JDCS
디지털콘텐츠학회논문지
Journal of Digital Contents Society
Vol. 27, No. 10, pp. 2753-2765, Oct. 2026

Check for updates

구조적·함축적 맥락 반영 생성형 이모지 서비스 제안

서 유 빈[1] · 윤 현 경[1] · 김 건 동[2*]
[1]홍익대학교 디자인컨버전스학부 학부과정
[2]홍익대학교 디자인컨버전스학부 교수

A Proposal for a Generative Emoji Service Incorporating Structural and Implicit Contexts

Yu-Bin Seo[1] · Hyeon-Gyeong Yun[1] · Geon-Dong Kim[2*]

[1]Undergraduate Program, School of Design Convergence, Hongik University, Sejong 30016, Korea
[2]Professor, School of Design Convergence, Hongik University, Sejong 30016, Korea

[요 약]

본 연구는 생성형 AI의 도입으로 새롭게 서비스가 확장되는 생성형 이모지 영역에서, 사용자 맥락 반영 생성형 이모지 서비스 프로토타입 제작을 통한 초기 가이드라인 제안을 목표로 한다. 이를 위해 Dey와 Winograd의 맥락 정의 관련 선행연구를 바탕으로 맥락을 구조적·함축적 맥락으로 재정의한 후 생성형 이모지의 생성 체계에 적용하였다. 이후 채팅 플랫폼 이용자 55명을 대상으로 한 설문조사와 퍼소나 워크샵을 통해 구조적 맥락 유형 하위 사람, 장소, 사물, 그리고 함축적 맥락 유형 하위 행위, 화법, 상호작용의 총 6개 퍼소나를 도출하였다. 이어서 10명 대상 GPTs 기반 Think-Aloud 방법 병행 심층 인터뷰를 진행한 후 3단계 구조적·함축적 맥락 이모지 생성 시나리오 기반 디자인 프로토타입 제작과 유저 테스트를 진행하였다. 본 연구는 새롭게 열리는 생성형 이모지 서비스 시장 환경에 맞춰 사용자 맥락 반영 생성형 이모지 제작 가이드라인을 제안한 점에 의의가 있다.

[Abstract]

This study proposes initial guidelines for a user context-driven generative emoji service by developing a prototype for the generative emoji domain, a newly emerging sector driven by the introduction of generative AI. Building on Dey and Winograd's definitions of context, this study redefines context into two categories: structural context and implicit context, and applies this framework to the design of a generative emoji creation system. A survey of 55 chat platform users and a persona workshop were conducted, resulting in six personas: Person, Place, and Object under the structural context category, and action, speech, and interaction under the implicit context category. Subsequently, in-depth interviews using a GPT-based think-aloud method were conducted with 10 participants. Based on these findings, a design prototype employing a three-stage emoji generation process based on structural and implicit contexts was developed and evaluated through user testing. This study is significant in that it proposes design guidelines for user-context-driven generative emoji services tailored to the emerging generative emoji market.

색인어 : 생성형 이모지, 생성형 AI, 구조적 맥락, 함축적 맥락, 맥락 반영
Keyword : Generative Emoji, Generative AI, Structural Context, Implicit Context, Context-Driven

http://dx.doi.org/10.9728/dcs.2026.27.10.2753

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-CommercialLicense(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 01 July 2026; **Revised** 27 July 2026
Accepted 24 August 2026

***Corresponding Author; Geon-Dong Kim**
Tel: ※ 개인정보 표시제한
E-mail: geon705@gmail.com

Ⅰ. 서 론

1-1 연구 배경 및 목적

이모지(Emoji)는 디지털 환경에서 쓰이는 그래픽 문자로, 일본어의 그림에 해당하는 일본어 발음 기준 '에'에 준하는 'e'(絵), 그리고 문자를 의미하는 '모지'에 해당하는 'moji'(文字)가 합쳐진 '에-모지'에서 유래하여 점차 영어식 발음으로 '이모지'로 변천한 시각 기호이다[1]. 이와 같은 이모지는 텍스트 중심의 소통 환경에서 감정과 의도를 시각적으로 전달하기 위한 수단으로 발전해 왔으며, 얼굴의 표정, 손짓, 사물, 그리고 상황을 간단한 그림으로 표현하며 의미를 빠르게 전달하고 감정의 뉘앙스를 보완하는 데 효과적이다.

이모지의 역사를 보면, 1980년대 초반부터 역사가 시작되는 이모티콘(Emoticon) 대비, 이모지는 1999년 일본 통신사 NTT DoCoMo에서 처음 도입된 이후, 2010년대 유니코드(Unicode)의 통합과 스마트폰 기본 키보드 자판 내에 특수문자 형태로의 탑재를 거쳐 현재 전 세계적으로 널리 사용되고 있다[2]-[4]. 이후 Apple, Microsoft, Samsung 등 주요 기업의 운영체제(OS;Operating System)에서 피부색, 성별, 장애와 같은 사용자 정체성을 반영한 이모지를 포함함으로써 이후 더 많은 표정, 상황, 장면을 표현할 수 있도록 지속 업데이트 중이다. 더 나아가 현재 이모지는 소셜미디어, 메신저와 같은 디지털 소통 환경에서 직관적으로 의미를 전달할 수 있는 보편적 시각 기호로 널리 사용되고 있다.

최근 생성형 AI(Generative AI)의 발전으로 기존에 없던 장면이나 감정을 생성형 이모지(Generative Emoji)로 표현할 가능성이 열리고 있다. 지금까지의 이모지는 운영체제가 제공하는 목록 내에서 선택하는 방식이었다면 생성형 이모지가 등장하면서 사용자는 본인이 원하는 이모지를 프롬프트로 생성할 수 있게 되었다. 특히 2024년 12월 출시된 Apple의 AI 기반 이모지 생성 기능인 Genmoji는 키보드에서 "디제잉을 하는 곰젤리"를 입력하면 DJ 데크로 음악을 믹싱하는 빨간 곰젤리, 푸른 배경의 액자 속 선명한 색감의 곰젤리, 한 손에 하프를 들고 혀를 내밀며 웃는 곰젤리와 같은 이모지를 3-4개 생성하여 제안한다[5].

이렇듯 생성형 AI의 발전으로 표현의 가능성은 열렸으나, 현재의 생성형 이모지는 외형 묘사에 초점을 두고 있다. 하지만 실제 소통 환경에서는 상대와의 관계, 대화의 분위기, 감정의 강도 같은 맥락적 요소의 반영이 상당히 중요하다. 따라서 본 연구는 이와 같은 여러 맥락 관련 선행 연구 중 Anind K. Dey의 맥락 정의를 바탕으로 사람(person), 장소(place), 사물(object)을 중심으로 사용자의 위치, 시간, 활동, 관계, 그리고 주변 환경과 같이 관찰 및 측정 가능한 정보를 구조적 맥락(Structural Context)으로 정의하였다. 또한, Terry Winograd의 사용자의 행동과 해석을 가능하게 하는 조건으로서의 맥락 개념과, 숨은 의도가 상호작용의 의미를 결정한다는 관점을 바탕으로, 문장에 직접 드러나지 않은 의도, 기대 행동, 관계적 의미와 같은 맥락들을 함축적 맥락(Implicit Context)으로 정의하였다. 본 연구에서는 이러한 맥락 정의를 생성형 이모지의 맥락 생성 체계에 반영하여 연구하고자 한다. 이를 통해 단순히 새로운 이미지를 생성하여 이모지에 적용하는 관점을 넘어, 감정과 관계, 대상자와의 맥락적 소통 상황과 같은 실제 채팅 소통에서의 맥락을 반영하는 생성형 이모지의 방향성을 제시하고자 한다. 이를 위해 본 연구는 메신저와 SNS 등 각종 대화 기반 플랫폼의 대화 환경에서 사용자의 대화 내용과 대화 맥락에 맞춰 Dey의 구조적 맥락과 Winograd의 함축적 맥락이 반영된 생성형 이모지 예시 프로토타입 제작을 통해 초기 맥락 반영 생성형 이모지 생성 가이드라인을 제안하는 데 목적을 둔다.

1-2 연구의 범위 및 방법

본 연구는 향후 각종 채팅 플랫폼에 응용될 수 있는 맥락 반영 생성형 이모지의 생성 원리와 소통 경험을 확장해 줄 맥락 응용 방식을 제안하기 위한 디자인 프로토타입 연구로 본 연구의 범위는 기획에서 예시 디자인 프로토타입 제작에 이르는 전체 과정을 포함한다. 실제 프로그래밍 기반의 작동 구현은 해당 콘셉트의 기술 구현 가능성에 대한 검토까지 진행하며, 본 연구 범위에 포함하지 않는다. 연구 방법은 다음과 같다. 첫째, 선행연구를 바탕으로 이모지의 개념, 발전 흐름과 시장 현황, 생성형 AI의 맥락 분석 기술, 실시간 생성 기술을 분석하고, 주요 플랫폼에서의 활용 사례를 검토한다. 둘째, 맥락이 반영된 생성형 이모지 생성을 위해 Dey와 Winograd의 맥락관련 선행연구를 분석한다. 셋째, 생성형 이모지에 대한 양적 사용자 조사 후 퍼소나 워크샵과 질적 태스크 병행 심층 인터뷰를 진행한다. 넷째, ChatGPT, Krea AI, Higgsfield 등 생성형 AI 툴을 접목하여 리얼타임 생성 및 사용자 간 상호작용이 가능한 생성형 이모지를 디자인하고, 시나리오 유형이 반영된 디자인 프로토타입을 제작한다. 마지막으로 제작한 프로토타입의 유저 테스트를 진행하고, 사용자 피드백을 수렴하여 연구의 핵심 결과와 한계, 향후 발전 방안을 제시한다.

Ⅱ. 관련 연구 및 사례 분석

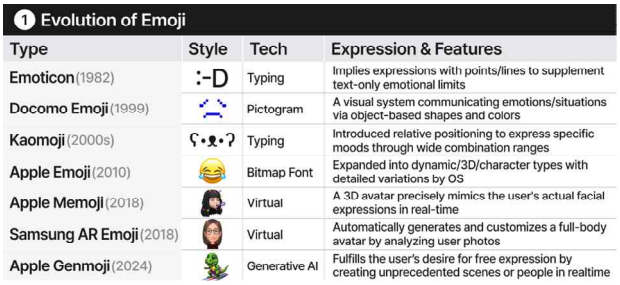
2-1 이모지의 개념, 역사 및 사례

1) 이모지의 개념

이모지는 그림과 문자의 결합으로 이루어진 디지털 기호로 일본어 '絵文字(그림문자)'에서 유래했으며 감정과 상황을 그림으로 표현해 글과 함께 사용한다[1]. 이모지는 장식 목적뿐 아니라 감정 표현, 의도 강조, 분위기 조성과 같이 디지털 소통에 있어 중요한 기능을 하며, 유니코드 컨소시엄(Unicode Consortium)에 의해 전 세계적으로 통일된 코드 체계를 갖고 있다[6]. 여기에서 유니코드란 정보 교환의 표준 부호로,

문자마다 고유 번호를 매겨 컴퓨터가 똑같이 이해할 수 있도록 만든 공통 언어 규격[7]이며 스마트폰, SNS, 메신저처럼 여러 플랫폼에서 동일 시각 표현으로 전달된다[8]. 이모티콘(Emoticon)이 텍스트의 조합으로 얼굴을 표현한 초기 형태라면, 이모지는 이를 그림으로 구현한 것이다.

2) 이모지의 역사



① Evolution of Emoji

Type	Style	Tech	Expression & Features
Emoticon (1982)	:-D	Typing	Implies expressions with points/lines to supplement text-only emotional limits
Docomo Emoji (1999)		Pictogram	A visual system communicating emotions/situations via object-based shapes and colors
Kaomoji (2000s)	ʕ•ᴥ•ʔ	Typing	Introduced relative positioning to express specific moods through wide combination ranges
Apple Emoji (2010)		Bitmap Font	Expanded into dynamic/3D/character types with detailed variations by OS
Apple Memoji (2018)		Virtual	A 3D avatar precisely mimics the user's actual facial expressions in real-time
Samsung AR Emoji (2018)		Virtual	Automatically generates and customizes a full-body avatar by analyzing user photos
Apple Genmoji (2024)		Generative AI	Fulfills the user's desire for free expression by creating unprecedented scenes or people in realtime

그림 1. 이모지의 진화 타임라인
Fig. 1. Timeline of emoji evolution

그림 1은 이모지의 진화 과정을 정리한 내용으로 디지털 소통 환경에서 시각 요소를 활용해 감정과 의도를 전달하려는 표현 방식의 변화가 담겨 있다. 이모지의 초기 형태로 볼 수 있는 이모티콘은 1982년 미국의 컴퓨터 과학자 Scott Fahlman이 ':-)'와 같은 기호를 사용하면서 시작되었다[2]. 이는 얼굴 전체를 묘사하기보단 눈, 코, 입을 닮은 점과 선으로 표정을 암시하며, 온라인상에서 글만으로는 감정을 전달하기 어려운 한계를 보완하기 위한 시도였다. 1999년에 일본의 이동통신사 NTT 도코모는 12*12 픽셀로 구성된 176개의 그림 기호를 문자 메시지 시스템에 도입하였다[3]. 이모지를 처음 디자인한 Shigetaka Kurita는 전화기, 웃는 얼굴, 날씨 기호와 같이 구체적인 사물과 감정을 표현하는 도형과 색채를 사용하여 감정과 상황을 시각적으로 전달할 수 있는 기호 체계를 만들었다[9]. 이후, 2010년에 유니코드 컨소시엄은 이모지를 문자 체계에 공식 통합하였고, 2011년 애플이 iOS에 이모지 키보드를 기본 탑재하며 이모지가 세계적으로 통용되는 시각 기호가 되었다[10],[11].

2010년대 이후 Apple, Microsoft, Samsung과 같은 주요 기업들이 유니코드 기반 이모지를 자사 플랫폼과 디자인 시스템에 맞게 시각화하였다. 동일 코드라도 운영체제에 따라 색상, 형태 및 세부 표현이 달라졌으며, 이 시기부터 이모지는 모션이 적용된 동적인 이모지, 3D 렌더링된 입체적인 이모지, 사용자의 얼굴 구조와 신체적 특징을 반영한 캐릭터 이모지와 같이 그 유형이 확장되었다. 2010년대 후반부터는 TrueDepth, Face Tracking과 같은 기술의 발전으로 사용자의 얼굴을 닮은 3D 아바타가 실시간으로 움직이며 사용자의 표정을 따라 할 수 있게 됨에 따라 Apple의 Memoji, 삼성의 AR Emoji로 확장되었다[12],[13]. 최근에는 생성형 AI의 기술이 발전하면서 이모지의 생성과 사용 방식에도 새롭고 신선한 창의적 시도를 할 기회가 열리고 있다. 이런 측면에서 2024년 Apple은 생성형 이모지 기능인 Genmoji를 iOS 18.2 업데이트에 도입하였다. Genmoji는 사용자가 원하는 이모지나 사진을 프롬프트와 조합하여 새로운 상황, 공간, 감정이 반영된 이모지를 생성하는 Apple의 기능이다[14]. 지금까지의 이모지는 유니코드로 정형화된 기호 중에서 원하는 이모지를 찾아내고, 선택하여 사용하는 방식이었다면, 생성형 AI가 도입되면서 기존에 없던 장면을 실시간으로 만들 수 있게 되었고, 사용자나 주변 인물도 Genmoji로 생성할 수 있게 되었다.

3) 생성형 이모지 사례

그림 2는 Apple의 Genmoji 사례로, 사용자의 입력 정보가 생성형 이모지에 어떻게 반영되는지를 나타낸 것이다. Apple의 Genmoji에서 생성형 이모지를 만드는 세 개 방법 중 첫 번째는 텍스트 프롬프트만으로 생성하는 방법이다. 이 경우 사용자가 입력한 텍스트 프롬프트에 대해 Apple의 OS Emoji 스타일 그대로 생성한다[15]. 그림 2의 1-1의 A처럼, "DJ octopus" 프롬프트를 입력하면 디제잉하는 문어 Genmoji를 생성할 수 있다. 두 번째는 1-2의 사례와 같이 이미지와 텍스트 프롬프트를 함께 활용하여 사용자를 닮은 Genmoji를 생성하는 방법이다[15]. 1-2의 B는 사용자의 얼굴 사진을 입력하는 과정으로, 얼굴형, 눈, 코, 머리색과 같은 외형 특징이 Genmoji 생성의 기초 정보로 활용된다. C는 텍스트 프롬프트를 입력하는 단계로 이 사례에서는 "Race car driver" 입력을 통해 얼굴 사진으로부터 추출한 외형 정보에 레이싱카 운전사 복장이 적용되게 생성한다. 세 번째는 기존의 서로 다른 이모지를 결합하여 새로운 Genmoji를 생성하는 방법으로 1-3과 같이 전구 이모지와 나무늘보 이모지를 결합하여, "한 손에 전구를 들고 있는 나무늘보" 모습을 새롭게 생성할 수 있다[16]. 이처럼, 생성형 AI 도입으로 기존 이모지에 사용자의 목적, 맥락과 취향에 맞게 실시간으로 생성할 수 있는 생성형 이모지 체계로의 전환이 이루어지고 있다.

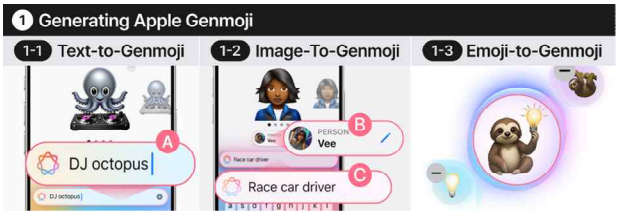


그림 2. Apple의 Genmoji 생성 방식
Fig. 2. Apple Genmoji generation methods

2-2 생성형 AI 기술 사례 분석

1) AI의 맥락 이해 기술

최근 AI의 문맥 이해 기술은 언어의 표면 의미를 처리하는 수준을 넘어, 문맥 안에 담긴 의미를 추론하고 복잡한 뉘앙스를 해석하는 방향으로 발전하고 있다. 문맥 기반 기계 번역 관련 선행연구를 보면 신경망 모델이 언어의 의미를 이해하기

위해 복잡한 뉘앙스를 학습해야 하며, 잘 사용하지 않거나 처음 보는 표현 또한 주변 문맥을 통해 의미를 추론할 수 있다고 한다[17]. 대형 언어 모형에 대한 공감 추론 관련 선행연구에서는 사고 과정 추론(CoT; chain of thought)과 공감 추론(CoE; chain of empathy)같은 추론 기반 프롬프팅을 통해 AI 모델이 사용자의 감정 상태와 그 원인이 되는 상황 요인을 단계에 따라 해석할 수 있음을 보여주고 있다[18]. 특히 CoE는 감정과 상황 단서를 함께 추론한 뒤 응답을 생성하기에 맥락을 인지하고 공감하여, 정서적으로 적절한 반응을 가능하게 한다[19]. 또한, 비전-언어모델(VLM; Vision-Language Models)과 같은 멀티모달 인공지능 기술이 발전함에 따라 텍스트 외에 이미지 내부의 중요한 객체와 단서를 포착하여, 성별, 나이, 직업, 교육 수준과 같은 다양한 개인 속성을 추론할 수 있게 되었다[20]. 이처럼 생성형 이모지 서비스가 텍스트와 주변 상황 단서, 그리고 사용자의 감정 상태 추론을 통해 이모지를 생성하고 제안할 수 있도록 기술이 발전하고 있다.

2) 리얼타임 생성 사례

생성형 이모지를 실제 소통 상황에 적용하기 위해서는 사용자의 감정, 의도, 상황과 같은 맥락에 실시간으로 반응하는 시각 표현 방식이 이모지 사용자 경험에 중요한 역할을 한다[21]. 이모지는 감정과 분위기 표현, 장식 요소, 문장 보완과 강화 같은 여러 목적으로 활용되며, 사용자의 생각과 감정을 전달하는 중요한 비언어 단서로 기능하기 때문이다[22]. 특히 채팅 환경에서는 메시지가 짧고 빠르게 오가며, 감정의 변화나 맥락 전환이 매우 빠르게 이루어지고, 현재의 생성과 출력 방식은 일정 대기 시간을 필요로 하여, 사용자의 심리 상태나 상황 변화가 빠르게 일어나는 실제 채팅 맥락을 실시간으로 반영하기 어려우므로, 실시간 피드백이 가능한 리얼타임(Realtime) 생성 기술의 도입이 필요하다.

리얼타임 생성은 사용자의 텍스트 입력, 참조 이미지, 또는 외부 시각 자료를 바탕으로 AI가 생성한 이미지를 즉시 업데이트하는 기술로, 대화의 흐름을 방해하지 않고 자연스러운 인터랙션이 가능하다[23]. 그림 3의 1은 Krea AI 사례로 사용자의 입력에 실시간으로 반응하여 이미지를 생성하는 툴로, 1-1처럼 텍스트 프롬프트를 입력하거나 1-2처럼 드로잉을 하여 입력하면 입력이 반영되자마자 실시간으로 시각화 결과가 출력된다. 잠재 일관성 모델(LCMs; Latent Consistency Models)은 차세대 이미지 생성 모델로, 기존 이미지 생성 방식과 달리 최소한의 추론만으로 고해상도 이미지를 빠르게 생성할 수 있도록 설계된 모델이다[24]. 따라서, LCMs 기술을 활용하면 텍스트와 시각 입력을 프레임 단위로 처리할 수 있으며, 사용자는 별도의 생성 버튼을 누르거나, 기다리지 않아도 생성 과정과 모습을 실시간에 가깝게 바로 확인할 수 있다[24]. 또한 리얼타임의 일종인 Motion Stream은 생성형 AI의 실시간 처리 기술을 활용하여 위 리얼타임 대비 움직임, 시선, 그리고 제스처와 같은 입력값을 실시간으로 더욱 자세하게 반영하는 인터랙티브 콘텐츠 기술이다[25]. 2-1과 같은 Long-Form 영상도 실시간 생성이 가능해져서 모션이나 장면 전환이 지속해서 이루어지는 콘텐츠에서도 일관된 품질과 흐름을 유지할 수 있다. 또, 2-2처럼 드래그 기반 모션 조절 기능을 통해, 카메라의 시점이나 영상 속 물건의 위치를 조정할 수도 있다. 이러한 리얼타임 기술은 AI로 무언가를 생성할 때, 생성 과정을 알 수 없이 기다려야 하는 기존의 방식을 벗어나 사용자가 실시간으로 생성에 참여할 수 있다.

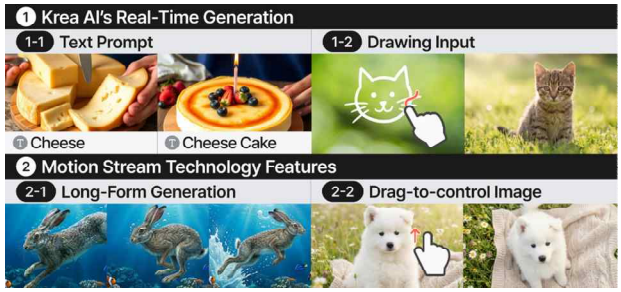


그림 3. 리얼타임 기술 사례
Fig. 3. Examples of real-time generation technologies

2-3 맥락의 개념과 재정의

맥락(Context)은 어떤 정보가 특정 상황에서 어떻게 해석되고 사용되는지를 결정하는 조건으로, 배경 정보의 역할 뿐만 아니라 상호작용 시 의미를 형성하는 핵심 요소이다. Dey는 맥락을 어떤 개체의 상황을 특정 짓는 데 사용할 수 있는 모든 정보라고 정의하였다[26]. 여기에서 개체란 사람(person), 장소(place), 사물(object)로 구성되며, 각각의 개체는 정체성(Identity), 위치(Location), 활동(Activity), 시간(Time)과 같은 정보로 설명될 수 있다. 가령, 첫 번째 사람은 사용자가 누구인지, 어디에 있는지, 무엇을 하고 있는지, 언제 해당 행동을 하는지로 특정될 수 있으며, 두 번째, 장소는 어떤 공간인지, 어디에 위치하는지, 그곳에서 어떤 활동이 일어나는지, 어느 시간대의 장소인지로 설명될 수 있다. 마지막으로 사물은 어떤 사물인지, 어디에 놓여 있는지, 어떻게 사용되고 있는지, 언제 사용되는지에 따라 그 상황이 특정된다. 즉, Dey 관점의 맥락은 사람, 장소, 사물의 상태를 정체성, 위치, 활동, 시간 정보를 통해 설명하는 상황의 객관적 뼈대이자 측정 가능한 정보라는 측면에서 본 연구는 구조적 맥락(Structural Context)이라 정의하였다. 반면, Winograd는 맥락을 단순한 배경 정보가 아닌 사용자의 행동과 해석을 가능하게 하는 조건의 개념으로 정의했다[27]. 따라서 단순히 사전에 정의된 주어진 정보의 집합으로 보기보다 인간의 활동 과정에서 형성되는 의미 구조로 보았으며, 그 측면에서 사용자의 숨은 의도를 파악하는 것이 중요하며 대화가 일어나는 상대방과의 관계와 그 속에서 상호작용이 일어나는 행위적 측면의 해석이 중요하다[27]고 하였다. 따라서 Winograd의 맥락은 고정된 정보가 아니라 특정 행동과 발화가 어떤 의미로 이해되는 조건이 될 수 있도록 목적(Purpose), 행동

(Action), 상호작용(Interaction), 해석(Interpretation)의 의미 형성 체계가 필요하다. 본 연구는 이와 같은 Winograd의 맥락 정의를 목적에 따른 의도와 행위를 함축하고 있다는 측면에서 함축적 맥락(Implicit Context)이라 정의하였다. 이처럼 맥락은 생성형 이모지 생성에 중요한 요소이며 사용자 니즈와 상황에 맞게 구조적, 함축적 맥락을 절충해서 적용하는 것이 중요하다. 본 연구는 선행 연구에서 정의한 맥락의 개념을 생성형 이모지 생성 서비스에 반영하고자 한다.

Ⅲ. AI 기반 생성형 이모지 서비스 기획

3-1 양적 설문조사를 통한 사용자 맥락 유형화

그림 4는 채팅 환경에서 맥락 요소를 도출하기 위한 설문조사와 사용자 행동 원형 도출(Behavioral archetype mapping)을 위한 퍼소나 워크샵 결과이다. 설문은 1에 정리된 내용처럼 2025년 5월 3일부터 5월 10일까지 Google Form을 통해 진행했으며, 모바일 채팅 및 이모지 이용 경험이 있는 사용자 55명이 참여하였다. 응답자는 남성 27명(49.1%), 여성 28명(50.9%)으로 구성되었으며, 연령대는 20대 36명(65.5%), 10대 11명(20%), 30대 5명(9.1%), 40대 3명(5.4%)이다. 문항은 채팅 시 사용자가 어떤 맥락 정보를 활용하여 메시지를 이해하는지 파악하기 위해, Dey와 Winograd의 관점에 기반해 설문 문항을 구성하였다. 먼저, Dey의 관점에 따라 대화 상황에서 사람, 장소, 사물의 개체가 각각 어떤 정체성을 가지며, 어디에 위치하고, 어떤 활동과 관련되며, 어느 시간대에 이루어졌는지를 파악할 수 있는 질문을 포함하였다. 이를 통해 대화가 이루어지는 상황에 따른 관찰 가능한 구조적 맥락을 도출하고자 하였다. 또한, Winograd의 맥락 관점을 바탕으로 특정 행동과 발화가 어떤 의미로 해석되는지 파악하기 위해 대화의 목적, 사용자의 행동, 직전 대화의 흐름, 화법, 그리고 상대의 기대 반응 및 실제 반응 차이 분석을 위한 질문을 포함하였다. 2는 퍼소나 워크샵 결과로 설문 응답에서 반복적으로 나타난 해석 단서와 표현 행동을 Dey와 Winograd 관점 기준에서 상황에 영향을 받는 정도, 정보를 포함하는 정도, 말투를 조절하는 정도, 간접 표현 성향의 정도, 상대와의 친밀도의 다섯 개 행동 변수를 설정하였다. 대표 답변을 보면 상대방의 메시지를 이해할 때 시간, 장소, 관계, 말투, 이모지, 숨은 의도와 같은 맥락을 함께 고려하고 있었으며, 동시에 감정을 직접 말하지 않고 상대와의 관계에 따라 말투를 조절하거나, 상황에 따라 답장 방식이 달라지는 반복된 행동 표현을 파악하였다. 이후 각 사례를 변수 축 위에 배치하고 유사한 패턴끼리 그룹핑한 뒤, Dey와 Winograd 관점에 따른 2개 상위 퍼소나를 도출하였다. 2-1은 Dey의 맥락 관점에서 사람, 장소, 사물과 같은 관찰 및 측정 가능한 요소의 구조 기반 구조적 타입이며, 2-2는 Winograd 맥락 관점에서 행동, 화법, 상호작용에 담긴 메시지의 숨은 의미 해석 중심 함축적 타입이다. 마지막으로 3은 이 2개 상위 퍼소나 답변 중 반복하여 나타난 메시지 해석 방식을 바탕으로 각각 3개 하위 퍼소나를 유형화한 내용으로 총 6개 퍼소나 유형을 도출하였다. Dey 관점 구조적 맥락 타입 기준 답변을 그룹핑하여 사람, 장소, 사물의 하위 세 개 퍼소나를 도출하였으며, 각각 객체의 정체성, 위치, 활동, 시간 기반 답변을 정리하여 퍼소나의 특징을 부여하였다. 마찬가지로 Winograd 관점 함축적 맥락 타입 기준 답변을 그룹핑하여 행동과 움직임을 통해 의미를 추론하는 행동 중심, 말투와 뉘앙스를 중심으로 해석하는 화법 중심, 친밀도와 상호작용 속에 담긴 숨은 의도를 파악하는 상호작용 중심의 하위 세 개 퍼소나를 도출하여 각 퍼소나의 특징을 부여하였다. 그 외 설문 답변을 종합한 결과 사용자의 실제 의도가 텍스트만으로 충분히 전달되기 어려우며, 또한, 메시지를 주고받을 때 텍스트 외에 말투, 상황, 관계와 같은 맥락의 중요함에 대한 공통 인식이 있었다. 그 예시로 '화가 났을 때 친한 사이라면 우회적으로 돌려서 표현한다', '화났지만 시험 기간이라 이해하고 넘어갔다', '속상한 마음을 드러내고 싶어서 일부러 말투를 더 딱딱하게 바꾸어 말했다.'와 같이 담긴 숨은 의도를 파악하게 하는 함축적 맥락 측면의 답변이 다수 있었다.

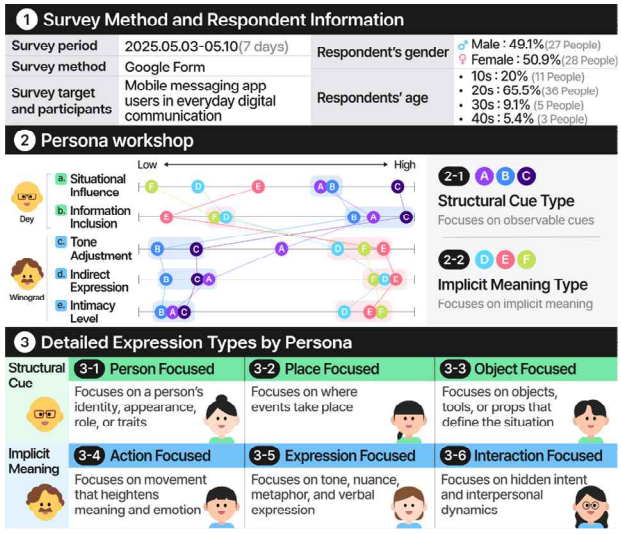


그림 4. 설문조사 기반 퍼소나 워크샵
Fig. 4. Survey-based persona workshop

3-2 Think-Aloud 기반 이모지 생성 관련 질적 실험과 결과

1) Think-Aloud 기반 이모지 생성 실험 설계

2025년 4월 10일부터 4월 12일까지 피험자 10명을 대상으로 피험자당 평균 40분 이상 충분히 의견을 들을 수 있도록 Think-Aloud 기반 질적 실험연구를 진행하였다. 실험 목적은 사용자가 메시지를 전달하고자 하는 이모지에 대한 본인이 그린 손 스케치와 이 메시지를 프롬프트화 하여 생성한 이미지와의 비교를 통해 구조적, 함축적 맥락이 어떻게 생성형

이모지에 반영되는지를 파악하기 위함이다. 실험을 위해 Think-Aloud 기법을 적용한 GPTs 기반 이모지 생성 실험을 설계하였다. 여기에서 Think-Aloud 기법은 피험자가 사전에 제시된 태스크를 수행하면서, 그 과정에서 어떤 생각과 판단을 거쳤는지를 동시에 말로 설명하도록 하는 방법론이다 [28]. 본 실험에서는 피험자에게 자신이 의도한 생성형 이모지의 모습을 직접 스케치하도록 하는 과업을 제시하였으며, Think-Aloud 절차에 맞춰 진행하였다. 그림 5의 1은 피험자 10명의 인구통계학적 정보 정리 내용으로, 채팅 환경에서 하루에 이모지를 최소 1회 이상 사용하는 대상자로 선정하였다. 2는 실험 과정을 정리한 내용이다. 실험에 앞서, 피험자에게 테스트할 이모지 생성 툴을 GPTs를 통해 구현하였다. 2-1의 A, B, C는 피험자에게 제시한 서로 다른 각각의 상황 시나리오이다. A는 월급날 통장 잔고를 확인한 후 친구에게 자랑하는 상황이며, B는 온라인 미팅 중 마이크가 켜진 줄 모르고 말실수를 한 상황, C는 지각으로 인해 상사에게 메시지를 보내야 하는 상황이다. 이때 피험자는 2-2처럼 각 상황에 보내고 싶은 메시지를 각각 GPTs에 입력한다. 가령, A의 상황에서 '나 오늘 월급 받음! 내가 커피 쏠게.'와 같은 메시지를 GPTs에 입력하면 해당 메시지가 프롬프트가 되어 이모지가 생성된다. 이모지가 생성되는 동안 피험자는 2-3처럼 자신이 기대하는 생성될 이모지의 모습을 스케치한 후, 실제 생성된 이모지와 자신의 스케치를 비교하며 인터뷰를 진행하였다.

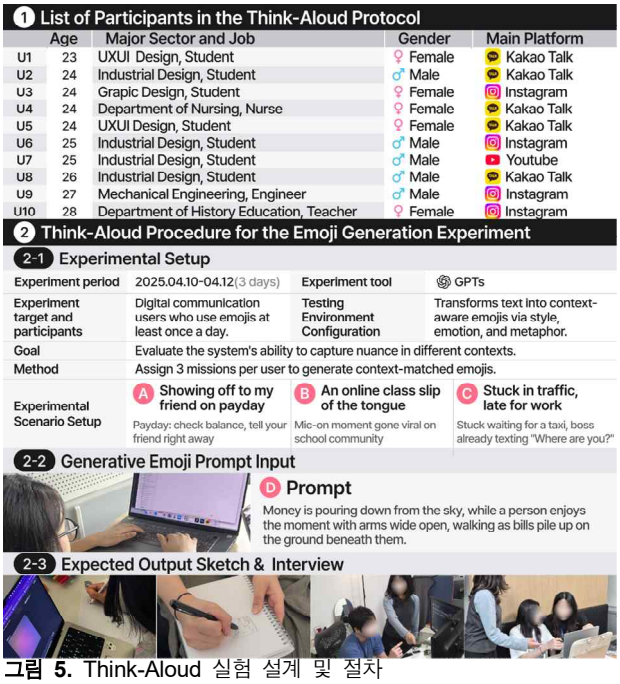


그림 5. Think-Aloud 실험 설계 및 절차
Fig. 5. Think-Aloud experiment design and procedure

2) Think-Aloud 기반 이모지 생성 실험 결과

그림 6은 Think-Aloud 실험의 분석 결과로 10명의 피험자가 주어진 세 개 상황에서 도출한 총 30쌍의 생성형 이모지

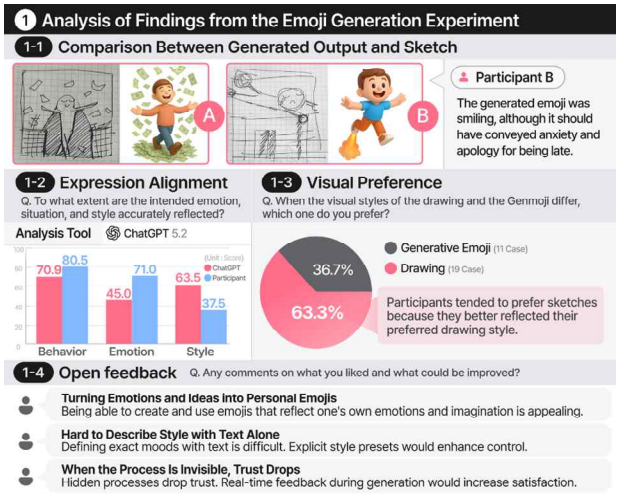


그림 6. Think-Aloud 실험 결과 및 사용자 피드백
Fig. 6. Think-Aloud experiment results and user feedback

와 스케치 간 비교를 정리한 결과이다. 1-1은 생성된 결과물과 피험자의 스케치를 비교한 예시 사례로, A는 월급날 친구에게 자랑하는 상황이다. 생성된 이모지는 휘날리는 돈다발, 위풍당당한 걸음걸이와 같이 구체적인 사물 단서나 행동 표현을 통한 맥락 표현 전달로 피험자의 스케치와의 일치도가 비교적 높았다. 다만 B처럼 늦었을 때 초조하고 미안한 마음과 빠르게 가겠다는 의도를 함께 전달하는 상황의 경우 '발에서 로켓처럼 불이 나오며 날아가려고 하는 모습'으로 이모지는 생성되었지만, 늦은 상황에 적절하지 않게 웃는 표정으로 생성되었다. 실제 피험자 또한 '늦어서 안절부절하고 미안한 마음이 담겼으면 좋겠는데 웃고 있어서 아쉽다.'고 답했다. 1-2는 ChatGPT 5.2를 사용해 생성형 이모지와 피험자 스케치의 표현 일치도를 분석한 결과이다. 채점 척도는 행동, 감정, 스타일의 각 항목을 100점 만점 기준으로 환산하는 방식으로 설정하여 점수가 높을수록 생성형 이모지와 피험자 스케치 간 표현 일치도가 높은 것으로 정의하였다. 또한, 채점에 사용한 프롬프트의 경우, 평가를 위해 ChatGPT 5.2에 지침 프롬프트를 입력하여 진행하였고, 행동 유사도는 동작과 상징을 통해 동일한 상황과 행위가 표현되었는지, 감정 유사도는 감정의 종류와 강도, 복합 감정이 전달되었는지, 스타일 유사도는 생성형 이모지와 손 스케치의 매체 차이를 고려하여, 그림체, 과장 정도와 같은 시각적 표현 방식이 얼마나 유사한지를 기준으로 분석하였다. 또한, AI의 평가 결과가 피험자의 판단과 얼마나 일치하는지 확인하기 위해 동일한 행동, 감정, 스타일 기준을 적용한 피험자 인간 평가와 함께 비교하였다. 평가 결과, ChatGPT의 평가는 행동 일치도가 70.9로 가장 높았고, 감정 일치도는 45.0점, 스타일 일치도는 50.6점으로 나타났다. 반면 피험자 평가는 행동 일치도 80.5점, 감정 일치도 71.0점, 스타일 일치도 37.5점으로 나타났다. 반복 측정 시 편차를 확인하기 위해 동일 프롬프트로 3회 반복 평가를 수행하였으며, 반복 평가 간 표준편차는 행동 0.5점, 감정 1.8

점, 스타일 3.4점으로 나타났다. 주로 언급한 행동 일치도, 감정 일치도 외에 스타일 일치도 측면에서 선의 굵기, 질감, 색감과 같은 시각 표현 요소는 주어진 세 가지 상황 전반에서 비교적 일관되게 유지되었으나, 피험자들이 기대한 그림체와는 다르게 생성되는 경우가 많았다. 1-3은 생성형 이모지와 스케치의 비주얼이 다를 경우 어떤 표현 방식을 더 선호하는지에 대한 질문의 결과로, 전체 비교 사례 중 63.3%가 직접 그린 스케치를 선호하는 것으로 나타났다. 그 이유로 직접 그린 스케치가 의도한 감정을 더 잘 담아내며, 개인의 취향에 부합하는 그림체나 스타일을 바로 반영할 수 있어 보인다는 답변이 많았다. 1-4는 실험에 대한 피험자들의 종합적 평가와 개선 의견을 정리한 내용이다. 피험자들은 자신의 감정과 상상이 반영된 이모지를 직접 구현하고 사용할 수 있다는 점을 주요 장점으로 뽑았으나 텍스트만으로는 원하는 시각적 스타일을 구체적으로 반영하기 어려워 별도의 스타일 선택 기능이 필요하다고 언급하였다. 또한, 한 번 텍스트를 입력하면 결과물이 생성되기까지의 과정을 확인할 수 없는 점에 대해 불편함을 느꼈다고 답하였다.

실험 결과를 종합하면, 첫째, 생성형 이모지는 A의 돈다발처럼 직접 드러나는 구조적 맥락 요소를 반영하는 데에는 일관성을 보였으나, 주어진 상황에서 사용자의 말투, 감정 강도, 뉘앙스를 포함한 함축적 맥락의 표현에는 한계를 드러냈다. 특히, 충분한 맥락 정보가 제공되지 않을 경우, 생성 결과는 입력값을 깊이 있게 해석하기보다는 표면적으로 드러난 단서를 중심으로 구성되는 경향을 보였다. 둘째, 피험자들은 입력한 텍스트가 해당 이모지로 생성된 이유와 논리를 이해하고자 했으며, 생성된 결과물의 세부 요소와 표현 방식을 직접 조정할 수 있기를 원했다. 이는 생성형 이모지가 획일적인 생성 결과가 아니라, 글쓰기처럼 사용자의 표현 습관과 선호 스타일에 맞춰 구조적, 함축적 맥락을 절충하는 개인화 서비스가 될 필요가 있음을 보여준다. 또한, 생성형 이모지의 생성 근거 제시, 감정 강도, 대상, 시각 요소를 사용자가 직접 수정할 수 있는 조정 인터페이스를 함께 제공할 필요가 있다.

3-3 생성형 이모지 생성 시나리오 및 기술 프로세스

1) 생성형 이모지 생성 시나리오

그림 7은 본 연구에서 제안하는 채팅 상황에서의 이모지 생성 시나리오와 사용한 생성형 AI 툴 활용 프로세스 정리 내용이다. 시나리오는 크게 기본 생성형 이모지, 맥락 반영 심화 이모지, 추가 세부 이모지 생성 조정 및 전송의 세 단계로 구성된다. 먼저 1은 기본 생성형 이모지 생성 단계를 정리한 내용으로 1-1의 A는 대화 상대인 친구가 사용자에게 네 장의 사진을 보내오며 '프로필 사진 바꾸려는데 어떤 사진이 좋을까?'라고 질문한 상황이다. 사용자가 A'처럼 사진을 탭하면 썸네일 사진이 확대되고, 네 장 사진을 좌우로 스와이프하며 비교한 후 4번 사진이 제일 잘 나온 것 같다고 판단한다. 이때 사진 '4번 사진이 좋은데~'와 같이 텍스트 메시지로 단순하게 답변할 수 있지만, 친구에게 더욱 흥미롭게 생성형 이모지를 만들어 보내고자 B처럼 이모지 생성 아이콘을 탭하면 팝업이 활성화되고 선택 사진이 생성을 위한 이미지로 등록된다. C는 생성형 이모지의 그림체를 선택하기 위한 과정으로 D처럼 원하는 스타일을 선택한 뒤 D'와 같이 완료 버튼을 누르면 E처럼 선택한 그림체 스타일로 이모지가 생성된다. 1-2는 프롬프트 입력을 통해 이모지가 실시간으로 생성되는 과정을 보여준다. F처럼 '4번째 사진이 완전 아이돌 같아!!!'라

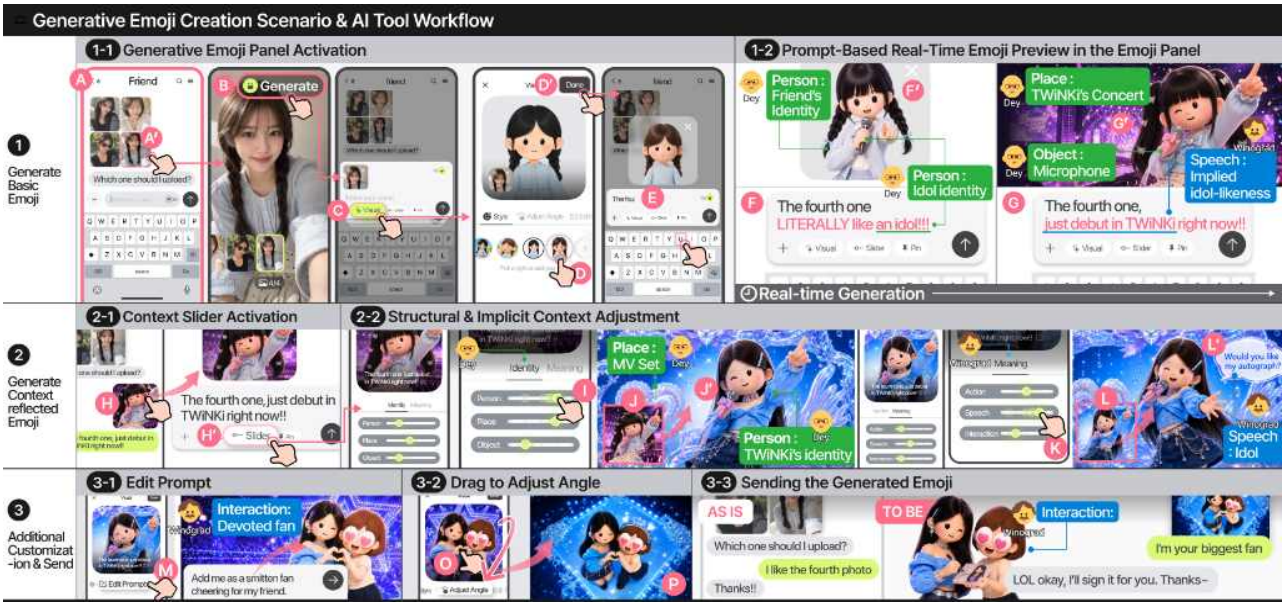


그림 7. 이모지 생성 시나리오 및 AI 툴 활용 프로세스
Fig. 7. Generative emoji scenario & AI tool workflow

고 입력하면 F'와 같이 4번째 사진 속 친구의 양갈래 머리와 메이크업 스타일, 그리고 무대를 마친 엔딩 요정처럼 마이크를 들고, 카메라를 향해 엔딩 포즈를 취하는 아이돌스러운 표현이 Dey의 맥락 관점 중 '사람' 기준의 정체성 반영 지점이다. G는 친구가 평소 좋아하는 가상의 아이돌 그룹인 트윙키(TWiNKi)를 염두해 '당장 TWiNKi 데뷔해!!'라는 프롬프트를 입력하는 상황으로 이 경우 G'처럼 Dey 관점 중 트윙키의 콘서트 공연장인 '장소', 그리고 해당 의상과 헤어스타일, 리본 마이크를 착용한 아이돌 스타일이 적용된 '사물' 기준 정체성이 반영되어 실시간 생성된다. '당장 TWiNKi 데뷔해!!'의 표현은 Winograd의 맥락 관점 중 4번 사진이 데뷔할 정도로 최고라는 의미를 화법으로 함축해 전달하고 있다.

2는 맥락 반영 심화 이모지 생성 단계로 2-1은 이모지에 반영될 Dey, Winograd의 구조적, 함축적 맥락의 비율을 조절하는 슬라이더(Slider)를 활성화하는 과정이다. H처럼 편집 화면 진입 후 H'와 같이 슬라이더 버튼을 탭하면, 2-2처럼 구조적 관점인 Dey의 대표 3개 Identity와 함축적 관점인 Winograd의 대표 3개 Meaning 반영 슬라이더가 활성화된다. 먼저 Dey의 관점인 Identity의 세부 슬라이더는 사람, 장소, 사물로 구성되며, I처럼 세부 항목을 조절할 수 있다. 사용자가 슬라이더를 조절하면 J에서, 트윙키의 정체성과 뮤직비디오 세트장을 반영한 J'로 변경된다. Winograd의 Meaning 세부 슬라이더는 행위, 화법, 상호작용으로 구성되며 마찬가지로 K처럼 조절할 수 있으며 L의 이모지가 L'와 같이 아이돌스러운 함축된 화법이 반영된 말풍선이 추가된다.

3은 세부 이모지 생성 조정 및 전송 단계이다. M처럼 사용자가 "친구를 보고 팬이 된 내 모습을 넣어줘"라는 프롬프트를 입력하면 열성팬이 된 사용자가 아이돌이 된 친구 이모지와 함께 함축적 맥락 의미 요소의 포즈를 취하고 상호작용하는 장면이 생성된다. O는 이후 음악 방송 같은 구도를 연출하는 과정으로 앵글 조정 버튼 탭 후, 드래그로 P처럼 원하는 구도로 이모지를 생성할 수 있다. 또한, 3-3처럼 상대방의 이모지를 활용해 팬 사인회처럼 "고마워. 내가 사인해줄게~"의 프롬프트로 다음 상호작용을 더욱 흥미롭게 이어갈 수 있다.

2) 이모지 생성의 기능 요구사항 및 기술 프로세스

그림 8은 앞서 제시한 사용자 시나리오 기반의 기능 요구사항과 기술 처리 과정의 정리내용이다. 1은 사용자의 사용 프로세스로, 1-1의 텍스트와 이미지 입력 후 1-2의 이모지 생성 결과 확인, 1-3의 슬라이더, 프롬프트 편집을 통한 조정, 1-4의 이모지 전송, 그리고 1-5의 전송 후의 상호작용으로 구성되어 있다. 2는 이에 대응하는 이모지 생성 프로세스로 2-1은 GPTs를 활용해 입력된 텍스트와 이미지의 구조적, 함축적 맥락을 분석하는 단계, 2-2는 Nano Banana를 활용해 이모지를 생성하는 단계이다. 2-3은 Krea Realtime을 활용해 사용자의 조정값을 생성 결과에 실시간 반영하는 단계로, 사용자의 확인과 이모지 생성과 실시간 조정은 사용자의 확인과 수정에 따라 반복하여 이루어진다. 2-4는 생성 결과를 전송하고, 저장하는 결과 관리 단계, 2-5는 Higgsfield를 활용한 전송 이모지 간 상호작용 결과 생성하는 단계이다.

A부터 E는 각 단계에서의 기능 요구사항이다. A는 입력 범위 선택과 맥락 분석, B는 구조적, 함축적 맥락 기반 이모지 생성, C는 생성 결과의 실시간 확인과 조정, D는 결과의 확정, 전송, 저장, E는 전송된 이모지 간 상호작용 생성을 포함한다. 그림 하단의 F부터 H는 기능 구조 설계 기반 기술 구현 방향으로, F는 각 단계별로 적합한 AI 모델을 선택하고 연계하는 AI 다중 모델 통합 기술[29], G는 네트워크와 AI 처리 지연 최소화를 위해 요청 통합, 미리보기 생성, 실시간 갱신 및 최종 생성의 순서로 연산을 분리하는 실시간 생성 파이프라인 StreamDiffusion[30]이다. H는 시스템 경량화를 위해 활용하는 캐시 기반 경량화 기술인 DeepCache[31]로, 변경된 데이터만 갱신하고 기존 생성 결과를 재사용하여 이모지를 생성하는 과정을 경량화 하여 구현할 수 있다.

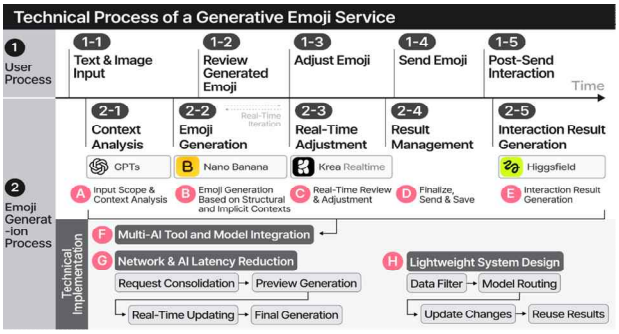


그림 8. 생성형 이모지 서비스의 기술 프로세스
Fig. 8. Technical process of a generative emoji service

3-4 서비스 핵심 가치 기반 디자인 프로토타입

1) 이모지의 감정 맥락 반영 정도 조절

그림 9는 슬라이더를 활용해 구조적, 함축적 맥락이 생성형 이모지에 반영되는 정도를 웹툰 그림체의 적용을 통해 강조하는 추가 시나리오이다. 1은 Dey 관점의 Identity 슬라이더를 사용해 사용자의 감정 상태를 사물 정체성으로 구조적 맥락 표현 강화 사례이다. 사용자가 "그 피규어 품절됐대.. 완전 망했어"라는 문장을 입력하면, 오랜 기다림 끝에 사고 싶었던 피규어가 품절된 상황에서 허탈감을 담은 '망연자실한 표정의 이모지'가 생성된다. 이를 과장된 웹툰처럼 표현하기 위해, A의 비주얼 버튼과 B의 스타일 프리셋 버튼을 탭, 그리고 C의 웹툰 한 장면을 스타일 프리셋으로 업로드 후 C'의 완료 버튼을 누르면, D처럼 이모지에 웹툰 그림체의 넋이 나가는 표현을 적용할 수 있다. 또한, 1-2의 E처럼 슬라이더 버튼을 탭한 후 Dey 관점의 Identity 탭을 선택하면 사람, 장소, 사물의 구조적 맥락 요소를 슬라이더로 조절할 수 있다. F는 빠져나간 영혼의 존재감을 키우기 위해 사물 슬라이더를 높인 모습이다. 2는 Winograd 관점의 Meaning 슬라이더를 활용해 요즘 Z 세대의 밈(Meme)이며, '오늘 만나기 좀 피곤한

데'를 표현하는 '우리... 오늘 만나?'라는 표현을 반영한 사례이다. 2-1처럼 메시지창에 사용자의 사진을 넣고, '우리... 오늘 만나?'를 입력 후, G의 슬라이더 활성화 후 H 버튼을 탭해 '오늘 만나기가 좀 피곤한데'라는 프롬프트를 입력하면, 2-2의 I처럼 침대에서 쉬는 모습이 생성된다. 이후 Meaning에서 화법 맥락을 더 높이면 '오늘 만나지 말고 쉬자'는 함축된 의도가 J처럼 간절한 눈, 생각 벌룬과 함께 시각화된다. 이처럼 슬라이더로 이모지의 감정 표현을 세밀하게 조정할 수 있다.



그림 9. 슬라이더를 활용한 맥락 반영 정도 조절
Fig. 9. Context-driven control using sliders

2) 사용자 간 이모지 인터랙션

앞서 진행한 Think-Aloud 실험에서 기존의 선택형 이모지 방식만으로 맥락을 충분히 반영하기 어렵다는 점을 확인하였다. 이에 본 제안 서비스는 그림 10처럼 이모지 간 결합, 반응 기반 변화, 드로잉 기반 표현을 통해 하나의 장면으로 확장할 수 있는 상호작용 방식을 반영하였다. 1은 이모지 간 조합 과정으로 1-1의 A처럼 사용자가 대화 맥락에 맞는 이모지를 선택하고, 1-2의 B와 같이 조합하고 싶은 이모지에 드래그 앤 드롭하면 각각의 이모지가 하나로 결합되어, 1-3의 C처럼 상대방의 이모지 스타일이 적용된 새로운 표현 방식을 반영할 수 있다. 2는 드래그 인식 기술과 접촉 감지 기술 기반의 반응형 인터랙션 장면이다. 2-1은 상대방이 사용자에게 삐져 등 돌린 이모지를 보낸 상황이며 사용자가 상대방의 웃음을 유도하려는 의도를 담아 강아지풀 이모지로 상대방의 이모지를 간지럽혀 화를 풀어주는 장면이다. 2-2는 상대 이모지의 자세와 표정이 변화하는 과정으로 D처럼 드래그 제스처를 통해 토라진 상태의 이모지는 상대의 인터랙션에 반응하여 자세와 감정 표현이 변화한다. 이를 통해 생성형 이모지는 대화 상대의 개입과 반응에 따라 변화하는 반응형 이모지로의 역할 확장이 가능하다. 3은 드로잉 인터랙션 과정으로 3-1의 E처럼, 이모지에 선글라스를 드로잉하면, 선글라스를 착용한 모습으로 연출된다. 또한, 3-2처럼 Pin 기능을 통해 이모지의 스타일을 고정해 3-3과 같이 이모지 캐릭터의 제스처, 소품과 같은 표현을 응용할 수 있으며, 동적으로 확장도 가능하다. 이처럼 본 제안 서비스는 이모지 간 결합, 상대방의 반응 인터랙션, 드로잉 연출 기능로 확장 생성이 가능하다.

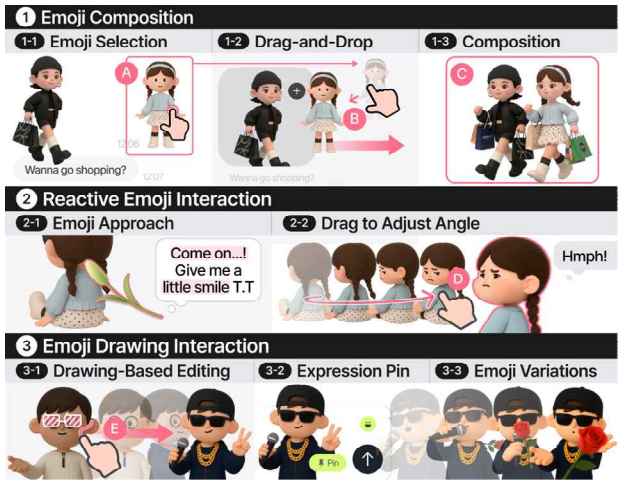


그림 10. 이모지 결합, 반응, 편집 인터랙션
Fig. 10. Emoji composition, reaction, editing interaction

3) 생성형 이모지의 확장 및 활용 구조

그림 11은 생성형 이모지 생성 후, 여러 플랫폼에 확장 활용되는 구조를 정리한 것이다. 1-1은 사용자가 유튜브 유형의 플랫폼에 공유된 생성형 이모지의 특정 스타일을 자신의 이모지에 적용하기 위해 다운로드하는 과정이다. A처럼 공유가 허용된 이모지를 롱-탭하면, B처럼 스타일이나 프롬프트를 가져올 수 있는 선택 창이 팝업되고 스타일 가져오기 기능 선택 시 C처럼 해당 이모지 스타일이 스타일 목록에 추가된다. 이후 해당 스타일을 선택하면, C'처럼 자신의 생성형 이모지에 동일 화풍의 표현 방식이 적용된다. 2는 생성형 이모지의 비즈니스 모델로 2-1처럼 제휴된 특정 브랜드가 제품 이미지, 캐릭터, 로고를 본 제안 서비스에 제공하면 이를 생성형 이모지 프리셋으로 변환 활용하여 해당 브랜드 인지도를 강화할 수 있다. 2-2는 특정 브랜드의 텀블러 제품을 생성형 이모지로 변환한 사례로, 채팅 및 콘텐츠로 사용할 수 있다. 선행 코카콜라사가 OS 이모지의 콜라 비주얼과 닮은 컵을 통해 성공적인 마케팅 효과를 거둔 사례처럼, 생성형 이모지 또한 브랜딩 도입을 통한 마케팅 시도가 가능하다.

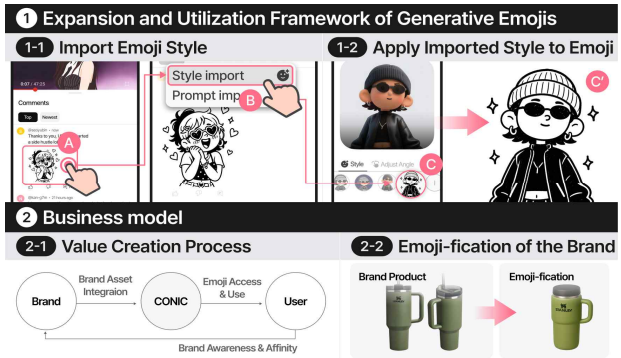


그림 11. 생성형 이모지의 확장 및 활용
Fig. 11. Generative emoji expansion and utilization

Ⅳ. 프로토타입 기반 사용자 테스트와 구조적, 함축적 맥락 기반 가이드라인

4-1 프로토타입 기반 사용자 테스트

그림 12는 제안 서비스의 사용성을 검증하기 위해 진행한 사용자 테스트와 인터뷰를 정리한 내용이다. 1은 사용자가 실제 모바일로 프로토타입을 체험하는 장면으로, 이모지를 직접 생성하고 활용하는 과정을 경험할 수 있도록 사용자 테스트를 구성했다. 2는 인터뷰 방법과 참여자 정보로, 인터뷰는 2025년 11월 29일부터 12월 2일까지 4일간 대면으로 진행되었으며, 기존 OS 제공 이모지와 Apple의 Genmoji의 사용 경험이 있는 사용자 30명을 대상으로 수행하였다. 응답자의 성별은 남성이 12명, 여성이 18명이었으며, 연령대는 10대 6명, 20대 22명, 30대 2명으로 구성되었다. 2-1은 인상 깊은 기능에 대한 질문으로 “대화 맥락을 이해해 두 사용자의 스타일이 함께 반영되는 점이 인상적이다.”, “사용자 이모지끼리 만나는 방식이 흥미롭다.”는 답변을 받았으며, 사용자가 대화 맥락 반영과 이모지 결합 개념에 주목하고 있음을 확인할 수 있었다. 2-2는 기존 이모지와의 차이에 대한 질문으로 “대화 내용에 맞춰 바로 이모지를 생성하여, 이모지 목록을 찾아볼 필요가 없다.”라는 의견처럼 검색 피로를 줄일 수 있다는 점이 장점으로 언급되었다. 2-3에서 Apple의 Genmoji와 같은 생성형 이모지와 비교했을 때는 “기존 Genmoji는 특정 운영체제에 묶여 있어 호환성이 낮다.”, “프롬프트로 이모지를 만들 수는 있지만, 채팅 상황의 맥락이 깊게 반영되지는 않아 보인다.”와 같이 답하였다. 마지막으로 2-4의 경쟁력 관련 질문에서는 “다른 사람의 이모지 스타일도 쓸 수 있는 점이 가장 설득력 있는 차별 요소 같다.”, “브랜드를 연결하는 비즈니스 구조가 재밌고 이해하기 쉽다.”는 의견과 같이, 제안 서비스가 기존 이모지와 차별성이 있음을 확인할 수 있었다.

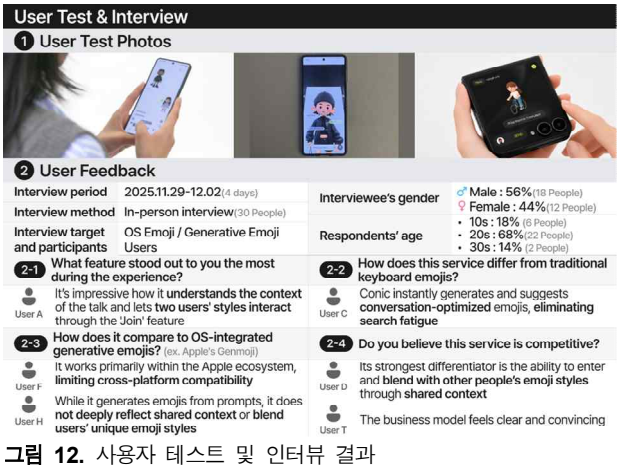


그림 12. 사용자 테스트 및 인터뷰 결과
Fig. 12. User test and interview results

4-2 구조적, 함축적 맥락 기반 생성형 이모지 가이드라인

그림 13은 구조적, 함축적 맥락 기반 생성형 이모지 제작 가이드라인이다. 1은 Dey의 구조적 맥락 가이드라인으로 직관적 이해를 돕기 위해 맥락층을 레이어로 구분하였다. 맥락층은 A의 인물의 외형과 정체성을 구성하는 사람 레이어, B의 공간 맥락을 포함하는 장소 레이어, 그리고 C의 소품의 형태, 크기, 위치를 포함하는 사물 레이어로 구성되어 있다. 구조적 맥락은 개별 요소 간 핵심 맥락이 흐려지지 않도록 주의해야 한다. 구조적 맥락의 세부 가이드라인은 다음과 같다. A-1처럼 사람 슬라이더의 비중을 높여 장소와 사물을 유지하면서 인물의 정체성과 외형을 조정할 수 있으며 이때 A-2와 같이 표정, 포즈를 조합할 시 대화자와 대화 맥락에 맞는 인물의 정체성 반영 표현이 중요하다. B-1처럼 장소 슬라이더의 비중을 높이면 인물과 사물을 유지하면서 대화 맥락에 맞는 공간과 분위기를 조정할 수 있다. 이때는 B-2처럼 배경, 조명 조합이 가능하며, 장소 정체성을 세부 조절할 수 있다. C-1과 같이 사물 슬라이더의 비중을 높이면, 인물과 장소를 유지하면서 특정 사물의 형태 조정이 가능하며 C-2처럼 사물의 형태, 크기, 인물 활용의 조합이 중요하다.

2는 Winograd의 함축적 맥락의 가이드라인으로 D는 인물의 행위를 구성하는 행위 레이어, E는 발화 내용, 표현 방식을 구성하는 화법 레이어, 그리고 F는 인물 간 관계와 반응을 구성하는 상호작용 레이어이다. 함축적 맥락 이모지를 생성할 때는 사용자의 숨은 의도, 관계와 뉘앙스를 담도록 하는 것이 중요하다. 함축적 맥락 기반의 세부 가이드라인은 다음과 같다. D-1과 같이 행위 슬라이더의 비중을 높이면 화법과 상호작용을 유지하면서 인물의 동작을 조정할 수 있다. 이때 D-2와 같이 포즈, 손동작과 같은 요소를 조합해 행위의 의미가 전달되어야 한다. E-1처럼 화법 슬라이더의 비중을 높여 행위와 상호작용은 유지하고 발화 방식, 감정의 강도를 조정할 수 있으며 E-2와 같이 말풍선, 서체를 조합해 반어, 비유 등 화자 특유의 세부 화법 반영이 중요하다. F-1처럼 상호작용 레이어의 비중을 높이면 행위와 화법을 유지하면서 인물 간 관계와 반응을 조정할 수 있으며 F-2와 같이 시선, 거리를 조합하여 상대와 상호작용하는 모습의 반영이 중요하다.

3은 구조적, 함축적 맥락의 결합 가이드라인으로, 전달 목적에 따라 필요한 레이어를 조합할 수 있도록 시각 표현의 우선순위를 정리한 것이다. 3-1의 핵심 의미 전달은 C의 사물, A의 사람, D의 행위의 순서로, 무엇을 대상으로 누가 어떤 행동을 하는지에 대한 시각 표현 순서가 필요하다. 3-2의 의도 구체화는 A의 사람, D의 행위, E의 화법 순서로, 누가 어떤 태도로 어떻게 말하는지를 통해 발화의 의도와 뉘앙스를 표현하는 것이 중요하다. 마지막으로 3-3의 장면 맥락 보완은 C의 사물, B의 장소, F의 상호작용을 순서로 하며, 무엇이 어디에서 어떤 관계에 놓이는지 순차 조율할 필요가 있다.

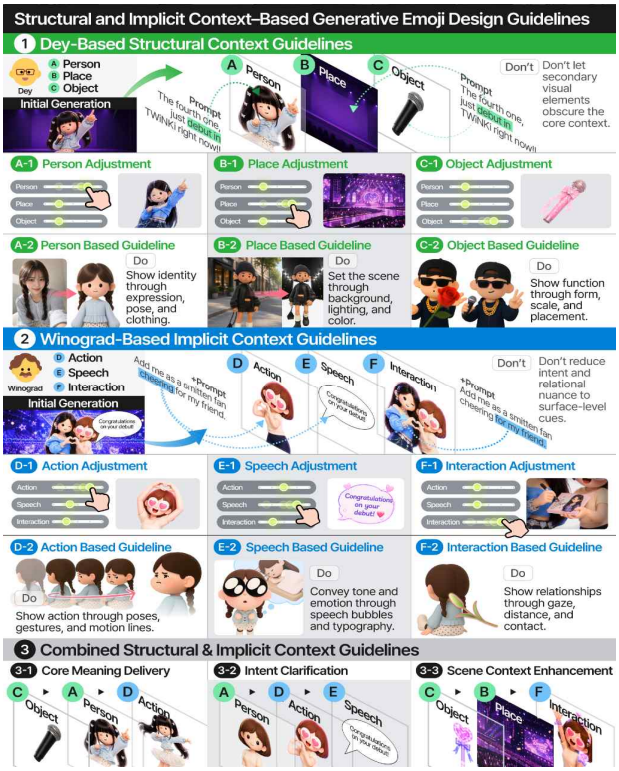


그림 13. 구조적, 함축적 맥락 기반 가이드라인
Fig. 13. Structural and implicit context guidelines

V. 결 론

생성형 AI는 최근 Apple의 Genmoji 사례처럼 이모지도 원하는 대로 직접 생성해서 사용할 수 있는 표현의 영역을 확장시켰다. 하지만 현재는 프롬프트 기반 외형 표현에 집중되어 있어 장소, 관계, 상황, 말투와 같이 실제 소통 상황에서 중요한 역할을 하는 맥락 요소의 반영은 부족하다. 이에 본 연구는 생성형 이모지에 적용 가능한 맥락 반영 체계를 정의한 후 Dey의 구조적 맥락과 Winograd의 함축적 맥락을 중심으로 맥락 반영 생성형 이모지 서비스의 제작 방향성을 제안하고자 했다. 본 연구의 결과는 다음과 같다. 첫째, 생성형 이모지의 대표 상용화 사례인 Apple의 Genmoji를 분석한 결과, 결과물이 주로 텍스트 기반 외형 특징 묘사에 의존하는 한계가 있었다. 둘째, 문맥 기반 기계 번역과 CoT, CoE 기반 감정 추론 연구를 분석한 결과, 최근 AI는 텍스트로 사용자의 의도와 상황을 추론할 수 있음을 확인하였다. 또한, Krea AI, Motion Stream과 같은 리얼타임 생성 기술 분석을 통해, 생성형 이모지에 맥락을 실시간으로 반영하는 기술적 가능성을 확인하였다. 셋째, Dey와 Winograd의 맥락 개념을 재정의하여 구조적 맥락과 함축적 맥락을 생성형 이모지의 맥락의 기준으로 반영하였다. 넷째, 모바일 메신저를 이용하는 사용자 55명 대상 설문 조사 후, 퍼소나 워크샵을 진행해 메시지를 해석할 때 중요하게 활용하는 단서를 기준으로 행동 원형을 유형화했다. 그 결과 관찰 가능한 정보로 메시지를 이해하는 구조적 단서 해석 중심 유형과, 행동과 표현 속 숨은 의미로 이해하는 함축적 의미 중심 유형의 두 가지 유형을 도출하였다. 이후 각 유형의 반복되는 요소를 분석해 각각 사람, 장소, 사물, 그리고 행위, 화법, 상호작용의 각 세 개의 세부 해석 단서를 도출하여 생성형 이모지의 맥락 기준으로 활용하였다. 다섯째, 생성형 AI의 구조적, 함축적 맥락 반영 가능성을 검증하기 위해, 이모지를 매일 1회 이상 사용하는 10명을 대상으로 Think-Aloud 기법을 활용한 이모지 생성 실험과 심층 인터뷰를 진행하였다. 그 결과, 돈, 건물 및 걷는 모습과 같은 직접적인 사물 및 행위 단서는 생성형 AI와 피험자 답변이 거의 일치하였지만, 감정 일치도는 피험자 평가 71.0점, AI 평가 45.0점으로 26점의 차이가 나타나 '초조함'과 '미안함'처럼 복합적인 감정과 의도가 주로 포함된 함축적 맥락 표현 과정 시에는 더 많은 변인에 대한 세밀한 조정과 검토가 필요함을 파악하였다. 또한, 사용자는 생성 결과의 근거와 세부 표현의 변경에 대한 니즈가 있음을 확인하였다. 여섯째, 생성형 이모지의 구현 가능성을 검증하기 위해 ChatGPT, Krea AI, Higgsfield 3종의 생성형 AI와 Figma를 연계 활용하여 프로토타입을 제작하였다. 프로토타입에는 사용자가 슬라이더로 구조적, 함축적 맥락의 반영 정도를 조절하여 추가 생성할 수 있게 설계하였으며, 이모지 결합, 이모지 간 반응, 드로잉 커스텀과 같은 기능들을 적용하였다. 마지막으로 사용자 테스트를 통해 생성형 이모지의 사용성에 대한 30명 대상 답변을 수집하고 가이드라인 초안을 제작하였다. 본 연구는 콘셉트 기획부터 디자인 단계까지의 프로토타입 제안 연구이나 이후 구현 방향을 구체화하기 위해 사용자 입력, 맥락 분석, 이모지 생성, 실시간 조정, 결과 및 상호작용의 기술 프로세스를 함께 정리하였다. 단계별 기능 요구사항은 입력 범위 선택, 맥락 분석, 구조적 맥락과 함축적 맥락 기반 실시간 이모지 생성, 전송, 이모지 간 상호작용으로 구분하였다. 기능 구조 설계 기반 기술 구현 방향으로는 다중 AI 모델 통합 구조와 네트워크와 AI 지연을 줄이기 위한 데이터 선별, 변경 값 갱신, 기존 결과 재사용을 통해 시스템의 부담을 줄이는 방향을 제시하였다. 다만, 이는 프로토타입 기반의 제안으로 실제 모델의 성능과 처리 지연 감소 효과 및 시스템 경량화 성능은 후속 구현을 통해 검증할 필요가 있다. 이외에도 실시간 생성 과정에서 발생할 수 있는 윤리적 가이드라인와 정책 보완은 병행되어야 한다. 본 연구는 기존 생성형 이모지와 달리 사용자의 맥락을 생성 기준으로 설정하고, 이 Dey와 Winograd의 구조적, 함축적 맥락 관점에서 생성 가이드라인으로 제안했다는 점에서 차별점이 있다. 또한, 사용자가 직접 맥락의 반영 정도를 슬라이더로 조절할 수 있으며 이모지 결합, 반응형 인터랙션, 스타일 공유로 생성형 이모지를 상호작용형 소통 도구로 제안하였다는 점에 의의가 있다. 본 연구가 사용자 맥락 반영 생성형 이모지의 초기 가이드라인으로 후속 연구에 도움이 되기를 바란다.

감사의 글

이 논문은 2025학년도 홍익대학교 디자인컨버전스학부 졸업 전시 작품을 바탕으로 추가 연구를 진행하였습니다.

참고문헌

[1] S. Aljasir, “Emoji as a Social Presence Tool Among Arab Digital Media Users: Do the Demographic Variables of the Sender Play a Role?,” *Social Science Computer Review*, Vol. 42, No. 1, pp. 270-284, February 2024. https://doi.org/10.1177/08944393231181638

[2] S. E. Fahlman. The Birth, Spread, and Evolution of the Smiley Emoticon [Internet]. Available: http://www.cs.cmu.edu/~sef/Smiley2021.pdf.

[3] J. McCurry. The Inventor of Emoji on His Famous Creations – and His All-Time Favorite [Internet]. Available: https://www.theguardian.com/technology/2016/oct/27/emoji-inventor-shigetaka-kurita-moma-new-york-text.

[4] Emojipedia. Unicode 6.0 Emoji List [Internet]. Available: https://emojipedia.org/unicode-6.0.

[5] Apple. Apple Intelligence Now Features Image Playground, Genmoji, and More [Internet]. Available: https://www.apple.com/ci/newsroom/2024/12/apple-intelligence-now-features-image-playground-genmoji-and-more/.

[6] E. E. Manganari, “Emoji Use in Computer-Mediated Communication,” *The International Technology Management Review*, Vol. 10, No. 1, pp. 1-11, January 2021. https://doi.org/10.2991/itmr.k.210105.001

[7] M. Kim, “A Basic Study on the Construction of a SOLARS Unicode-Based System,” *Bulletin of the Seoul National University Library*, Vol. 127, pp. 4-36, December 2004.

[8] Unicode Consortium. Unicode Technical Standard #51: Unicode Emoji [Internet]. Available: https://www.unicode.org/reports/tr51/.

[9] The museum of Modern Art (MoMA). Inbox: The Original Emoji, by Shigetaka Kurita [Internet]. Available: https://www.moma.org/calendar/exhibitions/3639.

[10] A. Pardes. The WIRED Guide to Emoji [Internet]. Available: https://www.wired.com/story/guide-emoji/.

[11] H. S. Jung, A Study on Emoji Communication Effect for Visual Storytelling, Master’s Thesis, Hongik University, Seoul, 2017.

[12] CNET Korea. Memoji Exactly Like Me: Early Meeting with Apple iOS 12 Beta [Internet]. Available: https://www.youtube.com/watch?v=mr3-e_gOBGI.

[13] Samsung Newsroom Malaysia. [Interview] “AR Emoji Can Be a Visual AI Assistant” – Developers on the Galaxy S10 AR Emoji [Internet]. Available: https://news.samsung.com/my/interview-ar-emoji-can-be-a-visual-ai-assistant-developers-on-the-galaxy-s10-ar-emoji.

[14] Apple Support. Create Your Own Emoji with Genmoji on iPhone [Internet]. Available: https://support.apple.com/ko-kr/guide/iphone/iph4e76f5667/ios.

[15] Apple. Apple Intelligence Now Features Image Playground, Genmoji, Writing Tools Enhancements, Seamless Support for ChatGPT, and Visual Intelligence [Internet]. Available: https://nr.apple.com/dA2t5F6hg2.

[16] Apple. New Features Available with iOS 26 [Internet]. Available: https://www.apple.com/kr/os/pdf/All_New_Features_iOS_26_Sept_2025.pdf.

[17] M. Fadaee, Understanding and Enhancing the Use of Context for Machine Translation, Ph.D. Dissertation, University of Amsterdam, Amsterdam, Netherlands, November 2020.

[18] J. Wei, X. Wang, D. Schuurmans, M. Bosma, B. Ichter, F. Xia, ... and D. Zhou, “Chain-of-Thought Prompting Elicits Reasoning in Large Language Models,” in *Proceedings of the 36th International Conference on Neural Information Processing Systems (NeurIPS 2022)*, New Orleans: LA, pp. 24824-04837, December 2022. https://doi.org/10.48550/arXiv.2201.11903

[19] Y. K. Lee, I. Lee, M. Shin, S. Bae, and S. Hahn, “Chain of Empathy: Enhancing Empathetic Response of Large Language Models Based on Psychotherapy Models”, *Korean Journal of Cognitive Science*, Vol. 35, No. 1, pp. 23-48, 2024. https://doi.org/10.19066/cogsci.2024.35.1.002

[20] B. Tömekçe, M. Vero, R. Staan, and M. Vechev, “Private Attribute Interface from Images with Vision-Language Models,” in *Advances in Neural Information Processing Systems 37 (NeurIPS 2024)*, Vancouver, Canada, pp. 103619-103651, 2024. https://doi.org/10.52202/079017-3291

[21] J. H. Ahn and C. Lee, “A Study on the Methodology of Emoji Automatic Authoring Based on Communication Data,” *A Journal of Brand Design Association of Korea*, Vol. 21, No. 3, pp. 31-42, September 2023.

[22] J. Gu, A Study on Sentiment Analysis of Unicode Emoji Using Deep Learning, Master’s Thesis, Graduate School of Information and Telecommunications, Konkuk University, Seoul, August 2024.

[23] Krea AI. Realtime [Internet]. Available: https://docs.krea.ai/user-guide/features/realtime.

[24] S. Luo, Y. Tan, L. Huang, J. Li, and H. Zhao, “Latent Consistency Models: Synthesizing High-Resolution

Images with Few-Step Inference," arXiv:2310.04378, 2023. https://doi.org/10.48550/arXiv.2310.04378

[25] J. Shin, Z. Li, R. Zhang, J.-Y. Zhu, J. Park, E. Shechtman, and X. Huang, "MotionStream: Real-Time Video Generation with Interactive Motion Controls," in *Proceedings of the International Conference on Learning Representations (ICLR)*, 2026.

[26] A. K. Dey, "Understanding and Using Context," *Personal and Ubiquitous Computing*, Vol. 5, No. 1, pp. 4-7, February 2001. https://doi.org/10.1007/s007790170019

[27] T. Winograd and F. Flores, *Understanding Computers and Cognition: A New Foundation for Design*, Norwood, NJ: Ablex Publishing Corporation, 1986.

[28] S. M. Mun and K. I. Nam, "Effects of Training Conditions of Think-Aloud and Prior Knowledge on Think-Aloud Responses in Reading Comprehension and Expository Text Processing," *Asian Journal of Education*, Vol. 11, No. 4, pp. 1-26, December 2010. https://doi.org/10.15753/aje.2010.11.4.001

[29] Y. Shen, K. Song, X. Tan, D. Li, W. Lu, and Y. Zhuang, "HuggingGPT: Solving AI Tasks with ChatGPT and Its Friends in Hugging Face," in *Advances in Neural Information Processing Systems 36 (NeurIPS 2023)*, New Orleans: LA, pp. 38154-38180, December 2023. https://doi.org/10.52202/075280-1657

[30] A. Kodaira, C. Xu, T. Hazama, T. Yoshimoto, K. Ohno, S. Mitsuhori, ... and K. Keutzer, "StreamDiffusion: A Pipeline-level Solution for Real-Time Interactive Generation," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, Honolulu: HI, pp. 12371-12380, October 2025. https://doi.org/10.1109/ICCV51701.2025.01150

[31] X. Ma, G. Fang, and X. Wang, "DeepCache: Accelerating Diffusion Models for Free," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle: WA, pp. 15762-15772, June 2024. https://doi.org/10.1109/CVPR52733.2024.01492



서유빈(Yu-Bin Seo)

2021년～현　재: 홍익대학교 디자인컨버전스학부 재학

※관심분야 : UX Design, Interface design, Interaction design, Generative AI



윤현경(Hyeon-Gyeong Yun)

2022년～현　재: 홍익대학교 디자인컨버전스학부 재학

※관심분야 : UXUI Design, Interface design, Visual Design



김건동(Geon-Dong Kim)

1998년 : 홍익대학교 대학원 (미술학석사)

2007년 : Rhode Island School of Design, USA (MFA, 미술학석사)

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, Generative AI, Meaning Making 등