

An Amicable Divorce: Programmatic Derivation of Faceted Data from Library of Congress Subject Headings for Music

Casey A. Mullin

Cataloging & Classification Quarterly, ISSN 1544-4554, Vol. 56, n.7, 2018, p. 607-627

La disponibilidad de los nuevos vocabularios facetados de la Library of Congress y las capacidades actualizadas de los formatos MARC han favorecido el desarrollo de los metadatos de música. La mayor oposición a esta política viene de los datos heredados, que vuelven los esfuerzos manuales retrospectivos imprácticos. Este artículo describe los esfuerzos de un equipo de bibliotecarios de música para separar encabezamientos de género/forma musicales de los LCSH en su terminología facetada correspondiente. Para lograrlo se utilizó la ingeniería inversa sobre la sintaxis que los catalogadores de música habían usado para crear estos encabezamientos. El resultado son unos descriptores facetados que se relacionan con los vocabularios facetados de la LoC y designan campos específicos en los registros bibliográficos de MARC. La guía de buenas prácticas de la Music Library Association ofrece un marco para articular esta actividad y evaluar los resultados. El proyecto comenzó analizando los encabezamientos de género/forma de LCSH, muy consistentes para identificar instrumentos y voces. El grupo decidió realizar un algoritmo basado en la sintaxis y puntuación usada en los encabezamientos. Una vez identificados los componentes, pueden repetirse en el campo 382 con sus correspondientes \$a para cada instrumento. El algoritmo se ve

más limitado al describir voces y bandas. Con las especificaciones para el campo 382 establecidas, la siguiente fase consistió en desarrollar especificaciones para la generación de otras facetas bibliográficas que incluyen género/forma, datos codificados, lugares asociados, características de la audiencia y características del creador/colaborador. Mapear estas facetas es más sencillo, ya que las correspondencias son habituales. Tras desarrollar todo el conjunto de especificaciones de algoritmos, se procedió a probar los resultados para detectar errores. Así se comprobó que algunos encabezamientos no generaban campos de datos facetados, lo que se pudo solventar refinando el algoritmo. Con este estudio se pretende crear basad de datos bibliográficas que permitan enriquecer lotes de registros musicales de una manera semiautomática. Aunque el proyecto se ha centrado en MARC, también debería ser útil en otros formatos como MODS, BIBFRAME o Performed Music Ontology.

Resumen elaborado por Antonio Rodríguez Vela