

1. 担当 PM

稲見 昌彦（東京大学 先端科学技術研究センター 教授）

2. クリエータ氏名

大原 嶺（東京大学 大学院学際情報学府 学際情報学専攻）

青田 香菜子（東京大学 大学院学際情報学府 学際情報学専攻）

藤井 樹里（東京大学 大学院学際情報学府 学際情報学専攻）

3. 委託金支払額

2,736,000 円

4. テーマ名

文脈に基づいた emoji 推薦とその選択インタフェースの開発

5. 関連 Web サイト

- Web サイト：<https://emolingual.com/>
- 成果物の PV：<https://www.youtube.com/watch?v=2PR8b69nZE4>
- 本プロジェクトを取り上げて戴いたポッドキャスト yancanfm vol.21：
<https://anchor.fm/yancanfm/episodes/21-erg64j>

6. テーマ概要

本プロジェクトでは、文脈から情報を抽出することで文章に合致する emoji（絵文字）を推定・提示し、ユーザに選択をさせるインタフェースを開発する。現状 emoji の入力手段としては、OS やアプリケーションに備わっている emoji の一覧から入力する方法と、文字変換を利用したものが存在する。前者は目的の emoji を探し選択する時間的コストが大きく、後者は英語圏など文字変換の無い言語圏で使用することができないといった問題点がある。

そこで本プロジェクトでは、入力した文章について文脈に合致した emoji を自動で推定し、ユーザに選択させるシステムを開発する。既に Word2Vec を用いた試作があるが、これは各単語の情報にしか着目していない。例えば逆説や二重否定などにも対応できるよう、文構造も含めて情報を抽出し、ある種文章を emoji に翻訳するようなタスクをこなすことを目指した。そして、推薦された

emoji を人間にわかりやすいように可視化し、直感的に入力することのできるユーザーインターフェースを開発した。

emoji が世界中で使われ出してからまだ 10 年も経っておらず、最近になって emoji にまつわる研究が学際的に行われるようになってきた。その中で emoji の入力インターフェースに焦点を当て、成果を出した研究はまだ存在しないため、本システムは様々な OS やアプリケーションに組み込まれる可能性がある。本システムにより、emoji 利用のハードルが下がり、多種類の emoji の利用が促進されることで、emoji を使ったよりリッチなコミュニケーションの創発が期待される。

7. 採択理由

現代に蘇った絵文字・象形文字ともいえる「emoji」は日本発祥で仮名漢字以上に世界で広く使われる文字である。絵文字やピクトグラムは手書きする手間が大きいですが、情報技術により簡便に記述可能となった。提案者らは emoji を愛し、emoji 入力のための専用キーボードを過去に開発している。本プロジェクトは文脈や文化差を考慮しつつ emoji の入力をより気軽に行うためのシステムの提案である。プロジェクト期間中に試行錯誤しつつ新たな emoji 体験を提示することを期待し採択した。

8. 開発目標

本プロジェクトは、emoji 入力インターフェースの課題解決を前提としながら、開発者やユーザ向けに社会実装することで「emoji×tech ブームの礎を築く」ことである。本プロジェクトで自己完結するのではなく、多くの人的リソースや技術を取り込む契機を作ることで、長期的な emoji 文化の発展に貢献することを目指した。

9. 進捗概要

ユーザが入力するであろう emoji を文脈より推定するシステム、それを応用した直感的かつ楽しい emoji 入力インターフェースを開発した。以下、それぞれの開発内容について詳述する。

■ Text2Emoji : 文脈に基づいた emoji 推薦システム

本プロジェクトではまず、文章を入力するとその内容に合致した emoji を出力するシステム、Text2Emoji を開発した (図 1)。実装は以下の 3 工程で行われた。

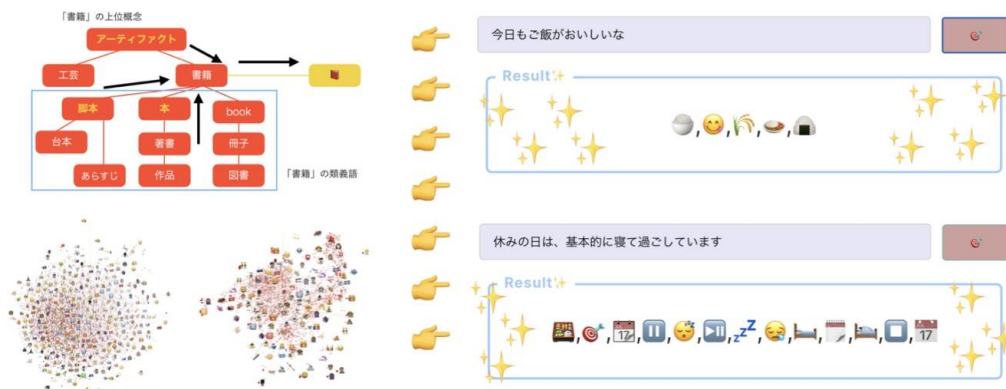


図 1 : Text2Emoji の実装手法と結果

(左上) 概念辞書 WordNet に emoji を付与し、推薦アルゴリズムを考案した。

(左下) 汎用言語表現モデル BERT を用いて、文化差や流行も考慮した emoji の定量化を行った。

(右) 推薦結果

(1) 概念辞書に対する emoji の付与

単語の情報が構造化されたデータベースである概念辞書 WordNet に対し、emoji を付与する拡張を行った。従来の emoji 検索機能は刺激語と emoji にタグ付けされた単語との照合を行っているだけであった。本システムでは刺激語の類義語や上位概念を参照することによって、広範な emoji 推薦を可能とした。

(2) 汎用言語表現モデルによる文化差や流行も考慮した emoji の定量化

「🍳」は海外において“Tofu on fire”と解釈され、感情が昂ぶった際に使用される。このように、emoji の用法には文化差や流行等が反映されることがある。以上に対応した推薦を実現するため、多様な emoji の用法を網羅した EmojiNet の文章を、汎用言語表現モデル BERT に学習させた。学習によって得られた 1024 次元の emoji ベクトルを評価すると、例えば「🍳」と「🔥」の類似性が高いという良好な結果を得た。これら emoji と emoji の関係性を用い、データベースを再拡張した。

(3) RESTful API としての実装

「emoji×tech ブームの礎を築く」という本プロジェクトの目的を達成するため、Text2Emoji を無償で利用可能な RESTful API として実装した。

▪ emolingual editor / emolingual IME : 新しく楽しい emoji 入力インタフェース

新たな emoji 入力インタフェースを開発した (図 2)。プロトタイプの開発を繰り返し、入力の楽しさを重視した 2 種類の emoji 入力インタフェースを実装した。以下ではこれらの開発について述べる。



図 2：開発した emoji 入力インタフェース

(左) 文章入力と emoji 入力の時空間的統合を図った emolingual editor👉

(右) フリックにより感情 emoji を直感的に入力できる emolingual IME👉

▪ **emolingual editor👉文章入力と emoji 入力の時空間的統合／emoji との出会い創出**

emolingual editor👉は、emoji の溢れた文章を快適に作成できるテキストエディタである。「アイデアは降ってくる」というコンセプトのもと、提示される emoji は上部から下部へと雨滴のように落ちていく。emoji の推薦は、文章入力部分に重畳する形で、かつ文章入力と同時に行われる。このように、文章入力と emoji 入力の時空間的統合を図ることで、emoji 入力インタフェースの課題を解決した。推薦アルゴリズムは、

- [1] Text2Emoji👉による文脈に基づいた推薦
- [2] emoji の使用履歴に基づいた推薦
- [3] 無作為な推薦を複合しており、それぞれの重みはユーザが決定

[3] 無作為な推薦では、新しい emoji に出会う体験や突飛な emoji を使用する体験を創出した。

▪ **emolingual IME👉フリックにより感情 emoji を直感的に入力可能なインタフェース**

emolingual IME👉は、スマートフォン向け emoji 入力インタフェースである。SNS 等においてインスタントメッセージを入力する際は、「😂」等の感情 emoji が多用される。emolingual IME👉は、興奮冷めやらぬうちに、フリックという身体動作によって感情 emoji を入力することのできるインタフェースである。

先行研究より感情を表す様々な形容詞を抽出し、Text2Emoji へと入力することで emoji のカテゴリ化・感情強度定義を行った。例えば「😄」のボタンをフリックすると、そのフリック距離に応じて「😄」→「😊」→「😁」→「😂」→「😃」といった emoji を入力することが可能となる。

10. プロジェクト評価

当プロジェクトのクリエイターらは本事業への応募前から東京大学学生の制作展にて emoji 入力専用キーボード『emolingual』を自作、出展し、ソーシャルメディアで話題を集めていた。本プロジェクトは、三千種類以上と常用漢字を超えた文字数が emoji にはあるにも関わらず、普段使う種類が限られていることに対する疑問を起点とし、各種 emoji 入力機能を実装したものである。こちらでもクリエイターらの居ても立っても居られないような emoji 愛に駆動されたプロジェクトであった。

コロナ禍によりオンラインのコミュニケーションが増える中で、テキストでのやり取りでも微妙な情動を伝えることを補助する「表感文字」(emograph)としての emoji を広めるものであるとの評価を、東京大学で漢文学を研究する齋藤希史教授から受けた。そのようなきっかけから、令和 2 年度東京大学学位記授与式総長告辞にて当プロジェクトが紹介されるなど、学術的・社会的にも興味深い取り組みとして認知されつつある。

本プロジェクトでは emoji の定量化、emolingual editor[👉]、emolingual IME[👉]の 3 点が成果物である。カナ emoji 変換を用いた入力システムとしてはプロジェクト期間中に iOS の機能として実装された。しかしながら本プロジェクトはオープン性、入力そのもののエンタテイメント性、新たな emoji との偶有性の 3 点で優れていると判断する。しかしながら、多くのユーザからのフィードバックを得るまでには至っておらず、今後の展開に期待したい。

PM としては、他のプロジェクトのクリエイターたちにもいつの間にか emoji が広がり、成果報告会でもスライドに emoji が例年にも増して使われているように、コロナ禍でのクリエイター間のつながりを感じ、笑みをたたえた次第である。

11. 今後の課題

Text2Emoji[👉] API は、すでに利用申請者に対する運用を行っている。本 API は無償で提供を行っているという点からも、継続的な運用と広報活動によって様々な Web サービスで使用される可能性がある。emolingual editor[👉]と emolingual IME[👉]に関しては、ユーザのフィードバックを取り入れた改善を行うようである。また、本成果物に留まらず、emoji 入力のデファクトスタンダードを目指したさらなる開発も続けるとのことである。

なお、本プロジェクトの成果も含むクリエイターの活動は、エンジニア向けポッドキャスト“yancanfm”に取り上げられた。このように技術者の共感を得た点

は、「emoji×tech ブーム」の第一歩とすることができる。メディア掲載や学会発表等を契機として emoji 愛好家コミュニティと連携し、今後も製品開発とその広報活動を継続的に行う必要があると考える。