

IME edition

© まろん。 / @rin_montblank

自己紹介



まろん。

ネタ駆動開発をしています。

Portfolio: montblank.fun

GitHub: @otnc

Twitter: @rin_montblank

キッカケ

最近便利なIMEを見つけた。

IMEとは

Input Method Editorの略。

読みを入力して実際の文字に変換するためのソフト。

e.g. ohayou → おはよう

Meltype

日本語と英語を見分けて、入力中に変換してくれる。

打った文字	変換結果
kyouhagoogledekensaku	今日はgoogleで検索
ashitanomeetingwotsuika	明日のmeetingを追加

半角/全角キーで切り替えずに、日本語と英語を混ぜて打てる。
macの標準の変換がイメージに近いかも。

yksr-melt/Meltype

Meltypeの仕組み

言語の判定と、かなを漢字にする処理を組み合わせている。

- ローマ字や文章の並びや辞書などを使って、一定のルールで日本語か英語かを判定する
- 日本語はMicrosoft IME / Mozcなどで漢字に変換する
- 入力中の表示や候補、確定する文字を管理する

判定や変換はPCの中で処理している。

おもしろい

何かできないか

変換結果を少しいじって面白くできそう。

昔作ったライブラリを組み込んでみる

いちいち変換器にコピペせずに打っているそばから変な文字になるIME。面白そう。

高校生の頃に作ったライブラリの紹介

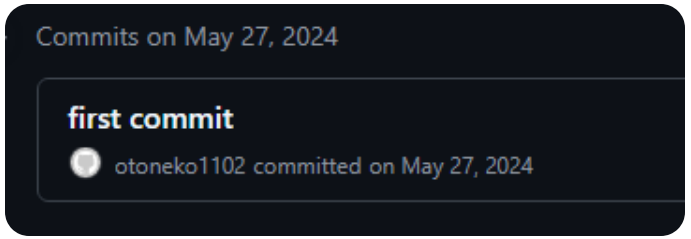
たまたまギャル文字変換用のライブラリが手元にあった。

当時、 cjp という怪しい翻訳ライブラリをするライブラリに影響を受けて作った。

```
import { galjp } from 'galjp';

galjp('こんにちは！'); // 変換結果: ⊇ωﾆ乞レ£.′
```

最初のコミットが2年半前



otnc/galjp

galjpはどう変換している？

文字の置き換えと、漢字の部品への分解を組み合わせる。

対象	方法	例
かなや英数字、記号	変換表で別の文字に置き換える	は → £
漢字	部品に分けて、それぞれを変換する	信 → 亻言
旧字体にできる漢字	旧字体に置き換える	学 → 學

漢字の構造にはKanjiVGを使う

漢字の筆画や部品の構造をSVGに記録したデータ。

- galjpでは部品の構造から漢字を分割するためのデータを生成する
- 分けた部品の書き方はgaljp側の変換表で決める

信での例
↓
KanjiVGの構造データ:亻 と言に分ける
↓
galjpの変換表:亻 をｲ(半角カナ)に置き換えて、ｲ言にする

KanjiVG

galjp + Meltype = Galtype

Meltypeの変換結果に、galjpを適用してみた。

oto-lab/Galtype

ギャル文字変換を呼びだす試み

Meltypeは.NET(C#)なので、galjpをそのまま組み込めなかった。
→Node.jsを別のプロセスで動かし、標準入出力でやり取りする。

言語	方法
Galtype / C#	変換したい文字列と設定をJSONで送る
Node.js / JavaScript	galjpを呼び、結果をJSONで返す

デモ

余談: ギャル文字変換後、文字数が変わる問題

元の文字列情報を内部的に保持して削除等の動作時に使用する。

活用例: **校** の変換後は **木交** の2文字になるが、BackSpaceでの削除時はまとめて消える

ご清聴ありがとうございました

oto-lab/Galtype

※windowsのみ対応。他の環境は自力でビルドしてください。

参考: Meltype