Research on ethical issues and improvement tasks related to Big Data analysis and utilization of Artificial Intelligence

Se Eun Jeong

ABSTRACT

As Artificial Intelligence(AI)-based data analysis technology develops, it fails to keep up with general ethical standards and gaps arise between them, leading to ethical issues. When AI technology is used to analyze and utilize Big Data, it is no longer a threat to ethical influence, but can be a good turning point that turns a crisis into an opportunity. It is necessary to examine the ethical influence and other issues that these technologies have on the Big Data analysis process.

In this paper, it is first examined the relationship between AI and Big Data, and looked at how AI uses Big Data to learn data through Machine Learning algorithms and classify or predict new data.

Next, it is investigated how data is collected, what algorithms are used, and how the recommendation system is improved through real-time learning through a case study on Netflix. In relation to this, it is studied issues such as bias and lack of transparency that arise when examining ethical influences of AI technology on the Big Data analysis and utilization process.

Here, when the AI decision-making process is opaque, problems such as lack of trust, responsibility issues, difficulty in risk management, and conflicts of interest arise.

As a solution to these problems, both legal and technical solutions were presented. As such, Artificial Intelligence is in the process of developing very good technology in itself, but the problem of using Artificial Intelligence for Big Data cannot but be considered.

Therefore, if issues such as ethical impact are reviewed and evaluated in advance, and solutions and development tasks are sufficiently prepared, it can be expected to minimize the risks that may arise and maximize the advantages of Artificial Intelligence technology in using Big Data.

☞ keyword : Artificial Intelligence, Artificial Intelligence technology, Big Data, Big Data utilization, ethical implications.

1. Introduction

Today, in the era of the 4th Industrial Revolution, the amount of data generated by the Internet and IoT (Internet of Things) is increasing explosively. Furthermore, the development of data storage and analysis technologies such as cloud computing and Machine Learning has made it possible to effectively process large amounts of data. Moreover, Artificial Intelligence and Big Data are in a mutually complementary relationship, and as they develop together, they are leading to innovation in various industries.

However, since these are expected to have a great impact on our lives from an ethical perspective in the future, this paper fully recognizes that problems may be raised under this research background, and focuses on issues such as the ethical impact of Artificial Intelligence technology on Big Data analysis and utilization.

Recently, the development of data analysis technology based on Artificial Intelligence has not kept up with general ethical standards, so there is a tendency for a gap to arise between them. In particular, the ethical implications of Big Data analysis related to Artificial Intelligence technology and other controversial issues should be carefully reviewed, not only because of the ethical nature of Big Data analysis, but also because it should consider the ethical implications of Artificial Intelligence technology on the daily lives of the public and the growth and innovation of the national economy.[1]

Consequently, if Artificial Intelligence technology can be handled well, the technology will no longer be a threat to ethical implications, but rather a good turning point that turns crises into opportunities. Therefore, this paper focused on analyzing the various issues and ethical implications that may arise due to Artificial Intelligence technology in the future, and suggested improvement tasks such as legal regulations, technical solutions, and ethical standards.

2. The relationship between Artificial Intelligence and Big Data

2.1 Ethical Issues in a Big Data Environment

2.1.1 Current situation

The task of selecting and combining information appropriately is moving toward being solved in the 21st century. In addition, the development of Big Data technology that draws in a huge amount of information and processes it quickly has made all of this possible, thanks to

information storage and processing technology such as Big Data.

Big Data plays a very important role in strengthening the learning ability of Artificial Intelligence, and they are closely connected. Therefore, despite expectations and concerns about its importance and influence as an inevitable huge trend in today's information society, countries around the world are promoting policies to activate Big Data.

The development of the Big Data industry inevitably brings with it the issue of personal information protection, which creates a regulatory dilemma where both the leakage of personal information and excessive protection are problematic. In other words, in order to activate Big Data, it is necessary to have a legal environment where personal information is appropriately protected while also allowing such information to be freely utilized.

The development and activation of the Big Data industry bring with it the issue of personal information protection, which creates a regulatory dilemma where both the leakage of personal information and excessive protection are problematic. Strengthening formal prior regulations such as the current Personal Information Protection Act in Korea may only result in weakening the competitiveness of Big Data-related industries without providing actual personal information protection.

2.1.2 Ethical Issues and Privacy Issues in Big Data Environments

2.1.2.1 Big Data Environment and Ethical Issues

Key ethical issues in the Big Data environment include the fairness of data, the reliability of data sources, the responsibility for predictions and decisions, and ethical issues such as automation and jobs. If data is biased, the analysis results may be disadvantageous to a certain group. This risks deepening social inequality. There are also issues regarding the method of collecting data and whether the source is trustworthy. This is because incorrect data can lead to incorrect conclusions.

In particular, when making decisions through Big Data analysis, there is a need for discussion on who is responsible for the results of data analysis. In addition, as automation progresses due to the development of Big Data and AI, social and ethical issues such as job losses may arise. These ethical issues are becoming more important as Big Data technology develops, and it is necessary to discuss them and find solutions.

2.1.2.2 Big Data Utilization and Personal Information Protection Issues

Under the current law, personal information includes not only personally identifiable information but also non-identifiable information that cannot identify a specific individual by itself but can be easily combined with other information to identify the individual. The information utilized by Big Data is mainly non-identifiable information, and if such information has the potential to identify an individual, it becomes subject to protection under the Personal Information Protection Act. In principle, the consent of the Information Subject must be obtained in order to collect and use it.

Even if consent is obtained, only the minimum information must be collected within the necessary scope, and if it is to be used beyond the purpose of collection, separate consent must be obtained from the Information Subject, and personal information that has passed the retention period or achieved its purpose must be destroyed without delay.

However, Big Data is not simply a collection of data, but there is a significant possibility that meaningful information will be extracted from the data and used for purposes that were not thought of at the time of personal information collection. If the Personal Information Protection Act is applied to such Big Data as it is, the problem of obtaining consent from the Information Subject not only at the stage of collecting the information but also at every stage of analyzing it may arise. In addition, if the determination of personal information is made based on the unclear standard of identifiability, the possibility of legal disputes arising due to the Information Subject's reluctance to provide the information itself or the claim of personal information leakage will greatly increase.

In this regard, case law states that identifiability does not presuppose that all other information that can be combined with the relevant information is held by the same person, and does not mean that other information can be easily obtained, but that it means that the relevant information and other information can be easily combined without special difficulty to identify a specific individual, regardless of whether it is easy or difficult to obtain.

However, determining the possibility of personal identification due to easy combination with other data is profiling used in the Big Data processing process. Or, it will be a very difficult issue for lawyers who make normative judgments as to how to legally understand the technical difficulty of data mining technology.

Meanwhile, consumers have no idea whether their activities are being monitored online unless they are notified

of the existence of the advertising network of the website they visited and that data is being collected. In addition, since the processing of Big Data is fundamentally about the connectivity between data, even if the information does not have personal identification at the data collection stage, there is a significant possibility that specific individuals will be identified later during the data processing process. In this case, it will also be a problem how to protect personal information and what measures to take.

Therefore, the development and activation of the Big Data industry are accompanied by issues of personal information protection, and both the leakage of personal information and excessive protection are causing a regulatory dilemma. It should be noted that the strengthening of formal pre-regulation, such as the current Personal Information Protection Act in Korea, can only weaken the competitiveness of big data-related industries without substantial personal information protection.

Accordingly, some argue that the concept of personal information should be reduced, that is, that non-identification information should be treated differently from personal identification information, and the level of protection should be different or excluded from the object of protection at all.[2]

Therefore, it is judged more effective to ease regulations on information with a low level of personal information protection so that it can be easily used in the Big Data industry, but strictly establish the obligation to prevent abuse or infringement of information collectors or information processors for such information. This is because most Information Subjects do not check the personal information processing policy, it is very rare for them to carefully check the personal information consent form, and almost no individuals in the Big Data field can accurately understand the personal information processing policy and other related matters and consent to the provision of information to third parties.

Personally thinking, in a situation where information is being transferred and utilized between countries, the current Personal Information Protection Act focuses on protecting personal information from infringement and misuse of personal information by domestic companies. In the future, it is necessary to establish relevant regulations on information collection and infringement by global companies and devise measures to protect personal information in practice.[3]

2.1.3 The relationship between Artificial Intelligence and Big Data

Artificial Intelligence (AI) uses Big Data to perform pattern recognition, predictive modeling, and decision support. Big Data includes a large amount of diverse data, and AI analyzes this data to extract meaningful insights. Machine Learning algorithms learn patterns from learning data and use this to classify or predict new data. In this way, AI and Big Data can be seen as complementary.[4]

AI is a technology with human-like thinking abilities, and Big Data is a technology that analyzes various data. However, Big Data is essential for AI to make excellent decisions, and it is also important to have excellent computational technology to analyze Big Data and process it into meaningful information.

Therefore, Big Data is an inevitable huge trend in the current information society, and countries around the world are promoting policies to activate Big Data amid expectations and concerns about its importance and influence. In order to create new benefits and values through the use of Big Data, a large amount of data must be accumulated and integrated through various digital devices of AI, and it is true that the possibility of personal information leakage and abuse increases significantly.

2.2 Case study on Big Data analysis and utilization of Artificial Intelligence

2.2.1 Data collection and utilization, personalized customized content recommendations

Here is a real-world example of the combination of AI and Big Data: Netflix's recommendation system. Netflix collects a vast amount of data, including the content users watch, their ratings, search history, and viewing time. This data is used to understand and analyze individual user preferences, thereby improving user experience and providing personalized content recommendations.

Netflix records what content users watch, how often, and at what time, and collects information such as keywords and genres that users search for within the platform. This is used to identify user interests and recommend related content. It also stores a list of movies or TV shows that users have watched, which is used to improve the personalized recommendation system and suggest new content similar to content that users previously liked.

Netflix also conducts surveys asking users to rate or provide feedback on content. This feedback is used to improve content production and recommendation algorithms. By monitoring user activity or trends on social media platforms, it can analyze popular content or user

preferences. Through these data collection methods, Netflix provides personalized user experience, increases the accuracy of content recommendations, and improves the overall quality of its services. Compliance with privacy policies is also an important factor in the process of data collection and utilization.

2.2.2 Utilizing recommendation algorithms, analyzing various contents

Netflix uses several recommendation algorithms. The main algorithms are as follows: First, collaborative filtering is a method of making recommendations based on the behavior of users with similar tastes. For example, if A and B like similar movies, A's favorite movie can be recommended to B. Second, content-based filtering is a method of recommending similar content by analyzing the characteristics of content that a user has previously watched. For example, if a user watches a lot of romance movies, other romance movies of a similar genre can be recommended. Third, Netflix uses Deep Learning models to analyze complex patterns and provide more sophisticated recommendations to users. For example, it can make recommendations by considering various factors related to a specific genre of movies, such as actors, directors, and topics.

2.2.3 Utilizing real-time learning system, providing customized content

Netflix's recommendation system continuously improves its recommendation system by reflecting user behavior data in real time. This system adjusts the recommendation algorithm immediately as users watch or rate new content, providing a more personalized experience. Netflix collects data in real time, such as what content users watch, how long they watch it, and what content they skip, and analyzes user preferences based on this. It uses Machine Learning algorithms to identify users' preferences and compare patterns with other users with similar tastes.

The analyzed data is used to operate the recommendation algorithm, which provides personalized content recommendations to encourage users to spend more time on the platform. When users watch recommended content, feedback is reflected back into the system, continuously improving the algorithm. This real-time learning system allows Netflix to provide customized content to users, which plays a vital role in increasing user satisfaction and subscriber retention.

2.2.4 Use A/B group testing to continuously improve services

A/B group testing is a controlled experiment where observation data is divided into groups A and B, that is, the control group and the experimental group. The control group is divided into customers who received the service before changing the setting elements, and the experimental group is divided into customers who received the service after the change. Through this, the behavior of the control group and the experimental group is observed, hypothesis testing is performed based on the observed data, and the differences between the two groups are statistically analyzed.[5]

Netflix uses A/B testing to verify the effectiveness of its recommendation system. It experiments with various recommendation methods to analyze which method is most effective for users. Then, it collects user behavior data such as the number of clicks, playback time, and bounce rate from each group.

It analyzes the collected data to compare the performance of the two groups, and uses statistical methods to evaluate the significance of the results. Based on the results of the A/B test, the optimal option is selected, and successful changes are applied to all users. This plays an important role in Netflix providing customized user experiences and continuously improving its services.

2.2.5 Discussion

In short, Netflix's recommendation system combines Artificial Intelligence and Big Data to provide customized content to users, thereby maximizing user experience and increasing customer retention. Thanks to this system, Netflix has gained the competitiveness to retain many users worldwide.

3. Ethical Issues Arising from Big Data Analysis and Utilization of Artificial Intelligence Technology

3.1 Side Effects and Risk Factors

3.1.1 Side effects due to misuse

AI can be misused to create forgeries, crimes, and fake news. In fact, it can be exploited politically and spread false information, so it seems that a humanistic solution should be prepared for this area. In fact, AI that has been trained differently based on skin color and gender has developed

incorrect perceptions. In order to prevent the misuse of technology, it should also be prepared for the development of AI that can identify fake news or images. If it is inevitable in the age of AI, it is necessary to coexist, be ready for it, and prepare for the future.

In this way, AI can bring incalculable advantages and value to humans, but on the other hand, it can also cause privacy and personal information infringement issues. Apple is struggling between protecting customer privacy and developing AI, and has promoted the development of AI technology in accordance with its customer privacy protection policy.

Apple CEO Tim Cook presented a privacy policy on the company's website in September 2014, stating, "Apple is clearly opposed to marketing using user information." Therefore, Apple has no choice but to analyze data centered on iPhone data rather than iCloud, so Apple's technical challenge is to figure out how to increase user convenience with AI while protecting customer privacy.

In a survey of respondents who were either AI professionals or non-professionals regarding possible side effects of AI, respondents expressed greater concerns about side effects caused by AI misuse or malfunction than about job losses due to the introduction of AI. The respondents were found to have a great fear that AI could be misused for international crimes such as drugs or terrorism or that errors in algorithms that perform important functions of AI could lead to serious disasters.

The subjects of the survey were the Industry Analysis Team of Institute of Information & Communications Technology Planning & Evaluation (IITP), which analyzed the benefits and side effects that AI could bring. The significance of this survey lies in the fact that it was conducted to diagnose and secure basic data on how our country should develop and utilize Artificial Intelligence in the future.

3.1.2 Risk factors for security management systems

Since Artificial Intelligence is based on computer systems, even if it utilizes cutting-edge Artificial Intelligence technology, it is not safe from various hacking threats such as ransomware. For example, on July 21, 2015, an SUV traveling at 70 miles per hour on the St. Louis Highway in the United States suddenly stopped all operations and fell into a ditch on the side of the road.

The analysis of the cause of the accident revealed that the electronic control system of the SUV in question was hacked remotely 10 miles away from the accident site. The hackers infiltrated hacking programs to remotely control the steering wheel, and used the existing adaptive automatic cruise control to reduce the speed while simultaneously using the lane departure prevention function to change the direction of the car. After that, they made the accident prevention radar, which continuously sounds a warning sound when a dangerous object approaches, ignore the electronic safety control function and the ABS (Anti-lock Braking System), causing the car to fall into a ditch. Therefore, as seen in these cases, the possibility of hacking into the autonomous vehicle system is certainly there, as various wired and wireless lines are intertwined to control the autonomous vehicle.

3.1.3 Discussion

In short, recent Artificial Intelligence such as AlphaGo has shown amazing performance by utilizing the Internet of Things, Big Data, and Deep Learning, but it cannot respond to completely new variables such as changes in the rules of the game, and it has shown unpredictable behavior in cases where it has not been learned or cannot be inferred. It is true that Artificial Intelligence has technical limitations and risk factors because it has not been able to understand and express the philosophical and subjective concepts of human society to date.

In this way, the limitations of Artificial Intelligence and the risk factors for side effects were considered. If humans, who can critically and rationally view any fact or situation and overturn existing knowledge and discover new facts, can quickly analyze and infer specific facts using the excellent tool called Artificial Intelligence, they will be able to appropriately respond to changes such as the '4th Industrial Revolution' in the future.

3.2 Ethical Issues

3.2.1 Invasion of privacy

The collected data may be used for purposes other than the original purpose. Data collected for marketing purposes may be misused to evaluate an individual's credit. In addition, user data may be collected or analyzed without permission, which may infringe on an individual's privacy, and if sensitive information is leaked, it may cause serious damage to the individual. In the process of analyzing and predicting an individual's behavior or choices, location data or Internet usage history may be excessively collected.

3.2.2 Illegal surveillance

When AI technology is used for surveillance purposes, it can infringe on the freedom and rights of individuals. In particular, surveillance in public places can raise ethical issues. When AI manipulates or interferes with an individual's choices, it can infringe on the autonomy of the individual. For example, a recommendation system can force a specific product or opinion. In order to solve these problems, legal and social standards related to data ethics should be established, and transparent data use and AI development should be carried out.

3.2.3 Information asymmetry

There is an information asymmetry between data collectors and users, which may lead users to not understand how their data is processed.

3.2.3.1 Differences in information ownership

Data collectors have a variety of data collected from users and can analyze it to derive insights. On the other hand, users may lack information about how their data is collected and utilized.

3.2.3.2 Imbalance in data utilization

Data collectors can provide customized services based on the collected information and generate revenue through this. However, users may lack a clear understanding of how their data is used and for what purpose.

3.2.3.3 Reliability issues

Information asymmetry can create trust issues between users and data collectors. Users may be concerned about whether their data is being managed safely or whether it could be used illegally.

3.2.3.4 Asymmetry of decision-making power

While data collectors can have a decisive influence on marketing strategies or product development based on the results of data analysis, users may have limited choices about how their data is used. In short, this information asymmetry complicates the relationship between data collectors and users, and there is an ongoing debate about the importance of providing users with more transparency and choice.

3.2.4 Data security threats

The collected data may be at risk of hacking or leakage, which may result in personal information being exposed to outside parties.

3.2.4.1 Data breach threat

As AI systems process large amounts of data, security vulnerabilities can arise. There is a greater risk that hackers will infiltrate the system and leak personal or sensitive data. This can result in a violation of personal privacy and a loss of trust.

3.2.4.2 Unauthorized access

As AI technology analyzes data, there is a possibility that users or systems without access to the data may access the information. Such unauthorized access threatens the safety of the data and, if exploited, can lead to serious consequences.

3.2.4.3 The Complexity of Data Management

The AI models used in Big Data analytics are often complex and opaque. This makes it difficult to clearly understand how data is collected and stored, which makes security management difficult. Uncertainty in data management poses a challenge to complying with data protection regulations.

3.2.4.4 The problem of user consent

There may be cases where the user's explicit consent is not obtained during the data collection process. When AI analyzes and utilizes data, data security issues may arise if the user does not understand how their data is being used.

3.2.4.5 Ethical Responsibility

Companies or organizations that use AI technology have ethical responsibilities related to data security. They need to consider how they will be held accountable in the event of a data leak or misuse. In short, these issues are becoming more complex as Big Data analysis and AI technology advance, and ethical considerations are needed.

3.3 Bias problem

3.3.1 Employment discrimination

If the dataset that the AI model learns from is biased, unfair decisions or predictions can be made as a result.[6] If the AI model excessively reflects data from a specific

group, prejudice against that group can be strengthened. The bias of AI algorithms can cause various social problems, and can be thought of largely focusing on the effects related to racism and gender discrimination.[7]

If an AI-based recruitment system learns based on data biased against a specific race or gender, it can evaluate or exclude applicants from that group unfavorably. Therefore, this can potentially lead to an imbalance in opportunities. This bias problem, which can lead to employment discrimination, will mainly occur due to the following factors.

3.3.1.1 Data Bias

If the training data is imbalanced against certain groups (e.g. gender, race, age, etc.), AI models will make predictions based on this data, which may reinforce unfavorable judgments or evaluations of certain groups.

3.3.1.2 Outcome Bias

If past hiring decisions are already biased, models trained on this data will likely repeat the same bias. For example, if applicants of a certain race or gender were not selected in the past, the model is more likely to evaluate this group lower.

3.3.1.3 Feature Selection Bias

If an AI model emphasizes or ignores certain characteristics, groups related to those characteristics may be discriminated against. For example, if characteristics such as education or career are overemphasized, other important factors may be overlooked, which may disadvantage applicants from certain groups.

3.3.1.4 Opaqueness of the Decision Process

If the decision process of an AI model is opaque, employers may rely on the model's predictions, which may lead to unreasonable discrimination. This may result in unfair decisions made by simply accepting the model's results.

In short, to address these issues, it is important to build datasets that consider diversity and inclusiveness, regularly review the model's results, and apply algorithmic approaches to reduce bias. In addition, transparency of the AI's decision process should be increased.

3.3.2 Crime Prediction

If a crime prediction algorithm is based on historical crime data, it may reinforce bias toward certain races or regions. This could lead to excessive police surveillance or crackdowns on certain groups. If the dataset that the AI model learns from is biased, bias can appear in crime prediction in the following ways:

3.3.2.1 Data Bias

Crime data is based on past crime records for certain regions or groups. This can lead to over- or under-representation of groups with certain races, genders, or socioeconomic backgrounds. For example, if a certain area has a high crime rate due to intensive police crackdowns, it may incorrectly predict that residents of that area are more likely to commit crimes.

3.3.2.2 Outcome Bias

As an AI model learns past crime data, it repeats decisions that are already biased. For example, if a certain race has a high rate of crimes in the past, the model may evaluate that race as more likely to commit crimes based on this. This can exacerbate discrimination.

3.3.2.3 Overgeneralization

If an AI model becomes overly reliant on specific characteristics or patterns, it will ignore the characteristics of individual cases or individuals. This may lead to an unreasonably high evaluation of the likelihood of individuals belonging to a certain group committing a crime.

3.3.2.4 Opaqueness of the Decision Process

If it is unclear how the AI prediction results were derived, the reliability of the crime prediction may decrease and lead to incorrect judgments. This may lead law enforcement agencies to uncritically accept the results of the AI.

3.3.2.5 Social instability

Biased crime prediction models may lead to distrust and social conflicts toward certain groups. This may have a negative impact on crime prevention and social integration.

In short, to solve these problems, it is necessary to compose datasets in a diverse and comprehensive manner and apply various techniques to reduce the bias of the algorithm. In addition, it is important to increase transparency of the prediction results and ensure that

judgments are made considering the social context.

3.3.3 Credit Rating

If there is a bias against a specific race or gender in the credit rating process in which AI assigns credit scores, unfair loan terms or credit restrictions may occur.

3.3.3.1 Data Bias

Credit rating models learn based on various data such as past credit history and loan repayment history. If a group with a specific race, gender, or socioeconomic background is not sufficiently reflected or is evaluated negatively, the credit score of that group may be underestimated.

3.3.3.2 Outcome Bias

If past credit ratings are already biased, the AI model may learn based on this data and repeat the same bias. For example, if a specific group is not able to receive a loan or is charged a high interest rate, the credit score of that group may be continuously evaluated low.

3.3.3.3 Feature Selection Bias

Variables used in credit ratings may work against a specific group. For example, young people or immigrants without credit history may receive low credit scores because of their lack of credit history. This situation may limit their access to credit.

3.3.3.4 Opaqueness of the decision-making process

If the decision-making process of an AI model is opaque, it may be difficult to understand and appeal the credit rating results. This makes it difficult for consumers to complain or appeal their credit scores.

3.3.3.5 Social inequality

Biased credit rating models can disadvantage certain groups and exacerbate social inequality. This leads to discrimination in access to financial services, which limits economic opportunities.

In short, to solve these problems, it is necessary to build a dataset that reflects diverse population groups and monitor and correct bias in the algorithm. In addition, it is important to increase the transparency of credit ratings and establish a system that allows consumers to understand and appeal their credit scores.

3.3.4 Accessibility to Healthcare Services

If AI-based diagnostic systems lack data on a specific group, incorrect diagnoses or treatment methods may be selected, which may exacerbate health inequalities. If the datasets that AI models learn from are biased, this may have various negative effects on accessibility to healthcare services.

3.3.4.1 Data Bias

If medical data is collected disproportionately for a specific population group (e.g., gender, race, region), AI models learn based on this data. For example, if medical records for a specific race are lacking, the diagnosis or treatment outcomes for that race may be inaccurately evaluated.

3.3.4.2 Diagnosis and Treatment Bias

AI models may underestimate or misinterpret symptoms or diseases of a specific group. This may prevent patients of a specific race or gender from receiving the diagnosis or treatment they need. For example, heart disease symptoms in women may present differently than in men, and failure to reflect these differences may delay the diagnosis of female patients.

3.3.4.3 Inequity in the distribution of medical resources

If AI models underestimate the medical needs of certain groups, medical resources or services may be distributed insufficiently for those groups. This has a greater impact on patients from low-income or underprivileged communities.

3.3.4.4 Inequity in Access to Healthcare

Biased AI systems can impede access to healthcare services. If certain groups are discriminated against in AI-based healthcare services, they may not receive the necessary treatment or may be delayed.

3.3.4.5 Decreased Trust

If the decision-making process of an AI system is opaque, patients may lose trust in the system. This may lead to disregarding AI-based diagnoses or treatment recommendations, and weaken trust between healthcare providers and patients.

In short, to address these issues, it is important to build

a comprehensive dataset that reflects diverse populations and continuously monitor and correct model biases. In addition, policies and systems should be established to increase accessibility to healthcare services and ensure that all patients receive fair treatment.

3.3.5 Social Media and Information Filtering

When algorithms filter information based on personal interests, there is a possibility that the voices of certain groups will be underrepresented or distorted. This can lead to social conflict. Bias in social media and information filtering can manifest itself in the following ways:

3.3.5.1 Information Bias

When AI algorithms learn to prefer certain types of content or opinions, the information provided to users is limited. For example, if content with certain political views or social ideologies is overemphasized, users lose the opportunity to access other perspectives or information.

3.3.5.2 Filtering Effect

Information filtering on social media platforms is based on user preferences. If the learning data is biased, users will only be exposed to information that matches their own, which can lead to a 'mirror room of information' phenomenon, reducing exposure to diverse opinions or information.

3.3.5.3 Social Divide

If biased information about a certain group is continuously exposed, conflict or misunderstanding between that group and other groups can be exacerbated. This can lead to social division and undermine the integrity of the community.

3.3.5.4 Spread of misinformation

Biased algorithms can spread misinformation or conspiracy theories more. As information favorable to a certain group is emphasized, the risk of spreading untrue content to users increases.

3.3.5.5 Impact on identity formation

Information that users encounter on social media has a significant impact on their identity and value formation. Biased information can distort users' worldview, which can have a negative impact on their personal behavior or social attitudes.

In short, to solve these problems, it is necessary to use a comprehensive dataset that includes data from various sources and continuously monitor the bias of the algorithm. In addition, it is important to build a system that provides users with diverse perspectives and prevents the spread of misinformation through fact-checking. Social media platforms need to adjust their algorithms so that users can access diverse opinions and make efforts to maintain a balance of information. In addition, efforts are needed to increase the transparency and fairness of the algorithm and minimize bias by utilizing various data sources.

3.4 Lack of Transparency

The following problems can arise when the decision-making process of AI is opaque:[8]

3.4.1 Lack of Trust

If users do not understand the decision-making process or reasons of AI, their trust in the results may decrease. This can hinder users from adopting AI systems. The problem of lack of trust that can arise when the decision-making process of AI is opaque appears in various aspects. The main contents are as follows:

3.4.1.1 Difficulty in interpreting results

If users cannot understand the results of AI systems, their trust in the results decreases. For example, if the basis for AI's judgment is unclear in important decisions such as medical diagnosis or credit rating, users may have difficulty accepting the results.

3.4.1.2 Lack of Consistency in Decisions

When it is unclear how an AI model made a decision, its decisions in similar situations may be inconsistent. Users will lose trust in AI as they experience this inconsistency.

3.4.1.3 Concerns about bias and discrimination

If the decision-making process of AI is opaque, users may suspect that the model is biased or may make discriminatory decisions against certain groups. This is especially true in situations where discrimination based on race, gender, or age is a concern.

3.4.1.4 Uncertainty of responsibility

If AI makes a decision incorrectly, it may be unclear who is responsible for it. Users may not know how to respond when AI makes a wrong judgment, which may lead to distrust in the system.

3.4.1.5 Impaired user participation

Lack of trust prevents users from actively utilizing AI systems. For example, even if a company introduces AI-based decision-making, employees may be reluctant to participate in the process or provide opinions. This may result in hindering collaboration and innovation within the organization.

In short, to solve these problems, it is important to make the decision-making process of AI transparent and help users understand the process. For example, it is necessary to introduce Explainable AI (XAI) technology to clarify the reasons for decision-making and build trust with users.

3.4.2 Responsibility Issues

If the entity responsible for decisions made by AI becomes unclear, responsibility for wrong decisions may become ambiguous. This may lead to legal and ethical issues. The main issues of responsibility that may arise when the decision-making process of AI is unclear are as follows:

3.4.2.1 Uncertainty of Responsibility

If an AI system makes a wrong decision, it is unclear who is responsible. For example, if a medical AI makes a wrong diagnosis, there may be controversy over who should be held responsible among the medical provider, the AI developer, or the data provider.

3.4.2.2 Legal Responsibility Issues

There is a problem of how to determine legal responsibility for damages caused by AI decisions. Since the current legal system generally determines responsibility for human actions, it may be difficult to respond sufficiently with the existing legal framework in situations where AI is involved.

3.4.2.3 Lack of Transparency

If the decision-making process of AI is opaque, it is difficult for external parties to review or object to the decision. This will also be a major obstacle in the process of holding those responsible for wrong decisions accountable.

3.4.2.4 Decreased Trust

If responsibility is unclear, users and society will lose trust in AI systems. This may hinder the introduction of AI technology and hinder innovation and development.

3.4.2.5 Ethical Issues

When AI decisions affect human life or safety, the issue of responsibility becomes more serious. For example, ethical discussions will be needed on who should be held responsible when a self-driving car causes an accident.

In short, in order to solve these responsibility issues, it is important to increase the transparency of AI systems and develop legal and ethical frameworks that can clearly identify the location of responsibility. There is a need for a mechanism that can understand how AI's decision-making process takes place and clarify responsibility for it.

3.4.3 Difficulty in risk management

If the AI's decision-making process is opaque, it becomes difficult to predict and prevent system errors or failures in advance. This can cause safety-related problems. The difficulty in risk management that can occur when the AI's decision-making process is opaque appears in various aspects. The main contents can be said to be as follows.

3.4.3.1 Difficulty in risk identification

If it is unclear how the AI system makes decisions, it is difficult to identify potential risk factors. For example, if it is not known what data a specific algorithm uses to make decisions, it will be difficult to predict the negative consequences of that decision.

3.4.3.2 Absence of risk assessment

If the decision-making process is opaque, it is difficult to evaluate the results of AI's decisions. This will be a major obstacle for organizations or individuals to properly assess and manage risks that may arise from the use of AI systems.

3.4.3.3 Difficulty in establishing response strategies

An opaque AI decision-making process makes it difficult to establish an appropriate response strategy when a problem occurs. For example, if AI makes an incorrect recommendation, it will be difficult to devise effective corrections or improvement measures if the cause cannot be

identified.

3.4.3.4 Difficulty in building trust

If there is a lack of trust in AI systems, users or organizations will have doubts about whether the system can be used safely. This will make them hesitant to introduce AI, and ultimately hinder the development and use of technology.

3.4.3.5 Regulatory and compliance issues

If the decision-making process of AI is opaque, it may be difficult to comply with legal regulations or ethical standards. This may be a major risk factor for companies to assume legal responsibility.

3.4.3.6 Inefficiency in disaster management

When AI systems make decisions in disaster management or emergency situations, if the process is opaque, it may be difficult to respond quickly and effectively. This may increase casualties or property damage.

In short, to solve these difficulties in risk management, it is essential to increase the transparency of AI systems and strengthen the explainability of the decision-making process. In addition, it is very important to establish a risk management framework to establish a system that can identify and evaluate potential risks due to the use of AI in advance.

3.4.4 Conflict of Interest

If the decision-making process is opaque, there is a high possibility that companies or organizations will use AI to manipulate or make distorted decisions based on their interests. Conflicts of interest that may arise when the decision-making process of AI is opaque can appear in various aspects. The main contents are as follows.

3.4.4.1 Disagreement between stakeholders

If the decision made by the AI system does not meet the interests of specific stakeholders (e.g., companies, consumers, government, etc.), conflicts may arise. For example, if an AI model that prioritizes the interests of a company makes a decision that overlooks the safety or welfare of consumers, the interests of the two groups will conflict.

3.4.4.2 Lack of Trust Due to Lack of Transparency

If the decision-making process of AI is opaque, stakeholders may question the fairness or objectivity of the system. This may cause each stakeholder to oppose or distrust AI decisions in order to protect their own interests.

3.4.4.3 Conflict of Ethical Standards

If AI decisions conflict with the ethical standards or values of stakeholders, conflicts may arise. For example, if the method of data collection and utilization is judged to infringe on an individual's privacy, there may be a conflict of interest between the individual and the company.

3.4.4.4 Unfairness of Decision-making

If the AI system favors a specific stakeholder, other stakeholders will be disadvantaged. This may lead to dissatisfaction or conflict, which may negatively affect collaboration or communication within the organization.

3.4.4.5 Unclear Responsibility

If AI decision-making is unclear, conflicts may arise because it is unclear who is responsible for a wrong decision or result. Stakeholders may try to avoid responsibility or blame each other.

3.4.4.6 Confusion in policy and regulation

If AI decisions cause conflicts between stakeholders, it may be difficult for governments or regulatory agencies to establish appropriate policies. This may lead to social confusion or conflict, which may hinder the effectiveness of policy implementation.

In short, in order to resolve such conflicts of interest, it is important to make the decision-making process of AI systems transparent and establish a system for collecting opinions from stakeholders. In addition, AI development and operation that considers fairness and ethics are necessary, and efforts are required to promote communication and cooperation between stakeholders. In particular, it is important to make the decision-making process of AI transparent and develop a system that can be explained to users.

4. Solutions and Improvements

4.1 Legal Solutions[9]

4.1.1 Strengthening the Rights of Information Subjects

It is necessary to clarify the rights of individuals to request access to, correction, deletion, and restriction of processing of their data, and simplify the procedures to make it easier to exercise these rights.

4.1.1.1 Right of Access of Information Subjects

Information Subjects have the right to check how their personal data is being processed. This allows individuals to demand transparency regarding their data.

4.1.1.2 Right to Correction

Information Subjects have the right to have their personal data corrected if it is inaccurate or incomplete. This will play an important role in preventing damage caused by incorrect information.

4.1.1.3 Right to erasure(Right to be forgotten)

Information Subjects have the right to have their personal data deleted under certain conditions. This will greatly contribute to strengthening the protection of individuals' privacy.

4.1.1.4 Right to Restrict Processing

Information Subjects have the right to restrict the processing of their personal data. For example, they can request that the processing of their data be suspended while the accuracy of the data is verified.

4.1.1.5 Right to data portability

Information Subjects have the right to transfer their personal information to other service providers. This should strengthen the ownership of data and allow individuals to manage their data more freely.

4.1.1.6 Right to automated decision-making

Information Subjects have the right to object to decisions made in an automated manner. This is important for ensuring fair processing.

In conclusion, these rights will greatly help Information Subjects to strengthen their control over their personal information and prevent unfairness in the data processing process. In addition, these provisions of the Personal Information Protection Act contribute to respecting individual rights and clarifying corporate responsibilities, thereby increasing the transparency of data use.

4.1.2 Increased transparency of data processing

Legal obligations should be imposed on companies and organizations to increase transparency in the way they collect and process personal information, so that Information Subjects can clearly understand the use of their data. Increased transparency of data processing is an important element of personal information protection and data use, allowing Information Subjects to understand how their personal information is collected, processed, and stored.

4.1.2.1 Clear Privacy Policy

Companies or organizations should write their privacy policy in clear and easily understandable language. This policy should include the purpose of data collection, processing method, retention period, and whether or not data will be provided to third parties.

4.1.2.2 Obligation to provide information

When collecting data, necessary information should be provided to the Information Subject. This should be made clear at the time of data collection, and users should be able to know how their data will be used.

4.1.2.3 Clarity of consent

Consent for the collection and processing of personal data should be voluntary and clear, and users should be provided with an easy way to withdraw consent. This should help users exercise their rights.

4.1.2.4 Recording of data processing history

Companies should record and manage the details of personal data processing. This can be used to prove the lawfulness of data processing, and the information should be available to the Information Subject upon request.

4.1.2.5 Regular audits and reviews

Regular audits and reviews of the data processing process should be conducted to check compliance with the privacy policy. This will contribute to increasing the transparency of

the organization.

4.1.2.6 Education and awareness raising

It is important to provide personal information protection education to employees to raise awareness of the importance of data processing and related laws. This will also help create a transparent data processing culture internally.

In conclusion, these measures will contribute to individuals feeling a sense of control over their own data and building trust in the data processing process. Therefore, transparent data processing will play an important role in increasing the reliability of the company and fulfilling social responsibility for personal information protection.

4.1.3 Strict Consent Request

Consent for the collection and processing of personal information should be more strictly required, and regulations should be put in place to prevent data collection without prior consent. Strict consent requirement is a measure to require clear and specific consent before collecting, processing, and using personal information. This aims to ensure that individuals fully understand and agree to how their information will be used.

4.1.3.1 Legal Basis

In many countries, data protection laws require that companies or organizations obtain explicit consent to collect personal data. For example, the General Data Protection Regulation (GDPR) strictly requires consent from Information Subjects, and requires that consent be freely given, specific, unambiguous, and informed.[10]

4.1.3.2 Requirements for Consent

Consent should be clearly stated in understandable language, and individuals should be able to clearly understand what they are consenting to. Individuals should also be able to give consent freely, without coercion or disadvantage, and they should have the right to withdraw their consent.

4.1.3.3 Specificity

Consent should only be given for data collected for specific purposes, and broad consent is not permitted.

4.1.3.4 Impact

Such strict consent requirements will require companies to take a more transparent and responsible approach to data collection and processing. In addition, individuals will be able to better understand and exercise their rights regarding their data.

In conclusion, strict consent requirements are an important measure to strengthen personal information protection, contributing to protecting individual rights and increasing transparency in data use.

4.1.4 Strengthening penalties for data breaches

Legal penalties for personal information leaks or breaches should be strengthened to encourage companies or organizations to take a more responsible approach to data protection. A data breach is a situation where an individual's personal information is accessed, leaked, altered, or destroyed without authorization. In the event of such a breach, it is necessary to strengthen the penalties for companies or organizations according to relevant laws.

4.1.4.1 Legal Basis

Many countries' personal information protection laws specify penalties for data breaches. For example, under GDPR, companies can face significant financial penalties in the event of a data breach, with fines of up to 20 million EUR or 4% of annual global turnover, whichever is higher.

4.1.4.2 Financial fines

In addition to compensation for damages caused by a data breach, high fines may be imposed depending on the legal standard.

4.1.4.3 Criminal penalties

In the case of a serious data breach, criminal penalties may be imposed on the person responsible. This may include imprisonment or fines.

4.1.4.4 Administrative sanctions

Administrative sanctions such as suspension of operations or restrictions on data processing may be imposed on companies that have committed a data breach.

4.1.4.5 Impacts

Strengthening penalties will make companies more mindful of data protection and strengthen their responsibility for protecting personal information. In addition, individuals

will be able to have greater confidence that their information is protected more safely.

In conclusion, strengthening the punishment for data breaches is an important measure to emphasize the importance of personal information protection and strengthen corporate responsibility. This will ultimately contribute to protecting individual rights and making the data environment safe.

4.1.5 Mandatory Appointment of Data Protection Officer

Companies of a certain size or larger should be required to appoint a data protection officer to ensure a professional management system for personal information protection. One measure to strengthen the Personal Information Protection Act and legal regulations on data use is the mandatory appointment of a Data Protection Officer (DPO). This measure includes several important aspects.

4.1.5.1 Strengthening Responsibility

By designating a Data Protection Officer, responsibility for personal information protection can be clearly established within the organization. The DPO is responsible for supervising data processing activities and leading the organization to comply with personal information protection-related laws and regulations.

4.1.5.2 Legal Compliance Support

The DPO understands laws and regulations related to personal information protection laws and helps the organization comply with them. This should reduce legal risks and provide a foundation for responding quickly when legal issues arise.

4.1.5.3 Regular training and awareness raising

The DPO is responsible for educating employees within the organization on personal data protection and making them aware of the importance of personal data protection. This will help establish a personal data protection culture throughout the organization.

4.1.5.4 Monitoring data processing activities

The DPO monitors the organization's data processing activities and checks whether the data protection policy is being properly implemented. This will help prevent data leaks or personal data breaches in advance.

4.1.5.5 Communication and cooperation

The DPO plays a communication role with external stakeholders (e.g. regulators, customers, etc.). When questions or issues related to data protection arise, the DPO can act as a mediator, thereby building trust.

4.1.5.6 Risk assessment and management

The DPO plays a key role in assessing the risks associated with data processing and establishing measures to manage them. This will help increase the effectiveness of data protection.

4.1.5.7 Increased Transparency

The presence of a DPO can increase transparency about an organization's data processing activities and provide customers and users with confidence that their data is being protected. This can help convey the message that their personal data is being processed safely.

For these reasons, the mandatory designation of a Data Protection Officer can play an important role in strengthening personal data protection and ensuring compliance with legal regulations. This can ultimately contribute to protecting the rights of individuals and increasing the reliability of data processing.

4.1.6 Strengthening International Cooperation

In order to respond to global data flows, it is necessary to establish international data protection standards through cooperation with other countries and ensure compliance with them.

4.1.6.1 Setting Global Standards

Since personal data protection laws and regulations differ in various countries and regions, it is important to set international standards to harmonize the laws of each country. This should facilitate the flow of data across borders and help companies comply with legal requirements in various markets.

4.1.6.2 Information sharing and best practice exchange

International cooperation can be used to share information and cases related to personal information

protection. By cooperating with regulatory agencies, companies, and researchers from each country to exchange effective data protection methods and best practices, the level of personal information protection can be raised worldwide.

4.1.6.3 Establishing a legal cooperation system

It is necessary to establish a legal cooperation system that can effectively respond internationally when a data leak or personal information infringement incident occurs. This can enable a rapid response through legal support and information exchange between countries.

4.1.6.4 Concluding bilateral and multilateral agreements

By concluding bilateral or multilateral agreements for personal information protection, it is possible to unify personal information protection standards among countries and establish regulations for data movement. Such agreements provide an internationally applicable legal framework to strengthen data protection.

4.1.6.5 Forming a network among regulatory agencies

It is important to form a network among personal information protection regulatory agencies from each country to promote mutual cooperation and information exchange. This will enable regulators in each country to share their experiences and work together to address common challenges.

4.1.6.6 Education and Awareness Raising

International cooperation can help raise awareness of privacy protection at a global level by developing education programs on privacy protection and providing them to stakeholders in each country.

4.1.6.7 Technical Cooperation

Joint research and development of data protection technologies and solutions can increase the effectiveness of data protection and promote innovation. This will contribute to strengthening privacy protection along with technological advancements.

In conclusion, these measures to strengthen international cooperation can play an important role in strengthening privacy protection, increasing consistency, ensuring the safe flow of data, and protecting the rights of individuals in a global society.

4.2 Technical Solutions

4.2.1 Explainable AI

This is a technology that allows us to understand the decision-making process of an AI model. With the discovery of new explainable technologies and the establishment of standards for AI technology, solutions can be sought in terms of Artificial Intelligence bias, but efforts to preemptively respond to them from an ethical point of view are needed. Here, Explainable AI (XAI) is a technical solution that allows us to understand and interpret the decision-making process of an AI model.

4.2.1.1 Model Interpretation Method

LIME (Local Interpretable Model-agnostic Explanations) is used to explain the output of a model for a specific prediction. It explains the local behavior of the model by perturbing sample data. SHAP (SHapley Additive exPlanations) numerically evaluates the impact of each feature on the model prediction and visualizes the contribution of each feature.

4.2.1.2 Model Simplification

Instead of complex models, interpretable models such as decision trees and linear regression should be used to make the results easier to understand.

4.2.1.3 Visualization tools

Visualize the decision boundaries, feature importance, etc. of the model to help users intuitively understand the results.

4.2.1.4 Human-centered design

Design a way to explain the results considering the user's level of understanding. For example, it should be explained in terms that non-experts can understand.

4.2.1.5 Feedback loops

By allowing users to provide feedback on the model's predictions, the model's interpretability can be increased and continuously improved. In short, these methods can play an important role in making the decision process and results of the AI model clearer and building a system that users can trust.

4.2.2 Model interpretation tools

Tools that visualize the internal structure of the model or evaluate the importance of the characteristics can be used to analyze how much any input affects the results. Model interpretation tools are technical methods that help users understand and explain the prediction results of an AI model. These tools play an important role in helping users understand the internal workings of the model and increase confidence in the results. The main model interpretation tools include the following:

4.2.2.1 LIME (Local Interpretable Model-agnostic Explanations)

LIME is a method that explains the output of a model for a specific prediction. In order to understand the prediction of the model, samples are generated around the original data points and the model is trained on these samples. Through this, it will be possible to visually represent the effect of each characteristic on the prediction.

4.2.2.2 SHAP (SHapley Additive exPlanations)

SHAP is a method based on game theory that numerically evaluates the degree to which each feature contributes to the prediction of the model. This allows us to quantitatively explain the impact of each feature on the prediction results, and it can be intuitively understood when used with visualization tools.

4.2.2.3 Feature Importance

This is a method to evaluate how important each feature of the model is to the prediction. In random forests or boosting models, characteristic importance can be calculated to easily identify the most influential characteristics.

4.2.2.4 PDP (Partial Dependence Plots)

This is a graph that visually shows the average impact of a specific feature on the model prediction. It can analyze multiple features simultaneously, so it is useful for understanding the interaction between features.

4.2.2.5 ICE (Individual Conditional Expectation) Plots

Similar to PDP, but it shows the prediction change for each individual data point. This will allow for a more detailed analysis of the impact of certain characteristics on model prediction.

In short, these model interpretation tools play an important role in increasing the transparency of the algorithm and enhancing user trust in the AI system. These tools will help users better understand the model's prediction results and improve the model as needed.

4.2.3 Interactive Visualization

Interactive visualization tools can be developed to help users easily understand how the algorithm works and the results by visually expressing them. Interactive visualization is a technical method that plays an important role in increasing the transparency of the algorithm. This method visually expresses data and allows users to directly interact with it, making it easy to understand the information. The main features and advantages of interactive visualization are as follows:

4.2.3.1 Intuitive data representation

It helps users easily understand complex data and model results by visualizing them in graphs, charts, maps, etc. Visual elements clearly reveal patterns and relationships in the information.

4.2.3.2 User interaction

It allows users to select or filter specific data points, allowing them to conduct in-depth analysis on the areas of interest. For example, users can adjust variable values using sliders and check the changes in results in real time.

4.2.3.3 Real-time feedback

Interactive visualization provides immediate feedback to help users understand the model's behavior. The visualization changes dynamically according to the values entered by the user, which allows the user to intuitively experience the model's prediction process.

4.2.3.4 Analysis of various scenarios

Users can simulate various scenarios and compare the results for each scenario. This will allow them to discover the strengths and weaknesses of the model and gain insights necessary for decision-making.

4.2.3.5 Improved comprehension

Interactive visualizations are designed to be easily

understood by non-experts, making them useful for explaining the complex operating principles of AI models. User-friendly interfaces can reduce the learning curve and contribute to building trust in the model.

In short, such interactive visualizations are important tools for increasing the transparency of the algorithm and enhancing communication with users, thereby increasing trust in AI systems. Through this tool, users will be able to understand the relationship between data and models more deeply and make better decisions.

4.2.4 Simplification of algorithms

Using simple models that are easy to interpret instead of complex models can increase transparency. Simplification of algorithms is one of the important technical methods to increase transparency of AI models. This approach should reduce the complexity of the model to make it easier to understand, and as a result, make it easier for users to understand how the model works.

4.2.4.1 Simple model selection

Use interpretable models such as decision trees, linear regression, and logistic regression instead of complex models (e.g., Deep Learning). These models are relatively intuitive, and it will be easy to understand the effect of each feature on the results.

4.2.4.2 Feature selection

This is the process of reducing the number of features (variables) used in the model. Simplifying the model by removing unimportant or redundant features will make the results easier to interpret and reduce the problem of overfitting.

4.2.4.3 Simplification of model structure

Models with complex structures can be difficult to interpret. For example, using a neural network with a small number of layers and neurons is a good example. Keep the model structure simple, and the prediction results will be better understood.

4.2.4.4 Rule-based models

Using a rule-based approach (e.g., decision trees) can clearly explain how each prediction was made. Such models will provide intuitively understandable rules, which will increase the transparency of the predictions.

4.2.4.5 Leveraging model explanation tools

Even in simplified models, model interpretation tools such as LIME and SHAP can be used to explain the impact of each feature on the results. These tools provide additional interpretation even in simple models, helping users better understand the results.

4.2.4.6 Feedback loops

User feedback can be used to continuously improve and simplify the model. By adjusting the model based on the user's understanding, transparency can be further increased.

In short, simplifying the algorithm plays an important role in increasing the interpretability of the model and enabling users to trust the model's prediction results. This will enhance the transparency of the AI system and support better decision-making.

4.2.5 Model Validation and Audit

The process of verifying and auditing the results of an algorithm will ensure that the decisions of the algorithm are fair, consistent, and understandable. Model validation and auditing are important technical methods to increase the transparency of the algorithm and are essential to ensuring the reliability and accuracy of the AI model. This process helps to ensure that the model is operating correctly and that the results are fair. The main elements of model validation and auditing are as follows:

4.2.5.1 Model Validation

Performance Evaluation is to evaluate how well the model works based on real data. Typically, metrics such as cross-validation, accuracy, precision, recall, and F1 score are used to measure the performance of the model. Use of Test Data is to evaluate the generalization ability of the model using separate test data for validation. This is important to ensure that the model is not overfitted to the training data.

4.2.5.2 Fairness Check

It evaluates whether the prediction results of the model are biased against a specific group. For example, this is a process to ensure that the results are not unfairly displayed based on gender, race, age, etc. This will enable us to fulfill our social responsibility and build ethical AI systems.

4.2.5.3 Interpretability Assessment

It analyzes the characteristics on which the model's predictions are based so that the results can be explained. Tools such as LIME or SHAP should be used to quantitatively evaluate the impact of each characteristic on the model's predictions.

4.2.5.4 Audit Process

It is to document all decisions and results during the model development and operation process so that they can be reviewed later. This increases the transparency of the model and helps stakeholders clearly understand how the model works. Regular audits ensure that the model continues to be reliable. This helps the model adapt to changing environments and prevent performance degradation.

4.2.5.5 Feedback Mechanism

This is a method to increase the reliability and understandability of the results by building a system that receives feedback from users or stakeholders on the results of the algorithm and continuously improves the model. This will allow the model to understand how it works in real situations and adjust it as needed.

In short, model validation and auditing play an important role in increasing the transparency of the algorithm and building trust in the AI system. This will allow users to trust the model's prediction results and provide clearer grounds for decision-making.

4.2.6 Generating Understandable Explanations

Utilizing technology to generate natural language explanations for the results of the algorithm, the results are delivered so that even non-experts can understand them. Generating understandable explanations is an important technical method to increase the transparency of the algorithm, and focuses on delivering the prediction results of the AI model to users in a clear and intuitive manner. This approach helps users understand how the model works and trust the results. The main elements are as follows:

4.2.6.1 Natural Language Explanations

This is a method to explain the model's prediction results in natural language. For example, for a specific prediction requested by a user, "This customer is a low credit risk. This is because they have a good previous payment history and a high income level." The results are explained in a way that is simple and clear so that even non-experts can understand them.

4.2.6.2 Visual Explanations

Using data visualization techniques, information related to the model's predictions are expressed graphically. For example, feature importance is visualized in a bar chart, or decision boundaries are visually represented to help users easily understand the results.

4.2.6.3 Case-based explanation

This is a method of explaining by presenting similar cases based on a specific prediction. For example, "Customer A with similar characteristics was approved for a loan, and Customer B was rejected." In this way, the validity of the prediction should be strengthened through real cases.

4.2.6.4 Explaining the working principle of the model

Simple explanations are provided so that users can understand how the model makes decisions. For example, "This model evaluates based on age, income, and credit score." Provide explanations that connect the inputs and outputs of the model.

4.2.6.5 Adjusting the level of explanation

Provide customized explanations to users, adjusting the depth of the explanation to the level that each user can understand. Provide more technical and detailed information for experts, and simple and intuitive explanations for general users.

4.2.6.6 Feedback and improvement

Continuously improve the explanation method based on feedback received from users. Analyze which explanations are easy to understand and which parts are confusing so that better explanations can be created.

In short, generating understandable explanations plays a vital role in making the results of AI models transparent and building trust with users. This will allow users to better understand the model's predictions and make rational decisions based on them. These methods help users and stakeholders have better trust by making the algorithm's operation transparent.

4.3 Improvement tasks such as setting

ethical standards

4.3.1 Transparency

The operation method and decision-making process of the AI system should be clearly explained and should be understandable to users.

4.3.1.1 Understandability of algorithms

A clear explanation is needed of how the AI system works and what data and algorithms are used. Users should be able to understand the decision-making process and results of the AI, and this will enable them to build trust.

4.3.1.2 Disclosure of data sources

By disclosing the source and characteristics of the data that the AI model learns, the quality and bias of the data should be assessed. This will help increase the reliability of the data and prevent unfair decisions or predictions.

4.3.1.3 Interpretability of results

An explanation should be provided for the decisions or predictions made by the AI system. Users should be able to understand how the results were derived, and this will enable them to increase their trust in the AI.

4.3.1.4 Accountability

Responsibility for the decisions or actions of the AI should be clearly defined. When transparency is guaranteed, it becomes clear who should be held accountable for wrong decisions or results, which promotes ethical use.

4.3.1.5 User Participation

In the development process of AI systems, stakeholders and users should be provided with opportunities to participate and provide opinions. This helps increase trust in the system and reflect diverse perspectives.

In conclusion, transparency is an essential element for ethical use of AI, contributing to building trust and creating fair AI systems. This will increase the acceptance of users and society as a whole.

4.3.2 Fairness

Algorithms should be designed to be non-discriminatory and ensure that all users are treated equally and fairly. This is important to minimize data bias. The following are the main aspects of fairness:

4.3.2.1 Bias Removal

If the data that AI models learn contains bias against a specific group, unfair decisions or predictions may result. In order to secure fairness, it is necessary to increase the diversity and representativeness of the dataset and identify and correct biased data.

4.3.2.2 Equal Access

AI technology should be equally accessible to all users. It is important to ensure that certain groups or individuals do not have access to the technology or are discriminated against. For this, design and policies that take accessibility into account are necessary.

4.3.2.3 Fair Results

Decisions made by AI systems should not be disadvantageous to certain groups. For example, AI should be designed not to discriminate based on race, gender, age, etc. during loan screening or hiring processes.

4.3.2.4 Continuous Monitoring

Continuous monitoring and evaluation are necessary to maintain the fairness of AI systems. Since data and environments may change over time, the system should be regularly inspected and modified when necessary.

4.3.2.5 Stakeholder Participation

Efforts to improve fairness are necessary by reflecting the opinions of various stakeholders (users, experts, social groups, etc.). This will allow for consideration of various perspectives and more comprehensive decision-making. Among the attributes of individuals, attributes that should not cause discrimination, such as gender, race, age, and religion, should be defined as sensitive attributes, and conditions should be specified that the judgments made by AI models should not change significantly depending on the sensitive attributes.

In this way, the final decision made by AI for a given individual can be used to approve or reject a loan, and the risk or stability of crime can be predicted by AI. In addition, the predicted credit score or predicted risk of recidivism can be determined through the score predicted by the AI model for a given individual. In addition, the actual outcome for a given individual can be determined through whether the loan is actually repaid or whether the individual actually commits a recidivism.

In conclusion, fairness is an essential ethical standard for AI technology to have a positive impact on society and provide equal opportunities to all users. By building and operating a fair AI system, trust and social acceptance can be increased.

4.3.3 Responsibility

Responsibility for decisions or actions of AI should be clarified, so that it is clear who is responsible when a problem occurs. This responsibility means taking clear responsibility for decisions or actions made by AI systems during the development and use of AI.

4.3.3.1 Clarification of Responsibility

It should be clarified who is responsible for the results of AI systems. There may be various entities such as developers, users, and companies, and the roles and responsibilities of each entity should be specifically identified.

4.3.3.2 Transparent explanation of results

When an AI system makes a specific decision, it should be able to explain the process and results. It is considered an important factor in increasing accountability to enable users to understand the decision-making process of AI.

4.3.3.3 Ethical Considerations

When an AI system makes decisions that affect human lives (e.g. medical diagnosis, legal judgment, etc.), it should comply with ethical standards. This is necessary to ensure that the system can make the right decisions.

4.3.3.4 Response when problems occur

A system must be established to respond quickly and effectively when an AI system makes a wrong decision or a problem occurs. This will minimize damage and demonstrate an attitude of acceptance of responsibility.

4.3.3.5 Continuous improvement

In order to strengthen accountability, the performance and results of the AI system must be continuously monitored and, if necessary, improved. This will contribute to increasing the reliability of the system and meeting the expectations of users and society.

In this way, efforts should be made to minimize possible damage by establishing a responsible entity in the process of developing and utilizing AI. To this end, the responsibilities among AI designers and developers, service providers, and users must be clearly defined.[11]

In conclusion, accountability is an essential ethical standard for AI technology to have a positive impact on society. Responsible AI development and utilization can build trust and increase acceptance by users and society.

4.3.4 Privacy protection

Privacy protection is one of the most important ethical standards when developing and utilizing AI. The collection and processing of user data must comply with the Personal Information Protection Act and be based on the user's consent. This means respecting the privacy of individuals, minimizing misuse of personal information, and ensuring safe processing.[12]

Only the minimum data necessary for the operation of the AI system should be collected and processed, and unnecessary data collection should be avoided. In addition, clear consent should be obtained from users for the collection and use of personal data, and users should have the right to manage their own data at any time. Collected data should be safely stored and processed, and protected from illegal external access or leakage.

This requires technical measures such as encryption and access control. Meanwhile, information on how data is collected and used should be clearly provided to users, and trust should be built through explanations of the data processing process. Users should have the right to request access to, correction, and deletion of their data, and such requests should be processed in a timely manner.

In conclusion, by complying with these privacy protection standards, AI systems should be able to process personal data safely and ethically, and can contribute to building trust between users and society.

4.3.5 Safety

AI systems should be designed to be safe and have mechanisms to prevent unexpected results or accidents. Safety is ensuring that AI systems operate in a predictable manner and do not cause harm to users and society. In April 2019, the European Commission (EC) listed technical 'safety' as one of the requirements for developing and using trustworthy AI.

AI systems should operate consistently and provide reliable results even in exceptional circumstances. Potential risks of AI should be identified in advance and measures should be taken to minimize them. In addition, mechanisms should be in place to safely terminate or resolve issues in

case the system behaves in an unexpected manner. The AI should be transparent so that users can understand how it operates and the decision-making process, and users should be educated about the limitations of the AI system and how to use it safely, thereby promoting proper use. In conclusion, by complying with these safety standards, AI systems will be able to have a positive impact on society.

4.4 Ethical Dilemmas and Challenges of Artificial Intelligence

4.4.1 Introduction

The use of services or devices equipped with Artificial Intelligence is becoming increasingly familiar, and the scope of interaction between digital devices and humans is continuously expanding. This trend shows that the prediction that various entities such as Artificial Intelligence robots will appear in the post-human era is no longer just science fiction. The question of how to respond to ethical issues that arise when humans interact with new entities other than humans will be in a transitional position that goes beyond the level of extending existing ethics to new entities and moving toward presenting new ethics.

However, even if the development of Artificial Intelligence is useful in our society, if it is inevitable to provide relief to victims, it is limited to respond with the current legal system, such as the tort system, and ethical issues will arise accordingly. Therefore, this paper will examine the ethical issues necessary to establish an ethical and social system focused on responsibility for Artificial Intelligence.

4.4.2 Development of Artificial Intelligence and Ethical Dilemmas

4.4.2.1 Floridi's Information Ethics

The technological development brought about by the revolution of Artificial Intelligence has a duality that protects human life while also having the potential risk of violating legal interests. This duality can provide an opportunity for theoretical discourse that can establish a rational and moral criminal responsibility system depending on how the ethical dilemma of Artificial Intelligence is resolved.

In this way, ethical issues related to Artificial Intelligence were raised in earnest by Floridi (Luciano Floridi). Here, Floridi proposed information ethics, that is, macroethics centered on the passive and the existential, as this transitional ethics. Floridi attempted to move beyond the anthropocentric perspective and grant moral status to the existence of information itself and to construct moral principles from the perspective of information. In this way, naturalistic attempts can be seen as an essential process in the discussion of dehumanization. This is because, in order to overcome anthropocentrism, it is necessary to derive moral evaluations from nature, which is distant from the human perspective.[13]

Based on this complex network of relationships, Floridi proposed a new concept called 'distributed moral action'. Morally important results cannot be reduced to the morally important actions of some individuals, but must be distributed among multiple actors. In this case, the question of who should be responsible for distributed moral actions and how much remains, but the argument that ethical responsibility should be attributed to all distributed actors is very meaningful in the era of Artificial Intelligence.[14]

4.4.2.2 Ethical dilemmas and guilt issues

Programming ethical algorithms for autonomous vehicles is based on research on the relationship between the most fundamental human value system and industrial technology. Recently, car manufacturers have been struggling with ethical dilemmas. For example, Google is focusing on pedestrians, who are vulnerable, and Mercedes-Benz has stated that it will focus on protecting humans in vehicles by evaluating the survival probability of drivers and passengers highly. Setting different standards like this can be quite problematic in assessing criminal responsibility. In addition, since it is impossible to predict all of the judgments of Artificial Intelligence, the question of how to do ethical programming will ultimately be a difficult one.

In particular, one of the issues that may be problematic in relation to self-driving cars is the part related to the Trolley Dilemma of Artificial Intelligence.[15] If ethical programming is instilled in advance to infringe on other legal interests in order to avoid illegal results, the responsibility of the programmer or manufacturer behind the scenes becomes an issue. For example, if the Artificial Intelligence installed in a self-driving car saves the life of the driver and crashes into a pedestrian, resulting in a punishable result, it will be necessary to examine whether the programmer or manufacturer who implemented ethical programming can be held liable for intentional injury or murder.

First, it shall be examined whether it can be considered an act out of Necessity. Necessity can be exempted from illegality within a reasonable scope in order to protect a superior legal interest of higher value than oneself, but as

explained above, the case of the Trolley Dilemma of self-driving cars can be said to be an issue in which equally valuable legal interests are at issue. Therefore, in cases of equal legal interests or when it is difficult to determine legal interests, there is no possibility of expectation, so it can be said that there is a possibility of responsibility excuse.

Second, the perpetrators did not intentionally program the driver or pedestrian to imagine or kill them. For this reason, it is necessary to determine whether the perpetrators had a duty to protect the life or body of the driver or pedestrian, or whether the crime of intentional omission can be established.[16]

In short, the dilemma itself means that it is impossible to program that can simultaneously structure the legal interests of both the driver and the pedestrian. Accordingly, there is a possibility that the illegality can be excused due to a conflict of duties, and furthermore, if the perpetrators inevitably loaded the AI with ethical programming in a realistically conflicting situation, it is believed that there is no possibility of expectation, so responsibility can be excused.

4.4.2.3 Review of ethical dilemmas by field

① Operation service field

It is very important to strike a balance between judgment based on AI technology and human judgment, and the balance and the relationship between the two sides are expected to change in the future, and accordingly, ethics are expected to change. Ethical review is necessary when emotions, beliefs, and behaviors are manipulated without the user's knowledge due to services provided by AI technology, or when priorities or selection are required.

It is necessary to examine how human relationships will change as AI technology expands human cognition and behavior. Other visions should be reconsidered, such as the acceptance and change of creative values in cooperation with Artificial Intelligence technology and the way each person relates to Artificial Intelligence technology. Here, the Trolley Dilemma is an experiment on what criteria should be used to make judgments if AI algorithms directly affect human life, and whether such judgments can be left to the algorithm. It is one of the ethical tests called the 'Trolley Dilemma' first proposed by British ethical philosopher Philippa Foot in 1967.[17]

The technology of self-driving cars is not much different from that of humans driving, in that a single vehicle perceives, controls, and drives on its own. In the case of the Trolley Dilemma, self-driving cars may be able to make decisions faster than humans, but this will not solve the fundamental problem.[18]

In order for self-driving cars to become popular, there is an ethical dilemma that must be resolved first. In an extreme situation where loss of life cannot be avoided, who should the self-driving algorithm be designed to sacrifice? If the self-driving algorithm is programmed to save 10 people in a crowd first, there will be few consumers who will readily purchase a car that cannot protect the owner in an extreme situation.

On the other hand, if the self-driving algorithm is programmed to save the owner first, utilitarians may focus their moral criticism on the car manufacturer and the owner who intentionally chose the manufacturer. Of course, it will not be easy to simply determine right or wrong in any choice. However, before millions of self-driving cars are fully deployed on the road, there should be sufficient social consensus on this ethical dilemma.

② Artwork Creation field

When Artificial Intelligence that has learned from past data accumulated by humans creates new works, the social evaluation of the skills and abilities of masters and craftsmen that have been highly regarded so far will change.

In the current situation where the purpose and target of using and utilizing Artificial Intelligence can be set by humans, the question arises as to whether the objects that can be manufactured using Artificial Intelligence should be explicitly limited. There is a possibility that it will become impossible to distinguish authenticity through Artificial Intelligence, and it will take a lot of effort and money to prove that it is a genuine work created by a human. If it is a work created by a human, it may be impressive, but if it is discovered that it is a work created using Artificial Intelligence, people will doubt that the Artificial Intelligence is impressing them, which will also raise ethical issues.

③ Medical diagnosis field

As the use and utilization of AI technology makes it possible to accurately estimate health conditions, it will be possible to predict diseases or possibilities that can occur before getting sick, thereby preventing illness. On the other hand, if the future of oneself feels too negative, one's hope for life may decrease.

In addition, if the disease-related prediction diagnosis by AI is too accurate, there may be fewer opportunities to have hope, take risks, or challenge life. Therefore, medical ethics issues such as the patient's right to know the results of the diagnosis, the right not to want to know, and the doctor's duty to convey them should be reestablished.

④ Conversation and communication field

In the case of problems caused by maliciously written conversational agents, conversational agents can have natural conversations with people using natural language and facial expressions. However, ethical issues such as how to ensure ethics to prevent negative effects on people and whether it is an insult to human dignity for AI that is indistinguishable from humans to communicate with humans by pretending to be humans will arise.

There may also be ethical issues raised, such as whether AI should always clearly state that it is AI, or to what extent it is acceptable to use AI to influence or manipulate human minds and emotions, or to use AI in dating businesses.

4.4.3 Suggestions

Intentional discrimination can be created by AI algorithms. However, as AI moves from the era of knowledge to the era of data, the problem of discrimination or distortion by data can become more serious. For example, if the creator or user of data used for Deep Learning is mostly from advanced countries or is male, unexpected distortions can occur.[19]

In addition, in order to solve these problems, I would like to suggest the establishment of a system such as the so-called "AI Ethics Judgment Standards Committee" to regularly check developers for ethics and the possibility of distortion and discrimination in software and data, and to fairly evaluate and analyze it with external experts.

Furthermore, with regard to the legal system, the appropriate distribution of risks and the subject of responsibility should be clearly defined. The National Information Technology Basic Act should be comprehensively revised to include design obligations such as the obligation to prevent service malfunctions for developers or users of AI technology and the "kill switch" system that can be forcibly terminated from the outside.

Here, the 'Kill Switch' is a modified punishment method specific to similar forms of AI such as death penalty or imprisonment, that is, it can completely delete AI program data,[20] or ultimately cut off the energy switch, which is practically equivalent to the death penalty by sentencing, and it can also take initialization measures such as reprogramming the AI program like imprisonment. Moreover, it can cause functional disabilities to the robot by taking various programming measures for the body of the AI itself.[21]

Meanwhile, various sanctions such as community service, such as requiring the robot to perform individual service activities for the community free of charge, can be considered.

4.4.4 Discussion

Ethical programming can be largely divided into which priority is given in the conflict between the driver's right to life and the pedestrian's right to life. For example, in the case of a self-driving car that enforces utilitarian programming, if there are many pedestrians, the AI will choose to place greater value on protecting the many pedestrians than on protecting the passengers. In that case, there is a risk that the driver himself may suffer the consequences of infringement of legal interests at any time. It is possible to show a contradictory attitude of supporting the utilitarian view while opposing riding in a utilitarian autonomous vehicle. In this respect, a view that supports programming that values the driver's right to life, unlike utilitarianism, can be derived.[22]

The technological development of Artificial Intelligence should fundamentally minimize the conflict between humans and Artificial Intelligence. More ethical considerations about Artificial Intelligence are needed, and such considerations should be carried out in the harmonious development of related industries and human life. The case of the autonomous vehicle explained above clearly shows the conflict between humans and Artificial Intelligence and the ethical dilemma of Artificial Intelligence. In other words, what choice will the Artificial Intelligence make when an emergency situation such as a brake failure occurs? In a way, the technical problem is directly connected to the ethical or moral problem. Therefore, it will be necessary to establish some standards or norms and legislate them.

In the case of the United States, in January 2017, AI researchers, lawyers, and ethicists gathered in Asilomar, California, USA, and announced the 'Asilomar AI Principles', consisting of a total of 23 articles that are declarative norms on the purpose of AI development, ethics and values, mid-term issues, and ethical issues such as AI safety, job replacement, legal responsibility, and securing human uniqueness. In Japan, the Society for Artificial Intelligence established ethics guidelines, and Canada, France, and other countries are also establishing related regulations at the government level.

In relation to this, Isaac Asimov announced the Three Laws of Robotics in the short story 'Runaround' in his book 'I, Robot' in 1942. The first law is that a robot must not injure a human being or ignore a situation where harm would be done to a human being. The second law is that

a robot must obey human orders unless it would conflict with the principles. The third principle is that robots must protect themselves as long as it does not violate the first and second principles mentioned above. Therefore, I believe that it is necessary to standardize this principle as a minimum rule.[23]

In the case of Korea, the Ethics Charter and the Ethics Guidelines for the Intelligent Information Society were announced in 2018. The contents presented four major principles of publicness, accountability, controllability, and transparency as ethics guides for the intelligent information society. However, there is a need to further refine the specific service models that AI can provide and the level of discipline that follows.

AI technology should be developed in a way that respects human welfare and rights and assists human decisions. In addition, AI technology should be developed and utilized in a sustainable manner, considering the impact on the environment and society. In addition, there are various other standards, and additional ethical guidelines may be required depending on the characteristics of each organization or company. As of now, the level is limited to theoretical and macroscopic discussions, so it should move forward to developing sophisticated specific models and establishing systems.

Technically, the need is being raised for designing systems that reduce the possibility of unpredictable and unintentional behaviors to ensure the safety of Artificial Intelligence, developing debugging tools and test environments that make it easy to understand and operate advanced Artificial Intelligence systems, and developing technologies to stop and fix the system in the event of a failure or malfunction.

5. Conclusion

In today's era of rapid development of Artificial Intelligence technology, this paper has examined the ethical implications of these technologies on Big Data analysis and other issues. When utilizing Artificial Intelligence technology for Big Data analysis, it is crucial to consider the implications it will have. Among these, consideration related to ethics, which involves moral value judgments regarding behavior, is essential.

For this purpose, I first examined the correlation between Artificial Intelligence and Big Data. Specifically, I looked at what Artificial Intelligence does by utilizing Big Data, and how Machine Learning algorithms learn data and classify or predict new data. Through this, it was able to confirm the complementarity between Big Data and Artificial Intelligence.

Next, I investigated how data is collected, what algorithms are utilized, and how they learn in real time to improve their recommendation system through a case study on Netflix. Netflix is moving toward becoming globally competitive by combining Artificial Intelligence and Big Data, such as providing customized experiences to users through A/B testing for an effective recommendation system. In examining the ethical implications of Artificial Intelligence technology on Big Data analysis, I studied the issues of privacy protection, bias, and lack of transparency.

Personal information protection issues may include lack of consent from the Information Subject, data misuse, information asymmetry, and data security issues. In addition, issues of bias such as employment discrimination, crime prediction, credit rating, accessibility to medical services, social media, and information filtering should be considered. On the other hand, when the AI decision-making process is opaque, issues such as lack of trust, responsibility issues, difficulty in risk management, and conflicts of interest arise.

Solutions to these issues can be categorized as follows: First, strengthening legal regulations through strengthening the rights of the Information Subject, increasing transparency in data processing, strict consent requirements, strengthening punishment for data infringement, mandating the appointment of a data protection officer, raising education and awareness, and strengthening international cooperation.

Second, technical solutions include methods such as explainable AI, model interpretation tools, interactive visualization, simplification of algorithms, model verification and audit, and generation of understandable explanations.

Third, ethical standards can be set when developing and utilizing AI to promote transparency, fairness, accountability, privacy protection, and safety.

Although AI is in the process of developing excellent technology in itself, the problems in utilizing AI for Big Data cannot be ignored. If these ethical issues are reviewed and evaluated in advance, and sufficient solutions are prepared, Big Data can be analyzed by maximizing the advantages of Artificial Intelligence technology while minimizing potential risks. Even if Big Data utilized by AI technology does not make any ethical judgments inherently, it can be interpreted externally as if it is making ethical judgments. In other words, AI is creating norms or ethical standards in our society, whether intentionally or unintentionally. A method should be developed to judge and verify whether AI or Big Data is following social norms.

Even if decisions are simply made by AI, sensitive issues can be raised socially. Therefore, user ethics should be

added to AI technology that analyzes and utilizes Big Data.

For the future of AI, such 'AI Ethics Committees' should be operated more transparently and openly, as in the case of Google, and companies, governments, and the public should engage in joint discussions. The direction of this research should be to understand how people make judgments in various situations in a more ethical and socially consensual manner.

In short, through sufficient social discussion and debate in the future, it will be able to examine issues such as the ethical impact of AI technology on Big Data analysis. If an ethical foundation is established through this process, it will be possible to create a virtuous cycle that fosters the development of better AI technology.

참고문헌(Reference)

[1] Edward Lee, "Technological Fair Use", SOUTHERN CALIFORNIA LAW REVIEW(Vol. 83:797), p.798, 2010.

[2] Kim Soo-yeon, "Review of improving personal information protection regulations for activating the Big Data industry", Korea Economic Research Institute, p.13-14, 2015.

[3] Hyundai Economic Research Institute, "AI Era, Where is Korea Now? - Examination of Domestic AI (Artificial Intelligence) Industry Base," 16(8), p.1, 2016.

[4] Yoo Seong-min, "The Impact of Big Data on Artificial Intelligence," Journal of the Korea Information Technology Society, 14(1), Korea Information Technology Society, p.32, 2016.

[5] Jo Eun-ji, Kim Do-hyun, "Development of A/B Testing Method Considering Data Distribution," Proceedings of the Fall Conference of the Korea Quality Management Society, Korea Quality Management Society, p.59, 2020.

[6] Son Eun-ji, A Study on the Constitutional Risks of Artificial Intelligence and the Establishment of a Responsive Normative System, Ph.D. Dissertation, Sookmyung Women's University, p68, 2023.

[7] Hong, Seong-wook, "Artificial Intelligence Algorithms and Discrimination," 2018 STEPI Fellowship, Science and Technology Policy Institute, p.13, 2018.

[8] Yang Jong-mo, "The Impact of Bias and Opaqueness of Artificial Intelligence Algorithms on Legal Decision-making and Regulatory Measures", Law Association, 66(3), p.75, 2017.

[9] Personal Information Protection Commission, "Policy Direction for Safe Use of Personal Information in the Era of Artificial Intelligence", Press Release, p.4, 2023.

[10] Park Gwang-bae, Chae Seong-hee, Kim Hyeon-jin, "Processing system of generated information in the Big Data era" - A study on the regulations and improvement measures of our personal information protection law regarding the processing of inferred information -, Korea Information Law Society, 21(2), p.185, 2017.

[11] Ministry of Science and ICT, National AI Ethics Standards (Draft), p.8, 2020.

[12] Ministry of Science and ICT, "Human-Centered Artificial Intelligence (AI) Ethics Standards", p.4, 2020.

[13] Kim Yu-min, Mok Gwang-su, "Reconstructing Luciano Floridi's naturalism in information ethics: Beyond the critique of naturalistic errors", Pan-Korean Philosophy, (100), Pan-Korean Philosophy Society, p.453, 2021.

[14] Lee Jung-won, "Can Artificial Intelligence Be Responsible?", Philosophy of Science, 22(2), p.82, 2019.

[15] Chul Kang, "Trolley Problem and Three Basis for Moral Judgment", Ethics Research, 90(0), Korean Ethics Society, p.147, 2013.

[16] Kim Young-joong, "A Study on the Significance of Emergency Evacuation," Ph.D. dissertation, Seoul National University Graduate School, p.9, 2013.

[17] Gabriel Andrade, "Medical ethics and the trolley Problem", Journal of Medical Ethics and History of Medicine, p.140, 2019.

[18] Byun Yong-wan and Jang Jae-ok, "Civil Liability Principles for Distributing Responsibility in the Environment of Artificial Intelligence", Sports Entertainment and Law, 23(3), p.139, 2020.

[19] Han, Sang-ki, "Social Issues and Ethical Problems of Artificial Intelligence Technology," Journal of the Korean Multimedia Society, 20(3), p.43. 2016.

[20] Lee Won-sang, "Cybercrime Theory", p.170, Park Young-sa, 2021.

[21] Shin Song-i and Kim Je-wan, "Legal Issues on Recognition of Legal Personhood for Artificial Intelligence", Korea Business Law Association, 2023

Summer Joint Academic Conference, p.62-63, 2023.
[22] Oh Do-bin, Yoo Ji-won, Yang Gyu-won, and Koo Bo-min, “The Limits of Uniform Programming of Autonomous Vehicles from a Post-Paternalistic Perspective,” Korea Law Review, (84), p.270-272, 2017.
[23] Asimov, Isaac, “Runaround”, ISBN 978-0-385-42304-5, p.40, 1950.

◐ Author Biography ◑

Se Eun Jeong
Korea University, BS in Law
Indiana University Bloomington School of Law(, MS
Korea University, Ph.D. in Law
aSSIST, MS in AI Bigdata
CJ Feed&Care, Legal Team
Research Interest : AI, Big Data, intellectual property rights
E-mail : sejng@hanmail.net