多言語自動翻訳掲示板の利活用に関する実践研究

1J20F037 奥村 飛悠　　　　　　　　　　　　　　　　　　　　　　　　　　　　（指導教員）　西村 昭治

1. はじめに

インターネットやソーシャルメディアの台頭に伴い、掲示板はユーザー間の活発な議論の場として重要な役割を果たしている。しかし、掲示板におけるコミュニケーションは主に同一言語間で行われている。そのため、異なる言語を使用するユーザーは翻訳ツールの利用や専用のスレッド、言語コミュニティの利用が一般的である。この背景には、翻訳技術の品質が不十分であったことが挙げられる。翻訳品質はユーザーのコミュニケーションに大きな影響を持つため、十分な翻訳品質を提供できない技術ではコミュニケーションが制限されていたと考えられる。

一方で、翻訳サービスの精度は日々向上している。この進展の背景には、ニューラル機械翻訳の進化がある。多くの翻訳サービスは、開発者向けに便利なAPIやライブラリを提供している。過去には、日本語と韓国語の翻訳機能を持つ掲示板サービス「enjoy Korea」や日本語とウイグル語間の翻訳掲示板システム、ユーザインターフェースのカスタマイズ性に着目した複数言語翻訳システム「CustomChat」などが開発されたが、これらのシステムを利用するユーザーのデータ収集やその分析は行われていなかった。

これらの背景を踏まえ、本研究では、掲示板のグローバル化を目指し、近年の高精度な翻訳サービスを利用した多言語自動翻訳掲示板の開発とその利活用に関する実践的な研究を行った。

1. 掲示板について

本研究では、多言語自動翻訳掲示板「The Channel」というWebアプリケーションを開発した。この掲示板では、世界中のユーザーが自分の言語で投稿できる。投稿は閲覧するユーザーが選択した言語に翻訳されて表示されるため、好きな言語でコンテンツを読むことができる。主な使用技術は、WebサーバーとしてのNginx、フロントエンドにはTypeScript、React、Next.js、MaterialUI、Sass、バックエンドにはPython、FastAPI、データベースにはMySQLを使用した。コンテナ化技術にはDockerを採用し、コード管理にはGitとGitHub、翻訳技術にはGoogle Trans APIを使用した。サーバーはさくらのVPSを利用した。

1. 利活用についての分析

本研究で開発した多言語自動翻訳掲示板の利活用を分析するため、多様な言語を使用する人々を集めることが必要であった。特に、日本語以外を使用するユーザーの獲得は、多言語間コミュニケーションの分析において重要な要素である。しかし、ユーザー数の分析から、日本語以外のユーザー獲得には成功していないことが明らかになった。

|  |  |
| --- | --- |
| 国名 | ユーザー数 |
| 日本 | 206 |
| 中国 | 10 |
| アメリカ | 7 |
| インドネシア | 3 |
| 台湾 | 1 |

表1.主な国別ユーザー数

この原因として、ユーザーの多くが著者や著者が所属する研究室からの宣伝を受けて集まり、Twitterなどのソーシャルメディアからの流入が想定以下であったこと、開発に多くの時間がかかり、宣伝活動に十分なリソースが割り当てられなかったこと、そして掲示板上での会話よりも単発的な投稿が多く持続的な会話や議論が少なかったため、ユーザーの定着が難しい状況が生まれていたことが挙げられる。

本研究で明らかになった課題に対して、私たちは以下のような対応策を提案する。アニメや漫画など、国際的に認知されている日本文化に関連するトピックを中心に据えることで、異なる言語圏のユーザーを引き寄せる可能性がある。さらに、掲示板における対話を促進する機能を導入することにより、ユーザー間の交流が活性化し、単発的な投稿を超えた意見交換や議論が期待される。これらの対策によって、日本語以外のユーザーを獲得し、持続的な会話や議論によるユーザーの定着を図ることができるのではないだろうか。