



CDS Announcement

Upcoming Record Cleanup Projects

The Library of Congress will soon be undertaking several data cleanup projects on bibliographic records:

1) CIP records older than 5 years for which the Library did not receive published copies will be systematically purged from the LC database. MDS-Books subscribers should expect approximately 35,000 delete records as a result of the older CIPs cleanup effort.

2) All bibliographic records containing the Alif character will be updated. The MARC 21 character encoding for the Alif was revised in March 2005 to change the Unicode mapping for the Alif from U+02BE to U+02BC. In an effort to clean up records to include the correct Alif value, many records with romanized text for Arabic, Persian, Hebrew, Korean, and many other languages will be updated and redistributed. Subscribers should expect approximately 300,000 records to be redistributed from this effort.

The cleanup efforts listed above will begin on or about June 30, 2009. Other cleanup projects will occur periodically resulting in a larger number of records than normal in future MDS issues.

Questions or concerns may be directed to:

Policy and Standards Division

email: policy@loc.gov

June 22, 2009

For more information about LC's MARC Distribution Services, contact:
Library of Congress ■ Business Enterprises ■ Cataloging Services
Washington, DC 20541-4910 USA ■ Phone: (202) 707-6100 ■ Fax: (202) 707-1334
Email: cdsinfo@loc.gov ■ Web: www.loc.gov/cds