This text is from the freedb howto, that can be downloaded from the Download/Misc section.

THE FREEDB FILE FORMAT

Database entries must be in the US-ASCII, ISO-8859-1 (the 8-bit ASCII extension also known as "Latin alphabet #1" or ISO-Latin-1) or UTF-8 (Unicode) character set. Lines must always be terminated by a newline/linefeed (ctrl-J, or 0Ah) character or a carriage return character (ctrl-M, or 0Dh) followed by a newline/linefeed character. All lines in a database entry must be less than or equal to 256 characters in length, including the terminating character(s). Database entries with lines that are longer will be considered invalid. There must be no blank lines in a database entry.

Lines that begin with # are comments. Comments should appear only at the top of the file before any keywords. Comments in the body of the file are subject to removal when submitted for inclusion to the database. Comments should consist only of characters in the set:

{ tab (09h); space (20h) through tilde (7Eh) inclusive }

Comments should be ignored by applications using the database file, with several exceptions described below.

The beginning of the first line in a database entry should consist of the string "# xmcd". This string identifies the file as an xmcd format CD database file. More text can appear after the "xmcd", but is unnecessary.

The comments should also contain the string "# Track frame offsets:" followed by the list of track offsets (the # of frames from the beginning of the CD) obtained from the table of contents on the CD itself, with any amount of white space between the "#" and the offset. There should be no other comments interspersed between the list of track offsets. This list must follow the initial identifier string described above. Following the offset list should be at least one blank comment, even though database entries without such a blank comment are also considered valid.

After the offset list, the following string should appear:

"# Disc length: N"

where the number of seconds in the CD's play length is substituted for "N". The number of seconds should be computed by dividing the total number of 1/75th second frames in the CD by 75 and truncating any remainder. This number may not be rounded.
Any string, such as "seconds", may be appended to the line provided there's at least one white space between the amount of seconds and the string. An application must be able to parse the line correctly at all times.

Note for Windows programmers:

The disc length provided by the Windows MCI interface should not be used here. Instead, the lead-out (address of the N+1th track) should be used. Since the MCI interface does not provide the address of the lead-out, it should be computed by adding the length of the last track to the offset of the last track and truncating (not rounding) any remaining fraction of a second. Note that the MCI interface yields an incorrect track offset which must be

corrected by adding one frame to the total frame count when performing the disc length computation. For more information see the DISCID Howto.

After the disc length, the following string should appear:

"# Revision: N"

where the database entry revision (decimal integer) is substituted for "N".

Files missing a revision are assumed to have a revision revision level of 0. The revision is used for database management when comparing two entries in order to determine which is the most recent. Client programs which allow the user to modify a database entry should increment the revision when the user submits a modified entry for inclusion in the database.

After the revision, the following string should appear:

"# Submitted via: client_name client_version optional_comments"

where the name of the client submitting the entry is substituted for "client_name", the version of the client is substituted for "client_version", and "optional_comments" is any sequence of legal characters. Clients which allow users to modify database entries read from the database should update this string with their own information before submitting.

The "client_version" field has a very specific format which should be observed:

[leading text]version_number[release type][level]

Where:

> Leading text: is any string which does not include numbers.
> Version number and level: is any (possibly) decimal-separated list of
>   positive numbers.
> Release type: is a string of the form:
>   alpha, a, beta, b, patchlevel, patch, pl

For example:

> release:2.35.1alpha7
> v4.0PL0
> 2.4

The only required portion of the version field is the version number. The other parts are optional, though it is strongly recommended that the release type field be filled in if relevant. Strict version checking may be applied by software which evaluates the submitter revision, so it is wise to make it clear when a release is beta, etc.

Following the comments is the disc data. Each line of disc data consists of the format "KEYWORD=data", where "KEYWORD" is a valid keyword as described below and "data" is any string consisting of characters in the set:

{ space (20h) through tilde (7Eh) inclusive; no-break-space (A0h) through
  y-umlaut (FFh) inclusive }

or an UTF-8 encoded string.

Newlines (0Ah), tabs (09h) and backslashes (2Fh) may be represented by the two-character sequences "\n", "\t" and "\\" respectively. Client programs must translate these sequences to the appropriate characters when displaying disc data.

All of the applicable keywords must be present in the file, though they may have empty data except for the DISCID and DTITLE keywords. They must appear in the file in the order shown below. Multiple occurrences of the same keyword indicate that the data contained on those lines should be concatenated; this applies to any of the textual fields.
Keywords with numeric data should not have a comma after the last number on each line. Valid keywords are as follows:

DISCID: The data following this keyword should be a comma-separated list of
        8-byte disc IDs. The disc ID indicated by the track offsets in the
        comment section must appear somewhere in the list. Other disc IDs
        represent links to this database entry. Note that linking entries is
        now deprecated and should not be used by submitting programs!
        The algorithm for generating the disc ID is described in the freedb.howto.

DTITLE: Technically, this may consist of any data, but by convention contains
        the artist and disc title (in that order) separated by a "/" with a
        single space on either side to separate it from the text. There may be
        other "/" characters in the DTITLE, but not with space on both sides,
        as that character sequence is exclusively reserved as delimiter of
        artist and disc title! If the "/" is absent, it is implied that the
        artist and disc title are the same, although in this case the name
        should rather be specified twice, separated by the delimiter.
        If the disc is a sampler containing titles of various artists, the disc
        artist should be set to "Various" (without the quotes).

DYEAR: This field contains the (4-digit) year, in which the CD was released.
        It should be empty (not 0) if the user hasn't entered a year.

DGENRE: This field contains the exact genre of the disc in a textual form
        (i.e. write the genre here and do not use e.g. simply the MP3 ID3V1
        genre code). Please note that this genre is not limited to the
        11 CDDB-genres. The Genre in this field should be capitalized, e.g.
        "New Age" instead of "newage" or "new age".

TTITLEN:There must be one of these for each track in the CD. The track
        number should be substituted for the "N", starting with 0. This field
        should contain the title of the Nth track on the CD. If the disc is a
        sampler and there are different artists for the track titles, the
        track artist and the track title (in that order) should be separated
        by a "/" with a single space on either side to separate it from the text.

EXTD: This field contains the "extended data" for the CD. This is intended
        to be used as a place for interesting information related to the CD,
        such as credits, et cetera. If there is more than one of these lines
        in the file, the data is concatenated. This allows for extended data
        of arbitrary length.

EXTTN: This field contains the "extended track data" for track "N". There
        must be one of these for each track in the CD. The track number
        should be substituted for the "N", starting with 0. This field is
        intended to be used as a place for interesting information related to

the Nth track, such as the author and other credits, or lyrics. If there is more than one of these lines in the file, the data is concatenated. This allows for extended data of arbitrary length.

PLAYORDER:
This field contains a comma-separated list of track numbers which represent a programmed track play order. This field is generally stripped of data in non-local database entries. Applications that submit entries for addition to the main database should strip any data from this keyword (i.e. add an empty "PLAYORDER=" line).

A minimal database entry is as follows. A "[...]" indicates repetition.

```
# xmcd
#
# Track frame offsets:
#       150
[ ... 21 frame offsets omitted ]
#       210627
#
# Disc length: 2952 seconds
#
# Revision: 1
# Submitted via: xmcd 2.0
#
DISCID=270b8617
DTITLE=Franske Stemninger / Con Spirito
DYEAR=1981
DGENRE=Classical
TTITLE0=Mille regretz de vous abandonner
[ ... 21 TTITLEN keywords omitted ]
TTITLE22=L'arche de no
EXTD=Copyright (c) 1981 MCA Records Inc.\nManufactured f
EXTD=or MCA Records Inc.
EXTT0=Des Prez\nYez
[ ... 21 EXTTN keywords omitted ]
EXTT22=Schmitt: A contre-voix \n(excerpt)
PLAYORDER=
```

Please note that the EXTD section above is split in 2 pieces and contains a \n. It should be displayed to the user as:

Copyright (c) 1981 MCA Records Inc.
Manufactured for MCA Records Inc.