

# Towards the Disintermediation of Creative Music Search: Analysing Queries To Determine Important Facets.

Charlie Inskip

Dept of Information Science  
City University London  
c.inskip@city.ac.uk

Andy MacFarlane

Dept of Information Science  
City University London

Pauline Rafferty

Dept of Information Studies  
University of Aberystwyth

## Abstract

Creative professionals search for music to accompany moving images in films, advertising, television. Some larger music rights holders (record companies and music publishers) organise their catalogues to allow online searching. These digital libraries are organised by various subjective musical facets as well as by artist and title metadata. A facet analysis of a number of queries is discussed in relation to the organisation of the music in these bespoke search engines. Subjective facets such as Mood and Genre are found to be highly important in query formation. These findings are discussed in relation to disintermediation of this process. It is suggested that there are a number of barriers to this, both in terms of classification approaches and also commercial / legal factors.

## 1. Introduction

Music Owners, such as record companies and music publishers, attempt to exploit the recordings and compositions they control by encouraging their use in films, TV programs, commercials, websites and corporate presentations. The process of adding music to a moving image is known as synchronisation. Many Owners' collections are digitised and act as digital music libraries. Music Users such as ad agency creatives and music supervisors search for music for synchronisation. They generally deal direct with a number of expert intermediaries employed by the Owners, interpreting the query and performing searches of their own catalogues on the Users' behalf.

A number of Owners operate online search tools which are designed to disintermediate this process. The metadata used by these bespoke Music Search Engines (MSEs) (Nanopoulos et al, 2009) was investigated and it was shown that they employ both bibliographic and descriptive terms in their music classification schemes, some of which are outside the traditional music classification paradigm, such as Subject and Mood (Inskip et al 2009).

Recent important studies in music user information need (Bainbridge et al, 2003; Cunningham et al, 2007; Kim and Belkin 2002; Lee et al., 2007) have focussed on consumers. This paper is part of ongoing research into a group of creative professionals who may have different information needs than recreational consumers. They are choosing music on behalf of others, not to listen for

pleasure. The search for music to accompany moving images is frequently an unknown item search. Users do not know specifically what they are looking for but they seem to have very clear ideas of what elements are important, such as Mood, Genre and Structure. In advertising they are also often looking for a suitable 30 second element, not the whole song.

The contribution of this paper is to analyse a selection of queries in order to inform systems development (Downie & Cunningham 2002; Lee, 2007). In the next section the Methodology is presented. Section 3 discusses Findings, focusing on descriptive and bibliographic facets and additional ways of constructing and clarifying synchronisation queries. This is followed by a discussion on factors affecting whether the process may be disintermediated. Conclusions are summarised in Section 5.

## 2. Methodology

During the course of a wider research project looking at searching for music to accompany moving images (Inskip et al, 2008a, 2008b) a number of written queries ('briefs') were collected. These briefs came from creative music searchers who are employed by advertisers or brands to find music to be used in advertising, on websites, in corporate presentations etc.. A selection of these briefs also related to TV trailer use. Briefs are often sent by email to Music Owners when a music search is taking place and they attempt to encapsulate the search criteria for an unknown piece of music which will match existing footage. They are a rich source of information regarding the semantics of music search. They are often up to one page in length and include a range of detail explaining the Users search criteria to the catalogue Owner.

The aims of this approach were:

- a) to investigate the semantics of creative music search;
- b) to relate this to knowledge organization in existing bespoke music search engines;
- c) to make observations on whether the process may be disintermediated.

The metadata used by a selection of MSEs had been examined previously (Inskip et al 2009). The facets by which these systems are organised is divided according to whether it is bibliographic (eg Artist, Writer, Title) or descriptive (eg Mood,

Subject, Genre). It was suggested that these systems are organized by taking a domain analytic approach – the Owners classifying the documents by purpose (Hjørland & Nissen Pedersen, 2005).

In pursuit of substantiation of this proposal the queries were examined for links with and differences to the organization of the MSEs. The briefs were imported into NVivo 8 software. This flexible Computer Assisted Qualitative Data Analysis (CAQDAS) package allows texts to be searched, coded, annotated and queried. Each brief was analysed word by word and phrase by phrase and coded according to facets derived from the MSE analysis (Artist, Title, Year, Genre, Subject, Mood etc). As new facets arose these were iteratively added to the coding list. When all of the 25 briefs had been coded, the facets were ranked by frequency of appearances within the set of briefs. (see Appendix).

The sections that had been coded were then examined in turn in order to consider the words that had been used within each code. For example, in the code 'Artist' there were 11 references in total. It is a simple process with NVivo to isolate these references and analyse their role within the discourse of the query. This approach enables the researcher to consider not just the words that are coded (eg 'Pink Floyd') but also the words *around* the coded words:

*"i.e. a classic piece of music like Pink Floyd - We don't need no education - a track which is about rebellion."* (Brief 001)

The value of this discourse analytic approach (Paltridge, 2006) is that it is the words on either side of the coded words that help to explain the context of the reference. In the example above, if 'Pink Floyd' were the only words considered, we would not appreciate that this is partially a similarity request rather than a known item request. The User is not solely asking for Pink Floyd's 'We don't need no education' (sic – the correct song title is 'Another Brick In The Wall Pt 2'). Pink Floyd are partly being used as a similarity metric to give context to the request, although the User is also asking whether this song (or a version of it) is available.

A word frequency count was also performed on the entire query set. Irrelevant terms such as 'should', 'can', 'his', 'you' etc were discarded from the list. Facet terms and musical descriptors were concentrated on. Analysing this list gave deeper insight into the regularity of appearance of some key concepts. Again, the terms could be drilled down on using the software to reveal their context so they were not considered in isolation from the rest of the surrounding text.

## 3. Findings

### 3.1. Descriptive Facets

By far the main facet used to describe the music being sought is that of Mood, which featured in 80% of the briefs. Positive descriptors such as 'charming', 'beautiful', 'fresh', 'playful', 'quirky', 'exciting' far outweighed negative terms such as 'dark', 'uncertain', 'anxious', 'sinister'. Notably these negative terms are mainly used as 'Exclude' terms in order to instruct the intermediary that music matching these would not be relevant for the search in question ("*Please do not pitch music with an overtly sinister, dark, or serious feel*"). Although a larger sample of queries could generate more 'negative' mood criteria it seems likely that as these queries are focussed on finding music for advertising, the users are looking for positive music moods to associate with and enhance consumers' opinions of their products. Mood has been used to match music to images since the time of silent movies (Rapee, 1924) and advertising theorists are well aware of its value in selling products (Alpert, 1990). As a subjective facet it is not an ideal way to describe the elements of music that the user is looking for, as it is difficult to match the users meaning with that of the intermediary or, indeed, the system. However the use of Mood (this is specified either for the sought music or to describe the desired 'feel' of the finished advert) far outweighs any other term employed in this set of queries.

Unsurprisingly Genre ("*heavy metal*", "*rock*", "*pop*", "*lofi*", "*folky*", "*classical*", "*jazz*") is mentioned in many queries as a guide to the music being sought. Although Genre definitions are fuzzy, generally they are agreed by a community and can be applied more successfully than more subjective terms such as Mood. Genre is a useful way for the search to be narrowed down, at least, and its widespread long term use in describing especially different forms of popular music as well as the three main genres (Art (Classical), Popular (Pop) and Folk) indicates it is an extremely valuable approach to music classification. The use of Genre terms can help as codes in describing music for a particular audience (products aimed at the youth market are often associated with contemporary pop), or which instruments would be appropriate (electric guitars do not often figure in classical).

Given the short length of the TV commercial, it is rare that a whole piece of music is used to accompany the footage. The users are looking for a specific element of a piece that they can use to convey their message to the viewer. They may discuss (in Music Structure) what musical elements they are looking for that may appear within a song: "*should have some quieter moments*", "*music evaporates into nothing*", "*build to a swelling, string-soaked chorus*", "*...with a crescendo in it*". The word, *build*, in particular appears regularly in these queries and in other discourses surrounding these

practices. These criteria are important to the users on two levels. They help to convey the mood and the message to the viewer, and they also allow important Extra-Musical Factors (such as sound effects or voice overs) to be used successfully. The importance of Extra-Musical Factors in the search is reinforced by these regularly appearing in briefs.

### 3.2. Bibliographic Facets

While the use of subjective facets seemed to be key in communicating Users' music needs to Owners, a equal number of bibliographic facets are also employed. The benefit of factors such as Date/Period (of recording being sought), key words required in Lyrics, Tempo, Instruments featured in the track and Chart Position is that they are easily attached to music documents as metadata and can be more reliable search parameters.

The value of Date/Period is that it can be matched to target Audience demographics, as well as being used to refine a general description of a style of music. There are relatively frequent references to finding music that is "contemporary", while other briefs refer to decades rather than particular years:

*"Please avoid 80s electronica, retro tracks, or anything that could be considered 'old skool'."* (Brief 011)

*"Instinctively we think that the track we need is probably from the 50's or 60's, maybe the 70's."* (Brief 012)

Songs that include particular lyrics are discussed. Examples of these include:

*"We are looking for lyrics which need to be relevant and carry the ad. Think along the lines of ideas / imagination / optimism / growth / design / drive / movement etc etc..."* (Brief 007)

*"Lyrics and choruses involving sleep, eyes, waking, dreaming, touch, or some other direct link to the narrative, would be great."* (Brief 012)

However lyrics are not always important and often Instrumentals (no vocals) are requested. This use of instrumentals not only gives space to voice overs (VO) and sound effects (SFX) but recognises the creative nature of advertising and sophistication of the viewers who are well-versed in interpreting these short messages without relying on lyrical reinforcement.

Musical facets such as tempo and instruments are occasionally mentioned in this sample of briefs. It is interesting to note that by far the most frequent tempo descriptor is 'upbeat', a term indicating a positive mood as well as a faster than average tempo. This particular combination here of affective

and structural facets into one descriptor is very effective shorthand which appears so frequently in interviews on the subject as to become a cliché (Inskip et al 2009). Users also mention key instruments (piano, guitar, strings, percussion) they wish to appear in the selected music. Again, this information can be added to metadata for ease of searching.

Artist name is occasionally used, mainly as a similarity guide rather than to describe a known item search:

*"We are looking for a recognisable song that reflects a 'Happy Goodbye'. Think 'My Way' as performed by Frank Sinatra."* (Brief 023)

In fact it would not be easy for these MSEs to match items by similarity. They can only search catalogue they control and the example may not be within that control. Intellectual Property (IP) legislation can prohibit them from including material not under their ownership, restricting their ability to develop this type of search functionality.

Chart position, on the other hand, is easily available to Owners and is a simple way to measure 'familiarity'. If a User requests a familiar tune this means it is likely to have appeared in the sales charts so searching by chart position can be used in a 'familiarity' search.

Although they are often the most important factor in deciding whether a piece of music is used (or not), budgets are rarely revealed in queries:

*"..budget can be stretched."* (Brief 001)  
*"..without being very costly!"* (Brief 017)  
*"Don't worry about budget please."* (Brief 024)

The expert interpretation of these Budget facets along with a knowledge of the brand's budgeting history and an understanding of which elements of the catalogue command different rates can lead to certain types of music being offered.

### 3.3. Visuals Facets

Although most queries in this sample focussed on advertising, a small number were concerned with looking for music for websites or TV trailers. Mentioning the Format (ad, tv, website) in the query gives richer detail to the intermediary about the eventual use of the music being sought and is an additional clue to specific facets which may be of interest. These would include length (TV ads are normally 30 seconds long, while website uses may require the whole song, or a loopable section) and raise the issues of licensing – using a piece for a TV ad would require different permissions than web or TV trailer use. These may help the intermediary in narrowing down the search results to manageable

levels. Other Visuals Facets, such as Project Title, Visuals Subject, Brand, Visuals Function and Visuals Available are also incorporated into the queries. These provide detailed contextual information for the intermediary and help to clarify the query further.

### 3.4. Query clarification

There are a number of phrases within the queries where the Users attempt to clarify their query by discussing the role of the music within the finished advert. Music Function appears frequently. The Users describe how they wish the music to interact with the visuals, or how they want the music to enhance their message:

*"...juxtapose against this theme..."* (Brief 001);  
*"The music needs to complement the visuals without being too cold"* (Brief 003);  
*"...reflect the charm and playful nature of the spot"* (Brief 004);  
*"tidily juxtapose with the childlike imagery of the animatic"* (Brief 007)  
*"reflect the gliding motion of the journey"* (Brief 009).

The value in matching music to moving images is that one can enhance the other (in the case of advertising the music is designed to enhance the visuals, while with music videos it is the visuals that are designed to enhance the music). It is not clear from the queries how this is evaluated, and while other research indicates this is often a 'gut feeling' decision based on experience and creativity, there is also a wealth of literature which discusses music and its use with film since the end of the 19<sup>th</sup> century from which users may draw. Clearly this type of criterion can only be evaluated once the music is played simultaneously with the image. 'Demo' versions of the final advert are frequently supplied along with the query in order to help the intermediaries in their search.

While the bulk of the text of the queries describes what the users are looking for, they often also clarify what would not suit. These Exclude elements again are designed to guide the search by narrowing down the results set:

*"we want to avoid anything that makes it feel like a middle class day out at the shops"* (Brief 019);  
*"avoid anything too folky or dreamy", "something known will be tossed"* (Brief 025),  
*"it is important not to emotionalise things too much"* (Brief 026)

although careful interpretation by the intermediary is again required.

For the purposes of query clarification other intertextual references may be used, such as Films that use a particular type of music, Similarity to other suitable pieces or artists, and the option to the intermediary to offer a piece that does not appear to match the query but may be appropriate: this Left Field element reflects the subjective nature of this type of searching, and allows the expert intermediary to make a contribution to the search using their own experience and interpretation of the visuals:

*"Please also feel free to suggest anything off brief that you think works well to picture, we are open to suggestion no matter how left field."* (Brief 007)  
*"But feel free to send other suggestions if you think they work."* (Brief 012)

There are many anecdotal examples of music being used in commercials that did not meet the original brief and were supplied by the intermediary as a 'Left Field' suggestion:

*"She just threw that in as a kind of a random idea, and they went for it."* (Music Supervisor)  
*"Sometimes you have to come up with something that's completely off the wall that the director won't have thought about"* (Music Publisher)

### 4. Whither disintermediation?

We have seen that music briefs describing music which will accompany moving images incorporate a range of music facets and additional contextual detail. Some of these can be matched with bibliographic metadata. However a large number of subjective facets are used, relying on a shared understanding of the meaning between the User and Owner. The expert intermediaries employed by the Owners are well-versed in interpreting these queries, may have the opportunity to discuss them in detail with the Users and sometimes view a copy of the piece of film in question to help in their search.

Building a searchable digital library that suits this verbose and subjective type of request is not an easy task. Granted, some of the bibliographic facets can be dealt with by applying suitable metadata fields, but these have to be accurate if they are to be of any value. Songs are often classified by inexperienced humans leading to inconsistent and unreliable metadata (Inskip et al 2009). With regard to the bibliographic facets this is a solvable problem. However it appears from our analysis that these searches rely more on descriptive metadata than factual bibliographic detail. This means that if the process is to be satisfactorily disintermediated the focus needs to be on successfully matching query terms with relevant and suitable metadata.

## 4.1. Mood

Let us consider Mood. Users employ a wide range of words:

*'charming', 'beautiful', 'fresh', 'playful', 'quirky', 'exciting', 'dark', 'uncertain', 'anxious', 'sinister'.*

The MSEs also use a wide range of Mood descriptors from controlled vocabularies, which are presented to the User for them to select the most appropriate to their search:

aggressive ambient angry angst anthemic  
atmospheric bittersweet brooding calm carefree celebratory confident dark  
depressed desire **dramatic dreamy** driven  
driving **dynamic** energetic ethereal  
euphoric exuberant fiery funky funny graceful  
happy **high** intense joyous jubilant light longing  
melancholy mellow **passionate**  
reflective **rousing sad**  
**sentimental spirited**  
time upbeat

Figure 1 MSE moods (Inskip et al 2009)

Encouraging a User to select a Mood from a controlled vocabulary, rather than asking them to input it as a keyword means it is the Owner who is also involved in encoding the query to match the system, rather than the User encoding the query to match the music being sought. This can remove the creative element of the search from the User and may dissuade them from attempting to perform their search online. Clearly, if the Mood choices are to be presented to Users then it is important to investigate how they determine whether a piece is *'charming'*, *'beautiful'* or *'fresh'* and developing the controlled vocabulary accordingly. This applies equally to Genre, although as previously stated, Genre is less subjective than Mood. The variation in interpretations of the cultural meanings of music reinforces the value of taking a domain analytic approach when developing music search tools (Abrahamson, 2003).

## 4.2. Music Structure

It is notable that these searches are focusing on a small element of lengthier works. The viewer only hears up to 30 seconds of music in a TV commercial, probably less. Most popular songs last around 3 minutes. It is important in the search process that the part of the piece that matches the search criteria is found. Songs may vary in mood, tempo, may have crescendos at the middle or the end. The whole song has to be listened to in order

to find out whether it includes a relevant section. Ad creatives have little time and would benefit from being presented with the 30 second element of a song that matches the query rather than be forced to listen to an entire song. When the human intermediary does the search s/he may know where in a piece the *'build'* takes place and possibly will direct the User to this section of the music. Disintermediation should perform a similar function. Tools that search music signals for Music Structural facets such as crescendos, solos, and specific instrument onsets may be of particular value in this area.

## 4.3. Copyright and competition

It may be that a 'global' search system may help the searchers who are time poor by saving them doing the same search on multiple services. However there are legal and business hurdles to this solution. Record companies may not use lyrics without permission – they are controlled by music publishers. Conversely, music publishers may not own the recordings of the compositions they control and would need permission to include them in their online catalogues. If we combine this problem with the fact that these services are designed for the exploitation of catalogue in a highly competitive industry then collaboration between Owners is difficult. However it is not unsurpassable. There is currently one service ([www.ricall.com](http://www.ricall.com)) which unifies a selection of client catalogues in an attempt to simplify the problematic element of the search process relating to Users having to use numerous interfaces for one search.

## 5. Conclusion

Anecdotal evidence suggests that, historically, although there is the will to disintermediate the process, there have not always been the resources or the technology. A number of MSEs were developed up to five years ago in a rush to compete for business. They have not all been updated to keep up with web and search engine technology, although there are exceptions with some services currently in re-development and not available for analysis.

It appears that although the Music Owners are designing search tools for Users who wish to search online (Inskip et al 2009) the possible mismatch between the Users' approach and that of the Owners must be considered. If successful disintermediation of these services is to take place then the Users and their contexts have to be considered in detail.

## 6. Acknowledgements

We would like to thank all the participants in our research for providing us with briefs and being so free with their valuable time. Charlie Inskip gratefully acknowledges financial support from AHRC for this PhD research.

## 7. References

- Abrahamson, K.T. (2003) Indexing of Musical Genres: An Epistemological Perspective Knowledge Organization 3 / 4 pp 144-169
- Alpert, J. & Alpert, M. (1990) Music Influences on Mood and Purchase Intentions, Psychology and Marketing 7(2) pp 109-133
- Brief 001 – 025 (2009) Private documents supplied by anonymous creative music searchers
- Cunningham, S., Bainbridge, D. and McKay, D. (2007). Finding new music: a diary study of everyday encounters with novel songs. Proceedings of the 8th International Conference on Music Information Retrieval (ISMIR), Sep 23 – 27, Vienna, Austria
- Downie, S. and Cunningham, S. (2002) Toward a theory of music information retrieval queries: system design implications. Proceedings of the 3rd International Conference on Music Information Retrieval (ISMIR), Oct 13 – 17, Paris, France.
- Hjørland, B. and Nissen Pedersen, K. (2005) A substantive theory of classification for information retrieval. Journal of Documentation 61(5) pp582-597.
- Inskip, C., Macfarlane, A. & Rafferty, P. (2008a) Meaning, communication, music: towards a revised communication model. Journal of Documentation, 64(5) pp 687-706
- Inskip, C., Macfarlane, A. & Rafferty, P. (2008b) Music, Movies and Meaning: Communication in Film-Makers Search for Pre-Existing Music, and the Implications for Music Information Retrieval. Proceedings of the Ninth International Conference on Music Information Retrieval, Philadelphia, PA, 14-18 Sep 2008
- Inskip, C., Macfarlane, A. & Rafferty, P. (2009) Organizing Music for Movies. Proceedings of International Society for Knowledge Organization (UK) Content Architecture conference, London, UK, 22-23 Jun 2009
- Kim, J. & Belkin, N. (2002), Categories of Music Description and Search Terms and Phrases Used by Non-Music Experts, Proceedings of the Third International Conference on Music Information Retrieval, Paris, France, Oct 13-17 2002.
- Lee, J., Downie, J.S.; Jones, M.C. (2007) Preliminary Analyses of Information Features Provided by Users for Identifying Music, Proceedings of the Eighth International Conference on Music Information Retrieval, Vienna, Austria, Sep 23-27 2007.
- Nanopoulos, A., Rafailidis, D., Ruxanda, M, & Manolopoulos, Y. (2009) Music search engines: Specifications and challenges, Information Processing and Management 45(3) pp 392-396
- Paltridge, B. (2006) Discourse Analysis. Continuum, London and New York.

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

## 8. 8. Appendix

Music Facets	References	Type
Mood	130	Subjective
Genre	39	Subjective
Music Structure	21	Subjective
Date / Period	20	Objective
Audience	16	Objective
Lyrics	14	Objective
Artist	11	Objective
Tempo	10	Objective
Instrument	9	Objective
Extra-musical	9	Objective
Song Title	7	Objective
Chart position	5	Objective
Budget	4	Objective
Version	3	Objective
Music Style	3	Subjective
Length	3	Objective
Instrumental	3	Objective
Clearable	3	Objective
Vocal	1	Objective
Territory	1	Objective
Song Subject	1	Subjective
Other intertextual refs	1	Subjective
Exploitation	1	Subjective

Visuals Facets	References	Type
Format (ad, film, tv)	25	Objective
Project Title	18	Objective
Visuals Subject	16	Objective
Brand	13	Objective
Visuals Function	6	Objective
Visuals Available	6	Objective

Query clarification	References	Type
Music Function	47	Subjective
Exclude	22	Subjective
Film title	9	Objective
Similarity	6	Subjective
Left Field	4	Subjective