The Legal Rewards and Risks of Artificial Intelligence in Social Services Operations

By Paula Mallory Engel, Chief Welfare Attorney, Onondaga County
NY Public Welfare Association, Winter Conference 2025

Introduction

The adoption of artificial intelligence (AI) across various sectors marks a turning point in how organizations approach problem-solving and decision-making. In the field of social services, AI promises to improve efficiency, accuracy, and resource allocation. However, the use of AI also introduces significant legal, ethical, and operational risks, particularly concerning privacy, bias, and accountability. This paper explores the concept of AI, the regulatory landscape governing its use, its potential benefits, and its associated risks. The focus is on how local departments of social services (LDSS) can integrate AI responsibly while ensuring a high level of public trust and remaining compliant with federal and state laws.

Of the many ways you might measure the potential value of AI on governments, one statistic jumps out. According to Gartner®, the annual spend on AI software by use case, digital government services, is projected to reach $41.8 billion in 2027. That tops all other industry sectors, with banking coming in second at $28.2 billion.[1] This represents a significant shift in priorities, as governments recognize the potential of AI to enhance public sector efficiency, transparency, and citizen engagement.

In particular, the USDA released its first official guidance on SNAP in early 2024[2] to respond to the "widespread excitement" from SNAP agencies to use AI to address capacity and resource constraints in an environment with constant pressure to meet metrics for timely processing and payment accuracy.[3]

I. Understanding Artificial Intelligence

1. Definition and Scope

AI refers to systems capable of simulating human intelligence by using machine learning (ML), natural language processing (NLP), and other technologies to process data, learn patterns, and perform tasks autonomously. The federal statutory definition of Artificial Intelligence is "a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments."[4]

A more palatable definition might be that AI is like a digital assistant that helps organize data to help make informed decisions.[5]

AI includes various subfields:

- **Machine Learning (ML):** Uses algorithms to process data and improve performance through experience.
- **Generative AI:** Creates original content (e.g., text, images) using advanced neural networks like GPT (Generative Pre-trained Transformers). Examples include ChatGPT, DALL-E, and Google Bard.
- **Deep Learning:** Mimics human neural networks to process complex datasets and enhance prediction accuracy.

2. **Distinguishing true AI from basic digital assistants[6]**

Basic digital assistants use predefined intent categories to determine what the user wants (from their inputs), search for a matching response in a database, and produce a response. On the other hand, AI virtual assistants utilize natural language processing to interpret the user input more accurately and leverage machine learning and deep learning algorithms to generate a response or perform a specific task. That said, AI-powered chatbots also use NLP to process user inputs.

Chatbots and digital virtual assistants may use decision trees to predict or classify outcomes. However, most chatbots have a simple decision tree or algorithm to process user inputs, classify them, and determine a suitable response. The rule-based chatbot's decision tree is preprogrammed, meaning it can only process inputs within its scope, resulting in a more focused functionality. AI virtual assistants utilize more capable machine learning algorithms for prediction and classification. Such ML models can adapt to new data, enabling digital virtual assistants to have broader functionalities, including managing dynamic user interactions.

A more user-friendly definition of Generative AI (GenAI), like ChatGPT is that it "acts like a creative advisor, not only assisting in regular tasks but also capable of drafting legal documents, designing educational materials, or creating public service announcements, showcasing its ability to produce new and original content when you ask it questions, otherwise known as Prompts."[7]

3. **Evolution of AI in Social Services**

Early applications of automation, like basic chatbots, relied on simple rules-based algorithms. Modern AI systems use NLP and ML to understand context, make predictions, and provide tailored responses. For example, AI-powered call centers now use conversational agents to handle inquiries efficiently, reducing wait times and improving accessibility.

As one researcher reported, social work agencies are adopting AI increasingly to do various important functions, such as to "conduct risk assessments, assist people in crisis, strengthen prevention efforts, identify systemic biases in the delivery of social services, provide social work education, and predict social worker burnout and service outcomes, among other uses."[8] Among the ethical challenges of this adoption, top issues include

"informed consent and client autonomy; privacy and confidentiality; transparency; client misdiagnosis; client abandonment; client surveillance; plagiarism, dishonesty, fraud, and misrepresentation; algorithmic bias and unfairness; and use of evidence-based AI tools."[9]

Specific concerns regarding AI's potential for bias and inaccurate outputs, which issues might significantly impact SNAP recipients' rights or safety, the USDA has identified some low-level risk AI applications in SNAP processing:

- Interactive Voice Recognition (IVR) technology that uses voice recognition to assist callers in navigating menus and routing calls;
- Optical Character Recognition (OCR) that transcribes information from uploaded documents or paper forms;
- Chatbots using natural language processing to understand questions, with human-coded, logic-based preset outputs, not generative AI responses;
- Sentiment analysis/natural language processing that categorizes themes and trends in unstructured text for customer experience and customer satisfaction surveys, helpdesk tickets, or social media posts referencing a benefits programs;
- Creation of synthetic data for testing information technology systems; and AI-enabled search tools that answer questions about program requirements or policies by directing caseworkers to the relevant section of an official policy manual or other primary source.

A few local districts are already working with OTDA to develop and implement some of the call center possibilities. Human oversight remains critical to maintaining internal and external trust in the AI performance.

II. Legal Framework for AI Deployment

1. Federal Regulations and Guidance

At the writing of this paper, the federal framework for implementing AI in the public sector is rapidly changing. The advent of President Donald Trump's second administration is already setting in motion new rules in keeping with his stated belief that "AI development should be rooted in free speech and human flourishing."[10]

a. Executive Order 14110 (2023):[11] Signed by President Biden, this order outlines principles for the safe and ethical deployment of AI, prioritizing:
   - Privacy and civil liberties.
   - Equity and protection against bias.
   - Responsible innovation to support American workers and consumers.

There are eight guiding principles and priorities set forth by President Biden in WO 14110 in setting the framework for federal agencies' adoption of AI:

(1) **ensure safe and secure AI technology;**

(2) **promote responsible innovation, competition, and collaboration;**
(3) **support American workers;**
(4) **advance equity and civil rights;**
(5) **protect American consumers, patients, passengers, and students;**
(6) **protect privacy and civil liberties;**
(7) **manage the federal government's use of AI; and**
(8) **strengthen US leadership abroad, promoting safeguards so that AI technology is developed and deployed responsibly**

Currently, this set of mandates only applies to federal agencies. However, as the LDSS is the administrative/operational arm of several federal programs through NYS law, county and city districts must be compliant with the structures imposed. "When undertaking the actions set forth in this order, executive departments and agencies (agencies) shall, as appropriate and consistent with applicable law, adhere to these principles, while, as feasible, taking into account the views of other agencies, industry, members of academia, civil society, labor unions, international allies and partners, and other relevant organizations."[12]

b. Background Executive Orders underpinning EO 14110

- Fair Information Practice Principles (FIPPs), 1973 Advisory Committee to US Department of Health, Education and Welfare: "*Records, Computers and The Rights of Citizens*"[13]
- EO 13960 – "*Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government*"[14] (Biden, October 2020)
- EO 13985 – "*Advancing Racial Equity and Support for Underserved Communities Through the Federal Government*"[15] (Biden, February 2023, repealed by Trump, January 2025) – rescinded by President Trump on January 20, 2025.

c. HHS/ACF Reports (2022-2024): Studies conducted by the Department of Health and Human Services (HHS) outline the risks and benefits of AI in federal programs, emphasizing data privacy and equity in decision-making.[16]

d. US Department of Homeland Security Framework for Safe Deployment of AI: Identifies, among other things, the public sector roles and responsibilities in ensuring safety of AI deployment.

On Nov. 14, 2024, the Department of Homeland Security released a set of recommendations (handout). This first-of-its kind VOLUNTARY resource was developed by and for entities at each layer of the AI supply chain: cloud and compute providers, AI developers, and critical infrastructure owners and operators – as well as the civil society and public sector entities that protect and advocate for consumers. The Artificial Intelligence Safety and Security Board ("Board"), a public-private advisory committee established by DHS Secretary Alejandro N. Mayorkas, identified the need for clear guidance on how each layer of the AI supply chain can do their part to ensure that AI is deployed safely and securely in U.S. critical infrastructure…vulnerabilities introduced by the

implementation of this technology may expose critical systems to failures or manipulation by nefarious actors. Given the increasingly interconnected nature of these systems, their disruption can have devastating consequences for homeland security.[17]

Among the critical needs identified by DHS in this framework is the focus on respect for civil rights. The use of AI in critical infrastructure and its corresponding costs and benefits will vary depending on the specific application, the context of the sector and use case, and many other factors. Nevertheless, the consideration of privacy, civil rights, and civil liberties is foundational and must be carried across all AI systems. Accordingly, this Framework makes safeguarding civil rights, identifying disparate impacts, and mitigating harm shared responsibilities across the full AI ecosystem that supports the development and deployment of AI in critical infrastructure.[18]

§

e. Use of AI by US Department of Health and Human Services

HHS developed its AI Strategic Plan in late 2024 and released it on January 15, 2025. (Initially, the report was publicly available, but less than a week later, was no longer freely accessible on the internet, with no explanation.) One of that plan's identified current uses of AI is in the Children Welfare Information Gateway - a hotline for answering questions or requesting information on resources. Another current use in gathering and answering collective bargaining questions. In all., the US Government Accountability Office reported in 2023 that HHS had identified 271 AI use cases across thirteen of the Department's agencies.[19]

HHS/ACF released its opportunities, challenges, and risk assessment report on AI in September, 2022:[20]

The agency was tasked in 2020 to do a study with three objectives:

(1) **Understand AI** and how HHS could leverage it;
(2) Understand existing and potential **barriers, facilitators, risks and benefits;**

(3) Identify **options and opportunities** to address and mitigate the existing and potential risks, as well as promote benefits of using AI in mission work.

The September 2022 report addressed Objective #3: Risks and Benefits. The following is an excerpt from that report:

Seven categories of challenges were identified from the case studies and literature reviewed:

(1) User confidence and trust.

Trust in the reliability of AI models is an overarching challenge, as nearly every other challenge identified relates to some aspect of trust. Specific challenges primarily focused on the ability for decision makers to adequately interpret, evaluate, and act on information provided by a model ("justified trust") and trust by citizens, including those who may not realize they are being affected by decisions based on a model or a model's outputs ("public trust").

(2) Model performance.

Challenges were identified related to the fit-to-use, accuracy, or robustness of predictive AI models related to putting a model into practice. These challenges relate to whether the model accomplishes what it was designed to do, and issues related to addressing changes in the model or in the deployment setting over time.

(3) Maintaining privacy.

Weighing privacy risks against potential benefits can be a key challenge in deciding whether to pursue the development or use of AI. Maintaining privacy is seen as essential for people to maintain some degree of autonomy. Promoting and using strong privacy practices can help to build trust in AI. At the same time, strong privacy protections can have tradeoffs. For instance, prioritizing privacy may contribute to less transparency in the models and their underlying data or can introduce limitations into models' performance.

Challenges associated with privacy concerns were identified at all stages of the AI lifecycle and were related to data gathering practices, data protections and security (particularly for sensitive data), and model use.

(4) Bias.

The National Institute of Standards and Technology has identified three categories of bias in how AI is designed, developed, and used:

a. statistical and computational biases that occur when a model's underlying data is not representative of the population the model is addressing;

b. systemic biases related to how data can capture or reflect historical and ongoing inequities; and

c. human bias in model development and results interpretation.

Explicit examples of statistical and computational biases and systemic biases were identified in the case studies. Human bias was not explicitly identified in the case studies but was implicit in some of them.

(5) Data and dataset quality.

Challenges related to data and dataset quality occur when data and datasets are incomplete, incorrect, nonrepresentative, or outdated. Poor data quality can contribute to insufficient model performance and bias.

(6) Transparency and explainability.

A lack of transparency and explainability can contribute to decreased trust, decreased capability in determining how a model can be used, or decreased capability for testing and evaluating models and identifying limitations. In the context of our case studies, "transparency" refers to the actions and operations surrounding a model and its outputs being visible to and understandable by desired parties. "Explainability" is sometimes used interchangeably but is treated as a distinct concept that refers specifically to the ability to understand how a model arrives at a particular outcome given a certain input or set of inputs.

(7) Capacity.

Challenges related to capacity included the usability of systems, limitations on computational resources and other computing infrastructure necessary to implement AI, and limited expertise and workforce available to develop, use, and govern AI

(Each one of these concerns is addressed in the best practices section, below, other than capacity, which is left to the operation and informational technology subject matter experts.)

2. State-Level Statutory Framework and Guidance

a. New Article 4 of the State Technology Law enacted on December 21, 2024.

The landscape of New York AI regulation is also rapidly evolving. At the end of 2024, New York State adopted a statutory scheme representing one of the most far reaching sets of requirements and restrictions on the adoption of AI in the public sector.

The law (new Article 4 to the NY State Technology Law) necessitates state agencies to conduct thorough reviews and report their AI usage. These reports are not only submitted to the governor and legislative leaders but are also made publicly available online to ensure transparency.

Moreover, the legislation imposes restrictions on the use of AI in making certain critical decisions—such as those involving unemployment benefits and childcare assistance—without human oversight. This step is vital to prevent potentially biased algorithms from causing discriminatory outcomes. Additionally, it safeguards state employees from AI-driven changes in their job duties or reductions in work hours, reflecting the law's attention to human-centric protections in the workplace.

According to one media source, "The push for this legislation also stems from broader public and legislative scrutiny on AI technologies. Public concerns about AI's impact on transparency, accountability, and job security have driven state legislators to act. This law represents a proactive step to balance the potential of AI to enhance governmental efficiency while safeguarding public interests and ensuring ethical use of technology in decision-making. By implementing these measures, New York sets a precedent for AI governance that could inspire similar regulations in other jurisdictions."[21]

In urging Governor Hochul to sign the legislation, the Public Employees Federation, one of the state's largest unions and which represents more than 500K members, cited "significant errors" in existing AI-powered state work. The union's president, Wayne Spence, said, "Unregulated and nontransparent expansion of AI systems into government decision making processes – like social services, unemployment insurance benefit determinations, and workers' compensation claims — is completely irresponsible and potentially dangerous."[22]

These mandates (reporting and definitions are effective on December 21, 2025 [Sections 401, 403-404], meaningful review requirements are effective on December 21, 2025 [Section 402]), were enacted to safeguard individuals from adverse effects of AI decision-making. This law requires:

1. **Human Oversight to be in place by 12/21/2025:**[23] Prohibiting fully automated decisions without meaningful review by qualified individuals.[24] State agencies – including districts (LDSS) or entity operating on behalf of state agencies (ditto) – requires "continued and operational meaningful human oversight" when using AI which directly or indirectly -

    (a) relates to the delivery of any public assistance benefit;
    (b) will have a material impact on the rights, civil liberties, safety or welfare of any individual within the state;
    (c) affects any statutorily or constitutionally provided right of an individual.[25]

The statute defines "**meaningful review**" as follows:

"review, oversight and control of the automated decision-making process by one or more individuals who understand the risks, limitations, and functionality of, and are trained to use, the automated decision-making system and who have the authority to intervene or alter the decision under review, including but not limited to the ability to approve, deny, or modify any decision recommended or made by the automated system."[26]

2. **Impact Assessments (eff. 12/21/2024):** Regular evaluations to ensure AI systems operate lawfully and without discriminatory outcomes.[27]

   Before adopting any new AI system, and every two years thereafter, agencies must submit reports to the Governor and the legislature heads an impact assessment report for each AI system, which outlines the following:

   (a) a description of the objectives of the automated decision-making system;

   (b) an evaluation of the ability of the automated decision-making system to achieve its stated objectives;

   (c) a description and evaluation of the objectives and development of the automated decision-making including:

      (i) a summary of the underlying algorithms, computational modes, and artificial intelligence tools that are used within the automated decision-making system; and

      (ii) the design and training data used to develop the automated decision-making system process;

(d) testing for:

      (i) accuracy, fairness, bias and discrimination, and an assessment of whether the use of the automated decision-making system produces discriminatory results on the basis of a consumer's or a class of consumers' actual or perceived race, color, ethnicity, religion,
national origin, sex, gen der, gender identity, sexual orientation, familial status, biometric information, lawful source of income, or disability and outlines mitigations for any identified performance differences in outcomes across relevant groups impacted by such use;

      (ii) any cybersecurity vulnerabilities and privacy risks resulting from the deployment and use of the automated decision-making system, and the development or existence of safeguards to mitigate the risks;

      (iii) any public health or safety risks resulting from the deployment and use of

the automated decision-making system;

(iv) any reasonably foreseeable misuse of the automated decision-making system and the development or existence of safeguards against such misuse.

Section 404 on the filing/publication of these reports also addresses agencies which have already deployed AI systems prior to December 21, 2025. The law requires those agencies to file initial impact statements regarding those systems no later than December 21, 2025.[28]

3. **Protection of public employees** from job replacement and impact from the deployment of AI.[29]

State workers will be shielded from having their hours or job duties limited because of AI under the law.

b. **NYS Information Technology Policy:** Issued by the Office of Information Technology Services (OITS), this policy includes a robust risk assessment framework covering security, privacy, and bias.[30] (handout)

This State IT Policy sets recommended guidelines for the use of artificial intelligence and is intended as a "tool to aid" state entities in adopting new AI systems. The policy was disseminated in keeping with OITS statutory and administrative authority.[31]

The policy already recommended many of the mandates that are now part of the new Article 4 of the State Technology Law.

4.2 Human Oversight (*now*, NY State Tech 402)
4.3 Fairness and Equity, and Explain Ability [*sic*] (*now*, NY State Tech 402)
4.4 Transparency (§ 402)
4.5 AI Risk Assessment and Management (§ 402)
4.6 AI Inventory (§ 403)
4.7 Privacy (Existing Law)
4.8 Security (Existing Law)
4.9 Technology (encouraging the use of open standards, model lifecycle management, and regular re-training of the AI systems)
5.10 Intellectual Property (recommending communication with counsel's office about using, for example, copyrighted materials as inputs or AI-generated outputs that contain copyrighted elements.[32]

Specific Highlights of the Acceptable Use Policy:

NYS OITS has identified the importance of "trustworthiness" in any AI system the government uses. The National Institute of Standards and Technology (NIST), a division of the US Department of Commerce, has published many useful tools to develop, rollout,

manage, explain, and test AI systems so that they are trustworthy. "Trustworthy AI" is defined as having the characteristics that it is "valid, reliable, safe, secure, resilient, accountable, transparent, explainable, interpretable, privacy-enhanced, and fair with human bias managed."[33]

4.2 **Human Oversight**

What does human oversight really mean? In the context of identified AI possible uses in SNAP administration, the USDA[34] has identified these examples of humans in the loop:

AI Function or Output	Human Oversight
AI-enabled tool is used to create, summarize, or transform (such as rewriting in plain language) public facing program materials	A human with program expertise holds accountability for the output and reviews for validation
Translating vital documents and/or public facing program materials	A skilled human translator validates and holds accountability for the output
Providing auto captioning	Auto-captioning functionality is offered in addition to having a live American Sign Language interpreter
Converting scans of physical documents into machine readable formats for further analysis and/or improved accessibility	A human with program expertise provides validation for both input (original document) and outputs of (document in a machine-readable format)

4.4 **Transparency**

The OITS Acceptable Use Policy and the new Article 4 of the State Technology Law both require that the AI systems be "transparent," which essentially as three components: algorithmic transparency, interaction transparency and social transparency.

Algorithmic transparency means that you can explain where the inputs come from (data labeling), how the data is fed into the AI system, how the AI system's algorithms give resulting output, and how you test for bias and competency in the outputs. Transparency can help mitigate issues of fairness, discrimination, and trust.[35]

However, as published research in the Harvard Business Review pointed out, "it is becoming clear that disclosures about AI pose their own risks: Explanations can be hacked, releasing additional information may make AI more vulnerable to attacks, and disclosures can make companies more susceptible to lawsuits or regulatory action."[36]

The researcher called it AI's "transparency paradox,"[37] requiring AI management to focus not only on the identified risk areas, but also how detailed their reports are on the algorithm so the reports does not outline easy avenues for bad actors to disrupt government operations with malware and ransomware— while generating more information about AI might create real benefits, it may also create new risks.

Interaction transparency. The second level of transparency is on the user interface. The same Harvard Business Review article provided the following example, which can highlight the risks to the "public trust" by not being transparent about when and how they are interfacing with an AI system:

> "In 2018, one of the largest tech companies in the world premiered an AI that called restaurants and impersonated a human to make reservations. To "prove" it was human, the company trained the AI to insert "umms" and "ahhs" into its request: for instance, "When would I like the reservation? Ummm, 8 PM please."
>
> The backlash was immediate: journalists and citizens objected that people were being deceived into thinking they were interacting with another person, not a robot. People felt lied to."

Interaction transparency deals with the communication and interactions between users and AI systems. It involves making exchanges more transparent and understandable. Businesses can achieve this by creating interfaces that communicate how the AI system operates and what users can expect when interacting with it.[38]

There are several ways to address interaction transparency, but they include opting it or out banners or buttons or messages; quick and easy exits to a human; and information on how the conversation/interaction will be stored and used.

Social transparency extends beyond the technical aspects and focuses on the broader impact of AI systems on society as a whole. This level of transparency addresses the ethical and societal implications of AI deployment, including potential biases, fairness, and privacy concerns.[39] The required Impact Assessment reports under the new State Technology Law § 403 prompts the state entity to explain the risk assessment and mitigate those issues to the satisfaction of the Governor and Legislature.

The law requires that if an AI system fails to generate unbiased and non-discriminatory outcomes, the state entity must immediately discontinue its use until the issues are identified and resolved.[40]

The advantages of AI transparency include the following:[41]

- **Builds trust with the public and stakeholders;**
- **Promotes accountability and responsible use of AI;**
- **Detects and mitigates data biases and discrimination;**
- **Improves AI performance; and**
- **Addresses ethical issues and concerns.**

4.5 **Requirement of a Risk Assessment:**

NYS AI Policy requires that state entities (a) frame the risk; (b) assess the risk; (c) respond to the risk; and (d) monitor the risk of using AI systems in their work.

Components of an acceptable risk assessment process include addressing risks associated with

(a) Security and Privacy/Confidentiality
(b) Legal and Reputational Issues (including bias)
(c) Competency

A. Security and Privacy/Confidentiality

An acceptable AI policy should outline measures to protect data privacy and secure AI systems from cyber threats. Many free AI products train their large language model computers to evolve with every use, so information used in prompts is not private.

Foremost among concerns is a security or data privacy breach. If sensitive information — such as customer data or confidential business information — is put into a generative AI platform that is not secure, the information could be offered somewhere else or be incorporated into training, which would effectively make it public.[42]

B. Legal and Reputational Issues, including Bias

A highly publicized whistleblower from Google's Ethical AI Team claimed she was fired after she wrote a paper highlighting bias in AI. Timnit Gebru, who was a co-leader of Google's Ethical A.I. team, said she was fired after criticizing Google's approaches to minority hiring and the biases built into today's AI systems.[43]

The New York Times article noted that researchers "worry that the people who are building artificial intelligence systems may be building their own biases into the technology. Over the past several years, several public experiments have shown that the systems often interact differently with people of color — perhaps because they are underrepresented among the developers who create those systems."[44]

As another example in the healthcare industry, underrepresenting data of women or minority groups can skew predictive AI algorithms. For example, computer-aided diagnosis (CAD) systems have been found to return lower accuracy results for African American patients than white patients.[45]

C. Competency

Underpinning all of the laws and regulations surrounding public sector use of AI is the need for the public and stakeholders to have faith that the outputs of AI are (1) accurate, (2) robust; (3) fair; (4) able to handle diverse situations, (5) protective of confidential and/or protected information; and (6) as free as possible from bias or other ethical implications.

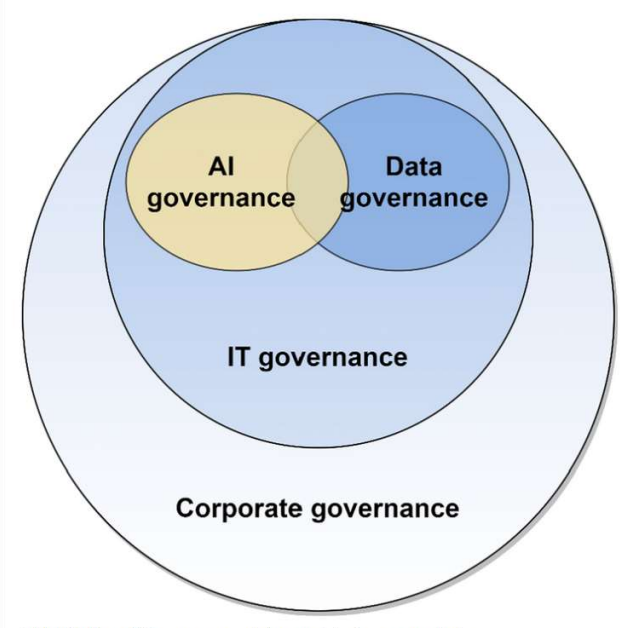
As one scholar put it, "we all know that the quality of the data used for AI learning and customized use matters, but we still have no means to judge data quality and, thus, AI validity and transparency."[46]

Accordingly, federal law (EO 14110) and state law (NY State Technology Law Article 4) both require that AI systems be tested regularly and heavily with reliable measurements and evaluations of their underlying technologies and use. NIST is consistently advancing the measurement science of AI, so having the appropriate in house and externally competent AI experts checking for and utilizing the latest tools to assess competency of the outputs.[47]

III. The Requirement of Having a Sound AI Governance Structure

How do we ensure that our AI system does what we intended it to do, in the way we intended it, and is giving us "trustworthy" outputs? This is the result of building a sound AI governance structure PRIOR to developing and implementing any AI system. "Artificial intelligence governance refers to the **policies, regulations, and ethical guidelines** that govern the development, deployment, and use of AI technologies. It encompasses a range of issues, including data privacy, algorithmic transparency, accountability, and fairness."[48]

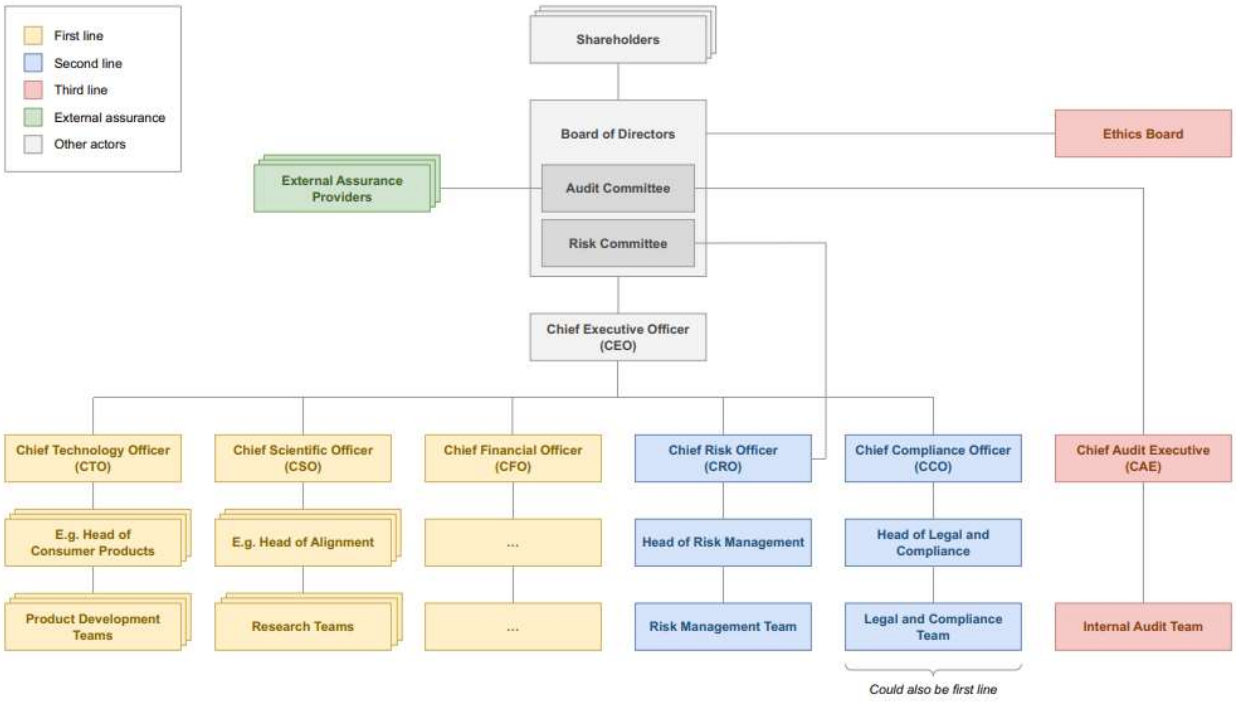
An AI governance structure not only addresses all of the issues of compliance, security, privacy, transparency, competency, etc., through policies and procedures, it also involves setting up a **clear management organizational chart**, with individuals clearly identified as to their roles and tasks in the governance structure. Below is an illustration of the overlapping charters of AI, data, IT and management personnel (insert "responsible public officials" for "Corporate Governance" roles):



Artificial intelligence (AI) governance as part of an organization's governance structure

Source: Mantymaki, M., et al. "Defining Organizational AI Governance."[49]

Here is a sample organizational chart of an AI governance structure:[50]



The benefits of a clear organizational chart and defined responsibilities include efficient coordination of different people from different teams with different responsibilities. Without a clear process for decision-making and evaluation, risk areas can go unidentified, or one team could think the other team had handled that issue.

Importantly, it puts the stakeholders at the top of the chart. For public sector agencies, stakeholders include the public (stakeholder committee) and the entity's governing body (elected legislature), and the structure ensures that accountability is woven into every decision made with respect to the use of AI in government operations.

A word about **stakeholder engagement committees**: The public trust in the government depends on the accountability, transparency and responsiveness of concerns of the residents. A well-constituted stakeholder engagement committee ensures that every level of concern with AI use is heard and addressed on an ongoing basis. Here are the key elements of an AI stakeholder engagement committee:[51]

Diverse membership (IT, vendors, legal, end users, key operational staff, community leaders, perhaps academia, external regulators)

Identified Stakeholders; analyzing and identifying all relevant stakeholders, understanding their needs, impacts and level of influence

Communication: establishing an open and transparent communication loop with scheduled meetings, surveys, workshops and feedback mechanisms

Ethical considerations: actively discussing risks related to bias, privacy, transparency, accountability and fairness in AI development

Risk Assessment (discussed above)

IV. Benefits of AI in LDSS

1. Operational Efficiency

AI can automate routine administrative tasks, such as processing applications and reviewing eligibility, freeing up staff to focus on complex cases. For instance, predictive analytics help agencies forecast demand for services during crises like economic downturns or natural disasters. Particular cases can be made for the use of AI in the public sector's social work mission in the following areas: call centers, cybersecurity and program integrity (recipient and vendor fraud), inventory management for shelter, housing, child care accessibility, supply chain operations (staffing models, purchase of services contract needs), and recruitment (assessing whether candidates meet minimum qualifications, but not other civil service or retention goals – see NY State Tech Law Article 4).[52]

2. Fraud Detection

AI's ability to analyze large datasets allows for more effective detection of fraudulent activities. By identifying anomalies in financial transactions or benefit claims, AI minimizes resource waste and ensures benefits are distributed equitably.

3. Improved Access and Personalization

AI-powered tools enable LDSS to provide personalized services, such as matching individuals with housing or childcare resources based on specific needs. AI systems also facilitate language translation, ensuring non-English speakers can access critical information.

V. Risks of AI in LDSS

1. Bias and Discrimination

As discussed above, AI models often inherit biases present in their training data, leading to discriminatory outcomes. For example, studies have shown that algorithms used in child welfare systems can disproportionately flag minority families for investigation. LDSS must actively test and mitigate biases to prevent systemic harm.

2. Privacy and Security Concerns

AI systems process sensitive personal data, including financial, health, and familial information. Without robust safeguards, these systems become vulnerable to breaches, potentially exposing intimate details of individuals' lives.

3. Accountability Challenges

The complexity of AI algorithms makes it difficult to attribute responsibility for errors or unintended consequences. This "black box" nature can erode trust and make compliance with transparency requirements challenging.

VI. Use of AI by Lawyers, Generally

There is a growing number of lawyers from New York who are facing possible disciplinary action after using ChatGPT or other GenAI systems to draft

court filings or create evidence that was found not to be "competent", i.e., to have references to previous cases that did not exist. Here are a few recent cases:

1. Counsel has an affirmative obligation to disclose the use of AI in evidence submitted to court, for review at a *Frye* hearing. Matter of Weber as Tr. of Michael S. Weber Tr., 220 N.Y.S.3d 620, 635 (N.Y. Sur. 2024) (Saratoga County).

   "In what may be an issue of first impression, at least in Surrogate's Court practice, this Court holds that due to the nature of the rapid evolution of artificial intelligence and its inherent reliability issues that prior to evidence being introduced which has been generated by an artificial intelligence product or system, counsel has an affirmative duty to disclose the use of artificial intelligence and the evidence sought to be admitted should properly be subject to a *Frye* hearing prior to its admission, the scope of which should be determined by the Court, either in a pre-trial hearing or at the time the evidence is offered." Matter of Weber as Tr. of Michael S. Weber Tr., 220 N.Y.S.3d 620, 635 (N.Y. Sur. 2024)

2. If evidence includes statistical analysis (gene typing) that was done by AI, then both the analyst who entered the data and parameters, as well as the AI expert who can explain the algorithm, must testify and be subject to cross-examination, at the *Frye* hearing. People v. Wakefield, 38 N.Y.3d 367, 386, 195 N.E.3d 19, 32 (2022) (Appeal from Schenectady County case).

   A report on genetic stereotyping was admissible after *Frye* hearing, where "both the analyst who performed the electrophoresis on the DNA samples and Dr. Perlin, who fully understood the parameters and methodology of the TrueAllele software in its DNA interpretation processes, testified at trial and were subject to cross-examination." People v. Wakefield, 38 N.Y.3d 367, 386, 195 N.E.3d 19, 32 (2022), relying in part on the President's Council of Advisors on Sci. and Tech., Exec. Office of the President, Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature–Comparison Methods, at 80 [2016]

https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf [published after the *Frye* hearing was held]), and NIST, DNA Mixture Interpretation: *A NIST Scientific Foundation Review,* at 3 [June 2021 Draft report] *https://nvlpubs.nist.gov/nistpubs/ir/2021/NIST.IR.8351–draft.pdf*).

3. In Mata v. Avianca, plaintiff's attorneys were found to have acted in bad faith and each were given a $5,000 penalty when they submitted "non-existent judicial opinions with fake quotes and citations created by […] ChatGPT," then stood by them, even under the Judge's questioning. No. 22-cv-1461 (PKC), 2023 WL 4114965 (S.D.N.Y. June 22, 2023).

   In an article on this case, the reporter noted that the "misunderstood nature of ChatGPT was made clear for the umpteenth time this weekend when news broke that US lawyer Steven A. Schwartz had turned to the chatbot to find supporting cases in a lawsuit he was pursuing against Colombian airline Avianca. The problem, of course, was that none of the cases ChatGPT suggested exist. […]

   Schwartz claims he was "unaware of the possibility that [ChatGPT's] content could be false," though transcripts of his conversation with the bot show he was suspicious enough to check his research. Unfortunately, he did so by asking ChatGPT, and again, the system misled him, reassuring him that its fictitious case history was legitimate.[53]

4. Attorney who submitted a document to Surrogate's court with fictional or erroneous citations from GenAI is subject to sanctions. Will of Samuel, 82 Misc. 3d 616, 622, 206 N.Y.S.3d 888, 892 (N.Y. Sur. 2024) (Kings County), *citing* 18 NYCRR 130-1.1(a).

5. *Pro se* plaintiff uses AI legal software to submit memorandum of law that is replete with fictitious citations, receives warning. Dowlah v. Pro. Staff Cong., 227 A.D.3d 609, 610, 213 N.Y.S.3d 13, 14 (2024), *leave to appeal denied*, No. 2024-592, 2025 WL 84301 (N.Y. Jan. 14, 2025) (appeal from NY County case).

   Memorandum of law with several nonexistent cases were "the result of research using "legal software applications" that deploy artificial intelligence." Dowlah v. Pro. Staff Cong., 227 A.D.3d 609, 610, 213 N.Y.S.3d 13, 14 (2024), leave to appeal denied, No. 2024-592, 2025 WL 84301 (N.Y. Jan. 14, 2025).

VII. Best Practices for AI Implementation

1. Strategic Planning

- Develop a roadmap outlining goals, decision-making hierarchies, and oversight mechanisms.

- Ensure compliance with legal mandates, such as regular impact assessments under NYS Tech Law § 403.

2. Stakeholder Engagement

- Involve diverse stakeholders, including community representatives, to identify potential risks and solutions during the planning phase.
- Establish feedback loops to address public concerns about AI's role in service delivery.

3. Transparency and Oversight

- Clearly label AI-generated outputs and explain how decisions are made.
- Create oversight boards to review AI system performance and address grievances.

4. Continuous Monitoring

- Conduct post-launch audits to evaluate the accuracy and fairness of AI outputs.
- Utilize dashboards to track performance metrics, ensuring compliance with privacy and equity standards.

VIII. Conclusion

Key Takeaways, summarized:

(A) BENEFITS OF USING AI IN YOUR LOCAL DISTRICT

Artificial Intelligence has the potential to:

(1) Increase efficiency and accuracy
(2) Free up staff to tackle more complex/nuanced work (*but not replace)*
(3) Reduce fraud, waste and abuse
(4) Help deploy resources quickly where really needed

(B) RISKS OF USING AI IN YOUR LOCAL DISTRICT

Artificial Intelligence has the potential to:

(1) Be a tremendous drain on resources while in development
(2) Require implementing new structures, policies and training, reporting
(3) Need to be constantly checked (bias, accuracy)
(4) Make our districts more vulnerable to cyber-attacks and data theft, the more we rely on machine learning to do the work

The integration of AI into LDSS operations presents an opportunity to enhance efficiency, reduce costs, and improve service delivery. However, this potential can only be realized through adherence to legal frameworks, proactive risk management, and a commitment to transparency and accountability. By adopting best practices and involving stakeholders at every stage, LDSS can harness AI's capabilities while safeguarding the rights and trust of the communities they serve.

Endnotes

[1] Gartner, "Compare AI Software Spending in the Government Industry," 2023-2027, By Daniel Snyder, James Ingham, Inna Agamirzian, 27 March 2024, *accessed* 21 January 2025.

[2] The FNS guidance can be found here: https://www.fns.usda.gov/snap/advanced-automation

[3] American Public Human Services Association (APHSA). "AI-Powered SNAP Modernization: An Introduction to Current and Potential Uses of AI in SNAP Case Processing - Digital Government Hub." Digital Government Hub, 18 Dec. 2024, digitalgovernmenthub.org/library/ai-powered-snap-modernization/. *Accessed* 22 Jan. 2025. (handout)

[4] 15 USC 9401(3), 3 C.F.R. 14110 (2023).

[5] Primer for Counties: The Transformative Power of Artificial Intelligence. (2024). National Association of Counties. https://www.naco.org/resource/primer-counties-transformative-power-artificial-intelligence, *accessed* 22 January 2025.

[6] *Excerpted from* The Upwork Team. (2024, August 22). AI Chatbot vs. AI Virtual Assistant: What's the Difference? Upwork.com; Upwork. https://www.upwork.com/resources/ai-chatbot-vs-virtual-assistant, *accessed* 21 January 2025.

[7] *Id.*

[8] Reamer, F. (2023). "Artificial Intelligence in Social Work: Emerging Ethical Issues." International Journal of Social Work Values and Ethics, 20(2), 52-71. https://doi.org/10.55521/10-020-205, at p. 53. (handout)

[9] *Id.*

[10] Press, A. (2024). The Outlook Is Uncertain for AI Regulations as the US Government Pivots to Full Republican Control. US News & World Report; U.S. News & World Report. https://www.usnews.com/news/politics/articles/2024-11-28/the-outlook-is-uncertain-for-ai-regulations-as-the-us-government-pivots-to-full-republican-control, *accessed* 07 January 2025.

[11] Exec. Order No. 14,110, 3 C.F.R. 14110 (2023).

[12] *Id*.

[13] Federal Privacy Council. (2024). FPC.gov. www.fpc.gov. https://www.fpc.gov/resources/fipps/, *accessed* 22 January 2025.

[14] Federal Register. (2020, December 8). Federal Register. https://www.federalregister.gov/documents/2020/12/08/2020-27065/promoting-the-use-of-trustworthy-artificial-intelligence-in-the-federal-government, *accessed* 17 January 2025.

[15] Federal Register. (2021, January 25). Advancing Racial Equity and Support for Underserved Communities Through the Federal Government. Federal Register. https://www.federalregister.gov/documents/2021/01/25/2021-01753/advancing-racial-equity-and-support-for-underserved-communities-through-the-federal-government, *accessed* 07 January 2025.

[16] HHS/ACF, **AI Strategic Plan** (2024), accessed 07 January 2025, now firewalled and no longer publicly available at https://www.childwelfare.gov, *accessed* 07 January 2025, attempted again 21 January 2025.

[17] Groundbreaking Framework for the Safe and Secure Deployment of AI in Critical Infrastructure Unveiled by Department of Homeland Security | Homeland Security. (2024). U.S. Department of Homeland Security. https://www.dhs.gov/archive/news/2024/11/14/groundbreaking-framework-safe-and-secure-deployment-ai-critical-infrastructure, *accessed* 07 January 2025.

[18] *Id.*

[19] U.S. Government Accountability Office. (2023, December 12). Artificial Intelligence: Agencies Have Begun Implementation but Need to Complete Key Requirements | U.S. GAO. Www.gao.gov. https://www.gao.gov/products/gao-24-105980, *accessed* 07 January 2025.

[20] Options and Opportunities to Address and Mitigate the Existing and Potential Risks, As Well As Promote Benefits, Associated with AI And Other Advanced Analytic Methods. (2023). hhs.gov. https://www.acf.hhs.gov/opre/report/options-opportunities-address-mitigate-existing-potential-risks-promote-benefits, *accessed* 07 January 2025.

[21] Justin, R. (2024, December 24). New law will require state agencies to monitor use of generative AI. *Times Union*. https://www.timesunion.com/capitol/article/new-law-require-state-agencies-monitor-use-19999749.php, *accessed* 17 January 2025.

[22] *Id.*

[23] N.Y. State Tech. Law Article 4 (Added L.2024, c. 674, § 2, eff. Dec. 21, 2024). Interestingly, Westlaw and other research sites list the effective date as 12/21/2025. However, the legislation that Governor Hochul signed on 12/21/2024 states at § 4 thereof: "This act shall take effect immediately, provided that section two of this act shall take effect one year after it shall have become a law. NY LEGIS 674 (2024), 2024 Sess. Law News of N.Y. Ch. 674 (S. 7543-B) (McKinney's).

[24] N.Y. State Tech. Law §§ 402 and 403(2) (2024).

[25] N.Y. State Tech. Law § 402(1) (2024).

[26] N.Y. State Tech. Law § 401(2) (2024).

[27] N.Y. State Tech. Law § 402 (2024).

[28] N.Y. State Tech. Law § 404(3) (2024)

[29] N.Y. State Tech. Law § 402(3) (2024)

[30] New York State IT Policy NYS-P24-001: Acceptable Use of Artificial Intelligence Technologies. Albany, New York, NYS Office of Information Technology Services, 8 Jan. 2024, its.ny.gov/system/files/documents/2024/01/, *accessed* 7 January 2025.

[31] N.Y. State Tech. Law § 103(10) (2024).

[32] *See* NY State IT Policy NYS-P-24-001, which also references guidance from the federal government to prevent unauthorized use of protected intellectual property: https://www.federalregister.gov/documents/2023/03/16/2023-

05321/copyright-registration-guidance-works-containing-material-generated-by-artificial-intelligence

[33] NIST. "AI Risk Management Framework." NIST, 12 July 2021, www.nist.gov/itl/ai-risk-management-framework, *accessed* 22 January 2025.

[34] *See*, endnote 2.

[35] Wren, Hannah. "What Is AI Transparency? A Comprehensive Guide." Zendesk, 18 Jan. 2024, www.zendesk.com/blog/ai-transparency/, *accessed* 20 January 2025.

[36] Burt, Andrew. "The AI Transparency Paradox." Harvard Business Review, 13 Dec. 2019, hbr.org/2019/12/the-ai-transparency-paradox, *accessed* 12 December 2024.

[37] *Id.*

[38] Wren, "What is Transparency?", *supra*.

[39] *Id.*

[40] N.Y. State Tech. Law § 403(2) (2024), effective December 21, 2025.

[41] *Id.*

[42] "AI in the Workplace: Navigating Benefits, Security and Ethics."www.jamf.com, www.jamf.com/blog/ai-in-the-workplace-balancing-benefits-and-security/, *accessed* 17 January 2025.

[43] Metz, Cade, and Daisuke Wakabayashi. "Google Researcher Says She Was Fired over Paper Highlighting Bias in A.I." The New York Times, 3 Dec. 2020, www.nytimes.com/2020/12/03/technology/google-researcher-timnit-gebru.html, *accessed* 12 December 2024.

[44] *Id.*

[45] IBM Data and AI Team. "Shedding Light on AI Bias with Real World Examples." Ibm.com, IBM, 16 Oct. 2023, www.ibm.com/think/topics/shedding-light-on-ai-bias-with-real-world-examples, *accessed* 12 December 2024; *see also*, McDade, Maria, and Anthony Testman. "Tackling Bias in AI." IBM.com, IBM, Inc., 29 Aug. 2024, www.ibm.com/think/insights/tackling-bias-in-ai, a*ccessed* 21 Jan. 2025.

[46] Ebert, Christof, and Ulrich Hemel. "Grow Your Artificial Intelligence Competence." Computer, vol. 57, no. 10, Oct. 2024, pp. 144–150, https://doi.org/10.1109/mc.2024.3436168. *accessed* 18 Oct. 2024.

[47] "AI Test, Evaluation, Validation and Verification (TEVV)." NIST, 5 July 2022, www.nist.gov/ai-test-evaluation-validation-and-verification-tevv, *accessed* 18 Oct. 2024.

[48] Dr. Frank Appiah. "What Is AI Governance? The Reasons Why It's so Important." apus.edu, American Military University (AMU), 5 July 2024, www.amu.apus.edu/area-of-study/information-technology/resources/what-is-ai-governance/, accessed 22 January 2025.

[49] Mäntymäki, Matti & Minkkinen, Matti & Birkstedt, Teemu. (2022). Defining organizational AI governance. AI and Ethics. 2. 1-7. 10.1007/s43681-022-00143-x, *accessed* 12 December 2024.

[50] Schuett, Jonas. (2023). Three lines of defense against risks from AI. AI & SOCIETY. 1-15. 10.1007/s00146-023-01811-0, *accessed* 18 October 2024.

[51] "Who Is Accountable for AI: The Role of Stakeholder Engagement in Responsible AI | Lumenova AI" Lumenova AI, 2024, www.lumenova.ai/blog/responsible-ai-accountability-stakeholder-engagement/, *accessed* 22 January 2025.

[52] Haan, K. (2024, June 15)."24 top AI statistics & trends in 2023 – Forbes advisor," (R. Watts, Ed.). Forbes. https://www.forbes.com/advisor/business/ai-statistics/,m *accessed* 21 January 2025.

[53] Vincent, James. "OpenAI Isn't Doing Enough to Make ChatGPT's Limitations Clear." The Verge, 30 May 2023, www.theverge.com/2023/5/30/23741996/openai-chatgpt-false-information-misinformation-responsibility, *accessed* 12 December 2024.