chiat Naun, Steven Folsom, Sarah Ross, Ardeen White

Cornell cataloguers for pilot: Yen Bui, Roswitha Clark, Sung Ok Kim, Yelena Kurbanova, Apikanya McCarty, Teresa Mei
Also: Frances Webb (Cornell Discovery & Access team)

Outline
- Motivations
- Players
- LCSH vs FAST
- Discovery implications
- Cataloguing experiences
- Vocabulary development
- Future possibilities
Why FAST? Why Us? Why Now?

Two drivers:

Discovery and Access
- Improved genre/form access and faceting
- Minimal level cataloging
- Controlled subject access

Discovery and Access Initiative
- Charged to provide integrated access
- Multiple data sources
- Improved user experience
- Representation across the library
- Library Technical Services members
- Project developer
- Blacklight system
- Open source
- Runs off triple store
- Supports faceted interface
- Full local control over Solr index

Cornell Pipeline for Blacklight

Data Sources
- MARC21 (from ILS and elsewhere)
- Non-MARC

Integration Layer
- XML
- RDF

Mapping to Solr Fields for Blacklight

Indexing
Library Technical Services

- No-backlog policy
- Minimal-level cataloguing
- About 25% of original cataloguing output
- Subject keywords
- Non-MARC cataloguing
- Emphasis on cross-functional training
- Internal Batch Processing/Automation unit

Timeline

- **September 2013**
  - Cornell makes initial inquiry to OCLC about genre/form algorithm
  - OCLC announces plan to incorporate the FilmR hetero FAST
- **September 2013**
  - FAST driven on FAST
- **January 2014**
  - Cornell and OCLC agree to FAST conversion project
- **February 2014**
  - FAST cataloguing pilot commences (will become MLC standard in December)
- **Spring 2014**
  - Batch conversion testing
  - Vocabulary development discussions (local extensions, problematic mappings)
- **October 2014**
  - FAST adds converted records
  - Initial efforts to map FAST headings into Voyager records
- **Early 2015?**
  - FAST implemented in discovery

LCSH Limitations for Faceting

- Inconsistent terminology (place names)
- Implied concepts (events)
- Obsolete coding (genre/form)
- Nonspecific, nonparsable dates
- Weak hierarchy
### Current Subject Faceting in Blacklight

<table>
<thead>
<tr>
<th>Language</th>
<th>Subject/Gene</th>
<th>Subject Region</th>
<th>Subject Era</th>
<th>Facets</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>Gambling</td>
<td>Nevada, Las Vegas</td>
<td>20th century</td>
<td>Fiction/Non-Fiction</td>
</tr>
<tr>
<td>English</td>
<td>Casinos</td>
<td>United States, Small Business Administration</td>
<td>20th century</td>
<td>Fiction/Non-Fiction</td>
</tr>
<tr>
<td>English</td>
<td>Minority business enterprises</td>
<td>United States</td>
<td>20th century</td>
<td>Fiction/Non-Fiction</td>
</tr>
<tr>
<td>English</td>
<td>United States</td>
<td>United States</td>
<td>20th century</td>
<td>Fiction/Non-Fiction</td>
</tr>
<tr>
<td>English</td>
<td>Bildungsromans</td>
<td>United States</td>
<td>20th century</td>
<td>Fiction/Non-Fiction</td>
</tr>
</tbody>
</table>

### Faceting: Before and After

**Before**

- **Subject/Gene**: Gambling, Casinos, Minority business enterprises, United States, Small Business Administration, Bildungsromans
- **Subject Region**: Nevada, Las Vegas, United States, Arizona, Black Mesa (Navajo County and Apache County, Ariz.)
- **Subject Era**: 20th century

**After (aspirational)**

- **Subject/Topic**: Gambling, Casinos, Minority business enterprises, United States, Small Business Administration
- **Subject Region**: Nevada, Las Vegas, Arizona, Black Mesa (Navajo County and Apache County, Ariz.)
- **Subject Era**: 1900-1999
- **Genre**: Bildungsromans

### What about all those duplicated headings?

Apply conditional logic in Solr mappings

- Prefer FAST for faceting
  - Use $2XX with $2 $fast

- Prefer LCSH in record view
  - Suppress FAST from individual record view if LCSH is present.
Individual record view

BEFORE

AFTER

Cataloguing with FAST

- For minimal-level cataloguing only
  - Replaces use of uncontrolled 653

- Pilot since February 2014
  - Using http://fast.oclc.org/searchfast/

Cataloguing with FAST

- Introductory presentation
- Characteristics of FAST
- Strengths and weaknesses
- Cheat sheet
- FAST “mindset”
FAST heading types

Based on LCSH, but separated into facets:
- Topical (650)
- Geographical (651)
- Event (641)
- Genre/form (655)
- Name (600, 610)
- Chronological (648)
- Uniform titles (630)

FAST strengths

Easier to learn
Better controlled
- All headings have authority ID (linked data)
Better suited to faceted discovery systems
- Individual facets more meaningful than 6XX subfields
Backwards-compatible
- Can be converted from LCSH
- OCLC now adding FAST

FAST weaknesses

Not as specific as LCSH, e.g.
- United States – Foreign relations – China
- Hugo, Victor, 1802-1885 – Musical settings
Some irregularities caused by source data
Editing tools not yet integrated
Long-term commitment by OCLC?
Key training points

Use what you find (all headings established)

Subjects do not cross facets

Observe distinction between topical and genre/form facets

Fewer application rules
  • No constraints on combinations of topical and geographical terms
  • Order of headings is not significant

Dates can be whatever you need to assign.
FAST vs LCSH: Round Two

1987: Subtitle, Susan C., 1957-
voice in the band: this doctor, her patients, and how the outlook on AIDS care changed from 1987 to 1999.

1987: Beginning of project:
- 1993: Two more years, or indefinite.

Cataloguing: state of play

1,500+ titles catalogued to date
Generally well-accepted by cataloguers
Will adopt as minimal-level standard in December
Further steps
- Maintenance
- Improvements to editing tools?
- LCSH vs FAST again

Vocabulary issues

Unestablished headings
Problematical mappings
Missing references (e.g., vernacular)
Scope notes
Event headings
What happens if a heading is not in FAST?

Extensions

For manual cataloging:
- Not in LCNAF:
  - 600 17 $a Folsom, Steven $2 fast/NIC
- In LCNAF but not in FAST:
  - 600 17 $a Childress, Eric $2 fast/naf $0 (DLC)no2005043559
- Legal but not explicitly established in LCSH:
  - 650 _7 $a Elephants $x Weight $2 fast/lcsh

For batch conversion:
- Accept headings that failed to validate, but add separate FAST extension:
  - 650 _7 $a Roriquez Rivera, Virginia $2 fast/NIC/NAC

Batch load

For loading into Voyager:
- Retain existing LCSH, but merge FAST headings into records that don’t have them.
- Convert $x form subdivisions to FAST 655s except for post-2000 LC records.

For loading to WorldCat:
- Add $2 fast/NIC, fast/naf, and fast/lcsh headings to WorldCat master records
- Do not add $2 fast/NIC/NAC to WorldCat
- OCLC Research will harvest $2 fast/naf and fast/lcsh (but not fast/NIC) and add to the FAST vocabulary

Maintenance

In theory:
- For new records lacking FAST, rely on OCLC master record updates
- OCLC will supply report of changed headings
- IDs should facilitate automation
- Not restricted to ILS data
FAST change report

Mapping subdivisions
Description and travel → Travel
Death and burial → Death and burial of a person
Economic conditions → Economic history
Pictorial works → Pictorial works (655)
Congress → Conference proceedings (655)
Relations → ???

Problematical mappings
“Relations” – OCLC solution

Relations [countries] → International relations
Foreign relations → Diplomatic relations
Foreign economic relations → International economic relations
Military relations → Military relations
Relations [religions] → Interfaith religions

Event headings

FAST conversion output:
- 001 _a Trenton, Battle of (New Jersey: 1776) $2 fast
- 003 _a Trenton $2 fast
- 004 _a Trenton, Battle of (New Jersey: 1776) $2 fast

FAST event authority:
- 008__a Trenton, Battle of (New Jersey: 1776)
- 010__a Trenton, Battle of (New Jersey: 1776)
- 020__a Trenton, Battle of (New Jersey: 1776)
- 022__a Trenton $2 fast
- 023__a Trenton $2 fast

034__d W0744435 $e W0744435 $f N0401301 $g N0401301 $2 geonames
043__a n-us-nj
046__s1776
053__0 American Revolution (1775-1783) $0 (OCoLC)fst01351668
053__0 American Revolution (1775-1783) $0 (OCoLC)fst01351668
053__0 American Revolution (1775-1783) $0 (OCoLC)fst01351668
053__0 American Revolution (1775-1783) $0 (OCoLC)fst01351668
053__0 American Revolution (1775-1783) $0 (OCoLC)fst01351668

Implied data: other possibilities?

016 _7 _a fst0103376152 OCoLC
040 _2 Sa OCoLCb engSc OCoLCsf fast
053 _0 Sa DD253Sb DD255Sc Germany
150 _2 Sa National socialism
450 _2 Sa National socialism
551 _2 Sa Germany $0 (OCoLC)fst01210272
Recap of Benefits
Better designed terms for faceting
Eliminates split values in facets
Potential to parse out facet types, e.g. subject vs genre
Better control of subjects in minimal level cataloguing
Aligns with linked data goals

IDs and Linkages

Two name forms
LC: Rembrandt Harmenszoon van Rijn, 1606-1669
ULAN: Rembrandt van Rijn
ULAN

VIAF: ULAN to LC

LC to FAST
Closing thoughts
Cataloguers addressing discovery
Value of research-driven organization
Editorial process
Having control over your tools
Getting there from here

Resources
Ed O’Neill and Lois Mai Chan, FAST: Faceted Application of Subject Terminology. Libraries Unlimited, 2010 (OCLC #240255331)
Kelley McGrath, “Facet-Based Search and Navigation with LCSH: Problems and Opportunities” http://journal.code4lib.org/articles/23
OCLC FAST web site: http://oclc.org/research/activities/fast.html
Cornell cheat sheet: http://lts.library.cornell.edu/lts/pp/cat/127FAST
Cornell FAST introductory presentation: http://ecommons.cornell.edu/bitstream/1813/34435/2/FAST_C%26M.pptx

Questions?