



Unlocking Creator-AI Synergy: Challenges, Requirements, and Design Opportunities in AI-Powered Short-Form Video Production

Jini Kim*

jinik@andrew.cmu.edu

Carnegie Mellon University
Pittsburgh, PA, USA

Hajun Kim*

hajune28@neowiz.com

Neowiz
Gyeonggi-do, Korea

ABSTRACT

The emergence of AI-Powered Short-Form Video Generators (ASVG) has showcased the potential to streamline production time and foster creative ideas. Despite their widespread adoption, research has underexplored ASVG, especially from creators' perspectives. To evaluate the role of ASVG as creator-centered collaborators, we conducted mixed-method research: (1) interviews ($N = 17$) and (2) a participatory design workshop ($N = 12$) with short-form video creators. In our interviews, we investigated creators' production process and challenges in creating short-form videos. In participatory workshops, short-form video creators envisioned AI-powered video tools, addressing their requirements and AI collaboration perceptions. Our findings indicate ASVGs can provide various advantages including inspiration, swift access to video sources, and automated highlight generation. To put things in perspective, we also underscore concerns arising from AI collaboration, including potential creator identity dilution, reduced creative output, and information bubble. We also discuss design considerations when designing ASVG to retain their creative values.

CCS CONCEPTS

- Human-centered computing → Empirical studies in collaborative and social computing; Empirical studies in HCI.

KEYWORDS

Short-form video generation, Human-AI Collaboration, Algorithmic experience, Creators, Creative economy, Participatory design

ACM Reference Format:

Jini Kim and Hajun Kim. 2024. Unlocking Creator-AI Synergy: Challenges, Requirements, and Design Opportunities in AI-Powered Short-Form Video Production. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*, May 11–16, 2024, Honolulu, HI, USA. ACM, New York, NY, USA, 23 pages. <https://doi.org/10.1145/3613904.3642476>

1 INTRODUCTION

Online platforms such as YouTube, Instagram, and TikTok are expanding the size of the creator economy, offering short-form video creators opportunities to generate revenue and gain recognition through their creative work [2, 9, 60]. The mobile-friendly nature of short-form videos has recently drawn significant interest from the public [18, 37] with YouTube Shorts having over 2 billion users [45] and TikTok reaching over 1.5 billion users [23]. Notably, over 83% of TikTok users have engaged in both the watching and uploading of videos [15], highlighting active participation of not only professional creators but also ordinary users stepping into the realm of creators.

However, despite its immense popularity, short-form video production still poses many challenges to short-form video creators. They experience highly expedited cycle of content consumption compared to traditional long-form videos [63], facing difficulty in constantly having to come up with new ideas for a next round of captivating videos. Furthermore, maintaining video quality for consumers necessitates multiple iterative steps of optimization within a limited time frame, while users rapidly swipe through contents and seek additional contents that draw their attention. Nevertheless, there is a scarcity of research addressing the practical challenges faced by short-form video creators during their production, and even fewer studies focusing on supporting them.

Recent works in Human-Computer Interaction (HCI) and Computer-Supported Cooperative Work (CSCW) investigate the impact of AI (Artificial Intelligence) on creative work [27, 29, 75, 76]. They suggest that judicious AI intervention can enhance human capabilities in work environments and, potentially, their creativity, responsibility, and self-efficacy [26, 57]. These factors are crucial components of short-form video creators' motivation [3, 13], indicating the potential for improving their overall work cycle through collaboration with AI. Surprisingly, there has been minimal research on the design of AI collaborative tools to support them.

In our research, we explore design opportunities for AI short-form video generators to elicit synergistic effects between short-form video creators and AI throughout the short-form video production process. We conducted a mixed-method study: (1) semi-structured interview and (2) participatory workshops. Through semi-structured interviews ($N = 17$), we investigated the short-form video production process and grouped it into the three distinct short-form video creator types: (1) Participatory Entertainer, (2) Knowledge Sharer, and (3) Highlight Capturer. Following the framework of creative process [9], we analyzed the difficulties specifically associated with each type of short-form video creators.

*Both authors contributed equally to this research.



This work is licensed under a Creative Commons Attribution International 4.0 License.

CHI '24, May 11–16, 2024, Honolulu, HI, USA

© 2024 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0330-0/24/05

<https://doi.org/10.1145/3613904.3642476>

We held participatory design workshops with short-form video creators ($N = 12$) to co-design an AI collaborator tool that meets their specific needs. However, recent studies [36, 72, 73] have shown that end-users often struggle to conceptualize AI systems due to their unpredictable capabilities and outputs. To address this, we employed various stimuli to help participants envision collaborative AI tools. Throughout a series of four workshops, participants generated concrete design proposals to address the difficulties identified during interviews. Analyzing these design proposals allowed us to identify potential opportunities and considerations for designing AI-powered tools for short-form video content creation.

In summary, our research contributes an in-depth understanding of the production cycle of short-form video creators, highlighting their unique types, and challenges they encounter. It also uncovers design opportunities, concerns, and considerations for designing AI-powered short-form video generators to facilitate more effective integration of AI in the content creation process. A wide range of stimuli were utilized to help short-form video creators envision AI features through the participatory design process. Overall, our findings democratize AI collaborative tool design by involving end-users, ensuring their insights are thoroughly reflected in the design process.

2 RELATED WORK

In this section, we provide an overview of the existing literature pertaining to the short-form video creator economy and the challenges creators face. Furthermore, we explore the HCI literature on the integration of AI in content production tools and discuss user-centered design methodologies tailored to AI tools.

2.1 Potential of AI Short-Form Video Generators for Creator Economy

Online platforms offer creators opportunities to generate revenue and gain reputation through their creative activities. The expansive and widespread use of such platforms positions the creator economy as a core part of modern market and culture [2, 9, 60]. Notably, short-form video contents play a critical role in the advancement of the creator economy. Due to their rapid growth in traffic and increased public interests [16, 18, 37], creators are increasingly creating more short-form videos, and more industries are using it for marketing [12]. Their works are garnering recognition for creativity, enriching our culture [53] and communication [66].

Despite their growing impact, short-form video creators face significant challenges in producing their videos. Such challenges include time constraints of video often being less than a minute [21], constantly maintaining fresh ideas amid quickly fading popularity [63], and a lack of motivation [3, 13, 60]. These various challenges often force them to give up and leave the short-form contents platform, with previous studies underscoring the gravity and the need to address the issues. [9, 13, 60].

Meanwhile, many previous works deal with the development of tools and technologies to assist video production in such platforms [58, 67, 78]. In particular, studies have revealed the potential of AI to innovate the editing experience [32, 58, 62]. Soe proposed a comprehensive study focused on automating each stage of the video editing workflow [59]. However, those studies mainly focused

on long-form videos, overlooking the unique characteristics of short-form videos. Although some research on AI-Powered Short-Form Video Generators (ASVG) do exist, they mainly focused on the technical or commercial aspects [39, 68], not considering the practical experiences and challenges faced by real-life short-form video creators in video production.

To create a human-centric, AI-powered video production tool for content creators in the realm of short-form videos, we underscore the need to understand and address their challenges. Therefore, our research explores the process of short-form video production through interviews with real-life short-form video creators, identifying the practical challenges they encounter during production. Then, we conduct a participatory design workshop to collaborate with them to understand the requirements and design opportunities for future ASVG.

2.2 Impact of AI Integration on Content Creation

The HCI community has actively explored ways to integrate AI into applications to augment experts' capabilities [75, 76]. This includes technical guidelines [1, 35], systems [8, 50], and interfaces [25, 74]. This phenomenon has recently gained additional momentum with the emergence of generative models such as Large Language Models (LLMs) [4, 6, 61] and Large-scale Text-to-Image Generation Models (LTGM) [54, 55]. These models can generate high-quality outputs in various modalities such as text and image for every request [42] and handle both unseen and open domain tasks with only an instruction and/or few-shot examples [6, 27]. Empowered by these opportunities, studies have suggested leveraging generative models to automate experts' workflow and expand their creativity. Kim et al. introduced LMCanvas, a block-based interface that empowers writers to interact seamlessly with LLMs, enhancing their writing skills [27]. Similarly, Ko et al. conducted interviews with visual artists to establish design guidelines for integrating LTGMs, aiming to support the creative process and communication within their workflow [29].

For content creators, improving the quality of their work and supporting their ideas and self-expression are the key motivators for producing content [3, 13]. Recently, HCI researchers have found that an appropriate level of AI intervention can enhance humans' abilities while also ensuring their sense of self-efficacy [26, 57]. Some studies have investigated the impact of AI integration in real-world content production scenarios. We-toon [28] is a tool designed to facilitate communication between writers and artists in the creation of webtoon – digital comics optimized for mobile screens. The tool enables writers to easily revise sketches using high-level controls. On the other hand, Choi and Kang et al. have highlighted that YouTube creators faced constraints in their creative work and felt frustrated losing control due to the uncertainty of YouTube recommendation algorithm [9].

The advancements in AI capabilities imply AI's immense potential to support content creators. AI can collaborate with content creators to enhance the quality, creativity, and self-efficacy of their work. Although the advent of generative models underscores the need for further research into AI's role in assisting content creators, studies focusing on the immensely popular short-form contents and

their creators, remain underexplored. Thus, our goal is to identify areas where AI can assist in innovating the workflow of short-form video creators, and collaborate with them to explore AI-powered tools that satisfy their needs and preserve balanced intervention.

2.3 Participation-Driven Design to Envision User-Centric AI

HCI researchers have widely employed participatory design and co-design methods to design digital technology from the perspective of the end-users [20, 52, 77]. These methods are effective for identifying unexplored users' challenges at the early stages and offering solutions tailored to their needs. These approaches significantly lowers the barrier to user adoption [51, 56]. By involving them in the design process, researchers can integrate their domain-specific insights into design, making the technology more user-friendly.

Recent research has shed light on the challenges of conceptualizing AI systems. These challenges arise from the unpredictable capabilities and complex outputs of AI [36, 72, 73], as well as other significant issues. These issues include lack of understanding the intersection between AI technology and human needs [47], the tendency to overestimate AI capabilities [30], difficulties in prototyping [22], and managing risks, with a special focus on ethics and social impact [14, 69]. While the process of conceptualizing AI may add complexity to the design process, consistent emphasis is being placed on the necessity of involving end-users in the user-centered design of AI systems [17, 34, 40, 48, 65]. In addition, the HCI community is actively exploring various approaches to better understand and leverage AI capabilities empowered by the use of stimuli [19, 24, 31, 43] such as the abstraction of AI capabilities into 40 features across various domains [76], the Creativity Matrix that utilizes action verbs to represent human capabilities [11], and the Insight-Actions-Outcomes cards for formulating AI ideas [41]. While previous attempts had some effectiveness, they had limitations in addressing the full spectrum of AI conceptualization aspects. In our participatory workshop, we used various methods to help short-form video creators envision AI concepts for enhancing their content creation: We provided a list of challenges in short-form video production, allowing creators to conceptualize AI solutions tailored to their specific challenges. Leveraging the AI capabilities framework [76], we introduced AI capability cards to simplify complex AI functionalities. Inspired by Zhang [77], we integrated video editing tool interventions and a working AI-infused editing tool prototype. These elements showcased how humans and AI can interact, encouraging the short-form video creators to imagine detailed AI-powered tools and their potential impact. We used the Idea Summary to ensure short-form video creators understood AI capabilities and to help them organize and clarify ideas on addressing specific challenges. Through facilitated discussions, we not only explored AI's potential in the video production process but also addressed user concerns regarding the ethical and social impact of AI integration.

3 STUDY 1: SHORT-FORM VIDEO CREATOR INTERVIEWS

To understand their creative process, motivation, and challenges, we conducted semi-structured interviews with short-form video

creators. In this section, we first describe background information about the choice of platforms and focus of our research. We also detail our recruitment process, participant dynamics, interview protocol, and results of the data analysis.

3.1 Context of Research

Our study demonstrates YouTube Shorts, Instagram Reels, and TikTok as the research context, considering the scale and maturity of these platforms, which are among the most widely used short-form video platforms worldwide [7, 46, 64]. We investigate the experiences of short-form video creators on these platforms. These platforms have varying regulations for video length; typically, the short-form video content consumed on these platforms is around 1 minute in length. YouTube, known for its long-form content as well, permits up to 1-minute videos in its Shorts section. While Instagram Reels and TikTok allows content up to 3 minutes, it is notable that the average length of content consumed by users is 33.0 seconds on Instagram Reels, 41 seconds on YouTube Shorts and 61.3 seconds on TikTok [49]. All platforms enable creators to generate revenue through their creative activities. YouTube Shorts has the YouTube Partner Program (YPP)¹, Instagram has the Reels Play Bonus Program², and TikTok operates the Creator Fund³. Additional income can be made through brand sponsorships, advertising, and donations. However, all these fundamentally depend on video views, necessitating consistent creative work to maintain views and grow their accounts. As the short-form market continues to expand [12, 44], short-form video creators are increasingly challenged to stay competitive in this evolving landscape. They face difficulties in producing diverse content within short cycles [63]. To alleviate the burden on short-form creators and keep them motivated, we conducted an in-depth exploration of their practices and challenges in short-form video production with participants across three different short-form video platforms. While each platform offers unique experiences, our preliminary study focuses on investigating commonalities in the creative process of short-form video production and identifying shared challenges faced by creators.

3.2 Recruitment

To recruit participants with substantial experience in short-form video creation, we invited individuals who met the following criteria: short-form video creators (1) who have actively managed TikTok, Instagram, or YouTube channels for at least one year, (2) who have a minimum of 1,000 subscribers on their YouTube or Instagram channels, or 10,000 followers on TikTok channels, (3) who generate revenue from their channels, including on-platform sources like view-based earnings, ad revenues, and subscriber donations, as well as off-platform income like lecturing about content or production, (4) who have consistently uploaded short-form videos at least once a month over the past year, and (5) who have created a minimum of 20 short-form videos. Participants were recruited through four methods: we posted recruitment announcements in (1) online communities tailored for video creators (e.g., KTUBE

¹https://youtube.com/intl/en_us/creators

²<https://creators.instagram.com/earn-money>

³<https://tiktok.com/creators>

Participant ID (Gender, Age)	Channel Category	Duration of Production	Platforms Worked	Subscribers (Main Platform)	Commitment Level
P1 (M, 25)	Dance	1 year	Tiktok	432K	Part-time
P2 (F, 38)	Beauty, Lifestyle	2 years	Instagram	7.8K	Full-time
P3 (M, 34)	Work out	1 year	Instagram	3.5K	Part-time
P4 (M, 27)	Humor, Magic	2 years	Youtube	1.1M (Team) 27.3K (Solo)	Part-time
P5 (F, 23)	Life Hack	4 years	Tiktok	13K	Full-time
P6 (M, 40)	Family Vlog	3 years	Youtube	2.47K	Full-time
P7 (M, 28)	Fishing	7 years	Youtube	54.1K	Full-time
P8 (M, 30)	Cooking	1 year	Youtube	130K	Part-time
P9 (M, 30)	Video Editing	10 years	Youtube	13.3K	Part-time
P10 (M, 36)	Car Reviews	4 years	Youtube	41.9K	Full-time
P11 (M, 36)	Video Editing	5 years	Youtube	79.7K	Full-time
P12 (M, 26)	Fashion	4 years	Youtube, Instagram	45.3K	Full-time
P13 (F, 26)	Travel, Photos	2 years	Instagram	80K	Full-time
P14 (M, 25)	Cooking	2 years	Youtube	308K	Full-time
P15 (M, 33)	Meme	2 years	Youtube	26K	Part-time
P16 (M, 27)	Photograph	5 years	Instagram	29K	Part-time
P17 (F, 27)	Dance, Humor	1 year	Instagram, Tiktok	5K	Part-time

Table 1: Interview participants' demographic details, channel information, platforms were working on (indicated by YouTube for YouTube Shorts, Instagram for Instagram Reels, and TikTok), subscribers, and their levels of commitment.

⁴, Iamyouuber⁵), (2) social media platforms like Instagram and Facebook, (3) various universities' online community websites, and (4) direct invitations sent via email to short-form video creators who had provided business emails on their channels. To ensure a diverse pool of participants in terms of their experience as short-form video creators, we included a link to our pre-screen survey in the recruitment announcements and invitation emails. This pre-screen survey gathered information on whether they worked as part-time or full-time, the types of content they created, the tools they used, and the number of hours dedicated to content production.

3.3 Participants

We had a total of 17 short-form video creators (4 female, 13 male). Among them, 9 exclusively operated YouTube Shorts channels, 4 on Instagram Reels, 2 managed a TikTok channel, and 2 operated on multiple platforms simultaneously. In terms of platform preference, 10 participants primarily focused on YouTube Shorts, 5 predominantly used Instagram Reels, and 2 primarily utilized TikTok. These channels covered a wide range of topics, including dance, beauty, drama, movies, cooking, fashion, travel, lifestyle tips, and fitness. The average age of the participants was 28.16 years ($SD = 5.15$), and their production experience ranged from 1 year to 10 years. Of the participants, 9 were full-time creators, and 8 were part-time creators (See Table 1).

⁴<http://ktube.kr/>

⁵<https://cafe.naver.com/MyCafeMain.nhn?clubid=29579937>

3.4 Interview Protocol

We conducted semi-structured interviews remotely via Zoom video conferencing. All sessions lasted for 60 minutes and were organized into three main sections. In the first part of the interview, we introduced the study's focus, ranging from the challenges creators faced during production, their production process, and to their well-being. In the second part, we asked about their personal motivation to produce the short-form video contents, their overall experience of content production, and the impact such experiences have had on their lives. In the third part, we focused on understanding the process of the short-form video production, participants' challenges during production and using existing tools, and the strategies they employed to overcome these challenges. Prior to asking these questions, we reviewed channels run by the participants to tailor our inquiries to the characteristics of their short-form videos and the tools they used. All interviews were conducted in Korean and recorded with participants' consent; all participants ran their channels in Korean. Participants were compensated with 40,000 KRW (Approx. 30 USD) for each session.

3.5 Interview Analysis

All interviews were recorded via Zoom and transcribed using the Clova Note⁶, automatically converting voice to text. The quoted statements in this paper were translated into English. Following

⁶<https://clovanote.naver.com/>

a qualitative thematic analysis approach [5], the two researchers conducted an open coding on a subset of the transcripts, and independently generated initial codes by focusing on the content creation procedures, challenges, and perceptions of existing tools supporting the short-form content production. This process was conducted using the qualitative analysis tool Atlas.ti⁷. During this phase, discussions regarding the format and granularity of the initial codes took place in team meetings. In cases of conflicting opinions, we revisited the transcripts to gain additional context from the source text and defined codes only when both researchers reached a consensus. After the team reached consensus on initial codes, two researchers independently coded to all transcripts and reconvened to resolve any disagreements on coded data through discussion. We then classified the codes into themes and sub-themes through iterative processes, defining the production process types based on the repeated mentions of their work processes and content characteristics by the participants. Subsequently, we defined challenge themes by grouping the challenges commonly mentioned within the creative processes. Lastly, the codes were labeled and reviewed for consistency within themes and sub-themes. Through this process, we identified three types of short-form video production processes and categorized the challenges into four creative process (planning, performance, editing, uploading & management), which we will describe in the following section.

4 PRODUCTION PROCESSES AND CHALLENGES OF SHORT-FORM VIDEO CREATORS

4.1 Short-Form Video Creators' Production Process

To create an AI tool that facilitates short-form video production, it is crucial to explore the production process that short-form video creators go through and identify the challenges they face. Previous work [71] broadly classified short-form videos into two types: (1) original short-form video contents where short-form video creators shoot their own videos, and (2) remix short-form video contents and where they use videos made by others. However, the results of our analysis based on their creative process [9] (planning, performance, editing, uploading & management), content characteristics and short-form video creator values, revealed that the two types could further be refined into three distinct types. Below, we describe the short-form video content types derived from our analysis and their respective production processes.

- Type 1 - **Participatory Entertainer** : This type involves short-term idea conceptualization and filming through active participation. Type 1 creators draw inspiration mainly from everyday life events. They actively present themselves as main subjects in the video such as dance challenges and role-playing (Figure 1-1).
- Type 2 - **Knowledge Sharer** : This type aims to convey new and insightful knowledge to viewers. To effectively communicate the information, Type 2 creators invest a substantial

amount of time and effort in learning new knowledge, collecting information and visual materials, followed by script writing for clarity, video editing, and post-production voice-over narration. Creators of this type primarily feature knowledge-sharing content, such as scientific discoveries and tutorials (Figure 1-2).

- Type 3 - **Highlight Capturer** : This type edits and reprocesses videos that have been edited and published before. This type mainly features highlight capturing content such as TV shows and movies (Figure 1-3).

These three types are organized into four major levels from the perspective of the production process: Planning, Performance, Editing, and Uploading & Management. There are notable differences within each process. In this paper, we have named the existing types of creators as **Participatory Entertainers**, **Knowledge Sharers**, and **Highlight Capturers** for easier recall.

4.1.1 Planning. In the initial stages of short-form content creation, all three creator types typically follow a common sequence of steps. They begin by gathering subscriber feedback for content inspiration, which is crucial for understanding their audience's preferences. Following this, they analyze trends to gain algorithm-favored insights, using these to identify topics. Lastly, they analyzed the factors behind the high-viewed short-form videos on the platform, using these as a reference for their content strategies.

However, each creator type also exhibits unique behaviors and priorities in the planning phase. Type 1 creators aim to capture and depict everyday life events as part of their planning process, creating engaging storytelling. For instance, they might document their morning routine, showcasing daily rituals and moments to connect with the audience. Type 2 creators distinguish themselves by learning new domain knowledge for content planning; they invest time in expanding their expertise and collecting visual materials to aid viewer understanding. They also prioritize checking the accuracy of the data, ensuring the information they present is reliable. Type 3 creators are notable in their ability to capture the most interesting and eye-catching parts in existing long videos like movies. In content planning, they excel at identifying and showcasing captivating segments, creating concise and engaging content that appeals to their audience.

4.1.2 Performance. When it comes to recording videos and narration, Type 1 and Type 2 need to consider various factors such as framing, lighting, and object placement to ensure their content remains visually appealing and understandable on vertical small screens, particularly in fast-paced scenarios. Type 1 creators pay special attention to framing their shots accurately, since they have to capture dynamic scenes while dancing or walking on the road. On the other hand, Type 2 creators usually streamline the recording process but place more emphasis on creating a well-structured script and narration to enhance the delivery of information, making it more understandable, engaging, and dynamic.

4.1.3 Editing. Short-form video editing normally considers video lengths of around 1 minute, optimization for mobile viewing, and portrait orientation. Type 1, 2, and 3 all adhere to these conditions. Unlike Type 2 and Type 3 creators, Type 1 creators typically

⁷<https://atlasti.com/>

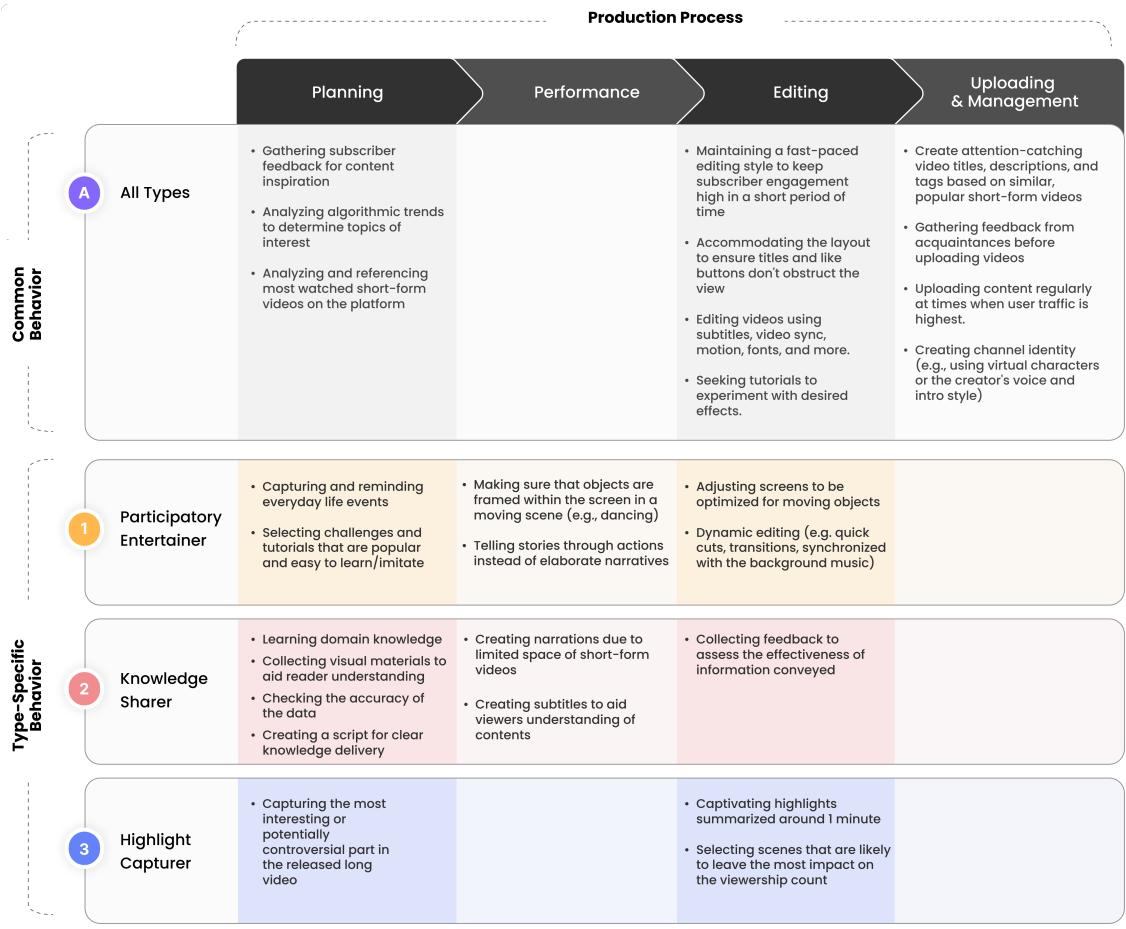


Figure 1: Short-form video creators' production process, showing common behaviors (A) and distinctive behaviors across three types: (1) Participatory Entertainers, (2) Knowledge Sharers, and (3) Highlight Capturers.

perform straightforward edits on their mobile phones immediately after recording. Type 1 creators optimize screens by focusing on dynamic editing for moving objects within the small mobile frame, employing techniques such as quick cuts, effects, and synchronization with music. Moreover, Type 2 creators prioritize knowledge delivery, assessing the effectiveness of conveying information to viewers and whether their videos bore viewers, prompting them to switch to another video. Type 3 Creators concentrate on condensing lengthy content into one-minute highlights and selecting scenes with the potential to have the greatest impact on viewership counts.

4.1.4 Uploading & Management. Across Type 1, 2, and 3, creators share a common step of crafting video titles and descriptions, followed by uploading to their chosen platforms. This process involves managing channel identity, reviews, and scheduling.

4.2 Challenges in Short-Form Video Production Process

4.2.1 Planning: Challenges in Initial Idea Generation.

Difficulties in Searching for and Archiving Inspirational Contents. Short-form video creators often face challenges in organizing their ideas and capturing inspiration due to their more frequent content production compared to long-form videos. However, they lack efficient tools for preserving inspiring moments. For instance, P12 mentioned archiving inspiring materials using platform's "Watch Later" feature but feels overwhelmed and finds it challenging to find resonating videos. P4 said that he relied on photos and memos but sometimes forgot to convert them into creative ideas. Short-form video creators typically assess follower response through metrics such as view counts, likes, and comments. However, P8 encountered difficulty in reading all feedback and translating it into new ideas.

Limited Understanding of Algorithm-Favored Contents. Short-form video creators struggle to keep up with rapidly changing trends driven by algorithms. In response, participants mentioned employing a strategy of benchmarking popular short-form videos or analyzing trends. However, they also express concerns about these strategies, recognizing potential side effects. P3 was concerned that indiscriminate benchmarking might have hindered channel growth and failed to enhance their sense of achievement.

P8 also pointed out that these strategies often resulted in a flood of similar videos, including topics, scripts, and editing styles. However, there is a lack of assistance in tailoring these trends to align with their individual short-form video content and the need to allow creators to maintain their unique identity while adapting to algorithmic trends.

Difficulties in Knowledge Delivery. Knowledge Sharers aim to convey new knowledge to viewers but face several challenges in each production iteration. Firstly, they invest a significant amount of time in seeking and comprehending knowledge from various domains while verifying its reliability. P2 expressed concerns about potential damage to channel credibility if misinformation is not properly filtered out. Additionally, Knowledge Sharers encounter difficulties when trying to convey content within the 1-3 minute time constraint. They often struggle with planning ideas, selecting key points, revising scripts repeatedly, speaking rapidly, and eliminating silences to address this issue. Lastly, Knowledge Sharers invest a substantial amount of time in gathering visual materials to enhance content understanding. Participants pointed out instances where existing visual materials may not fully align with the intended message, yet they are used because the creator lacks the skill to create and customize visuals according to the content.

4.2.2 Performance: Challenges in Shooting and Recording Short-Form Videos.

Challenges in Smartphone Filming and Framing. Participatory Entertainers use their smartphone cameras for the convenience and efficiency of shooting short-form content, swiftly editing it for online sharing. They prefer the rear camera over the front camera for recording high-resolution videos. However, participants who need to appear on screen face difficulties in framing their shots as they cannot see themselves, especially when filming alone without assistance from others. Moreover, the participants complained that they cannot know in advance if the shot subject is overlapped by 'like' or 'comment' buttons in the platform layout while filming.

Difficulties in Conveying Creator Identity through Vocal Narration. Knowledge Sharers invest time in clear narration, addressing factors like noise, wind, and interruption sounds. While some use Text-to-Speech (TTS) for efficiency, it falls short in conveying diverse emotions and content identity. P8 also pointed out that the unadjustable speech speed in TTS systems directly determines the length of dialogues in videos, consequently affecting the overall dynamic and duration of the content. P10 avoided TTS due to concerns about identical virtual voices negatively impacting brand and channel identity. Participants also struggled with auto-generated subtitles, unable to reproduce speech dynamics (e.g., dialect). Furthermore, participants emphasized providing subtitles and narration in multiple languages for international viewers due to platforms offering only automatic translation for subtitles, lacking voice translation capabilities.

4.2.3 Editing: Limitations of Short-Form Editing Tools and Lack of Quality Output.

Difficulties in Capturing and Sustaining Viewer Attention. On short-form video platforms, viewers can easily move to the next video with a finger swipe. As a result, short-form video creators face the challenges of capturing followers' attention right from the beginning and keeping them engaged until the end of the video.

Most of them try to spark viewers' curiosity within the first 3 seconds. To achieve this, participants employed various tactics. Some inserted engaging scenes or spoiled the ending in advance. P16 aimed to maintain 'liveliness' and a fast 'tempo' in their videos. Similarly, others frequently zoomed in and out of the video, cut out silences and irrelevant scenes, and stitched together only the essential scenes. However, this repetitive process has caused chronic fatigue among short-form video creators.

Inconvenience in Adjusting Layout. Short-form video creators frequently encounter challenges in adapting horizontal visual materials, especially videos, to fit a vertically rectangular format. They must either enlarge the content or add top and bottom margins to accommodate the format. Participants across the three types commonly expressed the exhaustion of dynamically adjusting subtitles to align with object movements in videos. P5 highlighted that although adjusting these details is time-consuming, they could not take it lightly as it significantly impacted video quality.

Limitations of Conventional Editing Tools: Lack of Interaction and Procedural Efficiency. Editing tools are claimed to enhance the overall workflow and reduce time spent in the production process through strong interaction. However, short-form video creators argue that the tools provide features in only a fixed, one-way manner, which often don't align well with their unique contexts and needs. For instance, while templates provide ease of use with their pre-designed layouts with built-in editing features, there is no way to effectively communicate their requirements back to the template, leading to a constrained editing experience. P13, who recently achieved quicker uploads thanks to pre-designed 'Instagram template', wished for a feature to insert her music into an auto-generated, personalized template. P6, a Highlight Capturer, also pointed out a similar issue with the existing auto-cut feature. It's designed to generate cuts based on changes in screen pixels, which showcases its limitation in providing only one-directional interaction. This often leads to the unintended cutting of conversations in videos.

4.2.4 Uploading & Management: Challenges in Uploading and Maintaining Contents.

Challenges of Anticipating User Reaction. The title, description, and tags of short-form videos also provide the initial impression to viewers, influencing whether they continue to watch the video. However, there is a need to assist short-form video creators in writing these elements in an intriguing way. P6 mentioned, "I am an old man, and it's hard for me to draw younger generations' attention. The titles I use are far from trendy words or memes." P14 and P15 also expressed frustration at not having colleagues to discuss matters with, including feedback, reaction, and unexpected ethical issues on their content before uploading their videos.

Lack of Reward System for Repetitive Tasks in Platform Upload Processes. Short-form video creators engage in the competitive uploading of trendy topics, and the popularity of their videos tends to fade quickly. They often struggle with maintaining consistent motivation for uploads. While P17 diligently strives for daily uploads, some, like P3, end up abandoning their channels due to the continuous pressure of uploads. Participants expressed demotivation due to the lack of assistance and motivation. This struggle

Workshop Session	Participant ID (Gender, Age)	Channel Category	Platform Worked	Subscribers (Main Platform)	Duration of Production	Commitment
W1	P1 (F, 30s)	Vlog, Parenting	YT, IG	74.9K	5 yrs	Full
	P2 (F, 20s)	Travel	IG	81K	2 yrs	Part
	P3 (M, 20s)	Humor, Fashion	YT, IG	5.81K	4 yrs	Full
W2	P1 (M, 20s)	Video Production	YT	17K	3 yrs	Part
	P2 (M, 30s)	Cooking	YT, IG	217K	1 yr	Part
	P3 (M, 30s)	Science	YT	213K	4 yrs	Full
W3	P1 (M, 30s)	Video Editing	YT, IG	85.7K	1 yr	Part
	P2 (M, 40s)	Movie, Drama	YT	173K	1 yr	Full
	P3 (F, 20s)	Counseling Talk	YT	1.1M	4 yrs	Full
W4	P1 (M, 20s)	Motion Graphic	YT	13.6K	3 yrs	Part
	P2 (M, 20s)	K-POP	TT	428K	2 yrs	Part
	P3 (F, 20s)	Vlog, Dance	IG, YT, TT	8K	4 yrs	Part

Table 2: Workshop participants' demographic details, channel information, platforms were working on (indicated by YT for YouTube Shorts, IG for Instagram Reels, and TT for TikTok), and their level of commitment.

is intensified for short-form video creators active on multiple platforms, requiring them to repetitively check and upload content on each platform.

5 STUDY 2: PARTICIPATORY DESIGN WORKSHOP

We conducted participatory design workshops involving short-form video creators to generate design solutions for the challenges identified in Section 4.2. In the workshop, we encouraged participants to contextualize these challenges into design solutions and provided materials to support their understanding and exploration of the potential for AI integration in short-form video production processes.

5.1 Recruitment

We adopted a process and recruitment criteria similar to those outlined in Study 1 (Section 3.2). We employed a pre-screen survey to gather data on participants' employment status (part-time or full-time), content categories, their usage of short-form video production tools, and the amount of time they dedicated to content production on a weekly basis. This comprehensive approach enabled us to encompass all three identified process types of short-form video production as detailed in Section 4.1.

5.2 Participants

In total, we recruited 12 short-form video creators in our participatory workshop. Since some of them were sensitive about exposing their personal information to other participants, we allowed them to share less sensitive personal details. Accordingly, the recruited participants covered a wide age range from their 20s to 40s, with 4 female and 8 male participants. Furthermore, the participants in this study represented a diverse range of backgrounds and engagement levels. Out of the participants, 5 were full-time content creators, and 7 engaged in content creation part-time. Among them, 5 exclusively operated YouTube channels, 1 on Instagram Reels, 1 managed a TikTok channel, and 5 operated on multiple platforms simultaneously. In terms of platform preference, 9 participants primarily focused on YouTube Shorts, 2 predominantly used Instagram Reels, and 1

primarily utilized TikTok. Detailed information about the participants can be found in Table 2. Participants were compensated with 50,000 KRW (approximately 38 USD) and all workshop sessions lasted up to 90 min remotely over Zoom. The facilitators refrained from giving tips, advice, or opinions to the participants and focused on facilitating the idea generation process according to the defined protocol.

5.3 Workshop Protocol

We conducted a participatory design workshop to collaboratively design solutions with the participants. The overall sequence of the workshop is presented in Figure 2. Each session involved three participants, and the workshop was facilitated by two researchers. One researcher specialized in AI research, and the other in the UX of AI research, with both having a strong background in Human-Computer Interaction (HCI). Our workshop structure drew inspiration from past participatory design studies [77] that focused on collaborative workshops for generating new solutions. Our participatory design workshop aimed to collaboratively explore solutions for enhancing short-form video creators' happiness by co-creating with AI. The workshop was conducted in 4 sessions. We organized sessions that specifically tailored to each group: one session for "Participatory Entertainers," another for "Knowledge Sharers," a third for "Highlight Capturers," and a fourth session that brought together creators of different types. This approach aimed to facilitate focused problem-solving for each type's challenges and solution designs. Each session comprised an orientation and three phases.

Introduction and Ice-Breaking. We introduced the purpose of our research, the objectives of the workshop, and workshop procedures. To foster rapport among participants and enhance mutual understanding, participants were asked to introduce themselves and provide information about their channels, content, and experiences in short-form video production.

Phase 1: Rating Challenges. Based on the challenges identified in previous interviews, we provided participants a survey including those challenges to participants. We first listed up the challenges and related scenarios (See Appendix A.1). We then formulated a

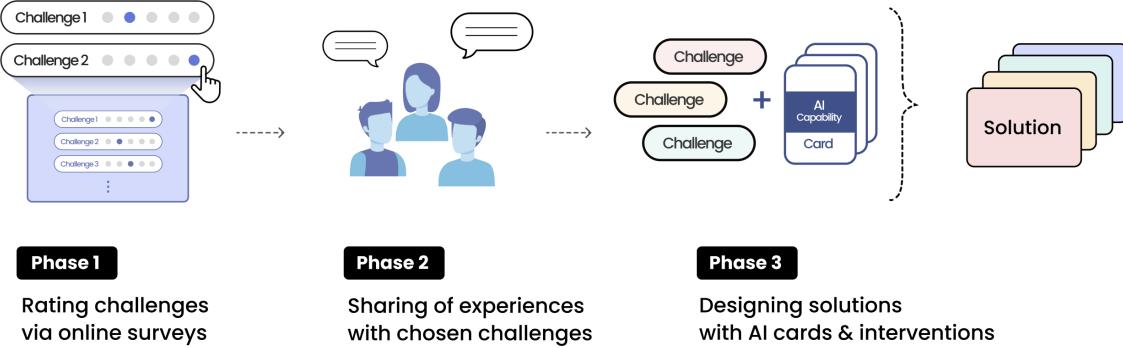


Figure 2: The participatory design workshop consisted of three phases: (1) participants rated the challenges they deemed most relevant and in need of solutions, (2) they shared their experiences related to the selected challenges, and (3) they collaboratively developed solutions using workshop materials to address these challenges.

5-point Likert-scale self-evaluation questionnaire to ask participants to select the most relevant challenges with a quantifiable analytical framework. As the final goal of the workshop was to design solutions for a minimum of three challenges, the researchers reviewed the results of the survey and selected three challenges to be addressed at first.

Phase 2: Sharing of Experiences with Challenges. We introduced the three selected challenges to participants one by one as missions. Participants shared their experiences and reflections on why they considered the specific challenge most in need of solutions. This sharing allowed participants to gain insights and inspiration for innovative solutions from others' experiences.

Phase 3: Designing Solutions to Solve the Challenges. To assist participants in generating solutions, we provided four workshop materials: (1) AI capability cards, (2) interventions for scaffolding AI concepts, (3) AI-infused editing tool, and (4) idea summary template. The intention was not to design the solutions, but rather to analyze the types of solutions and needs participants envisioned when combining solutions with AI.

AI Capability Cards: Inspired by [76], we presented AI capabilities in card format to enhance participants' understanding of AI technology and facilitate their brainstorming of specific solution types with the aid of AI (See Appendix A.2). In collaboration with AI researchers and UX designers, we designed AI capability cards in an easily comprehensible format for non-experts. Participants were asked to choose one or two cards and develop solution ideas based on them. While participants were encouraged to use the provided cards, they were also allowed to use any AI capability they believed to be effective. During the workshop, when participants had questions, the AI researcher promptly offered explanations about AI technology, accompanied by real-world use cases. The researcher was instructed to maintain a neutral tone, avoiding overly positive or negative nuances regarding the AI technology. Additionally, the researcher was encouraged to use plain language, opting for general terms (e.g., generating images) over technical terms (e.g., multimodal).

Interventions for Scaffolding AI Concepts: This materials were inspired by previous work [77]. We also presented three example

intervention types: (1) video editing tool features, (2) third-party video editing applications, and (3) collective information sharing platforms. These intervention types were chosen to align with the kinds of editing apps and social media platforms that creators typically incorporate into their work. Including three interventions aimed to facilitate idea generation among short-form video creators.

Example of AI-Infused Editing Tool: During our research, we identified the lack of functional AI short-form video production tools. To stimulate creative thinking, we developed a prototype tool (See Appendix A.3) showcasing AI-powered short-form video editing with deep learning capabilities to identify highlight scenes automatically. We surveyed 37 users, comparing professionally edited and AI-edited videos on a Likert scale of 5. It was found that there was no statistically significant difference between them in terms of engagement, enjoyment, clarity, and editing completeness. Notably, it was observed that AI-edited videos that generated by our tool received the same average score for engagement (4.09) as the professionally edited videos. In terms of editing completeness, AI-edited videos scored marginally higher on average compared to the professionally edited ones (4.8 vs. 4.69) (See Appendix A.4). Moreover, the qualitative feedback indicated that participants were unable to distinguish between AI-edited videos and professionally edited videos. They also acknowledged the user-friendly and efficiency-enhancing potential of AI-infused editing tools. This affirmed our prototype's effectiveness as a catalyst for participants to envision features enhancing their short-form video production experiences. Providing a pseudo-realistic sample artifact helped them better envision the necessity of integrating AI into their workflow, fostering imagination about the practical applications of ASVG.

Idea Summary Template: We included a writing section for participants to organize, clarify, and summarize their ideas. Inspired by [41], the Insight-Action-Outcome templates were adapted in our approach. These templates were designed to help participants reflect insights about the challenging situations, define the AI actions to address those challenges, and document the corresponding outcomes. During the discussion, there was an opportunity for participants to share and discuss each other's templates, further allowing them to develop these ideas. While participants were organizing their ideas,

facilitators played a crucial role in guiding participants to carefully consider their proposed solutions by asking questions pertaining to the benefits and potential side effects of integrating AI.

Through these four materials and facilitators' explanations, participants could comprehend AI technologies and conceive design solutions. Those facing challenges in visualizing the abstract concept of AI technology (e.g., text generation) as a collaborative tool (e.g., AI writing a short-form video contextualized script based on user's input) were able to envision interactive features by exploring examples of interventions and an AI-infused editing tool.

5.4 Data Analysis

The average duration of each workshop session was approximately 86.6 minutes (min: 82, max: 90). All design workshop sessions were recorded via Google Meet. The quoted statements in this paper were translated into English. The recordings were transcribed via Clova Note⁸, and workshop results on the Google Slides were documented in Google Docs. Similar to Study 1, we conducted thematic analysis [5]. Two researchers conducted open coding on a subset of transcripts, identifying recurring patterns and thematic sentences. Once the team reached a consensus on the initial code format and granularity through discussions, two researchers proceeded to apply the codes to all transcripts. We frequently reconvened to share and address any discrepancies in the coded data through discussion. After coding workshop data, we generated high-level themes by grouping the design solutions by the purpose of solution design. For example, an initial code might be "W3P2 suggested AI facilitating co-planning of scripts to instantly engage viewers." Then, the two researchers iteratively grouped the solutions into higher-level themes. For instance, one higher-level theme might be "Inspiring Creators to Captivate Viewers' Attention in 3 Seconds." During the process, we also defined challenge themes by grouping the challenges and concerned that participants mentioned. We then classified the codes into themes and sub-themes through iterative processes. Finally, we categorized themes to identify overlapping content across the four workshop sessions or unique insights within each context. This rigorous process helped distinguish universal solutions from those specifically effective for certain production processes. In our findings section, we present our categorized findings based on the short-form creative production process.

6 FINDINGS

We categorized our findings into two main sections: (1) design opportunities for AI-Powered Short-Form Video Generators (ASVG) that enhance creators' production processes and (2) concerns related to AI usage. In Section 6.1, we introduce creator-centered AI solutions, collaboratively developed with creators, aimed at enhancing the short-form production process across various phases: (1) planning, (2) performance, (3) editing, and (4) uploading & management. We present design solutions that are beneficial for all types of creators, as well as those specifically tailored to certain types. In Section 6.2, we address cases where AI intervention might adversely affect user perception.

⁸<https://clovanote.naver.com/>

6.1 Design Opportunities in Leveraging AI Short-Form Video Generators

In this section, we introduce design solutions that benefit all types of creators, as well as solutions that are specifically tailored to certain types. We have identified various design opportunities for ASVG during content production, including offering creative inspiration through trend analysis, simplifying the recording process, streamlining labor-intensive editing tasks, preserving the creator's distinct identity, and supporting secure channel management and motivation (See Figure 3).

6.1.1 Planning: Mitigating the Burden of Content Idea Formulation Process. Participants perceived the potential for collaboration with AI in the production process of short-form videos. They expected AI to play a role of a friendly colleague that can plan and develop ideas together. Following the findings from the workshop, we discovered four primary design opportunities for ASVG in content planning.

Inspiring Creators to Captivate Viewers' Attention in 3 Seconds. Participants aimed to capture viewers' attention quickly in short-form videos, adopting the "3-second strategy" to impress viewers within a limited time frame. They structured videos differently from traditional formats, placing the most interesting part at the beginning. Seeking AI assistance, participants expressed dissatisfaction with existing tools such as LLM (ChatGPT⁹, WRTN¹⁰), mentioning limitations in understanding short-form video uniqueness. W3P2 proposed an AI tool that is able to organize the content structure and facilitate co-planning of scripts with short-form video creators. For instance, once a user inputs their video into AI-infused tool, the algorithms can analyze the video content to identify the most attention-grabbing moments within the first 3 seconds. Subsequently, the AI can suggest an engaging video script to be placed at the beginning of the content. This tool aims to enhance video content planning, considering engaging openings, content delivery within time constraints, and retaining viewer interest.

Harmonizing Trends with Creators' Unique Identity. Participants aimed to adapt to the rapid turnover of short-form video trends when creating content. To achieve this, (W2P1, W4P1, W4P3) expressed a desire for AI to explore platform algorithm preferences, viewer interests, and statistical data to suggest content ideas aligned with current trends. Additionally, participants were concerned about losing their channel identity while incorporating trends without considering their unique characteristics. Thus, W2P2 and W2P3 proposed an algorithmic recommendation feature that blends their unique style with current trends, ensuring the preservation of creators' individual styles. For example, one participant mentioned, "Let's assume I am a psychology YouTuber, and AI understands trending topics, my channel, and recent posts. Then, it could suggest an idea based on a trending topic like 'The Glory' – a drama about school violence – to discuss the impact of such experiences on youth mental health, particularly focusing on depression and anxiety."

Fostering Trusted Creator-Viewer Relationships in Content Ideation. Participants expressed a desire for algorithms to analyze

⁹<https://chat.openai.com/>

¹⁰<https://wrtn.ai/>

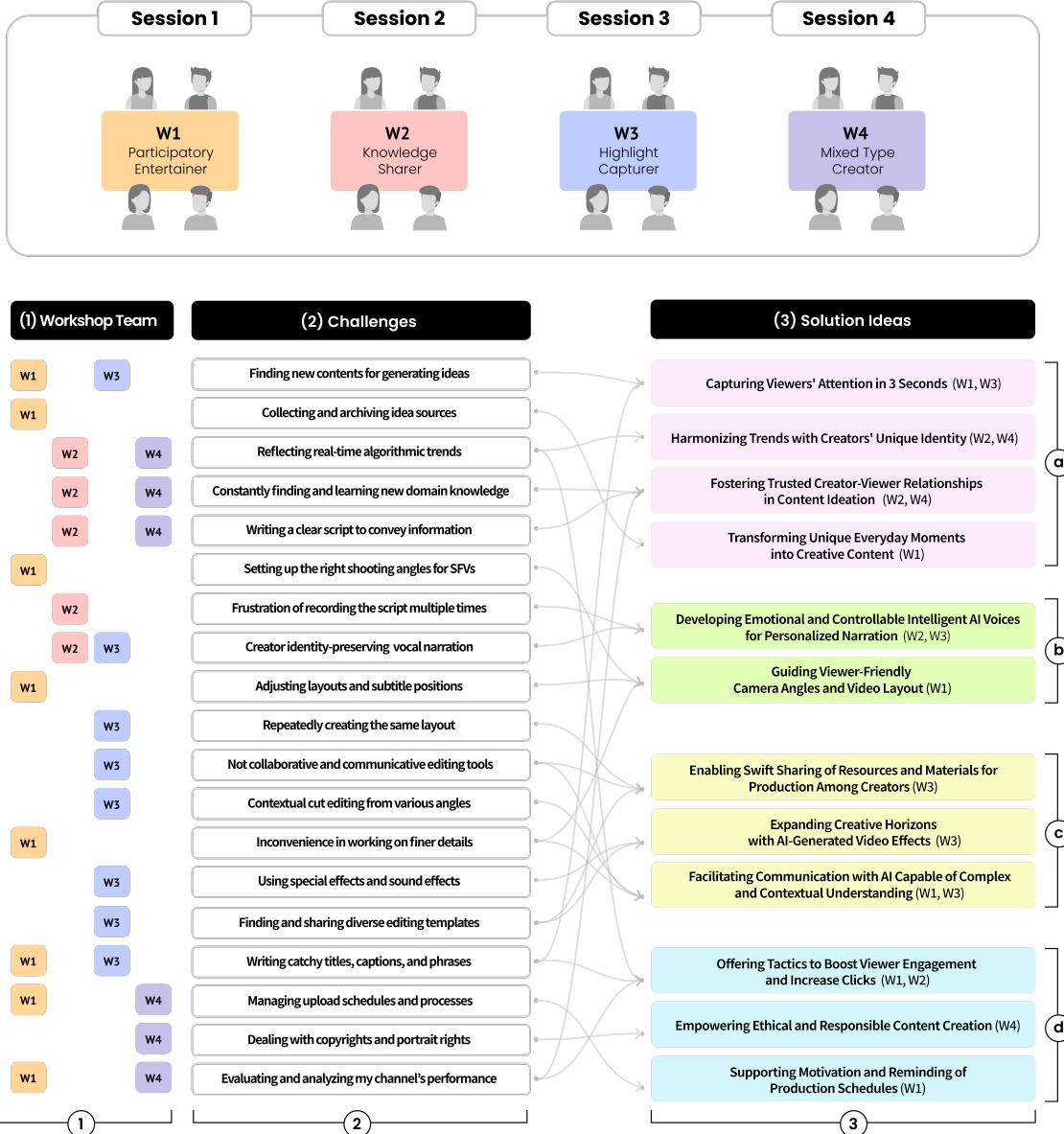


Figure 3: Thematic analysis map of findings: (1) team challenge selection, challenges chosen by workshop teams, (2) selected challenges, and (3) design suggestions from workshop sessions. Design suggestions were categorized based on the production process: (a) planning, (b) performance, (c) editing, and (d) uploading & management

viewer feedback, helping them discover new topic ideas and consider their channel identity. For instance, W4P1 and W4P2 proposed algorithms that automatically analyzed comments, offering insights about new content ideas and feedback on their performance. Participants believed such algorithms could enhance bond with viewers and potentially cultivate a loyal viewer community. Insights from the W2 session highlighted specific content planning needs for Knowledge Sharers, who had to continuously seek and summarize information within short video time frames. They envisioned AI as an intelligent collaborator that provides prompt responses and

clear explanations of information, incorporating features for fact-checking to prevent the dissemination of inaccurate information and false content.

Transforming Unique Everyday Moments into Creative Content. Participatory entertainers from the W1 and W4 sessions offer unique insights. They seize opportunities for potential content ideas from everyday events and reinterpret them to produce daily content. However, they struggle to extract these captures from their past records, such as memos or photos, and organize multiple captures into a single idea. Therefore, participants proposed an algorithm

that could automatically filter records based on specific keywords or situational descriptions and suggest new content ideas. For example, once short-form video creators connect their mobile phone albums to AI and ask AI to find a funny episode from their archive, AI can suggest videos with scenarios of high interest based on laughter sounds or laughing faces.

6.1.2 Performance: Recording Viewer-Friendly Short-Form Videos. The majority of participants expressed a strong desire to create viewer-friendly short-form videos, emphasizing the importance of layout and narration. Upon analyzing the workshop results and transcripts, two primary requirements for ASVG in the context of recording and performance were identified.

Providing Emotional and Controllable Intelligent AI Voices for Personalized Narration. In short-form video production, participants frequently use AI-generated voices for quick, clear narration. However, they identified shortcomings such as monotony and a lack of emotional range. To address this, W3P3 proposed an algorithm that creates intelligent voices capable of expressing diverse emotions and personas. Also, concerns arose about losing channel uniqueness by using generic AI voices. Addressing this, W3P1 suggested an AI short-form video narration tool that learns and mimics a creator's unique vocal style, thereby enhancing the channel's distinctiveness and maintaining its identity. Lastly, participants proposed including options for customizing specific aspects of the AI's voice, such as modifying the pitch and voice speed to introduce more dynamism.

Guiding Viewer-Friendly Camera Angles and Video Layout. Participants recognized the importance of a well-structured screen arrangement to enhance the visual appeal and overall aesthetics of short-form videos. However, they expressed a desire for AI to assess the suitability of these layouts during the recording process, considering the limited horizontal width of the screen and potential overlap with platform features like "like" buttons. For instance, while short-form video creators are recording the video, AI could identify problems in advance and suggest visually appealing layouts, enabling creators to effectively adjust their layouts when recording short-form video content.

6.1.3 Editing: Enhancing Editing Efficiency and Video Quality. The quality of short-form videos is significantly influenced by the editing process. In this context, participants suggested the use of AI to enhance the efficiency and engagement of the editing processes.

Enabling Swift Sharing of Resources and Materials for Production Among Creators. Although platforms like Instagram Reels provide templates, participants still found themselves limited and often had to create effects or music templates from scratch. To combat this, W3P1 proposed a feature for sharing self-made templates among creators with similar styles. Furthermore, there was a desire for algorithms that could learn a creator's editing style and suggest compatible templates from a shared archive. This approach aims to streamline content production, enhance channel exposure, and cultivate community among short-form video creators.

Expanding Creative Horizons with AI-Generated Video Effects. AI has the potential to enable creators to venture into unexplored territories and unleash richer creativity. Participants expressed their desire for AI that could understand simple descriptions and automatically generate video effects tailored to the frame of

short-form videos. For example, a participant from W3P3 mentioned, "If I ask AI to add sparkling effects in the middle of this video, then it can automatically create and add effects." They even envisioned adaptive algorithms that could suggest a variety of effects suitable for different genres or channels, enabling creators to apply them as needed. For example, the adaptive algorithm could analyze the content, recognize the genre, and suggest specific effects tailored for each. For a travel vlog, it might recommend vibrant and scenic overlays, while for a comedy sketch, it could suggest humorous animations or text overlays.

Facilitating Communication with AI Capable of Complex and Contextual Understanding. In Workshop 3, all participants unanimously expressed frustration with the passive nature of existing video editing tools and discomfort with the opaque characteristics of algorithms. Thus, they proposed AI-infused editing tools incorporating interactive prompts that enable users to interact with specific commands. Participants also envisioned AI going beyond mere scene identification, highlighting the limitations of the AI-editing tools provided during the workshop, which only cut scenes based on subjects. Instead, they wished for AI to understand the key content within each scene by analyzing emotions or situational contexts. For instance, W3P2 envisioned a scenario like "I can ask AI to find me a scene where character 461 in the Squid Game gets angry!".

6.1.4 Uploading & Management: Boosting Short-Form Video Creators' Motivation and Publishing Skills. Participants anticipate that AI will aid in the final stage of short-form video creation by facilitating the publication of their videos, understanding algorithm preferences, ethical rules, and streamlining the process of uploading to various social media platforms.

Offering Tactics to Boost Viewer Engagement and Increase Clicks. Short-form video creators highlighted the significance of video titles, thumbnails, and topic statements at the beginning of short-form videos, stressing their role in shaping the initial impression and capturing viewers' interest. They expressed a desire for AI to enhance their intuition by providing suggestions for concise yet curiosity-evoking titles, thumbnails, and topic statements. W1P2 proposed an AI that leverages current trends, memes, and statistical data to recommend directions, reducing their reliance on intuition and helping them produce more engaging content.

Empowering Ethical and Responsible Content Creation. Participants emphasized the importance of steering clear of copyright infringement, sharing experiences of removing content or even having to shut down channels due to such violations. They expressed a wish for AI to provide guidance on copyright law, helping short-form video creators enhance their understanding and compliance. Furthermore, they hoped for algorithmic features that could assist them in crafting 'ethical and responsible' content. Recognizing their responsibilities as content creators, they wanted AI to serve as a reminder to consider ethical values and to analyze content for potential issues such as false information, ethical concerns, and copyright infringements, offering suggestions for necessary modifications.

Supporting Motivation and Reminding of Production Schedules. What if algorithms managed production schedules for short-form video creators? Maintaining a consistent upload schedule is

crucial for building trust with viewers, but it can be challenging without motivation. Participants suggested AI could help creators by systematically organizing their uploads. For example, by analyzing a short-form video creator's uploading frequency, AI can identify when they take longer to produce or miss their timing. In such cases, AI can remind them of important deadlines by stating, "Shall we double-check the schedule?" or informing them, "You have two days left until the next upload schedule." Furthermore, AI can provide motivation by analyzing positive comments from viewers and offering compliments such as 'You are a dedicated creator'. This support can help creators stay enthusiastic and capture viewers' interest.

6.2 Challenges in Leveraging an AI Short-Form Video Generators

Despite the potential benefits of integrating AI into the creative process, participants also expressed concerns and challenges associated with this integration. Based on our research findings, we observed that short-form video creators expressed three primary concerns when collaborating with AI in the production of short-form videos: (1) Difficulty in retaining channel-specific identity, (2) Lack of empathetic communication between humans and AI, and (3) Challenges in supporting trust and transparency for the viewers. In this section, we describe three prominent negative perceptions associated with collaborating with algorithms.

6.2.1 Retaining Channel-Specific Identity. Short-form video creators have expressed valid concerns about maintaining the integrity of their channel's identity while integrating algorithms into their content creation process. For instance, W2P1, who operates a channel as a knowledge sharer, mentioned "Using AI tool like ChatGPT is incredibly convenient, but there are times when I prefer to keep some distance because I have the desire to come up with ideas on my own. I worry that relying too much on the ideas generated by AI could potentially disrupt the flow and character of my channel. So, it would be great if AI could respect my identity and help me maintain my creative habits." Additionally, W2P2 has voiced worries that "I am worried that my content might gradually become too similar to that of other creators." They struggle to convey their uniqueness to algorithms, fearing these systems may not preserve the distinct style that sets their channels apart. To address this, it is crucial to develop algorithms that comprehend and secure a creator's unique identity, fostering a productive partnership between content creators and AI tools.

6.2.2 Providing Empathetic Communication Between Human and AI. Participants expressed discomfort with algorithms struggling to grasp intended emotions and context. For instance, W3P3 shared that subscribers preferred human voices with emotional inflections over emotionless TTS voices. Creators highly value emotional connections with their audience, finding it challenging when algorithms fall short in this aspect. As AI technology advances, addressing these concerns by developing algorithms capable of understanding and replicating human emotions and context becomes essential. Additionally, participants found existing short-form video creation tools unfriendly and challenging to interact with. W1P1

highlighted the absence of features for explaining and requesting modifications to specific parts of the content. Furthermore, artistic expressions, like emotional conveyance or humor, were often misunderstood by AI, necessitating the integration of human senses into content creation. To establish trust as collaborators, ASVG should provide easy user interaction and enhanced editability.

6.2.3 Supporting Trust and Transparency. Participants faced challenges related to the transparency of AI decision-making processes, making it difficult for them to understand how AI reaches conclusions. These difficulties heightened concerns about trust. For instance, W2P3 used ChatGPT to generate a script for new science technology but was disappointed due to its inaccuracy and lack of timeliness. Concerns were also expressed about the ambiguity surrounding AI's data sources and operational methods, making it hard to trust the system. Participants in W2 were aware of the possibility of AI providing incorrect or biased information and were uncomfortable with the idea that their subscribers might be exposed to such content. Therefore, they independently selected news from reputable sources, entrusting AI only with specific tasks like script-writing. This underscores the importance of addressing transparency and fairness issues in AI to foster trust in collaborations.

7 DISCUSSION

Based on our research findings, we observed that short-form video creators expressed three primary concerns when collaborating with AI in the production of short-form videos: (1) Difficulty in retaining channel-specific identity, (2) Lack of empathetic communication between humans and AI, and (3) Challenges in supporting trust and transparency for the viewers. In Section 7.1, we explore how future ASVG should be designed to address these concerns through various creative solutions. Furthermore, in Section 7.2, we draw lessons from the participatory workshop experiences, particularly in the design of AI-Creator collaboration tools.

7.1 Balancing Needs, Trust, and Acceptance in Collaborative Short-Form Video Content Creation

Understanding the creative process, motivation, and challenges of short-form video creators is crucial for developing effective support interventions. Through interviews and workshops, we identified four key outcomes: (1) the short-form video production type and each process, (2) challenges at different stages, (3) solutions to address these challenges, and (4) perceptions and concerns regarding AI collaboration. This findings provide insights for designers and developers to tailor tools to each type of production process, considering unique needs. Furthermore, we discovered the essential role of balancing trust, acceptance, and needs for a more productive collaboration between creators and AI. In our study, we found that complete reliance on AI may not always be positively perceived by creators in terms of acceptance. Short-form video creators tried to balance their collaboration with AI due to concerns about potential long-term impacts on their identity and creativity. They were concerned that excessive dependence on AI-generated ideas might lead them to lose the distinctiveness of their channels. They sought AI

tools capable of suggesting tailored content and results that align with their channel identities and unique styles. To maintain their independence and creative spark, future research in AI-Powered Short Video Generators (ASVG) should aim to provide short-form content creators with diverse interaction modes for Human-AI collaboration. For example, we could offer two distinct interaction styles: human-generated and AI-validated, and AI-generated and human-validated. In the first scenario, creators could initially produce their own content, such as videos or scripts, which AI would then refine. Conversely, AI could first generate content, which creators would subsequently edit. Importantly, we have also discovered that trust in AI plays a crucial role in this collaborative process. Short-form video creators expressed concerns about transparency and reliability of AI because biased information generated by AI could confine creators to information bubbles and lead viewers to develop inappropriate fixed beliefs. Therefore, it is crucial for developers and designers to create AI systems with a focus on transparency and ethical data use. This means building AI tools that not only generate unbiased, accurate information but also clearly explain their decision-making processes. By considering creators' needs, acceptance, and trust, we believe that technology developed to facilitate AI-creator collaboration will avoid potential problems and foster a harmonious partnership between creators and AI, which can be applied to various creative fields requiring such collaboration.

7.2 Takeaways from Co-Designing AI-Creator Collaboration Tool with Short-Form Video Creators

We engaged in co-designing AI concepts with short-form video creators, employing participatory design research—an extensively utilized approach in the HCI community [10, 33]. Our goal of our participatory workshops was to involve short-form content creators, leveraging their experiences for AI solutions tailored to their needs. Reflecting on the process, we identified the necessity for three methodological supports to aid participants in the formulation of AI concepts: 1) stimuli for AI understanding, 2) materials for creator-contextualized challenges, and 3) a facilitator's guide for considering AI advantages and concerns. Firstly, to aid participants unfamiliar with AI capabilities, we supplied stimuli like AI capability cards, illustrative AI concept examples, and summary notes. AI researchers and facilitators addressed participants' queries about AI capabilities using layman terms, such as "conversational abilities," rather than technical jargon like "large language model." Secondly, to ensure participants considered AI alongside the challenges of content creation within limited ideation time, we proposed presenting them with pre-defined challenges and missions. By utilizing a survey inspired by value-centered design methods and aligned with creators' values, participants swiftly empathized with the problem space, leading to the generation of relevant solutions. Thirdly, facilitators offered clear guidelines to encourage participants to weigh both the advantages and concerns of AI concepts. By prompting participants to contemplate how these solutions would collaborate with users, interact in various scenarios, and impact their lives, they were able to consider both the positive and negative impacts of AI collaboration. In conclusion, we found that structured support proved essential in guiding domain experts to develop concrete

solutions through participatory design workshops. From a broader perspective, the insights gained from our study are anticipated to have applicability in participatory design (PD) with various creators such as writers and composers. This can facilitate HCI researchers in understanding their production processes, challenges, and potential solutions, contributing to the development of creator-AI collaboration tools that provide a more user-friendly environment to support their creative endeavors.

8 LIMITATIONS AND FUTURE WORK

We acknowledge several research limitations. First, our sample, focusing on Korean short-form video creators, may lack representativeness for the general population, potentially limiting result generalizability. The characteristics and habits of Korean creators may differ from those in other regions. Future research aims to explore short-form video creators from diverse cultural backgrounds for more comprehensive solutions.

Second, the research participants, mainly creators in their 20s to 40s, possessed basic knowledge of systems and AI concepts. We provided materials and guides to enhance their understanding of AI capabilities for envisioning AI-infused collaboration tools, resulting in successful incorporation of this understanding and generation of design solutions. However, these materials may pose challenges for creators with limited technological literacy. Future studies, particularly within the HCI community, can explore designing tools and stimuli for AI workshops targeting users with little to no prior knowledge of using tools or AI.

Third, our results capture the current state of the rapidly evolving short-form content ecosystem, but may not fully address future processes, challenges, and opportunities for new short-form video creators. This study is the first empirical exploration of short-form content creation and its challenges based on current creator characteristics. However, the dynamic short-form content landscape consistently introduces new creator types and trends, prompting ongoing research to adapt to this environment. Future research aims to illuminate the processes of emerging creators and explore evolving challenges and opportunities in the short-form content field.

Fourth, this research centered around short-form video creators exclusively from three prominent platforms, which may have introduced bias and restricted insights into the challenges encountered by creators on alternative platforms. Furthermore, the recruitment of participants was based on platform user statistics [38, 70], potentially skewing the sample towards a larger number of YouTube users. While efforts were made to encompass diverse platform experiences, it is possible that certain findings may not be universally applicable. Therefore, future research should adopt a more comprehensive approach and validate its findings across various platforms.

Finally, the study aimed to comprehend the shared challenges faced by short-form video creators on the three platforms and to identify design opportunities. However, they may adopt distinct strategies influenced by platform characteristics, such as content length rules. Subsequent research will delve into platform-specific traits and challenges, offering customized guidelines for short-form video creators on each platform.

9 CONCLUSION

Our research focused on designing AI-Powered Short-Form Video Generators (ASVG) to enhance creators' production and collaboration experiences with AI. Our investigation began by examining the challenges faced by short-form content creators during the production process. We conducted semi-structured interviews and categorized short-form video production into three distinct creator types, identifying the common challenges as well as those unique to each creator type. To address these challenges, we facilitated participatory design workshops involving short-form video creators. The primary goal of these workshops was to generate potential design solutions tailored to the needs of these creators using AI-powered short-form video generators. During these workshops, we uncovered several benefits that ASVG could offer, including providing creative inspiration based on trends, preserving the creator's unique identity, simplifying the recording process, streamlining labor-intensive editing tasks, and supporting secure channel management and motivation. However, our observations revealed potential conflict between creators' needs and concerns, which could impact their trust and acceptance of AI. In conclusion, our findings emphasize the importance of platforms considering the intricate interplay between creator needs, acceptance, and trust when developing ASVG to create a more creator-centric environment which enables creators to maintain their unique identities and sustain their creative activities. In summary, our research contributes to a growing understanding of how AI support creators' work experiences and how creators envision technology design aligned with their values and production processes.

ACKNOWLEDGMENTS

We thank our reviewers and study participants for their time and feedback. We also would like to express sincere gratitude to Professors Jodi Forlizzi and Hong Shen for their invaluable support and guidance throughout this study. Finally, we extend our thanks to the members of the Snowman who contributed to the development of the AI-infused editing tool.

REFERENCES

- [1] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N Bennett, Kori Inkpen, et al. 2019. Guidelines for human-AI interaction. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–13.
- [2] Hemant K Bhargava. 2022. The creator economy: Managing ecosystem supply, revenue sharing, and platform design. *Management Science* 68, 7 (2022), 5233–5251.
- [3] Xiang Bi and Cunchen Tang. 2020. Research on the motives affecting the behavior of short video's creators. *IEEE Access* 8 (2020), 188415–188428.
- [4] Sid Black, Stella Biderman, Eric Hallahan, Quentin Anthony, Leo Gao, Laurence Golding, Horace He, Connor Leahy, Kyle McDonell, Jason Phang, et al. 2022. Gpt-neox-20b: An open-source autoregressive language model. *arXiv preprint arXiv:2204.06745* (2022).
- [5] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.
- [6] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems* 33 (2020), 1877–1901.
- [7] Valerie A Canady. 2022. Psychiatrists collaborate with YouTube to develop MH content. *Mental Health Weekly* 32, 20 (2022), 7–7.
- [8] Kabdo Choi, Hyungyu Shin, Meng Xia, and Juho Kim. 2022. Algosolve: Supporting subgoal learning in algorithmic problem-solving with learnersourced microtasks. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–16.
- [9] Yoonseo Choi, Eun Jeong Kang, Min Kyung Lee, and Juho Kim. 2023. Creator-friendly Algorithms: Behaviors, Challenges, and Design Opportunities in Algorithmic Platforms. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–22.
- [10] Fernando Delgado, Solon Barocas, and Karen Levy. 2022. An uncommon task: Participatory design in legal AI. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW1 (2022), 1–23.
- [11] Accenture The Dock. 2020. *Designed Intelligence*. Retrieved August 10, 2023 from <https://youtu.be/pkFOcgD8AnA?t=551>
- [12] Xuebing Dong, Hong Liu, Nannan Xi, Junyun Liao, and Zhi Yang. 2023. Short video marketing: what, when and how short-branded videos facilitate consumer engagement. *Internet Research* (2023).
- [13] Zhiwen DONG and Tian Xie. 2022. Why People Love Short-Form Videos? The Motivations for Using Tiktok and Implications for Well-Being. *The Motivations for Using Tiktok and Implications for Well-Being* (2022).
- [14] Graham Dove and Anne-Laure Fayard. 2020. Monsters, metaphors, and machine learning. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–17.
- [15] Brandon Doyle. 2023. TikTok Statistics – Updated Sep 2023.
- [16] Xin Du, Toni Liechty, Carla A Santos, and Jeongeun Park. 2022. 'I want to record and share my wonderful journey': Chinese Millennials' production and sharing of short-form travel videos on TikTok or Douyin. *Current Issues in Tourism* 25, 21 (2022), 3412–3424.
- [17] Andrea Gasparini, Ahmed Abdi Mohammed, and Gabriele Oropallo. 2018. Service design for artificial intelligence. In *Linköping Electronic Conference Proceedings*. Linköping University Electronic Press, 1064–1073.
- [18] Xudong Gong, Qinlin Feng, Yuan Zhang, Jiangling Qin, Weijie Ding, Biao Li, Peng Jiang, and Kun Gai. 2022. Real-time Short Video Recommendation on Mobile Devices. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*. 3103–3112.
- [19] Katerina Gorkovenko, Daniel J Burnett, James K Thorp, Daniel Richards, and Dave Murray-Rust. 2020. Exploring the future of data-driven product design. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [20] Christina Harrington, Sheena Erete, and Anne Marie Piper. 2019. Deconstructing community-based collaborative design: Towards more equitable participatory design engagements. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–25.
- [21] Jianchao He, Miao Hu, Yipeng Zhou, and Di Wu. 2020. LiveClip: towards intelligent mobile short-form video streaming with deep reinforcement learning. In *Proceedings of the 30th ACM Workshop on Network and Operating Systems Support for Digital Audio and Video*. 54–59.
- [22] Matthew K Hong, Adam Fourney, Derek DeBellis, and Saleema Amershi. 2021. Planning for natural language failures with the ai playbook. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–11.
- [23] Mansoor Iqbal. 2023. *TikTok Revenue and Usage Statistics* (2023). Retrieved Oct 31, 2023 from <https://www.businessofapps.com/data/tik-tok-statistics/#:~:text=TikTok%20reached%202015%20billion%20users,increase%20on%20the%20previous%20year>.
- [24] Xiaoneng Jin, Mark Evans, Hua Dong, and Anqi Yao. 2021. Design heuristics for artificial intelligence: inspirational design stimuli for supporting UX designers in generating AI-powered ideas. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–8.
- [25] Jeongyeon Kim, Daewon Choi, Nicole Lee, Matt Beane, and Juho Kim. 2023. Surch: Enabling Structural Search and Comparison for Surgical Videos. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–17.
- [26] Tae Soo Kim, DaEun Choi, Yoonseo Choi, and Juho Kim. 2022. Stylette: Styling the web with natural language. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–17.
- [27] Tae Soo Kim, Arghya Sarkar, Yoonjoo Lee, Minsuk Chang, and Juho Kim. 2023. LMCanvas: Object-Oriented Interaction to Personalize Large Language Model-Powered Writing Environments. *arXiv preprint arXiv:2303.15125* (2023).
- [28] Hyung-Kwon Ko, Subin An, Gwanmo Park, Seung Kwon Kim, Daesik Kim, Bokyung Kim, Jaemin Jo, and Jinwook Seo. 2022. We-toon: A Communication Support System between Writers and Artists in Collaborative Webtoon Sketch Revision. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. 1–14.
- [29] Hyung-Kwon Ko, Gwanmo Park, Hyeon Jeon, Jaemin Jo, Juho Kim, and Jinwook Seo. 2023. Large-scale text-to-image generation models for visual artists' creative works. In *Proceedings of the 28th International Conference on Intelligent User Interfaces*. 919–933.
- [30] Rafal Kocielnik, Saleema Amershi, and Paul N Bennett. 2019. Will you accept an imperfect ai? exploring designs for adjusting end-user expectations of ai systems. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [31] Peter Kun, Ingrid Mulder, and Gerd Kortuem. 2018. Design enquiry through data: appropriating a data science workflow for the design process. In *Proceedings of the 32nd International BCS Human Computer Interaction Conference* 32. 1–12.

- [32] Mackenzie Leake, Abe Davis, Anh Truong, and Maneesh Agrawala. 2017. Computational video editing for dialogue-driven scenes. *ACM Trans. Graph.* 36, 4 (2017), 130–1.
- [33] Min Kyung Lee, Daniel Kusbit, Anson Kahng, Ji Tae Kim, Xinran Yuan, Alissa Chan, Daniel See, Ritesh Noothigattu, Siheon Lee, Alexandros Psomas, and Ariel D. Procaccia. 2019. WeBuildAI: Participatory Framework for Algorithmic Governance. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 181 (nov 2019), 35 pages. <https://doi.org/10.1145/3359283>
- [34] Fan Li and Yuan Lu. 2021. Engaging end users in an ai-enabled smart service design—the application of the smart service blueprint scape (SSBS) framework. *Proceedings of the Design Society* 1 (2021), 1363–1372.
- [35] Tianyi Li, Mihaela Vorvoreanu, Derek DeBellis, and Saleema Amershi. 2022. Assessing Human-AI Interaction Early through Factorial Surveys: A Study on the Guidelines for Human-AI Interaction. *ACM Transactions on Computer-Human Interaction* (2022).
- [36] Q Vera Liao, Hariharan Subramonyam, Jennifer Wang, and Jennifer Wortman Vaughan. 2023. Designerly understanding: Information needs for model transparency to support design ideation for AI-powered user experience. In *Proceedings of the 2023 CHI conference on human factors in computing systems*. 1–21.
- [37] Dong Kyun Lim. 2021. A Study on the Experiential Response of Short-Form Video Users. *The International Journal of Advanced Smart Convergence* 10, 4 (2021), 273–277.
- [38] Seongho Lim. 2023. *7 out of 10 people watch short-form videos. Usage rate increases by more than 10 percent in one year.* Maeil Business Newspaper. Retrieved March 1, 2023 from <https://stock.mk.co.kr/news/view/52351>
- [39] Qin Lin, Nuo Pang, and Zhiying Hong. 2021. Automated multi-modal video editing for ads video. In *Proceedings of the 29th ACM International Conference on Multimedia*. 4823–4827.
- [40] Daria Loi, Thomas Lodato, Christine T Wolf, Raphael Arar, and Jeanette Blomberg. 2018. PD manifesto for AI futures. In *Proceedings of the 15th Participatory Design Conference: Short Papers, Situated Actions, Workshops and Tutorial-Volume 2*. 1–4.
- [41] Eoin Ó Loideáin. 2019. *Towards algorithmic equality.* Retrieved August 10, 2023 from <https://medium.com/design-voices/towards-algorithmic-equality-118e6ba10b74>
- [42] John Maeda. 2022. . Retrieved August 10, 2023 from <https://twitter.com/johnmaeda/status/151397684735328272>
- [43] Nirav Malsattar, Tomo Kihara, and Elisa Giaccardi. 2019. Designing and Prototyping from the Perspective of AI in the Wild. In *Proceedings of the 2019 on Designing Interactive Systems Conference*. 1083–1088.
- [44] Omar Marzouk, Joni Salminen, Pengyi Zhang, and Bernard J Jansen. 2022. Which message? Which channel? Which customer? Exploring response rates in multi-channel marketing using short-form advertising. *Data and Information Management* 6, 1 (2022), 100008.
- [45] Ivan Mehta. 2023. *Google says 2 billion logged in monthly users are watching YouTube Shorts.* Retrieved July 26, 2023 from <https://techcrunch.com/2023/07/25/google-says-2-billion-logged-in-monthly-users-are-watching-youtube-shorts/>
- [46] Devadas Menon. 2022. Factors influencing Instagram Reels usage behaviours: An examination of motives, contextual age and narcissism. *Telematics and Informatics Reports* 5 (2022), 100007.
- [47] Steven Moore, Q Vera Liao, and Hariharan Subramonyam. 2023. fAIlureNotes: Supporting Designers in Understanding the Limits of AI Models for Computer Vision Tasks. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–19.
- [48] Iohanna Nicenboim, Elisa Giaccardi, Marie Louise Juul Søndergaard, Anuradha Venugopal Reddy, Yolande Strengers, James Pierce, and Johan Redström. 2020. More-than-human design and AI: in conversation with agents. In *Companion publication of the 2020 ACM designing interactive systems conference*. 397–400.
- [49] OpenSurvey. 2023. *Comparing the Trends of Reels, Shorts, and TikTok - From Content Themes to Length.* <https://blog.opensurvey.co.kr/article/socialmedia-2023-2/>
- [50] Eunji Park, Yugyeong Jung, Inyeop Kim, and Uichin Lee. 2023. Charlie and the Semi-Automated Factory: Data-Driven Operator Behavior and Performance Modeling for Human-Machine Collaborative Systems. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–16.
- [51] Sebastiaan Theodorus Michaël Peek, Eveline JM Wouters, Katrien G Luijkx, and Hubertus JM Vrijhoeve. 2016. What it takes to successfully implement technology for aging in place: focus groups with stakeholders. *Journal of medical Internet research* 18, 5 (2016), e5253.
- [52] Jennifer Pierre, Roderic Crooks, Morgan Currie, Britt Paris, and Irene Pasquetto. 2021. Getting Ourselves Together: Data-centered participatory design research & epistemic burden. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–11.
- [53] Victor Potrel. 2022. *Five Insights Into The Popularity Of Short-Form Video Content.* <https://www.forbes.com/sites/forbescommunicationscouncil/2022/09/06/five-insights-into-the-popularity-of-short-form-video-content/?sh=294ac2e579e5>
- [54] Aditya Ramesh, Prafulla Dhariwal, Alex Nichol, Casey Chu, and Mark Chen. 2022. Hierarchical text-conditional image generation with clip latents. *arXiv preprint arXiv:2204.06125* 1, 2 (2022), 3.
- [55] Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen, and Ilya Sutskever. 2021. Zero-shot text-to-image generation. In *International Conference on Machine Learning*. PMLR, 8821–8831.
- [56] Isabella Scandurra and Marie Sjölander. 2013. Participatory design with seniors: Design of future services and iterative refinements of interactive eHealth services for old citizens. *Medicine 2.0* 2, 2 (2013).
- [57] Ben Shneiderman. 2020. Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human-Computer Interaction* 36, 6 (2020), 495–504.
- [58] Than Htut Soe. 2021. AI video editing tools. What editors want and how far is AI from delivering? *arXiv preprint arXiv:2109.07809* (2021).
- [59] Than Htut Soe. 2021. Automation in Video Editing: Assisted Workflows in Video Editing. In *AutomationXP@ CHI*.
- [60] Wondwesen Tafesse and Mumin Dayan. 2023. Content creators' participation in the creator economy: Examining the effect of creators' content sharing frequency on user engagement behavior on digital platforms. *Journal of Retailing and Consumer Services* 73 (2023), 103357.
- [61] Romal Thoppilan, Daniel De Freitas, Jamie Hall, Noam Shazeer, Apoorv Kulshreshtha, Heng-Tze Cheng, Alicia Jin, Taylor Bos, Leslie Baker, Yu Du, et al. 2022. Lamda: Language models for dialog applications. *arXiv preprint arXiv:2201.08239* (2022).
- [62] Anh Truong, Floraine Berthouzoz, Wilmot Li, and Maneesh Agrawala. 2016. Quickcut: An interactive tool for editing narrated video. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. 497–507.
- [63] Baptist Vandersmissen, Frédéric Godin, Abhimeshwar Tomar, Wesley De Neve, and Rik Van de Walle. 2014. The rise of mobile and social short-form video: an in-depth measurement study of vine. In *Workshop on social multimedia and storytelling*. Vol. 1198. Citeseer, 1–10.
- [64] Arantxa Vizcaíno-Verdú and Crystal Abidin. 2023. TeachTok: Teachers of TikTok, micro-celebrification, and fun learning communities. *Teaching and Teacher Education* 123 (2023), 103978.
- [65] Greg Walsh and Eric Wronsky. 2019. AI+ co-design: developing a novel computer-supported approach to inclusive design. In *Conference companion publication of the 2019 on computer supported cooperative work and social computing*. 408–412.
- [66] Changyu Wang, Jhi Yan, Lijing Huang, and Ningyue Cao. 2023. Helping middle-aged and elderly short-video creators attract followers: a mixed-methods study on Douyin users. *Information Technology & People* (2023).
- [67] Hongcheng Wang, Ning Xu, Ramesh Raskar, and Narendra Ahuja. 2007. Videoshop: A new framework for spatio-temporal video editing in gradient domain. *Graphical models* 69, 1 (2007), 57–70.
- [68] Yanheng Wei, Lianghua Huang, Yanhai Zhang, Yun Zheng, and Pan Pan. 2022. An Intelligent Advertisement Short Video Production System via Multi-Modal Retrieval. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 3368–3372.
- [69] Allison Woodruff, Sarah E Fox, Steven Rousso-Schindler, and Jeffrey Warshaw. 2018. A qualitative exploration of perceptions of algorithmic fairness. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–14.
- [70] Matthew Woodward. 2023. *TikTok User Statistics.* Search Logistics. Retrieved September 20, 2023 from <https://www.searchlogistics.com/learn/statistics/tiktok-user-statistics/>
- [71] Jihoon Yang. 2018. *Attempts to Improve the Quality of Short Video Content in China and the U.S. Media Issue & Trends Issue No. 9.* KCA Korea communications Agency. 51–52 pages. <http://www.moazine.com/article/detail.asp?articleid=1422295> Public organization under the Ministry of Science and ICT of Korea.
- [72] Qian Yang, Alex Scuito, John Zimmerman, Jodi Forlizzi, and Aaron Steinfeld. 2018. Investigating how experienced UX designers effectively work with machine learning. In *Proceedings of the 2018 designing interactive systems conference*. 585–596.
- [73] Qian Yang, Aaron Steinfeld, Carolyn Rosé, and John Zimmerman. 2020. 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*. 1–13.
- [74] Saelene Yang, Changyoon Lee, Hijung Valentina Shin, and Juho Kim. 2020. Snapstream: Snapshot-based interaction in live streaming for visual art. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–12.
- [75] Nur Yildirim, Alex Kass, Teresa Tung, Connor Upton, Donnacha Costello, Robert Giusti, Sinem Lacin, Sara Lovic, James M O'Neill, Rudi O'Reilly Meehan, et al. 2022. How Experienced Designers of Enterprise Applications Engage AI as a Design Material. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [76] Nur Yildirim, Changhoon Oh, Deniz Sayar, Kayla Brand, Supritha Challa, Violet Turri, Nina Crosby Walton, Anna Elise Wong, Jodi Forlizzi, James McCann, et al. 2023. Creating Design Resources to Scaffold the Ideation of AI Concepts. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. 2326–2346.

- [77] Angie Zhang, Alexander Boltz, Chun Wei Wang, and Min Kyung Lee. 2022. Algorithmic management reimagined for workers and by workers: Centering worker well-being in gig work. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–20.
- [78] Jamie Zigelbaum, Michael S Horn, Orit Shaer, and Robert JK Jacob. 2007. The tangible video editor: collaborative video editing with active tokens. In *Proceedings of the 1st international Conference on Tangible and Embedded interaction*. 43–46.

A APPENDIX

A.1 Challenge List

Production Process	Challenges	Scenarios
Planning	Finding new contents for generating ideas	"I am struggling to find videos that resonate with me, as well as to determine which segments to use for creating compelling content."
	Collecting and archiving idea sources	"I usually collect and archive materials that are similar to my short-form videos or good for reference. But it is a hassle to find each one, and it is hard to pull it back up later."
	Reflecting real-time algorithmic trends	"To get favored by the algorithm, I have to create content that a lot of people are interested in. But it is quite challenging to pinpoint what trends that resonate with people."
	Constantly finding and learning new domain knowledge	"Every time I make content to share knowledge, I have to learn new background information. It is not easy to quickly understand terms and concepts. I am also worried about disappointing my subscribers if I use incorrect information."
Performance	Writing a clear script to convey information	"To clearly convey information and intention to people, a good script is necessary. But it is hard to figure out how to make content that viewers can easily understand and trust."
	Setting up the right shooting angles for SFVs	"I want a main subject or object to be in the middle of mobile vertical frame. So I pay attention to that while filming. But things still fall out of the frame, or get overlapped by layout such as 'like', 'comments' buttons. it really kills the mood."
	Frustration of recording the script multiple times	"I record the narration at least 10 times to get a natural sound. It takes a long time to distinguish the most attractive sound considering noises like beeping and wind. I also struggle to get the recording to sound with the atmosphere or emotion I want."
	Creator identity-preserving vocal narration	"I use a virtual voice actor on my channel, but I am not the only one using it. There are elements that make my channel unique to viewers, maybe like a brand? But if everyone starts using the same voice, it might dilute my channel's identity"

Editing	Adjusting layouts and subtitle positions	"Adjusting horizontal videos to fit a vertical frame is a hassle. And even when I do, sometimes the 'like' and 'comment' buttons cover the content. The video also gets cut off depending on the phone model."
	Repeatedly creating the same layout	"It is hard to systemize things like the caption, motions, transitions that I use all the time. It is challenging to create the same template over and over."
	Not collaborative and communicative editing tools	"I want to tell the tool specifically what I am looking for, but I am not sure how to convey that. I also do not know if it understands me well."
	Contextual cut editing from various angles	"Current tools do not understand various situations and contexts. I used the automatic screen editing feature of my tool, and while it did a good job with the screen transitions, it cut off my narration. How great would it be if I could clearly convey what I want."
	Inconvenience in working on finer details	"Editing details is so cumbersome. Zooming in, removing stutters, adding pinpoint subtitles, identifying scene transition points, inserting background music, color correction, cutting out silences to make it feel short, adding mosaic effects, all of these require an immense amount of labor. How long will I have to keep doing these repetitive tasks? Is there a faster way to get them done?"
	Finding and sharing diverse editing templates	"Thanks to reels templates, life is often easier! But most tools lack diverse templates. So, I have to create them all myself. It would be great if I could share templates from other creators and also share my own."
	Using special effects and sound effects	"There are many special visual and sound effects I want to try, but my limited skills make it hard to create diverse content."
	Writing catchy titles, captions, and phrases	"How can I write titles and body that stop people moving on to the next short-form video? I have been copying the titles and content from YouTube videos with high views that are similar to mine... Well.. I have also had good responses when I followed memes or trends."

Uploading & Management	Managing upload schedules and processes	"To get favored by the algorithm, I need to upload daily or consistently, but it is hard to manage that schedule."
	Dealing with copyrights and portrait rights	"Occasionally I make mistakes, but they are critical ones. If I get reported for copyright infringement or invasion of privacy, I have to delete the video. It is sad to think that the channel I have worked so hard on could disappear due to a mistake"
	Evaluating and analyzing my channel's performance	"When I am filming, planning, and editing by myself, there are times I have to judge whether the result is good enough and will resonate with people. I do not have friends to ask, or they are indifferent. It is very difficult to assess the state and performance of my channel, and what needs to be improved."

Table 3: The list of challenges consists of related scenarios. We provided participants with a survey that included these challenges, based on the ones identified in previous interviews.

A.2 AI Capability Cards

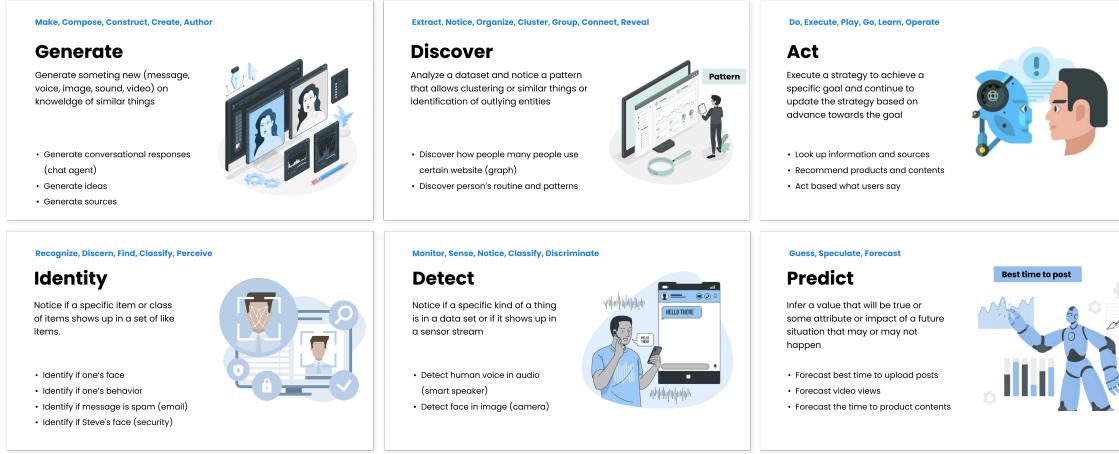


Figure 4: Examples of AI capability cards. To improve participants' comprehension of AI technology, we offered AI capabilities in a card format, drawing inspiration from the work of [76].

A.3 Example of AI-infused Editing Tool

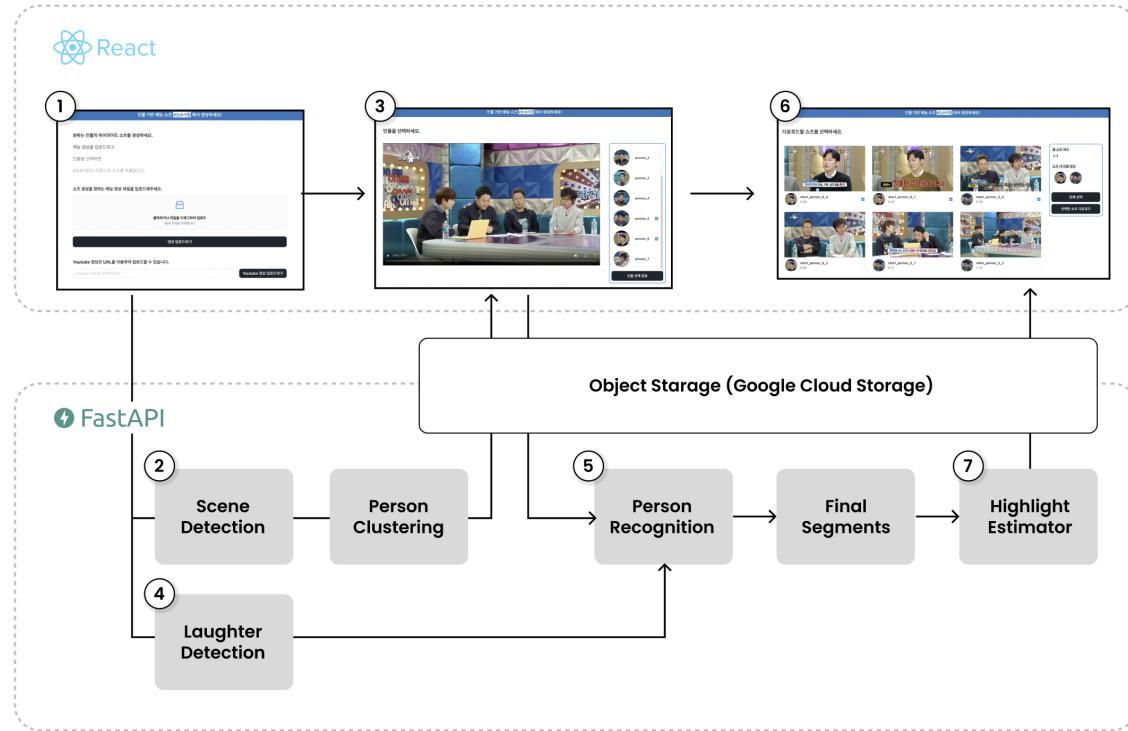


Figure 5: A brief overview of how we constructed the AI-infused editing tool to serve as a stimulus for the participatory workshop

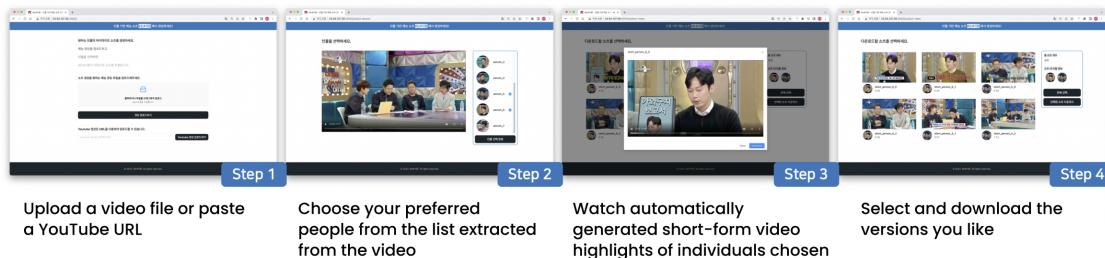


Figure 6: The process of how to use the example of the AI-infused editing tool during the participatory workshop

A.4 Comparative Analysis between AI-Edited Videos and Professionally Edited Videos

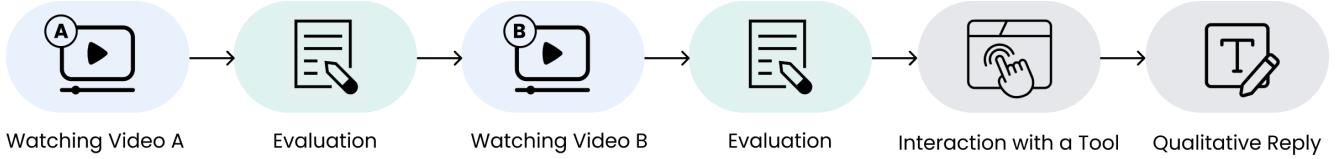


Figure 7: The evaluation process compares AI-edited videos with professionally edited ones – the baseline. This evaluation served as a functional assessment of our prototype’s effectiveness as workshop materials for participants to envision AI-infused editing tool features. Users assess their experiences after viewing two versions of videos. The baseline, a high-engagement short-form video, was selected from YouTube based on top views, likes, and comments. The evaluation concludes with an qualitative survey of each participant, collecting feedback on the AI-infused editing tools.

Survey Items	<i>AI-edited</i>		<i>Human-edited</i>		<i>Sig</i> <i>p-value</i>
	μ	σ	μ	σ	
(1) <i>The editing was done with a high level of completeness.</i>	4.8	0.47	4.69	0.79	.49
(2) <i>The length of the edited video appears appropriate.</i>	4	1.14	3.94	1.10	.54
(3) <i>The content of the edited video was easily comprehensible.</i>	4.02	0.467	4.45	0.877	.14
(4) <i>The edited video was engaging and enjoyable.</i>	4.09	0.94	4.09	1.01	.64

Table 4: To evaluate the functionality of an editing tool and gather user feedback, we conducted an experiment with 37 users, using a within-group design with counterbalancing. Participants evaluated the baseline and AI-edited videos using a 5-point Likert scale. The table contains mean scores, standard deviations, and paired t-test scores (95% Confidence Interval). The result indicated that there was no statistically significant difference between the two types of videos. Notably, it was observed that both achieved the same mean score for engagement (4.09). In terms of editing completeness, AI-edited videos had a marginally higher mean score compared to the baseline (4.8 vs. 4.69). Participants’ qualitative feedback revealed that they were unable to distinguish between videos edited by AI and professionally edited videos. It was also suggested that AI-infused editing tools are user-friendly and have the potential to streamline workflows.