

MCM-JP CG会合 オムロンソフトウェア 文字入力ソフトウェア（IME）ご紹介

2024年7月25日

オムロンソフトウェア株式会社 ITソリューション事業部 営業統括部

出野 健太郎

目次

- 文字入力システム概要
- 辞書データについて
- IMEをハブとしたメディア連携

文字入力システム概要

言語処理技術 「Wnn（うんぬ）」

Wnnシリーズ・・・低スペック～高スペック機器まで、様々な機器へ快適な言語入力を提供

30年以上に渡り培ってきた文字入力技術「Wnn」は、国内外数百社以上の様々な機器に搭載され、たくさんの人々に利用されています。
スマートフォンやタブレット、カーナビ、複合機、テレビなど、文字入力が必要とする様々な機器に搭載され、用途に合わせてUIデザインや辞書のカスタマイズも可能です。

また、日本語以外にも多数の言語に対応しており、国内外のメーカーから高い評価を受けています。



日本語IMEの特徴

日本語IMEは欧米等のIMEとは異なる、以下の特徴がある

①「変換」処理

- ・読み文字列(ひらがな)を別の表記文字列に変換する。
 (“変換”する言語は、日本語、中国語、韓国語のみ)
- ・なおかつ、漢字・ひらがな・カタカナ・英数字が混在する。

②文節区切り

- ・読み文字列から、文節を適切に区切って変換する必要がある。
 (連文節変換)
- ・どこで区切るかによって変換結果も変わる。

例) 今日は医者に行く。／今日歯医者に行く。

③同音異義語

- ・同音異義語が多くあるため、適切に変換する必要がある。

例) ×健康施工、○健康志向

④語尾の活用

- ・語尾が活用変化するので、適切な予測・変換候補を出す必要がある。

例) 走(はし)→走る、走れば、走った、走って、...

Android向け入力ソフト iWnn IME for Android

Android向けの文字入力ソフトウェア。ソフトウェアKBDのUIと変換処理エンジン・辞書をセットで提供

機能例①

状況適応予測



季節・時間帯・時制などのシチュエーションに合わせて最適な予測候補を判断して表示することが可能

機能例②

ワイルドカード予測

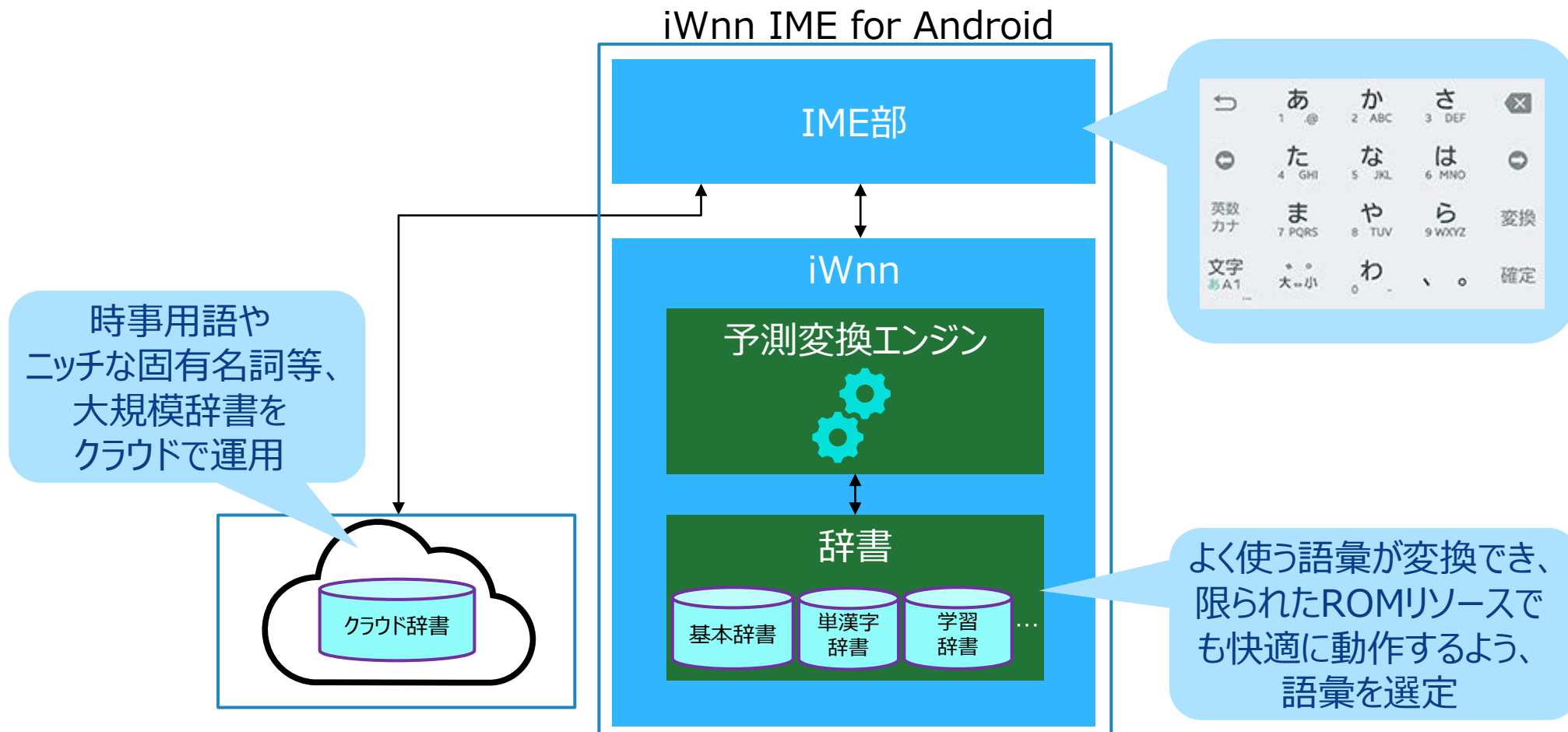


冒頭の文字のみ入力し、残りの文字長から予測変換の候補を絞り込み



iWnn IME for Androidの概略構造

変換エンジン+辞書のiWnnとUI部のIME部のセットで構成される



組込辞書とクラウド辞書の変換例

使用頻度が高い語彙：組込辞書
使用頻度が低い語彙：クラウド辞書

G 換気

換気	歓喜	乾期	乾季	寒気	勘気	喚起
神吉	カンキ	かんき	かんき	かんき	かんき	200,022
200022	200022	200022	200022	200022	200022	kanki
kanki	KANKI	KANKI	KANKI	KANKI	KANKI	Kanki
Kanki						

外部変換

組込辞書からの
変換語彙

G 換気

換気口

換気	歓喜	乾期	乾季	寒気	勘気	喚起
神吉	カンキ	かんき	かんき	かんき	かんき	200,022
200022	200022	200022	200022	200022	200022	kanki
kanki	KANKI	KANKI	KANKI	KANKI	KANKI	Kanki
Kanki						

桓騎	間期	寛喜	刊記	環起	神木	甘輝	韓琦
韓軌	咸熙						

クラウド辞書からの
変換語彙



出典元：集英社 キングダム コミックス28巻P.122
引用サイト：https://youngjump.jp/kingdom/vote_meigen/index2.html

辞書データについて

辞書開発プロセス

専用の辞書開発チームにて年間通して辞書データを作成

① 語彙収集

Linux
寄生獣
科学忍者隊ガッチャマンF
佐倉綾音
スバル・360
スバル・R1
篠原信一
湯江健幸
ガンダムビルドファイター・・・etc

② 語彙選別

読み・表記チェック

→品質チェック

不快語チェック

→品質チェック

語彙の選別

→品質チェック

③ 語彙収録

辞書コンテンツ

新語

住所
地名

ニュース・
時事辞書

ダイエット・
健康辞書

ファイナンス
辞書

食品辞書

旅行・ホテル
辞書

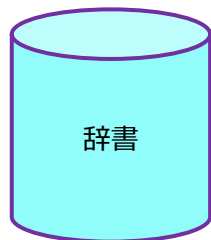
子ども辞書
・・・etc

定期的に語彙を収集 ⇒ 各語彙の品質をチェック
各語彙を選別 ⇒ 語彙をジャンルごとに収録

辞書データ構造／頻度値と候補表示

個々の収録語は、以下の属性情報を持つ

- ・読み文字列
- ・表記文字列
- ・品詞
- ・利用頻度値
- ・(固有名詞の場合)カテゴリ

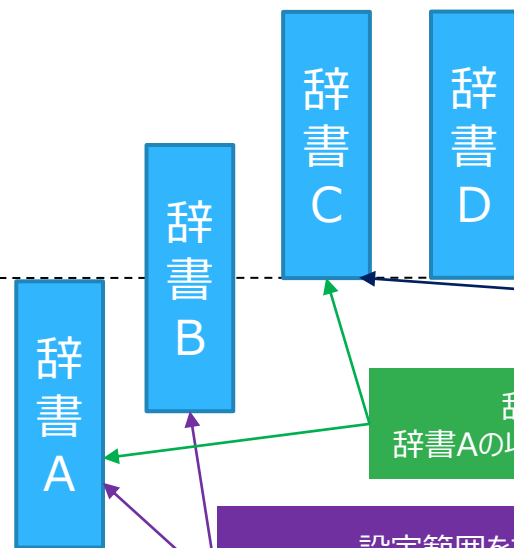


上記の属性情報を効率的に検索できるよう、独自の構造でバイナリ化

収録語単位の頻度値に加え、辞書自体の頻度値で予測変換候補の表示順を制御

辞書自体に頻度値を設定し、各辞書の優先順位関係を指定可能

高
辞書頻度値
低



設定範囲が同じであれば、辞書収録語の頻度値情報のみで優先度を判断

辞書Cの収録語は、辞書Aの収録語よりも必ず優先される

設定範囲を重ねることで、辞書B収録語のうち、優先度の高い物だけを辞書A収録語より優先させる

IMEをハブとしたメディア連携

ユーザ接点観点でのIMEの特長

「スマホユーザ」と最大の接点を持つアプリ



「スマホ」から発信されるコミュニケーションの玄関口となるのがIME

放送連携IMEの開発事例

IMEをユーザ接点として活用する、放送データと連携したIMEをNHK様の実証テーマとして開発



いつもの入力操作から、ユーザ意図を起点とした放送連携サービス利用

NHK試作デモ画面

