2024 Urban Fire Forum (UFF) Position Statement

Safe and Ethical Use of Artificial Intelligence

Introduction:

Artificial intelligence (AI) holds extraordinary potential for both promise and peril. Responsible AI use has the potential to help solve urgent challenges while making our world more prosperous, productive, innovative, and secure. At the same time, irresponsible use could exacerbate societal harms such as fraud, discrimination, bias, and disinformation; displace and disempower workers; stifle competition; and pose risks to national security. Harnessing AI for good, particularly for the fire service and across public safety, and realizing its myriad benefits requires mitigating its substantial risks.[1]

As a body of knowledge, AI is a rapidly evolving, disruptive technology. AI technologies can drive inclusive economic growth and support scientific advancements that improve the conditions of our world. AI technologies, however, also pose risks that can negatively impact individuals, groups, organizations, communities, society, the environment, and the planet. Like risks for other types of technology, AI risks can emerge in a variety of ways and can be characterized as long- or short-term, higher low-probability, systemic or localized, and high- or low-impact.[2]The safety, security, and ethical concerns posed by AI must be factored into how local fire service and public safety decides to adopt and implement it within local fire departments. Many industries advocate for clearer frameworks to better understand AI's role, and public safety is at the forefront of this discussion. We must be vigilant in our efforts to successfully adopt and implement best practices that promote positive outcomes and limit negative impacts.

Artificial Intelligence must be safe and secure. Meeting this goal requires robust, reliable, repeatable, and standardized evaluations of AI systems, as well as agency-level policies, and as appropriate, other

[1] White House. (30 October 2023) “Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence” <https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/>

[2] National Institute of Standards and Technology. (January 2023) “NIST AI Risk Management Framework” < https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf>

mechanisms to test, understand, and mitigate risks from these systems before they are put to use. It also requires addressing AI systems' most pressing security risks — including with respect to cybersecurity, critical infrastructure, and other security dangers — while navigating AI's opacity and complexity. Testing and evaluations, including post-deployment performance monitoring, will help ensure that AI systems function as intended, are resilient against misuse or dangerous modifications, are ethically developed and operated in a secure manner, and are compliant with applicable laws and policies. [3]

This position paper outlines the key safety and ethical considerations broadly surrounding the use of AI and advocates for a framework that ensures the responsible development and deployment of AI technologies among the local fire service. Our commitment to the safe and ethical use of AI is not a one-time pledge but rather an ongoing conversation and effort to monitor and adapt technology to the needs of the fire service.

Definitions:

We begin by defining the key terms necessary for understanding AI's role in governance and the key types of AI currently in use.

Artificial Intelligence 15 U.S.C. 9401(3): a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. Artificial intelligence systems use machine- and human-based inputs to perceive real and virtual environments; abstract such perceptions into models through analysis in an automated manner; and use model inference to formulate options for information or action.[4]

AI Model means a component of an information system that implements AI technology and uses computational, statistical, or machine-learning techniques to produce outputs from a given set of inputs.[5]

Machine Learning means a set of techniques that can be used to train AI algorithms to improve performance at a task based on data. It is also considered to be a field within artificial intelligence, focuses on the ability of computers to learn from provided data without being explicitly programmed for a particular task.[6]

Adversarial Machine Learning is the process of extracting information about the behavior and characteristics of an ML system and/or learning how to manipulate the inputs into an ML system in order to obtain a preferred outcome.

Natural Language Processing falls under the fields of computer science, linguistics, and artificial intelligence. NLP deals with how computers understand, process, and manipulate human languages. It can involve things like interpreting the semantic meaning of language, translating between human languages, or recognizing patterns in human languages. It makes use of statistical methods, machine learning, neural networks and text mining.

Generative AI is a technology that can create content, including text, images, audio, or video, when prompted by a user. Generative AI systems create responses using algorithms that are trained often

---

[3] Ibid

[4] 15 U.S.C. 9401. <https://uscode.house.gov/view.xhtml?req=(title:15%20section:9401%20edition:prelim)>

[6] National Institutes of Standards & Technology. "Artificial Intelligence: Adversarial Machine Learning." < https://www.nccoe.nist.gov/ai/adversarial-machine-learning>

on open-source information, such as text and images from the internet.[7] An example of generative AI includes tools that can change the language someone speaks in videos to reach multiple audiences without the need for translation services. This would be useful in fire prevention, education, and community risk reduction.[8]

Predictive AI blends statistical analysis with machine learning algorithms to find data patterns and forecast future outcomes. It analyzes historical data trends to assign weights to models that help predict certain events, such as solutions to classification problems and anomaly detection. Examples of predictive AI are neural networks, various forms of linear and logistic regression, clustering algorithms, and decision trees. Each of these can be applied to support communities and public safety through analysis of population growth patterns, as well as supporting critical infrastructure investments.

Synthetic Data Generation is a process in which seed data are used to create artificial data that have some of the statistical characteristics of the seed data.[9]

Data Governance is everything accomplished to ensure data is secure, private, accurate, available, and usable. It includes the actions people must take, the processes they must follow, and the technology that supports them throughout the data life cycle.

Data Governance is the framework of policies, definitions, and metrics by which data is secured, accurate, available and usable within an organization. Data Governance models establish authority, management and decision-making parameters related to the data produced or managed by the enterprise.[10]

Goal Statement

This position paper establishes a preliminary framework for the fire service's safe and ethical use of AI. It sets forth five guiding principles to ensure the deployment of AI aligns with the fire service's values, legal obligations, and standards while protecting the public's trust in our organizations.

Use of AI

At a minimum, fire service organizations shall create and adopt a data governance and data policies, including an AI governance policy. Both policies must be grounded in current proven practices and reflect the principles described below. Once the policy has been created and adopted, the fire service organization shall be trained to understand it and ensure compliance. Integral to this process shall be an evaluation of all AI systems currently used by the entity to ensure compliance and adherence to these principles. It should be noted that the local fire service would benefit greatly from the development and availability of a "Fire Department Data Governance" template and supporting templates for model data policies and AI-specific governance models and policies. Such templates would serve as a baseline for use by local fire departments nationwide and would drive the implementation of proven effective practices consistently nationwide.

---

[7] Government Accountability Office. "Science & Tech Spotlight: Generative AI." < https://www.gao.gov/assets/830/826491.pdf>

[8] Ibid

[9] Ibid

[10] National Institute of Standards and Technology. "Computer Security Resource Center." <https://csrc.nist.gov/glossary/term/data_governance>

Fire service organizations should ensure the AI application has clear value before deployment. Simply adding AI to a process does not necessarily make that process better; the use of AI must be evaluated to ensure that it improves service delivery to stakeholders.

How, why, where, and to what extent a fire department uses AI must be governed, documented, and able to be plainly described to stakeholders.

Government employees must not rely solely on output from generative AI tools to make decisions that could adversely impact individuals or communities, including civil rights, civil liberties, or physical or mental health.

Safe and Ethical Principles

It is incumbent upon the fire service to develop an understanding of the application, outputs, and limits of AI technologies. The fire service confirms its validity and safe application to our communities when using AI. Fire departments must be knowledgeable, informed consumers capable of understanding and communicating how and why their organization uses and interacts with AI. This requires not necessarily deep technological and theoretical knowledge but a thoughtful and risk-averse application. It is essential for the entire fire service, including the vendor community that services it, to ensure that the trust the fire service has earned within its communities is preserved.

Fire service organizations must develop the necessary information, training, structures, and guidance to support fire departments in engaging with AI. Training should include an overview of AI's foundational aspects, including how it functions and the underlying data it uses to inform its processes. In support of this, believe the following five principles are critical to the effective use of AI in public safety:

1. **Safety and Security**

The deployment of AI systems for the fire service requires stringent safety measures. AI must be rigorously tested to ensure it does not cause harm, whether through malfunction, misuse, or unintended consequences. Additionally, AI systems must be secure against cyber threats, ensuring that malicious actors cannot easily manipulate or exploit them.

In addition to the aforementioned areas, we are mindful of the climate surrounding AI. As the UFF, we seek input from relevant government agencies to assist with the safe application of AI within the fire service. To this end, we are monitoring and actively implementing strategies from best practices surrounding ethical AI, including the National Institute of Technology (NIST)'s AI risk management framework[11] and the International Organization for Standardization (ISO) 42001 standard. Additionally, we encourage all fire service organizations to monitor relevant organizations dedicated to ensuring that we responsibly use AI and technologies to protect our firefighters, emergency medical professionals, and communities.

2. **Privacy and Data Protection**

AI systems often rely on vast amounts of personal data to function effectively. This creates significant concern about privacy and data security; therefore, data governance policy shall include appropriate direction and requirements. Ethical use of AI mandates that data privacy is respected, with data collection and processing conducted with explicit consent from the individual. Data should be

[11] Ibid

anonymized where possible, and robust security measures must be in place to protect against unauthorized access and breaches.

3. Fairness and Non-Discrimination

AI has the potential to exacerbate existing biases if not carefully managed. Algorithms trained on biased data (e.g., data labeled by humans) can produce discriminatory outcomes, perpetuating inequalities in areas like hiring, policy enforcement, and access to services. Developing and deploying AI systems that are fair, inclusive, and designed to minimize bias is crucial. This requires ongoing monitoring, ensuring the use of diverse data sets, and inclusive development and implementation teams that reflect the diversity of the populations we serve, as we are all affected by AI decisions.

4. Transparency and Accountability

AI technologies shall be transparent in their operations, allowing users and stakeholders to understand how decisions are made. This is crucial to ensure AI is used responsibly and build stakeholder trust. Developers and organizations must be accountable for AI-driven actions, particularly when these technologies are involved in critical decisions related to service delivery to our communities. Clear documentation and explanation of AI decision-making processes are essential to prevent misuse and to address potential biases. Any community member, internal, or external stakeholder interaction with a generative AI tool must include notice that the response may contain AI-generated content or data.

Any use of AI should contemplate the potential answer to an open records request and explain how or why the organization utilized the technology's output. As an organization, we advocate for the development of metrics necessary to assess the monitoring, sourcing, and quality of training datasets for AI models, and we also must advocate for understanding the limits of AI technologies.

5. Autonomy and Human-Centered Design

AI should enhance human autonomy rather than undermine it. Systems must be designed to support human decision-making, providing users with the tools and information needed to make informed choices. This includes the ability to override AI decisions when necessary. AI should not be designed to manipulate or control human behavior in ways that compromise individual autonomy or well-being. For example, we do not advocate for AI systems to replace vital labor. In turn, we expect all fire service organizations to carefully track AI's impact on their employees/volunteers, management, and the surrounding community. The goal should be to augment and support the fire service workforce while understanding the limitations of AI models.

Summary

Fire service organizations should confidently yet cautiously approach Artificial Intelligence as a valuable technology to speed, leverage, and enhance current practices and decision-making tools. We must avoid solutionism[12] while ensuring that the use and application improve or enhance the ability of a fire service organization to provide service to and for the communities they protect. Most importantly, the local fire service agencies must take deliberative action to establish a strong

[12] Solutionism: A term coined by the technology critic Evgeny Morozov, technological solutionism is the mistaken belief that we can make great progress on alleviating complex dilemmas, if not remedy them entirely, by reducing their core issues to simpler engineering problems.

foundation of agency-level data governance and policy as a precursor to implementing any artificial intelligence model, method, or system.

Lastly, as identified above, the local fire service needs appropriate resources and guidance on data governance and policy, such as templates, and a mechanism to share and exchange the latest best practices and lessons learned in AI implementation and usage across agencies.