디지털 환경에서
아동·청소년 권리 보장을
위한 정책토론회

플랫폼 안전설계를 위한 위험평가 제도를 중심으로

# 디지털 환경에서 아동·청소년 권리 보장을 위한 정책토론회

## 플랫폼 안전설계를 위한 위험평가 제도를 중심으로

**일 시** 2026. 9. 1. (화) 13:30

**장 소** 국회의원회관 제2간담회의실

**주 최** 더불어민주당 국회의원 조인철 · 한준호 / 국민의힘 국회의원 최형두 / 초록우산

# 디지털 환경에서 아동·청소년 권리 보장을 위한 정책토론회

## 플랫폼 안전설계를 위한 위험평가 제도를 중심으로

**일 시** 2026. 9. 1. (화) 13:30

**장 소** 국회의원회관 제2간담회의실

**주 최** 더불어민주당 국회의원 조인철 · 한준호 / 국민의힘 국회의원 최형두 / 초록우산

# 목 차 Contents

## 개 회 사


**조인철 국회의원** (더불어민주당, 광주 서구갑) ····· 06

**한준호 국회의원** (더불어민주당, 경기 고양시(을)) ····· 08

**최형두 국회의원** (국민의힘, 경남 창원시마산합포구) ····· 10

**황영기 회장** (초록우산) ····· 12


## 주제발표


**플랫폼 위험평가 제도 - 해외 사례와 국내 적용 과제** ····· 15

윤금낭 센터장(디지털산업정책연구소 디지털안전센터)


## 종합토론


좌장 **김경희 교수**(한림대학교 디지털콘텐츠융합스쿨)

토론 **성욱제 연구위원**(정보통신정책연구원 미디어전략그룹) ····· 33

**김현기 과장**(초록우산 옹호사업본부) ····· 37

**한승혜 리더**(한국인터넷기업협회 디지털경제연구원) ····· 43

**최진응 입법조사관**(국회입법조사처 과학방송통신과) ····· 47

**김우석 과장**(방송미디어통신위원회 디지털유해정보대응과) ····· 49

## 개 회 사

**조인철 국회의원** (더불어민주당, 광주 서구갑)

**한준호 국회의원** (더불어민주당, 경기 고양시(을))

**최형두 국회의원** (국민의힘, 경남 창원시마산합포구)

**황영기 회장** (초록우산)

디지털 환경에서
아동·청소년 권리 보장을 위한 정책토론회

# 개 회 사



조인철 국회의원
(더불어민주당, 광주 서구갑)

안녕하십니까. 광주 서구갑 국회의원 조인철입니다.

먼저, 우리 아이들의 건강하고 안전한 디지털 미래를 위해 뜻깊은 자리를 마련해 주신 초록우산 황영기 회장님과 관계자 여러분께 깊은 감사를 드립니다. 아울러 귀중한 발제를 준비해 주신 디지털산업정책연구소 윤금낭 센터장님과, 이어지는 토론에서 풍성한 지혜를 나눠주실 전문가 여러분께도 진심으로 감사드립니다.

오늘날 디지털 플랫폼은 우리 아동과 청소년들에게 단순한 여가 공간을 넘어, 세상과 소통하고 배우며 성장하는 '일상의 무대' 그 자체가 되었습니다. 하지만 아이들이 마주하고 있는 디지털 환경의 이면에는 결코 가볍지 않은 위협들이 도사리고 있습니다.

생성형 AI와 숏폼 콘텐츠가 일상화되면서 스마트폰 과의존 위험군이 지속적으로 증가하고 있으며, 서비스의 알고리즘과 설계 방식 자체가 아이들의 과몰입을 유도하고 유해 콘텐츠에 반복적으로 노출시키는 등 새로운 '구조적 위험'을 낳고 있습니다.

그동안 우리 사회의 디지털 아동 보호 체계는 유해 정보나 불법 콘텐츠가 발생한 뒤에야 사후적으로 삭제하거나 접근을 차단하는 수동적인 방식에 머물러 있었습니다. 그러나 노출과 확산이 빛의 속도로 이루어지는 작금의 플랫폼 환경에서는, 이러한 사후 대응만으로 우리 아이들을 온전히 지켜내기 어렵습니다.

저 역시 국회에서 AI 윤리 기준을 바로 세우고 플랫폼 내 안전장치를 마련하는 등, 기술 발전의 이면으로부터 아동·청소년의 권리와 안전을 지키기 위한 의정활동에 주력해 왔습니다. AI 기술의 안전성 확보와 소셜미디어 생태계 정화를 논의하며 얻은 명확한 결론은, 서비스 기획과 설계 단계에서부터 플랫폼의 예방적 책임을 대폭 강화해야 한다는 것입니다.

그런 의미에서 오늘 심도 있게 논의될 '플랫폼 사전위험평가제도'는 우리 아이들을 위한 디지털 안전망을 재구축하는 핵심 열쇠가 될 것입니다. 이제 우리나라도 선도적인 국제적 흐름에 발맞추어, 아동의 관점이 오롯이 반영된 구체적인 위험평가 지표와 촘촘한 안전 체계를 국내 플랫폼 생태계에 굳건히 뿌리내려야 할 때입니다.

오늘 이 자리에서 모아주신 소중한 고견과 연구 결과가 단순한 논의에 그치지 않고 실효성 있는 정책으로 구현될 수 있도록, 저 조인철이 국회에서 최선을 다해 뒷받침하겠습니다. 서비스 설계 단계부터 아동의 권리를 최우선으로 보장하는 '사전위험평가제도'가 확립될 수 있도록 관련 법안과 제도를 꼼꼼히 살피겠습니다.

다시 한번 오늘 뜻깊은 토론회 개최를 진심으로 축하드리며, 우리 아이들이 안전한 환경에서 마음껏 꿈을 펼칠 수 있도록 헌신해 주시는 모든 분들의 건강과 행복을 기원합니다.

고맙습니다.

디지털 환경에서
아동·청소년 권리 보장을 위한 정책토론회

# 개 회 사



한준호 국회의원
(더불어민주당, 경기 고양시(을))

안녕하십니까, 더불어민주당 경기 고양시(을) 국회의원이자 국회 과학기술정보방송통신위원회 간사를 맡고 있는 한준호입니다.

디지털 환경에서 아동·청소년의 권리를 보장하기 위한 정책토론회에 함께해 주신 모든 분들께 감사 인사를 드립니다. 특히 오늘 토론회를 함께 준비해 주신 조인철 의원님, 최형두 의원님, 그리고 오랜 시간 아동의 안전한 성장을 위해 애써오신 초록우산 황영기 회장님과 관계자 여러분께 깊은 감사의 말씀을 전합니다. 아울러 귀한 연구 성과를 발제해 주실 디지털산업정책연구소 윤금낭 센터장님과 현장의 다양한 시각을 나눠주실 토론자 여러분께도 감사드립니다.

이제 온라인 플랫폼은 아동·청소년에게 일상의 공간이 되었습니다. 그만큼 공간의 설계와 운영이 아이들의 삶에 미치는 영향도 커지고 있습니다. 그동안 우리 사회는 유해정보가 발견된 이후 이를 삭제·차단하는 사후적 대응을 중심으로 아동을 보호해 왔습니다. 물론 이러한 사후 대응도 앞으로 계속 필요한 조치입니다. 다만 콘텐츠의 확산 속도가 갈수록 빨라지는 디지털 환경 속에서 서비스 설계와 운영 단계에서부터 위험을 예방하려는 보완적인 노력 역시 필수적입니다.

오늘 토론회는 우리 아이들을 디지털 환경으로부터 안전하게 지킬 방안에 대해 함께 고민하는 자리라는 점에서 뜻깊습니다. 해외 여러 나라의 제도와 경험을 살펴보고, 우리 현실에 맞는 방안을 모색해 나가는 과정은 아동의 권리를 지키는 소중한 발걸음이 될 것입니다.

저 역시 국회 과학기술정보방송통신위원회에서 오늘 나와주신 전문가와 관계자 여러분의 고견을 국회에서의 논의에 성실히 담아내도록 노력하겠습니다. 아동·청소년의 안전이라는 공동의 목표 앞에서, 오늘 이 자리가 서로의 지혜를 모으는 장이 되기를 바랍니다.

다시 한번 오늘 토론회를 준비해 주신 관계자 여러분과 발제자, 토론자, 그리고 참석해 주신 모든 분들께 감사드립니다. 아이들이 더 안전한 디지털 환경에서 자라날 수 있도록, 앞으로도 관심을 가지고 함께하겠습니다. 고맙습니다.

디지털 환경에서
아동·청소년 권리 보장을 위한 정책토론회

# 개 회 사



최형두 국회의원
(국민의힘, 경남 창원시마산합포구)

안녕하십니까. 국회 과학기술정보방송통신위원회 국민의힘 간사, 국회 과학기술원자력우주항공소위원장, 마산합포 국회의원 최형두입니다.

우리 아이들의 건강하고 안전한 디지털 환경을 조성하기 위해 마련된 『디지털 환경에서 아동·청소년 권리 보장을 위한 정책토론회』 개최를 매우 뜻깊게 생각하며, 공동주최자로서 자리를 함께 해 주신 모든 분을 진심으로 환영합니다.

먼저 아동·청소년의 디지털 권리 보호에 뜻을 모아주신 존경하는 조인철 의원님, 한준호 의원님, 그리고 아이들의 행복을 위해 현장에서 늘 헌신해 오신 초록우산 관계자 여러분께 깊이 감사드립니다. 또한, 오늘 토론회에서 혜안을 보태주실 발제자와 토론자 여러분께도 따뜻한 환영의 인사를 전합니다.

오늘날 디지털 플랫폼은 단순히 정보를 검색하는 도구를 넘어, 우리 아이들이 학습하고 소통하며 자라나는 핵심적인 일상 공간이 되었습니다. 그러나 이에 비례하여 유해 콘텐츠 노출, 과몰입, 개인정보 침해 등 아동이 마주하는 위험의 크기와 깊이도 구조적으로 심화되고 있습니다.

그동안 우리의 아동보호 정책은 문제가 터진 뒤에야 삭제하고 차단하는 '사후 대응'에 의존해 온 것이 사실입니다. 그러나 소 잃고 외양간 고치는 식의 대응만으로는 유기적으로 진화하는 디지털 위험으로부터 아이들을 온전히 보호하기 어렵습니다. 이제는 플랫폼의 기획·설계 단계부터 위험 요소를 사전에 식별하고 줄여나가는 '예방적 안전 체계'로 패러다임을 전환해야 합니다.

이러한 시점에 「정보통신망법」 개정을 통한 '플랫폼 위험평가 제도' 도입 논의와, 이를 뒷받침할 초록우산의 연구 성과 공유는 매우 시의적절하고 귀중한 발걸음입니다. 해외 선진 사례를 바탕으로 우리 실정에 맞는 실효성 있는 평가 체계를 구축하는 것이야말로 아이들의 안전과 기업의 기술 혁신이 상생하는 길입니다.

저 최형두 역시 국회 과방위 간사이자 오늘 토론회 공동주최자로서 선제적이고 실질적인 입법 결실을 맺도록 앞장서겠습니다.

발의된 법안들이 현장에서 제대로 작동할 수 있도록 제도적 완성도를 높이고, 플랫폼 기업들이 스스로 안전한 환경을 만들 수 있도록 법·제도적 지원을 아끼지 않겠습니다.

오늘 토론회가 아동·청소년의 안전한 디지털 권리 보장을 위한 명확한 제도적 이정표를 제시하는 소중한 자리가 되기를 기대합니다.

참석해 주신 모든 분의 건승과 행복을 기원합니다.

매사 뜻대로 만사형통, 마산형통하십시오!

감사합니다.

디지털 환경에서
아동·청소년 권리 보장을 위한 정책토론회

# 개 회 사



황영기 회장
(초록우산)

안녕하십니까. 초록우산 회장 황영기입니다.

'디지털 환경에서 아동·청소년 권리 보장을 위한 정책토론회'에 함께해 주신 모든 분께 감사드립니다. 또한, 토론회를 위해 힘써주신 조인철 의원님, 한준호 의원님, 최형두 의원님, 발제를 맡아주신 디지털산업정책연구소 윤금낭 센터장님, 토론에 참여해 귀한 의견을 나눠주실 전문가 분들께도 감사드립니다.

아동·청소년에게 디지털 플랫폼은 정보를 얻고 친구와 소통하며, 여가를 보내는 일상의 공간입니다. 이러한 디지털 공간의 서비스 설계와 운영 방식은 아동의 삶과 안전에 직접적인 영향을 미칩니다. 특히, 어떠한 콘텐츠가 우선 노출되는지, 다음 영상을 계속 보도록 어떻게 유도하는지 등 설계와 운영 방식에 따라 더욱 쉽게 위험에 노출될 수 있는 취약성이 존재합니다. 실제 초록우산이 올해 1월 숏폼 플랫폼을 이용하는 만 14세 이상 아동·청소년 1,000명을 조사한 결과, 응답자의 53.4%가 유해 콘텐츠에 노출된 경험이 있다고 답했습니다. 유해 콘텐츠를 접한 아동·청소년 가운데 80.3%는 그 경로로 추천 알고리즘을 꼽았습니다. 이는 유해 콘텐츠 노출이 일부 이용자의 우연한 경험에 그치지 않으며, 플랫폼의 콘텐츠 선별 및 추천 구조가 위험의 확산과 밀접하게 연결돼 있다는 것을 보여줍니다.

그간 디지털 공간에서의 아동보호는 유해 콘텐츠를 발견한 뒤 삭제하거나 신고하고, 이용자의 접근을 제한하는 방식에 주로 의존해 왔습니다. 그러나 이미 노출이 발생한 이후의 조치만으로는 추천과 확산이 빠르게 이루어지는 플랫폼 환경에 충분히 대응하기 어렵습니다. 아동에게 어떤 위험이 발생할 수 있는지를 서비스의 설계·운영 단계에서 미리 확인하고, 그 가능성을 낮추는 예방적 접근이 필요한 이유입니다.

국제사회는 이미 이러한 예방적 접근에 나서고 있습니다. 호주, 영국, 유럽연합 등은 '안전설계(Safety by Design)'를 통해 디지털 공간의 아동 위험을 사전에 평가하고 완화하는 제도를 마련하고 있습니다. 플랫폼이 서비스를 제공하기에 앞서 아동에게 미칠 수 있는 위험을 평가하고, 확인된 위험을 줄이기 위한 조치를 마련하도록 제도를 정비하는 것입니다. 이는 보호의 책임을 아동과 보호자의 주의에만 맡기지 않고, 서비스를 설계하고 운영하는 플랫폼도 함께 부담하도록 정책의 중심을 옮기는 것으로 볼 수 있습니다.

초록우산은 국내에도 디지털 공간의 아동보호를 위한 플랫폼 위험평가 제도가 필요하다고 지속적으로 제안해 왔습니다. 동시에 제도가 실효성 있게 작동할 수 있도록 어떤 서비스와 기능을 평가할지, 아동에 미치는 위험을 어떤 기준으로 판단할지 등 구체적인 설계 작업이 필요하다고 판단했습니다. 이번 토론회에서 공유될 「디지털 플랫폼에서의 사전 위험평가 체계 도구 개발을 위한 기초연구 – 아동보호 관점을 중심으로」는 이러한 배경에서 진행되었습니다. 이 연구는 해외 주요국들의 위험평가 제도와 평가도구를 비교·분석하고, 이를 토대로 국내 플랫폼 환경에 적용할 수 있는 위험평가 체계의 기초 구조를 검토했습니다. 또한, 제도의 필요성을 주장하는 데서 한 걸음 더 나아가, 국내 플랫폼 환경의 실제 평가체계를 어떻게 구성할 것인지 등 보다 구체적 논의를 제안하고 있습니다.

현재 국회에서는 「정보통신망 이용촉진 및 정보보호 등에 관한 법률」 개정을 통해 플랫폼 위험평가 제도를 도입하기 위한 논의가 진행되고 있습니다. 오늘 토론회에서 초록우산의 연구 결과를 바탕으로 평가 대상과 기준, 평가의 검증과 공개, 플랫폼의 이행 방식 등 실효성 있는 제도 마련을 위한 실질적 논의가 이뤄지길 기대합니다. 아울러 디지털 공간에서의 아동 위험에 대해 사후대응이 아닌, 플랫폼의 설계 단계부터 아동의 권리와 안전을 보장하는 제도로 나아가는 시작점이 되길 바랍니다.

다시 한번 오늘 토론회를 위해 애써주신 국회의원님들, 디지털 아동보호를 위해 노력하고 계신 모든 전문가 분들께 깊이 감사드립니다. 디지털 환경에서 모든 아동이 행복하고 안전하게 성장할 수 있도록 앞으로도 초록우산은 언제나 어린이 곁에서 함께하겠습니다.

## 주제발표

**플랫폼 위험평가 제도**
**- 해외 사례와 국내 적용 과제**

윤금낭 센터장(디지털산업정책연구소 디지털안전센터)

2026. 9. 1.

# 아동·청소년의 디지털 안전을 위한 새로운 보호체계 모색

-해외 사전위험평가제도 사례와 국내 적용 과제-

윤금낭 센터장

디지털산업정책연구소 디지털안전센터

iDIP 디지털산업정책연구소 Institute of Digital Industry & Policy

© 2026. (디지털산업정책연구소) all rights reserved.

목 차

01.
**아동·청소년 디지털 이용환경 현황 및 문제점**

02.
**해외 사례 및 시사점**

03.
**플랫폼 사전위험평가 체계의 국내 적용 방향 및 과제**

© 2026. (디지털산업정책연구소) all rights reserved. 

01

# 아동·청소년 디지털 이용환경 현황 및 문제점

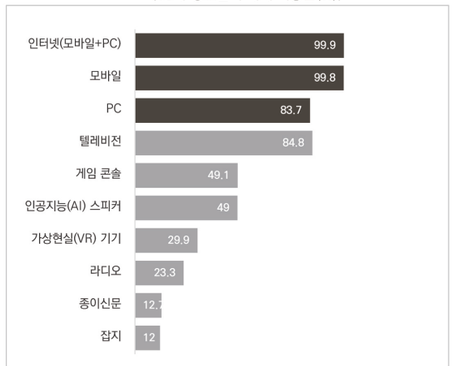
© 2026. (디지털산업정책연구소) all rights reserved. 



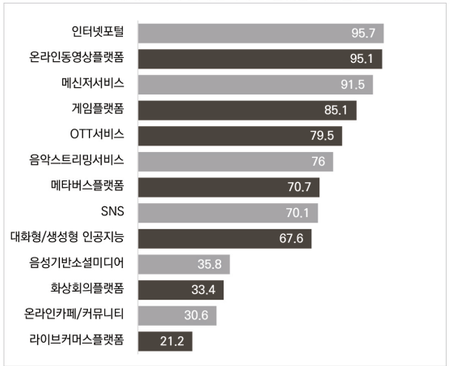
## 1-1. 디지털 환경에서 아동·청소년이 직면한 위험의 구조적 특성

- 국내 10대 청소년의 인터넷 이용은 99.9%로 일상 매체를 뛰어넘은 필수 매체로 작용함
- 일상화된 온라인 서비스 이용은 유해정보 노출 · 과몰입 · 개인정보 침해 등 위험의 발생 가능성을 높임
- 온라인동영상(95.1%), 게임(85.1%), 메타버스(70.7%) 플랫폼 이용에 이어 대화형/생성형 인공지능(67.6%)도 일상적인 서비스로 자리 잡고 있음

〈10대 청소년의 매체 이용률(%)〉




〈10대 청소년의 온라인서비스/플랫폼 이용률(%)〉



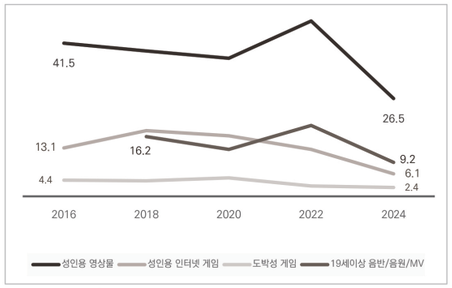

출처: 한국언론진흥재단(2025). 10대 청소년 미디어 이용 조사 (n.2,674); Q: 귀하는 지난 일주일 동안 다음 각 매체를 이용한 적이 있나요?, 천혜선(2026)에서 재인용

© 2026. (디지털산업정책연구소) all rights reserved. 

## 1-1. 디지털 환경에서 아동·청소년이 직면한 위험의 구조적 특성

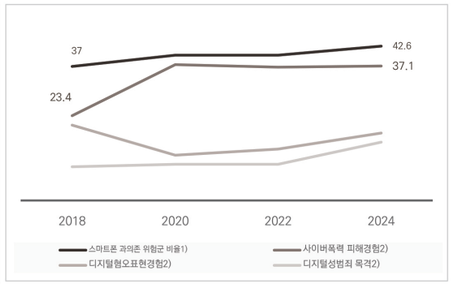
- 최근에는 유해정보 노출에 그치지 않고, 온라인 그루밍 위험이나 스마트폰 과의존 및 과몰입과 같은 새로운 유형의 위험 양상이 확대되는 추세
  - 성인용 영상물 이용률: 2016년 41.5% → 2024년 26.5%로 대폭 감소
  - 스마트폰 과의존 위험군 비율: 2018년 37.0% → 2024년 42.6%까지 증가
- 이전 인터넷상의 아동 위험은 불법유해정보로 인해 발생하는 위험이었다면, 현재는 동의 없는 아동의 개인정보 처리, 콘텐츠 추천·노출 알고리즘으로 인한 과몰입 등과 같이, 콘텐츠 자체의 속성과 상관없이 **서비스의 설계와 이용 맥락에 의해 발현되는 새로운 유형의 위험임**

〈청소년의 불법유해정보 경험률(%)〉




출처: 성평등가족부 「청소년 매체이용 및 유해환경 실태조사」, 각 년도

〈청소년의 온라인상 사회적 관계기반 불법·유해행위 경험률(%)〉




1) 2025년 스마트폰 과의존 실태조사 보고서, 2) 2025년 사이버폭력실태조사, 천혜선(2026)에서 재인용

© 2026. (디지털산업정책연구소) all rights reserved.




## 1-1. 디지털 환경에서 아동·청소년이 직면한 새로운 위험의 구조적 특성

- 온라인 플랫폼의 알고리즘 · 추천 시스템 등 플랫폼 설계 자체가 위험을 심화시키는 구조
- 이용자의 체류시간, 자극, 참여를 극대화하도록 최적화된 설계가 **아동의 과몰입과 유해 콘텐츠 반복 노출을 유발함**
- 영국은 아동이 직면하는 온라인 위험을 콘텐츠, 접촉, 행위위험이라는 3C로 구분함
- 이들 위험은 개별 콘텐츠가 아니라 플랫폼의 설계 · 작동 방식에서 구조적으로 발생한다는 특징을 가짐

| **콘텐츠 위험** (Content Risk) | **접촉 위험** (Contact Risk) | **행위 위험** (Conduct Risk) |
|---|---|---|
| 유해하거나 연령에 부적합한 콘텐츠에 노출되어 피해를 입을 위험. 콘텐츠를 수동적으로 소비하는 과정에서 발생<br>주요 위험요인<br>• 폭력, 성적 콘텐츠<br>• 자해, 자살, 섭식장애 조장<br>• 아동 성 착취물(CSAM)<br>• 범죄 조장 콘텐츠 | 위험한 이용자와 상호작용(접촉)을 하는 과정에서 발생하는 위험. 아동과 타인 간 상호작용에서 발현<br>주요 위험요인<br>• 온라인 그루밍(Online Grooming)<br>• 성 착취 목적 접근<br>• 낯선 성인과의 접촉<br>• 원치 않는 연락(DM)<br>• 개인정보 요구, 오프라인 만남 유도 | 이용자 간 상호작용, 행동 과정에서 위해행위의 대상이 되거나 참여함으로써 정신적, 정서적 피해를 입을 위험<br>주요 위험요인<br>• 사이버불링(Cyber-bullying)<br>• 온라인 괴롭힘(Harassment)<br>• 혐오 표현, 협박, 모욕<br>• 알고리즘에 의한 과몰입, 과의존 |

자료: 호주 온라인안전법(OSA), 영국 Ofcom의 Content, Contact, Conduct(3C) 위험분류체계

© 2026. (디지털산업정책연구소) all rights reserved.




## 1-2. 현행 국내 아동 · 청소년보호 체계의 한계

- **청소년 보호 책임자 지정**
  - 일일 평균 이용자 수 또는 매출액 등이 일정 기준*에 해당하거나, 청소년유해매체물을 제공·매개하는 정보통신서비스 제공 사업자는 청소년 보호 책임자를 지정해야 함(「정보통신망법」 제42조의3 제1항)

    * 전년도 말 직전 기준 3개월 일일 평균 이용자 10만 명 이상 또는 정보통신서비스부문 전년도 매출액 10억 원 이상
  - 「정보통신망법 시행령」 제26조는 청소년보호책임자의 업무를 ① 연 1회 청소년보호계획 수립, ② 청소년유해정보에 대한 청소년 접근 제한 및 관리조치, ③ 정보통신업무 종사자에 대한 청소년 보호교육, ④ 청소년유해정보로 인한 피해상담 및 고충처리, ⑤ 그 밖의 청소년 보호에 필요한 조치로 구체화하고 있음
- **청소년유해정보 관리**
  - 「청소년 보호법」에 따른 청소년유해매체물 결정·확인 및 유통 제한과 「정보통신망법」에 따른 표시·광고의 제한·접근 관리가 연계된 구조
  - 청소년유해매체물에 대한 청소년의 접근 제한 의무를 부여하여 접근 제한 조치를 수행함
- **청소년보호책임자 지정 및 보호업무 이행 확인**
  - 서면 자율점검 및 청소년보호조치 모니터링, 일부 사업자에 대한 현장 점검을 실시
    - 방송미디어통신위원회 자율점검표: 청소년보호책임자 지정, 청소년보호계획 수립·이행, 청소년 접근 제한 및 관리, 종사자 교육, 피해상담과 고충처리 등의 이행여부 확인

© 2026. (디지털산업정책연구소) all rights reserved.




## 1-2. 현행 국내 아동 · 청소년보호 체계의 한계

① **사후 대응 중심 체계의 구조적 공백**
이미 유해매체물로 결정·확인되었거나 유통 중인 정보의 차단·관리에 중점을 두어, 위험을 사전에 식별·평가하는 독립적 절차가 없음

② **신고, 삭제, 차단 위주의 피해 구제 중심 운영**
청소년보호책임자 업무 중 피해상담 · 고충처리 등은 피해가 발생한 뒤 이루어지는 사후 대응에 해당함

③ **서비스 설계 단계의 위험 예방 수단 부재**
신규 서비스 출시, 추천 알고리즘 · 연락 기능 · 자동재생 등 주요 기능 변경 전 아동 위험을 평가하도록 하는 의무가 없음

④ **사업자 자율 안전관리 역량 점검 도구 및 이행 실적 공개, 보고 절차 부재**
사업자의 위험 예방이나 대응 역량을 진단하는 도구가 부재하며, 청소년보호 계획의 이행실적 및 보호조치의 효과를 보고하는 법적 근거 부재함

© 2026. (디지털산업정책연구소) all rights reserved.

02

# 해외 사례 및 시사점

© 2026. (디지털산업정책연구소) all rights reserved.




## 2-1. 호주: 설계에 의한 안전(Safety by Design, SbD) - 설계 기반 예방 모델

- 호주는 **온라인 안전 위원회**(eSafety Commissioner)를 중심으로 예방 · 자율규율 · 감독 · 집행을 연계한 통합 거버넌스 체계를 구축함

- **설계에 의한 안전**(Safety by Design, SbD)을 통해 온라인 안전을 서비스 설계 · 개발 · 운영 전 과정에 내재화하도록 유도함
  - 사업자 책임성, 이용자 권능 · 자율성, 투명성 · 책임성을 3대 원칙으로 제시하고, 사업자가 잠재적 위해를 사전에 식별하여 이에 상응하는 예방 · 완화조치를 마련하도록 함

- 기본 온라인 안전 기대사항(BOSE) · 산업규범(Industry Codes)을 통한 사업자 자율규율과 정부 감독을 연계한 공동규제 체계를 구축함

- 2024년 법 개정을 통해 16세 미만 이용자의 소셜미디어 계정 이용 제한을 플랫폼의 법적 의무로 확대함

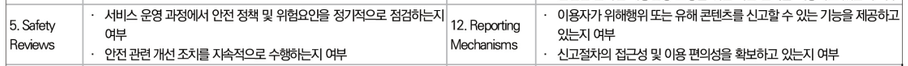
© 2026. (디지털산업정책연구소) all rights reserved.




## 2-1. 호주: 설계에 의한 안전(Safety by Design, SbD) - 설계 기반 예방 모델

- **사업자 자가진단 도구**(Start-up / Enterprise Assessment Tool)를 통해 서비스 특성과 안전관리 역량을 스스로 점검하고 개선하도록 지원함
  - 사업자 규모에 따라 스타트업이나 엔터프라이즈 평가 도구(Start-up / Enterprise Assessment Tool)로 구분하며, 위험 발생 여부를 점수화하기보다 조직 · 절차적 안전관리 역량을 점검함
  - 평가는 **서비스 프로파일링 → 위험 인식 → 자가진단 → 건전성 체크리포트**의 단계로 진행함

**설계에 의한 안전** 사업자 자가진단 도구 흐름




  - 서비스 프로파일링: 이용자 연령, 아동 · 성인 간 상호작용, 콘텐츠 공유 · 소통 기능 등을 확인하여 서비스의 기본 위험 특성을 파악함
  - 위험인식: 온라인 위해(Online Harms) 학습을 통해 유해 콘텐츠 노출과 같은 **콘텐츠 위험**, 성인 · 아동 간 상호작용 등에 따른 **접촉 위험**, 괴롭힘·학대 등 **행위 위험**을 포함한 주요 온라인 위해 가능성을 인식하도록 함
  - 자가진단: Safety Culture, 책임체계, 내부정책, 콘텐츠 중재, 신고 · 대응, 이용자 보호, 투명성 등 14개 영역의 안전관리 체계와 대응 역량을 점검함
  - 결과 환류: 평가 결과를 점수 · 등급이 아닌 건전성 체크리포트(Health Check Report)로 제공하여 취약 영역과 우선 개선사항, 관련 가이드 · 참고자료를 제시함

© 2026. (디지털산업정책연구소) all rights reserved.




## 2-1. 호주: 설계에 의한 안전(Safety by Design, SbD) - 설계 기반 예방 모델

**Safety by Design Assessment Tool 평가항목**

| 단계 | 주요 평가 항목 | 단계 | 주요 평가 항목 |
|---|---|---|---|
| 1. Safety Culture | · 조직 내 온라인 안전 관련 정책 또는 원칙을 보유하고 있는지 여부<br>· 이용자 안전을 조직의 핵심 가치로 반영하고 있는지 여부<br>· 온라인 안전과 관련된 조직문화가 형성되어 있는지 여부 | 8. Internal Triaging | · 신고 또는 위해사건 발생 시 사안의 심각도에 따라 우선순위를 분류하는 절차를 운영하는지 여부<br>· 위해유형별 대응 프로세스를 마련하고 있는지 여부 |
| 2. Safety Leads | · 온라인 안전을 담당하는 책임자 또는 전담 인력을 지정하고 있는지 여부<br>· 안전 관련 의사결정 및 책임소재가 명확하게 규정되어 있는지 여부 | 9. Reporting to Authorities | · 법집행기관 또는 관계기관에 신고가 필요한 경우 관련 절차를 보유하고 있는지 여부<br>· 이용자 안전과 관련된 사건 발생 시 외부기관과 협력체계를 구축하고 있는지 여부 |
| 3. Staff Training | · 임직원 및 계약직 근로자를 대상으로 온라인 안전 관련 교육을 실시하는지 여부<br>· 불법 · 유해 콘텐츠를 처리하는 인력에 대한 지원체계(상담 등)를 운영하고 있는지 여부 | 10. Community Standards | · 이용자 행동규범을 마련하고 있는지 여부<br>· 허용행위 및 금지행위를 이용자에게 명확히 고지하고 있는지 여부 |
| 4. Internal Policies | · 이용자 보호를 위한 내부 정책 및 절차를 보유하고 있는지 여부<br>· 온라인 위해 발생 시 대응 절차를 마련하고 있는지 여부 | 11. Social Contracts | · 회원가입 단계에서 서비스 제공자와 이용자가 준수하여야 할 책임과 의무를 명시하고 있는지 여부<br>· 안전한 온라인 이용환경 조성을 위한 기본 원칙을 제공하고 있는지 여부 |
| 5. Safety Reviews | · 서비스 운영 과정에서 안전 정책 및 위험요인을 정기적으로 점검하는지 여부<br>· 안전 관련 개선 조치를 지속적으로 수행하는지 여부 | 12. Reporting Mechanisms | · 이용자가 위해행위 또는 유해 콘텐츠를 신고할 수 있는 기능을 제공하고 있는지 여부<br>· 신고절차의 접근성 및 이용 편의성을 확보하고 있는지 여부 |
| 6. Moderation Practices | · 이용자 게시물 및 활동에 대한 모니터링 또는 중재 절차를 운영하는지 여부<br>· 위해 콘텐츠 또는 위해행위 발생 시 적절한 대응체계를 마련하고 있는지 여부 | 13. User Empowerment | · 이용자가 자신의 온라인 활동을 통제할 수 있는 기능을 제공하고 있는지 여부<br>· 차단, 신고, 개인정보 설정 등 이용자 보호수단을 제공하고 있는지 여부 |
| 7. Illegal Content | · 불법 콘텐츠를 식별 · 신고 · 삭제하기 위한 정책 또는 절차를 운영하는지 여부<br>· 불법 콘텐츠를 발견 시 관련 기관 또는 이해관계자와 협력할 수 있는 체계를 마련하고 있는지 여부 | 14. Transparency and Accountability | · 이용자 보호정책 및 안전조치에 관한 정보를 공개하고 있는지 여부<br>· 안전 관련 의사결정 및 운영에 대한 책임성을 확보하고 있는지 여부 |

© 2026. (디지털산업정책연구소) all rights reserved.

## 2-2. 영국: 아동 위험 평가(CRA, Child Risk Assessment) - 위험기반 평가 의무화 모델

- **영국은 「온라인안전법 2023**(Online Safety Act 2023, OSA)」 기반으로 Ofcom 중심의 아동 온라인 안전체계를 구축하고, 아동이 접근할 가능성이 있는 온라인 서비스에 **아동 위험평가를 법적 의무**로 부과함
- 아동 접근성 평가를 선행하고, 접근 가능성이 확인된 서비스에 아동 위험평가를 연계하는 단계적 구조를 적용함

**아동 위험평가 절차** (Ofcom 권고)

**아동 접근성 평가** (아동 접근 가능성 판단) → **아동 위험평가** (유해 콘텐츠 노출·영향, 위험수준) → **아동 보호 안전조치** (Codes of Practice 준수)

- 아동 위험 평가를 통해 개별 유해 콘텐츠뿐 아니라 **서비스 기능·이용자 상호작용·아동 이용 특성이 위험을 발생·증폭시키는 방식까지 종합적으로 평가**하고, **평가 결과를 실제 보호조치와 연계**함
- **아동 위험평가 결과를 Ofcom의 아동 보호 실무 규범과 연계**하여 구체적인 보호조치로 전환함
- Ofcom의 아동 보호 코드(Codes of Practice) 준수 시 관련 법적 의무를 준수한 것으로 간주하며, 동등한 보호 수준을 확보하는 대안적 조치도 허용함

© 2026. (디지털산업정책연구소) all rights reserved.




## 2-2. 영국: 아동 위험 평가(CRA, Child Risk Assessment) - 위험기반 평가 의무화 모델

- 아동 접근성 평가를 통해 아동의 접근 가능성이 확인된 서비스는 Ofcom의 가이드라인과 아동 위험 프로필(Children's Risk Profiles)에 따라 4단계 아동 위험평가를 수행함
  - **1단계) 위험 식별**: **최우선 순위, 우선 순위, 미지정 유해 콘텐츠를 확인**하고, 서비스 특성 · 기능을 바탕으로 **콘텐츠**(Content) · **접촉**(Contact) · **행위**(Conduct) **위험을 포함한 아동 위해요인을 식별**함

| 최우선 순위 콘텐츠 (PPC) | 우선 순위 콘텐츠 (PC) | 미지정 유해 콘텐츠 (NDC) |
|---|---|---|
| 포르노·자살·자해·섭식장애 조장 | 혐오·폭력·괴롭힘·위험 챌린지 | 지정 유형 외 중대 위해 콘텐츠 |

  - **2단계) 위험 평가:** 각 위해요인에 대해 아동의 노출 가능성과 피해 영향을 평가하고, 서비스 기능 및 아동 이용 특성을 종합하여 위험수준을 산정함

| 구분 | 평가항목 |
|---|---|
| 서비스 특성(Service Characteristics) | • 서비스 유형, 이용자 규모, 아동 이용 가능성, 이용자 간 상호작용 기능, 콘텐츠 공유 기능 |
| 위험요인(Risk Factors) | • Content Risk, Contact Risk, Conduct Risk |
| 기능별 위험성(Functional Risks) | • 추천시스템, 검색기능, 라이브 스트리밍, 다이렉트 메시지, 콘텐츠 공유 기능, 그룹 및 커뮤니티 기능 |
| 아동 이용 특성(Child User Factors) | • 이용자 연령, 이용자 취약성, 아동 이용 비중, 위해 노출 가능성 |
| 위험완화 조치(Mitigation Measures) | • 연령확인 조치, 콘텐츠 관리조치, 신고 및 차단 기능, 이용자 보호설정, 모더레이션 체계 |

  - **3단계) 조치 결정·실행:** 위험수준에 따라 연령확인, 콘텐츠 관리, 신고·차단, 이용자 보호설정, 모더레이션 등 적절한 위험완화 조치를 결정·이행함
  - **4단계) 검토·업데이트:** 평가 결과와 조치의 효과를 지속적으로 점검하고, 최소 연 1회 및 서비스의 중대한 변경 시 위험평가를 검토·업데이트함

© 2026. (디지털산업정책연구소) all rights reserved.




## 2-3. EU: 디지털서비스법(DSA, Digital Services Act) - 시스템 위험 기반 법적 규제 모델

- EU는 「디지털서비스법(Digital Services Act, DSA)」을 기반으로 플랫폼의 개별 콘텐츠 규제를 넘어 **시스템 위험**(Systemic Risk)**을 사전에 평가 · 관리하는 규제체계**를 구축함
  - EU 집행위원회와 회원국 디지털서비스조정관이 역할을 분담하며, 초대형 온라인 플랫폼 (Very Large Online Platform, VLOP) · 검색엔진(Very Large Online Search Engine, VLOSE)에 강화된 위험관리 의무를 부과함
- **시스템 위험평가는 콘텐츠 자체뿐 아니라 플랫폼 설계 전반을 평가 대상에 포함함**
  - 추천 시스템, 광고 시스템, 콘텐츠 조정, 이용약관 집행, 데이터 처리 등이 시스템 위험에 미치는 영향을 종합적으로 검토함
  - 특히 DSA 제34조는 아동의 권리와 미성년자의 신체적 · 정신적 안녕에 대한 부정적 영향을 핵심 위험요소로 포함함
- **위험평가 결과를 구체적인 완화조치로 연계**하는 **'평가–조치' 구조**를 법적으로 의무화함
  - 위험평가(제34조) → 위험완화조치(제35조)를 연계하여 서비스 설계, 추천·광고 시스템, 콘텐츠 조정 절차 등을 개선하도록 함

**위험평가–완화조치 연계 및 미성년자 보호**

**제34조 시스템 위험평가** (아동권리 등 부정적 영향의 식별·분석·평가) → **제35조 완화조치** (합리적·비례적·효과적 완화조치 이행)

© 2026. (디지털산업정책연구소) all rights reserved.




## 2-3. EU: 디지털서비스법(DSA, Digital Services Act) - 시스템 위험 기반 법적 규제 모델

- 사업자의 **자체 위험평가에 독립 감사와 투명성 보고를 결합**하여 위험관리의 책임성과 외부 검증을 강화함
  - VLOP · VLOSE에 자체평가 → 위험완화 → 독립감사 → 투명성 보고가 연계된 관리 체계를 적용함
  - 미성년자가 접근 가능한 플랫폼에는 제28조에 따라 비례적 보호조치와 미성년자 대상 프로파일링 광고 금지 등을 별도로 적용함

**제28조 미성년자 보호**

**비례적 보호조치** | **프로파일링 광고 금지** | **연령적합 이용약관**

  - EU는 위험평가 의무는 법제화했으나 위험평가 방법론(진단도구)은 표준화하지 않아, 영국, 호주와 달리 위험평가를 지원하는 별도의 표준화된 진단도구는 제공하지 않음

© 2026. (디지털산업정책연구소) all rights reserved.

## 2-4. 해외 사례 비교 및 국내 시사점

- 국내에 비해 콘텐츠 사후 삭제나, 차단에 대한 강력한 청소년 보호 체계를 가지고 있지 못했던 해외에서는 **설계·운영 단계의 사전 예방적 위험관리**로 전환하면서 **위험기반 접근·사업자 책임성 · 비례성 원칙**을 공통으로 적용함
- 국내는 유해 콘텐츠의 삭제 차단 등 콘텐츠 중심으로 구축되어 왔으나, 새로운 유형의 디지털 위험에 대응하기 위해서는 서비스 설계 운영 단계에서 위험을 사전에 파악 예방하는 접근으로의 보완이 필요
- **국내 적용 가능 요소**: 자가진단 도구 구조(호주), 위험유형 분류 및 단계적 평가체계(영국), 시스템 위험 개념과 평가–완화 연계 구조(EU)를 국내 환경에 선택적으로 적용할 수 있음

| 구분 | 호주<br>Safety by Design | 영국<br>Child Risk Assessment | EU<br>Digital Services Act |
|---|---|---|---|
| **규제 성격** | 자율 · 공동규제 (권고) | 법적 의무 (위험기반) | 법적 규제 (시스템 위험) |
| **평가 주체** | 사업자 자가진단 | 사업자 (Ofcom 감독) | 초대형 플랫폼 (VLOP) |
| **적용 대상** | 전 온라인 서비스 | 아동접근 가능 서비스 | VLOP, VLOSE 중심 |
| **핵심 특징** | SbD 3원칙, 자가진단 도구 | 콘텐츠위험 3단계 분류 | 위험평가(§34) → 완화(§35) |
| **감독·집행** | eSafety Commissioner | Ofcom | EU 집행위 · DSC |

© 2026. (디지털산업정책연구소) all rights reserved.


# 03
# 플랫폼 사전위험평가 체계의 국내 적용 방향 및 과제

© 2026. (디지털산업정책연구소) all rights reserved.




## 3-1. 사전위험평가 체계의 필요성과 기본 방향

- **디지털 플랫폼 위험의 공통화**
  - 이용자 간 상호작용, 개인정보 활용, 서비스 설계 등으로 발생하는 아동 위험은 특정 국가에 국한되지 않는 플랫폼의 구조적 위험
  - 국내 아동 역시 글로벌 플랫폼을 포함한 유사한 디지털 서비스 환경을 이용하고 있어, 해외에서 확인된 위험유형과 위험관리 경험을 국내에서도 검토할 필요가 있음

- **글로벌 스탠다드와의 정합성 확보**
  - 호주, 영국, EU 등은 사후 콘텐츠 규제 체계가 미흡한 상황에서 유해 콘텐츠 관리의 상당 부분을 사업자에게 위임하는 방식으로, 서비스의 설계, 운영 단계에서 위험을 사전에 식별, 평가, 관리하는 체계를 발전시키고 있음
  - 국내는 청소년유해정보의 삭제 차단 등 콘텐츠 중심의 보호체계를 구축해 왔으나, 플랫폼의 구조적이고 복합적 위험에 대응하기 위해 서비스 설계 운영 단계에서 위험을 사전에 파악 예방하는 체계로의 보완이 필요
  - 특히 국내 아동의 글로벌 플랫폼 이용 비중이 높은 환경을 고려할 때, 해외에서 확산(글로벌 스탠다드)되고 있는 사전적 위험기반 접근과의 정합성을 고려한 국내 보호체계 구축이 필요
  - 각 사례는 자율진단, 법적 위험평가, 시스템 위험관리 등 서로 다른 접근을 통해 평가기준과 운영경험을 축적하고 있어 국내 제도 구축의 참고 모델로 활용이 가능

© 2026. (디지털산업정책연구소) all rights reserved.




## 3-1. 사전위험평가 체계의 필요성과 기본 방향

- **국내 사전위험평가 체계의 적용 가능성**
  - 국내에는 아동의 플랫폼 이용 과정에서 발생할 수 있는 위험을 서비스 단위에서 체계적으로 식별, 평가하는 독립적 위험평가 체계가 아직 부재
  - 해외에서 공통으로 확인되는 위험유형, 평가절차, 안전관리 요소를 활용하면 국내 평가체계 구축의 시행착오를 줄이고 평가의 체계성과 객관성을 확보할 수 있음
  - 해외 사례의 **공통원칙과 실효적 요소를 추출**하여 **국내 법제도 및 플랫폼 환경에 맞게 재구성**할 필요가 있음

- **해외에서 도출할 수 있는 핵심 요소**
  - (호주) 사업자가 스스로 안전관리 수준을 점검하는 자가 진단형 평가도구
  - (영국) 아동 접근 가능성을 선행 판단하고 위험을 평가하는 단계적 위험평가 절차
  - (EU) 위험평가 결과를 실제 완화조치와 연결하는 위험평가 – 완화의 연계구조
  - (UNICEF) 사업자 중심 평가를 보완하는 아동 권리, 아동참여 관점

✓ 플랫폼의 구조적 위험에 대응하기 위한 **위험기반 접근**은 해외 주요 사례에서 공통적으로 확인되는 정책 흐름
✓ 해외 제도를 그대로 도입하기보다는 **검증된 공동 요소를 국내 플랫폼 환경에 맞게 선별, 재구성하는 접근이 필요**

© 2026. (디지털산업정책연구소) all rights reserved.

## 3-2. 사전위험평가 체계 적용 방향

- 사전위험평가 체계의 국내 도입을 위해 고려할 주요 축으로 아래 네 가지를 선정
  **① 평가 대상 및 적용 범위, ② 평가 영역 및 핵심지표*, ③ 평가 절차 및 운영방식, ④ 운영체계 및 거버넌스**
  * 평가 영역 및 핵심지표에서는 해외 사례와 연구 결과를 토대로 사전위험평가 체크리스트(안)을 구성 · 제안함

**주요 고려사항**

**① 평가 대상 및 적용 범위**
누구에게 어느 수준의 평가 의무를 부과할 것인가?

**② 평가 영역 및 핵심지표**
무엇을 평가할 것인가?

**③ 평가 절차 및 운영 방식**
어떻게 평가할 것인가?

**④ 운영체계 및 거버넌스**
누가 기준을 마련하고, 이행을 지원 점검할 것인가?

© 2026. (디지털산업정책연구소) all rights reserved.




## 3-2. 사전위험평가 체계 도입 방안 - ① 평가 대상 및 적용범위

- **서비스 유형**과 **아동 접근 가능성**을 기준으로 사전위험평가 적용대상을 설정
  - 아동이 실질적으로 접근 이용할 가능성이 있는 플랫폼을 기본 적용대상으로 검토

- 모든 사업자에게 동일한 의무를 부과하기보다 **서비스 위험 특성, 아동 노출 수준, 사업자 규모에 따라 평가범위와 수준을 차등화**
  - 기본평가와 심화평가를 구분하여 위험수준에 비례한 평가의무 체계 구축

- 사전위험평가 의무화는 모든 플랫폼에 동일한 부담을 부과하는 방식이 아니라, 아동접근 가능성과 서비스 위험수준에 따라 차등적으로 적용이 필요

© 2026. (디지털산업정책연구소) all rights reserved.






## 3-2. 사전위험평가 체계 도입 방안 - ② 평가 영역 및 핵심지표

- '유해콘텐츠'에서 **'아동의 서비스 이용과정'** 중심으로 평가범위 확대
- **알려진 위험에 대해 공통으로 평가해야 할 최소 위험 영역과 핵심지표**를 설정
- 다만, 서비스별 위험 특성을 고려하여 세부 평가항목은 탄력적으로 적용

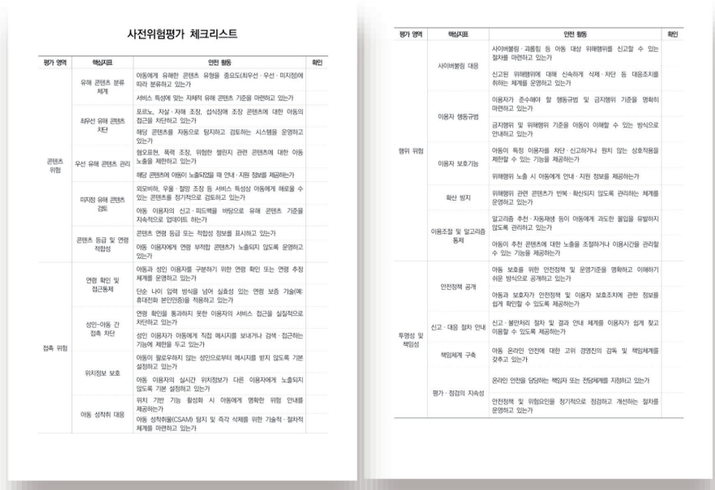
| 평가 영역 | 핵심 지표(안) | 대표 점검 문항(예시) |
|---|---|---|
| **콘텐츠 위험**<br>Content Risk | • 유해 콘텐츠 분류 체계 운영<br>• 최우선·우선 · 미지정 유해 콘텐츠 관리<br>• 콘텐츠 등급 · 연령 적합성 관리 | 포르노, 자살·자해 조장, 섭식장애 조장 콘텐츠에 대한 아동의 접근을 차단하고 있는가 |
| **접촉 위험**<br>Contact Risk | • 연령 확인 · 접근통제<br>• 성인 · 아동 간 접촉 차단<br>• 위치정보 보호<br>• 아동 성착취 대응 | 성인 이용자가 아동에게 직접 메시지를 보내거나 검색·접근하는 기능에 제한을 두고 있는가 |
| **행위 위험**<br>Conduct Risk | • 차단 · 신고 및 상호작용 제한<br>• 위해행위 대응 안내<br>• 이용조절·추천 관리 | 아동이 추천 콘텐츠에 대한 노출을 조절하거나 이용시간을 관리할 수 있는 기능을 제공하는가 |
| **투명성 및 책임성**<br>Transparency and Accountability | • 안전정책 공개<br>• 신고·대응 절차 안내<br>• 책임체계 구축<br>• 평가·점검의 지속성 | 아동 안전에 대한 고위 경영진의 감독·책임체계를 갖추고 있는가 |

© 2026. (디지털산업정책연구소) all rights reserved.




## 3-2. 사전위험평가 체계 도입 방안 - ② 평가 영역 및 핵심지표

사전위험평가 체크리스트 (예시)



사전위험평가 체크리스트

| 평가 영역 | 핵심지표 | 안전 항목 | 확인 |
|---|---|---|---|
| 콘텐츠 위험 | 유해 콘텐츠 분류 체계 | 아동에게 유해한 콘텐츠 유형을 중요도(최우선 · 우선 · 미지정)에 따라 분류하고 있는가 | |
| | 최우선 유해 콘텐츠 식별 | 서비스 특성에 맞는 자해위 유해 콘텐츠 기준을 마련하고 있는가 | |
| | | 포르노, 자살 · 자해 조장, 섭식장애 조장 콘텐츠에 대한 아동의 접근을 차단하고 있는가 | |
| | 우선 유해 콘텐츠 관리 | 해당 콘텐츠를 자동으로 탐지하고 검토하는 시스템을 운영하고 있는가 | |
| | | 썸오프된 목록 조정, 위험한 챌린지 관련 콘텐츠에 대한 아동 노출을 제한하고 있는가 | |
| | 미지정 유해 콘텐츠 검토 | 해당 콘텐츠에 아동이 노출(조회)될 때 안내 · 차단 정보를 제공하는가 | |
| | | 외모비하, 무물 · 집단 조장 등 서비스 특성상 아동에게 해로울 수 있는 콘텐츠를 정기적으로 검토하고 있는가 | |
| | | 아동 이용자의 신고 · 피드백을 바탕으로 유해 콘텐츠 기준을 지속적으로 업데이트 하는가 | |
| | 콘텐츠 등급 및 연령 적합성 | 콘텐츠 연령 등급 또는 적합성 정보를 표시하고 있는가 | |
| | | 아동 이용자에게 연령 부적합 콘텐츠가 노출되지 않도록 운영하고 있는가 | |
| 접촉 위험 | 연령 확인 및 접근통제 | 아동과 성인 이용자를 구분하기 위한 연령 확인 또는 연령 추정 체계를 운영하고 있는가 | |
| | | 단순 나이 입력 방식을 넘어 실효성 있는 연령 보증 기술(예: 휴대전화 본인인증)을 적용하고 있는가 | |
| | 성인-아동 간 접촉 차단 | 연령 확인을 통과하지 못한 이용자의 서비스 접근을 실질적으로 차단하고 있는가 | |
| | | 성인 이용자가 아동에게 직접 메시지를 보내거나 검색 · 접근하는 기능에 제한을 두고 있는가 | |
| | 위치정보 보호 | 아동이 팔로우하지 않는 성인으로부터 메시지를 받지 않도록 기본 설정하고 있는가 | |
| | | 아동 이용자의 실시간 위치정보가 다른 이용자에게 노출되지 않도록 기본 설정하고 있는가 | |
| | 아동 성착취 대응 | 위치 기반 기능 활성화 시 아동에게 명확한 위험 안내를 제공하는가 | |
| | | 아동 성착취물(CSAM) 탐지 및 즉각 삭제를 위한 기술적 · 절차적 체계를 마련하고 있는가 | |

| 평가 영역 | 핵심지표 | 안전 항목 | 확인 |
|---|---|---|---|
| 행위 위험 | 사이버불링 대응 | 사이버불링 · 괴롭힘 등 아동 대상 위해행위를 신고할 수 있는 절차를 마련하고 있는가 | |
| | | 신고된 위해행위에 대해 신속하게 삭제 · 차단 등 대응조치를 취하는 체계를 운영하고 있는가 | |
| | 이용자 행동규범 | 이용자가 준수해야 할 행동규범 및 금지행위 기준을 명확히 마련하고 있는가 | |
| | | 금지행위 및 위해행위 기준을 아동이 이해할 수 있는 방식으로 안내하고 있는가 | |
| | 이용자 보호기능 | 아동이 특정 이용자를 차단 · 신고하거나 원치 않는 상호작용을 제한할 수 있는 기능을 제공하는가 | |
| | | 위해행위 노출 시 아동에게 안내 · 지원 정보를 제공하는가 | |
| | 확산 방지 | 위해행위 관련 콘텐츠가 반복 · 확산되지 않도록 관리하는 체계를 운영하고 있는가 | |
| | 이용조절 및 알고리즘 통제 | 알고리즘 추천 · 자동재생 등이 아동에게 과도한 몰입을 유발하지 않도록 관리하고 있는가 | |
| | | 아동이 추천 콘텐츠에 대한 노출을 조절하거나 이용시간을 관리할 수 있는 기능을 제공하는가 | |
| 투명성 및 책임성 | 안전정책 공개 | 아동 보호를 위한 안전정책 및 운영기준을 명확하고 이해하기 쉬운 방식으로 공개하고 있는가 | |
| | | 아동과 보호자가 안전정책 및 이용자 보호조치에 관한 정보를 쉽게 확인할 수 있도록 제공하는가 | |
| | 신고 · 대응 절차 안내 | 신고 · 불만처리 절차 및 결과 안내 체계를 이용자가 쉽게 찾고 이용할 수 있도록 제공하는가 | |
| | 책임체계 구축 | 아동 온라인 안전에 대한 고위 경영진의 감독 및 책임체계를 갖추고 있는가 | |
| | 평가 · 점검의 지속성 | 온라인 안전을 담당하는 책임자 또는 전담체계를 지정하고 있는가 | |
| | | 안전정책 및 위험요인을 정기적으로 점검하고 개선하는 절차를 운영하고 있는가 | |

## 3-2. 사전위험평가 체계 적용 방향 - ③ 평가 절차 및 운영방식

| | |
|---|---|
| Step 1. **서비스 프로파일링** | • 서비스 유형 및 주요 기능 파악, 아동 접근 가능성 확인, 서비스별 주요 위험유형 파악, 기본심화 등 적용 평가수준 결정 |
| Step 2. **위험인식 및 식별** | • 사업자가 자사 서비스에서 발생 가능한 위험유형과 위험 시나리오를 사전에 이해<br>• 콘텐츠 위험뿐만 아니라 접촉위험, 과몰입, 알고리즘에 따른 반복노출 등 실사용 환경에서 발생 가능한 위험까지 폭넓게 인식 |
| Step 3. **위험평가 수행** | • 공통평가영역 및 서비스별 세부지표를 활용하여 위험수준 평가<br>• 위험요인과 기존 안전조치의 적정성 점검 |
| Step 4. **위험완화 및 개선** | • 평가결과를 통해 필요한 안전조치 마련<br>• 위험수준에 상응하는 서비스 기능 운영 개선 |
| Step 5. **기록 및 점검, 재평가** | • 평가결과와 위험완화조치를 기록<br>• 일정 주기 또는 중대한 서비스 변경 시 재평가 |

➢ 평가가 형식적 절차에 그치지 않도록 평가 결과와 위험 완화 조치를 연계

© 2026. (디지털산업정책연구소) all rights reserved.




## 3-2. 사전위험평가 체계 적용 방향 - ③ 평가 절차 및 운영방식

- 알려진 위험을 객관적으로 평가할 수 있는 활용도구로서, 구체화된 **'사전위험평가 수행도구' 마련**이 필요
  - (서비스 정보 입력) 서비스 유형, 기능, 아동 접근 가능성 확인
  - (평가범위 결정) 기본평가 또는 심화평가 적용
  - (위험영역별 평가) 콘텐츠 접촉, 행위, 투명성 및 책임성
  - (위험수준 및 취약영역 확인) 위험요인과 기존 안전조치 적정성 확인
  - (완화조치 및 개선) 필요한 개선조치 도출 및 이행
  - (기록 재평가) 평가결과와 조치를 기록하고 주기적으로 갱신

▪ **수행도구의 역할**
  - 사업자가 의무를 실제로 이행할 수 있도록 표준적인 평가기준과 절차 제공
  - 형식적인 규제수준 체크리스트보다 실질적인 위험 발견 완화에 초점
  - 사업자 규모와 서비스 특성에 따라 문항과 평가수준을 차등 적용
  - 평가결과를 서비스 개선 및 재평가와 연계

➢ 평가의무만 부과하는 것이 아니라, 사업자가 무엇을 어떻게 평가해야 하는지 알 수 있는 표준화된 수행도구를 함께 제공
➢ 의무화의 실효성을 위해 **평가기준, 지표, 절차의 명확성 확보가 중요함**

© 2026. (디지털산업정책연구소) all rights reserved.




## 3-2. 사전위험평가 체계 적용 방향 - ④ 운영체계 및 거버넌스

**공적 기관의 역할**
- 평가기준, 핵심지표, 수행도구의 개발 및 지속적 갱신
- 사업자가 위험평가 의무를 이행할 수 있도록 가이드라인 및 기술적 지원 제공
- 새로운 위험과 서비스 변화에 대한 평가기준 보완

**사업자의 책임**
- 자사 서비스에 대한 사전위험평가 수행
- 평가결과에 따른 위험완화조치 마련 이행
- 평가결과 및 개선조치의 기록 관리
- 중대한 서비스 변경이나 위험환경 변화 시 재평가

**감독·검증체계**
- 사업자의 평가 실시 여부와 절차 준수 여부 확인
- 필요시 평가결과 및 위험완화조치에 대한 보고체계 마련
- 고위험 서비스 또는 대규모 플랫폼에 대해서는 강화된 검증방식 검토

**참여환류·거버넌스**
- 정부 · 사업자 · 아동 · 전문가 등이 평가체계 개선에 참여
- 아동참여형 점검 및 실제 적용 결과를 평가지표와 안전기준 개선에 환류

➢ 공적 기관의 역할은 개별 위험을 대신 평가하는 것이 아니라, 공통 평가기준을 마련하고 사업자의 위험관리 책임이 실질적으로 이행되는지 지원 점검하는 것임
➢ 민간단체는 다양한 아동의 실제 이용경험과 보호 필요를 반영하여 안전활동 및 보호조치의 유형을 발굴 제안하고, 아동의 관점이 평가도구와 제도 설계에 반영될 수 있도록 협력 지원하는 역할 필요
➢ 아동의 실제 이용환경을 고려하여, 이미 알려진 위험에 대해 아동이 스스로 대응하고 주도성과 안전을 확보할 수 있도록 구체적인 안전활동 목록을 마련할 필요

© 2026. (디지털산업정책연구소) all rights reserved.




## 3-3. 향후 과제 및 제언

▪ **사전 위험평가 수행 도구 고도화**
  - 플랫폼이 실제 평가를 수행할 수 있는 구체적인 도구를 개발하는 후속 작업 필요
  - 서비스 기능 및 유형, 아동의 이용환경에 따라 적용 가능한 다양한 안전조치 옵션을 발굴하고, 이를 토대로 아동의 실제 이용환경을 반영한 안전설계의 세부 기준을 구체화

▪ **아동 참여구조의 단계적 구현**
  - 유니세프 사례*와 같이 아동 참여 방식을 포함하여, 당사자인 아동 · 청소년의 경험과 관점을 수렴하는 절차 마련

* 「디지털 환경 관련 아동권리 영향평가(Child Rights Impact Assessment in relation to the Digital Environment, D-CRIA)」로서, 디지털 기술을 개발 · 배포 · 사용하는 기업이 아동권리 관점에서 인권 실사를 수행할 수 있는 자발적 평가 프레임워크. 아동·청소년이 직접 참여하는 절차를 핵심 요소로 포함하여 권리 당사자로서의 의견을 수렴토록 함

▪ **거버넌스 구축 및 이해관계자 협력**
  - 정부, 사업자, 시민사회, 학계와 함께 아동을 주요 당사자로 포함하는 협의구조를 마련하여, 다양한 이해관계자의 의견을 도구 개발과 제도 설계에 반영하는 등 제도의 실효성과 수용성 확보가 필요

© 2026. (디지털산업정책연구소) all rights reserved.

# 감사합니다

© 2026. (디지털산업정책연구소) all rights reserved.


## 종합토론

**좌 장**

**김경희 교수**(한림대학교 디지털콘텐츠융합스쿨)

**토 론**

**성욱제 연구위원**(정보통신정책연구원 미디어전략그룹)

**김현기 과장**(초록우산 옹호사업본부)

**한승혜 리더**(한국인터넷기업협회 디지털경제연구원)

**최진응 입법조사관**(국회입법조사처 과학방송통신과)

**김우석 과장**(방송미디어통신위원회 디지털유해정보대응과)

# 종합토론

1 **성욱제 연구위원**

정보통신정책연구원 미디어전략그룹

# 토론문

성욱제 연구위원 (정보통신정책연구원 미디어전략그룹)

오늘 발제는 아동·청소년의 온라인 안전 문제를 기존의 유해 콘텐츠 규제나 사후적 피해구제의 관점에서만 볼 것이 아니라, **서비스를 기획하고 설계하는 단계에서부터 위험을 미리 파악하고 예방하는 방향으로 전환할 필요가 있다**는 점을 잘 보여주셨다고 생각합니다.

저도 이러한 기본적인 문제의식에는 동의합니다. 특히 최근 온라인상의 위험은 단순히 어떤 유해한 콘텐츠가 존재하느냐의 문제만은 아닌 것 같습니다. 추천 알고리즘이나 자동재생, 다이렉트 메시지, 위치정보, 이용자 간 접촉기능처럼 플랫폼의 설계와 기능 자체가 아동·청소년의 위험을 발생시키거나 증폭시킬 수 있기 때문입니다. 발제에서도 이러한 점을 기존 국내 제도의 중요한 한계로 지적하고 있습니다.

저는 오늘 발제와 관련해서 크게 세 가지 정도 말씀드리고 싶습니다.

첫 번째는 **Safety by Design(안전중심설계)이라는 개념을 어떻게 이해할 것인가** 하는 부분입니다. Safety by Design은 호주에서 처음 만들어진 고유한 개념이라기보다는, 제품이나 서비스를 만든 다음 문제가 발생하면 안전장치를 추가하는 것이 아니라, **기획과 설계 단계에서부터 위험을 예상하고 안전을 내재화하자는 보다 일반적인 원칙**으로 보는 것이 적절하다고 생각합니다. 특히 디지털 분야에서 이러한 'by design' 접근이 국제적으로 확산되는 데에는 Privacy by Design의 영향이 상당히 컸다고 볼 수 있습니다. 1990년대 캐나다에서 Privacy by Design이 체계화됐고, 사전 예방, 개인정보 보호를 기본값으로 설정하는 것, 설계 자체에 개인정보 보호를 내재화하는 것, 투명성과 이용자 중심성 등의 원칙이 제시됐습니다. 이후 2016년 EU GDPR 제25조에서 'Data Protection by Design and by Default'가 법제화됐습니다. 호주 eSafety Commissioner의 Safety by Design은 그 이후인 2018년부터 본격적으로 추진됩니다. GDPR이 호주의 Safety by Design에 영향을 줬다고 단정하기는 어렵겠지만, Privacy by Design, Security by Design 등과 같이 **이미 확산돼 있던 사전예방적 설계 원칙이 온라인 안전 분야로 확장됐다고 이해하는 것은 충분히 가능하다**고 생각합니다. 호주가 중요한 이유는 Safety by Design이라는 말을 처음 사용해서라기보다, 온라인 안전 분야에서 이를 구체적인 정책 프레임워크로 발전시킨 대표적인 선도국가라는 점에 있다고 생각합니다.

두 번째로 말씀드리고 싶은 부분은 **호주와 영국, EU의 모델을 어떻게 비교할 것인가**입니다. 세 모델은 디테일에는 상당한 차이가 있지만 큰 방향에서는 공통점이 있습니다. 한마디로 말하면 **"문제가 발생한 뒤 처리하지 말고, 서비스의 설계와 운영 단계부터 위험을 생각하라"**는 것입니다. 그런데 실제 규제방식에서는 상당한 차이가 있습니다. 호주의 Safety by Design Assessment Tool은 기본적으로 **사업자가 스스로 자신의 안전 관리 체계를 진단하고 개선하도록 돕는 자율적인 도구**입니다. 그래서 Safety Culture, Safety Lead, 직원교육, 내부정책, 위험평가, 콘텐츠 모더레이션, 신고체계, 이용자 권한, 투명성과 책임성 등 굉장히 폭넓은 항목을 봅니다. **"우리 회사가 안전을 중요한 가치로 생각하고 있는가, 그리고 그것을 조직과 제품개발 전반에 제대로 반영하고 있는가"**에 가깝습니다.



반면 영국의 Online Safety Act는 훨씬 법적인 접근입니다. 아동이 이용할 가능성이 있는 서비스에 대해서는 Children's Risk Assessment(CRA)를 실시하도록 하고, 어떤 콘텐츠 위험이 있는지뿐 아니라 추천시스템, 검색, 라이브스트리밍, 다이렉트 메시지 같은 **구체적인 서비스 기능이 아동에게 어떤 위험을 발생시키는지 평가하도록 합니다.** 그리고 위험이 확인되면 그것을 실제로 완화하도록 하는 의무(보호 조치)까지 연결됩니다. EU DSA도 규제 기법으로 보면 호주보다는 영국에 더 가깝다고 생각합니다. 특히 초대형 플랫폼 등에 대해서 **시스템적 위험을 평가하고, 그 결과에 따라 위험완화조치를 취하도록 하는 구조**입니다. 필요한 경우 서비스의 설계나 기능, 알고리즘 등을 변경하는 것까지 포함합니다. 발제에서도 EU의 특징을 위험평가와 위험완화가 법적으로 연계되어 있다는 점으로 잘 정리하고 있습니다. 시기적으로 보면 호주가 2018년 Safety by Design을 본격적으로 추진했고, 영국은 2019년 Online Harms White Paper, EU는 2020년 DSA 제안, 2022년 DSA 채택, 그리고 영국은 2023년 Online Safety Act 제정으로 이어집니다. **2010년대 후반 이후 온라인 규제의 패러다임 자체가 '사후적인 콘텐츠 규제'에서 '사전적이고 시스템적인 위험관리'로 이동한 흐름**으로 이해하는 것이 적절하다고 생각합니다.

세 번째가 제가 오늘 가장 말씀드리고 싶은 부분인데요. **그렇다면 이것을 우리나라에는 어떻게 적용할 것인가** 하는 문제입니다. 발제에서는 호주의 자가진단 도구, 영국의 단계적 위험평가, EU의 평가와 완화조치 연계 구조를 선택적으로 결합하는 방안을 제안하고 있습니다. 기본적인 방향에는 저도 동의합니다. 다만 한 가지 고민이 있습니다.

호주의 Safety by Design Assessment와 영국이나 EU의 법적 위험평가는 성격이 조금 다릅니다. 예를 들어 호주에서 이야기하는 Safety Culture나 Safety Lead, 직원교육은 기업이 안전을 얼마나 중요한 조직적 가치로 받아들이고 있는가를 보는 것입니다. 반면 영국 CRA에서 "추천 알고리즘이 13세 아동에게 어떤 위험을 발생시키는가", "성인이 아동에게 DM을 보낼 수 있는 기능이 어떤 위험을 만드는가"를 평가하는 것은 특정 서비스와 기능의 위험수준을 평가하는 것입니다. 전자는 조직의 안전관리 역량에 관한 것이고, 후자는 서비스의 위험에 관한 것입니다. 둘 다 중요하지만 이것을 하나의 법정 위험평가 체크리스트에 모두 집어넣으면 무엇을 평가하는 제도인지 조금 불분명해질 수 있습니다.

그래서 저는 국내에서는 오히려 두 층위로 구분하는 방안을 검토해 보면 어떨까 생각합니다. 첫 번째 층위는 **Safety by Design을 온라인 서비스 전반에 적용되는 기본원칙으로 설정하는 것**입니다. 기업의 안전문화, 경영진의 책임, 안전담당자, 직원교육, 내부정책, 이용자의 신고·차단수단, 투명성 같은 부분입니다. 이런 영역에는 호주의 Safety by Design Assessment Tool이 좋은 참고가 될 수 있습니다. 그리고 이 부분은 처음부터 강한 법적 의무를 부과하기보다는 **자가진단이나 가이드라인, 공동규제 방식으로 시작할 수도 있을 것입니다. 두 번째 층위에서는 아동이 실질적으로 이용하는 일정 규모 또는 일정 위험수준 이상의 서비스에 대해서 별도의 아동 위험평가를 법적인 의무로 부과하는 방안**을 검토할 수 있습니다. 여기에서는 콘텐츠 위험뿐 아니라 접촉위험, 행위위험, 추천 알고리즘, 자동재생, DM, 위치정보, 검색기능 등이 연령별 아동에게 어떤 영향을 미치는지를 평가하도록 하는 것입니다. 그리고 여기서 중요한 것은 **평가만 하고 끝나서는 안 된다는 것**입니다. 위험이 발견됐다면 추천시스템을 바꾸거나, 기본설정을 변경하거나, 성인과 아동의 접촉기능을 제한하거나, 신고·차단 기능을 개선하는 등 실제 위험완화조치로 연결돼야 합니다. 이런 점에서는 EU의 'Risk Assessment → Risk Mitigation' 구조가 좋은 참고가 될 수 있다고 생각합니다. 발제에서도 이 부분을 중요

하게 지적하고 있습니다.

다시 말해, Safety by Design을 상위의 기본 원칙으로 두고, 그 아래에 법적인 Child Risk Assessment를 두고, 다시 그 평가 결과를 실제 Risk Mitigation으로 연결하는 구조입니다. 이렇게 보면 호주와 영국·EU의 모델을 서로 다른 층위에서 절충 가능하다고 생각합니다.

그리고 마지막으로 한 가지 더 강조하고 싶은 부분이 있습니다. 아동·청소년의 온라인 안전을 강화하는 과정에서 안전이 곧 접근제한을 의미해서는 안 된다는 점입니다. 사업자 입장에서는 위험을 가장 쉽게 줄이는 방법이 아동의 서비스 이용 자체를 막거나 강한 연령 확인을 적용하는 것일 수도 있습니다. 하지만 아동에게는 보호받을 권리뿐 아니라 정보에 접근하고, 표현하고, 다른 사람과 소통하고, 디지털 사회에 참여할 권리도 있습니다. 따라서 위험평가 과정에서도 단순히 "얼마나 위험한가"만 볼 것이 아니라, 위험을 줄이기 위한 조치가 아동의 권리나 최선의 이익을 과도하게 제한하지 않는가도 함께 볼 필요가 있다고 생각합니다. 그런 의미에서 발제에서 UNICEF의 아동권리 영향평가나 아동참여를 보완적인 요소로 제시한 부분도 의미가 있다고 봅니다.

정리하겠습니다. 오늘 발제의 가장 중요한 의미는 아동·청소년 온라인 보호정책의 질문을 "어떤 유해 콘텐츠를 삭제할 것인가"에서 "어떤 온라인 환경을 설계할 것인가"로 전환하고 있다는 점이라고 생각합니다. 다만 국내 제도를 설계할 때, 호주, 영국, EU의 사례를 병렬적으로 연결시키기보다는 Safety by Design은 폭넓은 기본 원칙으로, 실질적으로 아동에게 높은 위험을 초래할 수 있는 서비스에는 비례적인 위험평가와 위험 완화 의무를 부과하는 방식이 어떨까 싶습니다. 또한 아동·청소년을 디지털 환경으로부터 배제하는 것이 아니라, 아동·청소년이 보다 안전하게 디지털 환경에 접근하고 참여할 수 있도록 만드는 것이 궁극적 목적이어야 한다는 점도 다시 한번 강조하고 싶습니다.

이상으로 토론을 마치겠습니다. 감사합니다.

## 종합토론

**2 김현기 과장**

초록우산 옹호사업본부

# 토론문

김현기 과장(초록우산 옹호사업본부)

## 1. 들어가며: 아동이 마주하는 위험

*“유튜브에 폭력적이고 자극적인 영상이 화면에 보일 때가 있어 불편해요.”*
*“뉴스나 웹툰을 자주 보는데, 선정적인 광고가 너무 많이 떠서 보기 민망해요.”*

지난해 초록우산 아동권리옹호단[1]이 펴낸 제7차 유엔아동권리협약 심의 대응을 위한 아동보고서[2] 「클릭 후, 우리가 겪는 진짜 이야기」에 담긴 아동들의 말입니다. 아동들은 본인의 의사와 상관없이, 한 번 시청한 영상과 유사한 콘텐츠가 알고리즘을 통해 추천되면서 점점 더 수위가 높은 콘텐츠에 노출된다고 말하고 있습니다.

이는 아동보고서를 집필한 일부 아동만의 이야기가 아닙니다. 초록우산이 숏폼 플랫폼 이용 경험이 있는 전국 만 14세 이상 아동 1,000명을 대상으로 실시한 설문조사[3]에서도, 절반이 넘는 53.4%가 유해콘텐츠에 노출된 경험이 있다고 응답했습니다. 그리고 그중 80.3%는 추천 알고리즘을 통해 유해콘텐츠를 접한 것으로 나타났습니다.

설문조사의 숫자를 넘어, 아동이 디지털 플랫폼에서 마주하는 위험이 구체적으로 어떤 것들인지 직접 확인해 봤습니다. 만 14세 아동을 기준으로 설정한 가상계정을 3개 숏폼 플랫폼에서 각각 약 1시간씩 관찰했습니다.

| 구분 | 내용 |
|---|---|
| 조사 시점 | 2025년 12월 |
| 계정 설정 | 2011년생(조사 당시 만 14세) 기준 가상계정 12개 |
| 유형 구분 | 유해콘텐츠 유형(성, 자살·자해, 섭식장애, 마약·도박) 별로 분리 운영 |
| 대상 플랫폼 | 숏폼 플랫폼 3개 |
| 관찰 방법 | 간접·직접 키워드 탐색, 추천 피드 관찰, 추천 피드 해시태그 추적 등 |
| 관찰 시간 | 계정당 약 1시간 |
| 기록 방식 | 화면녹화 |

1) 자신과 타인의 권리를 이해하고, 아동의 권리를 증진시키기 위해 필요한 법과 정책을 개선하는 데 자신의 목소리를 내는 아동참여조직

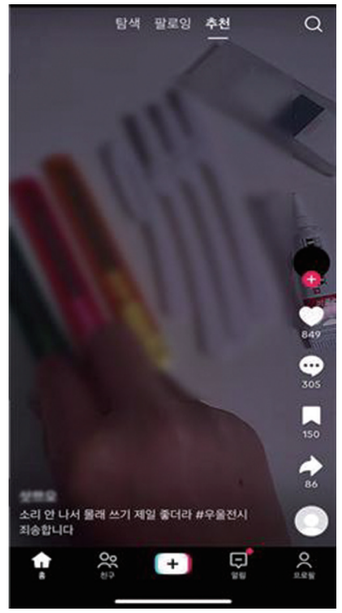
2) 유엔아동권리위원회가 유엔아동권리협약 이행 상황을 심의하는 과정에서 당사국 아동의 목소리를 듣기 위해 제출되는 문서

3) 조사기간: 2025년 12월 22일~29일 / 조사항목: 플랫폼 이용형태, 유해콘텐츠(성, 자살·자해, 섭식장애, 마약·도박) 노출 경험 및 영향, 플랫폼 아동보호장치 인지도·이용경험·효과성 등 총 17문항

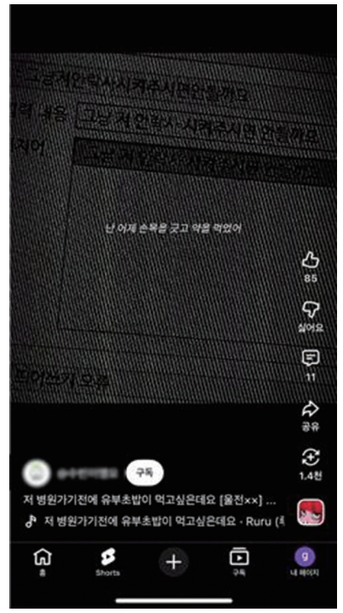


그 과정에서 마주한 위험들은 다음과 같았습니다.

‘슬픔’, ‘우울’과 같은 키워드에서 출발했을 때, A플랫폼에서는 약 34분 만에 자해 도구(커터칼, 눈썹칼, 반창고)와 함께 “소리 안나서 몰래 쓰기 제일 좋더라”라는 피드가 추천되었습니다. 댓글은 305개가 달려있었고, 댓글에는 어떤 도구를 써야 덜 아픈지, 부모님 몰래 할 수 있는지, 어떤 약을 발라야 빨리 낫는지 등 해당 플랫폼이 커뮤니티 가이드라인에서 허용하지 않는다고 밝혔거나, 아동에게 노출시키지 않겠다고 한 내용들이 가득했습니다. B플랫폼은 약 48분 만에 자살·자해 암시 표현을 담은 노래를 배경음악으로 한 콘텐츠가 추천되었습니다. 이 노래는 성인 인증을 거쳐야 들을 수 있는 곡이었음에도 숏폼 배경음악으로 아무 제한 없이 사용되고 있었습니다.




A플랫폼 ‘자살·자해 도구 추천’




B플랫폼 ‘자살·자해 암시 표현 노래’

설문조사에서 가장 높은 노출 비율을 보였던 성 관련 유해콘텐츠[4]의 경우, A플랫폼에서는 약 23분 만에 성 암시적 콘텐츠가 추천되었고, C플랫폼에서는 약 54분부터 57분까지 노골적인 유해콘텐츠가 연속으로 추천되었습니다. ‘데이트’, ‘용돈’과 같은 일상적인 키워드를 검색하는 과정에서 신체를 평가하거나 성적 의미를 강화하는 콘텐츠가 노출되었고, 별도의 우회 없이 불법 도박 사이트 연관 키워드로 탐색이 확장되는 경우도 확인했습니다. 더 나아가 A플랫폼에서는 아동 계정의 추천피드에 연령에 적합하지 않은 대출광고와 성인웹툰 광고까지 노출되었습니다.

4) 유해콘텐츠 유형별 경험: 성 42.7% / 섭식장애 18.8% / 마약·도박 18.6% / 자살 17.2% / 자해 16.5%




A플랫폼 '대출광고'




A플랫폼 '성인 웹툰 광고'

아동이 디지털 환경에서 겪는 위험은 단순히 숫자로 설명할 수 있는 문제가 아니었습니다. 그 위험은 갈수록 복잡해지고, 무엇보다 플랫폼의 기능과 설계에 크게 영향을 받고 있었습니다. 아동이 무엇을 검색했는지보다, 검색 이후 플랫폼이 어떤 콘텐츠를 추천하고 노출하도록 설계되어 있는지가 아동이 마주하는 위험의 크기를 좌우하고 있었습니다.

## 2. 안전설계를 위한 위험평가의 필요성

가상 계정 테스트에서 문제가 된 것은 유해콘텐츠 하나하나가 아니라, 아동을 그 콘텐츠로 데려간 구조였습니다. 추천 알고리즘이 유사한 콘텐츠를 이어서 노출시켰고, 그 콘텐츠에 붙은 해시태그가 다음 콘텐츠로 가는 통로가 되었습니다. 여기에 콘텐츠가 무한 재생되는 숏폼 서비스의 특성이 더해집니다. 아동이 무언가를 결정하는 지점은 점점 사라지고, 화면은 계속 다음 것을 보여줍니다. 한 번 그 흐름에 들어서면 빠져나오기 어렵고, 그 사이 노출되는 콘텐츠의 수위는 점점 높아집니다.

아동들도 이 구조의 문제점을 인지하고 있었습니다. 초록우산 설문조사에서 아동이 유해콘텐츠를 시청하게 되는 데 누구의 책임이 있다고 보는지 물었을 때, 71.8%의 아동이 '플랫폼 회사'를 책임 있는 주체로 꼽았습니다.[5] 이는 유해콘텐츠를 게시한 사람뿐 아니라, 해당 콘텐츠가 아동·청소년에게 도달하도록 추천·노출한 플랫폼에도 책임이 있다고 본 것입니다.

5) 유해콘텐츠 시청 시, 책임 소재 인식(1~3순위/중복응답): 콘텐츠를 만들거나 올린 사람 80.9%, 플랫폼 회사 71.8%, 국가 28%



그러나 현재의 대응은 콘텐츠를 찾아 삭제하는 데 머물러 있습니다. 이 방식으로는 아동이 이미 그 화면에 노출된 뒤에야 대응할 수 있고, 그마저도 삭제된 영상의 자리는 곧 다른 영상이 채웁니다.[6] 오늘 발제에서도 국내 아동·청소년 보호체계가 사후 대응 중심으로 운영되고 있다는 점을 상세히 짚어주셨습니다. 위험이 서비스의 설계와 운영에서 발생한다면, 점검도 그 단계에서 이루어져야 합니다.

이것이 초록우산이 안전설계를 위한 위험평가 제도의 도입을 촉구하는 이유입니다. 현재 국회에는 위험평가 제도를 도입하는 정보통신망법 개정안(대표발의 조인철 의원/의안번호 2216321)이 계류되어 있습니다. 현행법으로는 서비스의 설계 구조나 알고리즘에 내재된 유해성을 선제적으로 관리할 수 없다는 문제의식에서 출발한 법안입니다. 아동이 위험에 노출된 뒤에 대응하는 것이 아니라, 노출되기 전에 점검하는 체계로 옮겨가는 첫 단추입니다.

## 3. 개선과제

위험평가 제도가 실질적인 사전예방 장치로 작동하기 위해서는 제도의 도입과 더불어 두 가지가 마련되어야 한다고 생각합니다.

### 첫째, 아동참여 문항으로 실효성을 높인 평가도구 개발

오늘 발제에서는 사업자 자가진단 도구를 단계적으로 고도화해나가면서, 그 과정에 아동참여 구조를 함께 구현해 나가는 방향을 제안해 주셨습니다. 실제 플랫폼에 "아동이 추천 콘텐츠 노출을 조절할 수 있는 기능을 제공하고 있는가?"라고 묻는다면, 대부분의 플랫폼은 "그렇다"고 답할 것입니다. 실제로 기능은 존재합니다. 다만, 설문조사에서는 추천 알고리즘 새로고침, 중단과 같은 기능을 이용해 본 적이 없다고 응답한 아동은 62.1%였고, 가장 큰 이유(56.1%)는 '기능이 있는지 몰라서'였습니다. 또한, "플랫폼이 유해콘텐츠로부터 아동을 실제로 보호하고 있다고 느낀다"는 문항에 긍정적으로 응답한 아동은 27.4%에 그쳤습니다. 기능은 존재하지만 인지되지 않고, 보호받고 있다는 체감으로 이어지지 않는 것입니다. 따라서 평가도구의 문항은 '기능을 제공하는가'가 아니라, '아동이 그 기능을 찾고, 이해하고, 사용할 수 있는가'를 묻는 방향으로 설계되어야 합니다.

### 둘째, 실효성 확보를 위한 감독체계 마련과 제재 수단 강화

위험평가 결과가 플랫폼 내부에만 머무른다면, 평가가 제대로 이루어졌는지 확인할 방법이 없습니다. 플랫폼은 위험평가 결과와 그에 따른 주요 위험 요인, 개선·완화 조치의 내용과 이행 현황을 관할 정부·규제기관에 제출하고, 핵심 사항은 대중에게 공개하도록 하는 장치가 필요합니다. 규제기관은 그 조치가 실제로 아동 보호에 기여하고 있는지를 지속적으로 확인하고, 필요한 경우 추가적인 개선이나 시정을 요구할 수 있어야 합니다. 이는 해외에서 이미 시행되고 있는 방식이기도 합니다. 영국 Ofcom은 위험평가 제도 도입 후 1년차 보고서[7]를

6) 가상계정 테스트 과정에서 유해콘텐츠를 신고하여 삭제하였으나, 다음날 똑같은 유해콘텐츠가 같은 계정에 재업로드됨.

7) Ofcom, 「Year 1 Online Safety Risk Assessments: Report on standards and improvements」, 2025. 12. 4.

통해 사업자별 개선 영역을 제시하고 사업자별 맞춤형 피드백을 전달했으며, 필요한 개선이 이뤄지지 않을 경우 제재조치를 취하고 있습니다. EU에서는 정부뿐 아니라 200여 개 시민사회단체와 연구자들이 대형 플랫폼의 위험평가 보고서를 직접 분석해 부족한 부분을 짚는 1년차 보고서[8]를 발간하며 규제당국의 집행을 보완하고 있습니다. 정부의 감독과 사회적 감시가 함께 작동할 때, 위험평가는 형식적 절차에 그치지 않습니다.

그리고 이러한 절차를 담보하는 것이 제재입니다. 위험평가를 이행하지 않거나 허위 또는 부실하게 수행하는 경우, 그리고 평가 결과에 따른 개선·완화 조치를 이행하지 않는 경우에 대해 과징금 부과와 시정명령, 반복 위반 시 가중 제재와 같은 단계적인 수단이 마련되어야 합니다. 영국의 경우 주의의무 미이행 등 OSA 2023의 규정을 위반한 서비스 제공자에게는 최대 1,800만 파운드(한화 약 340억 원) 또는 연간 전 세계 매출액의 10% 중 더 큰 금액의 벌금을 부과할 수 있습니다. 위험평가를 이행하지 않는 데서 오는 이익이 제재로 인한 부담보다 크다면, 그 제재는 작동하지 않습니다. 평가가 그 자체로 목적이 아니라 실제 위험을 줄이기 위한 절차라면, 이행을 담보할 수 있는 수준의 제재가 전제되어야 합니다.

## 4. 마무리

유엔아동권리위원회는 일반논평 제25호 「디지털 환경과 아동권리」에서 국가가 기업으로 하여금 아동에게 미칠 수 있는 위험을 미리 점검하고, 아동의 권리를 지킬 책임을 다할 것을 요구하고 있습니다. 특히, 아동에 대한 디지털 환경의 차별적이고 때로는 심각한 영향을 특별히 고려하여 아동권리 영향평가를 수행하고 이를 공개하도록 요구해야 한다고 명시하고 있습니다. 위험을 미리 점검하는 일은 기업의 선의에 맡겨둘 선택지가 아니라 국가가 만들어야 할 제도라는 것이며, 이것이 이미 국제사회가 합의한 기준입니다.

위험평가 제도의 도입을 포함해 디지털 환경에서 아동의 권리가 실질적으로 보장될 수 있도록 제도적 기반이 이번 국회에서 마련되기를 기대합니다. 감사합니다.

8) DSA Civil Society Coordination Group, 「Initial Analysis on the First Round of Risk Assessments Reports under the EU Digital Services Act」, 2025. 3. 17.



종합토론

3 한승혜 리더

한국인터넷기업협회 디지털경제연구원

# 토론문

한승혜 리더(한국인터넷기업협회 디지털경제연구원)

현재 국내외 입법 논의의 상당 부분은 특정 연령 미만의 이용 자체를 제한하거나, 플랫폼에 연령 인증 의무를 부과하는 방식에 집중되어 있습니다. 프랑스 헌법위원회는 올해 8월 「15세 미만 SNS 금지법」에 위헌 결정을 내렸습니다. 청소년을 단일 집단으로 묶어 일률적으로 제한하고 부모의 개별 판단 자율권을 박탈하며, SNS 접속을 표현의 자유 영역에서 배제한 것이 비례 원칙에 위반된다는 이유였습니다. 호주는 2025년 12월 16세 미만 SNS 이용 금지법을 시행했지만, 시행 수개월이 지나도록 대다수 청소년은 여전히 접속하고 있다는 보도가 이어지고 있습니다. 연령 인증 역시, 비대면 환경에서 특정 연령층의 접근을 제한하려면 전체 이용자의 신원을 사전 검증해야 하는 구조이며, 이는 청소년뿐 아니라 성인의 프라이버시까지 광범위하게 침해하고 플랫폼의 개인정보 수집 범위만 확대합니다.

이러한 맥락에서 발제가 제시하는 사후 콘텐츠 규제에서 서비스 설계 단계의 위험 평가로의 전환은 주목할 만한 방향입니다. 서비스 설계 단계에서 위험을 미리 파악하고 관리하는 체계는 사업자에게 예측 가능한 기준을 제공하고 실질적인 안전 개선으로 이어질 수 있습니다.

다만 발제와 관련하여 몇 가지 보완이 필요한 지점들을 말씀드리고 싶습니다.

## 1. 세 가지 위험의 성격 차이와 평가 설계의 정합성

발제에서 아동·청소년이 직면한 온라인 위험을 콘텐츠 위험(Content Risk), 접촉 위험(Contact Risk), 행위 위험(Conduct Risk)의 3C로 분류하고 있습니다. 그런데 이 세 가지 위험의 성격은 본질적으로 다릅니다. 유해정보의 노출은 콘텐츠 관리의 문제이고, 온라인 그루밍이나 성 착취는 범죄 예방의 문제이며, 과몰입과 알고리즘 기반의 강박적 이용은 서비스 설계의 구조적 문제입니다. 이 셋은 원인도 해법도 책임 주체도 서로 다른 문제입니다.

이를 하나의 규제 틀로 묶어 일괄 적용할 경우, 위험 유형에 맞지 않는 수단이 선택되거나 규제의 초점이 분산되어 어느 문제도 충분히 다루지 못하는 결과로 이어질 우려가 있습니다. 이행 비용은 가중되는 반면 실질적 안전 개선은 불분명해집니다.

사전위험평가 체계가 실효적으로 기능하기 위해서는 대상 서비스가 어떤 위험을 주로 발생시키는지를 먼저 특정하고, 그에 맞는 평가 항목과 완화 조치를 연결하는 방식이어야 합니다. 위험 유형이 다르면 해결의 주체와 수단 또한 달라져야 할 필요가 있습니다.

## 2. 청소년에 한정된 접근이 가진 한계

발제문에 제시된 사전위험평가 체계 안에서, 위험 유형에 따라 연령 구분의 실효성은 다르게 나타납니다. 그 중에서도 알고리즘 기반 과몰입 같은 영역의 경우, 이용자의 연령을 구분하여 적용하는 것의 실효성은 낮은 편입니다.

무한 스크롤, 자동 재생, 알림 시스템은 체류 시간 극대화를 위한 설계 요소로, 이용자의 연령을 가리지 않고 작동합니다. 이 구조를 청소년만을 대상으로 보정하는 방식은 과몰입 완화라는 목적에 적합한 해결 방안이라 보기 어렵습니다. 오히려 실효성은 낮으면서 연령 확인과 별도 적용을 위한 이행 부담만 가중되었던 게임 셧다운제의 경험을 되풀이할 우려가 큽니다.

호주 Safety by Design이나 EU DSA가 초대형 플랫폼 전체에 위험완화 의무를 부과하는 방식을 택한 것도 이 맥락이라 볼 수 있습니다. 전체 이용자를 대상으로 한 서비스 설계 기준을 기반에 두고, 청소년에 대한 추가 보호를 그 위에 얹는 구조가 더 실효성 있는 평가 체계가 될 수 있을 것으로 보입니다.

## 3. 의무화 중심의 평가 체계가 디커플링으로 귀결될 우려

발제는 사전위험평가를 법적 의무 기반의 제도로 설계하고, 평가 대상·지표·절차·거버넌스를 구체적으로 제시하고 있습니다. 이 방향 자체는 의미있는 접근이라고 생각합니다.

다만 의무화를 중심축으로 삼는 구조에서는, 형식적 준수와 실질적 개선이 분리되는 것이 구조적으로 자연스러운 귀결입니다. 감독기관이 확인할 수 있는 것은 체크리스트가 작성되었는지와 절차가 이행되었는지에 한정되고, 그 결과로 실제로 어떻게 달라졌는지를 측정하고 평가하기는 어렵습니다. 이 구조에서 사업자의 합리적 대응은 실질적 위험 발굴보다 소극적 의무 이행으로 귀결됩니다.

발제에서도 '형식적인 규제수준 체크리스트보다 실질적인 위험 발견 완화에 초점'을 두어야 한다고 명시하고 있습니다. 그러나 이 목표가 의무화 구조 안에서 어떻게 달성될 수 있는지, 평가 결과가 서비스 개선으로 연계되는 메커니즘이 어떻게 작동하는지에 대해서는 아직 과제로 남겨져 있습니다.

여기에 더해, 광범위한 체크리스트를 단일하게 적용하는 방식은 이 문제를 심화합니다. 평가 항목이 많을수록 서비스에 실제로 해당하지 않는 항목에서도 검토 부담이 발생하고, 사업자는 지표 너머의 실질적 위험 해소보다는 눈앞의 항목에 대한 대응에 더 집중할 수밖에 없습니다.

## 4. 제언

청소년을 선별하여 보호하겠다는 접근은, 의도와 달리 구조적 한계가 생각보다 큽니다. 연령 인증은 전체 이용자의 신원 검증으로 이어지고, 과몰입을 유발하는 알고리즘 설계는 연령을 가리지 않고 작동합니다. 이 논리를 일관되게 따라가면, 결국 플랫폼에 한정한 논의를 넘어서 다양한 해법의 모색이 필요하다는 방향으로 귀결됩니다. 청소년의 선별적 보호를 위해서는 디바이스·OS 수준의 관리와 실질적인 청소년 여가시간에 대한 논의

까지도 확장하여 고민해야 합니다.

그럼에도 플랫폼에 한정해서 사전위험평가체계를 논의할 때는 “누구를 막을 것인가”보다 “서비스를 어떻게 설계할 것인가”에 집중하는 편이 더 올바른 접근이라고 생각합니다. 모든 서비스에 동일한 체크리스트를 일괄 적용하기보다는 특정 서비스에서 나타날 수 있는 위험 유형에 따라 더 효과적인 지표를 선별 적용하는 방식이 필요합니다. 위험 유형이 다르면 기준과 책임 주체도 달라져야 합니다. 이와 함께 정부·사업자·시민사회·아동/청소년이 기준을 함께 개선해 가는 협력적 거버넌스 구조가 병행될 때 평가 체계는 법적 이행의 영역을 넘어 실질적 위험 개선의 도구가 될 수 있습니다.

오늘 토론회가 아동·청소년의 디지털 안전을 위한 실효성 있는 제도 설계로 이어지는 데 기여하길 바랍니다. 감사합니다.

## 종합토론

### 4 최진응 입법조사관

국회입법조사처 과학방송통신과

## 토론문

최진응 입법조사관(국회입법조사처 과학방송통신과)

□ **아동·청소년의 디지털 안전을 위한 사전위험평가를 사업자의 법적 의무로 할 것인지, 자율규제로 할 것인지 명확히 규제 방향을 설정할 필요가 있음**

○ 호주·영국·EU에서는 아동·청소년의 온라인 안전 보호를 위한 사전위험평가가 온라인플랫폼 사업자의 법적 의무이고, 이를 이행하지 않을 경우 사업자는 법정 제재를 받게 됨

○ 하지만 국내법에서는 아동·청소년의 온라인상 안전을 보호하기 위한 사전위험평가는 법적 의무가 아니기 때문에 현행법에서 사전위험평가의 도입은 사업자에게 강제적 효과는 없는 자율규제 형태의 가이드라인 형태가 되지만, 강제적인 법적 의무를 부과하고자 한다면 법률 개정이 필요한 사항임

□ **아동·청소년 보호의 실효성 확보를 위해 사전위험평가를 사업자의 법적 의무로 할 경우, 기본권 침해 최소화 측면에서 영국과 EU 방식의 사전적 통제 중심의 입법 체계로 현행 정보통신망법을 전면 개편하는 것이 필요함**

○ 영국의 온라인안전법, EU의 디지털서비스법은 청소년유해매체물에 대해 정부가 직접 삭제를 강제하는 시스템이 아니라 민간 사업자에게 미성년자 보호를 위한 포괄적인 사전적 법적 책임을 부여하고 이를 이행하지 않을 경우 정부가 사후에 관여하는 법정 제재하는 시스템임

○ 반면 호주의 온라인안전법은 청소년유해매체물에 대해 정부가 온라인플랫폼사업자에게 삭제 명령과 같은 사후적 통제가 있고, 이에 더하여 정부의 온라인안전기본요건과 민간의 산업규약을 통해 민간사업자에게 아동·청소년 보호를 위한 사전위험평가를 이행하고 불이행 시 법정 제재하는 사전적 통제도 구축하고 있음

○ 우리나라의 정보통신망법도 호주와 유사하게 청소년의 접근을 차단하지 않은 채 유통되는 청소년유해매체물을 불법정보로 규정하여 정부가 삭제 명령하는 사후적 통제 방식을 갖고 있는데, 여기에 호주와 같이 플랫폼 사업자의 사전적 통제 조치를 강화하는 경우 기본권을 과도하게 제약할 수 있음

○ 즉 사전적 통제방식인 영국과 EU의 법제와 달리 호주 같은 입법 체계는 온라인 안전을 목적으로 사전적 통제와 사후적 통제를 동시에 하는 시스템으로 온라인상 표현의 자유 침해 및 인터넷 사업자의 부담이 가중될 우려가 큼

# 종합토론

## 5 김우석 과장

방송미디어통신위원회 디지털유해정보대응과

# 토론문

김우석 과장(방송미디어통신위원회 디지털유해정보대응과)

# 메 모