

# 我が国在外公館における業務効率化を目的とした YouTube 上における対日論調分析の支援ツール

原田武夫<sup>1</sup>

<sup>1</sup>株式会社原田武夫国際戦略情報研究所  
takeo.harada@haradatakeo.com

## 概要

我が国においては少子高齢化が加速度的に進捗し、とりわけ国の根幹を成す中央省庁における職員の確保、さらには知識伝播と技術伝承が課題であり続けている。そうした中で SNS を通じた論調の拡散は時に国境を越えたインパクトを与えている。これらの状況については外交実務の現場も例外ではなく、ここでは特に YouTube 上の動画拡散による対日論調の展開は今後より一層注目すべきでありつつ、同時に効率的な業務の遂行という要求にも応える必要がある。以上を踏まえ、本研究では我が国在外公館において経験の乏しい若手の末端担当職員が往々にして担当している YouTube 上の対日論調分析に関し、効率的な支援ツールを提案するものである。

## 1 はじめに

少子高齢化が加速度的に進む中、本稿執筆段階で我が国政府は「AI 基本計画」の策定を進めようとしているとの報道がある[1]。同報道によれば人工知能の利活用を推し進めるべく、公的機関での積極活用を進めるべしとの方針がそこでは盛り込まれる見込みであるが、そもそもこうした方針が定められること自体、我が国の行政の現場においては人工知能の利活用が遅れていることを如実に表している。ましてや中央省庁の中でも外交を司る外務省においては、そうした人工知能の社会実装に向けた目に見える努力の形跡が見られないことが[2][3]で論じられてきている。

そこで本研究は以上の問題状況を踏まえつつ、我が国外交の現場においても利活用が可能な、大規模言語モデルをベースとする YouTube 上の対日報道分析の支援システムを提案するものである。具体的には (1) 我が国外交、更には中央省庁における行政全体としての人工知能の導入状況につき振り返りつつ、(2) ソーシャル・メディア (SNS) の中で

も YouTube 上の動画に対する分析に関する先行研究を概観し、次に (3) 生成 AI を中心とする人工知能の外交実務における利活用に関する先行研究につき考察する。以上を踏まえた上で (4) 我が国外交実務の中でも在外公館における任国での YouTube 上の対日関連動画を網羅的に分析するためのシステムを提案し、(5) 具体的にはドイツ語圏を対象に現地の日本大使館での実務を念頭におきつつ、本件システムを稼働させ、出力した結果について検討する。そして (6) 本件システムが、外交実務における業務効率化の観点から有効であることを示し、さらにこれを我が国外交実務の現場へと落とし込んでいくにあたり、今後検討すべき諸課題について述べることにしたい。

## 2 先行研究

ソーシャル・メディア (SNS) 全般と、外交の中でもいわゆるパブリック・ディプロマシーとの関係性を論じたのが[4]である。その中で Di Martino はパブリック・ディプロマシーにおけるソーシャル・メディアを通じた世論の把握に際し、5つのスペクトラムがあると論じている。具体的には「能動的傾聴 (active listening)」「戦術的傾聴 (tactical listening)」「聞き入ること (listening in)」「背景に関連し、カジュアルな形で行う傾聴 (background/casual listening)」「秘密傾聴 (surreptitious listening)」の5つである。我が国中央省庁の一つである外務省における実務上、課題となるのはソーシャル・メディア (SNS) におけるいわゆる「炎上」である。そしてこれを抑止すると共に、仮に予防的措置で間に合わなければ適時的確に反論することが実務上は最も重要なのであって、その意味では「戦術的傾聴」こそが最大の課題であると言えることが出来る。

ソーシャル・メディア (SNS) の中でもとりわけ YouTube に関連し、研究が進んでいるのが投稿

された動画に対するコメントに関する感情分析である。外交実務の現場に当てはめてみると、個別の外交政策を展開するにあたり、例えば首脳会談や国際会議などの現場の画像を公式 YouTube 上で配信をした場合、それに対して視聴者がどの様に受け取ったのか、もっといえば外務当局の側が意図したとおりにそこでは受け取られているのかが大変気になる場所である。[5]は YouTube 上のコメントについての感情分析の手法について述べたものであり、また[6]はこの分野における先行研究をまとめたレビュー論文である。しかしこれらはいずれも外交分野あるいは外交実務を念頭においた研究ではない点で本研究にとって先行研究とは言えない。また、本研究がスコップとしているのは、外交当局が発信した YouTube 動画に対する一般公衆からのコメントではなく、その意味でもこれらは先行研究とは見なし得ないものである。

本研究が念頭に置いているのは、我が国の外交実務の現場の中でも在外公館において、任国で YouTube 上に投稿された動画における対日論調に関し、可能な限り仔細に分析する人工知能を用いた実務支援システムの構築とその効果検証である。そのシステムでは古典的な機械学習のアルゴリズムではなく、生成 AI を用いることを念頭に置いている。そこで次に生成 AI を中心とした人工知能の外交実務における活用に関する研究を検討する。

[7][8]は生成 AI を念頭に置きつつ、外交と人工知能の関係性について俯瞰する研究である。とりわけ前者では 1988 年から 2023 年の間に、外交と人工知能の関係性について論ずる研究が加速度的に増大してきたと述べている。また後者ではとりわけパブリック・ディプロマシーにつき、生成 AI の登場によって大きく変化が生じているとし、最新の研究動向を俯瞰する。しかしこれらの研究はいずれも概観を述べるに止まり、結果としてどの様なアルゴリズムを用いて、外交実務における個別具体的な負担を軽減することが出来るのかといった技術的な論点が欠如しているため、本研究にとっての先行研究とは言えない。これに対して[9]は外交実務における個別の局面で人工知能の中でもこういったモデルが適用可能なのかにつき、より詳細に述べるものがある。しかし本研究が個別具体的に対象とする、ソーシャル・メディア (SNS) の中でも YouTube を取上げ、そこにおける対日論調の詳細な分析といった、より実務に即した議論までカバーしているわけ

ではないので、同様に先行研究とは言えないものである。

また[10][11]は生成 AI と外交との関係性についてより絞り込んだ形で論じたものである。この内、前者は生成 AI が外交に与える影響には 3 つのシナリオがあり得ると論じる。また後者は生成 AI の投入により、リアルタイムでの情報収集がより精緻に可能になったことにも触れ、生成 AI の外交における活用可能性について論じるものである。しかしこれらについてはいずれも、一般論を抽象的に論じるにすぎず、特に業務効率性の観点から個別の外交実務における生成 AI の活用を述べるものとは言えないため、本研究の先行研究とは言えない。同様のレベルでの研究としては[12]もある。

本研究が論じる我が国外交実務に対する支援システムの利用は、経験の乏しい若年担当職員にとってはトレーニングとしての意味合いを持つ。[Phan 25]は外交官に対する教育訓練と人工知能の関係性について論じている。その際、感情分析についても触れるが、生成 AI をも用いてソーシャル・メディア (SNS) 、とりわけ YouTube に特化した形でそこでの情報拡散について外交実務当局者の観点から分析するためのより詳細なツールを取上げてはいない。したがってこの観点から言うと、本研究と視点として重なる部分が無いわけではないが、先行研究とは言えないものであると言ふべきである。

### 3 採用するシステム、データ及び評価方法

本研究で提案する業務支援システムは我が国の在外公館における任国の YouTube 上での対日論調の分析を支援するためのものである。ここでは利用場所を、一例としてドイツ連邦共和国を管轄する在ベルリン日本国大使館と仮定した。コーディングするにあたり、Google Colabratry 上の Python3.12.11 にて環境構築を行った。YouTube とは API キーを用いて連携させた。その上で OpenAI 社の大規模言語モデル gpt-4o を用いるべく、同じく API キーを用いて連携させた。その際、OpenAI 社の gpt-4o を用いたのは実験段階における性能の日本語環境における技術的な比較優位性による[2][3]。実装実験は 2025 年 6 月に実施したが、そこから遡って 180 日の間にドイツ連邦共和国内において YouTube にアップロードされた動画を対象に対日論調の動向分析

のための検索を行った。その際、まず検索ワードとしては以下のとおり設定した（この時期に開催されていた「大阪・関西万博」を意識した設定）。

```
search_query = 'Japan OR Tokio OR Ishiba OR
Kaiser OR "Osaka Expo" OR "Kansai Expo"
relevant_terms= ["Japan", "Tokio", "Osaka",
"Ishiba", "Kaiser"]
```

次に検索でヒットした動画について ID を取得し、ドイツ語がそこで用いられているのかを判定する。その後、字幕を取得し、根拠文を抽出した上で感情分析を実施し、かつ要約文を作成した。そして根拠文と要約の両方について日本語への翻訳を実施したが、後者については本件業務支援システムを用いることが想定されている若手の末端担当職員から、より忙しい上長に対して簡潔に報告が行われるべきであるという観点から 200 語以内という制限をかけた。

結果については Python 上で各動画について「ドイツ語で記されたタイトル」「URL」「感情分析の結果 (positive/negative)」「感情分析の根拠文 (日本語訳)」「日本語要約文」「視聴回数」「チャンネル登録者数」「投稿日」を並べて DataFrame の形で示すようにした。

そして最終的に提示される各動画の最終的な日本語要約文については、感情分析の一覧を見て特にチェックすべきと考えられる動画に関し、上長が仔細に見る可能性が最も高い項目であるとの観点から、次の視点からの評価を大規模言語モデル (gpt-4o) 自身に 10 段階評価で行われた (1 が最低点であり、10 が最高点)。

- ▶ 文脈の関連性 (context relevance)
  - ▶ 回答の忠実性 (answer faithfulness)
  - ▶ 回答の関連性 (answer relevance)
  - ▶ ノイズへの頑健性 (noise robustness)
  - ▶ 回答不可能な質問の却下 (negative rejection)
  - ▶ 情報の統合 (information integration)
  - ▶ 反事実への頑健性 (counterfactual robustness)
- これら 7 つの指標の内、前三者は検索拡張生成について一般的な性能評価を行う際に用いるべきものとして提案されているものである[13][14]。また後四者は検索拡張生成の構成要素となる大規模言語モデルに対し、求められる要素として提案されている指標である[13][15]。

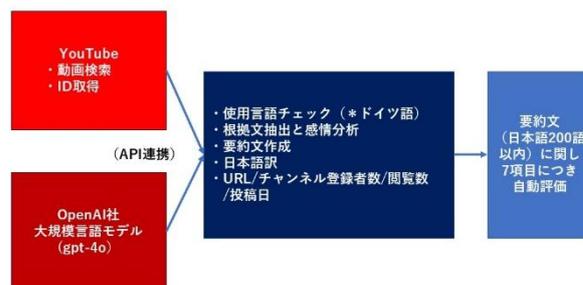


図 1 本研究が提案する業務支援システムのフロー

## 4 実験計画及び結果

前章で示した業務支援システムを我が国外交実務における実装を想定しながら、2025 年 6 月に大規模言語モデルに何らの事前学習をさせることなく稼働させたところ、全部で 27 本の動画を抽出することが出来た。出力に際して要した時間は約 18 分であった。概要を示すならば以下のとおりである。

### ・感情分析

Positive は 23 件、negative は 4 件であった。

### ・動画の内容

我が国における個人的な観光体験に関するものは 20 件、社会・文化に関するものは 5 件、法制度に関するものは 1 件、政治に関するものは 1 件であった。

### ・動画の発信者

ドイツより個人が発信していた動画が 19 件、企業によるものが 3 件、公営放送局が 2 件であった。他方でドイツ以外の国からの発信が 3 件（我が国が 1 件、米国が 1 件、パキスタンが 1 件）あった。

### ・動画の要約に対する自動評価

7 つの評価項目のそれぞれについて抽出された 27 本に付された要約の平均評価値 (10 点満点) は次のとおりであった。

- ▶ 文脈の関連性：9.4
- ▶ 回答の忠実性：9.3
- ▶ 回答の関連性：9.5
- ▶ ノイズへの頑健性：9.3
- ▶ 回答不可能な質問の却下：9.2
- ▶ 情報の統合：8.7
- ▶ 反事実への頑健性：9.1

### ・動画の視聴回数及びチャンネル登録者数平均値

実験実施段階での視聴回数の平均値は約 788952 回、チャンネル登録者数の平均値は約 1388390 人であった。

## 5 考察及び今後の課題

実験の対象となった動画のアップロード期間には折々における我が国の政治経済状況等に関する報道も大手メディアによって多々なされていたはずであるが、そうした個別報道についての網羅的なピックアップは今回の業務支援システムでは出来なかった。また、動画のアップロードを行った国をドイツ連邦共和国とコード上設定したにも拘わらず、対日論調関連とはいえ、それ以外の国を発信源とする動画が3件も含まれていた。

他方で抽出された動画の要約作成については自動評価上、概ね評点が9前後であり、外交経験の乏しい末端担当職員がこのシステムを稼働させ、その出力結果を経験値の高いマネージャーがこれをチェックするに際し、叩き台としては十分利用することが可能なレベルであったと考える。また、公営放送局による動画2件が抽出されていたが、その内1つはドイツ第2公営放送(ZDF)が我が国の刑務所制度について批判的な観点から検証した番組であった。我が国在外公館としては正に対日論調に歪みがないかチェックすべき対象である動画を適切に抽出することが出来たと言える(その際、感情分析はnegativeと判断しており、マネージャーによるチェックはしやすい出力結果となっている)。他方で公営放送局によるものではなかったが、感情分析上negativeとの評価を受ける形で我が国の財政状況について詳細の論じる動画を民間企業がアップロードしたのも本業務支援システムは抽出出来ていた。この 이슈も我が国の国家財政に対する対外的な信頼を維持するという観点で外交実務の現場において無視出来ないものである。実験段階での閲覧数は44500名に止まっていたが、実務的にはその拡散の度合いとスピードについて、要経過観察とすべきと考える。

他方で自ら在外公館に勤務した経験を持つ筆者にとってすら驚きであったのは、時に数十万人以上のチャンネル登録者数を抱える男女のドイツ人ブロガーたちが我が国を旅行者として訪問した時の印象をつぶさに動画に収め、コメントし、それを同じく数十万人以上の視聴者が見ているという事実を確認出来たことであった。在外公館における主たる任務は外務本省から指示される「我が国の公的見解」を正確に相手国政府に伝達し、また相手国政府からも同じレベルでのメッセージを受け取り、東京に伝達することである。しかし昨今はこれに加えて、パブリ

ック・ディプロマシーの観点が不可欠であることは前述のとおりなのであって、しかも数十万人以上のパワーYouTuberたちが日々アップロードしている「ニッポン旅行記」こそが、一般公衆レベルでの対日論調を形成していることも容易に想定されるのである。したがって政策的なインプリケーションとしては民間レベルでしばしば行われる「インフルエンサー・マーケティング」と同様にこうしたパワーYouTuberたちに対して、ステルス・マーケティングとの誇りを事後的に受けることのないよう十分に配慮しつつ、本研究で提案する業務支援システムによって俎上に上がってきた動画に関し、在外公館の広報文化担当より積極的にアプローチするといった、伝統的な外交では考えられなかったようなきめの細かな対応が必要不可欠であることがうかがわれるのである。

今後の課題としてはまず在外公館での使用を前提にする以上、本業務支援システムで抽出される動画について、当該任国以外からアップロードされたものを完全に排除するような仕組みにすべきである。また動画系SNSはYouTubeだけではなく、それ以外の同種の媒体についても網羅的に取り扱うシステムの構築が求められる。さらに本研究で提示した業務支援システムにおいては前述のとおりOpen AI社の大規模言語モデルであるgpt-4oを利用したが、我が国外交の最前線である在外公館において実装するに際しては、秘密保全の観点からこの点に十分配慮した「国産大規模言語モデル」を利用することが最終的には望ましい。

その上で対日論調は一時的に形成されて終わるものではなく、むしろ定点観測をしながら中長期的にその動向をウォッチすべきものでもある。したがってこの様な観点に立ちながら、「定点観測」を異なる時期で行い、例えばそこで抽出される動画において用いられるテキストにつき自然言語処理を施し、対象国における対日論調のトレンドをより巨視的な観点から抽出する中で我が国外交実務の現場における実装という観点で有意な成果を上げられるのかどうかも、今後取り組むべき課題と考える。

## 参考文献

1. 読賣新聞. 政府、「世界で最もAIを開発・活用しやすい国」目指して「反転攻勢」...基本計画策定に着手,

<https://www.yomiuri.co.jp/science/20250912-OYT1T50071/>, 2025-9 閲覧.

2. **原田武夫**. 国会答弁案作成システムにおける生成 AI の活用 「対北朝鮮外交」 を題材とした実装の検討. 人工知能学会全国大会論文集 第 38 回, 2024.
3. **原田武夫**. 大規模言語モデルを活用した我が国の対北朝鮮外交実務支援システム. 人工知能学会全国大会論文集 第 39 回, 2025.
4. **Di Martino Luigi**. Conceptualising public diplomacy listening on social media: Place Branding and Public Diplomacy 16.2, 2020.
5. **Bindhumol M., Tripty Singh, Payel Patra**. Sentiment Analysis using YouTube Comments: 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT). IEEE, 2024.
6. **Meghana Karnam**. Artificial intelligence and sentiment analysis in YouTube comments: A comprehensive overview: 2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT). IEEE, 2024.
7. **Bano Muneera**, Zahid Chaudhri, Didar Zowghi. The role of generative AI in global diplomatic practices: A strategic framework: arXiv preprint arXiv:2401.05415, 2023.
8. **Bano Muneera**, Zahid Hafeez Chaudhri, Didar Zowghi. "Mapping the Scholarly Landscape on ai and Diplomacy: *The Hague Journal of Diplomacy* 20.2, 2025.
9. **Mostafaei Hamidreza**, Kordnoori Shirin, Ostadrahimi Mohammadmohsen, Banihashemi Saeed Seyed Agha. Applications of artificial intelligence in global diplomacy: A review of research and practical models: *Sustainable Futures*, 2025.
10. **Bjola Corneliu, Ilan Manor**. Digital diplomacy in the age of technological acceleration: three

impact scenarios of generative artificial intelligence: *Place Branding and Public Diplomacy*, 2024.

11. **Bjola Corneliu, Ilan Manor**. First contact: integrating generative AI into digital diplomatic intelligence: *Place Branding and Public Diplomacy*, 2024.
12. **Sevin Efe, M. Evren Eken**. Yet another turn? prioritising the needs of diplomacy over the capabilities of generative AI: *Place Branding and Public Diplomacy*, 2024.
13. **Gao Yunfan, Xiong Yun, Gao Xinyu, Jia Kangxiang, Panm Jinliu, Bi Yuxi, Dai Yi, Sun Jiawei, Wang Haofen**. Retrieval-augmented generation for large language models: A survey.", arXiv preprint arXiv:2312.10997, 2023.
14. **Es Shahul, James Jithin, Espinosa Anke Luis, Schockaert Steven**. Ragas: Automated evaluation of retrieval augmented generation.", arXiv preprint arXiv:2309.15217, 2023.
15. **Chen Jiawei, Lin Hongyu, Han Xianpei, Sun Le**. Benchmarking large language models in retrieval-augmented generation. *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 38. No. 16. 2024.