

AI Governance in the Context of the EU AI Act

BYEONG-JE KIM¹, SEUNGHOO JEONG², BONG-KYUNG CHO³, AND JI-BUM CHUNG⁴

¹The Institute of Social Data Science, Pohang University of Science and Technology, Pohang 37673, Republic of Korea

²Research and Development Planning and Policy Department, Korea Railroad Research Institute, Uiwang 16105, Republic of Korea

³Strategy Group, INTERX, Seoul 08503, Republic of Korea

⁴Department of Civil, Urban, Earth, and Environmental Engineering, Ulsan National Institute of Science and Technology, Ulsan 44919, Republic of Korea

Corresponding author: Ji-Bum Chung (learning@unist.ac.kr)

This work was supported in part by Korea Institute of Public Administration (KIPA) Research Funding Grant; in part by the Research and Development Program of Korea Railroad Research Institute under Grant PK2503A1; in part by the Ministry of Trade, Industry, and Energy (MOTIE), South Korea, under the Global Industrial Technology Cooperation Center (GITCC) Program supervised by Korea Institute for Advancement of Technology (KIAT) under Grant P0028468; and in part by Korean National Research Foundation under Grant RS-2025-00517091.

ABSTRACT The rapid advancement of artificial intelligence (AI) has brought about significant societal changes, necessitating robust AI governance frameworks. This study analyzed the research trends in AI governance within the framework of the European Union Artificial Intelligence Act (EU AI Act). This study conducted a bibliometric analysis to examine the publications indexed in the Web of Science database. Our findings reveal that research on AI governance, particularly concerning AI systems regulated by the EU AI Act, remains relatively limited compared to the broader AI research landscape. Nonetheless, a growing interdisciplinary interest in AI governance is evident, with notable contributions from multi-disciplinary journals and open-access publications. Analysis of publications per country revealed that while the United States and China dominate AI governance research, European countries, along with the United Kingdom, also contribute significantly with a focus on specific systems restricted by the Act. Dominant research themes include ethical considerations, privacy concerns, and the growing impact of generative AI, such as ChatGPT. Notably, education, healthcare, and worker management are prominent application domains. Keyword network analysis highlights education, ethics, and ChatGPT as central keywords, underscoring the importance of these areas in current AI governance research. Subsequently, a comprehensive literature review was undertaken based on the bibliometric analysis findings to identify research trends, challenges, and insights within the categories of the EU AI Act. This review revealed critical gaps in research concerning regulated AI systems, highlighting the need for more focused research aligned with the Act's regulatory framework. The findings provide valuable insights for researchers and policymakers, informing future research directions and contributing to developing comprehensive AI governance frameworks beyond the EU AI Act. Crucially, the study identifies a significant lag between AI technological advancement and the development of policy and regulation, especially concerning specific AI systems categorized as high-risk by the EU AI Act.

INDEX TERMS Artificial intelligence, bibliometric analysis, EU AI Act, governance, research trend, Web of Science.

I. INTRODUCTION

Advancements in artificial intelligence (AI) technology have progressed at an unprecedented pace. The proliferation of generative AI models, such as ChatGPT, is accelerating

The associate editor coordinating the review of this manuscript and approving it for publication was Derek Abbott¹.

transformations in daily life. Public interest initially centered on industrial capabilities, such as efficient computation and process optimization. Generative AI has gained prominence, sparking both excitement and worries about its effects on society [1].

While AI offers numerous benefits, including increased industrial efficiency and enhanced convenience, it also poses

novel risks, such as privacy violations and ethical dilemmas [2], [3]. Given the extensive influence and reach of AI, experts increasingly warn that its unregulated proliferation, absent adequate societal oversight, could precipitate irreversible societal disruption [2], [4].

According to Maslej et al. [5], the growing number of AI misuse and abuse cases worldwide underscores the escalating risks AI poses to society. These risks encompass a range of concerns, including bias and discrimination, security breaches, privacy violations, and overreliance [2], [4], [6]. A balanced approach, integrating policy and societal discussions alongside technological development, is therefore essential. Consequently, the need for AI governance, which involves the development of technical, policy, and social solutions to address these risks, is rapidly increasing [4], [7]. Mäntymäki et al. [8] have defined AI governance as “a system of rules, practices, processes, and technological tools that are employed to ensure an organization’s use of AI technologies aligns with the organization’s strategies, objectives, and values; fulfills legal requirements; and meets principles of ethical AI followed by the organization.” AI governance, therefore, plays a critical role in minimizing risks and ensuring ethical, legal, and social responsibility throughout the AI system lifecycle, from development to deployment and use.

The rapid evolution of AI has dramatically amplified its impact across all societal sectors, fueling the demand for robust ethical AI standards and regulatory frameworks. Reflecting this urgency, international organizations such as the United Nations have underscored the critical need for global AI governance, calling for intensive international action [7], [9]. In line with these global calls, leading nations in AI technology, including the United States, Canada, and the United Kingdom, are intensifying their efforts to establish regulations for AI’s safe development and application. In the United States, the introduction of the Digital Equity Act (2020) and Blueprint for an AI Bill of Rights (2023) represent significant steps toward building a framework for trustworthy and secure AI systems. The European Union (EU) has adopted a proactive approach by proposing the AI Act in April 2021, a legislative framework aimed at guiding technological advancements to mitigate the potential societal implications of AI. Finalized in the second quarter of 2024, the Act’s core principle is the restriction or prohibition of AI systems that pose risks to human rights and ethical standards.

Despite the implementation of policies for the development of safe and reliable AI systems, academic research on AI has predominantly concentrated on technological development and model optimization [4]. The technology-centric nature of the field has resulted in relatively limited researcher interest in safe and ethical AI. This imbalance could widen the gap between technological advancements and governance, potentially creating a misalignment between technology and societal norms.

To foster responsible and ethical AI development, research into AI policy is urgently needed [10]. However, most studies on its research trends still focus on technical aspects, with

TABLE 1. Classification and description of prohibited and High-Risk AI systems under the EU AI Act.

AI system	Type	Description
Prohibited AI Systems	P1	Subliminal manipulation
	P2	AI systems that exploit people’s vulnerabilities
	P3	AI systems that evaluate or classify people based on their social behavior or personal traits
	P4	AI systems that predict a person’s risk of committing a crime
	P5	Untargeted scraping of facial images from the internet or CCTV (Closed-Circuit Television) footage
	P6	AI systems that infer emotions in the workplace or educational institutions
	P7	AI systems that categorize people based on their biometric data
	P8	Real-time remote biometric identification in publicly accessible spaces
High-risk AI Systems	H1	Biometric identification and categorization
	H2	Critical infrastructure management
	H3	Educational and vocational training
	H4	Employment, worker management, and access to self-employment
	H5	Access to essential private and public services
	H6	Law enforcement
	H7	Border control and migration management
	H8	Administration of justice and democratic processes

insufficient consideration for governance and policy implications. This gap hinders the development of effective and ethical AI regulations. This study aims to bridge this gap between AI technology and social norms by conducting a bibliometric analysis combining a literature review of AI governance research performed within the framework of the EU AI Act. It specifically aims to find the dominant research themes and trends within the relevant studies. The findings will help researchers in various fields understand the trends and key issues in AI governance research, providing a foundation for further research on responsible AI development, and insights into how AI-related governance and policies have evolved in response to regulatory frameworks.

II. THE EU AI ACT

The EU AI Act (Regulation (EU) 2024/1689) aims to ensure the trustworthy development and deployment of AI systems for EU citizens while fostering a robust AI ecosystem. Moreover, the Act strives to protect individual safety and fundamental rights, strengthen transparency and accountability, promote innovation, and maintain a stable market environment. The EU AI Act defines an AI system as a machine-based system designed to operate with varying degrees of autonomy and adaptability, and produces outputs such as predictions, content, recommendations, or decisions, influencing physical or virtual environments. The Act adopts a risk-based regulatory framework, categorizing AI systems

based on their potential risks to health, safety, and fundamental rights. AI systems with unacceptable risks are classified as “prohibited AI systems,” while systems with higher but acceptable levels of risk are categorized as “high-risk AI systems.” These two categories form the primary focus of the regulation. The risk-based categorization within the EU AI Act aligns conceptually with emerging international AI risk management frameworks, such as the trustworthiness criteria outlined in the NIST AI RMF [11]. However, a key distinction lies in the Act’s binding legal status versus typically voluntary guidelines. Table 1 summarizes the types of AI systems classified as prohibited or high-risk under the EU AI Act and provides descriptions of each.

The prohibited systems, which are subject to the strictest regulation, include those that deploy subliminal techniques to manipulate human behavior, exploit human vulnerabilities, or violate fundamental rights. These systems contravene the EU’s core values of human dignity, freedom, equality, non-discrimination, democracy, and respect for the rule of law. Examples of prohibited systems include AI systems that manipulate individuals through subliminal techniques beyond their conscious awareness, systems exploiting specific vulnerabilities (e.g., age, disability, or socio-economic status), general-purpose social scoring systems (e.g., systems analogous to China’s Social Credit System), and real-time remote biometric identification systems in publicly accessible spaces. Additionally, systems that infer sensitive information from biometric data, predict criminal risks, indiscriminately collect biometric data from non-targeted individuals, or automatically recognize the emotions of workers or students are included in the prohibited category.

The EU AI Act clearly distinguishes between the concepts of Foundation Models and General-Purpose AI Systems. Following the so-called “ChatGPT shock,” the regulation of foundation models emerged as a priority. However, the EU chose to regulate not the models themselves but the risks associated with the AI systems and applications that integrate such models. Accordingly, the EU AI Act explicitly incorporates foundation models within its risk-based framework, focusing on assessing and regulating risks at the system level rather than targeting the models directly. This approach provides a theoretical basis for this study’s analysis of governance trends at the AI system level.

Globally, there is also a growing movement toward AI regulation, including UNESCO’s recommendation on the ethical considerations of AI [3], [10]. In October 2023, the United States issued an executive order aimed at the development and risk management of AI systems, with state-level regulations under active discussion. In California, home to leading AI innovators, legislative proposals targeting AI developers are being actively debated. However, following a change in United States’ leadership in January 2025, the Trump administration swiftly repealed this executive order, signaling a shift towards accelerating AI innovation with minimal regulatory constraints. In China, the government has implemented direct and strict regulations, including algorithm registration and

review systems, AI content censorship, and data verification. Specifically, China has implemented interim measures for generative AI and adopted a vertical regulatory approach. The European Union, in contrast, has adopted a structured, risk-based strategy through its AI Act. These diverse approaches reflect distinct philosophical underpinnings: market-driven innovation in the United States, state control and economic growth in China, and ethical considerations in the EU. Japan, on the other hand, announced AI Governance Guidelines for Business Version 1.0 (April 2024) and has established an organization named “AI Safety” to review evaluation and risk assessment methods for ensuring AI safety. South Korea, on December 26, 2024, passed a comprehensive AI legislative framework through its National Assembly plenary session, becoming the second in the world after the EU to establish such an overarching AI regulatory act. This global landscape underscores the challenge of balancing AI advancement with responsible governance on an international scale.

Based on the aforementioned background, this research aims to analyze the trends in academic research related to AI governance. Specifically, the study will focus on the research trends for core AI systems regulated by the EU AI Act. While the Act’s legal jurisdiction is limited to the EU, its significant international implications and role as a leading framework make it a valuable lens for examining global AI governance research. To achieve this purpose, the following research questions are posed:

1) What is the extent of progress made in research on both overall AI governance and governance of prohibited and high-risk AI systems covered by the EU AI Act?

2) Which countries are leading in researching AI governance related to the EU AI Act?

3) Through text mining, what are the predominant research topics in the field of AI governance?

This study seeks to contribute by providing answers to these questions and gaining insights into the current state and challenges of AI governance research, as well as identifying potential future research directions.

III. METHODOLOGY

Bibliometric analysis offers a data-driven approach to comprehensively understand a research field. It is effective for exploring interdisciplinary and convergent research areas [12], [13]. Quantitatively assessing the metadata of publications allows for the systematic organization of core concepts and research methods embedded within a large number of academic publications. It also facilitates the effective sharing of synthesized scientific knowledge among researchers. The sharing of systematically organized and analyzed bibliographic metadata inspires other researchers, serving as a driving force for knowledge dissemination [14], [15], [16].

This study employed text mining-based keyword analyses to examine which research topics have been assessed in the context of AI governance research. The research trends of AI governance studies relevant to the EU AI Act were analyzed

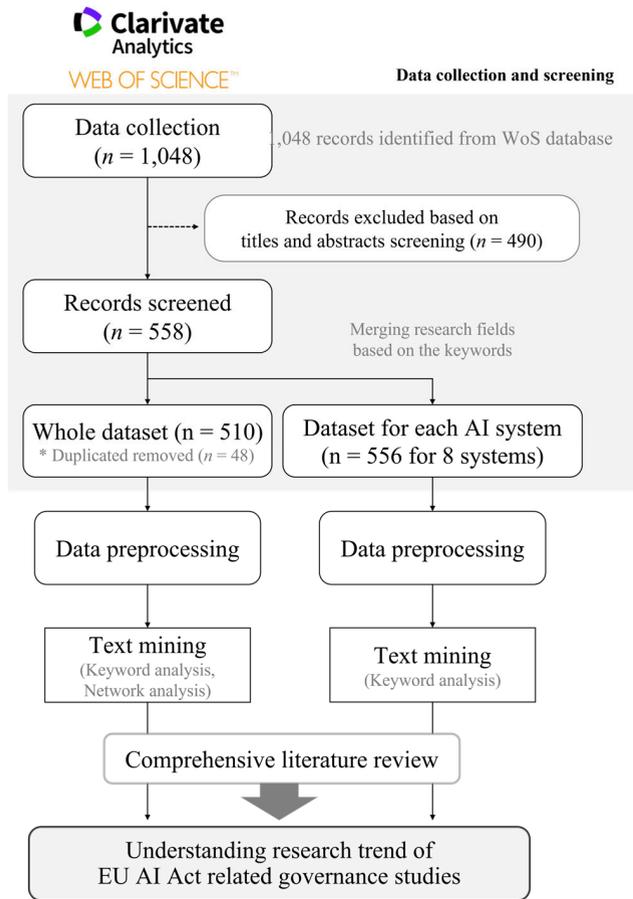


FIGURE 1. Research flow summary.

based on the research framework shown in Fig. 1. This study consists of: (1) data collection, (2) data screening, (3) data preprocessing, (4) text mining, and (5) understanding of the research trends.

Subsequently, a comprehensive literature review was undertaken based on the bibliometric analysis findings to identify research trends, challenges, and insights within the categories of the EU AI Act.

A. DATA COLLECTION AND SCREENING

Relevant publications were collected from the Web of Science (WOS) Core Collection database, utilizing the “topic” search option. The “Topic” field in the Web of Science encompasses the title, abstract, author keywords, and Keywords Plus. The search was conducted on October 11, 2024, with no restrictions placed on the publication period. Using the WOS Core Collection facilitates robust bibliometric analysis due to its citation quality standards. Nonetheless, it should be acknowledged that reliance on this single database restricts coverage, particularly for non-English literature and preprints often indexed in databases like Scopus or arXiv.

Search keywords were structured as “(AI OR “artificial intelligence”) AND (regulation OR governance) AND *System-Specific Keywords* to ensure the retrieval of studies

related to AI governance. These keywords were established by referring to prior research and the provisions of the EU AI Act [6], [17]. Upon reviewing the search keywords for the 16 types of AI systems restricted by the AI Act, it was found that the keywords for P7 and P8 of the Prohibited AI Practices, and H1 of the High-Risk AI systems were identical. Consequently, this study consolidated similar AI systems into nine distinct categories. The research trends for each redefined system were then searched. The specific search keywords for each AI system are detailed in Table 2.

While the initial search aimed to include studies highly relevant to AI governance based on the keywords, the collected data might include studies with varying degrees of relevance. Therefore, a screening process was conducted to exclude studies with low relevance. The authors assessed the relevance of the retrieved publications to relevant AI systems based on titles and abstracts. Following this process, the final papers for analysis were selected.

TABLE 2. WOS search keywords for Redefined AI Systems.

Redefined AI System	EU Definition	Search keywords
A. Subliminal manipulation	P1	(subliminal OR manipul* OR decept*)
B. Exploiting vulnerabilities	P2	(child* OR disab* OR elder*)
C. Social scoring	P3	("social behavior" OR "personal trait")
D. Predictive policing	P4	(crime OR criminal)
	H6	("law enforcement")
	H7	("border control" OR visa OR asylum OR immigra* OR emigra*)
E. Biometric identification and categorization	H8	("judicial authority" OR election OR referendum OR vote)
	P5	("facial recognition")
	P6	("emotion infer*" OR "emotion recog*" OR "emotion identif*")
	P7 p8	("biometric categor*" OR "biometric identif*")
F. Critical infrastructure	H2	(critical infrastructure)
G. Education and training	H3	("education" OR "vocational")
H. Worker management	H4	(employment OR employee* OR worker* OR recruit*)
I. Essential private and public services	H5	("private services" OR "public services" OR "essential services" OR "triage" OR "credit score" OR creditworthiness OR insurance)

B. DATA PREPROCESSING

The present work performed data preprocessing to ensure that the collected bibliographic information was appropriate for text analysis. Data preprocessing was performed using the R programming language. This process included lemmatization and the removal of repetitive phrases and stopwords.

Text refinement and analysis were conducted on the titles, abstracts, and keywords, respectively. Each word was lemmatized to its standardized form using the `textstem` package in R. This process ensures that variations of a word, such as plurals or different verb tenses, are treated as a single term, thereby enhancing the efficiency of the analysis. Repetitive phrases, such as copyright notices from publishers, were also removed. Furthermore, a list of stopwords was added to the basic stopwords list in the `tm` package to eliminate terms related to AI technologies and models (Table 3). As this study does not focus on AI research tools and techniques, additional stopwords were incorporated based on prevalent AI research tools and techniques found in existing studies [1], [6].

TABLE 3. List of Additional Stopwords.

technology, good, new, analysis, internet of thing, systematic review, literature review, scope review, large language model, natural language, machine learn, neural network, deep learn, federate learn, reinforcement learn, learn, model, ai, artificial, intelligence, differential, equation, dynamic, datum, system, use, can, result, show, train, novel, algorithm, linear, graph, ann, network, deep, code, real, world, e.g., u.s., et al, state, art, time, large scale, open source, fine tune, long term, high, task, dataset, recent, github, http, https, com, url, support vector machine, due, play, role, make, base, much, one, two, work, may, demonstrate, need, will, first, many, far, within, aim, end, predict, optimization, via, towards, propose, process, approach, method, provide, paper, research, study, also, present, however, include, aai fss, proceeding, challenge

C. TEXT MINING

For descriptive analysis, statistics on the annual publication trends and the countries of corresponding authors were compiled. Text analysis was performed by calculating keyword frequency. Furthermore, to determine the context of key terms, network analysis was conducted on the author-provided keywords. Network analysis is a bibliometric method that interprets the connections and interactions among various actors in a network structure, analyzing relationships formed by interactions within complex systems [18]. This study employed Gephi, an open-source network analysis software, to perform the network analysis [19].

A network consists of nodes and edges (links). In the context of author keyword networks, nodes represent keywords used in the studies, and edges represent the co-occurrence frequency of keywords within the same study. The edge weight is proportional to the number of co-occurrences between keywords. For instance, if two keywords appear together in one paper, an edge weight of one is assigned. The network was designed as undirected, assuming no directionality between keywords. The resulting network was analyzed using centrality measures and a community detection algorithm.

Centrality is a widely used method to assess the importance of a node relative to others in the network [20]. Among various centrality measures, this study focused on degree centrality to gauge the overall research interest in specific topics. Degree centrality indicates the number of connections a node

has with other nodes. A higher degree centrality value suggests that the node appears more frequently with other nodes. Analyzing centrality values helps identify which research topics are considered important in the network. Specifically, degree centrality quantifies direct research interest in topics, offering a robust indicator of thematic prominence. To explore connections between themes, betweenness centrality was analyzed, which measures a keyword's role as a 'bridge' based on its position on shortest paths between other nodes, thereby pinpointing keywords that act as vital hubs linking different research domains. Identifying these bridging keywords is particularly useful for highlighting the interdisciplinary intersections inherent in AI governance research. Important keywords were identified based on centrality metrics. By examining the edge weights, which signify the co-occurrence frequency between keywords, the main research themes and the relationships between them in AI governance research were identified.

Community clusters within the keyword network were identified using modularity analysis, a technique that measures the strength of network division into modules [21], [22]. Modularity analysis was performed in Gephi, allowing for the interpretation of each cluster and their interrelationships based on their constituent keywords.

IV. RESULTS

A. DESCRIPTIVE ANALYSIS

As of October 2024, a search of the Web of Science database using the keywords (“AI” OR “artificial intelligence”) yielded over 220,000 studies. When the search was refined to “(AI OR “artificial intelligence”) AND (governance OR regulation),” approximately 6,000 studies were identified.

A search for research trends related to the 16 AI systems defined in the EU AI Act identified 1,048 studies. After a review of titles and abstracts, 490 studies with low relevance were excluded. This process resulted in 558 studies related to governance and regulation relevant to the EU AI Act. After removing duplicates between each system, 510 unique documents were identified. This represents approximately 7% of the research related to AI governance and regulation. The 510 unique publications were analyzed for the overall research trend analysis and network analysis. To understand trends in each field, analyses were conducted separately for each AI system.

The number of publications for each system is presented in Table 4. The field with the most relevant literature was Type G (186 studies), which involves the use of AI in education or vocational training and learner assessment. This was followed by Type A systems (86 studies), which influence individual choices through AI. Type B systems, which target vulnerable populations, and Type H systems, which are related to worker management, also featured prominently, with 64 studies each.

Despite the recent surge in interest in AI-related research, search results show that studies on Type C system were relatively scarce. Only two studies on this system were identified

TABLE 4. Number of publications for each AI system after screening.

Redefined AI System	AI System (EU AI Act)	Records
A. Subliminal manipulation	P1	75
B. Exploiting vulnerabilities	P2	64
C. Social scoring	P3	2
D. Predictive policing	P4, H6, H7, H8	47
E. Biometric identification and categorization	P5, P6, P7, P8, H1	40
F. Critical infrastructure	H2	28
G. Education and training	H3	186
H. Worker management	H4	64
I. Essential private and public services	H5	52

in this research. Due to this limited number of documents, the authors determined that a keyword-based bibliometric analysis would not be effective for this category. Thus, for the system-specific analysis, research trends were examined using 556 studies across the eight remaining AI systems.

AI governance-related literature began to be published in significant numbers from 2017 onwards, with a notable surge in the number of related studies in the 2020s (Fig. 2). Despite the search results for 2024 spanning only from January to October, the number of published studies in most fields has already surpassed that of the previous year. Types A, B, G, and H systems, in particular, have shown a consistent trend of annual increases in the number of published studies. Notably, Type B systems saw a significant jump in publications, with 28 papers published up to October 2024, compared to only 8 in the previous year. Similarly, for Type G systems, the number of papers published more than doubled year over year, from 37 to 95.

Table 5 presents the journals and publishers with the highest number of publications related to this research. *Sustainability* published the most studies, with a total of 14. *Computer Law & Security Review* (12 articles), *IEEE Access* (7 articles), and *Government Information Quarterly* (7 articles) also demonstrated significant interest in AI governance research. These results indicate that research on AI governance is being conducted across various disciplines, including not only technical journals but also multidisciplinary journals, policy-focused journals, and journals specializing in education and healthcare. Notably, the proportion of open-access journals and publishers was particularly high.

Table 6 displays the number of publications per country based on the corresponding author's affiliation. The analysis of publication counts across countries reveals variations in research directions and approaches to AI regulation and governance, contingent upon specific national contexts.

The analysis of the number of publications by country generally shows that AI governance research is most active in the United States (U.S.), followed by China, the United Kingdom (U.K.), and Australia. The U.S. has published the most publications in eight of the nine AI system categories, the

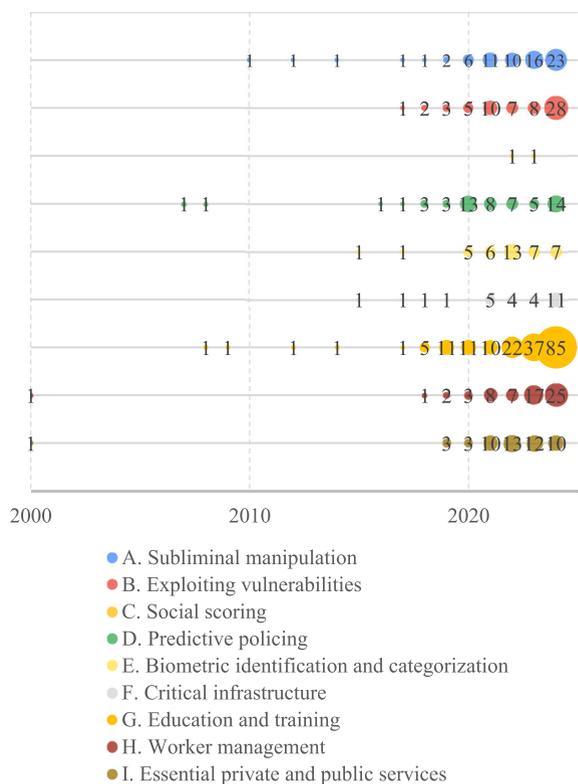


FIGURE 2. Number of annual publications by AI Systems.

TABLE 5. Top journals and publishers publishing relevant research.

Journal (Publisher)	Number of articles
<i>Sustainability</i> (MDPI)	14
<i>Computer law & security review</i> (Elsevier)	12
<i>IEEE Access</i> (IEEE)	7
<i>Government information quarterly</i> (Elsevier)	7
<i>British Journal of Educational Technology</i> (Wiley)	6
<i>Frontiers in Public Health</i> (Frontiers)	6
<i>Learning Media and Technology</i> (Taylor & Francis)	6
<i>Technology in Society</i> (Wiley)	5
<i>Applied Sciences-Basel</i> (MDPI)	4
<i>Education and Information Technologies</i> (Springer)	4
<i>International Journal of Educational Technology in Higher Education</i> (Springer)	4
<i>Journal of Computer Assisted Learning</i> (Wiley)	4
<i>Journal of Medical Internet Research</i> (JMIR)	4
<i>Policy And Internet</i> (Elsevier)	4

exception being Type G. China shared the top position with the U.S. for Type B and E systems and published the most studies in Type G systems. The U.K. is one of the countries with a high level of interest in AI research in general. It tied with the U.S. for the highest number of publications on Type D systems and had the third-highest number of publications on Type B, H, and I systems, after the U.S. and China. For Type G systems, the order of publication count was China,

TABLE 6. Number of publications per country based on the Corresponding Author’s Affiliation.

AI System	Top publication countries (n. of publications)	Records
A. Subliminal manipulation	USA (11), China (9) , Italy (5), England, Germany, South Korea (4), Malaysia, Netherlands, Scotland, Spain, Wales (3), Australia, Canada, Israel, Japan, Russia (2)	75
B. Exploiting vulnerabilities	China (13), USA (13) , England (8), Greece (4), Canada, Italy, Netherlands, Norway, Portugal, Singapore, Sweden, Wales (2)	64
C. Social scoring	Germany, USA (1)	2
D. Predictive policing	England, USA (7) , China, Netherlands (4), Russia, Spain (3), Germany, South Korea (2)	47
E. Biometric identification and categorization	China, USA (7) , Belgium, Canada, England, Netherlands, Portugal (2)	40
F. Critical infrastructure	USA (5) , China, India (3), Italy (2)	28
G. Education and training	China (35), USA (32) , Australia, England (13), Germany (9), Canada, Italy, South Africa, Sweden (5), Greece, India, Scotland, Spain, Switzerland (4)	186
H. Worker management	USA (10), China (7) , England, Spain (4), Australia, Poland, Taiwan (3), Germany, Romania, Serbia, South Africa, Ukraine (2)	64
I. Essential private and public services	USA (7), Australia, England (6) , India, Italy, Taiwan (3), China, Estonia, Finland, France, Germany, Netherlands, Spain (2)	52
Total	USA (84), China (76), England (40) , Australia (27), Germany (20), Italy (18), Spain (15), Netherlands (14), Canada (12), India (12), Taiwan (9)	510

the U.S., Australia, and the U.K., which is consistent with the findings of previous research on AI in education [23]. Regarding Type I systems, the U.S. published seven studies, while China published only two, indicating a difference in the level of interest. In the case of Type I systems, several studies were also published in Australia and the U.K., in addition to the U.S.

B. RESEARCH TRENDS OF AI SYSTEMS

1) KEYWORD ANALYSIS

To identify the primary areas of interest in the relevant research, this study analyzed frequently occurring keywords. Table 7 presents the list of the most frequent keywords in titles, abstracts, and author-provided keywords after the removal of stopwords. As revealed in the search results of the relevant literature, studies related to education were prevalent across titles, abstracts, and keywords. Key AI governance-related keywords identified include risk, ethic(al), policy, and privacy.

The prominence of terms such as generative AI and ChatGPT indicates that alongside the expanding application scope of AI systems, there is a growing body of research focused on the ethical considerations and potential side effects of their deployment. In particular, the keyword analysis highlighted a significant interest in AI ethics, alongside numerous

education-related keywords. Furthermore, frequent mentions of the EU’s General Data Protection Regulation (GDPR) and related terms, as well as transparency, were observed in relation to potential privacy infringements associated with AI technology use.

TABLE 7. Most frequent words in Titles, Abstracts, and Keywords of AI Governance studies related to the EU AI Act.

Rank	Title		Abstract		Author keywords	
	Word	Freq	Word	Freq	Word	Freq
1	education	72	education	358	ethic	27
2	future	29	development	276	education	26
3	public	29	human	265	chatgpt	23
4	human	28	application	234	big_data	14
5	application	26	potential	233	generative_ai	13
6	health	26	digital	230	chatbot	13
7	digital	25	public	221	high_education	13
8	framework	24	risk	210	privacy	13
9	healthcare	21	service	208	data_protection	12
10	ethical	20	ethical	206	gdpr	9
11	service	19	health	204	transparency	9
12	generative	18	policy	204	learn_analytics	8
13	policy	18	framework	201	self-regulate_learn	8
14	social	18	design	197	self-regulation	8
15	chatgpt	17	healthcare	196	child	7
16	design	17	decision	194	explainable_ai	7
17	self	17	develop	187	innovation	7
18	law	16	tool	181	public_service	7

Table 8 presents the top 10 author keywords for each system type, excluding common keywords (e.g., AI, regulation, governance) and system-specific keywords (e.g., insurance, education, employee) used for the search.

In Type A systems (Subliminal Manipulation), keywords such as human, market, design, and image appeared frequently. This suggests active research into marketing applications of this technology and image-focused techniques for influencing the subconscious. The frequent appearance of the keyword ethical also indicates that the ethical considerations of this technology are being raised.

For Type B systems (Exploiting Vulnerabilities), keywords such as healthcare, social, and design appeared frequently, reflecting a high level of interest in medical technologies for vulnerable populations, including the elderly and children. Although not shown in Table 8 due to the removal of the search terms, the primary subjects of research in this area were children (appearing 84 times) and the elderly (appearing 31 times). The main research topics also included AI-based education methods for children and personalized services for vulnerable populations. The keyword risk reflects concerns about the potential dangers and safety of AI systems targeting vulnerable groups.

System A: Subliminal manipulation, B: Exploiting vulnerabilities, D: Predictive policing, E: Biometric identification and categorization, F: Critical infrastructure, G: Education and training, H: Worker management, I: Essential private and public services.

TABLE 8. Frequently used terms in abstract by AI Systems.

	System A		System B		System D		System E	
	Word	Freq	Word	Freq	Word	Freq	Word	Freq
1	human	65	care	60	legal	50	ethical	26
2	market	41	social	54	risk	36	privacy	23
3	development	38	design	47	framework	31	social	23
4	activity	35	development	44	digital	30	human	21
5	ethical	34	healthcare	41	public	28	information	21
6	potential	34	health	38	human	27	development	20
7	design	33	human	38	decision	26	law	19
8	application	31	policy	37	policy	26	service	19
9	image	31	risk	35	social	26	application	18
10	control	30	develop	34	development	24	right	18
	System F		System G		System H		System I	
	Word	Freq	Word	Freq	Word	Freq	Word	Freq
1	digital	31	student	175	human	57	health	50
2	service	25	chatgpt	129	digital	48	government	43
3	solution	21	development	108	management	48	healthcare	43
4	control	20	application	106	health	35	patient	40
5	transformation	20	potential	106	potential	32	sector	37
6	public	19	healthcare	95	safety	32	risk	34
7	quality	19	ethical	89	job	31	information	28
8	information	18	support	89	labor	30	policy	28
9	integrate	17	practice	87	decision	28	enable	27
10	decision	16	tool	83	framework	28	framework	27

System A: Subliminal manipulation, B: Exploiting vulnerabilities, D: Predictive policing, E: Biometric identification and categorization, F: Critical infrastructure, G: Education and training, H: Worker management, I: Essential private and public services.

In Type D systems (Predictive Policing), keywords such as legal, risk, framework, digital, and policy were frequently used. This relates to active discussions on the ethical and legal risks, including decision-making and legal accountability when using AI in law enforcement and judicial systems. It also reflects scholarly interest in the digital transformation of the judicial system, policy changes, and societal impacts.

For Type E systems (Biometric Recognition), keywords such as ethical and privacy were predominantly featured. This indicates ongoing discussions regarding potential privacy violations, ethical issues, and legal regulations related to the use of biometric recognition technologies. The frequent appearance of keywords like social and human suggests that emotion recognition technology is mainly used for analyzing social interactions.

In Type F systems (Critical Infrastructure), keywords such as digital, public, and service appeared frequently, indicating that this system category includes research related to digital transformation such as smart cities, public service improvement, and infrastructure management. Research has primarily focused on the application of AI systems in public policy, public services, and urban infrastructure.

In Type G systems (Education), keywords such as student, chatgpt, and tool appeared frequently. This suggests that AI systems in this field are mainly used in AI-based educational tools, learner assessment, and personalized learning

support. Notably, the keyword healthcare appeared frequently, demonstrating significant interest in the application of AI technology in medical education. The keyword ethical also appeared frequently.

In Type H systems (Worker Management), keywords such as human, digital, health, safety, and management ranked highly, indicating active research on AI-based management systems for worker safety, health management, and work efficiency improvement.

Finally, in Type I systems (Essential Private and Public Services), keywords related to healthcare, such as health, healthcare, and patient, appeared frequently. The frequent appearance of these healthcare-related terms indicates a high level of interest in the application of AI in essential public services, particularly in healthcare and insurance. It also suggests that patient privacy protection is a major issue in the deployment of such AI systems.

2) NETWORK ANALYSIS

To provide a comprehensive view of the knowledge structure across the AI governance research, a network analysis based on author keywords was conducted (Fig. 3). Due to the insufficient number of studies in each specific field, publications from all system types were combined for analysis. The co-occurrence analysis of author keywords reveals that AI ethics and privacy concerns related to the use of AI technology, along with research in the field of education, are major areas of interest in the related research. Analyzing 510 publications (excluding duplicates between systems) identified 4,730 co-occurrence relationships among 1,478 keywords. The network was visualized based on relationships between keywords that co-occurred at least twice, to simplify the network structure and highlight the most important keywords. As a result, a network of 75 keywords and 132 connections were derived, with a maximum co-occurrence frequency of five.

In the network, the size of the nodes was set to be proportional to their weighted degree centrality. The thickness of the edges was configured to be proportional to the co-occurrence frequency of the two nodes.

The top keywords based on degree centrality were education, ethic, ChatGPT, regulation, governance, privacy, chatbot, and GDPR (Table 9). Excluding the keywords used in the search, ChatGPT, privacy, chatbot, GDPR, big data, transparency, and generative AI exhibited high degree centralities. Notably, big data and ethic were the keywords with the highest betweenness centrality (0.073 and 0.067, respectively), indicating they are central themes in recent AI governance research. High betweenness centrality identifies keywords that bridge different research areas, emphasizing the interdisciplinary connections in AI governance.

There are ten pairs that showed a co-occurrence frequency of four or more. The most frequent co-occurring keyword pairs were AR-VR (Augmented Reality-Virtual Reality) (5 times), ChatGPT-Education (5), Generative AI-ChatGPT (5), regulation-ethic (5), GDPR-privacy (4), ChatGPT-ethic (4),

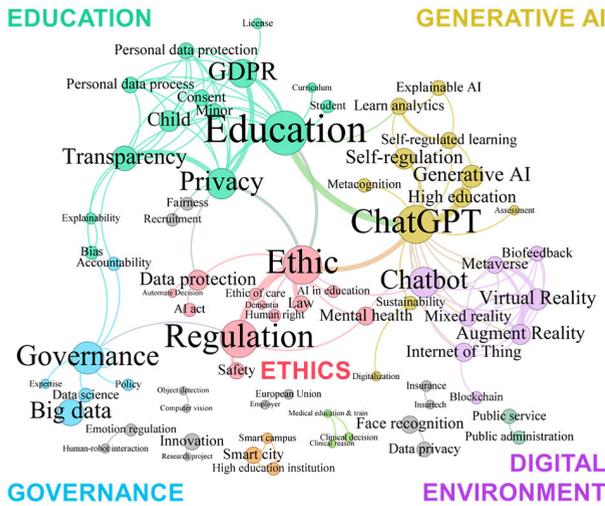


FIGURE 3. Keywords network of AI governance publications. Node size represents weighted degree centrality, and edges represent co-occurrence relationships. Node color represents each cluster detected by modularity: green (Education), red (Ethics), yellow (Generative AI), purple (Digital environment), and blue (Governance).

TABLE 9. Top 20 keywords based on weighted degree centrality (WDC).

Rank	Keyword	DC	WDC	BC	EC
1	education	102	121	0.097	1.000
2	ethic	82	103	0.067	0.546
3	chatgpt	74	101	0.062	0.825
4	regulation	91	99	0.074	0.117
5	governance	75	83	0.064	0.083
6	privacy	60	77	0.027	0.893
7	chatbot	64	76	0.034	0.734
8	gdpr	58	70	0.038	0.806
9	big data	63	64	0.073	0.005
10	transparency	46	59	0.016	0.839
11	generative ai	44	51	0.029	0.157
12	virtual reality	34	47	0.006	0.536
13	self-regulation	43	46	0.023	0.130
14	child	38	46	0.016	0.795
15	data protection	40	44	0.021	0.188
16	augment reality	32	44	0.005	0.536
17	high education	35	40	0.016	0.127
18	covid-19	39	39	0.021	0.000
19	internet of thing	31	39	0.024	0.545
20	mental health	34	37	0.009	0.182

DC: Degree centrality, WDC: Weighted degree centrality, BC: Betweenness centrality, EC: Eigenvector centrality.

high education-ChatGPT (4), law-ethic (4), metaverse-VR (4), and transparency-privacy (4). ChatGPT showed high co-occurrence with keywords such as education, generative AI, ethic, and high education, suggesting active discussions related to the use and regulation of generative AI technology in those fields.

Community analysis, based on modularity, identified five major clusters. In Fig. 3, each keyword is color-coded

according to its cluster. The main clusters were related to education, ethical, and generative AI (ChatGPT), which correspond to the keywords with the highest weighted degree centrality. Additional clusters were identified related to AI-related digital environments, including chatbot, AR, VR, and IoT, as well as a governance-related cluster.

The “education” cluster was centered around the keyword education, the most frequently occurring keyword, and included terms such as GDPR, privacy, transparency, and consent. This indicates that the main topics of discussion in the application of AI in education include student privacy and the need for informed consent.

In the “ethics” cluster, keywords used in the search, such as regulation, data protection, and mental health, appeared. Although the connectivity between the ethics and other keywords in this cluster was low, it suggests that data protection, regulatory compliance, and the protection of vulnerable populations are being considered in relation to AI ethics. The connection between the keywords AI Act and data protection indicates that while direct mentions of the EU AI Act in current research are relatively infrequent, when it is discussed, it is in relation to data protection.

The “generative AI” cluster, led by ChatGPT, included keywords such as generative AI, higher education, and self-regulated learning, indicating a close relationship with the field of education. This suggests that generative AI is being actively used in education, particularly to support self-regulated learning among higher education students.

The “digital environment” cluster mainly featured technologies enabling digital environments, such as chatbot, VR, AR, and metaverse. These keywords were closely linked to ChatGPT, indicating they appeared in the context of generative AI implementation.

Finally, the “governance” cluster was relatively small, despite governance being included as a search term for constructing the dataset. This cluster included keywords such as policy and data science, suggesting discussions on policies and methodologies for AI governance. Additionally, the accountability appeared alongside these terms, which was in turn connected to the transparency. This implies that key issues in AI governance include ensuring the transparency and explainability of AI systems and clarifying the accountability for the impacts caused by these systems.

V. DISCUSSION

This study comprehensively analyzes trends in AI governance research based on the EU AI Act framework. A search of the Web of Science database revealed that studies addressing AI governance and regulation are relatively scarce, particularly in the context of policies related to AI systems regulated by the EU AI Act. This supports the global discourse that policy and institutional discussions lag behind the rapid pace of AI technological advancement [9]. Following the risk-based approach outlined in the EU AI Act, the study categorized AI systems by type and analyzed research trends for each category. Findings indicate that interdisciplinary journals and

TABLE 10. Summary of literature review by AI systems.

EU AI Act Category	Key Research Areas	Major Concerns	Future Research
A. Subliminal Manipulation	Vulnerable group protection Care bot ethics Consumer transparency & Conversational AI	Unintentional manipulation Algorithm transparency Deepfake misuse	Tailored regulations (vulnerable groups, chatbots, deepfakes) Enhance AI transparency & explainability Multi-stakeholder governance collaboration
B. Law Enforcement	Biometric ID (border control) Predictive policing AI in court & democracy	AI bias & human rights Judicial authority & political misuse False information & privacy	New regulations for biometric ID & predictive policing Transparency & accountability of algorithms Ethical, legal, & accountability research on AI crime & military AI
C. Biometric Identification	Human-computer interaction Emotion recognition	Privacy & data misuse Security & explainability	Enhance reliability & public acceptance Robust data protection & legal frameworks
D. Critical Infrastructure	Digital infrastructure Transportation Energy Healthcare	Cyber security Human surveillance	Governance-focused risk analysis AI impact assessment tools Risk assessment aligned with EU AI Act
E. Education & Training	Higher education Medical training Generative AI (ChatGPT)	Academic misconduct of Generative AI Hindered critical thinking Transparency of AI based assessment	Analysis of AI's effects on learning Utilizing AI in vocational training Ensuring assessment transparency
F. Worker Management	Automated HR decisions (hiring, promoting, firing)	Ethical issue in hiring Equity and transparency Worker consent & rights Workplace data	Worker-AI interactions Protection of worker rights Development of safety standards in labor law
G. Essential Services (Public & Private)	Public service (e-government) Private service (insurance, finance, healthcare)	Reliability and ethics issues Data protection Trust & citizen engagement	Institutional measures for ethical use & reliability for AI services Ethical and legal regulation for AI use Balance tech & human interaction Public-private collaboration

open-access platforms have published a significant number of AI governance studies, highlighting the field's cross-disciplinary nature and the importance of rapid dissemination of research.

Furthermore, while the U.S. and China dominate AI governance research, cultural and national backgrounds were found to influence research trends. The observation that the absolute publication volume from EU member states is not predominant, despite the EU AI Act being a leading regulatory framework, raises an interesting question. This finding suggests that the regulation may not have significantly boosted related research activities in the short term, or perhaps that it might have even inadvertently hindered certain academic explorations. However, it is crucial to acknowledge that the effects of the EU AI Act might not yet be fully reflected in publication metrics due to potential time lags. Furthermore, this interpretation does not account for baseline differences in national research capacities, underscoring the need for more nuanced analysis of national research trends in future studies.

Text mining and network analysis identified key themes such as ethical considerations, data protection (e.g., GDPR),

and generative AI (e.g., ChatGPT), underscoring growing academic interest in ethical issues and privacy concerns. Notably, the prominence of educational applications and the strong network centrality of ChatGPT signal that generative AI is a focal point in recent governance research.

Building upon these general findings, the following subsections present a more granular qualitative literature review of research pertaining to each specific AI system category, extracting valuable insights and identifying future research directions for each.

A. SUBLIMINAL MANIPULATION AND EXPLOITATION OF VULNERABILITIES

The findings of this study indicate that current research on AI systems focuses primarily on safeguarding vulnerable groups, such as children, seniors, and individuals with disabilities, from subliminal manipulation and exploitation [24], [25], [26], [27], [28], [29], [30]. Researchers have emphasized the need for stringent regulations for AI systems targeting vulnerable groups, ethical considerations in the design and use of care bots [25], [31], transparency in consumer

decision-making support systems [32], and potential risks associated with conversational AI [33], [34]. Furthermore, the possibility of AI systems unintentionally manipulating humans has been raised [35], underscoring the urgent need for sustainable regulatory measures to enhance algorithm transparency [30], [36].

Several key concerns have been raised regarding the potential harms of AI systems. Studies have highlighted the negative impacts of media and AI exposure on children [26], underscoring the need for AI literacy education tailored to young children [29]. In healthcare, research emphasizes the importance of incorporating the perspectives of child and adolescent patients when implementing AI technologies [27]. Further research has stressed the need for ethical guidelines on using intelligent assistive technology for the elderly and individuals with disabilities, particularly addressing issues of autonomy, data management, and distributive justice [28]. The negative potential for deepfake technology, especially by facilitating online abuse and violating women's rights, has also been identified as a significant concern [30].

Overall, to minimize the risks of subconscious manipulation and exploitation of vulnerabilities, AI systems must be designed with greater transparency and explainability [30], [36]. Furthermore, it is essential to adhere to principles of personal data protection and non-discrimination. Tailored regulations are also needed to address the specific characteristics of different fields, including protecting vulnerable populations, establishing ethical guidelines for chatbots, safeguarding consumers, and regulating deepfakes. To achieve this, collaboration among various stakeholders, including technology developers, policy makers, ethicists, and citizen groups, is imperative in order to reach a social consensus and establish a sustainable AI governance system.

B. LAW ENFORCEMENT AND POLICING

In this study, we comprehensively reviewed the fields of AI-based law enforcement and policing (including predictive policing). This is because public safety and law enforcement discussions exhibit parallel trends and often encompass continuous processes.

The most prominent issues in the field of public safety were the potential ethical problems such as AI bias and human rights violations [37], [38], [39]. These discussions were frequently raised as issues in the field of remote biometric identification (RBI) systems using AI [40]. AI-based biometric identification technology is most commonly used in border control, such as passport screening for foreigners. However, using AI for border control has also raised ethical concerns [41], [42], [43]. To address these concerns, new regulations should be created and AI algorithms should be made transparent and accountable [44].

Biometric identification and predictive policing using AI naturally lead to subsequent law enforcement issues, and this discussion can be divided into optimistic claims that utilizing AI can improve the effectiveness and fairness of

court rulings [45], [46], [47], and concerns about exacerbating inequality, weakening judicial authority, increasing the possibility of political exploitation, and strengthening digital authoritarianism [47], [48]. This eventually extends to the impact of AI systems on political systems, that is, the impact on democratic systems. Analyses of experimental democratic systems using AI technology [49], [50], [51] show the possibility of a new democratic methods using AI technology, but there may also be various side effects such as the problem of false information in the election process [48], [52], and illegal use of personal information [53]. It shows that AI can be an important opportunity for future democracy while also posing a serious threat, which will be a very important research opportunity in the future.

Last but not least, there is a field that is highly underexplored. It is the field of malicious use of AI. This refers to crimes using AI (AIC; Artificial Intelligence Crime) [54], and discussions on the military use of AI [55], [56]. Technologies utilizing AI to attack other individuals or nations are being secretly researched in many countries and are even being used in actual battlefields, but there is a lack of academic research on this topic. Areas such as autonomous weapon systems (AWS) without human intervention, AI-based surveillance systems, and cyber warfare that maliciously utilize AI will occupy a very important domain of future national security, echoing concerns from the UN about their legal, security, and humanitarian implications [9]. However, academic discussions on related ethical issues, accountability, and international law are relatively lacking. Moreover, even the EU AI Act, known as the world's strongest law, does not address these issues. In the future, academia needs to approach this problem more proactively. To ensure safer utilization of AI technology in future society, it is essential to establish a new normative system for the malicious use of AI.

C. BIOMETRIC IDENTIFICATION AND CLASSIFICATION

AI-based biometric recognition and classification technologies are advancing human-computer interaction (HCI) by enabling the recognition and analysis of user facial features or emotions. Key applications include assisting individuals with autism in understanding others' emotions [57], enabling customer service to respond based on identified customer emotions [58], and developing affective tutoring systems that adapt to learners' emotions to enhance motivation [59].

However, the use of AI-based biometric recognition raises significant ethical and privacy concerns due to its reliance on sensitive personal data. Insufficient security, transparency, reliability, and explainability in processing sensitive information like emotions and facial features can lead to privacy violations and data misuse [60]. Collecting biometric data in public spaces poses particular challenges, such as difficulties in fulfilling user notification requirements and potential public resistance. The retraction of a facial recognition system for security by a Japanese railway company due to privacy concerns [61], and the potential for misuse of

biometric identification exemplified by the EU's iBorderCtrl system [62], underscore these concerns.

While AI-based biometric recognition technologies are being actively developed across various fields, in-depth discussions regarding their application in public sectors and policy-making remain scarce. This is reflected in the reservations of regulatory experts who oppose the adoption of these technologies due to concerns over privacy violations and ethical issues [63]. Consequently, to facilitate the smooth integration of AI-based biometric systems into the public sphere, efforts are needed to enhance their reliability and public acceptance. In conclusion, resolving the ethical challenges of AI-based biometric systems and ensuring their safe deployment in the public sector urgently requires advancements in explainable AI (XAI) to ensure transparency and accountability, the establishment of robust personal data protection mechanisms, and the development of clear legal regulations on the scope and methods of biometric data utilization.

D. CRITICAL INFRASTRUCTURE

The EU AI Act designates AI systems designed for safety components in critical infrastructure sectors, including digital infrastructure, transportation, energy, and healthcare, as high-risk.

In the digital infrastructure field, there is a strong focus on incorporating AI into the cyber world [64], developing control systems for biometric authentication in infrastructure security [65], implementing generative AI in the service industry [66], and addressing AI security in cloud environments [67]. These topics primarily pertain to the adoption of AI and the corresponding security concerns in the context of digital transformation. The transportation industry has placed particular emphasis on utilizing AI for automation and conducting research on cyber security. This includes the implementation of AI for monitoring intelligent railway infrastructure [68], the implementation of autonomous shuttle services [69], the utilization of autonomous aviation systems [70], and the adoption of cyber security standards to protect infrastructure [71]. In the energy sector, research has been conducted on the use of AI for transmission network design [72], power quality management [73], and protection of infrastructure against extreme weather conditions [74]. The medical field has focused on researching the implementation of AI in public healthcare [75] and AI-based elderly healthcare assistance [76].

The study of AI governance within critical infrastructure has received less attention compared to other fields, and has focused more on exploring the potential applications of AI rather than regulation. Particularly, the intelligent railway infrastructure monitoring system in the transportation sector poses a risk of unintended human surveillance. Therefore, further research is necessary to establish a governance-based framework for analyzing potential risks associated with AI use in critical infrastructure, develop AI impact assessment tools, and conduct risk assessment aligned with EU AI Act.

E. EDUCATION AND TRAINING

The use of AI in education and vocational training has been one of the most active areas of research among the AI systems covered in this study. Notably, higher education [77], [78] and the medical field [79], [80] have shown significant interest in AI technology.

As revealed in the text analysis, in higher education, studies have investigated student perceptions of AI adoption [78] and explored how AI can enhance learning outcomes and improve educational governance [77]. Alongside these investigations, there is a strong interest in utilizing generative AI, particularly ChatGPT. Over 30 studies mentioned ChatGPT or generative AI, with a significant portion focusing on how these technologies can be used to enhance students' self-regulated learning abilities [81], [82]. However, concerns have also been raised regarding the potential for generative AI to facilitate academic misconduct and hinder the development of critical thinking skills [83]. Therefore, further research is needed to analyze the multifaceted effects of AI technology on learning outcomes, motivation, and attitudes.

In the medical field, studies have aimed to improve the efficiency of training involving complex and advanced interpretive skills (e.g., X-ray, computed tomography (CT), and magnetic resonance imaging (MRI) image interpretation) through the use of AI [79], [80]. In contrast, research on AI applications in vocational training and workplace education has been relatively scarce. With few exceptions, such as research on AI-based safety training in manufacturing [84], vocational training was generally mentioned as one of possible applications of AI on improving industrial process efficiency.

Furthermore, only a few studies mentioned generative AI tools other than ChatGPT (e.g., Perplexity, Claude, and Gemini). The multimodal capabilities of generative AI were also rarely mentioned. This lack of mention reveals the perspective gap between governance research and technical research.

The EU AI Act regulates the use of AI in education, particularly in areas such as admissions, grading, educational assessment, and plagiarism detection. While few of the analyzed papers directly addressed the specific regulatory scope of the EU AI Act, many covered indirectly related topics, such as enhancing educational outcomes through AI and personalizing education. However, these areas are indirectly associated with assessments of grades and educational attainment. Thus, future efforts must focus on securing transparency in assessment processes as stipulated by the EU AI Act, developing specific measures to enhance the reliability of AI systems. Furthermore, it is crucial to perform detailed risk assessments on various generative AI, taking into account their multimodal capabilities and active use in the field of education.

F. WORKER MANAGEMENT

AI has been used in managing employees to make decisions like hiring, promoting, and firing automatically [85], [86], [87]. However, it is crucial to establish governance protocols

to safeguard workplace data [85] as there are potential ethical issues, such as discrimination. In particular, establishing an audit framework is required to enhance the equity and transparency of AI-based hiring systems [86]. Additionally, concerted efforts and strategies are necessary to mitigate potential human biases throughout the hiring process [87].

Research has demonstrated that AI-based management can lead to negative psychological and ethical consequences for workers [88]. Therefore, it is imperative to prioritize worker consent and rights when implementing AI [89]. In the era of AI, discussions are underway on the direction of changes in labor laws to protect workers' rights [90]. In addition, there are proposals to include obligations for labor groups to participate in the implementation process of AI in the workplace and develop safety standards in labor law [91].

AI tools can contribute to increased worker productivity and demonstrate the potential for human-AI collaboration [92]. While AI-based human resource management system can enhance organizational performance [93], its acceptance may vary based on the employee's locus of control [94]. Employees with an internal locus of control may view the implementation of AI as a potential threat to their employment.

In conclusion, AI-based worker management necessitates discourse on multifaceted aspects, including ethical decision-making, worker-AI interactions, protecting worker rights, and productivity enhancements. Comprehensive research on legal and institutional frameworks is crucial to mitigate ethical and rights-infringement concerns.

G. ESSENTIAL PRIVATE AND PUBLIC SERVICES

AI research in essential public services centers on improving quality of public services, enhancing citizen interactions, addressing ethical and reliability concerns in the public sector, and exploring legal and policy implications. AI has the potential to improve the effectiveness of public services [95] and enhance data-driven decision-making [96]. AI-based electronic government systems can strengthen local democracy [97]. In European countries, research has been conducted on the implementation of AI-based public services [98] and the trust and acceptance of citizens [99]. While the introduction of AI can bring about positive impacts, it is crucial to establish institutional measures to ensure ethical use [100] and reliability [49]. Furthermore, it is important to establish legal responsibility and framework for the use of AI [101], [102].

Active research explores the applications of AI in critical private services, including insurance, finance, and healthcare. AI has demonstrated enhanced reliability and customer satisfaction in insurance and financial services [103], [104]. Technological advancements enabled by AI have contributed to improved efficiency within the insurance and finance sectors [105]. In healthcare, AI holds significant promise for service optimization [106], patient flow management [107], and advancements in disease prevention and treatment [108].

The utilization of AI technology is rapidly increasing in both the public and private sectors; however, it is imperative to address critical concerns such as reliability, ethics, and data protection. For electronic government and AI-based public services, a policy-based foundation is crucial to enhance trust and citizen engagement. Ignoring security and ethical issues could hinder the private sector's long-term growth. It's essential to balance technology and human interaction, enforce ethical and legal regulations, and encourage collaboration and complementary development strategies between the public and private sectors.

VI. LIMITATIONS OF THE STUDY

Despite its contributions, this study has several limitations that simultaneously highlight key areas for future research. First, the scarcity of research on Type C systems (social scoring) precluded an in-depth analysis of this specific system. This may be attributed to the limited scope of the definitions of "social behavior" and "personal characteristics" within the EU AI Act. Future research should employ more comprehensive search terms to capture a wider range of relevant studies on this topic. Furthermore, the limited research on social scoring could be due to factors beyond the definitional scope, such as its political sensitivity, ethical concerns (e.g., China's Social Credit System), or potentially the EU AI Act's prohibition limiting academic work. It remains an open question for future investigation whether this scarcity is mainly due to regulatory discouragement or the methodological obstacles associated with studying forbidden technologies.

Second, the reliance on the Web of Science database means that studies indexed in other databases, such as Scopus and arXiv, were not included. Although WOS is recognized for maintaining citation quality standards, it provides limited coverage of publications not in English as well as preprints. Consequently, our findings might be skewed towards peer-reviewed research originating predominantly from Western countries, potentially failing to capture emerging developments from non-EU regions or within rapidly advancing technical fields. Future research should utilize multiple databases to provide a more holistic view of research trends. This is particularly important in the rapidly evolving field of AI, where many studies are shared as preprints rather than through traditional peer-reviewed outlets [6].

Third, because the regulatory landscape for AI is constantly evolving, future research could supplement the static perspective offered by bibliometric analysis with real-time monitoring of policy shifts, utilizing resources such as OpenAI's policy tracker or EU Commission updates. Furthermore, as noted earlier, the reliance on the peer-reviewed database might have limited the ability of this study to capture the most recent trends in AI research. Building on the implications of this study, future research should leverage preprint databases like arXiv along with real-time policy tracker to analyze the latest developments in AI and thereby contribute to practical policy-making.

VII. CONCLUSION

This study employed bibliometric analysis and text mining techniques to comprehensively analyze recent research trends in AI governance, specifically within the context of the emerging EU AI Act framework. The analysis reveals that research on AI governance and regulation constitutes a relatively small portion of the overall AI research landscape. Notably, studies focusing on policy aspects related to AI systems regulated under the EU AI Act are even more limited. This indicates that the development of relevant policies and regulations is lagging behind the rapid pace of AI technological advancement. Despite this, the study confirms that AI governance research is expanding across various academic disciplines, with ethics, privacy, and generative AI-related topics emerging as major areas of interest. Although our analysis primarily detected discussion surrounding ChatGPT in the existing literature, the AI landscape now includes other significant models like Gemini, Claude, Grok, and DeepSeek; considering their unique capabilities and societal influence, these emerging models, alongside the established presence of ChatGPT, are expected to be central to future governance considerations.

Furthermore, the prominence of multidisciplinary journals and open-access publications in AI governance research underscores the importance of rapid dissemination and sharing of findings in this nascent field.

The analysis also reaffirms that while the U.S. and China are leading in AI governance research, variations exist in research themes and approaches, influenced by national contexts, cultural factors, and specific policy priorities. Our findings highlight significant contributions from European countries, like the U.K., Germany, and Italy. This suggests a potential catalytic effect of the Act in fostering AI governance research, particularly given the comparatively fewer publications from Europe in broader AI research fields, as observed in prior studies [1], [23]. This study holds significance as it systematically analyzes AI governance research trends based on the concrete regulatory framework of the EU AI Act and provides a comprehensive overview of major research themes, keywords, and national research trends.

This study aimed to contribute to the existing literature by adopting the EU AI Act's classification system as an analytical framework. This approach allowed us to identify core areas where policy and governance discussions lag behind rapid AI advancements and to propose urgent tasks for future AI governance research. These tasks underscore the growing importance of ensuring transparency, explainability (XAI), and reliability in AI systems amidst increasing technological complexity. Furthermore, they highlight the urgent need for research into specific technical and institutional solutions, such as algorithm audits, fair assessment process design, and robust data quality management.

Specifically, our analysis diagnosed a relative scarcity of in-depth governance research focused on high-risk AI systems designated by the EU AI Act, particularly when compared to the academic attention given to data protection and

privacy issues. We also confirmed that legal/regulatory discussions and international consensus concerning the potential for malicious use of AI remain underdeveloped. Consequently, there is an urgent need for research focused on developing risk assessment methodologies, establishing regulatory compliance measures, and analyzing the potential societal impacts of AI systems within these high-risk domains. Moreover, as rapidly evolving generative AI, including multimodal capabilities, can amplify the risks of malicious use proportionate with its performance and functional improvements, research that evaluates the potential and risks of these new technologies in a balanced manner and seeks preemptive and adaptive governance measures emerges as a critical future task.

Furthermore, this study suggests the need for critical reflection on the potential tension between strengthening AI ethics and promoting the utilization of AI technology for vulnerable groups. The EU AI Act maintains a notably cautious stance regarding AI applications targeting these populations, which could lead towards restricting technology deployment for protective reasons. Yet, the potential for AI technology to help these groups overcome their specific vulnerabilities is also evident. Therefore, critical examination and societal discussion are required regarding whether stringent protective regulations might inadvertently restrain relevant research or limit opportunities for vulnerable groups to benefit from AI technologies. This remains an important question to be continuously explored within the broader context of AI's ethical implications and human rights protection.

Finally, our findings suggest a persistent gap in research concerning the design of governance frameworks and institutional mechanisms aimed at fostering and supporting practical cooperation and communication among diverse stakeholders including technology developers, policymakers, users, and civil society. This is a crucial element for establishing effective AI governance, as governance should be delivered through an ecosystem of critical functions where stakeholder engagement is a key component [7]. Addressing these challenges through in-depth research and concerted societal efforts is essential for cultivating a responsible and trustworthy AI ecosystem.

REFERENCES

- [1] R. Dwivedi and L. Elluri, "Exploring generative artificial intelligence research: A bibliometric analysis approach." *IEEE Access*, vol. 12, pp. 119884–119902, 2024, doi: [10.1109/ACCESS.2024.3450629](https://doi.org/10.1109/ACCESS.2024.3450629).
- [2] L. Weidinger, M. Rauh, N. Marchal, A. Manzini, L. Anne Hendricks, J. Mateos-Garcia, S. Bergman, J. Kay, C. Griffin, B. Bariach, I. Gabriel, V. Rieser, and W. Isaac, "Sociotechnical safety evaluation of generative AI systems," 2023, *arXiv:2310.11986*.
- [3] UNESCO, "Ethical impact assessment: A tool of the recommendation on the Ethics of artificial intelligence," *UNESCO*, Paris, France, 2023.
- [4] S. Fosso Wamba, R. E. Bawack, C. Guthrie, M. M. Queiroz, and K. D. A. Carillo, "Are we preparing for a good AI society? A bibliometric review and research agenda," *Technological Forecasting Social Change*, vol. 164, Mar. 2021, Art. no. 120482, doi: [10.1016/j.techfore.2020.120482](https://doi.org/10.1016/j.techfore.2020.120482).
- [5] N. Masle, "The AI index 2024 annual report," AI Index Steering Committee, Stanford, CA, USA, Tech. Rep., Apr. 2024.

- [6] M. Juric, A. Sandic, and M. Brcic, "AI safety: State of the field through quantitative lens," presented at the 2020 43rd Int. Conv. on Inf., Opatija, Croatia, Sep. 2020, pp. 1254–1259.
- [7] UN System Chief Executives Board for Coordination. (2024). *Summary of Deliberations: Chief Executives Board for Coordination, 1st Regular Session of 2024: Addendum*. [Online]. Available: <https://digitallibrary.un.org/record/4064466?v=pdf>
- [8] M. Mäntymäki, M. Minkkinen, T. Birkstedt, and M. Viljanen, "Defining organizational AI governance," *AI Ethics*, vol. 2, no. 4, pp. 603–609, Feb. 2022, doi: [10.1007/s43681-022-00143-x](https://doi.org/10.1007/s43681-022-00143-x).
- [9] United Nations. (2024). *Governing AI for Humanity*. [Online]. Available: https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf
- [10] UNESCO, Paris, France. (2022). *Recommendation on the Ethics of Artificial Intelligence*. [Online]. Available: <https://unesdoc.unesco.org/ark:/48223/pf0000381137>
- [11] *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, Standard NIST AI 100-1, National Institute of Standards and Technology, 2023, doi: [10.6028/NIST.AI.100-1](https://doi.org/10.6028/NIST.AI.100-1).
- [12] A. Khan, N. Choudhury, S. Uddin, L. Hossain, and L. A. Baur, "Longitudinal trends in global obesity research and collaboration: A review using bibliometric metadata," *Obesity Rev.*, vol. 17, no. 4, pp. 377–385, Feb. 2016.
- [13] F. Demiroz and T. W. Haase, "The concept of resilience: A bibliometric analysis of the emergency and disaster management literature," in *Local Disaster Management*. Evanston, IL, USA: Routledge, 2020, ch. 3, pp. 16–35, doi: [10.4324/9781003036234-3](https://doi.org/10.4324/9781003036234-3).
- [14] S. Laengle, J. M. Merigó, J. Miranda, R. Słowiński, I. Bomze, E. Borgonovo, R. G. Dyson, J. F. Oliveira, and R. Teunter, "Forty years of the European journal of operational research: A bibliometric overview," *Eur. J. Oper. Res.*, vol. 262, no. 3, pp. 803–816, Nov. 2017, doi: [10.1016/j.ejor.2017.04.027](https://doi.org/10.1016/j.ejor.2017.04.027).
- [15] J. Mingers and Q. L. Burrell, "Modeling citation behavior in management science journals," *Inf. Process. Manage.*, vol. 42, no. 6, pp. 1451–1464, Dec. 2006, doi: [10.1016/j.ipm.2006.03.012](https://doi.org/10.1016/j.ipm.2006.03.012).
- [16] S. R. Rosas, J. M. Kagan, J. T. Schouten, P. A. Slack, and W. M. K. Trochim, "Evaluating research and impact: A bibliometric analysis of research by the NIH/NIAID HIV/AIDS clinical trials networks," *PLoS ONE*, vol. 6, no. 3, Mar. 2011, Art. no. e17428.
- [17] M. Comunale and A. Manera. (2024). *The Economic Impacts and the Regulation of AI: A Review of the Academic Literature and Policy Actions*. [Online]. Available: <https://www.imf.org/en/Publications/WP/Issues/2024/03/22/The-Economic-Impacts-and-the-Regulation-of-AI-A-Review-of-the-Academic-Literature-and-546645>
- [18] S. Wasserman and J. Galaskiewicz, *Advances in Social Network Analysis: Research in the Social and Behavioral Sciences*. Newbury Park, CA, USA: Sage, 1997.
- [19] M. Bastian, S. Heymann, and M. Jacomy, "Gephi: An open source software for exploring and manipulating networks," in *Proc. Int. AAAI Conf. Weblogs Social Media*, vol. 3, 2009, pp. 361–362.
- [20] S. Oldham, B. Fulcher, L. Parkes, A. Arnatkevičiūtė, C. Suo, and A. Fornito, "Consistency and differences between centrality measures across distinct classes of networks," *PLoS ONE*, vol. 14, no. 7, Jul. 2019, Art. no. e0220061.
- [21] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *J. Stat. Mechanics: Theory Exp.*, vol. 2008, no. 10, Oct. 2008, Art. no. P10008.
- [22] A. Willis, A. Fisher, and I. M. Lvov, "Mapping networks of influence: Tracking Twitter conversations through time and space," *Participations, J. Audience Reception Stud.*, pp. 494–530, May 2015. [Online]. Available: <http://oro.open.ac.uk/43281/>
- [23] M. Afzaal, X. Shanshan, D. Yan, and M. Younas, "Mapping artificial intelligence integration in education: A decade of innovation and impact (2013–2023)—A bibliometric analysis," *IEEE Access*, vol. 12, pp. 113275–113299, 2024, doi: [10.1109/ACCESS.2024.3443313](https://doi.org/10.1109/ACCESS.2024.3443313).
- [24] A. McStay and G. Rosner, "Emotional artificial intelligence in children's toys and devices: Ethics, governance and practical remedies," *Big Data Soc.*, vol. 8, no. 1, Mar. 2021, Art. no. 205395172199487, doi: [10.1177/2053951721994877](https://doi.org/10.1177/2053951721994877).
- [25] G. C. K. Yew, "Trust in and ethical design of carebots: The case for ethics of care," *Int. J. Social Robot.*, vol. 13, no. 4, pp. 629–645, Jul. 2021, doi: [10.1007/s12369-020-00653-w](https://doi.org/10.1007/s12369-020-00653-w).
- [26] A. Jordan and N. Natarajan, "From TV to social media to 'ambient' AI: Insights from 30 years of children's media policy in the United States," *J. Children Media*, vol. 18, no. 3, pp. 378–385, Jul. 2024, doi: [10.1080/17482798.2024.2345530](https://doi.org/10.1080/17482798.2024.2345530).
- [27] L. Lee, R. K. Salami, H. Martin, L. Shantharam, K. Thomas, E. Ashworth, E. Allan, K.-W. Yung, C. Pauling, D. Leyden, O. J. Arthurs, and S. C. Shelmerdine, "'How i would like AI used for my imaging': Children and young persons' perspectives," *Eur. Radiol.*, vol. 34, no. 12, pp. 7751–7764, Jun. 2024, doi: [10.1007/s00330-024-10839-9](https://doi.org/10.1007/s00330-024-10839-9).
- [28] E. Mitsea, A. Drigas, and C. Skianis, "Digitally assisted mindfulness in training self-regulation skills for sustainable mental health: A systematic review," *Behav. Sci.*, vol. 13, no. 12, p. 1008, Dec. 2023, doi: [10.3390/bs13121008](https://doi.org/10.3390/bs13121008).
- [29] J. Su, "Kindergarten Teachers' perceptions of AI literacy education for young children," *Int. J. Technol. Design Educ.*, vol. 34, no. 5, pp. 1665–1685, Nov. 2024, doi: [10.1007/s10798-024-09876-8](https://doi.org/10.1007/s10798-024-09876-8).
- [30] A. Yadin-Segal and Y. Oppenheim, "Whose dystopia is it anyway? Deepfakes and social media regulation," *Convergence, Int. J. Res. into New Media Technol.*, vol. 27, no. 1, pp. 36–51, Feb. 2021, doi: [10.1177/1354856520923963](https://doi.org/10.1177/1354856520923963).
- [31] E. Portacolone, J. Halpern, J. Luxenberg, K. L. Harrison, and K. E. Covinsky, "Ethical issues raised by the introduction of artificial companions to older adults with cognitive impairment: A call for interdisciplinary collaborations," *J. Alzheimer's Disease*, vol. 76, no. 2, pp. 445–455, Jul. 2020, doi: [10.3233/jad-190952](https://doi.org/10.3233/jad-190952).
- [32] L. Bjørlo, Ø. Moen, and M. Pasquine, "The role of consumer autonomy in developing sustainable AI: A conceptual framework," *Sustainability*, vol. 13, no. 4, p. 2332, Feb. 2021, doi: [10.3390/su13042332](https://doi.org/10.3390/su13042332).
- [33] L. Rosenberg, "The metaverse and conversational AI as a threat vector for targeted influence," in *Proc. IEEE 13th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Mar. 2023, pp. 504–510.
- [34] S. D. Conca, "The present looks nothing like the jetsons: Deceptive design in virtual assistants and the protection of the rights of users," *Comput. Law Secur. Rev.*, vol. 51, Nov. 2023, Art. no. 105866, doi: [10.1016/j.clsr.2023.105866](https://doi.org/10.1016/j.clsr.2023.105866).
- [35] M. Carroll, A. Chan, H. Ashton, and D. Krueger, "Characterizing manipulation from AI systems," presented at the Equity Access Algorithms, Mech., Optim., Oct. 2023.
- [36] S. Lu, "Data privacy, human rights, and algorithmic opacity," *California Law Rev.*, vol. 110, pp. 2087–2147, Dec. 2022, doi: [10.15779/Z38804XM07](https://doi.org/10.15779/Z38804XM07).
- [37] C.-P. Yen and T.-W. Hung, "Achieving equity with predictive policing algorithms: A social safety net perspective," *Sci. Eng. Ethics*, vol. 27, no. 3, p. 36, Jun. 2021, doi: [10.1007/s11948-021-00312-x](https://doi.org/10.1007/s11948-021-00312-x).
- [38] C. Arnett, "Black lives monitored," *UCLA L. Rev.*, vol. 69, Apr. 2022, Art. no. 1384.
- [39] F. Odilla, "Unfairness in AI anti-corruption tools: Main drivers and consequences," *Minds Mach.*, vol. 34, no. 3, Jul. 2024, Art. no. 28, doi: [10.1007/s11023-024-09688-8](https://doi.org/10.1007/s11023-024-09688-8).
- [40] E. Ö. Kurtpınar, "Privacy's sky-high battle: The use of unmanned aircraft systems for law enforcement in the European union," *J. Intell. Robotic Syst.*, vol. 110, no. 3, Jul. 2024, Art. no. 99, doi: [10.1007/s10846-024-02071-w](https://doi.org/10.1007/s10846-024-02071-w).
- [41] A. Beduschi, "International migration management in the age of artificial intelligence," *Migration Stud.*, vol. 9, no. 3, pp. 576–596, Dec. 2021, doi: [10.1093/migration/mnaa003](https://doi.org/10.1093/migration/mnaa003).
- [42] A. Shachar and A. Mahmood, "The body as the border: A new era," *Historical Social Res.*, vol. 46, no. 3, pp. 124–150, Jan. 2021. [Online]. Available: <https://www.jstor.org/stable/27075120>
- [43] P. Jo Pesch, D. Dimitrova, and F. Boehm, "Data protection and machine-learning-supported decision-making at the EU border: ETIAS profiling under scrutiny," in *Proc. Annual Privacy Forum*. Cham, Switzerland: Springer, 2022, pp. 50–72.
- [44] B. Lorch, N. Scheler, and C. Riess, "Compliance challenges in forensic image analysis under the artificial intelligence act," in *Proc. 30th Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2022, pp. 613–617, doi: [10.23919/EUSIPCO55093.2022.9909723](https://doi.org/10.23919/EUSIPCO55093.2022.9909723).
- [45] N. Deng, X. Chen, and C. Xiong, "Design and construction of intelligent decision-making system for marine protection and law enforcement," in *Proc. Int. Conf. Broadband Wireless Comput., Commun. Appl.* Cham, Switzerland: Springer, 2020, pp. 828–837.

- [46] X. Feng, A. Asante, E. Short, and I. Abeykoon, "Cyberstalking issues," in *Proc. IEEE 15th Int. Conf. Dependable, Autonomic Secure Comput., 15th Int. Conf. Pervasive Intell. Comput., 3rd Int. Conf. Big Data Intell. Comput. Cyber Sci. Technol. Congress (DASC/PiCom/DataCom/CyberSciTech)*, Nov. 2017, pp. 373–376, doi: [10.1109/DASC-PiCom-DataCom-CyberSciTech.2017.78](https://doi.org/10.1109/DASC-PiCom-DataCom-CyberSciTech.2017.78).
- [47] P. C. Stern, T. Dietz, and M. P. Vandenberg, "The science of mitigation: Closing the gap between potential and actual reduction of environmental threats," *Energy Res. Social Sci.*, vol. 91, Sep. 2022, Art. no. 102735, doi: [10.1016/j.erss.2022.102735](https://doi.org/10.1016/j.erss.2022.102735).
- [48] A. Yücel, "Hybrid digital authoritarianism in turkey: The 'Censorship Law' and AI-generated disinformation strategy," *Turkish Stud.*, vol. 26, no. 1, pp. 1–27, Aug. 2024, doi: [10.1080/14683849.2024.2392816](https://doi.org/10.1080/14683849.2024.2392816).
- [49] M. Kuziński and G. Misuraca, "AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings," *Telecommun. Policy*, vol. 44, no. 6, Jul. 2020, Art. no. 101976, doi: [10.1016/j.telpol.2020.101976](https://doi.org/10.1016/j.telpol.2020.101976).
- [50] K. M. U. Mannonov and S. Myeong, "Citizens' perception of blockchain-based E-voting systems: Focusing on TAM," *Sustainability*, vol. 16, no. 11, p. 4387, May 2024, doi: [10.3390/su16114387](https://doi.org/10.3390/su16114387).
- [51] E. Stockinger, J. Maas, C. Talvitie, and V. Dignum, "Trustworthiness of voting advice applications in Europe," *Ethics Inf. Technol.*, vol. 26, no. 3, Sep. 2024, Art. no. 55, doi: [10.1007/s10676-024-09790-6](https://doi.org/10.1007/s10676-024-09790-6).
- [52] C. Marsden, T. Meyer, and I. Brown, "Platform values and democratic elections: How can the law regulate digital disinformation?" *Comput. Law Secur. Rev.*, vol. 36, Apr. 2020, Art. no. 105373.
- [53] P. A. Assibong, I. A. P. Wogu, M. A. Sholarin, S. Misra, R. Damasevičius, and N. Sharma, "The politics of artificial intelligence behaviour and human rights violation issues in the 2016 U.S. presidential elections: An appraisal," in *Advances in Intelligent Systems and Computing*, 2020, pp. 295–309.
- [54] T. C. King, N. Aggarwal, M. Taddeo, and L. Floridi, "Artificial intelligence crime: An interdisciplinary analysis of foreseeable threats and solutions," *Sci. Eng. Ethics*, vol. 26, no. 1, pp. 89–120, Feb. 2020, doi: [10.1007/s11948-018-00081-0](https://doi.org/10.1007/s11948-018-00081-0).
- [55] D. Garcia, "Lethal artificial intelligence and change: The future of international peace and security," *Int. Stud. Rev.*, vol. 20, no. 2, pp. 334–341, Jun. 2018, doi: [10.1093/isr/viy029](https://doi.org/10.1093/isr/viy029).
- [56] S. Ams, "Blurred lines: The convergence of military and civilian uses of AI & data use and its impact on liberal democracy," *Int. Politics*, vol. 60, no. 4, pp. 879–896, Aug. 2023, doi: [10.1057/s41311-021-00351-y](https://doi.org/10.1057/s41311-021-00351-y).
- [57] R. Liu, J. P. Salisbury, A. Vahabzadeh, and N. T. Sahin, "Feasibility of an autism-focused augmented reality smartglasses system for social communication and behavioral coaching," *Frontiers Pediatrics*, vol. 5, Jun. 2017, Art. no. 145.
- [58] A. P. Henkel, S. Bromuri, D. Iren, and V. Urovi, "Half human, half machine—augmenting service employees with AI for interpersonal emotion regulation," *J. Service Manage.*, vol. 31, no. 2, pp. 247–265, Jul. 2020, doi: [10.1108/josm-05-2019-0160](https://doi.org/10.1108/josm-05-2019-0160).
- [59] M. A. Hasan, N. F. M. Noor, S. S. B. A. Rahman, and M. M. Rahman, "The transition from intelligent to affective tutoring system: A review and open issues," *IEEE Access*, vol. 8, pp. 204612–204638, 2020, doi: [10.1109/ACCESS.2020.3036990](https://doi.org/10.1109/ACCESS.2020.3036990).
- [60] K. Siau and W. Wang, "Artificial intelligence (AI) ethics," *J. Database Manage.*, vol. 31, no. 2, pp. 74–87, Apr. 2020, doi: [10.4018/jdm.2020040105](https://doi.org/10.4018/jdm.2020040105).
- [61] A. Ozaki, "The use of live facial recognition technology by private entities in public places: A study of recent case in Japan," in *Proc. IFIP Int. Conf. Hum. Choice Comput.*, Aug. 2022, pp. 27–35, doi: [10.1007/978-3-031-15688-5_3](https://doi.org/10.1007/978-3-031-15688-5_3).
- [62] J. Sánchez-Monedero and L. Dencik, "The politics of deceptive borders: 'Biomarkers of deceit' and the case of iBorderCtrl," *Inf., Commun. Soc.*, vol. 25, no. 3, pp. 413–430, Feb. 2022, doi: [10.1080/1369118x.2020.1792530](https://doi.org/10.1080/1369118x.2020.1792530).
- [63] M. C. Horowitz and L. Kahn, "What influences attitudes about artificial intelligence adoption: Evidence from U.S. local officials," *PLoS ONE*, vol. 16, no. 10, Oct. 2021, Art. no. e0257732.
- [64] M. Ahn, C. Huang, P.-C. Huang, X. Zhong, J. Himmelreich, K. Desouza, and R. Knepper, "Cyber-physical innovations: Cyber-infrastructure for research, cyber-physical architecture for real-time applications, autonomous vehicle (AV) governance and AI artifacts for public value," in *Proc. 22nd Annu. Int. Conf. Digit. Government Res.*, Jun. 2021, pp. 590–592.
- [65] G. Á. Werner, "Fuzzy logic adapted controller system for biometrical identification in highly-secured critical infrastructures," in *Proc. IEEE 10th Jubilee Int. Symp. Appl. Comput. Intell. Informat.*, May 2015, pp. 335–340.
- [66] R. Gupta and B. Rathore, "Exploring the generative AI adoption in service industry: A mixed-method analysis," *J. Retailing Consum. Services*, vol. 81, Nov. 2024, Art. no. 103997.
- [67] V. Agarwal, R. Bar-Haim, L. Eden, N. Gupta, Y. Kantor, and A. Kumar, "AI-assisted security controls mapping for clouds built for regulated workloads," in *Proc. IEEE 14th Int. Conf. Cloud Comput. (CLOUD)*, Sep. 2021, pp. 136–146.
- [68] M. J. Pappaterra and F. Flammini, "A review of intelligent infrastructure surveillance to support safe autonomy in smart-railways," in *Proc. IEEE 26th Int. Conf. Intell. Transp. Syst. (ITSC)*, Sep. 2023, pp. 5603–5610.
- [69] A. Bucchiarone, S. Battisti, A. Marconi, R. Maldacea, and D. C. Ponce, "Autonomous shuttle-as-a-service (ASaaS): Challenges, opportunities, and social implications," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 6, pp. 3790–3799, Jun. 2021.
- [70] R. Gettler, "Augmenting the human pilot—AI/ML influences on one pilot to many unmanned aerial system flight," in *Proc. Integr. Commun., Navigat. Surveill. Conf. (ICNS)*, Apr. 2022, pp. 1–7, doi: [10.1109/ICNS54818.2022.9771489](https://doi.org/10.1109/ICNS54818.2022.9771489).
- [71] V. Santoro and L. Pensato, "Critical infrastructure protection: The need for evolving standards: Mutating cyber-space and security issues in ITS," in *Proc. 5th IEEE Int. Conf. Models Technol. Intell. Transp. Syst. (MT-ITS)*, Jun. 2017, pp. 219–222.
- [72] M. Adnan, Y. Ghadi, I. Ahmed, and M. Ali, "Transmission network planning in super smart grids: A survey," *IEEE Access*, vol. 11, pp. 77163–77227, 2023.
- [73] A. R. Singh, M. Dashtdar, M. Bajaj, R. Garmsiri, V. Blazek, L. Prokop, and S. Misak, "AI-enhanced power quality management in distribution systems: Implementing a dual-phase UPQC control with adaptive neural networks and optimized PI controllers," *Artif. Intell. Rev.*, vol. 57, no. 11, Sep. 2024, Art. no. 311.
- [74] J. Nyangon, "Climate-proofing critical energy infrastructure: Smart grids, artificial intelligence, and machine learning for power system resilience against extreme weather events," *J. Infrastructure Syst.*, vol. 30, no. 1, Mar. 2024, Art. no. 03124001.
- [75] S. Joshi, M. Sharma, R. P. Das, J. Rosak-Szyrocka, J. Żywiołek, K. Muduli, and M. Prasad, "Modeling conceptual framework for implementing barriers of AI in public healthcare for improving operational excellence: Experiences from developing countries," *Sustainability*, vol. 14, no. 18, p. 11698, Sep. 2022.
- [76] M. Jovanovic, G. Mitrov, E. Zdravovski, P. Lameski, S. Colantonio, M. Kampel, H. Tellioglu, and F. Florez-Revuelta, "Ambient assisted living: Scoping review of artificial intelligence models, domains, technology, and concerns," *J. Med. Internet Res.*, vol. 24, no. 11, Nov. 2022, Art. no. e36553.
- [77] C. K. Y. Chan, "A comprehensive AI policy education framework for university teaching and learning," *Int. J. Educ. Technol. Higher Educ.*, vol. 20, no. 1, Jul. 2023, Art. no. 38, doi: [10.1186/s41239-023-00408-3](https://doi.org/10.1186/s41239-023-00408-3).
- [78] S. Chatterjee and K. K. Bhattacharjee, "Adoption of artificial intelligence in higher education: A quantitative analysis using structural equation modelling," *Educ. Inf. Technol.*, vol. 25, no. 5, pp. 3443–3463, Sep. 2020, doi: [10.1007/s10639-020-10159-7](https://doi.org/10.1007/s10639-020-10159-7).
- [79] M. P. Recht, M. Dewey, K. Dreyer, C. Langlotz, W. Niessen, B. Prainsack, and J. J. Smith, "Integrating artificial intelligence into the clinical practice of radiology: Challenges and recommendations," *Eur. Radiol.*, vol. 30, no. 6, pp. 3576–3584, Jun. 2020, doi: [10.1007/s00330-020-06672-5](https://doi.org/10.1007/s00330-020-06672-5).
- [80] Q. Waymel, S. Badr, X. Demondion, A. Cotten, and T. Jacques, "Impact of the rise of artificial intelligence in radiology: What do radiologists think?" *Diagnostic Interventional Imag.*, vol. 100, no. 6, pp. 327–336, Jun. 2019, doi: [10.1016/j.diii.2019.03.015](https://doi.org/10.1016/j.diii.2019.03.015).
- [81] J. G. Borger, A. P. Ng, H. Anderton, G. W. Ashdown, M. Auld, M. E. Blewitt, D. V. Brown, M. J. Call, P. Collins, S. Freytag, L. C. Harrison, E. Hespings, J. Hoysted, A. Johnston, A. McInnery, P. Tang, L. Whitehead, A. Jex, and S. H. Naik, "Artificial intelligence takes center stage: Exploring the capabilities and implications of ChatGPT and other AI-assisted technologies in scientific research and education," *Immunology Cell Biol.*, vol. 101, no. 10, pp. 923–935, Nov. 2023, doi: [10.1111/imcb.12689](https://doi.org/10.1111/imcb.12689).

- [82] D. H. Chang, M. P.-C. Lin, S. Hajian, and Q. Q. Wang, "Educational design principles of using AI chatbot that supports self-regulated learning in education: Goal setting, feedback, and personalization," *Sustainability*, vol. 15, no. 17, p. 12921, Aug. 2023.
- [83] M. H. Elnagar, S. Yadav, S. R. Venugopalan, M. K. Lee, M. Oubaidin, S. Rampa, and V. Allareddy, "ChatGPT and dental education: Opportunities and challenges," *Seminars Orthodontics*, vol. 30, no. 4, pp. 401–404, Sep. 2024, doi: 10.1053/j.sodo.2024.03.004.
- [84] S. Mao, B. Wang, Y. Tang, and F. Qian, "Opportunities and challenges of artificial intelligence for green manufacturing in the process industry," *Engineering*, vol. 5, no. 6, pp. 995–1002, Dec. 2019, doi: 10.1016/j.eng.2019.08.013.
- [85] A. Todolí-Signes, "Algorithms, artificial intelligence and automated decisions concerning workers and the risks of discrimination: The necessary collective governance of data protection," *Transfer: Eur. Rev. Labour Res.*, vol. 25, no. 4, pp. 465–481, Nov. 2019.
- [86] E. Kazim, A. S. Koshiyama, A. Hilliard, and R. Polle, "Systematizing audit in algorithmic recruitment," *J. Intell.*, vol. 9, no. 3, p. 46, Sep. 2021.
- [87] M. Jibril and T. A. Florentina, "Governing AI in hiring: An effort to eliminate biased decision," in *Proc. JSAI Int. Symp. Artif. Intell.*, 2024, pp. 49–63.
- [88] P. Liu, L. Yuan, and Z. Jiang, "The dark side of algorithmic management: Investigating how and when algorithmic management relates to employee knowledge hiding?" *J. Knowl. Manage.*, vol. 29, no. 2, pp. 342–371, Oct. 2024.
- [89] S. Chowdhary, A. Kawakami, M. L. Gray, J. Suh, A. Olteanu, and K. Saha, "Can workers meaningfully consent to workplace wellbeing technologies?" in *Proc. ACM Conf. Fairness, Accountability, Transparency*, Jun. 2023, pp. 569–582.
- [90] A. Bernhardt, L. Kresge, and R. Suleiman, "The data-driven workplace and the case for worker technology rights," *ILR Rev.*, vol. 76, no. 1, pp. 3–29, Jan. 2023.
- [91] S. Gutsu and H. Spitsyna, "Labor legislation in the era of artificial intelligence," in *Proc. Conf. Integr. Comput. Technol. Mech. Eng.-Synergetic Eng.*, Dec. 2024, pp. 252–270.
- [92] J. Cranefield, M. Winikoff, Y.-T. Chiu, Y. Li, C. Doyle, and A. Richter, "Partnering with AI: The case of digital productivity assistants," *J. Roy. Soc. New Zealand*, vol. 53, no. 1, pp. 95–118, Jan. 2023.
- [93] S. Chowdhury, P. Dey, S. Joel-Edgar, S. Bhattacharya, O. Rodriguez-Espindola, A. Abadie, and L. Truong, "Unlocking the value of artificial intelligence in human resource management through AI capability framework," *Human Resource Manage. Rev.*, vol. 33, no. 1, Mar. 2023, Art. no. 100899.
- [94] B. Cheng, H. Lin, and Y. Kong, "Challenge or hindrance? How and when organizational artificial intelligence adoption influences employee job crafting," *J. Bus. Res.*, vol. 164, Sep. 2023, Art. no. 113987.
- [95] C. van Noordt and G. Misuraca, "Artificial intelligence for the public sector: Results of landscaping the use of AI in government across the European union," *Government Inf. Quart.*, vol. 39, no. 3, Jul. 2022, Art. no. 101714.
- [96] Z. Engin and P. Treleaven, "Algorithmic government: Automating public services and supporting civil servants in using data science technologies," *Comput. J.*, vol. 62, no. 3, pp. 448–460, Mar. 2019.
- [97] C. Vrabie, "E-government 3.0: An AI model to use for enhanced local democracies," *Sustainability*, vol. 15, no. 12, p. 9572, Jun. 2023.
- [98] C. van Noordt and G. Misuraca, "Exploratory insights on artificial intelligence for government in Europe," *Social Sci. Comput. Rev.*, vol. 40, no. 2, pp. 426–444, Apr. 2022.
- [99] Y.-F. Wang, Y.-C. Chen, and S.-Y. Chien, "Citizens' intention to follow recommendations from a government-supported AI-enabled system," *Public Policy Admin.*, vol. 40, no. 2, May 2023, Art. no. 09520767231176126.
- [100] R. Baeza-Yates and U. M. Fayyad, "Responsible AI: An urgent mandate," *IEEE Intell. Syst.*, vol. 39, no. 1, pp. 12–17, Jan. 2024.
- [101] S. Hu and Y. Li, "Policy interventions and regulations on generative artificial intelligence: Key gaps and core challenges," in *Proc. 25th Annu. Int. Conf. Digit. Government Res.*, Jun. 2024, pp. 1034–1036.
- [102] P. Hacker, "The European AI liability directives-critique of a half-hearted approach and lessons for the future," *Comput. Law Secur. Rev.*, vol. 51, Nov. 2023, Art. no. 105871.
- [103] G. Pisoni and N. Díaz-Rodríguez, "Responsible and human centric AI-based insurance advisors," *Inf. Process. Manage.*, vol. 60, no. 3, May 2023, Art. no. 103273.
- [104] C. W. L. Ho, J. Ali, and K. Caals, "Ensuring trustworthy use of artificial intelligence and big data analytics in health insurance," *Bull. World Health Org.*, vol. 98, no. 4, pp. 263–269, Apr. 2020.
- [105] D. Lanfranchi and L. Grassi, "Translating technological innovation into efficiency: The case of US public P&C insurance companies," *Eurasian Bus. Rev.*, vol. 11, no. 4, pp. 565–585, Dec. 2021.
- [106] S. K. Marimekala, J. Lamb, R. Epstein, and V. Bhupathi, "Using AI and big data in the HealthCare sector to help build a smarter and more intelligent HealthCare system," in *Proc. IEEE World AI IoT Congr. (AllIoT)*, May 2024, pp. 356–362.
- [107] F. M. Dawoodbhoy, J. Delaney, P. Cecula, J. Yu, I. Peacock, J. Tan, and B. Cox, "AI in patient flow: Applications of artificial intelligence to improve patient flow in NHS acute mental health inpatient units," *Heliyon*, vol. 7, no. 5, May 2021, Art. no. e06993.
- [108] W. W.-W. Hsiao, J.-C. Lin, C.-T. Fan, and S. S.-S. Chen, "Precision health in taiwan: A data-driven diagnostic platform for the future of disease prevention," *Comput. Structural Biotechnol. J.*, vol. 20, pp. 1593–1602, Apr. 2022.



BYEONG-JE KIM received the Ph.D. degree in disaster management engineering from Ulsan National Institute of Science and Technology (UNIST), Ulsan, Republic of Korea, in 2022.

He is currently a Research Assistant Professor with the Institute of Social Data Science, Pohang University of Science and Technology (POSTECH), Republic of Korea. His research interests include disaster management policy, risk perception, and safety education.



SEUNGHOO JEONG received the Ph.D. degree in urban infrastructure engineering from Ulsan National Institute of Science and Technology (UNIST), Ulsan, Republic of Korea, in 2020, with the dissertation title of "Structural health monitoring of cables in cable-stayed bridges."

From March to May 2020, he was a Postdoctoral Researcher with the School of Civil, Architectural Engineering and Landscape Architecture, Sungkyunkwan University, Suwon, Republic of Korea. Since June 2020, he has been a Senior Researcher with the Research and Development Planning and Policy Department, Korea Railroad Research Institute (KRRRI), Republic of Korea. His research interests include structural health monitoring, smart construction, and disaster/risk management.



BONG-KYUNG CHO received the master's degree in public policy from the Graduate School of Public Administration, Seoul National University, Seoul, Republic of Korea, in 2012, and the Ph.D. degree in urban planning from the Graduate School of Environmental Studies, Seoul National University, in 2021.

She is currently the Team Leader of the Strategy Planning Team with INTERX, Republic of Korea. Her current research interests include AI regulation, AI policy, governance, dataspace, carbon regulation, and urban planning.



JI-BUM CHUNG received the Ph.D. degree in urban planning from Yonsei University, Seoul, Republic of Korea, in 2007.

He is currently an Associate Professor with the Department of Civil, Urban, Earth, and Environmental Engineering, Ulsan National Institute of Science and Technology (UNIST), Republic of Korea. His research interests include risk communication/management and disaster management.

...