

グラフィカルアブストラクト生成のプロンプトエンジニアリングに関する考察 Considerations on Prompt Engineering for Generating Graphical Abstracts

大橋 美紗希[†] 岡 愛美[‡] 小澤 賢司[†]
Misaki Ohashi Manami Oka Kenji Ozawa

1. はじめに

グラフィカルアブストラクト (GA: Graphical Abstract) は、学術論文の要点を視覚的に表現し、一目で内容を簡潔に伝えることを目的とした図解である。山梨大学は2024年度に文部科学省によるオープンアクセス加速化事業に採択され、「真の意味での万人への情報提供の実現」を取組の1つの柱としている [1]。論文のオープンアクセスにあたり、GAを付与することは一般読者への訴求に有効であると考えられる。しかし、そのためだけにプロのイラストレータに外注することは、時間・費用の観点から憚れる。そこで、論文の Abstract から生成 AI により GA を自動生成するプロンプトエンジニアリングに取り組んだ。

2. 実験方法

2.1 プロンプトの設計と GA の生成

プロンプトとして、以下の3種を準備した (詳細は[2])。Pro ① 論文の Abstract の内容を表す画像を生成させる。Pro ② 論文の Abstract を約 20 単語で要約させた後で、その内容を表す画像を生成させる。Pro ③ 論文の Abstract を理解させ、細かな要件を指定して、その内容を表す画像を生成させる。

これらのプロンプトを利用して、GA を生成した。生成 AI としては Copilot (無償版) を用い、2025 年 1 月に実施した。文献 [3] について実際に生成した GA を図 1 に示す。

2.2 評価方法

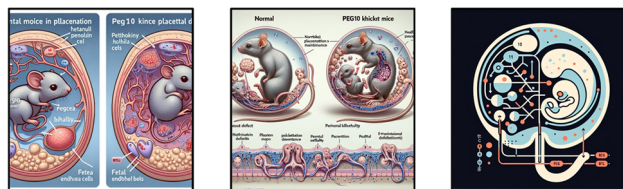
生成 AI により生成した GA の有効性を検討するため、以下の2つの観点で評価を行った。

- (1) **内容推測度**: 論文の書誌情報のみ/あわせて GA を見た後で、論文内容を推測可能と感ずる程度
- (2) **関心喚起度**: GA を見ることで、その論文を読みたいという気持ちになる程度

被験者は、山梨大学工学部コンピュータ理工 (CS) 学科の卒研究生 (5 名) と同コースの修士院生 (5 名) の計 10 名である。4 名の研究者について各 5 編の論文を対象としたが、うち 2 名は CS 学科の教員であり、残り 2 名は他学科 (工学部電気電子工学科と生命環境学部生命工学科) の教員である。1 編の論文に対する、具体的な評価手順は以下のとおりである。

- 書誌情報のみを見て、内容推測度を 5 点満点で評価
- 横並びで提示される 3 種の GA を見て、それぞれについて関心喚起度を 5 点満点で評価
- 改めて、内容推測度を 5 点満点で評価

以上をランダム順 (被験者間では同一順) で全ての論文に対して行った。ただし、評価の安定性を観測するため、最



(a) Pro ① (b) Pro ② (c) Pro ③

図 1 文献 [3] について 3 種のプロンプトにより生成したグラフィカルアブストラクト

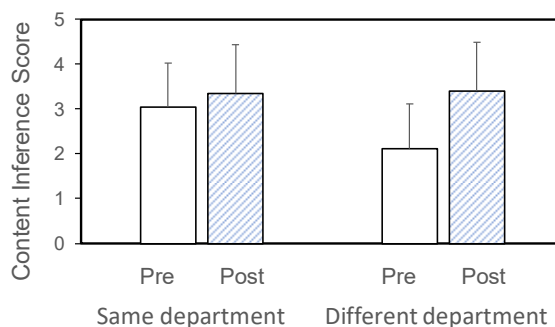


図 2 著者の所属に依存した GA 閲覧前後での内容推測度の平均値 (誤差棒: 10 論文間の標準偏差)

初に評価対象となった論文を、最後に改めて評価対象とした。それゆえ、被験者は 21 編の論文について評価を行った。

なお、実験は、被験者に PowerPoint のファイルを配付し、各人が日頃から利用しているディスプレイを用いて実施した。スライド 1 枚あたり 1 編の論文について、まず書誌情報のみが提示され、(a) の評価を行った。次に、アニメーションで 3 枚の GA が同時に出現し、被験者は各 GA について (b) の評価を行った (3 種のプロンプトに関する並び順は論文ごとにランダムである)。最後に、(c) を評価した。

3. 実験結果

3.1 評価の安定性

前節に述べたとおり、同じ論文を最初と最後に 2 回評価した。分散分析の結果、両評価値に有意差はなかったため、実験実施中に評価の基準が変わることはなかったと考えた。そこで、最初の評価は練習と見なし、最後の評価を採択して、20 論文を対象とした評価値を分析することとした。

3.2 論文内容の推測しやすさ

図 2 に結果を示すとおり、GA を見る前後での論文内容の推測度の変化は、著者属性 (自学科/他学科) により傾向が異なる。すなわち、「自学科の教員」による論文についての変化は小さいが、「他学科の教員」の論文については大きな変化が見られる。2 要因分散分析 (被験者内計画)

[†] 山梨大学研究推進・社会連携機構 Org. for Promotion of Research and Social Collaboration, Univ. of Yamanashi

[‡] IT フリーランス IT Freelance

表 1 内容推測度に関する分散分析表

要因	SS	df	MS	F	p
個人差(S)	38.67	9	4.32		
GA の閲覧前後(A)	61.62	1	61.62	212.48	***
S × A	18.85	9	2.09		
論文(B)	166.25	19	8.75	30.18	***
S × B	149.08	171	0.87		
A × B	30.13	19	1.59	5.48	***
S × A × B	40.90	171	0.29		

*** $p < .001$

[4] の結果を表 1 に示すとおり、GA 閲覧の効果は論文と交互作用がある。下位検定の結果、GA 閲覧の単純主効果は、「自学科」については有意な論文は無く、「他学科」では論文 10 編中 9 編で有意であった。

これは、自学科の教員の研究については、論文題目・雑誌名という書誌情報のみである程度まで論文内容を推測できたことを示唆する。逆に、題目・雑誌名だけでは内容推測が困難な「読者にとって馴染みのない分野の研究論文」については、GA を添えることの効果があったといえる。

3.3 論文への関心の喚起

全論文について、3 種のプロンプトに対する関心喚起度の平均値を図 3 に、分散分析表を表 2 に示す。図 3 からは Pro ② が若干優れているが、論文×プロンプトの交互作用が有意であるため、論文ごとに単純主効果を検討した。その結果、20 論文中の 12 論文(自学科: 6, 他学科: 6)でプロンプトの効果がありであった。それらの 12 論文について LSD 多重比較検定を行ったところ、例えば図 1 に示した 3 枚の GA については、Pro ① = Pro ② > Pro ③ という有意性が見られた。

プロンプトの優劣を考えるため、有意性の星取を計測することとした。図 1 の場合には、Pro ① > Pro ③ かつ Pro ② > Pro ③ なので、Pro ① と ② は各 1 勝、Pro ③ は 2 敗とカウントする。この要領で、12 論文について計測した結果を表 3 に示す。この結果から、Pro ② が負けにくいこと、つまり無難な GA を生成する可能性が高いことが示唆された。

4. 考察

被験者の選定にあたり、本報告の第 3 著者が CS 学科教員を兼務しているため、周囲の学生に依頼しやすかったという理由で当該学科の学生に偏った。しかし、その副産物として、馴染みのない分野の論文については、GA の提示が内容推測度の向上に寄与することを示すこととなった。これは、プロジェクトとしての目標である「真の意味での万人への情報提供の実現」、すなわち専門家ではない方に学術論文への興味を喚起することに、AI で生成した GA が寄与することを示唆している。

一方で、ここで取り上げた 3 種のプロンプトでは、20 論文中の 8 論文については関心喚起度に有意差が生じなかった。また、単純主効果のあった 12 論文についても、必ず Pro ② が優れているというほど強い傾向を示すことはできなかった。プロンプトエンジニアリングとしては、さらなる検討の余地があるものと考えている。

ところで、対象論文の著者に、AI が生成した GA の評価を求めたところ、好意的なコメントが得られる場合が多か

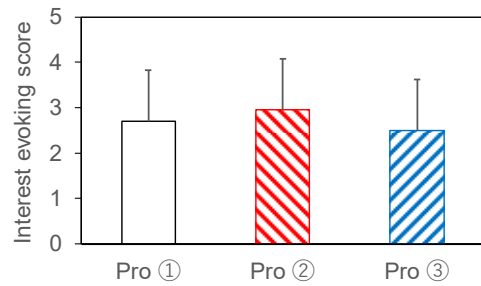


図 3 プロンプトによる関心喚起度の平均値
(誤差棒: 20 論文間の標準偏差)

表 2 関心喚起度に関する分散分析表

要因	SS	df	MS	F	p
個人差(S)	93.53	9	10.40		
プロンプト(Pro)	20.33	2	10.17	12.58	***
S × Pro	21.97	18	1.23		
論文(B)	127.43	19	6.71	8.31	***
S × B	150.20	171	0.88		
Pro × B	172.07	38	4.53	5.61	***
S × Pro × B	276.30	342	0.81		

*** $p < .001$

表 3 プロンプト間の星取の集計表

	勝	負	得点
Pro①	8	-5	3
Pro②	8	-1	7
Pro③	2	-12	-10

った。しかし、論文によっては「意味不明です」という評価もあった。また、「図中に謎のアルファベット文字列が出現するのは不適切に感ずる」という感想もあった。これらの意味においても、さらに GA 生成に適切なプロンプトを検討することが必要である。

5. おわりに

本研究では、AI による GA の自動生成を目指した。その結果として、Abstract を 20 語に要約した後に GA を生成させることで無難な GA が生成されることが示された。しかし、プロのイラストレータの作品には比べるまでもない品質である。今後は、「正確な情報—おおまかな印象」と「対象はナイーブ層—専門家」という 2 次元的な調整までできるプロンプトの開発を目指したい。

参考文献

- [1] 山梨大学オープンアクセス加速化事業成果報告 <https://www.yamanashi.ac.jp/social/49033, <2025/05/21>>.
- [2] 研究のポイントがわかる! グラフィカルアブストラクトとは? https://www.scrs.yamanashi.ac.jp/research/how_to_make_graphical_abstract/, <2025/05/21>.
- [3] H. Shiura, et al., "PEG10 viral aspartic protease domain is essential for the maintenance of fetal capillary structure in the mouse placenta," *Development* **148**, dev199564 (2021).
- [4] 田中敏, 山際勇一朗, ユーザのための教育・心理統計と実験計画法, 教育出版, pp. 105-127 (1992).