Music Information Retrieval Research

1) Introduction

There is a long history of music librarianship in the domain of printed Western classical music. Special schemes have been developed to aid in the organisation and retrieval of musical works, and existing schemes have been widely used to include these types of documents in larger physical library collections. However the advent of digital consumer technology in the form of MP3 players and mobile phones, combined with the enormous impact of the internet and the digitisation and ease of compression of audio files has brought new formats and types of user interaction to the fore. This has led to an explosion in music information retrieval research, concentrating on how to most beneficially use computers to organise, search, and retrieve music information and recordings from large digital collections.

Many of us today carry around music collections of thousands of digitised music recordings and access all manner of types of music on the web but still are unsure what to listen to next – the enormous size of these collections and the instant accessibility of 8 million Western pop, classical, jazz and folk songs can cause confusion and trepidation. Where the classical music researcher would previously consult academic texts and visit a specialist music library, or the post-rock listener would read NME and visit the Rough Trade shop for advice on what's coming up, now we access music through hand-held devices and laptops. The issue is no longer 'I hope I can find that Velvet Underground live album somewhere this year, I wonder what it sounds like' but 'which Velvet Underground live track shall I read about / download / stream now?'.

Efficient and effective automated retrieval tools can help users access globally distributed and personal digital music libraries. Recent research in this area generally involves extracting musical features from audio files and using these features to make comparisons between pieces of music to determine whether or not they are similar. Enormous data resources from internet useage and unfeasibly large music collections are used for machine learning to enable autotagging. The web is trawled for socio-cultural information, and algorithms are developed and tested by a thriving scientific community. Although some of this work is hampered by the fear of breaking copyright regulations great steps are being made to allow users to engage more actively with digital music recordings in almost utopian ways.

This chapter examines how the special nature of music impacts on the retrieval of digital audio and provides a critical overview of developments in the area of music information retrieval. Important musical facets are introduced and discussed in relation to the communication of musical meaning. This discussion leads to an analysis of various metadata schemas. Prevalent retrieval tools and systems are examined and key approaches to retrieval are identified. The importance of the music user as both a key source of research data and the ultimate participant of the musical communication process is discussed, and existing approaches to the evaluation of music retrieval tools and systems are presented and considered.

2) What is music information?

The very name, 'music information retrieval' (MIR) suggests that there is information within music that can be used for retrieval purposes – a user may wish to retrieve a particular piece of music information (‘*find me all the appearances in this song of a G7 chord / sad refrain / lyric to do with excitement’)* or to retrieve music containing that information (*‘find me all the songs with a G7 etc’)*. The multi-faceted nature of music includes pitch, tempo, harmony, timbre, editorial, textual, and bibliographic facets (Downie, 2003a). Music’s lack of the word elements found in text causes problems in identifying ‘units of meaning’ (Byrd and Crawford, 2002). Looking closely at this idea should help to shed some light on the aims and approaches of MIR research, and highlight any areas where this could be improved in some way. Along with ‘information’ comes ‘communication’. Indeed, successful communication relies on information getting from one place to another without too much interference. Identifying how musical meaning is conveyed or transmitted from the composer or performer to the listener should help determine what it is about music that we need to access to enable successful retrieval techniques. A range of communications models of increasing sophistication have been devised to interpret this process (including Shannon & Weaver, 1948; Hall, 1980; Tagg, 1999; Inskip et al, 2008) . These models generally include:

> *“a sender, a channel, a receiver, a relationship between sender and receiver, an effect, a context in which communication occurs and a range of things to which messages refer” (McQuail & Windahl, 1993:5).*

Successful communication relies on these elements all working together with minimal interference.

As the saying goes, ‘music is organised sound’. However this definition could also apply to speech. We need to examine organised sound more deeply if we are to determine fruitful paths to follow that will lead to successful retrieval approaches. The ethnomusicologist Bruno Nettl (2006) writes that music is an art made by combining sounds, a set of physiological processes and a form of communication. Blacking (1973) states that listeners who share cultural experiences are likely to demonstrate similar responses to music, and that if the listener shares the experiences of the composer / performer then musical meaning is more likely to be communicated. This will be affected by socio-cultural contexts. In fact, with no listener and no cultural context notes are just a jumble of noise and music does not exist. In other words, it’s not just the notes that matter. Cultural criteria also impact on successful musical communication.

This double-headed aspect, the content and context of music, are central to its understanding. Although it may seem common sense that without the notes there is no music, it is not so obvious that without listeners, and their attendant cognitive and socio-cultural mechanisms, music does not exist. If music can only happen by being listened to, then it necessarily follows that humans are as important in the process of musical communication as the notes themselves.

Music has a number of facets. These facets vary within and between pieces, contributing towards their identifiability and uniqueness. A song may have been performed many times by one performer. Recordings of these performances are likely to vary. Recordings of classical works, which are generally interpreted within a relatively strict framework expressed by notation and historical tradition, may only vary minutely. Recordings of performances by Bob Dylan of his own material may be almost unrecognisable from one recording to another. Identifying a song, therefore, relies on the listener knowing something about its content. This can be broken down into key facets at varying levels of complexity. Different sounds can be identified by their pitch (how high or low they are), their intensity (their loudness), and their timbre (the variation between a sound with the same pitch and intensity from two different sources). However pitch, intensity and timbre do not make music (although perhaps some Futurists and Avant Noise aficionadi may disagree). To become music, a collection of sounds have to be organised musically: some may have a beat (tempo), tonality (an agreed relationship between the notes) is likely to be important, a time signature and a key signature may indicate some structure, and possibly lyrics and a title may be used to connect the sounds in a coherent way. Music has both horizontal nature, unfolding through time, and a vertical nature, expressed by the relationship between its notes. Notes have different onsets – a drum attacks a note while a violin may fade in. They have different lengths and different endings. This multitude of musical elements, and this list does not attempt to be comprehensive, illustrates the special nature of music – and gives some indication of the problems inherent in its organisation. Combine this multi-faceted nature with a plethora of representations – including written notation, MIDI, guitar tabs, and audio containers ranging from wax cylinders to mobile phones and the information management and retrieval challenges presented by music are enormous.

The way we engage with music also highlights some key attributes which are negotiated and agreed by society. After the important bibliographic categories artist, title, composer, work etc the over-arching genres, art (or classical), pop, and folk have enormous influence on music organisation. Art music requires training to appreciate it and has a known composer, folk music has no known composer and evolves through use in the community, and pop music is defined by the relationship between performer and listener and gauged in terms of commercial success (Brackett, 2000). These ideas are a step forward from the more traditional definitions of the three musics, which focussed on their geographic sources (Redfern, 1978) and reflect the global nature of these genres. Indeed, Brackett's definitions may be applied to music from any culture, and not exclusively the usual focus on Western music. However once these three genres are broken down into the more specific hierarchical types enjoyed by record company marketing campaigns they tend to become more fuzzy as they become more granulated. The difference between pop and rock, dance and R&B, roots and traditional are likely to vary between listeners. Genre can be used as a cultural identifier (*'I like punk and garage bands therefore I am a rebellious outsider', 'I enjoy Romanticism but Modernists get my goat'*) and is traditionally the way to explore CDs in a store. Despite genre's flexible nature causing some problems in organising materials they are the predominant category used in many universal digital and physical music collections. Musical context can also be identified by the mood the music attempts to communicate, or the mood it generates in the listener, it's novelty, availability or even its use. A tag cloud of musical moods (Figure 1) illustrates the plethora of emotional attributes that may be applied to music. Finding shared definitions of these can be extremely difficult.

Figure 1: Music moods (Inskip et al, 2009a)

It is occasionally suggested that the subject of lyrics would be a useful access point. This may arise in specialist catalogues designed to find music to accompany moving images, for example, or user generated tags on social networking sites such as last.FM may refer to '*songs about puppies / existentialism / cities etc*' (last.FM, 2010).

It is the combination of all of these attributes that makes music information retrieval a challenging and exciting area. If we are able to successfully identify important facets suitable for analysis and categorisation this will take us one step closer to retrieval systems that meet user information needs.

3) Music and its organisation

Historically, in pursuit of successfully organising large collections of music, general schemes such as Dewey, Library of Congress, Bliss, Brown and Colon all have varying abilities in accommodating the special nature of music. However these schemas were originally designed for printed texts and do not consider some issues relating particularly to modern popular music, such as multiple authors (composers, performers, producers or remixers), the likelihood of the performer also being the author, the extraordinary range of genres, and the difficulty in identifying music's subject. Eric Coates' specialist scheme, British Catalogue of Music (BCM), is based on Ranganathan's Colon classification. It was the dominant notated Western classical music classification scheme in music libraries from 1957, but its focus on printed literature and printed scores has meant it has limited application for recorded music.

As large digital collections are so widespread and tend to serve different user bases it has been suggested by the International Association of Sound and Audiovisual Archives (IASA, 2009) that it is not relevant to call for a 'discographic' metadata standard but rather, adopt metadata infrastructures that are versatile, extensible and sustainable, with modularity, granularity, liquidity, openness and transparency and are relational (2009:15). They suggest that schemas should be informed by the key categories of descriptive (content, artist, title, composer, performer), structural (CD number, track number) and administrative metadata (format, barcode, catalogue number). Applying the principles of Functional Requirements of Bibliographic Records (FRBR, 2009) (Work, expression, manifestation, item; Person, corporate body; Concept, object, event, place) to any such scheme should ensure that it would be sufficiently comprehensive to meet a wide range of user information needs and this approach is flexible enough to reflect the varying requirements of different collections.

This lack of existing standards has led to a proliferation of schemas in music organisation including ID3 tags, MPEG-7, Music XML, Music Ontology, Music Vocabulary, free DB and MusicBrainz. An analysis of the metadata fields offered by these schemas (Corthaut et al, 2008) concurred with IASA, recommending that the scheme used should be the one that meets user information needs. Some of these schemas, notably MPEG-7 (MPEG-7, 2004) and Music Ontology (Music Ontology, 2010) are extremely comprehensive, attempting to incorporate all relevant musical concepts, while others (those employed by freeDB (freeDB, 2010) and MusicBrainz (MusicBrainz, 2010), for example) are more focussed on consumer users and do not include extensive amounts of technical administrative information.

Perhaps one of the currently most widely used schemas are ID3v2 tags, used in conjunction with MP3 files. More than 70 fields can be completed in cataloguing a piece of music using these tags, which follow the 'descriptive, structural, administrative' and FRBR principles, including Album/Movie/Show title, Composer, Media type, Publisher, Artist/Performer etc. (Nilson, 2000). This approach informs the metadata employed by iTunes, the worlds premier legitimate deliverer of downloadable audio and is, perforce, adopted by consumer-users of these products. These editable fields are populated manually by rights holders and thence automatically by commercial database services such as Gracenote (Gracenote, 2010) when CDs are copied to libraries using iTunes software. Users have administrative editing access once the files have been ingested into their hard drive.

While ID3 and others lean towards identifying the bibliographic elements of music, predominantly from an administrative slant, there are also instances of schemas that approach music from another angle. The Pandora 'Music Genome Project' (Pandora, 2010) reportedly considers such musical content features as Instrumentation, Feel, Structures and Influences in their cataloguing. This content-based focus is designed to allow Pandora's internet-radio service to make song recommendations to listeners based on their listening behaviour. A seed song or artist is chosen by the user and tracks with matching facets are presented on the user interface on a player. This approach to music similarity attempts to incorporate some contextual elements as well as musical content and has proven to be very successful in the US marketplace. Similar services such as last. FM are informed by user behaviour but instead of the expert cataloguers employed by Pandora rely on playcounts and user-generated tags to make their recommendations.

In terms of comprehension, the Music Ontology, developed using RDF/XML, uses three levels of 'expressiveness': editorial information, music creation information and even decomposition (Music Ontology, 2010). The use of RDF allows the Music Ontology to bring in other ontologies that may be required. This means that a range of pre-existing vocabularies may be combined to reflect the complex nature of music, and recognises the impossibility of making an all-encompassing music ontology without access to existing ideas which are so freely available on the semantic web including FOAF, FRBR and others. A close look at the terms used in this approach reveals the highly bibliographic nature of this ontology. Combining Classes, Properties and Individuals is indeed comprehensive and, because it is based on identifiable and widely agreed and confirmable facets allows users access to a large selection of content and contextual information drawn from the semantic web.

Specialist schemes have also been employed in music industry business-to-business services to aid the exploitation of recordings and compositions. The use of music in films, for example, requires catalogues to be searchable not by artist and title but by subject, tempo and mood. One of the earlier schemes is drawn from a collection of mainly classical romantic pieces bound together in a volume and supplied to cinemas for musicians to accompany silent movies. These pieces are organised according to the mood or subject of the action on the screen, including Battle, Birds, Chase, Chatter, Fire-fighting, Grotesque, Humorous, Misterioso, National, Neutral, Orgies, Passion, Pulsating, Purity, Race, Railroad, Sea-storm, Western (Rapee, 1924). Identifying these types of highly subjective facets requires deep insights into the ways potential users think about music and these contextual ideas are likely to change over time and across cultures.

In summary, although general schemes may be sufficient to incorporate printed notation and music textbooks a more specialist approach is required to accommodate the complexities of music audio. The enormous variety of music information required according to the nature of the collection and the information needs of the users means that it may be more appropriate to use flexible cataloguing approaches rather than rely on rigid standards.

4) MIR Systems

The enormous value of digital audio is that the user can access a wide range of material quickly, rather than being reliant on physical collections which may not be as comprehensive or relevant. Music collections, either held locally on a users hard drive, or remotely accessible via the world wide web can be accessed by specialist software which interrogates metadata and presents the user with an organised scheme appropriate for their particular use. A wide range of applications are used by consumers and professionals, including: music library/encyclopedia, personal collection management, commerce and transactions, music editing/production, music playback, music recommendation, music retrieval, musical notation (Corthaut et al 2008). Many of these applications can be served simply by using text retrieval approaches to metadata. Finding a track to download on iTunes or stream using Spotify or Napster, for example, is a simple process if the listener knows the artist and the title of the track or album they are intending to listen to, while finding information about an artist is relatively intuitive using web services such as All Music Guide. What is of particular interest today, however, are issues surrounding problems such as comprehensive automatic cataloguing (*how a rights holder can tag 8 million songs accurately, quickly and cheaply?*) and unknown item search (*how can a listener find music they are going to like but haven't heard yet?*).

In the late 1990s peer-to-peer networks took immense advantage of the portability of MP3 audio and the interconnectivity of the world wide web. Napster software allowed users to 'share' music on their hard drives with other users. Owing to the inherent copyrights of recorded music this prompted a landmark legal case in 2000 which forced Napster to cease operations in that form. Major players in the record industry then gradually agreed with Apple that they should make digital recordings available for download at a price on the internet via the iTunes music store. Other services have since arisen across the internet, mainly serving the sale and distribution of commercial Western popular music and employing schemas focused on bibliographic metadata.

Although the term 'music information retrieval' had first been coined in the 1960s by Michael Kassler, who was working on a system to perform Schenkerian analysis on classical notated music (Kassler, 1966), it was not until the Napster case was at the forefront of the media in 2000 that the first International Symposium on Music Information Retrieval (ISMIR) was convened in Plymouth, Mass., USA (Downie et al, 2009). The event, which since then has taken place annually in various settings around the world, is the primary focus for academics working on music information retrieval issues and has been extremely influential in providing a serious academic forum in this area. ISMIR encourages a multi-disciplinary approach to MIR research and aims to attract academics and professionals from the areas of computer science, library and information management, musicology, psychology and sociology, as well as musicians and other types of users. The main strand of interest is in the development of computer-based tools such as algorithms to determine various aspects of music which may be used in systems which are designed to enhance user engagement with digital audio. Key areas of research include content-based querying, classification (particularly by genre, style and mood), recommendation and playlist generation, fingerprinting and digital rights management, score following and audio alignment, transcription and annotation, tempo induction and beat tracking, summarisation, streaming, text and web mining, optical music recognition, and database systems and indexing and query languages (Dixon, 2008). The tools under investigation are tested by their developers and then submitted to a formal evaluation process, known as MIREX, the results of which are presented at each annual conference. These evaluations are discussed later on in this chapter.

The work of ISMIR is so wide-ranging and fast-moving it would be a disservice to attempt to summarise it. Some important themes have arisen throughout the ten years of the conference which are particularly relevant to the field of information retrieval.

a. Musical Relevance

Information Retrieval is concerned with the searching of large (text) collections, normally evaluated by relevance, which is evaluated by Precision and Recall. These measures rely on an agreed ground truth – a document is either relevant or not relevant. Experimental relevance of text documents is generally determined by whether or not it meets the needs of the user, or solves their anomalous state of knowledge (Belkin et al, 1982). In MIR this approach presents some difficulties. Firstly the concept of musical relevance is not as clear-cut as it may be in text. Although in known-item searching it is easy to present the 'right' result from a search using artist and title, once the facets such as genre, subject, mood and similarity are driving a query then relevance becomes much more difficult to pin down. Precision and Recall are equally difficult to determine with these subjective facets. These difficulties become paramount in the annual MIREX evaluations where it is unusual for precision and recall to be used to evaluate tools under submission, 'accuracy' and a range of statistical measures are brought in.

b. Musical focus

Much of the research into MIR to date has been undertaken in Western universities, predominantly in USA and Europe although more recently Japan and China, by PhD students in their twenties and thirties. It should not come as too much of a surprise, therefore, to find that the focus in such research is on recorded Western popular music and, to a lesser degree, on notated Western classical music. Although some work has been performed on music from other cultures and Western folk music these are far outweighed by more popular consumer and educational genres. Jazz is rarely mentioned. This reflects the culture of the community and is similar to the situation in text retrieval, which has tended to concentrate on English language texts.

c. The impact of social media

Over the last ten years much work has been done on extracting musical features, such as tempo, rhythm, timbre, melody, and using these features to make decisions about music recordings in terms of their similarity with others. It is generally assumed that recordings with the same features as others will be deemed to be similar by listeners. It has been found with this content-based approach that there are glass ceilings in terms of accuracy. More recently researchers have been given access to large amounts of user-generated data, most particularly tags, and are now including these data in their approaches (Barrington et al, 2009). Combining this more contextual approach with content-based algorithms has led to a noticeable improvement in evaluation results. Partly tagged collections are being used in machine learning in order to develop 'auto-tagging' systems (Mandel et al, 2010; Coviello et al, 2010), and automatic playlist generators based on this combination are becoming more reliable (Barrington, et al 2009).

It is only fair that in the early years of the discipline the MIR community focused on the essential technical development of specific tools for their work. A lot of work needed to be done to exploit the rapid recent changes in technology and attendant computer processing power. The recent ten year anniversary of ISMIR gave the community opportunity to pause and reflect on their progress and may in the future be seen to be the pivotal moment when MIR shifted from primarily focussing on systems to incorporating users into the experience and from focussing on tools to development of human usable systems (Downie et al, 2009).

5) Without a listener…

Listeners, or music users, can be drawn from three main areas of interest: casual or recreational listeners, professional users such as musicians and other people using music such as DJs, and scholars and theorists. Although they are all engaging with music on some level these users may have different information needs, illustrated by a variety of types of query:

- *Find me a song that sounds like this*
- *Given that I like these songs, find me more songs that I may enjoy*
- *I need to organize my personal collection of digital music (stored in my hard drive, portable device, MP3 player, cell phone, . ..)*
- *Retrieve musical works that have a rhythm (melody, harmony, orchestration) similar to this one*
- *I am looking for a suitable soundtrack for . . .*

Query examples such as those above tend to relate to searches for unknown items. The user often frames their query in terms of similarity. This 'query by example' is a suitable way to narrow down large collections of broadly similar items. It removes the problem of using textual metadata, which are likely to either be missing (in terms of not knowing the artist or title) or vague and subjective. It can be a lot easier for someone to say '*I'm looking for a song like this*' than '*I like songs that have a slow build in the first thirty seconds, are in ¾ time, have a sax solo in the middle eight, lyrics about puppies and were recorded in the middle 1980s*'. The difficulty in framing music queries verbally and the proliferation of mood, genre and other textual terms combined with the variation of musical knowledge of users does create difficulty in solving unknown item search requests.

Although it is possibly an unsolvable problem, identifying emotion and mood in music is an important element of MIR. If we return to the definition of music as an art combining sounds then it is important to understand that music requires a listener, and that although specific motivations to listen to music may be many and varied, music is a social process that appears across history and cultures. As it is the ability of music to carry and enhance emotion and mood that makes it an art form rather than a craft (Bicknell, 2009) then musical mood is an important facet. Although this facet is not easily categorised and can vary amongst listeners there are some generalisations that are applicable to music, such as major key for happy, minor key for sad, heavy beats for triumphant, soft for relaxed. One useful tool widely used in MIR research to identify mood is the Valence / Arousal model (Figure 2) which attempts to model human experience of emotion in a two (or, here, three) dimensional scale (Russell, 1980).

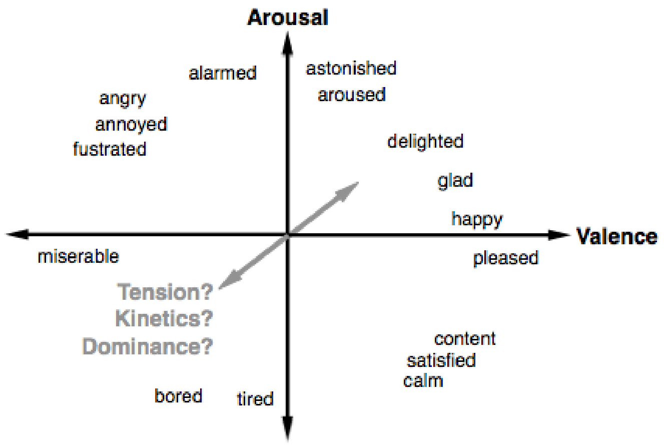

Figure 2 Valence-Arousal space (from Kim et al 2010)

Listeners under investigations using this space appear to have similar experiences when listening to music and mapping their moods as vectors allows them to be manipulated by computers as numerical vectors (Kim et al, 2010). It is vital to involve listeners in mood research – matching musical and contextual features to qualities such as delight, happiness, boredom and frustration requires substantive input from research into human cognitive behaviour. The need for a multi-disciplinary approach to MIR research is freely and regularly acknowledged by the community. The problems of collecting large amounts of reliable data and the need for a 'ground truth' for robust evaluation are being met not only by interviews and focus groups but also creatively by the community accessing large datasets from last.FM's API and the design of web-based games which collect data from participants listening to music and tagging tracks (Law et al, 2009).

The human side of music listening does not necessarily only involve extremely personal experiences. Many listeners attempt to communicate their interest in music by using the world wide web on social networking sites such as MySpace, last.FM, Youtube and Facebook. This proliferation of human-generated data has been of great value in MIR research. Users seem happy to publish their listening habits and personal music categorisations (as tags) either online through last.FM or via proprietary download services such as iTunes' Genius application. This habitual sharing of data has been widely accessed by researchers and marketers attempting to gain deeper insights into human musical behaviour and the uptake of playlists generated by these systems is starting to influence listening behaviour (Barrington et al, 2009).

Taking a holistic view (Ingwersen and Jarvelin, 2005) of information retrieval systems, they should include not only the tools and machines holding and searching the documents but also the humans using them to satisfy their information needs. Gathering data from listeners is therefore integral to MIR research. This view is supported by the recent development of, particularly, auto-tagging and automatic playlisting systems by the MIR community, which combine elements of both the content and context of music in order to satisfy human listening needs.

6) Evaluation

The MIR community's evaluation, known as Music Information Retrieval Evaluation eXchange (MIREX) of the tools in development has been informed by a close study of the text retrieval evaluation approaches of TREC. The recommendations of a detailed TREC consultation in 2003 focus on three main issues. At the time there was:

> *"1. no standard collection of music against which each team could test its techniques; 2. no standardized sets of performance tasks; and, 3. no standardized evaluation metrics." (Downie 2003b).*

The MIREX team have since been continuously developing a robust evaluation framework for MIR tools. They have collected a large amount of recordings (unique pieces of music that have been used in evaluations: 143,817, individual audio files: 228,480) of music drawn from a wide range of styles and genres. To accommodate intellectual copyright regulations, which prevent MIREX from supplying these recordings to developers as test-beds these are held in one place, accessible only to the MIREX team. Algorithms are submitted annually by researchers for competition after being tested on the researchers test collections. a number of tasks have been devised, changing every year according to interest from the community. These tasks most recently included: music similarity (audio and symbolic), audio cover song identification, auto tag classification, query by tapping, humming and singing, and audio tempo estimation. Algorithms designed to satisfy these tasks are applied to collections by the MIREX team and ranked results are published at ISMIR. As the tasks vary conceptually, so their evaluation cannot be standardised. It is not possible or indeed suitable to evaluate them all using text retrieval's Precision and Recall. The main evaluation metric in 2004 and 2005 was accuracy. Through encouragement from the organisers, precision and recall were more widely used thereafter, alongside other statistical measures such as p-score, f-measure, ANOVA (Downie, 2008). It is only through a satisfactory definition of musical relevance that precision and recall may be more widely used. This is partly resolved by Downie's suggestion that:

> *"there should be enough information contained within the query records that reasonable persons would concur as to whether or not a given returned item satisfied the intention of the query." (2003c)*

As the MIREX evaluations move forward the organisers and the participants become more experienced and continually learn from previous years of competition. Human volunteers are recruited from the community to generate ground truth data and, most recently, larger numbers of anonymous participants have been involved through web-based services such as Amazon Mechanical Turk (Amazon Web Services, 2010). Although this commitment to gathering human experience of music can be expensive and takes time and is not always accurate (Downie, 2008) it shows a strong commitment to reliable evaluation in the discipline, acknowledging the vital involvement of the listener in successful musical communication.

7) ***Conclusion***

Although music is sometimes called a language and can communicate certain ideas to the listener, it has many facets that are dissimilar to those of text. These facets are not just to be found within the music itself, in the form of melody, harmony, tempo and timbre, but are interpreted by the cognitive processes of the listener within frameworks of culturally agreed rules, such as genre, style and mood. The last ten years have seen a rapid development in the portability and availability of recorded music leading to a valuable opportunity to engage computer processing power to search through millions of songs to find the ones we wish to listen to, for whatever reason. There is acknowledgement within the MIR community that applying theories and standards developed through the lengthy history of text information retrieval and music librarianship should inform a professional and rigorously academic music information retrieval research discipline. However there are some issues, such as relevance and a lack of ground truth, that mean new and creative approaches are required in the development of suitable evaluation of tools and systems. If the next ten years of MIR is to see the widespread adoption of systems that use the tools currently being developed by the community then the continuation of a multi-disciplinary approach and a holistic view of music retrieval is key to this success.

8) References

Amazon Web Services, 2010 *Amazon Mechanical Turk (beta)* available online at <http://aws.amazon.com/mturk> last accessed [01 Nov. 10]

Belkin, N., Oddy, R. & Brooks, H. (1982). ASK for Information Retrieval. *Journal of Documentation*, 38 61-71 (part 1) & 145-164 (part 2).

Barrington, L., Oda, R., and Lanckriet, G. (2009) Smarter Than Genius? Human evaluation of music recommender systems. *Proceedings of International Conference for Music Information Retrieval, Kobe, Japan, 26-30 Oct, 2009*

Bicknell, J. (2009) *Why Music Moves Us*, Palgrave Macmillan: Basingstoke, UK.

Blacking, J. (1973). *How musical is man?* University of Washington Press, Seattle & London

Brackett, D. (2000). *Interpreting Popular Music.* University of California Press, Berkeley

Byrd, D. & Crawford, T. (2001) Problems of music information retrieval in the real world, *Information Processing & Management,* Volume 38, Issue 2, March 2002, Pages 249-272

Corthaut, N., Govaerts, S., Verbert, K. and Duval, E. (2008) ‘Connecting the Dots: Music Metadata Generation, Schemas and Applications’ *Proceedings of International Conference for Music Information Retrieval, Philadelphia, 14-18 Sep 2008*

Coviello, E., Barrington, L., Chan, A. and Lanckriet, G. (2010) Automatic Music Tagging With Time Series Models, *Proceedings of International Conference for Music Information Retrieval, Utrecht, Netherlands, 9-13 Aug 2010*

Dixon, S. (2008), Correspondance to Music-IR email list, Dec 2008

Downie, J.S., Byrd, D. and Crawford, T. (2009) Ten Years of ISMIR: Reflections On Challenges And Opportunities, *Proceedings of International Conference for Music Information Retrieval, Kobe, Japan, 26-30 Oct, 2009*

Downie, J. S. (2003a), Music information retrieval, in ed. Cronin, B. *Annual Review of Information Science and Technology 37*, Medford, NJ: Information Today, Chapter 7, pp 295-340.

Downie, J.S. (2003b). Interim Report on Establishing MIR/MDL Evaluation Frameworks: Commentary on Consensus Building, *The MIR/MDL Evaluation Project White Paper Collection Edition #3* (2003) 43-44 Available at: http://www.music-ir.org/evaluation/wp.html (accessed 23 March 2010).

Downie, J.S. (2008). The Music Information Retrieval Exchange (2005-2007): A Window Into Music Information Retrieval Research. *Acoustic Science and Technology* 29 (4) 247-255.

Downie, J.S. (2003c). Toward the Scientific Evaluation of Music Information Retrieval Systems, *Proceedings of International Symposium on Music Information Retrieval, Baltimore, USA, Oct 27-30 2003*.

FreeDB (2010) FreeDB.org, available online at < http://www.freedb.org/en/> last accessed [27 Oct. 10]

Functional Requirements for Bibliographic Records (1997, *IFLA*, available from (http://www.ifla.org/VII/s13/frbr/frbr.pdf) last accessed 20/10/10

Gracenote (2010) *Apple iTunes and Genius* available online at <http://www.gracenote.com/casestudies/itunes/> last accessed [27 Oct. 10]

Hall, S. (1980). *Encoding / Decoding* in ed Hall, S., Hobson, D., Lowe, A., Willis, P. (1980) *Culture, Media, Language*. Hutchinson, London. Chapter 10, pp 128-138

IASA (2009) *IASA TC-04 Guidelines on the Production and Preservation of Digital Audio Objects 2nd Edn*. ed Bradley, K. IASA, Auckland Park, South Africa.

Ingwersen, P. and Jarvelin, K. (2005). *The Turn*. Springer, Dordrecht.

Inskip, C., Macfarlane, A. & Rafferty, P. (2008). Meaning, communication, music: towards a revised communication model, *Journal of Documentation* 64(5) 687-706

Inskip, C., Macfarlane, A. & Rafferty, P., (2009a). Organizing Music for Movies, *Proceedings of International Society for Knowledge Organization (UK) Content Architecture conference, London, UK, 22-23 Jun 2009*

Kassler, M (1966). Towards Musical Information Retrieval. *Perspectives of New Music* 4, pp 59-67

Kim, Y, Schmidt, E., Migneco, R., Morton, B., Richardson, P., Scott, J., Speck, J. and Turnbull, D. (2010) Music Emotion Recognition: A State of the Art Review, *Proceedings of International Conference for Music Information Retrieval, Utrecht, Netherlands, 9-13 Aug 2010*

Last.FM (2010) *Tag results for 'songs about something'* available online at <http://www.last.fm/search?q=songs+about+something&type=tag> last accessed [27 Oct. 10]

Law, E., West, K., Mandel, M., Bay, M. & Downie, J.S. (2009) *Evaluation Of Algorithms Using Games: The Case Of Music Tagging Proceedings of International Conference for Music Information Retrieval, Kobe, Japan, 26-30 Oct, 2009*

Mandel, M.. Eck, D. and Bengio, Y. (2010) Learning Tags That Vary Within A Song, *Proceedings of International Conference for Music Information Retrieval, Utrecht, Netherlands, 9-13 Aug 2010*

MPEG-7 (2004) *MPEG-7 Overview* available online at =< http://mpeg.chiariglione.org/standards/mpeg-7/mpeg-7.htm> last accessed [27 Oct. 10]

MusicBrainz (2010) *Welcome to MusicBrainz*, available online at < http://musicbrainz.org/> last accessed [27 Oct. 10].

Music Ontology (2010) *Music Ontology Specification*, available online at <http://musicontology.com/> last accessed [27 Oct. 10]

Nettl, B (2006). *Music*, [internet] Grove Music Online ed. L. Macy (Accessed [27 Nov 2006]), <http://www.grovemusic.com>

Nilsson, M. (2000) *ID3 tag version 2. - Frames* from internet <http://www.id3.org/Frames> last accessed [27 Oct. 10]

Pandora (2010) *The Music Genome Project* available online at <http://www.pandora.com/mgp.shtml> last accessed [27 Oct. 10]

Rapee, E. (1924) *Motion Picture Moods for Pianists and Organists*. Arno Press, New York, 1974 reprint

Redfern, B. (1978). *Organising Music in Libraries, Vol 1: Arrangement and Classification*. Clive Bingley, London.

Russell, J. (1980) A Circumplex Model of Affect, *Journal of Personality and Social Psychology* 39 (6) pp 1161-1178

Shannon, C. and Weaver, W., (1949). *The Mathematical Theory of Communication*. University of Illinois Press

Tagg, P. (1999). *Introductory notes to the Semiotics of Music, version 3, unpublished*, [internet] available from <http://www.tagg.org/xpdfs/semiotug.pdf> [last accessed 5 July 2010]