Mississippi Department of Employment Security

ARTIFICIAL INTELLIGENCE GOVERNANCE POLICY

PURPOSE: The purpose of this policy is to establish a set of standards, procedures, and guidelines which govern the requisition and implementation of Artificial Intelligence within the Mississippi Department of Employment Security (MDES technological environment. As AI continues to expand its presence within modern enterprise environments, MDES has begun to consider expansion of its offerings to include platforms and applications which leverage automation to improve scalability and decrease reliance upon human interaction with end users. MDES recognizes the necessity that any automation leveraging AI must be implemented responsibly and ethically, and that AI is a complement to its existing human-based services, as opposed to a replacement.

DEFINITIONS:

Artificial Intelligence (AI): Artificial Intelligence (AI) -has the meaning set forth in 15 U.S.C. §9401(3) (section 5002(3) of Pub. L. 116-283) that states "a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI 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.

Convolutional Neural Networks (CNN): Convolutional Neural Networks are a subset of Machine and Deep Learning that utilizes convolutional layers to analyze and classify images and recognize objects within them. CNNs are primarily focused on image recognition and identification.

Deep Learning (DL): Deep Learning is a subset of Machine Learning (ML) that utilizes multilayered neural networks to simulate complex decision-making including image recognition, natural language processing, and speech recognition. Deep learning focuses on prediction and classification.

Foundation Model (FM): A foundation model is a large Machine-Learning model that forms the foundation for many downstream tasks and applications. Large foundation models are pre-trained on massive amounts of data and require a lot of computing to build. This allows them to learn vast amounts of knowledge and patterns. The FM can then be fine-tuned for specific tasks like language translation, text summarization, image recognition, or domain-specific purposes.

Generative Artificial Intelligence (GenAI): Generative AI refers to AI that is capable of and used to produce new content, including, audio, code, images, text, and video, according to the data inputs and machine learning model on which it is trained. This type of AI can create the same content when prompted by a user. GenAI may include both Large Language Models (LLM) and other types of ML models to allow them to perform two functions: (1) to perform the tasks of a LLM and also (2) to analyze images and recordings, or create similar items in response to prompts from a user. GenAI requires training to focus on the creation of new content.

Large Language Model (LLM): A Large Language Model is a type of AI model specifically designed to process and generate human language. LLMs are trained on vast amounts of text data, enabling them to understand, generate, and translate text, as well as answer questions and perform other language-related tasks.

Machine Learning (ML): Machine Learning is a subset of AI that involves the development of algorithms and models that enable computers to learn from and make predictions or decisions based on data. Rather than being explicitly programmed for specific tasks, ML systems identify patterns within data and use those patterns to improve their performance over time.

Natural Language Processing (NLP): Natural Language Processing is a type of AI model specifically designed to understand and process human language, written and spoken. NLPs utilize structure rules to translate analyze, and manipulate human language to compute, extract, and perform tasks related to language.

Robotic Processing Automation BOT (RPA BOT): Robotic Processing Automation BOTS are software robots that perform specific tasks based on user-defined rules and workflows. The following characteristics distinguish RPA BOTs: they require human intervention to make decisions and to adapt to new information; they are incapable of “learning”, and they are not designed for autonomy.

OBJECTIVES: This policy is designed to ensure that the integration of AI into MDES assets, processes, and offerings is governed via an enforceable set of standards and guidelines which define responsible use, the safeguarding of data proprietary to the agency and its end users, and ethical standards relevant to any existing or future AI-based projects.

SCOPE: This policy shall be applied to all internal implementations of AI within MDES, and, by extension, any business partner or third-party vendor from which this agency procures technologies that employ AI or support for such products and services.

This policy shall not be applied to internal implementations of RPA BOTs. Essentially, RPA BOTs follow strict instructions provided by human programming. These AI applications do not analyze complex information, learn from data, or make autonomous decisions in dynamic situations without human intervention. RPA BOTs must be reviewed to determine if any special functions such as analysis, learning, or autonomous decision-making are being utilized, that would require the application of this policy.

The only authorized AI tools are LLMs, CNNs, DLs, LLMs, MLs, or, NLPs, or any other later-developed AI tools. These AI tools must have been developed in the United States, by entities authorized to do business in the United States, or be approved for use by the MDES Executive Committee after a description of the AI from the AI Officer.

ROLES AND RESPONSIBILITIES:

Formal designations shall be established and documented which define authorities within MDES responsible for the execution, review, and maintenance of this policy. This shall include the following:

1. **AI Officer** (or equivalent as defined by MDES): This person shall provide oversight of AI governance, ensuring all delegated roles are performing their responsibilities effectively. This shall include:

a. Planning with Agency Executives to develop strategic decisions related to AI governance, ethical alignment, and regulatory compliance.
b. Leading and coordinating the AI governance team, ensuring collaboration and alignment with organizational goals.
c. Serving as the authoritative voice in AI project approvals, policy updates, and key decisions. Preparing high-level reports for senior management and the Executive Director.

d. Overseeing the collection, preparation, and management of data used in the AI system, ensuring data quality, consistency, and readiness for AI processes.
e. Providing guidance and engaging with senior management to aid in the procurement of AI services and products.
f. Coordinating the revision of this AI policy if the Governor's executive order, state or federal law or regulations affect or nullify any provisions herein.

2. **AI Governance Team**: A designated group of staff which report to the AI Officer will be responsible for the day-to-day management and operational support of AI systems within the organization. The team's key responsibilities include:

a. Managing the setup, integration, and ongoing maintenance of the AI application.
b. Continuously monitoring the AI model's performance and troubleshooting any technical issues that arise.
c. Engaging with the AI Officer to conduct regular reviews of AI integration and implementation within the MDES environment.
d. Providing training to staff regarding this AI policy and staff interactions with each specific integrated AI application.
e. Developing a process to identify potential use cases to capture common problems and scenarios where AI can be used beneficially. Also, maintaining a use case inventory of these identified use cases.
f. Establishing and updating a methodology for the prioritization and review of the AI use cases to include, but not be limited to, factors such as potential cost savings, improved service delivery, enhanced customer experience, improved staff quality of life, and increased efficiencies.

The above listed AI Officer and AI Governance Team may be hired to exclusively perform this role or be existing staff who have been designated by the Executive Director or his designee to perform this role.

INTERNAL AI GOVERNANCE:

MDES recognizes that although AI integration into its current services will aid in expediting and automating its customer service experience, the sensitive nature of client data the agency collects, transmits, and stores must be considered and safeguarded at all times.

MDES has taken the steps necessary to identify data which falls within the purview of the relevant compliance standards including IRS 1075 and NIST 800-171. To ensure that current and future AI integration does not undermine technical and administrative controls as well as any executive order, state or federal law or regulations currently imposed or later issued, enacted, promulgated, or adopted on sensitive data within the MDES environment, the following standards must be adhered to:

1. **Data Protection**: Any integration of AI services and products provided by third-party vendors must not interact with sensitive data in any capacity:

    a. AI technologies managed or supported by an external party must not be provided any sensitive data unless the vendor provides specific safeguards designed for utmost security for data used in artificial intelligence applications. This shall include sharing of sensitive data with the vendor itself, or importing sensitive data into a language model supported, managed, or provided by a third-party vendor.
    b. Sensitive data shall be defined as staff and customer personal identifiable information (PII), Federal Tax Information (FTI), and/or any data governed by the legal standards in section 1(f) above.
    c. Integration of sensitive data into a LLM or other AI repository effectively renders the entirety of the model as sensitive through sensitivity inheritance. This must be avoided where possible and secured using the highest-level safeguards available to ensure the protection of the data.

2. **Monitoring**: All AI integration within the MDES environment and its services will be continuously monitored to ensure data accuracy and integrity. A required component of modern AI integration involves human oversight of results generated by machines.

a. Appropriate baselines must be established, documented, and regularly reviewed for all AI services. This shall include metrics such as data throughput, processing time, and input sanitization to include responses to malicious prompts.
b. A member of the AI Governance team shall be tasked with consistent monitoring of all AI applications, services, and products. This will include systematic "spot checking" at defined intervals to identify the following:
    1. Algorithm deviation
    2. Unintended responses derivative of a LLM (hallucination)
    3. Exposure or introduction of sensitive data
    4. Tracking and moderation of undesirable and/or harmful content.

3. **Reporting**: Any perceived or realized importation of sensitive data to a third-party managed AI service or product must be reported immediately:

    a. All staff must be aware of a requirement to alert the AI Officer in the event of a perceived or realized disclosure of sensitive data immediately.
    b. The AI Officer will work with appropriate staff to determine whether an incident has occurred.
    c. If it is determined that an incident has occurred, the AI Officer will coordinate with relevant staff to initiate the agency's AI Incident Response plan.

4. **Human Oversight**: Human oversight and intervention must be maintained at all times as AI solutions are requisitioned, installed, and employed within the MDES environment.

    a. AI systems should not be solely responsible for making critical decisions without human intervention. Important decisions, particularly those impacting sensitive data or significant agency operations, must involve human review and approval.
    b. Human oversight is necessary to validate and interpret AI-generated outcomes, ensuring they align with organizational goals and ethical standards.
    c. Integration of AI into MDES' current offerings is intended to complement or enhance human-designed and managed services. AI should be leveraged to support critical information system and

business processes which are reliant upon human interaction and intervention so long as the responsibility for a decision or final determination rests with an Agency staff person and not an AI application.

EXTERNAL AI GOVERNANCE:

MDES does not currently intend to develop AI, CNN, DL, LLM, ML, NLP platforms and applications internally. As such, the agency will rely upon a third-party vendor to procure and support the integration of AI functionality into its existing services. All external parties which MDES procures AI technology from are subject to the following:

1. **Disclosure of AI Tooling**: All vendors must acknowledge and disclose tooling necessary to the function of the requisitioned service to the extent that MDES understands how the tool handles its proprietary information.

    a. Any functions provided through a party external to the vendor must be disclosed prior to acquisition of the product or service.
    b. MDES understands the constraints of intellectual property (IP) considerations and all disclosures should balance necessary transparency and sensitive data issues with IP concerns of vendors.

2. **Isolation**: Any service that is managed or supported by an external entity to MDES and its network perimeter must ensure that any data exchanged between vendor and client is isolated physically or logically from other clients. Neither MDES data that is used to tune an FM utilized by the Agency or the tuned FM itself may be shared with model providers, or used to improve base models. FMs utilized by MDES should be private copies of the FM ensuring both physical and logical isolation of MDES data and tuning.

    This shall include transmission, processing, and storage of any data proprietary to MDES beyond its network perimeter.

3. **Data Sharing**: Any intent or need to share data between the vendor and an external party aside from MDES must be disclosed prior to acquisition.

    No data provided by MDES shall be transmitted or stored remotely by the vendor without expressed, documented permission from the Agency.

4. **Internal Control**: Vendor must make available to MDES a centralized platform with which to monitor the AI, LLM, or ML AI, CNN, DL, LLM, ML, NLP service whilst in use within the MDES environment.

    a. The platform must provide MDES the ability to control the AI environment, to include starting and stopping the service as needed.
    b. If the AI platform is considered an integral part of critical infrastructure or core business, then the vendor must provide MDES with a "kill switch" that will gracefully transfer control back to humans tasked with oversight.

5. **Regulatory Compliance**: In the event that a proposed AI project will interface directly or indirectly with any data which MDES manages, the vendor must ensure compliance with all relevant regulations and standards as imposed by the agency.

    This includes adherence to legal requirements specific to data protection and privacy, ensuring that the vendor's practices align with MDES' compliance obligations.

DATA QUALITY AND INTEGRITY:

1. **Accuracy**: All vendors must develop processes in coordination with MDES that helps to ensure that the data used to train its AI systems is accurate and reflective of the real-world scenarios it aims to model. Any detected inaccuracies must be corrected promptly.

2. **Relevance**: All vendors must develop processes in coordination with MDES to help ensure that data used for the training of AI models is current, and to set the required intervals to refresh data sets to maintain their relevance. Outdated data that could affect decision-making or predictions must be identified and updated regularly.

    All data used must be relevant to the specific AI application, ensuring that only necessary data is included to reduce noise and potential biases.

3. **Validation**: Regular validation of data must be conducted to identify and correct errors, anomalies, or inconsistencies. Vendors should utilize automated and manual validation techniques to ensure data accuracy.

4. **Auditing**: Periodic audits must be performed by vendors to assess the integrity of data throughout its lifecycle, from collection to processing and storage. Audit results should be shared with MDES to ensure transparency and accountability. The Vendor is encouraged to offer integrated and automatic audit functions.

RISK MITIGATION AND ASSESSMENT:

1. **Security Controls**: Vendors must implement robust security measures to protect sensitive data, including encryption, access controls, and regular security testing. These measures must align with MDES's internal security policies.

   Vendor must agree to employ security controls necessary to meet the expectations of the legal standards imposed on MDES for all data which it handles., be it technical, administrative, or physical.

2. **Data Handling**: Vendor must agree to and demonstrate the ability to employ data handling protocols commensurate with MDES' legal requirements. This includes the methods with which data is collected, stored, transmitted, and processed. These protocols must minimize the risk of unauthorized access or data breaches.

3. **Risk Assessments**: The integration of AI applications must be incorporated into MDES's existing Risk Assessment Plan. This includes identifying specific AI-related risks, such as model bias, data leakage, and system vulnerabilities, and developing strategies to mitigate these risks.

   Both MDES and vendors providing AI products or support must perform annual risk assessments of such integrations. These assessments should evaluate potential threats to data security, privacy, and ethical standards, as well as the effectiveness of existing controls.

4. **Training and Awareness**: To ensure effective risk management and mitigation, MDES will implement a comprehensive training program for all relevant staff and stakeholders.

    a. AI and Data Privacy Training shall be conducted annually for all senior management and staff which comprise AI Governance Team, including the AI Officer.

    b. Training course content will include an overview of AI technologies, data privacy principles, and specific legal standards that apply to the Agency's operations.

ETHICAL AI USE AND TRANSPARENCY:

MDES recognizes the need to ensure that any AI integration is designed to abide by certain ethical standards to avoid generating biased or otherwise harmful responses. As such, each vendor must agree to and comply with the following:

1. Vendors and MDES must ensure that any integrated AI applications respond to prompts fairly, without discriminating against individuals or groups based on characteristics such as race, gender, age, or religion.

2. The use of AI for activities that could cause physical, psychological, or social harm is strictly prohibited. This includes using AI for surveillance, profiling, or decision-making that negatively impacts individuals or communities without their informed consent. This requirement requires oversight to ensure the AI uses only legal methods of profiling and identification of information and any decision generated from said AI usage regarding a claimant or employer is made by an Agency staff person.

3. AI projects must undergo an ethical review process prior to implementation. This review will assess potential ethical risks and ensure that AI applications align with MDES's values and ethical standards.