

## データベース名称： 音声翻訳実証実験固有名詞対訳辞書

### データの概要

平成 21 年度「地域の観光に貢献する自動音声翻訳技術の実証実験<sup>1</sup>」で採択された 5 つのプロジェクト、「来道外国人観光客の満足度向上を支援する多言語自動音声翻訳技術の実証実験」、「山梨県における観光振興に貢献する多言語音声翻訳技術の実証実験」、「中部広域観光圏・多言語観光促進プロジェクト」、「関西圏・多言語観光プロジェクト」、「九州域多言語自動音声翻訳を活用した観光・ホスピタリティプロジェクト」において、日・英・中・韓国語の固有名詞辞書を収集した。

このデータを NICT で整備した辞書である。

### 関連発表

安田圭志 他、電子情報通信学会論文誌 D Vol.J95-D, No.1, pp.19-29, 「音声翻訳システム実利用データを用いたシステム改善手法」

### データ内容

日本語、英語、中国語、韓国語の 4 ヶ国語固有名詞辞書

辞書は EXCEL ファイルで見出し語は 30,818 語、項目は「地域」、「日本語」、「読み」、「カテゴリ」、「英語」、「中国語」、「韓国語」からなる。

項目のうち「地域」にはその単語を収集した地域を示す 01\_Hokkaido、02\_Kanto、03\_Chuubu、04\_Kansai、05\_Kyuushu のいずれかが記載されている。

また、「カテゴリ」はその単語のカテゴリとして、以下の表のいずれかが記載されている。

地名	レストラン	姓名	土産
交通路	店名	人格	アトラクション
地形名	日姓	組織	日時
国名	日名	団体名	文字列
宿泊施設	外姓	会社名	引用
建物	外名	飲食物	その他

以下に辞書の例を挙げる。

地域	日本語	読み	カテゴリ	英語	中国語	韓国語
01_Hokkaido	「北の国から」 資料館	キタノクニ カラシリョ ウカン	建物	Kita no Kunikara Museum	“北国之恋”資料 館	《북쪽의 나라에서》 자료관
01_Hokkaido	あいの里公園	アイノサト コウエン	交通路	Ainosato koen	爱里公园	아이노사토 고엔
01_Hokkaido	あいの里公園 駅	アイノサト コウエンエ キ	交通路	Ainosato-koen Station	爱里公园站	아이노사토 고엔 역
...	...	...	...	...	...	...

<sup>1</sup> [http://www.soumu.go.jp/menu\\_news/s-news/02tsushin03\\_000002.html](http://www.soumu.go.jp/menu_news/s-news/02tsushin03_000002.html)