Enriching WorldCat with FAST

OCLC is systematically adding FAST (Faceted Application of Subject Terminology) headings to many WorldCat bibliographic records.

This document provides information about:

- What is FAST?
- OCLC’s addition of FAST to WorldCat bibliographic records
- Why OCLC is adding FAST to WorldCat
- How the LCSH-to-FAST conversion process works

What is FAST?

FAST (Faceted Application of Subject Terminology) is a fully enumerative faceted subject heading schema derived from the Library of Congress Subject Headings (LCSH), a widely-used subject-access vocabulary published and maintained by the Library of Congress. The development of FAST has been a collaboration of OCLC and the Library of Congress with advice from members of the ALCTS/SAC/Subcommittee on FAST.

The individual terms in the FAST vocabulary are divided into eight distinct categories or facets. In MARC, each facet is recorded in the appropriate 6xx field.

<table>
<thead>
<tr>
<th>FAST Facet</th>
<th>MARC 21 Field</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chronological</td>
<td>648</td>
</tr>
<tr>
<td>Corporate Names</td>
<td>610</td>
</tr>
<tr>
<td>Events</td>
<td>611</td>
</tr>
<tr>
<td>Form/Genre</td>
<td>655</td>
</tr>
<tr>
<td>Geographic Names</td>
<td>651</td>
</tr>
<tr>
<td>Personal Names</td>
<td>600</td>
</tr>
<tr>
<td>Titles</td>
<td>630</td>
</tr>
<tr>
<td>Topics</td>
<td>650</td>
</tr>
</tbody>
</table>

FAST facets are designed to be used as a set, but each may also be used independently.
The rules of FAST application are very simple: a cataloger or other user applies personal judgment to determine which attributes (e.g., topic, geographic name, event) are primary for the resource being described, whether the resource should be categorized as a particular form or genre, selects the appropriate FAST terms, and includes them in the bibliographic record. As with LCSH, the most specific but encompassing term is preferred.

FAST has been developed in large part to attempt to meet the perceived need for a general-use subject terminology scheme which is:

- simple to learn and apply;
- faceted-navigation-friendly; and
- modern in its design (e.g., useable in a linked data environment)

The full development of FAST has required several years and resulted in an eight facet vocabulary with a universe of approximately 1.7 million headings across all facets. Documentation has been made available in the form of a published monograph¹, and FAST has been made available by OCLC in various online tools and published as linked data².

Additionally, OCLC Research has made significant use of FAST for various research activities and in several prototypes, and these uses have led to improvements in FAST.

**OCLC is adding FAST headings to WorldCat records**

Beginning in September 2013, OCLC is systematically adding FAST headings to WorldCat bibliographic records. The records affected will already have LCSH assigned, and the routines will be applied only to records that are attributed as being English-language-of-cataloging.

WorldCat records enriched by FAST headings will display an “array” (i.e. group of headings from appropriate, complementary facets which taken together present key attributes of the content of a resource) of FAST headings such as this array assigned to a book (WorldCat #818953763) about the 2012 United States presidential election:

```
600 17 Obama, Barack. +2 fast +0 (OCoLC)fst00348231
600 17 Romney, Mitt. +2 fast +0 (OCoLC)fst00355923
650  7 Presidential candidates. +2 fast +0 (OCoLC)fst01075708
650  7 Presidents +x Election. +2 fast +0 (OCoLC)fst01075747
651 7 United States. +2 fast +0 (OCoLC)fst01204155
```


² See [http://fast.oclc.org/searchfast/](http://fast.oclc.org/searchfast/)
FAST-enriched WorldCat records will include “OCLCF” in the MARC 040 subfield “d” (ǂd), have one or -- more likely -- several 6xx fields with a second indicator of “7” and subfield “2” (ǂ2) that contains the string, “fast.” Additionally, subfield “zero” (ǂ0) will be present in the machine-process-added FAST headings -- this provides the FAST authority record number.

The subfield “zero” provides a machine-friendly means to link the heading to its corresponding authority record. This significantly improves the ability of systems to keep assigned headings correct and up-to-date, to provide easier machine indexing (indexes can be keyed directly to the identifier), and makes it easier for systems to express this identifier-explicitly-associated data in a variety of forms including as linked data. (Note that one facet, Chronological -- expressed in MARC as a 648 field -- is not controlled as it relies on syntax rules primarily, and this FAST facet has only a small number of authority records for certain named periods.)

The processes performing this FAST enrichment work in WorldCat will run as background processes and should not adversely impact system performance. OCLC will initially prioritize the enriching records originally created by the Library of Congress and then will apply the FAST enrichment process to records from other sources. Initially, records will be excluded from FAST enrichment until at least six months from date of entry have elapsed (this is intended to reduce the occurrence of after-initial-cataloging LCSH changes causing the FAST enrichment-assigned FAST headings to be out of synch with LCSH assigned to the respective record).

OCLC anticipates that the FAST enrichment of record batch process will take at least one year, and that FAST headings will be added to approximately 100 million records during the enrichment process. Following the initial enrichment run, OCLC will put in place maintenance routines to keep FAST headings updated in the respective records and enrich selected records newly added to WorldCat.

WorldCat records containing member-assigned FAST headings will not be modified by the FAST enrichment process. And members may modify FAST headings assigned by the batch enrichment processes just as they would any other headings appearing in the master record.

**Why is OCLC systematically adding FAST to WorldCat Records?**

OCLC Research and other agencies³ have been experimenting with using FAST for a variety of purposes and in a range of application for many years. FAST has repeatedly proven itself to be remarkably effective for clustering, indexing, analysis, and navigation of WorldCat data. And the release of FAST as linked data and in data dumps (including MARC authorities) under an

---
open license by OCLC, beginning in late 2012, has stimulated additional interest in and use of FAST by third parties.

Presently, FAST appears in several OCLC applications including:

- WorldCat Identities (FAST is used for the subject tag cloud that appears at the bottom of each WorldCat Identities Page)
- OCLC Classify (FAST terms are used for the “subject” index)
- In the experimental “xFinder” applications (e.g., Cookbook Finder) and some related works-view-based applications such as Kindred Works
- WorldCat.org linked data, and,
- OCLC has experimented with leveraging FAST as a collection analysis approach through the FAST/VIAF Centers work by OCLC Research.

These applications have access to FAST headings associated with WorldCat records in an internal, enhanced copy of WorldCat maintained by OCLC Research.

Going forward OCLC has plans to make FAST available in its cataloging tools. In particular, a FAST assignment feature will be released in early 2014 in an updated version of OCLC WorldShare Metadata Record Manager, which is currently available to WMS libraries and will be available to all cataloging subscribers in the future. Additionally, OCLC Research has begun research work to determine if FAST may potentially demonstrate value as a means to cross-vocabulary mapping.

Taken together all of this activity with FAST has demonstrated significant value and offers the promise of more. But it is challenging for OCLC and other interested parties to efficiently leverage FAST across more applications without having FAST headings present in WorldCat master records.

Enriching the shared, production version of WorldCat with classification and subject headings is an ongoing OCLC community activity. These enrichments are often done record-by-record by individual catalogers, though certain corrections and enrichments are done by machine processes. While errors will occur, overall the impact is a welcome, richer more useful WorldCat. And many records contain headings and classification assignments from a variety of schemes.

With this FAST enrichment project, FAST will join the many schemes enriching WorldCat records.
How does the LCSH-to-FAST conversion work?

OCLC Research has developed a set of conversion routines that can process assigned LCSH headings into an equivalent array of FAST headings.

In a series of steps, valid LCSH headings are read, and the corresponding FAST headings are generated. This initial FAST array is then further processed to remove any duplicate headings that may be generated.

Because FAST constrains each tag/field to be a single facet, a single, complex LCSH heading may map to several FAST headings, and in some cases, the shift from LCSH to FAST may present an equivalent concept slightly differently.

For example, a resource about a work by a personal author (or corporate author, or a work known simply by its title) would all map to the Titles facet in FAST. To illustrate, here is an example of an LCSH-to-FAST conversion of an LCSH Name-Title entry for Macbeth [in MARC21 markup]:

- **LCSH: 600 10 $a Shakespeare, William, $d 1564-1616 $t Macbeth.**
- **FAST: 630 7 $a Macbeth (Shakespeare, William) $2 fast #0 (OCoLC)fst01356136**

To illustrate a somewhat more complex conversion, here is a sample LCSH-to-FAST conversion (with MARC 21 tags supplied):

<table>
<thead>
<tr>
<th>LCSH</th>
<th>FAST</th>
</tr>
</thead>
<tbody>
<tr>
<td>600 10 $a Washington, George, $d 1732-1799</td>
<td>600 17 $a Washington, George, $d 1732-1799 $2 fast #0 (OCoLC)fst00178100</td>
</tr>
<tr>
<td>650 0 $a Trenton, Battle of, Trenton, N.J., 1776 $v Juvenile literature.</td>
<td>611 7 $a Trenton, Battle of (New Jersey : 1776) $2 fast #0 (OCoLC)fst01404297</td>
</tr>
<tr>
<td>651 0 $a Delaware River (N.Y.-Del. and N.J.).</td>
<td>648 7 $a 1776 $2 fast</td>
</tr>
<tr>
<td></td>
<td>651 7 $a New Jersey $z Trenton $2 fast #0 (OCoLC)fst01207908</td>
</tr>
<tr>
<td></td>
<td>651 7 $a United States $z Delaware River $2 fast #0 (OCoLC)fst01310316</td>
</tr>
<tr>
<td></td>
<td>655 7 $a Juvenile works $2 fast #0 (OCoLC)fst01411637</td>
</tr>
</tbody>
</table>

As can be seen, the conversion process has parsed the LCSH (Personal Name, Topic, Geographic Name) and mapped it to FAST Personal Name, Event, Chronological, Geographic Name, and Form/Genre facets.

[Note: This particular example illustrates some important, but nuanced differences between LCSH and FAST: 1) FAST has a concept of “events” for named events, 2) Rivers that cross

---

4 Note: FAST Personal and Corporate Name facets are drawn from the LC Name Authority File, and FAST Form/Genre headings represent an evolving selection of LCSH (including x55, $v, selected x50 terms) and LCGFT terms – as LoC shifts LCSH Form/Genre to LCGFT, FAST will be adapted to draw more heavily upon LCGFT.
multiple political units are reassigned to the next highest order unit that encloses the full length of the river, 3) The order of FAST headings - unlike LCSH - is not considered meaningful (OCLC’s LCSH-to-FAST conversion routine orders by its output by MARC tag as a convention).]

For a sample WorldCat record exhibiting an OCLC automated LCSH-to-FAST conversion, see #794359167.

**For more information about FAST**

Additional information about FAST is available here:
http://www.oclc.org/research/activities/fast.html