

メタデータ、多言語テキスト関連度判定サーバを提供開始 ～辞書不要！で素早く現地語の連想検索サーバを構築可能

メタデータ株式会社（所在地：東京都文京区 代表取締役社長：野村直之）は、辞書作成が原則不要のため、新規言語対応の高精度な関連・類似検索機能を1日で開発可能な**多言語テキスト関連度判定サーバ**を開発しました。本日から提供開始いたします。この**多言語テキスト関連度判定サーバ**の活用により、海外クチコミを分析する必要のあるグローバル日本企業はもちろん、自社のソーシャル・プラットフォームや、応用サービスを海外展開する企業にとって、現地語による注目記事に関連した情報や、共通の話題を含むクチコミを素早く一気に収集・分類することが可能となり、クチコミの分析が飛躍的に効率化、高品質化します。既に11カ国語に対応済みで、8カ国語の混在したデータベースで検証済みです。注目記事に関連した表現を多く含む順に、当該言語の記事を速やかに上位にランキングすることができます。今後は、本技術をアライドアーキテクト株式会社のプラットフォームに適用し、海外対応のソーシャル×ビッグデータ活用の市場開拓を支援していく予定です。

■背景

ビッグデータがビジネスを制する時代といわれ、大量の顧客の声、ソーシャルメディアの書き込みを前にしながら、現場ではそれらを読みこなさきれず、活用できていない、といわれます。特にそれが英語以外の外国語で記述されている場合、現地語を母国語とする人々に、分析ノウハウや日本流のレポーティングを1から教え込む莫大な教育コストがかかり、なかなかビジネスを海外展開できない、という悩みがありました。

■内容

本日から提供する**多言語テキスト関連度判定サーバ**の活用により、収集した外国語ビッグデータ（テキスト）をデータベースに取り込むだけで、高精度で、関連性・類似性を自動判定できるようになります。

The screenshot shows the 'Similarity Search' interface. On the left, there is a search form with 'Question' set to 'tweet' and 'lang' set to 'vi' (Vietnamese). The search results are displayed in a table titled 'Listing text_entries'. The table has columns for 'Question', 'lang', 'Text', and 'Relevance'. The first result is a tweet in Vietnamese with a high relevance score. The interface also includes a search bar for the question and a filter for the language.

Question	lang	Text	Relevance
tweet	vi	Tôi đã thích video http://youtu.be/_KXyout.0N7a Từ đồng theo dõi Facebook tự đồng theo dõi facebook Cập nhật mới	High
tweet	vi	Thông báo về lễ tưởng niệm 40 năm ngày Trung Quốc xâm chiếm Hoàng Sa tại Hà Nội vào Chủ Nhật 19/1/2014. http://fb.me/i1hnpRoL	Medium
tweet	vi	CUỐN TIỂU SỬ MỠI MỪNG SÁNH NHẬT	Low

例えば、上図は、ブラウザ操作で読み込んだ8カ国語のツイート（ツイッターへの書き込み）を自動的に解析しつつデータベースに取り込んだものに、最上位にランクされたベトナム語のツイートとの関連性・類似性を自動判定させた画面です。関連性を比較する基準となるベトナム語のツイート自身が最上位にランクされ、以下、関連した表現を多く含み、且つそれがデータベースの中でユニークな度合いに応じてランキングされていきます。

このデータベースには、アラビア語、ヒンズー語、インドネシア語、韓国語、ロシア語、タイ語、フィリピン語（タガログ語）、ベトナム語の8カ国語を混在させています。特定言語の基準記事（文書）であれば、その言語の記事が上位にならび、言語混在した基準記事については、各言語の記述の関連の強さ、ユニークさ総合点の順に、言語混在した上位ランキングとなります。これらの母国語話者人口の多い言語への対応を済ませ、また、既に弊社製品 **アンケート分析Pro** で対応した日本語、英語、中国語にも対応しています。さらに新規言語を追加するのに、原則1日以内で対応可能です。

本 **多言語テキスト関連度判定サーバ** は、ソーシャル関連サービスの提供企業、利用企業とも、海外展開で言語の壁に阻まれていた企業にとって福音となります。今後、本技術をアライドアーキテクト社のプラットフォームに適用し、海外対応のソーシャル×ビッグデータ活用の市場開拓を支援していく予定です。

メタデータ株式会社について

メタデータ株式会社は、2005年12月にメタデータ活用技術、セマンティック技術の応用ソフトウェア会社として設立。適合型自然言語解析エンジン『**メタパーザ**』を駆使した意味解析技術により、5W1H抽出によるコンテンツ連携や個人情報自動匿名化や、評判分析（ネガポジ・感情解析）、風評発見・監視ソリューション、Web対話ロボットなど、SaaS、クラウド向けアプリケーションを開発してまいりました。これにより、顧客企業様のソーシャル活用推進、CRM局面での情報セーフティネットの構築を行っています。また、資料送信まで自動で行える対話ロボット“Web受付嬢”により、Webサイトを親しみやすく、ITを使いやすくすることを支援しています。

本件に関するお問い合わせ先

メタデータ株式会社

担当：野村 nomuran@metadata.co.jp

TEL:03-3813-5447

FAX:03-3813-5447 <http://www.metadata.co.jp>

