Music Records Markup Language (MuReML)

1. Status of this memo

2. Abstract

   Many music libraries, music centers, authors societies, music
   publishers, music shops, broadcasting companies and public need to
   share bibliographic musical records. No standard format exists and
   exchanging musical records involves an important pre- and/or
   post-processing of these data.

   Searching, sorting and cataloging music bibliographical records does
   not currently follow any standard. The researcher needs, in each
   library and each music information center, to use different
   procedure. In some cases, it is just impossible to obtain the
   targeted result.

   This document defines the requirements for a standard musical

bibliographic format.

. **Introduction**

        Many music libraries, music centers, authors societies, music
        publishers, music shops broadcasting companies and public need to
        share bibliographic musical records. No standard format exists and
        exchanging musical records involves an important pre- and/or
        post-processing of these data.

        Searching, sorting and cataloging music bibliographical records does
        not currently follow any standard. The researcher needs, in each
        library and each music information center, to use different
        procedure. In some cases, it is just impossible to obtain the
        targeted result.

        The following format may be the base of a standard for musical
        bibliographic records.

        It designed as an XML application, as defined by W3C in
        REC-xml-19980210 accessible at http://www.w3.org/TR/REC-xml.html.

        It features all properties of XML metalanguage : a structural
        extensibility, validity controls, independence of data and
        formatting, and it allows heterogeneity of data sources and targets.
        XML will likely be supported by the main web browsers in a short
        future.

        This format fits the goals and recommendations of RFC-2413 (Dublin
        Core Metadata for Resource Discovery) :

            - Simplicity of creation and maintenance
            - Commonly understood semantics
            - Conformance to existing and emerging standards
            - International scope and applicability
            - Extensibility
            - Interoperability among collections and indexing systems

        Organizations may automate to any degree (or not at all) both the
        creation of these records (about their own publications) and the
        handling of the records received from other organizations.

        This format is designed to be simple, for people and for machines,
        to be easy to read ("human readable") and create without any special
        programs.

        The focus of this format has been into many aspects of digital
        libraries including searching and accessing techniques that do not
        necessarily use bibliographic records (for example, natural language
        processing, automatic and full-text indexing). However, the

continued use of bibliographic records is expected to remain an
important part of the library system environment of the future and
its use is an important link between the servers of records and the
clients on site, on line or using a distributed media.

The use of this format is free and encouraged. There are no
limitation on its use.

[4](). Formal Syntax

The following syntax specification uses the Extensible Markup
Language (XML) 1.0.

4.1. Character set

The characters set used is defined by ISO-10646 (coding characters
on 32 bits) and permits the use of symbols and non latin alphabets.
It is preferable, but not mandatory to explicitly declare it.

   <?xml version=1.0" encoding='ISO-10646' ?>

All entities defined by ISO-10646 are permitted but "&" and "<".
They can be used, as any other entity, on pointing the referring
ISO-10646 code or the pre-defined XML entities with the standard XML
syntax :

   & or & for "&"
   $#60; or < for "<"

4.2. Languages

XML's support for multiple human languages, using the "xml:lang"
attribute, handles cases where the same character set is employed by
multiple human languages. In consequence, MuReML is a multi-language
format. It gives the possibility to labellize the chosen language
for each field, and the default language of the record. XML syntax
applied to ISO-639 (for language) and optionally to ISO-3166 (for
regional linguistic particularities) may be used.

   <media_title xml:lang='en-GB'>My Foot!</media_title>
   <media_title xml:lang='fr'>Mon Œil!</media_title>

4.3. Specific formattings

Data of each field may refer to any adequate ISO norm for its
representation. According to XML specification, this norm will be
declared in the opening tag.

```
  <broadcasting_country code="ISO-3166">F</broadcasting_country>
  <world_creation_date format="ISO-8601">1999-10-02T20:30:00Z
   </world_creation_date>
```

4.4. Cases

   Data are case-sensitive.

4.5. White space and End-of-line handling

   The Music Records Markup Language, as a subset of XML, has the same
   white space handling and end-of-line handling as specified in
   Extensible Markup Language (XML) 1.0 (W3C Recommendation
   10-February-1998).

4.6. Grammar

   XML has been chosen because it is a flexible, self-describing,
   structured data format that supports rich schema definitions, and
   because of its support for multiple character sets. XML's self-
   describing nature allows any property's value to be extended by
   adding new elements.

   This format is a "tagged" format with self-explaining alphabetic
   tags. It should be possible to prepare and to read bibliographic
   records using any text editor, without any special programs.

   It is very easy to adapt any database for reading and writing this
   format. Converters may be developed to transform such data from this
   format to plain text or HTML for example.

   As an XML application, the lay-out and the design of the formatted
   data may be freely cosen by normalized style sheet mechanism like
   Cascading Style Sheet (CSS1, CSS2) or Extensible Stylesheet Language
   (XSL).

   Since linear records are unable to efficiently manage the relation
   between the different kinds of information involved in music records
   management, the relational aspect of cataloguing must be maintained.

   Each element has a descriptive name intended to convey a common
   semantic understanding.

   Each packet of data considered in this format contains all
   significant information regarding a specific aspect of a record.
   This involves that several linked tables with several fields are
   necessary. Some fields are mandatory to insure integrity of the
   records and consistency of the links.

Each packet starts with an indentifier (eventually random). This
identifier is to check the relative identity of each packet and to
make it traceable. A community of users may use it to identify its
own packets.

4.7. The tables

The various tables must follow the format described below. This
diagram constitutes the minimum requirement of the format. It can be
extended with other tables for particular uses. To fit the needs of
musical records management (for example : highest hierarchy of
tables, unnecessary differentiation of the various contributors...),
this structure lightly diverges from the recommendations of the
Dublin Core Metadata Element Set.

Some tables as one-to-many relationship with others. It involves
that some tables may be repeated as needed, for example for works
with several rights-holders (composer, author, arranger, publisher,
sub-publisher...) or for media with including several versions.
Tables are also optional. They may appear in any order inside a
particular packet.

MEDIA
 Records relative to the supports of the versions.

OCCURRENCE
 Records relative to the occurrence of a particular version in a
  particular format.

RAPPORT
 Records relative to the rapport between a particular version
 and a rights-holder.

RIGHTS-HOLDER
 Records relative to the rights-holders of the works
 (composers, librettist, arranger, publisher, sub-publisher...).

VERSION
 Records relative to the instrumental versions of the works.

WORK
 Records relative to the original works.

4.8. The fields

The various fields should follow the format described below. These
diagrams constitute the minimum requirement of the format. They can
be extended with other fields for particular uses. These
complementary fields names (or tags) have to be built in accordance

of XML requirements.

These fields are repeatable. A missing mandatory field invalidates
the packet.

Each field tag name begins with the parent table name, followed
by an underscore. For example : <work_title>Monochrone</work_title>

[M] means Mandatory; a record without it is invalid.

[O] means Optional (here to illustrate the extensibility of MuReML)

[L:FIELD] designs the table targeted by a link. Just the fields are
parts of this format. Links will be optionally used in the database
systems to optimize the data management and the consultation of the
records.

```
PACKET
-----
[M] id

MEDIA
-----
[M] id
[M] title
[M] type
[O] producer
[O] localization
[O] keywords
[O] notes

OCCURRENCE
---------
[M] id
[M] id_version  [l:version]
[M] id_media    [l:media]
[O] keywords
[O] notes

VERSION
-------
[M] id
[M] id_work     [l:work]
[M] specificity
[O] opus
[O] start_composition_date
[O] start_composition_place
[O] end_composition_date
[O] end_composition_place
[O] keywords
[O] notes
```

```
WORK
----
[M] id
[M] title
[O] original language title
[O] US-english title
[O] start_composition_date
[O] start_composition_place
[O] end_composition_date
[O] end_composition_place
[O] duration
[O] citations
[O] keywords
[O] notes

RAPPORT
-------
[M] id
[M] id_rights-holder  [l:rights-holder]
[M] id_version        [l:version]
[M] status
[O] keywords
[O] notes

RIGHTS-HOLDER
-------------
[M] id
[M] name
[O] type
[O] contact
[O] keywords
[O] notes
```

4.9. Meta Format and DTD

   MuReML is an open format. Communities of users may enlarge it to
   their own needs. The minimal elements needed in a DTD to fit the
   MuReML specifications are :

```
<!ELEMENT packet (packet_id, media*, occurrence*, version*,
         work*, rapport*, rights-holder*) >
<!ELEMENT packet_id (#PCDATA) >
<!ELEMENT media (media_id, media_title, media_type) >
<!ELEMENT media_id (#PCDATA) >
<!ELEMENT media_title (#PCDATA) >
<!ELEMENT media_type (#PCDATA) >
<!ELEMENT occurrence (occurrence_id, occurrence_id_media,
         occurrence_id_version) >
```

```
<!ELEMENT occurrence_id (#PCDATA) >
<!ELEMENT occurrence_id_media (#PCDATA) >
<!ELEMENT occurrence_id_version (#PCDATA) >
<!ELEMENT version (version_id, version_id_work,
          version_specificity) >
<!ELEMENT version_id (#PCDATA) >
<!ELEMENT version_id_work (#PCDATA) >
<!ELEMENT version_specificity (#PCDATA) >
<!ELEMENT work (work_id, work_title) >
<!ELEMENT work_id (#PCDATA) >
<!ELEMENT work_title (#PCDATA) >
<!ELEMENT rapport (rapport_id, rapport_id_rights-holder,
          rapport_id_version) >
<!ELEMENT rapport_id (#PCDATA) >
<!ELEMENT rapport_id_rights-holder (#PCDATA) >
<!ELEMENT rapport_id_version (#PCDATA) >
<!ELEMENT rights-holder (rights-holder_id, rights-holder_name) >
<!ELEMENT rights-holder_id (#PCDATA) >
<!ELEMENT rights-holder_name (#PCDATA) >
```

5. Example

```
--------------------- Begin of Example -------------------

<?xml version='1.0' encoding='ISO-10646' standalone='no' ?>

<packet xml:lang='en-GB'>
   <packet_id>AVK990127223015</packet_id>

   <media>
      <media_id>BS542187935</media_id>
      <media_title>Two works for four hands</media_title>
      <media_type>music sheet</media_type>
      <media_producer>Big Deal Publishing</media_producer>
      <media_notes>Produced by the publisher</media_notes>
   </media>

   <occurrence>
      <occurrence_id>a12354879654-12</occurrence_id>
      <occurrence_version_id>PF2H0001</occurrence_version_id>
      <occurrence_media_id>BS542187935</occurrence_media_id>

      <occurrence_id>a12354879655-13</occurrence_id>
      <occurrence_version_id>PF2H0002</occurrence_version_id>
      <occurrence_media_id>BS542187935</occurrence_media_id>
   </occurrence>

   <version>
      <version_id>PF2H0001</version_id>
      <version_work_id>00000001</version_work_id>
```

```xml
        <version_specificity>piano four-hand</version_specificity>
        <version_opus>102</version_opus>
        <version_notes>ordered by the publisher and dedicated
         to AmF</version_notes>

        <version_id>PF2H0002</version_id>
        <version_work_id>00000002</version_work_id>
        <version_specificity>piano four-hand</version_specificity>
        <version_opus>102</version_opus>
        <version_notes>ordered by the publisher</version_notes>
</version>

<work>
        <work_id>00000001</work_id>
        <work_title xml:lang='it'>La bella Postina</work_title>
        <work_start_composition_date format="ISO-8601">1999-02-05
         </work_start_composition_date>
        <work_duration>00:12:30</work_duration>
        <work_keywords>modules, rehearsals, repetitive, composer's
         introduction</work_keywords>

        <work_id>00000002</work_id>
        <work_title>Jazz</work_title>
        <work_start_composition_date format="ISO-8601">1999-01-15
         </work_start_composition_date>
        <work_end_composition_date format="ISO-8601">1999-01-30
         </work_end_composition_date>
        <work_duration>00:09:00</work_duration>
</work>

<rapport>
        <rapport_id>5478985251454117</rapport_id>
        <rapport_rights_holder_id>BE_ED001<rapport_rights_holder_id>
        <rapport_version_id>PF2H0001<rapport_version_id>
        <rapport_status>original publisher</rapport_status>

        <rapport_id>5478985251454117</rapport_id>
        <rapport_rights_holder_id>BE_CP001<rapport_rights_holder_id>
        <rapport_version_id>PF2H0001<rapport_version_id>
        <rapport_status>composer</rapport_status>

        <rapport_id>5478985251454117</rapport_id>
        <rapport_rights_holder_id>BE_CP002<rapport_rights_holder_id>
        <rapport_version_id>PF2H0002<rapport_version_id>
        <rapport_status>composer</rapport_status>
</rapport>

<rights-holder>
        <rights-holder_id>BE_ED001</rights-holder_id>
        <rights-holder_name>Big Deal Publishing</rights-holder_name>
        <rights-holder_type>publisher</rights-holder_type>
```

```
        <rights-holder_contact>Alain Van Kerckhoven
         </rights-holder_contact>
        <rights-holder_keywords>post-modernism, classical, Devreese,
         Lachert, Lysight, Mendes, Pelecis</rights-holder_keywords>
        <rights-holder_notes>Founded in 1989</rights-holder_notes>

        <rights-holder_id>BE_CP001</rights-holder_id>
        <rights-holder_name>Lachert, Piotr</rights-holder_name>
        <rights-holder_type>composer</rights-holder_type>
        <rights-holder_contact>Lachert, Piotr</rights-holder_contact>
        <rights-holder_keywords>post-modernism, computer music,
         letters music</rights-holder_keywords>

        <rights-holder_id>BE_CP002</rights-holder_id>
        <rights-holder_name>Lysight, Michel</rights-holder_name>
        <rights-holder_type>composer</rights-holder_type>
        <rights-holder_contact>Lysight, Michel</rights-holder_contact>
        <rights-holder_keywords>post-modernism, repetitive music,
         minimalism</rights-holder_keywords>
    </rights-holder>

</packet>

------------------------ End of Example -------------------

Indentations and line-breaks are for convenient visualization.

This example illustrates a music sheet (MEDIA BS542187935) titled
"Two works for four hands". It includes one version (PF2H0001 and
PF2H0002) of two different works : "La bella Postina" (00000001) and
"Jazz" (00000002). The first one is published and has two
rights-holders : the publisher Alain Van Kerckhoven (BE_ED001) and
the composer Piotr Lachert (BE_CP001). The second one is unpublished
and has been reproduced with a simple agreement of the composer, who
has the all rights : Michel Lysight (BE_CP002).

For reference, the above example contains 3,405 characters.
```

6. Mandatory fields description

PACKET_ID
Any value (random, sequential or inductive) marking the beginning
and the end of each packet, in order to avoid merging of packets
in case of a media default.

MEDIA_ID
Identifies the media record and is used in management of these
records.

MEDIA_TITLE
Main title of the media. If necessary, sub-titles or translations of
the title have to fill other fields.

MEDIA_TYPE
Type of support : music sheet, CD, CD-ROM, DVD... Formats of the
information can be described in other fields (encoding, file type,
standard, compression...).

occurrence_ID
Identifies the occurrence of a version in a media.

occurrence_VERSION_ID
Points to a specific version.

occurrence_MEDIA_ID
Points to a specific media.

VERSION_ID
Identifies the version record and is used in management of these
records.

VERSION_WORK_ID
Points to a specific work.

VERSION_SPECIFICITY
Main information making this version different from the other
versions of the same work. It will often contain formation data :
clarinet and piano, flute and piano etc.

WORK_ID
Identifies the work record and is used in management of these
records.

WORK_TITLE
Main title of the media. If necessary, sub-titles or translations
of the title have to fill other fields.

RAPPORT_ID
Identifies the rapport between a rights-holder and a version.

RAPPORT_ID_RIGHTS-HOLDER
Points to a specific rights-holder

RAPPORT_ID_VERSION
Points to a specific rights-holder.

RAPPORT_STATUS
Describes the status of the rights-holder regarding the pointed
version. A composer may be an arranger, and a publisher may be
a librettist.

```
RIGHTS-HOLDER_ID
Identifies a rights-holder record and is used in management of these
records.

RIGHTS-HOLDER_NAME
Name of the rights-holder (person of company). This includes
composers, publishers, sub-publishers, librettists, transcribers,
illustrators, arrangers, orchestrators etc.
```

7. Security Considerations

The Music Records Markup Language, as a subset of XML, has the same
security considerations as specified in [RFC-2376].

8. Acknowledgments

This document has benefited greatly from the luminous suggestion by
Mark Needlaman to move my first format proposition
(draft-avk-bib-music-rec-01.txt) into a XML application.

Thanks to John Stracke for introducing the Dublin Core to me.

Thanks to Steve Coya and to IESG for critics of the first release of
this memo.

Thanks to the "lazy bits" of Brussels. You know who you are.

Thanks to Mireille.

9. References

[1] Alvestrand, H., "Tags for the identification of languages", RFC
    1766, UNINETT, March 1995.
    <http://www.faqs.org/rfcs/rfc1766.html>

[2] Berners-Lee, T., and D. Connolly, "HyperText Markup Language
    Specification - 2.0", RFC 1866, MIT/LCS, November 1995.
    <http://www.faqs.org/rfcs/rfc1866.html>

[3] W3C XML Working Group (WG), "Extensible Markup Language (XML)
    1.0", W3C Recommendation, February 1998.
     <http://www.w3.org/TR/1998/REC-xml-19980210>

[4] Weibel, S., Kunze, J., Lagoze, C., Wolf, M., "Dublin Core
     Metadata Element Set" <http://purl.org/metadata/dublin_core>

[5] Weibel, S., Kunze, J., Lagoze, C., Wolf, M., "Dublin Core

         Metadata for Resource Discovery" RFC-2413
         <http://www.faqs.org/rfcs/rfc2413.html>

    [6] Date and Time Formats (based on ISO 8601), W3C Technical Note,
         <http://www.w3.org/TR/NOTE-datetime-970915.html>

    [7] Ohta, M., "Character Sets ISO-10646 and ISO-10646-J-1", RFC
         1815, Tokyo Institute of Technology, Juy 1995.
         <http://www.faqs.org/rfcs/rfc1815.html>

    [8] ISO-639, "Code for the representation of names of languages.",
         International Standards Organization, 1988

    [9] ISO-3166, "Codes for the Representation of Names of Countries",
         International Standards Organization, May 1981.

10. **Author's Address**

   Alain Van Kerckhoven
   avenue Broustin 110
   B-1083 Brussels
   Belgium

   Phone: +32 2 420.21.21
   Fax : +32 2 420.05.05
   EMail: alain@avk.org

11. **Full Copyright Statement**