

《Press Release》

2024年6月21日
株式会社時事通信社

◎汎用版「綴」(カタカナ表記推定 AI システム)、利用スタート

株式会社時事通信社(代表取締役社長:境克彦、以下「時事通信」)は、外国人名のアルファベット表記のカタカナ訳を一瞬で推定・翻訳する AI システム、汎用版「綴」の運用を始めました。

「綴」は、短期間に数万人規模の外国人名のカタカナ翻訳が必要な五輪パラリンピックなどの大規模イベントを念頭に、名古屋大学大学院の佐藤理史研究室と時事通信との共同研究の成果として構築されました。地元開催の東京 2020(コロナ禍のため 2021 年開催)では参加全選手名の日本語化を実現し、時事通信の新聞、テレビ・ラジオ、インターネット向け報道サービスを下支えしました。

共同研究は 2015 年に始まり、佐藤研究室が AI システムの設計と開発、時事通信はデータ整備の役割を担いました。東京 2020 では、206 カ国・地域から参加した 13,032 人の選手すべてについて「綴」を使った翻訳を実施、それぞれの固有言語の発音に沿ったカタカナ表記を作成しました。「綴」で翻訳されたカタカナ表記は、人手による確認、修正などを経て実際に利用されますが、最終的な採用率は 90 パーセント程度と高い数字を示しました。すべてを人手で行う場合に必要な人的リソースの確保、人件費コストなどから考えると、極めて利用価値の高いツールとなりました。

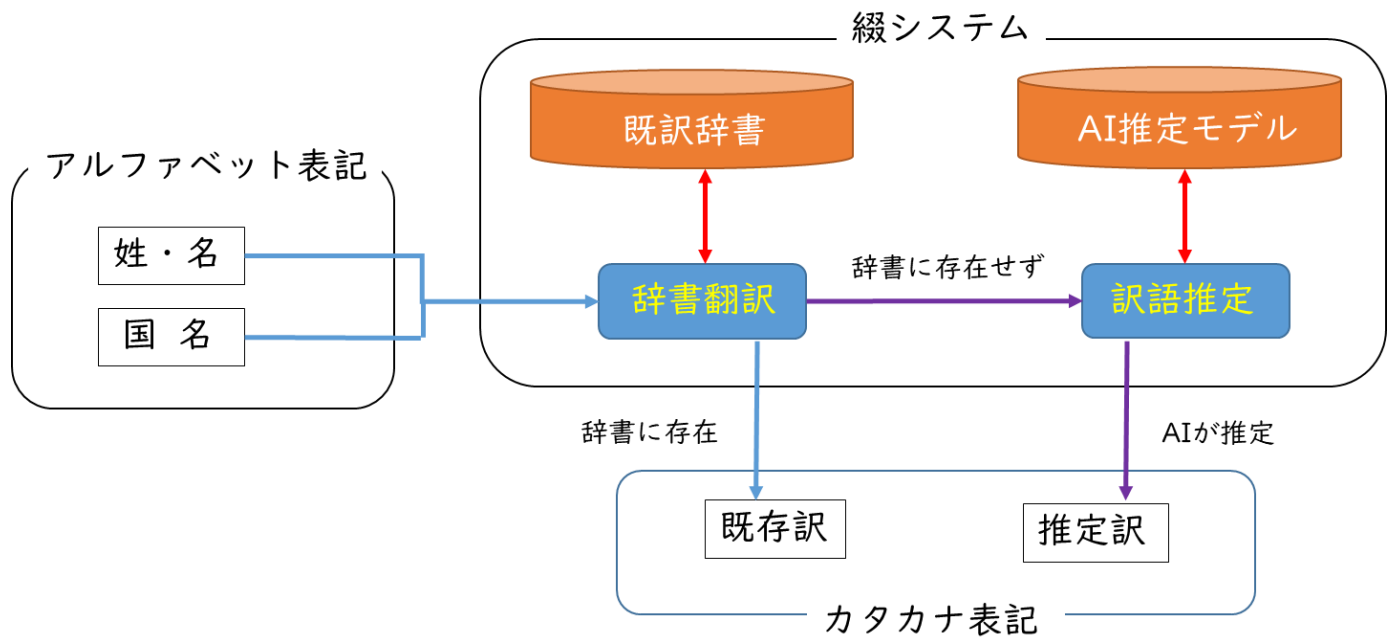
今回運用を開始した汎用版「綴」では、データを追加して精度アップを図ったほか、新データを容易にシステム反映できる仕組み、excel ファイルなどから簡単に一括翻訳を実行できる機能などを加え、汎用(はんよう)性を向上させました。

今夏のパリ五輪パラリンピックで活用しますが、スポーツ以外の報道現場での利用および業務効率化、外部への展開にもつなげたいと考えています。

システム構成や東京 2020 での活用の実際については、言語処理学会の機関誌「自然言語処理」vol.30 No.2(2023 年 6 月)に掲載された佐藤教授の論文「2020 東京オリンピック参加者名簿の翻訳」に詳述されています。下記 URL をご覧ください。

<https://doi.org/10.5715/jnlp.30.748>

「綴」システムの概念図



「綴」のインターフェース

英字1: michael × 英字2: Schumacher × 国コード (IOCコード): GER (ドイツ) × クリア 翻訳

* 英字1・2のいずれかは必須です 最大選択可能数: 制限なし

推定結果	英字 1		英字 2	
入力	michael		Schumacher	
国コード (IOCコード)	GER	all	GER	all
国名	ドイツ	unknown	ドイツ	unknown
一致 1	ミヒヤエル GER,AUT,SUI,ITA			
推定 1	ミヒヤエル GER,AUT,SUI,ITA	マイケル USA,GBR,CAN,IRL,NZL, AUS,ANT,BAH,JAM,MLT, PNG,BIZ,RSA,KEN,NGR, ZAM,BAR,CMR,CRC,CUB, GUM,MHL,PAK,SSD,SUR	シューマッハー	シューマッハー

【問い合わせ先】

☎ : 03-3524-6014 (スポーツ事業推進室)

☎ : 03-3524-6840 (システム開発局)

✉ : mt-project@jiji.co.jp