Redefining the "S" in ISMIR: Visualizing the Evolution of a Domain

Richard P. Smiraglia

Palmer School of Library and Information Science
Long Island University
720 Northern Blvd.
Brookville NY 11548 USA
(516) 299-2174 voice
(516) 299-4168 fax

ABSTRACT: Music Information Retrieval is an evolving and highly productive domain. In 2008 the domain assumed the formal structure of a named society of scholars. Visualization of the domain occurs through author co-citation analysis of conference papers and journal articles from 2000 to 2008.

This paper is about the continuing evolution of a new global scientific domain, and about how information science can use the tools of domain analysis of a discourse community to understand the expansion and contraction of theoretical nodes. As an ongoing case study, the music information retrieval community is analyzed.

1.0 Visualizing the Evolution of a Global Domain

Music Information Retrieval (MIR) is a vibrant, evolving, and global domain. At the center of the movement is the interdisciplinary search for solutions to the problems of storage and retrieval of music. The domain has been highly productive, yielding test collections, a shared bibliography, a systems evaluation laboratory "IMIRSEL," an *ARIST* review (Downie 2003) and a note in *The Economist* ("Music recognition software" Oct. 17, 2002). Since 2000, annual international conferences have attracted a rich mix of scholars focused on the problems of music information retrieval. Papers presented at the International Symposium on Music Information Retrieval, or ISMIR, cover the gamut from copyright and subject headings to deconstruction of musical sounds and classification of the parts for information storage and retrieval. That is, the ISMIR attracts librarians and musicologists as well as information scientists and engineers. The group has not only been successful at producing music retrieval systems, it has self-consciously constructed the parameters of its own domain. At the close of 2008 the "S" in ISMIR was changed from "Symposium" to "Society," as the formerly nascent interdisciplinary domain assumed the formal structure of a named society of scholars (Fingerhut 2008), incorporated in Canada.

Techniques for domain analysis have many uses in information science, many related to the generation of knowledge organization systems. The literature of a domain--the record of its accomplishments as well as the chief vehicle by which domain knowledge is officially communicated—can be used to generate analyses that help us visualize the intellectual function of the domain as well as its internal social networks. Information scientists use a variety of techniques to "visualize" literatures; that is, to reveal the contours of progress over time, to identify research fronts and evolving theoretical paradigms. These techniques, described fully by White and McCain (1997), may yield a variety of maps that enhance the understanding of a domain. In particular, visualization is helpful to identify evolving or emerging domains. Hjørland (2002) enumerates steps that provide information about a domain, including bibliometric studies and empirical analyses focused of terminology. Tennis (2003) describes two axes for the operationalization of a domain: 1) areas of modulation—definitions that modulate the domain by stating its extension; and, 2) degrees of specialization, including "focus" and "intersection"—ways to qualify a domain by increasing its intension.

But we also can learn about the communication of knowledge within, between, and among discourse communities through the use of domain analytical tools. One interesting question is how exactly does an interdisciplinary domain evolve? The answer lies tantalizingly close to the visualization of the growth of MIR as a domain. There are two essential and related components. The first is the intense focus on a technological development. Hjørland (2003) refers to the importance of specific technologies as catalysts for developments in knowledge organization; if we understand the conversion of musical signals from aural information to classified data for retrieval, we can see how the technological development has driven the evolution of MIR as a domain. But we also can understand this technological development as constituting the essential problem-set for the structure and life cycle of a domain. To extend our visualization in this direction we can appeal to the work of Collins (1998), who suggests the vitality of a domain lies in the richness of its research front, especially when the problem-base continues to yield new problems that generate new or rival approaches. Collins also suggests that a law of small numbers operates to restrict the extension of a domain, such that the number of theoretical nodes is limited, which leads not only to rivalry but also to synthesis. Finally, Collins also suggests that domains concretize and become conservative as a means of self-preservation.

The CAIS 2009 call for papers is focused on "Mapping the 21st Century Information Landscape." The CFP asks "Has globalization produced a "sea-change" in our understanding of the role and function of information institutions?" What better insight than the visualization of the evolution of a global domain? We will see that MIR continues to evolve around its core technological challenge, and that the extension of the domain both concretizes and shifts as a part of this evolution. And we can understand, the 2008 shift from "symposium" to "society" as the conservative, concretizing, self-preserving action of the domain. The analysis of this global, successful, scientific domain, can yield clues for mapping the 21st century information landscape.

2.0 The Growth of MIR

An initial "visualization" of MIR appeared in Downie's 2003 *ARIST* review. Perhaps the most visible, and arguably the most productive venue for MIR is the ISMIR annual conference. Scholars from diverse methodological and intellectual backgrounds came together to shape the new domain by outlining diverse strategies for achieving their common technological goal. Their success is mirrored in the papers accepted for the annual conferences, which are subject to double-blind refereeing. The number of papers submitted and accepted each year has grown dramatically from the beginning, from 10 in 2000 to 127 in 2007. Table 1 shows the distribution of papers across time, and Figure 1 gives a visualization of the dramatic growth.

Conference	*Contributions*
2000	10
2001	20
2002	31
2003	22
2004	59
2005	115
2006	61
2007	127
2008	24

Table 1. Distribution of ISMIR Conference Papers

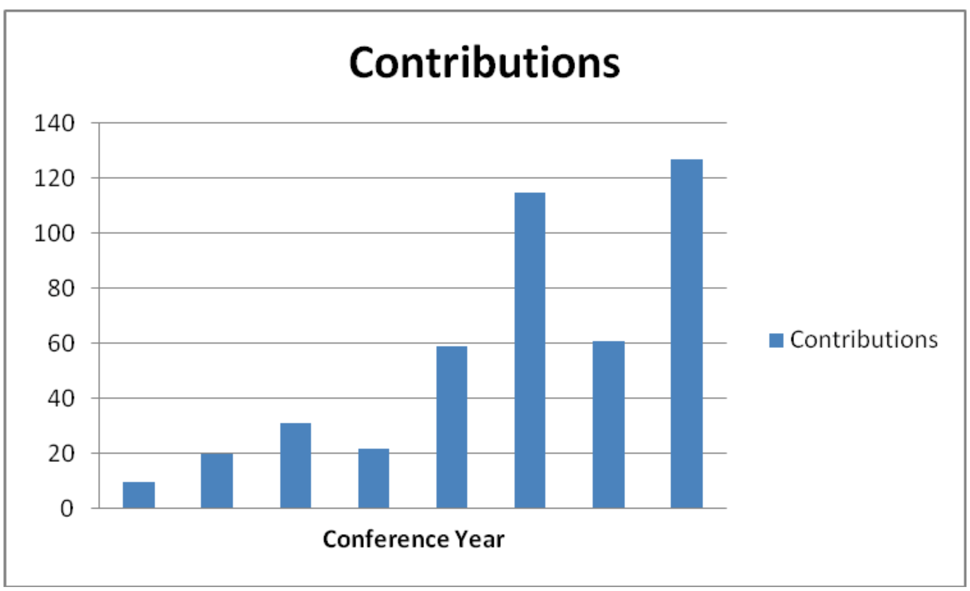

Figure 1. Growth of ISMIR Contributions over Time

Immediately apparent is the dramatic jump in papers that took place in 2004, when the number more than doubled. Also apparent is the small number of full-length papers in 2006 and the larger totals for both 2005 and 2007. At the 2007 conference papers were accepted in both “long paper” and “short paper” categories, which likely accounts for this disparity. Overall, the trend as the domain expanded was a doubling of contributions in the first year and then every third year after. What we see here is the repeated tug and pull of Collins’ notion of expansion and contraction of problem sets. As the number of theoretical nodes increases the domain contracts around those that are most promising. But following each contraction is a rapid expansion as new and rival approaches are introduced. In 2008 (not pictured in the figure) the number of conference papers was constrained dramatically—to 24. But there were 25 presentations in a session called "Late Breaking Demos," and the number of poster sessions also expanded dramatically to 18 sessions with 105 posters.

3.0 Thematic Granularity as the Domain Expands and Contracts

In 2006 co-word analysis of ISMIR conference proceedings revealed a thematic map of the MIR domain (Smiraglia 2006). In that paper we saw rapidly increasing granularity of thematic content. From simple musical and information retrieval terminology used in the first conference in 2000 to the incredible complexity of "beat-boxes" and "harmonic temporal clustering" in 2005, we watched the domain tend to its technological and scientific challenges in a direct way. This co-word analysis revealed information retrieval and sound as the axes of modulation defining the extension of the domain. Two technologies, classification and system design, then revealed the degrees of specialization, which are the axes of intension in the domain. The simple taxonomy in Table 2 was derived from the 2000 papers.

- Automat*
- Classification, Instrument
- Digitization
    - Digitization project
- Information Retrieval,
    - Information retrieval, Audio
    - Information retrieval, Text
- Instrument segmentation
- MIR
- Modeling
    - Modeling, Language
    - Modeling, Music
- Monophonic music
- Music Content
    - Music Content Description
- Optical Music Recognition System
- Orchestral music
- Polyphonic Music
- Prototype
- Searching
- Structure
- Transcription
- Web (WWW)
- XML

Table 2. ISMIR 2000 initial taxonomy

By 2005 the taxonomy had grown so dramatically that it was too large to display in a single table. By 2006 ISMIR had developed its own meta-level taxonomy by which papers were assigned to sessions. The latest iteration of this taxonomy from 2008 appears in Table 3.

Analysis	
	Music summarization
	Music transcription, Automatic
	Performance analysis
Computational musicology	
Content-based retrieval	
	Categorization
	Similarity
Data exchange	
	Archiving
	Evaluation
Feature representation	
Harmony	
Knowledge representation	
	Metadata
	Tags
	Web-mining
Melody	
MIR	
	Methods
	Platforms
MIREX	
Music recommendation	
	Organization
	Recognition
	Visualization
Musical expression and meaning	
OMR	
	Alignment
	Annotation
Rhythm and meter	
Social and music networks	
Timbre	
User interfaces and visualization	

Table 3. ISMIR 2008 taxonomy

We see here confirmation of the expansion and contraction of the field. As theoretical nodes become fixed (as evidenced in the constriction of accepted full-length papers) the taxonomy becomes more sophisticated. But this is accompanied by an expansion of rival approaches to problem solving. This was first observed in 2003 (Smiraglia 2006, 5) where word frequency analysis identified twelve new themes representing approximately 60% of the terms in the distribution. We see it again in 2008 in the dramatic increase in the number of "late-breaking" and poster presentations.

4.0 Author Co-Citation Analysis

The present paper pushes forward by employing the techniques of c author co-citation analysis. Citation analysis helps us understand the interrelationships among researchers in a domain, because in theory, their citations reveal the intellectual infrastructure in operation. Citations serve as markers that point to the most recent, relevant research, as well as to the classic tomes that define a domain. The number and currency of citations can be used to classify a domain as scientific, humanistic, or somewhere in between. Citation practice suggests ISMIR works in the manner of a hard science. With the exception of occasional review papers, the majority of ISMIR papers have relatively few and predominantly recent citations. The mean number of citations per paper was 17.2, with a range from 5 to 66, and a mode of 12. Most of the works cited were relatively recent; the mean age of citation was 7.3 years, with a range from .5 to 17.1. Interestingly, the number of citations fell to a mean of 3.5 in 2003 from the initial 2000 mean of 5.6; by 2005 the mean had returned to the same level, 5.7. This drop coincided with the year in which the granularity of conference content increased so dramatically. In 2008 the mean age of citation had dropped again to 4.5. Also, the majority of papers at ISMIR are collaborations; there are few papers by individual authors—only 2 in 2008—which is also indicative of a hard science.

Co-citation, which occurs when an author is cited by two or more authors in common, reveals clusters, sometimes called "constellations" (White and McCain 1989, 146). Such clusters demonstrate consensual views of key authors, and provide evidence of influence that has been brought to bear, for whatever reason, on the citing works. Essentially a qualitative method, co-citation analysis begins with the construction of a coherent set of essential papers. For the present analysis, two approaches were taken: a0 author co-citation analysis among the ISMIR conference papers; and b) author co-citation analysis in the journal literature for the same set of authors.

For the first stage, the twenty-seven authors who were identified as most cited overall were used. For each paper, every citation was copied to a spreadsheet for analysis. Because the domain is so small and barely emergent, self-citation was not eliminated, as is often done to rule out self-promotion. Co-citation was recorded in a spreadsheet. In the end (to produce a legible plot while retaining the shape of the domain) 12 authors are represented in the multi-dimensional scaling (MDS) plot produced by SPSS in Figure 2 below.

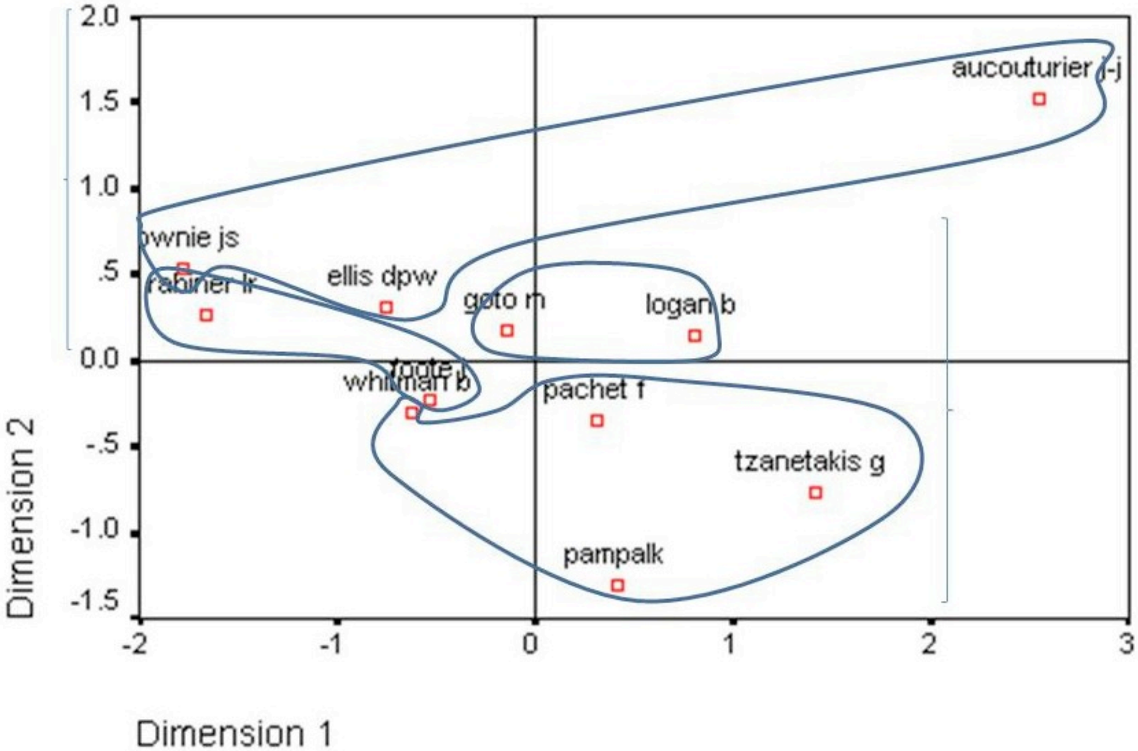

Figure 2: Co-citation *by* the core authors *of* the core authors within the conference papers

It is important to remember when viewing an author co-citation plot that what is visualized is the perception of similarity by the citing authors, in this case, authors of conference papers over the span 2000-2008. As might be expected in a narrowly-defined theoretical domain, co-citation of the core authors is abundant, and the proximity of the members of the clusters are resonant with the narrow technological focus of MIR. The Dendrogram indicates that the plot shows two large clusters, each containing two smaller clusters. At the right in the center and below we have Goto and Logan clustered together with Pachet, Whitman, Pampalk and Tzanetakis. In the upper part of the plot we have Rabiner and Foote clustered together with Downie, Ellis, and Aucouturier.

At opposite polls left and right we see J. Stephen Downie and George Tzanetakis. Downie is the nominal founder of the domain, having called together the first ISMIR in 2000 and having led the drive to develop the running MIREX (Music Information Retrieval Evaluation eXchange) evaluation project. An information scientist at the University of Illinois at Urbana-Champaign, he was the author of the ARIST review in 2003. Much of the citation of Downie is classic reference to his original catalytic paper. Tzanetakis is a computer scientist at the University of Vancouver, who set an outline for audio analysis in 1999, and who was involved in the drive to develop query-by-humming. His distance from the cluster is also a reflection of deference by the domain to his agenda-setting early work, but his proximity alludes to his continued work with music visualization and query-by-humming. We also see some distance from the central cluster represented by Jean-Julien Aucouturier (a computer scientist at Temple University's Japan campus, with experience at Sony labs, artificial intelligence, and robot manufacture), here separate from his frequent co-author François Pachet (an engineer with Sony CSL in Paris, working on electronic music distribution, spatialization, and musical perception). Aucouturier and Pachet are authors of a frequently-cited article from 2004 titled "Improving timber similarity, How high's the sky?" Dan Ellis (Columbia University), wrote a frequently-cited paper on chord-segmentation in 2001, and more recently has been involved with melody transcription; he also is associated with a MIR special section and a special issue of *IEEE Transactions Audio, Speech, and Language Processing*. Together with Rabiner and Foote these upper clusters represent the classic basis of MIR. The two lower clusters, anchored by Tzanetakis, represent the engineering component of the domain.

For the second stage of analysis the same authors were searched in Thompson-Reuters *Web of Knowledge*^{SM} and co-citation figures were plotted. A very different map emerged, as seen in Figure 3 below. For one thing, because co-citation was much less frequent in the journal literature, the number of authors represented dropped to 26. The remaining of plot is a good fit for the data, with low stress and a high R-squared (Stress = .10507 and RSQ = .97048).

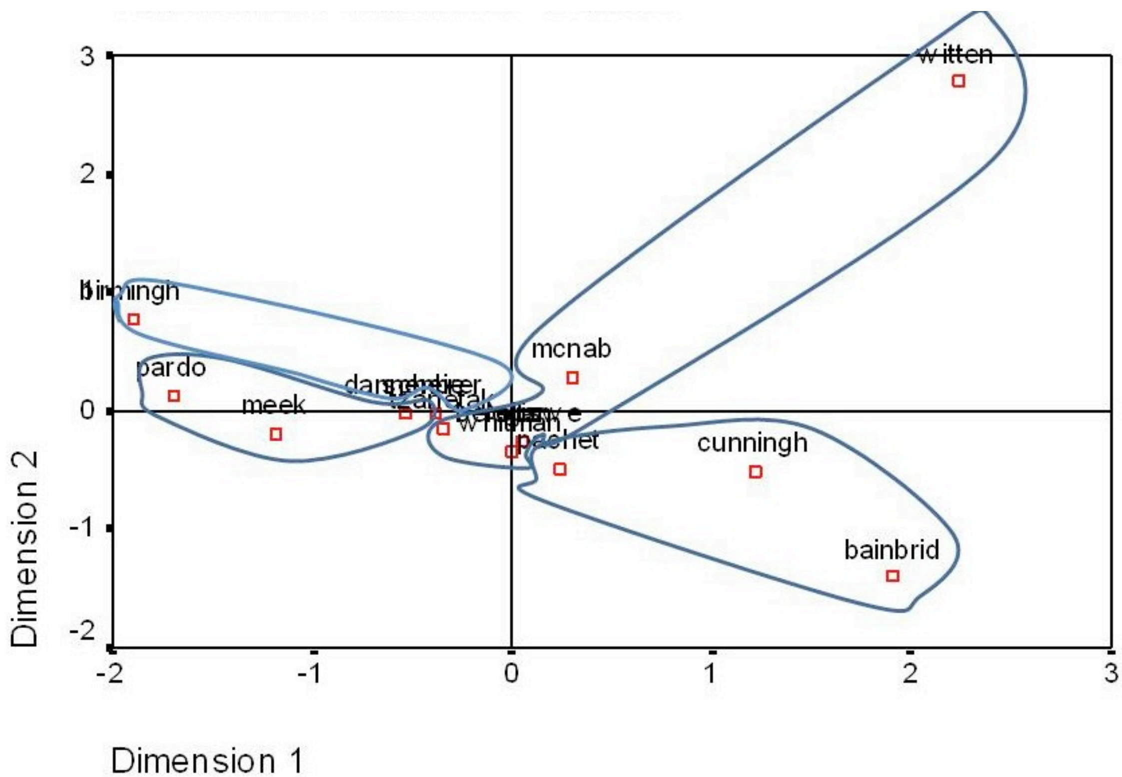

Figure 3: Co-citation of the core authors in journal literature

The accompanying Dendrogram reveals four distinct regions in the plot. Remembering that we are looking at the perceptions of this community held by those who cite them, we returned to the *Web of Knowledge*^{SM} to analyze the source of this perceived similarity. Curiously, authors such as McNab and Cunningham, who were dropped from the conference plot for low co-citation, appear in this plot. Beginning in the lower right quadrant, we see a cluster that includes Pachet with Bainbridge and Cunningham. Bainbridge is well-known for work on digital music libraries and Cunningham for multi-lingual music information seeking. This quadrant is characterized by information seeking and the design of digital libraries. To the far left is a small cluster that includes Pardo and Meek but stretches toward the center to include Tzanetakis. Pardo is known for work on search engines and on the test-bed; he is a frequent co-author with Tzanetakis. Meek is known for work on sung music queries. This cluster has some of the largest co-citation totals, and its focus is query-by-singing. Immediately above it is another small cluster including Birmingham, Dannenberg and Scheirer. Birmingham is also known for work on query-by-singing, Dannenberg for the query-by-humming test-bed, and Scheirer from the Bose Corportion gave an oft-cited paper on metadata. This cluster is about evaluation of the query-by-humming test-bed. The final cluster slashes from the upper right toward the center, stretching from Witten to McNab, Whitman, and Berenzweig in the center. Interestingly, within this cluster Berenzweig and Witten are not co-cited, nor are Whitman and McNab. This cluster is about the topology of music, including segregation algorithms: in other words, music information storage for retrieval.

The research front is closely knit around a set of singular problems that converge on the technology of query-by-humming. The difference between the journal and conference paper plots suggests both the volatility of the new emerging domain and the traditional scientific core of the more public segment, represented by peer-reviewed journals. In Figure 2 we are looking at the collective perceptions of the core ISMIR authors; the result is a tight cluster focused on development of MIR systems, with classic reference to Downie's original catalytic paper. With Figure 3 we see the domain as it appears to the larger scientific community, for whom it represents the image of a cutting-edge research front, but with more thematic differentiation. We also have confirmation here of the constriction of the domain toward a small number of core theoretical positions around which new developments can take place.

5.0 Conclusions

At the most basic level we have seen the success of the MIR community reflected in these analyses. In a very short time—less than a decade—a group of interdisciplinary problem-solvers was called together and created a technology that promises to revolutionize the digital music milieu. The appearance of the Shazam™ application on the iPhone is evidence of the coming ubiquity of MIR in the popular sector, and testimony to the success of the domain (Hilton 2008). Furthermore, the domain has a firm foothold in the information retrieval domain. Wise leadership swiftly implanted the essential ingredients of bibliography, interdisciplinary discourse, test-bed development, and evaluation trials to see that the technological problem was met with workable, if theretofore unimaginable, solutions. In 2006 I asked whether MIR represented the invisible substrate of information science, and with this paper we have a better answer—indeed it does. MIR lies at the concurrence of at least three disciplines—musicology, engineering, and information science—and while it is undoubtedly bound in its extension by musicology and engineering, the foundation of the domain comes from information retrieval.

Perhaps of more interest for information science is the continued monitoring of the evolving MIR domain. Two prior studies (Downie 2003, and Smiraglia 2006) surveyed the opening years of the effort and identified the essential theoretical positions—classification and similarity, segmentation, query, and evaluation. The current paper shows how those positions have expanded and contracted as the domain first solved its technological quest, and then burst out twice into a panoply of new developments. We see, then, a reflection of Collins' (1998) law of small numbers, which suggests that no domain can handle more than six positions at a time. Whenever the number increases, conservative contraction takes place to reduce the number of positions to a manageable level. We also see the dichotomy between the incredible richness of the MIR research front and the concretizing conservativeness of the theoretical base. This was noted explicitly in an email from the MUSIC-IR listserv in 2007; the author (Berman 2007) wrote: "I tend to agree that there is a bias in MIR circles that favors the algorithmic complexity of the solution over the interestingness of the musical problem." Finally, we see the self-preservation instinct in the concretizing of the domain by turning it from a wide-open interdisciplinary symposium to a bylaws-and-officers formal society.

What does this tell us about the 21st century information landscape? First that it is highly volatile, with developments taking placed rapidly in increasingly constrained timeframes. Second that it is incredibly rich and rife with problems in need of solution. As we move further and further into the reality of the virtual world made possible through the evolution of digital technology we will see more and more of these new domains exploding on the scene with energy and potentially lucrative industry contracts. But we also see that human patterns of social networking in intellectual discourse are not different in the digital world. Collins was writing about networks of philosophers, but his ideas provide perfect analogies for the evolution of this particular information community. It will behoove us as a discipline—information science—to engage in more and more intense monitoring of the developments of new cutting-edge domains.

References

Berman, Tamar. 2007. RE: MUSIC-IR: QBT. Post to MUSIC-IR listserv April 18, 2007.

Collins, Randall. 1998. *The sociology of philosophies: a global theory of intellectual change*. Cambridge: Belknap.

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

Fingerhut, Michael. 2008. The "S" in ISMIR. Post to MUSIC-IR listserv October 21, 2008.

Hjørland, Birger. 2002. Domain analysis in information science: eleven approaches – traditional as well as innovative. *Journal of documentation* 58: 422-62.

Hjørland, Birger . 2003. Fundamentals of knowledge organization. *Knowledge organization* 30: 87-111.

Tennis, Joseph T. 2003. Two axes of domains for domain analysis. *Knowledge organization* 30: 191-5.

Smiraglia, Richard P. 2006. Music information retrieval: An example of Bates' substrate? In *Information Science Revisited: Approaches to Innovation: Proceedings of the Canadian Association for Information Science annual conference June 1-3 2006, ed. Haidar Moukdad*. http://www.cais-acsi.ca/search.asp?year=2006.

Hilton, Robin. 2008. Top 5 iPhone Music Applications. All songs considered; NPR Music blog, 7-23-2008. http://www.npr.org/blogs/allsongs/2008/top_5_iphone_music_application_1.html accessed 17 April 2009.

White, Howard D. and Katherine W. McCain. 1997. Visualization of literatures. In *Annual review of information science and technology* 32, Medford N.J.: Information Today, pp. 99-168.