MuTaTeD'II: A SYSTEM FOR MUSIC INFORMATION RETRIEVAL OF ENCODED MUSIC

Abstract

MuTaTeD'II started in November 1999, building on the results of the MuTaTeD project. Its aim is to design and implement a music information retrieval system with delivery/access services for encoded music. The prototype service will provide a user friendly, web-based search/browse/query interface to access musical content.

Introduction

The Objectives of the project MuTaTeD were: to integrate SMDL - ISO/IEC DIS 10743 (Newcomb, 1991) as the Model with NIFF as one possible View, and establish a standard Meta-DTD for music tagging languages, which could be used by the wide user community. Additionally, it was to research into the development and integration of a SMDL DTD for the wider music user community. The System in MuTaTeD was implemented using LEX/YACC. The work also heavily influenced the "Structured Music MPEG7 proposals" (Boehm/Hall, 1999) which were proposed in order to ensure an SMDL-compliant standard.

Besides proving the concept of integrating these two standards, the use of the LEX/YACC technology in MuTaTeD enabled us to implement the converters, but to certain extent also restricted our flexibility. This restriction was the major reason for implementing the follow-up system in the project MuTaTeD'II in a different way. The MuTaTeD'II team decided to use the Groveminder Technology, which, besides other advantages, caters for the reading-in of different DTDs.

NIFF, being binary, also posed some problems of accessing and searching the content. The fact that any compilation/search procedure has to always be preceded by a reading in or out of the whole binary data in the case of binary formats, has consequences in the design of intuitive content music retrieval, specifically for displaying dynamically parts of music. So although NIFF has proved to be a powerful format, we are contemplating using other text based representation languages for future implementations. Additional information on the two projects can be found at their relevant websites. (MuTaTeD1 WWW 1999, MuTaTeD2 WWW 2000)

MuTaTeD'II

In MuTaTeD'II, we will use Groveminder's support for tagged based languages, specifically SGML based languages. The prototype service will provide a user friendly, expandable web-based search/browse/query interface to access musical content.

Both high level searches and more complex low-level searches will be catered for. High level searches will include searches for specific metadata regarding composer or date of composition. As SMDL supports the storage of a certain set of bibliographic metadata, this will be utilised within this system. Low level searches will concentrate on musical aspects within each score or across a pre-selected collection of scores. This low level searching will enable the user to search for a certain pitch, note or articulation patterns. In the next version of the system, it is planned to provide any combination of pitch, note and articulation searches, and other SMDL elements, such as lyrics will be supported. This will seamlessly transform a search activity into a tool for music analysis.

Two main API's are used, both written in C++. The first is the Groveminder system itself which is used to validate, parse and search SMDL files. The second is GNU Cgicc. This is a gnu open source API which provides functions for talking to a HTTP server and also creating HTML

on the fly. The system internal steps involved in the successful indexing for the searching process include:

1. Opening and parsing a SMDL document for validation, thus building the grove.
2. Parsing the resulting tree/grove to locate the information searched for.
3. Create containers and store search specific data in the containers.
4. Apply the search functions to the containers and return the results.
5. Store results and return to client via web page.

Summary

In the past the availability of music information on the net has been hampered for a variety of reasons, one of the main ones being the propriety nature of the file formats used in most score notation packages. Throughout the Mutated projects, we have tried to overcome some of these restrictions. In MuTaTeD! the task to make SMDL files viewable was undertaken by creating SMDL to NIFF converters. This opened the door for the next obvious step which was carried out in MuTaTeD'II. SMDL being a tag-based language was one of the obvious choices for transporting music information over the internet.

While the process of creating a fully interactive online music information search and analysis tool is still in its embryonic stage it is immediately obvious that a solution of this nature could have a variety of benefits to the music community.

1. high-level to low-level music search facilities
2. adaptable, expandable analysis tools
3. use of an underlying powerful, standardized description language, which does not necessarily restrict the handling of purely encoded music
4. the expansion possibilities into areas of audio/video tagging

Finally, the expansion of the web and the general increase in tag-based languages like XML and SGML mean that, in future, there will be a greater provision for developing applications that deal in these languages. The next generation of browsers (i.e. Netscape 6) will be fully XML compliant and, given time, SGML (or something of a similar level of complexity) will be more commonly used. This drive towards easier interchange of information is inevitable and provides an opportunity to put data under the control of users rather than leaving them prone to the fickle trends of commercial application developers who more often than not have been responsible for the bottleneck in information interchange.

Author Information

Donald MacLellan
Department of Music
University of Glasgow
don@music.gla.ac.uk
tel: +44 (0) 141 330 6065

Carola Boehm
Department of Music
University of Glasgow
carola@music.gla.ac.uk
tel: +44 (0) 141 330 4903

Suggested Readings

Cover, Robin. The XML Cover Pages Website. Http://www.oasis-open.org/cover Especially sections on SMDL, groves and property sets.

Boehm, Carola and MacLellan, Donald. MuTaTeD'II website. http://www.pads.ahds.ac.uk/mutated2.html

Hall, Cordy and Boehm, Carola. MuTaTeD website http://www.pads.ahds.ac.uk/mutated

Newcomb, Steven. 1991 Standards: Standard Music Description Language Complies with Hypermedia Standard. IEEE Computer 24/7/91: p76-79. ISSN:0018-9162