Real music libraries in the virtual future: for an integrated view of music and music information

Michael Fingerhut
Director, Multimedia Library and Engineering Bureau
IRCAM – Centre Pompidou

Digitization and networking have brought in their wake the ubiquity of information and of knowledge together with quasi-instant gratification. While on the one hand text digitization has progressed, here and there, with various degrees of success due to the immensity of the task at hand[1], recorded music has mushroomed on the internet (uncontrollably in most cases, to the chagrin of the multiplicity of rights owners), first on FTP[2] servers and now through peer-to-peer services, principally due to the fact that compact discs obviate the need to digitize their contents. More importantly: recorded music lends itself much more easily to dematerialization than written text: in contradistinction to the latter, it needs analogue or digital "prosthetic devices" (turntables, compact disc players); computers, together with their storage and networking capacity can provide the same sensory perception than that of a compact disc, locally or at a distance.

But music is far from being only recorded music. It consists of musical scores, of books and other forms of publications; of live events and on information about them. While some of these objects will remain physical for quite a long time, others already lend themselves to digitization, and hence to their juxtaposition in specialized information systems. Libraries – and music libraries in particular – have had to handle this increased variety of contents and information for their public, in addition to their traditional holdings.

Be it for text or for music, the search-and-retrieve models used in libraries and on the internet are vastly different. For one thing, the very nature of a library collection – its development policy as well as its intellectual and physical organization – has little to do with that of the sum of all the comparable documents on the internet, originating from unrelated amateur and professional sources alike and finding themselves on the common virtual shelves just by the virtue of networking. Moreover, the sheer size of the internet introduces much noise and imprecision in domain-specific searches (such as: musical work, musical genre, instrumentation and the like). While this may bring up some surprisingly interesting documents, the need to be able to search efficiently across domain-specific sources can be felt[3].

---

[1] The recent announcement from Google about the digitization of ca. 15 millions books within a few years has rocked the world. Not only the relatively small one of librarians: in France, where literature (and culture in general) is a major political issue reflecting on the Gallic self-image of prestige, even the French President Jacques Chirac took a stand and recommended that there be a French counter-project, much as his vaunted predecessor, Charles de Gaulle, had reacted to the perceived American political domination in Europe by kicking NATO out of France and by developing the French nuclear arsenal. Why is it that such a project, of a technological nature, should evoke such a widespread and deeply-felt reaction even (or especially) among those who do not use the internet to read books? One reason is that it probably nurtures the fantasy of the universal Borgesian library, in which every book ever written will be on its virtual shelves, within a single click of everyone's finger as epitomized by the Google "I feel lucky" button. The mind can indeed be boggled by this cataclysmic collapse of all of the world's past and present culture into our immediate present, and by the final realization of the far-removed yearning to possess the fruits of the Tree of Knowledge. The end of history, if it were indeed true.

[2] Acronym for "File Transfer Protocol", a mechanism for exchanging files on the Internet, well before the emergence of the Web.

[3] The Z39.50 protocol allows for distributed searches across library catalogues. Yet its sheer complexity of implementation is a probable cause of its lack of overall success. Functionally, OAI (see later in the paper) provides comparable functions in a much lighter form.

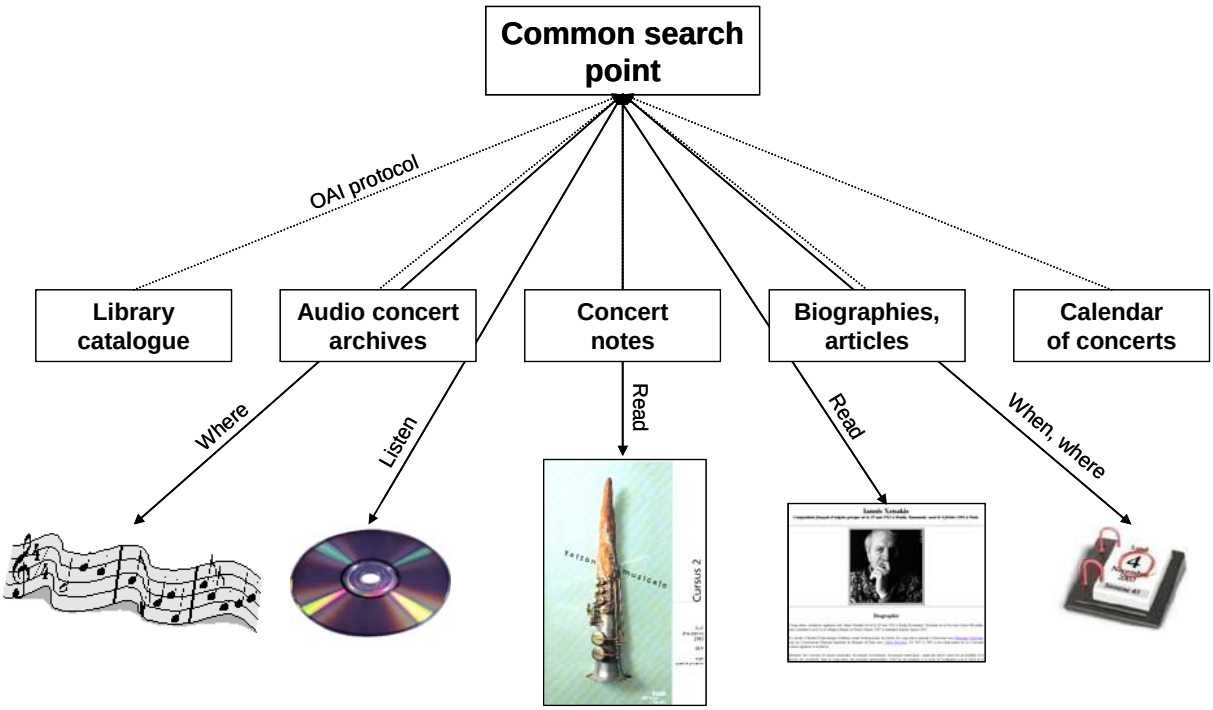

Figure 1. Sharing information sources with OAI

Yet digitization offers unheard-of advances, in addition to the ubiquity of the documents, such as the dynamic reorganization of collections according to various criteria (such as user preferences), the federation of contents across physically distinct places, the machine-assisted extraction and production of "meaning" from arbitrary aggregates of documents as well as from within individual documents and its combination with man-produced meaning, the production of new meaning by the combination of existing documents and the potential of annotating documents without altering the original ones.

Several projects currently under way at IRCAM build upon some of these advantages of digitization in the context of the extended music library.

Sharing information sources

In addition to its catalogue of books, scores and periodical, the IRCAM Multimedia Library holds the archive of the digitized sound archives and concert notes of its public concerts, supplemented by a collection of compact discs of mostly contemporary music, both available online; a separate database, BRAHMS[4], provides biographies of hundreds of living composers and detailed information about their works. Another database contains online versions of the publications of its researchers in all its domains of inquiry (science, technology, musicology and related fields). An online calendar lists the events of its musical season, a year of concerts which culminate in its spring festival, Agora.

Faced with this multiplicity of informational sources, the question arises as to how to provide the patron with an efficient way to access all these hencetofore distinct sources. This is the gist of a project currently under way at IRCAM[5], whose goal is to provide a single access

[4] Base Relationnelle d'Articles Hypertextes sur la Musique du 20e Siècle (relational database of hypertext articles about 20th century music), available here: <http://brahms.ircam.fr/>.
[5] David Denocq, followed by Baptiste Bouillot, Vincent Gourson and the author.

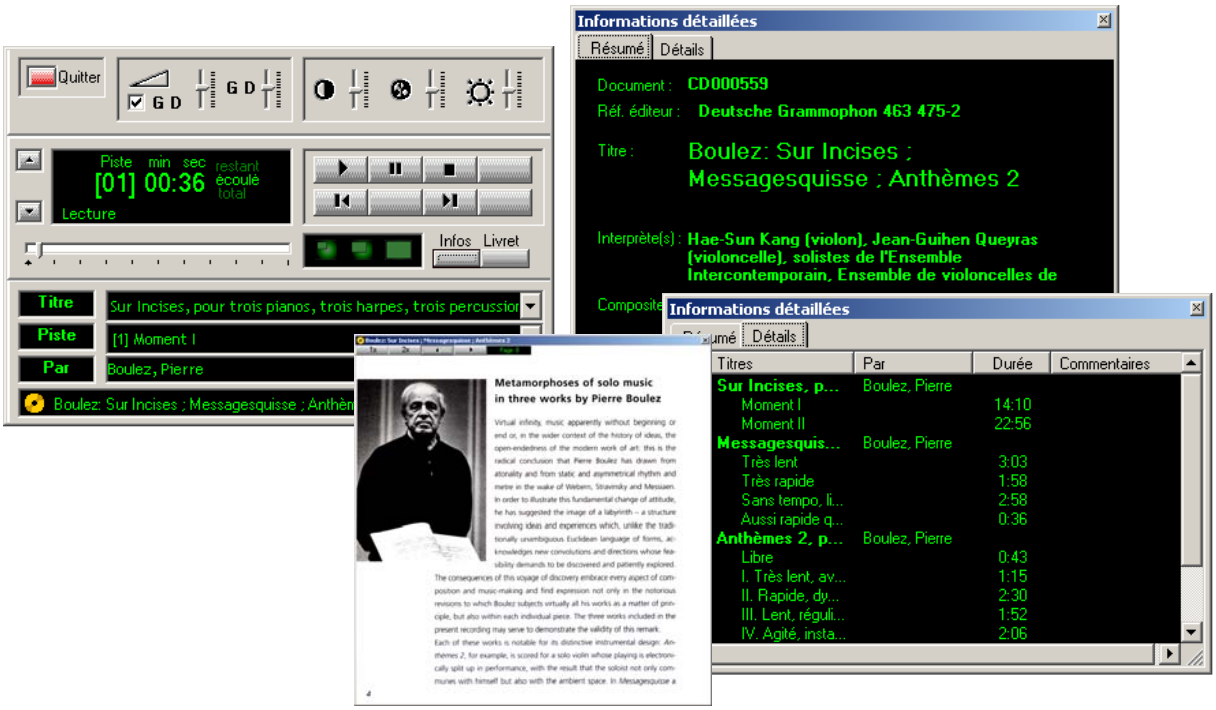

Figure 2. Audio player showing the musical structure

point for a federated search through these sources, using the OAI-PMH[6] protocol to collect the metadata records from each individual source, transforming them into a common model and allowing the user to search within this common pool through specific indexes as well as through full text – in the metadata records and, if available online, in the document itself[7] (see Figure 1).

While the common records are less detailed than those found in each individual database, they provide enough information to satisfy most queries, and include music-specific information, such as instrumentation of musical works. Thus, upon searching, say, for a specific music work by composer and title, this system will return the location of the (physical) scores of this piece and of monographies about it, provide online access to available recordings from past concerts together with the digitized concert notes from those events as well as to online notes about the piece from the BRAHMS database, list dates of future concerts where this piece will be played, etc. This system also allows the user to access the original records as found in the specific databases, if he so wishes.

To implement this system, we have been using SDX[8], which not only allows for the indexation and retrieval of XML documents (including the use of one or several thesauri), but also comprises an OAI module, which is used to query all the relevant databases. This requires

[6] Open Archive Initiative for Metadata Harvesting, a protocol used to *harvest* (collect) metadata records (coded in XML) from possibly heterogeneous *repositories*, usually so as to provide a combined data store as a single search point for all harvested sources. This protocol is much "lighter" to implement and use than z39.50, but a comparison is out of the scope of this paper. See <http://www.openarchives.org/> and <http://www.loc.gov/z3950/agency> for details.

[7] Full-text search in metadata, while very powerful, is not good enough to provide relevant replies to precise queries (such as, say, instrumentation). It is thus useful to combine it with searches within specific indexes.

[8] System for Documentation in XML, an open source search engine and publishing framework for XML documents. See <http://adnx.org/sdx/> for details.

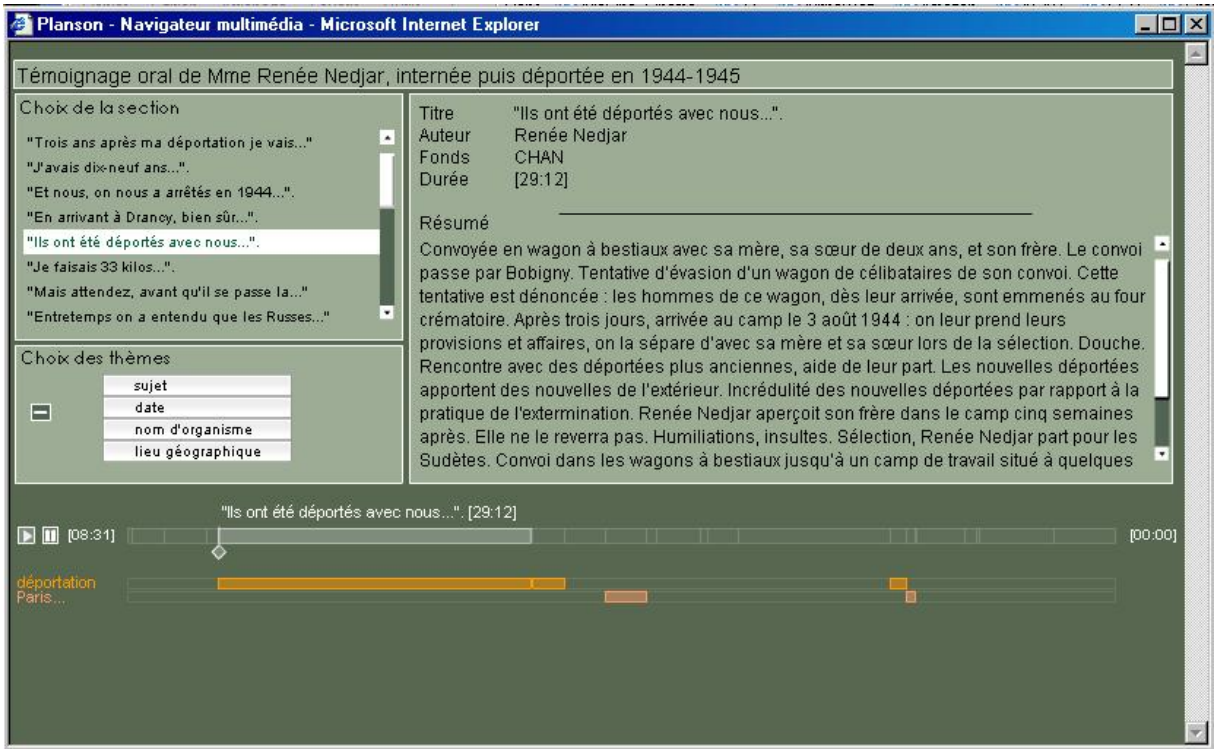

Figure 3. Multimedia player with table of contents, keywords, abstracts

a usually minor development[9] to enable them to respond to these queries, as the protocol is relatively simple to implement.

As this system uses open standards, it can include external sources (which it clearly identifies as such), provided they respond to OAI queries. One interesting such source to include would be the Grove dictionary. Another use of this system which we envision is its extension to operate as a common portal to French contemporary music resources.

Navigating within single documents

Music has structure. Like a book subdivided in chapters, it can be composed of movements (acts and scenes in opera). At a deeper level, it may have a formal structure too (subject, countersubject… etc.). This is indeed true of any audio document (e.g., spoken word). While it is easy to browse through a text document, as its structure is may be apparent to the eye, and as hypertext may be used to highlight it, it is not the case for sound without visual cues. On CDs, for instance, a table of contents on the cover allows for associating some structure with the audio contents (tracks), while current players show a one-dimensional view of the contents.

At IRCAM, we have been faced since 1995 with the need to provide our users with such mechanisms to access our online sound archives. While several emerging standards allow for describing the structure of multimedia documents (such as METS[10]) and for providing so-called hypermedia navigation capabilities "into" those objects, such as (such as SMIL[11]), this

[9] Provided the database software is extensible (i.e., that its software can be extended). Where this is not the case (such as in our library system), alternative ways of exposing the metadata records can be developed.
[10] Metadata Encoding & Transmission Standard), used to encode metadata regarding objects in a digital library. See <http://www.loc.gov/standards/mets/> for details.
[11] Synchronized Multimedia Integration Language, an HTML-like language allowing for the authoring of interactive audiovisual presentations. See <http://www.w3.org/AudioVideo> for details.

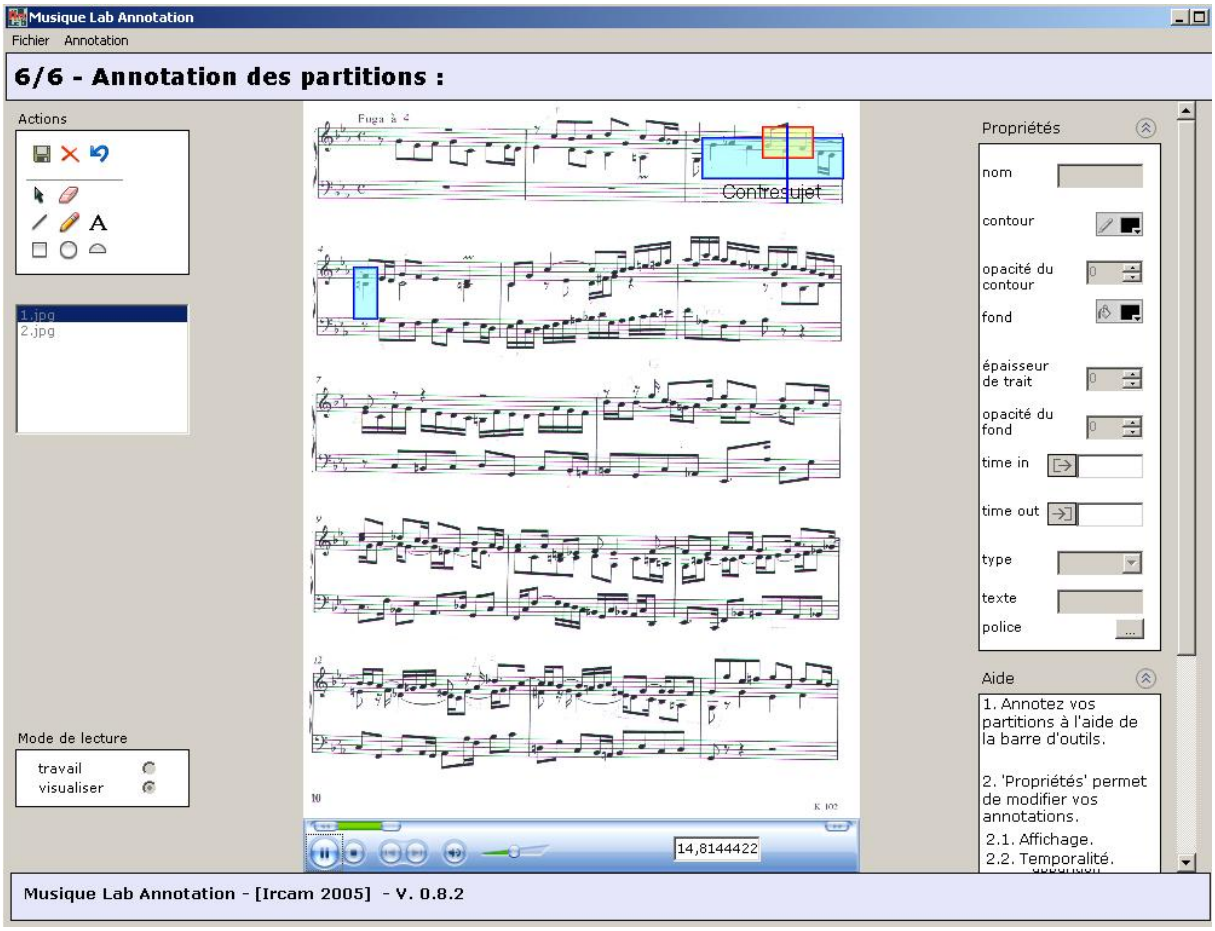

Figure 4. Synchronized and annotated score (only those annotations including the cursor are shown at play time)

was not the case then. So we designed our own simple structured metadata schema representation of pieces and their movements, and developed a stand-alone player (actually, an elaborate "skin" of Windows Media Player, see Figure 2) capable of displaying this structure and allowing the listener to move from piece to piece or movement to movement[12].

More recently, we[13] have been working on the design and development of a network of servers hosting heritage sound collections belonging to French institutions[14] and which have been digitized under the aegis of the Ministry of Culture. While each server allows for direct access to the documents belonging to its organization, there is also a portal providing common search and access to all collections (using SDX on each node to handle the local metadata and OAI to exchange metadata with the portal, in much the same principle as above).

While most collections contain short documents with a single content (a song, a tale) for which a standard audio player could be used, one collection contains particularly lengthy ones: these are recorded interviews of Holocaust survivors, some lasting well over an hour. It was thus necessary to provide the user with a tool to navigate within such document. The holding institution, the French National Archives, uses EAD[15] to describe each individual

[12] Vincent Gourson and the author.
[13] Ludovic Gaillard, followed by Xavier Sirven, under the direction of the author.
[14] Musée des Arts et Traditions Populaires, Musée de l'Homme, Maison Méditerranéenne des Sciences de l'Homme, Archives Nationales.
[15] Encoded Archival Description, an XML schema used to encode archival finding aids. See <http://www.loc.gov/ead/> for details.

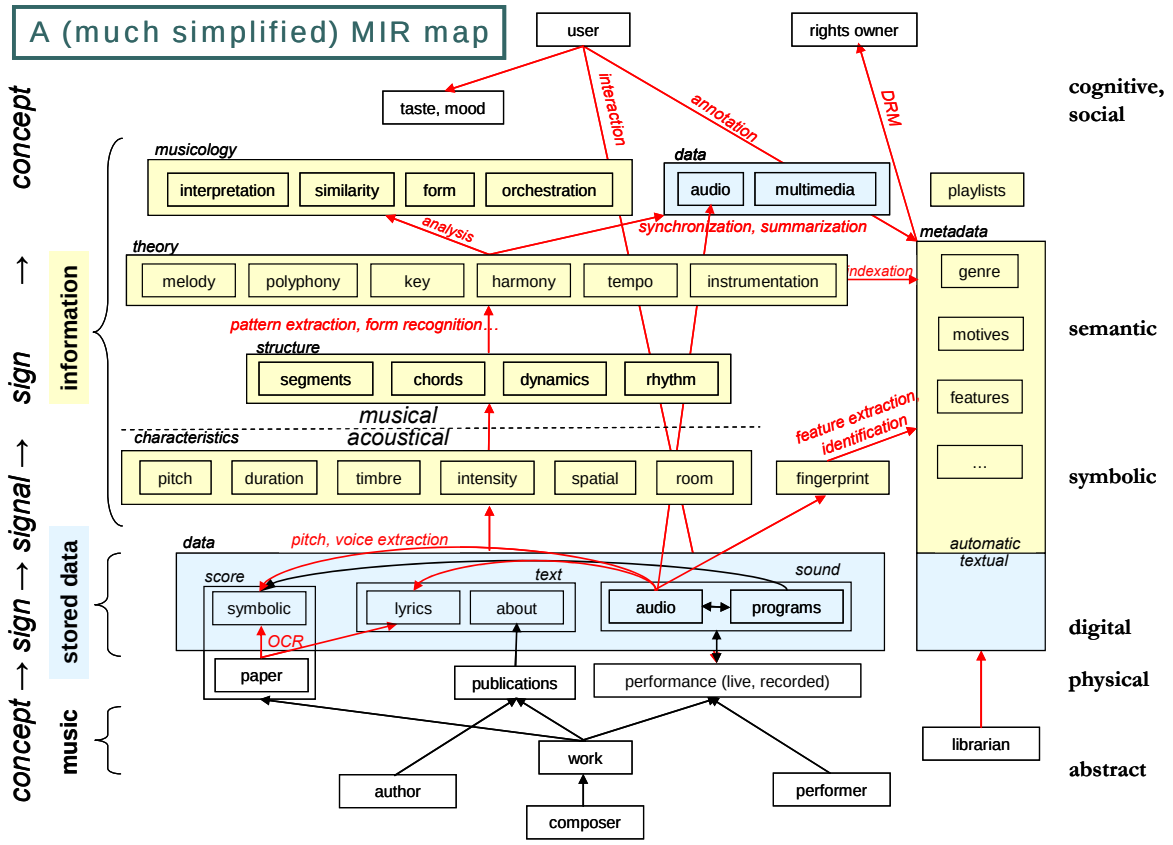

Figure 5. What is music information retrieval all about

document: topical segmentation with incipits, indexation (dates, places…), abstract. Ludovic Gaillard developed an application which transforms those descriptions into a modified SMIL object, including its own navigation capabilities and which can be viewed by an enhanced FLASH[16] player. These objects are integrated in the overall system along with the other, "atomic", audio files, and their "internal" metadata (used to describe their structure and content) is indexed along with the external metadata (used to describe the collections and the files as single elements). Needless to say, this can be used for any kind of sound recording, including classical music.

Non-destructive annotation

Careful reading is an activity which involves, more often than not, annotation. While many tools exist to perform this action for text, IRCAM has been working on developing a toolbox for music annotation, to be used both for teaching and studying music[17] (see Figure 4). This application allows one to synchronize a sound file (typically: a recording of a performance of a piece of music) together with one or several images (typically: the score; but it could be any still – or moving image), and then add annotations of various shapes (rectangles, ovals, lines, images…), colors and textual content, which can appear and disappear at set times when the music plays or stay throughout the piece. Several layers of annotations can be added, and selectively shown at play time. Even the score can be selectively displayed. Several different

[16] See <http://www.macromedia.com/software/flash/> for details.

[17] This development, involving Cedric Godefroy followed by Olivier Zeller, is part of the Music Lab 2 project (under the technical supervision of the author) on behalf of the French Ministries of Culture and Education, whose goal is the development of a suite of software applications for teaching music at the levels of high school and music conservatories.

recordings (e.g., performances) can be synchronized with one single annotated score and thus compared.

This process is non-destructive, as it only references the sound and image files. The annotations are kept in a separate file (in XML). It thus can be safely added on top of the existing digital music library and make full use of the audio documents it contains.

Automatic summaries

Music information retrieval (see Figure 5) is a young cross-discipline (which emerged in the late 1990s), whose main goal is to automatize the extraction of information on music from stored data (recorded sound, midi files, scanned scores, text about music, etc.)[18]. One of its domains of fruitful research is the production of automatic short "summaries" from long audio recordings; while this is a complex problem (what are "meaningful" excerpts, how are they identified, collected and collated?), some useful algorithms have already been developed in various contexts.

A project which is about to start in June aims at using the results of research performed by Geoffroy Peeters at Ircam in order to generate automatic summaries (of 30-45 s.) for all our online audio archives. This will serve two purposes: providing a tool to facilitate the browsing through lengthy online recordings, on the one hand; providing *some* idea of how the piece sounds, for those who access the archive from the net, as the full audio contents are accessible only from within the Library, because of restrictions imposed by intellectual property rights.

In summary

Technology extends the reach of libraries, and in particular of music libraries: it allows to reach more interested people and to provide them with better information, by integrating physical and digital resources, by integrating text, notation, image and sound, and by providing a wider range of useful tools to use this information. These are stimulating times for music librarians.

[18] Michael Fingerhut: *Music Information Retrieval, or how to search for (and maybe find) music and do away with incipits*. IAML-IASA 2004. Oslo (Norway), 2004. Available here: <http://mediatheque.ircam.fr/articles/textes/Fingerhut04b>. The web site of the yearly International Conference on Music Information Retrieval provides access to the cumulative proceedings of all past conferences. See here: <http://www.ismir.net/>.