

# コミュニティの成長過程分析のための LLM を用いた Slack 履歴分析手法の検討

## Considering Method for Analyzing Slack Logs using LLM toward Investigating Community Growth Process

佐野 正太郎<sup>1\*</sup> 高木 彩翔<sup>1</sup> 橋本 慧海<sup>1</sup> 白松 俊<sup>1</sup>  
三枝 大祐<sup>2</sup> 三井 実<sup>3</sup>

SANO, Shotaro<sup>1</sup> TAKAGI, Saito<sup>1</sup> HASHIMOTO, Ekai<sup>1</sup> SHIRAMATSU, Shun<sup>1</sup>  
SAEGUSA, Daisuke<sup>2</sup> MITSUI, Minoru<sup>3</sup>

<sup>1</sup> 名古屋工業大学

<sup>1</sup> Nagoya Institute of Technology

<sup>2</sup> 塩尻市振興公社

<sup>2</sup> Shiojiri City Promotion Public Corporation

<sup>3</sup> ミテモ株式会社

<sup>3</sup> mitemo Co., Ltd.

**Abstract:** To facilitate citizens' co-creative activities for addressing social issues, growth process of innovative community is worth to be analyzed and clarified. We aim to develop a method for analyzing slack logs toward clarifying how communities grow and become innovative using a large language model (LLM). Concretely, we deal with Slack logs of "Sunaba," which is a innovation hub established by Shiojiri-city, Nagano-prefecture. Our analysis focuses on ratio transition in the number of statements that use themselves as the subject and the community as the subject. The analysis in this paper itself is only a preliminary study and needs to be interpreted in combination with qualitative analysis in order to obtain useful findings. As a future work, we are planning more detailed qualitative analysis through interviews with the Sunaba management team.

## 1 はじめに

地域課題に関する市民の共創的活動を促進するには、活発なコミュニティが形成されるプロセスの解明が非常に重要である。しかし、そのコミュニティ形成にどのようなコミュニケーションや要素が有効かは未だ明らかになっていない。そこで、本研究では成功事例として長野県塩尻市のイノベーションハブ「スナバ」を中心としたコミュニティを取り上げる。具体的には、スナバ運営チームが 2018 年から利用している Slack のデータを、LLM (Large Language Model; 大規模言語モデル) を用いて分析する。

本稿は、たった 1 つのコミュニティに関するケーススタディではある。しかし、スナバは 5 年間で延べ 35 の新規事業を生み出す [三枝 24] など、非常に活発なコミュニティであるため、そのようなコミュニティが形成されるプロセスを明らかにできる知見があると考えられる。また、我々の最終的な目的は、コミュニティの活発化を促すシステムや仕組みを作ることであり、本稿はそのための予備的検討と位置付けられる。

## 2 先行研究

本研究に関連する研究として、Slack の会話データから、そのコミュニティの孤立ユーザと非活性グループを発見することを目的としたシステムについての研究について述べる。

\* 連絡先: 名古屋工業大学  
名古屋市昭和区御器所町  
E-mail: s.sano@srmtlab.org

## 2.1 Slack の会話データから孤立ユーザと非活性グループを発見するシステム

本研究と類似した研究として、瀧澤らが作成した、Slack の会話データを分析することによって、その Slack を利用しているコミュニティ内において、コミュニティと疎遠になっているユーザや、コミュニティ全体としての活動が少ないグループを発見することを目的としたシステムについての研究が挙げられる [瀧澤 23]。

この研究では、Slack の情報を Slack API などを用いて取得して、得られた定量的な指標を用いて分析することによって、実際のコミュニティの構造について明らかにすることを目的としている。なお、ここで定量的な指標とは、各ユーザの発言数や得られたリアクション数、メッセージに付与されたタイムスタンプなどを指すものとする。

以上の手法では、発言数やタイムスタンプなど、発言に関する情報を用いての分析は可能だが、それぞれの発言の内容そのものに立ち入った分析が難しい。

対して本研究では、従来手法で利用されていた定量的な指標による分析に付け加えて、大規模言語モデルを用いた会話内容についての分析も行っている。

## 3 提案手法

### 3.1 分析項目

本研究では、Slack データから設立当初から現在に至るまでの組織の変化を LLM で分析したい。LLM としては、OpenAI の GPT-4[Achiam 23]（事前学習モデル gpt-4-0125-preview）を用いる。分析項目については、スナバ代表の三枝氏による「立ち上げ時は一人称の自分語りが多く、コミュニティの成長に従って地域やスナバを主語とする発言が増えていったのではないか」という仮説を踏まえ、以下のように定めた。

- 主語の変化  
投稿内容の主語が自分なのかスナバなのかを分類する
- 感情極性  
肯定的な意見を述べているか、否定的な意見を述べているか、あるいは中立的な意見を述べているかといった情報から -1 から 1 までの値で評価されたもの
- 心理的安全性  
組織の中で自分の考えや気持ちを誰に対して

でも安心して発言できる状態のこと

- 会話のかみ合い度  
前後の会話の中で、その投稿が会話の流れにかみ合っているか
- 発言者の偏り  
投稿者の発言数に偏りがいないか
- 「スナバ」のカルチャーに関わる発言  
スナバにおける 5 つのカルチャーに関連した発言の数

3 つ目の心理的安全性と 4 つ目の会話のかみ合い度については、前後の会話内容が重要になってくる。しかし、Slack のデータが膨大でありすべてのデータを一度に LLM に渡すことができない点と、LLM の評価の一貫性を保つことができないという点から、この 2 つの評価項目については今後の課題とする。また、スナバのカルチャーに関連した発言の数についても、本稿では分析が間に合わなかったため、今後の課題とする。

### 3.2 分析手法

Slack API を用いて、スナバの 3 か月ごとの Slack データを取得する。この 3 か月は、各年の四半期で分割されるようにする。API の期間指定できるパラメータでは、UNIX 時間形式のタイムスタンプが用いられており秒数まで指定できるが、期間の境界は両方含むか両方含まないかの 2 通りのみ指定できる仕様となっている。そのため、奇数月から始まる期間では両境界の時間を含み、偶数月では含まない設定になっている。これを各チャンネルのデータで行う。そして、同じ期間における異なるチャンネルのデータを一つにまとめる。この時投稿データの順序は、その投稿日時の時系列順にする。これを全期間で行う。基本的にこのまとめたデータ群を使用する。細かい分析手法について、各分析項目ごとに以下に示す。

#### 3.2.1 主語の変化

まとめたデータに格納された投稿を、1 つずつ GPT への入力とし主語を判定する。主語は「投稿者自身」と「スナバ」「地域」「その他」の 4 つに分類する。1 つの投稿に対して 1 つの主語にのみ分類手法と、1 つの投稿に複数のメッセージが含むと判断した場合にそのメッセージごとで主語の分類を行う 2 つの手法をとった。分類後、各期間における各主語の数と割合を計算する。

### 3.2.2 感情極性

まとめたデータに格納された投稿を、1 つずつ GPT に渡し感情極性を評価する。感情極性とは、肯定的な意見を述べているか、否定的な意見を述べているかを-1 から 1 までの数値で評価したものである。ネガティブな内容ほど-1 に近く、ポジティブな内容ほど 1 に近い値をとり、ちょうどその中間の中立的な内容は 0 となる。各投稿ごとにこの感情極性を評価し、その後期間ごとに感情極性の平均をとる。

### 3.2.3 発言者の偏り

まとめたデータに格納された投稿それぞれの、発言者のラベルを各期間でカウントする。その後、各ユーザがその期間で発言した割合を確率とみなして、エントロピーを計算する。

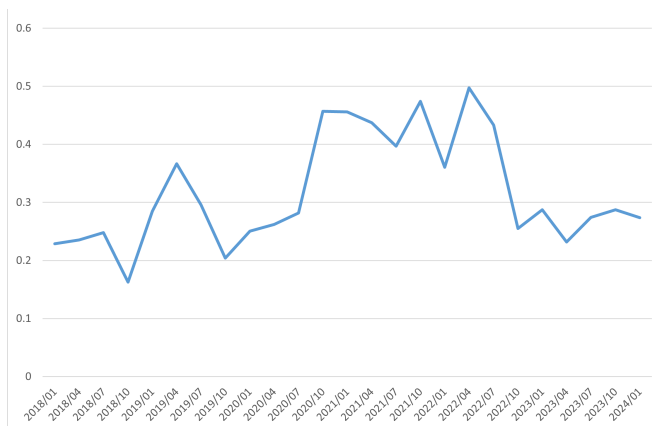


図 3 各期間における感情極性の平均値の推移

## 4 分析結果と考察

### 4.1 主語の変化

各主語の投稿数の推移は以下の図 1 のようになった。error は 4 つの分類以外の出力が出た回数である。全投稿の中で 4 回の予期しない出力がされ、ハルシネーションが起きてしまっていると考えられる。その出力内容の多くは、other\nother\nsunaba などの主語の分類の間に \n が入ったものであった。 \n は改行コードであるため、その投稿内容が複数の主語に分類される、または 1 つの投稿に複数のメッセージが含まれていると判断された可能性がある。今後の課題として、投稿ごとだけでなく投稿内の各メッセージごとの分析も必要だと考えられる。

次に図 2 の各主語の割合の推移を見ると、スナバを主語とした投稿の割合が、2018 年から 2019 年の 2 年間で増加しその後緩やかに減少していく様子が確認できる。逆に、地域を主語とした投稿は、設立当初は 8 割近くを占めていたが、その後 2 年間で半分ほどに減少し、その後は徐々に増加している。

この結果は、「立ち上げ時は一人称の自分語が多く、コミュニティの成長に従って地域やスナバを主語とする発言が増えていったのではないか」という仮説を単純に支持するものではなかった。よって、今後はスナバ運営チームとのヒアリングを通じ、より詳細に定性的な分析をあわせて上記分析結果を解釈していく必要がある。

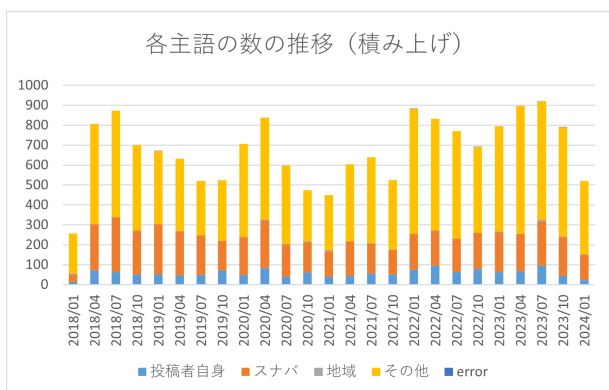


図 1 各主語の数の推移 (積み上げ)

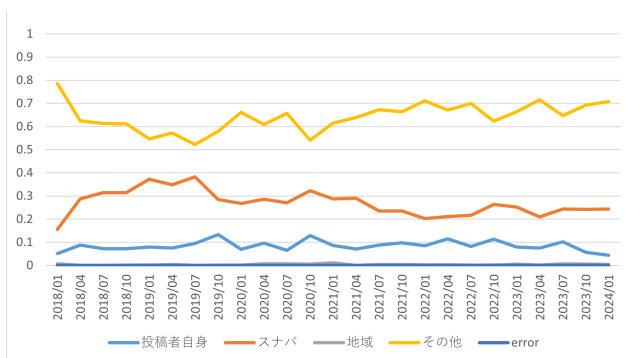


図 2 各主語の割合の推移

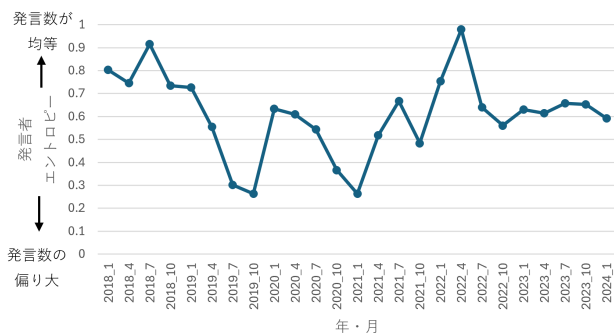


図4 発言者エントロピーの推移

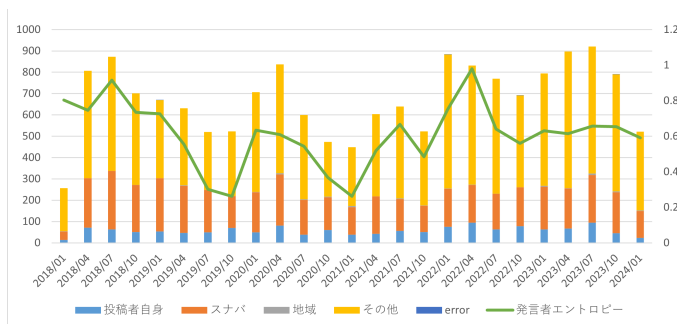


図5 主語の積み上げグラフと発言者エントロピーの比較

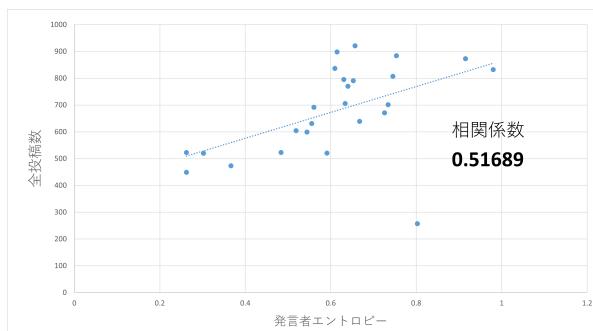


図6 投稿数と発言者エントロピーの関係

## 4.2 感情極性

得られた各期間における感情極性 (-1 から 1) の平均値の推移を図3に示す。投稿ごとで見ると感情極性が負の値であるものもあったが、平均すると全期間で正の値に偏っており、ポジティブな発言が多いことがわかる。特に、2020年からのコロナ禍の時期にポジティブな発言が増えているが、これが偶然なのか、あるいはコロナ禍のような社会情勢と関係があったのかは、今後定性的に分析する必要がある。

## 4.3 発言者の偏り

各ユーザがその期間で発言した割合を確率とみなして計算したエントロピー（発言者エントロピー）は以下の図4のように推移した。発言者エントロピーが小さいほど投稿者の発言数に偏りがあり、大きいほど均等に発言されていることを示している。ここで、図2つのグラフを1つにまとめたところ図5のようになった。このグラフから、全体の投稿数と発言者エントロピーには相関があるように見える。これを確認するため、図6に投稿数と発言者エントロピーの散布図を示す。ピアソン相関係数は+0.517となり、中程度の正の相関があることが確認できた。つまり、投稿数が多くコミュニケーションが活発な時期ほど、各参加者が均等に発言する傾向がある。

ただし、2023年春から夏にかけて、投稿数が増えているにもかかわらず発言者エントロピーが増加していない例外的な時期が確認できる。これについても、今後定性的な分析が望まれる。

## 5 まとめと今後の展望

本稿では、長野県塩尻市のイノベーションハブ「スナバ」のSlackデータをLLMで分析し、活発なコミュニティがどのような成長過程を辿るのかを定量的に分析することを試みた。主語の割合遷移や感情極性に着目した定量的分析を行ったが、本稿の分析は予備的検討に留まっている。ここから有用な知見を得るには、定性的分析を併用して解釈する必要がある。今後は、スナバ運営チームへのヒアリングを通じ、より詳細な定性的分析を試みる。心理的安全性や会話のかみ合い度、スナバのカルチャーに関連した発言の定量的分析についても更に実装を進め、定性的分析とあわせて解釈を試みる予定である。

また、Slackだけでなく音声で行われる会議のデータも記録することで、より詳細な分析を可能にしたい。さらに、スナバ以外のコミュニティについても同様の分析を試行することで、より汎用的なコミュニティ分析手法を模索する予定である。

## 謝辞

本研究の一部は、JST CREST (JPMJCR20D1) および NEDO (JPNP20006) の支援を受けたものです。Slackのデータをご提供頂いたスナバの皆様に感謝します。

## 参考文献

- [Achiam 23] Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., et al.: GPT-4 technical report, *arXiv preprint arXiv:2303.08774* (2023)
- [三枝 24] 三枝 大 祐. シビック・イノベーション拠点スナバがみ出すインパクトの最化に向けての考察 – 中間援組織の視点から分析する地域イノベーション創出の成功要因と今後の可能性 –, 修士論文, 長野県立大学 ソーシャル・イノベーション研究科 (2024)
- [瀧澤 23] 瀧澤亮佑, 本松大夢, 中尾一心, 谷中健介, 瀧口 諒久, 荒川豊. 組織 Slack における孤立ユーザと非活性グループの発見を目的としたコミュニケーション分析システムの開発, 第 31 回マルチメディア通信と分散処理ワークショップ論文集, pp. 106–113 (2023)