A Semantic Map of the last.fm Music Folksonomy

Biberstine, J. [1], Börner, K. [1], Duhon, R. J. [1], Hardy, E. [1], Skupin, A.[2]

[1]School of Library and Information Science, Indiana University, Bloomington, IN 47405

[2]Department of Geography, San Diego State University, San Diego, CA 92040, USA
Email: skupin@mail.sdsu.edu

What does the world of music look like? We present a visualization derived from a repository of user-generated tags attached to more than one million items within the social music website last.fm (http://www.last.fm/). The site enables users to discover new music based on their listening history and – crucial for our study – users can annotate music-related items such as artists and songs with arbitrary tags, ranging from categories like “rock” or “jazz” to event-related attributes (e.g. “seen live”) and affective utterances (e.g., “songs I absolutely love”). Tags also vary in scope, from rather broad categories, like “classical”, to finer distinctions, like “britpop” or “female fronted metal”.

The map offers viewers a mix of recognition, surprise, and discovery. Viewers have reported appreciation of the coherent patterns of hierarchical relationships among musical styles. The map is also notable for offering opportunities for discovering new musical categories, from the various flavors of “metal” to such niche areas as “shoegaze” or “drone”.

The project presented here advances the visual analysis of social media on a number of fronts. Together with a recent study visualizing two million biomedical documents, which generated several key technology elements used in our project, this is to our knowledge one of the largest self-organizing maps ever created in a single process (e.g., not hierarchically built), and certainly one of the most comprehensive semantic depictions of the world of music. A project of this scale would not have been possible without parallelization and supercomputing resources. Geographic Information Science concepts are at the core of how the project advances the analysis of social media data, which has traditionally been dominated by a network paradigm. Instead, our use of SOM leverages a view of social media items as existing in a high-dimensional attribute space, which calls for very different conceptual and computational approaches. Geographic metaphors and GIS technology features prominently in the conceptualization and rendering of map elements.