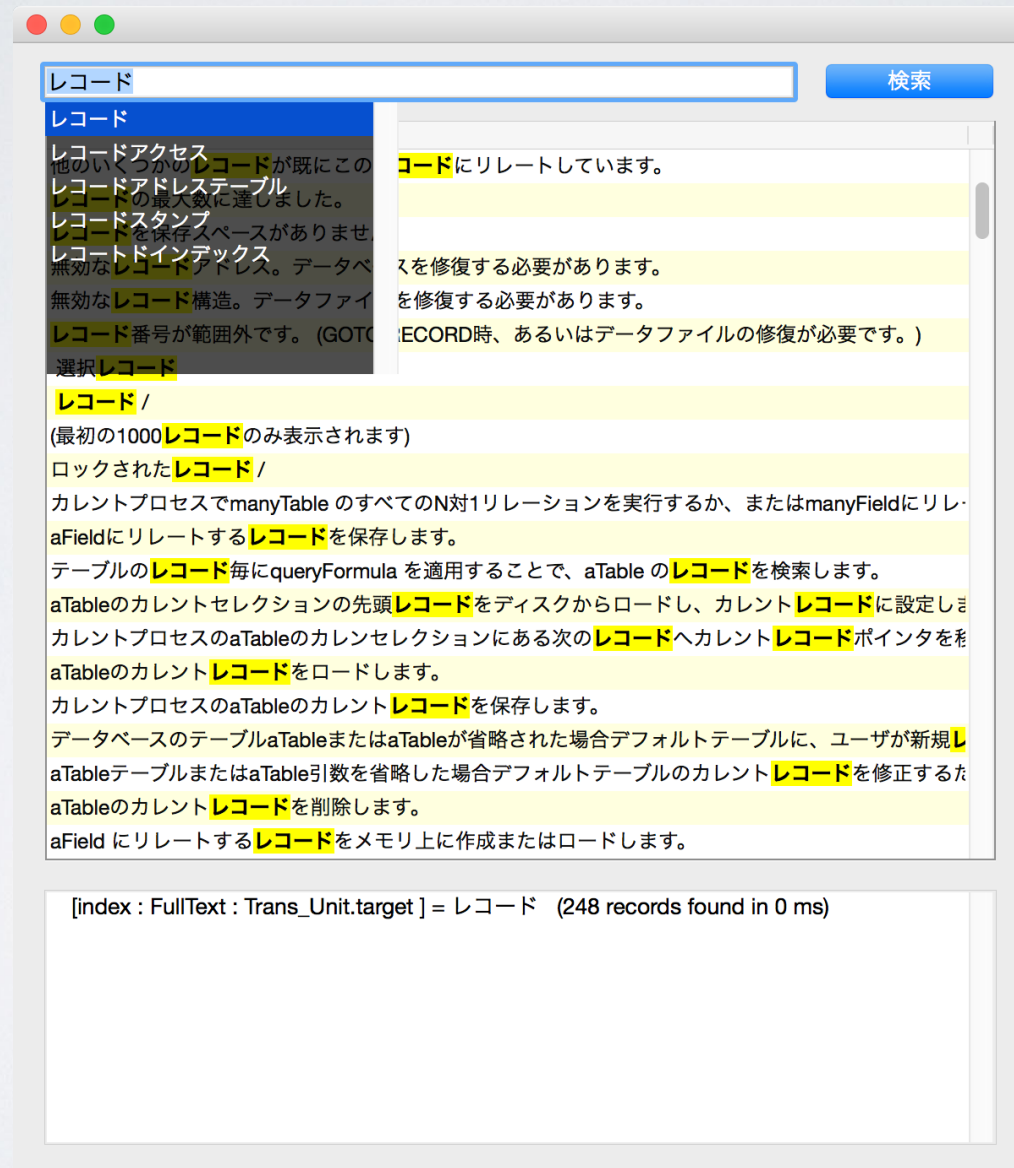


キーワード検索とスタイル付きテキストで簡単に実現!

日本語全文検索



デモンストレーション

検索エリアに入力したキーワードをハイライト表示したい



実装ガイド

技術	バージョン	備考
日本語キーワード検索	14	
スタイル付きテキスト《背景色》	14 [*]	
Find in sorted array	14R4	<i>optional</i>

* …v12 …v14


日本語キーワード検索

mecab

概要: 形態素解析エンジン (<http://ja.wikipedia.org/wiki/MeCab>)

インストール: 日本語版4D/Server/Unlimited Desktopはインストール済み

場所:  ./Resources/mecab/mecab.dll

 /Contents/Resources/mecab/mecab.bundle (パッケージ内)

バージョン: 0.996

辞書: jumandic 7.0-20130310/utf-8

重要: ビルド 180557以降を使用すること!

日本語キーワード検索

データベース設定

highlight-query - データベース設定

一般 インターフェース コンパイラー データベース 移動 バックアップ C/S Web SQL PHP セキュリティ 互換性

データストレージ メモリ

一般設定

- ☐ ストラクチャーファイルが読み込みのみのときに警告を表示しない
- ☐ 読み込みのみのデータファイルの使用を許可する

テンポラリフォルダーの場所

現在:

次回起動時: ☐ システム ☒ データファイルフォルダー ☐ ユーザー指定

...

テキスト比較

文字列の途中に含まれる@はワイルドカードとして扱わない

現在のデータ言語:

- ☐ 非文字・非数字のみをキーワード区切り文字とする
- ☒ 旧バージョン互換の文字列比較を使用する

キャンセル OK

データベース言語: 日本語

非文字区切りモード: 無効

日本語キーワード検索

ルール#1: 動詞・名詞・形容詞以外はキーワードに含まれない

pos-id	type	sub-type
2	指示詞	副詞形態指示詞
3		名詞形態指示詞
4		連体詞形態指示詞
5	助詞	格助詞
6		終助詞
7		接続助詞
8		副助詞
9	助動詞	
23	特殊	括弧始
24		括弧終
25		記号
26		句点
27		空白
28		読点
29	判定詞	
12	接頭辞	ナ形容詞接頭辞

日本語キーワード検索

ルール#2: 特定の接頭辞はキーワードに含まれる

pos-id	type	sub-type
14	接頭辞	名詞接頭辞
12		ナ形容詞接頭辞

+

pos-id	type	sub-type
31	名詞	サ変名詞
32		形式名詞
33		固有名詞
34		時相名詞
39		普通名詞
40		副詞的名詞

お隣の元アイドルが再デビュー

日本語キーワード検索

ルール#3: 特定の接尾辞はキーワードに含まれる

pos-id	type	sub-type
2	指示詞	副詞形態指示詞
3		名詞形態指示詞
4		連体詞形態指示詞
5	助詞	格助詞
6		終助詞
7		接続助詞
8		副助詞
9	助動詞	
23	特殊	括弧始
24		括弧終
25		記号
26		句点
27		空白
28		読点
29	判定詞	

pos-id	type	sub-type
17	接尾辞	動詞性接尾辞
22	動詞	

+

pos-id	type	sub-type
15	接尾辞	形容詞性述語接尾辞
16		形容詞性名詞接尾辞
17		動詞性接尾辞
19		名詞性特殊接尾辞

怒りっぽい人は友達ができにくいよ

今日限りで辞めさせていただきます。

それはできません

+

pos-id	type	sub-type
9	助動詞	

キーワードに分解

GET TEXT KEYWORDS

GET TEXT KEYWORDS("見た目どおりに優れたコンピュータです。";\$words)

mecab
見た目
どおり
優れた
コンピュータ

ICU
見
た
目
ど
お
り
に
優
れ
た
コンピュータ
で
す

非文字のみ
見た目どおりに優れたコンピュータです

キーワードリストを取得

DISTINCT VALUES

標準インデックス: セレクション内の**重複しない値**のリスト

キーワードインデックス: セレクション内の**キーワード**のリスト

キーワードリストを取得

キーワードインデックスの有無

```
C_BOOLEAN($keyword)
```

```
Begin SQL
```

```
  SELECT KEYWORD  
  FROM _USER_INDEXES  
  WHERE INDEX_ID =  
  (SELECT INDEX_ID  
   FROM _USER_IND_COLUMNS  
   WHERE TABLE_ID = :$tableId AND COLUMN_ID = :$fieldId)  
  INTO :$keyword  
End SQL
```



キーワードリストを取得

キーワードインデックスの有無

```
$Table:=Table($tableId)
$Field:=Field($tableId;$fieldId)
$temp:="$"+Generate UUID
ARRAY TEXT($words;0)
START TRANSACTION
CUT NAMED SELECTION($Table->;$temp)
CREATE RECORD($Table->)
Case of
: (Type($Field->)=Is text) | (Type($Field->)=Is alpha field)
  $Field->:="a,b"
  SAVE RECORD($Table->)
  DISTINCT VALUES($Field->;$words)
: (Type($Field->)=Is picture)
  $svg:=DOM Create XML Ref("svg")
  $rect:=DOM Create XML element($svg;"rect";"x";0;"y";0;"width";1;"height";1)
  C_PICTURE($image)
  SVG EXPORT TO PICTURE($svg;$image;Own XML data source)
  $Field->:=$image
  CONVERT PICTURE($Field->;".jpg")
  SET PICTURE METADATA($Field->;IPTC keywords;"a,b")
  SAVE RECORD($Table->)
  DISTINCT VALUES($Field->;$words)
End case
$keyword:=(Size of array($words)=2)
USE NAMED SELECTION($temp)
CANCEL TRANSACTION
```


配列内の高速サーチ

Find in sorted array

あらかじめ配列を並び替えておく

```
Sort Array($Keys->,>)
```

```
For ($i;1;Size of array($wordsInText))  
  If (Find in sorted array($Keys->($wordsInText{$i}+"@" );>,$first,$last))  
    For ($j;$first;$last)  
      Append to Array($keywords;$Keys->{$j})  
    End for  
  End if  
End for
```

1回のFind in sorted arrayで
要素の範囲を取得できる

配列内のサーチ

Find in array

Repeat~UntilでFind in arrayを繰り返し実行 (-1が返されるまで)

```
For ($i;1;Size of array($wordsInText))  
    $find:=1  
    Repeat  
        $find:=Find in array($Keys->,$wordsInText{$i}+"@";$find)  
        If ($find#-1)  
            APPEND TO ARRAY($keywords;$Keys->{$find})  
            $find:=$find+1  
        End if  
    Until ($find=-1)  
End for
```

検索フレーズの切り出し

Match regex

開始カッコまたは『』

終了カッコまたは『』以外

```
$i:=1
```

```
While (Match regex("( [\\p{Ps}\\"] ( [^\\p{Pe}\\"]+ ) [\\p{Pe}\\"] )"; \\  
$text;$i;$pos;$len))  
  APPEND TO ARRAY($words;Substring($text;$pos{2};$len{2}))  
  $text:=Delete string($text;$pos{1};$len{1})  
  $i:=$pos{1}  
End while
```

検索ワードの切り出し

Match regex

ワード（非文字・非数字）

```
$i:=1
```

```
While (Match regex("(\\w+)";$text;$i;$pos;$len))  
  APPEND TO ARRAY($words;Substring($text;$pos{1};$len{1}))  
  $i:=$pos{1}+$len{1}  
End while
```


検索フレーズのハイライト

セレクションリストボックス《式》

プロジェクトメソッド

プロパティリスト	
(Result)	
オブジェクト	
オブジェクト名	Result
データソース	
式	Styled ([Trans_Unit]target;"QueryWords")
データタイプ	文字列
選択リスト	<なし>
関連付け	リスト項目の値
自動挿入	<input type="checkbox"/>
座標とサイズ	
幅	545
最小幅	10
最大幅	32000
すべて	

フィールドの代わりに《式》を入力

《式》は表示に**必要な行数のみ**評価される (On Display Detail)

検索フレーズのハイライト

Position

```
$styleText:=$1
$word:=$2

If (Length($word)#0)
  $plainText:=ST Get plain text($styleText)
  C_LONGINT($start;$length)
  $start:=1
  Repeat
    $start:=Position($word;$plainText;$start;$length)
    If ($start#0)
      $end:=$start+$length
      ST SET ATTRIBUTES($styleText;$start;$end;\
        Attribute bold style;1;\
        Attribute text color;0;\
        Attribute background color;0x0FFFFFF00)
      $start:=$end
    End if
  Until ($start=0)
End if

$0:=$styleText
```

開始位置を指定 (v11)

Length (\$word) と同一
であるとは限らない

検索フレーズのハイライト

ST SET ATTRIBUTES

```
$styleText:=$1  
$word:=$2
```

標準テキストで位置を計算

```
If (Length($word)#0)  
  $plainText:=ST Get plain text($styleText)  
  C_LONGINT($start;$length)  
  $start:=1  
  Repeat  
    $start:=Position($word;$plainText;$start;$length)  
    If ($start#0)  
      $end:=$start+$length  
      ST SET ATTRIBUTES($styleText;$start;$end;\n\n      Attribute bold style;1;\n      Attribute text color;0;\n      Attribute background color;0x0FFFFFF00)  
      $start:=$end  
    End if  
  Until ($start=0)  
End if  
  
$0:=$styleText
```

標準テキストで位置を指定

半透明の背景色

Resourcesフォルダー

```
<?xml version="1.0" encoding="UTF-8" standalone="no" ?>
<svg
  xmlns="http://www.w3.org/2000/svg"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  viewBox="0 0 24 24">

  <g>
  <rect
    fill="#000000" fill-opacity="0.7" stroke="none"
    x="0" y="0" width="24" height="24"
  />
</g>

</svg>
```

相対パス (Resources)

リストボックス背景色 《なし》

ライブラリピクチャー

