

#taggenerator: NovelAI 를 이용한 이미지 생성 서비스의 텍스트 프롬프트 탐색 및 태그 생성 행태 연구

#taggenerator: How an Online Community Designs an Interface for Text Prompt Inputs Using a NovelAI's Diffusion Model

김선욱
Seonuk Kim
울산과학기술원
UNIST
d02reams@unist.ac.kr

권유상
Yousang Kwon
울산과학기술원
UNIST
yk7244@unist.ac.kr

이경호
Kyungho Lee
울산과학기술원
UNIST
kyungho@unist.ac.kr

요약문

Dall-E 이후 거의 무한대의 주제 조합이 가능하며, 인간 능력을 상회하는 높은 수준의 이미지 생성이 가능한 확산모델(diffusion model)이 등장해 많은 관심을 끌고 있다. 이들 모델은 text-to-image 형식에 기반한 입력을 요구하나 사용자의 입장에서 어떤 언어 구문의 조합이 자신이 원하는 주제와 스타일의 그림을 생성하는데 적합한지 실행해 보기 전에는 알 수 없다는 한계를 갖는다. 이를 극복하기 위해 사용자들은 온라인 커뮤니티 활동을 통해 수많은 시행착오를 거치며 적합한 텍스트 프롬프트를 발견하고, 공유하고, 개선하기 위한 체계적인 노력을 기울이고 있으며, 그를 통해 여러 텍스트 프롬프트 엔지니어링 툴도 집단지성을 통해 개발해 나가고 있다. 본 논문에서는 어떻게 사용자들이 어떠한 과정이나 흐름을 거쳐 이처럼 거대한 잠재공간을 가진 모델을 탐색하고, 사용하기 위해 노력하는지 분석하였다. 이를 위해 AI 를 이용한 만화풍 이미지 생성에 대해 활발한 논의가 진행되고 있으며 Tag Generator 라는 텍스트 프롬프트 입력을 위한 툴을 만든 아카라이브 AI 그림 채널을 연구장소로 선정하였다. 연구대상으로는 78,911 개의 게시글을 수집했고, 근래 확산모델의 프롬프트 실행에 관한 시행착오와 인터페이스의 디자인과 밀접히 관련된 게시글 1,068 개를 추출하여 분석하였다. 이를 바탕으로 주제분석 방법을 통해 사용자의 행동 패턴과 관련된 8 개의 유형과 25 개 범주로 분류했다. 또, 사용자에 의해 소개된 이미지 생성 AI 보조 서비스가 8 개의 행동 패턴 유형 중 어떤 유형에 해당하는지 분석했고, 이를 바탕으로 대규모 확산모델 AI 에 근간한 이미지 생성 서비스의 UI/UX 개선을 위한 실천적 방안을 제안하였다.

주제어

NovelAI Image Generator, 만화풍 이미지 생성 AI, 태그, 프롬프트, 서비스, text-to-image 모델

1. 서론

AI 를 이용한 생성적 이미지를 만드는 연구에 있어서는 그 목적과 결과의 퀄리티 등 다양한 방향으로 연구되어 Generative Adversarial Network (GAN), Variational Auto Encoder (VAE), Flow 기반 모델 등이 이용되고 있다. 거대언어모델 GPT-3 기반의 DALL-E 2 가 화제가 된 후, 특히 확산모델을 기반으로 이미지와 텍스트의 조합에 따른 무한대의 자유도를 가지는 이미지 생성 능력과 인간 상상력을 상회하는 듯한 결과물에 대한 관심이 한창이다.

확산모델은 무작위로 선택된 픽셀들로 이루어진 왜곡된 이미지를 이용해 학습한다. 모델은 이렇게 왜곡된 이미지를 원래 형태로 복원하는 방법을 배운다. 확산모델은 임의의 픽셀을 골라 왜곡된 이미지를 만들고, CLIP (Contrastive Language-Image Pre-training, 언어-이미지 대조 기반 사전 훈련)의 도움을 받으며 입력된 텍스트에 부합하도록 복원하는 과정에서 완전히 새로운 이미지를 만든다. 입력되는 언어구문인 텍스트 프롬프트(prompt) 속의 아주 작은 형식 및 입력 절차의 변화에도 모델은 민감하게 반응한다는 특징을 가진다. 모델의 간단한 모식도는 [그림 1]과 같다.

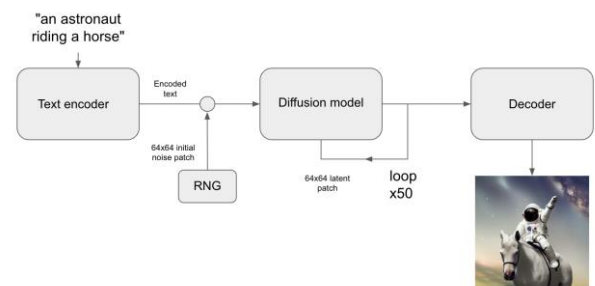


그림 1. 안정적인 확산모델의 구성과 동작방식을 보여주는 그림 (A walk through latent space with Stable Diffusion by Ian Stenbit)

이미지 생성 작업에서의 확산모델에 가능성과 그 이용에 관한 연구에 있어서는 Ho et al(2020)는 확산모델의 가능성에 대한, Radford et al(2021)과 Rombach et al(2022)는 확산모델의 성능 및 결과물의 품질 향상을 위한 방향성을 제시하였다 [1-3].

한편, Yang et al(2020)은 Human-AI Interaction 기반의 제품 및 서비스 과정에서 AI는 그 이용에 있어 완전히 새로운 패러다임전환이 필요하고, 동작방식에 대한 배경지식과 경험을 요구한다는 점에서 UX 디자인에서 사용하기 어려운 '새로운 재료'라 논의한 바 있다 [4]. 이와 같은 측면에서 DALL-E 2, Midjourney, Stable Diffusion 등 확산모델 기반 text-to-image 서비스가 나오고 있는 지금, 많은 사용자가 본인의 의도와 바램을 담아 특정한 주제와 형식을 갖춘 이미지를 만드는데 어려움을 겪고 있다.

그 이유로는 (1) 입출력과 생성 과정에 있어 중요한 역할을 하는 확산모델의 잠재 공간(latent space) 속의 정확한 동작기전에 대해 밝혀진 바가 크지 않기 때문이며, (2) 기존의 이미지 생성을 위한 어플리케이션과 전혀 다른 언어 태그 형태의 입력을 요구하기 때문이다. 또, (3) 이러한 방식 외에도 이미지 생성의 세부사항을 조절하는 파라미터(parameter)의 입력 방식 등이 직관적이지 않다. 이러한 불합리함을 개선하기 위해 모델 내부를 분석하여 잠재공간 속 동작기전을 설명하려 하거나, 예측 가능한 이미지를 생성할 수 있도록 출력 결과를 아는 프롬프트를 판매하는 등 다양한 시도가 이루어지고 있다 [5-6].

그러나, 확산모델 기반의 이미지 생성 AI 서비스를 이용함에 있어 입력할 프롬프트를 도출하는 과정에서 사용자가 겪고 있는 불편은 무엇이고, 서비스 구성에 있어 필수적인 UI 요소와 서비스 디자인 개선에 대한 연구는 잘 이루어지고 있지 않다. 이러한 문제점을 해결하고자, 본 연구에서는 이미지 생성 AI 서비스 디자인의 개선 방향을 제시하여 서비스 이용 과정에서 발생하는 불필요한 절차를 줄이고, 사용자가 원하는 이미지를 생성할 수 있도록 하는 것을 목표로 하였다. 연구자들은 이러한 새로운 디자인 재료로서의 생성모델을 응용한 인터페이스 디자인의 개선 방향을 제시하기 위해 사용자의 행동 패턴과 서비스 간의 관계 파악이 선행되어야 한다고 고려하였다. 연구자들은 첫째, 이미지 생성 AI 서비스를 이용 중에 텍스트 프롬프트를 도출하기 위해 사용자는 어떤 행동 패턴을 보이며, 둘째, 각 행동 패턴은 어떤 관계가 있고 어떤 일련의 과정을 거치는지, 셋째, 특정 행동 패턴을 촉진하거나 불필요한 절차를 줄이기 위해 사용자가 개발한 보조 서비스를 바탕으로 어떤 개선된

서비스가 필요한가라는 세 가지 연구질문을 선정했다.

이를 탐구하기 위해 최근 AI를 이용한 만화풍 이미지 생성에 대해 활발하게 논의되고 있으며 Tag Generator라는 텍스트 프롬프트 입력 툴을 만든 아카라이브 AI 그림 채널을 연구대상으로 선정하였다. 연구자료로 게시글 78,911 개를 수집했고, 커뮤니티에서 가치를 인정받아 추천수 10 이상인 게시글 1,068 개에 주목해 분석하였다. 주목한 게시글은 확산모델의 프롬프트 실행과 시행착오 과정을 잘 보여주고, 인터페이스 디자인과도 밀접히 관련되어 있었다.

분석에 있어서는 주제 분석(thematic analysis) 방법을 통해 세 명의 다른 연구자가 사용자의 행동 패턴과 관련된 8 개의 유형과 25 개 범주로 분류했다. 또, 커뮤니티의 사용자에 의해 반복적으로 언급되거나 사용된 이미지 생성 AI 보조 서비스는 연구자들이 도출한 8 개의 행동 패턴 유형 중 어떤 유형에 해당하는지 분석해 연구결과가 향후 인터페이스를 분석하는 카테고리로서 사용될 수 있는지, 실제 디자인을 진행함에 실천적인 지침으로 삼을 수 있는지 그 가능성을 제시하였다.

본 연구에서 정의한 사용자 행동 패턴은 일련의 과정에서 디자인을 개선함에 따라 사용자가 어떻게 행동할지를 예측하는데 유용한 판단 근거가 될 수 있을 것으로 생각되며, 제시한 방향성을 바탕으로 이미지 생성 AI 서비스 디자인을 개선한다면, 앞서 언급한 인간중심 AI 관점에서 보다 설명력 있고 직관적인 인간-AI 사용자 경험을 제공할 수 있을 것이다.

2. 이론적 배경 및 현황

2.1 Stable Diffusion

Stable Diffusion은 Stability AI에서 개발한 확산모델(diffusion model) 기반 이미지 생성 AI다. 컴퓨팅 자원을 적게 소비하면서도 오픈소스라는 특징 덕택에 재학습(fine-tuning)에 용이했고, 많은 Stable Diffusion 기반 이미지 생성 AI가 새롭게 등장하였다. 특히 만화풍 이미지 생성 AI에 큰 영향을 끼쳐 Waifu Diffusion, NovelAI Diffusion 등의 모델이 개발되었고, NovelAI Image Generator 같은 상업적 AI 이미지 생성 서비스가 출시되기도 했다 [5].

Stable Diffusion의 text-to-image 모델은 오토인코더의 원리를 바탕으로 텍스트로부터 이미지를 생성한다. 오토인코더는 인코더와 디코더로 구성되는데, 인코더는 사용자가 입력한 텍스트를 AI가 이해할 수 있도록 벡터화하고, 디코더는 임베딩 벡터를 사람이 이해할 수

있는 이미지 등의 형태로 출력한다. Stable Diffusion 에서는 이미지를 벡터화시킨 후 잡음(noise)을 입히고 복구(denoise)하는 과정에서, 원본 이미지의 정보와는 관계없이 의도적으로 사용자가 입력한 정보를 반영시키기 위해, 인코더를 통해 만들어진 임베딩 벡터를 단서로 제공하여 이미지를 출력하도록 한다. 즉, AI가 복구했다고 믿는 이미지는 실제로는 사용자가 의도한 정보를 반영하여 창작된 완전히 새로운 이미지인 것이다 [3, 9].

2.2 NovelAI Image Generator

NovelAI Image Generator는 Anlatan에서 개발한 만화풍 이미지 생성 AI 서비스로, Stable Diffusion을 기반으로 각 만화풍 이미지와 텍스트를 한 쌍으로 학습시킨 NovelAI Diffusion 모델을 통해 품질 높은 만화풍 이미지를 생성하여 많은 관심을 받았다. 실제로 [그림 2]와 같이 텍스트 프롬프트를 입력하면 같은 구도로도 다양한 스타일의 이미지를 어색함 없이 생성할 수 있다. 반면, NovelAI 측이 공식 트위터에서 불펌 이미지가 많은 Danbooru로부터 학습 데이터를 수집했다고 밝혀 저작권 문제에 대해 비판을 받기도 했다. 이후 해킹 공격으로 NovelAI Diffusion 모델의 소스 코드가 유출됨에 따라, 로컬 환경에서 같은 조건을 입력하면 기존 서비스와 거의 동일한 그림을 생성할 수 있게 되었다. 각종 인터넷 AI 그림 커뮤니티에서 이미지 생성, 프롬프트 공유, 모델 재학습 등 다양한 논의가 더 활발해졌지만, 방송에서 작업중인 그림을 무단 복제하여 AI로 완성한 후 저작권을 주장하거나 AI 그림을 상업적으로 판매하는 등 일부 사용자가 논란을 일으키며 양면성을 보여주고 있다.

사용자 일부는 구독료에 대한 부담, 커스터마이징의 자

유도 등의 이유로 Stable Diffusion web UI를 기반으로 로컬 환경에서 NovelAI Diffusion 모델을 구동하고 있다. <표 1>에서는 아카라이브 AI그림 채널에서 통용되는 정보를 기반으로 이미지 생성을 위해 필요한 Stable Diffusion web UI 설정 관련 용어를 정리하였다 [7].

3. 연구 방법 및 결과

3.1 연구 방법 및 범위

아카라이브 AI 그림 채널은 2만 명이 넘는 구독자를 가진 대형 AI 그림 커뮤니티이다. 2022년 9월 8일에 첫 게시글이 올라온 이후, 2022년 11월 14일에 약 8만 개의 게시글이 올라왔을 정도로 사용자 간의 활발한 소통이 이루어지고 있다. 초기에는 다양한 이미지 생성 AI에 대해 다뤘지만, 현재는 대부분의 게시글이 만화풍 이미지 생성 AI에 대해 다루고 있다. 2022년 11월 14일을 기준으로 추천 10개 이상을 받은 게시글 1068개 모두를 연구대상으로 다음과 세 가지 단계로 연구를 진행했다. 첫째, 선정된 게시글을 읽고 만화풍 이미지 생성 AI 서비스 사용자에게 대한 사례연구를 진행하였다 [7]. 둘째, 사례연구를 기반으로 원하는 이미지 생성을 위해 사용자가 입력할 프롬프트를 찾는 과정을 파악하고, 주제 분석(thematic analysis)을 통해 행동 패턴의 범주(category)와 유형(theme)을 정의하였다 [10]. 셋째, 편의를 위해 사용자가 자발적으로 만들거나 제안한 서비스가 어떤 유형의 행동 패턴 유형을 보조하기 위한 것인지 분석하고, 도출된 유형과 각 서비스의 장단점을 바탕으로 추가 분석하여 사용자가 원하는 그림을 생성하는 경험을 극대화할 수 있는 방안을 제시하였다.



그림 2. NovelAI Image Generator로 같은 설정에서 작가 태그만 변경하여 생성한 이미지. 각 이미지 하단은 설정한 작가

표 1. Stable Diffusion web UI 설정 관련 용어 [7]

용어	설명
Model	Stable Diffusion 기반 AI 모델(.ckpt)
Hypernetwork	U-Net 의 각 Step 에서 Query, Key, Value 값 중 텍스트와 연관된 Key, Value 가중치(.pt)
Prompt	유저가 AI 에 입력하는 텍스트 전체. Main prompt 에 입력된 정보를 기반으로, Negative prompt 에 입력된 정보를 배제하는 방향으로 이미지 생성
Tag	Prompt 내 각각의 단어. 쉼표(.)로 구분
Embedding	AI 모델에서 이미지 모음을 학습시켜 새로운 개념을 매칭시킨 특정 태그
Sampling	이미지를 생성하는 과정. 보통 step 이 높을수록 그림 품질이 높아지지만 시간이 많이 걸리고, 일정 수치 이상일 때는 품질이 오히려 떨어짐
CFG Scale	사용자가 넣은 프롬프트 입력을 얼마나 따를지 판단하는 척도. 높을수록 프롬프트 입력을 지키면서 그림을 생성하고, 낮을수록 AI 가 해석의 자유를 가지고 그림을 생성
Seed	최초의 노이즈 이미지 생성에 사용되는 값(-1 일 경우 난수 자동 입력)

3.2 프롬프트 탐색 행동 사례연구

각 게시글을 하나의 사례로 간주하고 내용을 기반으로 게시글의 중심 내용을 요약하고 URL 과 함께 문서화했다. 프롬프트를 탐색하기 위해 사용자가 취하는 행동에 집중하여 문서화된 내용을 바탕으로 범주화를 시도했다. 범주화 과정에서 각 범주의 근거가 되는 내용을 다시 추출하는 과정을 거쳐 25 개의 범주를 도출하였다. 도출된 범주는 <표 2>의 범주와 같다.

3.3 프롬프트 탐색 행동 패턴 유형 및 범주 정의

25 개의 범주를 바탕으로 <표 2>과 같은 8 가지 행동 유형을 정의했다. 범주에서의 행동이 결과적으로 어떤 행동으로 일반화될 수 있는가에 대해 집중하였다. 단, 의미를 탐색하는 행동과 의미를 임의로 강화-약화하는 행동은 접근 방식에 큰 차이가 있기에, 의미적 접근과 수치적 접근으로 구분하여 한 번 더 분류했다.

표 2. 정의된 프롬프트 탐색 행동 패턴 유형 및 범주

유형(themes)	범주(categories)
소비	다른 사용자의 프롬프트를 모방
	다른 사용자에게 질문 커뮤니티 내부 정보 검색
탐색	의미적 접근
	정식 명칭 탐색 등 의미적 접근
	태그 간의 공통점과 특징을 바탕으로 태그 간의 궁합 판단
	의도한 요소가 나오도록 비슷한 위치나 특징을 공유하는 태그로 대체적 접근
수치적 접근	세분화한 요소를 태그로 사용해 예측 가능한 범위의 이미지 생성
	이중부정, 약화된 태그 부정으로 이미지를 생성하는 술어논리적 접근
	반복 입력을 통한 태그 강화
	가중치 설정 기반 태그 강화 및 약화
생산	다른 사용자와 프롬프트를 공유
	보편적 품질 향상을 위한 태그 제시 커뮤니티 외부 정보 검색 및 공급
정리	프롬프트 관련 정보 정리
	그림의 주요 요소를 분류하여 프롬프트 내의 태그 형식과 순서 정형화
역설계	임베딩을 통한 태그 정의
	모델 재학습, 서비스 개발 등 프롬프트 외적 개척 활동으로 프롬프트 탐색을 위한 인사이트 획득 및 제공
	가정
	AI 작동 과정을 인간이 문장을 처리하는 사고 과정에 대입하여 접근 AI 모델 특징 탐구 및 학습 데이터 등 기반 프롬프트 제시
실험	커뮤니티에서 제기된 가설을 검증
	여러 태그를 조합한 프롬프트를 생성된 이미지를 보고 경험적으로 접근하여 태그 간의 궁합 판단
	대조실험을 통한 프롬프트 선별
	시행착오를 통한 휴리스틱 탐색
혁신	그림의 주요 요소 구조를 재구성
	이모지, 이모티콘, 유니코드 문자 등을 태그로 활용하는 비언어적 접근

또한, 일련의 게시글을 분석하는 과정에서 각 행동 패턴 유형들이 [그림 3]과 같은 관계를 지닌다는 사실을 확인할 수 있었다.

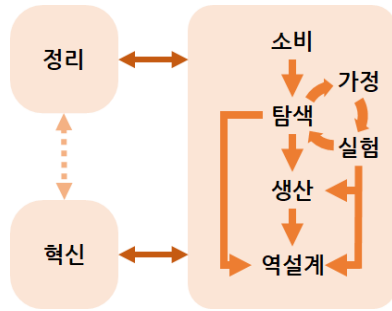


그림 3. 행동 패턴 유형 간의 관계 도식

3.4 이미지 생성 AI 보조 서비스 분석

사례연구를 진행하며 확인된 이미지 생성 AI 보조 서비스는 <표 3>과 같다.

표 3. 사용자가 커뮤니티에 소개한 AI 보조 서비스[4]

서비스	설명
Snowshell	이미지의 화질을 높여주는 업스케일러
Illuminarty	이미지가 AI가 그린 그림인지 판단
AI 이미지 EXIF 뷰어	이미지 EXIF 정보 기반 프롬프트 추출 및 디러닝 기반 제안
NovelAI 태그 생성기	프롬프트에 활용할 태그 분류, 프롬프트 저장, 복제 보조
WebUI 태그 자동완성	태그 자동완성 및 Danbooru 태그 수를 기준으로 정렬하여 추천
AIBooru	AI 생성 그림 및 프롬프트 정보 공유
Prompt Search	AI 생성 그림 및 프롬프트 정보 공유, 추천 기준별 정렬, EXIF 뷰어, 입력한 프롬프트 기준으로 공유된 그림 중 가장 유사한 그림 제시
NovelAI Tag Experiments	태그 분류 및 세부 태그 정리, 이미지 및 프롬프트 공유
Quick NAI	Web UI 문법 지원, 맞춤형 태그 추천 시스템, 즐겨찾기
태그 가중치 변환기	Web UI 및 NovelAI Image Generator 문법 상호 변환
Anime-Face-Detector	만화풍 이미지에서 캐릭터 얼굴 부분만 자동 인식 및 수집 보조

<표 3>의 서비스 외에도 이미지와 태그를 매칭한 후 학습 데이터용으로 정규화된 이미지를 다운로드 받을 수 있게 하는 서비스, 원활한 학습을 위한 서버 공유 서비스, 리터칭 서비스 등에 대한 수요가 있음을 확인했다. 이미지 생성 AI 보조 서비스와 연관된 프롬프트 탐색 행동 패턴 유형을 [그림 4]와 같다.

4. 논의 및 제언

[그림 4]를 통해 간접적으로 혁신과 역설계를 보조하는 서비스가 부족하며, 가정과 실험, 정리, 수치적 접근 기반 탐색을 보조하는 서비스 역시 보완될 필요가 있음을 확인할 수 있다. 이 중 보완의 효율성이 클 역설계, 실험, 탐색 보조 서비스의 증진을 위해 제언하고자 한다.

4.1 역설계 관련 제언: 이미지 정규화 기능이 있고 직관적인 UI를 가진 데이터베이스 서비스

역설계는 AI에 대한 정보를 바탕으로 프롬프트 이외의 요소로 모델을 개선한다는 특징 때문에, 접근 난이도가 매우 높은 편에 속한다. [그림 3]과 [그림 4]를 함께 비교해볼 때, 소비에서 생산 단계에 도달하는 것보다 다른 단계에서 역설계 단계로 도달하는 것이 더 힘들다는 것을 유추해볼 수 있다. 역설계는 필연적으로 학습 데이터가 필요한데, 이미지-태그 매칭 그리고 학습 데이터에 맞게 이미지 크기의 정규화가 필요하다. 따라서, 생성된 데이터를 학습 데이터화해서 자유롭게 다운로드 받을 수 있는 데이터베이스 서비스가 필요할 것으로 보인다. 또한, 접근성을 향상시킬 수 있도록 직관적인 드래그 앤드롭(Drag and Drop) 등의 형태로 UI 개선도 필요하며, 많은 사용자가 역설계 단계에 도달할 수 있게 이전 단계의 서비스의 전반적 개선이 필수적일 것이다.

4.2 실험 관련 제언: 실시간 이미지 생성 과정(step) 시각화 및 챌린지(Challenge) 서비스

대조실험은 프롬프트의 효과를 규명할 때, 가장 신뢰도 높은 방법이지만, (1) Seed에 따라 다른 결과가 나오는 등 재현이 안 될 가능성 (2) 결과물을 얻기까지 긴 소비 시간 (3) 너무 많은 경우의 수로 인해 낮은 일반화 가능성 같은 큰 단점이 존재한다. 만약 이미지 생성 중 각 Step의 결과를 시각화해준다면 가정적 접근으로 영감(insight)을 얻는 사용자가 생길 수 있고, 이를 통해 만들어진 가설을 검증하면서 실험이 더 활성화될 수 있을 것이다. 또한, Kaggle과 비슷한 챌린지 서비스가 생긴다면 제시된 문제로부터 영감을 얻어 새로운 실험 문제를 제시하고, 초보자는 해결된 문제로부터 사용법을 연습하여 더 쉽게 역설계 단계에 도달할 수 있도록 발판을 마련하는 효과도 기대할 수 있을 것이다.

	소비	탐색		생산	정리	역설계	가정	실험	혁신
		의미	수치						
Snowshell				●					
Illuminarty	●					●	●	●	
AI 이미지 EXIF 뷰어	●	●	●	●				●	
NovelAI 태그 생성기	●	●		●	●				
WebUI 태그 자동완성	●	●		●			●		
AIBooru	●	●	●	●	●		●	●	
Prompt Search	●	●	●	●	●		●	●	
NovelAI Tag Experiments	●	●	●	●	●		●		
Quick NAI		●	●	●	●		●	●	
태그 가중치 변환기		●	●	●					
Anime-Face-Detector						●			

그림 4. 그림 생성 AI 보조 서비스가 보조하는 프롬프트 탐색 행동 패턴 유형 정리

4.3 탐색 관련 제언: 프롬프트 인터페이스 정형화 및 가중치 수치 UI 개선

의미적 접근 기반과는 다르게 수치적 접근 기반 탐색은 시도가 적고 일반화가 잘 이루어지지 않고 있다. 이는 순서, 단어의 가중치, 슬어논리적 접근 등 상황에 따라 변화무쌍한 파라미터가 너무 많기 때문이다. 만일 프롬프트를 텍스트를 입력하는 형태가 아닌 태그와 각 태그의 가중치 수치를 지정된 수치로만 선택할 수 있게 하면 경우의 수가 줄어들어 수치적 접근 기반 탐색의 여러 요소에 대한 일반화 가능성이 높아질 수 있다. 즉, 인터페이스를 정해진 수치의 범위 안에서 사용자가 조절할 수 있는 UI 같이 사용성을 체계적이고 합리적인 방법으로 개선하면 대조실험적 접근이 용이해질 것이다.

5. 결론

본 연구는 게시글 기반 주제 분석을 통해 사용자가 만화풍 AI 이미지 생성을 위해 프롬프트 탐색을 하는 행동 패턴 유형 8가지를 도출하였다. 또한, 사용자가 소개한 이미지 생성 AI 보조 서비스를 분석하여 어떤 유형을 보조하는지 파악하고, 서비스가 더 필요하거나 개선되어야 할 유형을 무엇인지 확인할 수 있었다. 이 중 역설계, 실험, 탐색 유형을 선정하여, 최종적으로 관련 서비스의 개발이나 개선점을 제안했다. 한편, 본 연구는 아카라이브 그림 AI 채널의 게시글만을 대상으로 진행하여 다양한 커뮤니티의 의견을 반영하지 못했다는 한계점이 있으며, 추천 기준으로 과하게 필터링했을 가능성 또한 존재한다. 향후 직접 서비스를 개발한 후 데이터를 수집하거나, 인터뷰와 같이 양적 연구가 동반한 사용자 조사를 통해 구체적인 불편사항과 피드백을 분석한다면, 이미지 생성 AI 서비스 디자인을 더 사용자 친화적인 방향으로 개선할 수 있을 것이다.

참고 문헌

- Ho, J., Jain, A., and Abbeel, P. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33, pp. 6840–6851. 2020.
- Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., and Sutskever, I. Learning transferable visual models from natural language supervision. In *International Conference on Machine Learning*. PMLR, pp. 8748–8763. 2021.
- Rombach, R., Blattmann, A., Lorenz, D., Esser, P., and Ommer, B. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10684–10695. 2022.
- Yang, Q., Steinfeld, A., Rosé, C., and Zimmerman, J. Re-examining whether, why, and how human-AI interaction is uniquely difficult to design. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–13. 2022.
- Bahng, H., Chun, S., Yun, S., Choo, J., and Oh, S. J. Learning de-biased representations with biased representations. In *International Conference on Machine Learning*. PMLR, pp. 528–539. 2020.
- PromptBase, <https://promptbase.com> November 14, 2022.
- 아카라이브 AI그림 채널, <https://arcalive/b/aiart> November 14, 2022.
- NovelAI, <https://novelai.net> November 14, 2022.
- Alammar, J. The Illustrated Stable Diffusion <https://jalammar.github.io/illustrated-stable-diffusion/> November 14, 2022.
- Braun, V., and Clarke, V. Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), pp. 77–101. 2006.