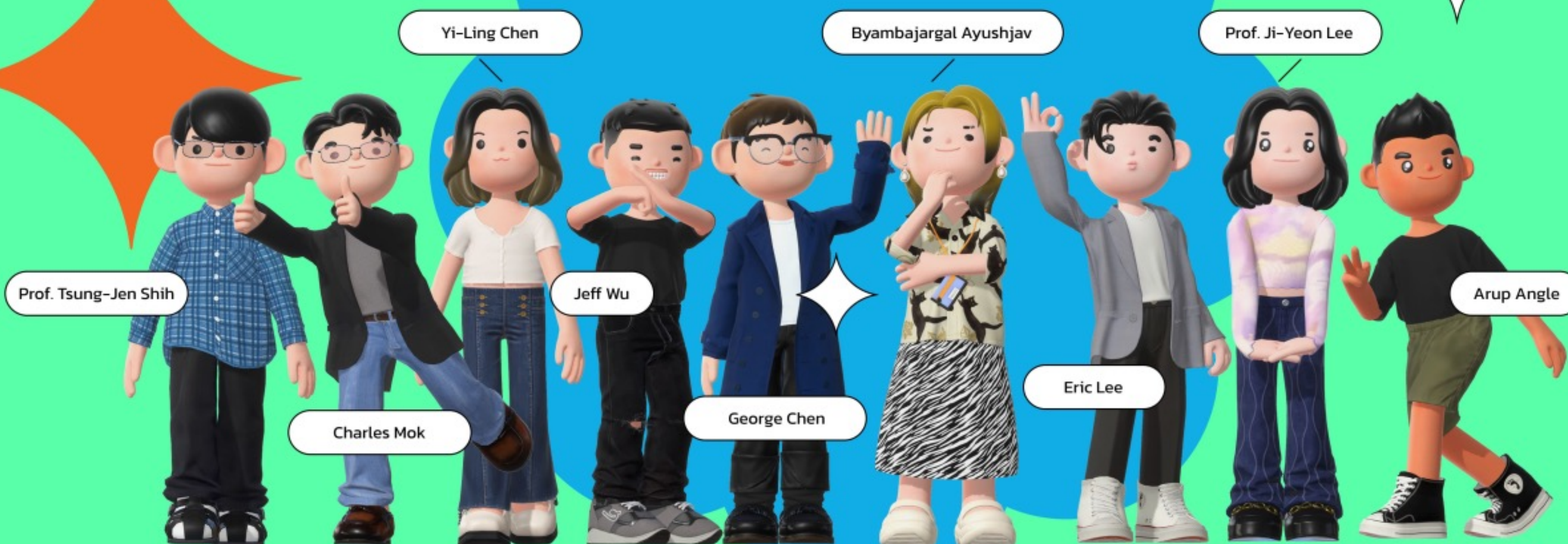


# Safety by Design, Empathy by Design

**Ji-yeon Lee, Ph. D., L. P. (Licensed Psychologist) in N. Y., U.S.A**  
**Dean of Student Affairs**  
**Professor, Counseling·UX Psychology**  
**Hankuk University of Foreign Studies**



# Bondee Trust & Safety Advisory Board



## META

### Women's safety partners

In partnership with Meta's Global Women's Safety Expert Advisors, we have developed policies, resources and tools to protect women from online abuse and allow them to control their experience online.

Our commitment to women's safety is longstanding. Over the years, we've sought the help of experts in the field to ensure our platforms are safe for women and have:

- Developed [strong policies](#) to protect women from online abuse
- Built [tools](#) to empower users with control over their experience online
- Launched [technology](#) to combat the sharing of non-consensual intimate images
- Created the Women's Safety hub as a centralized location for our online safety resources

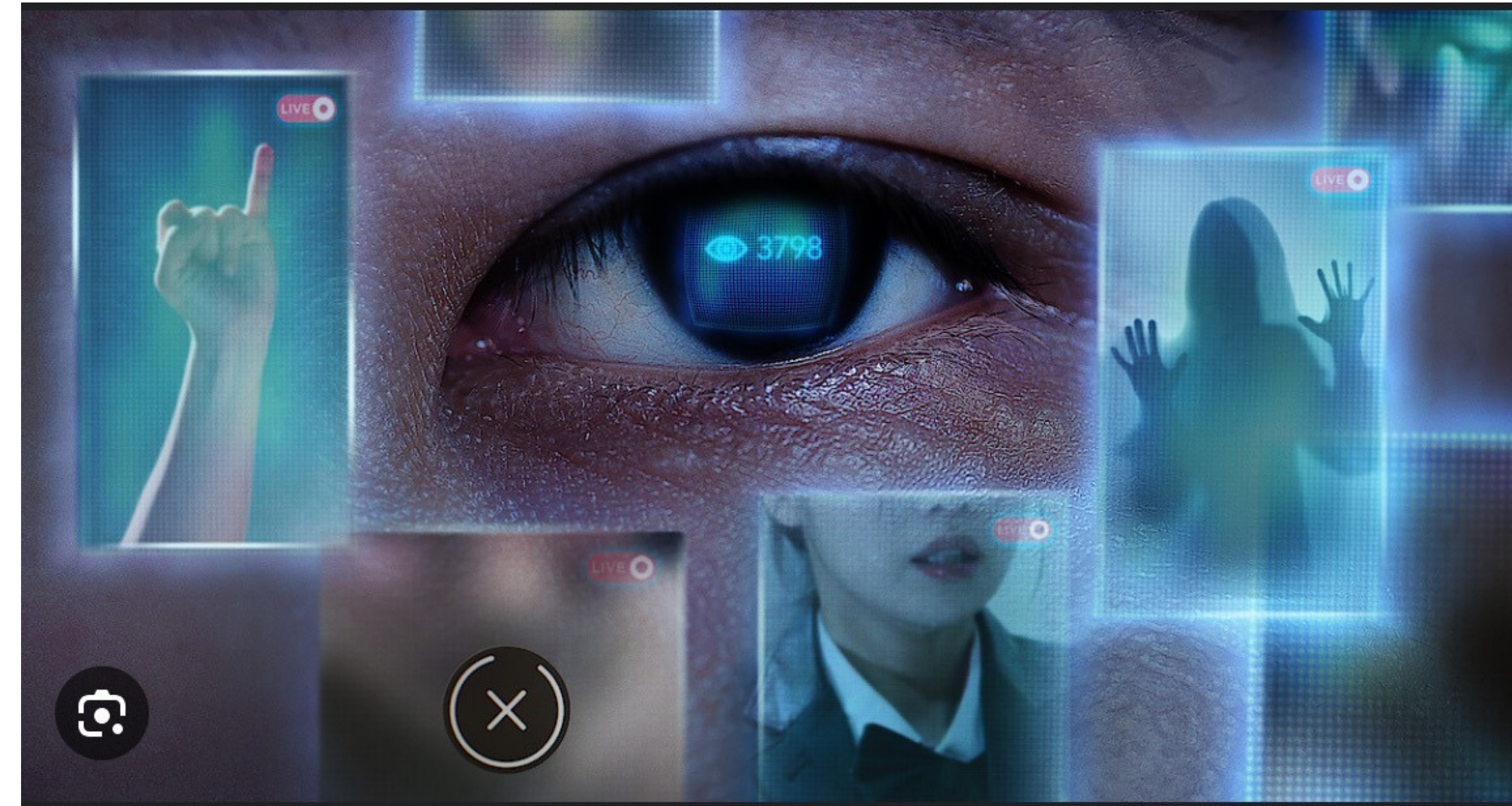
After working with over 400 women's safety organizations and experts across the world, we established a specific group of advisors. These experts have contributed to furthering the safety of women both online and off and are distinguished in the field. The group partakes in our quarterly meetings dedicated to advancing the safety of women online.

### Global Women's Safety Expert Advisors

Our current group of global advisors includes:

- Bishakha Datta, Point of View (India)
- Enrica Duncan, NOSSAS / Mapa do Acolhimento (Brazil)
- Asher Flynn, Monash University (Australia)
- Audace Garnett, National Network to End Domestic Violence (US)
- Caitríona Gleeson, Women for Election (Ireland)
- Margarita Guillé Tamayo, Interamerican Network of Women Shelters (Mexico)
- Neema Iyer, Pollicy (Uganda)
- Dr. Ji-yeon Lee, Hankuk University of Foreign Studies (Korea)
- Kalliopi Mingeirou, UN Women (Global)
- Lisa Moore, The Women's Foundation (Hong Kong)
- Tunggal Pawestri, Yayasan Hivos (Indonesia)
- Mariane Dorothy Rosario, Youth Activist (Philippines)
- Jyoti Vadehra, Centre for Social Research (India)
- Stephanie Willman Bordat, MRA Mobilising for Rights Associates (Morocco)

# Hidden Camera in South Korea; Nth room incidents



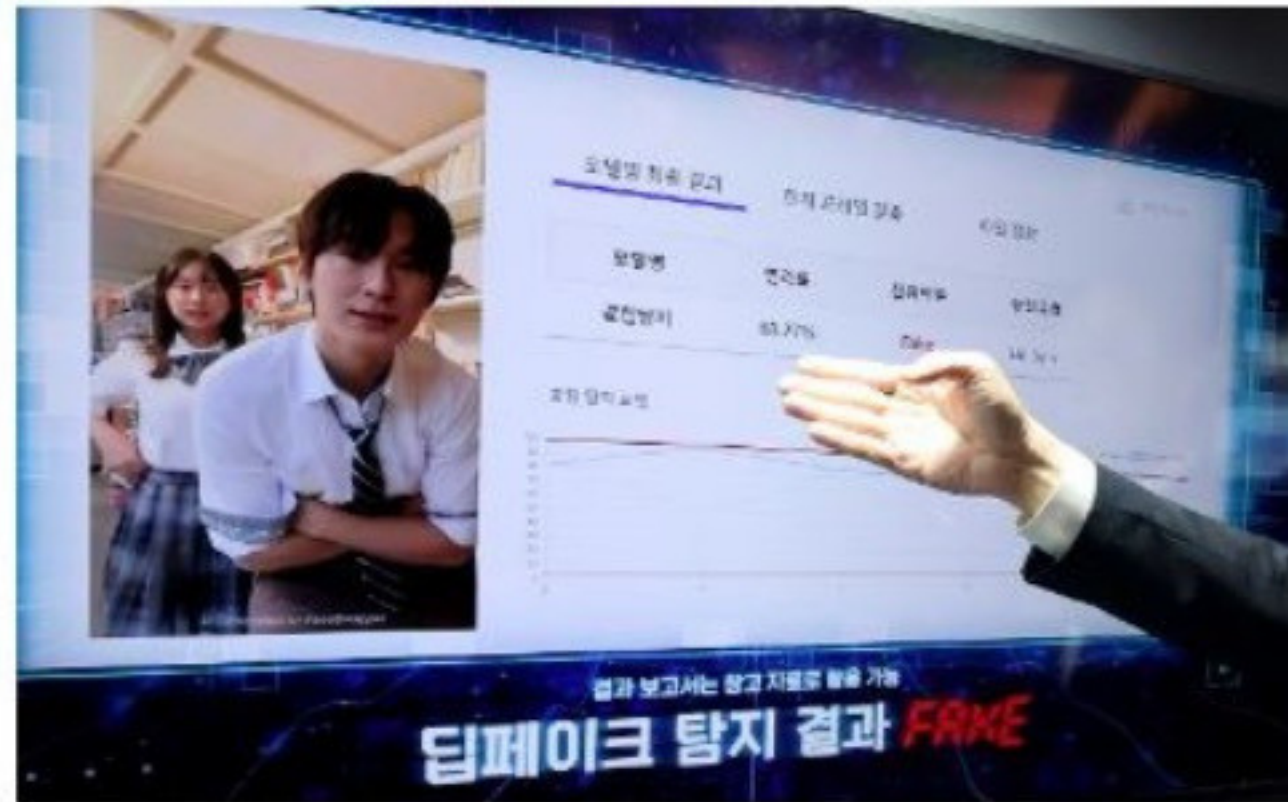
A hidden camera search squad scan a restroom near a public pool.

# DEEPFAKES IN SOUTH KOREA

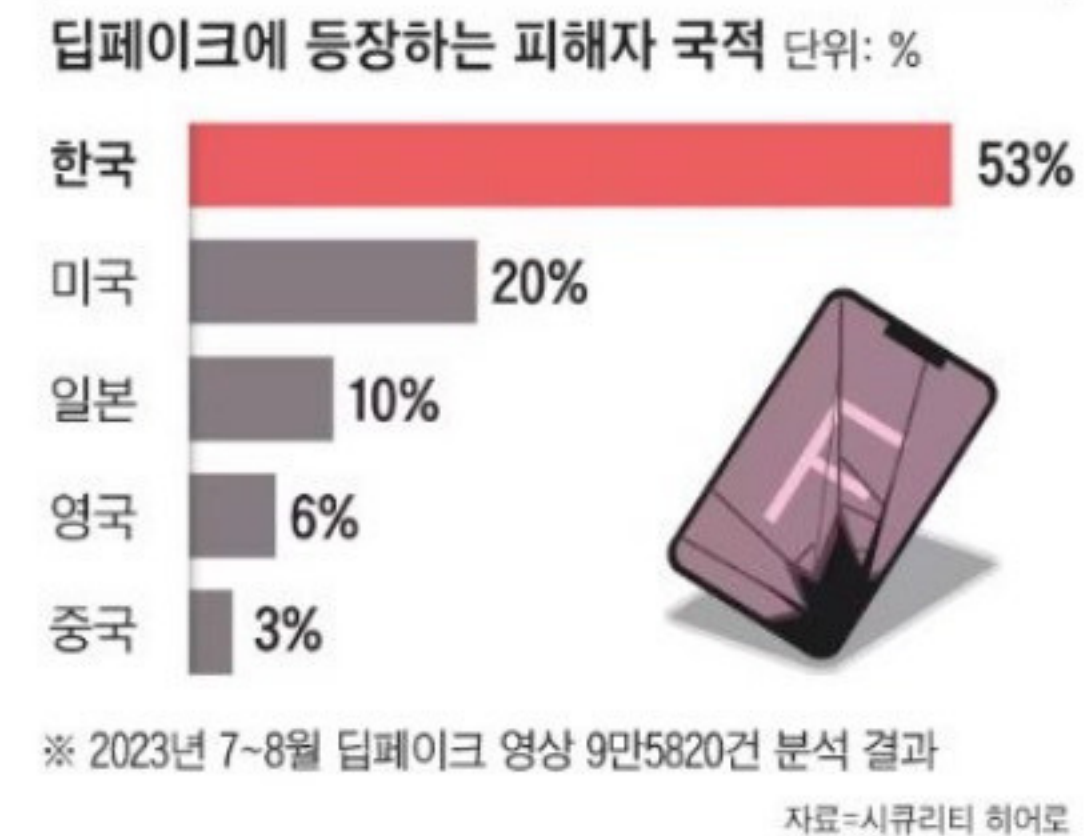
## 53% of Deepfake Victims Are Korean, the Highest in the World

윤진호 기자 김경필 기자

입력 2024.08.31. 01:20



More than half of the victims of deepfake(AI-generated fake content) sexual exploitation are Korean, according to a recent report. The U.S. cybersecurity firm 'Security Hero' recently released its "2023 DeepfakeStatus" report, which revealed that "half of the individuals appearing in deepfake pornographic content are Korean." The report analyzed 95,820 videos posted on ten deepfake pornographic sites and 85 video-sharing platforms such as YouTube over two months, from July to August of last year. In response to the growing spread of deepfake pornographic crimes, the South Korean government is considering criminalizing the possession of deepfake pornographic material and allowing undercover police investigations to catch the creators. Current sexual violence laws allow for punishment if someone's face, body, or voice is edited, synthesized, or manipulated in a lewd manner without their consent, but it must be proven that there was an intention to distribute the material.



그래픽=이철원

According to Security Hero's report, 53% of individuals appearing in deepfake pornographic content were Korean. This is more than double the number of victims in the United States (20%), the second-largest group. Japan followed with 10%, the UK with 6%, China with 3%, India with 2%, Taiwan with 2%, and Israel with 1%.

Security Hero stated, "Korean singers and actresses, in particular, have been frequent targets of deepfake content." The report listed the top 10 most-targeted individuals in deepfake pornographic content, eight of whom were Korean singers. Positions 1 through 7 and 9 were held by Korean singers, while a Thai singer was 8th, and a British actress was 10th. The most-targeted Korean singer appeared in 1,595 deepfake sexual exploitation videos, which garnered a total of 5.61 million views. The Wall Street Journal, citing this report, recently stated, "This suggests that South Korea is at the epicenter of the global problem of generating and distributing fake pornography." The BBC also reported, "South Korea is facing a deepfake pornography emergency."

Domestic statistics also show a sharp increase in the number of underage victims. According to data submitted by lawmaker Yang Bu-nam of the Democratic Party from the National Police Agency, 59.8% (315 individuals) of the 527 deepfake victims reported to the police from 2021 to 2023 were teenagers.

# DEEPPAKES IN SOUTH KOREA

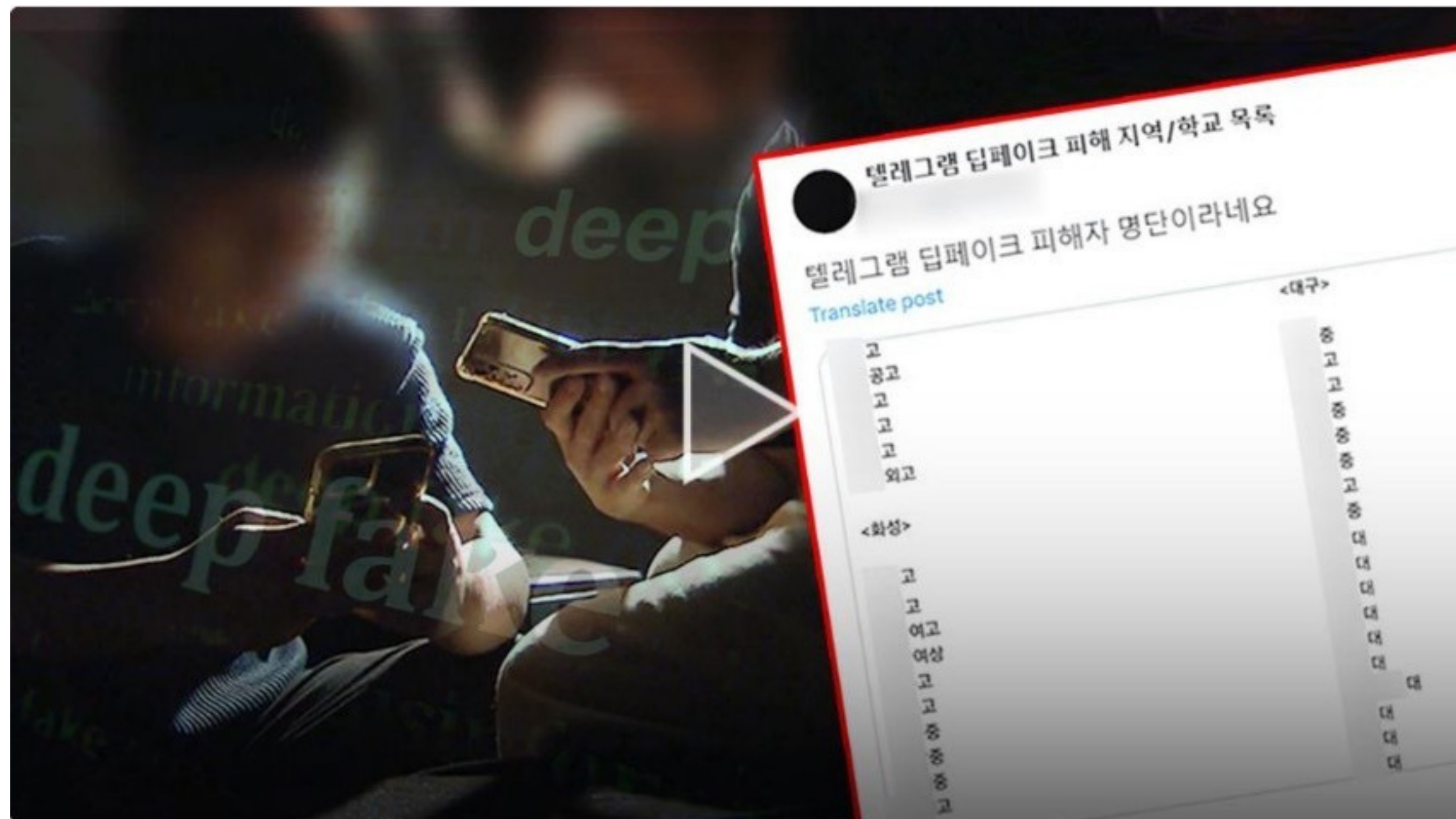
## **'Seoul National University Deepfake' Accomplice Sentenced to 5 Years in Prison... "Victims' Dignity Completely Destroyed"**

A defendant involved in the production and distribution of illegal deepfake videos of female alumni by graduates of Seoul National University has been sentenced to five years in prison. This is the first sentencing among those involved in the Seoul National University deepfake group sex crime case. On the 28th, Judge Kim Yoo-rang of the Seoul Central District Court's Criminal Division 14 sentenced Park (28), who was indicted and detained on charges of violating the Act on Special Cases Concerning the Punishment of Sexual Crimes (illegal filming), to five years in prison. Judge Kim stated, "The content of the recordings and the fabricated videos is so humiliating and disgusting that even an ordinary person would find it hard to describe," adding, "Considering the duration of the crimes, the number of videos posted, and the relationship with the victims, the severity of the crimes is extremely serious."

It was confirmed that Park committed illegal filming crimes for about eight months starting in 2017 and, over the span of four years starting in 2020, edited and synthesized fake photos and videos that could provoke sexual desire or shame. He then posted or sent these through platforms like Telegram. Park recorded a victim urinating while visiting his residence and then synthesized and edited this into photos and videos, combining the victims' faces with naked bodies. In total, he created 419 fake photos and videos aimed at causing sexual humiliation for 12 victims. Furthermore, he posted or sent 1,735 photos and videos of 17 victims through platforms like Telegram. Park was an accomplice in the so-called "Seoul National University Deepfake Sex Crime" case, where Seoul National University graduates produced and distributed illegal synthesized materials targeting female alumni.

# DEEPFAKES IN SOUTH KOREA

## "Our School Was Also Humiliated" —List of 150 Middle and High Schools Causes Uproar




A list of schools where either victims of "deepfake sex crimes" or the perpetrators are reportedly enrolled has been rapidly shared and circulated through platforms like Telegram and X (formerly Twitter). An account titled "[Telegram DeepfakeVictim School List](#)" has posted the names of around 150 middle and high schools, as well as 40 universities. The operator of the account has been adding schools to the list while sharing chatroom conversations and reports related to so-called "acquaintance humiliation" crimes committed through deepfakes. Although the actual scope or details of the crimes have not yet been confirmed through investigations, the fear that "my photo could also become 'pornographic' and circulate online" is causing considerable anxiety among students and parents. In some cases, students from specific schools are reportedly creating their own awareness campaigns about the situation, urging fellow students to remove their selfies or other photos from social media.

# PROTESTS

**여성혐오폭력 규탄 공동행동**  
@endfemicide2023

Thank you to everyone who attended today's rally and supported it by changing your profile pictures. Thanks to your support, we were able to complete the rally successfully.

When we initially filed the assembly report, we set a target of 500 participants, but we concluded with a remarkable turnout of 6,000 people on-site.



오전 12:56 · 2024년 9월 22일 · **127.3만** 조회수

2,731 재게시 100 인용 8,294 마음에 들어요 205 북마크

**Woman of Leisure** @WomenLeisure

Dear our international friends. Korean women are under a serious attack. This message was sent to me from Korea just now. They need your help. The Korean media has turned a blind eye. There is no news of any kind. They are seeking your help to bring this to light. Please RT!

게시물 번역하기

**we are korean students**

pleas help us. we are a korean students. no woman is safe in korea. spacial 12~22 students. and i am 18. Deepfake Room combines girls' faces and pornography to spread the footage

South Korea's male population aged 15-29 is 4.3 million.  
220,000 Telegram Deepfake Room Users

So, about 5%  
One out of every 20 people is in this room

can't punishing them for the laws of the Republic of Korea...  
They're saying, "Enjoy it more because you're not punished."

For some reason, the Korean media is trying to subside this case and not report it.  
pleas help us.....

오후 8:37 · 2024년 8월 26일 · **139.5만** 조회수

**[현장] "성착취물 떠도는 것 알고 자퇴 고민...꼭 살아 있어 달라"**

'딥페이크 엄벌 촉구' 6년 만에 혜화역 시위

기자 김채은

수정 2024-09-22 20:18 등록 2024-09-21 19:08

기사를 읽어드립니다 4:40 ▶ 🔊



21일 오후 서울 지하철 4호선 혜화역 앞 대학교에서 열린 '딥페이크 성착취물 엄벌 촉구 시위' 현장. 만든놈, 판놈, 본놈 모조리 처벌하라' 집회에서 참석자들이 구호를 외치고 있다. 김채은 기자

"만든놈, 판놈, 본놈 모조리 처벌하라!"

21일 오후 3시, 서울 지하철 4호선 혜화역 앞 대학교에 검은 옷을 입고 검은 마스크를 낀 여성들이 끝없이 늘어섰다. 수만명의 여성이 모였던 '2018년 혜화역 시위' 뒤 6년 만이다. 그때는 불법촬영, 지금은 딥페이크(불법합성)로, 구호만 바뀌고 '그대로'인 한국 사회 성범죄에 분노한 이들이 뭉쳤다.

# From Prosecution to Sentencing: Gaps in Digital Crime Justice

- NCII cases are likely to be dropped at the prosecution level ; the rejection rate is high
- Even in prosecuted cases: most of cases **suspended** or **fined**.

## Legal debate on Possession

What counts as possession in the digital era?

- URL links? Cache files?
- Cloud-stored content?



# Silenced by FEAR

Impact on **Personal Freedom** and **Expression**  
→ limit SNS Usage (self-censoring)



**80% of women**  
reported high levels of daily **fear** about  
**hidden cameras and deepfakes**

real safety isn't just about protection.  
It's about freedom — the freedom to exist, to speak,  
to be visible without fear



# REFRAIN FROM SNS

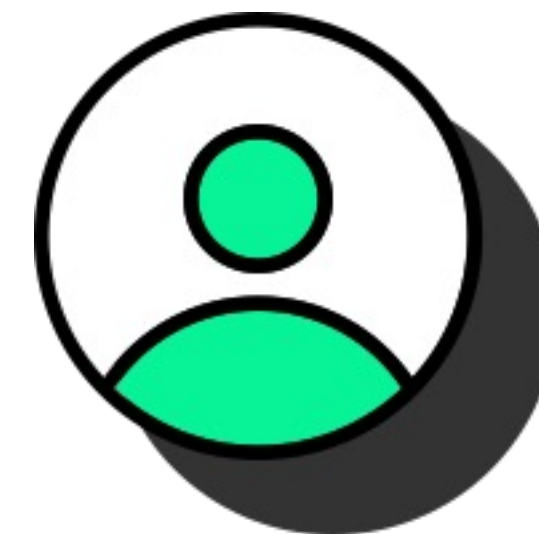
I deleted all my pictures from fb and deleted my profile picture on kakaotalk. however, simply not using SNS doesn't solve the issue. Even official documents like graduation photos or student IDs can be used to create deepfakes, meaning women have to live as if wearing an "online hijab," it feels as though women's very existence is being erased.



---

# CHANGES IN SOCIAL RELATIONS (Micro-aggression)

It feels like society suppresses the ability to talk about feelings of anxiety of deepfake, even add to the problem by making hurtful comments, saying things like, "**You won't go through that, don't worry,**" trivializing others' fears. people make degrading comments about others' appearances during this process. So I ended friendships over comments like these.



# From Reaction to Prevention

How do we design a digital world  
where women no longer need to disappear to feel safe?

# From Bystander to UPSTANDER

shift from teaching women to protect themselves  
to teaching everyone to respect others

# Stop NCII, Helping a friend

# User-centered Outreach/Safety by Design



## Session

50 minutes  
Training



## Peer Mentors

3 Digital Safety  
Ambassador  
(university students)



## Participants

18 High school  
Student  
-Peer Counselors



## Mentor Eligibility

Must take Users'  
digital Safety course in  
Uni.  
& Digital Safety  
Tool Training



## Approach

Trauma-informed  
  
Interactive

# From Bystander to Upstander; Trauma-Informed Education against NCII

01



## Limitation of Self-Protection-Centered Approach

Traditional approaches may assume students as potential victims or perpetrators->trigger unnecessary gender conflict in mixed-gender groups.

Our model: focuses on shared responsibility & collective safety.

02



## Advantages of Peer-Support Education

Enhances participation, empathy, and inclusivity → fosters learning engagement.

Training evokes positive emotions (*pride, confidence*) rather than shame or anxiety.

Certification (e.g., *Digital Ambassador Certificate*) strengthens motivation.

03



## Beyond victim vs perpetrator

Frames digital safety not as an issue for specific groups, but as part of everyday life for everyone

Bystander -> Upstander

# Safety by Design

## Peer-led & relatable



Messages delivered by university students help reduce psychological distance (and defensive attitude) when addressing sensitive topics.

## Skill building, not just information giving



Focuses on role-play-based training rather than one-way information delivery, Experiential learning

Helping a  
Friend  
experiencing  
NCII.

## Psychology based (Theory of Planned Behavior)



Promotes supportive attitudes, positive subjective norm, and behavioral control (by cultivating communication skills & teaching digital safety tools) to helping a friend experiencing NCII

## Sustainable Model: Linking High School & University



Most high schools run **Peer Counselor Clubs**. Club activities are officially **counted as credits** for university entrance. Our model: **connects these extracurricular activities with university curricula (user safety course at HUFS)** → creating a sustainable model

# Project Timeline

## Training University students as Mentors



Trauma-informed  
Communication skills,  
Using Digital Safety tool

## Program on High school Students



The program's  
effectiveness has been  
empirically validated.

## Program on University Students



The program's  
effectiveness has been  
empirically validated.

## Teachers Workshop



The program's  
effectiveness has been  
empirically validated.



Culture

Media

Local and national policies

What needs to shift to create and maintain safety?

Education

Tech platforms

It's about rebuilding trust —  
trust in technology,  
trust in institutions,  
and trust in one another.

# STOP NCII user



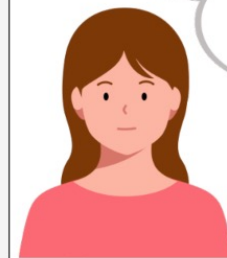
- Name: A
- Gender / Nationality: Female / South Korea
- Age: 21
- Situation: After breaking up with her ex, she told him she had a new boyfriend. In retaliation, he said, "I've already spread all the naked pictures you sent me on the internet." She is now experiencing extreme anxiety and fear and is unable to ask for help from anyone around her. She wants to find and delete these photos as quickly as possible.
- How she found STOPNCII: While searching online for a solution to her problem, she stumbled upon the site.

## Does STOP NCII meet user's needs? - Problems in the exploration stage

### User's (A) Needs

1. A wants to feel reassured and calm their anxiety and fear while exploring the website.
2. A wants to be confident that the website is trustworthy.
3. A wants to intuitively understand the steps involved in the illegal content removal process.

Does STOP NCII meet the first user's need?



## Does STOP NCII meet user's needs? - Problems in the engagement stage

### Problems in the Case Creation Process



#### 1) Absence of a Back Button

Since there is no back button, if the user accidentally selects the wrong option, they must restart the "Create Your Case" process from the beginning, causing inconvenience.

#### 2) Language That May Cause psychological Burden for the User

During the application process, the use of direct language such as "Nude/Semi-Nude" and "Are you engaging in sexual activity?" may impose a psychological burden on the user.

## The composition of the stages in a User Journey Map

- Awareness → Exploration → Engagement → Action → Follow-up Support → Revisit and Recommendation

Stage	User Actions	Emotions
Awareness	Realizes that illegally leaked photos have been distributed.	Shock, anxiety, anger
Exploration	Searches for solutions by looking up terms like "how to delete illegal photos."	Confusion, fear, slight relief
Engagement	Attempts to request photo deletion on the website.	Anxiety, hope
Action	The website tracks and deletes the illegal post.	Anticipation, nervousness
Follow-up	After confirming deletion, considers psychological counseling or legal action.	Relief, or continued anxiety
Retention & Advocacy	Revisits the site if another issue arises or recommends it to other victims.	Trust, caution

## Does STOP NCII meet user's needs? - Problems in the exploration stage

### UI that can impose a psychological burden on the user



#### 1) Illustrations and language used on the website

For A, who is unable to tell anyone about her situation and is feeling fearful and anxious, illustrations depicting nudity or covered eyes can impose a psychological burden.

Likewise, direct words such as "victim of abuse" and "threat" can have the same effect.

## Does STOP NCII meet user's needs? - Problems in the engagement stage

### Problems in the Case Creation Process



#### 3) Inefficient Navigation Structure

The user sees this screen after selecting the "18 or over" tab. However, at this stage, when clicking the "reporting guide" hyperlink, the page does not immediately display information about the report guide. Instead, it first shows guidance for helping minors under 18 (less relevant information).

This can lower the user's expectations of the website and reduce trust.

The "Search" button does not work → further decreases trust.