

ISMIR (International Symposium on Music-Information Retrieval)
Panel proposal: **Musical Data as Intellectual Property**

Barcelona, 10-14 October 2004 (panel: afternoon of 13 October)

General Background

Emerging music technologies confront uncertain prospects in relation to copyright and other forms of intellectual property protection articulated during eras of more limited capabilities. Most ideas underlying existing legislation are based on conceptions of a century ago. Each music technology of the twentieth century—piano rolls, recordings, radio, television, film, and multimedia—raised new issues and was met with further elaboration of rules designed to protect “musical works.”

Two developments of recent years have carried discussions of copyright into realms in which pre-existing mandates provide little direction to would-be providers of music technology. One of these is the atomization of musical information into codes that may represent specific features of music works without necessarily representing the works themselves as they are perceived on the page or the recording. The other is the use of the internet as a medium of the distribution of musical works in many media—sound files, MIDI files, notation codes, analytical codes. These can extend to graphical representations of structures, re-edited and recomposed music, digitally restored recordings and graphical images of musical manuscripts, and new musical compositions generated by algorithms which are imperceptible to the listener.

Schemes for the protection of “intellectual property” abound in discussions of hardware, firmware, and “digital-rights-management” software, but most proposed means for protecting physical products with musical content also diminish the ability of researchers to advance music technologies. For example, optical recognition software for musical scores works best on computer-typeset music, but almost all computer-typeset music is, by default, under copyright.

Music Retrieval and Intellectual Property

In principle, search- and query-researchers (who make up a substantial portion of the ISMIR community) should be unaffected by the fuzziness of legal provisions for electronic data. Many are engaged only in scientific projects, and what they provide to users is hypothetically limited to metadata fields, short snippets from musical works, and so forth.

Problems quickly accrue to this simplistic view when real-life scenarios are discussed. For the audio-research community, a fundamental issue is that protections on sound recordings vary widely in their terms and conditions, and also in their relationship to protections of printed music, from country to country. For the symbolic-data-research

community, uncertainties pertaining to how a “musical work” (in legal terms) relates to data representing that work (given that the data may be less complete, more complete, or significantly altered in relation to “the original” to suit a specific application environment). For the data-interchange communities concerned with either audio or symbolic data, the “core data” of a musical work seems more akin to the (patentable) root-stock of a plant than to the (copyrightable) irreducible substance of a work of art. Computer participation in the generation of new musical works raises even more profound and still less easily answered questions of musical identity.

On the surface, query researchers seem unlikely to be a threat to anyone. Their most immediate goal is the production of methodologies for locating specific musical content. At present much “music” search technology is actually based on textual metadata. Like optical recognition, this presupposes the use of recently manufactured sources (usually recordings) under copyright, for pre-digital media did not contain machine-readable metadata. There is no metadata, apart from what exists in library and manufacturer catalogues, for earlier media.

Much commercially funded research is motivated by schemes for data delivery via “push” technologies or music-recommendation services. Researchers cannot easily develop such technologies without ample access to the materials in their fullest form, for at this juncture there is much uncertainty as to the most effective and efficient methods for music retrieval, the most characteristic features for summarization, and so forth. Also, many strategies that work relatively well in a small, selected repertory are found to work less well in a larger, more varied data sample. It will be almost impossible to establish robust methods of musical retrieval without extensive access to large amounts of data.

Many fundamental questions pertain to the electronic use of materials in libraries and archives. There is, for example, no easy way to ascertain the copyright status of recordings without labels, editions of music without publication information (e.g., those produced in the former Deutsche Demokratische Republik prior to 1979). Much material which is relatively old (e.g., outdated in its editorial presentation) may still be under copyright. It may have been out of print for decades, but public use is still restricted.

Future Contributions

While many questions of intellectual property remain unresolved, music technologies promise to put many new tools at the service of those intent on protecting digital content. Digital-rights-management tools are appearing at a rapid rate. Perhaps more significantly, music-query capabilities promise to streamline music copyright by making content searches feasible at the time a new work is created. They promise to give prospective users easy access to out-of-copyright materials. They promise to simplify cataloguing tasks for both publishers and librarians. They can facilitate internet commerce by non-verbal means. Many researchers hope they will even make music searchable by “mood”.

A Hypothetical Case

To bring focus to this discussion, we pose a hypothetical case intended to represent some recurrent needs of the music-query community. Here is the case:

A loosely knit group of researchers (some working in industry, some in academic institutions) wish to form a repository of “musical works” for mutual access. Their aim in establishing this repository is to be able to compare their strategies by relating them always to the same database. The ultimate goals of their projects vary individually and range from providing descriptive metadata to “serving” complete digital scores and recordings to the public.

Each panelist is invited to give a **five-minute overview of his particular interests and a ten-minute presentation containing advice to this group** from the perspective of his or her area of expertise. He/she may recommend restrictions on content and terms of use, may suggest alternative methods for conducting the same research, may focus on one aspect of the scenario best suited to his/her expertise, and so forth. Discussion and questions from the floor will follow.

Total time allotted to panel: 2 hours.

Panelists

Charles Cronin, Law Librarian (and Musicologist), Columbia University, New York (US). Special interest in music plagiarism. Convener of the first Virtual Scores conference (Kernochan Center for Law and the Arts, Columbia University, May 2003).

Enric Enrich, Enrich Advocats, Barcelona (ES). Recording industry and film issues, e.g., in relation to Spanish dance music. Chairman of the Copyright section of the Barcelona Bar Association.

Masataka Goto, National Institute of Advanced Industrial Science and Technology, Tsukuba (JP). Originator of the RWC Music Database (2003).

Moderator:

Eleanor Selfridge-Field, Consulting Professor, Music and Symbolic Systems, Stanford University; administrator of the *MuseData* Database of electronic scores (www.musedata.org).