

## 숏폼 콘텐츠 크리에이터를 위한 생성형 AI 기반 차세대 숏폼 제작 플랫폼

조 동 희<sup>1</sup> · 김 민 서<sup>1</sup> · 김 성 민<sup>1</sup> · 김 건 동<sup>2\*</sup><sup>1</sup>홍익대학교 디자인컨버전스학부 학부과정<sup>2</sup>홍익대학교 디자인컨버전스학부 교수

## A Generative AI-based Next-generation Short-form Creation Platform

Dong-Hee Jo<sup>1</sup> · Min-Seo Kim<sup>1</sup> · Sung-Min Kim<sup>1</sup> · Geon-Dong Kim<sup>2\*</sup><sup>1</sup>Undergraduate Program, School of Design Convergence, Hongik University, Sejong 30016, Korea<sup>2</sup>Professor, School of Design Convergence, Hongik University, Sejong 30016, Korea

### [요 약]

숏폼 시장의 빠른 성장에 따라 숏폼 시청 경험은 꾸준히 증가하고 있지만 숏폼 제작 환경 및 게시 경험은 부족한 실정이다. 이에 본 연구는 숏폼 콘텐츠 아이디어 기획부터 제작 및 편집까지 돕는 생성형 AI 기반 모바일 숏폼 제작 플랫폼을 제안하는 데 목적이 있다. 사례 조사로 3개의 숏폼 플랫폼과, 2개의 LLM 서비스, 4개의 생성형 AI 기반 이미지 및 영상 생성 서비스, 그리고 3개의 프롬프트 기반 영상편집 서비스를 분석하여 숏폼 크리에이터의 작업 단계별 생성형 AI 적용 가능 지점을 도출하였다. 79명 대상 사용자 조사를 통해 니즈를 파악한 후, 기술 리서치를 통해 LLM 모델에서 T2I와 T2V 모델로 연결되는 영상 제작 및 편집 알고리즘을 재구성하였다. 4개의 숏폼 제작 단계별 유저 플로우 기반 생성형 AI 반영 핵심 프로토타입 제작을 통해, 숏폼 콘텐츠 크리에이터들이 숏폼 아이디어 기획에서 편집까지 좀 더 쉽고 직관적으로 사용할 수 있는 서비스를 제안하였다.

### [Abstract]

With the rapid growth of the short-form market, the experience of watching short-form content has increased steadily. However, the environment for creating and publishing short-form content remains insufficient. To address this, this study proposes a generative AI-based mobile creation platform that assists users in planning, producing, and editing short-form content. To this end, we analyzed three short-form platforms, two large language model services, four generative AI-based image and video creation services, and three prompt-based video editing services. Based on the analyses, we identified key opportunities for applying generative AI to each step of a short-form creator's workflow. Based on a survey involving 79 participants, we leveraged technological research to reconfigure video creation and editing algorithms from large language models to text-to-image and text-to-video models. By developing core prototypes that integrate generative AI into the four stages of short-form content creation based on user flows, we propose a service that enables short-form content creators to plan and edit their ideas easily and intuitively.

**색인어** : 숏폼, 생성형 AI, 프롬프트, 세로형 타임라인, 숏폼 제작 플랫폼**Keyword** : Short-form, Generative AI, Prompts, Vertical Timeline, Short-form Creation Platform<http://dx.doi.org/10.9728/dcs.2024.25.7.1701>

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 02 June 2024; Revised 02 July 2024

Accepted 18 July 2024

\*Corresponding Author; Geon-Dong Kim

Tel: 

E-mail: geon705@gmail.com

## I. 서론

### 1-1 연구배경 및 목적

숏폼(Short-form)은 길이가 짧게는 15초, 길게는 10분 이내의 짧은 영상 콘텐츠로[1], 스마트폰 기반 모바일 환경에서 MZ 세대를 중심으로 활발히 소비되고 있다. 이와 같은 숏폼 콘텐츠의 사용 추이를 보면 오픈서베이의 ‘소셜미디어·검색포털 트렌드 리포트 2023’ 기준 15세에서 59세 대상자 10명 중 7명은 숏폼 콘텐츠를 경험할 만큼 시청자는 가파르게 증가하고 있다[2]. 그 중 MZ 세대의 숏폼 활용도가 높은 경향을 보이는데 MZ 세대의 콘텐츠 소비 특징 중 하나는 시간 효율성 추구이다. 가령, 여가에 1시간이 넘는 예능 한 편을 보는 것보다 핵심 부분이 집약된 영상 콘텐츠를 짧게 몰입하여 시청하는 것을 더 선호하고, 그 잔여 시간을 또 다른 관심 콘텐츠를 이어서 보거나 찾아보며 효율성을 추구한다[3]. 다른 특징으로는 영상 콘텐츠를 공유하고자 한다. 시청한 영상을 친구와 함께 공유하거나 직접 제작한 영상 콘텐츠를 게시하고, 그 반응을 보는 과정에서 즐거움을 느낀다. 이처럼 숏폼 콘텐츠의 소비가 빠르게 증가하며 주 사용자인 MZ 세대의 숏폼 콘텐츠 활용 특징도 연구되고 있지만 실제 숏폼의 업로드 수를 살펴보면 숏폼 콘텐츠의 공급이 수요를 따라가지 못하고 있다. 또한, 현 숏폼 편집 툴도 기존 롱폼(Long-form) 편집 방식을 유지하는 경우가 많아 MZ 세대의 모바일 이용 성향과 트렌드를 반영하기 어려운 실정이다. 따라서, 본 연구는 생성형 AI 기술을 모바일 기반 숏폼 콘텐츠 제작 플랫폼에 적용해 숏폼 크리에이터의 창의적인 영상 콘텐츠 기획과 제작을 돕는 것에 목적을 둔다. 이를 위해 최근 여러 분야에 활발하게 적용되고 있는 생성형 AI 기술을 활용하여 프롬프트를 기반으로 아이디어 보드에서 숏폼의 시나리오 스크립트를 작성한 후 통 숏폼을 생성하거나 기존 타임라인 기반 영상편집 방식과 생성형 AI 기반 장면 및 효과 편집기능의 효율적 기능 연계를 통해 숏폼 크리에이터의 성향과 트렌드에 부합한 차세대 숏폼 제작 플랫폼의 프로토타입을 제안하고자 한다.

### 1-2 연구 범위 및 방법

본 연구는 모바일에서 생성형 AI 기반 숏폼 제작이 가능한 서비스 플랫폼 제안 연구로, 숏폼 콘텐츠 크리에이터가 숏폼 제작 및 편집을 하기 위해 진행하는 기획에서 프로토타입 제작까지의 숏폼 제작 과정이 담긴 플랫폼의 프로토타입 디자인 제안을 주 연구 범위로 한다. 연구 방법은 다음과 같다. 첫째, 선행 대표 숏폼 제작 플랫폼에서 콘텐츠 제작 시 직면하는 문제를 분석한다. 둘째, 숏폼 크리에이터의 작업 과정을 정의하고 단계별로 적용 가능한 생성형 AI 관련 기술을 분석한다. 셋째, 선행 생성형 AI 서비스 사례와 프롬프트 중심의 영상편집 서비스 분석을 통해 숏폼 콘텐츠의 이미지 및 영상 생성 관련 기술 조사 후 그 알고리즘 체계를 도출한다. 넷째,

MZ세대 중 숏폼 소비 경험이 있는 이들을 대상으로 숏폼 소비 관련 설문 조사를 진행하고, 전체 피험자 그룹 중 특히 직접 숏폼 제작을 경험한 대상자들에게 숏폼 콘텐츠 기획에서 제작에 이르기까지 숏폼 콘텐츠 제작 툴의 문제점 및 니즈를 살펴본다. 넷째, 선행사례 분석 및 설문 조사의 분석 결과를 바탕으로 숏폼 콘텐츠 제작 및 편집에 적합한 숏폼 제작 플랫폼의 차별화된 핵심 개념트를 도출한 후 프로토타입을 제안한다. 마지막으로 본 연구의 결론 및 한계점을 제시한다.

## II. 관련 연구 및 선행사례 분석

### 2-1 숏폼 서비스 현황분석

#### 1) 숏폼 콘텐츠의 정의 및 특징

숏폼은 짧다(Short)의 의미와 형식(Form)의 합성어로 초기 숏폼 콘텐츠는 예고(Teaser)나 쿠키 영상(Post-credits scene), 그리고 프롤로그(Prologue)와 에필로그(Epilogue) 형태나 핵심 장면들을 집약한 하이라이트(Highlights)와 같이 축약된 영상을 의미하는 단어였다[1], [4]. 반면 최근 숏폼은 모바일 숏폼 플랫폼 내에서 화면을 위아래로 스와이프(Swipe)하며 간편하게 콘텐츠를 이어서 소비하는 몰입도 높은 시청 환경을 의미한다[5]. 숏폼은 롱폼 대비 촬영 및 편집이 간단한 특징을 바탕으로 MZ세대 사이에서 쉽게 생산 소비되는 동영상 콘텐츠의 한 형식으로 자리 잡으며[4] 인기를 이어가고 있다.

#### 2) MZ세대의 숏폼 콘텐츠 이용 실태

숏폼 콘텐츠 시청 경험은 2022년 56.5% 대비 2023년 68.9%로 12.9% 증가하였다. 하지만 숏폼 콘텐츠 게시 경험은 2022년 4.6% 대비 2023년 5.3% 정도로 10대에서 20대 기준 약 10명 중 1명만 숏폼 콘텐츠 영상을 직접 만들어 올리고 있어 시청 경험 대비 업로드 경험은 매우 낮은 것을 확인할 수 있었다. 시청한 영상 콘텐츠를 친구와 함께 공유하거나, 직접 제작한 영상 콘텐츠를 게시하고 반응을 보는 과정 자체에서 즐거움을 느끼는 MZ세대 특징에도 불구하고 현 숏폼 제작 플랫폼은 아직 사용자에게 진입장벽이 있음을 알 수 있다. 따라서 본 연구는 이와 같은 문제점을 바탕으로 해당 세대가 모바일에서 숏폼 콘텐츠 제작에 어려움을 겪는 이유를 분석한 후 대상 사용자들이 더 쉽게 숏폼 콘텐츠를 제작하고 공유할 수 있는 숏폼 제작 서비스 플랫폼을 제안해 보고자 한다.

#### 3) 선행 숏폼 플랫폼의 연계 분석

그림 1은 2023년 2월 기준, 국내 숏폼 시장을 주도하는 이용자 수가 가장 많은 3대 플랫폼을 분석 대상으로 선정 후 연계 분석한 내용이다. 플랫폼은 순서대로 유튜브 숏츠(YouTube Shorts), 인스타그램 릴스(Instagram Reels), 틱

특(Tik Tok)이다. 분석 도구는 숏폼 시청 화면, 촬영 화면, 편집 화면의 환경과 기능, 전체 서비스의 장점 및 문제점을 정하였다. 첫 번째로, 그림 1-1은 각 플랫폼의 숏폼 시청 화면에 위치되어 있는 영상 재제작 기능을 분석한 내용이다. 유튜브 쇼츠는 우측 하단의 리믹스 기능을 이용해 원본 영상을 재제작하는 방법을 제공하고 있다. 인스타그램 릴스는 미리 디자인된 영상 레이아웃이 담긴 템플릿 기능을 제공한다. 숏폼에 사용된 음악과 자막, 영상 특수효과를 그대로 사용할 수 있어 사용자가 직접 촬영한 영상과 이미지를 갈아 끼워 동일 포맷의 영상으로 제작할 수 있다. 틱톡은 숏폼 시청 화면에서 우측 하단의 'Send to' 버튼으로 진입한 후 한 화면 안에 원본 영상과 사용자가 촬영한 영상 붙이기인 'Duet'과 원본 영상 뒤에 사용자 영상 붙이기인 'Switch'를 통해 동영상 재제작이 가능하다. 두 번째로, 그림 1-2는 각 플랫폼의 편집 화면의 환경과 기능을 분석한 내용으로 세 플랫폼 모두 숏폼을 편집하면서 많이 활용되는 필터나 효과를 모바일에서 직접 제작하거나 검색할 수 있는 기능은 따로 제공하지 않았다. 유튜브 숏츠를 제외한 두 플랫폼의 사용자가 원하는 필터나 효과를 제작하려면 데스크탑 PC를 켜고 전문적 AR 필터 편집 플랫폼을 통해 제작한 후 인증절차를 거쳐 모바일에 연동되는 구조이며, 유튜브 숏츠는 이 서비스 또한 제공하고 있지 않았다. 인스타그램 릴스는 필터 및 효과 제작을 위한 외부 필터로 모회사인 메타(Meta)에서 제공하는 Meta Spark Studio를 사용하고 있으며 틱톡은 자체 AR 필터 제작 플랫폼인 Effect House에서 필터 제작 후 모바일 숏폼 플랫폼에 적용할 수 있다. 세 번째로, 그림 1-3은 전체 서비스의 장점과 문제점을 분석한 내용이다. 유튜브 숏츠는 유튜브와 연결되어 크리에이터가 보다 경쟁력 있게 성장할 수 있는 생태계를 제공한다는 장점이 있다. 하지만 편집 화면에서 제공하는 영상 편집 기능이 적어 다른 숏폼 플랫폼에 비해 콘텐츠 제작 환경이 좋지 않음을 알 수 있다. 마찬가지로 인스타그램 릴스도 메타 소속의 서비스인 페이스북과 연동되어 크리에이터의 긍정적인 성장 환경을 제공한다는 장점이 있다. 또한, 인스타그램 릴스는 홈 화면의 피드 또는 탐색 탭에서 릴스로 노출될 수 있어 다른 사용자와의 접근성이 유용하다는 장점이 있다. 하지만 시청 환경에선 원하는 카테고리의 영상을 볼 수 없다는 어려움이 있으며, 제작 환경에 있어서 틱톡 대비 유튜브 숏츠와 마찬가지로 영상편집 기능이 적은 불편함이 있다. 마지막으로, 틱톡의 장점은 숏폼 시청 알고리즘을 사용해 사용자에게 최적화된 몰입감 있는 시청 환경을 제공한다는 점이다. 하지만 편집 화면은 기능이 너무 많아 복잡하고 적용하기 어려운 문제점이 있다. 그 외 숏폼 기획 기능 제공 여부 측면에서의 세 개 플랫폼을 분석해 보면 모두 숏폼 콘텐츠를 기획할 수 있는 기능을 제공하지 않아 특히 초보 크리에이터들에게는 숏폼 콘텐츠 스토리 구성 기획을 포함한 앞단 도입 단계의 어려움이 있음을 알 수 있다. 이상의 분석 내용을 종합해 보면 크리에이터가 모바일 환경에서 숏폼 콘텐츠를 제작할 때 다음의 한계점을 발견할 수 있다. 첫째, 숏폼 시청 화면에

위치되어 있는 콘텐츠 재제작 기능은 기존의 영상을 가져와서 붙이거나 같은 템플릿을 복사하는 정도의 제한된 기능만 제공하여 크리에이터의 창의적 숏폼 콘텐츠 기획방법 시도의 공간이 부재하다. 둘째, 편집 화면에서는 편집기능과 함께 제공된 다수의 필터 및 특수효과를 통해 영상 콘텐츠 제작에 반영은 가능하나, 모바일 서비스에서는 아직 자체적으로 원하는 필터 및 효과를 제작할 수 없어 PC를 거쳐 모바일로 연계해야 하는 환경이다. 따라서 본 연구는 창의적 숏폼 콘텐츠 기획과 아이디어 구성이 가능한 공간을 마련하고자 하며, 생성형 AI를 도입하여 PC를 거치지 않고 모바일 서비스에서도 자체적으로 원하는 필터 및 효과를 프롬프트 기반으로 쉽고 직관적으로 제작, 적용할 수 있는 서비스를 제안하고자 한다.

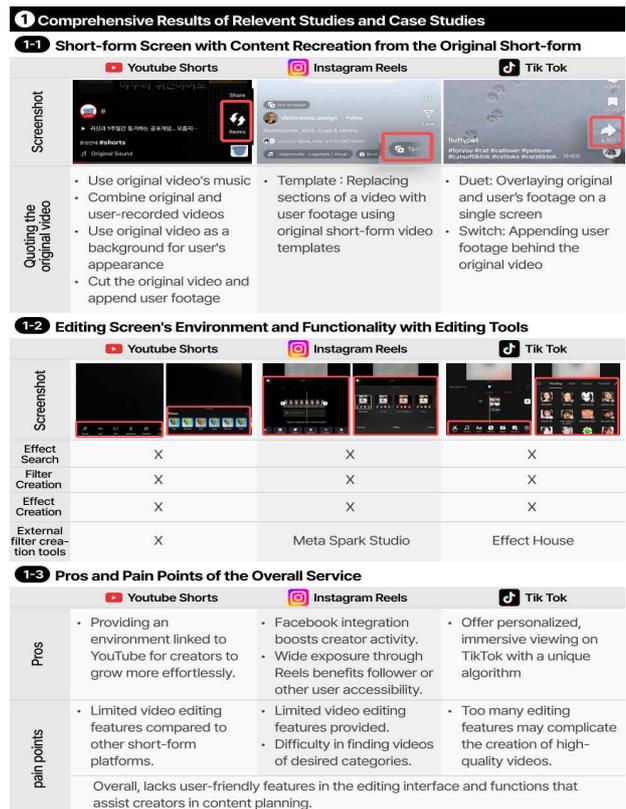


그림 1. 대표 숏폼 3개 플랫폼 사례 분석  
Fig. 1. Analysis of three major short-form platforms

#### 4) 숏폼 크리에이터를 위한 숏폼 제작 플랫폼 기획

콘텐츠 크리에이터는 본인이 원하는 장소에서 모바일 기기로 콘텐츠를 촬영하며 자유로운 포맷으로 영상을 제작, 배포하는 사람들을 일컫는다. 최근 많은 콘텐츠 크리에이터가 1인으로 활동하고 있으므로[6], 본 연구는 이에 맞춰 개인 사용자도 쉽고 직관적으로 숏폼을 제작할 수 있는 숏폼 제작 플랫폼을 기획하고자 한다. 이를 위해 영상 제작 전문 지식이 없더라도 1인 크리에이터가 모바일에서 쉽게 숏폼 영상을 제작할 수 있도록 아이디어 기획 시점에서 타임라인 Cut과 Scene의 적용 과정을 직관적 UI로 제공하고, 또 각 단계에 적합한

생성형 AI의 적용을 통해 숏폼 제작 플랫폼의 진입장벽을 낮추고자 하였다. 그림 2는 한상욱의 연구와[7] 이수진의[8] 연구를 참고하여 숏폼 크리에이터의 작업 단계를 재구성해본 내용으로 다음 단계를 포함한다. 첫째, 어떤 콘텐츠를 제작할지 구상하는 기획 단계는 콘셉트를 구상하고 스크립트를 작성한다. 둘째, 콘텐츠를 촬영하는 촬영 단계에서는 스크립트 기반 숏폼 프리뷰를 참고하며 장면을 촬영한다. 셋째, 편집 단계는 영상에 자막을 추가하거나 특수효과를 적용하며 편집하는 단계이다. 본 연구가 제안하는 모바일 기반 숏폼 제작 플랫폼에 적용하고자 하는 생성형 AI는 사전 학습된 모델(pre-trained model)의 인공지능망을 통하여 사용자의 특정한 요구에 맞춰 능동적으로 새로운 결과를 생성해 내는 기술로[9] 위 분류 단계에 적합한 생성형 AI의 모델을 기술하면 다음과 같다. 먼저 숏폼 콘텐츠의 기획 단계는 자연어 처리 작업에 특화된 인공지능 모델인 LLM 모델(LLM, Large Language Model)이 주로 사용될 수 있다. 이는 사용자가 입력한 텍스트를 이해하여 결과값으로 텍스트를 생성하고 번역, 요약, 질의응답 하는 다양한 언어 관련 작업이 가능하므로 크리에이터가 기록한 아이디어의 문맥을 이해하고 조합하여 새로운 아이디어를 추천하는 형태의 서비스가 가능하다. 또한, 기획 단계에서도 글 외에 이미지 생성을 통한 아이디어 발상할 수 있는데 이 경우에는 이미지 및 영상 생성 서비스에 사용되는 T2I 모델(Text-to-Image), T2V(Text to Video) 모델의 사용을 통해 명령어 입력에 따른 이미지 및 영상 생성이 가능하다. 이 T2I, T2V는 세 번째 제작 및 편집 단계의 주요 생성 방식으로 활용될 수 있다. 두 번째 촬영 단계 시점의 경우 본 연구 제안 범위에서는 생성형 AI 역할이 크지 않으나 크리에이터가 직접 출연하여 촬영할 때 조명 조도 포함 각종 영상 장비의 스마트 장비 컨트롤에도 적용될 수 있다.

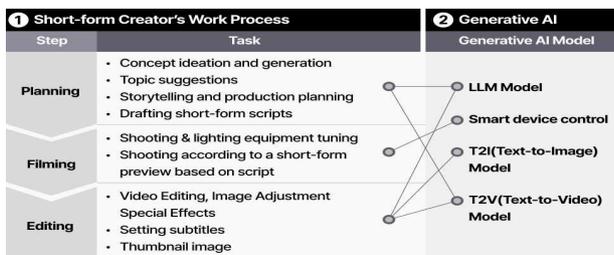


그림 2. 숏폼 크리에이터의 작업 단계별 생성형 AI 모델  
 Fig. 2. Short-form creator's work steps using gen. AI

2-2 생성형 AI 연계 텍스트, 이미지, 영상 생성 서비스 사례

1) 숏폼 콘텐츠 기획을 위한 LLM 모델 및 서비스 분석

숏폼 콘텐츠 기획을 도울 수 있는 사례로 앞서 언급한 LLM 모델기반 생성형 AI 서비스를 분석하였다. 분석 대상으로 대화형 AI인 오픈 AI의 챗 GPT(Chat GPT 4)와 마이크로소프트의 코파일럿(Copilot)을 선정하였고, 서비스의 특징과 장단점을 분석 도구로 분석하였다. 2개의 LLM 서비스 모두

대화형 인터페이스를 제공하며, 입력창에 프롬프트를 입력하면 해석된 생성 답변을 얻을 수 있다. 먼저 챗GPT는 오픈 AI가 출시한 2023년 1년간 가장 많은 관심을 받은 AI 서비스로 [10] 플러그인을 활용하여 실시간 정보를 포함한 결과를 얻을 수 있으며, 정확도가 높고 응용 범주가 다양하다. 장점으로 상담과 멘토링, 컨설팅처럼 피드백을 받을 수 있는 직업군과 대화하며 앞뒤 맥락을 연결한 뒤 결과를 도출해 논리적인 답변을 얻을 수 있다[11]. 하지만 학습된 데이터로만 맞춰진 결과물을 제공할 수도 있는 환각 오류(hallucination)와 같이 사실이 아닌 답변을 얻을 수 있는 단점이 있다. 마이크로소프트의 대화형 AI인 코파일럿은 오픈 AI의 LLM 모델인 GPT 4 기반의 프로메테우스 모델을 사용하는 인공지능 검색엔진이다. 장점으로 실시간으로 복잡한 다단계 웹 검색을 수행하여 논리가 필요한 질문에 답변할 수 있는 부분과 대부분 기능이 무료인 장점이 있다[11]. 반면, 챗GPT와 마찬가지로 잘못된 답변을 얻을 수 있는 단점이 있다.

2) 생성형 AI기반 이미지 및 영상 생성 서비스 분석

숏폼 콘텐츠 제작 과정에 생성형 AI를 활용하면 텍스트 프롬프트로 입력해가며 생성된 이미지를 통해 영감을 받는 것은 물론 세부 프롬프트 작성으로 생성된 이미지를 부분 수정(Fine-Tuning)해가며 실제 활용 가능한 이미지를 제작할 수 있다. 이에, 이미지 및 영상 생성 AI 서비스의 사례 분석을 통해 특징과 장단점을 분석하였다. 분석 대상은 총 네 개로, 이미지 생성형 AI 서비스인 어도비 파이어플라이(Adobe Firefly)와 이미지 및 영상 생성형 AI 서비스인 런웨이(Runway Gen2.0), 그리고 영상 생성 AI 모델인 루미에르(Lumiere)와 소라(Sora)이다. 첫 번째 어도비 파이어플라이는 어도비에서 출시한 이미지 생성 AI 서비스로 크리에이티브 클라우드(Creative Cloud) 제품에서 사용할 수 있다. 그중 포토샵에서는 텍스트 프롬프트를 이용해 기존 이미지를 자연스럽게 확장하거나 삭제하며 이미지를 편집할 수 있는 생성형 채우기(Generative Fill) 기능을 제공한다. 하지만 일부 신체를 표현하는 프롬프트의 결과물이 실제와 다르게 왜곡 표현되는 한계가 있다. 두 번째로 런웨이는 영상 생성 AI 서비스로 텍스트 프롬프트뿐만 아니라 I2V(Image to Video)와 V2V(Video to Video)를 사용해 영상을 제작할 수 있다. 텍스트와 이미지를 함께 혹은 각각 첨부하여 영상 제작이 가능하다. 세부 조정을 위한 설정값으로 카메라의 움직임(Camera Motion)과 모션이 적용되는 범위를 설정할 수 있는 모션브러쉬(Motion Brush)를 사용할 수 있다. 또한, 3D 만화나 일러스트와 같은 영상 스타일 설정이 가능하다. 몇 개의 단어나 문장으로 영상 제작이 가능하다는 장점 대비 이후 16초로 영상의 길이를 연장할 수 있지만 처음 제작된 영상의 길이가 4초 내로 제한된다는 점과, 높은 품질의 영상을 제작하려면 상세한 프롬프트와 옵션값을 조절해야 하므로 초심자에게는 콘텐츠를 제작하기에 부담스러울 수 있는 단점이 있다. 하지만, 최근에는 영상 생성 AI 기술의 발전으로 더 자연스러

운 영상 제작이 가능해졌다. 생성형 AI로 영상 제작 시 TSR 모델(Temporal Super-Resolution)이 영상의 시작과 끝에 주요 키 프레임 이미지를 생성한 뒤 키 프레임 사이 빈 곳에 겹치지 않는 이미지의 데이터를 생성해 프레임에 이어 붙여 영상으로 제작하는 방법이[12] 사용되고 있다. 세 번째, 구글의 루미에르 모델은 기존의 영상 생성 방식과는 다르게 결과물을 한 번에 생성해 내는 시공간 U-넷(Space-Time U-Net)모델을 사용하여 사전 훈련된 텍스트-이미지 모델을 활용해 비디오 속 사물의 위치와 움직임을 동시에 고려하여 [13] 더욱 자연스러운 영상 제작이 가능해졌다. 네 번째, 오픈 AI의 영상 생성형 AI 모델인 소라 또한 T2V 방식으로 프롬프트를 입력하면 최대 60초의 영상을 제작할 수 있다. 소라는 전체 동영상을 한 번에 생성하는 방식을 사용하며 생성된 영상 길이를 확장할 수도 있다. 기존의 영상 생성 AI 모델보다 훨씬 길어진 영상의 길이와 자연스러운 결과물을 제공한다는 장점이 있다[14]. 하지만 현재까지 발표된 모델들은 앞에는 한 입 먹은 과자가 뒤에는 그대로 남아있는 것과 같이 현실의 물리적 인과관계를 반영한 결과물을 제대로 생성하지 못하는 경우도 생기는 단점이 존재한다. 앞선 분석을 바탕으로 본 연구는 영상을 생성할 때 구체적인 프롬프트를 작성해야 하는 입문자의 어려움을 보완하고자, LLM 모델을 활용해 숏폼의 내용을 스크립트 형태의 프롬프트로 제공하고, 이를 바탕으로 영상 생성 AI를 활용해 그 프롬프트를 중심으로 영상을 제작 및 편집할 수 있도록 지원하고자 한다.

### 3) 텍스트 프롬프트 기반 영상 편집 서비스

최근 프롬프트를 입력하여 더욱 쉽게 영상 편집이 가능한 툴들이 속속 출시되고 있다. 대표 사례로 영상 생성과 편집이 가능한 소라, LTX Studio, 그리고 어도비 프리미어 프로(Adobe Premiere Pro)의 특징을 분석하였다. 그림 3의 1-1과 같이 앞서 언급한 소라는 프롬프트를 입력하여 인풋 비디오의 환경과 스타일을 자연스럽게 변경할 수 있는 V2V editing 기능과 1-2와 같이 전혀 다른 두 개의 인풋 비디오 A와 인풋 비디오 B를 자연스럽게 합친 영상을 생성할 수 있는 Connecting videos 기능을 제공한다. 2의 LTX Studio 서비스는 스크립트부터 편집까지 간소화된 생성형 영상 제작 프로세스를 경험할 수 있는 PC 전용 플랫폼이다. 2-1과 같이 Scene과 Cut으로 나누어 영상 제작 및 편집을 할 수 있으며, Cut마다 프롬프트를 수정하면 세부 편집이 가능하다. 또한, 2-2에서 볼 수 있듯이 영상의 타임라인에서 생성된 배경음악과 나레이션, 음악 효과를 편집할 수 있다. 마지막으로, 프리미어 프로는 어도비의 영상 편집 전문 툴로 프롬프트와 타임라인 중심의 편집이 가능하다. 음성 인식 자막 기능을 통해 스크립트를 자동생성하며 타임라인에서 텍스트 프롬프트로 영상편집이 가능하다. 3-1과 같이 스크립트에서 텍스트를 선택하면 타임라인에서 동일 부분이 활성화되어 영상 클립을 편집할 수 있다. 3-2의 Generative Extend는 영상 생성 AI 모델인 런웨이, 피카(Pika), 프리미어 프로에서 끊어진 이후

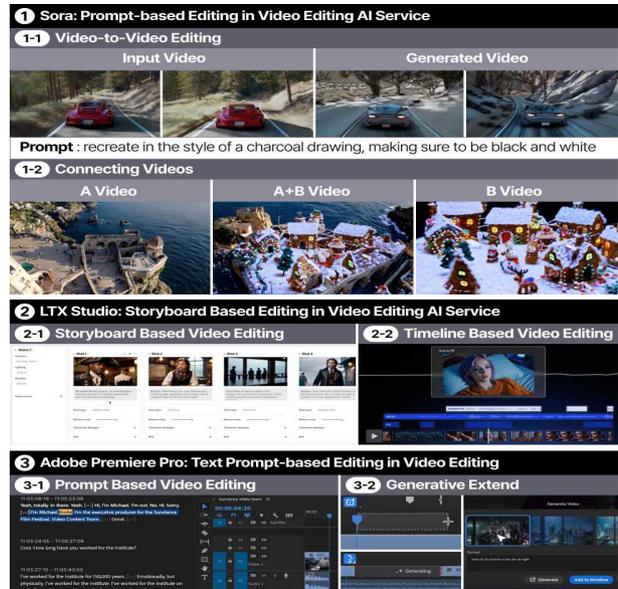


그림 3. 프롬프트 기반 편집 및 영상 합성 서비스 예시  
Fig. 3. Prompt-based video editing and compositing examples

영상을 앞 영상에 자연스럽게 연장 생성하는 기능이다.

### 2-3 관련 연구 종합 분석

관련 연구를 종합 분석해 본 내용은 다음과 같다. 첫째, 숏폼은 MZ세대 사이에서 인기가 많지만, 실제 제작 및 업로드 경험은 시청 경험을 따라가지 못해 이에 대한 니즈가 있는 대상자들을 위한 접근성 개선이 필요하다. 둘째, 기존 대표 세계 숏폼 플랫폼을 연계 분석한 결과, 숏폼 시청화면에서 다른 크리에이터의 영상 재제작 기능을 통해 콘텐츠 기획의 도움 받기는 가능하나, 기존 영상 템플릿에 사용자 영상을 갈아 끼우는 방식으로 사용자의 창의성을 담기엔 한계가 있다. 또한, 편집 화면에서 다수의 필터 및 특수효과를 제공하고 있지만, 효과 제작 기능과 검색 기능의 부재로 원하는 편집 효과를 적재적소에 배치하기에는 어려움이 있다. 셋째, 기존 영상 제작 툴에 익숙하지 않은 대상자를 위해 생성형 AI의 도움을 받아 콘텐츠를 효율적으로 제작할 수 있도록 숏폼 크리에이터의 작업 단계별 생성형 AI 모델 기술 도입 환경 조성이 필요하다. 넷째, 텍스트, 이미지, 영상 생성 기반 선행 서비스 분석 결과 숏폼 콘텐츠 기획을 위한 LLM 텍스트 모델의 경우 앞뒤 맥락을 고려한 논리적인 답변을 얻을 수 있다는 장점 대비 어떻게 결과를 도출했는지에 대한 명확한 설명을 제공하지 못하는 한계가 있음을 알 수 있다. 또한, 이미지 및 영상 생성 서비스는 최근 발표된 구글의 루미에르, 오픈 AI의 소라, LTX Studio와 어도비 프리미어 프로의 경우 T2V 방식으로 혁신적 영상 편집의 가능성을 열고 있으나 아직 상용화 전인 서비스가 많다. 본 연구는 이상의 종합 분석 내용을 바탕으로 숏폼 제작 및 업로드 경험 개선을 위해 기존 영상 템플릿 갈

아끼우기 방식이 아닌 선행 숏폼의 창의적 인사이트를 본인 숏폼 콘텐츠 기획에 연동되도록 기획하고자 한다. 또한, 숏폼 제작 단계에 맞춰진 Scene과 Cut, 그리고 개체 편집에 이르기까지 스크립트 편집과 연동된 생성형 AI 프롬프트로 숏폼 크리에이터에게 적재적소 편집에 최적화된 모바일 숏폼 제작 플랫폼을 제안하고자 한다.

### III. 생성형 AI 기반 프롬프트를 활용한 차세대 모바일 숏폼 제작 플랫폼 제안

#### 3-1 설문 조사를 통한 사용자 니즈 발견

그림 4는 숏폼 콘텐츠를 제작하는 프로세스와 그 과정에서의 불편함을 파악하기 위해 2023년 5월 24일부터 5월 27일까지 3일간 구글 폼(Google Form)을 이용한 숏폼 소비 및 제작 관련 온라인 설문 조사를 진행한 결과의 정리 내용이다. 대상자는 숏폼 콘텐츠를 한 번이라도 시청한 경험이 있는 10대에서 30대 79명 대상이며, 성비는 남성이 21명, 여성이 58명이다. 연령대는 20대가 중심으로 72명, 30대가 4명, 10대가 3명이며, 응답자의 직업은 학생이 61명으로 가장 많고, 직장인, 프리랜서 순으로 뒤를 이었다.

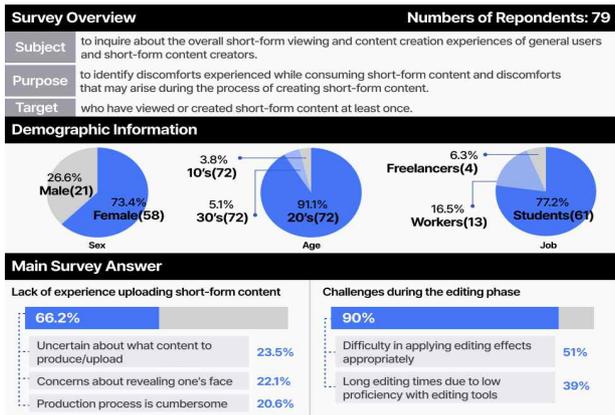


그림 4. 설문 조사 개요 및 피험자 구성  
Fig. 4. Survey overview and participant demographics

#### 1) 숏폼 콘텐츠 소비 및 제작 경험의 설문결과 분석

숏폼 콘텐츠를 한 번이라도 시청한 경험이 있는 피험자를 대상으로 하였으나 그중 87.2%의 응답자가 직접 숏폼 콘텐츠를 업로드해 본 경험은 없다고 답하였다. 또한, 응답자의 66.2%가 숏폼 콘텐츠 업로드에는 관심은 있지만, 선택 도전은 어렵다고 응답하였다. 이들이 숏폼 제작 도전 경험과 업로드 경험이 없는 주요 이유는 다음과 같다. 첫째, 어떤 콘텐츠를 제작하고 업로드 할지 모르겠다는 '콘텐츠 기획 및 스토리 구성의 어려움'과 관련된 의견이 23.5%로 가장 많았고 촬영 등으로 얼굴이 공개되는 게 부담스럽다는 의견이 22.1%로

다음 순위였으며, 제작하는 과정이 번거롭다는 의견이 20.6%로 모바일에서 편집 툴을 활용하는 것에 어려움을 느끼고 있었다. 또한, 숏폼 제작 유경험자의 답변 중 콘텐츠를 만들기 위해 가장 먼저 하는 일은 40%가 인기 있는 콘텐츠 조사를, 그리고 20%가 컨셉 설정으로 응답했으며, 제작 과정 중 편집과 후보정 단계에서 가장 많은 시간이 소요된다고 답한 응답자는 60%였다. 아이디어를 구상할 때 어려움을 겪었다고 답한 응답자 모두가 콘텐츠의 기획과 제작 사이의 간극을 느꼈으며 아이디어를 창의적으로 발상하는 데 어려움을 느끼고 있었다. 그 외 창의적인 아이디어를 구상하고 구체화하는 과정에서 간단한 영상을 먼저 만들어 볼 수 있는 도구나 도움이 필요하다는 의견도 있었다. 또한, 아이디어 전개 단계에서 다른 크리에이터들의 의견 또는 아이디어를 참고해 도움을 받길 원하는 응답자가 75%로 대다수 응답자가 스스로 콘텐츠를 구상하고 진행하는 데 어려움을 겪고 있음을 알 수 있었다. 촬영 단계에서는 응답자의 67%가 어려움을 겪고 있었으며, 모바일 편집 안내를 받는다면 촬영 경험이 개선될 것 같다는 추가 의견을 확인할 수 있었다. 편집 단계에서는 응답자의 72%가 적재적소에 맞는 필터 및 편집 효과를 찾는 어려움을 느낀다고 답하였고, 숏폼 콘텐츠를 편집하는 과정에서 툴의 숙련도가 낮고 편집에 능숙하지 않아 시간이 많이 소요되므로 편집을 쉽게 도전하기에는 어려움을 느끼고 있었다.

#### 2) 설문결과 해석 및 주요 인사이트

설문 조사를 바탕으로 다음과 같은 인사이트를 도출할 수 있었다. 첫째, 숏폼 콘텐츠의 기획에 어려움을 느끼는 대상자가 많으므로 사용자가 창의적 숏폼 콘텐츠 아이디어 발상을 도울 수 있는 아이디어 기획 보드와 같은 기능 제공이 필요함을 알 수 있었다. 둘째, 숏폼 콘텐츠 제작에 관심 있지만 숏폼 업로드 경험이 없는 대상자들은 아이디어 기획 단계에서부터 다른 크리에이터들의 도움을 받길 원한다는 의견을 반영하여 크리에이터 간 숏폼 콘텐츠 및 제작 팁 자료를 공유하는 오픈 소스 생태계 도입의 필요성을 발견할 수 있었다. 또한, 창의적 주제 선정 및 스토리 기획의 어려움이 많아 이를 개선할 수 있는 서비스 도입이 필요함을 발견하였다. 셋째, 촬영 경험의 개선을 원하는 응답자의 의견을 반영해 촬영이 익숙하지 않은 초보 크리에이터의 촬영 경험을 긍정적으로 개선하기 위해 조명, 카메라 각도, 앵글 등 촬영 과정에서 도움을 줄 수 있는 장비 조절 및 촬영 가이드 메시지의 필요성을 알 수 있었다. 또한 '얼굴이 공개되는 것이 부담스럽다'라는 의견을 반영하여 사용자의 얼굴 노출 없이 콘텐츠를 제작할 수 있는 기능, 또한, 사용자 선호에 맞는 각종 스타일 변경 필터 기능 제공이 필요함을 알 수 있었다. 넷째, 기존 숏폼 플랫폼 사용자는 숏폼 편집 과정에서 어려움을 느끼고 있었으며, 특히 툴의 숙련도가 낮고 편집에 능숙하지 않아 원하는 편집 효과 적용의 어려움에 따른 시간 소요가 큼을 알 수 있었다. 이에 영상 편집에 전문적인 지식이 없고, 영상편집 인터페이스가 익숙하지 않은 대상자도 숏폼 기획 및 제작 단계별 생성형 AI 도움

을 통해 쉽고 직관적으로 편집이 진행될 수 있게 숏폼 편집 UI의 개선이 필요함을 발견하였다.

### 3-2 서비스 기획 및 콘셉트 도출

#### 1) 브랜드 네임 기획

본 서비스의 이름은 와티(WHATI)로, 무엇(What)과 아티스트(Artist)의 뜻을 더해 '무엇'이든 만들 수 있는 크리에이터들의 공간이라는 뜻을 담고 있으며, 와티에서 숏폼을 제작하는 숏폼 콘텐츠 크리에이터를 와티스트(WHATIST)라고 지칭한다. 또한, 숏폼을 기획하는 창의적 메모장 형태의 보드를 와티보드(WHATI BOARD)로 명명하였다.

#### 2) 서비스 다이어그램

2장 선행사례 분석과 설문 조사의 주요 인사이트를 바탕으로 본 제안 서비스의 핵심 콘셉트를 프롬프트와 타임라인 편집을 연계하여 숏폼 콘텐츠의 기획부터 완성에 이르기까지 도와주는 생성형 AI 기반 차세대 숏폼 제작 플랫폼으로 정하였다. 그림 5는 숏폼의 제작 과정을 기획하기, 촬영하기, 편집 및 제작하기로 나누어 정리한 본 제안 연구의 서비스 다이어그램이다. 숏폼을 제작하기 위해 와티보드의 메모를 조합하여 스크립트를 작성한 후 미리보기 영상을 만들어 촬영을 진행하거나, 편집하는 과정을 거쳐 숏폼을 완성할 수 있다.

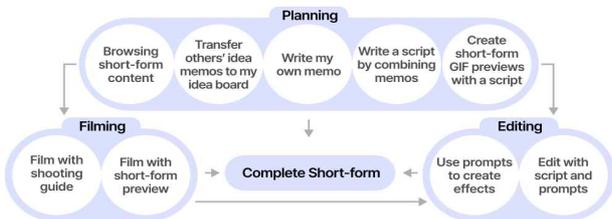


그림 5. 서비스 과정 다이어그램  
Fig. 5. Service flow diagram

#### 3) 서비스 핵심 가치 기반 UI

그림 6은 본 서비스의 핵심 가치 기반 UI이다. 메인 기기는 갤럭시 Z 폴드 4이며, 웹툰 스크롤 인터페이스의 인사이트를 바탕으로 새로운 세로형 숏폼 Scene 및 Cut 영상편집에 특화된 인터페이스 제안에 중점을 두었다. 1의 좌측 커버디스플레이에서는 하단에 축약된 핵심 편집기능을 제공하며, 2의 언폴드 화면에서는 좌우 2분할 화면을 통해 좌측 숏폼 영상의 Cut 단위 세부 편집이 가능한 타임라인 인터페이스와 우측 생성형 AI 프롬프트 기반 Scene 단위 전체 숏폼 시나리오 스크립트 창을 비교하며 작업하기에 편리하게 배치하였다. 우측 스크립트 창의 스크립트는 숏폼 스토리 기획 보드인 와티보드에서 사용자가 작성한 스크립트의 내용으로 A와 같이 현 작업 중인 Scene의 하단에 있는 프롬프트 입력창 기입을 통해 전체 Scene의 배경, 등장인물, 오디오 효과 및 대사를 A'와 같이 타임라인 편집 화면에 생성시킬 수 있다. 이후 B처럼

Scene의 하위 구성요소인 Cut에서 원하는 개체 이미지를 선택한 후 B'의 Cut 프롬프트 입력 창을 통해 Cut단위 개체의 상세 수정 생성을 할 수 있다. a, b, c는 각각 Scene 단위의 이미지, 사운드 효과, 대사 생성에 관한 내용으로 좌측 타임라인의 a', b', c'와 연동되어 있다. 선행사례와의 비교 실험을 통해 그중 a는 런웨이, 피카, 프리미어 프로의 Generative Extend 기능과 LTX studio의 컷 편집 및 생성 기능, 소라의 A+B영상 조합 기능을 본 연구의 폴더블 UI에 최적화시킨 Generative Extend 기능이며 끊긴 이후 영상을 앞 영상과 자연스럽게 연장하기, 컷 간의 연결 및 생성, A+B 영상 조합 후 연장 생성하기 옵션을 사용자가 선택하여 생성할 수 있다. 3은 이와 같은 장면 생성 및 편집에 대한 알고리즘을 선행 연구인 구글의 루미에르[13]와 오픈 AI 소라의 영상 생성 모델 [14]을 바탕으로 재구성한 내용이다. 먼저, 프롬프트를 입력하면 LLM 모델에 제공되며 re-captioning technique을 통해 짧은 프롬프트를 상세 프롬프트로 묘사한 text caption을 산출하여 T2I와 T2V 모델에 전달한다. 이후 영상의 전처리 단계 없이 여러 시각 데이터 배열을 체계적으로 처리할 수 있는 Spacetime Latent Patches로 전환된 후 비디오 시퀀스 전체의 공간적, 시간적 관계의 이해를 바탕으로 객체가 3차원 공간에서 가려지거나 프레임 밖으로 벗어나는 경우와 같이 여러 인과관계가 반영된 영상을 제작 및 편집할 수 있다.

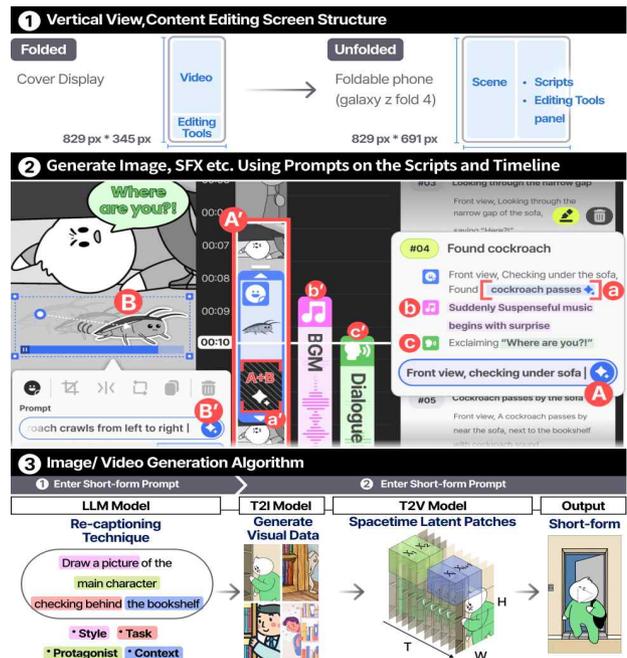


그림 6. 서비스 핵심 아이디어  
Fig. 6. Service core ideas

#### 4) 고객 여정 지도

그림 7은 고객 여정 지도(Customer Journey Map)로 본 연구의 제안시나리오를 구체화하기 위해 숏폼 제작 과정에 따라 작성한 내용이다. 앞선 그림 2의 숏폼 크리에이터의 작

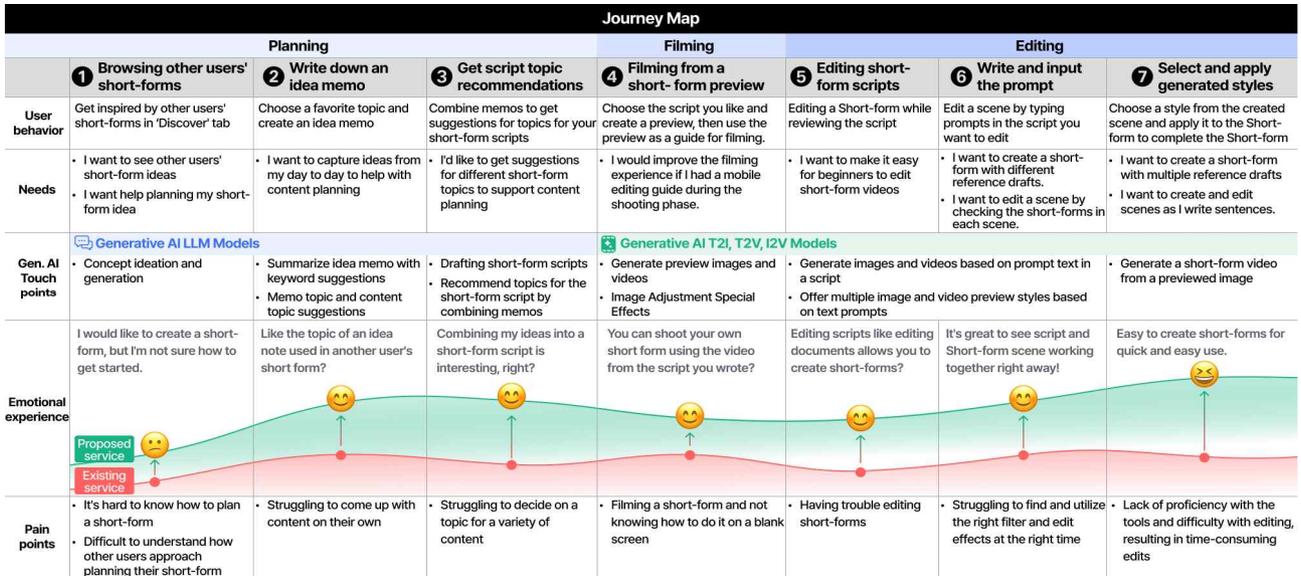


그림 7. 고객 여정 지도  
Fig. 7. Customer journey map

업 단계에 따라 기획, 촬영, 편집으로 나눈 후 생성형 AI 사용자 터치포인트 및 숏폼 제작 과정별 페인 포인트(Pain Points)와 니즈(Needs)를 도출하였다.

5) 서비스 주요 기능 정의 및 유저 플로우 기반 시나리오

그림 8은 앞선 선행사례 분석 및 설문 조사 결과에 따른 인사이트를 바탕으로 제작한 본 서비스 주요 기능의 유저 플로우(User Flow)기반 시나리오이다. 주요 기능은 크게 숏폼 콘텐츠 기획하기와 제작하기로 나뉜다. 1은 기획 단계로 콘텐츠의 아이디어선과 기획, 그리고 기록에 도움을 주는 본 연구가 명명한 '와티보드'활용 단계이다. 2는 숏폼 제작 단계로, 2단

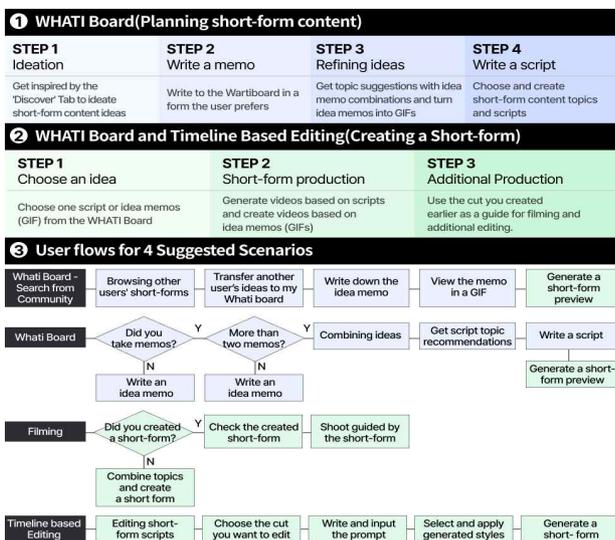


그림 8. 숏폼 기획과 제작과정의 유저 플로우 기반 4개 시나리오  
Fig. 8. 4 suggested scenarios for planning and creating short-forms based on user flows

계와 같이 작성한 스크립트 또는 아이디어 메모를 선택하여 소라, 루미에르와 같이 프롬프트 기반으로 숏폼 제작도 가능하며, 3단계는 언급한 LTX Studio, 그리고 프리미어 프로 사례처럼 Scene과 Cut의 타임라인 조절과 생성을 연계하여 쉽게 제작할 수도 있다. 3은 앞선 설문 조사를 바탕으로 유저플로우 기반 4개 시나리오를 정리한 그림이다. 4개 사용자 유형은 다음과 같다. 첫째, 다른 크리에이터의 숏폼을 참고해 아이디어를 기록하고, 아이디어 메모로 만들어진 GIF와 같은 형식의 4초 이내의 짧은 영상을 숏폼으로 제작하고 싶은 사용자이다. 둘째, 아이디어 메모를 작성하여 숏폼을 생성하고 싶은 사용자이다. 셋째, 생성한 숏폼을 참고해 직접 촬영하고 싶은 사용자이며, 넷째, 생성한 숏폼의 프롬프트를 편집하며 타임라인 기반 숏폼을 제작하고 싶은 사용자이다. 이를 바탕으로 4가지 사용 단계별 유저 플로우 기반 시나리오를 각 핵심 기능이 담긴 프로토타입으로 제작하였다.

6) 브랜드 디자인 콘셉트

그림 9는 본 서비스의 브랜드 디자인 콘셉트이다. 로고는 'H'의 스템(Stem)을 괄호로 표현하여 '아이디어'를 입력할 수 있는 프롬프트 입력창을 형상화하였으며, 크로스바(Crossbar)를 '◆'로 표현하여 생성형 AI를 통한 숏폼 콘텐츠 기획 및 제작을 할 수 있음의 콘셉트를 전달하고자 하였다. 본 서비스의 브랜드 경험 키워드는 세 가지로, 다채로운 콘텐츠 제작 경험 제공 측면의 '활기찬(Dynamic)'과, 다른 와티스트와의 즐거운 숏폼 콘텐츠 공유 경험 제공 측면의 '즐거움을 주는(Entertaining)', 그리고 더 쉬운 숏폼 제작 프로세스를 통한 창작 경험을 제공한다는 측면의 '창작하다(Create)'이다. 메인 컬러는 생성형 AI의 이미지에서 떠오르는 창의성과 성장을 표현할 수 있는 네온 그린 컬러로 선정하였다. 컬러

모드는 숏폼 제작과 시청 몰입 유도를 위한 다크 모드를 주 모드로 하였으나, 편집 화면의 텍스트 기반 숏폼 스크립트 창은 가독성을 위해 라이트 모드를 함께 사용하였다.



그림 9. 브랜드 디자인 콘셉트  
Fig. 9. Brand design concepts

### 3-3 유저 시나리오 기반 서비스 핵심 기능 인터페이스디자인

#### 1) 다른 유저의 숏폼 콘텐츠 참고가 가능한 오픈 생태계를 활용한 메모 기능 기반 숏폼 기획 단계

그림 10의 1은 프롬프트를 바로 작성하기에 어려움이 있고 사용자가 '와티보드'에서 토글을 켜거나 버튼을 탭하여 아이디어를 기록하는 데 도움을 받을 수 있는 둘러보기 모드에 진입하는 장면이다. '둘러보기'는 피그마(Figma)와 같이 오픈 커뮤니티 생태계이며 사용자는 오픈된 다른 사용자의 아이디어 발상과 숏폼 콘텐츠 기획 과정을 참고할 수 있다. 2는 MZ세대에서 인기 있는 유머/개그 카테고리[2] 중 2023년 8월 24일 업로드 이후 약 110만 조회수를 기록한 여름날 끊임없이 들리는 모기소리로 잠을 못 이루어 모기를 잡는 에피소드에 관한 숏폼 콘텐츠 예시로, 본 연구에서는 이를 오픈 커뮤니티의 하나의 콘텐츠로 가정하고 프로토타입을 진행했다. 이처럼 '둘러보기' 메뉴에서 아이디어와 기획이 흥미로운 오픈 숏폼 콘텐츠를 검색 후 해당 숏폼을 재생하며 작성된 스크립트와 스크립트 작성 시 사용된 아이디어 메모를 통해 인사이트를 얻을 수 있다. 가령, 모기 에피소드의 경우 '최근 나를 괴롭히는 또 다른 것은 무엇일까?'로 작성한 메모를 통해 새로운 소재와 주제에 대해 발상을 시작해볼 수 있다. 3은 사용자가 둘러보기를 통해 접하게 된 모기 에피소드 숏폼 스토리에 영감을 받고, 이를 바탕으로 자신의 에피소드를 숏폼 스토리 기획에 담아보는 시나리오이다. 가령, A는 메모 하단의 '와티보드로 옮기기' 버튼을 탭한 상황으로 다른 와티스트의 오픈 아이디어 메모를 옮겨와 4의 B와 같이 '최근 나를 괴롭히는 또 다른 것'을 바퀴벌레로 설정하고, 최근 이사한 자취방에 바퀴벌레가 출몰해 서로 대치했던 상황을 글과 생성형 그림으로 메모에 기록하는 과정이다. 이때 작성 완료된 메모는 원문 사용자의 저작권 보호 차원에서 아이디어 메모를 참고했음에 대한 '원본 보기' 문구가 표시된다. 4는 새로 정한 바퀴벌레 에피소드를 LLM 모델을 통해 키워드 생성 후 숏폼 프리뷰를 만드는 과정이다. 가령, '바퀴벌레가 방 안에서 출몰한다', '방안에 나타난 바퀴벌레를 슬리퍼로 때리다'와 같은 메모가 있을 경우 R. Mihalcea, P. Tarau의 선행 논문을 바탕으로 컨텍스트에서 순위가 높은 키워드 인식 알고리즘을 적용하여 [15], 식별된 주요 키워드인 '바퀴벌레', '방 안', '슬리퍼', '때

리다'는 C와 같이 메모 하단에 표시된다. 영감을 통한 아이디어 메모 작성 및 주제 기획이 중심인 이 단계에서도 D의 '숏폼 프리뷰 생성' 기능을 통해 간단히 프리뷰로 메모와 관련된 진행중인 생성 이미지를 확인할 수 있으며, 프리뷰 영상이 마음에 들면 메모 하단의 '추가 편집하기' 버튼을 탭한 후 타임라인에서 세부 편집 및 수정을 통해 숏폼으로 생성할 수 있다.

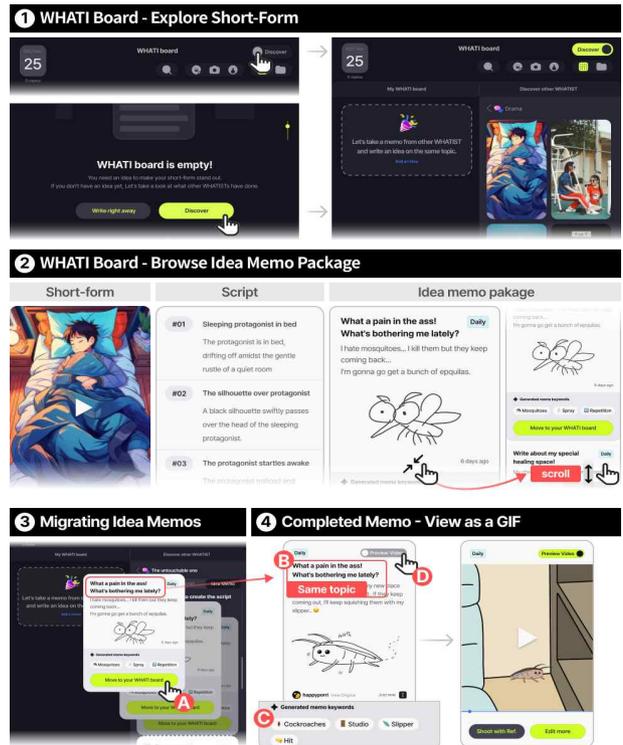


그림 10. 다른 유저의 숏폼을 참고해 아이디어 작성하기  
Fig. 10. Use other users' short-forms for inspiration

#### 2) 2개 이상의 메모 기능 조합을 통해 사용자의 창의적 주제 고도화 시도 단계

그림 11의 1은 창의적 주제 선정과 스토리 기획의 어려움이 많은 사용자를 위해 와티보드에서 작성한 2개 이상의 메모를 조합한 후 생성된 스크립트 작성을 통해 창의적인 주제를 선정하는 단계이다. 1의 A는 메모 조합의 사용 예로, 사용자가 앞서 작성했던 메모인 '바퀴벌레 출몰' 메모 위에 '내 방에서 섬뜩한 곳은?'이라는 주제의 메모를 드래그 앤 드롭하는 장면이다. 이처럼 2개의 메모가 링크되면 앞서 언급한 키워드 인식 알고리즘을 [15] 통해 A'처럼 '바퀴벌레, 방안, 슬리퍼, 책장, 때리다'와 같이 조합된 메모의 키워드가 생성되고, 사용자는 이를 주제와 스토리 기획용 스크립트에 활용할 수 있다. B는 사용자가 조합을 통해 받은 주제 중 '바퀴벌레와의 승부'가 마음에 들 경우의 예시로 C의 '스크립트 스타일 설정하기' 드롭다운 버튼을 통해 장르, 등장인물 수, 대사 방식, 영상 길이, 결말 방식과 같은 세부 조건을 설정하는 장면이다. D의 '스크립트 작성하기' 버튼을 탭하면 E처럼 선정 제목 하

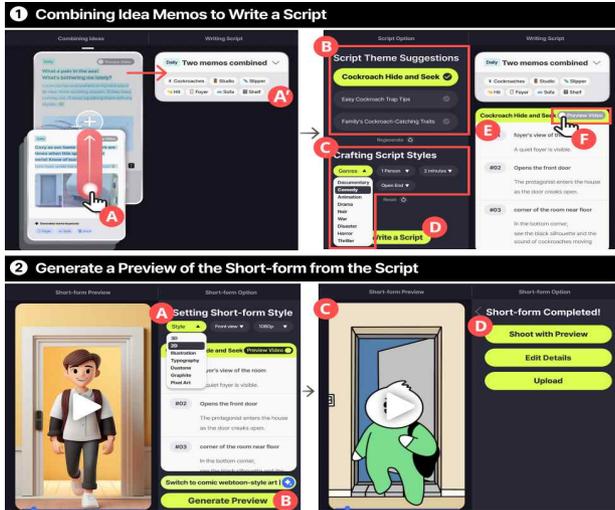


그림 11. 메모 조합 후 스크립트를 작성하여 숏폼 프리뷰 만들기  
 Fig. 11. Combine memos to create short-form preview

단에 Scene 단위로 생성된 세부 스크립트의 내용을 확인하고 수정할 수 있다. 이는 앞서 와티보드에서 단일 메모로 제작한 숏폼 프리뷰와 달리 위 언급한 세부 조건 설정 기능으로 좀 더 정교한 프리뷰 생성이 가능하며 F의 ‘숏폼 프리뷰 생성’을 통해 프리뷰 확인이 가능하다. 2는 클레이모피즈 스타일로 생성된 프리뷰를 코믹 웹툰 스타일 프리뷰로 생성하는 과정이다. A와 같이 ‘숏폼 스타일 설정하기’ 드롭다운에서 ‘2D’를 선택한 뒤 하단 프롬프트 입력창에 ‘코믹한 웹툰 스타일로 바꿔줘’라고 입력하고 B의 ‘숏폼 프리뷰 보기’ 버튼을 탭하면 C처럼 새롭게 생성된 프리뷰를 확인할 수 있다. 이후 D의 ‘참고해서 촬영하기’와 ‘추가 편집하기’, ‘바로 업로드하기’의 세 개 버튼을 통해 사용자 니즈별 서비스를 제공 받을 수 있다.

3) 와티보드를 통해 생성한 숏폼을 참고해 직접 촬영 후 숏폼 제작을 시도하고자 하는 단계

그림 12는 촬영을 통해 숏폼을 제작하고 싶은 사용자가 숏폼 콘텐츠를 촬영 후 반영하는 장면이다. 그림 12의 1은 촬영하기 화면으로, 사용자는 앞서 제작한 숏폼 프리뷰 영상을 가이드 삼아 숏폼 촬영이 가능하다. 좌측 화면은 촬영 화면이고 우측화면은 숏폼 가이드 프리뷰와 촬영하려는 Scene의 스크

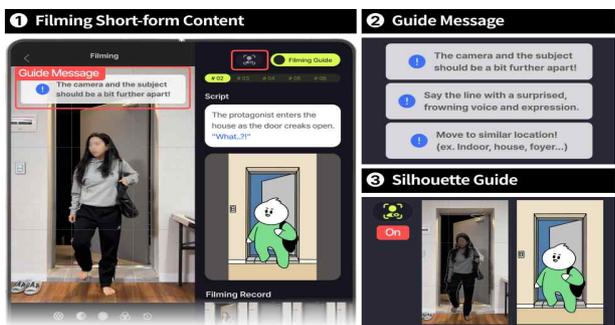


그림 12. 숏폼 촬영 화면  
 Fig. 12. Short-form filming screen

립트로 구성된다. 2는 촬영 가이드 메시지로 해당 장면을 촬영하기에 적합한 구도, 추천 장소 및 이동 안내가 제공되어 촬영에 익숙하지 않은 사용자도 원활한 촬영이 가능하도록 돕는다. 3은 ‘실루엣’ 가이드 기능으로 사용자가 ‘실루엣’ 버튼을 탭하면 구분 동작 가이드를 통해 자세 가이드 포함 프리뷰상 등장인물의 동작을 참고해 촬영할 수 있다.

4) 생성한 숏폼을 타임라인에서 좀 더 정교하게 편집하며 제작하는 단계

그림 13은 생성한 숏폼을 스크립트 창의 프롬프트와 함께 타임라인에서 좀 더 정교하게 편집하며 제작하는 단계이다. 그림 11의 2중 D에 있는 3개 버튼 가운데 ‘추가 편집하기’를 선택하면 그림 13과 같이 타임라인 편집을 시작할 수 있다. 단계별 주요 편집기능은 다음과 같다. 첫째, 그림 13의 1은 장면 추가의 예로, 사용자가 시야에서 벗어난 바퀴벌레를 찾아다녔던 경험을 Scene으로 추가하는 내용이다. 우측 Scene 하단의 프롬프트 입력창에 ‘정면으로 소파 밑을 살펴보는 장면’과 같이 추가하고 싶은 장면의 내용을 입력한 후 생성하기 버튼을 탭하면 작성한 내용이 우측 스크립트 창과 좌측 장면에 함께 추가된다. 둘째, 2는 장면 삭제의 예로, 스크립트 상단에 있는 쓰레기통 버튼을 탭하여 장면 전체를 통 삭제하거

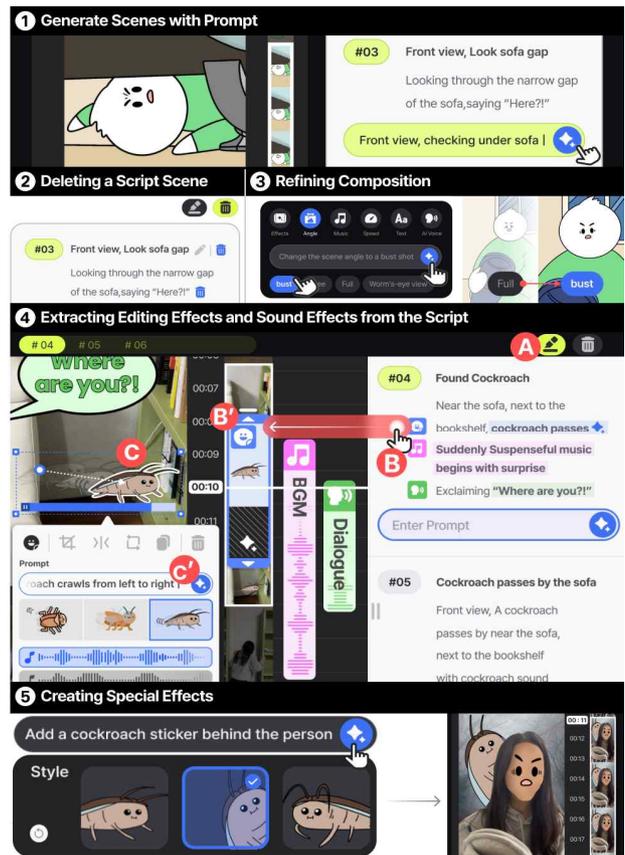


그림 13. 숏폼 콘텐츠 편집 플로우  
 Fig. 13. Short-form content editing flow

나 스크립트 텍스트 옆 개별 쓰레기통 아이콘을 탭하여 관련 스크립트 내용만 삭제할 수 있다. 셋째, 3은 장면 구도 수정의 예이다. 가령, 바퀴벌레를 잡은 줄 알았지만, 시야에 보이지 않아 당황한 인물의 얼굴 클로즈업을 통해 해당 Scene의 공포스런 분위기를 더욱 강조하고 싶다면, 편집 도구의 '구도' 기능을 통해 원하는 구도 프롬프트를 직접 입력하거나 '클로즈업', '바스트 샷', '풀 샷'과 같이 제공되는 구도 옵션 중 원하는 것을 선택해 장면 구도를 수정할 수 있다. 넷째, 4는 스크립트를 기능별로 분리 하이라이트 할 수 있는 A의 형광펜 아이콘을 선택한 장면으로 블루색은 짧은 사운드 효과가 함께 담긴 시각 기반 무비클립, 핑크색은 배경음과 같은 청각 기반 사운드, 그린색은 각종 다이얼로그인 텍스트로 표현된다. 가령, 직접 촬영이 어려운 '바퀴벌레 움직임'은 스크립트 창에서 해당 텍스트를 좌측 타임라인에 B와 같이 드래그 앤드롭 하면 움직임과 효과음이 함께 담긴 무비클립으로 적용된다. 이렇게 생성된 C의 바퀴벌레 무비클립을 탭하면 C'와 같이 상세 편집 팝업창이 뜨고, 더 으스스한 분위기로 연출하고 싶다면 바퀴벌레 모습과 기어가는 효과음을 비교하며 선택 후 수정할 수 있다. 마지막으로 5는 특수효과 제작의 예이다. 촬영한 장면을 조금 더 유머스럽게 바꾸고 싶다면 편집 도구에서 '바퀴벌레 스티커가 인물 뒤에 쭉 나타나는 효과'와 같은 프롬프트를 입력하여 웹툰 스타일의 익살스러운 표정과 함께 특수효과 필터가 적용된 장면을 제작할 수 있다.

### 3-4 프로토타입 기반 전시 현장 의견 수렴

그림 14는 2023년 12월 14일부터 18일까지 오프라인 전시를 통해 본 제안 서비스의 유용성과 사용성을 확인하고, 개선 사항의 의견을 수렴한 장면이다. 피드백 수렴 방법은 관람이 끝난 후 의견 수집에 동의하는 관람객을 대상으로 준비된 문항을 인터뷰하는 방식을 사용하였다. 피험자는 총 60명으로 연령대는 20대 52명, 30대 4명, 10대 2명, 50대와 60대 각각 1명 순으로 많았으며, 성비는 여성이 37명, 남성이 23명이었다. 개선점을 이야기한 2명을 제외한 58명의 관람객에게 서비스에 대한 긍정적인 의견을 수렴할 수 있었다. 주요 긍정 의견으로는 '누구나 쉽게 도전할 수 있어 숏폼 제작의 진입장벽이 낮아질 것 같다.', '숏폼을 편집하는 것이 번거롭고, 어렵게 느껴졌는데, 쉽게 제작할 수 있어서 나와 같은 숏폼 입문자들에게 유용한 툴이 될 수 있을 것 같다.', '평소 떠오르는 생각을 메모장에 기록하고 방치하는 편인데, 와티에서는 이를 조합하고 활용해 숏폼으로 재탄생시킬 수 있을 것 같아 재미있어 보인다.' 등이 있었다. 한편 개선점으로 '현재는 초보자에게 초점이 맞춰져 있는데, 숏폼 영상 제작에 익숙한 사용자를 위한 프로 버전도 필요해 보인다'와 '크리에이터와 공동으로 아이디어 작업을 할 수 있는 공유 아이디어보드 기능이 있으면 유용할 것 같다'의 의견이 있었고, '폼팩터인 갤럭시 z폴드의 이점을 제대로 활용할 수 있는 것 같아 비즈니스

스 모델을 구체화한 뒤 기업에 제안해 봐도 좋을 것 같다'라는 의견처럼 상용화 가능성에 대해서도 긍정적인 인사이트를 얻을 수 있었다.



그림 14. 제안 서비스의 오프라인 전시 장면  
Fig. 14. Offline exhibition photos of proposed services

## IV. 결 론

숏폼은 최근 MZ세대 사이에서 인기가 많아 관련 시장이 빠르게 성장하고 있지만, 실제 숏폼 제작 및 업로드 경험은 시청 경험을 따라가지 못하는 실정으로 이의 개선을 위해 본 연구가 시작되었다. 본 연구를 통해 도출한 연구 결과는 다음과 같다. 첫째, 관련 연구 분석을 통해 MZ세대의 숏폼 콘텐츠 소비 특성을 파악하였고, 선행 대표 3개 숏폼 플랫폼의 연계 분석을 통해 크리에이터의 콘텐츠 기획을 도와주는 기능의 부족함과 편집 화면에서의 제작 어려움에 대한 인사이트를 발견하였다. 이의 개선을 위해 숏폼 크리에이터의 숏폼 제작 단계를 생성형 AI 모델의 적용 가능 기준으로 구분하였다. 기술 리서치로 2개의 LLM 서비스, 4개의 생성형 AI 모델을 활용한 이미지 및 영상 제작 서비스, 그리고 3개의 텍스트 프롬프트 기반 T2I와 T2V 영상편집 모델을 분석하여 프롬프트를 활용한 영상 생성 서비스의 특징 및 한계와 텍스트 기반 편집 서비스의 특징을 파악하였다. 둘째, 79명 대상으로 진행한 사용자 조사를 통해 숏폼 아이디어 및 기획의 어려움, 기획과 제작 사이의 간극, 가이드 없는 촬영의 어려움, 제작 및 편집 시점의 툴 사용 어려움이 많음을 확인하였다. 특히 기획 단계에서 사용자들 간 창의적인 아이디어 발상을 살펴보거나 참고하면서 스크립트를 작성할 수 있는 기능에 대한 니즈를 발견하였고, 촬영 과정에 도움을 줄 수 있는 메시지 제공 및 숏폼 세부 편집 과정의 어렵지 않은 UI의 개선 필요성을 확인하여 이 부분들이 생성형 AI 도입 필요 지점임을 파악할 수 있었다. 셋째, 웹툰 스크롤 인터페이스의 인사이트를 바탕으로 본 연구의 메인 UI 콘셉트를 사용자가 스크립트를 편집하면서 동시에 숏폼 영상을 편집할 수 있도록 생성형 스크립트와 타임라인이 좌우 연동된 세로형 폴더블 모바일 기반 영상편집 방식으로 정의하였다. 넷째, 기술 리서치를 통해 영상 제작 및 편집 알고리즘은 LLM 모델에서 re-captioning technique을 통해 T2I와 T2V 모델에 전달하는 방식인 구글의 루미에르와 오픈 AI의 소라의 영상 생성 모델을 바탕으로 재구성하여 제시하였다. 다섯째, 제안 시나리오를 구체화하기 위해 숏폼 제작 과정을 고객 여정 지도로 작성하였다. 마지막

으로 4개 유형의 유저 플로우 기반 서비스 사용 단계 프로토타입을 제작한 후 전시 관람객을 대상으로 한 인터뷰를 통해 본 서비스의 핵심 가치의 유용성과 사용성 기반 답변을 수렴하였다. 이를 통해, 본 연구는 사용자가 적극적으로 숏폼 콘텐츠를 제작할 수 있도록 생성형 AI를 연계한 숏폼 제작 단계 기반 차세대 숏폼 제작 플랫폼을 제시하였다는 점에 의의가 있음을 확인할 수 있었다. 본 연구의 한계점으로는 79명의 사용자 조사와 60명의 전시 현장 관람객 대상 인터뷰를 진행하였으나 이후 본 서비스의 상용화와 심화를 고려한다면 좀 더 폭넓은 사용자 피드백을 수렴하기 위해 포괄적인 데이터를 양적, 질적으로 수집한 후 분석을 강화할 필요가 있다. 또한, 현 생성형 AI 기술 기반 알고리즘을 리서치한 후 각 디자인 제안 단계에 적용 시도를 하였으나 디자인 제안 연구이다 보니 실제 개발을 위해서는 기술 구현 측면에서의 검토가 필요하다. 오픈 AI의 소라 서비스 발표 후 생성형 AI기반 영상 편집 툴이 새롭게 속속 출시되는 현시점에 맞춰 기술 트렌드에 부합되게 본 연구의 부족한 부분을 계속 보완하여 다른 크리에이터와 함께 숏폼 콘텐츠를 제작할 수 있는 협업 숏폼 제작 플랫폼에 관한 후속 연구를 이어갈 계획이 있다. 숏폼 제작에 관심은 많으나 실제 제작 및 업로드 시도의 제약이 많은 현 숏폼 크리에이터들의 어려움에 맞춰 본 연구의 제안 서비스는 유튜브 숏츠, 인스타그램 릴스, 틱톡 등 현 모바일 숏폼 제작 서비스들의 숏폼 제작상의 페인포인트와, 생성형 AI기반 영상 편집 기능은 적극적으로 도입되고 있으나 데스크탑 PC 기반의 전문 이용자를 위한 툴 기반인 프리미어 프로, LTX Studio와 같은 생성형 영상 편집 툴의 페인포인트와 소라, 런웨이, 피카의 영상편집 방식을 분석하여 개선방안을 제안했다는 점, 그리고 그 결과로서 모바일 폴더블 폼팩터 기반의 생성형 스크립트와 세로형 타임라인 편집의 직관성을 연계하여 숏폼 크리에이터들이 아이디어 기획에서 편집까지 좀 더 쉽고 직관적으로 숏폼 콘텐츠 제작에 도움이 되는 서비스를 제안하였다는 점에 본 연구의 의의를 두고자 하며, 생성형 AI 기술을 연계 시도한 초기 연구로서 이후 숏폼 산업의 성장 및 발전을 위한 후속 연구에 도움이 되길 바란다.

## 감사의 글

이 논문은 2023학년도 홍익대학교 디자인컨버전스학부 졸업 전시 작품을 바탕으로 추가 연구를 진행하였습니다.

## 참고문헌

[1] M. Kim and D. H. Byun, "A Study on the Change of Editing Style in YouTube Short-From Content," *Trans-*, Vol. 13, pp. 59-90, July 2022. <https://doi.org/10.23086/trans.2022.13.04>

[2] Opensurvey. Comparing Content Topics, Lengths, and Trends between News, Shorts, and TikTok [Internet]. Available: <https://blog.opensurvey.co.kr/article/socialmedia-2023-2/>.

[3] H. J. Kim, S. O. Oh, and S. B. Jo, "Short-Form Contents Trend," *Excellence Marketing for Customer*, Vol. 54, No. 7, pp. 60-69, July 2020.

[4] J. Lee, "A Study on Types of Short-Form Video Contents," *Humanities Contents*, No. 58, pp. 121-139, September 2020. <https://doi.org/10.18658/humancon.2020.09.121>

[5] Mezzomedia. 2023 Mezzomedia Short Form Marketing Guide [Internet]. Available: [https://www.mezzomedia.co.kr/data/insight\\_m\\_file/insight\\_m\\_file\\_1546.pdf](https://www.mezzomedia.co.kr/data/insight_m_file/insight_m_file_1546.pdf).

[6] J. Kim, A Study on Analysis of Performance by Feature of Internet Personal Broadcasting, Master's Thesis, Ewha Womans University, Seoul, August 2016.

[7] S. W. Han, A Study on Community Space for Digital Content Creator -Focused Working Environments and Collaboration-, Master's Thesis, Hongik University, Seoul, February 2020. <https://www.doi.org/10.23174/hongik.00000024870.11064.0000305>

[8] S. Lee, Youtuber's Work, Profitability, and Autonomy, Master's Thesis, Seoul National University, Seoul, August 2018. <https://www.doi.org/10.23170/snu.000000152576.11032.0001094>

[9] Y. H. Lee, "Present and Future of the Generative AI Era," *Chungbuk Issue & Trend*, Vol. 52, pp. 30-35, June 2023.

[10] BIGDATAnews. ChatGPT, Overwhelmingly Tops Domestic and International AI Service Interest in 2023... Followed by Iruda-Google Bard [Internet]. Available: [https://www.thebigdata.co.kr/view.php?ud=20240116160909859696085a06ef\\_23](https://www.thebigdata.co.kr/view.php?ud=20240116160909859696085a06ef_23).

[11] ITWorld. The Way You Use MS Co-Pilot is Wrong [Internet]. Available: <https://www.itworld.co.kr/news/309583>.

[12] Byline Network. Google Unveils 'LUMIERE'... What is the Future and Present of Video Creation AI? [Internet]. Available: [https://byline.network/2024/01/240130\\_003/](https://byline.network/2024/01/240130_003/).

[13] Google Research. LUMIERE: A Space-Time Diffusion Model for Video Generation [Internet]. Available: <https://lumiere-video.github.io/>.

[14] OpenAI. Video Generation Models as World Simulators [Internet]. Available: <https://openai.com/research/video-generation-models-as-world-simulators>.

[15] R. Mihalcea and P. Tarau, "TextRank: Bringing Order into Text," in *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, Barcelona, Spain, pp. 404-411, July 2004.



**조동희(Dong-Hee Jo)**

2018년~2024년: 홍익대학교 디자인컨버전스학부 학사  
※ 관심분야 : UX Design, UI Design, Interaction Design 등



**김민서(Min-Seo Kim)**

2019년~현 재: 홍익대학교 디자인컨버전스학부  
※ 관심분야 : UX Design, UI Design, Branding



**김성민(Sung-Min Kim)**

2019년~현 재: 홍익대학교 디자인컨버전스학부  
※ 관심분야 : UX Design, UI Design



**김건동(Geon-Dong Kim)**

1998년 : 홍익대학교 대학원 (미술학석사)  
2007년 : Rhode Island School of Design, USA (MFA, 미술학 석사)  
2017년 : 서울대학교 대학원 (Doctor of Design, 디자인학 박사)

2001년~2004년: 엔씨소프트  
2007년~2009년: Tellart, Interaction Design Consultancy, USA  
2019년~2019년: Visiting Scholar, Duke University, USA  
2009년~현 재: 홍익대학교 디자인컨버전스학부 교수  
※ 관심분야 : Information Design, Interface Design, UX Design, Meaning Making 등