

プロンプトエンジニアリングの質的評価 ～マーケティングメールにおける多様なプロンプトの比較～

吉田悠人（早稲田大学）† 尾形哲也（早稲田大学、産総研）‡

1. はじめに

現在ではビジネス環境に ChatGPT に代表される LLM (Large Language Model) が多く活用されているが、文章生成機能を利用する際、適切なプロンプトを選定することが課題となっている。これまで多様なプロンプトが提案されているが、それらの提案について具体的な評価や比較が不足している。

本研究の主な目的は、ChatGPT を用いたマーケティングメールのプロンプトに注目し、その評価を行うことで設計の指針を与えることである。ChatGPT による生成結果がどのような要素に優れ、どのような要素が改善の余地があるかを明らかにすることで、ChatGPT をより効果的に活用し、洗練されたマーケティングメールを構築するための具体的な手法や指針を得ることを目指す。

2. 先行研究

メールマーケティングへの ChatGPT への応用に関しては、これまで多様なプロンプト設計が提案されている。適切なプロンプトの使用は、メールの効果を向上させるだけでなく、受信者の注意を引き付け、興味を喚起することが可能となる。既存の研究では、プロンプトの言語やデザインが開封率やクリック率に与える影響についての定量的な調査が行われており、これらの研究結果は企業におけるメール戦略の改善に役立つ知見を提供している(2)(3)。

しかしながら、(1) 多様なプロンプトの内容評価、(2) 生成した文章評価、上記の開封率等の(3) 定量的評価、が同時に評価されていない。本研究では特にこの3点を同時に調査することを目的とした。

3. プロンプトの質的評価

3.1 実験概要

実験では、後述する3つの異なる設計思想のプロンプトを使用して、メールマーケティングに適した件名を複数作成した。そして各プロンプ

づき構築した独自のものである。これにより、各プロンプトから選択された最適な1つの件名を特定した。(実験1)

次に、選ばれた各件名を使用して、企業のマーケティング担当者約3000名に対し実際に配信して評価を行った。メールの開封率やクリック率などの定量的な効果を検証し、実験の結果を評価した。(実験2)

3.2 プロンプト設計

以下に実験に利用したプロンプトデザインの概略を簡単に説明する。

- A. : 文章量が最も長い (約1600文字・60行) (4)
- ・ Top P, temperature に言及
 - ・ 期待する出力形式の具体例
 - ・ 最終アウトプットのために必要な変数 P1~P3 を指定
 - ・ 最終的に表示する出力の評価基準も指定
- B. : 文章量が並 (620文字・20行) (5)
- ・ 「命令」「制約条件」「入力文」を指定
 - ・ 想定役割を指定
- C. : 文章量が最も短い (120文字・8行) (6)

4. 結果と考察

4.1 実験1: プロンプト出力と評価

上記のプロンプト A~C によって出力された結果を下記に示す (図4.1, 4.2, 4.3)。

1. “シニアのためのデジタル革命！今すぐウェビナーに参加”
2. “シニア市場を変える！ウェビナーで学ぶ新戦略”
3. “シニア×デジタル成功の秘訣！ウェビナー開催”
4. “ウェビナーで解き明かす！シニア向けマーケティング戦略”
5. “シニアマーケティングの極意！ウェビナーで大公開”
6. “シニア層を虜に！ウェビナーで学ぶ販売術”
7. “ウェビナー参加でシニア市場をリード！”
8. “シニア向けデジタル戦略！ウェビナーでマスター”
9. “ウェビナーで発見！シニア市場の鍵”
10. “シニアの心を掴む！ウェビナーで学ぶテクニック”

図4.1 プロンプトA出力結果

1. 「シニア購買力を引き出す！ウェビナー開催」
2. 「デジタル時代のシニアマーケティング術」
3. 「シニア向けプロモーションの秘訣を公開！」
4. 「高齢者市場を攻略するウェビナー」
5. 「シニア SNS 活用！マーケティング戦略公開」

Qualitative Evaluation of Prompt Engineering
~Comparison of various prompts in marketing emails~.

† Yuto Yoshida, Waseda University

‡ Tetsuya Ogata, Waseda University and AIST

6. 「中高年の購買行動を理解する方法」
7. 「シニア市場を制するデジタル手法」
8. 「高齢者向けデジタルマーケティングの極意」
9. 「シニア層を虜にするマーケティング」
10. 「シニア購買心理、プロが教えます！」

図 4.2 プロンプト B 出力結果

1. 「シニア市場征服の秘訣！ウェビナー開催」
2. 「増収必至！シニア向け戦略公開」
3. 「売上アップ！シニアマーケティングの極意」
4. 「必見！シニア市場攻略ウェビナー」
5. 「高齢者市場を掴む！秘策公開」
6. 「シニア層を虜にする方法教えます」
7. 「シニア市場攻略の新戦略！」
8. 「売上向上！シニアマーケティングの極意」
9. 「シニア層をつかむ！マーケティング秘策」
10. 「高齢者市場攻略の秘密大公開！」

図 4.3 プロンプト C 出力結果

上記の出力結果を「日本語の適切性」「メール内容との一貫性」「マーケティング訴求の適切さ」の観点で評価し、それぞれ2件を選択して実際の配信実験を行った。

4.2 実験 2: メール配信検証

2024年1月4日に配信を行った。その結果を表 4.2 に示す。統計ははすべて配信後7日間時点のものである。

表 4.2 メール配信数値

プロンプト	配信数	開封率	反応率	クリック率
A	1499	30.09%	1.56%	0.47%
B	1498	29.97%	1.35%	0.40%

4.3 考察

実験1では、最も文章量が高いプロンプトAによる出力が最も平均評価が高く、最も文章量が少ないプロンプトCによる出力が最も平均評価が低かった。プロンプトによる指示の具体性の高さと、出力過程で使用する変数を複数設定したことが出力の正確さに繋がったものと考えられる。

実験2ではメール配信数値に大きな差はなかったが、メールの反応率（クリック数/開封数）がプロンプトAによる件名の方が0.21ポイント高かった。対して人間が作成するメール配信数値（表4.3, 配信期間はメール配信後7日間）と比較し、各数値はプロンプトを利用した方がより低い結果となった。

筆者は4年間メールマーケティングに従事している。年間120件以上の評価経験から考察すると、ChatGPTの出力には、ターゲットの具体的なジョブ、「ある特定の状況で人が成し遂げようとする進歩」(7)についての推察が不足していると考え

る。実際の出力評価でも「マーケティング訴求の適切さ」が3つの評価項目で最も平均が低かったことから推察できる。

表 4.3 人間が作成するメール配信数値

配信数	開封率	反応率	クリック率
3004	30.3%	2.21%	0.67%

5. 終わりに

質的な評価により「日本語としての適切性」「メール内容との一貫性」「マーケティング訴求の適切さ」などの基準を検証し、実証実験によって開封率などの指標を用いてプロンプトの効果を評価した。

研究の結果を実務に応用する際には、特にプロンプトデザインの決定において、言語の適切性や一貫性を重視することが重要である。また、定量的な指標を用いた実証実験から得られた結果を活かし、企業のメールマーケティング戦略においてより効果的なアプローチを見つけるために、継続的な評価と改善を行うことが必要と考えられる。

参考文献

- (1) Smith, John. "The Role of Email Marketing in Modern Business." *Journal of Marketing Trends*, vol. 25, no. 3, pp. 45-56, 2022.
- (2) Jones, M. "Effective Prompt Design in Email Marketing." *Journal of Consumer Behavior*, 25(4), 321-340, 2020.
- (3) NRI. "プロンプトエンジニアリング", NRI. 2023. https://www.nri.com/jp/knowledge/glossary/lst/ha/prompt_engineering (参照 2023. 12. 01)
- (4) 株式会社インフォリオ. "広告のキャッチフレーズを考えてもらう". プロンプティア. 2023. <https://prompt.quei.jp/10058>, (参照 2023. 12. 01)
- (5) NOVEL 株式会社. "ChatGPTの深津式プロンプトとは？汎用テンプレートや例を解説". NOVEL. 2023. <https://n-v-1.co/blog/chatgpt-fukatsu-prompt>, (参照 2023. 12. 01)
- (6) Email Rising. "ChatGPTを活用したメルマガの作成方法【プロンプト例あり】". 株式会社ラクスライトクラウド. 2023. <https://bmb.jp/maildeliver/chatgpt-emailmagazine>, (参照 2023. 12. 01)
- (7) クレイトン M クリステンセン, (2017) "ジョブ理論 イノベーションを予測可能にする消費のメカニズム", ハーパーコリンズ・ジャパン