

研究速報

## ChatGPTを活用した研究データの作成事例

### A Case Study of Research Data Preparation Using ChatGPT

新原俊樹<sup>1\*</sup>, 甲斐尚人<sup>2</sup>, 小柏香穂理<sup>3</sup>, 船越幸夫<sup>4</sup>

Toshiki SHIMBARU<sup>1\*</sup>, Naoto KAI<sup>2</sup>, Kahori OGASHIWA<sup>3</sup>, Sachio FUNAKOSHI<sup>4</sup>

1 西南学院大学情報処理センター

Information Processing Center, Seinan Gakuin University

〒814-8511 福岡県福岡市早良区西新 6-2-92

E-mail: shimbaru@seinan-gu.ac.jp

2 大阪大学附属図書館

University Library (Research and Development Unit), Osaka University

〒560-0043 大阪府豊中市待兼山町 1-4

3 お茶の水女子大学教学 IR・教育開発・学修支援センター

Center for Institutional Research, Educational Development, and Learning Support, Ochanomizu University

〒112-8610 東京都文京区大塚 2-1-1

4 記録管理学会事務局

The Records Management Society of Japan

〒267-0066 千葉県千葉市緑区あすみが丘 5-54-8

\*連絡先著者 Corresponding Author

研究活動におけるChatGPTの有効な活用方法として、比較的短い文章の集合から研究データを作成する方法を提案した。事例研究として、4つの学会が発行する学術雑誌に掲載された239編の論文のタイトルを対象とし、各論文のタイトルから内容を推定するためのプロンプトをChatGPTに与え、得られた回答を整理した。ただし、研究データには高い再現性が求められることから、同一のプロンプトをChatGPTに10回繰り返し与え、回答を積算して集計することで誤判定の影響を低減させる工夫を施した。その結果、同じ手法で作成した2つの研究データを比較すると、データ全体の97%が同じ結果となり、高い再現性を確認することができた。一方、ChatGPTに与えるプロンプトの僅かな表現の違いに応じて、得られる回答も変わり得ることが明らかになった。ChatGPTから目的に相応しい研究データを作成するための最適なプロンプトの表現について、さらに検証する必要がある。

As an effective use of ChatGPT in research activities, the authors proposed a method to prepare research data from a set of relatively short sentences. The titles of 239 articles published in four journals were used as a case study. Prompts were given to ChatGPT to infer the content from each title, and the

obtained responses were organized. Therefore, the authors reduced the effect of erroneous judgments by giving the same prompt ten times to ChatGPT and summing up the answers. As a result, when comparing two data sets generated by the same method, 97% of the data sets showed the same results, indicating high data reproducibility. However, depending on the slight difference in the wording of the prompt given to ChatGPT, the response obtained could also change. We need additional thought on the optimal wording of the prompts to get a correct response from ChatGPT.

キーワード: 対話型生成AI, ChatGPT, プロンプト, 研究データ作成

Keywords : Interactive Generative AI, ChatGPT, Prompt, Research Data Preparation

## 1 はじめに

ChatGPT[1]は OpenAI 社が開発した大規模言語モデルに基づく対話型生成 AI(以下, 生成 AI)のサービスであり, 2022 年 11 月の公開以来, 大きな注目を集めている. この生成 AI を活用することで, 誰でも手軽に大規模言語モデルを用いて自然言語処理の手法で問題の解決に取り組むことが可能になった. 文章の生成や校正, 翻訳, プログラミングをはじめ, 生成 AI の具体的な活用方法については数多くの提案がなされている. 研究活動の現場でも生成 AI の活用は広がっていくと期待される. 特に, 言語モデルである ChatGPT は自然な文章を生成することに強みを持つため, 論文執筆の支援機能としての役割への期待は大きい. しかし, それ以外の場面で生成 AI がどのように貢献し得るのか, 具体的な提案はまだなされていない.

こうした中で, 本研究は, 研究活動における ChatGPT の有効な活用方法として, 比較的短い文章(以下, 短文)の集合を研究対象とする場合に, これらの短文の生データから解析用の研究データを作成する工程での活用について提案したい. ここで, 研究データとは, 短文別に, それらが含む意味や内容, 特徴の有無を表形式で整理したものを想定している.

この工程を既存の自然言語処理の手法を用いて行う場合, 例えば, Sentence BERT[2]などの手法が考えられるが, 本研究ではこれらの手法の代わりに GPT-4[3]を搭載した ChatGPT を活用して研究データの作成を試みた. プログラミングによらず手軽に扱える手法で実現することができれば, 研究作業の効率化に資するだけでなく, 自然言語処理の手法を用いた研究分野の裾野の拡大にも貢献し得る.

ただし, このとき, ChatGPT から得た回答の再現性の確保に注意しなければならない. 生データに基づき研究データを生産させるための指示(以下, プロンプト)が同じでも, ChatGPT から得られる回答は毎回異なり得る. 信用される研究成果を得るためには, 第三者の追試でも同じ結果になるように再現性の高い研究データが得られなければならない. そこで, 本研究の提案手法で得られる研究データの再現性についても, 併せて検証した.

## 2 生データ(論文タイトル)の収集

本研究では, 研究データを作成するための生データとして, 学術雑誌に掲載された多数の論文タイトルを用いることとした. 記録管理やアーカイブズ分野を研究分野とする 4 つの学会を対象として, 各学会が概ね過去 20

表 1 各学会誌の論文タイトルの収集状況

	学会名(学会誌名)	設立年	巻・号(期間)	論文数
A	記録管理学会 (レコード・マネジメント)	1989	48～84号[4] (2004年10月～2023年3月)	99
B	アート・ドキュメンテーション学会 (アート・ドキュメンテーション研究)	1989	9～30号[5] (2001年3月～2022年5月)	80
C	日本アーカイブズ学会 (アーカイブズ学研究)	2004	3～36号[6] (2005年11月～2022年6月)	39
D	デジタルアーカイブズ学会 (デジタルアーカイブ学会誌)	2017	2～7号[7] (2018年1月～2023年3月)	21

年間に発行した学会誌に掲載された合計 239 編の論文のタイトルを収集した。各学会誌の論文タイトルの収集状況を表 1 に示す。

ここで収集の対象としたのは、査読付の論文(及び、これに準じた査読付の研究ノートなど)とし、全国大会や研究会の予稿については対象外とした。各論文のタイトルから内容を推定して研究データとして整理し、各学会が掲載する論文の内容にどのような違いがあるか検証する一連の工程の中で、ChatGPT を活用して研究データを作成する。

### 3 ChatGPTに論文タイトルから内容を判断させる手法

収集した論文のタイトルを ChatGPT に提示し、それぞれの内容を推定させた。このとき、ChatGPT に与えたプロンプトと、ChatGPT から得た回答を図 1 に示す。プロンプトの主要部は冒頭の 2 行であり、論文タイトルの一覧と A~L の 12 種類の研究内容(以下、「観点」)を特徴づける複数の単語を提示し、個々の論文がそれぞれの単語に関係する(1)か否か(0)、判定した結果を表形式で回答させている。このプロンプトを受けた ChatGPT は、単にタイトル内に指定された単語が含まれるか否かに留まらず、単語の意味も踏まえて両者を比較する。これにより、例えば、「大学」の語を含むタ

#### (ChatGPT に与えたプロンプト)

下の方の番号が付いた文章のタイトルを見て、次の A~L の各単語に 1 つ以上関係する場合は 1、不明な場合は 0 として、表の形式で回答してください。各番号のタイトルの内容は独立していますので、1 つずつ、考え直してください。表の左に番号を付けてください。タイトルは不要です。L 列の回答が空白にならないように注意してください。

A)近世,近代,戦前,戦中,戦後  
 B)電子,デジタル,情報,データ  
 C)政府,行政機関,自治体,行政文書,公文書,公文書館  
 D)ビジネス,企業,民間  
 E)写真,画像,映像,動画  
 F)アーカイブ  
 G)記録管理,資料整理,文書分類,ファイリングシステム  
 H)事例,調査  
 I)教育,学校  
 J)外国,国際標準  
 K)図書館,博物館,文化財  
 L)美術館,絵画,芸術,アート

— —

1,国立大学法人における卒業論文の扱い—著作物でもある法人文書等の公開(利用)について  
 2,技術・技能保存継承ツールとしての発電所建設記録映像のデジタルアーカイブズ

#### (ChatGPT から得た回答)

No.	A	B	C	D	E	F	G	H	I	J	K	L
1	0	0	1	0	0	0	0	0	1	0	0	0
2	0	1	0	0	1	1	0	0	0	0	0	0
3	0	0	1	0	0	0	1	0	0	1	0	0
4	0	1	0	0	0	1	0	1	0	0	1	0
5	0	0	0	0	0	1	0	1	1	0	0	0
6	0	0	0	0	0	0	1	0	0	0	0	0
7	0	0	1	1	0	0	1	1	0	1	0	0
8	0	0	0	1	1	0	0	0	0	0	0	0
9	0	0	0	0	0	1	0	0	1	0	0	0
10	0	0	0	0	0	0	1	0	0	1	0	0
11	0	0	0	0	1	0	0	0	0	0	1	0
12	0	1	0	0	0	1	0	0	0	1	0	0
13	0	1	0	0	0	1	0	0	0	0	1	0
14	0	0	0	0	0	1	0	0	1	0	0	1
15	0	0	0	0	0	1	0	0	1	0	0	0

図 1 ChatGPT に与えたプロンプト(上), ChatGPT から得た回答(下)

イトルは観点 I(教育,学校)に,「アメリカ」「韓国」などの語を含むタイトルは観点 J(外国,国際標準)に関係すると判定されるようになる。

#### 4 回答の再現性を高める工夫

ただし,全く同一のプロンプトを繰り返し与えても ChatGPT の回答は所々で判定が異なり,毎回厳密に同じ結果にはならない。研究データの再現性を確保できなければ,その解析結果も信頼が得られない。API 経由で直接 GPT-4 を動作させる場合は,パラメータを設定することで回答の再現性を高める工夫も可能である。しかし,ChatGPT ではこれらのパラメータの設定ができない。そこで,ChatGPT から得られる回答の再現性を高めるため,同じプロンプトを 10 回繰り返し与え,「ChatGPT に関係あり(1)と判定された回数」を集計し直した。なお,同じプロンプトを繰り返し与える際,各回の回答がそれ以前の回答の影響を受けないように,プロンプトの入力履歴を ChatGPT に学習させない設定とした。

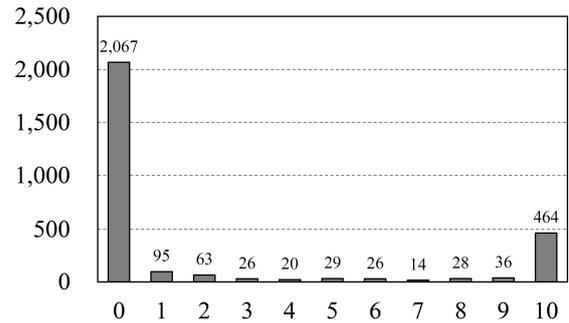


図 2 研究データの各要素の値別の度数分布

得られた回答の全要素(239 編×12 観点=2,868 要素)について,回数別の度数分布を図 2 に示すとともに,観点 A(近世,近代,戦前,戦中,戦後)に関して 1~9 回となった論文タイトルの一覧を表 2 に整理した。また,表 2 の各タイトルを確認し,観点 A に関係するとみられる単語に下線を引いた。

図 2 から,0 回や 10 回となった要素,すなわち,全ての試行で同じ判定結果になった要素は 2,531 に上り,全体の 88%を占めている。また,表 2 を見ると,8 回や 9 回となった論文タイトルには,いずれも観点 A に関する単

表 2 観点 A に「関係あり(1)」と判定された回数が 1~9 回の論文タイトル一覧

論文タイトル	観点A
対馬藩の御内書、老中奉書の選別— <u>18世紀後期</u> における文書管理の転換	9
<u>明治期</u> の書籍における「題辞」の様態—美術書を事例として—	9
<u>戦後</u> 日本企業の記録管理におけるアーカイブ機能の形成に関する一考察—「山一証券資料」を素材にして—	8
『 <u>日本占領下</u> のマラヤ1941—1945』におけるアーカイブの利用	8
ドイツと <u>明治政府</u> の記録管理:レジストラトゥーラを伝えたお雇い外国人	7
「デジタル化」とアーカイブズ—“正倉院文書データベース”と近代史料のデジタル化を通して—	7
<u>19世紀</u> フィンランドにおける情報アクセスのための実務—アルヴィドソンら「アーキビスト」による成果について—	6
米国におけるバイタル記録管理の歴史的展開:NAREMCOを中心に	2
<u>浅草文庫旧蔵「蘭書」</u> の行方—記録・史料による追跡	2
国立大学法人における卒業論文の扱い—著作物でもある法人文書等の公開(利用)について	2
東日本大震災の事例から見えてくる震災アーカイブの現状と課題	2
情報公開法施行前の国立大学における文書管理規程と文書移管—東京大学を事例に—	1
歴史が行う記録管理—東日本大震災からの教訓—	1
神奈川県立近代美術館美術情報システムの構築—事例報告	1
東日本大震災に関する記録・証言などの収集活動の現状と課題	1
銀座・三原橋地下街と都市の輪郭—オーラル・ヒストリーと資料からの考察—	1
ボズナーの『日本のアーカイブズ施設とレコードの集積の仮リスト』—「 <u>ウォーナーのリスト</u> 」との関係—	1
身装画像データベースのデータソースとしての新聞連載小説挿絵— <u>明治中期</u> を対象として	1

語が含まれている一方、1回や2回のタイトルの多くは観点 A との関係が見出しづらい。

そこで、あらためて、8回以上であれば「関係あり(1)」、2回以下であれば「関係なし(0)」, その間の 3~7回 は「不明(0.5)」とし、これを研究データとした。

この研究データを同じ方法で2回作成したところ、2,868要素のうち2,788要素(全体の97%)が同じ結果となった。また、異なる結果となった80要素についても、関係なし(0)→関係あり(1), 又は、関係あり(1)→関係なし(0)に変化した要素は全くなかった。このように、上記の手法で作成した研究データは再現性の高いものになっていることが確認できた。

## 5 まとめと今後の課題

研究活動の場面で生成 AI を活用するための一事例として、生データである多数の短文から解析用の研究データを作成する作業を ChatGPT に行わせるプロンプトを提案した。この研究データには高い再現性が求められることから、同一のプロンプトを10回繰り返し与え、判定結果を積算することで誤判定の影響を低減させ、再現性を高める工夫を行った。その結果、同じ手法で作成した2つの研究データを比較すると、データ全体の97%が同じ結果となり、高い再現性を確認することができた。本研究が提案した手法は、プログラミングの必要がなく手軽に実践できるため、研究作業の効率化に資するほか、自然言語処理の手法を用いた研究活動の活性化にも貢献できる。

一方、ChatGPT に与えるプロンプトの表現を僅かに変化させると、得られる結果が変わる場合がある。特に、各観点を特徴づける単語の順序には留意が必要である。例えば、観点 I では、「教育,学校」の順に並べた場合と「学

校,教育」の順に並べた場合で判定が異なるケースが増えることが確認されている。具体的には、「教育,学校」のように大きな概念を表す単語を先に置くほど「関係あり(1)」と判定されるタイトルの数が増える傾向にある。的確な判定を得るための最適な単語の配置について、より詳細に検証する必要がある。

本研究が提案する研究データの作成手法は、論文タイトルに限らず、例えば、社会調査で実施されるアンケートに対する自由記述の回答文など、比較的短い文章を研究対象としたものであれば幅広く活用できる可能性がある。今後、具体的な事例で検証を重ね、提案手法の汎用性を確認していきたい。

## 6 データ利用可能性宣言

本研究中に作成された、分析に用いたデータセットは、妥当な理由があれば責任著者から入手可能である。

## 参考文献

- [1] OpenAI: “Introducing ChatGPT,” <https://openai.com/blog/chatgpt> (2023年8月21日参照).
- [2] Nils, Reimers; Iryna, Gurevych: “Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks,” arXiv, 1908.10084, 2019.
- [3] OpenAI: “GPT-4 Technical Report,” arXiv, 2303.08774, 2023.
- [4] 記録管理学会: 「RM レコード・マネジメント コンテンツ一覧(34号〜)」, <https://www.rmsj.jp/recordsmanagement-bn/> (2023年8月21日参照).
- [5] アート・ドキュメンテーション学会: 「アート・ドキュメンテーション研究」, <http://www.jads.or>

g/pub/kenkyu.html(2023年8月21日参照).

[6] 日本アーカイブズ学会:「アーカイブズ学研究」, <http://www.jsas.info/?cat=7>(2023年8月21日参照).

[7] 国立研究開発法人科学技術振興機構:

「J-STAGE デジタルアーカイブズ学会誌」, <https://www.jstage.jst.go.jp/browse/jsda/-char/ja>(2023年8月21日参照).

(2023年9月4日 受付)

(2023年9月26日 採択)

(2023年11月30日 J-STAGE早期公開)