

# **Standards for Music Information**

darkness

Version 0.7.1

Last revised: September 13, 2006

Generated: September 13, 2006

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Assumptions and limitations . . . . .	3
<b>2</b>	<b>Definitions</b>	<b>3</b>
2.1	Catalog . . . . .	3
2.2	Sets, Collections, and Tracks . . . . .	3
2.3	Fields . . . . .	4
<b>3</b>	<b>Standard fields</b>	<b>4</b>
3.1	Artist name . . . . .	4
3.2	Extra artists . . . . .	4
3.3	Modifying artists . . . . .	4
3.4	Compilation producer . . . . .	5
3.5	Set title . . . . .	5
3.6	Set subtitle . . . . .	5
3.7	Country of release . . . . .	5
3.8	Media type . . . . .	5
3.9	Release date . . . . .	5
3.10	Collection title . . . . .	5
3.11	Collection subtitle . . . . .	6
3.12	Collection position and collection total . . . . .	6
3.13	Genre . . . . .	6
3.14	Track title . . . . .	6
3.15	Track subtitle . . . . .	6
3.16	Modification subtitle . . . . .	6
3.17	Track position and tracks total . . . . .	7
<b>4</b>	<b>Acquisition</b>	<b>7</b>
4.1	Authoritative sources of information . . . . .	7
4.1.1	Collection/Set packaging . . . . .	7
4.1.2	Web sites . . . . .	8
4.1.3	Electronically distributed or bootleg music . . . . .	8
4.2	Compilation sets . . . . .	9
4.2.1	Artist names . . . . .	9
4.2.2	Obtaining information from other sets . . . . .	9
4.3	Pre-release sets . . . . .	9
4.4	Bootleg sets . . . . .	9
4.5	Bonus tracks . . . . .	10
4.5.1	Track title . . . . .	10
4.5.2	Track position and tracks total . . . . .	10
4.6	Censored fields . . . . .	11
4.7	Set titles and collection titles . . . . .	11
4.8	Set subtitles and collection subtitles . . . . .	11
4.9	Genre . . . . .	11

4.10	Release date for music transferred to a new media . . . . .	11
4.11	Unknown information . . . . .	12
4.12	Consistency of information within a set . . . . .	12
<b>5</b>	<b>Formatting</b>	<b>12</b>
5.1	Tags . . . . .	13
5.2	General formatting . . . . .	13
5.2.1	Fields with multiple values . . . . .	13
5.2.2	Capitalization . . . . .	13
5.2.3	Punctuation and other symbols . . . . .	14
5.2.4	Dates . . . . .	15
5.2.5	Geographic locations . . . . .	15
5.3	Pre-release sets . . . . .	15
5.4	Censored (“Clean”) sets . . . . .	15
5.5	Artist name . . . . .	15
5.6	Collection subtitle . . . . .	16
5.7	Satisfying uniqueness requirements . . . . .	16
5.7.1	Artist name . . . . .	16
5.7.2	Set title . . . . .	16
<b>6</b>	<b>Acknowledgments and comments</b>	<b>17</b>
6.1	Acknowledgments . . . . .	17
6.2	To-Do . . . . .	17
6.3	Contacting the author . . . . .	18
<b>A</b>	<b>Trusted cataloging sites</b>	<b>18</b>
<b>B</b>	<b>Trusted retail sites</b>	<b>18</b>
<b>C</b>	<b>Acknowledgments and comments</b>	<b>19</b>
C.1	Acknowledgments . . . . .	19
C.2	To-Do . . . . .	19
C.3	Contacting the author . . . . .	19
<b>D</b>	<b>Revision history</b>	<b>20</b>

# 1 Introduction

This document specifies standards for acquisition and formatting of music information. “Music information” includes album titles and artists’ names, among other pieces of information that might be associated with music. The primary goal for this standard is to provide for consistency across catalogs of music information.

## 1.1 Assumptions and limitations

This document was written with rules for American English in mind.

The motivation for writing this document was to provide a standard for music information that would then be embedded within digitally encoded music; in particular, this digitally encoded music would be in the form of compressed files, usually split up by track. Additionally, only contemporary music methods of composition and distribution were seriously considered. Thus, for genres such as Classical music or music not distributed in a typical album-with-tracks set-up, this document may prove highly inadequate.

# 2 Definitions

This section defines terms that will be used throughout this document with special meaning. Other terms may be introduced as they are used, at which time their definitions will be made clear.

## 2.1 Catalog

A *catalog* is a collection of music information. Often this might be called a “collection” (as in “a collection of music”), but use of this term would be in conflict with the definition found in section 2.2. For the purposes of this document *catalog* will always be used to refer to the music information for all tracks in a person’s library.

## 2.2 Sets, Collections, and Tracks

A **set** is a group of *collections* that are distributed together. For example, an “album” containing two CDs is a two CD *set*. It is possible for a *set* to contain only one *collection*.

A **collection** is a group of *tracks* (see below). The most common type of *collection* is a single piece of physical media, such as a single CD. It is possible for a *collection* to contain only one *track*.

A **track** is a single element of a *collection*. In most cases, it is safe to think of a *track* as a single song from a *collection*.

## 2.3 Fields

Information about a track can be split up into small, isolated units called *fields*. A *field* is a single, specific piece of information about a track. Fields have a full name, and are also given “short field identifiers” which are usually a three to four letter abbreviation for a field. Either of these names can be used in this or other documents to refer to a specific field and its accompanying formatting as defined in this document.

Fields are said to be *unset* before they have been assigned a value, or *set* after information has been acquired, formatted, and assigned to the field. Fields may also be *formatted fields* or *unformatted fields*. A *formatted field* is one that contains some sort of text formatting (for example, all lowercase) that is clearly intentional and artistically important. An *unformatted field* is a field that has no intentional or meaningful text formatting.

## 3 Standard fields

This section defines a set of commonly used fields. Information about the acquisition and formatting of these fields can be found in sections 4 and 5.

### 3.1 Artist name

Short field identifier: ARTT

*Artist name* is the name of the primary performing artist for a track. This field must be unique in a catalog.

### 3.2 Extra artists

Short field identifier: XART

The *extra artists* field is used for the name(s) of performing artist(s) not listed in the artist name field. For example, a featured artist would have their name listed in the extra artists field.

### 3.3 Modifying artists

Short field identifier: MART

*Modifying artists* holds the name of any artist(s) who modified the track from its original form. The artist(s) named in this field does not qualify for the *extra artists* field since they did not contribute to the original performance version, but did modify it after recording. For example, if an artist were credited as remixing a track, their name would be placed in this field.

### **3.4 Compilation producer**

Short field identifier: CMPT

*Compilation producer* holds a list of one or more people involved in putting together a compilation set. For example, the name of the DJ that mixed a compilation set would be put in this field.

### **3.5 Set title**

Short field identifier: SETT

*Set title* is the title of the set that includes the track. The pair of fields set title and set subtitle must be unique among all other set title/subtitle pairs within an artist name.

### **3.6 Set subtitle**

Short field identifier: SETS

*Set subtitle* is the subtitle of the set that includes the track.

### **3.7 Country of release**

Short field identifier: CNTR

*Country of release* describes the country in which the track's set was released.

### **3.8 Media type**

Short field identifier: MEDA

*Media type* specifies the original type of media a particular set (or copy of a set) was released on.

### **3.9 Release date**

Short field identifier: RELD

*Release date* is the date this track's set was released to the general public (date of main/major distribution).

### **3.10 Collection title**

Short field identifier: COLT

*Collection title* is the title of the collection on which the track appears. The pair of fields collection title and collection subtitle must be unique among all other collection title/subtitle pairs within a set title/subtitle pair.

### **3.11 Collection subtitle**

Short field identifier: COLS

*Collection subtitle* is the subtitle for the collection the track appears on.

### **3.12 Collection position and collection total**

Short field identifier: COLP, COLZ

*Collection position* is the position within a set occupied by this collection. *Collection total* is the total number of collections in the set. Both of these numbers must always be greater than or equal to one, and both must be integers.

For example, for a two CD set the first CD (collection) in the set would have a *collection position* of 1, and both discs would have *collection total* set to 2. The second CD in the set would, of course, have a *collection position* of 2.

### **3.13 Genre**

Short field identifier: GENR

*Genre* is the style of music that this track belongs to.

### **3.14 Track title**

Short field identifier: TRKT

*Track title* is the name of the track.

### **3.15 Track subtitle**

Short field identifier: TRKS

*Track subtitle* is the subtitle of the track.

### **3.16 Modification subtitle**

Short field identifier: MODS

*Modification subtitle* holds an additional subtitle specifically for a subtitle added by an artist after modifying a track from its original recorded version. For example, a remix subtitle would be placed in this field.

### 3.17 Track position and tracks total

Short field identifier: TRKP, TRKZ

*Track position* and *tracks total* are analogous to *Set position* and *set total*. *Track position* is the track's position within the collection this track is found on. *Tracks total* is the number of tracks in the collection. Both of these numbers must be greater than or equal to one, and both must be integers.

## 4 Acquisition

This section deals with some issues regarding the acquisition of data for fields.

Note that much of music information acquisition requires proper judgment on the part of the acquirer. There are two general rules to help guide this judgment, in order of descending importance:

1. Keep information common between different tracks consistent wherever possible. (Note that “consistent” does not necessarily mean “the same”; for example, see information about artist name DJ mixes.)
2. Preserve the original artist's intent.

Any other concerns may be considered only after these two rules do not result in a clear conclusion.

### 4.1 Authoritative sources of information

This section deals with sources for music information that are considered trustworthy and thus authoritative. Sources not listed in this section of the document must be judged relative to the quality and consistency of these sources, and must only be used as a last resort. If no trustworthy source of information can be found, the acquisition guidelines for unknown fields in section 4.11 must be followed with all fields initially unknown.

The following sources are listed in order of descending preference. The first source from this list to offer information must be selected as the *primary information source*. Missing information must be acquired from sources of lower preference than the primary information source, but only after the primary information source has been checked. More details on the proper acquisition of music information for specific fields follow this section.

#### 4.1.1 Collection/Set packaging

Original collection or set packaging is usually straightforward to read. However, crosschecking other sources of lower preference is recommended, and required when the packaging is of questionable quality, origin, or a mistake (such as a printing error) is suspected. In cases of questionable packaging, other sources

of lower preference can be selected as the primary information source. Missing information must be filled in from another source.

The following are some additional guidelines and tips for reading collection or set packaging. In many cases, if any of the below items are encountered the information must be crosschecked as described above.

- Take care to notice missing articles. For example, articles are commonly omitted from a band's name for aesthetic purposes. Other sets or collections from the same artist(s) may include this article while a few will not. This article should usually be presumed to be part of the artist's name.
- A long title may actually be a title with a subtitle.
- Watch for the difference between a set title and a collection title, if such a distinction is made.

#### **4.1.2 Web sites**

A common source of information for many albums, especially those that were obtained without original packaging, would be web sites. These are subcategorized, in order of descending preference:

**Artist's or label's websites.** The authenticity of these sites must be ensured before acquiring information from them.

**Trusted cataloging sites.** See appendix A for a list of sites currently considered trusted at the time of this document's publication.

**Trusted retail sites.** See appendix B for a list of sites currently considered trusted at the time of this document's publication.

**Fan sites, or other sites.** These sites must possess longevity, have good design, current content, and accurate information (as verified using any known good information from other sources); in other words, high-quality sites. This category of sites must only be used as a last resort.

Checking multiple sites for information is suggested. Also beware of cataloging nuances of some web sites, especially those in the first three categories. For example, some sites might elide articles or leave off featured artists.

#### **4.1.3 Electronically distributed or bootleg music**

In cases of electronic-only (i.e., Internet), bootleg, or other unauthorized distribution of music, often times the data files (the packaging) distributed with the music must be used as the primary source of information. If simple spelling, grammar, or punctuation errors are suspected with reasonable certainty, these errors must be corrected.

## 4.2 Compilation sets

Any set containing original works from multiple artists, modified from their original form or not, is considered a *compilation set*. For example, this might include soundtracks, “best of” albums, remixes, or DJ recordings.

### 4.2.1 Artist names

In order of descending preference, one of the following methods must be used to determine the artist name for a track that is part of a compilation set:

1. The artist name from the original version of the track. For example, if the track is a remix, the artist name will be the original artist’s name rather than the artist who remixed the track. This method may only be used for sets where up to 10% of artist names unknown, with the rules in section 4.11 being used for the unknown artist name fields. If more than 10% of artist names are unknown, another method must be chosen.
2. The compilation performer, editor, or producer if credited prominently on the album. For example, for a DJ recording where the track’s original artist is not known, the DJ’s name must be used.
3. The string “Various Artists” may be used for the artist name.

Exactly one of these methods must be used for an entire compilation set.

### 4.2.2 Obtaining information from other sets

Since compilation sets are composed of tracks probably found elsewhere, if you can find the musically identical track on another set this other set can be used to acquire missing information. However, extra care must be taken to ensure the tracks are musically identical.

## 4.3 Pre-release sets

*Pre-release sets* are sets that are released before the final public distribution of a set has commenced. Pre-release sets must have their music information after the official release where cases of misspelling, punctuation errors, or similar mistakes were included in the pre-release packaging. Track ordering or widely different track names must remain unchanged.

## 4.4 Bootleg sets

For bootleg sets without set titles, the set title must be set as follows:

Bootleg, <Location>

Where “<Location>” is the geographic location where the bootleg was performed. If the location is not known, that portion of the set title must be left off, leaving only “Bootleg”.

All other fields for a bootleg set follow normal rules.

## 4.5 Bonus tracks

Some sets will contain tracks that are not documented on the packaging. These are frequently called “hidden tracks” or “bonus tracks.” If no information on these tracks is available from any authoritative sources, the following rules must be used to derive certain field values for the tracks. These rules are to be applied in combination with, but take precedence over, the rules for unknown fields (see section 4.11).

### 4.5.1 Track title

If no value for the track title field is available, the value “**Bonus track**” should be used. Note that this value should also be used for those tracks which are separate, but which you would consider “hidden tracks.”

### 4.5.2 Track position and tracks total

The track position assigned to the track depends on where the track occurs in its collection:

- If the bonus track occurs before any other documented tracks in the collection, the bonus track will be assigned a track position less than that of the first documented track in the collection. If necessary, this will increase (“shift”) the track position of all tracks including and following the track position that was supplanted.
- If the bonus track occurs after any other documented tracks in the collection, the bonus track will be assigned a track position greater than that of the last documented track in the collection.
- If the bonus track occurs in the middle of the collection, it should be assigned a track position less than<sup>1</sup> or equal to the position of the track that immediately follows it. If necessary, the track position of the immediately following track, and all tracks after that position, will increase.

Sequences of bonus tracks should be kept in order, and track positions of other tracks (i.e., documented tracks) should be increased as necessary. For example, a collection where the first documented track has track position one, but is preceded by three bonus tracks, would have its three bonus tracks numbered one, two, and three in the order they occur, followed by the first documented track taking on track position four.

The tracks total field must reflect all documented tracks as well as bonus tracks.

---

<sup>1</sup>This provides for some sort of obscure case where the track position numbering may have gaps in the numbering.

## 4.6 Censored fields

Sometimes field contents are obviously modified to remove objectionable content. For example, a track title such as *Bulls!@#* has most likely been censored. In such cases the censored content will remain.<sup>2</sup>

## 4.7 Set titles and collection titles

In cases where a distinction between set title and collection title is not clear, the two fields are filled as follows:

- The set title is given the prominent title.
- The collection title is left unset.

## 4.8 Set subtitles and collection subtitles

Beware of cases where a title is listed with a parenthetical name afterwards. For example if a song title was listed on an album cover was “**Foxes Running (Kill the Guys)**”, “**Kill the Guys**” would be the set subtitle.

## 4.9 Genre

Genre can be very subjective. To increase the chance of consistency of genres between different catalogs, very broad genres should be used. Wherever possible, the genre field should be filled with data acquired from an authoritative source of information, or derived from information acquired from an authoritative source of information. For example, using “**Alternative**” when an online retail site categorizes the music as “**Alternative Rock**”.

Also note that the genre field must be consistent for all tracks within a collection, as per section 4.12. Additionally, if possible and appropriate, the genre field should have the same value for all tracks within a set.

## 4.10 Release date for music transferred to a new media

Sometimes an older work is transferred to a new media without any change beyond track ordering. In this case, the release date must be the release date of the first release of the set. Note that the transferred set must contain the exact same tracks that the original work did, or else the release date must be set to the release date of the new media version of the set.

---

<sup>2</sup>Ideally the censorship could be reversed where the original content was clear and the artist was unhappy with the censorship. Obviously determining the intent of the artist, and in some cases even determining the original content of the field, could become very subjective and lead to many divergent versions of the field among copies of the same work. This is unacceptable.

## 4.11 Unknown information

In some cases proper information for a field cannot be acquired, but a value is needed for the field regardless. The following guidelines cover what values a few fields must take on if no information can be acquired for the field. Unset (unknown) fields not covered below must remain unset.

**Artist name** must be set to “Unknown Artist”.

**Track name** must be set to “Unknown”.

**Media type** must be set to “Collection”.

## 4.12 Consistency of information within a set

Some fields for tracks within a set must be consistent across all tracks in that set. Additionally, some fields for tracks within a collection must be consistent across all tracks in the collection. Both of these categories are covered in table 1.

Table 1: Consistent fields

Field	Consistent for tracks within
Artist name	Collection <sup>a</sup>
Set title	Set
Collection title	Collection
Collection subtitle	Collection
Collection position	Collection
Collection total	Set
Tracks total	Collection
Release date	Set
Genre	Collection
Country of release	Set
Media type	Collection

<sup>a</sup>For compilation sets, “consistency” means following the rules in section 4.2.

## 5 Formatting

After the data has been acquired, some fields may require additional formatting. Wherever applicable, the rules in this section must be followed to ensure consistent presentation of field data.

Note that formatted fields cannot have their formatting altered or removed by the rules below. Adding tags to formatted fields is permitted.

## 5.1 Tags

A *tag* is a specific method for appending information to a field. Tag data is not part of field data, but is extra information added to qualify or otherwise augment acquired field data. Tags are used, for example, to satisfy uniqueness for two fields that would otherwise be identical.

A tag is separated from the end of the actual field data by a single space. All tag data for a field is enclosed in one pair of square brackets. The tag data cannot contain leading or trailing spaces. Any character is valid in tag data except for a semi-colon (“;”). A semi-colon is used to separate multiple pieces of tag data. For a field with no tag, the square brackets are placed as described and then the desired tag data is placed between the square brackets. For fields with existing tags, a semi-colon followed by a space is placed at the end of the existing tag data, and then the new tag data is appended.

For example, assume there is a field with the following contents:

```
Some Field Data
```

If you were to add the tag “Some Tag Data” to this field, the field would now have the following contents:

```
Some Field Data [Some Tag Data]
```

If you were to add another tag containing “Some Other Tag Data” to this field, the field would have the following contents:

```
Some Field Data [Some Tag Data; Some Other Tag Data]
```

## 5.2 General formatting

This section presents general formatting guidelines. These rules must be followed for all fields, unless superseded by a rule given for a specific field later in the formatting section, or unless overridden by the presence of a formatted field.

### 5.2.1 Fields with multiple values

When a field has multiple values, those fields must be separated by a comma followed by a space. For example, for a track which has “The Foos” and “Bar Brady” as values for the extra artists field, the final content of the field should be “The Foos, Bar Brady”.

### 5.2.2 Capitalization

Capitalization for unformatted fields is to follow normal American English rules for capitalization wherever possible. A field will be treated like a title for purposes of capitalization.<sup>3</sup> One noted exception to the rules of capitalization are

---

<sup>3</sup>This is an important distinction to make so that the rule for capitalizing the first word in the field is within the rules of American English capitalization.

values or parts of values that are explicitly supplied in this document, such as the content of a tag; such content must appear as supplied.

For reference, a few major capitalization rules from American English are:

- Articles, prepositions, and coordinating conjunctions are always in lowercase. These may include: a, an, and, or, the, but, of, from, for, by, on, so, nor, and yet.
- Subordinate conjunctions must be capitalized. Example of subordinate conjunctions may be: as, because, although, before, after, when, if, where, since, that, and who.
- The first letter of a field must be capitalized.
- Acronyms must be all uppercase.
- Proper nouns, pronouns, adjectives, verbs, and adverbs must have the first letter of each word capitalized.

There are a few special cases which may not be clearly covered in American English reference works:

- Composites of words such as “Decksanddrumsandrockandroll” must be capitalized based on the first word in the composite. In the example given, “Decks” is the first word, so the word would be capitalized. No words internal to a composite of words are capitalized.
- If a word is unknown, such as an indecipherable composite of words, it must be capitalized as long as it is clear the word is not functioning as an article or conjunction.
- Words that begin with non-alphabetic characters, such as “4get” must be treated as composites whose first component is the non-alphabetic sequence of characters. No capitalization inside these words may be used.

### 5.2.3 Punctuation and other symbols

The ampersand character (&) when used as a conjunction must be replaced with the word “and”. The exception to this rule is for ampersands in acronyms, in which an ampersand must not be changed.

Any series of two or more dots that is thought to represent an ellipsis must be replaced with a proper ellipsis, “...”. There must be no space before an ellipsis, and a single space after an ellipsis if text follows the ellipsis.

A single hyphen must be used for hyphenation, ranges, and limits. Two hyphens must be used for an em dash (“—”).

#### 5.2.4 Dates

Dates must be in ISO 8601 date format. This format is:

YYYY-MM-DD

Where YYYY is the four-digit year, MM is the two digit zero-padded month, and DD is the two digit zero-padded day of the month. For example, February 14, 2002 would be written as 2002-02-14. The fields in a date may be removed from right to left if the value is not known. This means that YYYY-MM and YYYY are also valid date formats, but YYYY-DD or MM-DD is not valid. Unless otherwise noted, as much of a date as is known must be recorded.

#### 5.2.5 Geographic locations

Geographic locations are given as:

Specific Information, CC

Where CC is the ISO 3166-1 Alpha-2 country code and **Specific Information** is any information about state, province, or other designator that indicates a subset of the country given by CC.

Unless otherwise noted, geographic locations must be given in the broadest form possible. In most cases, for example, the **Specific Information** can be left off. In this case the field appears as just “CC”.

### 5.3 Pre-release sets

A pre-release set that cannot conclusively be proven to be musically identical to the final released version of the set must have the tag “**Pre-release**” added to the set title.

### 5.4 Censored (“Clean”) sets

Some albums are released in two versions: an “explicit” version which is the unmodified work by the artist, and a “clean” version which has any profanity or unacceptable content removed. For example, a few large retail chains in the US sell only the clean version of an album. In the event that an album is found to be the clean version, the set title must include the tag “**Clean**”. Note that if an album is censored (clean) but there is no explicit version in distribution this section does not apply.

### 5.5 Artist name

The artist name must always be in “natural format”, meaning first name first (as opposed to last name first). For example, “**Fiona Apple**” not “**Apple, Fiona**”.

Leading articles in an artist’s name must be relocated to the end, separated from the rest of the name by a comma. For example, “**The Beatles**” must be changed to “**Beatles, The**”.

## 5.6 Collection subtitle

When collection total is greater than one, a tag containing the media type field, exactly one space, and the collection position must be added to the collection subtitle. For example, for the second CD in a three CD set, a tag containing “CD 2” must be added to the collection subtitle. Note that this requirement ensures the uniqueness of collection title/subtitle, as required in their definition.

## 5.7 Satisfying uniqueness requirements

The fields artist name and set title/subtitle have uniqueness requirements. This section covers all permitted methods for resolving uniqueness conflicts in these fields.

Where multiple methods for resolution are listed, these methods must be applied in order, and their tags added in the order listed, unless otherwise specified. The fewest methods needed to satisfy a field’s uniqueness requirements must be used. Any methods applied must be applied to the field or fields for all tracks involved in the conflict.

Note that, when entering music information for a track, the utmost care for future uniqueness of data must be taken. Sources of data should be checked for other artist names or versions of albums, for example, and the precautions below should be taken when a duplicate is possible, even when such a duplicate does not exist in your catalog.

### 5.7.1 Artist name

Artist names that are non-unique within a catalog must apply one or more of the below methods:

- Adding a geographical location tag corresponding to the country of origin of the artist. This is the country where the artist is considered to come from.
- Adding a tag containing the date of the artist’s first major public release.

In the extreme case where none of the above methods suffice, the only tag added by this section must be an upper case Roman numeral to each conflicting artist name in whatever orders the cataloger sees fit.

### 5.7.2 Set title

If both set title and set subtitle are duplicated for an artist name, one or more of the following methods must be used to make the set title unique:

- For live recordings of a previously recorded non-live set, a tag must be added to the set subtitle in the following format:

Live in <Location>

Where “<Location>” is the geographic location of the performance.

- If different versions of the set were made for different geographical location (for example, different releases for different countries), a tag containing the geographical location of the album’s release must be added to the set subtitle.
- If different versions of the set were released on different media types, a tag specifying the media type of the track must be added to the set subtitle. For example, a tag containing “CD” would be added to the set subtitle for the Compact Disc version of a set meeting these conditions.
- If the same set title/subtitle were used for albums with significantly different or released/rerecorded material, a tag containing the date of release of the set must be added to the set subtitle.
- For bonus or special edition sets, including sets containing extra media (for example, bonus CDs), the tag “Special Edition” must be added to the set subtitle.

If none of the above methods are sufficient to make the set title/subtitle unique, an upper case Roman numeral tag must be added to the set subtitle field of each track involved in the conflict. The Roman numerals must be in order of set release date. The above methods must be used first to make the smallest number of sets in conflict, only after which must the Roman numeral tag be added.

## 6 Acknowledgments and comments

### 6.1 Acknowledgments

Thanks to...

- introvert, ardent, Didel, SoulFly, and anyone on the moo that I forgot.
- The nice folks at <http://www.id3.org/>. The ID3v2.4.0 standard was used for design hints and a starting point for standard fields.
- ISO, and all the nice people who put capitalization tips out on the Internet.
- Google, without which we would all be lost.

### 6.2 To-Do

- A list of allowed genres, or better source for genres would be nice. As things stand right now, genre selection for a piece of music is too subjective.

- Addition of more fields, particularly those supported by ID3v2.4.0. In reflection this may not be good nor necessary. If we feel we have captured enough useful information, we should probably stop. Additionally I've read of a bit of backlash against ID3v2.4.0 WRT it having too many fields.
- Media type list, with acronyms or allowed names for referencing a piece of media of a given type.
- Maybe definitions of “must” and “should,” a la that one RFC.
- Add more examples.
- Fix grammar and punctuation errors in places. For example, commas inside quotes wherever it wouldn't make the expected values for fields unclear. (I just fixed one in a to-do item above: comma outside quotes where I think it should be inside.)

### 6.3 Contacting the author

The author of this document can be contacted at [darkness@codefu.org](mailto:darkness@codefu.org). Please contact me with any grammar or spelling issues, omissions, or parts that are not clear. Interpretation questions would also be welcome. If you simply don't agree with something in this document, and you're sure you understand it correctly, you can contact me; I just don't guarantee that I'll care, unless you present me with a well-reasoned argument.

## A Trusted cataloging sites

This is a list of acknowledged trusted cataloging web sites as of the revision date of this document. These sites seem to store reliable, consistent information; they make no over commercial measures, other than perhaps ads and even links to retail sites; and they often times include information specifically formatted for programmatic or otherwise standard music cataloging data.

<http://www.allmusic.com/>

## B Trusted retail sites

This is a list of acknowledged trusted retail web sites as of the revision date of this document. These are large, trusted retail sites that have been selling music for quite some time on the Internet. They seem largely consistent, have a large amount of musical information (much of probably gleaned directly from the label), and are usable sites.

<http://www.amazon.com/>

## C Acknowledgments and comments

### C.1 Acknowledgments

Thanks to...

- introvert, ardent, Didel, SoulFly, and anyone on the moo that I forgot.
- The nice folks at <http://www.id3.org/>. The ID3v2.4.0 standard was used for design hints and a starting point for standard fields.
- ISO, and all the nice people who put capitalization tips out on the Internet.
- Google, without which we would all be lost.

### C.2 To-Do

- A list of allowed genres, or better source for genres would be nice. As things stand right now, genre selection for a piece of music is too subjective.
- Addition of more fields, particularly those supported by ID3v2.4.0. In reflection this may not be good nor necessary. If we feel we have captured enough useful information, we should probably stop. Additionally I've read of a bit of backlash against ID3v2.4.0 WRT it having too many fields.
- Media type list, with acronyms or allowed names for referencing a piece of media of a given type.
- Maybe definitions of “must” and “should,” a la that one RFC.
- Add more examples.
- Fix grammar and punctuation errors in places. For example, commas inside quotes wherever it wouldn't make the expected values for fields unclear. (I just fixed one in a to-do item above: comma outside quotes where I think it should be inside.)
- Not quite certain that capitalization like a title is right for all fields. For example, custom genres. Should probably think about this more.

### C.3 Contacting the author

The author of this document can be contacted at [darkness@codefu.org](mailto:darkness@codefu.org). Please contact me with any grammar or spelling issues, omissions, or parts that are not clear. Interpretation questions would also be welcome. If you simply don't agree with something in this document, and you're sure you understand it correctly, you can contact me; I just don't guarantee that I'll care, unless you present me with a well-reasoned argument.

## D Revision history

### Version 0.1:

- Initial release.

### Version 0.2:

- Rephrased several things.
- Moved sections around to make things more logical.
- Added tags description.
- Added revision date to cover page.
- Added bootleg set titles to acquisition.
- Added set subtitle and modified subtitle fields.
- Completely rewrote rules for non-unique artist name and set title resolution.
- Removed article relocation for titles.
- Made sure to mention that genres should be broad.
- Probably a few other things I forgot.

### Version 0.3:

- Removed stray ‘‘Tags’’ section under definitions.
- Renamed ‘‘Modified subtitle’’ to ‘‘Modification subtitle’’.
- Added Roman numerals as last resort to set title/subtitle for uniqueness.
- Changed special edition tag from ‘‘Bonus’’ to ‘‘Special’’.
- Changed title from ‘‘Standards for Music Information Formatting’’ to just ‘‘Standards for Music Information’’.
- Fixed several typos, bad phrasings, etc.
- Changed a bunch of ‘‘should’’ to ‘‘must’’.
- Added contact section.
- Added natural ordering requirement for artist name field to formatting section.
- Moved collection title/subtitle uniqueness directly under formatting: now it is a requirement to add ‘‘CD 1’’ and such whenever collection total is greater than one.
- Tags added for uniqueness of set title/subtitle have been moved from set title to set subtitle.
- Added section about set and collection subtitles enclosed in parentheses.

### Version 0.4:

- Changed unknown field values from an option (‘‘may’’) to a requirement (‘‘must’’).
- Fixed a grammar/spelling error or two.

### Version 0.5:

- Added compilation producer field.

Version 0.6:

- Added ‘‘Clean’’ tag.
- Added section on censored fields.
- Added capitalization rules for composite and unknown words.
- Added prepositions to the list of examples of words not to be capitalized.
- Described comma-separated fields briefly.
- Added acquisition rules for bonus tracks.
- Updated to-do.
- Changed contact information.
- Removed CDNow from trusted retail sites, since it’s now totally Amazon.
- Changed capitalization rules to follow that of titles instead of sentences.

Version 0.7:

- Added requirement on ordering by release date to roman numerals for ensuring set title uniqueness.
- Added article relocation for artist name.
- Changed unknown information section to indicate that the given values ‘‘must’’ be used for an unset field, instead of ‘‘may’’ be used.
- Added rule for a release date on an album which is merely a transference of an older work to a new media.
- Added some texttt in examples, like ‘‘Fiona Apple’’.

Version 0.7.1:

- Andy corrected my grammar: parenthesis->parentheses in many places.

\$Revision: 1.4 \$