

---

libnmz  
API

「libnmz API」

Windows版 - 初級編 -

(Version 2.0.19 対応版)

第4版(Version 0.0.0.4)

アカネイロソフトウェア

発行: 2009/05

---

---

---

## 目 次

---

---

|       |                              |    |
|-------|------------------------------|----|
| 1.    | はじめに.....                    | 4  |
| 2.    | 初級編について.....                 | 5  |
| 3.    | libnmz (nmz.dll) のライセンス..... | 5  |
| 4.    | libnmz について.....             | 6  |
| 5.    | インデックス .....                 | 7  |
| 1.    | NMZ.w.....                   | 7  |
| 2.    | NMZ.t.....                   | 7  |
| 3.    | NMZ.field.xxxx .....         | 7  |
| 6.    | ヘッダーファイル .....               | 8  |
| 7.    | ライブラリファイル.....               | 8  |
| 8.    | コンパイル .....                  | 9  |
| 9.    | 手順.....                      | 10 |
| 10.   | サンプルコード.....                 | 11 |
| 11.   | 定数.....                      | 14 |
| 1.    | BUFSIZE.....                 | 14 |
| 2.    | nmz_stat.....                | 14 |
| 3.    | nmz_sortmethod .....         | 15 |
| 4.    | nmz_sortorder.....           | 16 |
| 12.   | 構造体 .....                    | 17 |
| 1.    | nmz_data.....                | 17 |
| 2.    | _NmzResult.....              | 18 |
| 13.   | 主要関数一覧 .....                 | 19 |
| 1.    | libnamazu.h .....            | 19 |
| 2.    | field.h .....                | 19 |
| 3.    | hlist.h .....                | 19 |
| 4.    | idxname.h .....              | 19 |
| 5.    | i18n.h .....                 | 19 |
| 6.    | search.h.....                | 20 |
| 7.    | codeconv.h .....             | 20 |
| 14.   | 関数.....                      | 21 |
| 15.   | 「Namazu の開発日記」メルマガについて.....  | 32 |
| 15.1. | 「Namazuの開発日記」メルマガについて .....  | 32 |
| 15.2. | メルマガ登録/解除 .....              | 32 |

---

---

---

---

|     |          |    |
|-----|----------|----|
| 16. | 免責.....  | 33 |
| 17. | 著作權..... | 33 |
| 18. | 履歷.....  | 34 |

sample

---

---

## 1. はじめに

libnmz は、日本語全文検索システム Namazu の namazu.exe や namazu.cgi.exe で利用している検索用ライブラリです。本書はこの libnmz の使用方法を簡単にまとめたものです。

本書は Windows での利用が前提となっています。Solaris, Linux, FreeBSD 等の UNIX 系 OS で利用される方は、「libnmz API UNIX 版」をお読みください。

Namazu for Windows のバイナリでは Visual C++ 2005 SP1 でコンパイルした libnmz のスタティックリンクライブラリをリンクしています。Search-Namazu では MinGw + gcc でコンパイルした libnmz のスタティックリンクライブラリをリンクしています。

libnmz のダイナミックリンクライブラリは、nmz.dll という名前で **Namazu for Windows** のバイナリのアーカイブに含まれています。Namazu for Windows 2.0.19 以降のアーカイブには、Visual C++ 2005 SP1 でコンパイルした x86 版、x64 版、IA64 版が含まれています。

### ■ Namazu for Windows 2.0.18 では

x86 版 libnmz のダイナミックリンクライブラリは、nmz.dll という名前で Namazu for Windows のバイナリのアーカイブに含まれています。(nmz2.0.18.001-win32.zip 以前のアーカイブに含まれる nmz.dll の代わりに libnmz2.0.18.001-win32.zip に含まれる nmz.dll を使用してください。nmz2.0.18.001-win32.zip より新しいバイナリアーカイブに含まれる nmz.dll はそのまま使用することができます。)

Win64(IA64, x64)版 libnmz は nmz2.0.18.001-win64.zip 他に含まれる IA64 版 nmz.dll, x64 版 nmz.dll をご利用ください。(Win64 版 libnmz については動作確認ができていません。あらかじめご了承ください。)

Windows で libnmz を利用するプログラムを作成する場合は、手軽に利用できるダイナミックリンクライブラリの nmz.dll をご利用ください。本書でも nmz.dll を利用することを前提に説明を行っています。

---

---

---

Namazu は常に最新版をインストールしてご利用ください。最新バージョンはセキュリティ上の問題を含め、多くのバグ修正、性能向上が行われています。

## 2. 初級編について

本書は初級編として、Perl モジュールの Search-Namazu が利用している libnmz の機能に限定して解説しています。

libnmz は Namazu の検索機能を切り出してライブラリ化したものですが、設計が悪くライブラリとして見た場合、汎用性に欠ける関数等も多く存在します。このため、全ての機能を説明するには、namazu.exe や namazu.cgi.exe の内部処理についてより具体的に解説する必要があり、非常に難しい内容となってしまいます。

Search-Namazu では検索に必要な最小限の libnmz の機能を使って作られているため、これらに絞って解説することで必要最小限の libnmz の機能を比較的簡単に理解し、利用できるようになるものと思います。

## 3. libnmz (nmz.dll) のライセンス

libnmz(nmz.dll) のライセンスは LGPL ではなく、GPL ライセンスです。スタティックリンクであろうと、ダイナミックリンクであろうと、libnmz(nmz.dll) を使用するアプリケーションは GPL ライセンスと矛盾しないライセンスでなければなりません。

libnmz(nmz.dll) を使用したアプリケーションを配布する場合は、注意してください。

---

---

## 4. libnmz について

- libnmz を用いると Namazu 2.X 形式のインデックスファイルから検索を行うアプリケーションを比較的簡単に作成することができます。
- libnmz は単語検索、部分一致検索、正規表現検索、フィールド検索等、namazu.exe や namazu.cgi.exe がサポートする全ての検索式をサポートしています。
- 検索式は Shift\_JIS, EUC-JP のどちらでも指定することができます。しかし、UTF-8 はサポートしていません。
- 検索結果はスコア順/日付順/フィールド順にソートして取り出すことができます。また、昇順/降順の指定も可能です。
- 各文書のフィールド値を取り出すことができます。ただし、漢字コードは EUC-JP です。
- URI は HTML エンコードされたまま取り出されます。必要な場合は、アプリケーション側でデコードしてください。
- URI の置換機能はサポートしていません。必要な場合は、アプリケーション側で対応してください。
- 検索語の強調表示機能はサポートしていません。必要な場合は、アプリケーション側で対応してください。
- その他、libnmz は namazurc(namazurc) を読み込みませんので、これらの設定ファイルで設定する項目は全てサポートしていません。
- 複数のインデックスに対応しています。複数のインデックスをまとめて一度に検索することができます。
- libnmz にはインデックスを作成する機能は含まれていません。インデックスは mknmz.bat で作成したものを使用してください。

---

---

## 5. インデックス

### 1. NMZ.w

単語ファイル。

漢字コードは EUC-JP で記録されています。

### 2. NMZ.t

タイムスタンプが記録されているファイル。

-1 はその文書が削除されたことをしめします。

日付ソートのキーとして使用します。

### 3. NMZ.field.xxxx

1 文書当たり 1 行で、文書のフィールド情報が記録されているファイル。

フィールド名をファイル名に含みます。

漢字コードは EUC-JP で記録されています。

#### 主なフィールド名

|            |             |
|------------|-------------|
| from       | 著者          |
| date       | 日付          |
| size       | 文書ファイルのサイズ  |
| subject    | サブジェクト、タイトル |
| summary    | 要約          |
| message-id | メッセージ ID    |
| newsgroup  | ニュースグループ    |
| to         | 宛先          |
| uri        | URI         |

---

---

## 6. ヘッダーファイル

libnmz を利用するアプリケーションでは、次のヘッダーをインクルードしてください。

```
#include "namazu/libnamazu.h"
#include "namazu/field.h"
#include "namazu/hlist.h"
#include "namazu/idxname.h"
#include "namazu/il8n.h"
#include "namazu/search.h"
#include "namazu/codeconv.h"
```

Namazu のヘッダーファイルは標準では C:\¥namazu¥include にインストールされます。

## 7. ライブラリファイル

libnmz を利用するアプリケーションでは、nmz.lib をリンクしてください。

Namazu のライブラリファイルは標準では C:\¥namazu¥lib にインストールされます。

---

---

## 8. コンパイル

libnmz を用いたプログラムをコンパイルするには、Microsoft Visual C++ 2005 SP1 以降が必要です。(無償で提供されている Visual C++ 2005 Express Edition SP1 でも可)ここでは x86 版の libnmz を用いたプログラムを Visual C++ 2005 でコンパイルする場合を例に示します。

Visual C++ 2005 SP1 をコマンドラインで使用する場合は、コマンドプロンプトを開き、vcvars32.bat を実行します。

```
C:> "C:\Program Files\Microsoft Visual Studio 8\VC\bin\vcvars32.bat"
```

ヘッダーファイルが C:\%namazu%\include ライブラリが C:\%namazu%\lib にインストールされていると仮定すると次のように /I, /link オプションを指定します。

```
C:> cl program.c /I:C:\%namazu%\include /link C:\%namazu%\lib\nmz.lib
```

作成した program.exe を実行するには nmz.dll が必要です。nmz.dll を program.exe と同じディレクトリにコピーするか、システムディレクトリにインストールしてください。

### ■ Microsoft Visual C++ 2005 SP1 再頒布可能パッケージ

Visual C++ でコンパイルされたアプリケーションを Visual C++ 2005 がインストールされていない PC で実行するためには、Visual C++ ライブラリのランタイムコンポーネントが必要になります。

Microsoft Visual C++ 2005 SP1 再頒布可能パッケージは、実行に必要な Visual C++ ライブラリのランタイムコンポーネントで、Microsoft のダウンロードサイトから無償で入手することができます。このパッケージはライセンスに従えば再頒布が可能です。

Visual C++ ライブラリのランタイムコンポーネントには、x86 版、x64 版、IA64 版があります。使用している OS の種類に合ったものをインストールしてください。

---

---

## 9. 手順

簡単な libnmz の使用手順を以下に示します。

### 1. 準備

|                    |               |
|--------------------|---------------|
| nmz_set_lang       | 言語設定          |
| nmz_set_sortmethod | ソートの種類を設定     |
| nmz_set_sortorder  | 昇順/降順の設定      |
| nmz_set_sortfield  | フィールドソート時の文字列 |
| nmz_set_maxhit     | 最大ヒット数の設定     |
| nmz_add_index      | インデックスの追加     |
| nmz_codeconv_query | 検索式のコード変換     |

### 2. 検索

|            |    |
|------------|----|
| nmz_search | 検索 |
|------------|----|

### 3. 結果

|                    |           |
|--------------------|-----------|
| nmz_get_field_data | フィールド値の取得 |
|--------------------|-----------|

### 4. 終了処理

|                   |          |
|-------------------|----------|
| nmz_free_hlist    | ヒットリスト解放 |
| nmz_free_internal | 内部メモリ解放  |

---

---

## 10. サンプルコード

```
#include "namazu/libnamazu.h"
#include "namazu/field.h"
#include "namazu/hlist.h"
#include "namazu/idxname.h"
#include "namazu/i18n.h"
#include "namazu/search.h"
#include "namazu/codeconv.h"

#include <stdio.h>
#include <stdlib.h>
#include <string.h>

int main()
{
    int i;
    char buff[BUFSIZE] = "";
    NmzResult hlist;

    putenv("LANGUAGE=");
    putenv("LC_ALL=");
    putenv("LC_MESSAGES=");
    putenv("LC_CTYPE=");
    putenv("LANG=");
    nmz_set_lang("ja_JP.eucJP");

    /* ソート順 */
    nmz_set_sortmethod(SORT_BY_DATE);

    /* インデックスの指定 */
    nmz_add_index("C:¥¥namazu¥¥var¥¥namazu¥¥index");
```

---

---

## 11. 定数

libnmz で使用する主要な定数をしめします。  
これらはヘッダーファイル内で定義されています。

### 1. BUFSIZE

BUFSIZE は、バッファサイズを示す定数です。

### 2. nmz\_stat

nmz\_stat は、ステータス定数です。

```
enum nmz_stat {
    FAILURE = -1,
    SUCCESS,
    ERR_FATAL,
    ERR_TOO_LONG_QUERY,
    ERR_INVALID_QUERY,
    ERR_TOO_MANY_TOKENS,
    ERR_TOO_MUCH_MATCH,
    ERR_TOO_MUCH_HIT,
    ERR_REGEX_SEARCH_FAILED,
    ERR_PHRASE_SEARCH_FAILED,
    ERR_FIELD_SEARCH_FAILED,
    ERR_CANNOT_OPEN_INDEX,
    ERR_NO_PERMISSION,
    ERR_CANNOT_OPEN_RESULT_FORMAT_FILE,
    ERR_INDEX_IS_LOCKED,
    ERR_OLD_INDEX_FORMAT
};
```

---

---

## 12. 構造体

libnmz で使用する主要な構造体をしめします。  
これらはヘッダーファイル内で定義されています。

ここに示す構造体は、libnmz を利用するアプリケーションでは読み出し専用です。値を書き変えてはいけません。

### 1. nmz\_data

検索結果 1 件(1 文書)単位の情報が格納されています。

```
struct nmz_data {  
    int score;  
    int docid; /* document ID */  
    int idxid; /* index ID */  
    int date; /* document's date */  
    int rank; /* ranking data for stable sorting */  
    char *field; /* field's contents for field-specified search */  
};
```

|       |  |
|-------|--|
| score | スコア  |
| docid | ドキュメント ID  |
| idxid | インデックス ID  |
| date  | 日付(タイムスタンプ)<br>1970 年 1 月 1 日 00:00:00 UTC からの経過時間 |
| rank  | ランク  |
| field | フィールドソート時に一時的に利用するもの<br>( <u>読み書き禁止</u> )          |

---

---

## 13. 主要関数一覧

libnmz の主要関数の一覧です。

### 1. libnamazu.h

```
extern void nmz_set_sortmethod ( enum nmz_sortmethod method );
extern void nmz_set_sortorder ( enum nmz_sortorder order );
extern void nmz_set_maxhit ( int max );
extern void nmz_free_internal( void );
```

### 2. field.h

```
extern void nmz_get_field_data ( int idxid, int docid, const char
*field, char *data );
```

### 3. hlist.h

```
extern void nmz_free_hlist ( NmzResult hlist );
extern void nmz_set_sortfield ( const char *field );
```

### 4. idxname.h

```
extern enum nmz_stat nmz_add_index( const char *idxname );
```

### 5. i18n.h

```
extern char *nmz_set_lang ( const char *lang );
```

---

---

---

---

## 14. 関数

|   |        |  |
|---|--------|--|
| <pre>extern void nmz_set_sortmethod ( enum nmz_sortmethod method );</pre> |        |  |
| 引数  | method | <p>SORT_BY_SCORE または SORT_BY_DATE または SORT_BY_FIELD を指定します。</p> <p>SORT_BY_SCORE ... スコア順<br/>SORT_BY_DATE ... 日付(タイムスタンプ)順<br/>SORT_BY_FIELD ... フィールド順</p>                         |
| 戻り値   |        | なし   |
| 説明  |        | <ul style="list-style-type: none"><li>● 検索結果のソート方法を指定します。</li><li>● ソート順は nmz_set_sortorder で指定します。</li><li>● SORT_BY_FIELD を指定した場合は nmz_set_sortfield でソートキーを指定する必要があります。</li></ul> |

---

---

## 15. 「Namazu の開発日記」メルマガについて

### 15.1. 「Namazuの開発日記」メルマガについて

「Namazu の開発日記」メールマガジンは、Namazu に関する様々な情報をお知らせするメールマガジンです。Namazu の TIPS や、次世代 Namazu に関する機能案の紹介、Namazu for Windows で利用する ActivePerl 用の PPM に関する情報等も紹介しています。

多くは「ナマズのブログ」<http://namazu.asablo.jp/blog/> の内容を再編集した記事が元になっており、「ナマズのブログ」のダイジェストの役割を果たしています。

メルマガにはアフィリエイト広告が含まれています。広告収入は Web サイトの管理運営費の一部に充てています。

### 15.2. メルマガ登録/解除

「Namazu の開発日記」メールマガジンは、下記 URL から登録、解除が可能です。(無料)

<http://mg7.com/cgi-bin/hp.cgi?id=16084>

---

---

## 16. 免責

「libnmz API」は有用とは思いますが、頒布にあたっては、市場性及び特定目的適合性についての暗黙の保証を含めて、いかなる保証も行なっていません。

## 17. 著作権

「libnmz API」は著作物です。本書の内容のすべてまたは一部を許可なく配布することを禁じます。

Copyright (C) 2008-2009 TADAMASA TERANISHI

Copyright (C) 2008-2009 AKANEIRO SOFTWARE. All rights reserved.

sample

---

---

## 18. 履歴

- 2008/07/14 初版
- 2008/07/15 第 2 版
  - コマンドラインでのコンパイルオプションの間違いを修正
- 2008/07/30 第 3 版
  - 説明文のコンパイルオプションの訂正漏れを修正
  - サンプルに LC\_CTYPE を追加
  - サンプルの解説を追加
  - 「libnmz API UNIX 版」について追記
  - 開発環境を追記
  - Win64 版 libnmz について追記
- 2009/05/23 第 4 版
  - Namazu for Windows 2.0.19 対応
  - Microsoft Visual C++ 2005 SP1 再頒布可能パッケージの説明を追加

## 索 引

### —

\_NmzResult ..... 18

### A

ASCENDING ..... 16

### B

BUFSIZE ..... 14

### D

DESCENDING ..... 16

### E

ERR\_CANNOT\_OPEN\_INDEX ..... 15

ERR\_CANNOT\_OPEN\_RESULT\_FORMAT\_FILE ..... 15

ERR\_FATAL ..... 15

ERR\_FIELD\_SEARCH\_FAILED ..... 15

ERR\_INDEX\_IS\_LOCKED ..... 15

ERR\_INVALID\_QUERY ..... 15

ERR\_NO\_PERMISSION ..... 15

ERR\_OLD\_INDEX\_FORMAT ..... 15

ERR\_PHRASE\_SEARCH\_FAILED ..... 15

ERR\_REGEX\_SEARCH\_FAILED ..... 15

ERR\_TOO\_LONG\_QUERY ..... 15

ERR\_TOO\_MANY\_TOKENS ..... 15

ERR\_TOO\_MUCH\_HIT ..... 15

ERR\_TOO\_MUCH\_MATCH ..... 15

### F

FAILURE ..... 15

### I

IA64 ..... 4

### N

NMZ.field.xxxx ..... 7

NMZ.t ..... 7

NMZ.w ..... 7

nmz\_add\_index ..... 28

nmz\_codeconv\_query ..... 31

nmz\_data ..... 17

nmz\_free\_hlist ..... 26

nmz\_free\_internal ..... 24

nmz\_get\_field\_data ..... 25

nmz\_search ..... 30

nmz\_set\_lang ..... 29

nmz\_set\_maxhit ..... 23

nmz\_set\_sortfield ..... 27

nmz\_set\_sortmethod ..... 21

nmz\_set\_sortorder ..... 22

nmz\_sortmethod ..... 15

nmz\_sortorder ..... 16

nmz\_stat ..... 14

NmzResult ..... 18

### S

SORT\_BY\_DATE ..... 15

SORT\_BY\_FIELD ..... 15