

프롬프트의 우선순위가 생성형 AI 이미지에 미치는 영향 연구

생성형 AI 작품사례를 중심으로

A study of how the prioritization of prompts affects generative AI images

Focus on examples of generative AI works

주 저 자 : 심수연 (Shim, Su Yeon)

중앙대학교 예술대학원 디자인학과 객원교수
ssy0208@hanmail.net

<https://doi.org/10.46248/kidrs.2024.4.901>

접수일 2024. 11. 25. / 심사완료일 2024. 11. 30. / 게재확정일 2024. 12. 09. / 게재일 2024. 12. 30.

Abstract

Despite its unlimited potential and ability to create, generative AI is limited by its inability to accurately reflect human intent. This limits its usefulness, which is why a comparative experiment was conducted to improve its usefulness by performing planned tasks according to the creator's intention. This study is a case study of the hypothesis that prioritizing and arranging prompts can lead to results that clearly reflect the creator's intent. We used Midjourney, Gemini, and Co-Pilot to generate the images. The keywords were ranked on a 10-point scale, with the most important keywords at the front and the least important at the back, and their alignment with the researcher's topic. The success rate of prioritizing keywords was (single) 4.5 : 10 points for the keyword ranking experiment / (mixed) 4 : 10 points for the keyword + prompt ranking experiment / (single) 6.5 : 10 points for the prompt ranking experiment. The results confirm the hypothesis that the prioritization of prompts leads to clearer results, and that the information in the initial part of the prompt plays an important role in determining the clarity of the results. This study provides basic knowledge for creators and researchers who want to use generative AI.

Keyword

Generative AI(생성형 AI), Priority(우선순위), Prompt(프롬프트), Keyword(키워드)

요약

생성형 AI는 무한한 잠재력과 창작 능력에도 불구하고 인간의 의도를 정확히 반영하지 못하는 한계가 있다. 이것은 활용도의 제한을 의미한다. 때문에 창작자의 의도대로 계획된 작업을 수행하고, 활용도를 높이고자 비교실험을 하게 되었다. 이 연구는 '프롬프트에 우선순위를 부여하여 배열하면 창작자의 의도가 명확히 반영된 결과물을 얻을 수 있다'는 가설의 사례실험 연구이다. 이미지 생성은 미드저니, 제미니, 코파일럿을 사용하였다. 앞쪽이 중요한 키워드, 뒤쪽이 중요도가 떨어지는 키워드 순서이며 연구자가 작성한 주제의 부합에 따라 10점을 척도로 평가하였다. 키워드의 우선순위 성공률은 (단일형) 키워드 순위 비교실험 4.5점 : 10점 / (혼합형) 키워드+프롬프트 순위 비교실험 4점 : 10점 / (단일형) 프롬프트 순위 비교실험 6.5점 : 10점이다. 결과적으로 '프롬프트의 우선순위 배열이 명확한 결과물을 얻을 수 있다'는 가설을 입증하였고 프롬프트의 초기 부분의 정보가 결과물의 명확성 결정에 중요한 역할을 하였다. 이는 생성형 AI를 사용하고자 하는 창작자와 연구자에게 기본지식을 주는 기초자료가 되었다.

목차

1. 서론

- 1-1. 연구 배경과 목적
- 1-2. 연구 방법

2. 이론적 배경

- 2-1. 인공지능의 역사와 생성형 AI의 기본구조
- 2-2. 초거대 AI와 생성형 AI

2-3. 프롬프트 엔지니어링

2-4. 생성형 AI의 윤리적 고려사항

3. 프롬프트 우선순위에 대한 실험

- 3-1. 프롬프트 우선순위 디자인 및 규칙
- 3-2. 키워드 및 프롬프트 우선순위 비교실험
- 3-3. 키워드 및 프롬프트의 영향도 분석

4. 결론

1. 서론

1-1. 연구의 배경 및 목적

생성형 AI인 인공지능이 대량의 데이터 학습을 기반으로 새로운 콘텐츠를 창출하는 기술이다. 대부분 자연어(NLP)만으로도 창의적인 작업을 지원해준다. 이를 통해 창작자는 문학, 예술, 음악, 디자인 같은 창의적인 분야와 마케팅 및 광고 애플리케이션 등 다양한 분야에서 혁신적인 결과물을 만들어 내고 있다. 즉 AI 시스템에 의해 새로운 형태의 창작 세계가 펼쳐지고 있다.

생성형 AI의 가장 큰 이점은 자연어(NLP)만으로 창의적인 결과물을 생성해준다는 점이다. 이제 소수의 전문가 집단이나 권한이 있는 집단만의 전유물이 아니라 누구나 쉽게 사용이 가능한 지식과 기술의 민주화가 실현된 것이다.

생성형 AI 기술의 도입은 다양한 분야에서 혁신을 가져왔으나, 그 이면에는 해결해야 할 과제들이 산재해 있다. 생성형 AI를 사용하다 보면 인공지능은 마치 고도로 숙련된 전문가 같지만, 사용자의 의도대로 움직여 주지 않는다. 예술 영역의 경우 이미지를 생성하는 데 있어 창작자의 의도가 반영된 명확한 이미지가 아닌 인공지능이 외적으로만 매력적인 이미지를 생성하여 창작자에게 혼란을 준다. 창작자는 이를 수정하는 데 많은 시간이 소요되기 때문에 차라리 AI가 생성하는 결과에 맞추어 중요한 콘셉트를 변경하는 주객전도 현상도 발생할 수 있다. 이런 현상은 생성형 AI를 디자인업 무나 창의적 활동을 하는데 창작활동을 저해하고, 나아가 창작자의 주도권을 약화시키는 결과를 초래할 수 있다. 이는 유능한 인재를 효과적으로 활용하지 못하고 그에 따라 주도권을 상실하는 상황과 유사하다. 그러므로 우리는 이러한 문제점을 검토하고 한편으로는 신속하게 대응하는 자세가 필요하다. 본 연구는 이러한 문제점에 접근하여 생성형 AI의 작용과 부작용을 이해하고, 올바른 활용 방법을 제시하는 것을 목적으로 한다.

1-2. 연구 방법

본 연구는 생성형 AI에 대한 이론적 이해를 포함하여, Midjourney, Gemini, Copilot 등의 도구를 활용한 주도적이고 명확한 이미지 생성 연구에 초점을 맞춘다. 창작자의 의도를 제대로 반영하고 명확성을 높이

기 위해 키워드와 프롬프트의 우선순위가 중요하다고 판단되어 순서에 대한 구체적인 실험을 통해 이미지를 제어하는 방법을 실험하여 제시하고자 한다.

이 실험은 생성형 AI로 이미지를 생성하는 과정에서 '키워드의 배치와 순서가 명확한 이미지 생성에 주요한 영향을 미칠 것이다' 라는 가설하에 설계하였다. 이를 위해 두 가지 주요 비교실험을 설계하였다. 우선순위를 무시한 키워드 & 프롬프트 이미지 vs 우선순위에 따라 배열한 키워드 & 프롬프트 이미지를 생성하여 비교 분석한다. 이 실험을 통해 키워드 및 프롬프트 배열이 결과물에 미치는 영향을 평가하고, 이미지의 수렴함보다 창작자의 의도와 일치하였는지를 주요하게 검토할 예정이다. 그리고 우선순위를 무시한 키워드+프롬프트 혼합형 vs 우선순위에 따라 배열한 키워드+프롬프트 혼합형 결합 전략을 각각 적용하여 그 유효성을 분석한다.

이러한 연구 방법을 통해 키워드와 프롬프트의 우선순위가 명확하고 주도적인 이미지 생성에 미치는 영향을 명확히 밝히고 효과적인 전략을 제시한다. 아울러 생성형 AI의 활용에 있어 연구자 및 디자인 전문가, 학습자들도 적은 노력으로 간편하고 효과적으로 도구를 사용하여 다양한 분야에서도 생성형 AI의 활용이 활발하게 진행될 것으로 기대한다.

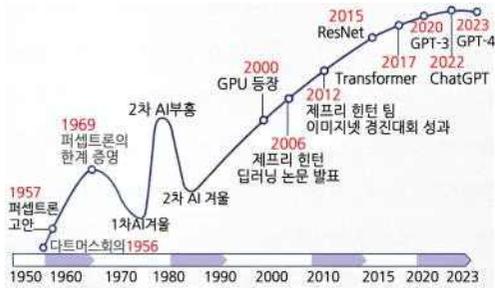
2. 이론적 배경

2-1. 인공지능의 역사와 생성형 AI의 기본구조

인공지능이라는 개념은 1950년 앨런 튜링이 '기계 가 생각할 수 있을까?'라는 논문에서 시작되었다. 이후 1956년 다투머스 회의에서 인공지능이라는 용어를 처음 사용하였다.¹⁾ 앨런 튜링의 생각하는 기계처럼 인간은 컴퓨터가 스스로 생각하여 인간의 일을 대신 처리해 주는 미래를 상상하였는데 인간의 뇌 기능을 모델화한 퍼셉트론 고안을 시작으로 인공지능은 여러 차례의 황금기와 침체기의 역사를 겪으며 발전했다. 이후 GPU의 등장으로 컴퓨팅 성능이 향상되자 대규모 데이터 세트가 기하급수적으로 증가하여 풍부한 데이터 속

1) 황현정, 'AI 리터러시 구성개념 도출에 관한 연구', 건국대학교 박사학위 논문, 2023, p.21.

에서 알고리즘을 더욱 잘 학습하게 되어²⁾ 현재와 같은 진화를 이뤘었다.



[그림 1] 인공지능의 역사 출처 : 박태웅의 AI 강의

AI는 사전에 정해진 규칙에 동작하는 반면, 생성형 AI는 사용자의 지시에 따라 창의적인 결과물을 생성해 낸다. 생성형 AI의 핵심은 적대적신경망(Generative Adversarial Networks, GAN)이다. 2014년 이안 굿펠로우(Ian Goodfellow)에 의해 제안되었으며 그는 이 개념을 다음과 같이 설명한다. 위조범은 생성 모델로 위조 화폐를 만드는 팀과 유사하며, 경찰은 판별 모델로 위조 화폐를 탐지하려는 것과 유사하다. 이 게임에서 경찰과 위조범이 서로 경쟁을 한다. 위조범은 위조 화폐가 진짜와 가짜를 구별할 수 없을 때까지 방법을 개선하고³⁾ 경찰은 위조 화폐를 탐지할 때까지 화폐를 가려낸다. 생성형 AI는 이러한 원리로 기존 작가의 스타일을 진짜와 구별할 수 없을 때까지 학습하고 경쟁하는 구조에서 새로운 문학, 미술, 음악, 디자인, 엔터테인먼트 같은 다양하고 창의적인 콘텐츠와 스타일을 생성한다.

2-2. 초거대 AI와 생성형 AI

우리는 생성형 AI와 초거대 AI를 같은 개념으로 보는데 이는 서로 다르다. 초거대 AI는 ChatGPT와 같이 매우 큰 규모의 매개변수를 가지고 방대한 데이터를 처리하거나 문제를 해결하는 반면, 생성형 AI는 미리 학습한 데이터를 기반으로 창의적인 아이디어를 생성해

2) 레이 커즈와일, 『특이점이 온다』, 김영사, 2007, p.31.

3) Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio, 'Generative Adversarial Nets', Université de Montréal, Montréal, QC H3C 3J7, 2014, p.1.

내는 인공지능이다.⁴⁾ 이 연구에 사용될 이미지 생성기는 생성형 AI인 Midjourney, 초거대 AI Gemini, Copilot을 사용한다.

2-2-1. 미드저니(Midjourney)

미드저니(Midjourney)는 2022년 7월에 출시되어 디스코드라는 채팅앱 기반의 이미지 생성 전용 인공지능이다. 미드저니는 디스코드봇을 통해서만 액세스할 수 있는 장벽이 있지만 복잡한 코딩이나 언어 없이 자연어만으로 창의적인 이미지 생성이 가능하다. 2022년 8월 미국 콜로라도주 '주립 박람회 미술대회'에서 1등을 차지하여 큰 화제가 되어 대중에게 알려졌다.⁵⁾ 따라서 생성형 AI의 대표성이 있어 이 실험에서 이미지 생성에 사용한다.

2-2-2. 제미니(Gemini), 코파일럿(Copilot)

제미니(Gemini)는 2023년 12월 구글과 딥마인드가 개발한 멀티모달(LMM) 생성형 인공지능 모델이다. 텍스트뿐만 아니라 오디오, 이미지, 비디오와 같은 다양한 입출력을 지원한다.⁶⁾ Imagen3 모델을 사용하고 있으며 멀티모달 쪽에서 특히 우수한 성능을 보여 이 연구에서 프롬프트 및 이미지 생성에 사용한다.

코파일럿(Copilot)은 2023년 2월 마이크로소프트에서 제공하는 GPT-4를 기반으로 한 '프로메테우스' 모델 기반의 AI이다.⁷⁾ 출시 이후 모델의 답변이 정확해지고 있어 미드저니와 더불어 이미지 생성에 사용된다. Bing 크리에이터와 같이 DELL-E3 모델을 사용한다. 이 실험에서 프롬프트 및 이미지 생성에 사용한다.

2-3. 프롬프트 엔지니어링

프롬프트 엔지니어링은 AI에게 명확하고 구체적인 지시를 내려 원하는 결과가 나오도록 AI를 효과적으로 쓰는 방법이다. 최상의 답변을 끌어내기 위해 최적의 단어, 입력값의 조합을 찾는 작업이다.⁸⁾ ChatGPT와

4) 김재필, 브라이언 곽, 『챗GPT 새로운 기회』, 한스미디어, 2023, p.173.

5) 심수연, '생성형 AI 이미지 향상을 위한 방법론' 중앙대학교 박사논문, 2024, p.40.

6) 나무위키, (2024.11.16.). [https://namu.wiki/w/Gemini\(%EC%9D%B8%EA%B3%B5%EC%A7%80%EB%8A%A5%20%EB%AA%A8%EB%8D%B8\)](https://namu.wiki/w/Gemini(%EC%9D%B8%EA%B3%B5%EC%A7%80%EB%8A%A5%20%EB%AA%A8%EB%8D%B8))

7) 나무위키, (2024.11.16.). <https://namu.wiki/w/Microsoft%20Copilot>

8) 장민, 안재관, 『누구든 시작하라 프롬프트

같은 챗봇이 지능적으로 설계되었다고 해서 대략적으로 질문하면 응답의 품질이 떨어진다. 질문의 명확성과 상세성은 좋은 응답을 얻는데 중요한 역할을 한다.⁹⁾ 최근 대중들이 생성형 AI에 대한 관심이 높아지자 다양한 인공지능 애플리케이션이 쏟아져 나오고 있다. 그러나 이와 동시에 잘못된 정보를 사실인 것처럼 거짓 답변하는 ‘환각 증상(할루시네이션)’ 같은 치명적인 한계점도 보여주고 있어 인공지능에 대해 우려가 크다. 이러한 ‘환각 증상(할루시네이션)’ 같은 오류를 줄이기 위해 Prompt Engineering의 영역이 더욱 중요해지고 있다.¹⁰⁾ 프롬프트 엔지니어링의 일환으로 여기서는 응답의 품질을 높이기 위해 3-1. 프롬프트 우선순위에 대한 실험에서 구성 요소 규칙을 자세히 살펴본다.

2-4. 생성형 AI의 윤리적 고려사항

한국저작권위원회에서는 저작권법상 ‘저작물’은 ‘인간의 사상 또는 감정을 표현한 창작물’을 의미한다. 즉 AI가 만든 산출물은 저작권을 인정하지 않으며 ‘저작자’ 역시 자연인, 인간만이 저작자가 될 수 있다는 뜻이다. 다만 AI 작품에 인간의 창작성이 부가된 경우나 인간이 창의적인 ‘추가 작업’을 한다면 그 부분만 저작물성이 인정되는 경우 저작권 등록이 가능하다.¹¹⁾ 예를 들어 인간의 아이디어로 스케치하고 DELL-E와 같은 도구로 완성하여 인간이 수정하고 마무리한다면 저작물로 인정될 여지가 있다는 뜻이다. 왜냐하면 인간이 창의적 아이디어를 내고 작품의 전 과정에 참여하기에 작품을 수행하였기에 창작성이 인정된다.

인공지능의 불충분한 자료에 의한 학습이나 인간의 고의적인 편향적 질문 등으로 할루시네이션도 문제가 되지만 최근 LLM 모델에서 사용자가 명령어나 질문을 악의적으로 변경하거나 조작하는 사례가 있어 더욱 문제가 되고 있다. 이를 ‘프롬프트 인젝션(prompt injection)’이라 하는데 예를 들어, 사용자가 특정 명령어 질문 등을 추가해 AI가 원래 의도된 방식과는 다른 응답을 하게 만드는 것이다. 이로 인해 AI가 사용자에게 허위 정보를 생성하고 유포할 수 있으며 불법적인 행위를 자행할 수 있다. 이는 AI에 내재되어 있는 데이터

보안의 취약점으로 작용할 수 있으며, 악의적인 의도로 AI를 오용하거나 잘못된 정보를 제공하는 상황을 유발하여 심각한 윤리적 문제를 일으킬 수 있다.¹²⁾ 이를 예방하기 위해서 사용자의 대응 방안은 다음과 같다. 1. AI가 생성하였더라도 신뢰할 수 있는 출처의 내용만을 사용한다. 2. 명확하고 구체적인 프롬프트를 작성한다. 3. 생성형 AI의 응답을 항상 비판적인 시선에서 검토하는 태도가 요구된다. 4. 민감한 주제들을 다룰 때는 편향되거나 악의적인 관점이 주입되지 않도록 각별히 주의한다.¹³⁾ 위와 같이 거대 언어 모델 기반 생성형 AI의 윤리적 문제는 언어 모델이 가진 특징으로 인해 해결이 쉽지 않아 사실상 교육을 통한 해결 방안 모색이 필요하다.

3. 프롬프트 우선순위에 대한 실험

3-1. 프롬프트 우선순위 디자인 및 규칙

명확하지 않은 프롬프트는 생성물의 품질을 저하하고, 외형적으로만 매력적인 결과물을 생성하여 창작자의 의도를 왜곡한다. 이에 창작자의 의도가 명확히 반영된 결과물을 생성하도록 다음과 같은 규칙이 있다.

- 핵심 사항 : 이 실험은 핵심은 ‘명확성’이다. 아무리 매력적인 결과물이 나오더라도 창작자의 의도 및 명확성이 떨어지면 높은 점수를 얻지 못한다.
- 실험 절차 : 3-2. 키워드 및 프롬프트 우선순위 비교실험에 따른다.
- 결과 분석 : 8개~10개의 키워드로 프롬프트를 제작하여, 우선순위로 생성되었는지를 10점 척도로 평가한다. 가설을 검증하고 결과를 요약한다.
- 추가 고려사항 : 다양한 생성 모델 (멀티모달(LMM) & 대규모 언어 모델(LLM))을 사용하여 실험을 진행해 일반화 가능성을 높인다.

엔지니어』. 알투스. 2023, p.22.

9) 심수연, Ibid., 2024, p.59.

10) 장민, 안재관, 2023, Ibid., p.21.

11) 삼성SDS, (2024.11.18.)
<https://www.samsungsds.com/kr/insights/ai-copyright-240813.html>

12) 허진웅, ‘생성형 AI의 윤리적 문제와 도덕 교육적 극복 방안’, 한국초등도덕교육학회, 2024. Vol.-No.88, p.7.

13) 허진웅, Op. cit., 2024, p.9.

[표 1] 프롬프트 디자인 및 규칙

항목	디자인 및 규칙
문제 제기	불명확한 프롬프트는 생성 결과의 질적 저하 및 창작자의 의도와의 괴리를 야기하여, 결과물의 만족도를 떨어뜨리고 추가적인 수정 작업을 필요로 한다.
가설	프롬프트에 우선순위를 부여하여 배열한다면 창작자의 의도가 명확히 반영된 고품질의 결과물을 얻을 수 있을 것이다.
프롬프트 구성 기준	1. 우선순위 프롬프트 : 가장 중요하다고 생각되는 우선순위 프롬프트를 앞쪽에 놓는다. 2. 명확한 지시 : 생성형 시가 수행할 작업을 명확하고 구체적으로 지시한다. 3. 배경지식 : 생성형 시의 이해를 돕기 위해 주변 사항 또는 배경지식을 추가한다. 4. 반복적 지시 : 중요한 사항은 반복적인 지시를 한다.

3-2. 키워드 및 프롬프트 우선순위 비교실험

본 연구는 키워드와 프롬프트의 우선순위가 이미지의 품질에 미치는 영향을 분석하고자 수행되었다. 주제는 연구자가 작성하였고, 네이버 앱을 활용하여 랜덤하게 혼합된 키워드를 사용하였다, 키워드와 프롬프트 모티브는 연구자가 직접 작성 또는 코파일럿, 제미니에게 모티브를 받아 텍스트를 생성하였다. 앞쪽이 중요한 키워드, 뒤쪽이 중요도가 떨어지는 키워드 순서이며 연구자가 작성한 주제의 부합에 따라 10점을 척도로 평가한다.

실험 절차 : 실험의 공정성을 위해 실험① 단일형과 실험② 혼합형 키워드의 주제는 같다. 실험③ 키워드와 실험과 실험④ 프롬프트의 주제 역시 같다. 주제 선택과 모든 실험은 연구자가 진행하였고 이미지 생성은 다양성을 갖기 위해 제미니, 코파일럿, 미드저니를 사용하였다.

[표 2] 키워드 프롬프트 우선순위 비교실험

항목	비교실험 규칙
(단일형)키워드 우선순위 실험	무작위 순서 키워드 or 우선순위 키워드를 생성하여 비교한다. 창작자의 의도 반영률을 10점 척도로 평가한다.
(단일형)프롬프트 우선순위 실험	무작위 순서 프롬프트 이미지 생성 vs 우선순위 프롬프트로 이미지 생성. 창작자의 의도 반영률을 10점 척도로 평가한다.
(혼합형)키워드+프롬프트 우선순위 비교실험	무작위 순서 키워드+프롬프트로 이미지 생성 vs 우선순위 키워드+프롬프트로 이미지 생성한다. 창작자의 의도 반영률을 10점 척도로 평가한다.

[표 3] 실험① (단일형) 키워드 순위 비교실험

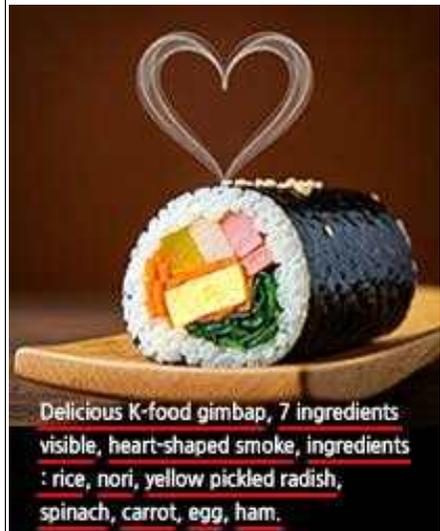
주제	주제는 김밥과 하트 모양의 연기이다, 요리는 재료와 방법이 정확히 들어가 창작자의 주문이 명확히 판단 가능하기 때문에 선택하였다.
핵심 의도	김밥 단면에 7가지 재료가 모두 들어가야 함.(2점) 하트 모양의 연기가 나와야 한다.(2점) 제대로 된 김밥 모양(1점) 10가지 키워드.(5점)
10개 키워드	K-김밥, 속재료는 모두 한 김밥 안에서 보여야 한다. heart-shaped smoke(하트연기), 밥, 김, 노란 단무지, 시금치, 당근, 계란, 햄 (제미니 사용).
	빨간 라인은 명확히 표현된 키워드이다. 앞쪽을 중심으로 이미지가 표현되는 것을 알 수 있다. 단면에 7가지 재료 : 부족 0점, 하트모양 연기 2점. 전체 모양 : 김밥이 겹쳐져 있어서 0점, 10개 키워드 5~6가지 키워드 표현되어 2.5점 = 총 4.5점

무작위 순위 키워드



우선순위로 배열하자 모든 키워드가 표현된 것을 알 수 있다. 속재료 7개, 하트모양 연기. 전체 모양 : 상, 10가지 키워드 모두 표현되어 총 10점

우선 순위 키워드



[표 4] 실험② (혼합형) 키워드+프롭트 순위 비교실험

주제	주제는 김밥과 하트 모양의 연기이다, 요리는 재료가 정확히 들어가 창작자의 주문이 명확히 판단 가능하기 때문에 선택하였다 (애니메이션 스타일 추가)
핵심 의도	김밥 단면에 7가지 재료가 모두 들어가야 함.(2점) 하트 모양의 연기가 나와야 한다.(2점) 제대로 된 김밥 모양(1점)과 10가지 키워드.(5점)
10개 키워드	K-김밥, K-김밥, 속재료는 한 김밥 안에서 보여야 함. heart-shaped smoke(하트연기), 밥, 김, 노란 단무지, 시금치, 당근, 계란, 햄. (미드저니 사용)
무작위 순위 키워드 + 프롭트	빨간 라인은 명확히 표현된 키워드이다. 앞쪽을 중심으로 이미지가 표현되는 것을 알 수 있으며 뒷부분은 거의 표현되지 않았다. 단면 속재료 0점, 하트모양 연기 0점. 전체 모양 0점, 10가지 키워드 2.5점 = 총 2.5점



Rice, spinach, carrots, K-food kimbap, yellow pickled radish, nori, ham, seven visible ingredients, egg, and heart-shaped smoke. Brush the nori with sesame oil to make it shiny. Draw a big, appetizing K-kimbap. The important thing is that it should be heart-shaped and have white smoke coming out of it. --niji

우선순위로 배열하자 모든 키워드가 표현된 것을 알 수 있다. 속재료 7개, 하트모양 연기. 전체 모양 : 10가지 키워드 모두 표현되어 총 10점 이 실험은 멀티모달을 사용하지 않았음에도 창작자의 의도가 잘 표현되는 것을 알 수 있다.



I need you to draw a delicious and tasty K-food kimbap with all 7 ingredients. The ingredients are yellow radish, rice, nori, ham, spinach, carrot, and egg. The important thing is that the heart-shaped white smoke is coming out of the rice, and all seven ingredients are inside. The nori is coated with sesame oil to give it a nice shine. Draw a big, appetizing K-kimbap. (Korean kimbap). --niji

우선 순위 키워드 + 프롭트

[표 5] 실험③ (단일형) 키워드 순위 비교실험

주제	청룡과 소녀는 마주 본다. 그들의 크기 대비는 압도적이다. 초현실적이며 안개가 낀 음산한 숲 배경. 트레버 헨더슨 스타일을 표방한다.
핵심 의도	트레버 헨더슨 스타일 이어야 한다. (1점) 청룡과 소녀는 마주 보고 있어야 하며, 크기의 대비는 청룡이 압도적으로 커야 한다. (1점) 8가지 키워드가 순차적으로 표현되어야 한다. (8점)
8개 키워드	트레버 헨더슨, 신비로운 파란색 청룡, 붉은 눈, 초현실적, 천상의 거대한 거인, 대비, 소녀, 안개 낀 숲. (미드저니 사용)
무작위 순위 키워드	무순위로 배열하자 소녀와 청룡이 친구같이 표현되었고 핵심 의도를 크게 벗어났으며 청룡과 소녀와 대비 효과도 없었다. 키워드가 뒤로 갈수록 잘 표현되지 않았다. 4개의 키워드만이 표현되어 총 4점



girl, celestial colossal giant, mystical blue dragon, surreal, contrast, foggy forest, red eyes, trevor henderson.

우선순위로 배열하자 트레버 헨더슨, 신비로운 파란색 청룡, 붉은 눈, 초현실적, 천상의 거대한 거인, 대비, 소녀, 안개 낀 숲, 10개의 키워드가 모두 표현되어 총 10점



trevor henderson, mystical blue dragon, red eyes, surreal, celestial colossal giant, contrast, girl, foggy forest.

우선 순위 키워드

[표 6] 실험④ (단일형) 프롬프트 순위 비교실험

주제	청룡과 소녀는 마주 본다. 그들의 크기 대비는 압도적이다. 초현실적이며 안개가 낀 음산한 숲 배경. 트레버 헨더슨 스타일을 표방한다.
핵심 의도	트레버 헨더슨 스타일 이어야 한다. (1점) 청룡과 소녀는 마주 보고 있어야 하며, 크기의 대비는 청룡이 압도적으로 커야 한다. (1점) 8가지 키워드가 순차적으로 표현되어야 한다. (8점)
8개 키워드	트레버 헨더슨, 신비로운 파란색 청룡, 붉은 눈, 초현실적, 천상의 거대한 거인, 대비, 소녀, 안개 낀 숲. (미드저니 사용)
무작위 순위 프롬프트	무순위로 배열하자 소녀가 달아난다. 소녀, 신비로운 파란색 청룡, 대비 0.5점, 초현실적, 붉은 눈, 안개 낀 숲 5개의 키워드만이 표현되어 총 6.5점  Girl, celestial giant was a mystical blue dragon in a surreal and contrasting misty forest, red eyes shining Trevor Henderson style. Japanese anime style.
우선 순위 프롬프트	우선순위로 배열하자 트레버 헨더슨, 신비로운 파란색 청룡, 붉은 눈, 초현실적, 천상의 거대한 거인, 대비, 소녀, 안개 낀 숲, 10개의 키워드가 모두 표현되어 총 10점  A mysterious blue dragon with red eyes in the style of Trevor Henderson. Surreal and celestial giant-like. A stark contrast to the girl standing in the foggy forest. Japanese anime style.

3-3. 키워드 및 프롬프트의 영향도 분석

[표 3] 실험① 과 [표 4] 실험②는 정확한 분석을 위해 주제가 같다, 요리는 들어간 재료가 정확히 드러나므로 재료를 평가의 기준으로 삼았다. 실험①의 기준은 일본 김밥이 아닌 한국식 김밥으로 7가지 재료가 모두 들어가 있어야 하며 하트 모양의 연기가 나오는 것이 중요하다. 제미니이는 구글과 딥마인드가 개발한 멀티모달(LMM) 이므로 이 실험에 사용하였다.

실험① 무작위 순서로 나열한 키워드 : 이미지는 몇 개의 김밥을 생성하였다. 첫 번째 김밥은 7개의 재료, 두 번째는 6개, 세 번째는 7개, 네 번째는 6개의 재료로 김밥 단면의 재료가 불규칙하고 부족하여 0점, 하트모양 연기 2점, 애매한 접시 모양 김밥이 표현되어 김밥 모양 점수 0점, 10가지 키워드 중 5~6개가 불규칙하게 표현되어 2.5점 총 4.5점을 받았다. 또한 빨간색 밀줄을 보면 앞쪽 키워드의 생성률이 높은 것을 확인할 수 있다. 따라서 중요한 키워드는 앞쪽으로 배치하는 것이 유리함을 확인하였다.

실험① 우선순위로 나열한 키워드 : 김밥 단면에 8개의 재료가 모두 들어가 2점, 하트모양 연기 2점, 정확한 김밥 모양 점수 1점, 10가지 키워드 모두 표현되어 5점, 총 10점을 받았다. 이로써 우선순위로 키워드를 배열하여 이미지를 생성하는 경우 창작자의 의도가 명확히 반영된 결과물을 생성되는 것을 확인하였다.



[그림 2] 실험① (단일형) 무작위 vs 우선순위 키워드 비교

실험② 무작위 순서로 나열한 키워드+프롬프트 : 맛 있어 보이고 스타일 좋은 K-김밥이지만 단면 속재료가 부족하여 0점. 또한 하트모양 연기는 전혀 생성되지 않았다 0점, 김밥 위에 계란이 올라가 김밥 모양 점수는 0점, 10가지 키워드를 전부 표현하지 못해 총 4점을 받았다. 그리고 빨간색 밀줄을 보면 앞쪽 키워드의

생성률이 높은 것을 확인하였다.

실험② 우선순위로 나열한 키워드+프롬프트 : 김밥 단면에 8개의 재료가 모두 들어가 2점, 하트모양 연기 2점, 제대로 된 김밥 모양 점수는 1점, 10가지 키워드 모두 표현되어 5점, 총 10점을 받았다. 이로써 키워드+프롬프트도 우선순위로 배열하여 이미지를 생성하면 창작자의 의도가 명확히 반영된 결과물을 생성되는 것을 확인하였다.



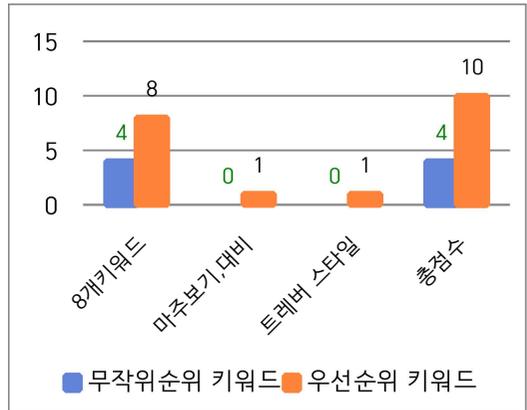
[그림 3] 실험② (혼합형) 키워드+프롬프트 순위 비교실험

[표 5] 실험③ 과 [표 6] 실험④는 정확한 분석을 위해 주제가 같다, 스토리가 있는 이미지 생성물은 점차 효용성이 높아지고 있다. 광고, 영상, 단편 영화, 디자인 등 다양한 목적으로 사용할 수 있다. 이들을 효율적으로 디자인하기 위해서는 매력적인 생성물도 중요하지만 창작자의 오더가 명확히 반영되었는지가 더욱 중요하다. 핵심 포인트는 트레버 핸더슨 스타일로 서로 마주 보아야 하며 이들의 크기가 천상의 거대한 거인과 인간이 마주하듯 압도적인 차이가 있어야 한다. 분위기는 초현실주의적이며 8개의 키워드가 모두 표현되어야 한다. ChatGPT와 같은 LLM 모델인 미드저니를 사용하였다.

실험③ 무작위 순서로 나열한 키워드 : 키워드는 네이버 앱을 사용하여 랜덤하게 혼합하였다. 트레버 핸더슨 스타일 0점, 거대한 청룡과 마주 보기, 압도적 크기 대비가 아닌 친구와 같이 비슷한 크기로 나란히 생성되어 0점, 8가지 키워드 중 4개가 순차적으로 생성되어 총 4점이다. 여기서도 앞쪽 키워드 생성률이 높은 것을 알 수 있다.

실험③ 우선순위로 나열한 키워드 : 트레버 핸더슨 스타일 1점, 거대한 청룡과 마주 보기, 압도적 크기 대

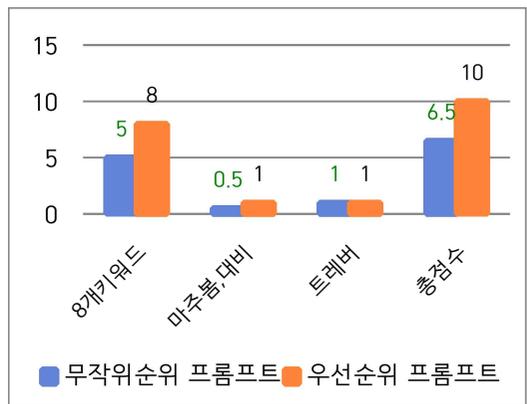
비 1점, 8가지 키워드 중 8개 모두 생성되어 총 10점이다. 이 실험 역시 우선순위로 키워드를 배열하면 창작자의 의도가 명확히 반영되는 것을 확인하였다.



[그림 4] 실험③ (단일형) 키워드 순위 비교실험

실험④ 무작위 순서로 나열한 키워드 : 트레버 핸더슨 스타일 1점, 거대한 청룡과 마주 보기, 압도적 크기 대비 0.5점, 8가지 키워드 중 5개 생성, 총 6.5점이다.

실험④ 우선순위로 나열한 키워드 : 트레버 핸더슨 스타일 1점, 거대한 청룡과 마주 보기, 압도적 크기 대비 1점, 8가지 키워드 순차적으로 생성되어 총 10점으로 반영률이 100% 올라가는 것을 확인하였다.



[그림 5] 실험④ (단일형) 프롬프트 순위 비교실험

앞선 실험을 통해 프롬프트의 우선순위가 생성 결과에 미치는 영향을 실험적으로 분석하였다. 그 결과, 프롬프트의 앞부분이 생성 모델의 출력에 더 큰 영향을 미치는 것을 확인하였다. 즉, 프롬프트의 초기 부분에

제시된 정보가 생성되는 결과물의 특징을 결정하는 데 중요한 역할을 수행한다는 것을 의미한다.

5. 결론

생성형 AI의 이미지 생성 능력은 이제는 놀라움보다 필연적인 기술 발전으로 인식되고 있다. 하지만 생성형 AI의 잠재력과 창작 능력에도 불구하고 인간의 의도를 정확히 반영하지 못하는 한계가 있다. 창작자의 의도와 필요에 부합하지 않는 잘 그려진 그림만을 그리기 때 문이다. 이것은 활용도의 제한을 의미한다. AI는 창작 자의 의도대로 사용할 수 있어야 다양한 분야에서 효 과적인 도구로 활용될 수 있기 때문이다.

따라서, 좀 인간의 의도가 잘 구현된 창의적인 작업, 계획된 디자인, 영상 작업 등에 활용도를 높이고자 이 와 같은 비교실험을 하게 되었다. 이 연구는 프롬프트 에 우선순위를 부여하여 배열하면 창작자의 의도가 명 확히 반영된 고품질의 결과물을 얻을 수 있다는 가설 하에 진행된 사례실험이다. 실험의 공정성을 위해 무순 위 키워드는 네이버 앱을 통해 랜덤하게 혼합한 키워 드를 사용하였다, 키워드와 프롬프트 모티브는 연구자 가 직접 작성 및 코파일럿, 제미니에게 모티브를 받 아 텍스트를 작성하였다. 앞쪽이 중요한 키워드, 뒤쪽 이 중요도가 떨어지는 키워드 순서이며 연구자가 작성 한 주제의 부합에 따라 10점을 척도로 평가하였다. 키 워드의 중요도 순서의 성공률은 (단일형) 키워드 순위 비교실험에서 4.5점 : 10점 / (혼합형) 키워드+프롬프 트 순위 비교실험 4점 : 10점 / (단일형) 프롬프트 순 위 비교실험 6.5점 : 10점으로 우선순위를 고려했을 때 50%의 높은 명확성을 보여주었다.

더불어 멀티모달 기능을 갖춘 Gemini와 Copilot, Midjourney 모델에서도 레퍼런스를 추가하여 우선순 위 실험을 진행하였으나 모든 모델에서 우선순위 규칙 을 준수하는 것이 레퍼런스 업로드 보다 만족스러운 결과를 주었다. 이는 프롬프트의 우선순위 규칙을 지켜 준다면 멀티모달인 레퍼런스 이미지를 업로드 하지 않 더라도 창작자의 의도가 명확히 반영된 결과물을 얻을 수 있음을 시사하였다.

본 연구의 한계점은 제한된 범위의 실험과 데이터를 기반으로 진행되어, 더욱 다양한 모델과 데이터를 활용 한 추가 연구가 필요하다. 또한, 프롬프트의 우선순위

설정 외에 다른 요소들이 이미지 생성 결과에 미치는 영향 및 키워드의 우선순위를 단계별로 체계화하여 이 를 디자인 프로세스에 적용할 수 있도록 구체적인 가 이드라인을 제시하는 연구도 필요하다.

이러한 연구가 앞으로 생성형 AI를 사용하고자 하는 연구자와 창작자에게 기본지식과 방향성을 주는 귀중한 기초자료가 되기를 기대한다.

참고문헌

1. 김재필, 브라이언 곽, 「챗GPT 새로운 기회, 한스미디어, 2023
2. 레이 커즈와일, 「특이점이 온다」, 김영사, 2007
3. 장민, 안재관 「누구든 시작하라 프롬프트 엔지니어」. 알투스. 2023
4. 허진웅, 생성형 AI의 윤리적 문제와 도덕 교육적 극복 방안, 한국초등도덕교육학회, 2024
5. Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio, Generative Adversarial Nets, Universite de Montr eal Montreal, QC H3C 3J7, 2014
6. 심수연, 생성형 AI 이미지 향상을 위한 방법론, 중앙대학교 박사논문, 2024
7. 황현경, AI 리터러시 구성개념 도출에 관한 연구, 건국대학교 박사학위 논문, 2023
8. namu.wiki
9. www.samsungsds.com
10. namu.wiki