DOI: 10.1111/risa.14038

ORIGINAL ARTICLE

Disaster risk and artificial intelligence: A framework to characterize conceptual synergies and future opportunities

Shital Thekdi[1] | **Unal Tatar**[2] | **Joost Santos**[3] | **Samrat Chatterjee**[4]

[1]Department of Analytics & Operations, Robins School of Business, University of Richmond, Richmond, Virginia, USA

[2]Department of Cybersecurity, University at Albany, State University of New York, Albany, New York, USA

[3]Department of Engineering Management and Systems Engineering, George Washington University, Washington, District of Columbia, USA

[4]Data Sciences and Machine Intelligence Group, Physical and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland, Washington, USA

Correspondence
Shital Thekdi, Robins School of Business, University of Richmond, 102 UR Drive, Richmond, VA, USA.
Email: sthekdi@richmond.edu

Abstract

Artificial intelligence (AI) methods have revolutionized and redefined the landscape of data analysis in business, healthcare, and technology. These methods have innovated the applied mathematics, computer science, and engineering fields and are showing considerable potential for risk science, especially in the disaster risk domain. The disaster risk field has yet to define itself as a necessary application domain for AI implementation by defining how to responsibly balance AI and disaster risk. (1) How is AI being used for disaster risk applications; and how are these applications addressing the principles and assumptions of risk science, (2) What are the benefits of AI being used for risk applications; and what are the benefits of applying risk principles and assumptions for AI-based applications, (3) What are the synergies between AI and risk science applications, and (4) What are the characteristics of effective use of fundamental risk principles and assumptions for AI-based applications? This study develops and disseminates an online survey questionnaire that leverages expertise from risk and AI professionals to identify the most important characteristics related to AI and risk, then presents a framework for gauging how AI and disaster risk can be balanced. This study is the first to develop a classification system for applying risk principles for AI-based applications. This classification contributes to understanding of AI and risk by exploring how AI can be used to manage risk, how AI methods introduce new or additional risk, and whether fundamental risk principles and assumptions are sufficient for AI-based applications.

KEYWORDS
artificial intelligence, machine learning, risk analysis, risk management

1 | INTRODUCTION

Risk events involving physical and cyber-infrastructure are becoming increasingly common and complex. In February 2021, a massive energy, water, and communication infrastructure failure in Texas, USA, resulted from a rare winter weather event. The perfect storm confluence of electricity grid failure, environmental, social, COVID-19, and political factors resulted in a humanitarian crisis. Increased economic interdependencies across regions and nations have also profoundly demonstrated the dire effects of disasters on supply chains. For example, the 2004 Indian Ocean Tsunami severely disrupted the production of cars in Japan and created a cascade of supply shortfall worldwide (McKenzie et al., 2014). Similarly, in May 2021, a cyber breach of Colonial Pipeline led to a widespread fuel outage in the mid-Atlantic United States, with the potential for massive shutdowns in the movement of goods, services, and people (Sanger & Perlroth, 2021). Furthermore, the May 2021 ransomware attack that caused the shutdown of JBS—the largest meat supplier in the world—has shown the vulnerability of the food industry sector to cyberattacks (Batista et al., 2021), and a subsequent Kaseya attack has shown the potential for more widespread outages in infrastructure

and commerce (Bobrowsky, 2021). While the precedence for these widespread attacks is apparent, there is potential for much more disastrous outcomes.

Artificial intelligence (AI) has the potential to be widely used to manage operations for a variety of systems, such as infrastructure, cybersecurity, and manufacturing. These physical, cyber, and cyber-physical systems provide the foundation for functioning societies. Errors or inadvertent misuse of AI for risk applications of these systems can result in devastating consequences.

The emerging use of AI has the potential to influence the likelihood of consequence of risk events for these systems. However, it is not clear whether AI methods would have prevented or reduced the consequences of the risk events described above. Existing research has not yet fully defined how AI models interact with the ability to manage risk; and how effective risk principles are when applied to systems operated by AI technologies. As failures of AI-driven systems can be severe, potentially exceeding the impacts of other types of disasters typically studied in the risk discipline, it is imperative to question and examine this relationship between AI and risk.

While the use of AI for infrastructure systems is relatively new and developing, the concept of AI has been in existence as early as the 1930s. The birth of AI as a scientific method can be attributed to Turing (1950), as well as the 1956 conference held in Dartmouth College, where it was more formally launched (Appenzeller, 2017). The recent popularity of AI is primarily driven by the increase in technological capabilities, while the algorithms and methods commonly used by AI have been in existence for decades. As described by Bini (2018), processing power valued at billions of dollars in the 1970s is equivalent to relatively inexpensive technologies today. To date, innovations and applications of AI methods have become pervasive, notably in the domains of computing, healthcare, and manufacturing. Despite its rising utility and popularity, several studies have asserted that AI remains underutilized and has not yet reached its full potential (Bhattacharya & Singh, 2020).

The United Nations Office for Disaster Risk Reduction defines disaster risk management as "the application of disaster risk reduction policies and strategies to prevent new disaster risk, reduce existing disaster risk and manage residual risk, contributing to the strengthening of resilience and reduction of disaster losses" (UNDRR, 2022b). Disaster risk applications may involve "immediate and localized" events with varying duration (UNDRR, 2022a). Disaster risk management is a complex activity with a need for fast analysis, regional considerations, and network complexities, such as supply chain implications. Given the automated and data-centric properties of AI, there is a large opportunity for AI to be widely used in recent disaster risk management applications (Sun et al., 2020). However, there is a need for a framework that can identify gaps and opportunities for using AI to analyze and manage disaster risk. Novel AI algorithms coupled with high-performance computing capabilities can be utilized to more accurately predict the geospatial and temporal patterns of disasters as they occur (Van Heteren et al., 2020). Indeed, a more in-depth study is needed to explore the potential benefits of more extensive deployment of AI methods in the field of risk analysis to address the rising magnitude and severity of natural and human-induced disasters.

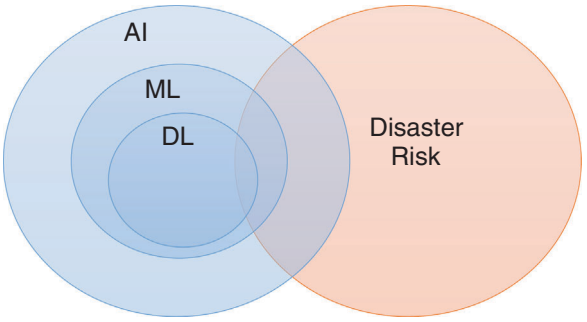

FIGURE 1 Comparison between artificial intelligence methods and risk principles

AI consists of algorithms that are implemented using robust datasets and technologies (IBM, 2022a). In comparison, risk is a broader multidiscipline consisting of principles, methodologies, and application areas (Society for Risk Analysis, 2018), as shown in Figure 1. The figure shows that AI methods include subsets of machine learning (ML) and deep learning. The figure also features the term "disaster risk," encompassing both "disaster risk management" and "disaster risk reduction," which are commonly used terms in the broader field of risk analysis. Disaster risk is a broad field containing many mathematical methods beyond the use of AI. There is growing interest in increasing the overlap between AI and disaster risk, as represented by the overlap in the AI and Disaster Risk circles in Figure 1 (Kuglitsch et al., 2022).

More specifically, AI consists of models and algorithms that can perceive data, such as video, text, and images, and act to achieve some goal. The subset of ML involves using algorithms and statistical models to find inferences and patterns from data without explicit instructions. The subset of deep learning involves using a multilayered approach that can be supervised, semi-supervised, or unsupervised, leveraging artificial neural networks to perform an ML task. In set-theoretic terms, AI methods are broadly considered to contain a subset of ML methods, with ML encompassing a large portion of AI usage. Further, ML contains subsets of supervised, unsupervised, reinforcement, and deep learning (IBM, 2022a, 2022b).

This study uses expertise from risk and AI professionals to understand the main characteristics of a complementary risk-AI approach that can leverage the strengths of both disciplines while also exploring future opportunities to increase the synergy between risk and AI. Using this input, we develop a framework to understand how to use AI to manage risk, what new or additional risk is introduced through AI methods, and whether fundamental risk principles and assumptions are sufficient for AI-based applications. This classification system will help the risk professional identify

inconsistencies between risk and AI, particularly in developing new methods and models that leverage AI. There is a need for a common framework to evaluate whether basic assumptions are being met in a way that is comparable to an analysis of assumptions and residuals in a statistical study.

This study builds on other recent work in the risk discipline. Most notably, Guikema (2020) describes the potential to leverage AI methods for risk analysis of natural disasters, as well as the associated research needs. This study further extends and validates the potential opportunities that can result from the integration of the fields of AI and disaster risk analysis (encompassing not only natural disasters but also human-induced disasters). There are two significant contributions rendered by the current study, which complement and extend the findings by Guikema (2020). First, we conducted a literature search analysis to document the different knowledge domains and application areas associated with AI and disaster risk analysis. In so doing, gaps and opportunities are more explicitly discovered at the intersection of AI and disaster risk topics. Second, the findings from the literature search enabled the development of a survey questionnaire that seeks to validate the benefits of synergizing the fields of AI and disaster risk analysis. As will be discussed in subsequent sections of this study, a primary criterion for the selection of the survey respondents is their direct experience and subject matter expertise in either the field of AI or disaster risk analysis, or both.

Hence, this study discusses the issues presented above by conducting the following tasks: Section 2 discusses how AI methods are used for risk-related applications. Section 3 discusses how AI is conceptually applicable for risk applications; and how risk principles can be applied to AI methods. Section 4 presents a framework to gauge compatibility between risk and AI, and applies that framework to a survey of risk and analytics experts. Section 5 discusses conclusions and opportunities for further research.

2 | HOW RISK-RELATED APPLICATIONS USE AI METHODS

This section describes a literature synthesis of how disaster risk applications use AI methods. First, we identify the volume of publications covering topics of AI and disaster risk by using a query-based keyword analysis. Second, we identify the topical categories of journals publishing papers in the areas of AI and disaster risk. Finally, we conduct an AI-based topic analysis (using a generative statistical model in the natural language processing domain) of published research article titles and abstracts to identify the main themes for AI and disaster risk research. This analysis is limited by the search engine used, the selected keywords that may not encompass all keywords used across applications, and does not intend to substitute a manual review, as could be conducted in a more formal review paper. Instead, the goal is to provide a data-driven overview of AI's use for disaster risk, identifying themes that can be used for the characteristics studied in Section 3.

A search of academic publications (Web of Science, 2021) using the query phrases ''*artificial intelligence* AND *disaster risk*'' yielded 88 records for the years 2016–2021. The number of records published each year grew from 2 in 2016 to 30 in the first nine months of 2021. Among these records, most articles were published in the *International Journal of Disaster Risk Reduction* (5) and *Sustainability* (5), with two publications in *Risk Analysis*. Other key journals include *Geomatics Natural Hazards Risk* (4), *Remote Sensing* (4), and *Water* (4). In terms of Web of Science research categories, the top three were *Environmental Sciences* (26), *Geosciences* (22), and *Water Resources* (22). Notable risk-focused publications include two perspective articles in the *Risk Analysis* journal (Cox, 2020; Guikema, 2020) discussing the potential of AI methods for risk and decision analysis. A primary benefit of AI was shown in the domain of consequence assessment. Challenges identified for AI included the need for large volumes of training data, validation, and uncertainty analysis.

Next, a search (Web of Science, 2021) using the query phrases ''*machine learning* AND *disaster risk*'' yielded 276 records for the years 2016–2021. The number of records published each year grew from 6 in 2016 to 96 in the first nine months of 2021. Among these records, most articles were published in the *International Journal of Disaster Risk Reduction* (28), with five publications in *Risk Analysis*. Other key journals include *Remote Sensing* (19), *Stochastic Environmental Research and Risk Assessment* (9), and *Water* (9). In terms of Web of Science research categories, the top three were *Geosciences* (113), *Water Resources* (86), and *Environmental Sciences* (76). The five articles in the *Risk Analysis* journal focus on bias correction for damage modeling (Wagenaar et al., 2021), misinformation monitoring (Hunt et al., 2020), perspectives on risk analysis for open-world novelty (Cox, 2020), resilience analytics (Eisenberg et al., 2019), and prediction of flash flood impacts (Terti et al., 2019).

Finally, a search (Web of Science, 2021) using the query phrases ''*deep learning* AND *disaster risk*'' yielded 95 records for the years 2016–2021. The number of records published each year grew from 1 in 2016 to 41 in the first nine months of 2021. Among these records, most articles were published in the *International Journal of Disaster Risk Reduction* (9), with no publications in *Risk Analysis*. Other key journals include *Remote Sensing* (5), *Water* (4), and *Applied Sciences Basel* (4). In terms of Web of Science research categories, the top three were again *Geosciences* (30), *Water Resources* (27), and *Environmental Sciences* (24).

While *Risk Analysis* journal does not appear in the third search result above, *International Journal of Disaster Risk Reduction* is at the top of all three search results. *Remote Sensing* and *Water* constitute 14.7%–20.2% of records identified. Moreover, across all search results above, the Web of Science categories of *Geosciences*, *Water Resources*, and *Environmental Sciences* collectively comprise 79.5%–99.6% of records identified. Search results with *machine learning* yielding more records (276) than *artificial intelligence* (88)

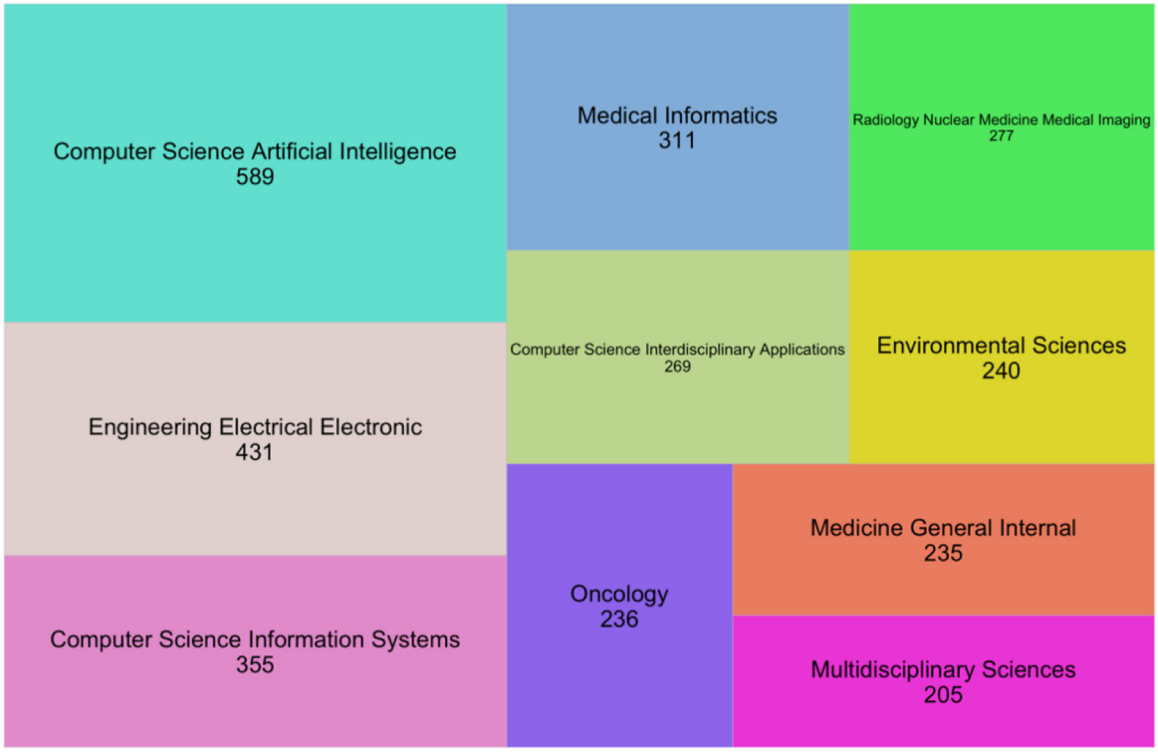

FIGURE 2 Topic areas for journals publishing risk and artificial intelligence (AI) research (Note: Topics here represent the top 10 Web of Science categories with the number of records returned under each category)

or *deep learning* (95) possibly suggest a more prevalent use of *machine learning* in the disaster risk literature.

In contrast to 88 records that were returned with the search query ''*artificial intelligence* AND *disaster risk*,'' results using a broader search query ''*artificial intelligence* AND *risk*'' yielded 4379 records for the years 2016–2021. This significant increase in records suggests broader use of the *artificial intelligence* terminology in the overall risk literature. Figure 2 presents the top 10 Web of Science topic categories for journals publishing risk and AI research for the years 2016–2021. The results show that much of the research is categorized under computer science and engineering domains, emphasizing healthcare applications. The journals publishing this work also indicate the healthcare emphasis where the top 10 (by frequency) venues include *Artificial Intelligence in Medicine*, *Journal of Medical Internet Research*, *International Journal of Environmental Research and Public Health*, and *Frontiers in Oncology*.

Following the query-based keyword analysis, we conducted a topic analysis with published research article titles and abstracts at the intersection of AI and disaster risk. Using the methods of natural language processing, the Latent Dirichlet Allocation (LDA) is a broadly used generative probabilistic model. LDA uses Dirichlet and multinomial distributions to estimate the probability distribution of a document corpus over latent topics (Blei et al., 2003). A key assumption is that text documents are considered as a mixture of topics, each of which are defined by a distribution of terms. Mathematically, defining a word w as a discrete data item, a document as a sequence of N words $w = (w_1, \ldots, w_N)$, and a corpus as a collection of M documents represented as $D = \{w_1, \ldots, w_M\}$, LDA assumes the generative process below for every document w in a corpus $\boldsymbol{D}$ (Blei, Ng & Jordan, 2003):

Step 1: Select $N \sim$ Poisson(ξ) with parameter ξ;
Step 2: Select $\Theta \sim$ Dirichlet(α) with parameter α;
Step 3: For each of N words w_n:
Step 3.1: Select a topic $z_n \sim$ Multinomial(Θ).
Step 3.2: Select a word w_n from $p(w_n|z_n, \beta)$, where β is per-topic-per-word probability.

The goal eventually is to estimate classification probability (via maximum likelihood estimation) of a corpus, $\boldsymbol{D}$ over latent topics by multiplying the marginal probabilities of single documents as:

$$p(D|\alpha, \beta) = \prod_{d=1}^{M} \int p(\Theta_d|\alpha) \left(\prod_{n=1}^{N_d} \sum_{z_{dn}} p(z_{dn}|\Theta_d) p(w_{dn}|z_{dn}, \beta) \right) d\Theta_d.$$

In this study, we utilized LDA to search, understand, and organize similar topics within a set of query results from a large electronic database, such as Web of Science. The LDA method was implemented using the "topicmodels" package in open-source *R* programming language for statistical computing and graphics (Grün et al., 2021; R Core Team, 2017).

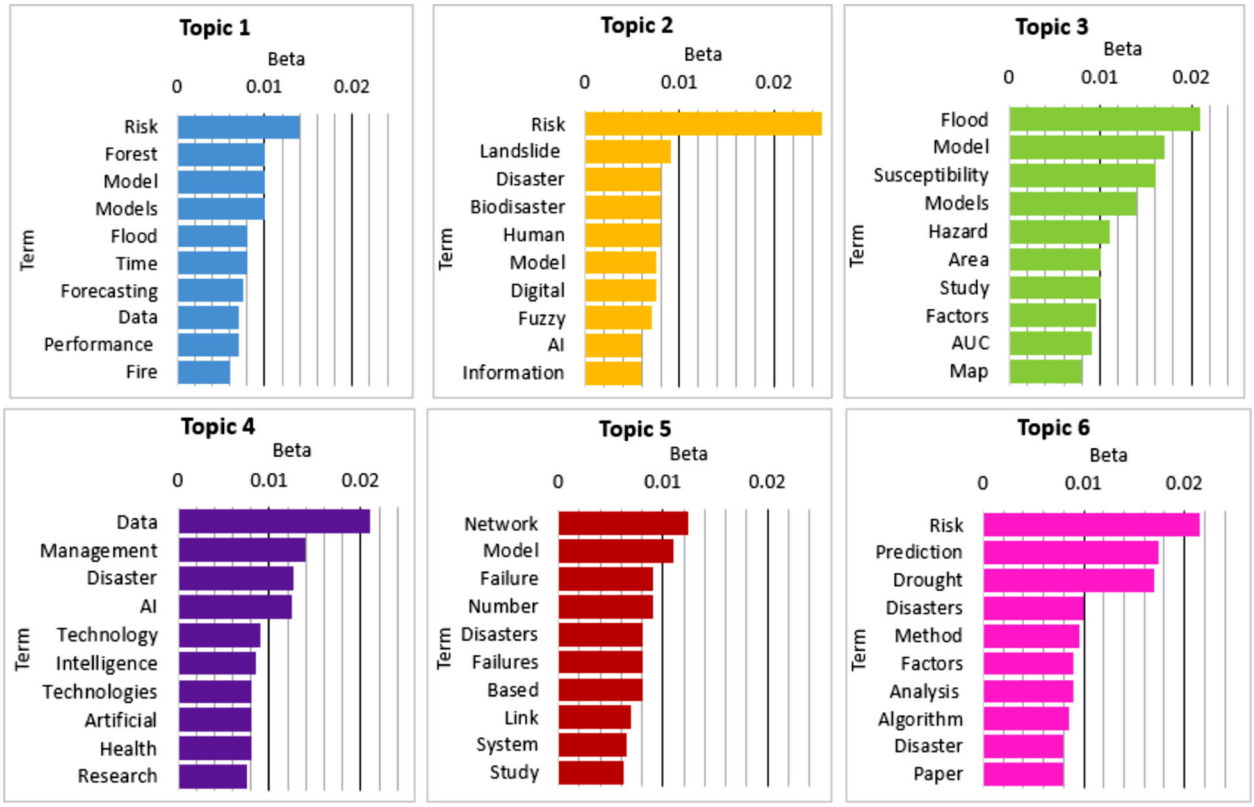

FIGURE 3 Topic analysis (using LDA method) with published articles titles and abstracts related to disaster risk and artificial intelligence (AI) (Note: Each of the six subplots refers to a topic with top-10 terms or words (*y*-axis) and beta (*x*-axis) refers to an estimate of per-topic-per-word probabilities)

While the LDA analysis includes uncertainties in the outputs, it is used to provide a data-driven understanding of themes that can be used to define the characteristics presented in Section 3 of this study.

Figure 3 shows the results of an LDA topic analysis with 88 document titles and abstracts (Web of Science, 2021) related to AI and disaster risk. Here, we present six topics with the top 10 words under each topic and their associated probabilities. Because the LDA analysis did not include word-stemming, terms like "model" and "models" are treated as different terms. Note that in this case, each record or document is a mixture among these six topics, and each topic comprises a distribution of terms or words. Synthesizing unique key terms from each topic, we summarize: *Topic 1* with emphasis on risk, modeling and forecasting, and fires; *Topic 2* with emphasis on risk, landslide, bio-disaster, and human impacts; *Topic 3* with emphasis on flood modeling; *Topic 4* with emphasis on data management and AI technology; *Topic 5* with emphasis on network and system modeling; and *Topic 6* with emphasis on risk, prediction, and droughts.

Figure 4 presents the frequency distribution of the most, second most, and third most likely topics across 88 documents. We observe that *Topic 4*, with emphasis on data management and AI technology, is most frequent at all likely topic levels. The next frequent topic in the most likely category is *Topic 5*, with emphasis on network and system modeling, second most likely category are *Topics 2 and 3*, with emphasis on flood modeling, risk, landslide, bio-disaster, and human impacts; and third most likely category is *Topic 6* with emphasis on risk, prediction, and droughts. Moreover, across 88 documents, we observe diverse ordering of likely topics estimated for each document (e.g., a publication titled "*Artificial Intelligence and Geospatial Analysis in Disaster Management*" has an estimated likely topic ordering from most to least likely as [*Topic 4*, *Topic 3*, *Topic 5*, *Topic 2, Topic 6, Topic 1*] which is consistent with AI and flood modeling emphasis).

The results above suggest that using document titles and abstracts, there is a blend of risk and AI terms; however, risk terms mostly indicate disaster types rather than specific elements from the risk triplet (Kaplan & Garrick, 1981), and AI terms indicate the use of data, models, and algorithms but not specific techniques. We also observe mention of disaster risk research in the domains of floods, landslides, fires, and droughts. It can be hypothesized that these types of disasters may involve a relatively high availability of historical data. In general, AI methods tend to be more mature for scenarios with high availability of geodatabases (Van Heteren et al., 2020). This may also be because natural hazards are relatively more predictable than human-induced disasters.

There was also an unexpectedly low volume of topics related to terrorism/human-induced/cyberattacks. This could be hypothesized as resulting from trade secrets held in private industry or confidential data related to national security. This

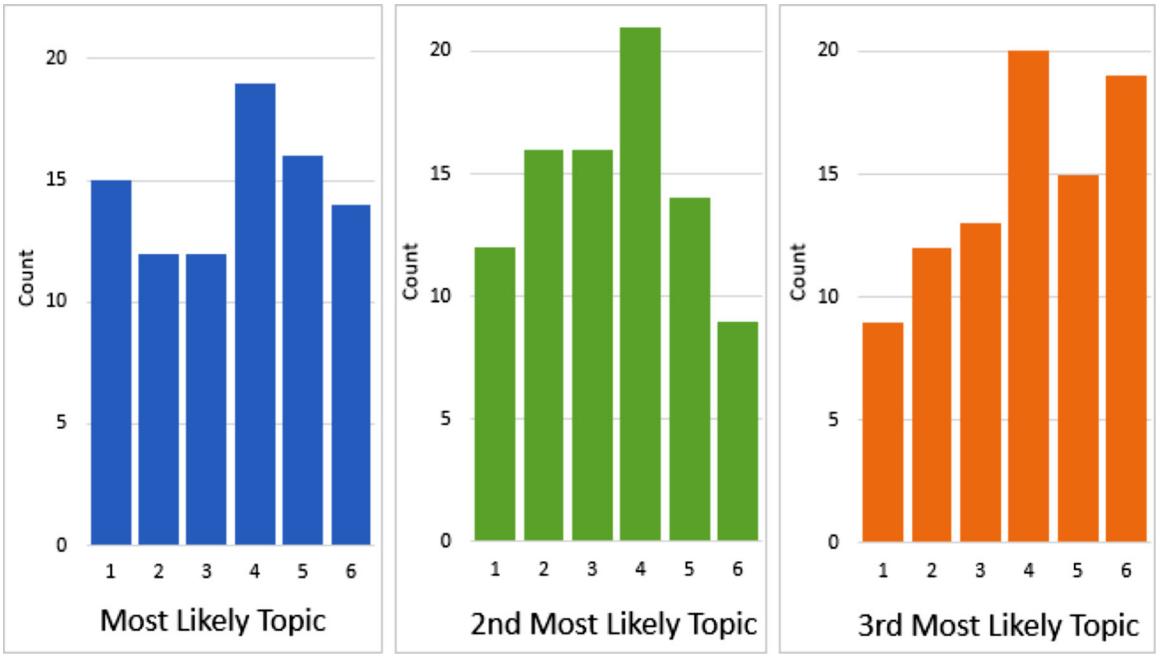

FIGURE 4 Frequency distribution of most likely topics across documents

is particularly notable for the academic community, as the innovations may be occurring in isolation from peer review. There also may be low availability of geospatial data and temporal data for these types of applications. Research on these topic areas may be primarily in developmental stages. It can also be observed from the results that the use of AI methods for disaster risk applications has become more prevalent. In contrast, the use of risk methods for AI applications is relatively sparse; hence, this provides an opportunity to synergize the strengths of these two fields.

Overall, the analysis in this section showed a surprising scarcity in the use of AI methods for broader disaster risk applications. In particular, there were relatively few articles using AI to study risk for areas of infrastructure and commerce, which are areas that can involve catastrophic consequences and would benefit most from an enhanced study of risk. Thus, there is a demonstrated need for increased momentum for leveraging a balance between AI and disaster risk.

3 | COMPLEMENTARY CHARACTERISTICS OF AI METHODS AND THE RISK DISCIPLINE

This section identifies themes related to the balance between risk and AI to develop a set of characteristics describing the use of AI methods for risk applications, how the risk field can be beneficial for the utilization of AI methods, and implications for increasing the adoption of AI methods in risk applications. The characteristics to be discussed in the following sections are sourced from the literature synthesis in Section 2, a review of other relevant literature, current practices in risk and AI, and recent events.

The following sections relate AI and risk. Section 3.1 describes how AI can be leveraged for risk applications, including elements related to evidence/data and automation. Section 3.2 describes how risk principles relate to the use of AI methods, including elements of uncertainty and the incorporation of domain knowledge. Section 3.3 describes the implications for increasing the adoption of AI methods in risk applications.

3.1 | AI as a method for risk applications

There are several benefits to using AI as a method for risk applications. These benefits are as follows:

A1: Leverage evidence/data: AI methods are useful because they offer fast processing of large quantities of data. This is important as risk applications are increasingly based on big data resources (Choi & Lambert, 2017; Nateghi & Aven, 2021). Tasks that require substantial amounts of labor can be automated with impressively accurate results. For example, AI can be used for tasks such as decoding themes in social media data, monitoring traffic flows, and balancing electricity supply and demand (Anderson et al., 2018; Kenyon, 2021; Wolfe, 2017). These types of tasks require some level of human supervision, but the supervision occurs at a managerial level. This type of automation enables effective situational awareness to predict disasters, manage operations during disasters, and manage recovery efforts. Guikema (2020) provides a comprehensive discussion of how AI is used in various disaster risk applications.

A2: Automate difficult and labor-intensive tasks: Tasks that are repetitive can be outsourced to automation (Chui et al., 2015). Additionally, AI methods can make these types of tasks even more efficient and cost effective (Viechnicki & Eggers, 2017). For example, AI methods for image detection are becoming faster and more accurate than human abilities. Consider a healthcare setting in which AI algorithms can detect certain cancers. This practice could have large implications for early detection of cancers, quality of care, and cost effectiveness of care. Also, consider the use of AI methods for disaster avoidance and response, such as detecting damage to infrastructures, prioritizing emergency response, and detecting unsafe conditions in new or dangerous environments. For example, recent drone and sensor technologies have the potential to detect wildfires with 99% accuracy (Hampson, 2021). As another example, consider the use of LIDAR and AI algorithms to evaluate road damage and conduct optimal routing after a hurricane (Foy, 2020). Thus, automation has the potential to allow human resources to be more efficiently dedicated toward tasks of human judgment, incorporation of values, supervision, decision-making, and communication.

A3: Use evidence to make objective assessments and decisions: AI and other algorithmic models demonstrate objectivity as decisions and actions are based on a set of data-driven rules and processes. This level of objectivity allows actions and decisions to be reproducible and transparent, which can promote trust and acceptance of these systems. An article by Kuglitsch et al. (2022) proposed a coupled AI and disaster risk reduction lifecycle framework featuring a feedback loop comprising data collection, model development, and model deployment. While this can be useful in some contexts, one can question whether a rule-based process is always the best choice, particularly in situations involving noncommensurate objectives or situations in which social issues, fairness, trust, reputation, and other nonquantitative objectives are critical for decision-making. For example, following much discussion of these issues, there is still no clear best practice for methods to address bias and fairness (Ntoutsi et al., 2019).

3.2 | Risk discipline for utilization of AI methods

Using AI to manage infrastructure systems is not without risk. Overreliance on AI methods can introduce a variety of risk issues. There continues to be a need for risk principles in AI applications. The following characteristics demonstrate how AI and risk can complement each other:

R1: Address uncertainty in analysis: AI methods benefit from the ability to recognize and embrace uncertainty. For example, there is a benefit to layering AI models with a risk management overview, such as in accordance with ISO's principle of understanding the effect of uncertainty on objectives (ISO, 2018). This type of judgment forces the users of AI methods to carefully consider the limitations of the available data, information not being captured in the data, and implicit assumptions being used in the AI methods. For example, because AI models can be self-learning, these models should consider how the applicability and accuracy of the methods change when the characteristics of the system change. The literature review found no systematic standards for the appropriate level of uncertainty-focused supervision and human intervention in this type of assessment.

R2: Utilize a holistic risk approach: Within any risk application, it is imperative to understand the context (ISO, 2018). It is possible for overreliance on AI to reduce human supervision and judgment opportunities, which may be critical in surprise or improbable event scenarios. AI-based system decision-making can allow for human judgment and to aid in structured decision-making processes, as seen in NHTSA's structuring of levels of automation in vehicles (Campbell et al., 2018). This type of coordinated AI and risk hybrid approach can be a new regime in human-automated decision-making partnerships that can make both the AI and risk domains stronger, synergistic, and more applicable to emerging application areas.

R3: Balance domain knowledge and data: AI methods benefit from improving how to incorporate domain knowledge and data effectively. Data represent a simplification of conditions, spaces, and environments, and data cannot encapsulate all essential aspects of a system. The risk discipline has emphasized the importance of incorporating varying levels of knowledge, insufficient knowledge, belief, and expert elicitation into quantitative models (Thekdi & Aven, 2016). The risk discipline has the opportunity to adapt existing principles to AI methods that are now trusted with performing real-time analysis and actions.

3.3 | Implications for increasing the adoption of AI methods in risk applications

The increasing adoption of AI has major implications for the risk discipline. It can be argued that the adoption of AI methods is introducing *new and additional* risks to systems while possibly reducing others, as explained below. The main themes are as follows:

B1: Balance automation with human supervision: There is a delicate balance between tasks performed by humans versus mathematical methods. While progress has been made to develop AI algorithms that can mimic and even improve the use of human judgment and decision-making, it is unclear whether algorithms perform effectively in new or surprising situations. This was seen in the tragic Boeing 737 Max disasters, resulting from the AI software malfunctioning as it interacted with damaged sensors (Nicas et al., 2019). There is potential to use extreme value theory (De Haan & Ferreira, 2007) and generative methods (Wang et al., 2017) to help reveal insights for autonomous operations while using limited data and partial knowledge of system behavior during rare events. As a result, analysts can consider the most appropriate level of automation. For example, consider the topic

of autonomous vehicles. Autonomy implies that the vehicle can self-correct and self-realize with minimal reliance on the human. However, there may be a preference to allow the human to have the ability to override, manage cruise control, or choose to autopilot, considering the six levels of autonomy (IEEE, 2021).

It is important to recognize that system designers control the level of human supervision or autonomy needed for a system. For example, consider the context of self-driving vehicles. There may be benefits from having some aspects of operation be automated while other aspects are controlled by the drivers. As another example, consider a cyber intrusion detection system. System designers may prefer not to have some aspects of the system be automated to ensure there is some human oversight in case of an attack.

B2: Utilize a stakeholder-based approach to explainability: Users and decision-makers may not understand the mechanics of the AI methods. Knowledge of mathematical methods, assumptions, and data exist primarily at the level of mathematicians and programmers. It is the programmers who are also developing the dashboards and managerial overview systems, which only share selected aspects of the system and computations. This type of cultural change is unprecedented. How does one manage risk using mathematical methods when the analysts and decision-makers do not fully understand the quantitative elements? This is an underlying issue that the risk field needs to address. The field can look to the topic of explainability, which recognizes that different stakeholders have different needs for understanding these complex AI models (IBM, 2021a).

B3: Conduct a knowledge-based assessment of whether relevant factors are reflected in the data: Because AI interprets and acts using data, these methods can only consider factors within the available data. We can equate nondigitized information to the risk concept of an ''unknown known,'' such that knowledge can be known by other systems and people but is not understood by AI methods. This could result from designers of the algorithms not predicting the need to incorporate some types of data or knowledge or the fundamental issue that not all factors can be reflected in a digitized dataset. Additionally, this type of data can introduce vulnerabilities, such as poisoned data or other data integrity issues.

B4: Support security for systems: There is insufficient evidence to justify whether AI models can be gamed or can manipulate other systems when used to manage risk, such as to detect cybercrime (Bilen & Ozer, 2021). Adversarial machine learning is a method whereby inputs are purposely developed by an attacker to cause an AI system to commit a mistake in its classification. For example, AI-driven systems could do harm when acting without human oversight, potentially causing harm to cyber-physical systems with physical repercussions (Ring, 2015). As another example, AI models can contribute to ongoing issues with misinformation, such as through deep-fake imagery, fake news, or filtering information in ways that mislead or restrict points of view (Osoba & Welser, 2017a).

B5: Manage trust and social impact for systems: Because AI methods are data driven and algorithm driven, there is varying trust in these systems, which can be attributed to issues already discussed in this section. In addition, there are documented issues with AI algorithms exhibiting biases in ways that can promote discrimination (IBM, 2021b ; Parikh et al., 2019) and can "punish citizens for crimes they haven't yet committed" (Asher & Arthur, 2017). Additionally, like any model, AI methods can be wrong. These methods can have false positives and false negatives, with potentially catastrophic repercussions of those errors. Furthermore, with inadequate validation by independent reviewers, the analysis or decisions from these AI methods can be manipulated to promote the biases of the designer (Shneiderman, 2016), or cater to the interest of the general public (Chu et al., 2020). In relation to overall definitions of risk, this issue of trust can influence abilities to evaluate the strength of knowledge in relation to a risk characterization, such as in estimating a likelihood or consequence of a risk event.

B6: Regulate systems: In their current state, public agencies lack "common metrics" to understand how trustworthy current AI methods are in practice (NSCAI, 2021). While carefully designed AI-driven systems can be tested for safety and reliability, there is no guarantee that there is a common set of oversight and standards. For example, in relation to AI-enabled autonomous weapon systems, there is a confidence that nations can be in compliance with International Humanitarian Law. However, there are concerns over whether all nations have procedures that are "responsibly designed and lawfully used" (NSCAI, 2021). As another example, there are privacy and ethical concerns over how the private sector may use AI-sourced information (Jeong, 2019) or impacts general liberties in cases of government surveillance (Osoba & Welser, 2017b). While overregulation of AI can stifle the development and innovations in AI technologies (Information Society Project, 2017), it is apparent that a lack of regulation has serious implications for risk. In a seminal article, Scherer (2015) discussed issues with regulating AI (e.g., discreteness, diffuseness, opacity, foreseeability) and proposed the creation of an agency to implement a rigid certification process for AI system development, as well as a liability system for its sale and use.

B7: Utilize a systems approach for AI-based operations: There is also a need to recognize that AI-driven systems cannot be seen in isolation. The future of AI involves AI-driven systems interacting with one another. It would be assumed that these interacting systems are built on a common set of standards for accuracy, precision, and assumptions. Much like a network model, there is a need to understand the implications of these interacting AI models on overall system risk, vulnerability, and resilience. The more AI systems are trusted to learn and act autonomously, the more these systems can be high-value targets for security breaches. Additionally, there may not be clear logic for how and when AI systems acting in coordination should rely on one another or when one system should override another. Similarly, organizations should have processes for determining when human intervention is

necessary; and train humans to be proficient in intervening in these situations (Cheatham et al., 2019).

B8: Address causality: The foundation of risk involves understanding "why." For example, risk analysts seek to understand why risk events occur, how causal factors influence system safety, or why a particular risk management decision works best for a specific situation. When adopting AI-driven systems, there may not be an adequate justification for why the system behaved in a specific way (Gil & Selman, 2019). Stakeholders may not easily understand the AI-system's explanation of behavior; or in a manner that aligns with human values, emotions, and reasoning.

B9: Manage legal and ethical implications of systems: Finally, there is a need to recognize the implications of AI methods that act in ways that counter the stakeholder interests. In the case of the Texas infrastructure incident, the repercussions resulted in a humanitarian disaster. In the case of drivers trusting AI technologies to manage speed on roadways, the implications for poorly behaved models could involve legal and safety consequences. As another example, consider the Volkswagen emissions scandal involving the misuse of software. In this case, employees acting in their professional roles faced criminal charges (Miller & Matussek, 2018). One can question whether malfunctioning AI within risk models could have legal repercussions for risk analysts acting on behalf of their employers. For more comprehensive discussions on this topic, see Greenblatt (2016) and Kingston (2018).

4 | A FRAMEWORK TO GAUGE COMPATIBILITY BETWEEN RISK AND AI

This section leverages the characteristics developed in Section 3 to define a framework to gauge the balance between AI methods for disaster risk applications. Section 4.1 describes the framework, Section 4.2 describes a data collection exercise for demonstration of the framework, and Section 4.3 presents an analysis of the data collection results.

4.1 | Framework

This framework is designed to be applicable for a new, proposed, or maintained AI-driven system tasked with analyzing or managing risk.

Consider the role of an analyst in choosing the most appropriate analytical method. The analyst typically considers steps related to understanding the context of the problem, selecting an analytical method, re-evaluating modeling assumptions, and then implementing the analysis. Figure 5 shows a framework that includes those steps and the additional step of understanding how the analytical method aligns with risk principles. Considering the most appropriate mathematical method happens concurrently, or in a loop manner, with an understanding of how those methods align with risk principles and assumptions. This framework significantly contrasts with practices that consider risk in a post hoc manner. The following methods will expand on Steps 2 and 3 of the framework in Figure 5 by gauging the previously defined characteristics related to the risk–AI balance.

Based on the discussion in the previous section, there is a strong argument for AI's potential to help address disaster risk challenges. However, there is also a need to consider the new or additional issues that can emerge from the use of AI. As with the use of any new analytical method, there is often concern over the adoption. For example, Apostolakis (2004) describes the phases in the adoption of quantitative risk assessment. First, the community may be skeptical, then familiarity develops, and with time, confidence can be built. One way to develop confidence in a new method is to systematically understand the positive and the negative features of the method, with the eventual goal of addressing any negative features of the method. Thus, we refer to this adaption of addressing negative features as a "balance" between AI and disaster risk.

To evaluate this balance between AI and risk, consider some existing or proposed analytical methods, collection of methods, systems, or practices utilizing AI methods. Analysts, decision-makers, or other stakeholders then form ratings associated with each characteristic presented in Section 3. No characteristic is assumed to be relatively more important than another. There is also flexibility in defining a rating scale, such as a Likert scale, 1–5 rating, or a descriptive scale. There is also an option to assign weights to the characteristics or incorporate the scale into multicriteria decision-making methodologies. However, for the presented framework, there is no assumed weighting. Overperformance for one characteristic may or may not compensate for underperformance in other characteristics.

This framework can further refine main assumptions and mathematical procedures with the intent to intervene if concerns arise. If one or more characteristics are poorly rated, one refers to system managers, owners, and decision-makers to determine whether an intervention is needed. System managers may choose to intervene using a variety of actions, such as by changing or adapting analytical methods, changing user interfaces, or changing the job duties or training for users interacting with these systems.

This framework can be a component of a more comprehensive monitoring and review process that contributes to continuous improvement and learning from issues that arise after system implementation. It may be necessary to regularly re-evaluate systems using this framework as part of more extensive Enterprise Risk Management activities to address new technologies, environments, uses, and operating conditions.

4.2 | Data collection and results

To demonstrate the framework presented in Section 4.1, we consider two different AI-based practices: (1) The current

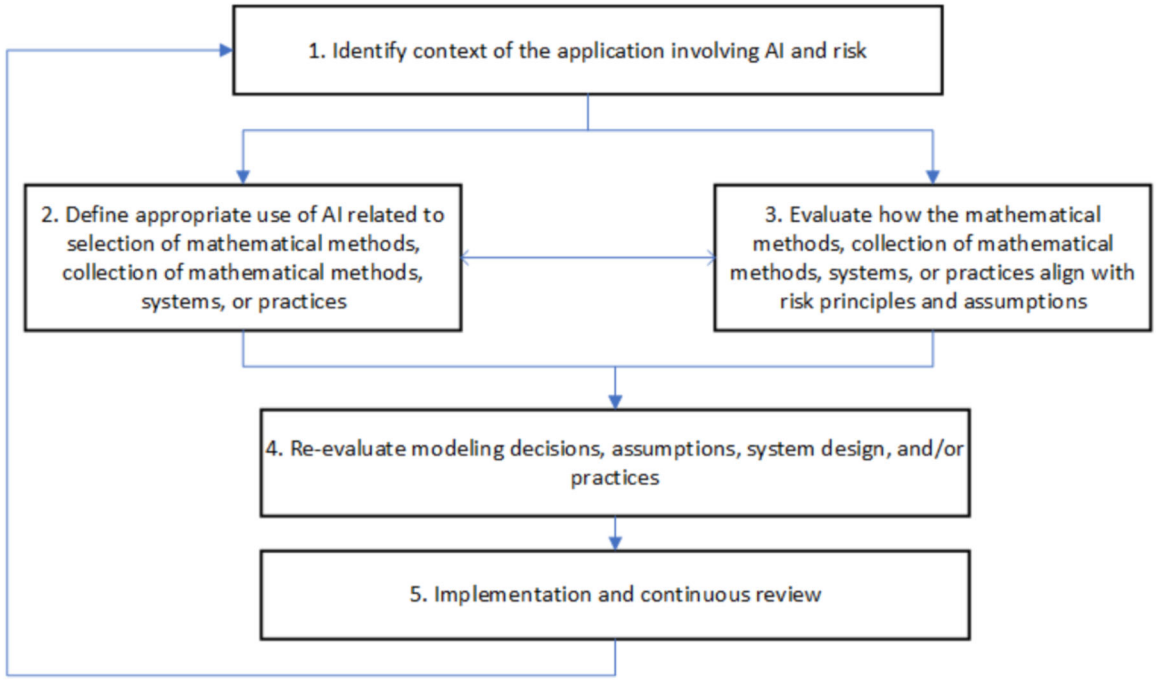

FIGURE 5 Basic framework presented to balance disaster risk and artificial intelligence (AI) applications

state of practice for AI and risk, based on recent research, applications, and recent events, and (2) the future state of practice for AI and risk, projecting over the next 10 years. To evaluate these practices, a questionnaire gauged the compatibility between risk and AI, using the characteristics presented in Section 3. Respondents were asked to rate the effectiveness of each studied characteristic for the two considered practices by considering (1) the current effectiveness for each characteristic and (2) the potential effectiveness for each characteristic. The questions asked were not application specific. They instead reflected the respondent's own opinions related to research and practice, thereby demonstrating how the framework of this study can be applied to various applications.

The current survey was conducted to elicit information from experts, and the results offer perspectives from N = 40 experts[1]. In this survey, the majority of the respondents (i.e., 80%) have a doctorate degree in a relevant field, while others have a master's degree, and the respondents are categorized as either "*Very knowledgeable*" or "*Extremely knowledgeable*" in at least one of the fields of disaster risk analysis and artificial intelligence. The respondents come from diverse sectors, including academia, critical infrastructure sectors, government, and national laboratories[2]. The qualifications and backgrounds of the respondents are sufficient to qualify this study as a suitable expert elicitation method to generate reasonable findings pertaining to the potential synergies between the fields of AI and disaster risk analysis.

Experts rated each characteristic (see Section 3) using an effectiveness scale of 1–5, with 5 denoting the highest level of effectiveness. These ratings are not stand-alone measurements, but rather a descriptive rating suggesting that higher numbers imply a stronger belief about effectiveness. The general results of the data collection are provided below. Additional discussion and analysis of this information will follow in Section 4.3.

Figure 6 shows the current (C) and potential (P) effectiveness ratings for each studied characteristic. The bottom of each bar shows the first quartile, the top of each bar shows the third quartile, the bottom line shows the minimum, and the top line shows the maximum. Figure 7 shows how the average current and potential effectiveness compare to one another. These results indicate a generally strong momentum toward leveraging evidence/data and automating processes using AI. However, there is also a substantial deficiency in enabling AI methods to address risk aspects related to uncertainty and the balance between domain knowledge and data.

Table 1 shows the summary statistics for the questionnaire responses. The column titled "Opp. Avg./Cur. Avg." shows the ratio between the opportunity (potential average – current average) and the current average. The results show relatively high optimism related to characteristics of *B.3: Conduct a knowledge-based assessment of whether relevant factors are reflected in the data; B.5: Manage trust and social impact for systems; and B.9: Manage legal and ethical implications of systems*.

[1] A sample size of N ≥ 30, which the current survey satisfies, is fairly common in statistical analysis. Notably, a sample size of 30 or more often increases the confidence associated with the survey findings.

[2] We adopted a similar approach that was utilized in a study published in the Risk Analysis journal with fewer respondents (Thekdi & Santos, 2019).

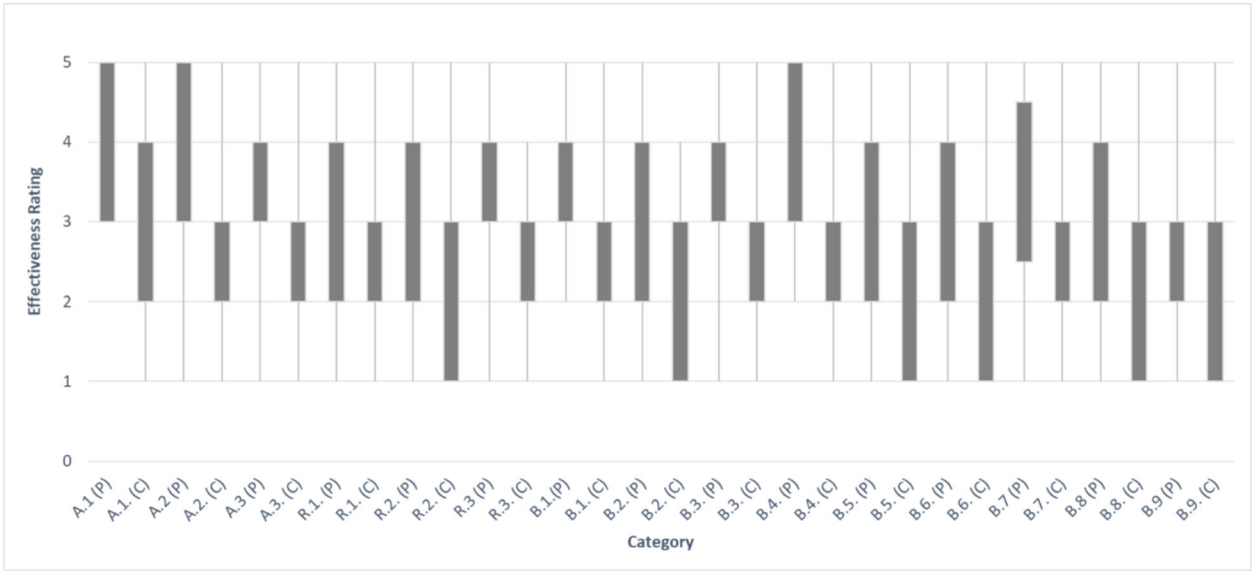

FIGURE 6 Effectiveness rating for each category, where "C" denotes "current" and "P" denotes "potential

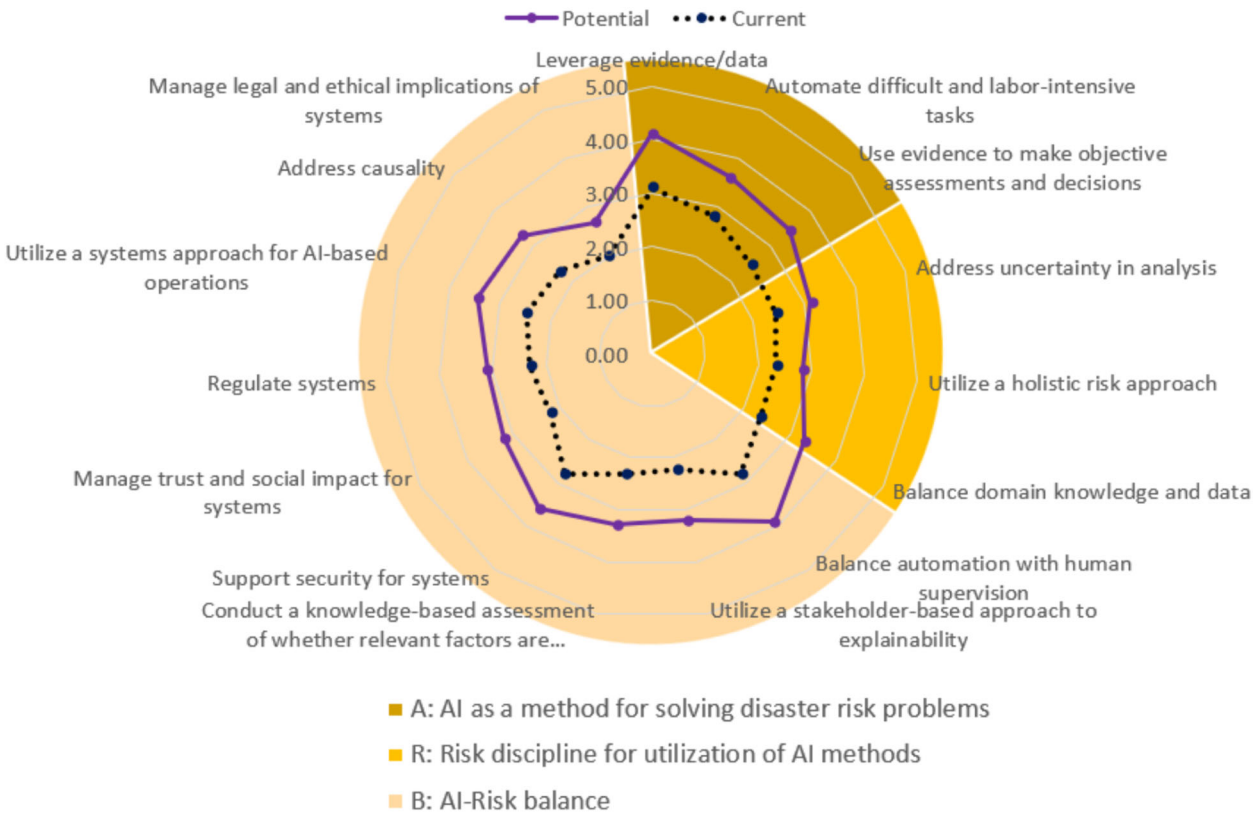

FIGURE 7 Current and potential effectiveness of each artificial intelligence (AI)–risk balance characteristic

Many of the Opportunity Avg. values are above a value of one. This suggests a firm belief in the potential for the use of AI in disaster risk management and vice versa. However, there is a relatively low Opportunity Avg. related to *R.1: Address uncertainty in analysis*, *R.2: Utilize a holistic risk approach, B.4: Support security for systems, and B.6: Regulate systems*. This is concerning, particularly because both the current and potential ratings are relatively low. Section 4.3 will further analyze the data collection results.

TABLE 1 Current and potential for compatibility between risk and artificial intelligence (AI), using a sample size of 40 (subject to rounding errors)

		Potential				Current				Opportunity	
	Category	**Avg.**	**Min**	**Max**	**St. Dev.**	**Avg.**	**Min**	**Max**	**St. Dev.**	**Avg.***	**Opp. Avg./Cur. Avg.***
A.1	Leverage evidence/data	4.14	3.00	5.00	0.82	3.07	1.00	5.00	1.00	1.07	0.35
A.2	Automate difficult and labor-intensive tasks	3.64	1.00	5.00	1.11	2.77	1.00	5.00	1.98	0.87	0.31
A.3	Use evidence to make objective assessments and decisions	3.49	1.00	5.00	1.10	2.49	1.00	5.00	1.09	1.00	0.40
R.1	Address uncertainty in analysis	3.10	1.00	5.00	0.92	2.35	1.00	5.00	1.03	0.75	0.32
R.2	Utilize a holistic risk approach	2.81	1.00	5.00	1.11	2.23	1.00	5.00	1.07	0.59	0.26
R.3	Balance domain knowledge and data	3.27	1.00	5.00	1.05	2.28	1.00	4.00	1.00	0.99	0.43
B.1	Balance automation with human supervision	3.81	2.00	5.00	0.88	2.72	1.00	5.00	1.01	1.09	0.40
B.2	Utilize a stakeholder-based approach to explainability	3.19	1.00	5.00	1.20	2.16	1.00	4.00	0.90	1.03	0.48
B.3	Conduct a knowledge-based assessment of whether relevant factors are reflected in the data	3.42	1.00	5.00	1.21	2.16	1.00	4.00	0.90	1.27	0.59
B.4	Support security for systems	3.63	2.00	5.00	1.13	2.74	1.00	5.00	1.09	0.89	0.33
B.5	Manage trust and social impact for systems	3.24	1.00	5.00	1.13	2.14	1.00	4.00	1.03	1.10	0.51
B.6	Regulate systems	3.05	1.00	5.00	1.15	2.26	1.00	5.00	1.14	0.79	0.35
B.7	Utilize a systems approach for AI-based operations	3.47	1.00	5.00	1.19	2.42	1.00	5.00	1.07	1.05	0.44
B.8	Address causality	3.29	1.00	5.00	1.16	2.28	1.00	5.00	1.11	1.01	0.44
B.9	Manage legal and ethical implications of systems	2.75	1.00	5.00	1.21	1.91	1.00	5.00	1.06	0.83	0.44

*Rounding is responsible for discrepancies in the Opportunity Avg. calculation.

4.3 | Analysis of results

The data collection and literature review results show that despite new and additional risks associated with the use of AI, there is overwhelming potential for AI as a method for disaster risk. AI has been and is expected to continue to effectively leverage large amounts of data and automate difficult, labor-intensive, and even dangerous tasks.

However, there are deficiencies in the ability of AI to address fundamental principles of risk. There is an opportunity to consider the following tasks to strengthen compatibility between risk and AI. There is an opportunity to increase the balance between human supervision and automation. For example, to address *B1: Balance automation with human supervision*, analysts can further investigate the most appropriate amount of human supervision and explore the use of pauses to enable human supervisors to oversee automated decisions and actions. This could also involve increased training for unplanned scenarios, ensuring humans are knowledgeable about how to override the system when problems occur. There is also an opportunity to address *R1: Address uncertainty in analysis*. For example, this could be addressed by further innovations in uncertainty algorithms in deep learning. From a broader perspective, there is a need for greater emphasis on understanding causality in modeling and how that causality integrates into the knowledge assessment and risk approaches. There is also potential for simulation-supported AI to study existing and what-if conditions (IBM, 2022c). When addressing *B2: Utilize a stakeholder-based approach to explainability*, it should be acknowledged that analysts may not fully understand the mechanisms of AI tools. Thus, there is a need to understand how to present the information to various stakeholders. There is also a need to address *B3: Conduct a knowledge-based assessment of whether relevant factors are reflected in the data*. System designers and managers recognize a tradeoff between benefits and efficiencies of automation versus simplifications made when modeling complex systems. The infamous quote indicates "All models are wrong, but some are useful" (Box, 1979). Finally, there is a need to address *B5: Manage trust and social impact for systems*. As with any new technology, there will be issues with mistrust. It is a healthy and useful task for developers of new technologies to prove trustworthiness over time. Risk analysts can benefit from regular training and certification related to new methods and technologies, which can lead to greater innovations while also addressing any emerging limitations to those new methods and technologies. This level of trust-building can permeate into broader adoption of new methods and technologies, which in turn, can increase the effectiveness of risk-based decision-making and communication activities.

Along with the needs described above, there is an opportunity for the risk field to further adapt to AI's increased use and potential. There are potential avenues to improve the ability of AI-based methods to appropriately consider risk. For example, consider using domain-aware-AI, allowing analysts to combine domain knowledge to inform model parameters, developing risk-based methods to learn governing equations from data (de Silva et al., 2020), and developing hybrid model-based or model-free approaches. With sparse data, the risk field can look to new techniques such as domain-shifting, allowing AI to learn from one domain area and use insights for another, or Bayesian deep learning.

At a foundational level, there is also a need for greater emphasis on understanding the integrity of data and evidence. A complex and tested model is only as reputable as the data and assumptions contained within the model. Understanding the relationship between poorly understood AI or ML models and the need for validation and verification is an essential next step in using these methods for risk applications (Lathrop & Ezell, 2017).

The deficiencies presented in these sections serve as a starting point to guide future risk theory and application.

5 | CONCLUSIONS AND OPPORTUNITIES FOR FURTHER RESEARCH

This study presented a framework to characterize synergies and incompatibilities between AI and disaster risk. This study is the first to develop a classification system for applying risk principles for AI-based applications, allowing for use across disciplines, such as business, infrastructure, engineering, and other analytical disciplines. When this framework is included in broader Enterprise Risk Management and process improvement activities, there is potential to make large strides in promoting the use of fundamental risk concepts within the use of AI.

The analysis of current literature showed that there is emerging published research that addresses the use of AI in broader disaster risk applications, including natural and human-induced events. For the research that currently exists, the application areas are primarily in the domains of floods and other natural hazard applications. There is a large opportunity to increase attention toward application areas related to infrastructure and business, as these areas have shown in recent history to have potentially disastrous consequences.

The analysis also suggests that while AI is being used to conduct risk tasks, these studies are not primarily being published in risk journals or are being conducted in private industries, independent of peer review in the academic discipline. One can ask whether risk principles are being sufficiently addressed. There is also a need to recognize that AI models are often used for the purposes of automating tasks in as-planned operating scenarios. The study of risk involves unplanned operating scenarios, suggesting that AI precludes risk. As discussed in this study, while AI methods can be very accurate for analytic purposes, they lose accuracy in new or surprise situations, times at which effective risk management is most critical. Finally, it is essential to note that aspects such as the value of lives, health and safety, fairness cannot be

easily quantified or modeled using AI methods. Therefore, a risk perspective should be used to evaluate these aspects.

The results also showed that there are serious concerns over the current status and potential for using the as-is fundamental principles for risk for AI-based disaster risk applications. There is an opportunity to innovate the risk and AI fields by promoting a balance between risk and AI concepts. There are many unanswered questions and avenues for further studies. For example, there is an opportunity to develop guidance for determining the compatibility between specific AI-based mathematical models and risk concepts. There is an opportunity to investigate the effective properties related to human supervision for risk studies, such as evidenced by ongoing studies on human supervision in autonomous vehicles (IEEE, 2021). Similarly, there is a need to understand effective practices and methods to address explainability for AI-based risk applications (IBM, 2021). There is also an opportunity to further explore the role of knowledge, or lack of knowledge, in assessments for data-driven applications related to AI.

Finally, the results of the framework demonstration showed the value of collecting data on the characteristics developed in Section 3. This information enables system managers to be guided through immediate needs for system interventions. The demonstration also showed the importance of including this framework as part of a more extensive dialogue among stakeholders. This dialogue can encourage stakeholders to recognize and address the risk associated with increased automation, changing technologies, and other uncertainties. While the demonstration showed that the studied characteristics might have varying importance for different systems, the system managers and stakeholders can be included in those discussions to understand which characteristics matter the most in particular situations.

The framework of this study, including the characteristics presented in Section 3, does not necessarily relate *only* to AI. They could apply to the use of any new method, such as related to the QRA adaption described by Apostolakis (2004). This framework can also be adapted to evaluate specific AI models as a supplement to the broad characterization presented in this study.

The framework presented in this study serves as a stimulus for additional research on the balance between risk and AI. Future work is needed to evaluate this balance using a more granular perspective. For example, there is a need to address the heterogeneity in the various AI models and understand how those models compare with one another from a risk-based perspective. In addition, those AI models can be interpreted differently based on the more specific risk characterization tasks, such as in simulating risk events, estimating probabilities, and estimating consequences (Kaplan & Garrick, 1981).

The last few years have been tumultuous, with many high-profile risk incidents, including war, infrastructure failures, cyberattacks, and riots. These incidents all involve a data trail that could have been studied using AI models. One can ask if those AI models could have prevented those incidents, reduced the impact, or improved resilience for the involved systems. Likely, the answer is yes. However, because these events were unprecedented, one can ask if risk principles would have allowed decision-makers and stakeholders to more effectively prepare and act.

If AI is trusted to conduct risk activities, there is concern that these methods are not following risk principles. This poses one of the largest challenges for the risk field. The risk field must recognize the use of AI models and adapt the principles to embrace the characteristics of AI approaches. There is concern that the risk and AI fields will evolve separately, decreasing the visibility of the risk science discipline. There is an urgent need for the risk field to foster research on AI topics and encourage greater accessibility of the risk field to AI applications. Priorities include conducting more workshops on AI methods and coordinating joint events with the risk and AI communities. Consequently, the risk field can modernize and form an "AI-infused risk thinking," or hybrid risk and AI process that allows the risk and AI domains to coordinate efforts, creating a stronger risk-focused society.

The framework presented in this study offers a starting point for this AI-infused risk thinking. Now is also the time for the risk field to act and adapt to this new regime.

ACKNOWLEDGMENT

The authors thank the anonymous reviewers who provided very helpful feedback that strengthened this study.

ORCID

Shital Thekdi https://orcid.org/0000-0003-4145-508X
Unal Tatar https://orcid.org/0000-0001-8233-9566
Joost Santos https://orcid.org/0000-0001-5092-543X

REFERENCES

Anderson, J., Rainie, L., & Luchsinger, A. (2018). Artificial intelligence and the future of humans. Pew Research Center. https://www.pewresearch.org/internet/2018/12/10/artificial-intelligence-and-the-future-of-humans/

Appenzeller, T. (2017). The AI revolution in science. *Science*, *357*, 16–17.

Apostolakis, G. E. (2004). How useful is quantitative risk assessment? *Risk Analysis: An International Journal*, *24*(3), 515–520.

Asher, J., & Arthur, R. (2017, June 13). *Inside the algorithm that tries to predict gun violence in Chicago*. The New York Times.

Batista, F., Hirtzer, M., & Dorning, M. (2021, May 31). All of JBS's U.S. beef plants were forced shut by cyberattack. *Bloomberg*. https://www.bloomberg.com/news/articles/2021-05-31/meat-is-latest-cyber-victim-as-hackers-hit-top-supplier-jbs

Bhattacharya, S., & Singh, A. (2020). Why the tremendous potential of uploading health educational material on medical institutions' website remains grossly underutilized in the era of the Fourth Industrial Revolution? *Journal of Education and Health Promotion*, *9*, 248.

Bilen, A., & Özer, A. B. (2021). Cyber-attack method and perpetrator prediction using machine learning algorithms. *PeerJ Computer Science*, *7*, e475.

Bini, S. A. (2018). Artificial intelligence, machine learning, deep learning, and cognitive computing: What do these terms mean and how will they impact health care? *The Journal of Arthroplasty*, *33*(8), 2358–2361.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993–1022.

Bobrowsky, M. (2021). Kaseya ransomware attack: What we know as REvil hackers demand $70 million. Wall Street Journal. https://www.wsj.com/articles/kaseya-ransomware-attack-11625593654

Box, G. E. (1979). *Robustness in the strategy of scientific model building. Robustness in statistics*. Academic Press.

Campbell, J. L., Brown, J. L., Graving, J. S., Richard, C. M., Lichty, M. G., Bacon, L. P., Justin. M. F., Hong, L., Diane, W. N., & Sanquist, T. (2018). Human factors design guidance for level 2 and level 3 automated driving concepts (No. DOT HS 812 555). https://trid.trb.org/view/1574671

Cheatham, B., Javanmardian, K., & Samandari, H. (2019). Confronting the risks of artificial intelligence. https://www.mckinsey.com/capabilities/quantumblack/our-insights/confronting-the-risks-of-artificial-intelligence

Choi, T. M., & Lambert, J. H. (2017). Advances in risk analysis with big data. *Risk Analysis*, *37*(8), 1435–1442.

Chu, L. C., Anandkumar, A., Shin, H. C., & Fishman, E. K. (2020). The potential dangers of artificial intelligence for radiology and radiologists. *Journal of the American College of Radiology*, *17*(10), 1309–1311.

Chui, M., Manyika, J., & Miremadi, M. (2015). Four fundamentals of workplace automation. *McKinsey Quarterly*, *29*(3), 1–9.

Cox, L. A., Jr (2020). Answerable and unanswerable questions in risk analysis with open-world novelty. *Risk Analysis*, *40*(S1), 2144–2177.

De Haan, L., & Ferreira, A. (2007). *Extreme value theory: An introduction*. Springer Science & Business Media.

de Silva, B. M., Higdon, D. M., Brunton, S. L., & Kutz, J. N. (2020). Discovery of physics from data: Universal laws and discrepancies. *Frontiers in Artificial Intelligence*, *3*, 25.

Eisenberg, D., Seager, T., & Alderson, D. L. (2019). Rethinking resilience analytics. *Risk Analysis*, *39*(9), 1870–1884.

Foy, K. (2020, April 21). With lidar and artificial intelligence, road status clears up after a disaster. MIT News. https://news.mit.edu/2020/lidar-and-ai-road-status-clears-after-disaster-0415

Gil, Y., & Selman, B. (2019). A 20-year community roadmap for artificial intelligence research in the US. arXiv preprint arXiv:1908.02624.

Greenblatt, N. A. (2016). Self-driving cars and the law. *IEEE Spectrum*, *53*(2), 46–51.

Grün, B., Hornik, K., Blei, D. M., Lafferty, J. D., Phan, X. - H., Matsumoto, M., Nishimura, T., & Cokus, S. (2021). Topicmodels: An R package for fitting topic models. R package version 0.2–12. https://cran.r-project.org/web/packages/topicmodels/topicmodels.pdf

Guikema, S. (2020). Artificial intelligence for natural hazards risk analysis: Potential, challenges, and research needs. *Risk Analysis*, *40*(6), 1117–1123.

Hampson, M. (2021). Drones and sensors could spot fires before they go wild. IEEE Spectrum. https://spectrum.ieee.org/drones-sensors-wildfire-detection

Hunt, K., Agarwal, P., & Zhuang, J. (2020). Monitoring misinformation on Twitter during crisis events: A machine learning approach. *Risk Analysis*, *42*, 1728–1748. https://doi.org/10.1111/risa.13634

IBM. (2021a). AI explainability 360. https://aix360.mybluemix.net/

IBM. (2021b). Tackling bias in AI. https://www.ibm.com/blogs/systems/tackling-bias-in-ai/

IBM. (2022a). Artificial intelligence (AI). https://www.ibm.com/cloud/learn/what-is-artificial-intelligence

IBM. (2022b). Machine learning. https://www.ibm.com/cloud/learn/machine-learning

IBM. (2022c). AI-enriched simulation. https://research.ibm.com/science/ai-enriched-simulation/

IEEE. (2021). Accelerating autonomous vehicle technology. https://spectrum.ieee.org/transportation/self-driving/accelerating-autonomous-vehicle-technology

Information Society Project. (2017). Governing machine learning. https://law.yale.edu/sites/default/files/area/center/isp/documents/governing_machine_learning_-_final.pdf

ISO. (2018). ISO 31000:2018(en) risk management—Guidelines. https://www.iso.org/obp/ui/#iso:std:iso:31000:ed-2:v1:en

Jeong, S. (2019, April 10). Insurers want to know how many steps you took today. The New York Times.

Kaplan, S., & Garrick, B. J. (1981). On the quantitative definition of risk. *Risk Analysis*, *1*(1), 11–27.

Kenyon, T. (2021). How are social media platforms using AI? https://aimagazine.com/ai-strategy/how-are-social-media-platforms-using-ai

Kingston, J. (2018). Artificial intelligence and legal liability. arXiv preprint arXiv:1802.07782.

Kuglitsch, M. M., Pelivan, I., Ceola, S., Menon, M., & Xoplaki, E. (2022). Facilitating adoption of AI in natural disaster management through collaboration. *Nature Communications*, *13*(1), 1–3.

Kuglitsch, M., Albayrak, A., Aquino, R., Craddock, A., Edward-Gill, J., Kanwar, R., Koul, A., Ma, J., Marti, A., Menon, M., Pelivan, I., Toreti, A., Venguswamy, R., Ward, T., Xoplaki, E., Rea, A., & Luterbacher, J. (2022). Artificial intelligence for disaster risk reduction: Opportunities, challenges, and prospects. *World Meteorological Organization*. https://public.wmo.int/en/resources/bulletin/artificial-intelligence-disaster-risk-reduction-opportunities-challenges-and

Lathrop, J., & Ezell, B. (2017). A systems approach to risk analysis validation for risk management. *Safety Science*, *99*, 187–195.

McKenzie, C., Barker, K., & Santos, J. (2014). Modeling a severe supply chain disruption and post-disaster decision making with application to the Japanese earthquake and tsunami. *IIE Transactions*, *46*(12), 1243–1260.

Miller, H., & Matussek, K. (2018, June 18). VW executive tally in diesel scandal grows with Stadler arrest. *Automotive News*. https://www.autonews.com/article/20180618/OEM02/180619742/vw-executive-tally-in-diesel-scandal-grows-with-stadler-arrest

Nateghi, R., & Aven, T. (2021). Risk analysis in the age of big data: The promises and pitfalls. *Risk Analysis*, *41*(10), 1751–1758.

National Security Commission on Artificial Intelligence. (2021). The Final Report. https://reports.nscai.gov/final-report/table-of-contents/

Nicas, J., Kitroeff, N., Gelles, D., & Glanz, J. (2019, June 1). Boeing built deadly assumptions into 737 Max, blind to a late design change. The New York Times.

Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M. E., Ruggieri, S., Turini, F., Papadopoulos, S., Krasanakis, E., Kompatsiaris, I., Kinder-Kurlanda, K., Wagner, C., Karimi, F., Fernandez, M., Alani, H., Berendt, B., Kruegel, T., Heinze, C., … Staab, S. (2019). Bias in data-driven artificial intelligence systems—An introductory survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *10*(3), e1356.

Osoba, O. A., & Welser, W. (2017a). *The risks of artificial intelligence to security and the future of work*. RAND.

Osoba, O. A., & Welser, W., IV (2017b). *An intelligence in our image*. RAND corporation.

Parikh, R. B., Teeple, S., & Navathe, A. S. (2019). Addressing bias in artificial intelligence in health care. *JAMA*, *322*(24), 2377–2378.

R Core Team. (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Ring, T. (2015). Connected cars—The next target for hackers. *Network Security*, *2015*(11), 11–16.

Sanger, & Perlroth, (2021, May 14). Pipeline attack yields urgent lessons about U.S. cybersecurity. New York Times. https://www.nytimes.com/2021/05/14/us/politics/pipeline-hack.html

Scherer, M. U. (2015). Regulating artificial intelligence systems: Risks, challenges, competencies, and strategies. *Harvard Journal of Law & Technology*, *29*, 353.

Shneiderman, B. (2016). Why algorithms urgently need independent oversight. *Proceedings of the National Academy of Sciences*, *113*(48), 13538–13540.

Society for Risk Analysis. (2018). Society for risk analysis glossary. https://www.sra.org/wp-content/uploads/2020/04/SRA-Glossary-FINAL.pdf

Sun, W., Bocchini, P., & Davison, B. D. (2020). Applications of artificial intelligence for disaster management. *Natural Hazards*, *103*, 2631–2689.

Terti, G., Ruin, I., Gourley, J. J., Kirstetter, P., Flamig, Z., Blanchet, J., Arthur, A., & Anquetin, S. (2019). Toward probabilistic prediction of flash flood human impacts. *Risk Analysis*, *39*(1), 140–161.

Thekdi, S. A., & Aven, T. (2016). An enhanced data-analytic framework for integrating risk management and performance management. *Reliability Engineering & System Safety*, *156*, 277–287.

Thekdi, S. A., & Santos, J. (2019). Decision-Making analytics using plural resilience parameters for adaptive management of complex systems. *Risk Analysis*, *39*(4), 871–889.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *49*, 433–460.

UNDRR. (2022a). Disaster. https://www.undrr.org/terminology/disaster

UNDRR. (2022b). Disaster risk management. https://www.undrr.org/terminology/disaster-risk-management

Van Heteren, A., Hirt, M., & Van der Veken, L. (2020, January 14). Natural disasters are increasing in frequency and ferocity. Here's how AI can come to the rescue. World Economic Forum. https://www.weforum.org/agenda/2020/01/natural-disasters-resilience-relief-artificial-intelligence-ai-mckinsey/

Viechnicki, P., & Eggers, W. D. (2017). How much time and money can AI save government? Cognitive technologies could free up hundreds of millions of public sector worker hours. https://www2.deloitte.com/us/en/insights/focus/cognitive-technologies/artificial-intelligence-government-analysis.html

Wagenaar, D., Hermawan, T., van den Homberg, M. J., Aerts, J. C., Kreibich, H., de Moel, H., & Bouwer, L. M. (2021). Improved transferability of data-driven damage models through sample selection bias correction. *Risk Analysis*, *41*(1), 37–55.

Wang, K., Gou, C., Duan, Y., Lin, Y., Zheng, X., & Wang, F. (2017). Generative adversarial networks: Introduction and outlook. *EEE/CAA Journal of Automatica Sinica*, *4*(4), 588–598.

Web of Science. (2021). https://webofknowledge.com

Wolfe, F. (2017). How artificial intelligence will revolutionize the energy industry. Harvard University Blog, Special Edition on Artificial Intelligence. https://sitn.hms.harvard.edu/flash/2017/artificial-intelligence-will-revolutionize-energy-industry/

How to cite this article: Thekdi, S., Tatar, U., Santos, J., & Chatterjee, S. (2023). Disaster risk and artificial intelligence: a framework to characterize conceptual synergies and future opportunities. *Risk Analysis*, *43,* 1641–1656. https://doi.org/10.1111/risa.14038