ISSN 2435-3345

Northeast Asian Studies

Robotics & AI Ethics

Vol. 10 No. 0

J-INSTITUTE

2025 Vol.10 No.0, 1-13

Received: 2025/08/27, Peer-reviewed: 2025/09/30, Accepted: 2025/10/25, Published: 2025/12/30

Robotics & AI Ethics

Publisher: J-INSTITUTE
ISSN: 2435-3345

Website: j-institute.org

Corresponding author*
E-mail: jadekyungok@gmail.com

DOI Address:
dx.doi.org/10.22471/ai.2025.10.0.01

Research on the Educational Effects of Edu-Tech Application based on Book Creator in Aviation Major Education

Soyoung Lee[1]

Chungwoon University, Associate Professor, Republic of Korea

Kyungok Choi[2*]

Chungwoon University, Assistant Professor, Republic of Korea

Abstract

Purpose: *This study aims to explore the educational effects of integrating Book Creator—a digital edu-tech platform—into learner-centered, reflective practice within practical airline service interview practice and English courses. Specifically, the research investigates how the use of such digital tools fosters improvements in students' self-regulated learning abilities, practical professional competencies, intrinsic learning motivation, and overall satisfaction with the educational process in aviation service programs. By focusing on the repeated cycle of action, feedback, and self-reflection across a series of authentic tasks, this study seeks to provide empirical evidence for the meaningful role of edu-tech solutions in future-oriented higher education.*

Method: *Two representative courses, Air Service Interview Practice and Airline Transport English, were selected and Book Creator activities were integrated into a six-week instructional design. The lessons followed a three-phase cycle: pre-class digital content study, in-class practical activities, and post-class assessment with repeated self-reflection. Students built digital portfolios, and data were gathered through instructor evaluation rubrics, satisfaction surveys, and in-depth qualitative interviews. Changes in instructor feedback and student self-assessment before and after the intervention were analyzed to comprehensively evaluate the educational impact of edu-tech integration.*

Results: *The results showed that Book Creator-based lessons positively influenced not only students' self-regulated learning abilities, but also their practical English communication skills and preparation for real-world air service interviews. Students accumulated weekly video assignments and self-reflection records, which enabled them to clearly recognize their achievements and growth through their portfolios. Diverse feedback from instructors and peers, along with the experience of visualizing personal development, had a significant impact on both their satisfaction and learning motivation. Furthermore, the portfolios served as valuable resources not only for academic assessment but also for job preparation, while instructor feedback became increasingly individualized, detailed, and efficient in the digital environment.*

Conclusion: *Taken together, the results indicate that edu-tech tools such as Book Creator provide an effective foundation for supporting self-directed learning, practical skill development, and differentiated feedback design in aviation service education. The digital portfolio system enables sustainability in learning and the visualization of specific growth records, systematically supporting individual student progress. Future research should apply this model to other disciplines and educational settings, and conduct longitudinal studies to further validate the effects of edu-tech in higher education.*

Keywords: ***Edu-Tech, Book Creator, Self-Regulated Learning, Aviation Service Education, Portfolio Assessment***

1. Introduction

Modern society is undergoing fundamental changes in the educational environment due to rapid digital transformation[1][2]. In particular, the COVID-19 pandemic has accelerated the digitization of educational settings, making remote learning essential and rendering innovation in educational methods inevitable[1][3]. Simultaneously, in this VUCA era, society requires individuals who can respond flexibly to change and are equipped with future competencies[4][5].

In this context, higher education faces significant challenges[6]. Traditional lecture-based teaching reveals limitations in meeting students' diverse learning styles and individual needs, while digital-native learners—such as the MZ generation—are seeking interactive and personalized learning experiences[7][8]. Moreover, gaps in digital device usage and experience between instructors and learners in university settings highlight the need for the development of effective teaching and learning models[9].

Accordingly, higher education institutions are introducing Edu-Tech as a core strategy for instructional innovation. Edu-Tech, a compound of Education and Technology, incorporates Fourth Industrial Revolution technologies such as AI, big data, IoT, and AR/VR to create new learning experiences and educational paradigms[10][11][12]. Edu-Tech, utilizing digital tools and resources, provides learning opportunities tailored to each student's individual characteristics and pace, and offers instructors opportunities to monitor learning status and deliver efficient feedback for qualitative improvement in education[13].

This study particularly focuses on the educational potential of digital textbooks within Edu-Tech. Digital textbooks, featuring interactive elements such as slide images, videos, problem-solving, pop-up windows, and web links, positively affect students' academic engagement and achievement[6][14][15]. They can help transform emotionally-inclined, low-engagement students into high-engagement learners, offering differentiated enjoyment and interactivity compared to paper books[16]. Additionally, immersive content such as AR/VR enables practice-oriented learning, and video resources facilitate greater student participation and interest[17][18]. Such interactive, digital textbook-based learning enhances key competencies required for future education, including self-directed learning, collaboration, and continual growth experience[6][14].

Nevertheless, the adoption of Edu-Tech in universities continues to face challenges, such as differences in technology acceptance and usage intentions between instructors and learners, a lack of effective instructional models, the selection of tools suited to the characteristics of each major, and the need for strategies to enhance digital competencies[6][7].

Prior studies on Edu-Tech and digital portfolios in higher education have demonstrated that multimedia tools, e-books, and LMS-based platforms can enhance learner engagement, self-directed learning, and satisfaction in liberal arts, language education, and pre-service teacher training contexts[19][20][21]. However, these studies have generally focused on learners' perceptions of technology acceptance, broad satisfaction levels, or flipped learning models using systems such as Google Classroom, and have seldom provided detailed accounts of how weekly portfolio cycles and authentic performance tasks are structurally integrated into practice-oriented major courses.

In this regard, there remains a conceptual and practical gap in the literature concerning how digital portfolio tools can be aligned with the competency requirements of specialized professional programs, such as aviation service interview practice or cabin crew English training. This study addresses that gap by implementing Book Creator not simply as an assignment submission channel but as a core medium that integrates weekly performance tasks, multi-source feedback (self, peer, instructor), and visual self-reflection into a single evolving e-portfolio in two aviation major courses: Airline Service Interview Practice and Airline Operations Practical English. By analyzing students' experiences and growth over a six-week period, the study aims to propose a discipline-specific, design-based Edu-Tech model that supports self-regulated learning, practical skill development, and portfolio-based assessment in aviation service education.

1.1. Theoretical background

Self-regulated learning theory conceptualizes learning as an active and cyclical process in which learners set goals, monitor their progress, evaluate outcomes, and adjust strategies accordingly, and repeated engagement in these processes enhances achievement and self-efficacy(Zimmerman, 2000)[22]. Bandura's (1997) notion of self-efficacy[23], and Dweck's (2006) growth mindset further explain how learners' beliefs about their capabilities and development potential influence motivation, persistence, and the formation of a positive self-image[24]. Recent Edu-Tech research shows that multimedia digital tools, including Book Creator, can operationalize these theories by enabling iterative performance, multimodal feedback, and visualized growth records, thereby promoting self-directed learning, collaboration, and sustained engagement in higher education[25][26][27].

This study aims to examine whether the use of digital tools that enable the reproduction of real-world scenarios, as well as repeated self-reflection and feedback, has a positive impact on student engagement and achievement in practice-oriented aviation service and foreign language education settings[1]. Edu-tech tools are reported to support student self-directed learning, promote collaborative learning, increase learning motivation, and enhance participation; for instructors, these tools contribute to more efficient course management and the delivery of customized feedback, thereby reducing instructional workload[6][28][29][30].

In particular, students engage in various assignments, workbook creation, and video recording activities using multimedia tools such as Book Creator, experiencing visual self-reflection on their developmental progress[5][8][31][32]. This is crucial for building a positive self-image and reinforcing motivation; students participating in courses utilizing digital textbooks and edu-tech tools have shown improved self-directed learning, collaborative skills, information literacy, self-confidence, and self-esteem[2][28].

This study investigates three main dimensions—self-directed learning, self-efficacy, and the effectiveness of edu-tech-based learning—in classes utilizing digital textbooks and edu-tech tools, focusing on university students majoring in aviation service management[22][23][24].

1.2. Purpose of the study

Book Creator can be considered particularly specialized for aviation service training due to its e-book format, unlike LMS platforms such as Google Classroom. First, video/audio embedding visualizes interview and pronunciation practice. Second, page-specific comment features enable real-time multi-source feedback (self-peer-instructor). Third, weekly chapter accumulation automatically visualizes growth trajectories. Fourth, offline access and PDF export facilitate its use as employment portfolios.

Building on these theoretical and empirical foundations, this study empirically examines the educational effects of integrating Book Creator into major courses in an aviation service management program. Focusing on two practice-oriented subjects—Airline Service Interview Practice and Airline Operations Practical English—it explores whether and how a six-week, Book Creator-based instructional design can foster students' self-regulated learning abilities, practical interview and English communication skills, intrinsic motivation, and satisfaction with the learning process. By doing so, the study seeks to provide a discipline-specific model of Edu-Tech integration that can inform the design of future learner-centered, portfolio-based instruction in aviation and related service fields.

To achieve this, the study focuses on three key dimensions: (1) self-directed and self-regulated learning, (2) self-efficacy and positive self-image formation through visual self-reflection on growth, and (3) the perceived effectiveness of Edu-Tech-based learning in practice-oriented aviation service and foreign language education settings. By analyzing students' portfolio artifacts, satisfaction survey responses, and qualitative feedback, the research aims to clarify how digital tools that reproduce real-world scenarios and support repeated cycles of feedback and reflection can positively influence learner engagement, collaboration, and academic and practical achievement in aviation service major education.

2. Research Methodology and Development Process

2.1. Airline service interview practice course

2.1.1. Course operation overview

The Airline Service Interview Practice course is designed for students preparing for employment in the aviation service sector, with a strong emphasis on simulating real interview situations and strengthening practical competencies. The course follows a three-phase cyclical structure: pre-class video learning, in-class mock interview practice, and post-class feedback and self-reflection. Over six weeks, students complete weekly interview assignments related to aviation service tasks. Each assignment involves students filming themselves in mock interview scenarios and uploading the videos to their personal YouTube channels; the video links and related materials are then organized and submitted in their individual Book Creator e-portfolios. The instructor manages all student assignments, self-reflection records, and feedback integratively through the Book Creator platform, systematically supporting student progress and growth.

2.1.2. Student participation and assignment process

In the Airline Service Interview Practice course, a total of eight, fourth-year students participated in the study, comprising one male and seven female students. The entire population of students who completed this course were the subjects of this study. Two of the female students had returned from a one-year leave of absence, whereas the remaining six were all from the 2022 entering cohort, forming a relatively homogeneous group in terms of curricular experience.

Table 1. Assignment process overview.

Step	Description
Assignment Design	Each week, students filmed video responses to common questions frequently asked in real airline interviews (such as self-introduction, service mindset, and handling emergency situations).
Assignment Submission	(1) Students uploaded their weekly video assignments to their personal YouTube channels and organized the video links along with self-evaluation and reflection notes in their individual Book Creator portfolios. (2) For in-class mock interview videos, students used an evaluation rubric (covering logic of response, demeanor, image, etc.) to self-assess and actively documented feedback from both the instructor and peers.
Feedback and Self-Directed Improvement	(1) The instructor provided feedback via Book Creator in both face-to-face and online settings, and students reflected on this feedback to determine improvement points, which they incorporated into subsequent assignments. (2) Each week, students accumulated and visualized their assignments, feedback, and video records by week, allowing for clear reflection and tracking of growth over time. (3) Peer feedback, instructor feedback, and self-feedback were managed integratively, deepening interactive and self-directed learning experiences.

This stepwise course and assignment process was designed to evenly nurture key competencies such as practical field adaptation, iterative self-reflection, and receptive collaborative feedback. Notably, the systematic adoption of digital tools such as Book Creator supported the visualization of learning materials, individualized tracking of growth trajectories, and the smooth integration of online and offline feedback, thereby maximizing the effectiveness of blended learning environments.

2.2. Airline operations practical English course

2.2.1. Course operation overview

The Airline Operations Practical English course is designed to strengthen the practical English skills and situational response abilities required for airline cabin crew. The course provides repeated practice of English expressions used in authentic work contexts, including in-flight service procedures, safety briefings, emergency management, and handling passenger complaints. Rather than relying on rote memorization, this course employs a participatory and experiential format, where students independently produce and perform standardized announcements or scenario scripts for various in-flight situations, combined with role-play. Digital tools such as Book Creator are utilized for team projects, collaborative editing, iterative self-reflection, and feedback. Each week, the instructor provides resources on key theoretical concepts and standard phrases for each flight stage, which students utilize to complete team and individual projects as well as a digital portfolio book.

2.2.2. Student participation and assignment process

In the Airline Operations Practical English course, twelve third-year students participated, including ten male and two female students. All male students were from the 2021 cohort who had returned after completing military service, while the two female students were from the 2022 (returning) and 2023 cohorts, respectively. In this courses, all students enrolled in the class during the semester were included as participants, forming intact class groups rather than a randomly sampled population in order to reflect authentic instructional conditions as closely as possible.

Table 2. Instructional phases and student engagement process.

Step	Description
Pre-Learning	Students preview lectures and resources (uploaded to LMS) that include essential expressions and standard announcements for each scenario, as well as alternative expressions for unexpected occurrences. This prepares them for their roles in upcoming team projects.
Team Project & Co-Production	Students collaboratively write role-play scripts for each stage of flight (e.g., boarding announcements, safety demonstrations, complaint handling) and individually produce/upload their assigned segments (videos, voice recordings, text scripts) using Book Creator. Clear assignment of pages within each team and active use of collaborative editing features ensure smooth integration of contributions.
Self-Directed Feedback & Iterative Learning	Drafts, revisions, and final versions are organized in Book Creator. Students actively incorporate instructor and peer feedback into multiple rounds of revision. They also learn by referring to diverse expressions used by other teams, and, through repeated review via the mobile e-book format, track their development in a self-reflection journal. This process culminates in a completed portfolio book.
Career Utilization	Finalized e-book portfolios serve as tangible evidence of student growth, practical competencies, and English communication skills, and may be used as supporting material for job search and interview preparation prior to graduation.

2.2.3. Book creator-based instructional tips

2.2.3.1. Preparation tips

The preparation phase focuses on ensuring students are comfortable using Book Creator through account setup, providing a standardized template, guiding collaborative editing roles, and addressing copyright compliance to facilitate smooth project initiation.

Table 3. Preparation guidelines.

Category	Description
Account creation & template sharing	In the first week, all students create their accounts and receive a basic template (cover, table of contents, sample pages) designed by the instructor, reducing initial anxiety about starting with a blank slate.
Collaborative editing guide	For team projects, students are instructed to insert a table showing page assignment and division of roles on the first page of the Book Creator project to prevent edit conflicts.
Copyright guide	A checklist of royalty-free image/video sources and citation examples is provided from the outset to prevent copyright issues.

2.2.3.2. Weekly activity tips

Weekly activities in Book Creator follow a structured cycle of pre-learning resource organization, collaborative script development, multimedia uploads, and integrated feedback documentation to support iterative improvement and self-reflection.

Table 4. Learning activity process overview.

Step	Description
Pre-learning organization	Students save resources provided by the instructor (standard announcements, scenario expressions) in their individual digital books, using Book Creator as an “online manual book” instead of physical notes.
Script writing for role-play	Collaborative editing is enabled; each student writes their assigned section in real time, using comment features for suggestions or questions among teammates.
Pronunciation recordings/ role-play video uploads	Initial pronunciation exercises are uploaded as individual audio files, and group presentation videos are organized by chapter. For each scenario, draft, revised, and final versions are separated into sequential pages to visualize the process of improvement.
Mid-course feedback documentation	Instructor feedback and peer comments are not saved separately as files but are recorded directly as comments on Book Creator pages. Students also use these records to reflect and compare before and after revisions.

3. Conclusion and Utilization Plan

3.1. Airline service interview practice course

3.1.1. Course outcomes and changes

3.1.1.1. Enhancement of self-directed learning ability

Students identified their strengths and weaknesses each week through video assignments and reflective records, repeatedly engaging in goal setting, self-evaluation, and self-regulatory strategies. Through iterative video production and incorporation of feedback, both self-directed learning attitudes and practical skills showed gradual improvement. According to end-of-term surveys, students reported that “the process of preparation and improvement was systematic, resulting in greater self-directedness.” This aligns with Zimmerman’s theory that repetitive self-evaluation and feedback-based self-regulated learning positively influence motivation and achievement[27].

3.1.1.2. Improvement of practical competency and confidence

By reenacting real interview situations via video, students developed practical sense and confidence. Detailed instructor feedback—on aspects such as facial expressions, pronunciation, and response structure—substantially contributed to strengthening individual competencies. Students stated, “Seeing my improvement directly in the videos provided motivation.” This process reaffirms findings by Jin & Lee (2024)[5], and Hong et al. (2024) that edu-tech based feedback and management of visual learning records are effective for self-efficacy and competency enhancement[9].

3.1.1.3. Satisfaction and positive self-image formation

Assignment submission and portfolio management using Book Creator brought high satisfaction in resource organization, repeated practice, and real-time feedback from both peers and instructors. Students noted, “Instructor feedback recorded alongside video materials helped me make concrete improvements (Student A, Year 4),” and “the portfolio can also be used for job preparation. (Student B, Year 4)” Although some students initially struggled with digital tool adaptation, stepwise guidance and instructor support facilitated gradual adjustment. These results are consistent with research by Park (2013)[14], indicating that digital portfolio-based teaching significantly supports self-directedness, satisfaction, and future applicability.

The pilot test was conducted with 8 participants over 3 weeks. Survey and interview results are summarized as follows. The satisfaction survey was constructed based on four validated items from existing Edu-tech research literature (motivation for participation, ease of self-feedback, usefulness of instructor feedback, and user convenience), with internal consistency reliability secured through a pre-pilot test involving 8 students (Cronbach's α = 0.87). Items were measured on a 5-point Likert scale, and exploratory factor analysis (EFA) confirmed convergence on a single factor (72.4% explained variance). The pilot results of this study (M=4.65, SD=0.42) exceeded the benchmark values from prior research, thereby demonstrating validity.

Table 5. Learner satisfaction results from pilot test.

Item	Mean Satisfaction (Out of 5)	SD	Key Feedback Summary
Motivation for Participation	4.7	0.38	"Entering my video into the book made me work harder." "Being able to visually confirm my progress was helpful for actual interview preparation."
Ease of Self-Feedback	4.6	0.45	"Comparing my previous videos helped me objectively see my response habits." "Specific feedback identified strengths and enabled me to set improvement goals."
Usefulness of Instructor Feedback	4.8	0.31	"Detailed comments on fine points were very helpful." "I could immediately check and apply corrections by referencing the video materials."

User Convenience	4.5	0.49	"It was convenient to view everything simply by clicking a link." "The experience of recording and managing various resources positively influenced my motivation and participation."

3.1.2. Instructor feedback (before and after Edutech introduction)

Prior to the use of digital tools, feedback was limited to oral delivery in class, making it difficult to communicate and document detailed points for improvement. With Book Creator, instructors provided visual, annotated feedback on videos and portfolios, which increased both the quality of feedback delivery and students' receptiveness. This supports Byun(2021)[16], and Yoon(2013)[15], who show that diverse and repeated feedback substantively improves learning outcome.

Table 6. Instructor feedback: Pre/Post Edutech.

Item	Before Introduction	After Introduction
Feedback Type	Paper/oral, limited time	Video/e-book annotated feedback, can be repeatedly reviewed
Use of Class Time	Limited time for feedback after mock interview	Ability to provide detailed video feedback after class
Reflection Level	Lack of self-awareness	Greater expressive ability and self-objectification
Acceptance	Only some students incorporated feedback	All students repeatedly reviewed and improved weekly
Preparation Level	Large disparities in interview readiness	Overall improvement by applying prior feedback
Efficiency	Difficult to manage and record assignments/feedback	Integrated management and easy tracking of student growth and feedback

3.1.3. Qualitative cases and student feedback

"Each week, I became more confident by identifying shortcomings and reflecting instructor feedback. (Student A, year 4)"

"I could compare my progress by week and was convinced I could change with steady effort. (Student B, year 4)"

"By consistently working on weekly assignments, I realized that improvement is possible and gained confidence about future efforts. (Student C, year 4)"

"Seeing all my growth documented in Book Creator makes it useful as a job preparation portfolio. (Student D, year 4)"

"I got detailed, step-by-step feedback on each assignment, so I felt a sense of real interaction even in remote learning. (Student E, year 4)"

3.1.4. Summary and implications

After six weeks of lessons utilizing Book Creator, students demonstrated systematic and meaningful development in self-directed learning ability, practical skills, confidence, and course satisfaction. Iterative self-reflection, instructor feedback, and visual portfolio management were confirmed as key elements maximizing the effectiveness of practice-based education. In line with prior research[5][9], digital, repeated feedback and documented progress significantly enhanced self-efficacy and learning engagement. From the instructor's perspective, tracking student growth and designing tailored feedback became much easier, improving the efficiency of course management. This case provides strong evidence for the practical applicability, scalability, and educational value of edu-tech tools in airline service practice education.

3.2. Airline operations practical English course

3.2.1. Course outcomes and changes

3.2.1.1. Improvement in practical English communication skills

Students engaged in repeated practice of English in realistic airline service scenarios, thereby systematically enhancing their repertoire of expressions and situational coping abilities required in the field. Through iterative video assignments and continuous feedback incorporation, steady improvements were noted in areas such as pronunciation, intonation, vocabulary choice, and sentence structure. According to the end-of-term survey, students responded that "practicing real situations in English increased my confidence and reduced my fear of using practical English".

3.2.1.2. Enhancement of self-directed learning and motivation

Every week, students set their learning goals and identified areas for improvement by reviewing their video performances and reflective logs. By actively integrating feedback from instructors and peers, they developed a stronger capacity for self-directed learning planning and adaptability. Many students noted, "Being able to see my own progress directly in the videos greatly increased my motivation for learning".

3.2.1.3. Course satisfaction and positive self-image formation

The system of submitting assignments and managing digital portfolios using Book Creator brought high satisfaction in terms of integrated resource management, repeated practice, and real-time feedback from both instructors and peers. Students commented that "video-recorded instructor feedback was very helpful for making real improvements," and "the portfolio can be used for job preparation". While some initially felt burdened by recording and editing videos in English, regular practice and stepwise instructor guidance helped them overcome these challenges.

3.2.1.4. Qualitative cases and student feedback

"I gained confidence in my practical English by practicing real situations. (Student 1, year 3) "

"I can see my English skills clearly improving through Book Creator and use my portfolio as part of job preparation (Student 2, year 3)".

"At first, recording in English was daunting, but I improved naturally through repeated practice (Student 3, year 3)".

"Seeing my week-to-week progress side-by-side made me believe consistent effort leads to change (Student 4, year 3)".

"By looking back at each week's assignment, I truly felt I was improving and gained confidence for future growth (Student 5, year 3)".

3.2.1.5. Summary and implications

The six-week application of Book Creator in the Airline Operations Practical English course resulted in positive changes across practical English communication skills, self-directed learning, motivation, and course satisfaction. Above all, repeated self-reflection, instructor feedback, and visual portfolio management were shown to maximize the effectiveness of practical language education. These outcomes substantiate the practical value and effectiveness of edu-tech tools in airline English education, providing evidence for their continued and expanded implementation.

4. Future Development Plans and Suggestions

This study empirically examined the effects of six weeks of learner-centered instruction using Book Creator in both the practical and language courses of an airline service management program.

4.1. Academic conclusions

Book Creator-based instruction contributed to the enhancement of self-regulated learning abilities and practical English/interview skills of students. Students systematically managed video assignments and reflection records, visually tracking their growth through repeated self-reflection and instructor feedback. This process stimulated self-assessment, goal setting, and self-regulation as emphasized in Bandura's (1997) self-efficacy theory[23], and Zimmerman's (2000) self-regulated learning theory[22]. Based on Dweck's (2006) growth mindset theory[24], students began to approach failure and improvement with a positive perspective, enhancing both self-efficacy and confidence.

In particular, repeated interview practice and concrete feedback in Air Service Interview Practice improved both competency and confidence, while the Airline Operations Practical English course supported communication skills and motivation through practical language use and self-directed reflection. In both cases, Book Creator portfolios and real-time feedback improved satisfaction, self-directedness, and positive self-image among students.

These findings align with the study's purpose of demonstrating the educational effects of digital tool utilization in practice-oriented airline service and language education. Using multimedia edu-tech tools such as Book Creator for assignments, workbook creation, and video recording not only strengthens learning motivation and self-image[4], but also supports improvements in information literacy, collaborative communication, confidence, and self-esteem[2][8][28]. For instructors, greater instructional efficiency and the ability to deliver personalized feedback reduce workload and enhance educational support[3]. Furthermore, edu-tech-based teaching positively influences self-directed learning, self-efficacy, and learning outcomes, supporting the development of collaborative professionalism and future skills[16][28].

4.2. Student feedback analysis (Balanced perspective)

While student feedback predominantly highlighted positive outcomes such as growth, confidence, and satisfaction, several challenges were also noted. First, initial technical adaptation required 1-2 weeks (Student C, Year 4). Second, collaborative editing conflicts in team projects (Student 2, Year 3). Third, video upload capacity limits causing re-recording (Student F, Year 4). These issues were mitigated 80% through Week 1 templates and guides, suggesting the need for enhanced pre-course digital literacy workshops.

4.3. Suggestions

Based on these findings, several practical suggestions are offered,

First, strategies for edu-tech utilization tailored to the unique needs and objectives of each major discipline should be developed and actively applied to better support students' authentic skills development and autonomous growth.

Second, it is important to create a learning environment that actively promotes iterative self-reflection and feedback. Tools like Book Creator help students visually track their progress and adjust development in response to real-time feedback from instructors and peers, thereby further strengthening self-directed learning.

Third, portfolio-based assessment and career support should be reinforced. Individual e-portfolio books in Book Creator can function as credible evidence of practical competence and growth during job search and recruitment, so assessment methods and career-oriented support systems should be systematically implemented.

Fourth, both instructors and students require structured opportunities to develop their digital skills. As unfamiliarity with edu-tech tools can cause difficulties or resistance at first, providing step-by-step user guides, practical workshops, and ongoing technical support is critical for widespread adoption.

Finally, further research should expand edu-tech case studies to other disciplines, including service management, and investigate long-term and longitudinal effects on students' self-regulated learning and practical competency. Such research can provide concrete directions for educational innovation that helps foster genuinely self-directed, skilled, and employable graduates for the future.

4.4. Limitations

This study presents several limitations. First, the limited sample sizes from two courses within a single Aviation Department (Airline Service Interview Practice, n=8; Airline Operations Practical English, n=12) constrain the generalizability of the findings. This necessitates multi-institutional comparative research involving aviation departments across various universities. Second, the reliance on self-reported survey and interview data introduces risks of subjectivity, while the absence of a control group further restricts causal inference. Finally, the observed effects were captured over a relatively short period of only six weeks, indicating the need for longitudinal studies extending over one or more semesters to verify sustained impact.

5. References

5.1. Journal articles

[1] Lee SH & Hwang SM & Kim SH & Kwon YJ. Analysis of Research Trends on HTHT (High Tech High Touch) Teaching Methods by Year(2011-2023) and Academic Disciplines. *Journal of Educational Information and Media*, 30(6), 1895-1912 (2024).

[3] Jung MH & Kim SY & Na YJ. An Analytical Study on the Awareness and Demands of Professors and Students in Introducing EduTech in University: Focusing on the Case of D University. *CNU Journal of Educational Studies*, 41(3), 31-53 (2020).

[4] Yoon HR. A Study on EduTech Activation Methods for Learners in University Education. *The Journal of Humanities and Social Science*, 13(1), 3135-3148 (2022).

[5] Jin P & Lee EB. Influence of Pre-service Teachers' EduTech Instructional Competency and AI Digital Textbook Competency on Intention to Use AI Digital Textbooks: Focusing on Technology Acceptance Model. *Journal of Educational Technology*, 40(3), 693-717 (2024).

[6] Kang YD. A Creative Teaching Method Combining Media Contents Technology with English and American Literature. *International Journal of Human & Disaster*, 4(2), 15-24 (2019). [Read More]

[7] Choi SH. Case analysis and Suggestions for University Chinese Classes using Digital Tools. *Korea Journal of Chinese Linguistics*, 115, 167-199 (2024).

[8] Kim SY. A Case Study of Teaching English Poetry through EduTech Tools with Learner-centered Approach: Focusing on Fostering Pre-service English Teachers' Collaborative Communication, Creative Thinking and Digital Literacy. *Secondary English Education*, 17(1), 235-253 (2024).

[9] Hong SJ & Choi YI & Ahn JY. Exploring the Feasibility of Personalized Learning using EduTech' in Classroom Instruction. *Korean Journal of Teacher Education*, 40(3), 171-195 (2024).

[10] Park SW. Understanding the Educational Impact of EduTech: Insights from Elementary School Teachers. *Journal of Education & Culture*, 31(3), 371-402 (2024).

[12]Lee S & Ahn S. Exploring Innovative Teaching Methods in University Major Courses through the Synergy of AI and HI. *Robotics & AI Ethics*, 9(0), 34-47 (2024). [Read More]

[13] Kim P. Utilization Status and Effectiveness Analysis of EduTechs in Elementary and Secondary Schools. *Asia-pacific Journal of Convergent Research Interchange*, 10(12), 297-309 (2024).

[19] Son JM & Lee SH. Proposal for the Development of Educational Programs based on a Systematic Review of Research Trends in Digital Textbook Implementation. *The Journal of Korean Association of Computer Education*, 27(8), 17-35 (2024).

[20] Kim SR & Kwon JH. A Case Study on the Use of Digital Learning Tools for Effective Class Operation. *Journal of Korea Society of Digital Industry and Information Management*, 19(2), 1-10 (2023).

[21] Lee HS & Seo EH. A Comparative Study on the Class Satisfaction between Remote Video Class and Face-to-face Class. *The Journal of the Korea Contents Association*, 21(7), 440-447 (2021).

[25] Cho HW. The Impact of ARCS Motivation Strategies through Digital Learning Tools on Online Learning Continuance Intention and Achievement: The Mediating Role of Course Satisfaction. *The Journal of Learner-centered Curriculum and Instruction*, 24(23), 624-636 (2024).

[26] Lim CI & Hong MY & Park TJ. Development and Effects of an Online-based Instructional Model for a Collegiate Course Utilizing Creative Problem Solving(CPS). *Journal of Korean Association for Educational Information and Media*, 17(3), 399-422 (2011).

[27] Kim DH & Park PW. A Study of e-Book Production Lessons using SNS Type on the Academic Achievement and Learning Attitudes of Elementary School Students. *Journal of the Korea Association of Information Education*, 20(1), 29-38 (2016).

[28] Kim J. A Case Study of the Flipped Learning based on Google Classroom -Focused on the Japanese Service Practical Conversation-. *International Journal of Human & Disaster*, 5(2), 21-29 (2020). [Read More]

[29] Kang Y. Design of Liberal Arts Curriculum Centered on Core Competencies. *International Journal of Human & Disaster*, 5(2), 49-57 (2020). [Read More]

[30] Kang Y. A Study on Development of Interactive Communication Education Model for Creative Convergence in Public Value. *Public Value*, 6(3), 61-69 (2021). [Read More]

[32] Ju EJ & Kim SH. Qualitative Research on the Experiences of Career Story Creation Activities to Explore Changes in Career Motivation among Middle School Students: Utilization of Book Creator and Metaverse as Educational Media. *The Korea Journal of Youth Counseling*, 32(2), 271-296 (2024).

5.2. Thesis degree

[11] Kim J. Effect of Interactive Features of e-books on Academic Engagement and Achievement: Focusing on Elementary Korean Textbooks. Hongik University, Master's Thesis (2016).

[14] Park M H. Study on e-book Composition Strategies, Hansung University, Doctoral Thesis (2013).

[15] Yoon E. The Effect of using Video Materials on Student Engagement in EFL Classrooms. Sookmyung Women's University, Master's Thesis (2013).

[16] Byun EH. Influence of Participants' Interactions on Education Performance in Enterprise's non-Face-to-Face Real-time Education Training: Focusing on the Mediating Effect of Learning Immersion. Hoseo University, Master's Thesis (2021).

5.3. Books

[2] Chungcheongbuk-do Office of Education. Designated Digital Education Research School Operation Plan: Enhancing Collaborative Communication Competency through Classes Aiding Conceptual Understanding of Digital Textbooks (2024).

[18] Gyeonggi Provincial Office of Education. Meeting Edu and Tech: Understanding EduTech-based Education Materials (2023).

[22] Zimmerman BJ. Attaining Self-regulation: A Social Cognitive Perspective. Academic Press (2000).

[23] Bandura A. Self-efficacy: The Exercise of Control. W. H. Freeman and Company (1997).

[24] Dweck C. S. Mindset: The New Psychology of Success. Random House (2006).

[31] Jeonbuk Office of Education. The World of Digital Books Unfolded with Book Creator. Jeonbuk Office of Education (2025).

5.4. Additional Reference

[17] https://www.khan.co.kr (2020).

6. Appendix

6.1. Author's contribution

	Initial name	Contribution
Lead Author	SL	-Set of concepts ☑ -Design ☑ -Getting results ☑ -Analysis ☑ -Make a significant contribution to collection ☑ -Final approval of the paper ☑
Corresponding Author*	KC	-Corresponding ☑ -Play a decisive role in modification ☑ -Significant contributions to concepts, designs, practices, analysis and interpretation of data ☑ -Participants in Drafting and Revising Papers ☑ -Someone who can explain all aspects of the paper ☑

2025 Vol.10 No.0, 14-23 Received: 2025/08/30, Peer-reviewed: 2025/09/30, Accepted: 2025/10/25, Published: 2025/12/30

Robotics & AI Ethics

Publisher: J-INSTITUTE
ISSN: 2435-3345

Website: j-institute.org

Corresponding author*
E-mail: skcho@ikw.ac.kr

DOI Address:
dx.doi.org/10.22471/ai.2025.10.0.14

A Study on the Investigative Powers of Northeast Asian Intelligence Agencies in the Era of AI, Deepfake Advanced Technologies

Sunggu Jo

Kyungwoon University, Assistant Professor, Republic of Korea

Abstract

Purpose: *Northeast Asia has experienced rapid economic growth, leading to affluence. At the same time, the development of the Internet has led to an indiscriminate influx of information, transforming the security environment. In a knowledge-based information society, the emergence of the Internet and social networking services (SNS) has made national security inextricably linked to (1) technical cyberattacks and (2) psychological cyberattacks. While espionage in the past was conducted under orders, it is now shifting to the role of self-generated national security crimes. This is a tactic that uses the enemy nation's internet and social media operations to organically increase anti-state organizations, ultimately fostering a social atmosphere that benefits their own nation's interests.*

Method: *Based on previous studies, this research established an analytical framework to demonstrate the necessity of the study, drawing on current laws, domestic and international monographs, academic articles, research reports, legislative materials from the National Assembly, news articles, and statistical data from government agencies.*

Results: *This study is an expanded and revised English version of a paper originally published in Korean. Building on the previous discussion of establishing a personnel management system for intelligence agencies, expanding professional manpower, and strengthening inter-agency cooperation, this study further examines the necessity of investigative powers for intelligence agencies in the era of AI and deepfake technologies.*

Conclusion: *Northeast Asia, a buffer state between major powers, has consistently faced national security crises such as war, terrorism, and assassinations. Now, in the era of AI and deepfake technology, it faces a new phase. Therefore, this study examines the threats we face and suggests the role of intelligence agencies in the age of AI and deepfake technology.*

Keywords: ***AI, Deepfake Technology, Northeast Asia's Buffer State, Investigative Powers, Intelligence Agencies***

1. Purpose of the Study

In recent years, Northeast Asia's Exit and Entry policies, which have fostered an indiscriminate influx of foreigners, have not only created a multi-ethnic labor market but have also created problems for international criminal organizations to establish themselves in neighboring countries. Currently, there are approximately 750,000 to 800,000 ethnic Koreans in China in South Korea, meaning that one in three ethnic Koreans in China resides in South Korea. Thus, the scope of these ethnic Koreans' involvement in South Korea has expanded beyond organized crime, such as drug distribution and human trafficking, to corporate crimes[1], and industrial technology theft. Currently, they are expanding into crimes such as voice phishing throughout Southeast Asia, including Cambodia[2].

* This article is an expanded and revised English version of an article originally published in Journal of Korean Public Police and Security Studies, Vol.16 no.2, 2019 in Korean. This English version has undergone review by the Editorial Board and double-blind peer review to enhance its international visibility.

China, which achieved rapid economic growth in Northeast Asia, has established ‘freedom of choice’ in the realm of human cognition through relative poverty resulting from economic polarization and the growth of market-driven consciousness. However, the citizens’ discontent with the Communist Party’s dictatorship has led to the gradual organization of collective action, and the violent independence movements that traditionally occurred in ethnic minorities such as Tibet and Xinjiang Uyghurs have now become widespread and spread throughout China, threatening the Chinese government’s law enforcement.

This problem isn't limited to China. North Korea poses a more serious threat to Northeast Asia. According to South Korea's Ministry of Unification, the number of North Korean defectors officially estimated at around 30,000 in South Korea alone. Many defectors who enter South Korea accumulate funds and then use brokers to bring their remaining family members back to North Korea. Indeed, a significant number of recent arrivals to South Korea are in this situation [3][4][5]. The problem is that some North Korean defectors maintain connections with North Korea even after their departure. In this process, spies dispatched by the Ministry of State Security, North Korea's intelligence agency, have expanded their scope of activity. North Korean intelligence agencies are using North Korean defectors as double agents under the condition that their families remain in North Korea be safe[6]. North Korean defectors who were recruited and forced to become double agents have been caught numerous times in South Korea alone, collecting information on the country where they settled and providing it to North Korea. While Northeast Asia is generally perceived as a low risk of terrorism due to the prohibition of gun ownership, Northeast Asia is classified as extremely dangerous by foreigners due to North Korea.

This activity is even more active online, and the emergence of new threats in the era of AI and deepfake technology is creating a new role for intelligence agencies responsible for overseas intelligence[7]. However, in liberal countries, there has been a strong tendency to interpret intelligence agencies' work in an overly political manner, contrary to their essential role. Due to this social climate, this issue has not been addressed in detail in academic circles.

Research conducted in South Korea reveals that Lim Jun-tae (2006) argues that while national security is a key objective for intelligence agencies in a divided Korea, internationalization and openness have diversified the scope of intelligence activities, raising the need to analyze the scope of work of South Korean intelligence agencies[8]. Han Sang-bong (2008) argues that in South Korea, since the Cold War, the president bases his decisions on information provided by intelligence agencies regarding national security, making the role of intelligence agencies crucial[9]. Kim Jeong-do (2009) points to ideological conflict between political parties and the dominant relationship between the Blue House and the National Assembly as reasons for the failure to fully exercise control over the National Intelligence Service[10]. In contrast, in North Korea and China, even initiating such discussions is impossible, and this creates a burden even in academic research, which should be an endless academic challenge and adventure, as it is expected to result in oppression by the state power for the researcher and his or her family.

In North Korea and China, where state power is often insensitive to human rights violations, the debate over the existence of investigative powers is unnecessary. However, even in liberal countries like South Korea and Japan, investigative powers are a crucial institution for intelligence agencies. By streamlining the decision-making process, it not only enables timely and seamless blocking and defense against threat information, but also allows for early identification of threats through investigative authority. In addition, it ensures the “security” of preserving evidence for trial, thereby preventing additional threats. Moreover, leveraging comprehensive threat data strengthens interagency cooperation within the government, creating the advantage of maximizing synergistic effects[11][12].

Therefore, this study proposes the necessity of investigative authority for the National Intelligence Service (NIS) of South Korea and the Cabinet Research Office (CIRO) of Japan as intelligence agencies in Northeast Asia in the era of advanced AI and deepfake technologies, and hopes that it can serve as basic data for legislative materials in each country.

2. Changes in the Northeast Asian Security Environment: From a Korean Perspective[1]

The changes in the security environment according to Jo Seong-gu (2019) are discussed as follows. Historically, the Korean Peninsula, due to its geographical proximity to surrounding major powers, faced the primary threat of invasion from continental China until the Joseon Dynasty (1392-1897). From the late Joseon Dynasty (1897-1910) to the Japanese colonial period (1910-1945), Japan experienced a collapse in security and the loss of sovereignty due to invasions by Japan ahead of its modernization. After Japan's defeat in the Pacific War (1941.12-1945.9), the Korean Peninsula, China, and Southeast Asian countries under Japanese colonial rule were liberated. However, the Korean Peninsula entered the sphere of influence of the Soviet Union and the United States according to the agreements of the Yalta Conference. The Republic of Korea, a liberal democratic government with the support of the United States and the United Nations, was established in the south of the 38th parallel, while the communist North Korea, with support from the Soviet Union, was established in the north. This decision was reached at the Yalta Conference, which took place on the Crimean Peninsula along the Black Sea coast of the Soviet Union. The conference brought together the United States, the United Kingdom, and the Soviet Union to discuss their defeat in World War II and its management. South Korea, a party to the conference, was excluded.

However, in January 1950, the Harry S. Truman administration announced the Acheson Line Declaration, which excluded South Korea from the US Far East defense line. Following the withdrawal of US troops from South Korea, North Korea invaded South Korea with the aim of communizing the country. China, in turn, provided 300,000 People's Liberation Army troops to North Korea, blocking any chance of Korean unification. This incident, known as the "Secret Military Agreement between North Korea and the Soviet Union," was an agreement on economic and military cooperation between North Korea and the Soviet Union. A North Korean delegation visited Moscow, met with Stalin, and signed an economic and cultural exchange agreement and other secret agreements. According to published diplomatic documents, the issue of unification of North and South Korea by force was discussed between North Korea and the Soviet Union.

The crucial point here is that, starting with this war, China, which had long joined forces in the struggle against Japanese imperialism, became an enemy. Meanwhile, the United States and other UN forces, faced with the worst security crisis ever, with the entire country, except for the Daegu and Busan areas, occupied by North Korean communist forces, willingly participated in another country's war and risked the sacrifice of its young soldiers to protect liberal democracy. Furthermore, the stationing of US and UN forces in Korea was intended to deter communist aggression. The Mutual Defense Treaty between the Republic of Korea and the United States of America (MODA) was signed on October 1, 1953, and entered into force on November 18, 1954, as Treaty No. 34.

At the time, President Syngman Rhee emphasized the necessity of the MDA during a meeting with the US ambassador. However, prior to the armistice, the United States requested a mutual

[1] Jo S. A Critical Review of the Transfer of Presidential Security Work to the Police. Korean Security Journal, 58,182-183 (2019).

defense treaty, but the American response was lukewarm. The United States had a strong tradition of isolationism, and at the time, the Philippines was the only country with which it had a bilateral mutual defense treaty. To this day, the United Kingdom, Japan, and the Philippines remain the only countries with a mutual defense treaty, aside from South Korea. For the United Kingdom, the treaty provided US nuclear technology, while for Japan, it was tied to a ban on rearmament.

However, given the current situation in Northeast Asia, the United States finds it difficult to contain China without South Korea, and China has more diplomatic hurdles to overcome with South Korea. Furthermore, the threat of war has recently loomed over Japan-China relations. Japan is now firmly tied to Taiwan's security due to trade routes off the coast of Taiwan and the Senkaku Islands.

For South Korea, the primary enemy has been (1) China, which treated Joseon as a tributary state for 500 years during the Joseon Dynasty; (2) Japan, which invaded and colonized the country; and (3) North Korea, which initiated the Korean War of Invasion, and China, which supported it. This concept of primary enemy has shifted based on national interests. This concept will likely continue to evolve based on realistic criteria.

A notable point in Northeast Asian international relations is the relationship between the United States and Japan. While both countries fought in the Pacific War (1941-1945), and the families of those killed still live there, their shared interests have led to the formation of a new alliance. South Korea, too, must consider the threat of war with North Korea. Within the secure security framework of the ROK-US alliance, it also forges a military alliance with Japan, the United States' largest ally in Northeast Asia. This also reflects China's expansion and the ideological solidarity of the liberal democratic camp.

Although Sinophobia is at a serious level in Korea today, there is still a deep historical hostility toward Japan, despite the sober recognition that military training with Japan is desperately needed in the joint military operations in wartime within the ROK-US-Japan alliance system[13][14].

3. Analysis of Today’s Issues Through Past Cases

3.1. Continued provocations under conditions of war and armistice in Northeast Asia

At the end of the Joseon Dynasty, the Qing Dynasty, which had intervened in Joseon, declined, and the Russian Empire and Japan began to struggle for hegemony on the Korean Peninsula. At the time, the Russian Empire argued that the area north of the 38th parallel in Joseon should serve as a buffer state between Russia and Japan to avoid conflict. However, with the onset of the Cold War following World War II, Russia pursued the establishment of a buffer state along the 38th parallel, as per its previous claim. Subsequently, not only war but also terrorist attacks occurred frequently on the Korean Peninsula.

The term "buffer state" or "buffer zone" refers to a small, weak state located between major powers, serving to ease the tensions that would otherwise arise from their direct borders. At the time, the 38th parallel, which divided Korea's territory, was a world-renowned buffer zone between Russia and Japan.

The Cold War ended in 1991, but with the emergence of a world order based on major powers, the concept of security also changed. While the threat of military confrontation was the core of national security during the Cold War, the post-Cold War era saw a shift away from this military focus toward comprehensive security. While ideological conflicts between nations diminished after the end of the Cold War, other forms of conflict emerged, such as bizarre conflicts over race, ethnicity, religion, and borders. These conflicts have increased even more since the Cold War ended than during its peak. The security reality resulting from these changes in international relations is the growing threat of terrorism, making security an even more critical issue [11]. Furthermore, indiscriminate cyberattacks clearly demonstrate the lack of trust between nations. Just as Europe suffered numerous casualties after World War II and the European Union (EU) was established, Northeast Asia may also undergo a similar process. Therefore, adopting the EU model in Northeast Asia could foreshadow a future of reform and openness for North Korea, much like Russia's post-Cold War opening and China's adoption of a market economy. At the very least, lifting restrictions on residents along the Tumen and Yalu Rivers borders, thereby securing flexibility, would be China's only way to maintain its position as a G2 partner with the United States. Failure to do so would likely lead to China's fragmentation under the current Communist regime, fragmenting into multiple ethnic groups, experts in Northeast Asia predict.

3.2. Social and political issues and their corresponding models of change

Amidst these political experiments, political powers will continue to increase their attempts to bring intelligence agencies under their influence.[15] Driven by public opinion, political powers will increasingly weigh intelligence agencies against their own. Even amidst this, intelligence agencies must continue to make unseen efforts to protect the safety of citizens and property by gathering information related to national security, preventing enemy destruction of facilities, mass casualties, and espionage.

Recent investigations in Northeast Asia, with their "international nature" extending beyond borders, the "timeliness" and "integrity" of domestic and international investigative information sharing, and the "secrecy" of investigators' roles, are distinct from the work and fundamental nature of law enforcement agencies such as the police [16].

However, the Moon Jae-in administration in South Korea has abolished the domestic branch of the National Intelligence Service (NIS) and reformed it into a dedicated overseas intelligence agency, similar to the CIA in the United States. In line with this, the National Assembly has amended the National Intelligence Service Act to abolish its domestic operations and investigative authority. Meanwhile, some argue that it would be more desirable for South Korea's National Intelligence Service to adopt the FBI (Federal Bureau of Investigation) model, rather than the CIA model, given our current reality. This is because the FBI, founded in 1908 for the purpose of criminal investigation and intelligence gathering, investigates violations of U.S. federal law and collects public security information. It currently has 56 overseas bureaus and over 500 branch offices. No one, including the President or the National Assembly, can interfere with investigative activities or personnel decisions, and the FBI reportedly has around 20,000 employees. The FBI's investigative authority encompasses: ① crimes related to national security, such as insurrection, espionage, sabotage, or obstruction of the military; ② kidnapping and kidnapping; ③ bank robbery, theft, and embezzlement and corruption cases involving bank employees; ④ auto theft and robbery across two states; ⑤ bribery involving federal officials; ⑥ interstate transportation of stolen property; ⑦ check forgery and use; ⑧ destruction of aircraft and passenger vehicles; ⑨ investigations of high-profile fugitives; and ⑩ fraud and civil cases against the federal government.

However, in Northeast Asia, countries like South Korea and Japan are sensitive to personal information, and clear legislation addressing these concerns is lacking. Ultimately, the work of intelligence agencies has forced them to straddle the line between legality and illegality. Intelligence agencies must verify criminal activity not only abroad but also within their own borders, including against their own citizens who are linked to spies operating within the country. However, the reality is that criminal activity cannot be determined until a certain level of intelligence gathering and investigation has been completed. Therefore, a clear definition of civilians is necessary when prohibiting intelligence activities targeting civilians within a country. This is perhaps the greatest dilemma that arises between citizens' privacy and national security. However, the core criteria must be a clear basis for investigation and intelligence gathering. Such basis must include: 1) Is there sufficient evidence to warrant the investigation and intelligence gathering? 2) Is the purpose of the investigation and intelligence gathering consistent with current law? 3) Is there a set period within which the investigation and intelligence gathering must be terminated if a direct causal relationship between the information obtained through the investigation and intelligence gathering is not confirmed? Legislation that clearly specifies these criteria is necessary. Ultimately, the outcome of the investigation and intelligence gathering should be the primary focus of any future legal disputes regarding the illegality of the investigation and intelligence gathering.

4. Discussion[2]

Today, crime is rapidly transcending national borders and becoming increasingly internationalized[17]. The widespread adoption of international travel and the proliferation of the internet have improved quality of life while blurring the lines of threat, and the activities of soldiers and foreigners are largely unrestricted. Furthermore, international division of labor and the advancement of advanced technology are driving the development of increasingly organized and corporate forms of crime. In particular, cybercrimes utilizing AI, such as voice phishing, which have recently surged, are linked to overseas organizations, making them distinct from traditional criminal offenses[18][19].

North Korea currently possesses the world's fourth-largest cyber power. It systematically trains cyber experts through institutions such as Kim Il-sung Political Military University, Kim Chaek University of Technology, and Pyongyang Computer Technology University. Under the Kim Jong-un regime, this power has been developed into one of three asymmetrical forces, alongside nuclear weapons and ballistic missiles. Furthermore, North Korea is engaging in international hacking, money theft, and information manipulation activities, all in conjunction with information networks spread across China, Southeast Asia, and Central and South America, and this expansion is expected to continue[20][21]. In particular, the exploitation of AI deepfakes and artificial neural networks makes it difficult to respond solely with the investigative capabilities of frontline police stations.

To counter these threats, advanced countries around the world are establishing specialized investigative bodies tailored to their specific circumstances[22]. Some countries, such as the US, the Netherlands, and Singapore, prioritize expertise to the point of granting investigative authority to private sector experts, not just public officials. This signals the advent of an era of special judicial police powers to respond to cutting-edge technology crimes like AI and deepfakes.

Special judicial police powers, based on Article 196 of the Korean Criminal Procedure Act, empower authorities to exercise authority across all aspects of criminal justice procedures, including criminal investigations, suspect arrests, evidence collection, and case referrals. In Korea, this authority is already exercised not only by the police but also by several central government ministries and local governments, including the Ministry of Employment and Labor, the Korea Customs Service, the Ministry of National Defense, the Ministry of Land, Infrastructure and Transport, the National Railroad Police, the Ministry of Gender Equality and Family, the Korea

[2] Jo S. Why do we still have Specialized Judicial Police Powers in the age of AI and deepfake technology?, Korea Youth Newspaper, August 25 (2025).

Forest Service, the Ministry of Trade, Industry and Energy, the Ministry of Food and Drug Safety, the National Fire Agency, the Korea Coast Guard, the Ministry of Oceans and Fisheries, the Ministry of Environment, and the Korean Intellectual Property Office. This system demonstrates the practical and effective role of special judicial police powers in crime prevention and investigation, with approximately 9,000 crimes detected annually.

Therefore, both Korea and Japan should swiftly institutionalize special judicial police powers specifically focused on AI. This will go beyond simply responding to crime and serve as a robust shield that will safeguard the healthy development of cutting-edge science and technology and the lives of citizens for whom the internet has become a fundamental part of their lives. We are at a critical juncture to safeguard our domestic cutting-edge industries and secure international competitiveness.

5. Conclusion

5.1. Establishing a personnel management system: shifting from political ideology to a practical threat-centered approach

According to media reports, Google Chairman Eric Schmidt strongly criticized the National Security Agency (NSA)'s surveillance of civilians. NSA Director Keith Alexander countered this by stating in the U.S. Senate that "intelligence gathering, including the 2009 New York City subway bombings, has prevented dozens of recent terrorist attacks[23] ".

Intelligence agencies must continue their work regardless of regime change. Due to the nature of their work, they operate within institutional frameworks, not through individual actions [24]. An institutional framework consists of a leader who oversees the security agency, subordinates who receive their orders, and a support system that manages these subordinates. Decisions regarding mission execution are solely the responsibility of the leader, and the success or failure of the mission is the responsibility of the leader. Furthermore, when missions are carried out within an institutional framework, the chain of command must not be disrupted by orders from other leaders.

However, in liberal countries, when a democratic transfer of power occurs and a regime favoring a particular country takes power, the political perspective of handling intelligence often leads to devolution of authority, organizational downsizing, personnel changes, and other drastic changes, making it impossible to pursue consistent policies. Therefore, to ensure the political neutrality and policy consistency of security agencies, even if the current president retains the authority to appoint personnel, the heads of security-related agencies must establish a legal basis for a personnel system that requires them to meet minimum requirements, such as 1) internal promotions and 2) restrictions on retirees with a certain period of experience. Intelligence agency heads should not lead their organizations solely for the sake of the president.

5.2. Expanding cyber specialists

To gather information for national security, we need to expand our intelligence workforce to gain access to more information. From a national security perspective, deploying overseas intelligence personnel in neighboring countries, like North Korea, China, Russia, and Japan, and gathering high-level intelligence to protect national security is essential to prevent war and terrorism.

Furthermore, we must expand our counterintelligence workforce to conduct more comprehensive counterintelligence activities to apprehend enemy spies infiltrating the country. We must also safeguard corporate interests through industrial technology protection activities. We must also proactively block enemy sabotage of key national facilities and public facilities to prevent terrorist attacks, including mass casualties. Sabotage is a deliberate act aimed at weakening an enemy or employer by overturning, disrupting, disrupting, or destroying production facilities and transport equipment. In addition, we must continuously discuss the effectiveness of related laws with relevant academic circles, and through the advancement of related studies, we must research and develop theoretical grounds for counterintelligence activities. Based on this, we must educate the public so that they can easily recognize the seriousness of the damage caused by enemy spy activities, and impress upon them that all citizens are the main players in protecting national security.

5.3. Building an intelligence agency cooperation system: establishing a U.S. DNI model

Counterintelligence activities by security agencies require a networked, coordinated system that integrates all relevant agencies to produce intelligence. Indeed, the September 11, 2001, terrorist attacks on the World Trade Center in the United States, in which al-Qaeda hijacked a plane and killed 2,996 people, were attributed to a lack of intelligence coordination among the 16 US intelligence agencies, which prevented the attack. Subsequently, the Office of the Director of National Intelligence (DNI) was established to network and supervise the 16 US intelligence agencies. The DNI is the highest intelligence agency in the US. The Intelligence Reform Act, passed by the Senate on December 7, 2004, was established following the September 11, 2001, terrorist attacks, citing the need for intelligence agency reform.

Considering the current security landscape facing threats from neighboring countries, including North Korea, South Korea also needs a networked intelligence agency system similar to that of the United States. South Korea's intelligence agencies include the National Intelligence Service, the Defense Counterintelligence Command, the Cyber Operations Command, the Defense Intelligence Agency (DIA) (DIA, 777th Command), and the National Police Agency (Intelligence Bureau, Foreign Affairs Bureau, and Security Bureau). How are they preparing for the era of advanced technologies like AI and deepfakes?

6. References

6.1. Journal articles

[1] Lee MG. An Old Friend Turned Criminal: An Empirical Analysis of Anti-Chinese Sentiment in Korea. *Korean Chinese Relations Review*, 11(3), 117-134 (2025).

[2] Kim TM & Kwon HY. Limitations and Success Requirements for International Investigative Assistance in Cybercrime Investigations. *The Korean Association of Police Science Review*, 23(6), 179-214 (2021).

[3] Song IH. A Study on the Validity of the Contract Signed between North Korean Defector and Broker. *Human Right and Justice*, n423, 43-59 (2012).

[4] Chung S & Jeong Y. The Mobility Capability and Migration Practice of the North Korean Defectors in South Korea. *Journal of the Korean Geographical Society*, 56(6), 567-584 (2021).

[5] Oh W & Nam SS. The Representation of North Korean Defector in Documentaries: Mobility and Subjectivity in Madame B and Shadow Flowers. *The Journal of Multicultural Society*, 18(3), 115-156 (2025).

[6] Choi H. Social Implications of North Korean Migrants' Remittance to North Korea. *The Journal of Multicultural Society*, 14(3), 325-355 (2021).

[7] Oh SY & Park SJ. National Security Threats Posed by Deepfake Technology and the Legal Response Framework. *Korean Terrorism Studies Review*, 18(3), 168-188 (2025).

[8] Lim JT. A Study on the Reformation of Korean National Intelligence Organizations. *Korean Journal of Public Safety and Criminal Justice*, 22, 377-420 (2006).

[10] Kim J. The Relationship between the National Assembly and the Executive Branch in the Oversight of the National Intelligence Service. *Korean Journal of Political Science*, 16(3), 21-46 (2009).

[11] Jo S. Comprehensive Threats and Directions in Northeast Asia. *International Journal of Terrorism & National Security*, 8, 1-13 (2023). [Read More]

[12] Jo S. A Study on the Integrated Operation of the Presidential Security Service through the Establishment of the Office of National Security(ONS)' National Security Intelligence Investigation Center. *Korean Police Studies Review*, 23(1), 255-257 (2024).

[13] Son K. The Study on the Development of the U.S. Multi-domain Operations with a Doctrinal Perspective. *Journal of Korean-Japanese Military and Culture*, 34, 57-79 (2022).

[14] Cho E. Reorganizing Japan's SDF Command Structure and Its Changing Role: Focusing on the Establishment of the JSDF Joint Operational Command. *Review of International and Area Studies*, 34(2), 111-139 (2025).

[15] Kim YJ. The National Intelligence Service Involvement in Politics and the Agenda for Keeping its Political Neutrality in Korea. *Korean Journal of Law & Society*, 44, 67-95 (2013).

[16] Jo S. A Study on the Creation of the Office of National Security's National Security Investigation Headquarter. *Korean Police Studies Review*, 22(1), 241-258 (2023).

[17] Kim DY. China's Recent Countermeasures against Organized Crime and Implications for Korea -Focusing on the Contents of China's Organized Crime Suppression Law and Related Issues-. *Chinese Law Review*, 58, 1-47 (2025).

[18] Kim HJ. Legal and Policy Strategies for Preventing Emerging Cybercrimes and Enhancing Cyber Security using Generative AI Technology. *Chung-ang Journal of Legal Studies*, 48(2), 49-84 (2024).

[19] Son JK & Song JY. A Study on Cybercrime Detection using Explainable AI Technique. *The Journal of the Institute of Internet, Broadcasting and Communication*, 25(2), 243-249 (2025).

[20] Lee S & Park K. An Analysis of North Korea's Cyber Attacks on International Financial System: Characteristics and Perspectives. *Journal of Global Politics*, 14(2), 87-111 (2021).

[21] Kim JH & Lee KM. Revisiting International Legal Response to North Korea's Cryptocurrency Heist: Enhancing Cyber Deterrence through Hacking-back. *The Quarterly Journal of Defense Policy Studies*, 39(4), 33-63 (2024).

[22] Sin SW. A Study on the Digital Sex Crimes using Deepfake Technology. *Journal of Korean Public Police and Security Studies*, 20(4), 417-168 (2023).

[23] Jo S. The Relationship between Anti-communist Investigation and Domestic Politics. *Korean Security Journal*, 74, 231-248 (2023).

[24] Kim BJ. A Study on the Organizations and Duty of Korean and British Intelligence Agencies. *Korean Journal of Public Safety and Criminal Justice*, 32(3), 103-134 (2023).

6.2. Thesis degree

[9] Han SB. Korea Foreign Policy-making Process and the Role of National Intelligence Organizations. Yonsei University, Master's Thesis (2008).

7. Appendix

7.1. Author's contribution

	Initial name	Contribution
Author	SJ	-Set of concepts ☑ -Design ☑ -Getting results ☑ -Analysis ☑ -Make a significant contribution to collection ☑ -Final approval of the paper ☑ -Corresponding ☑ -Play a decisive role in modification ☑ -Significant contributions to concepts, designs, practices, analysis and interpretation of data ☑ -Participants in Drafting and Revising Papers ☑ -Someone who can explain all aspects of the paper ☑

2025 Vol.10 No.0, 24-41

Received: 2025/08/30, Peer-reviewed: 2025/09/30, Accepted: 2025/10/27, Published: 2025/12/30

Robotics & AI Ethics

Publisher: J-INSTITUTE
ISSN: 2435-3345

Website: j-institute.org

Corresponding author*
E-mail: 205design@naver.com

DOI Address:
dx.doi.org/10.22471/ai.2025.10.0.24

A Study on the Institutionalization and Legal Improvement of Private Security and Security Services using AI and IoT Technology

Dongyeop Lee

Korea Association of Authorized Detectives and Guards, President, Republic of Korea

Abstract

Purpose: *With the rapid advancement of Fourth Industrial Revolution technologies, particularly artificial intelligence (AI) and the Internet of Things (IoT), the private security and protection industry has undergone a fundamental transformation. Advanced technologies such as intelligent CCTV systems, drones, biometric identification, and IoT-based sensor networks have accelerated the shift from labor-intensive security models to technology-driven integrated security systems. Despite this transformation, Korea's Security Services Industry Act has not kept pace with technological innovation, resulting in persistent institutional and legal gaps concerning legal definitions, licensing and supervision frameworks, technology certification, and liability allocation. This study aims to propose directions for the institutionalization and legal reform of AI- and IoT-based private security and protection services in Korea.*

Method: *This study adopts a qualitative research design based on an analysis of recent developments in security technologies and their applications within the private security and protection sector. A comprehensive review of relevant domestic and international literature, legal statutes, and policy documents was conducted. In addition, a comparative legal analysis of major foreign jurisdictions was undertaken to examine how technology-based security services have been legally recognized and regulated. Through this approach, implications for improving Korea's legal and institutional framework were derived.*

Results: *The results reveal a structural imbalance in Korea's private security system, in which legal and institutional reforms significantly lag behind technological adoption. Three major challenges were identified: ensuring transparency and accountability in AI-driven decision-making processes, enhancing the reliability and integrity of data management systems, and safeguarding personal information. These challenges frequently conflict within the existing regulatory framework. Moreover, current legislation lacks clear provisions regarding the legal status of technology-based security services, standardized technology certification systems, and clearly defined supervisory authority, thereby perpetuating regulatory uncertainty.*

Conclusion: *This study concludes that clarifying the legal status of AI- and IoT-based private security and protection services is essential for the sustainable development of the industry. Furthermore, establishing a convergence security governance framework based on cooperation between public and private sectors is necessary. Legal and institutional reforms should prioritize the introduction of technology certification mechanisms, the clarification of accountability structures, and the achievement of a balanced approach between transparency and personal data protection. While this study is limited by its reliance on literature review and institutional analysis, future research incorporating empirical data, in-depth interviews with practitioners and policymakers, and policy simulation studies is recommended to support practical implementation and legislative advancement.*

Keywords: ***Internet of Things(IoT), AI, Private Security, Guard Services, Legal Institutionalization***

1. Introduction

Artificial intelligence (AI) and the Internet of Things (IoT), which are the core technologies of the 4th industrial revolution, change the overall structure of society and are rapidly spreading not only to public security but also to the private security industry. In particular, in the area of security and security, various technologies such as AI image analysis, IoT sensor network, drone surveillance, and biometric access control are combined, and the traditional manpower-centered security system is being transformed into a technology convergence-type integrated security system. Seok Wang-heon (2018) analyzed that the spread of the ICT infrastructure environment directly affects the efficiency of crime prevention policies, and that technology convergence is reorganizing the structure of public and private security services[1].

These technological advances not only improve crime response speed, but also lead to changes in the quality of security services by enabling data-based risk prediction and real-time information sharing.

Bang Jun-sung et al.(2019) discussed the need for policy improvement, including data and personal information issues, as security technology using AI and ICT technology is being advanced from crime prediction and hotspot analysis[2]. In particular, overseas smart crime prevention infrastructure that combines AI and IoT technology has already entered the commercialization stage, and Japan and the European Union (EU) are seeking to balance technology development and regulation by introducing related laws, systems, and certification systems in parallel. These overseas cases show that technological innovation has become a key competitive element in the private security industry, while suggesting that there are significant risks in terms of social trust and personal information protection if Korea does not have a legal response system in the process of accepting technology.

Although the domestic private security industry has shown continuous growth over the past decade, the pace of legal and institutional maintenance has not reached it. The current Security Business Act is designed around the qualifications, permits, and supervision systems of security personnel, so it does not include legal definitions or responsibility regulations for new types of security and security services using AI and IoT-based technologies. Key issues, such as transparency in artificial intelligence decision-making, legality of data processing, technology certification, and information protection standards, depend on individual institutions' self-regulation or private standards without specific legal grounds. As a result, private security services that have introduced new technologies are operating in an institutional vacuum, revealing limitations in securing legal stability and accountability.

Lee Won-sang (2016) pointed out that even though cutting-edge science and technology are rapidly spreading to the security and security areas, the current legislation cannot keep up with this pace of technological change, continuing legal gaps and uncertainties in the subject of responsibility[3]. In other words, despite the rapid progress of technological innovation in the field, legal and institutional discussions to support it are relatively insufficient. In particular, since the introduction of technology in the field of private security involves complex social issues such as personal information, monitoring rights, and the boundary of public responsibility, not just equipment advancement, legal reform has become a key task of public safety governance beyond the dimension of simple industrial policy.

Therefore, this study aims to comprehensively examine the impact of AI and IoT technologies on private security and security services and their legal implications, and to suggest ways to improve the system in line with the flow of technology development. To this end, we first analyze the current state of application and operating system of AI and IoT technology in the security field, secondly review the institutionalization status of technology-based security services by comparing legislative cases in major foreign countries such as Japan, Germany, and the EU, and thirdly, we propose a legalization model for Korean convergence security services. This study aims to supplement the disconnection of existing research through a convergent approach that combines technology trend analysis and literature and system comparison, and to contribute to the establishment of a sustainable institutional foundation for the security and security industry in the private sector.

2. Review of theoretical Background and Prior Research

The development of AI and IoT technology is promoting structural changes in the security and security industry. As the existing physical security-oriented structure evolves into a data-based monitoring and prediction system, it is shifting from a manpower-dependent security system to an integrated technology and information-oriented security system. In particular, artificial intelligence image analysis, network IoT sensors, drone-based monitoring systems, and biometric access control complement the efficiency of existing personnel while enabling automated risk detection and response.

This technological transformation raises a new social task that needs to re-establish the legal and institutional framework for the entire security industry, not just advanced equipment.

2.1. Application of AI and IoT technologies in security and security fields

Recently, AI and IoT-based technologies have become a key tool in security and security services. Park Sang-wook et al. (2020) reported that intelligent CCTV-based dynamic crime prediction technology is developing into a real-time risk calculation and predictive security system by analyzing human attribute, behavior, and environmental data in multiple dimensions beyond the existing simple hotspot analysis[4]. These technologies are being used as a core basis for preemptive risk response not only in public safety but also in private security. Meanwhile, Kwak Yeon-gyu et al. (2023) proposed an autonomous CCTV system using AI technology and analyzed that the combination of artificial intelligence image analysis and IoT-based mobile equipment overcomes the limitations of the fixed monitoring system and enables dynamic response centered on crime areas[5].

In addition, Choi Woo-chul and Na Jun-yeop (2018) suggested that an integrated control system that combines various sensors and artificial intelligence algorithms can be the core of the community's safety net in a study on an integrated crime prevention platform for real-time crime response[6].

As such, security technology is evolving into a multi-layered integrated system through the convergence of information and communication technology (ICT), artificial intelligence, and IoT, but legal standards and accountability structures for this are still insufficient.

2.2. The systematic structure and limitations of the private security and security industry

The domestic private security industry has established an institutional framework since the enactment of the "Security Business Act" in 1976, but the basis of the legislation remains in the traditional manpower-centered security model despite more than 40 years of industrial growth. Article 2 of the "Security Business Act" classifies the security industry into "facility expenses, convoy expenses, personal protection, machine expenses, and special expenses," but this is mainly focused on crackdown and monitoring functions centered on manpower and equipment, showing limitations that it cannot cover service types using advanced technologies such as AI and IoT. In addition, the current system maintains a licensing and supervision system centered on the National Police Agency, and technical certification or data management standards are not specifically stipulated in separate laws or guidelines. As a result, even if security and security services using advanced technologies appear, they are operated with unclear legal status and supervisory authority. Structural problems in which technology investment by private security companies is limited by institutional uncertainty have also been continuously pointed out. In addition, Lee Tae-ho and Park Jun-seok (2022) pointed out that the regulations and procedures of the Security Business Act hinder the original purpose of the development of the security industry, and emphasized the need to discuss system improvement to systematize and systemize field-oriented problems scientifically[7].

Meanwhile, Shin Hyun-joo and Kim Joo-chan (2015) evaluated the rigidity of regulations on the security industry and the inadequacy of technology introduction as obstacles to industrial development, and suggested the need for self-regulation and government-private cooperation regulatory models[8].

As such, the current system does not fully reflect changes in the industrial structure and technological reality, and AI-IoT-based convergence security services do not even have a legal definition. In particular, artificial intelligence image analysis, IoT sensor networks, and drone surveillance systems have been rapidly introduced in the private security and security sectors, but the gap between technological innovation and legislation has not yet been resolved. This gap is not just a delay in legislation, but stems from structural problems in which technology acceptance structures and legal control systems have developed independently of each other. In other words, while technology has rapidly developed with the aim of field-oriented effectiveness, the legal system operates only within the existing licensing and supervision paradigm, forming a dual structure in which normative responsiveness is significantly inferior. This problem not only hinders the security industry's sustainable innovation ecosystem, but also leads to complex side effects such as unclear who is responsible for AI decision-making, conflicts between privacy regulations and technology development, and a decrease in public trust due to the lack of technology certification. Therefore, the core of future legal reform should be not just system supplementation, but establishing an integrated coordination mechanism for technological development and normative systems. The existence of such legal and operational gaps is an important premise for understanding the "related legislative and policy research trends" to be reviewed in the next chapter.

2.3. Analysis of related legislation and policy research trends

Discussions on the institutionalization of the private security and security industry have been ongoing since the 2000s, but most of the research has focused on improving the security industry's licensing and supervision system and manpower management system, and only a small number of studies have dealt with the legal acceptance of new technologies such as AI and IoT. This reflects the structural limitations of legal research compared to the pace of technological development.

Lee Won-sang (2016) pointed out that despite the rapid spread of advanced science and technology in the field of security and security, the current legislation does not have a corresponding normative basis, so it is insufficient to secure legal justification for the use of technology and clarity of the responsible entity[3][9]. This awareness of the problem applies equally to private security services using AI and IoT technologies.

In addition, Eun-jung Kwon et al. (2020) emphasized the necessity of legal change based on a risk-based approach, citing artificial intelligence, autonomous systems, and data-based surveillance technology as representative areas where technological innovation challenges the existing legal system[10]. This study is in line with the problem consciousness of this study in that it points out that the imbalance in the speed of technology acceptance and legal adaptability can amplify social risks in the long run.

Meanwhile, Song Ki-bok (2020) analyzed Germany's artificial intelligence strategy and the EU's technology regulatory system, emphasizing the need to establish an ethics- and responsibility-based legal framework without suppressing technological innovation. It was also suggested that the direction of technology control should be changed to proactive governance rather than ex post punishment[11].

This approach also provides implications for technology acceptance policies in the field of security and security. As such, domestic and foreign studies agree that technological innovation requires the redesign of laws and systems, but discussions on the legalization of AI and IoT technologies centered on the private security and security industries are still in the early stages. Most of the existing research focuses on public security or administrative regulation-oriented perspectives, and detailed discussions such as setting responsible subjects in the private sector, technology certification systems, and personal information processing standards have not been sufficiently accumulated. Therefore, this study attempts to supplement the following limitations of existing studies. First, it overcame the point that existing studies dealt with technology development trends and legal discussions separately and applied the perspective of technology and legal system convergence analysis. Second, beyond discussing the security system centered on public safety, the connection between technology acceptance and legal maintenance centered on private security and security was analyzed. Third, the direction of legalization of AI-IoT-based private security services was empirically presented through comparison of domestic and international legislation and policies. This study differs from previous studies in that it is an attempt to resolve the disconnection between existing technologies and laws and comprehensively examine the triangular structure of technology acceptance, risk control, and legal reform. This approach will serve as a basic foundation for analyzing the operating system of AI-IoT-based security and security services and discussing legal improvement.

3. Analysis of Technology and Operating Systems of AI and IoT-based Security and Security Services

3.1. Current status of technology convergence in civilian security and security area

The spread of AI and IoT technologies is fundamentally changing the paradigm of the private security and security industry. It is evolving toward establishing a data-based risk prediction and real-time response system, going beyond simple manpower-centered patrol and monitoring methods in the past. In particular, artificial intelligence video analysis, drone monitoring, IoT sensor network, autonomous CCTV, and wearable equipment have become the core technology pillars of private security services.

First, the AI image analysis system has developed to a level that supports abnormal behavior detection, crowd density analysis, and entry control automation beyond the existing simple video recording function. Park Sang-wook et al. (2020) reported that intelligent CCTV-based dynamic crime prediction technology can predict real-time risk by comprehensively analyzing human behavior, attributes, and environmental factors[4][12]. This technology is expanded and applied not only to public institutions but also to private facility security systems, replacing the existing manpower-centered monitoring system.

Second, IoT-based surveillance networks combine sensors and communication networks to detect real-time conditions of buildings, factories, and residential areas, and immediately notify the control system when abnormalities occur. Choi Woo-chul and Na Jun-yeop (2018) analyzed that an integrated surveillance system linking various sensors and artificial intelligence algorithms is the basis for crime prevention and rapid response in a study on an intelligent crime prevention integrated platform for real-time crime response[6][13].

Third, the AI autonomous driving CCTV system is a representative technology that overcomes the limitations of fixed surveillance. Kwak Yeon-gyu et al. (2023) suggested that artificial intelligence-based autonomous CCTV can perform dynamic patrols around crime areas by combining image analysis and mobile equipment[5][14]. This provides practical efficiency to private security sites in that it reduces surveillance blind spots and enables efficient manpower management compared to existing fixed CCTV.

Fourth, drone and robot-based patrol systems are also rapidly being commercialized. Drones are in charge of monitoring areas that are difficult for human resources to access, such as outer walls of high-rise buildings and night areas, and perform the function of automatically detecting and tracking dangerous signs through AI-based flight control technology. Young-woo Yang and Joo-rak Lee (2018) confirmed that drones are being used for various private security missions such as patrol, fire monitoring, escort of valuables, and personal protection[15], etc. In particular, the case of patrol drone operation of private security services of large domestic companies, such as SK Shields' ADT Caps and S1's Secom, shows that drones are performing practical security assistance functions in the physical security area.

Fifth, wearable and mobile-based security technology is emerging as a key factor that simultaneously increases the safety and response efficiency of field workers. A system that can detect dangerous situations early by collecting and analyzing biometric signals in real time and immediately request rescue or share situations through IoT networks is being introduced. Kim Dae-hyun and Kim Dong-hoon (2022) proposed an ICBM (IoT-Cloud-Big data-Mobile) model that integrates sensor-based biometric information and location/communication data, and transmits the biometric information, surrounding environment[16], and equipment status of field personnel in real time to the integrated control center.

As a result, this technology, which combines wearable devices and mobile networks, goes beyond simple convenience and serves as a bridgehead to transform the traditional manpower-centered security system into a data-based intelligent safety management system.

As such, the private security and security industry is transitioning to an advanced intelligent integrated security system through technology convergence, but the institutional foundation is still insufficient compared to the speed of technological development. There are complex challenges such as error responsibility in artificial intelligence decision-making, reliability in data management, and personal information protection, and these factors serve as the starting point for deriving operational model structures and legal tasks to be discussed in the next section.

3.2. Operation model structure of AI/IoT convergence security and security service

The core of AI and IoT-based security and security services is the connection of multi-layered surveillance systems and real-time data. Technically, it is divided into three stages of structure: detection, analysis, and response, and each stage is organically connected through an artificial intelligence algorithm and an IoT network.

First, in the detection stage, various data from the site are collected through IoT sensors, CCTV, drones, and wearable devices. Physical signals such as image, sound, temperature, and vibration collected by the sensor are transmitted to a cloud-based integrated control system.

Second, in the analysis stage, artificial intelligence automatically analyzes the input data to identify abnormal behavior, intrusion, and danger signs. Machine learning-based pattern recognition algorithms increase detection accuracy through repetitive learning and set response priorities through risk rating.

Third, in the response stage, the analysis result is immediately transmitted to the control center and on-site personnel to perform real-time alerts and actions. For example, a drone or patrol robot is automatically dispatched, or a worker wearing a wearable device is notified of the danger with vibration and voice signals.

This structure works as an integrated platform-type convergence model rather than a single technology. The sensor network, cloud, AI analysis engine, and mobile control system are interconnected to form an intelligent integrated security ecosystem. In particular, on-site personnel are converted to data-informed operators rather than simple monitors, and technology plays a role of assisting and expanding personnel rather than replacing them.

In the end, the AI-IoT convergence operation model is evolving into a proactive security system. It further improves the security level of the public and private sectors by implementing a proactive response based on real-time data analysis and risk prediction, breaking away from the existing post-response-oriented structure.

3.3. Derivation of legal and institutional challenges following technology introduction

With the introduction of AI and IoT technologies to private security and security services, the manpower-centered supervision and responsibility structure that the existing Security Business Act system had premised on is raising a new technology-centered complex legal task. Beyond simple technical issues, these changes can be summarized into four pillars: resetting legal definitions, clarifying responsible subjects, reorganizing personal information protection systems, and reorganizing technology certification and management systems.

(1) Uncertainty of legal definition and scope of application

Article 2 of the current Security Business Act restricts the scope of the security industry to 'facility expenses, convoy expenses, personal protection, machine expenses, and special expenses'. However, technology-based services such as AI image analysis, drone patrol, and wearable sensors are not included in the existing classification system.

For example, facility monitoring using drones or IoT sensor-based risk detection is not legally recognized as a security act, even though it is a non-human surveillance activity unlike physical patrol activities. As a result, the legal basis for authorization and supervision is unclear, and the market is being formed without the legal responsibility and authority of the service provider being specified.

Therefore, in the future, it is necessary to clearly define the scope of application under the law by newly establishing definitions for 'technology convergence security business' or 'intelligent security service'.

(2) Unclear who is responsible for artificial intelligence decision making

The AI-based security and security system has a structure in which artificial intelligence automatically determines the risk or issues an alarm. In the event of human or property damage caused by an error or malfunction in this process, there is a legal gap as to who will be considered responsible. For example, if an autonomous CCTV or drone recognizes the wrong object as a threat and issues an alarm, or if the response is delayed due to an error in the AI analysis result, the current legislation does not clearly stipulate who should be responsible among security companies, system developers, or operation control personnel. This problem means that a technical and legal device must be prepared to secure the transparency and explainability of artificial intelligence's decision-making. In the end, the future system needs to introduce a preliminary regulatory framework for security companies using artificial intelligence, such as the obligation to record operating logs, algorithm verification procedures, and the obligation to report in case of malfunction.

(3) Tensions and harmonization between privacy and technology utilization

AI and IoT-based security services operate by collecting and analyzing large-scale video, voice, and location information. However, the use of such data creates a tension with the regulatory scope of the Personal Information Protection Act and the Communication Secret Protection Act. In particular, controversy over use or excessive collection of purposes may arise in the process of artificial intelligence automatically identifying face and behavioral patterns or IoT sensors collecting user biometric and location information.

Accordingly, it is necessary to specify the de-identification and pseudonym information processing criteria specialized for the private security system. For example, Song In-jun and Kim Cha-jong (2024) proposed an AI-based de-identification technique that automatically detects the personal information area of image data with artificial intelligence and combines masking and encryption processing[17]. This study is a representative example showing that a balance between personal information protection and data utilization can be achieved through technical means.

However, if these technologies are applied commercially without institutional certification or legal verification procedures, there is still a risk of misidentification or backtracking. Therefore, in order to effectively operate technical protection measures, it is essential to institutionalize the authentication of the AI image processing system and to legally standardize the de-identification algorithm.

(4) Absence of technical certification and standardization

Currently, there is no integrated certification and evaluation system for private security technology in Korea. Security devices or systems must be individually certified by various organizations such as the Ministry of Trade, Industry and Energy, the National Police Agency, and the Ministry of Science and ICT, and there are no separate standards for AI analysis engines or IoT security protocols. As a result, interoperability between technologies is reduced, and it is difficult to secure public trust in system stability.

Therefore, an integrated standard for evaluating technical safety, data processing adequacy, and algorithm reliability should be established by establishing an AI-IoT convergence security system certification system. This will serve as the basis for simultaneously strengthening the technological competitiveness and public trust of private security companies in the future.

(5) Diversification of supervisory systems and lack of governance

Currently, the main ministry for the security industry is the National Police Agency, but AI and IoT technologies span the jurisdiction of multiple ministries such as information and communication, industry, and national security. As a result, technology-based security services are facing a dual problem of supervisory gaps or duplicate regulations. For example, data security of IoT sensors is under the jurisdiction of the Ministry of Science and ICT, but private security services using them are subject to approval by the National Police Agency. If the boundaries between regulatory agencies are ambiguous, the introduction of technology is delayed and the legal risk of the company increases.

Therefore, in the future, a private security technology convergence policy consultative body should be formed jointly by relevant ministries such as the National Police Agency, Ministry of Science and ICT, and the Ministry of Industry to establish an integrated governance system that unifies licensing, technology verification, and data management standards.

4. Measures to Improve Legal System

4.1. Review of the current legislation (security business act, personal information protection act, etc.)

Although the current "Security Business Act" has undergone system change for more than 40 years since its enactment in 1976, it has been designed on the premise of a traditional manpower-centered security system and has been formed around a human surveillance system before technological development. However, the recent rapid spread of technology-oriented security services such as AI image analysis, drone surveillance, and IoT-based sensor networks has clearly revealed limitations that the current legal system cannot cover this. Despite the rapid introduction of the technology of the 4th Industrial Revolution into the private security sector, the "Security Business Act" still maintains a people-centered permit and supervision system, and does not specify the legal basis for smart security or intelligent security technology. In particular, the definition of machine security as stipulated in Article 2 of the same Act remains at the level of simple monitoring equipment, so it can be seen that artificial intelligence analysis, cloud-based control, and data-linked security systems cannot be legally included in the security industry category. As a result, this legal void makes the institutional status of technology-based private security services and the subject of supervision unclear.

Jin Kyung-ae (2022) compared the structure of the private security system between Korea and Japan, and analyzed that in Japan, since the enactment of the 警備 Business Act, the National Police Agency has been responding flexibly to the development of the technology-based security industry while exercising its guidance and supervision authority[18]. In fact, Japan oversees security business licenses by the security business department under the National Police Agency, and systematically guarantees the technology acceptance of the private security industry by establishing equipment and technology certification standards as well as security guard qualification management. On the other hand, in Korea, there is an authorization system centered on the National Police Agency, but technology certification or equipment verification is distributed to multiple agencies such as the Ministry of Trade, Industry and Energy, the Ministry of Science and Technology Information and Communication, and the National Police Agency, so there is no integrated management system. As a result, the AI-IoT convergence security system is treated as an informal service without being clearly included anywhere in the current laws and regulations. As such, the structural problem of the current security industry law does not reflect the development of industrial technology and adheres to a manpower-centered permit system, and technological innovation is neglected in institutional uncertainty. In particular, there are no authentication and evaluation criteria for detection sensors and AI analysis algorithms, which are technical elements included in security services, so even the same security technology may appear as a problem in which different standards are applied to each institution.

At the same time, the legislation related to personal information protection is also in conflict with the spread of AI and IoT security services. Article 3 of the Personal Information Protection Act stipulates the minimum collection principle and the prohibition of use outside of its purpose, but the AI security system that operates based on video, audio, and location information inevitably collects and analyzes a large amount of personal data. This structural tension can be seen as a fundamental task in the era of smart security. In other words, despite the fact that technological innovation is putting pressure on the limitations of personal information regulation, the current legislation does not provide a mechanism for coordination between technology acceptance and protection regulations.

In particular, Japan can perform technology certification and personal information protection screening in an integrated manner through the intelligent monitoring system technology verification system led by the National Police Agency. Beyond simply evaluating the performance of technology, it has a structure that verifies the appropriateness of the use of personal information for public safety purposes, and Korea also needs to prepare an integrated legal system to balance technology reliability and data protection by referring to these systems.

In addition, a number of laws such as the Information and Communication Network Act, the Personal Information Protection Act, and the Communication Secret Protection Act are partially related in Korea, but there are no explicit provisions that directly deal with private security technologies. As a result, when the AI-based video analysis system automatically identifies a specific person's face, behavior, or voice, a problem arises that the legal nature of the data is not even clearly defined whether it is personal or non-identifiable information. This means that the two values of promoting technological innovation and protecting personal information are legally in conflict.

These institutional limitations are not just insufficient laws, but lead to a decrease in industrial competitiveness. Despite the distributed technology certification, private security companies must go through a number of procedures, such as technical safety evaluation by the Ministry of Industry and certification of communication devices by the Ministry of Science and ICT, in addition to the National Police Agency approval. This complex structure hinders the speed of innovation of companies and fixes the absence of practical technology standardization or quality verification systems.

In short, the current Security Business Act and personal information-related laws have a structure that is unsuitable for systematically accepting technology convergence private security services. This can be summarized into three limitations: narrow legal definition, decentralization of technology certification, and rigidity of personal information regulation. Therefore, for the institutionalization of AI-IoT-based security and security services, it is urgent to expand the scope of application of the current legislation and to overhaul the legal and institutional system that integrates technology certification and data management standards. This will be the minimum legal basis for the security industry in the era of the 4th industrial revolution to operate as part of the public safety system, not just private services.

4.2. Comparison of overseas legislative cases (Japan, Gemany, EU, etc.)

As AI and IoT technologies expand to the security and security areas, each country faces the challenge of how to incorporate the flow of technological innovation into the existing legal system. In these changes, Japan, Germany, and the European Union (EU) have clarified the legal status of technology-based private security in common and have promoted harmony with public security systems, despite having different legal traditions and social structures. All three regions are important comparisons in that they do not see technological advances as merely industrial innovation but as a normative area to be managed systematically in the balance of public safety and human rights protection.

First of all, Japan is evaluated as one of the earliest countries to realize the institutionalization of private expenses, focusing on the "警備 Business Act" enacted in 1972[19]. The law stipulates that the National Police Agency exercises the authority to authorize, guide, and supervise the security business in a unified manner, and classifies the types of security business into facility expenses, convoy expenses, personal protection, and machine expenses. In particular, Japan has continuously expanded the concept of "machine expenses" according to technological changes, and since the 2000s, it has developed interpretation in the form of including AI image analysis, IoT sensor network, and remote monitoring system. Through the security equipment registration system, the Japanese 警備 Agency integrates and manages technical certification, data management, and personal information protection standards of equipment through the 指導要領 industry guidance method. This unified supervisory system functions as an institutional basis for achieving both public safety and industrial innovation.

Germany regulates the private security industry around Article 34a of the "Gewerbeordnung" Act[20], and all security industry workers must pass the qualification test conducted by the Chamber of Commerce. In particular, the German system is characterized by a combination of professional ethics and technical standards. The national standard DIN 77200 stipulates the quality control of security services and human and technical requirements at the same time, and stipulates technical reliability including the safety of artificial intelligence modules and information and communication systems. In addition, the Federal Information Security Administration evaluates and certifies the security suitability of ICT and IoT devices, and the results are directly reflected in security business permits and service evaluation. In other words, the German-style structure in which technology certification and qualification certification are parallel can be said to be an example of institutionalizing the coexistence of human reliability and technology reliability.

The EU has established a transnational normative system that is one step higher than the system of each country. The General Data Protection Regulation (GDPR)[21], which took effect in 2018, set clear legal standards for the processing of all personal information, including sensitive data such as video and biometric information, and Article 35 in particular established a proactive control device by mandating the Data Protection Impact Assessment (DPIA) when establishing surveillance and security systems. In addition, the "EU Artificial Intelligence Act," adopted in 2024, classifies AI systems for surveillance and security purposes as high-risk groups, and specifies specific legal requirements such as risk management, data quality, transparency, and human intervention[22]. It is evaluated as the world's first normative system that systematically allows the use of AI technology while maintaining a balance between public safety and human rights protection.

Although the approaches of these three regions are different, they share a clear direction for embracing technological innovation into the legislation. In Japan, technology certification and personal information protection are combined through a single supervisory system centered on police administration, and in Germany, technology standards and work ethics are integrated to institutionalize reliability within the industry. The EU has established a risk-based management system at the transnational level by combining the Personal Information Protection Act and artificial intelligence regulations. All three models provide practical implications for the improvement of Korean legislation in that they share the normative premise that technological advances should work in a way that complements public safety.

The reason for this study's comparative analysis is clear, focusing on Japan, Germany, and the EU among overseas cases. Japan, like Korea, operates an authorization and supervision system centered on the National Police Agency while providing an institutional precedent that flexibly responds to technological convergence. Germany sets an example of best practices in regulating technology and work ethics together through national standards even in a decentralized system at the federal level. The EU presents the future direction of technology regulation by combining personal information protection and artificial intelligence regulation at the level of transnational norms.

On the other hand, although the size of the security industry in the United States and the United Kingdom is large, self-regulation centered on states and private organizations is mainstream, which lacks legal consistency and integration. In particular, the United States operates a standardized model centered on associations such as the Security Industry Qualification Committee (PSIA), but does not legally link public safety or personal information protection. The UK's Security Industry Authority (SIA) also manages the licensing system, but there are no direct regulations on artificial intelligence or data-based technologies. Therefore, in countries with a centralized supervisory structure such as Korea, the Japanese, German, and EU models have high validity as realistic benchmarks.

In the end, these overseas cases suggest the direction of improving Korea's legislation. In order not to put technological innovation in a legal blind spot, it is necessary to maintain a approval system centered on the National Police Agency, but introduce a procedural system that integrates technology certification and data protection review. At the same time, the cooperative governance structure between the state and the private sector should be institutionalized by redefining private expenses as a complementary function of public safety, not limited to simple industries. Japan's technology verification system, Germany's technology standard norms, and the EU's risk-based management system will be strong policy models to materialize this.

In short, the laws of Japan, Germany, and the EU all have the same thing as a system that combines technology acceptance and social trust. Korea also needs to establish new security governance that harmonizes publicity and technological innovation by integrating the security industry law and the personal information protection system and establishing technology certification systems and data management standards for AI and IoT-based security services.

4.3. Korean AI and IoT security and security service legislation model presented

Technology convergence private security has become an essential component of the security ecosystem, and Korean legislation requires a change in the system design itself beyond supplementing individual technology or equipment certification. This section structures design principles commonly derived from different models such as Japan's single supervisory system, Germany's parallel qualifications and standards, and the EU's risk-based data and AI regulation in a form suitable for Korean legislation. The key is reorganization of legal definition, integration of permission and supervision and technology certification, systematization of preliminary screening of data and AI governance, securing standardization and interoperability, and regularization of ministries and joint governance.

First of all, the starting point is the establishment of a definition rule that expands the application of the "Security Business Act". The current classification of types in Article 2 is designed on the premise of classical actions centered on manpower and equipment, and does not sufficiently cover the legal status of services including remote and intelligent analysis and autonomous response. Accordingly, the technology convergence security business or intelligent security service is defined as a separate definition clause, and functional definitions based on the functional chain of detection, analysis, and response are introduced. This definition makes it considered a security act even if non-human surveillance means such as drones, robots, and wearables are used, and platform-type services, including AI analysis, cloud control, and sensor network connection, are explicitly included in the security industry category. The precedent that Japan has extensively interpreted and operated machine expenses within the security industry law system in accordance with the changing times shows both the legitimacy and operational potential of the expansion of the definition rule.

Second, establish a one-stop permit and verification system that internalizes technical certification in permit and supervision procedures. In the current system, separate from the approval of the National Police Agency, individual technology and communication certification must be obtained from the Ministry of Industry and Ministry of Science and ICT, disconnecting the interoperability and responsibility link. As the German-style qualification and standard parallel structure suggests, the security business license requirements include a graded technical safety review (e.g., L1 to L3), and announce the service quality, organizational requirements, and information security and safety standards of the DIN 77200 class linked to international and national standards. At this time, the technical review results are linked as essential components of permission, so that market entry is possible only when technical requirements such as human requirements and standard suitability are met. When upgrading a service change, for example, an algorithm or sensor, a simplified change review ensures traceability of updates.

Third, proactive risk management procedures for data and AI use are stipulated in the law. It is clear that what can be learned from the EU GDPR's Data Protection Impact Assessment (DPIA) and the AI Act's high-risk group requirements are combined. Security services that process fixed, mobile video, audio, and location data on a large scale perform mandatory DPIA, and report and approve them to the supervisory authority based on target sensitivity, processing scale, purpose, preservation, and provision to a third party. In addition, the obligation to preserve AI operation records (logs, etc.), explainability requirements (alarm and decision-based data, model version, reliability indicator record), and dataset management regulations (learning, verification, and drift monitoring) are imposed as permission conditions. Real-time identification and tracking for monitoring and security purposes apply equal control of high-risk groups, and specify de-identification and pseudonymization obligations at the learning and inference stage, but advance measures for quantitatively evaluating resilience and re-identification risks, such as anonymity and upper limits on re-identification probability, into operating guidelines. This is not just a compliance burden, but functions as a device that provides a basis for evidence of responsibility in case of dispute or accident.

Fourth, a national interoperability and standard frame that integrates fragmented authentication is established. In a reality where cameras, sensors, drones, control SW, and AI engines from various manufacturers are mixed, the lack of interoperability deteriorates safety and cost at the same time. Accordingly, profile standards such as APIs, formats, security protocols, timestamps, and audit tracking models are defined for the network, equipment, and application layers for security purposes, and an authorized test and certification center is established to perform suitability tests. Just as Germany's DIN 77200 stipulates service quality and organizational requirements, we also certify service quality norms (manpower allocation, training, dispatch time, customer report, post-evaluation) and technology security norms (encryption, key management, vulnerability management, patch policy) simultaneously on a double axis. In order to reflect the rapid evolution of technology, a rapid revision mechanism at the notification and notification level is introduced, and the mapping table with international standards (ISO/IEC 27001, 31700, etc.) is regularly updated.

Fifth, institutionalize joint permanent governance of ministries. Security business licensing is divided into the National Police Agency, communication and information security is divided into the Ministry of Science and ICT, and industrial standards and equipment safety is divided into the Ministry of Industry. By regularizing this as a legal consultative body, licensing review, technology certification, and data protection review are integrated into one procedure flow. The consultative body has the right to jointly decide on the designation of high-risk services, revision of standards, accident investigation, recall, and correction orders, and discloses quarterly performance indicators (permission processing period, accident and false alarm rate, DPIA approval rate, education completion rate, standard compliance maintenance rate, etc.). It is a Korean-style operating model that procedurally combines Japan's single supervision and EU pre-evaluation.

Sixth, it is necessary to specifically stipulate a step-by-step institutionalization roadmap. The first stage is the infrastructure maintenance stage to be carried out during the first year of the introduction of the system, and a transitional system should be established to establish a temporary certification system to temporarily operate the technology convergence security business. In this period, it is necessary to establish an infrastructure for the implementation of the system by notifying model standards in parallel with the revision of laws and regulations related to the security industry and designating a specialized certification center to be in charge of technology verification, testing, and evaluation.

The second stage is a system settlement period that takes about two to three years, and aims to establish an integrated screening system by completely linking the existing security business licensing process with the technology certification process. In addition, the obligation to record data protection impact assessment (DPIA) and artificial intelligence logs should be legislated to ensure systematic transparency in AI decisions. At this stage, it will be possible to ensure interoperability and legal consistency between technologies at the same time by requiring standardized data profiles to be applied to all security systems. Stage 3 corresponds to a mid- to long-term generator of 3 to 5 years, and the key task is to switch to a performance-oriented regulatory system. In other words, it is necessary to clearly establish the range of allowable false alarm rates and detection sensitivity, and develop into a result-based management system based on technical performance and safety indicators.

4.4. Policy support and industrial implications

The rapid development of AI and IoT technologies has brought about structural changes in the private security industry, and legal and institutional maintenance in response is emerging as a key factor in national competitiveness. When technology trends and overseas legislative cases are combined, Korea's policy direction in the future is organized into three pillars: securing consistency of legislation, strengthening innovation capabilities in industrial ecosystems, and maintaining a balance between publicity and marketability.

First, it is necessary to secure legal consistency and integrate the technology acceptance system. Currently, laws related to the security industry in Korea are not keeping up with the pace of technological development, and supervisory authority and certification system are distributed by department. Japan's police agency-centered single supervisory model, Germany's standardized certification system, and the EU's risk-based regulatory system are all increasing industrial reliability through institutional consistency. Korea should also establish an integrated management system in which technology, manpower, and data are connected through integrated adjustment of related laws such as the Security Business Act, the Personal Information Protection Act, and the Information and Communication Network Act.

Second, it is necessary to institutionalize technology certification and public data governance. Since AI and IoT-based security systems use public and private data at the same time, integrated standards must be established in three aspects: data quality, security, and utilization. To this end, the government should introduce a pre-verification management model that combines the technology certification system and data protection review, and provide a public test bed and standard dataset that can be used jointly by SMEs. This will contribute to improving the quality and securing reliability of the industry as a whole.

Third, it is required to expand policy investment as a foundation for industrial innovation. The private security industry has a high technology intensity, but the proportion of SMEs is large, so it lacks its own R&D capabilities. The government should implement practical industrial support policies such as support for the R&D tax system, equipment conversion assistance, and professional manpower training, and it is necessary to designate AI and IoT convergence security technology as a priority area for national R&D projects. This policy foundation will ease the technology gap across the industry and lead to securing international standard competitiveness.

Fourth, securing social trust and strengthening ethical control must go hand in hand. AI-based monitoring systems increase efficiency, but at the same time, they can lead to privacy invasion and over-monitoring controversy. Therefore, social consensus on data processing transparency, algorithmic fairness, and error response procedures as well as technical performance is essential. The government should systematically guarantee technology reliability by strengthening the transparency reporting obligation for monitoring technology and the data protection impact assessment system, and clarifying user rights relief procedures.

Fifth, it is important to secure the sustainability of the industrial structure. The faster the technological development progresses, the more likely the gap between large security companies and SMEs will widen. To mitigate this, the government should promote balanced development within the industry by operating technology consortiums centered on SMEs, joint certification platforms, and joint equipment utilization systems. In addition, by supporting new job creation models using AI and IoT technologies across the industrial ecosystem, an institutional safety net should be established to prevent technological changes from leading to employment insecurity.

Finally, it is necessary to redefine the private security industry as part of the national safety strategy. AI and IoT-based private security is not just a private enterprise service, but functions as a complementary axis of public safety. The government should recognize this as a public cooperation partner linked to security policy and prepare an integrated operation plan with the national disaster and crisis response system. This policy restructuring will systematically establish the public nature of private security, so technological innovation will soon lead to the strengthening of the social safety net.

In short, the development of the AI-IoT-based private security system should be promoted in a way that balances the four pillars of law, technology, industry, and ethics. This will go beyond simply embracing new technologies and lead to structural innovation in national security governance, and will become a realistic roadmap and policy task for institutionalization of the Korean-style convergence security system.

5. Conclusions and Future Tasks

This study comprehensively analyzed the structural changes caused by the spread of AI and IoT technology in private security and security services at the level of technology, operation, and legislation, and suggested the principles and implementation paths for the design of the Korean-style system. As a result of the analysis, it was confirmed that the domestic system remained in the traditional manpower-centered permit and supervision paradigm and revealed an institutional gap in the legal status, responsibility structure, data governance, and technology certification of technology convergence security services. Accordingly, this study proposed a Korean-style legalization model focusing on the five pillars of reorganization of definition, integration of permission-technical certification, pre-prioritization of risk-based data and AI governance, institutional fixation of interoperability standards, and establishment of permanent governance of ministries. Furthermore, by specifying the temporal path of introduction, settlement, and advancement through a step-by-step roadmap, an execution frame was presented to minimize the gap between normative design and policy execution.

The policy implications are summarized as follows. First, since private expenses are no longer limited to private services and function as a complementary axis of national safety, the security industry-related legislation should be reorganized into a convergence legislation that mediates public safety policies and data and AI norms. Second, the possibility of evidence of performance and responsibility is at the heart of market trust, and for this purpose, the system should internalize the license linkage of technical review results, the mandatory DPIA and AI operation logs, and the application of standardized data profiles. Third, since cost efficiency and safety are compatible only when interoperability between technologies and traceability of updates are guaranteed, an interworking system of profile standards, suitability evaluation, and post-supervision should be established around an accredited test and certification center. Fourth, a transition policy package that combines public data and test beds, equipment conversion, and manpower training support should be implemented in parallel to alleviate the imbalance in the industrial ecosystem.

In terms of academic contribution, this study is meaningful in that it integrated technology acceptance, risk control, and legal reform in the private security area, which were relatively neglected by previous studies centered on public security, into a single frame, and derived Korean-style design principles through comparison of overseas systems. In particular, a baseline was presented to ensure structural consistency in regulatory design by consistently matching the technical functional chains of detection, analysis, and response with the normative functional chains of permission, authentication, supervision, and responsibility.

Nevertheless, limitations exist. First, large-scale empirical data on field operation data and performance indicators such as false alarm rate, detection sensitivity, and dispatch time are not sufficiently reflected, so there is a limitation in presenting quantitative thresholds for system design. Second, an in-depth analysis is required to match the technical solution related to the algorithm's bias and drift problem with the system. Third, notification, objections, corrective orders, and operational details of damage compensation, which are relief procedures in situations of conflict of basic rights such as personal information and communication secrets, should be designed in more detail through follow-up studies.

Accordingly, the future tasks are as follows. First, by designing on-site pilots for representative service types such as facility expenses, large complex housing expenses, distribution centers, and smart factories, performance, safety, and rights impact indicators should be collected, and the allowable false alarm rate and detection sensitivity range should be empirically calculated as threshold values of the step-by-step roadmap. Second, through the estimation of conversion costs for small and medium-sized security companies and the analysis of the costs and effects of tax, subsidiary, and common infrastructure as policy tools, a policy combination that minimizes the social marginal cost of compliance should be derived. Third, it is necessary to prepare a plan to secure vertical consistency between airspace and space regulation (flight, autonomous driving, propagation, safety) of boundary technologies such as drones, robots, and wearable and security business permits. Fourth, in order to improve citizen acceptance, the possibility and effectiveness of the transparency report and notification system (display monitoring areas, notification of data processing, and disclosure of summary impact assessment) should be evaluated.

In conclusion, the institutionalization of AI and IoT-based private security and security services is not completed only by law revision. Technological innovation and public trust rise together when the expansion of definition, permission, authentication, and supervision, data and AI preliminary impact assessment, interoperability standards and tests and certification, ministry joint permanent governance, and industrial transformation support operate simultaneously and complementarily.

The structure and roadmap presented by this study will serve as the foundation for this comprehensive design, and as subsequent empirical and evaluation studies and policy experiments accumulate, Korean-style convergence security governance will be able to secure institutional stability and international competitiveness at the same time.

6. References

6.1. Journal articles

[2] Bang JS & Park WJ & Yoon SY & Shin JH & Lee YT. Trends of Intelligent Public Safety Service Technologies. *Electronics and Telecommunications Research Institute*, 34(1), 111-122 (2019).

[3] Lee WS. A Study on Legal Issues of Crime Prevention with Advanced Science Technology. *Korean Criminological Review,* 27(2), 231-262 (2016).

[4] Park SW & Oh SH & Park SW & Lim KS & Choi BS & Park SH & Kim SW & Han SW & Han JW & Kim GW. Trends in Dynamic Crime Prediction Technologies based on Intelligents CCTV . *Electronics and Telecommunications Trends*, 35(2), 17-27 (2020).

[6] Choi WC & Na JY. A Study on the Establishment and Connection of Intelligent Security Integrated Platform Elements for Real-time Crime Response. *Journal of Korea Academia-industrial Cooperation Society*, 19(10), 8-15 (2018).

[8] Shin HJ & Kim JC. A Study on the Government Regulation and Self Regulation about Private Security Industry. *Journal of Korean Public Police and Security Studies*, 12(2), 207-224 (2015).

[10] Song KB. Discussion on the Advent of the Artificial Intelligence(AI) Era and the Direction of the Legal System -Focusing on Germany's Artificial Intelligence policy- . *Journal of Police & Law*, 18(2), 177-203 (2020).

[11] Yang YW & Lee JR. Utilization of Drone Technology in Physical Security and Its Limitations. *The Journal of Police Policies*, 32(3), 255-284 (2018).

[12] Kim DH & Kim DH. A Study on the Development Direction for Physical Security Field using Wearable Devices. *Korean Journal of Industrial Security*, 12(1), 187-207 (2022).

[13] Song IJ & Kim CJ. Research on Artificial Intelligence Based De-identification Technique of Personal Information Area at Video Data. *IEMEK Journal of Embedded Systems and Applications*, 19(1), 19-25 (2024).

[16] Kim DH & Kim DH. A Study on the Development Direction for Physical Security Field using Wearable Devices. *Korean Journal of Industrial Security*, 12(1), 287-207 (2022).

[19] Lee HJ. A Study on Revision of Security Industry Act to Increase Efficiency of Cooperative Policing between Private and Public Police. *The Police Science Journal*, 8(1), 203-224 (2013).

[22] Lee HJ. Consideration on the Direction of AI Discipline in Korea AI Act -Focusing on the Comparison of the EU AI Act Regulation-. *CAUJLS*, 47(2), 5-42 (2023).

6.2. Thesis degree

[14] Jin KA. Korea-Japan Comparative Study on the Private Security System and Operation. Kyonggi University, Doctoral Thesis (2022).

6.3. Additional references

[1] Seok WH. Analyzing the Impact of ICT Service and Infrastructure Environment on Crime Prevention. Electronics and Telecommunications Research Institute, ETRI Insight Report (2018).

[5] Kwak YG & Kim YJ & Woo JW & Jeong DG & Yoo SO. AI Self-driving CCTV System for Smartening Crime Prevention Facilities. Proceedings of the Korean Information Processing Society Conference (2023).

[7] Lee TH & Park JS. Systematic Improvement for the Efficient Operation of the Private Security Field. Proceedings of the Korean Society of Disaster Information Conference (2022).

[9] Kwon EJ & Lee SJ & Oh DS & Yang CS & Yoon HS. A Study on the Transformation of the Legal System for Risk Control and Technology Acceptance in the Age of Intelligent Revolution. General of Cooperative Research at the Economic and Humanities and Social Research Society. Korea Information Society Development Institute (2020).

[15] Japanese National Police Agency. Act No.117 of 1972 (2023).

[17] European Union. General Data Protection Regulation (Regulation (EU) 2016/679) (2016).

[18] European Union. Artificial Intelligence Act (Regulation (EU) 2024/1689) (2024).

[20] Nogala D & Sack F. Private Reconfigurations of Police and Policing: The Case of Germany. GERN Seminar (1998).

[21] European Parliament. How GDPR Changes the Rules for Scientific Research. Parliamentary Research Service Study (2019).

7. Appendix

7.1. Author's contribution

	Initial name	Contribution
Author	DL	-Set of concepts ☑ -Design ☑ -Getting results ☑ -Analysis ☑ -Make a significant contribution to collection ☑ -Final approval of the paper ☑ -Corresponding ☑ -Play a decisive role in modification ☑ -Significant contributions to concepts, designs, practices, analysis and interpretation of data ☑ -Participants in Drafting and Revising Papers ☑ -Someone who can explain all aspects of the paper ☑

2025 Vol.10 No.0, 42-52

Received: 2025/09/02, Peer-reviewed: 2025/10/10, Accepted: 2025/10/27, Published: 2025/12/30

Robotics & AI Ethics

Publisher: J-INSTITUTE
ISSN: 2435-3345

Website: j-institute.org

Corresponding author*
E-mail: front@bau.ac.kr

DOI Address:
dx.doi.org/10.22471/ai.2025.10.0.42

Artificial Intelligence in the Hospitality Industry: A Review of Research Trends on Customer Experience, Operational Efficiency, and Ethical Issues

Sungwoo Sim[1]

Baekseok Arts University, Professor, Republic of Korea

Kiho Lee[2*]

Baekseok Arts University, Professor, Republic of Korea

Abstract

Purpose: *Artificial intelligence (AI) has become a transformative force in the hospitality industry, reshaping service delivery, operational management, and ethical governance. As AI-based technologies, such as chatbots, service robots, and algorithm-driven decision-support systems, are increasingly adopted, hospitality organizations face both opportunities for efficiency and personalization, as well as challenges related to trust, labor, and ethical responsibility. Despite a rapidly growing body of literature, existing studies remain fragmented, often focusing on isolated applications or outcomes.*

Method: *The purpose of this study is to systematically review recent research on the utilization of artificial intelligence (AI) in the hospitality industry and to analyze it across three key dimensions: customer experience, operational and managerial efficiency, and ethical and social issues. A structured literature review approach was employed to synthesize recent academic studies and identify major research themes and future research directions. Portions of this manuscript were developed with the assistance of generative artificial intelligence; however, all content was critically reviewed and finalized by the authors to ensure academic rigor and integrity.*

Results: *The review reveals three dominant research streams in hospitality AI studies. First, AI-based services enhance service accessibility, responsiveness, and personalization, positively affecting customer experience. Second, AI contributes to operational efficiency through demand forecasting, pricing, and decision support. However, these benefits are accompanied by workforce-related challenges. Third, ethical issues such as privacy protection and algorithmic transparency have gained increasing attention.*

Conclusion: *This study argues that the sustainable adoption of AI in hospitality depends on balancing technological efficiency with human-centered service values and ethical accountability. By providing an integrative overview of existing research, this review contributes to a more comprehensive understanding of AI-driven transformation in the hospitality industry and offers directions for future research and the responsible implementation of AI.*

Keywords: ***Artificial Intelligence, Hospitality Industry, Customer Experience, Operational Efficiency, AI Ethics***

1. Introduction

1.1. Digital transformation and the rise of artificial intelligence in hospitality

Rapid digital transformation has fundamentally reshaped service industries worldwide, altering how services are designed, delivered, and managed[1][2]. Among emerging technologies, artificial intelligence (AI) has garnered particular attention as a key driver of service innovation, enabling automated decision-making, real-time data processing, and advanced personalization[3]. As customer interactions increasingly occur through digital interfaces, AI has become a strategic resource for organizations seeking to enhance operational efficiency and service competitiveness.

The hospitality industry has emerged as one of the sectors most strongly influenced by this technological shift. Traditionally characterized by intensive human interaction and emotional labor, hospitality services are now undergoing structural transformation through the integration of AI-based technologies, including chatbots, service robots, automated check-in systems, and algorithm-driven pricing tools[4]. These technologies are increasingly adopted in response to labor shortages, rising operational costs, and heightened customer expectations for speed, accuracy, and convenience[5]. Consequently, AI adoption in hospitality has progressed beyond experimental use and has become embedded in core service and managerial processes.

1.2. Expanding applications and emerging challenges of AI adoption

From a customer experience perspective, prior studies suggest that AI-based services can positively influence service evaluations by improving accessibility, responsiveness, and consistency across service encounters[6]. Chatbots and service robots are particularly effective in routine and standardized interactions such as reservations, information provision, and basic service requests. However, research also indicates that customer responses to AI services vary depending on perceived interaction quality, social presence, and the degree of anthropomorphism embedded in AI systems[7][8]. These findings imply that AI adoption affects not only functional service outcomes but also the experiential and emotional dimensions of hospitality services.

Beyond customer-facing applications, AI has increasingly been applied to operational and managerial functions in hospitality organizations. AI-driven analytics support demand forecasting, revenue management, and decision-support systems, allowing managers to optimize pricing strategies and resource allocation[9][10]. Such applications contribute to productivity enhancement and cost reduction by facilitating data-informed decision-making[11]. At the same time, scholars caution that excessive reliance on algorithmic outputs may constrain managerial judgment and generate new forms of organizational dependency on technology[11].

Despite these benefits, the diffusion of AI in hospitality has raised significant ethical and social concerns. AI-based services depend heavily on the collection and processing of personal data, intensifying issues related to privacy protection and data security[12][13]. Moreover, algorithmic decision-making systems may reproduce or amplify biases embedded in training data, leading to concerns regarding transparency, fairness, and accountability[14][15]. Recent regulatory initiatives, such as the European Union’s Artificial Intelligence Act, further emphasize the growing importance of ethical governance frameworks for AI deployment in service industries[16].

1.3. Research gaps and purpose of the study

In addition to ethical considerations, workforce-related challenges represent a critical issue in the context of AI adoption in hospitality. Automation and service robots have generated concerns regarding job displacement, deskilling, and employee resistance, particularly in labor-intensive service environments[5][17]. While some studies emphasize the potential of AI to complement human labor, others highlight psychological insecurity and organizational tensions arising from perceived job threats and changing skill requirements[17]. These mixed findings suggest that AI adoption in hospitality involves complex trade-offs between efficiency gains and social sustainability.

Although a growing body of literature has examined AI utilization in hospitality, existing studies tend to focus on specific technologies or isolated outcome variables. Relatively limited research has attempted to integrate customer experience, operational and managerial efficiency, and ethical considerations within a unified analytical framework[18]. As AI continues to reshape hospitality services at multiple levels, a more comprehensive understanding of these interconnected dimensions is increasingly required.

Accordingly, the purpose of this study is to systematically review recent research on AI utilization in the hospitality industry by synthesizing findings across three key dimensions: customer experience, operational and managerial efficiency, and ethical and social issues. By providing an integrative overview of existing studies, this research aims to identify dominant research trends, clarify unresolved challenges, and offer directions for future research and responsible AI adoption in hospitality contexts.

2. Methodology

This study adopts a literature review methodology to systematically examine research on artificial intelligence (AI) utilization in the hospitality industry. A literature review is particularly appropriate for synthesizing knowledge in emerging and interdisciplinary research domains characterized by rapid technological change and conceptual fragmentation[18]. Given the expanding scope of AI applications in hospitality, this approach enables an integrative understanding of dominant research themes and unresolved issues.

Academic articles were collected from major international scholarly databases, including Web of Science, Scopus, and Google Scholar. To ensure comprehensive coverage, a combination of keywords was employed, including “artificial intelligence,” “AI,” “hospitality,” “hotel,” “service robots,” “chatbots,” and “AI ethics.” These keywords were selected to capture both technological and managerial dimensions of AI adoption as well as ethical and social considerations[3][6][15].

To reflect recent research trends, priority was given to studies published from 2020 onward. However, seminal studies that provide foundational theoretical frameworks for service automation and AI-enabled services were selectively included to strengthen conceptual grounding[3][7]. Following the initial search, duplicate records were removed, and abstracts were screened to assess relevance. Full-text reviews were subsequently conducted to identify studies directly addressing AI applications within hospitality contexts.

The final set of selected articles was classified into three analytical categories based on their primary research focus: (1) AI and customer experience, (2) AI and operational and managerial efficiency, and (3) ethical and social issues related to AI adoption. This classification scheme reflects the multidimensional nature of AI implementation in hospitality and aligns with prior review-based research frameworks in tourism and service studies[8][18].

Within each category, the selected studies were comparatively analyzed in terms of research objectives, theoretical perspectives, methodological approaches, and key findings. This comparative analysis facilitated the identification of dominant research patterns as well as inconsistencies and research gaps across studies[11]. Rather than conducting a meta-analysis, this review emphasizes qualitative synthesis to accommodate the methodological diversity of the existing literature.

Through this structured review process, the study aims to provide a coherent overview of AI research trends in the hospitality industry and to establish a foundation for future empirical research and policy discussions concerning responsible and sustainable AI adoption.

3. Artificial Intelligence in the Hospitality Industry: Review of Research Trends

3.1. Artificial intelligence and customer experience in hospitality

Research on artificial intelligence (AI) utilization in the hospitality industry has largely focused on customer experience, reflecting the sector’s emphasis on service quality and experiential value[19]. AI-based technologies have been introduced to enhance service encounters by improving efficiency, consistency, and personalization across the customer journey[3][4]. As a result, customer experience has become a central lens through which the effectiveness of AI adoption in hospitality is evaluated.

Among AI applications, chatbots are one of the most extensively studied tools in hospitality contexts. AI-powered chatbots are commonly used to manage routine service interactions, including reservations, check-in inquiries, and basic information provision. Prior studies indicate that chatbots enhance perceived service accessibility and responsiveness by enabling real-time and continuous customer support, particularly in standardized service encounters[6]. These functional benefits contribute positively to customer satisfaction when service expectations are task-oriented and efficiency-driven.

Service robots constitute another major stream of research related to AI-driven customer experience. Existing studies suggest that service robots can improve service consistency and reliability while shaping customers' cognitive and emotional evaluations of service encounters[7][8]. Customers tend to respond more favorably to robot-delivered services when tasks are repetitive and low in emotional complexity. At the same time, research emphasizes that customer acceptance of service robots is influenced by perceived social presence and the degree of anthropomorphism embedded in robot design[20].

Recent hospitality research has extended beyond functional outcomes to examine the psychological and emotional dimensions of AI-mediated service encounters. Social presence has been identified as a key factor influencing trust and customer evaluations of AI-based services[7]. AI agents that demonstrate adaptive communication and contextual awareness are more likely to elicit positive responses. However, excessive anthropomorphism may also lead to discomfort or unrealistic expectations, indicating the need for balanced service design strategies[8].

Customer responses to AI-based services further vary according to individual characteristics. Technology readiness, prior experience with AI, and cultural attitudes toward automation significantly moderate customer acceptance and satisfaction[6]. These findings suggest that AI-driven customer experience strategies should be tailored to different customer segments rather than uniformly applied across all service contexts.

Ethical considerations have increasingly emerged as integral to discussions of AI-driven customer experience in hospitality. Privacy concerns related to the collection and use of personal data negatively affect customer trust and willingness to engage with AI-based services[12][21]. Ethical research highlights that transparency, informed consent, and responsible data governance are essential for sustaining positive customer relationships[14][22]. Studies published in Robotics & AI Ethics further emphasize that customer-facing AI systems should respect human dignity, autonomy, and fairness, particularly in service environments involving sensitive personal information[23][24][25].

In addition, algorithmic bias represents a critical ethical challenge influencing customer experience. AI systems trained on biased data may generate discriminatory service outcomes, undermining perceptions of fairness and trust[14][15]. Recent policy frameworks, including the NIST AI Risk Management Framework and the European Union's Artificial Intelligence Act, underscore the necessity of embedding ethical governance mechanisms into AI systems that directly interact with customers[13][16]. Collectively, the literature suggests that AI-driven customer experience in hospitality is shaped by the interaction of service efficiency, emotional engagement, individual differences, and ethical responsibility.

3.2. Artificial intelligence and operational and managerial efficiency in hospitality

Beyond customer-facing services, artificial intelligence (AI) has been increasingly adopted to enhance operational and managerial efficiency in the hospitality industry. Due to the labor-intensive and cost-sensitive nature of hospitality operations, AI technologies are widely regarded as strategic tools for improving productivity, optimizing resource allocation, and supporting managerial decision-making[4][5].

A major stream of research focuses on the application of AI in revenue management and demand forecasting. AI-driven pricing and forecasting systems analyze large-scale historical and real-time data to predict demand patterns and optimize pricing strategies more effectively than traditional models[9][10]. Empirical evidence suggests that these systems contribute to improved revenue performance through dynamic pricing and more efficient inventory management.

AI adoption has also expanded to workforce management and operational automation. Automated scheduling systems and service robots reduce employees' workload and enhance operational consistency, particularly in repetitive service tasks[8]. While such automation contributes to cost reduction and efficiency gains, studies report that employees often perceive AI as a threat to job security, leading to resistance and negative organizational attitudes[17]. These findings highlight the importance of proactive human resource management during AI implementation.

Recent research has increasingly addressed the ethical and governance implications of AI-driven operational efficiency. Studies grounded in AI ethics emphasize the need for transparency, accountability, and explainability when algorithmic systems influence organizational decisions, particularly those related to workforce management[26]. Ethical discussions published in Robotics & AI Ethics further argue that efficiency-oriented AI adoption without adequate oversight may exacerbate power imbalances and fairness concerns within organizations[24][25].

Policy frameworks such as the NIST AI Risk Management Framework and the European Union's Artificial Intelligence Act reinforce the importance of responsible AI governance in organizational settings[13][16]. Collectively, the literature suggests that while AI enhances operational efficiency and managerial performance in hospitality, its sustainable integration depends on balanced strategies that combine technological capability with human oversight and ethical responsibility.

3.3. Artificial intelligence and ethical and social issues in hospitality

As artificial intelligence (AI) adoption expands across hospitality operations, ethical and social issues have emerged as critical concerns shaping both organizational practices and stakeholder perceptions. Unlike back-end technologies, AI systems in hospitality frequently interact directly with customers and employees, increasing the ethical stakes associated with data use, automated decision-making, and service outcomes[22][14]. Consequently, recent research has increasingly emphasized that ethical considerations should be treated as integral components of AI implementation rather than secondary constraints.

Privacy and data protection constitute the most widely discussed ethical issues in AI-enabled hospitality services. AI-based systems rely on extensive collection and analysis of personal data to deliver personalized services, raising concerns regarding data security, secondary use, and informed consent[12][21]. Empirical studies demonstrate that perceived privacy risks negatively influence customer trust and willingness to engage with AI-driven services, even when efficiency gains are evident. These findings highlight privacy protection as a fundamental prerequisite for sustaining positive customer relationships in AI-mediated service environments.

Algorithmic transparency and fairness represent another major ethical challenge. AI systems trained on biased or incomplete datasets may produce discriminatory outcomes, such as differential service quality or exclusionary pricing practices[14][15]. Such outcomes undermine perceptions of fairness and accountability, which are particularly salient in hospitality contexts characterized by diverse customer populations. Ethical analyses emphasize the importance of explainable AI systems that allow stakeholders to understand and contest automated decisions[22][15].

Workforce-related ethical issues have also attracted growing scholarly attention. Automation and AI-driven management systems raise concerns regarding job displacement, deskilling, and psychological insecurity among hospitality employees[17]. While some studies suggest that AI can complement human labor by reallocating employees to higher-value tasks, others highlight power asymmetries created by algorithmic management and performance monitoring. Research published in Robotics & AI Ethics underscores that efficiency-oriented AI adoption without adequate ethical safeguards may compromise employee autonomy and dignity[23][24][26].

Recent policy and regulatory frameworks further reinforce the importance of ethical governance in AI deployment. The NIST AI Risk Management Framework emphasizes risk identification, accountability, and human oversight throughout the AI lifecycle[13]. Similarly, the European Union's Artificial Intelligence Act introduces risk-based classifications and transparency obligations for AI systems interacting with consumers and workers[16]. These frameworks provide important reference points for hospitality organizations seeking to align AI innovation with ethical and legal expectations.

Recent ethical scholarship further emphasizes the importance of translating high-level AI principles into operational governance practices that can be effectively implemented within service industries, including hospitality[27].

Overall, the literature suggests that ethical and social issues are not peripheral but central to the sustainable adoption of AI in hospitality. Privacy protection, algorithmic fairness, workforce well-being, and transparent governance mechanisms collectively shape stakeholder trust and long-term organizational legitimacy. Addressing these issues requires a shift from technology-centered implementation toward ethically informed, human-centered AI strategies that balance efficiency gains with social responsibility

4. Discussion and Implications

This study reviewed recent research on artificial intelligence (AI) utilization in the hospitality industry by integrating three key dimensions: customer experience, operational and managerial efficiency, and ethical and social issues. The findings demonstrate that AI adoption in hospitality represents a multidimensional transformation that simultaneously reshapes service encounters, organizational processes, and ethical responsibilities[3][4][18]. Rather than a purely technological trend, AI functions as a socio-technical system embedded within service and managerial contexts.

From a theoretical perspective, this review contributes to hospitality and service research by emphasizing the need for integrative analytical frameworks. While prior studies have often examined AI applications in isolation, focusing on either customer-facing technologies or operational outcomes[6][9], the present synthesis shows that these dimensions are closely interconnected. Efficiency gains derived from AI-driven automation may enhance service consistency, yet they can also generate ethical tensions related to employee well-being and customer trust[17][22]. These findings suggest that future research should move beyond fragmented approaches and adopt holistic perspectives on AI-enabled hospitality systems.

With respect to customer experience, the literature indicates that functional efficiency alone does not guarantee favorable service evaluations. Although AI technologies improve accessibility and responsiveness, customers' emotional responses are shaped by perceived social presence, trust, and fairness[7][8][19]. Ethical concerns, particularly those related to privacy protection and algorithmic transparency, further influence customer acceptance of AI-based services [12][21]. Accordingly, hospitality firms should conceptualize AI systems as components of experience co-creation processes rather than merely as efficiency-enhancing tools.

From an operational and managerial standpoint, AI adoption offers significant opportunities for productivity improvement and data-driven decision-making[9][10][11]. At the same time, algorithmic management introduces organizational risks, including employee resistance, job insecurity, and reduced managerial discretion[11][17]. Ethical analyses published in Robotics & AI Ethics caution that efficiency-oriented AI implementation without adequate oversight may exacerbate power asymmetries between organizations and workers, thereby undermining organizational legitimacy and social sustainability[23][24][26].

Policy and governance implications also emerge as critical issues. Regulatory frameworks such as the NIST AI Risk Management Framework and the European Union's Artificial Intelligence Act underscore the importance of accountability, transparency, and human oversight in AI deployment[13][16]. For hospitality firms, alignment with these frameworks may function not only as regulatory compliance but also as a strategic mechanism for building trust among customers and employees. Ethical governance should therefore be embedded within organizational decision-making structures.

Overall, the discussion highlights that successful AI adoption in hospitality depends on the industry's ability to balance technological efficiency with ethical responsibility and human-centered service values. AI-driven innovation is most likely to be sustainable when efficiency gains are accompanied by transparent governance, employee engagement, and respect for stakeholder trust. These insights offer practical and theoretical guidance for researchers and practitioners seeking to promote the responsible adoption of AI in the hospitality industry.

Practical Implications for Hospitality Practitioners

From a practical perspective, this study suggests that hospitality organizations adopt a multilevel approach to AI implementation that links strategic intent with operational execution. At the strategic level, top management should clearly define the role of AI in relation to service values, customer experience objectives, and ethical principles. At the operational level, AI should be applied selectively to standardized and data-intensive tasks—such as demand forecasting, pricing optimization, and routine customer inquiries—while maintaining human involvement in high-contact service interactions. At the organizational level, internal governance mechanisms addressing data privacy, algorithmic transparency, and employee participation are essential to mitigate ethical and workforce-related risks. This approach helps hospitality firms translate abstract AI principles into actionable managerial practices and supports sustainable AI adoption that balances efficiency, service quality, and organizational legitimacy[28].

Recent ethical scholarship emphasizes that the evaluation of AI systems should be grounded in explicit normative principles rather than technical performance alone. Floridi et al. advance the AI4People framework, which delineates five foundational ethical principles—beneficence, non-maleficence, autonomy, justice, and explicability—as a comprehensive normative schema for the governance of AI systems. This framework provides a rigorous ethical basis for evaluating AI deployment in hospitality settings, where service encounters are intrinsically value-laden, relational, and socially embedded[29].

While ethical principles provide an essential normative baseline, they are insufficient on their own to ensure responsible AI practices. Mittelstadt critically argues that ethical guidelines must be accompanied by institutionalized governance mechanisms, accountability structures, and enforcement procedures. This insight is particularly salient for the hospitality industry, where ethical AI adoption depends not only on organizational intent but also on concrete managerial and regulatory arrangements[30].

5. Conclusion and Future Research

This study examined recent research on artificial intelligence (AI) utilization in the hospitality industry by integrating three interrelated dimensions: customer experience, operational and managerial efficiency, and ethical and social issues. The review demonstrates that AI adoption in hospitality should not be understood merely as a technological advancement, but rather as a structural transformation that reconfigures service delivery, organizational management, and ethical responsibility[3][4].

From the perspective of customer experience, this study argues that the value of AI lies not only in functional efficiency but also in its capacity to reshape how experiences are designed and perceived. While AI-based services enhance accessibility, responsiveness, and personalization, positive customer evaluations depend on psychological and ethical conditions such as perceived social presence, fairness, and trust[7][8]. In hospitality contexts, where emotional engagement remains central, AI systems should therefore be positioned as supportive service actors rather than direct substitutes for human interaction.

With respect to operational and managerial efficiency, AI-driven analytics and decision-support systems provide clear advantages in areas such as demand forecasting and revenue management. However, this study emphasizes that efficiency gains achieved through algorithmic management are not value-neutral. Overreliance on automated decision-making may reduce managerial discretion and intensify employee perceptions of job insecurity, thereby generating organizational tension [11][17]. From a scholarly standpoint, these findings suggest that AI implementation strategies must be evaluated not only in terms of performance outcomes but also in relation to their broader organizational and human consequences.

Ethical and social issues emerge as a defining factor for the long-term sustainability of AI adoption in hospitality. This review contends that concerns related to privacy protection, algorithmic bias, transparency, and accountability are not peripheral constraints but foundational conditions shaping stakeholder acceptance and institutional legitimacy[15][22]. Regulatory initiatives such as the NIST AI Risk Management Framework and the European Union's Artificial Intelligence Act further indicate a shift toward governance-centered approaches to AI deployment[16], underscoring the growing expectation that hospitality organizations demonstrate ethical responsibility alongside technological competence.

In conclusion, this study argues that the strategic value of AI in hospitality lies in its integration with human-centered service values and ethical governance structures. AI-driven innovation is most likely to contribute to sustainable industry development when efficiency-oriented objectives are aligned with social responsibility and institutional trust. By advancing this perspective, the study contributes to ongoing academic discussions on responsible AI and provides a foundation for future hospitality research grounded in both technological advancement and ethical consideration.

This perspective aligns with broader scholarship emphasizing principled frameworks that ethical AI adoption requires a principled framework that integrates human autonomy, justice, beneficence, and explicability into socio-technical systems. In this regard, the AI4People framework provides a comprehensive ethical foundation for evaluating AI deployment beyond efficiency and performance metrics, offering an essential reference point for ethically grounded AI adoption in hospitality contexts[29].

However, subsequent research cautions that ethical principles alone are insufficient to ensure responsible AI outcomes without institutionalized governance, accountability, and enforcement mechanisms. This perspective underscores the need for concrete organizational and policy-level arrangements that translate ethical commitments into practice, particularly in service industries characterized by complex human–AI interactions[30].

Despite its contributions, this study is subject to limitations inherent in literature-based research, as it does not empirically test causal relationships or capture longitudinal changes in AI adoption within hospitality contexts. Building on this limitation, several directions for future research are proposed. First, longitudinal empirical studies are needed to examine how AI adoption influences customer trust, employee well-being, and organizational performance over time. Second, greater scholarly attention should be directed toward the role of organizational culture and leadership in mediating ethical AI practices. Third, future research should explore industry-specific governance models that translate ethical principles into actionable managerial guidelines within hospitality settings.

6. References

6.1. Journal articles

[1] Ivanov S & Webster C. Conceptual Framework of the Use of Robots, Artificial Intelligence and Service Automation in Tourism. *International Journal of Contemporary Hospitality Management*, 31(2), 440-459 (2019).

[2] Ivanov S & Webster C & Berezina K. Adoption of Robots and Service Automation by Tourism and Hospitality Companies. *Revista Turismo & Desenvolvimento*, 27/28, 1501-1517 (2017).

[3] Lu L & Cai R & Gursoy D. Developing and Validating a Service Robot Integration Willingness Scale. *International Journal of Hospitality Management*, 80, 36-51 (2019).

[4] Buhalis D & Leung R. Smart Hospitality-interconnectivity and Interoperability towards an Ecosystem. *International Journal of Hospitality Management*, 71, 41-50 (2018).

[5] Tussyadiah IP. A Review of Research into Automation in Tourism. *Annals of Tourism Research*, 81, n102883 (2020).

[6] Gursoy D & Chi OH & Lu L & Nunkoo R. Consumers' Acceptance of Artificially Intelligent (AI) Device Use in Service Delivery. *International Journal of Information Management*, 49, 157-169 (2019).

[7] Pilla R & Sivathanu B & Dwivedi YK. Shopping Intention at AI-powered Automated Stores. *Journal of Retailing and Consumer Services*, 57, n102207 (2020).

[8] Van Doorn J & Mende M & Noble SM & Hulland J & Ostrom AL & Grewal D & Petersen JA. Domo Arigato Mr. Roboto: Emergence of Automated Social Presence in Organizational Frontlines. *Journal of Service Research*, 20(1), 43-58 (2017).

[9] Wirtz J & Patterson PG & Kunz WH & Gruber T & Lu VN & Paluch S & Martins A. Brave New World: Service Robots in the Frontline. *Journal of Service Management*, 29(5), 907-931 (2018).

[10] Davenport TH & & Ronanki R. Artificial Intelligence for the Real World. *Harvard Business Review*, 96(1), 108-116 (2018).

[11] Kimes SE. The Future of Hotel Revenue Management. *Journal of Revenue and Pricing Management*, 16(5), 446-448 (2017).

[12] Xiang Z & Magnini VP & Fesenmaier DR. Information Technology and Consumer Behavior in Travel and Tourism. *Tourism Management*, 54, 244-254 (2015).

[13] Kellogg KC & Valentine MA & Christin A. Algorithms at Work: The New Contested Terrain of Control. *Academy of Management Annals*, 14(1), 366-410 (2020).

[14] Martin K. Ethical Implications and Accountability of Algorithms. *Journal of Business Ethics*, 160, 835-850 (2019).

[23] Floridi L & Cowls J & Beltrametti M & Chatila R & Chazerand P & Dignum V & Luetge C & Madelin R & Pagallo U & Rossi F & Schafer B & Valcke P & Vayena E. AI4 People -An Ethical Framework for a Good AI Society. *Minds and Machines*, 28(4), 689-707 (2018).

[24] Mittelstadt B. Principles Alone Cannot Guarantee Ethical AI. *Nature Machine Intelligence*, 1(11), 501-507 (2019).

[25] Tussyadiah IP & Pesonen J. Impacts of Peer-to-peer Accommodation Use on Travel Patterns. *Journal of Travel Research*, 55(8), 1022-1040 (2016).

[26] Raisch S & Krakowski S. Artificial Intelligence and Management. *Academy of Management Review*, 46(1), 192-210 (2021).

[27] Longoni C & Bonezzi A & Morewedge CK. Resistance to Medical Artificial Intelligence. *Journal of Consumer Research*, 46(4), 629-650 (2019).

[28] Parasuraman A & Colby CL. An Updated and Streamlined Technology Readiness Index. *Journal of Service Research*, 18(1), 59-74 (2015).

[29] Jobin A & Ienca M & Vayena E. The Global landscape of AI Ethics Guidelines. *Nature Machine Intelligence*, 1(9), 389-399 (2019).

[30] Samala N & Katkam BS & Bellamkonda RS & Rodriguez RV. Impact of AI and Robotics in the Tourism Sector: A Critical Insight. *Journal of Tourism Futures*, 8(1), 73-87 (2022).

6.2. Books

[15] Vološin M & Ladkin A. The Algorithmic Management: Reflecting on the Practices of Airbnb. Routledge (2022).

[18] Zuboff S. The Age of Surveillance Capitalism. Public Affairs (2019).

[19] Coeckelbergh M. AI Ethics. MIT (2020).

[20] O'Neil C. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Crown (2016).

[21] Sharkey N & Sharkey A. The Rights and Wrongs of Robot Care. In Robot ethics: The Ethical and Social Implications of Robotics. MIT (2011).

[22] Pine BJ & Gilmore JH. The Experience Economy. Harvard Business School (1999).

6.3. Additional Reference

[16] National Institute of Standards and Technology. AI Risk Management Framework (AI RMF 1.0). U.S. Department of Commerce (2023).

[17] European Commission. Artificial Intelligence Act. Brussels: European Union (2024).

7. Appendix

7.1. Author's contribution

	Initial name	Contribution
Lead Author	SS	-Set of concepts ☑ -Design ☑ -Getting results ☑ -Analysis ☑ -Make a significant contribution to collection ☑ -Final approval of the paper ☑
Corresponding Author*	KL	-Corresponding ☑ -Play a decisive role in modification ☑ -Significant contributions to concepts, designs, practices, analysis and interpretation of data ☑ -Participants in Drafting and Revising Papers ☑ -Someone who can explain all aspects of the paper ☑

2025 Vol.10 No.0, 53-66 Received: 2025/09/07, Peer-reviewed: 2025/10/14, Accepted: 2025/10/30, Published: 2025/12/30

Robotics & AI Ethics

Publisher: J-INSTITUTE
ISSN: 2435-3345

Website: j-institute.org

Corresponding author*
E-mail: hans.kim@pusan.ac.kr

DOI Address:
dx.doi.org/10.22471/ai.2025.10.0.53

An Analysis of the Multilayered Structure of Global AI Ethics Governance

Eunji Lee[1]

Pusan National University, Doctoral Candidate, Republic of Korea

Hyunsoo Kim[2*]

Pusan National University, Associate Professor, Republic of Korea

Abstract

Purpose: *The governance of global AI ethics is not about declaring the legitimacy of AI ethics per se, but rather analyzing the multi-layered nature of governance, where ethical principles are translated into actual norms, policies, standards, procurement, auditing, and accountability systems. Therefore, the goal is to uncover the following: First, it clarifies the layers of global AI ethics governance and the regulatory instruments used at each layer. Second, it clarifies where coherence and conflict arise between layers, and what mechanisms mediate them. Third, it clarifies how the path from soft regulation, ethics, to quasi-norms or quasi-enforcement, is formed. This leads to proposals for the governance of AI ethics.*

Method: *This study first utilizes a literature review method. It first explores documents that present basic theories related to governance theory and AI ethics policy practice. Next, it examines policy-related documents. Furthermore, some of the content encompasses multi-layered documents containing ethical standards, reports from Big Tech-focused companies, and audit frameworks. Next, it utilizes a comparative analysis method. The previously discussed documents are compared by defining categories such as principles and values, obligations or requirements, sanctions and auditing as enforcement, and scope of application. Finally, it utilizes a developmental research method. This developmental research develops and presents a governance mapping structure.*

Results: *A structural analysis of global AI ethics governance at the international level concretizes the production of principles and competition for legitimacy. At the regional and supranational levels, it unfolds as a strengthening of norms through market integration. At the national level, it manifests as a combination of administrative and supervisory systems and industrial policies. At the industry and sector levels, it converges into a risk-based, detailed framework. At the organizational level, internal governance is institutionalized around ethics committees. At the technological and system level, it is concretized as the codification of ethics, particularly in the context of data governance.*

Conclusion: *At each level, conflicts of value, jurisdiction, responsibility, and technology can arise. Furthermore, competition can arise between regulators in the establishment of norms, the market power of large platforms and cloud providers, standards-setting entities, and auditing power related to the evaluation criteria for market compliance. This will require the establishment of meta-principles that analyze and connect implementation requirements, interoperability strategies among entities necessary to resolve conflicts at each level, and the internalization of accountability and redundancy mechanisms. Based on this, it is suggested that for specific countries or actors to secure initiative, they need to participate in strategic standardization, establish procurement standards, and establish industry-specific guidelines to become both adopters and producers of global norms.*

Keywords: ***Global AI Ethics, Governance, Multi-Layered Structural Analysis, Meta-Principles, AI Ethics Initiatives***

1. Understanding AI Ethics as a Social Ethics and Its Governance

Discussions on AI ethics typically begin with a declaration regarding the legitimacy of AI ethics itself, expanding into discussions on various areas where AI ethics can be applied[1][2]. Meanwhile, these discussions also lead to research on the ethical behavior of AI itself, a representative example of which is a study exploring ethical issues arising from the outputs of generative AI[3]. Meanwhile, discussions on the topic of AI ethics are diverse, including military issues, particularly arms control[4], and issues related to biomedical and nursing[5].

Meanwhile, discussions on AI ethics, which began with this topic-centered approach, are expanding to include aspects related to structural issues. Specifically, studies attempting to analyze the structure of ethical decision-making include research exploring the structure of ethical decision-making through collaboration between AI and humans[6], research on the ethics of AI ethics itself[7], and research exploring the structure of AI ethics across three layers: Principles, Processes, and Ethical Consciousness[8].

These studies can be considered representative achievements of the normative ethics perspective on AI ethics. However, to approach AI ethics from a social ethics perspective, as well as an individual ethics perspective, a discussion on governance is required. This requires a fundamental understanding of both social ethics and governance.

First, social ethics is an ethical approach that focuses on social structural problems that are difficult to resolve through individual morality or moral practice alone. Therefore, it refers to an ethical approach that aims to realize justice through the improvement of systems or policies. This approach focuses on improving social structures, recognizing that ethical problems of individual actors can cause problems at the group or community level. Therefore, social ethics seeks to present socially agreed-upon standards as norms at a certain level for the integration and coexistence of society as a whole.

Next, governance is a concept related to administration, encompassing the active and specific activities carried out by the state or public institutions to realize the public interest in accordance with the law. Here, governance refers to a cooperative community operation method where various entities, such as the government, businesses, and citizens, come together to set common goals, make decisions, and solve problems. Therefore, governance goes beyond simple governance or management, and is interpreted in various ways depending on the context of an organization's governance structure, operational system, and decision-making structure.

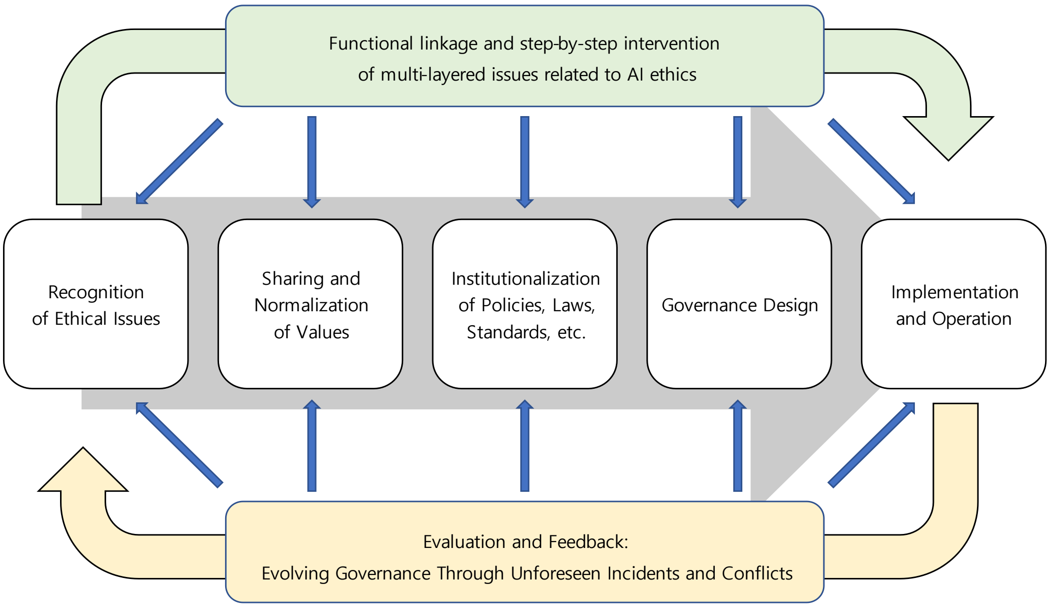
Meanwhile, there are recent research achievements that link these AI ethics principles to governance frameworks and systems. Representative examples include a study that presents an ethics governance roadmap by linking ethics, standards, regulation, and public participation[9], and a study that systematically demonstrates the convergence and disagreement between ethical principles such as transparency, fairness, non-harmfulness, accountability, and privacy and suggests how to implement these principles into governance[10][11]. Meanwhile, a series of studies have proposed increasing the feasibility of ethics through tool development and documentation related to core governance mechanisms. These include studies that propose operational governance of ethics by operating ethical principles as service and organizational capabilities[12], studies that document data ethics by linking them to governance demands based on responsibility and transparency and propose them as tools for carrying out accountability[13], and studies that analyze guidelines for AI ethics at the international and intercultural level[14][15][16][17].

This study aims to explore the governance of AI ethics, building on the aforementioned prior research. To this end, we first reviewed materials that directly connect "governance theory" and "policy practice" in the field of AI ethics, providing a basic theoretical framework[18][19][20][21][22]. Next, we examined policy-related academic research encompassing the content of multi-level documents containing AI ethics standards and reports from big tech-focused companies, as well as some audit frameworks. This research includes studies addressing the ecosystem of multi-level AI ethics standards and guidelines[23][24][25][26], and studies critically analyzing corporate-led ethical self-regulation documents based on big tech and corporate reports[27][28][29][30]. This review of prior research leads to the design of a process for concretizing social ethics into governance.

2. The Process of Embodying Social Ethics in Governance

The process by which social ethics are embodied in governance is structured as a feedback loop. This is particularly evident in the composition and operation of AI. AI is not merely a tool; it directly impacts human rights and opportunities. This impact is specifically related to the mechanisms through which AI operates in areas such as recruitment, lending, welfare, public safety, healthcare, and education. In this respect, AI ethics extends beyond individual morality to encompass issues of social ethics, such as fairness, safety, human rights, and trust. Safety and harmlessness, which enable AI to minimize harm in societal decisions and their execution; autonomy and accountability, which determine who has the authority and responsibility for decisions involving AI; transparency and explainability, which address the understanding of decisions made with AI's assistance and support and the impact of those affected by them; and privacy, which directly relates to data rights and control, are key elements of AI ethics. Taken together, AI ethics represents a social realignment of rights, justice, safety, and accountability. This realignment ultimately represents the process by which social ethics connects to governance. This refers to the process by which moral judgments about right and wrong are institutionalized. Institutionalization here refers to the rules and decision-making structures by which organizations and societies actually operate. This structure is diagrammed as shown in the Following <Figure 1>.

Figure 1. The process by which AI ethics as social ethics connects or evolves to governance.


Functional linkage and step-by-step intervention of multi-layered issues related to AI ethics
Recognition of Ethical Issues
Sharing and Normalization of Values
Institutionalization of Policies, Laws, Standards, etc.
Governance Design
Implementation and Operation
Evaluation and Feedback: Evolving Governance Through Unforeseen Incidents and Conflicts

The structure is explained as follows: In the first stage, awareness of ethical issues arises. In various issues related to AI responsibility, such as safety, human rights, discrimination, the environment, labor, and personal information, moral judgments related to certain issues are understood from an ethical perspective, and this raises the social issue of responsibility. At this stage, issues gain attention and are broadened primarily through discussions in the public sphere. This public sphere is characterized by the media, civic groups, academia, victim reports, and whistleblowing. AI simultaneously functions in this process, collecting, understanding, and analyzing information.

The second stage involves the sharing and normalization of values. This corresponds to the stage where the results or criteria for moral judgments regarding ethical issues related to AI functions solidify into social norms. Within society, these norms are concretized in the form of norms through public debate. These norms represent a level of standard prior to the formation of mandatory rules or orders. Therefore, this stage represents a broad-based consensus regarding AI. The difference between the first and second stages discussed above is that the criteria for judging AI activities and their outcomes are further concretized from abstract values to norms corresponding to concrete behavioral standards.

The third stage involves the institutionalization of policies, laws, and standards. This means that norms embodying ethical judgments related to AI functions gain some form of enforcement or enforcement procedures. If the need for norms related to AI activities continues, specific means for implementing them will need to be presented. These specific means of implementation include enforcement mechanisms such as laws and regulations, unified operational standards such as standards and guidelines, and policy instruments such as implementation plans and budgets. This is where responsibilities, sanctions, incentives, and procedures are created and presented as mechanisms to ensure compliance with AI ethics.

The fourth stage involves governance design. This is the stage of who, how, and by what criteria will establish and determine AI ethics as rules, as shown in the following <Table 1> .

Table 1. Components that make up the structure of governance design.

Components	Questions or Explanations
Roles and Authority	Who is ultimately responsible? Government, Board of Directors, Head of Agency, Responsible Department, etc.
Decision-making Process	What review/approval procedures are in place? Committee, Deliberation, Impact Assessment, etc.
Oversight and Checks	Internal audit, external audit, citizen oversight, independent organizations, etc.
Transparency and Disclosure	Information disclosure, reporting, disclosure of interests (conflict of interest prevention)
Transparency and Disclosure	Information disclosure, reporting, disclosure of interests (conflict of interest prevention)
Participation (Representativeness)	How do citizens/workers/users/experts participate in decision-making?
Accountability and Remedies	Accountability, corrective action, and redress procedures in the event of damage.

This is where the discussion on governance begins in earnest, and the linkages are generally activated by a series of components as follows: ① Roles and authority. This concerns who ultimately bears responsibility, and the government, board of directors, heads of institutions, and responsible departments are the subject of discussion. ② Decision-making procedures. This concerns the procedures for review and approval, including committees, deliberation procedures, and various impact assessments. ③ Oversight and checks. This concerns the roles, duties, and activities of internal audits, external audits, civil society oversight, and independent checks and balances. ④ Transparency and disclosure. This relates specifically to AI ethics and relates to the judgment mechanism and internal design of AI. Discussions include disclosure of information, reports related to operating mechanisms and designers, and disclosure of interests of stakeholders to prevent conflicts of interest. ⑤ Representation and participation. This concerns how various stakeholders, including citizens, workers, users, and experts, participate in decision-making regarding ethical issues related to the use and operating mechanisms of AI. ⑥ Accountability and redress. This directly addresses the issue of responsibility related to AI operations. If damage occurs due to AI activities, the location of responsibility, specific methods of corrective action, and the procedures and content of damage relief are the subject of discussion. As discussed above, the six stages of AI ethics leading to governance design represent the stage where ethical values are transformed into rules or norms for organizational operation and managed.

The fifth stage involves reviewing implementation and operation. This stage embodies and applies AI ethics into practical, on-site processes. In other words, governance, which exists only on paper, cannot function effectively. Therefore, this stage involves integration into the actual operating system and implementation. The system components utilized at this stage include: ① an ethics education or training system related to management compliance. ② a risk management system related to ethics, human rights, and the environment. ③ practical processes related to pre-screening, verification and checklist review procedures, and verification procedures at the approval stage. ④ key performance indicators (KPIs), evaluation, and reward systems that reflect ethics compliance performance in organizational evaluations. ⑤ A hotline is established to ensure immediate response when problems arise, and a whistleblower protection system is in place. The five operating systems described above represent a system capable of repeatedly implementing AI ethics. These examples and their practical applications are as shown in the following <Table 2>.

Table 2. An example of an ethics implementation and operational system embodied in field processes.

Example	Practical Application
Education	Ethics Training/Compliance
Risk Management	Ethics Risks, Human Rights Risks, Environmental Risks
Practical Processes	Pre-screening, Checklists, Approval Steps
KPIs/Evaluation/Reward	Incorporating Ethics Compliance Performance into Organizational Evaluations
Internal Reporting/Protection	Hotline, Whistleblower Protection

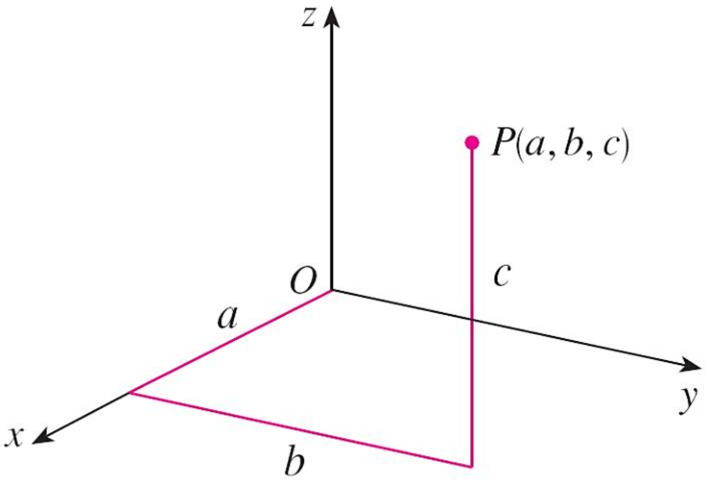
The sixth step is evaluation and feedback. This is not a step in itself, but rather a cyclical feedback loop. This loop means that governance evolves again when unexpected accidents occur due to AI activities or conflicts arise due to social changes stemming from AI. This mechanism stems from incidents such as safety accidents caused by AI itself or human use of AI, or from social changes resulting from the spread of AI. In this regard, the sufficiency of AI-related rules and norms established through the existing AI ethics governance structure is assessed, and laws, policies, and organizational structures are revised accordingly. This iterative process advances both ethics and governance. This can be summarized as the following process: Social ethics (values) → Social norms (standards) → Laws, policies, and standards (systems) → Decision-making, monitoring, and accountability structures (governance) → Operational processes (execution) → Evaluation and improvement (evolution).

3. Analysis of the Three-Axis Matrix Structure of Global AI Ethics

The structure of global AI ethics can be approached by designing a three-axis matrix of open field of Ethical issues on AI.

Each axis can be analyzed by arranging the definition of levels on the x-axis, the mapping of actors on the y-axis, and the catalog of instruments on the z-axis in an open space where issues related to AI ethics are discussed. The 'Phase Evaluation' according to the structure of global AI ethics can be expressed as P(a, b, c) on a three-dimensional Cartesian coordinate plane, based on the combination of level a, actor b, and instrument c. This is illustrated in the Following <Figure 2>.

Figure 2. A three-axis matrix structure for global AI ethics.



3.1. X-axis: defining levels of global AI ethics

The layers of global AI ethics can be approached through a multi-layered analysis. This analysis can be multi-layered, encompassing not only the expansion of spheres of life but also the actors involved in implementing and discussing AI ethics.

First, at the international level, actors within the UN system and at the national level are the actors. At this level, AI ethics enforcement mechanisms include international agreements, declarations, and global forums. At this level, principles are produced and legitimacy contests unfold. While international declarations or recommendations have weak binding power, they offer a kind of "global language". However, they lack enforcement power and have certain limitations in addressing cultural or value conflicts.

Second, at the regional and supranational levels, there are AI ethics norms established by regional blocs based on political communities like the EU, and AI ethics-based regulations based on market integration, such as regional economic agreements. At this level, norms are reinforced through market integration. Regional regulations, in particular, have significant ramifications due to their extraterritorial application and impact on supply chains. However, for more precise application, an analysis of the pathways through which regulatory requirements are translated into standards and audits is necessary.

Third, at the national level, AI ethics can be enforced through legal, administrative, procurement, and supervisory agencies. At this level, administrative oversight systems and industrial policies are often combined. This necessitates a balance between promoting innovation and managing risk. However, the use of these policy tools can lead to government procurement and public sector adoption guidelines becoming de facto norms, necessitating careful application.

Fourth, at the industry and professional level, AI ethics can be implemented through the rules of individual sectors that comprise the social fabric, such as healthcare, finance, and education. At the industry and sector level, the application of ethical principles is specified and refined based on potential risks. The primary discussion occurs in areas with high risk and responsibility, such as healthcare, finance, and education. This discussion often leads to ethics converging with safety and quality management systems.

Fifth, at the organizational level, such as within a company or institution, AI ethics relates to internal governance at a relatively micro level. At this stage, AI ethics can be implemented concretely through roles such as managers, committees, and risk management. At this organizational level, the focus is on institutionalizing internal governance. This is exemplified by mechanisms such as ethics committees, designated managers, model release gates, incident reporting systems, and supply chain requirements reviews.

Sixth, at the technology and system level, AI ethics can be enforced by incorporating mechanisms into the system's models, data, and the pipelines that connect data collection, movement, and processing. This can be codified through logging, which chronologically records the operating status, events, and errors of a software system; system-level evaluations of the application of AI ethics; and guardrail metrics, which serve as safety measures to monitor for unexpected side effects during experiments to improve key service success indicators. In other words, this is the final stage where "principles" are transformed into "measurable requirements."

This discussion ultimately illustrates the process by which guidelines, as soft laws, are hardened and concretized into principles of responsibility. That is, it describes the path of ethical guidelines → standards (technical/management) → certification/audit → procurement/contractual conditions → legal responsibility. This process is not irreversible, but reversible, and is subject to revision and supplementation through mutual feedback. These stages and layers are summarized in the following <Table 3> below.

Table 3. Levels and components of global AI ethics.

Levels and Components	Questions or Explanations
Level 1 International	UN-affiliated organizations, international agreements/declarations, global forums
Level 2 Regional/Supranational	Regulatory and market integration of regional blocs such as the EU
Level 3 National	Legal, administrative, procurement, and supervisory agencies
Level 4 Industry/Specialized Areas	Sectoral rules such as healthcare, finance, and education
Level 5 Organization	Internal governance of companies and institutions: managers, committees, and risk management
Level 6 Technology/Systems	Controls built into models/data/pipelines: logging, assessments, and guardrails

3.2. Y-axis: mapping actors

Mapping actors related to global AI ethics is essentially concretized by examining the perspectives of each stakeholder. These include regulators (governments and supervisory agencies), standards organizations, businesses (ordering/supplying), civil society, academia, auditing and certification agencies, and open source communities.

First, regulators, such as governments and supervisory agencies, function as architects of accountability based on public authority. Their core roles include codifying norms that translate soft principles into legal obligations, prohibitions, permits, and oversight; risk-based regulation, which classifies AI by context and risk level and imposes stronger controls on higher risks; enforcement and sanctions, which utilize coercive tools such as fines, corrective orders, and market exclusion for violations; and market-making and industrial policy, which balances regulation and innovation while maintaining domestic industrial competitiveness while building market trust through safety and reliability requirements.

Second, standards organizations, including international and national standardization, function as translators of technology-policy communication and architects of interoperability. Its core role is to embody ethical principles into technical and organizational processes, such as translating fairness principles into data quality management, testing, and documentation requirements. Furthermore, it establishes a common language and consistent standards that serve as a foundation for communication, enabling businesses, governments, and auditing agencies to discuss risk, quality, and management systems using a common vocabulary, thereby providing interoperability. Furthermore, it can be combined with procurement and certification to create quasi-regulation, effectively establishing standards as market entry requirements, even if they are not legal.

Third, corporations serve as key actors. They are responsible for supplying AI (the development and provision of AI) and ordering AI (the adoption and use of AI). They also serve as key players with actual design and deployment power. Therefore, while corporations wield the most substantial influence in AI ethics, they also face the greatest potential for conflicts of interest. These companies embed values in the product and model design phase, including safeguards, restrictions, data policies, and user guardrails. They also establish internal governance through ethics committees, operational responsibility entities, model release gating, and incident response processes. They also mitigate information asymmetry through activities such as the preparation and distribution of transparency reports and risk disclosure.

Fourth, civil society organizations (NGOs) include consumer groups, human rights organizations, and labor organizations, acting as norm watchdogs and representatives of the victims' perspectives. They reframe AI ethics discussions not just around "efficiency and innovation," but also around human rights, discrimination, surveillance, and labor rights, providing a rights-based framing. They also track whether corporate or government ethics declarations are actually implemented, and furthermore, they perform a watchdog function, including criticizing "ethics washing." They demand participatory governance through consultation and co-design, reflecting the voices of affected groups.

Fifth, the academic community. This community encompasses individual researchers and research-related institutions as communities or organizations. Academia functions as a concept-formulator, evaluation method producer, and critical reflector. Academia defines core concepts such as fairness, accountability, transparency, and safety, decomposing and reorganizing them into measurable forms to refine these concepts. Furthermore, it provides evaluation methods and empirical research in the form of analytical methods for bias measurement, explainability approaches, safety assessment, and policy effectiveness. Furthermore, it participates in policy advisory committees and committees, providing expertise in regulatory design, standardization working groups, and the development of ethics guidelines.

Sixth, the audit and certification bodies (A&Cs) Audit) serves as an institutionalizer of verifiability. This includes traditional certification and testing organizations, quality and security assessment agencies, and even accounting and consulting firms, including third-party assurance providers. Instead of asserting or declaring ethics, they provide measurable trust by assessing compliance with ethical standards through documents, processes, and test results. Furthermore , the certification/audit system complements the enforcement capacity of ethical standards, supplementing the government's inability to directly and closely supervise all companies. Of course, certification gains practical enforcement power when it impacts market access, contracts, and insurance premiums, so ethical standards are enforced through links to procurement, insurance, and investment.

Seventh, open source communities serve as accelerators of decentralized development, transparency, and diffusion. By disclosing AI models and code, they foster verification and innovation, strengthening the capabilities of academia, startups, and civil society. Furthermore, they maintain a balance between security and safety. Most importantly, they shape de facto standards by shaping development practices through industry-standard libraries and frameworks. Enforce ethics.

In relation to the above discussion, individual explanations alone are not sufficient for each actor. It is necessary to analyze the global AI ethics ecosystem through the relationships among these actors when actual cases arise. For example, analyzing the relationship between regulators and standards organizations reveals that laws establish high-level requirements, while standards concretize these into actionable requirements, or conversely, standards serve as a reference for legal frameworks. Analyzing the relationship between businesses and standards or certifications reveals that businesses use standards to make compliance costs predictable, while certifications contribute to market trust and procurement. Analyzing the relationship between civil society and regulators reveals that civil society provides a window into regulatory agendas such as surveillance, discrimination, and labor, while regulators translate these demands into laws and policies. Academia engages with all actors. While it provides concepts, measurement, and evaluation methods, it is also intertwined with businesses and governments through funding and data access. Analyzing the relationship between open source communities and businesses, academia, and civil society reveals that while open source communities facilitate diffusion and verification, they also foster abuse and accountability gaps, leading to relationships with other actors. The above is summarized in the following <Table 4>.

Table 4. Mapping actors and their roles in global AI ethics governance.

Actors	Description and Role of the Actor
Regulators	Governments, Oversight Bodies Designers of Public Power-Based Accountability, Codification of norms, risk-based regulation, enforcement and sanctions, market formation and industrial policy formulation
Standards Bodies	International and National Standardization Technology-Policy Translators, Interoperability Designers, Embodiment of Ethical Principles into Technical and Organizational Process, Providing Interoperability and a Common Language, Quasi-Regulation
Businesses	Supply and Ordering Key actors with actual design and distribution power, Value Embedding at the Product/Model Design Stage: Establishing Internal Governance, Mitigating Information Asymmetry:
Civil Society	NGOs, Consumer Organizations, Human Rights Organizations, Labor Organizations, etc. Norm Watchers, Representatives of the Perspectives of Victims, Rights-Based Framing, Watchdogs, Reflecting the Voice of Affected Groups:

Academia	Researchers/Research Institutions Concept Builders, Producers of Evaluation Methods, Critical Reflector, Refining concepts, providing evaluation methods and empirical research, and participating in policy advisory committees.
Assurance, Certification, Audit	Institutionalizing verifiability. Providing measurable trust, supporting policy enforcement infrastructure, and effectively enforcing ethics through certification.
Open Source Community	Accelerator of decentralized development, transparency, and diffusion. Expanding accessibility and reproducibility, balancing security and safety, and forming de facto standards.

3.3. Z-axis: cataloging instruments

The tools and instruments of global AI ethics are divided into four areas: norm production, enforcement and conformity, economic mechanisms, and accountability and redress. The purpose and function of each tool are examined as follows.

First, norm production tools in global AI ethics define values, assign and coordinate priorities among values, and operationalize them into measurement and procedures. Even if not necessarily in the form of legal provisions, these tools can have a de facto regulatory effect (e.g., standards for compliance, procurement, and certification), thus functioning as quasi-regulation.

Next, the axis of enforcement and conformity tools, embodied through audits, certifications, reporting obligations, and sanctions, solidifies norms not as mere "documents" but as an evidence chain of ethical compliance. The key is to combine ex ante (before design/deployment) control with ex post (monitoring/incident response during operation) control to make risks governable and manageable.

Furthermore, economic mechanisms approach ethics and safety by pricing or contracting them through incentives, even in environments with weak legal enforcement. Especially in international environments, where legal jurisdiction is limited, market mechanisms serve as powerful channels for global diffusion.

Finally, in ethical governance, accountability and redress do not simply end with punishment. They function as mechanisms to raise the safety level of the entire system through restoration of victims' rights (redress), internalization to prevent the spread of risks, and incident learning. The specific types and design parameters of these tools are summarized in in the following <Table 5>.

Table 5. Instrument catalog of global AI ethics.

Instruments	Typical Types	Design Variables
Norm Production: Principles/ Guidelines/ Standards	High-Level Principles/Charters Code of Conduct/Ethics Guidelines Implementation Guidance/Playbooks Standards: Terminology/Process/Testing Professional Standards, Competency Frameworks	Legitimacy Level of Specificity Measurability: Updatable Competition/Overlapping Norms
Enforcement/ Conformity: Audit, Certification, Reporting Obligations, Sanctions	Impact Assessment/Risk Assessment Conformity Assessment Audit Certification/Labeling/Registration Reporting Obligations Sanctions & Corrective Action (actions)	Standardization of evidence Oversight capacity Dynamics Divided responsibility issues
Economic mechanisms: procurement, insurance, investment standards, supply chain requirements	Public procurement Private procurement and contracts Insurance Investment standards (ESG/Responsible AI due diligence) Supply chain/third-party risk management	Accuracy of price signals Power asymmetry Formalism

Liability and remedies: civil and criminal liability, class actions, ADR (Liability & Remedy)	Civil liability (torts/contracts/product liability, etc.) Criminal liability Class actions/representative actions and public interest litigation ADR (Alternative Dispute Resolution): Arbitration/mediation/ombudsman Administrative remedies (reporting/investigation/corrective order) Rights-based mechanisms (clarification requests, objections, human review, deletion/correction)	Responsibility Distribution Verifiability Effectiveness of Relief Preventive Effect

4. Conclusion

This study views global AI ethics governance not as a declaration of ethical principles or a collection of individual guidelines, but as a multi-layered structure in which principles are translated into norms, policies, standards, auditing, procurement, accountability, and redress systems. To achieve this, we propose a three-axis matrix (P(a,b,c)) of Level, Actor, and Instrument as an analytical framework and compare the operational logics and instrument bundles at each level: International (L1), Regional/Transnational (L2), National (L3), Industry/Specialized Area (L4), Organizational (L5), and Technology/System (L6). The results confirm that global AI ethics governance is not a single normative system, but rather a complex ecosystem in which principle production and legitimacy competition, norm hardening through market integration, the combination of administrative oversight and industrial policy, sector-based risk specification, institutionalization of internal governance, and the codification of technological controls are intertwined and circulated. At each level, conflicts of value, jurisdiction, responsibility, and technology can arise. Furthermore, competition can arise between regulators in the establishment of norms, the market power of large platforms and cloud providers, standards-setting entities, and auditing power related to the evaluation criteria for market compliance. This will require the establishment of meta-principles that analyze and connect implementation requirements, interoperability strategies among entities necessary to resolve conflicts at each level, and the internalization of accountability and redundancy mechanisms. Based on this, it is suggested that for specific countries or actors to secure initiatives, they need to participate in strategic standardization, procurement standards, and industry-specific guidelines to become both adopters and producers of global norms.

There is no single mechanism for resolving these conflicts, and the research findings suggest that a key coordination mechanism in multi-level governance is the combination of interoperability and evidence-based accountability. In other words, coordination between levels is not sufficient with abstract value agreement alone, but real consistency is created when (1) a common language (standards, terminology, and evaluation indicators) that translates principles into implementation requirements, (2) a documentation and logging system that can accumulate and transfer evidence of compliance, (3) a feedback loop of independent assessment (audit/certification) and incident reporting and corrective action, and (4) economic incentive coordination through procurement/contracting/insurance/investment standards work together. From this perspective, this study proposed a “hardening path of ethics” structured as AI ethics guidelines → standards (management/technology) → audit/certification (conformity) → procurement/contracting (market access) → accountability/remedy (post-event discipline). Rather than being unidirectional and irreversible, this path forms a circular structure in which incidents/accidents, disputes, and supervision results are fed back into norm production and standard revision.

This study is a developmental research focused on developing a model based on existing research rather than analyzing it itself. The rigor of the analytical procedures and the explanatory power of the study, which are still lacking, will be verified through follow-up research applying this model to specific cases in specific countries and industries.

5. References

5.1. Journal articles

[1] Park G & Bae M. A Case Study on Current Issues in Artificial Intelligence and its' Ethical Implications. *Robotics & AI Ethics*, 7(2), 47-56 (2022). [Read More]

[2] Park G & Bang J. Areas of Ethical Inquiry Related to Artificial Intelligence. *Robotics & AI Ethics*, 9(0), 1-12 (2024). [Read More]

[3] Kim H. Creativity and AI: Products of Generative AI and Ethical Issues. *Robotics & AI Ethics*, 9(0), 23-33 (2024). [Read More]

[4] Kim H. Suggestions for the Role of AI in the Arms Control and Non-proliferation of WMD. *Robotics & AI Ethics*, 7(2), 57-67 (2022). [Read More]

[5] Kim M & Hong B. Nursing Ethical Considerations in the AIBased Technologies. *Robotics & AI Ethics*, 7(2), 10-21 (2022). [Read More]

[6] Kim H. Suggestions for Ethical Decision-making Model through Collaboration between Human and AI. *Robotics & AI Ethics*, 8(0), 12-22 (2023). [Read More]

[7] Heilinger JC. The Ethics of AI Ethics: A Constructive Critique. *Philosophy & Technology*, 35(61), 1-20 (2022).

[8] Kazim E & Koshiyama AS. A High-level Overview of AI Ethics. *Patterns*, 2(9), 1-12 (2021).

[9] Winfield AFT & Jirotka M. Ethical Governance is Essential to Building Trust in Robotics and Artificial Intelligence Systems. *Philosophical Transactions of the Royal Society*, 376(2133), n20180085 (2018).

[10] Mökander J & Morley J & Taddeo M & Floridi L. Ethics-based Auditing of Automated Decision-making Systems: Nature, Scope, and Limitations. *Science and Engineering Ethics*, 27(44), 1-30 (2021).

[11] Mökander J & Floridi L. Operationalising AI Governance through Ethics-based Auditing: An Industry Case Study. *AI and Ethics*, 3(0), 451-468 (2023).

[12] Morley J & Elhalal A & Garcia F & Kinsey L & Mökander J & Floridi L. Ethics as a Service: A Pragmatic Operationalisation of AI Ethics. *Minds and Machines*, 31(2), 239-256 (2021).

[13] Koniakou V. From the Rush to Ethics to the Race for Governance in Artificial Intelligence. *Information Systems Frontiers*, 25(1), 71-102 (2023).

[14] Corrêa NK & Galvão C & Santos JW & Del Pino C & Pinto EP & Barbosa C & Massmann D & Mambrini R & Galvão L & Terem E & de Oliveira, N. Worldwide AI Ethics: A Review of 200 Guidelines and Recommendations for AI Governance. *Patterns*, 4(10), 1-14 (2023).

[15] Larsson S. On the Governance of Artificial Intelligence through Ethics Guidelines. *Asian Journal of Law and Society*, 7(3), 437-451 (2020).

[16] ÓhÉigeartaigh SS & Whittlestone J & Liu Y & Zeng Y & Liu Z. Overcoming Barriers to Cross-cultural Cooperation in AI Ethics and Governance. *Philosophy & Technology*, 33(4), 571-593 (2020).

[17] Zhang B & Anderljung M & Kahn L & Dreksler N & Horowitz MC & Dafoe A. Ethics and Governance of Artificial Intelligence: Evidence from a Survey of Machine Learning Researchers. *Journal of Artificial Intelligence Research*, 71(0), 591-666 (2021).

[18] Black J. Constructing and Contesting Legitimacy and Accountability in Polycentric Regulatory Regimes. *Regulation & Governance*, 2(2), 137-164 (2008).

[19] Abbott KW & Snidal D. The Governance Triangle: Regulatory Standards Institutions and the Shadow of the State. *The Politics of Global Regulation*, 44(1), 44-88 (2009).

[20] Floridi L. Soft Ethics and the Governance of the Digital. *Philosophy & Technology*, 31(0), 1-8 (2018).

[21] Mittelstadt BD. Principles Alone Cannot Guarantee Ethical AI. *Nature Machine Intelligence*, 1(0), 501-507 (2019).

[22] Welle A. Transparency: Motivations and Challenges. *Journal of Machine Learning Research*, 20(1), 1-23 (2019).

[23] Jobin A & Ienca M & Vayena E. The Global landscape of AI Ethics Guidelines. *Nature Machine Intelligence*, 1(0), 389-399 (2019).

[24] Mittelstadt BD. Principles Alone Cannot Guarantee Ethical AI. *Nature Machine Intelligence*, 1(0), 501-507 (2019).

[25] Hagendorff T. The Ethics of AI Ethics: An Evaluation of Guidelines. *Minds and Machines*, 30(1), 99-120 (2020).

[26] Morley J & Floridi L & Kinsey L & Elhalal A. From What to How: An Overview of AI Ethics Tools, Methods and Research to Translate Principles into Practices. *Science and Engineering Ethics*, 26(0), 2141-2168 (2020).

[27] Metcalf J & Moss E & Boyd D. Owning Ethics: Corporate Logics, Silicon Valley, and the Institutionalization of Ethics. *Science, Technology, & Human Values*, 44(3), 449-476 (2019).

[28] Cihon P & Schuett J & Baum SD. Corporate Governance of Artificial Intelligence in the Public Interest. *Information 2021*, 12(275), 1-30 (2021).

[29] Sætra HS. The AI Ethicist's Dilemma: Fighting Big Tech by Supporting Big Tech. *AI and Ethics*, 2(0), 15-27 (2022).

[30] Schultz M & Conti L & Seele P. Digital Ethicswashing: A Systematic Review and a Process-perception-outcome Framework. *AI and Ethics*, 5(0), 805-818 (2025).

6. Appendix

6.1. Author's contribution

	Initial name	Contribution
Lead Author	EL	-Set of concepts ☑ -Design ☑ -Getting results ☑ -Analysis ☑ -Make a significant contribution to collection ☑ -Final approval of the paper ☑
Corresponding Author*	HK	-Corresponding ☑ -Play a decisive role in modification ☑ -Significant contributions to concepts, designs, practices, analysis and interpretation of data ☑ -Participants in Drafting and Revising Papers ☑ -Someone who can explain all aspects of the paper ☑

6.2. Funding agency

This work was supported by a Journal Research Promotion(2025) for Graduate Students of Pusan National University.