Applications of AI Alignment and Anticipatory Networks to Designing Industrial Risk Management Decision Support Systems

Andrzej M.J. Skulimowski
AGH University of Science and Technology
Decision Sciences Laboratory, Kraków, Poland *ams@agh.edu.pl*

Paweł Łydek
AGH University of Science and Technology
Decision Sciences Laboratory, Kraków, Poland *lydek@agh.edu.pl*

Abstract

This paper proposes new ideas to be implemented in the design of intelligent decision support systems (IDSSs) for the management of industrial risks and safety. This class of information systems links different departments of an enterprise in case of emergency and supports multiple threat management. It can also cover financial risk management and long-term resilience planning. The first idea consists of the active use of AI alignment principles in the design of IDSSs, while taking into account the best AI technologies available to the developer and expected adversarial technologies. According to the AI-alignment paradigm, the AI evolution is modelled to identify the most suitable techniques to solve security, safety and risk management problems within a given time frame. This approach is aimed at enterprise resilience building with system design and can be combined with DevOps. The other idea consists of ensuring optimal emergency action planning by adjusting the IDSS features to cyber-human capabilities and the technical background of an emergency response. This type of alignment applies anticipatory networks to model ad hoc organizational structures created to handle crisis situations, including natural disasters, technical accidents, as well as anthropogenic threats. The above ideas have been applied to the IDSS design for a large industrial plant.

Keywords: AI Alignment, Intelligent Decision Support Systems, Anticipatory Networks, AI-based Information Systems, Enterprise Risk Management

1. Introduction

Enterprise risk management (ERM) systems usually focus on financial risks, while industrial risk management (IRM) is a broader topic, focusing on controlling cyber-physical systems with video surveillance and other sensors. Both ERM and IRM systems also differ in the time frame of deployment: ERM is used for long-term financial planning without any pre-determined particular activity of deployment. On the contrary, the relevance of IRM is particularly evident in crisis situations when the IRM system governs short-term emergency response and middle-term recovery activities. In this article, we argue that due to the availability of new artificial intelligence (AI) methods, it is possible to integrate risk management into a multi-level intelligent decision support system (intelligent DSS, or IDSS), covering strategic and operational planning, as well as ad hoc emergency response. In particular, intelligent decision models will be supported by machine learning (ML) techniques, enabling autonomous threat model building from retrieved data streams and prior emergency action records.

Our ideas are inspired by the real-life case of an IDSS designed for a large industrial plant, where heterogeneous natural and anthropogenic threats can be mitigated with a combination of AI-based techniques, such as machine learning (ML) of threat models, sensor information fusion and understanding, and multicriteria decision making procedures. These techniques yield optimal response action plans to be processed and executed in real-time in case of an emergency. The preventive measures are undertaken as outputs from middle- and long-term planning that also utilize AI methods such as dynamic threat maps [13], recurrent-neural network based financial forecasting and natural disaster scenario planning. A real-life application of multiattribute preference models and other techniques derived according to the AI alignment principles is the IDSS for designing the emergency managerial response in an open-cast limestone mine in Poland.

An intelligent DSS that incorporates the above functionalities is capable of recommending situation-dependent risk mitigation actions, operations and strategies to ensure an optimal level of industrial safety to all planning horizons applicable in an enterprise. Such systems are termed industrial risk management DSSs or disaster resilience management DSSs (IRM DSSs or DRMSSs [12]). When designing an IRM DSS, its developers deal with complex information and knowledge management problems that should ultimately lead to optimal resolution of risk management issues and minimize related losses. To efficiently integrate risk management at different levels, within all relevant planning horizons, from immediate remedies to planning complex operations and long-term strategic resilience-building measures, we propose a causal model of threats, risks, crisis management decisions, and their consequences. The ultimate model will be implemented via an IRM DSS to support emergency and risk management decisions at all applicable levels. Optimal decisions are derived from all available threat-related information, such as sensor data, historical facts on past threats, and the ways and results of their handling. The constraints on decision rules are imposed either by law or by the internal regulations at the particular industrial plant. This system is developed according to the DevOps paradigm, while the experience learned through its operations is enhanced by links to external AI-foresight and AI-alignment modules. These support future-oriented development of subsequent IRM DSS releases.

Thus, the primary aim of this article is to propose a software architecture that allows decision makers responsible for crisis management to integrate the surveillance, signal processing, and decision support technologies into a holistic industrial security system. Another aim is the provision of an organizational scheme for the design of specialized information systems for solving industrial resilience and risk management problems. The functionalities that require intensive development of new AI techniques are the fusion of information obtained from various sources in real time and the development of optimal decision algorithms using this information. In addition, ML techniques, including both reinforcement and semi-supervised learning, may be applied to emergency situations as well as to define preventive and mitigation measures. A preliminary scheme of the AI-based enterprise risk management system is presented in the next section. In section 3, we propose an application of AI alignment principles, and in section 4, a causal-anticipatory model of risks and mitigating actions is presented. Conclusions are provided in section 5.

2. Industrial Risk Management Problem Statement and Related Research

In the industrial security problems presented in this paper, risk is attributed to external threats, to information processing procedures that can bias the data with errors, to human operational errors, and to systematic erroneous decisions that can be made during risk management. In addition, sensor measurements can be skewed due to their inaccuracy. The general threat transfer can be modelled as a network, where information loss and incidental operational and decision-making faults are sources of additional risks. This network is complemented by the risk management and optimization model involving decision algorithms, actions, and actuators to implement them. Both components of the model are coupled by feedback information, received by the sensors, compared to the values provided by the model and presented to the ML module and to the decision makers. The DSS engine sequentially uses the above model components and links them to a semi-supervised machine learning procedure [14]. Assessments of previous emergency situations serve to elicit labels of current unlabelled sensor data, threat characteristics, risk mitigation procedure parameters and managerial decisions.

Complex production processes are often related to specific risks caused by the potential occurrence of natural phenomena such as flooding or landslides, industrial accidents, or anthropogenic threats such as an intrusion or act of sabotage. These threats may simultaneously occur in multiple locations distributed over a large area. The industrial risk analysis quantifies the potential threat impacts, discerns the relationships between them and recommends optimal procedures of risk prevention and mitigation measures. The availability of information stored and processed in other enterprise information systems (EISs) [5] increases the efficiency of the IRM system. For example, inventory management in an enterprise resource planning (ERP) system provides clues as to where dangerous substances or expensive equipment are stored and exposed to specific risks. The maintenance support systems process information from sensors linked to the Industrial Internet of Things (IIoT) [8], while the predictive maintenance modules indicate vulnerable sites and their exposure to risks. Therefore, IRM systems should be interoperable and take into account data streams from heterogeneous crisis management processes.

The motivation behind applying advanced AI information processing methods and software results from the will to satisfactorily meet the above challenges by state-of-the-art monitoring and prevention technologies [6]. These include visual surveillance cameras, other sensors, as well as autonomous inspection robots. The ML-based automatic interpretation of images from the protected area, coupled with social media information [7], may indicate vulnerable sites, e.g., those threatened by flooding during heavy rain.

The information system community noted the relevance of IRM [9] relatively early. While AI is widely used in financial risk management systems, its deployment in natural and anthropogenic crisis management systems was rare up until recently [3], [12]. When designing our IRM DSS, we took into account the current development of crisis management systems with decision support functionalities. They emerged from early warning systems that evolved into cloud-based heterogeneous signal processing [7]. Various crisis management and DSS architectures were then proposed [3]. The diversity of threats, combined with feasible prevention and mitigation measures, showed an increased relevance of domain ontologies [1] and quality of information [10]. The evacuation of staff and equipment turned out to be the most important issue from the perspective of solving resilience problems in the open-cast mine [6]. Among the analytical models applied to model risks in information systems are system dynamics [4] and Petri nets [15].

3. AI Alignment Principles Applied to IRM DSS Design

This section provides a background on knowledge management in industrial risk modelling, referring to the real-life problems occurring in an enterprise that serves as a motivating example for our research. However, we see that the model presented will be useful in solving a broad family of related industrial risk and safety management problems in different sectors.

A general model of risk management processes data from a network of information processing modules, which includes sensors, information fusion, decision support, as well as automated decision-making nodes. Direct signals from threat factor measurement units, be they automated, supervised or human-operated, serve as inputs. This network is complemented by the risk management and optimization model involving decision algorithms, actions and actuators to implement them. Actionable feedback is directly provided to decision makers by the same information system and processed in the IDSS. The decision support model contains the following components [13]:

(i) sensors and other information gathering units on all threats, natural or anthropogenic
(ii) two types of information fusion nodes: those capable of fusing information of the same kind from different sensors (simple fusion), and complex fusion nodes capable of processing heterogeneous information
(iii) endangered humans, who may face threats to their health or lives, as well as equipment, vehicles, buildings and other objects linked with threat propagation relation
(iv) lower, middle and top-level decision units, artificial or human, and the decision support module applying causal and anticipatory decision models at the middle and top levels
(v) first responder teams, robots and other equipment that all implement emergency management decisions.

The model objects and relations between them can be represented as a dynamic directed multigraph, with three types of edges that denote information flow, threat propagation and impact, as well as decision transmission as commands.

The implementation of state-of-the-art AI techniques is particularly relevant to support the components (i), (ii) and (iv) of the IRM DSS architecture. The AI alignment problem, based on a prior needs analysis, has been formulated in [12] for the IRM DSS within the context of an integrated EIS via the following principles:

1. Identify the best AI technologies to be used in an EIS (including IRM DSS) and implement them in the enterprise to build resilience to external threats, caused by adversarial AI or not, and keep competitiveness at a satisfactory level.
2. Keep the AI implementation level in an EIS at a level that ensures at least an equally strong response of this enterprise to the challenges and threats created by external agents with adversarial AI.

For example, principle (2) requires that the capability of sensors to detect intrusions with intelligent drones outranks the ability of malicious AI used by intruders to hide themselves.

A general IDSS design procedure involving in-the-loop technology implementation is proposed in [12]. Here, we detail this approach to the design an IRM DSS for an open-cast mine with an AI-aligned portfolio of evacuation algorithms, information fusion, scene understanding and anticipatory decision models (cf. section 4). The alignment procedure uses expert Delphi, bibliometric trend analysis, and web scanning. A functional scheme of software architecture design based on AI alignment principles applied to the IRM DSS components (i), (ii) and (iv) for the above application is shown in Fig. 1.

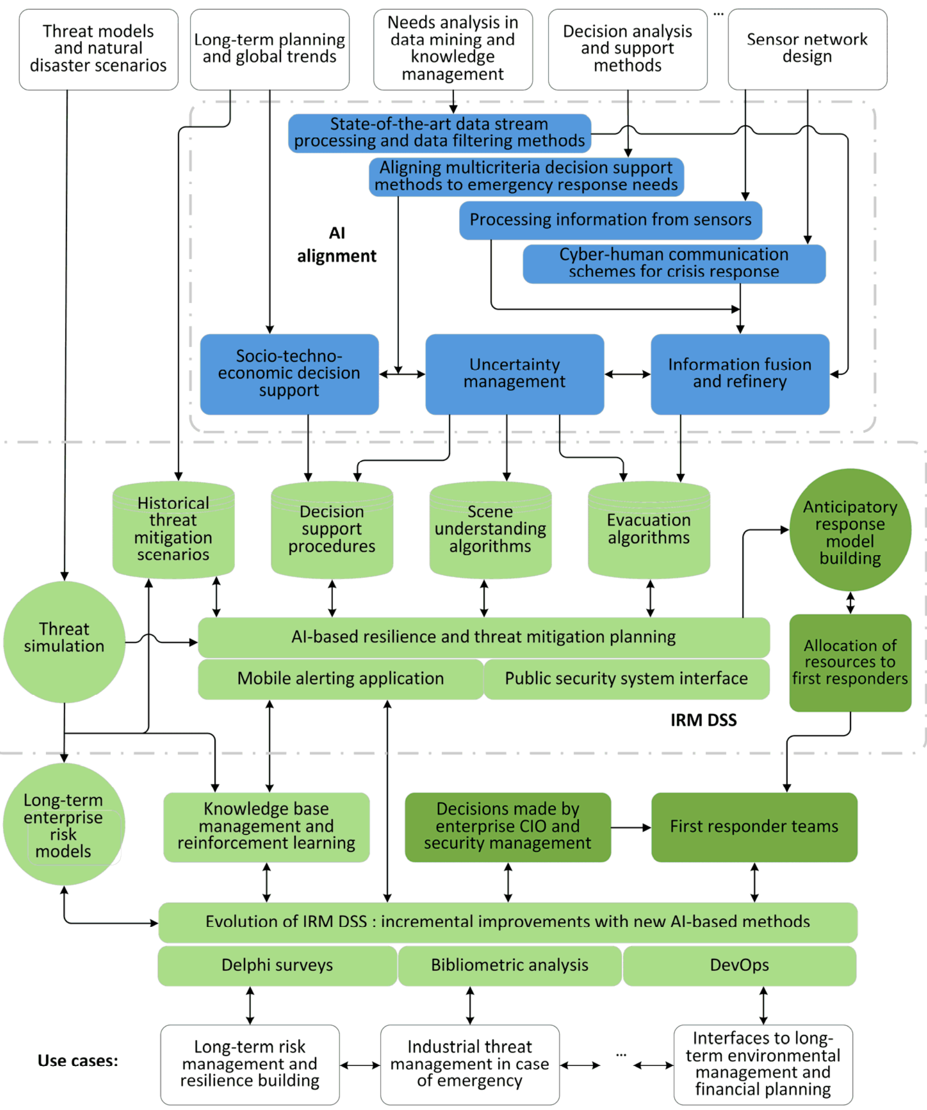

Fig. 1. A generic software architecture of the IRM DSS and an organization diagram of AI alignment

4. Application of Causal and Anticipatory Models in an IRM DSS

To reflect the complex character of decision making, in case of emergency, and to design efficient management procedures that could optimize quantitative criteria, we applied the anticipatory decision making scheme embedded in an anticipatory network (AN). According to the definition presented in [11], an AN is a directed multigraph without cycles in each of its components. An example of a realistic AN decision structure referring to the IRM DSS design for a limestone mine in Poland is shown in Fig. 2.

The nodes O_i in Fig. 2 correspond to the decision makers. The first graph (red edges annotated ψ_{ij}) is the impact relation, where O_i at higher decision level defines constraints on the decisions allowed to be made by O_j. This is denoted as $\psi_{ij}(u_{i,k})=\{u_{j,p1},\ldots,u_{j,p(k)}\}\subset U_j$, where $u_{i,k}$ is a certain decision made by O_i and U_j is the set of all admissible decisions of O_j. Unequivocal commands correspond to $p(k)=1$, but this is an exception, as the main idea behind the above structure is to allow subordinated decision makers a certain level of freedom in unexpected circumstances. This freedom is transferred to lower-level units, rescue teams and actuators. The second subgraph (blue edges annotated f_{ji}) models anticipatory feedback, specifically each f_{ji} defines the subset V_{ji} of U_j, which contains decisions solicited by O_i. V_{ji} rarely overlaps with any constraint $\psi_{ij}(u_{i,k})$, so O_i strives to choose a decision that substantiates the selection of an element of V_{ji} by O_j.

The IRM DSS design principle behind anticipatory modelling consists of designing a command structure $\psi_{i,j}$ so that anticipatory feedbacks f_{ji} are satisfied to the maximum possible extent, for example, by a maximum number of them. An optimal sequence of decisions at each level is then to be found. In-depth quantitative analysis of decision scenarios in ANs with different coordination levels expressed by the functions ψ_{ij} will be a subject of our future research.

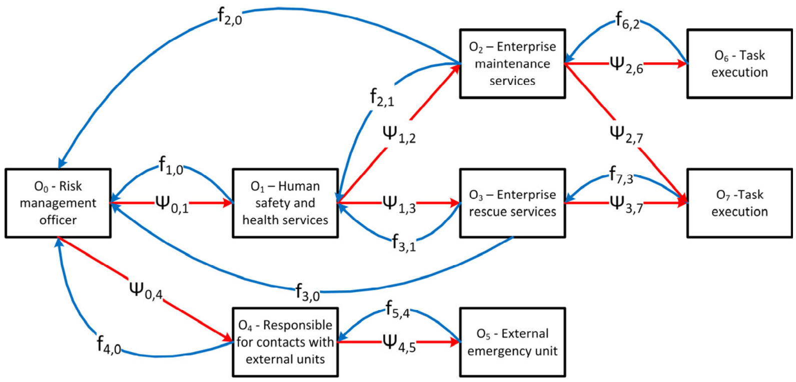

Fig. 2. The anticipatory network applied to the design of decision making procedures in an IRM DSS

5. Conclusions

Industrial risk management systems will no longer be stand-alone applications. When designing such a system for a large-scale industrial enterprise, several salient development trends have been detected. The transition from systems focused on information provision, visualisation and presentation to decision makers to intelligent, partly autonomous DSSs, has been identified from bibliographic scans, as global trends already indicated at the pre-design stage.

Further trends related to the uses of AI in an IRM DSS, which have been confirmed during a recent Delphi survey targeted at DSSs and IRM DSSs, and applied when approving the design process principles, are presented below.

- AI alignment principles can be combined with DevOps and sequentially used according to the scheme provided in section 3.
- Threat models and decision rules for the current emergency response can be learned within a semi-supervised ML procedure, with the data labels used for training inferred from the outcomes of prior solutions to similar problems.
- The proposed AI-based DSS design approach can provide viable implementations, capable of solving heterogeneous industrial threat management problems in real time.

- The deployment of anticipatory networks in IRM-DSSs ensures efficient and flexible control of threat mitigating action in case of an emergency.

In industrial plants, where there is a need to manage natural hazards with the plant's own capabilities, it is necessary to provide both appropriate hardware instrumentation and intelligent decision-making procedures. Further research on AI-based industrial risk assessment and mitigation will be focused on the analysis and selection of decision support methods and the final architecture of the IDSS. The specificity of threats and risks that may occur in an enterprise implementing an IRM DSS will require more penetrative risk modelling and risk response with appropriate preventive and mitigation actions.

References

1. Bayar, N., Darmoul, S., Hajri-Gabouj, S., Pierreval, H.: An immune based ontology implementation of hazard and operability analysis in production systems. Appl. Mech. Mater. 575, 884-894 (2014). DOI: 10.4028/www.scientific.net/AMM.575.884
2. Bevilacqua, M., Ciarapica, F.E., Paciarotti, C.: Business Process Reengineering of emergency management procedures: A case study. Safety Sci. 50 (5), 1368-1376 (2012). DOI: 10.1016/j.ssci.2012.01.002
3. Domdouzis, K.: Artificial-intelligence-based service-oriented architectures (SOAs) for crisis management. In: Handbook of Research on Investigations in Artificial Life Research and Development, pp. 79–95. IGI (2018). DOI: 10.4018/978-1-5225-5396-0.ch005
4. Garbolino, E., Chéry, J.-P., Guarnieri, F.: Systems dynamics applied to the analysis of risk at an industrial installation. In: Safety Dynamics, Evaluating Risk in Complex Industrial Systems, pp. 31-91. Springer, Advanced Sciences and Technologies for Security Applications (2019). DOI: 10.1007/978-3-319-96259-7_2
5. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research, MIS Quarterly 28 (1), 75-105 (2004). https://www.jstor.org/stable/25148625
6. Ibrahim, A.M., Venkat, I., Subramanian, K.G., Khader, A.T., De Wilde, P.: Intelligent Evacuation Management Systems: A Review. ACM Trans. Intell. Syst. Technol. 7, Art. 36, p. 27 (2016). DOI: 10.1145/2842630
7. Middleton, S. E., Middleton, L., Modafferi, S.: Real-time crisis mapping of natural disasters using social media. IEEE Intell. Syst. 29(2), 9-17 (2014). DOI:10.1109/MIS.2013.126
8. Porte, J., Briones, A., Maso, J.M., Pares, C., Zaballos, A., Pijoan, J.L.: Heterogeneous wireless IoT architecture for natural disaster monitorization. EURASIP J Wireless Com Network 2020, art. 184, p.27 (2020). DOI: 10.1186/s13638-020-01793-3
9. Rönnbäck, L., Holmström, J.: Running to stand still: Examining the role of information technology in industrial risk management. In: 16th European Conference on Information Systems (ECIS), 12 p. (2008)
10. Seppanen, H., Virrantaus, K.: Shared situational awareness and information quality in disaster management. Safety Sci. 77, 112-122 (2015). DOI: 10.1016/j.ssci.2015.03.018
11. Skulimowski, A.M.J.: Anticipatory Network Models of Multicriteria Decision-Making Processes. Int. J. Systems Sci. 45(1), 39-59 (2014). DOI: 10.1080/00207721.2012.670308
12. Skulimowski, A.M.J., Bañuls, V.A.: AI Alignment of Disaster Resilience Management Support Systems. In: Rutkowski, L. et al. (eds) Artificial Intelligence and Soft Computing. ICAISC 2021. Lect. Notes in Artif. Intell. 12855, pp. 354-366. Springer (2021)
13. Skulimowski, A.M.J., Łydek P.: AI-based design of decision support systems for industrial risk management. In: PP-RAI'2022: Proceedings of the 3rd Polish Conference on Artificial intelligence, April 25–27, 2022, Gdynia, Poland. Gdynia Maritime University Press, pp. 138-142 (2022)
14. van Engelen, J.E., Hoos, H.H.: A survey on semi-supervised learning. Mach. Learn. 109, 373–440 (2020). https://doi.org/10.1007/s10994-019-05855-6
15. Vernez, D., Buchs, D.R., Pierrehumbert, G.E., Besrour, A.: MORM - A Petri net based model for assessing OH&S risks in industrial processes: Modeling qualitative aspects. Risk Analysis 24(6), 1719-1735 (2004). DOI: 10.1111/j.0272-4332.2004.00562.x