Usability of a Music Digital Library: an OSM Case Study

Iain W. Connell
UCL Interaction Centre (UCLIC)
University College London
26 Bedford Way
London WC1H OAP, U.K.
+44 020 7679 5319
i.connell@cs.ucl.ac.uk

Ann E. Blandford
UCL Interaction Centre (UCLIC)
University College London
26 Bedford Way
London WC1H OAP, U.K.
+44 020 7679 7557
a.blandford@ cs.ucl.ac.uk

Thomas R.G. Green
Computer-Based Learning Unit
University of Leeds
Leeds LS2 9JT
U.K.
thomas.green@ndirect.co.uk

ABSTRACT

This paper indicates how one novel form of usability inspection method, namely Ontological Sketch Modelling or OSM, might inform the sorts of design decisions involved in creating a music digital library. OSM aims to identify any potential misfits between the designers' views of a product or system, embodied in the device itself, and those of its users. This is performed by describing how OSM's core concepts, namely entities, attributes, actions and relationships, are manifested in the both the system (device) and user models. The paper is a case study of the application of OSM in a real-world context. The object of the study was the New Zealand Digital Library (NZDL) Music Library of the University of Waikato Department of Computer Science. In support of the analysis, structured interviews were carried out with a number of musicians of varying computing and Music Information Retrieval (MIR) experience. The results show that OSM is a potentially valuable method for identifying the likely misfits between device and user models of a system such as a music library. The ways in which these misfits arose from the analysis are described and discussed, along with the implications for future music library design.

1. INTRODUCTION

Music digital libraries represent an interesting challenge to design because their potential users may need a higher level of both knowledge (musical, musicological) and experience (of computer hardware and software specialised for music playback) than do other digital libraries (DLs). Yet music DLs have been envisaged as a route towards a "truly popular" digital library, with appeal to both casual and professional users (Bainbridge et al 1999, McNab et al 1996). Evaluating the usability of music DLs would therefore require a more than usually accurate assessment of the match between designers' and users' expectations.

This paper indicates how one novel form of usability inspection method, namely Ontological Sketch Modelling or OSM, might inform the sorts of design decisions involved in creating a music DL. In an inspection method an experienced analyst systematically inspects a design or prototype for potential flaws, prior to testing with users. An advantage of inspection over testing is that the analyst can make an assessment of fundamental design issues at an earlier stage than may be achieved by testing alone.

OSM (Blandford & Green 1997, 1998, 2001) aims to identify any potential misfits between the designers' views of a product or system, embodied in the device itself, and those of its users. In the case of music digital libraries, we characterise the 'device' as both the digital library itself and the interface between library content and user which is represented on a host computer. We thus extend the scope of inspection to include more than just the characteristics of the user interface alone. In this case, we encompass wider issues such as the origins and organisation of a music DL, and the ways in which that DL may be queried on a host computer.

This paper is a case study of the application of OSM in a real-world context. The object of the study was the New Zealand Digital Library (NZDL) Music Library, in the form in which it existed between October 2001 and January 2002. The NZDL (see References for URL) allows access to a corpus of 14 'humanitarian and UN collections' and 17 'demonstration collections' including the Music Library (hereafter referred to as the NZDL ML). The NZDL is developed and maintained by the Department of Computer Science at the University of Waikato, Hamilton, New Zealand. The NZDL ML (McNab et al 1997, Bainbridge et al 1999, Bainbridge 2000) is developed at the same Department.

In order to inform the analyst's view of the knowledge and expectations of the users of a product or system, a user elicitation process may be undertaken with a sample of potential users drawn from an available population. In this case, interviews were carried out with a number of musicians of varying computing and Music Information Retrieval (MIR) experience. Excerpts from the transcripts of these interviews will be included. The interviewing approach used is described in Section 2.2.

OSM is still evolving, and its precise formulations are themselves under development. However, its core concepts, namely the focus on entities, attributes, actions and relationships, are established. These and the OSM approach itself are further described in the next Section.

1.1 Outline of OSM

Ontological Sketch Modelling (OSM) is an approach to usability evaluation which focuses on structures rather than tasks. It assesses the goodness of fit (and any resulting 'misfits') between the user's conceptualisation of a system and that which is implemented by the designer. OSM provides a common representation that supports reasoning about users, domains and devices (COSMEA publications: see References for URL). OSM is "a structured but informal representation of the ontology - the essential underlying structure - of a system, forming a basis for usability assessment" (Blandford & Green 1997).

The focus of OSM is thus on the concepts which underlie a system or product (in this case inherent in music DLs) rather than the tasks or processes involved in interacting with that system (here the steps involved in querying and playing back the contents of a music DL). OSM inverts the more usual concern with actions upon objects (typically those presented to the user via a graphical interface), instead focusing on the conceptual entities which a system embodies. It is this which makes OSM particularly suited for use as an inspection method.

OSM describes the entities, attributes and actions, plus any inter-relationships between these, that a user needs to work with when using a system. To first illustrate using the desktop application Word: an **entity** (e.g. a paragraph in Word) may be relevant to the domain (word processing) or the device (the application as presented to the user on a PC). The user may create or delete an entity (start a new paragraph by pressing <return>, or perform a cut on a paragraph). The user may also change the **attributes**, or properties, of an entity (e.g. by altering a paragraph's margin indents), but not create or delete attributes. Thus described are the **actions** involved in creating or deleting entities, or changing attributes. OSM also identifies any **relationships** or constraints which hold between entities, attributes and actions. For example, in Word a paragraph consists-of words, and cannot be indented below a certain width; reducing a paragraph's indent may alter the 'page-number' attribute of the entity 'cursor-position'.

As described above, OSM encourages a focus on issues outside the specifics of a particular interface. For example:

- In what ways will the potential users of the device find it matches (or fails to match) with their expectations and assumptions ?
- What aspects of the device-user interaction will cause difficulties because they are hidden, disguised, or transient ?
- What aspects of the device model are private (to the device) rather than shared between device and user ?

Any resulting **misfits** are identified via the following entity-attribute typology:

User-private: part of the user's domain knowledge, but not directly represented in the device. User-private entities are likely to lead to misfits because they either cannot be expressed via the device or have to be re-conceptualised by the user. For example, using Word to create overhead presentations forces a user view of pagination (freely re-ordered pages) onto a different device concept (Word's fixed page ordering).

Device-private: a concept from the device world that the user has to know about, but cannot change easily and may not be able to see. Device-private entities are hard for the user to learn about and may never be discovered independently. For example, the manipulation of Word style sheets remains undiscovered by many users, even though styles are visible on a standard toolbar.

Shared/device: explicitly represented in the device, but not part of the user's domain model (yet has to be manipulated by the user). Misfits are less likely with shared/device entities, though their use may need to be made explicit to novices. For example, Word file types are visible when first saving a document (and on subsequent use), but file manipulation is not a direct part of document creation.

Shared/domain: explicitly represented in the device, domain-relevant, and known to the user. Misfits are unlikely with shared/domain entities, since they are germane to both device and user. For example, a 'word' is the single most accessible entity in Word (even though comprised of the entity 'character'), and can be manipulated directly (by double-click and drag).

The intermediate representation of the output of an OSM is a set of tables of entities and their attributes, actions and relationships. Section 3.2 illustrates part of the results from the analysis of the NZDL music library. The future format of these tables will evolve along with the method itself.

1.2 Elicitation Approach

OSM requires that the analyst has insights into the knowledge, attitudes and assumptions which users bring to a product or system. The approach to acquiring such insights used in this study was to conduct structured interviews with a sample of potential users. In the context of OSM, the aim of interviewing is to illuminate the ways in which the entities and attributes which users bring to a system may be different from those embodied in the system itself. It is also possible to focus on particular aspects of the device which might be expected to cause users difficulties of interpretation, in order to identify any device-user misfits. Insights into the device model itself can be gathered in a number of ways including discussions with designers, device documentation, and inspection. Section 3.1 includes illustrative excerpts from the user interview transcripts.

1.3 OSM and MIR

This paper is an illustration of the way in which a novel evaluation approach can illuminate the design process, using MIR as an example. It is important to stress that OSM is not specifically intended as a design approach for DLs in general or MIR in particular. Indeed, the ways in which the insights which OSM brings to DLs are different to those for a wide range of other applications (including ticket vending machines, desktop applications, knowledge-based systems) are under investigation. However, a challenge for MIR is that it may rely on a more than usually specific knowledge and insights from its potential users. In this sense, MIR is a good subject for an OSM, not least because both it and OSM itself are at an early stage of development.

It is also important to comment on OSM's current status as an inspection method. While the contribution of OSM to the later stages of design and development are yet to be established, it is hoped that its potential as a predictive model for MIR usability can be investigated by a series of user tests on the NZDL ML, in both its present (late 2001 to early 2002) and later (2002) stages of development. It is expected that the insights gained from such comparisons between predicted and observed user-system interactions will be useful to both OSM and MIR.

2. METHOD

2.1 OSM: Modelling device-user entities

An in-depth inspection of the NZDL ML, in its then (October 2001 to January 2002) state of development, was carried out using the version linked to the NZDL home page (see References for URL). The aim was to identify the entities, attributes, actions and relationships embodied in the NZDL ML, and hence to detect any potential misfits between device and users. The output from this analysis was a set of tables, part of which are illustrated in Section 3.2. Inspection followed the following stages:

First the main device entities and attributes were identified. This was done by inspection, supported by additional information concerning the origins and development of the NZDL ML (kindly supplied by David Bainbridge of Waikato University). The main device entities consist of an extensive collection of melody files, organised into five sub-collections. The contents of each sub-collection can be queried via the internet, the results of a successful query being displayed on the host (user's) computer. Particular melody files can then be selected for playback. In the NZDL the constituents of melody files vary according to the sub-collection under view, one constraint being that only those files in the current sub-collection are available for playback. In common with most internet (world wide web) sites, both internal and external links (to other internet sites) are included. One feature of the NZDL ML is that a fifth sub-collection (not analysed in this study) is held at an external site (Rutgers University).

Next the main user entities were described. Initial analysis identified a potential set consisting of a tune (or melody) which was to be matched against a sub-collection contents, plus two text query modes (lyrics and titles). Possible descriptors of a 'tune' were later amplified by user interview (along with varying views of the hardware and software requirements for melody file playback). An unusual feature of the NZDL ML (unlike, for example, the VARIATIONS system described in Fuhrman et al 2001) is that tune matching can be performed on a user entry which is sung or played directly into the host computer. An additional requirement on the user is that she or he be aware of the implications for file storage and format for sung entries.

In the case of the NZDL ML, it proved necessary to further describe the four sub-collections analysed (as attributes of that entity). Having done so, the relationships between, and properties of, the device and user entities were further identified. Attention was also given to the difficulty envisaged in predicting the effect of a particular set of preferences settings, and in identifying the retrieved components of particular melody files. (These predicted usability issues were later supported by the user interviews, as discussed below.) Finally a detailed description of the necessary interface action sequences embodied in the NZDL ML was constructed. (These are not illustrated in Section 3.2.)

2.2 User elicitation: identifying user entities

In service of the user elicitation process introduced above, interviews were carried out with five potential users of a music DL. The aim was to identify any differences between the ways in which the device (here the NZDL ML) requires entities (here sub-collections of melody files) to be manipulated, and those which the user brings to her or his use of the DL (here sung or played melodies, text-based queries). All interviewees were musicians, three being lecturers or teachers of music at Middlesex University. The fourth is a MIDI user, the fifth a jazz musician. Experience of IT and digital libraries varied from high to nil; the third interviewee has little or no experience of DL and the fifth has little computing experience.

Interviewee 1 (M1) is a technician and MIDI player.

Interviewee 2 (M2) is a principal lecturer in computer science (educational technology) and part-time musician.

Interviewee 3 (M3) is a composer and arranger. He is programme manager for BA Music and Music and Arts Management (Middlesex University).

Interviewee 4 (M4) is a music teacher (PGCE, Masters in Music Technology) with a specialisation in the use of musical technology in education.

Interviewee 5 (M5) is a saxophonist in the second year of a BA in Jazz (Middlesex University).

Interview sessions were structured around a series of questions designed to assess the extent to which interviewees' knowledge and experience matched the requirements of the domain and the device (here music libraries in general, and NZDL ML in particular). Questions ranged from the general ("What is a tune ?") to the particular (interviewees were presented with a set of screen outputs from the NZDL ML and asked if they could identify the contents). Interviewees were allowed to speak as much as they wished; novel or unexpected insights were followed up, but otherwise the interviewer (the first author) did not intervene or direct the content of the responses.

Sessions were recorded on audio tape and later transcribed. Extracts from the transcripts appear in Section 3.1.

3. RESULTS

3.1 User elicitation: insights and inputs

The major implication arising from these interviews is good news for music digital libraries, in that most of the necessary user knowledge and experience required to make use of on-line music collections such as the NZDL ML appears to be held by this small sample. Even interviewee M5, a relative DL novice by comparison with the other four, expressed a positive attitude to the internet and on-line music access in general. As far as the NZDL ML itself is concerned, the interviews allowed us to identify some positive features for future developments of the interface to this and other music libraries, which will be discussed in Section 4.

Responses to each of the seven questions are summarised below.

Question 1: What is a tune, and how do you make use of tunes and scores ?

All but one of the interviewees emphasised the role of melody in a tune, but differed on the definition itself. However, for all interviewees a score was more than either a tune or a melody, comprising the lines for several instruments (and musicians), only one or some of which might carry the melody. M3 distinguished between Western and Eastern music in these two traditions' differing emphases on melody and rhythm.

Questions 2 and 3: How might a tune be represented on a computer (in digital form) in a way which may be different from a musical score ? What kind(s) of translations have to be done between a score on paper or a voiced recording in order to make it recognisable by a digital system ?

All interviewees except M5 confirmed that the representation of a tune on computer required a translation between the paper (scored) and computer forms, but there were different levels of both emphasis and knowledge. There was a difference of view between M1 and M4 regarding the directness with which sung-to-MIDI translations could take place. However, both M1 and M4 emphasised that the main conversion problem concerned pitch:

M1: "Audio is a completely different ball game altogether, you'd need [a MIDI file ?] to do it, if I could do it you'd have to devise some kind of pitch to MIDI converter, you'd have to bring it back to MIDI I think to make it make sense, because if you start [...] trying to work out recognition software, it's going to take you forever to try and do it that way. So really it's going to have to be converted into MIDI."

M4: "The main problem has always been pitch - MIDI wants to use twelve equal semitones, but when we sing we use lot of inflections - that's relative pitch. Actual pitch - in terms of trying to sing in one key and have it interpret it into another key; with complex scores, what bit of the score do you choose? Also knowing how the tune goes - dependent on cultural factors, memory. Also simplifying a recorded representation into a single tune - when you do this, it often doesn't sound like a real tune."

M3, while admitting to no knowledge of how such translations work, also emphasised cultural factors in tune identification:

M3: "[...] if it's the case of technology helping to identify a particular piece of music through the use of melody or a melodic fragment, that's fine, but for some pieces it might be quite difficult because not all of the music is the melody - this is a popular misconception about music."

Question 4: If you were using a digital library yourself, how would you expect to be able to access the tunes ? What kind of searching would you expect to be able to do ?

Various suggestions were made regarding search terms, as follows:

Artist (4 out of 5)
Composer (3 of 5)
Title (2)
Musical period (2)
Year/date of composition (2)
Orchestra (1)
Record label (1)
Featured artists (1)
Country of origin of tune (1)
Style (1)
Album name (1)

However, with such a small sample as we should be careful not to extrapolate to other user populations: for example, folk musicians might be more likely to choose title searching. This list only suggests some of the ways in which searching might be supported in a music DL.

Question 5: If you were using a digital library to match a tune in your head against those stored in the library, via a desktop computer, what steps would you expect to have to take to be able to input the tune in a way that could be recognised by the library ?

Four out the five interviewees understood that a sung tune might be translated into digital form (which could be used to match against a tune held in a library), but all also anticipated the ensuing problem of tune recognition. Again, M3 was most specific on the question of articulation:

M1: "... if you could actually play a bit of a tune in, that would be very good. You know, if you couldn't remember the song [or the] title all the way, that does happen, you hear a song and say 'how does that go, I know that', and you hum it, well if you could hum it into a microphone [and] it could find it for you, that would be nice. [...] but then if people can't sing in tune then you're going to have a few problems, so you'd have to have some kind of correction in there [...].
[...]
Researcher: "If you could whistle a tune, would that be easier to translate than trying to sing ?"
M1: "Yeah, it would be, na na na na na, you know, you would need some kind of filtering of some sort, so it could actually understand what you're doing, because if you miss a couple of notes I think you're going to be in quite a bit of trouble ..."

M2: "[...] if someone's got decent pitch they could whistle it or sing it, and that should be easy enough [...] that is quite straightforward. You could do it by pattern recognition of main motifs, [...] that's the best way to do it."

M3: "[...] I don't really know how good computers are at picking up voice - I would imagine that a hard consonant [...] would be more useful [...] if you go, em, we do a Christmas Carol [...] [hums 'Silent Night'], there's no articulation there, [but] if, as you say [la la lahs the melody] you can hear the articulative nature of the music. Or even a 't' [ta ta tahs it] the computer may be better at picking up the English notes. If you wanted me to create the space in between the notes, then that's clearly going to prevent a certain fluency [la la lahs the melody with spaces], you might lose the plot, mightn't you ?"

M3 again focused on the importance of rhythm, stressing the differences between Western and non-European music:

M3: "[...] Western music has relied so much on melody and pitch and so little on rhythm. In a sense, this is why you're here, asking me about melody in a databank, because you have come with that agenda [...] this is by definition a Euro-centric databank you're asking me about."

Question 6: What pre-conditions (if any) would have to be satisfied before you could successfully play back on a computer a tune which is stored in a digital library ?

Other than M5, the interviewees understood that this requires a combination of hardware and software (plus loudspeakers and online DL access). Both M1 and M2 distinguished between MIDI data and uncompressed formats such as WAV and RealTime. M2 correctly stated that playing back formats such as QuickTime and Windows Media Player requires an appropriate internet browser plug-in (when initiating playback from within a browser).

Question 7: Can you look at these [three separate] screen shots [on a single A4 page] from the [NZDL ML] system, and tell me if you recognise or understand the terminology ?

Figure 1 shows the combined extracts (on one page) as presented to interviewees. The first section ('DIGITAL MUSIC LIBRARY SEARCH PARAMETERS') shows the parameter setting options for search queries. The middle section ('DIGITAL MUSIC LIBRARY SOUND FORMATS FOR PLAYING BACK TUNES') is the list of sound format settings (made available at tune playback time if not already selected). The last section ('DIGITAL MUSIC LIBRARY PLAYING BACK FILES AND TUNE COMPONENTS' is an extract from a melody file, as displayed following successful retrieval from the fourth ('Midi Theme') sub-collection.

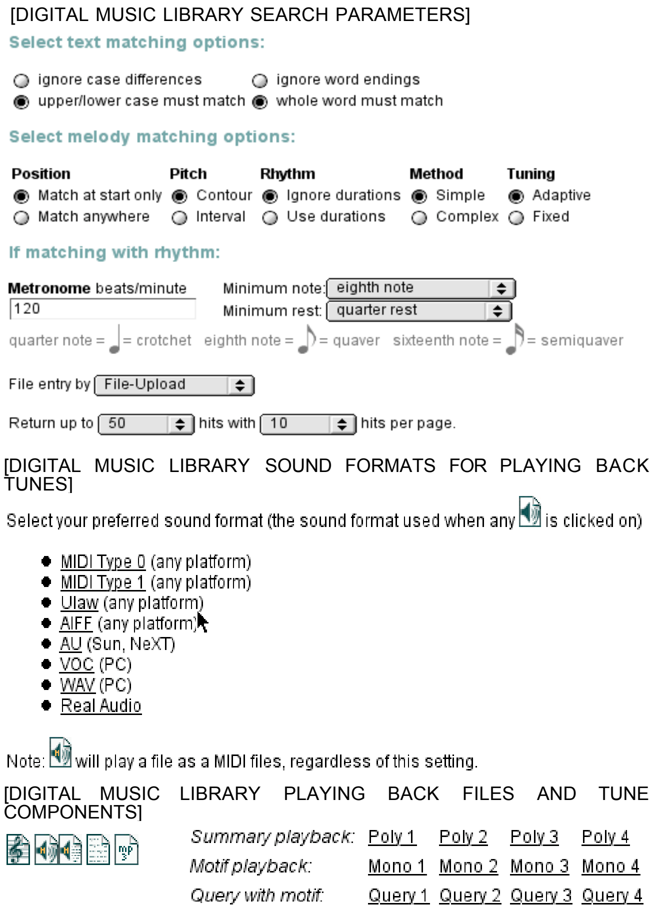

Figure 1. Combined screen extracts from the NZDL ML.

Little of the parameter settings terminology, sound formats and playback file icons from the NZDL ML (illustrated in Figure 1) was meaningful to these five interviewees. In particular, none of the interviewees understood the significance of the playback icons in the last section (other than the 'score' and 'mp3'), and none but M5 could interpret 'Poly', 'Mono' and 'Query'. However, M1 and M2 recognised most of the sound formats in the middle section, particularly MIDI, and M1 distinguished between MIDI Types 0 and 1.

M2: "Position, Match at start only, Match anywhere, I support that's somewhere within the actual note, the notation, the pitch, if there's say five notes in a melody you could start at note three, if you want [...]. Pitch Contour, I know an interval is between one note and another [...] going up, going down, I'm not sure what Pitch Contour means. [...] Method - Simple, Complex, don't know what that means [...]"

In summary, most interviewees were aware of the problems of translation involved in recognising sung input, including tune fragments, and of playing back melody files using local hardware and software. One raised the interesting issue of the Western tradition's emphasis on melody over rhythm. The main implications for the NZDL ML are the likely misfits between user and device views of search query parameters.

3.2 OSM: Findings and Misfits

The findings of the OSM analysis of the NZDL ML are represented in the form of tables. Figures 2 and 3 show the main device and user entities respectively. Note that for reasons of space, columns concerned with the creation, deletion and changing of entities and attributes have been omitted.

Entity	Type	Description	Notes
Collection	Shared/domain	(A database which) supports searches via the internet.	Can be both 'Internal' to this library site and 'external' to the client computer.

Attributes	Instances	Notes
Number of sub-collections	Currently 5 (4 'internal', 1 'external')	The number and composition of sub-collections is under NZDL control.
Type of sub-collection(s)	As described below	Ditto

Entity	Type	Description	Notes
Sub-collection	Shared/device	Consists of melody files. May contain sub-groups. Searchable (for a tune match). Browsable by title (A–Z).	Sub-collection contents and organisation are under NZDL control. It is not initially obvious that there are two query modes ('titles a-z' is in demoted to a menu option).

Attributes	Instances	Notes
Name	Folksong, Fake Book, MidiMini, MidiTheme, MidiMax	
Number of tunes	1,200 to 100,000.	The number of melody files in each sub-collection varies.
Type of tunes		Sub-collections contain different types of tunes from various sources.
Number of sub-groups (if any)	4 (Folksong) and 3 (Midi Mini)	Sub-group organisation is under NZDL control and may be difficult to predict.

Entity	Type	Description	Notes
Melody search (match) parameters	Shared/device	Describe the limits of each search. Entered (set) by the user in the sub-collection page and/or in the Preferences page.	Main parameters: set for each search. Preferences: unclear (it is very difficult to determine what the effects of each combination of parameters will be), and for how long they persist.

Attributes	Instances	Notes
See Figure 1	See figure 1	The determinants of the search parameter settings are held on a separate page to the main selections.

Entity	Type	Description	Notes
Melody file	Device-private	A data file whose contents include at least a digital representation of a melody in playback form.	Melody files can only be played back using software which is both appropriate to that format and suitable for the host computer.

Attributes	Instances	Notes
Type of components	Score. Melody. Text data: title, composer, lyricist, lyrics.	Scores are held as MIDI data, re-constructed as sheet music. Melodies are held as MIDI data. Text data are held as HTML pages extracted from the MIDI file. (Bainbridge et al 1999).
Number of components	3 to 5	The number of components in melody files varies according to the sub-collection and the original source(s).
How indexed	By title. Alpha-numerically.	Most NZDL Music Library melody files are currently indexed by title. The Midi Theme is ordered alphanumerically.
Playback (sound) format	MIDI type 0, MIDI type 1, AIFF, AU, VOC, WAV, Real Audio	Melody files can only be played back using software suitable for the host device computer. It is not clear which format is required for playback, nor for how long this setting persists.

Entity	Type	Description	Notes
Recorded tune	Shared/domain	The recording of the inputted tune (voiced or played) stored on the host computer.	The user must record the sung or played version on the host computer before it can be matched. This is clearly the same as the sung version.
Transcribed tune	Shared/device	The version of the sung or played melody as transcribed by MELDEX (the NZDL ML tune matching software).	MELDEX translates the recorded tune ('what you sang') into a version ('what I heard') which can be played back on the host computer. This is not clearly identical to the sung version.

Figure 2. NZDL ML Main Device Entities and Attributes

Entity	Type	Description	Notes
Tune	Shared/domain	A version of a melody, available in the user's working memory in a form that can be recorded (voiced or played) prior to being matched against sub-collection contents.	In order to be available for matching against a sub-collection, a learnt (user-private) tune must first be retrieved from long-term memory. A very simplified view of recalling or remembering is transfer from long-term to working memory.

Attributes	Notes
Melody	These potential descriptors for a 'tune' are derived from the series of interviews with musicians described in Section 3.1.
Pitch	
Notes	
Spaces	
Tempo	
Harmony	
Rhythm	
Tonality	

Entity	Type	Description	Notes
Lyrics	Shared/domain	A remembered part or whole of a song, held in the user's working memory (prior to being matched against a sub-collection).	A song will have a tune or melodic component (not necessarily associated with a particular lyric), while a tune may stand alone.
Title	Shared/domain	Held in working memory	NZDL ML text queries currently default to case-sensitive title inputs rather than lyrics or lyric fragments.

Attributes	Instances	Notes
Words, number of words	[As remembered]	Most interviewees did not specify either lyrics or titles as textual search terms.
Case	Upper, lower	

Figure 3. NZDL ML Main User Entities and Attributes

Comparison of Figures 2 and 3 reveals the following misfits between device and user:

1. The ways in which sub-collections are derived, organised and maintained is difficult for the user to elicit, and the results of a particular search will remain largely device-private. Further, one sub-collection is held externally to NZDL, its relationship to the others not being made obvious.

2. Though search parameters can be changed by the user, it is not clear the extent to which they persist, nor what their combined effects will be. Currently text searches default to case-sensitive title inputs; at present the effect of changing these settings is also unclear. In addition, most interviewees did not express a preference for either lyrics or titles in text searches.

3. Most of the musicians interviewed for this study gave a good account of the difficulties involved in translating a voiced recording into a form suitable for matching against a digital library. However, at present the version of the sung tune created by MELDEX (the NZDL ML tune transcription software - McNab et al 1997[1]) is not clearly identical to the recording. Further, the labels attached to the two versions - respectively 'What I heard' and 'What you sang' - are sufficiently ambiguous to engender a potential misfit between the recorded and transcribed versions.

4. Once successfully retrieved from a sub-collection, melody files can be played back on the host computer. However, this can only be done using software which is appropriate to both the type of file and the host computer (and which is stored locally). Playback (sound) settings can be selected prior to playback, but it is not clear which setting is appropriate, nor for how long the current setting persists. Further, the playback software may require an internet browser plug-in. Like the contents of melody files themselves (see Section 3.1), playback formats remain largely device-private.

5. The evidence from the five interviews is that the user-settable parameters for melody matching (illustrated in Figure 1) are likely to be only partially understood by even musically experienced users. In particular, the 'Pitch', 'Method' and 'Tuning' settings caused confusion due to different interpretations of the terminology. This suggests a further likely mismatch between device and potential users.

[1] This paper is linked from the Music Library web site.

4. CONCLUSIONS

This case study of the application of OSM (Ontological Sketch Modelling) to the evaluation of a digital music library (the NZDL ML) has shown that OSM is a potentially valuable method for identifying the likely misfits between device and user models of such a system. It has been emphasised that OSM is still evolving, and the results of this analysis will need to be compared with those from different types of system and other varieties of user interface. However, the combination of inspection using OSM and user elicitation via structured interviewing has thrown some interesting light on the kinds of demands which an on-line music library such as this one makes of its potential users. This is in contrast to the mainly lower-level emphasis of other inspection methods such as heuristic evaluation (Mack & Nielsen 1994), which focus more on the details of the user-system interaction than the wider device- and domain-related concepts which the user is required to address. It remains to be seen whether the insights brought by OSM can be as revealing with larger systems, as well as smaller desktop applications. So far, the evidence suggests that both a highly complex workplace environment such as an ambulance control centre (Blandford et al 2002) and a more simple personal schedule organiser (Blandford & Green 1998) are both amenable to OSM analysis. However, the extent to which OSM differs from other ethnological approaches such as activity theory (Nardi 1996) and cognitive work analysis (Vicente 1999), not to mention other usability inspection methods such as heuristic evaluation, remains to be established. In the shorter term it is hoped that the recommendations made to NZDL following this analysis can contribute to the success of that music library.

This analysis of the NZDL Music Library has indicated some potential misfits between device (here, the NZDL ML and its user interface) and potential user (here, musicians wishing to retrieve melody files from a music library by either tune or text matching). These are summarised below.

1. The organisation and contents of the five NZDL ML sub-collections were hard for the user to predict.

2. The range, effect and persistence of NZDL search parameters were difficult for the user to predict and assess.

3. There was sufficient ambiguity in the differences between recorded ('what you sang') and transcribed ('what I heard') tunes as processed by NZDL ML using MELDEX that users would have difficulty in acknowledging successful transcription of sung or played input.

4. It will be difficult for the user to predict both the appropriate software with which to play back retrieved melodies and the persistence of this selection once made.

However, we again emphasise that it was clear from the interviews carried out in support of the analysis that all of the musicians to which we talked would benefit considerably from music libraries such as the NZDL ML, and that all but one of them had the necessary knowledge and experience to make good use of such facilities in relatively short time.

Finally, this case study has identified several features of music digital libraries which would be profitable for future designers to concentrate upon:

- Ensuring that users have a clear idea of the contents of particular music collections (and any sub-collections), in terms of their type, origins, size, and organisation.
- Making clear to users what the effect, scope and persistence of particular search and tune matching parameter settings will be.
- Ensuring that the results of a successful match with sung or played inputs (in melody matching) can be recognised once retrieved and played back on the host computer.
- Making clear the range and suitability of available playback software on a particular host computer, and the persistence of such a setting once made.

5. ACKNOWLEDGMENTS

We are grateful for the information about the NZDL ML provided by David Bainbridge of Waikato University Department of Computer Science that supported this study, plus permission to use the extracts from the NZDL ML interface. The work reported in this paper has been supported by EPSRC grant no. GR/R39108.

6. REFERENCES

BAINBRIDGE, D. (2000). The role of Music IR in the New Zealand Digital Library project. In: *International Symposium on Music Information Retrieval (ISMIR)*, Plymouth, Mass., October 2000.

BAINBRIDGE, D., NEVILL-MANNING, C.G., WITTEN, I., SMITH, L.A. and McNAB, R.J. (1999). Towards a digital library of popular music. In: *Proceedings of the Fourth ACM Conference on Digital Libraries*, Berkely, Cal., 1999. New York: ACM Press.

BLANDFORD, A. E. & GREEN, T.R.G. (2001). From tasks to conceptual structures: misfit analysis. In: J. Vanderdonckt, A. Blandford & P. Gray (eds.), *People and Computers XV - Interaction Without Frontiers (Volume II)*, pp113-116. Joint proceedings of HCI 2001 and IHM 2001, Lille, September 2001. Toulouse: Cepadues-Editions.

BLANDFORD, A.E. & GREEN, T.R.G. (1997). OSM: An ontology-based approach to usability evaluation. In: *Workshop on Representations*, Queen Mary and Westfield College, July 1997.

BLANDFORD, A.E. & GREEN, T.R.G. (1998). Ontological sketch models of schedule organisers. In: *Proceedings of a Workshop on Understanding Work and Designing Artefacts*, University or York, September 1998.

BLANDFORD, A.E., WONG, B.L.W., CONNELL, I.W. & GREEN, T.R.G. (2002). Multiple viewp oints on computer supported team work: a case study on ambulance dispatch. Accepted for publication in HCI 2002 (*People and Computers XVI*).

COSMEA overview (April 2002): http://www.cs.mdx.ac.uk/staffpages/annb/OSM.html

FUHRMAN, M., GAUTHIER, D. & DILLON, A. (2001). Usability Test of VARIATIONS and DML Prototypes. University of Indiana Digital Music Library Project, 2001. http://dml.indiana.edu/pdf/VariationsTest.pdf

MACK, R.L. and NIELSEN, J. (1994). Executive summary. In: J. Nielsen and R.L. Mack (eds.), *Usability Inspection Methods*. New York: John Wiley & Sons.

McNAB, R.J., SMITH, L.A., BAINBRIDGE, D. and WITTEN, I.H. (1997). The New Zealand Digital Library MELody inDEX. In: *D-Lib Magazine*, May 1997.

McNAB, R.J., SMITH, L.A., WITTEN, I.H., HENDERSON, C.L. and CUNNINGHAM, S.J. (1996). Towards the digital music library: tune retrieval from acoustic input. In: *Proceedings of the first ACM International Conference on Digital Libraries*, 1996, pp11-18. New York: ACM Press.

NARDI, B.A. (1996). Activity theory and human-computer interaction. In: B.A. Nardi (ed.), *Context and Consciousness: activity theory and human-computer interaction*. Cambridge, MA: MIT Press.

NZDL (New Zealand Digital Library): http://www.nzdl.org/cgi-bin/library

VICENTE, K. (1999). Cognitive Work Analysis. NJ: Lawrence Erlbaum.