



# 생성형 인공지능의 연령별 등급제한 필요성과 기준 제안

## Proposed Criteria for and Need for Age-Based Ratings for Generative Artificial Intelligence

이대희<sup>†</sup> · 안성진<sup>††</sup>

Daehee Lee<sup>†</sup> · Seongjin Ahn<sup>††</sup>

### 요약

생성형 인공지능이 제공하는 콘텐츠가 실시간으로 달라지는 특성 때문에, 출시 시점에 심사하는 전통적인 완성형 콘텐츠(영화, 게임 등)의 등급 분류 체계만으로는 미성년자를 충분히 보호하기 어려운 실정이다. 본 연구에서는 생성형 AI에 특화된 등급 분류 체계를 마련하기 위하여 완성형 콘텐츠 등급 분류 체계를 참고하되, 전문가 심층 면접(FGI)을 통해 5개의 연령 구간과 14개의 분류 항목을 도출하였다. 설문조사를 통하여 기술통계분석, 상관분석을 진행하여 생성형 AI 등급 분류 체계를 검증하여 제안하였다. 이 기준은 기존의 허락되지 않은 미성년자에 대한 전면 금지 규제와 비교해 점진적 허용을 통한 접근성을 높이는 방법으로 규제를 완화하는 차별점을 가지고 있다. 이를 활용하여 미성년자를 보호하고 AI 기술이 사회적으로 적용 가능한 방향으로 발전할 수 있는 정책·기술적 함의를 제안하며, 향후 실제 서비스 환경에서 이 기준을 적용하기 위한 기술 개발 연구 및 제도적 입법 논의에 출발점이 될 수 있다.

주제어 인공지능, 생성형 AI, 연령별, 제한, 미성년자, 콘텐츠

### ABSTRACT

Because the content generated by Generative Artificial Intelligence changes in real time, it is difficult to adequately protect minors using only the traditional rating classification system—designed for completed content such as films and games, which is assessed at the time of release. In this study, in order to establish a rating classification system specialized for Generative AI, we referred to existing rating systems for completed content while conducting in-depth expert interviews (FGI). Through this process, we derived five age groups and fourteen classification items. Through a survey, we conducted descriptive statistical analysis and correlation analysis to validate and propose a generative AI rating classification system. Compared to the previous regulation, which imposed a blanket ban on minors without permission, this standard distinguishes itself by gradually allowing access, thereby easing restrictions and improving overall accessibility. Building on this, we suggest policy and technical implications for protecting minors and guiding the development of AI technology in a socially acceptable direction. It can also serve as a starting point for future technology development research and institutional legislative discussions to apply the guidelines in real-world service environments.

Keywords Artificial intelligence, generative AI, age, restrictions, minors, content

†정회원   성균관대학교 대학원 컴퓨터교육학과 박사수료  
††중신회원   성균관대학교 사범대학 컴퓨터교육과 교수(교신저자)  
논문투고   2025년 02월 07일  
심사완료   2025년 03월 04일  
게재확정   2025년 03월 05일  
발행일자   2025년 03월 19일

## 1. 서론

최근 인공지능(AI) 기술 중 하나인 생성형 AI(Generative AI)가 빠른 속도로 발전하면서, 글·그림·영상·음성 등 다양한 양질의 콘텐츠를 비전문가가 실시간으로 대량 생성·공급하는 사례가 늘어나고 있다. 이러한 기술 발전은 사용자에게 창의적인 아이디어를 제공하거나 생산성을 높이는 등 긍정적인 효과를 가져온다. 하지만 동시에, 생성형 AI가 만들어진 콘텐츠가 일부 사용자(특히 아동·청소년)에게 유해(모방 위험, 선정성, 심리적 영향, 신뢰 인증 등)하게 작용할 가능성도 함께 커지고 있다[1-4].

전통적인 완성형 콘텐츠(영화, 게임, 만화, 도서 등)는 콘텐츠가 완성된 이후 관련 법령을 기준으로 사전 심의를 거쳐 등급이 부여되므로, 연령별 또는 이용 환경별로 제한을 두는 체계가 비교적 확립되어 있다[5, 6]. 반면 생성형 AI는 사용자의 입력과 모델 학습 데이터에 따라 결과가 실시간으로 달라지므로, 기존의 일회성 심의·분류 방식만으로는 부적절한 콘텐츠 노출을 효과적으로 막기 어려운 실정이다[7].

세부적으로는 미성년자의 생성형 AI 사용을 제한하는 기준은 있지만, 점진적 허용에 관련된 해외 및 국내의 규제 제도와 연구자료를 거의 찾아볼 수 없는 상황이다. 이에 본 연구는 생성형 인공지능 사용자 규제가 미성년자를 효과적으로 보호할 수 있도록, 규제에 관한 정책적 관점을 제시하려는 목적에서 기획되었다. 이하에서는 우선 논의의 출발점으로서 생성형 AI의 연령별 등급 제한 필요성에 대해 살펴보고, 기존 완성형 콘텐츠 등급제도의 사례와 관련 법·제도(국내외)를 검토하고, 생성형 AI의 특성을 반영한 새로운 연령별 등급제한 기준을 제안하고자 한다.

## 2. 이론적 배경

### 2.1 생성형 AI 현황

#### 2.1.1 생성형 AI

생성형 AI란 학습데이터를 기반으로 사용자의 요구사항에 따라 글, 그림, 음성, 영상 등 다양한 형태의 콘텐츠 산출물을 짧은 시간에 대량으로 생성할 수 있는 기술이다 [8, 9]. 이는 날씨예측이나 이미지 분류 등 특정 기능만을 수행하는 AI 모델들과 구별되는 개념이다[10]. 생성형 AI는 사용자가 요구한 산출물을 생성하기 위해 스스로 데이터를 찾아서 학습하고 이를 조합하여 결과물을 제시한다 [11]. 이를테면 자연어처리 분야에서는 대규모 언어 모델(Large Language Models)을 통해 질문에 대한 답변이나 새로운 글을 생성(ChatGPT 등)할 수 있다[12]. 그림 분야 또한 딥러닝 기반으로 예술 작품이나 만화, 사진과 유사한 이미지를 만들어내는 모델(Midjourney 등)이 존재한다[13]. 음성 분야에서는 학습된 사람의 음성을 모방하여 주어진 텍스트를 읽거나, 제공된 가사에 어울리는 장르의 음악을 작곡(AIVA 등)한다[14]. 영상 분야에서는 썸

플 이미지와 텍스트로 장면을 설명하여 스토리보드 작성하면, 이를 준수하여 영상을 제작(Pictory 등)한다[15].

#### 2.1.2 생성형 AI 위험성

생성형 AI가 만들어내는 콘텐츠의 유연성과 창의성은 뛰어나지만, 동시에 사용자의 의도적이거나 잘못된 요청에 따라 비판적 사고가 성숙하지 못한 미성년자에게 부적절한 콘텐츠(모방위험, 선정성, 심리적 영향, 신뢰 인증 등)를 생성할 수 있다는 위험성을 안고 있다. 또한 학습 데이터 자체에 내재된 편향이나 윤리적 문제가 그대로 모델에 전이되어 산출된 콘텐츠에 포함될 수도 있다[1-4]. 따라서 노출이 적절치 않은 콘텐츠가 실시간으로 만들어져 제공될 가능성을 사전에 차단하거나 최소화할 방안이 필요하다.

### 2.2 완성형 콘텐츠 연령별 등급 제한 동향

완성형 콘텐츠의 연령별 등급 제한 제도는 이미 국내외에서 오랜 역사를 지니고 있다. 한국의 경우 영화·비디오물은 영상물등급위원회(KMRB)에서, 게임은 게임물관리위원회(GRAC)에서 각각 등급을 분류한다[16, 17]. 미국은 Motion Picture Association(MPA)과 Entertainment Software Rating Board(ESRB)를 운영 중이다[18, 19]. 브라질에는 Brazilian advisory rating system(ClassInd)에서 영화, 게임, TV 프로그램을 분류하고 있다[20]. 호주에서는 Australian Classification Board(ACB), 캐나다에서는 Entertainment Software-Rating Board(ESRB), 독일에서는 Unterhaltungs Software Selbstkontrolle(USK), 유럽과 이스라엘에서는 Pan-European Game Information(PEGI)가 담당하며, 국제적인 게임 등급 기준으로는 INTERNATIONAL AGENCY FOR RESEARCH ON CANCER(IARC)도 사용된다[21-24]. Fig. 1은 기관별 콘텐츠 제한 기준 중 일부를 연령 범위로 색상을 구분하여 나타낸 것이다.



Figure 1. Content restriction criteria by country

대부분 18세를 기준으로 완성형 콘텐츠를 제한하고 있지만, 한국(GRAC)의 경우 19세를 기준으로 하고 있음을 볼 수 있다. 고등학교 졸업 기준(19세)과 동일한 것으로

확인할 수 있는데, 최근 투표권 등을 만18세 기준으로 부여하는 것 보면 해당 나이에 합리적 의사 표현이 가능하다고 인정하고 있기에, 한국도 18세 기준으로 변경될 여지가 있다[25]. 특이한 점은 ACB의 경우 첫 등급이 15로 매우 넓은 편인데, 보호자의 관리가 필요한 콘텐츠와 그렇지 않은 콘텐츠로 명확하게 구분하여 규제하는 것을 볼 수 있다. 이들 제도는 완성된 콘텐츠를 배포 이전에 정해진 기준으로 심의하여 연령별 이용에 대한 허용·제한을 결정한다. Table 1은 다양한 기관별 2개 이상 중첩되는 항목들을 표로 나타낸 것이다. 체크가 되지 않은 분야도 대부분 유사한 항목에서 내포하여 점검하고 있다.

Table 1. Rating consideration factors[16-24].

Category	ACB	ClassInd	ESRB	USK	PEGI	GRAC	KMRB
Sexuality	<input checked="" type="checkbox"/>						
Violence	<input checked="" type="checkbox"/>						
Horror		<input checked="" type="checkbox"/>					
Crime	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Drug	<input checked="" type="checkbox"/>						
Language	<input checked="" type="checkbox"/>						
Gambling			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
Nudity	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>				
discrimination		<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>		
Online			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		
Purchases			<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>		
Gore			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			

### 2.3 생성형 AI 규제 동향

생성형 AI에 대한 제한의 공통적 특징은 허가과 금지로만 구분되는 것에 있다. 미국의 경우 아동 인터넷 개인정보 보호법(COPPA)에 따라 만 13세 미만 어린이들은 사용이 제한되며, 만13세 이상, 18세 미만의 학생은 부모 혹은 법적 보호자의 동의하에 사용할 수 있어서 대부분의 생성형 AI 서비스 제공 기업들은 이를 약관에 표시하고 있다. 우리나라는 생성형 AI 이용지침, 생성형 AI 윤리 가이드북, 생성형 AI 윤리 가이드라인 등을 제작하여 허가받지 못한 미성년자의 사용 제한을 안내하고 있다[26-28]. 이와 관련하여 한국지능정보사회진흥원에서는 분기마다 지능정보 윤리 리포트를 발간하여 지속적인 관심을 나타내고 있다[29]. Table 2는 생성형 AI 서비스 제공 기업별 미성년자 사용 기준을 정리한 것이다.

Table 2. Age available for generative AI[12-15].

Generative AI	Age Available	Parental Consent
ChatGPT	13 ~ 18	<input checked="" type="checkbox"/>
Midjourney	13 ~ 18	<input checked="" type="checkbox"/>

Generative AI	Age Available	Parental Consent
AIVA	13 ~ 18	<input checked="" type="checkbox"/>
Pictory	13 ~ 18	<input checked="" type="checkbox"/>

## 3. 생성형 AI 연령별 등급 제한 기준 연구

### 3.1 연구 방법

생성형 AI 연령별 등급 제한 기준을 연구하기 위하여 국내외 완성형 콘텐츠 연령별 등급 제한 기준, 연구기관의 연구 보고서, 관련 학술지 논문, 서비스 제공 기업 약관, 관련 사례 분석 등을 바탕으로 생성형 AI가 미성년자에게 미칠 수 있는 위험성과 악영향을 분석하였다. 분석을 통하여 정리된 내용을 중심으로 10인의 전문가 심층 면접(FGI: Focus Group Interview)을 2차에 걸쳐 진행하였다. FGI는 대면 및 비대면으로 진행하였다.

FGI를 통해 최종적으로 도출된 미성년자 보호를 위한 생성형 AI 사용을 제한하는 5개의 연령 구간과 14개의 분류 항목에 대한 상관관계를 검증하기 위해 41명의 이해 관계자를 대상으로 설문 조사를 진행하였다. 설문 조사를 활용하여 기술통계분석(Descriptives), 분류 항목별 상관분석(Correlation Analysis) 등을 진행하여 제안한 기준을 복합적으로 검증하고자 하였다.

### 3.2 전문가 심층 인터뷰(FGI)

전문가 심층 인터뷰에 참여한 전문가는 총 10명이었으며, 2차에 걸쳐 의견을 수렴하고 분류하여 기존 기준을 도출하였다. 참여한 전문가는 이 연구에 이해 관계자로 AI 분야 박사 3명, AI 분야 박사과정 2명, AI융합교육 석사학위 취득 교사 3명, 학부모 2명으로 구성하였으며 세부 정보는 Table 3과 같다.

Table 3. Information of experts group

Occupation	Number of Experts (%)	Career	Age Range
PhD in AI	3(30%)	16.x	fifties
PhD Student in AI	2(20%)	13.x	forties
Teacher	3(30%)	12.x	forties
parents	2(20%)	18.x	forties

Table 4. Age Restriction Criteria by Category

Category	Age	Description
Sexual (SX)	7+	Mild displays of affection
	12+	Indirect displays of romance or affection
	16+	Direct displays of affection or sexual suggestion
	18+	Explicit sexual behavior or depictions
Excp.		Artistic: sexual representation in a work of art Educational: Sexual content for sex education

Category	Age	Description
Violence (VL)	7+	Cartoonish or mild violence
	12+	Mild depictions of violence
	16+	Realistic violence but without blood or weapons
	18+	Explicit violence, including blood and weapons
	Excp.	Historical: the depiction of violence in historical events Education: Preventing School Violence
Horror (HR)	7+	Some element of surprise or suspense
	12+	Moderate horror elements
	16+	Strong horror or shocking scenes
	18+	Extreme horror or psychologically disturbing content
	Excp.	Cultural: Fear elements in legends or folk tales Educational: raising awareness of dangerous situations
Drugs (DR)	7+	Brief reference to the drug
	12+	Indirect description of drinking or smoking
	16+	Explanation or suggestion of drug use
	18+	Abuse and addiction of illegal drugs
	Excp.	Healthcare: Description of drug use for therapeutic purposes Education: Warning classes on the dangers of substance abuse
Language (LG)	7+	A bit of soft slang
	12+	Offensive language to the other person
	16+	Strong profanity or vulgar language
	18+	Very explicitly offensive language
	Excp.	Education: Lessons on Language Manners
Risk of Imitation (RI)	7+	Depiction of safe activities
	12+	Activities containing some element of risk
	16+	Suggestion of dangerous or illegal actions
	18+	Depiction of extremely dangerous or illegal acts
	Excp.	Educational: Explanation of risky behavior to emphasize safety rules
Nudity (ND)	7+	Light depiction of certain parts of the human body, depiction of animal bodies
	12+	Direct depiction of some parts of the human body, detailed depiction of animal bodies
	16+	Direct depiction of multiple parts of the human body
	18+	Explicit depiction of the human body
	Excp.	
Crime (CR)	7+	Depiction of minor rule-breaking
	12+	Minor crimes or moral dilemmas
	16+	Depiction of criminal or illegal activities
	18+	Depiction of serious crimes or violent illegal acts
Gambling Elements (GE)	7+	Simple chance-based games
	12+	Gambling elements using fake currency
	16+	Suggestion of gambling involving real money
	18+	Depiction or promotion of real gambling activities
	Excp.	Educational: Education on the dangers of gambling

Category	Age	Description
Discrimination (DS)	7+	Emphasis on diversity and inclusion
	12+	Indirect mentions of discrimination
	16+	Direct depiction of discrimination or prejudice
	18+	Hate speech or promotion of discrimination
User Data Collection (UDC)	7+	Minimal information collection
	12+	Collection of personal information with user consent
	16+	Collection and storage of personally identifiable information (PII)
	18+ Excp.	Collection of sensitive personal information or sharing with third parties
Psychological Impact (PI)	7+	Positive and straightforward moral lessons or messages
	12+	Content that helps understand social relationships and appropriate behavior
	16+	Content involving complex moral dilemmas or requiring value judgments
	18+	Content that can profoundly influence personal values or character, including intense ideological, political, or religious beliefs
	Excp.	Educational: Philosophical or ethical content that promotes critical thinking and self-reflection Counseling: Psychological advice beneficial for mental health and personal growth Cultural: Content that aids in understanding and embracing diverse cultures and worldviews
	All	Indicate the source of any original information used in the content. If the content is purely AI-generated with no identifiable source, label it with an "AI Mark."
Trust Certification (TC)	7+	trained on reliable and verified data
	12+	Results based on public institutions and other credible sources
	16+	Results from a variety of publicly accessible sources
	18+	Results that require user verification
Cost of Use (CU)	7+	Simple in-app purchases or advertisements
	12+	Increased prompts to buy additional content
	16+	Repeated purchase prompts or high-cost in-app transactions
	18+ Excp.	Excessive spending elements or gambling-like purchase inducements

1차 전문가 심층 인터뷰에서는 국내외 완성형 콘텐츠 연령별 등급 제한 기준을 각 기관에서 제공하는 문서를 정리하여 설명하였다. 또한 생성형 AI가 완성형 콘텐츠 등급 제한 기준을 적용할 수 없는 이유를 관련 문헌 분석 결과를 바탕으로 설명하고, 새롭게 제안하는 생성형 AI 연령별 등급 제한 기준과 항목을 배포한 뒤 기준에 문제점과 가능성, 고려사항 추가 항목 등을 도출하였다. 도출된 연령 범위는 7세~11세, 12세~15세, 16세~17세 18세 이상으로 IARC(국제 연령 등급 연합)의 연령 범위를 준용하기로 하였으며, 세부 항목은 기존 항목들을 종합하여 선정성, 폭력성, 공포, 약물, 언어, 모방위험, 노출, 범죄, 사행성 등을 기본으로 하고 추가로 생성형 AI 규제에 필요하다고 생각되는 차

별, 사용자 정보 수집, 심리적 영향(가치관 변화)등을 도출하였다. 또한 생성형 AI를 미성년자가 사용하기 위해서는 사전에 위험성 및 올바른 활용에 대한 교육이 필요하다고 의견을 모았다. 규제 방법은 기준에 일치하는 생성형 AI를 개발하는 것은 지양하고, 프로그램은 범용으로 사용하도록 용도에 맞는 검증된 데이터를 학습하거나, 산출 콘텐츠를 생성하는 시점에서 사용자의 연령에 맞는 기준의 콘텐츠를 생성하는 메커니즘을 개발사가 추가한 것을 점검하는 방식을 도입하기로 하였다.

2차 전문가 심층 인터뷰는 1차 인터뷰 결과를 사전에 배포한 후 진행하였다. 1차 때 연령 범위 외에 교육적, 문화적, 보안적, 역사적 등의 사유로 보호자의 동의가 있을 때 콘텐츠 생성에 대해 규제 예외를 적용하고자 하였는데, 인터뷰 결과 노출, 범죄, 차별, 사용자 정보 수집에 항목에 대해서는 예외를 두지 않기로 의견을 모았다. 사행성, 차별, 범죄, 폭력성은 나이별 제한 기준이 너무 관대하다고 판단되어 더욱 엄격하게 변경하였다. 그리고 생성형 AI의 고질적 문제인 환각에 대하여 비판적 사고가 없는 미성년자를 위해 신뢰 인증(출처 표기)을 항목에 추가하기로 하였다. 대체 적으로 학생 집단은 규제 완화를, 교사와 부모집단은 규제 강화를 요구하였으며, 이를 통해 규제의 영향을 받는 집단 간에 이해충돌이 생기는 것을 알 수 있었다. 특히나 교육청 지원으로 인공지능융합교육 석사학위를 취득한 교사 집단이 인공지능이 효과를 많이 경험한 만큼 위험성도 크게 인지하고 있어서 매우 엄격한 규제 강화를 요구하였다. Table 4 는 이를 정리한 표이다. 세부 의견으로는 연령별 규제가 아닌 인지발달 수준에 맞는 분류 기준을 추가하자는 의견이 있었으나 나이와 다르게 인지발달은 변화가 개별적 특성에 따라 다르기에 사용 때마다 검사해야 하는데, 효율적 어려움이 크다고 토의를 통해 판단되어 반영하지 않았다.

#### 4. 설문조사 분석

##### 4.1 조사 방법 및 대상

설문조사를 위한 문항은 전문가 심층 인터뷰에서 정리된 생성형 AI 연령별 등급 제한 기준을 검증 가능하도록 구성하였다. 미성년자 제한 연령 기준 구간을 7세, 12세, 16세, 18세로 하고 예외 조항을 포함하여 선정성, 폭력성, 공포, 약물, 언어, 모방위험, 노출, 범죄, 사행성, 차별, 사용자 정보 수집, 심리적 영향, 신뢰 인증, 사용료 등 14가지 분류 기준에 대한 적합성을 5점 척도(매우 그렇지 않다, 그렇지 않다, 보통, 그렇다, 매우 그렇다.)로 4점 이상일 경우 긍정 평가를 의미하는 것으로 구성하였다. 분류 별 연령 구간에 따른 기준에 대한 설문 결과를 상관관계 분석과 요소별 기술통계를 분석할 수 있도록 분리하여 진행하였다. 조사 대상은 인공지능 분야 전문가집단, 교사, 학부모를 고르게 선정하여 설문하였다.

Table 5. Demographic characteristics

Division		Frequency (persons)	Percent
Age Range	20~29	2	9.7%
	30~39	4	14.6%
	40~49	21	52.2%
	50~59	9	22.0%
	60~	1	2.4%
AI Usage time	Do not know	1	2.4%
	Know it exists	3	7.3%
	Have used it once	9	22%
	1 hour per week	4	9.8%
	2~4 hour per week	9	22%
	5 hour per week	4	9.8%
Field of Use	Paid subscription	11	25.8%
	Work	33	80%
	Daily Life	22	53.7%
Group	Not use	5	12.2%
	Subject Matter Experts(SMEs)	14	34.2%
	Teacher(Tr)	14	34.2%
	Parents(Pr)	13	31.7%

설문에 참여한 인원은 총 41명이며 인구통계학적 특성은 Table 5와 같다. 40대가 52.2%로 가장 많고, 생성형 AI 활용도는 67.4%가 주 1시간 이상을 사용하고 있었으며, 9.8% 주 5시간 이상을 사용하고, 25.8% 는 유료로 사용하는 생성형AI 서비스가 있었다. 80%가 업무에 활용하고 있었고, 53.7%는 일상생활에도 사용하고 있었다. 설문 참여자 중 사용하지 않는 비율은 단지 12.2%만 나타나는 것을 보면 생성형 AI가 우리 일상생활에 매우 밀접해졌다고 판단된다.

##### 4.2 구성요소별 기술통계 분석

Table 6. Component descriptive statistics

Category	CVR		mean		std	
	All	SMEs	All	SMEs	All	SMEs
		Tr		Tr		Tr
	Pr	Pr	Pr	Pr		
Sexual (SX)	0.71	0.43	4.12	3.64	0.71	0.63
		0.86		4.43		0.65
		0.85		4.31		0.63
Violence (VL)	0.56	0	4.17	3.5	0.77	0.52
		1		4.79		0.43
		0.69		4.23		0.73
Horror (HR)	0.56	0	4.07	3.5	0.98	1.16
		1		4.64		0.5
		0.69		4.08		0.86
Drugs (DR)	0.46	0	4.02	3.43	1.01	1.22
		0.86		4.5		0.65
		0.54		4.15		0.8
Language (LG)	0.46	-0.14	3.9	3.21	1.04	1.19
		0.86		4.5		0.65
		0.69		4		0.82



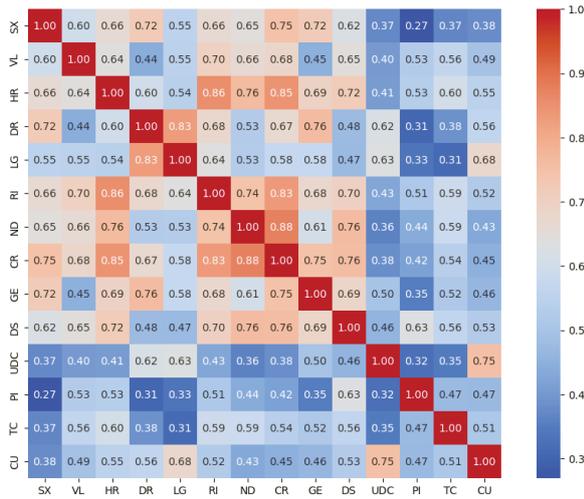


Figure 3. Overall Correlation Heatmap

### 4.4 연구결과

생성형 AI 기술이 급속도로 발전하면서 사용자에게 창의적인 아이디어를 제공하거나 생산성을 높이는 긍정적인 효과를 가져오고 있으나 아직 보호가 필요한 미성년자에게 규제 없는 콘텐츠 생산은 여러 위험성을 가지고 있다. 완성형 콘텐츠 제도의 연령 등급 제한 제도를 사용하고자 하지만 생성형 AI의 '실시간 유동적 콘텐츠 생성'이라는 특징 때문에, 출시 시점의 검증만으로는 유해 정보 노출을 막기 어렵다. 따라서 다층적 접근(학습 데이터 → 모델 접근 → 질문 입력 → 생성 결과 검수)이 필수적이며, 이 과정에서 기술적·법적 기반이 마련되어야 한다. 본 연구는 다양한 기관의 완성형 콘텐츠 연령 등급 제한 기준을 기반으로 국내외 생성형 AI 사용 가이드라인, 정책, 사용지침, 논문 등 다양한 문헌을 분석하여 미성년 사용자 보호를 중심으로 자료 분석과 2차의 FGI와 델파이 설문조사를 실시하여 기술 통계 분석, 요인 분석, 표본 적합도 분석, 신뢰도 분석, 상관관계 분석 등을 통해 생성형 AI 연령별 등급 제한 기준을 개발하였다. 전문가 심층 인터뷰와 설문 조사를 통해 완성된 AI 연령별 등급 제한 기준은 5단계의 연령 구간에 14개의 항목으로 구성되었다.

### 5. 결론

본 연구는 완성형 콘텐츠 등급제도를 기반으로, 생성형 AI의 특성에 맞게 연령별 등급제한 기준을 새롭게 구성하고 제안하였다. 선정성, 폭력성, 공포, 약물, 언어, 모방위험, 노출, 범죄, 사행성, 차별, 사용자 정보 수집, 심리적 영향, 신뢰 인증, 사용료 등 14개 항목을 마련하고, 미성년자의 연령은 7세, 12세, 16세, 18세 구간으로 분류하여 각각 제한 기준을 두는 방식으로 세분화하였다. 이를 통하여 생성형 AI에게 요청한 콘텐츠가 생성되는 시점에 검수 절차를 실행하여 사용자의 기준에 따라 필터링하거나 재생성 요청 혹은 공개 제한을 하는 구조로 콘텐츠 결과물을 제공하는 것으로, 동일

한 생성형 AI 모델의 프로그램이라도 성인에게는 최상의 성능과 표현의 자유를 제공하고, 미성년자에게는 점진적인 보호 기능을 제공하여 바람직한 생성형 AI 생태계를 구축하려는 시도이다.

그러나 본 연구는 설문조사 참여자와 반복 횟수가 부족하고, 기준을 준수하는 실제 검증 프로그램을 제작하여 테스트를 진행하고 결과를 분석하는 과정을 하지 못해 실제 적용되었을 때 다양한 상황에 대응이 다소 부족할 수 있는 완성도를 가진다고 할 수 있다. 규제 기준에 관한 연구이지만 대상자인 미성년자의 의견이 반영되지 못했다는 것도 미흡한 부분이다. 후속 연구에는 실제 생성형 AI 서비스에서 연령별 필터링 알고리즘을 적용하여 다양한 연령대 다수 사용자를 대상으로 콘텐츠 제작 경험을 조사하여 통계적 근거를 제시하는 연구를 진행해야 한다. 법·제도적 환경에서도 본 기준을 적용하려면 더 구체적이고 세분화시킨 기준이 추가로 필요하다.

그럼에도 불구하고 본 제안은 기존의 허가 및 금지로만 제한하는 일괄적 규제 방식에서 벗어나, 연령별로 점진적 허용 기준을 확대하여 미성년자의 활용 가능성을 높이고자 한다. 이는 생성형 AI 시대에 필요한 새로운 차원의 등급제도에 대한 기반을 마련했다는 데 의의가 있다. 향후 후속 연구를 통하여 AI 기업, 정부기관, 학계, 시민단체가 협력하여 유해 콘텐츠 방지와 창의적 활용이 균형을 이룰 수 있는 체계를 확립해 나가길 기대한다.

### 참고문헌

- [1] Oh, D. H. (2024, April 27). *An AI that keeps beating around the bush, but eventually reveals how to make a bomb*. newsis. [https://www.newsis.com/view/NISX20240415\\_0002700345](https://www.newsis.com/view/NISX20240415_0002700345)
- [2] Jeong, J. J. (2024, May 30). *Since early last year, the use of AI-generated images in spreading false information has surged*. KBS news. <https://news.kbs.co.kr/news/pc/view/view.do?ncd=7975913>
- [3] Choi, J. M. (2025, January 2). *Can It Even Control Human Decision-Making?... How Advanced Will AI Be in 2025?*, wikileaks-kr. <http://www.wikileaks-kr.org/news/articleView.html?idxno=161894>
- [4] Park, B. G. (2025, January 30). *Shocking Photo Turns Out to Be Me... The Culprit Was Just a Teenager*. herald. <https://biz.heraldcorp.com/article/10407820>
- [5] Ministry of Culture, Sports and Tourism. (2024, October 22). *Game Industry Promotion Act (Act No. 20485, Partially Amended)*.
- [6] Ministry of Culture, Sports and Tourism. (2024, May 1). *Act on the Promotion of Films and Videos (Act No. 19800, Partially Amended on October 31, 2023)*.
- [7] Kim, B. Y. (2024). *Legal Issues and Regulatory Discussions in Generative AI*. *Informatization Policy Academic Journal*, 31(3), 3-33.
- [8] Feuerriegel, S., Hartmann, J., Janiesch, C. et al. (2024).

- Generative AI. *Bus Inf Syst Eng*, 66(1), 111-126. <https://doi.org/10.1007/s12599-023-00834-7>
- [ 9 ] Jovanović, M. & Campbell, M. (2022). Generative Artificial Intelligence: Trends and Prospects, *Computer*, 55(10), 107-112. <https://doi.org/10.1109/MC.2022.3192720>
- [ 10 ] Hacker, P., Engel, A. & Mauer, M. (2023). Regulating ChatGPT and other large generative AI models. *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. <https://doi.org/10.1145/3593013.3594067>
- [ 11 ] Yang, J. & Yoon, S. (2023). Going beyond ChatGPT to the era of Generative AI: Media and Content Generative AI Service Case and How to Secure Competitiveness. *Media Issue & Trend Domestic Report*, 55, 62-70. [https://www.kca.kr/Media\\_Issue\\_Trend/vol55/KCA\\_ebook/index.html](https://www.kca.kr/Media_Issue_Trend/vol55/KCA_ebook/index.html)
- [ 12 ] OpenAI. (2025). *ChatGPT* [Large language model]. OpenAI. <https://chat.openai.com>
- [ 13 ] Midjourney. (2025). *Midjourney* [AI image generation tool]. <https://www.midjourney.com>
- [ 14 ] AIVA Technologies. (2025). *AIVA* [AI music composition tool]. <https://www.aiva.ai>
- [ 15 ] Pictory. (2025). *Pictory* [AI video creation tool]. <https://pictory.ai>
- [ 16 ] Korea Media Rating Board. (2025). *Video content rating system*. Ministry of Culture, Sports and Tourism. <https://www.kmr.or.kr>
- [ 17 ] Game Rating and Administration Committee. (2025). *Video content rating system*. Ministry of Culture, Sports and Tourism. <https://www.grac.or.kr>
- [ 18 ] Motion Picture Association (MPA). (2021, November 19). *Motion Picture Association age rating system*. <https://www.motionpictures.org>
- [ 19 ] Entertainment Software Rating Board (ESRB). (2013, August 2). *Entertainment Software Rating Board age rating system*. <https://www.esrb.org>
- [ 20 ] Brazilian Advisory Rating System (ClassInd). (2014, September 29). *Classificação Indicativa age rating system*. <https://www.classificacaoindicativa.gov.br>
- [ 21 ] Australian Classification Board (ACB). (2022, July 30). *Australian Classification Board age rating system*. <https://www.classification.gov.au>
- [ 22 ] Unterhaltungssoftware Selbstkontrolle (USK). (2009). *Unterhaltungssoftware Selbstkontrolle age rating system*. <https://www.usk.de>
- [ 23 ] Pan-European Game Information (PEGI). (2019, May 26). *Pan-European Game Information age rating system*. <https://www.pegi.info>
- [ 24 ] International Age Rating Coalition (IARC). (2023). *International Age Rating Coalition*. <https://www.globalratings.com>
- [ 25 ] National Election Commission. (2021, December 31). *Public Official Election Act* (Act No. 18588, Partially Amended). South Korea.
- [ 26 ] Park, J. H. (2023, August 29). *Guidelines for the Use of Generative AI by School Level*. Seoul Metropolitan Office of Education.
- [ 27 ] National Information Society Agency. (2024, January 1). *Generative AI Ethical Guidelines*. *Dysfunction Response Team*. <https://www.nia.or.kr/site>
- [ 28 ] Seoul Digital Foundation. (2023, November 13). *Seoul Metropolitan Government Generative AI Ethics Guidelines*. AI & Big Data Team. <https://sdf.seoul.kr>
- [ 29 ] National Information Society Agency. (2025, January 2). *Intelligent Information Ethics Issue Report, Winter 2024* (Vol. 5, No. 4). *Dysfunction Response Team*.
- [ 30 ] Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563-575. <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>



이대희

- 2011년 목원대학교 컴퓨터교육과 (학사)
- 2023년 성균관대학교 인공지능융합교육전공 (석사)
- 2025년 성균관대학교 컴퓨터교육학과 (박사수료)
- 2011년 ~ 현재 세명컴퓨터고등학교 정보-컴퓨터 교사

✚ 관심분야 : SW교육, 인공지능 활용, 교육과정 설계, 정보보안

✉ mail@leeday.net



안성진

- 1988년 성균관대학교 정보공학과 (학사)
- 1990년 성균관대학교 정보공학과 (석사)
- 1998년 성균관대학교 정보공학과 (박사)
- 1990년 ~ 1995년 KIST/SERI 연구원
- 1996년 정보통신기술사
- 1999년 3월 ~ 현재 성균관대학교 컴퓨터교육과 교수

✚ 관심분야 : 네트워크, 정보보안, SW교육, 인공지능 윤리

✉ sjahn@skku.edu

부 록

〈표 1〉 평가 고려 요소[16-24].

항목	ACB	ClassInd	ESRB	USK	PEGI	GRAC	KMRB
선정성	☑	☑	☑	☑	☑	☑	☑
폭력성	☑	☑	☑	☑	☑	☑	☑
공포		☑	☑	☑	☑	☑	☑
범죄	☑	☑	☑			☑	☑
약물	☑	☑	☑	☑	☑	☑	☑
언어	☑	☑	☑	☑	☑	☑	☑
사행성			☑	☑	☑	☑	
노출	☑	☑	☑				
차별		☑			☑		
온라인			☑	☑	☑		
사용료			☑		☑		
잔인성			☑	☑			

〈표 2〉 생성형 인공지능 사용 가능 연령[12-15].

생성형 인공지능	사용 가능한 연령	보호자 동의
챗GPT	13 ~ 18	☑
미드저니	13 ~ 18	☑
아이바	13 ~ 18	☑
픽토리	13 ~ 18	☑

〈표 3〉 전문가 그룹 정보

직업	전문가 수(%)	경력	연령대
AI 박사	3(30%)	16.x	50대
AI 박사과정	2(20%)	13.x	40대
교사	3(30%)	12.x	40대
부모	2(20%)	18.x	40대

〈표 4〉 항목별 연령 제한 기준

항목	나이	설명
선정성 (SX)	7+	가벼운 애정 표현
	12+	간접적인 로맨스 또는 애정 표현
	16+	직접적인 애정 표현 또는 성적 암시
	18+	노골적인 성적 행동 또는 묘사
	Excp.	예술적: 예술 작품에서의 성적 표현 교육적: 성교육을 위한 성적인 자료
폭력성 (VL)	7+	만화 같거나 가벼운 폭력
	12+	가벼운 폭력 묘사
	16+	현실적인 폭력이지만 피나 무기가 없는 폭력
	18+	피와 무기를 포함한 노골적인 폭력
	Excp.	역사: 역사적 사건에서 폭력 묘사 교육: 학교 폭력 예방 교육

항목	나이	설명
공포 (HR)	7+	약간의 놀라움 또는 긴장
	12+	중간 정도의 공포
	16+	강한 공포 또는 충격적인 장면
	18+	극도의 공포 또는 심리적으로 불안한 콘텐츠
	Excp.	문화적: 전설이나 민담에 등장하는 공포 요소 교육적: 위험한 상황에 대한 경각심 고취
약물 (DR)	7+	약물에 대한 간략한 언급
	12+	음주 또는 흡연에 대한 간접적인 설명
	16+	약물 사용에 대한 설명 또는 제안
	18+	불법 약물 남용 및 중독
	Excp.	건강 관리: 치료 목적의 약물 사용에 대한 설명 교육: 약물 남용의 위험성에 대한 경고 교육
언어 (LG)	7+	약간의 부드러운 속어
	12+	상대방에게 불쾌감을 주는 언어
	16+	강한 욕설 또는 지속적인 언어
	18+	매우 노골적으로 불쾌감을 주는 언어
	Excp.	교육: 언어 예절에 관한 수업
모방위험 (RI)	7+	안전한 활동 묘사
	12+	일부 위험 요소가 포함된 활동
	16+	위험하거나 불법적인 행위 묘사
	18+	매우 위험하거나 불법적인 행위 묘사
	Excp.	교육적: 안전 규칙을 강조하기 위한 위험한 행동에 대한 설명
노출 (ND)	7+	인체의 일부 부위에 대한 간접적 묘사, 동물의 신체 묘사
	12+	인체의 일부 부위에 대한 직접적인 묘사, 동물의 신체에 대한 상세한 묘사
	16+	인체의 여러 부위에 대한 직접적인 묘사
	18+	인체에 대한 노골적인 묘사
범죄 (CR)	7+	약한 규칙 위반 묘사
	12+	약한 범죄 또는 도덕적 딜레마 묘사
	16+	범죄 또는 불법 행위 묘사
	18+	심각한 범죄 또는 폭력적인 불법 행위 묘사
사행성 (GE)	7+	간단한 확률 기반 게임
	12+	가상 화폐를 사용한 도박 요소
	16+	실제 돈과 관련된 도박 권유
	18+	실제 도박 활동의 묘사 또는 권유
	Excp.	교육적: 도박의 위험성에 대한 교육
차별 (DS)	7+	다양성과 포용성에 대한 강조
	12+	차별에 대한 간접적인 언급
	16+	차별 또는 편견에 대한 직접적인 묘사
	18+	혐오 발언 또는 차별 조장
사용자 정보 수집 (UDC)	7+	최소한의 정보 수집
	12+	사용자 동의를 통한 개인정보 수집
	16+	개인 식별 정보(PII)의 수집 및 저장
	18+ Excp.	민감한 개인정보의 수집 또는 제3자 공유

항목	나이	설명
심리적 영향 (PI)	7+	긍정적이고 단순한 도덕적 교훈 또는 메시지
	12+	사회적 관계와 적절한 행동을 이해하는 데 도움이 되는 내용
	16+	복잡한 도덕적 딜레마나 가치 판단을 필요로 하는 내용
	18+	강렬한 사상적, 정치적, 종교적 이념을 포함하여 개인의 가치관이나 성격에 깊은 영향을 줄 수 있는 내용
	Excp.	교육용: 비판적 사고와 자기 성찰을 촉진하는 철학적 또는 윤리적 내용 상담: 정신 건강 및 개인적 성장에 도움이 되는 심리적 조언 문화: 문화: 다양한 문화와 세계관을 이해하고 포용하는 데 도움이 되는 내용
신뢰 인증 (TC)	All	콘텐츠에 사용된 원본 정보의 출처를 표시, 콘텐츠가 식별 가능한 출처 없이 순수하게 AI로 생성된 콘텐츠인 경우 "AI 마크"를 표시.
	7+	신뢰할 수 있고 검증된 데이터로 학습
	12+	공공 기관 및 기타 신뢰할 수 있는 출처에 기반한 결과물
	16+	공개적으로 액세스할 수 있는 다양한 출처의 결과물
사용료 (CU)	18+	사용자가 직접 검증이 필요한 결과물
	7+	간단한 인앱 구매 또는 광고
	12+	추가 콘텐츠 구매를 유도하는 메시지 증가
	16+	반복적인 구매 유도 또는 고가의 인앱 거래
18+ Excp.	과도한 지출 요소 또는 도박과 같은 구매 유도 요소	

〈표 5〉 인구통계학적 특성

구분	빈도(명)	비율
연령대	20~29	2 (9.7%)
	30~39	4 (14.6%)
	40~49	21 (52.2%)
	50~59	9 (22.0%)
	60~	1 (2.4%)
인공지능 활용	알지 못함	1 (2.4%)
	개념만 알고 있음	3 (7.3%)
	한번 사용해봄	9 (22%)
	주당 한시간	4 (9.8%)
	주당 2~4시간	9 (22%)
	주당 5시간	4 (9.8%)
사용 분야	유료 결제 사용	11 (25.8%)
	업무	33 (80%)
	일상 생활	22 (53.7%)
	사용 안함	5 (12.2%)
그룹	해당 분야 전문가(SMEs)	14 (34.2%)
	교사(Tr)	14 (34.2%)
	학부모(Pr)	13 (31.7%)

〈표 6〉 구성요소별 기술통계

항목	내용 타당도율		평균		표준편차	
	전체	전문가 교사 학부모	전체	전문가 교사 학부모	전체	전문가 교사 학부모
선정성 (SX)	0.71	0.43 0.86 0.85	4.12	3.64 4.43 4.31	0.71	0.63 0.65 0.63
폭력성 (VL)	0.56	0 1 0.69	4.17	3.5 4.79 4.23	0.77	0.52 0.43 0.73
공포 (HR)	0.56	0 1 0.69	4.07	3.5 4.64 4.08	0.98	1.16 0.5 0.86
약물 (DR)	0.46	0 0.86 0.54	4.02	3.43 4.5 4.15	1.01	1.22 0.65 0.8
언어 (LG)	0.46	-0.14 0.86 0.69	3.9	3.21 4.5 4	1.04	1.19 0.65 0.82
모방위험 (RI)	0.51	0.14 0.86 0.54	4.05	3.43 4.64 4.08	1.09	1.28 0.63 0.95
노출 (ND)	0.56	0 1 0.69	4.15	3.64 4.71 4.08	0.88	1.08 0.47 0.64
범죄 (CR)	0.56	-0.14 1 0.85	4.2	3.5 4.79 4.31	1.03	1.22 0.43 0.85
사형성 (GE)	0.51	0 1 0.54	4.15	3.64 4.64 4.15	0.94	1.15 0.5 0.8
차별 (DS)	0.51	-0.14 1 0.69	4.07	3.5 4.57 4.15	0.88	1.02 0.51 0.69
사용자 정보 수집 (UDC)	0.61	0.43 0.86 0.54	4.24	3.93 4.64 4.15	0.83	0.92 0.63 0.8
심리적 영향 (PI)	0.66	0.29 0.71 1	4.27	3.79 4.5 4.54	0.87	1.05 0.76 0.52
신뢰 인증 (TC)	0.71	0.29 1 0.85	4.44	4.14 4.93 4.23	0.74	0.95 0.27 0.6
사용료 (CU)	0.46	0.14 0.71 0.54	4	3.64 4.43 3.92	0.92	1.01 0.76 0.86

〈표 7〉 표본 적합도 및 신뢰도 분석

분석 유형	통계량 기준	값
KMO 측정	KMO 지수	0.82
등분산 검정	카이제곱	508.24
	p-값	0.000..
크론바흐 알파	신뢰도 계수	0.95