Prediction-Early Warning-Prevention: The Value Orientation, Future Dimension, and Practical Approach of Artificial Intelligence-Enabled Safety Management

Zhiqiang Wen, Chengyue Yang*

Tianjin Normal University, Tianjin, 300387, China
**Corresponding author: 18056108525@163.com*

Abstract: *With the development of digital technology and the appearance of many kinds of risks, safety management has moved from a reactive to a proactive mode. Many kinds of artificial intelligence have been used to improve both the accuracy and the speed of risk forecasting and early warning in recent years, so a closed-loop prevention system has begun to be built based on them. There are now problems of algorithmic bias and a crisis of trust in people and machines. Therefore, this paper puts forward a "three-dimensional collaboration" model to establish ethical norms for artificial intelligence, guarantee human control over important decisions, and promote public safety literacy education. In the future, more attention will be paid to the integration of technology, the expansion of application scenarios and governance innovation to promote intelligent and all-round upgrading of safety management.*

Keywords: *Artificial Intelligence, Safety Management, Prediction-Early Warning, Prevention*

1. Introduction

With the increase in demands from society for safety management, intelligent applications in the area of safety management have started to emerge. With the rapid development of the new round of technological revolution and industrial transformation, new technologies such as artificial intelligence are flourishing, and people's ability to understand and change the world has been greatly enhanced[1]. With the development of digitalisation and an increase in the complexity of risks, over the past few years, safety management systems around the world have begun to change from reactive to proactive. The old way of safety management was to rely on manual inspection, experience and fixed regulations, so it could not cope with all kinds of unforeseen risks caused by cyberattacks, emergencies and natural disasters effectively. The three main results of the application of artificial intelligence technology are:

a: Improve the Prediction Accuracy. Machine learning can collect a large amount of data, and with the help of artificial intelligence, it will learn how the risk has changed over time. Big Data technology and intelligent analysis models are employed in the community smart health service platform to offer personalised and convenient health management services for the community, and the prediction accuracy of health risks is 87%[2]. It will raise the general level of living and quality of life.

b: Prompt early warning. Together, calculation and data analysis can be employed to increase the speed of the early warning system. In the Design of an early warning system for an engineering project, one should consider both the scope and the uncertainty of the project, and build a multi-dimensional, multi-level early warning indicator system that can monitor all kinds of risks to the project[3]. It will not be worse.

c: Construction of a Closed-Loop Prevention System. Now, together they have started to move the focus of prevention from experience to data and are using digital twin technology and reinforcement learning to achieve this. A typical representative case is the community smart health service platform; it has been developed in a modular manner, and all the modules are connected to form a closed-loop health service system. Based on the above steps, the whole system of health service support will be built. Therefore, there will be fewer problems of risk events.

In the future, the following three will still be the main technical bottlenecks of artificial

intelligence-based safety management:

a: The construction of a multi-module perception network. Gather all kinds of information from all over, such as satellite remote sensing, drone inspection, public opinion on social media, etc., and build a multi-module association analysis engine.

b: Increase the Autonomy of Evolutionary Decision-Making Systems. Repeat the vaccination operation of the artificial intelligence model several times. A virus and antivirus software are in a continuous race, and this competition will train the defence system to be smarter with every battle; it will be able to handle all kinds of new attacks automatically.

c: New Ideas for Human-Machine Collaboration. The operation interface has clear buttons or voice prompts, and all major decisions need to be confirmed by a person before they are carried out, just as one has to press the close-door button when the elevator is full.

There are also some problems with the technology, such as algorithmic bias and discrimination, and some unfairness has occurred. There is a lack of trust in human-machine collaboration; people are not sure that the system will work properly. If too many people rely on the system, they will lose their own thinking and problem-solving abilities. Therefore, the model of "three-dimensional cooperation" proposed in the above section is presented here.

a: Issue ethical guidelines for the safe use of artificial intelligence, build an algorithm filing and impact assessment system, and deal with the corresponding problems more fairly.

b: The human expert should have the final say and the right of interpretation for serious matters to prevent arbitrariness in the system.

c: Introduce a national "safety literacy" education plan that includes AI risk awareness in the education system, but do not give it too much weight.

Artificial intelligence technology is altering how we create value and live in society by collecting data, running algorithms and employing intelligent management. Now there are better earthquake early warning systems, real-time detection of hidden dangers in major Internet security systems, risk simulation in smart cities, intelligent decision-making for emergency rescue, etc. Artificial intelligence is building an all-encompassing safety-protection network in the whole chain of "prediction - early warning - prevention". Strengthening the management of safety with artificial intelligence is a good way to improve the efficiency of technology and deal with the uncertainty of life in a new way. Efficiency and ethics need to go together in the current era, so we should build an intelligent safety ecosystem that can proactively find problems and issue early warnings to prevent accidents from happening in the first place. In the future, new research will be carried out on the topic of human-machine hybrid intelligence and multi-module learning to promote the development of safety management.

2. Values Orientation

Based on the above literature review, it can be seen that the application of artificial intelligence in safety management has moved from passive response to active prevention. In the future, artificial intelligence will be better at making independent judgments based on all kinds of data, and thus an effective system for human-machine cooperation in governance at all links of the "monitoring-response-governance" chain will be established. AI technology is being used in many places to change the basic ideas of safety management, and now its main function is to make safety management "data-driven" rather than "experience-driven". The three main values in current studies are:

2.1 Precision of Risk Prediction

Artificial intelligence has been collecting all kinds of data to enhance the accuracy of risk forecasting. The development and application of public data resources should be driven by market forces and regulated by policies from the government. By adding the data industry to the guidance catalog for industrial structure adjustment, we can promote technological innovation, cultivate high-end enterprises in the field of data element business, and support the development of social organisations and industry alliances, such as data industry associations and societies[4]. AI technology can be used to collect and organise all kinds of data, such as organised data, semi-organised data, etc., to help clarify data security and add attribute labels to the data[5]. Enhancement of the Interpretability Mechanism for

Generative AI Algorithms in Judicial Applications to Improve the Credibility of Judicial Judgments and Provide Transparent Support for Risk Prediction[6].

2.2 Agile Early Warning Response

Move the data processing from the central cloud to distributed computing and real-time analysis near the source of the data or the user to reduce the early warning delay. Power companies are now using AI large models to find problems with equipment in real time and can warn people about the problems a little earlier[7]. Zhang Jing[8] put forward the idea of a "one-network service" platform for the quick spread of risk information based on data sharing in ship safety management. Sang Xianjun and others[6] have also mentioned that large models can be used in the field of military security, and since the training data for these models contain relevant attributes of risk scenarios, they are suitable for all kinds of situations.

2.3 Closed-Loop Formation of the Prevention System

New-media techniques can be employed to build a 3D model that is very close to the real shape of an object. The model can reflect the actual operating conditions of a company in real time and make many kinds of forecasts, optimisations and decisions. Digital Twins and Reinforcement Learning technology can make the first kind of prevention data-driven. An intelligent + ecosystem for the management of laboratory safety in universities should be constructed, and artificial intelligence can be employed to simulate experimental risks and optimise emergency response plans, as shown in[9]. ChatGPT-type technology can be introduced in vocational education to improve the emergency response skills of students through virtual simulation training, and it has been shown that human-machine collaboration can enhance the prevention system[10].

Artificial intelligence technology has altered the idea of safety management, and now it is used to prevent problems rather than to deal with them. The main reason for this change is that artificial intelligence has altered the basic concepts of safety management, and the old "experience-driven" method must now be replaced by a "data-driven" one. Artificial intelligence can be employed to expand the scope and improve the accuracy of multi-modal data fusion in the management of safety, continually enhance autonomous decision-making, and explore more efficient and transparent models for human-machine collaboration and governance. The goal is to build an all-weather safety management system for the whole country, raise the general level of safety management at all levels through local optimisation and systematic quality improvement, and put forward new ideas for governing various sectors such as smart cities, critical infrastructure, industrial production and public security.

3. Road Ahead

In the future, artificial intelligence will be used more and more widely to enhance the safety of construction sites and bring new governance approaches. To address the problems of the current technology and build an all-weather safety management system that is complete and future-oriented is the aim. With the current progress of technology and practical demands, the following three directions will be the focus of future research, along with some new research ideas.

3.1 Technology Integration: From a Single Instrument to an Intelligent System

3.1.1 Collaborative Innovation in Multi-module Data Fusion and Edge Computing

At present, the problem is that only one source of data has been used for collection. In the future, a single integrated multi-module perception network will be established to gather all kinds of data from various sources in different areas, such as satellite remote sensing, drone inspection, Internet of Things sensors, public opinion on social media, etc., and then preprocess and analyze this data in real time using edge computing technology. At the same time, it should be possible to share data among the departments under the condition of protecting privacy to improve the accuracy and scope of risk prediction.

3.1.2 Progress in Autonomous Evolutionary Decision-Making Systems and Causal Reasoning

Most of the current artificial intelligence models can only determine whether there is a problem and

do not have the ability to reason causally or deal with new risks. In the future, we will need to construct a digital twin system according to the "vaccine-style training" method to simulate many risk development paths. Reinforcement learning can be used to have the AI try all kinds of attack and defence plans repeatedly in a simulated environment and then learn how to defend itself independently. An intelligent game system (adversarial network) has been developed in the field of cybersecurity; it consists of two neural networks that are in competition with each other, learn from this competition to generate new virus attack patterns, and at the same time, a defence system is independently trained to develop countermeasures. Present a way to calculate causal reasoning that can help artificial intelligence find causal links in risks and take preventive measures at the nodes of causation proactively; it should not be reactive.

3.2 Application Scenarios: From Vertical Fields to the Three-dimensional Expansion of Complex Systems

3.2.1 Multi-level Risk Co-governance in Smart Cities

A "smart city" has many kinds of systems, such as transport and energy. The safety management of these departments should be in line with the company's general risk management system. AI systems for vehicle-road coordination in intelligent transportation can be employed to find problems such as traffic congestion and bad weather in real time, and at the same time, these systems can be connected to the power grid to reduce the electricity consumption of streetlights and optimise emergency ambulance routes. Explain the idea of digital twin technology, build an urban risk simulation platform to model the spread of systemic risks in bad weather and public health crises, and issue early warnings about such disasters.

3.2.2 Cross-regional Cooperation on Global Risk Governance

With the spread of risks from climate change around the world and the loss of lives, we now need to cooperate on artificial intelligence governance at the border and in different areas. Artificial intelligence can be used to analyse medical data around the world to control the spread of epidemics, find new viruses in time, and cooperate with foreign countries to deal with them. Artificial intelligence monitoring technology in space has been employed to trace the spread path of transboundary pollutants in the ocean environment, and with the participation of other countries, an international system for data sharing, joint responsibility and cooperative governance has been established.

3.3 Governance Model: Paradigm Innovation in Human-Machine Collaboration for Human-Machine Co-governance

3.3.1 Dynamic Equilibrium of Ethical Governance and Technological Innovation

Problems of ethics in algorithms and data privacy need to be solved by the cooperation of technology and institutions. Build an objective algorithm, do not let bias in the data affect the results of technology-based decision-making, etc. Raise the ethical standards for the safe use of artificial intelligence in society and build a dynamic regulatory system of algorithmic transparency and risk control. Human judges should carry out an ethical review of the decisions made by generative artificial intelligence in the judicial process to ensure that legal fairness and public order and good morals are not violated by technology.

3.3.2 Joint Improvement of Public Safety Literacy and Artificial Intelligence Governance Ability

Increase the public's awareness of the risks of artificial intelligence and strengthen their participation in the governance of artificial intelligence in the future. Introduce the basic knowledge of artificial intelligence in the courses for young children and teenagers to help them learn algorithmic thinking and data security awareness. To solve the problem of an unsafe digital divide, AI safety tools that meet the needs of digitally disadvantaged groups, such as older people, need to be developed.

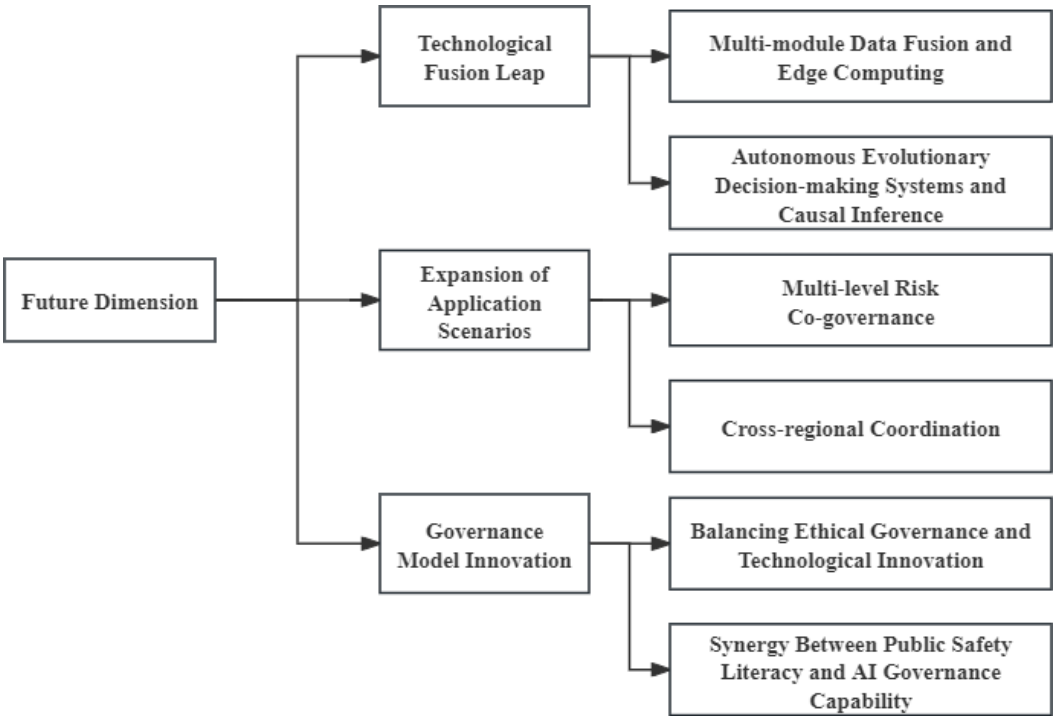

Figure 1 Schematic Diagram of the Three Major Directions of the Future Path

The following are some new research directions in the future: the cognitive collaboration mechanism of human-machine hybrid intelligence, capability enhancement of multi-module learning, and cross-domain risk management. At the same time, we also need to promote interdisciplinary research in artificial intelligence ethics, legal regulation and social governance to combine the rationality of technology with humanitarianism and safety.

4. Practical Approach

4.1 Technology Empowerment: Building an All-Weather Intelligent Safety Protection System

4.1.1 Construction of an Agile Early Warning Network Based on Edge Computing

Some of the data processing and early warning functions are performed at the network edge by way of edge computing. Data from meteorological satellites, geological sensors, social media disaster reports in the field of early warning for natural disasters can be combined, edge nodes are used to analyse abnormal signals in real time, and thus the time delay between data collection and the issuance of warnings is reduced. With the help of Edge AI, power companies have started to collect real-time vibration data of the transmission line and wind load information from wind power facilities. If the amplitude of oscillation is too large, an alarm will be triggered immediately and the location and reason for the fault will be notified to the maintenance department.

4.1.2 Closed-Loop Formation of a Digital Twin-driven Prevention System

Digital Twins are being developed for industrial production and smart cities, and at different times, the corresponding data from the real world is used to build models that show the current situation. University laboratories are building virtual experimental environments based on digital twin technology now, and artificial intelligence can predict potential risks in the experiments in advance, such as chemical leaks, equipment failure and sudden explosions; therefore, emergency plans will be prepared in advance and the management department will be informed.

4.2 Institutional Innovation: Establishment of a Rule System for Human-Machine Collaborative Governance

4.2.1 Ethics and Legal Regulations for the Safe Use of Artificial Intelligence

Hospitals and courts are high-risk places, so compulsory ethical norms have been issued to ensure

that the decisions of AI in these institutions are explainable. Most of the time, the industry has established self-regulation rules and encouraged companies to disclose the logic of their algorithms and data. For matters that may affect people's lives and property, such as medical surgery or the operation of a nuclear power plant, any advice provided by artificial intelligence must be confirmed by a qualified professional before it is carried out. A large-scale blackout operation needs to be carried out for the power outage decision system in smart grids, and then the dispatcher must verify biometrics.

4.2.2 Improvement of Interdepartmental Collaborative Governance Mechanism

Establish a special AI safety office under the emergency management department to organise the sharing of data and cooperation among the public security, transport, health and other departments. For example, in the event of a natural disaster, the plan for distributing rescue resources that has been created by artificial intelligence needs to be approved by all relevant departments after coordination to ensure that there are enough resources in all areas. To solve the problems of cross-border and cross-basin risks, regional artificial intelligence safety alliances will be set up to coordinate technical standards and data-sharing agreements.

4.3 Humanistic Cultivation: Building a Safety Culture Ecosystem for Human-Machine Symbiosis

4.3.1 Stepped Improvement of Public Safety Literacy

General education courses in "Artificial Intelligence and Safety" are available for students of primary and secondary school. ChatGPT can be employed to construct a virtual laboratory for teaching data security and algorithms, and then the students will learn by doing. Design of Age-Friendly AI Safety Products for Older Adults. Community classes and family pairings are being organised to help older people learn to use smart devices and avoid the problems of a digital divide and insecurity.

4.3.2 Two-way Construction of Human-Machine Trust Relationship

A kind of algorithm design that is easy for users to understand is one that shows the weights of the characteristic variables in a health risk prediction model, and a visual interface can be employed to help people understand and trust artificial intelligence. It is a closed-loop system of "user feedback - algorithm optimisation", and the public can submit objections to the safety decision of the artificial intelligence and propose corrections. To strengthen the sense of ownership among the public in the smart-city traffic signal optimisation algorithm, this algorithm will be adjusted according to changes in people's travel patterns.

With the help of technology, institutions and people can build an intelligent safety management system; this system will have a better sense of risk, issue timely warnings and preventive measures, and thus provide all-round and effective safety protection.

5. Conclusion and Outlook

The subject of this paper is AI-driven safety management, and with the changes in digitalisation and an increase in risks, artificial intelligence has begun to change the model of safety management from reactive response to proactive prevention. With the help of machine learning, edge computing and digital twins have been employed to provide early warnings of problems in time and build a closed-loop prevention system. At the same time, there are also problems of algorithmic bias and a crisis of trust in technology among the public. Therefore, this paper puts forward a "three-dimensional collaboration" model to establish an ethical norm system, ensure human control over serious decisions, and promote public safety literacy education. In the future, many new technologies will need to be developed in all kinds of places for the AI-driven safety management system; for example, data fusion, collaboration at the edge, autonomous decision-making, causal reasoning, switching from passive response to active intervention, etc., will also have to be added. In the future, all kinds of risk co-governance in smart cities and cooperation on global risks among different regions will also be established. The two should integrate ethics and new technologies to enhance the sense of security for the public and build trust in people and machines.

References

[1] Zhao Lei, Zhou Dawei, Pan Peng. Opportunities and Challenges of Artificial Intelligence Empowering Construction Safety Management. Journal of Engineering Management, 2025, 39(02):

100-105.
[2] Xu Qiang, Wang Xin, Wang Jie. Construction of a Community Smart Health Service Platform Based on Deep Big Data Profiling. Digital Communication World, 2025, (03): 76-78.
[3] Zhang Qian, Li Bo. Construction and Application of Engineering Project Risk Assessment and Early Warning Systems. Brick-Tile, 2025, (05): 123-125.
[4] Fan Lili, Zhang Xiang, Ye Zhipeng, Wu Xiaolong, Shen Feiwei, Dong Peng. A Joint Discussion on the Theoretical Explanation and Development Path of Accelerating the Development and Utilization of Public Data Resources. Journal of Chang'an University (Social Science Edition), 2025, 27(01): 1-37.
[5] Zhou Keke. Research on the Data Security Governance Framework and Its Application in Government Data Governance. Thesis Advisor: Ding Hongfa. Guizhou University of Finance and Economics, 2024.
[6] Sang Xianjun. Systematic Construction and Operation of Judicial Algorithm Interpretation Mechanisms. Modernization of Rule of Law, 2025, 9(03): 158-175.
[7] Zhuo Ying, Li Geng. Exploration and Practice of Multi-Region Collaborative Power Communication Safety Management. Electric Safety Technology, 2025, 27(04): 25-28.
[8] Zhang Jing. Research on Safety Management of Shanghai Inland River Ships from the Perspective of Technological Empowerment. Thesis Advisor: Rong Zhi. Shanghai Normal University, 2020.
[9] Wang Fan, Zhang Liucheng, You Tingting, Li Wantao. Analysis and Prospect of Safety Management in University Teaching Laboratories in the Digital Age. Laboratory Science, 2024, 27(04): 204-209.
[10] Guan Yongcai, Du Guobing, Liu Jiali, Yang Jingxin. Research on ChatGPT-Type Technologies Empowering High-Quality Development of Vocational Education. Journal of Xiangyang Vocational and Technical College, 2024, 23(05): 11-16.