NR&E Winter 2024: Environmental Health & Safety

Artificial Intelligence and the Future of Environmental Health and Safety Management

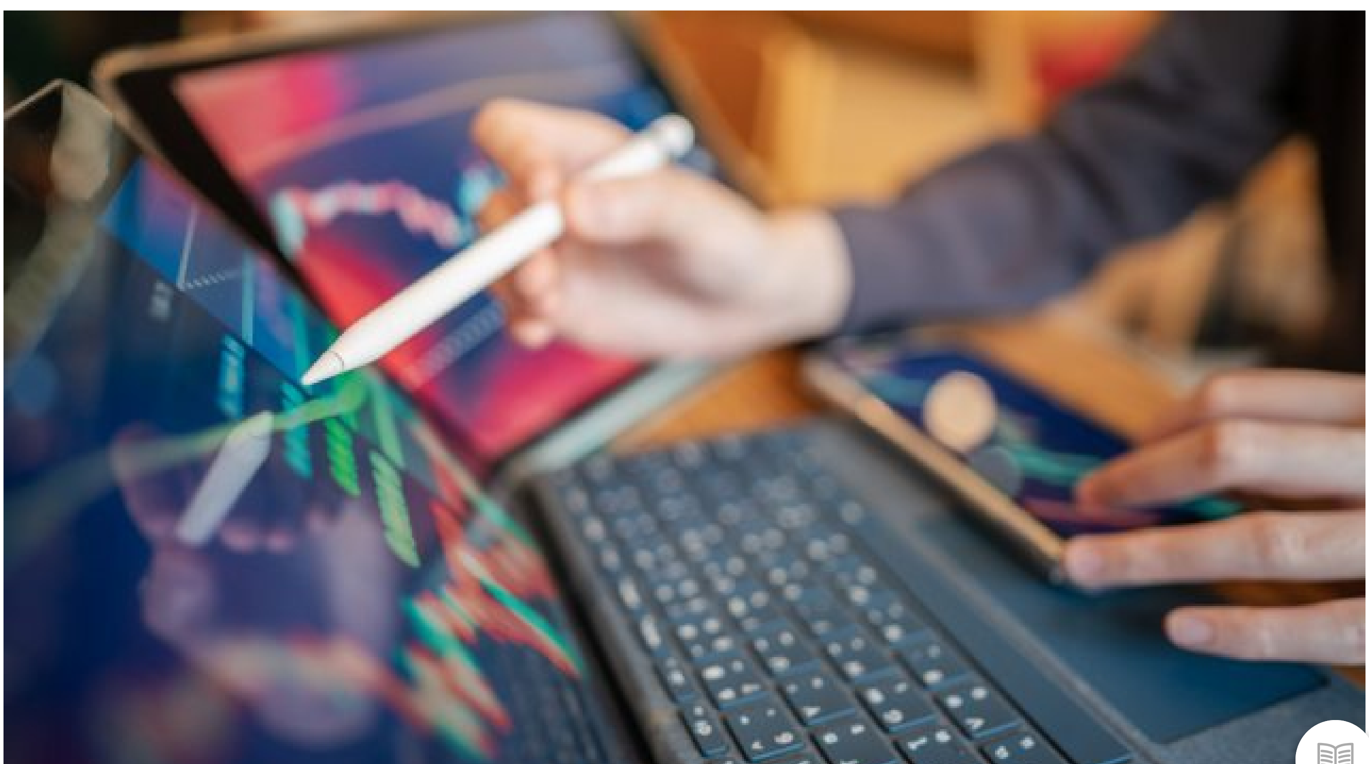
Heather Demirjian and Inga Caldwell

Feb 22, 2024 14 min read

Courts & Practice of Law | Health & Safety | Federal Agencies | Enforcement & Oversight

Summary

- Big data and artificial intelligence (AI) will drive the future of environmental health and safety.
- There are unique legal and governance risks associated with the use of AI systems.
- Environmental health and safety managers need to manage these risks at each step of the AI lifecycle.



primeimages via Getty Images

The phrase "artificial intelligence" (AI) seems to have suddenly surrounded us in the news and on the internet. And yet, the development and use of AI have been ongoing for decades. The same recent advances in the sophistication of AI technologies, such as ChatGPT, that have brought AI into mainstream awareness are also responsible for the growing interest in the benefits of AI technologies to worldwide business and industrial operations. According to the Global Artificial Intelligence Market Report 2023, the global AI market is rapidly expanding and is estimated to increase from $150.2 billion in 2023 to $1.34 trillion in 2030. The largest market share is in North America.

AI will revolutionize how industries around the world operate, including within the realm of environmental health and safety (EH&S). The use of AI in EH&S management is in its infancy but will quickly evolve due to the data-heavy nature of EH&S management. AI will likely impact EH&S management in the broadest sense regarding the protection of the health and safety of workers, the public, and the environment in industrial operations, and in achieving compliance with government regulations and requirements.

This article provides an overview of the rise of AI systems in EH&S and how they may evolve. We attempt to demystify the basics of AI technology and its uses. We also explore legal and governance challenges and risk management considerations associated with the incorporation of AI into EH&S management systems.

Types of Artificial Intelligence

AI is not a new technology. The term AI was first used in 1955 by Stanford University Professor John McCarthy. Professor McCarthy defined AI as "the science and engineering of making intelligent machines." The *Oxford Dictionary* defines AI more expansively as:

> [t]he theory and development of computer systems able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making and translation between languages.

Traditional computer programming involves step-by-step instructions that a computer follows to perform a defined task. In contrast, AI technologies use algorithms trained on datasets to make predictions and problem solve. Data are king in the world of AI.

AI systems differ in complexity and require different levels of human input. The simplest AI systems, referred to as "Weak AI" or "Narrow AI," are designed to perform only one task. An example of Narrow AI is an AI algorithm created to label images as those of cats versus dogs. These Weak AI system algorithms may be trained on a dataset of images that a human has pre-labeled as either "cat" or "dog." Once trained using the training dataset, the AI system will predict whether a new image it analyzes is that of a cat or dog. AI system predictions have varying degrees of accuracy depending on the characteristics of the dataset used to train the AI algorithm.

The majority of what has very recently been in the news about AI concerns complex AI systems such as ChatGPT. Some of these complex AI systems, referred to as Artificial General Intelligence or Strong AI, are designed to perform sophisticated tasks as quickly and easily as would occur with human intelligence. Even more complex systems, referred to as Artificial Super Intelligence, are systems that can improve themselves and perform much faster and more proficiently than would occur with human intelligence.

Both Artificial General Intelligence and Artificial Super Intelligence are forms of "machine learning" (ML). ML is a subfield of AI characterized by algorithms that allow the system to learn from the datasets it is trained on in order to perform predictive tasks without explicit programming. ML falls into three categories: supervised learning, unsupervised learning, and reinforcement learning. These three categories of ML incorporate various statistical methodologies, including logistic regression, weighted regression, kernel regression, decision trees, and random forests.

Supervised learning algorithms use statistical modeling and training datasets containing labeled inputs and outputs to predict an outcome. This is the type of algorithm described above that is trained on labeled images to distinguish an image of a cat from an image of a dog. In a business setting, a supervised learning algorithm can, for example, use an input dataset consisting of past emails with specifically labeled elements to map out an outcome label for the email as either "spam" or "not spam." Once the supervised learning algorithm is trained, the AI system can then predict whether a future email sent to an employee is or is not spam and appropriately classify that email in the employee's inbox.

Unsupervised learning algorithms can interpret and group unlabeled training datasets. These algorithms are given input variables but have no fixed set of outputs. Unsupervised learning algorithms cluster data into groups based on similar features. For example, if provided with a set of unlabeled emails, an unsupervised learning algorithm can learn identify common topics such as politics, news, and sports based on its analysis of words or phrases contained across those emails. The unsupervised learning algorithm can then label the emails based on topics, which may allow for more efficient email organization and management.

Reinforcement learning algorithms test actions to decide which actions work best. These algorithms are used for simulation and robotics-based applications. Another application of reinforcement learning algorithms is in advertising. For example, reinforcement learning algorithms can be used to test which email ad copy will attract more clicks and, therefore, more website traffic and sales. While traditional A/B testing involves sending “Email Ad A” 50% of the time and “Email Ad B” 50% of the time, a reinforcement learning algorithm will allocate increasingly more email traffic to the higher-performing email ad as it learns and confirms which ad is performing better.

What is known as “Deep Learning” is an even more complex AI technology. Deep Learning is essentially a very powerful type of unsupervised learning that relies on multiple layers of algorithms, called artificial neural networks, which are modeled after neurons in the human brain. Neurons in the human brain communicate with one another as they receive, process, and transmit signals to other neurons. Some Deep Learning systems are so sophisticated that computer scientists do not precisely understand how the system’s artificial neural network processes, analyzes, and transmits data throughout the network to reach predictions. Deep Learning systems learn and evolve over time based on the data they receive. These systems require a significant amount of training data, variation in the data, data storage capacity, and computational power. Applications of Deep Learning include self-driving cars and automatic facial recognition.

The more complex the AI system, the more likely the system will require the use of what is known as “Big Data” to adequately train the AI algorithm. Big Data includes datasets that are so large that traditional data management tools are insufficient. Datasets fall within the universe of Big Data not only based on scale or volume, but also when a dataset includes information of different data types and file formats (e.g., text, multimedia, audio, PDF, numeric), when the data are streaming (e.g., continuously flowing from various sources), and when the veracity or reliability of the data is unknown.

The use of Big Data comes with risks. Big Data is much more difficult to process and analyze. Given the nature of Big Data, its use can introduce bias and inaccuracies into an AI system’s predictions and outputs. But Big Data also comes with rewards. The use of Big Data with sophisticated AI algorithms can allow a business to answer questions more precisely and to answer questions related to the business that it did not have the capacity to answer before.

The Past, Present, and Future of AI in Environmental Health and Safety

EH&S management is inherently data and analysis driven, making it an ideal field to successfully incorporate AI systems as those systems evolve in sophistication. EH&S-related costs and manpower resources can be difficult for company executives to justify because EH&S activities do not directly translate into revenue. As a result, traditional EH&S management practices and operations were historically ripe for digitization and are now ripe for the incorporation of AI.

Traditional EH&S management practices were human resources and human intervention heavy. A company's EH&S knowledge and history were at risk of becoming siloed in situations where EH&S records were handwritten by one or two employees who may or may not have accurately recorded EH&S incidents or compliance-related data. The time to prepare EH&S records was considerable. If an EH&S employee left a company, then EH&S knowledge could be lost. The potential for errors and omissions in traditional EH&S management and compliance processes were sources of financial and legal risks to a company.

As computer technology advanced, EH&S management and compliance data became digitized. Many of these technologies were database driven and relied on inputs to spreadsheets created in software such as Excel. Digitization of EH&S management and compliance data improved a company's ability to identify and reduce EH&S risks and improve EH&S compliance rates, but these systems were still susceptible to knowledge silos and difficulty in implementing consistent EH&S practices and policies across company facilities. Digitization processes may have improved the organization and compilation of EH&S data, but they still required significant human resources, time, and intervention, which made them subject to human error.

As computer technology evolved, EH&S management moved beyond basic database- and spreadsheet-type systems into platform-based computer software that required less human input and improved EH&S practices. As AI systems are incorporated more and more into EH&S management, we will see even more powerful health, safety, and compliance tools. For example, EH&S computer systems have been developed using aspects of AI that can automatically collect data from multiple sources and provide real-time insights. These systems can already provide real-time safety guidance, answer an employee's EH&S questions, and identify patterns to predict future EH&S risks.

Legal and Governance Challenges in AI System Development

There are unique legal and governance risks associated with the development and introduction of an AI system into existing business operations. Making the use of AI systems even more challenging is the fact that there is no standard or reliable metric to assess and measure AI risks. The use of AI is beginning to be regulated, but these regulations are currently being introduced patchwork style at the state level in areas such as consumer data privacy, consumer rights, and employment practices. States are even using different and sometimes inconsistent definitions of AI in their regulations.

Legal and governance risks and challenges may be amplified even more when an AI system is developed for an EH&S application. These AI systems are often sophisticated and require enormous amounts of data to make accurate predictions. The consequences of errors and inaccuracies in the EH&S context can be legally and financially significant for a business. If implemented successfully, however, these AI systems can be used to predict human health and safety incidents, harm to the environment, and future noncompliance with government regulations and requirements.

To assist organizations with the responsible implementation of AI, in January 2023 the U.S. National Institute of Standards and Technology (NIST) issued an *Artificial Intelligence Risk Management Framework*. NIST recommends the use of "trustworthy AI" at each step of the AI life cycle. "Trustworthy AI" is AI that is valid and reliable, safe, secure and resilient, accountable and transparent, explainable and interpretable, privacy enhanced, and fair. Below, we consider the application of the elements of trustworthy AI to the different stages of AI system development for EH&S management.

Stages of AI System Development for Environmental Health and Safety

The first stage in AI system development is planning and design. This stage is critical for the successful deployment of an AI system for any business. For example, if a multifacility U.S. manufacturing company is interested in upgrading outdated traditional computing-based EH&S management software, key decision-makers at that company and within the EH&S department must work collaboratively to identify the specific company goals in order to develop an AI system that matches those goals. This may include designing an AI system that will minimize financial impacts resulting from the past use of EH&S management practices and systems that are more susceptible to human error, managing and mitigating key EH&S risk factors that could lead to liability exposure, ensuring consistency in EH&S policies and practices across all facilities, or eliminating avoidable violations of EH&S regulatory obligations that may result from missed deadlines or errors in accurately completing complex agency reporting submissions. This may also include whether the company's financial resources and technology expertise allow the company to build an AI model in-house or require the company to outsource the AI system to a third-party vendor.

In addition to articulating goals for an AI system, the company will also need to identify all relevant sources of data across each of the company's facilities, as well as public sources of relevant data. This may include historical handwritten data and information that predates a software-based recordkeeping system.

The ability to identify and collect adequate and reliable data is a key factor in selecting the type of AI system that will best fit the EH&S management needs and resources of a company. Identifying and collecting data are often time-consuming and can be inordinately expensive.

The more sophisticated and accurate an AI system is, the more it requires reliable and high-quality training data, resulting in higher data-related costs. However, AI system accuracy not only comes with a financial trade-off, but in general the more accurate the AI system is, the less explainable the AI system's decision-making process will be. This is called the "black box problem" and is particularly true with Deep Learning neural networks. In the EH&S context, balancing the need for accuracy with the need to have an explainable AI system may not be straightforward if the costs resulting from noncompliance due to inaccuracies are measured against the benefits in a litigation setting of being able to explain how an AI system made its decisions.

Finally, safety should be considered early in the planning and design stage. AI systems should not endanger human health, property, or the environment. Both internal and external safety concerns associated with EH&S need to be prioritized and made subject to the most stringent risk management processes. Safe operation of AI systems involves responsible decision-making from both the design team and the end users. AI systems can be designed to include metrics that sound alarms and solicit human intervention if the AI system is not operating as intended.

Collecting and Processing Data

After planning and design is complete, the next stage is to collect and process data. Going back to the example of a multifacility U.S. manufacturing company, that company may need to implement new processes to make EH&S data and information from old software programs across all facilities available to the new AI system. Other internal electronic sources of relevant EH&S data could include document repositories, email archives, and other software products, such as human resources software, which may not be centrally located. If the company also possesses relevant hardcopy records, those records will r to be collected from each facility and digitized consistently using standardized metadata fields. Procedures should also be developed for the ongoing collection of new EH&S data.

The large volume of data required to train an AI system may mean that a company does not possess sufficient internal data sources and must look to external publicly available EH&S data to fill the gap. Obtaining publicly available information is often expensive, and there may be no way to adequately evaluate the reliability of the data. AI systems are only as good as the quality of data the system is fed. A dataset containing biases (e.g., incomplete data, data containing errors or that are skewed) will likely introduce those same biases into the AI system's predictions and outputs. When utilizing publicly available EH&S data, it is important to obtain those data from a wide range of sources (e.g., Occupational Safety and Health Administration (OSHA), Environmental Protection Agency, state agencies, industry organizations) in a wide range of data types. If data are less reliable, using multiple layers of data may improve the accuracy of an AI system.

Transparency is important at all steps of the AI system life cycle, including data collection and processing. Data collection should be transparent, with all relevant data fed into the AI system. If data are censored during this stage, the AI system's decisions will be less accurate and potentially biased.

Building, Deploying, and Operating the Model

Once the relevant data are collected and are in a usable format, the building of the AI model can begin. During this stage, model engineers, data scientists, and developers create or select algorithms with a company's AI system goals in mind. The modeling team then uses a subset of the data to train the AI system. For example, if one of the goals of an EH&S AI system is to identify reportable incidents, then it would be important to include data relating to all past reportable incidents in the training dataset.

Privacy is an essential legal and governance consideration during model building. Key privacy considerations include who will own the data, where the data will be stored, and who will have access to the data. In the EH&S context, reports to the OSHA may include personally identifiable information that must be accessible for the AI system's analysis but not available to end users.

There are trade-offs between privacy-enhanced AI, which values anonymity and confidentiality, and transparent AI. Privacy-enhanced AI includes the use of privacy-enhancing technologies and de-identification and aggregation of model outputs. These privacy-enhancing techniques can sanitize the data to the point where they hinder the AI system from seeing relevant connections within the datasets and, as a result, are less accurate.

After building the AI model, the modeling team must then validate the model. Valid and reliable AI is AI that objectively meets the requirements for a specific intended use and performs as expected over the lifetime of the system. Validity and reliability are assessed by ongoing testing and monitoring.

After the AI model is trained and validated, the AI system is deployed. This is the stage where the challenges of explainability and interpretability of an AI system can have a crucial impact on its successful use. In our example of the multifacility U.S. manufacturing company, suppose that two end users separately query the EH&S AI system for the top five environmental risk factors at the company's Houston, Texas, facility. Both end users receive outputs with the same five risk factors; however, the factors are ranked differently. The end users would want to understand how the AI system made its ranking decision and why there was a difference in the outputs.

Explainable and interpretable AI helps end users understand both the purpose and potential impacts of the AI system. Explainability helps an end user understand how a decision was made by the AI system, while interpretability helps an end user understand why a decision was made. As mentioned above, there are trade-offs between transparency, explainability, interpretability, and accuracy, which a company's decision-makers must consider throughout AI system development.

The final stage in the AI system life cycle is operating and monitoring the model. Security is a critical risk during the long-term operation of an AI system. Security should be considered for both the integrity of the AI system and the underlying software and hardware. Secure AI systems are able to maintain confidentiality and prevent unauthorized access, which are often a required component of EH&S management.

Ongoing maintenance of the AI system is essential and may require several resources. Our hypothetical multifacility U.S. manufacturing company mentioned above should proactively schedule updates to its AI system to capture changes in EH&S regulations and in the internal company structure, such as when the company adds a new facility. If software products that interact with the AI system are updated, the modeling team needs to consider the impact of those updates. The company should also consider putting in place controls limiting the employees who are authorized to access the AI system to update its datasets. These are just a few examples of considerations to ensure that an AI system remains secure, and the outputs remain accurate.

The Future of AI and EH&S Is Now

This article gives only a high-level overview of how the use of AI systems in EH&S management may develop in the future. AI is predicted to impact almost every industry and business worldwide. While it is unlikely that AI systems will replace humans in EH&S or any other field, people who use AI are likely to replace those who do not. Now is the time to become familiar with AI technologies.

Authors

Heather Demirjian
Cole Schotz et al

...

View Bio →

Inga Caldwell

...

View Bio →

Related Content

Air & Climate | Chemicals | Courts & Practice of Law | Enforcement & Oversight | Tribes, Justice, & Community | Water

Vantage Point

Feb 22, 2024

Courts & Practice of Law | Enforcement & Oversight

The Need for Process Safety Management Regulatory Reform

Feb 22, 2024

Courts & Practice of Law | Enforcement & Oversight | Tribes, Justice, & Community

Coastal Resiliency and the Inflation Reduction and Infrastructure Investment and Jobs Acts

Feb 22, 2024

Air & Climate | Courts & Practice of Law | Energy | Enforcement & Oversight

Navigating to Net Zero: Regulatory Regimes and Market Trends Impacting Decisions

Aug 07, 2023

Chemicals | Courts & Practice of Law | Enforcement & Oversight | Resources & Land Management

PFAS in Food Packaging: A Snapshot in Evolution of PFAS Regulation and Litigation

Jun 21, 2023

Chemicals | Courts & Practice of Law | Enforcement & Oversight

Regulating PFAS at the Edge of Detection

Jun 21, 2023

American Bar Association | /content/aba-cms-dotorg/en/groups/environment_energy_resources/resources/natural-resources-environment/2024-winter/artificial-intelligence-and-future-environmental-health-and-safety-management