

Language Grid Toolbox: 多言語コミュニティ支援のための オープンソースソフトウェア

田仲 正弘[†] 村上 陽平[†] 稲葉 利江子[‡] 林 冬恵[†] 石田 亨[†]

情報通信研究機構 言語グリッドプロジェクト[†]

京都大学情報学研究科 社会情報学専攻[‡]

1. はじめに

国際交流支援団体を始めとする、多言語コミュニケーションを行うコミュニティでは、しばしば機械翻訳などのツールが利用される。しかし、それぞれのコミュニティでの活動やコミュニケーションの形態の多様さから、既製のツールでは各コミュニティの多様な要件を満たすことは困難である。

そこで、Language Grid Toolbox (以下 Toolbox) と呼ぶ多言語コミュニケーションツール群を開発した。翻訳支援機能・多言語掲示板・言語資源作成機能・多言語 Web 作成機能など、多言語コミュニティ支援のための基本的な機能を提供する Web アプリケーションである。これらの機能は、言語グリッド[1]上で提供される機械翻訳や辞書、形態素解析などの言語サービスを用いて実現されている。

Toolbox では、その全ソースコードをオープンソースで公開した上で、各機能の API を提供している。これにより、[2]など従来の多言語コミュニケーションツールと比較して、それぞれの多言語コミュニティでの要求にあわせたカスタマイズや、既存機能の組み合わせによる新たな機能がより容易になっている。

以降では、Toolbox が提供する多言語コミュニティ支援のための基本機能と、コミュニティごとのカスタマイズや機能開発を実現するためのアーキテクチャについて述べる。

2. 基本機能

Toolbox が多言語コミュニティ支援のために備えている機能を以下に挙げる。これらの機能は、それ自体が多言語コミュニケーションの支援機能であるが、各コミュニティでの要求に応じて機能の変更や拡張を行うことを想定したもので

あることから、基本機能と呼ぶ。

・ 多言語掲示板機能

参加者が自らの母言語で投稿したメッセージを、機械翻訳サービスを用いて多言語に翻訳することで、他のユーザは自らの母言語でメッセージを閲覧することができる。また、機械翻訳の結果を、他のユーザが修正することができる。

・ 翻訳支援機能

入力された文章の翻訳・折り返し翻訳を実行する。入力文・翻訳結果・折り返し翻訳結果で、対応する文をハイライト表示する機能を備える。

・ 多言語 Web 作成機能

Web ページの多言語化を支援するため、指定された Web ページを、HTML タグ構造を保ったまま機械翻訳を行う。複数のユーザが協力しながら多言語の Web を作成することを可能とするため、ユーザは典型的なタグ構造と翻訳結果をテンプレートとして登録し、機械翻訳と組み合わせる機能を用いる機能を備える。

・ 言語資源作成機能

多言語の対訳辞書や用例対訳集の作成と検索を行える。作成された対訳辞書は、多言語 BBS・テキスト翻訳・多言語 Web 作成において機械翻訳と連携することが可能である。機械翻訳と辞書との連携により、入力文に辞書中の語が含まれていた場合には、辞書に定義された訳語を用いて翻訳される。

また、Toolbox 上で実行される翻訳の設定のため、言語グリッドに登録された数種の機械翻訳サービスや辞書サービス、言語資源作成機能で作成された辞書を GUI 上で組み合わせる機能を備えている。

3. アーキテクチャ

Toolbox は、オープンソースの CMS である XOOPS¹ を拡張して言語グリッド上のサービスを利用可能とした Toolbox フレームワークと、その上に登録された各種の多言語コミュニケーション支援モジュールからなる。

Language Grid Toolbox: Open Source Software for Supporting Multilingual Communities

[†] Masahiro Tanaka, Yohei Murakami and Lin Donghui, Language Grid Project, National Institute of Information and Communications Technology (NICT)

[‡] Rieko Inaba, Toru Ishida, Department of Social Informatics, Kyoto University

¹ <http://xoopscube.jp/>

表 1 基本機能及びフレームワークの API

機能	API 例
多言語掲示板	メッセージの投稿・返信・一覧取得・編集・削除, メッセージカテゴリ階層の作成・一覧取得・編集・削除, メッセージ検索, 翻訳修正, 翻訳修正履歴の管理
翻訳支援機能	キャッシュ付き翻訳, キャッシュ付き折り返し翻訳
言語資源作成	言語資源の作成・一覧取得・編集・削除, レコード追加・一覧取得・編集・削除・検索, Web サービスとしての配備
Toolbox フレームワーク	機能ごとの翻訳設定の編集, 言語サービス呼び出し, 言語サービス一覧の取得

これらは全て PHP で開発されている。前章に挙げた各基本機能は、この多言語コミュニケーション機能モジュールとして実装されている。

Toolbox のアーキテクチャを図 1 に示す。

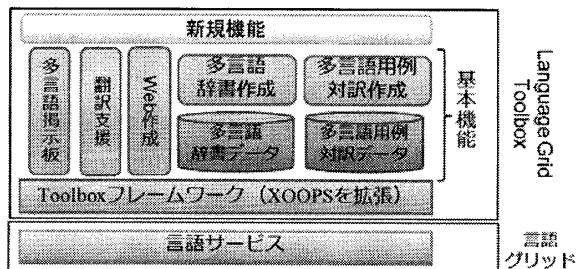


図 1 アーキテクチャ

各モジュールと Toolbox フレームワークは、新規機能の開発や、既存機能の拡張のための API を提供している。表 1 に、基本機能及び Toolbox フレームワークが提供する API の例を示す。

これらの API はいずれも PHP の関数として実装されており、他の多言語コミュニケーション支援モジュールから利用することができる。基本機能の API を用いることで、辞書検索インターフェースを加えた翻訳支援機能などの、複数の機能を組み合わせた新規機能を開発することができる。また、Toolbox フレームワークの API は、言語グリッドで Web サービスとして提供される言語サービスを、通常の PHP の関数呼び出しとして実行することを可能にする。このため、Web サービスの知識を持たない開発者でも、言語サービスを容易に利用できる。

新規機能は、その他の機能と独立した多言語コミュニケーション支援モジュールとして開発でき、Toolbox フレームワークとは別に配布することができる。実際にこの仕組みを利用して、京都市や京都大学によって、新たな基本機能や、基本機能の API を用いた機能を提供するモジュールが開発されている[3]。Toolbox を使用する

各コミュニティでは、様々な組織が開発したモジュールを、目的に応じて選択的に導入することが可能である。

4. おわりに

本稿では、多言語コミュニティにおけるコミュニケーション支援のためのツール群である Language Grid Toolbox について述べた。Toolbox はオープンソースで開発され、各機能の API が提供されているため、多言語コミュニティにおける様々な要求を満たすためのカスタマイズや新規機能の開発を容易に行うことができる。

今後、多言語コミュニケーションでの現場や、拡張機能の開発からのフィードバックを通じて機能の拡張や API の整備を行い、様々な組織が参加してモジュールの開発を行えるオープンソースコミュニティの構築を進めていく。

謝辞

本研究は、総務省戦略的情報通信研究開発推進制度 (SCOPE) の補助を受けた。

参考文献

- [1] 境 智史, 後藤雅樹, 村上陽平, 石田 亨. 言語グリッドブレイグラウンド: 軽量の構成部品を用いた異文化コラボレーション環境. ヒューマンインタフェース学会論文誌, Vol. 11, No. 1. pp. 115-123, 2009.
- [2] Toru Ishida. Language Grid: An infrastructure for intercultural collaboration. *IEEE/IPSJ Symposium on Applications and the Internet (SAINT-06)*, pp. 96-100, keynote address, 2006.
- [3] 稲葉利江子, 古白川亮太, 石松昌展, 嶋田雅彦, 後藤正浩, 甲斐信行, 村上陽平, 田仲正弘, 林冬恵, 辻正道, 植田浩司, 石田亨. 言語グリッドによる都市の多言語交流基盤の形成 - 京都 on 言語グリッド -. 情報処理学会第 72 回全国大会, 2010.