Richard P. Smiraglia
Palmer School of Library and Information Science, Long Island University,
Brookville New York 11548 USA

Music Information Retrieval: An Example of Bates' Substrate?

Abstract: Music Information Retrieval (MIR), and ISMIR annual conferences offer a rich panoply of intellectual and cultural diversity. We map the evolution of MIR using conference papers from 2000 through 2005. Results indicate tight thematic coherence in the domain around the problems of information retrieval and classification, and the locus of most research within computer science departments.

Résumé : Les conférences annuelles sur le repérage d'information musicale (MIR) et ISMIR offrent une riche panoplie de diversité culturelle et intellectuelle. Nous traçons le portrait de l'évolution du repérage d'information musicale en utilisant les communications des conférences de 2000 par 2005. Les résultats indiquent une correspondance thématique étroite dans le domaine touchant les problèmes de repérage et de classification d'information et dans la position de la plupart des recherches des départements d'informatique.

1. Introduction

Authors from Machlup and Mansfield (1983) to Bates (1999) have suggested that the basic structure of information science is inter- or pan-disciplinary, arising from divergent cultural origins in many distinct disciplines. Bates suggested that the intrinsic unity of information science lies in 'substrate'—that is, interest in the properties of information and its transmission that runs below the surface of all disciplines. One approach to the study of information science as an evolving field, therefore, is the analysis of evolving interdisciplinary domains.

The domain of Music Information Retrieval (MIR) is a relative newcomer, and one that illustrates cultural divergence, interdisciplinarity, and research activity. Is it possible that MIR represents the functioning of Bates' substrate? Bates suggested the proof would be functionality on problems of form-of-content, human information need, and informetric methodology, all across disciplinary boundaries and with a technology-driven focus. MIR has emerged from the interdisciplinary search for solutions to the problems of storage and retrieval of music. Centered around a sequence of conferences that began in 2000 as the International Symposium for Music Information Retrieval (ISMIR), the domain has spawned an active list-serv "music-ir," a smattering of test collections (http://php.indiana.edu/~donbyrd/MusicTestCollections.HTML), a shared bibliography (http://www.music-ir.org/research_home.html), a systems evaluation laboratory "IMIRSEL" (http://www.music-ir.org/evaluation/), a review in *ARIST* (Downie 2003), and a note in *The Economist* ("Music recognition software" Oct. 17, 2002).

ISMIR annual conferences offer a rich panoply of the intellectual and cultural diversity that constitute MIR. Musicologists, audio engineers, information scientists, librarians, and many others have worked through the challenges of MIR, often using the language of their own intellectual cultures. As their work has progressed, the shape of the independent contributions has evolved. A set of interesting questions, therefore,

surrounds the MIR domain, including the extent of interaction among participants, and the essential direction of the domain. For the field of information science, an even more interesting question is whether the interdisciplinary shape of MIR is representative in some way of the evolution of information science as a discipline, represented by the intersection of sometimes competing and sometimes cooperative discourse communities.

2. Mapping MIR

Several recent studies (e.g., Stephen 1999; Breitenstein 2003, Ram 2006) have used content analysis tools to seek out the parameters of evolving discourse communities. In particular, *WordStat* software has been used to generate comparative analyses to identify patterns in conceptual relationships across discourse communities. The software uses multidimensional scaling to generate images of nearness or distance among entities identified in a compiled taxonomy. The resulting analysis can be used to demonstrate divergence, convergence, or stasis, when applied across time within an emerging domain such as MIR.

The present study is a preliminary attempt to map the evolution of the complex MIR domain. A table of research papers accepted for the sequence of refereed MIR conferences from 2000 through 2005 was compiled. Invited papers, posters, and keynote addresses were excluded, yielding a collection of 257 research papers. The entire compilation was entered in an *Excel* spreadsheet, which was used to generate the initial analysis of author productivity across the six conferences. *WordStat* software was used to analyze titles of conference papers from the entire sequence of MIR conferences. Using word frequency statistics, the main themes of each conference were identified. To investigate the degree to which the domain might be representative of a substratum of information science, the institutional affiliations of the authors of representative papers were compiled. Finally, a tentative thematic map of MIR was produced, demonstrating the complexity of the domain.

3.0 ISMIR 2000-2005

Over the span of the six conferences a total of 257 papers were contributed. Most of these papers are available online at various ISMIR conference sites. The number of papers has increased dramatically over time, as has the breadth and complexity of the conferences. The breakdown is given in table 1.

Conference	***Contributions***
2000	10
2001	20
2002	31
2003	22
2004	59
2005	115

Table 1. Distribution of ISMIR Conference Papers

3.1 Most Productive Authors

In every domain, a small number of highly productive authors sets the agenda by producing a larger proportion of the research. Several power laws have been used to describe this phenomenon. Lotka's law of author productivity suggests that roughly 60% of the authors in a domain will produce only one paper of significance, while the remaining 40% will be divided into two groups, one of which is most prolific. In ISMIR as in every other domain this phenomenon is observable. The 257 papers from the six international conferences were arrayed by year, and each author was checked against the entire list to compile a list of authors who had contributed more than one paper. 68 authors were so identified. For each, each submission was pasted into a new array, which was then ordered by productivity in descending order. In cases of multiple authorship, all authors named were included; in some cases research teams were visible across the six conferences, but in other instances authors combined in different teams over time. Yet, roughly following Lotka's law, of the 68 authors identified as having presented multiple papers, the majority, 47 (69%), had submitted only 2 or 3 papers.

Twenty authors constituted the most productive group and these are identified along with their (self-attributed) academic affiliations in Table 2.

Author or Team	*# Papers*	*Academic affiliation*
Fujinaga, Ichiro	8	music
Tzanetakis, George	8	computer science
Pachet, François	8	computer science
Ellis, Daniel P. W.	8	electronic engineering
Downie, J. Stephen	7	information science
Raphael, Christopher	6	mathematics and statistics
Birmingham, William	5	electrical engineering and computer science
Pickens, Jeremy	5	computer science
Aucouturier, Jean-Julien	5	computer science
Martens, Jean-Pierre, Lesaffre M., Baets, De Meyer, Leman	5	electronics and information systems
Whitman, Brian, Adam Berenzweig	5	media lab
Pauws, Steffan	5	audio research
Rauber, Andreas	5	software technology
Gaël, Richard, Bertrand David	5	tele-communications
Bello, Juan Pablo	4	electronic engineering
Doraisamy, Shyamala	4	computing
Meek, Colin	4	electrical engineering and computer science
Dannenberg, Roger B.	4	computer science
Herrera-Boyer, Perfecto	4	audiovisual
Bainbridge, David	4	computer science

Table 2. Most Productive Authors/Collaborators with Academic Affiliations

As the table indicates, the majority come from computer science or engineering backgrounds, although the very top tier of productivity (above the shaded line) incorporates information science and mathematics as well. Three collaborative teams appear in this table indicating the strong productivity of their collaborations over time.

3.2 ISMIR Over Time : Taxonomy, Themes, Authors, and Collaboration

The content of ISMIR has change dramatically over time. From a small gathering in 2000 to the immense and complex 6th International Conference in 2005, the thematic content has intensified as has the breadth of the contributors' list. To capture this change over time, the six conferences were analyzed separately. Titles of papers were subjected to content analysis using *WordStat*. For each conference a taxonomy was constructed that was used as a data dictionary to filter thematic content. The change in the taxonomies is itself dramatic. Table 3 illustrates the initial taxonomy based on the papers from 2000 placed adjacent to the final taxonomy used in 2005. In order to save space, only a portion

ISMIR 2000 Taxonomy	***ISMIR 2005 Taxonomy***
Automat*	Acoustic index
Classification, Instrument	Algorithms, Geometric
Digitization	Audio front end
Digitization project	Audio signals
Information Retrieval,	Automat*
Information retrieval, Audio	Classification, Instrument
Information retrieval, Text	Classification, Audio
Instrument segmentation	Classification, Genre
MIR	Classification, Music
UF Music Information Retrieval	Cognitive models
Modeling	
Modeling, Language	…
Modeling, Music	
Monophonic music	
Music Content	
Music Content Description	Information Retrieval,
Optical Music Recognition System	Information retrieval, Audio
Orchestral music	Information retrieval, Text
Polyphonic Music	Information retrieval, Evaluation
Prototype	Information retrieval, Content-based
Retriev*	UF Performance study
Retrieval, Monophonic music	Information retrieval, Melody-based
Retrieval, Music	Instrument segmentation
Retrieving	
Searching	…
Structure	Modeling
Transcription	Modeling, Language
Web (WWW)	Modeling, Music
XML	…
UF Extensible Markup Language	Queries,
	Queries, Aural
	Queries, Humming
	SA Query by Humming
	Query by Voice
	Query-by-beat-boxing
	Recognition
	…
	Signal spotting
	Structure
	Long Term Structure
	Chord segmentation
	Microton*
	Testbed
	Theme
	Thematic extractor
	Tonality

	Transcription Voice characteristics Web (WWW) Works, Musical XML UF Extensible Markup Language

Table 3. Comparison of ISMIR Taxonomies from 2000 and 2005

of the 2005 taxonomy is displayed here. But the increased complexity of the offerings is readily apparent. In particular, the ten papers from the original symposium were focused on concepts such as "automatic music transcription" or "extensible markup language for MIR," all concepts easily described with simple terms. By 2005, the complexity of the vocabulary has produced 115 papers with topics such as "harmonic temporal clustering," "pattern extraction algorithm for abstract melodic representations" or "audio stream segmentation and classification." The increased complexity of the vocabulary is a sign of the richness of the domain, which in turn marks its success at the development of music retrieval mechanisms. What was a distant objective in 2000 has become a scientifically 'tweakable' reality by 2005.

WordStat was used to compile word frequency distributions for each year, which in turn were used to identify the main thematic threads of the domain. These appear in Table 4 (the proportions indicate the proportion of records, or titles, in which the term appears).

2000	***2001***	***2002***	***2003***	***2004***	***2005***
Music 80%	Music* 55%	Information retrieval 29.1%	Information retrieval 39%	Information retrieval 25.1%	Information retrieval 23%
Information retrieval 70%	Information retrieval 50%	Audio 19.4%		Automatic 16.7%	Audio 18%
Language 20%	Audio 15%	Automatic 19.4%		Audio 13.3%	Classification 15%
Modeling 20%	System 15%	Similarity 12.9%		System* 13.3%	
Instrument 10%		Polyphonic 9.7%		Polyphonic 11.7%	
		Analysis 9.7%			
		System 9.7%			
			all of the following 8.7% Acoustic, Content, QBH, Audio, Evaluation, Recognition, Automatic, Experiment, Similarity, Models, Classification, Polyphonic		Feature 8%
				Tempo 8.3%	Genre 7.1%

				Classification 6.7%	Polyphonic 7.1%
				Pattern, Search, Similarity 6.7%	Extraction, Similarity, Algorithm, Automatic 5.3%

Table 4. Conference Themes According to Word Frequency by Year

Two factors are immediately apparent. Most obvious, of course, is the growth of the breadth of the domain over time, and the rapid increase in complexity. "Music," and "Information retrieval," account for the majority of thematic content in the beginning, but by 2002 "music" is replaced by words that describe its replicable (and retrievable) components. This leads to the second obvious factor, which is the increase in granularity over time. This is most dramatic in 2003, when a very short upper tier is followed by a large cluster of a dozen terms all carrying the same word-frequency weight. Notice also that "polyphonic" replaces "music" as a key descriptive term over time—another sign of increasing granularity in ISMIR.

Finally, authors and collaborative teams associated with the highest-frequency terms in each year were identified together with their academic affiliations. These appear in table 5. A quick glance confirms that the academic affiliations are overwhelmingly in

2000	***20001***	***2002***	***2003***	***2004***	***2005***
Perry – digital libs	Birmingham – electrical engineering and computer science	Dannenberg et al - computer science	Downie – information science	Downie et al. – information science, supercomputing	Pickens and Iliopoulos – computer science
Pickens – computer science	Holger – computer science	Tzanetakis et al - computer science	Lesaffre et al – musicology, electronics and information systems, mathematics	Adams et al. – electrical engineering and computer science	Typke et al – computer science
Logan – Compaq	Nishimura – computing	Doraisamy et Rüger - computer science	Shifrin and Birmingham – electrical engineering and computer science	Zadel and Fujinaga - music	West and Cox – computing sciences
	Durey – electrical and computer engineering	Pickens et al - computer science, electrical engineering, music	Ukkonen et al – computer science	Taheri-Panah and MacFarlane - systems	Lidy and Rauber – software technology and systems
		Cooper and Foote - FX		Casey and Crawford - computation	Mandel and Ellis – electrical engineering
		Peeters et al - IRCAM		Yoshii et al. - informatics	Fiebrink et al. – music

					technology
		Wang et al – computer science		West and Cox – computing sciences	Pampalk et al. – artificial intelligence
		Pauws and Eggen - Philips		Lubiw and Tanur – computer science	Wei et al. – automation
				Typke et al. – information and computing sciences	Bray and Tzanetakis – computer science, music

Table 5. Major Thematic Contributors Academic Affiliations, by Year

computer science. However, this is a very subjective map of ISMIR, encompassing only approximately 1/8 of the total contributions. Nevertheless, from year to year, those contributing papers in the major thematic groups are predominantly from computer science, although (as we saw before) their emphasis is predominantly on information retrieval.

3.3 A Map of MIR

A preliminary map of MIR can be drawn based on the data in this study. To do so, the agglomerated data files were entered into *WordStat* for analyses similar to those demonstrated above. The results give us a tentative view of the domain. The word frequency map is led with "information retrieval" at 29.6%, "audio" at 16%, "automatic" and "classification" hover around 10% with lesser weights associated with the granular terms such as "polyphonic" and "similarity." Information retrieval remains the clear goal of the domain with sound the major focus. Classification has emerged as a central technique. The most prominent author-contributors are those named in table 2 (above). They work in diverse domains, primarily rooted in computer science and engineering, but with significant contributions from information science, mathematics, and music.

No evidence of international origins or cooperation has been presented in the present paper. However, the authors of the 257 papers come from every major industrial nation in the world. True international collaboration is rare—most teams come from a single institution or corporation—but not unheard of. ISMIR is a truly international enterprise.

WordStat has the capacity to produce via a three-dimensional map of the concept space under analysis. These maps have no fixed points in time, but they represent the relationships among concepts in a domain; in particular they show the probability of concept co-occurrence. This remarkably dense map appears below in Figure 1.

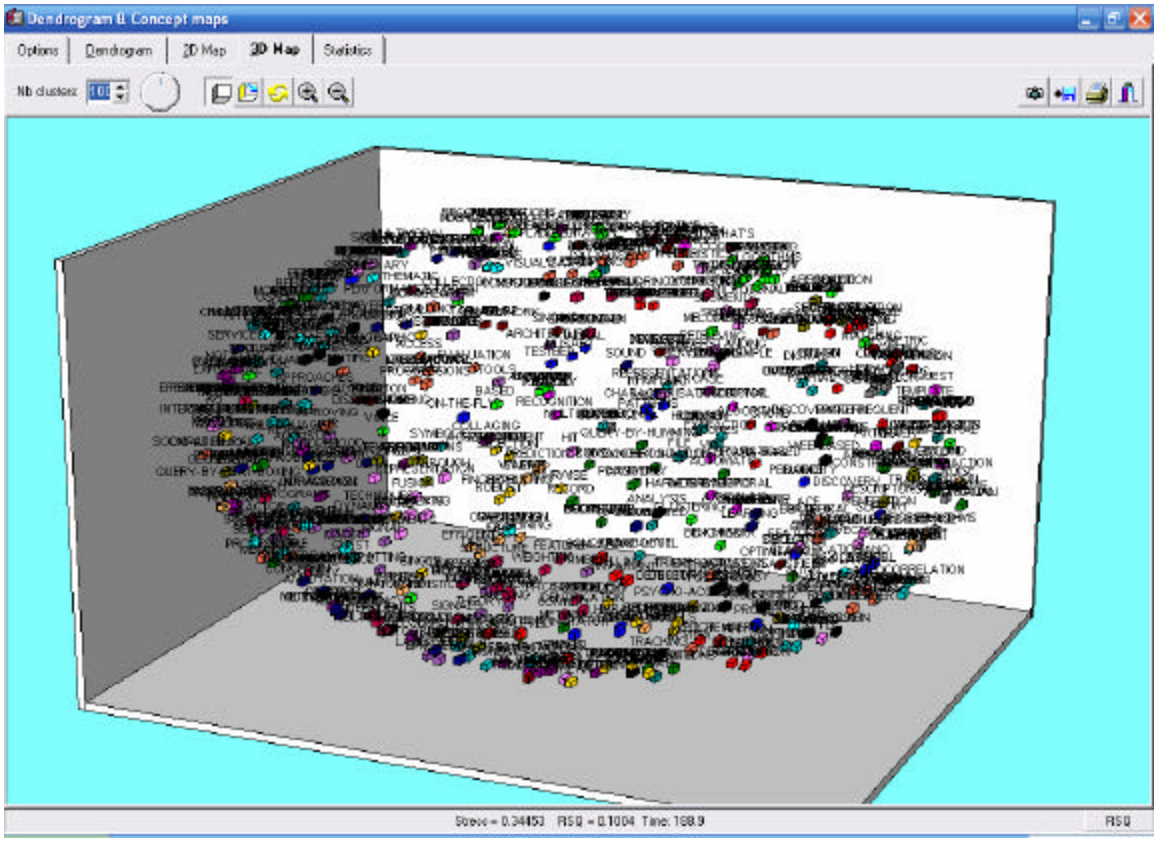

Figure 1. 3 -D Concept Map of ISMIR

No doubt the reader will not be able to make out distinctive points in the map, although the elliptical shape is clear. On the left are various MIR functions, such as "query-by-beat-boxing" or "disambiguating." On the right are fundamentals such as "template" or "discovery" or "autocorrelation," and in the vast center we find "query-by-humming," "analysis," "representation," and "recognition." On the upper edge we see "visualization" and on the lower edge "tracking." Thus the newest components appear at the left end, the persistent fundamental technologies on the right, and the highly valued functions in the center with experimental applications rotating around the edges. The impression of the whole domain, however, is blurred. This is important, because it is a reflection of the coherence of the domain. We have seen the methodical (if rapid) development of the domain by a process of classification, which leads to increasing granularity—one discovery begets a taxonomy, each element of which begets yet another discovery—and in turn the relationships beget hypotheses and experiments. And the experiments have been largely successful, with MIR systems now ub iquitous in public life. Yet the problems of MIR are far from all solved. At any rate, what we see in this diagram, as in the rest of this preliminary study, is a very coherent paradigm forming in this youthful domain.

4. Conclusion: Bates' Substratum?

The growth in the number and complexity of papers over time is an indication of the success of both the ISMIR conferences and the MIR movement. The shift from single contributors to a predominance of research teams places the domain among the hard sciences. The success of MIR—the contributors have achieved great progress in a short time—is mirrored in the increasing complexity of the contributions over time, as well as

in their increasing topical granularity. As major phenomena are identified, their components become topics of inquiry.

The majority of papers are focused on one or more aspects of information retrieval, which would place the domain squarely within the purview of information science. Another major theme is classification: both classification of empirical observation and classification of entities for retrieval. This would seem to place the domain within the realm of knowledge organization, one of the major components of information science. On the other hand, contributors to the domain are overwhelmingly from computer science. To be sure, they have significant partners from music and musicology, information science, mathematics, and engineering. But the academic locus of the contributors is computer science.

Future research is clearly called for. For one thing, this very tentative analysis has only begun to sketch the parameters of this developing domain. More analysis is needed including better coordination of the data dictionary. Content analysis would be richer if conducted with abstracts instead of titles. Some citation analysis might be useful to discover whether there is any inter-disciplinarity or whether the contributors are still keeping to their own domains. It seems early for co-citation analysis, with only six years of formal data. However, a promising question will be to what extent these scientists have begun to share information across their domains of origin.

Is MIR an example of the functioning of Bates' substrate? If so, the domain should be centered around problems of content of form, human information need, and methodological coherence. Results of this study suggest that, while it is too early to answer the question definitively, this domain bears watching. This tentative analysis suggests that topically the domain falls squarely within information science. And, the domain seems to be very coherently focused. Despite the increasing granularity we have seen in conference papers, the research front is closely knit around a set of singular problems. Downie projected this result (2003, 325) when he commented on the success of the ISMIR conferences in overcoming an observed disciplinary fragmentation. However, we have also seen that the majority of the contributors come from computer science and other disciplines. What does that tell us about this new domain? Is its locus within information science? Bates (1999, 1048) remarked on the catalytical nature of technological developments in information organization and access. A century ago, music information retrieval relied on the human ear and memory. Today, machines have begun to supplant that memory. The quantum leap represented by the library catalog over against the scholar's memory is replicated here—no longer must a set of designated humans memorize or transcribe all music to serve as intermediaries in retrieval. Now it is possible to hum a few bars and receive a variety of appropriate responses from a mechanical system. That is the success, preliminary as it is, of music information retrieval.

Works Cited

Bates, Marcia J. 1999. The invisible substrate of information science. *Journal of the American Society for Information Science* 50: 1043-50.

Breitenstein, Mikel. 2003. *Toward an understanding of visual literacy: Examination of conference papers of the International Visual Literacy Association.* PhD. dissertation, Long Island University.

Downie, J. Stephen. 2003. Music information retrieval. In *Annual review of information science & technology* 37, Medford N.J.: Information Today, pp. 295-342.

Machlup, Fritz and Una Mansfield. 1983. Cultural diversity in studies of information. In *The study of information: Interdisciplinary messages,* ed. Fritz Machlup and Una Mansfield. New York: J. Wiley, pp. 3-56.

Ram, Jacqueline L. 2006. *Human factors and ergonomics: analyzing synergy*. PhD. dissertation, Long Island University.

Stephen, Timothy. 1999. Computer-assisted concept analysis of HCR's First 25 Years. *Human communication research* 25: 498-513.