

# 地方自治体の青少年に向けたネットトラブル相談事業における 生成系 AI 利用時の課題の分析

池辺正典<sup>1</sup> 櫻井淳<sup>1</sup> 川合康央<sup>1</sup>

**概要**：近年の青少年のインターネット利用率は高い水準にあり、それと同時にネットトラブルも深刻化している。この状況に対応するために政府は青少年の安全なインターネット利用を目指した各種施策や総合的なこども支援のためにこども家庭庁を設置するといった様々な対策を行っている。このような流れを受けて青少年のネットトラブルに対する支援として情報倫理教育の充実やネットトラブルの相談窓口の整備が求められており、都道府県を中心として、その整備が行われている。しかし、ネットトラブルは様々な種類が存在する点やその内容が複雑化しており、さらには専門家による対応が必要なケースも多く見られ、対応を拡充することが困難な状況である。

このため、本研究においては、近年急速に発展している生成 AI を用いて、地方自治体のネットトラブルの相談窓口業務を支援することを想定した際に、どのような課題が発生するかを分析することを目的としている。具体的な分析としては、専門家の相談対応の内容と生成 AI の相談対応の内容を比較することで、一般的な生成 AI を相談業務に用いた際における相談内容のカテゴリ別の回答品質の測定を行う。そして、両者の回答で用いられている単語傾向を分析することで、大規模言語モデルの学習量が充分であるかを確認する。さらに、回答品質が低い事例の原因を分析することで、相談事業に特化したファインチューニングを行う際に必要となる追加学習の方向性を検討する。

**キーワード**：ネットトラブル、青少年、情報倫理、生成 AI、地方自治体

## Analysis of Issues When Using Generation AI in Local Government's Internet Trouble Consultation Service.

MASANORI IKEBE<sup>†1</sup> JUN SAKURAI<sup>†1</sup>  
YASUO KAWAI<sup>†1</sup>

### 1. はじめに

#### 1.1 青少年を取り巻く環境

近年の青少年を取り巻く情報利用環境は充実しており、各家庭ではスマートフォン、タブレット端末、携帯ゲーム機、テレビ等の様々な端末からのインターネット利用が可能となっている。こうした背景を受けて一般世帯における2022年度の青少年のインターネットの利用率としては2歳時点で6割以上、7~8歳で9割以上の利用率と高くなっている[1]。さらに、青少年がインターネットで日常的にトラブルに巻き込まれている。

こういった社会背景を受けて、政府は「青少年が安心してインターネットを利用できる環境の整備等に関する法律（平成20年法律第79号）[2]」を制定することで青少年が安心してインターネットを利用することができる環境の整備を目指している。そして、当該の法律は情報技術の進展にあわせて数度の改定が行われている。これに対応した具体的な施策としては、平成21年に青少年インターネット環境整備基本計画が策定され、3年毎の改定が行われている。直近の改定としては、令和3年に「青少年インテ

ーネット環境整備基本計画（第5次）[3]」が策定されている。この基本計画では青少年の安心安全なインターネット利用のための相談体制等の整備も示されており、公的機関による相談窓口が多く整備されるに至っている。

この他にも政府の取り組みとしては、こども基本法[4]を根拠として、こども政策の基本的な方針等を定めたこども大綱の策定を行い、その業務を円滑に進めるために、2023年4月にこども家庭庁が設置された。こども政策の中には「子育てに伴う喜びを実感できる社会の実現のためのサポートをすること」として相談窓口等の設置についても言及されている。そして、行政における青少年やその保護者を対象とした情報倫理教育の重要性の高まりと、実際にトラブルが発生した際のネットトラブル対応のための相談事業に対する需要が高まっている。

このような行政のネットトラブルの相談窓口は、専門家や警察への相談の前段階で行われることが多く、軽微な内容から深刻な内容まで多岐に渡る。特に深刻な問題については、警察等への相談を進めるケースが多いが、その対象はインターネット・ホットラインセンターで違法有害情報として定義されており、運用ガイドライン[5]に詳細が定められている。具体的には、「わいせつ関連情報4種類」、「薬物関連情報5種類」、「振り込め詐欺等関連情報2種類」、「不正アクセス関連情報2種類」の合計13種類の情報が対象と

<sup>1</sup> 文教大学  
Bunkyo University.

されている。これらの内容については、行政における相談窓口から最終的に警察に引き継ぎ、それ以外の内容については、行政の相談窓口において一定の解決に向けた流れが期待される。

## 1.2 行政の相談窓口について

行政の実際の相談窓口としては都道府県が設置しているケースが多く、最も大規模に事業を実施している東京都では、「こたエール[6]」という名称で相談事業が行われている。当該事業に関する統計情報等としては2011年以降のものが公開されており、ここ数年は年間1,700~2,800件程度の相談が寄せられており、多くの需要がある反面、人的作業による対応では多大なコストがかかる点や専門家の必要性といった課題もあり、多くの相談事例に対応することが困難である。このことから、簡単な相談事例についてはAI等を活用した相談業務支援に期待が集まっており、ネットトラブル対応でのAIによる回答品質を明らかにすることが期待される。

## 1.3 生成AIについて

現在は様々な生成AIのサービスが開始されており、その多くは、インターネット上の言語資源等を大規模に学習した大規模言語モデル(LLM)を利用した回答生成を行う。この場合の特徴としては、幅広い質問に対応が可能である反面、前述のネットトラブル相談対応のような、特定の話題を深く掘り下げる質問への回答品質が低いケースがある。こういったケースに対応するためには、ファインチューニング等による利用事例に応じたLLMへの追加学習等が必要となるが、質問の種類に応じて必要な学習量は異なると思われるために、多くの追加学習が必要なカテゴリを明らかにすることが重要となる。

## 2. 本研究について

### 2.1 研究の目的

本研究では、地方自治体等におけるネットトラブルの相談窓口について、AI技術による支援を行う際の課題を明確にすることを目的とする。具体的には、相談事業に生成AIを用いる際に、一般的なLLMでの対応が困難なカテゴリの特定や生成AIでの回答が苦手とする要因の分析を行う。

### 2.2 Webスクレイピングによる相談事例の収集

本研究では実際の相談事例にどのような種類があるかを確認するために、「こたエール」で公開されている相談事例を収集し、Webスクレイピングによる収集プログラムの開発を行った。収集対象の情報は以下の通りである。

- 相談の時期
- 相談内容
- 専門家によるアドバイス(回答)
- 専門家からのポイント(関連知識)

収集プログラムはPythonによる実装を行い、Webスクレ

イピングの処理を行うために、HTML DOMパーサーとしては、Beautiful Soup[7]を利用した実装を行い、16種類のカテゴリ別の事例として714件の情報を収集した。各カテゴリの件数は以下の通りである。

表1 取得した相談事例の各カテゴリの件数

Table. 1 Number of consultation cases in each category.

カテゴリ	件数	警告	違反
架空請求	114	33	6
ネットショッピング	13	0	0
ネットいじめ	53	1	0
迷惑メール	58	1	0
有害情報	22	2	1
著作権関連	19	0	0
依存	45	0	0
交友関係	55	0	0
性的トラブル	78	9	0
削除方法	98	0	0
料金関係	51	0	0
フィルタリング関係	11	0	0
取引トラブル	9	0	0
情報セキュリティ	10	0	0
個人情報関連	14	0	0
その他	64	0	2
合計	714	46	9

収集したデータのうち相談内容を利用することで、生成AIからも各相談内容に対して回答を得ることが可能となる。本研究では収集データに生成AIの回答をあわせることで、その内容を分析する。また、生成AIには、回答の生成に関するルールがあり、今回の対象とする相談内容が犯罪利用等の抑止のためのポリシー制約に違反と判定されることもある。実際には、714件中46件の警告と9件の違反という全体の7.7%がポリシー違反の対象となる可能性があった。警告や違反の場合にも現在は回答を得ることはできるが、今後は利用ができなくなる可能性もあるために、注意が必要である。

### 2.3 専門家の生成AIの回答の比較

#### 2.3.1 実験の条件

収集した714件の事例では、相談内容の他に、専門家による回答内容と関連する知識としてポイント解説が公開されている。本研究では、この質問文を利用することで、生成AIからの回答内容を取得し、専門家と生成AIの回答を比較する。今回の実験にあたり、生成AIはChatGPT[8]を用いた。また、回答生成に用いた言語モデルは、GPT-3.5である。

#### 2.3.2 カテゴリ別の単語傾向の比較

単語傾向の比較では、専門家の回答と生成AIの回答内容を対象として、形態素解析を行い、単語を取り出して、頻出単語の分布から回答の傾向を比較した。カテゴリ別の上位単語は次の通りである。

表 2 各カテゴリの頻出名詞上位 10 件

Table. 2 Frequent nouns for each category.

カテゴリ	専門家の回答	生成 AI の回答
架空請求 一致：1	電話, 契約, メール, 必要, 相手, 番号, 有料, 無視, <b>請求</b> , 同意	場合, 情報, サイト, 連絡, 確認, 的, 性, <b>請求</b> , 重要, 可能
ネットショッピング 一致：4	<b>者</b> , <b>保護</b> , 利用, <b>サイト</b> , <b>情報</b> , 相手, 規約, 未成年, 相談, 同意	場合, 連絡, <b>情報</b> , <b>者</b> , <b>サイト</b> , 的, 性, メール, 手続き, <b>保護</b>
ネットいじめ 一致：3	<b>相手</b> , 削除, <b>学校</b> , 先生, <b>相談</b> , インターネット, 人, SNS, 今後, 対応	場合, 重要, 問題, <b>学校</b> , 状況, 的, <b>相談</b> , 自分, 娘, <b>相手</b>
迷惑メール 一致：3	<b>メール</b> , <b>迷惑</b> , <b>メール</b> , アドレス, 利用, <b>サイト</b> , 無視, 電話, 設定, 登録, 人	<b>メール</b> , 場合, 情報, <b>迷惑</b> , 確認, 詐欺, 性, 個人, <b>メールアドレス</b> , 可能
有害情報 一致：4	<b>サイト</b> , <b>利用</b> , メール, 者, フィルタ, リング, 運営, <b>情報</b> , <b>確認</b> , 側	<b>情報</b> , 場合, <b>サイト</b> , <b>利用</b> , 子供, 的, 重要, 性, <b>確認</b> , 個人
著作権関連 一致：5	<b>著作</b> , <b>権</b> , <b>者</b> , 利用, 物, 性, 侵害, インターネット, 可能, 削除	<b>著作</b> , <b>権</b> , 場合, 的, 使用, 画像, 法的, <b>可能</b> , 性, <b>者</b>
依存 一致：6	<b>ルール</b> , <b>時間</b> , 利用, <b>制限</b> , 生活, <b>スマート</b> , <b>フォン</b> , 者, <b>保護</b> , 的	<b>時間</b> , <b>スマート</b> , <b>フォン</b> , 使用, <b>ルール</b> , <b>制限</b> , 重要, 的, 活動, 子供
交友関係 一致：4	<b>相手</b> , <b>情報</b> , 写真, 人, 今後, <b>個人</b> , 利用, <b>相談</b> , インターネット, 者	<b>相手</b> , 場合, <b>情報</b> , 重要, 性, <b>プライバシー</b> , <b>個人</b> , <b>相談</b> , 状況, 的
性的トラブル 一致：4	<b>相手</b> , <b>写真</b> , 児童, 警察, ポルノ, <b>相談</b> , <b>情報</b> , 性的, 画像, 行為	重要, 場合, <b>情報</b> , <b>相手</b> , 状況, <b>プライバシー</b> , <b>相談</b> , 的, <b>写真</b> , 性
削除方法 一致：3	<b>削除</b> , 投稿, 運営, <b>者</b> , 会社, 依頼, 本人, <b>情報</b> , サイト, 人	<b>削除</b> , 場合, <b>情報</b> , <b>プライバシー</b> , 連絡, 問題, アカウント, 的, <b>者</b> , 重要
料金関係 一致：3	<b>者</b> , 相談, 契約, 未成年, <b>会社</b> , <b>保護</b> , <b>ゲーム</b> , センター, 消費, 生活	課金, <b>会社</b> , 場合, 子供, 連絡, 問題, <b>ゲーム</b> , <b>者</b> , 重要, 確認
フィルタリング関係 一致：5	<b>制限</b> , <b>利用</b> , <b>ルール</b> , フィルタ, リング, 者, 保護, 機能, 子供, <b>設定</b>	<b>制限</b> , <b>利用</b> , <b>ルール</b> , 時間, <b>スマート</b> , <b>フォン</b> , 子供, アプリ, 使用, <b>設定</b>
取引トラブル 一致：2	<b>相手</b> , <b>アカウント</b> , 相談, ゲーム, 個人, トラブル, 契約, 利用, 話し合い, 行為	場合, <b>相手</b> , 連絡, 情報, 状況, 詐欺, 性, 報告, <b>アカウント</b> , 取引
情報セキュリティ 一致：4	<b>ウイルス</b> , 利用, <b>情報</b> , 者, 感染, 可能, 性, 電話, インターネット, <b>パスワード</b>	セキュリティ, <b>情報</b> , <b>パスワード</b> , 的, 場合, 性, 重要, <b>ウイルス</b> , 使用, ソフトウェア
個人情報関連 一致：3	<b>情報</b> , <b>相手</b> , <b>個人</b> , 特定, 写真, 自分, 削除, 対応, インターネット, 利用	<b>情報</b> , <b>個人</b> , 確認, <b>プライバシー</b> , 性, 場合, <b>相手</b> , 信頼, オンライン, 重要
その他 一致：3	<b>利用</b> , 者, <b>情報</b> , 相手, 保護, 自分, インターネット, 写真, <b>ルール</b> , <b>必要</b>	場合, <b>情報</b> , 重要, 性, 子供, 的, 可能, <b>必要</b> , <b>利用</b> , 使用

各カテゴリの回答の単語の取り出しでは、形態素解析機として Janome[9]を利用した。内包辞書はデフォルトの

mecab-ipadic-2.7.0-20070801 のままでの利用とした。そして、形態素解析の結果から、単語として名詞のみを抽出して、そこから意味をなさない名詞については除外を行った。

以上の手順により、その出現回数を集計したところ、カテゴリによって専門家の回答と生成 AI の回答で類似が見られるパターンと大きく異なるパターンに分かれる結果となった。類似するカテゴリとしては、「著作権関連」、「依存」、「フィルタリング関係」であり、これらのカテゴリについては、単語頻度上位 10 件のうち半数以上が同じ単語である。これらのカテゴリにおける回答の特徴としては、トラブルに対する対応方法が確立されている相談案件が多くインターネット上においても一般的な対応知識として掲載されていることも多いために、専門家の回答と生成 AI の回答があまり変わらないという状態になったと考えられる。

そして、「架空請求」、「取引トラブル」が一致する単語が少ないという傾向が見られた。「架空請求」に対しての専門家の回答は、悪質な業者への対応としては同意が不十分等のために契約が成立していないケースでは悪質業者に極力連絡をせずに無視するといった回答が多い。それに対して、生成 AI では悪質業者に連絡や確認を行って契約の解除等の相談をする方向に向けることが多かったために構成単語にも違いが生じてと思われる。

また、もう 1 個の一致単語の少なかったカテゴリである「取引トラブル」では個人間取引という概念について生成 AI の LLM にあまり情報がなく、企業と個人間における一般的な売買契約や個人間の通常の SNS のやり取りといったトラブル対応についての回答と考えられるものであった。以上のことから、これらの 2 個のカテゴリについては LLM の内容として不足している情報が原因ではないかと考えられるために、追加学習による精度向上が期待される。

### 2.3.3 回答品質の比較

次に専門家の回答と生成 AI の回答について、実際の相談内容に対する回答の品質を比較するために、両者の回答を目視により比較することで、以下の 4 パターンに分類を行った。

- A : 生成 AI の回答の方が詳細な説明もしくは参考になると思われるもの
- B : 生成 AI と専門家の回答が同程度に参考になると思われるもの
- C : 専門家の回答に比べ生成 AI の回答が抽象的等の理由により劣ると思われるもの
- D : 生成 AI の回答に大きな間違いもしくは不正確な内容が含まれるもの

本判定は各カテゴリの品質の確認を目的としているために、取得した 714 件の相談事例から、各カテゴリ 10 件（取引トラブルのみ 9 件）の 159 件を取り出して、目視確認による評価の対象とした。各カテゴリの回答比較

の結果を以下に示す。

表 3 回答品質の比較結果

カテゴリ	A	B	C	D
架空請求	0	1	6	3
ネットショッピング	0	5	4	1
ネットいじめ	0	7	2	1
迷惑メール	3	4	2	1
有害情報	1	8	0	1
著作権関連	2	7	1	0
依存	1	8	1	0
交友関係	3	5	1	1
性的トラブル	0	5	5	0
削除方法	2	6	2	0
料金関係	0	4	6	0
フィルタリング関係	2	6	2	0
取引トラブル	2	5	2	0
情報セキュリティ	2	5	3	0
個人情報関連	0	7	2	1
その他	2	4	4	0
合計	20	87	43	9
	12.6%	54.7%	27.0%	5.7%

全体の比較の分布としては、専門家と生成 AI が同程度の回答品質となるのが 54.7%と半分近い状態であった。さらに、生成 AI が優位なものが 12.6%、生成 AI の品質が低いものが 32.7%であった。この 32.7%が生成 AI 利用時の改善が可能な割合であるために、LLM のファイネーニングもしくは相談内容の構成単語のノイズ除去等の前処理を行うことで精度向上の余地があると考えられる。

さらに、カテゴリ別に回答品質を確認すると、相談カテゴリにより品質は大きく異なり、生成 AI が最も回答品質が低いのが「架空請求」、「ネットショッピング」、「料金関係」であった。これらについては、単語傾向の一致が少ないカテゴリでもあったことから LLM に追加学習が必要と予想される。また、「料金関係」においては、専門家の回答としては、契約の解除の可否を判断した上で、支払いや返金に関する内容を回答するという手順であったが、生成 AI の方では、主に返金や支払い停止についての回答が中心であり、相談対応としては不十分である印象が強かったために品質が低いと判断されている。

また、専門家よりも詳細な回答を返すカテゴリが「フィルタリング」や「削除方法」、「著作権関係」である。今回は件数が少ないためにあくまで参考程度であると考えられるが、学習の方向性としての参考情報として今後は件数を増やした調査を進めたいと考える。

### 2.3.4 特徴的な回答の考察

専門家の回答と生成 AI の回答は半数近くが同程度の回

答品質を示すこととなり、ファイネーニング前の LLM であってもある程度の相談内容に対応することは可能である。しかし、質問の内容によっては、専門家と生成 AI の回答が全く逆の方向を示すこともあったために、違いが大きかった 4 個の事例について考察を行う。最初の事例は、「架空請求」に関する事例で、その相談内容と両者の回答を以下に示す。

表 4 特徴的な回答の事例 1

Table. 4 First Case of a characteristic answer.

カテゴリ：架空請求，判定 D	
相談内容	子供がアダルトサイトで 18 歳以上かどうかの質問に「はい」をクリックしたところ、契約完了の画面が表示されたらしい。びっくりして契約取り消しのメールと電話をってしまったようだ。保護者と相談してからもう一度電話をしてほしいと言われたそうだが、どうしたら良いか。
専門家の回答	サイトにアクセスしたり年齢確認を押すだけでは、利用規約の同意や有料の契約に同意しているとは言えないため、契約は成立していないと考えられる。お金を請求されたとしても支払う必要はないと考えて、このまま無視をすると良い。しかし契約取り消しのためにメールや電話をしたことでメールアドレスや電話番号が相手に伝わり、今後は迷惑なメールや電話が増えるかもしれない。迷惑メール防止設定や電話の着信拒否などの対策をとってほしい。
生成系 AI の回答	このような状況で、以下の手順を踏むことをおすすめします。 1. 冷静に保護者と子供と話し合う：子供に起きたことや彼らの気持ちについて話し合い、保護者が全体の状況を把握できるようにしましょう。 2. 信頼できる成人と一緒に電話する：保護者や他の信頼できる成人と一緒に電話し、契約取り消しの手続きをするためのサポートを受けましょう。成人が子供をサポートすることで、適切な手続きを行うことができます。 3. サービスプロバイダーに連絡する：アダルトサイトのカスタマーサービスに連絡し、契約取り消しの手続きを依頼しましょう。問題を説明し、誤って契約を行ったことを伝えます。一般的に、カスタマーサービスはこのような問題に対して対応してくれます。 4. 書面で契約取り消しを要求する：電話での手続き後、契約取り消しの要求を書面で送付することも検討してください。メールやファックス、郵送などの方法で、契約取り消しの意思を明確に伝えることができます。書面での連絡は後の証拠となるため、保管しておくことをおすすめします。 5. 支払いに関する連絡を注意深く確認する：サイトからの請求書や支払い関連の連絡があった場合、保護者がそれらを注意深く確認しましょう。必要に応じて、不正な請求に関する警察への報告や法的な助言を求めるとも検討してください。 6. インターネットセキュリティについて教育する：この経験を通じて、子供にインターネットのリスクやセキュリティについて教育することが重要です。年齢制限があるコンテンツにはアクセスしないようにすることや、信頼できる成人と相談することの重要性を理解させましょう。 重要なことは、子供と保護者が連携し、問題を解決することです。保護者のサポートを受けながら、手続きを行い、必要な措置を講じてください。

事例 1 では、回答が大きく異なる理由として、専門家は相談内容から 18 歳未満の確認が不十分である等から契約が成立していない架空請求であるとの判断した上で、無視をする等の対応としている。しかし、生成 AI では事業者に向けて確認の連絡を入れるといった逆の対応が示されている。悪質な業者の場合は連絡をトリガーとして請求の連絡が増えること個人情報を知られることで被害が大きくな

るリスクがあると予想される。

表 6 特徴的な回答の事例 3

表 5 特徴的な回答の事例 2

Table. 5 Second Case of a characteristic answer.

カテゴリ：架空請求，判定 D	
相談内容	子供がパソコンで有料サイトに会員登録してしまった。請求画面がパソコンに表示されている。子供が慌てて、表示されている連絡先に電話をかけた。親から電話をするように言われたとのこと。相手には、氏名、住所などは知らせなかった。画面には金額、支払い期限、振込先口座が表示されているがどうしたらよいか？怖くてパソコンは触れていない。
専門家の回答	架空請求であり、支払いの必要はないことを説明し、保護者から電話をしないように助言した。相手から電話がくる可能性があるため、着信拒否をしようとよい。もしもかかってきても支払いには応じない。着信拒否しても別の番号から電話がくるのであれば、相手には警察へ相談しますと言おうとよい。また、振込先などの情報は、最寄りの警察署へ、不当請求業者の口座として情報提供できることを伝える。
生成系 AI の回答	このような状況では、以下の手順を踏むことをおすすめします。 1. 子供に落ち着いて待つように伝えましょう。子供が慌てずに待つことが重要です。 2. 親が電話をかけるようにしましょう。子供が連絡を取った相手には氏名や住所などの個人情報を伝えていないようですが、親が連絡を取ることで状況を把握できます。 3. 電話の相手には、子供が誤って会員登録してしまったことを伝えましょう。同時に、子供が未成年であり、誤って登録したことを強調します。 4. 金額や支払い期限、振込先口座などが表示されている場合は、子供の会員登録を取り消してもらうよう要求しましょう。未成年の子供が誤って登録した場合、一部のサイトでは取り消しの対応をしてもらえることがあります。 5. もしも相手を取り消しの対応をしてくれない場合や、詐欺的なサイトである可能性がある場合は、相手に個人情報を提供せずに、親が関係する団体や消費者機関に相談しましょう。彼らは適切なアドバイスを提供し、適切な対応をするための支援をしてくれるでしょう。 6. パソコンを触らずにしておくことも重要です。急いで問題を解決しようとするあまり、誤って追加の問題を引き起こしてしまう可能性があります。まずは専門家や関係機関の助けを借りることが重要です。 最終的には、親が適切な対応をすることが重要です。専門家や関係機関に相談し、子供の安全を確保するための適切な対策を講じましょう。

事例 2 は事例 1 と同じく「架空請求」に関しての類似する内容であるが、既に相談者が事業者に連絡済みである点が異なる。しかし、相談内容からやはり契約の成立していない架空請求であると判断可能であるために、専門家は事例 1 と同じく以後の連絡を行わず、業者からの連絡が続く場合には警察等への相談という回答になっている。これに対して、生成 AI では事例 1 と同じように業者に連絡して問題解決を図る対応が示されているが、架空請求業者であった場合には有効な手段とは言えない。

事例 1 や事例 2 で専門家と AI の判断が別れたポイントとしては、対象となる事業者が悪意のある架空請求業者であるという前提の有無が異なったためであると思われる。今回は明らかに架空請求業者であるために、生成 AI ではそういった事業者についての学習を追加で行うことで適切な回答が得られると考えられる。

Table. 6 Third Case of a characteristic answer.

カテゴリ：ショッピング，判定 D	
相談内容	掲示板でチケットを譲ってくださるという方とメールのやり取りをした。譲っていただくために個人情報を教えてしまった。個人情報を教えた途端連絡は途絶え、今は一切連絡が来ない。この個人情報は悪用されてしまうのか。お金は振り込んでいない。相手の個人情報は一切知らない。
専門家の回答	インターネット上の見知らぬ相手との個人間の取引には大きなリスクが伴う。これをしておけば 100%安全という対策はない。そして、相手に渡してしまった情報を取り戻すことは大変難しい。どのように使われるかは相手次第であり、その情報の利用を止めることも難しい。自分の手を離れてしまった以上は、自分でコントロールできるものではない。今後は、メールアドレスを教えたことで、迷惑メールが届く可能性もあるので、迷惑メール対策の設定を行うと良い。相手に住所を伝えているならば、保護者に伝えておいた方がよい。念のために、相手に伝えた情報、相手のことで知っている情報を整理しておく。相手とのやり取りのメールが残っているならば残しておく。
生成系 AI の回答	個人情報を知らせてしまった場合、悪用される可能性があるため、注意が必要です。ただし、具体的な状況によって異なる場合もありますので、以下の点に留意してください。 1. 悪意のある行動：相手が個人情報を悪用する可能性があります。クレジットカード情報や銀行口座情報など、金銭的な被害を受ける可能性もあります。この場合は、速やかに関連する金融機関に連絡し、対処する必要があります。 2. 個人情報の種類：共有した個人情報の内容によって、被害の程度が異なります。氏名や住所、電話番号、メールアドレスなど、個人を特定できる情報を知らせた場合、プライバシーに関するリスクが高まります。注意深く監視し、不審なアクティビティがないか確認してください。 3. 警察への連絡：状況が深刻で、不正行為の可能性のある場合は、警察に連絡し、詳細を報告することが適切です。警察は、不正行為を追及し、必要な対策を講じることができま。す。 4. 防衛策の強化：今後同様の被害を未然に防ぐために、オンラインセキュリティの対策を強化しましょう。パスワードの変更、二要素認証の有効化、信頼できるセキュリティソフトの使用などが有効な対策となります。 個人情報の悪用に関しては、具体的な状況により異なるため、法的な助言を求めることも考慮してください。詳細な情報を提供することで、専門家がより具体的なアドバイスを提供できるでしょう。

事例 3 は「取引トラブル」に関するものであり、個人間取引に関する相談内容である。そして、専門家による対応としては、個人間取引のリスク等の説明となっているが、生成 AI では個人情報の流出リスクが主題となっており、適切な回答とは言えなかった。この要因としては、相談内容から個人間取引という要素を判断することができなかつたために、LLM の追加学習により精度が向上する可能性がある。そして、質問文を構成する単語のうち、個人情報の流出の分類の方に導く特徴語は判定制度を下げる要因となる単語である可能性もあるために、こういった単語をノイズとして除去することで有効な精度向上に繋がるものについては、前処理としてノイズを除去することで精度の向上を見込むことができる可能性がある。

表 7 特徴的な回答の事例 4

Table. 7 Fourth Case of a characteristic answer.

カテゴリ：迷惑メール，判定 A	
相談内容	本人不在のため荷物を持ち帰ったという SMS (ショートメッセージサービス) が届き、メッセージ内に書かれていた URL を押してしまった。しかしフィルタリングでブロックされてサイトは開けなかった。あとでインターネットで調べて詐欺らしいとわかったが、サイトを開けなかったのが大丈夫か。
専門家の回答	宅配業者を装った偽の SMS の可能性が考えられる。書かれている URL にアクセスするとスマートフォンの中にある情報が抜き取られるなどトラブルに繋がる場合もあるのでクリックしてはいけない。もしもスマートフォンに見覚えのないアプリがインストールされていたら削除することを勧める。今後は、送り主がわからない SMS には返信せず、書かれている URL にもアクセスしないように注意してほしい。SMS を受信しないように設定するなどの対処法もあるので検討すると良い。
生成系 AI の回答	おそらく、URL をクリックしてもフィルタリングによってサイトがブロックされたため、問題はありません。フィルタリングが機能している場合、危険なウェブサイトへのアクセスを防止するため、そのようなサイトにアクセスすることはできません。ただし、セキュリティのためにいくつかの注意事項があります。次の点に留意してください： 1. フィルタリングが有効であっても、一部のウイルスやマルウェアはフィルタリングを回避する手法を持っている場合があります。したがって、アンチウイルスソフトウェアを最新の状態に保ち、システムをスキャンしてください。 2. フィッシング詐欺など、URL をクリックするだけでなく、個人情報を入力するよう要求する場合があります。したがって、自分の個人情報やアカウント情報を提供することは避けてください。 3. もし心配な場合は、疑わしいメッセージやサイトの情報をインターネット上で検索して、その詐欺やリスクについての情報を調べることをお勧めします。 最終的に、フィルタリングが機能していてサイトにアクセスできなかった場合、おそらくセキュリティ上のリスクは低いと言えます。ただし、注意深く行動し、必要に応じてセキュリティ対策を行うことをお勧めします。

事例 4 は生成 AI の回答品質が高いパターンである。この要因としては、相談事業に対応する専門家よりもソフトウェアの挙動の説明等では生成 AI の方が正確な回答をする傾向がある点や生成 AI はそういった内容については説明も丁寧である点が要因であると考えられる。これは LLM に対応する知識が学習済みであったからであり、情報関連の内容については一般的な LLM であっても十分な学習がなされているためと考えられる。

### 3. おわりに

本研究では、地方自治体における青少年向けの SNS トラブルを対象とした相談窓口において生成 AI の利用した際に発生する課題を分析するために、東京都の相談事例を参考として、生成 AI による回答と実際の相談例として掲載されている専門家の回答を比較した。比較した内容としては、回答内容における単語の出現傾向によって LLM が相談内容の各カテゴリの学習内容に不足があるかの確認を行い、生成 AI がファインチューニング前の一般的な LLM を利用した際の得意分野と苦手分野の確認を行った。その結

果として、「架空請求」や「取引トラブル」では相談内容から違うカテゴリのトラブル対応と誤認識することで、対応方法が適切ではないという事例が確認された。また、「迷惑メール」といった一般に多く見られるトラブル事例では AI の回答が高い品質を示した。

さらに、専門家の回答と生成 AI の回答が大きく異なるケースについて確認することで、生成 AI では回答が難しいケースの抽出や生成 AI が得意な相談内容についても確認した。本研究で得られた知見を活用することにより、青少年のネットトラブル相談事業で生成 AI を利用する際の精度を向上することが期待される。

今後は本研究で得られた知見を活かし、実際にファインチューニングを行うことで、当該問題に生成 AI を利用した場合に、どの程度の精度が向上するかという確認を行う予定である。また、ファインチューニング以外で回答精度を向上するために、相談内容からノイズとなる単語を前処理として除去することで、より良い回答を得られるかを検証するアプローチについても有効性を確認したいと考える。

**謝辞** 本研究は JSPS 科研費 21K02411 および文教大学共同研究費の助成を受けたものです。

### 参考文献

- [1] 総務省，青少年のインターネット利用環境実態調査，[https://www8.cao.go.jp/youth/kankyouto/internet\\_torikumi/tyousa.html](https://www8.cao.go.jp/youth/kankyouto/internet_torikumi/tyousa.html), (参照 2023-6-14).
- [2] 青少年が安全に安心してインターネットを利用できる環境の整備等に関する法律，<https://elaws.e-gov.go.jp/document?lawid=420AC1000000079>, (参照 2023-6-14).
- [3] 内閣府，青少年が安全に安心してインターネットを利用できるようにするための施策に関する基本的な計画（第 5 次），[https://www8.cao.go.jp/youth/youth-harm/suisin/pdf/dai5ji\\_keikaku.pdf](https://www8.cao.go.jp/youth/youth-harm/suisin/pdf/dai5ji_keikaku.pdf), (参照 2023-6-14).
- [4] こども基本法，<https://elaws.e-gov.go.jp/document?lawid=504AC1000000077>, (参照 2023-6-14).
- [5] インターネット・ホットラインセンター，運用ガイドライン，[https://www.internethotline.jp/pdf/guideline/guideline\\_rev15.pdf](https://www.internethotline.jp/pdf/guideline/guideline_rev15.pdf), (参照 2023-6-14).
- [6] 東京都，こたエール，<https://www.tokyohelpdesk.metro.tokyo.lg.jp/>, (参照 2023-6-14).
- [7] Leonard Richardson, BeautifulSoup, <https://www.crummy.com/software/BeautifulSoup/>, (参照 2023-6-14).
- [8] OpenAI, ChatGPT, <https://chat.openai.com/>, (参照 2023-6-14).
- [9] Tomoko Uchida, Janome, <https://mocobeta.github.io/janome/>, (参照 2023-6-14).