

痕跡を留める。ファイル形式仕様書

E1.00.00

この文書は、「痕跡を留める。」で使用されるレコード・ファイルの形式について記述します。

目次

痕跡を留める。ファイル形式仕様書.....	1
0 逃げ道と言いつ、そして・・・	3
0.1 本文書の対象.....	3
0.2 本文書の目的.....	3
1. レコード・ファイル群.....	4
1.1 統括ファイル.....	4
1.2 見出し語ファイル	5
1.3 管理ファイル.....	6
1.4 説明ファイル列挙ファイル.....	7
1.4.1 説明ファイル.....	8
1.5 データベース列挙ファイル.....	8
1.5.1 自己参照.....	9
1.5.2 データベース	10
1.6 関連データベース列挙ファイル	11
1.6.1 自己参照.....	11
1.7 被関連データベース列挙ファイル.....	12
1.7.1 自己参照.....	12
1.8 被参照データベース列挙ファイル.....	13
1.8.1 自己参照.....	13
1.9 参照語列挙ファイル.....	14
1.9.1 自己参照.....	14
1.10 被参照語列挙ファイル.....	15
1.10.1 自己参照.....	15
1.11 関連見出し語列挙ファイル.....	16
1.11.1 自己参照	16
1.12 関連ファイル列挙ファイル.....	16
1.12.1 セキュリティ.....	17
1.13 参考文献列挙ファイル.....	17
1.13.1 セキュリティ.....	17
2. 準拠の要件.....	18
2.1 レコード・ファイル群.....	18
2.2 文字コード	18
2.3 関連データベース	18

0 逃げ道と言いつ、そして・・・

この仕様は、データベースについて知識の足りない人間が一所懸命無知を絞って考えたものです。従って、知識のある方から見るとおかしな仕様となっている可能性が高いです。おかしな部分に気付いた方は、こっそりと教えて下さるとありがたいです。というより、意見は積極的に募集します。是非お知恵を拝借して下さい。

以上、作者からの言いつでした。

0.1 本文書の対象

本文書は、作成予定のカード型データベース「痕跡を留める。」で使用予定のレコード・ファイル形式について記述しています。従って、本文書は「痕跡を留める。」形式のレコード・ファイルを使用したデータベース・ソフトウェアを開発するソフトウェア開発者の方を対象にしています。

0.2 本文書の目的

本文書は、「痕跡を留める。」形式のレコード・ファイル仕様を規定するほかに、仕様に準拠したカード型データベース・ソフトウェアを普及させる目的で記述されています。

1. レコード・ファイル群

痕跡を留める。のレコード・ファイルは、下記のファイル群から構成されます。

- ・統括ファイル
- ・見出し語ファイル
- ・管理ファイル
- ・説明ファイル
- ・データベース列挙ファイル
- ・関連データベース列挙ファイル
- ・被関連データベース列挙ファイル
- ・被参照データベース列挙ファイル
- ・参照語列挙ファイル
- ・被参照語列挙ファイル
- ・関連見出し語列挙ファイル
- ・関連ファイル列挙ファイル
- ・参考文献列挙ファイル

以下に、各ファイルの詳細を記します。

1.1 統括ファイル

統括ファイルはレコード・ファイル群の書式を記述します。使用できる文字コードは Shift-JIS, EUC, BOM なし UTF-8 の 3 種です。

統括ファイルのファイル名は”index.idx”固定です。

統括ファイルで規定される書式を以下に記します。

■文字コード

レコード・ファイルを構成するファイル群の内容を記述する際に使用する文字コードを指定します。指定できる文字コードは Shift-JIS, EUC, BOM なし UTF-8 の 3 種です。統括ファイルを記述する文字コードと、統括ファイルで指定する文字コードは一致していなければなりません。

■仕様バージョン情報

データベース構築の元となった「痕跡を留める。」のレコード・ファイル仕様バージョンを記述します。

■データベース・バージョン情報

データベース自体のバージョンを記述します。記述形式は任意です。

■データベース名称

データベースの名称を記述します。記述形式は任意です。

■ ベンダ情報

レコード・ファイル群を構築したベンダの情報を記します。ベンダ情報は1~1024バイトに収めなければなりません。

■ 見出し語ファイル

見出し語ファイルを記述します。見出し語ファイルのファイル名は後述する通り `index.csv` 固定ですが、相対パスでの指定が可能です。フォルダの区切りは”¥”と”/”が使用可能です。

以下に、統括ファイルの記述例を示します。

```
Shift-JIS
E1.00.00
V1.00.00
レトロゲーム機便覧
電子計算機博物館
./index.csv
[EOF]
```

1.2 見出し語ファイル

見出し語ファイルは CSV 形式で記述されます。使用できる文字コードは統括ファイルで指定した文字コードです。ヘッダ行は存在しません。

各レコードの第 1 フィールドで見出し語、第 2 フィールドで管理ファイルのファイル名を指定します。

各フィールドは”(ダブル・クォート)で囲みません。従って、改行を含んだ見出し語を設定することは出来ません。また、半角カンマ(,)を含む見出し語も設定できません。

見出し語ファイルのファイル名は”`index.csv`”固定です。

以下に、見出し語ファイルの記述例を示します。

```
オデッセイ,Odyssey.csv
テレビテニス,TVtennis.csv
[EOF]
```

第 2 フィールドで指定する管理ファイルは相対パス/絶対パスでの記述が可能です。見出し語ファイルと同じフォルダに管理ファイルが存在する場合にのみ、ファイル名だけの指定が出来ます。フォルダの区切りは”¥”と”/”が使用可能です。

1.3 管理ファイル

管理ファイルで使用できる文字コードは統括ファイルで指定した文字コードです。

管理ファイルは9個のレコードから成り、各レコードは下記の情報を示します。

■説明ファイル列挙ファイルの情報

見出し語に対する説明を記述したファイルをフックしているファイルをフックします。

説明ファイル列挙ファイルについての詳細は後述します。

■データベース列挙ファイルの情報

見出し語を含むデータベース・ファイルをフックしているファイルをフックします。

データベース列挙ファイルについての詳細は後述します。

■関連データベース列挙ファイルの情報

見出し語に関連するデータベースをフックしているファイルをフックします。

関連データベース列挙ファイルについての詳細は後述します。

■被関連データベース列挙ファイルの情報

見出し語を含む「痕跡を留める。」のレコード・ファイルを関連データベースとしてフックしているデータベースをフックします。

被関連データベース列挙ファイルについての詳細は後述します。

■被参照データベース列挙ファイルの情報

見出し語を参照しているデータベースをフックしているファイルをフックします。

被参照データベース列挙ファイルについての詳細は後述します。

■参照語列挙ファイルの情報

見出し語の説明ファイルから参照している他の見出し語をフックしているファイルをフックします。

参照語列挙ファイルについての詳細は後述します。

■被参照語列挙ファイルの情報

当該見出し語を説明ファイルの中から参照している見出し語をフックしているファイルをフックします。

被参照語列挙ファイルについての詳細は後述します。

■関連見出し語列挙ファイル

当該見出し語に関連する見出し語をフックしているファイルをフックします。

関連見出し語列挙ファイルについての詳細は後述します。

■関連ファイル列挙ファイル

当該見出し語に関連するファイルをフックしているファイルをフックします。

関連ファイル列挙ファイルについての詳細は後述します。

以下に、管理ファイルの記述例を示します。

```
説明.csv
データベース.csv
関連データベース.csv
被関連データベース.csv
被参照データベース.csv
参照語.csv
被参照語.csv
関連見出し語.csv
関連ファイル.csv
[EOF]
```

各ファイルは相対パス/絶対パスでの記述が可能ですが、管理ファイルと同じフォルダに各ファイルが存在する場合にのみ、ファイル名だけの指定が出来ます。フォルダの区切りは”\”と”/”が使用可能です。

1.4 説明ファイル列挙ファイル

説明ファイル列挙ファイルは CSV 形式で記述されます。使用できる文字コードは統括ファイルで指定した文字コードです。ヘッダ行は存在しません。

各レコードの第 1 フィールドで説明ファイル名、第 2 フィールドで説明ファイルの文字コードを指定します。

第 1 フィールドで説明ファイル名として指定できるのはテキスト形式および HTML 形式のファイルのみです。HTML の対応バージョンは実装依存とします。

第 2 フィールドでは、第 1 フィールドでテキスト形式のファイルを指定した場合に、どの文字コードで記述されているかを指定します。指定できる文字コードは Shift-JIS, EUC, BOM なし UTF-8 の 3 種のみです。第 1 フィールドで HTML 形式のファイルを指定した場合、第 2 フィールドでの指定に優先して HTML のヘッダで定義された文字コードが使用されます。HTML のヘッダで文字コードが指定されていない場合、文字コードの解釈は実装依存です。

各フィールドは”(ダブル・クォート)で囲みません。

以下に説明ファイル列挙ファイルの記述例を示します。

```
説明 1.txt,Shift-JIS
説明 2.htm,EUC
説明 3.html,UTF-8
[EOF]
```

説明ファイルは相対パス/絶対パスでの記述が可能ですが、説明ファイル列挙ファイルと同じフォルダに説明ファイルが存在する場合にのみ、ファイル名だけの指定が出来ます。フォルダの区切りは”¥”と”/”が使用可能です。

1.4.1 説明ファイル

説明ファイルには見出し語の説明を記述します。記述形式およびファイル・フォーマットは本仕様にて規定しません。

1.5 データベース列挙ファイル

データベース列挙ファイルは CSV 形式で記述されます。使用できる文字コードは統括ファイルで指定した文字コードです。ヘッダ行は存在しません。

各レコードの第 1 フィールドでデータベースのレコード・ファイル名、第 2 フィールドでデータベースのレコード・ファイル形式を指定します。

データベース列挙ファイルで指定できるデータベースのレコード・ファイル形式についての詳細は後述します。

各フィールドは”(ダブル・クォート)で囲みません。

以下にデータベース列挙ファイルの記述例を示します。

```
データベース 1.csv,CSV
データベース 2.db3,DB3
データベース 3.xls,Excel
データベース 4.mcd,MCD
```

レコード・ファイルは相対パス/絶対パスでの記述可能ですが、データベース列挙ファイルと同じフォルダにレコード・ファイルが存在する場合にのみ、ファイル名だけの指定が出来ます。フォルダの区切りは”¥”と”/”が使用可能です。

1.5.1 自己参照

データベース列挙ファイルでは、自己参照(該当するデータベース列挙ファイルを含む「痕跡を留める。」データベース・ファイル群を参照すること)が可能です。また、参照しているデータベース内から自分自身(該当データベースを参照しているデータベース列挙ファイルを含む「痕跡を留める。」データベース・ファイル群)を参照している場合も、これを許可します。

よって、「痕跡を留める。」データベースのレコード・ファイルでは、循環参照が発生する可能性があります。循環参照が発生した場合の対応は、実装依存とします。

1.5.2 データベース

データベース列挙ファイルには、見出し語に関連するデータベースのレコード・ファイルを指定できます。ここで言うデータベースとは、カード型データベースおよびリレーショナル・データベースのことを指します。

「痕跡を留める。」では、下記の形式のレコード・ファイルに対応します。

ファイル形式	データベース列挙ファイルでの指定文字列
痕跡を留める。	KAT
EPWING 注 1	EPW
ONESWING 注 1	ONW
ePub 注 1	EPUB
PDIC 注 1	PDIC
CSV 注 2	CSV
Microsoft Excel 注 2	EXCEL
Database File	DB
SQLite	SQLITE
桐 9	TBL
Database File	DBF
dBASE III	DB3
MCardDB	MCD
ユーザ定義 1	USR1
ユーザ定義 2	USR2
ユーザ定義 3	USR3

注 1: EPWING, ONESWING, ePub および PDIC はカード型データベースまたはリレーショナル・データベースの規約ではありませんが、見出し語を検索できるファイル形式であるため、「痕跡を留める。」の仕様に採用しました。

注 2: CSV および Microsoft Excel はカード型データベースまたはリレーショナル・データベースの仕様ではありませんが、データベース的な運用が成されることの多いファイル形式であることから、「痕跡を留める。」の仕様に採用しました。

ユーザ定義 1～3 は、「痕跡を留める。」で未定義のレコード・ファイル形式を指定する場合に用います。ユーザ定義 1～3 の、実際の振る舞いは実装定義とします。

1.6 関連データベース列挙ファイル

関連データベース列挙ファイルは CSV 形式で記述されます。使用できる文字コードは統括ファイルで指定した文字コードです。ヘッダ行は存在しません。

各レコードの第 1 フィールドで関連語、第 2 フィールドでデータベースの名称、第 3 フィールドでデータベースのレコード・ファイル名、第 4 フィールドでデータベースのレコード・ファイル形式を指定します。

関連データベース列挙ファイルで指定できるデータベースのレコード・ファイル形式についての詳細は 1.5.1 データベースに準拠します。

各フィールドは”(ダブル・クォート)で囲みません。

関連データベース列挙ファイルには、見出し語に直接言及していなくとも、関連性のある別の見出し語を含むデータベースのレコード・ファイルを指定できます。

以下に関連データベース列挙ファイルの記述例を示します。

関連語 1,データベース 1,データベース 1.csv,CSV
関連語 2,データベース 2,データベース 2.db3,DB3
関連語 3,データベース 3,データベース 3.xls,Excel
関連語 4,データベース 4,データベース 4.mcd,MCD

レコード・ファイルは相対パス/絶対パスでの記述可能ですが、関連データベース列挙ファイルと同じフォルダにレコード・ファイルが存在する場合にのみ、ファイル名だけの指定が出来ます。フォルダの区切りは”\”と”/”が使用可能です。

1.6.1 自己参照

関連データベース列挙ファイルでは、自己参照が可能です。また、参照しているデータベース内から自分自身を参照している場合も、これを許可します。

よって、「痕跡を留める。」データベースのレコード・ファイルでは、循環参照が発生する可能性があります。循環参照が発生した場合の対応は、実装依存とします。

1.7 被関連データベース列挙ファイル

被関連データベース列挙ファイルは CSV 形式で記述されます。使用できる文字コードは統括ファイルで指定した文字コードです。ヘッダ行は存在しません。

各レコードの第 1 フィールドでデータベースの名称、第 2 フィールドでデータベースのレコード・ファイル名を指定します。

各フィールドは”(ダブル・クォート)で囲みません。

被関連データベース列挙ファイルには、他の「痕跡を留める。」形式のレコード・ファイル群で、当該レコード・ファイル群が関連データベース列挙ファイルに指定されている場合に、参照しているレコード・ファイル群の統括ファイルを指定します。

以下に、被関連データベース列挙ファイルの記述例を示します。

```
データベース 1,../データベース 1/index.idx  
データベース 2,../データベース 2/index.idx  
データベース 3,../データベース 3/index.idx  
データベース 4,../データベース 4/index.idx
```

レコード・ファイルは相対パス/絶対パスでの記述可能です。フォルダの区切りは”¥”と”/”が使用可能です。

1.7.1 自己参照

被関連データベース列挙ファイルでは、自己参照が可能です。

よって、「痕跡を留める。」データベースのレコード・ファイルでは、循環参照が発生する可能性があります。循環参照が発生した場合の対応は、実装依存とします。

1.8 被参照データベース列挙ファイル

被参照データベース列挙ファイルは CSV 形式で記述されます。使用できる文字コードは統括ファイルで指定した文字コードです。ヘッダ行は存在しません。

各レコードの第 1 フィールドでデータベースの名称、第 2 フィールドでデータベースのレコード・ファイル名を指定します。

レコード・ファイル名は”(ダブル・クォート)で囲みません。

被参照データベース列挙ファイルには、他の「痕跡を留める。」形式のレコード・ファイル群で、当該レコード・ファイル群の見出し語ファイルに記述された見出し語が他のデータベースから参照されている場合に、参照しているレコード・ファイル群の統括ファイルを指定します。

以下に、被参照データベース列挙ファイルの記述例を示します。

```
データベース 1,../データベース 1/index.idx  
データベース 2,../データベース 2/index.idx  
データベース 3,../データベース 3/index.idx  
データベース 4,../データベース 4/index.idx
```

レコード・ファイルは相対パス/絶対パスでの記述可能です。フォルダの区切りは”¥”と”/”が使用可能です。

1.8.1 自己参照

被参照データベース列挙ファイルでは、自己参照が可能です。

よって、「痕跡を留める。」データベースのレコード・ファイルでは、循環参照が発生する可能性があります。循環参照が発生した場合の対応は、実装依存とします。

1.9 参照語列挙ファイル

参照語列挙ファイルは CSV 形式で記述されます。使用できる文字コードは統括ファイルで指定した文字コードです。ヘッダ行は存在しません。

各レコードの第 1 フィールドで参照語、第 2 フィールドでデータベースの名称、第 3 フィールドでデータベースのレコード・ファイル名、第 4 フィールドでデータベースのレコード・ファイル形式を指定します。

参照語およびレコード・ファイル名は”(ダブル・クォート)で囲みません。

参照語列挙ファイルには、他のデータベースのレコード・ファイルに含まれる語(データ)を当該データベースから参照している場合に、参照しているデータベースのレコード・ファイルを指定します。対象となるレコード・ファイルが「痕跡を留める。」形式のレコード・ファイルの場合、その統括ファイルを指定します。

以下に、参照語列挙ファイルの記述例を示します。

```
参照語 1,データベース 1,../データベース 1/index.idx,KAT
```

```
参照語 2,データベース 2,../データベース 2/index.idx,KAT
```

```
参照語 3,データベース 3,../データベース 3/index.idx,KAT
```

```
参照語 4,データベース 4,../データベース 4/index.idx,KAT
```

レコード・ファイルは相対パス/絶対パスでの記述可能です。フォルダの区切りは”¥”と”/”が使用可能です。

1.9.1 自己参照

参照語列挙ファイルでは、自己参照が可能です。

よって、「痕跡を留める。」データベースのレコード・ファイルでは、循環参照が発生する可能性があります。循環参照が発生した場合の対応は、実装依存とします。

1.10 被参照語列挙ファイル

被参照語列挙ファイルは CSV 形式で記述されます。使用できる文字コードは統括ファイルで指定した文字コードです。ヘッダ行は存在しません。

各レコードの第 1 フィールドで被参照語、第 2 フィールドでデータベースの名称、第 3 フィールドでデータベースの統括ファイル、第 4 フィールドで被参照語を参照している見出し語を指定します。

被参照語および見出し語は”(ダブル・クォート)で囲みません。

被参照語列挙ファイルには、当該レコード・ファイル群の見出し語ファイルに記述された見出し語を他の「痕跡を留める。」形式のデータベースから参照されている場合に、参照しているレコード・ファイル群の統括ファイルと見出し語を指定します。

以下に、被参照語列挙ファイルの記述例を示します。

被参照語 1,データベース 1,..!データベース 1/index.idx,見出し語 1
被参照語 2,データベース 2,..!データベース 2/index.idx,見出し語 2
被参照語 3,データベース 3,..!データベース 3/index.idx,見出し語 3
被参照語 4,データベース 4,..!データベース 4/index.idx,見出し語 4

レコード・ファイルは相対パス/絶対パスでの記述可能です。フォルダの区切りは”¥”と”/”が使用可能です。

1.10.1 自己参照

被参照語列挙ファイルでは、自己参照が可能です。

よって、「痕跡を留める。」データベースのレコード・ファイルでは、循環参照が発生する可能性があります。循環参照が発生した場合の対応は、実装依存とします。

1.11 関連見出し語列挙ファイル

関連見出し語列挙ファイルは CSV 形式で記述されます。使用できる文字コードは統括ファイルで指定した文字コードです。ヘッダ行は存在しません。

各レコードの第 1 フィールドで関連見出し語、第 2 フィールドで関連見出し語を含むデータベースの名称、第 3 フィールドで関連見出し語を含むデータベース・ファイル(対象データベースが「痕跡を留める。」形式のレコード・ファイルの場合はデータベースの統括ファイル)、第 4 フィールドでデータベースの形式を指定します。

関連見出し語および各フィールドは”(ダブル・クォート)で囲みません。

関連見出し語列挙ファイルには、当該レコード・ファイル群の見出し語ファイルに記述された見出し語に関連する見出し語を含む他のデータベースのレコード・ファイルを指定します。

以下に、関連見出し語列挙ファイルの記述例を示します。

関連見出し語 1,データベース 1,../データベース 1/index.idx,KAT

関連見出し語 2,データベース 2,sample.xls,EXCEL

関連見出し語 3,データベース 3,sample.csv,CSV

関連見出し語 4,データベース 4,../データベース 4/index.idx,KAT

1.11.1 自己参照

関連見出し語列挙ファイルでは、自己参照が可能です。

よって、「痕跡を留める。」データベースのレコード・ファイルでは、循環参照が発生する可能性があります。循環参照が発生した場合の対応は、実装依存とします。

1.12 関連ファイル列挙ファイル

関連ファイル列挙ファイルは CSV 形式で記述されます。使用できる文字コードは統括ファイルで指定した文字コードです。ヘッダ行は存在しません。

各レコードの第 1 フィールドで関連ファイルのタイトル、第 2 フィールドで関連ファイルのファイル名を指定します。

関連ファイルのタイトルおよび各フィールドは”(ダブル・クォート)で囲みません。

関連ファイル列挙ファイルには、当該レコード・ファイル群の見出し語ファイルに記述された見出し語に関連するファイルを指定します。

以下に、関連ファイル列挙ファイルの記述例を示します。

関連ファイル 1,..../画像/kannren1.bmp
関連ファイル 2, http://www.kannren2.com/index.html
関連ファイル 3,..../文書/kannren3.docx
関連ファイル 4, http://www.kannren4.co.jp/index.html

1.12.1 セキュリティ

関連ファイル列挙ファイルに、悪意あるプログラム/マクロが指定されていた場合、または安全ではない URL が指定されていた場合、「痕跡を留める。」ではこれらの攻撃を防ぐ手段を提供できません。従って、関連ファイル列挙ファイルの中身を記述してレコード・ファイルを配布することは非推奨とします。

1.13 参考文献列挙ファイル

参考文献列挙ファイルは CSV 形式で記述されます。使用できる文字コードは統括ファイルで指定した文字コードです。ヘッダ行は存在しません。

各レコードの第 1 フィールドで参考文献のタイトル、第 2 フィールドで参考文献の著者、第 3 フィールドで参考文献のファイル名を指定します。

関連ファイルのタイトルおよび各フィールドは”(ダブル・クォート)で囲みません。

参考文献列挙ファイルは見出し語に連なるレコード・ファイル群を作成する際に参照した文献を列挙します。

以下に、参考文献列挙ファイルの記述例を示します。

参考文献 1,文献太郎,..../文献/bunnken.docx
参考文献 2,文献花子, http://www.bunnken.com/index.html

1.13.1 セキュリティ

参考文献列挙ファイルに、悪意あるプログラム/マクロが指定されていた場合、または安全ではない URL が指定されていた場合、「痕跡を留める。」ではこれらの攻撃を防ぐ手段を提供できません。従って、参考文献列挙ファイルの第 3 フィールドを記述してレコード・ファイルを配布することは非推奨とします。

2. 準拠の要件

本章では、「痕跡を留める。」レコード・ファイル仕様の準拠要件について記します。

2.1 レコード・ファイル群

「痕跡を留める。」仕様に準拠したレコード・ファイルでは、下記レコード・ファイル群を備える必要があります。

- ・統括ファイル
- ・見出し語ファイル
- ・管理ファイル
- ・説明ファイル列挙ファイル
- ・データベース列挙ファイル
- ・関連データベース列挙ファイル
- ・被関連データベース列挙ファイル
- ・被参照データベース列挙ファイル
- ・参照語列挙ファイル
- ・被参照語列挙ファイル
- ・関連見出し語列挙ファイル

2.2 文字コード

「痕跡を留める。」仕様に準拠したレコード・ファイルでは、Shift-JIS, EUC, BOM なし UTF-8 のうち、少なくとも1つの文字コードをサポートする必要があります。

2.3 関連データベース

「痕跡を留める。」仕様に準拠したレコード・ファイルでは、「痕跡を留める。」形式、CSV 形式の2形式をサポートする必要があります。