C3 TECHNOLOGY Advisors

RESPONSIBLE ARTIFICIAL INTELLIGENCE

ORGANIZATION NAME: ____

EFFECTIVE DATE: ____

POLICY

It is the policy of this organization to DISCOVER, MEASURE, GOVERN and CONTROL the PRIVACY of DATA through the application of the UNIFIED CONTROLS, to SENSITIVE OBJECTS, defined as in-scope, based on RISK, according to the UNIFIED CYBER STRATEGY.

SCOPE

This organization minimally includes, as in-scope, any DATA or INFORMATION risk-rated as EXTREME or HIGH, henceforth referred to as SENSITIVE DATA per APPENDIX A: UNIFIED SPECIFICATIONS.

RISK

This organization measures risk based on the assessed impact to CONFIDENTIALITY, INTEGRITY and/or AVAILABILITY of ALL organizationally relevant systems and data using ISO:30001-2018 to determine EXTREME, HIGH, MODERATE, and LOW risk levels per the UNIFIED CYBER STRATEGY.

CYBER PRECEDENCE

This organization has established SERVICE LEVEL OBJECTIVES and MONITORING METRICS to prioritize, react, respond, and remediate risk per the UNIFIED CYBER STRATEGY.

SEPARATION OF DUTIES

This organization separates the duties of individuals, as necessary, to prevent malevolent activity, distribute risk and ensure compliance. The STAKEHOLDERS of SENSITIVE OBJECTS are documented in the RISK & ATTACK SURFACE MANAGEMENT - APPENDIX X: RISK-BASED INVENTORY.

CHANGE CONTROL

This organization controls changes to in-scope systems and data through APPENDIX C: CHANGE CONTROL PROTOCOL of the CYBER STRATEGY AND RISK MANAGEMENT POLICY adapted from the INFORMATION TECHNOLOGY INFRASTRUCTURE LIBRARY (ITILv3) and NIST-800.53 CONFIGURATION CHANGE CONTROL (CM-3).

COMPLIANCE MONITORING

This organization monitors the compliance of SENSITIVE OBJECTS, and reports relevant variances and trends to STAKEHOLDERS as defined in the RISK & ATTACK SURFACE MANAGEMENT - APPENDIX X: RISK-BASED INVENTORY. This organization retains compliance monitoring records per the UNIFIED CYBER STRATEGY.

CONTINUOUS IMPROVEMENT

This organization routinely reviews the practices and outcomes associated with this policy as defined in APPENDIX F: UNIFIED CYBER CONTINUOUS IMPROVEMENT MONITOR.

POLICY EXCEPTIONS

Requests for exceptions to this policy shall be reviewed and approved by the PROGRAM OFFICER.

POLICY REVIEW

This organization ANNUALLY reviews the RISK, SCOPE, UNIFIED CONTROLS, SENSITIVE OBJECTS and PROGRAM STANDARDS associated with this policy.

POLICY APPROVAL

This organization has involved the appropriate Stakeholders, Executive Leadership, Governing Body and supporting departments such as Human Resources and Risk Management to ensure education, communication and revision is complete according to the organization's formal or informal policy adoption practices prior to approval and signature by an AUTHORIZED INDIVIDUAL.

RISK ACCEPTANCE

This organization has considered the SCOPE, RISKS, UNIFIED CONTROLS, SENSITIVE OBJECTS and PROGRAM STANDARDS and is determined to accept the risk associated with out-of-scope controls, objects, and standards. At the time organizational risk tolerances change formally (memo, executive order, or statutory regulation) or informally (funding, staffing, or tooling) the organization will re-evaluate and update this policy accordingly.

POLICY ADOPTION

This organization has evaluated this policy, the impact on resources, and has determined to adopt this policy.

AUTHORIZED INDIVIDUAL

or

CORPORATE OFFICER

Title:____________________________________

Name:____________________________________

Signature:________________________________

Date:____________________________________

UNIFIED SPECIFICATIONS

SYSTEM NAME	EXCEPTION LOG	APPROVED BY	REVIEW DATE	GAP LOG

UNIFIED SPECIFICATIONS

FUNCTION	SPECIFICATION	REF
Policy:		
□ Adopt and Approve Policy	Review ANNUALLY	PT-1
□ Perform Gap Analysis	Complete ANNUALLY	PT-1
□ Educate Stakeholders	Complete ANNUALLY	PT-1
□ Implement tools to ensure compliance	ANNUALLY, complete: □ Ad-hoc review □ C3 Technology Advisors Design Review □ C3 Technology Advisors Continuous Improvement Tracer □ 3rd party audit	
□ Retain Compliance Monitoring Data	Policy compliance monitoring data will be maintained for: □ 1 year □ 3 years □ 5 years □ 7 years	
	DATA INTEGRITY & ETHICS	
DATA INTEGRITY & ETHICAL RISK MANAGEMENT		
□ Legal and Regulatory Requirements	LEGAL and regulatory requirements involving GENERATED or PREDICTED, SENSITIVE DATA are discovered, understood, managed, and documented; specific to the intended SCOPE OF PRACTICE and the OPERATIONAL IMPLEMENTATION.	

UNIFIED SPECIFICATIONS

□ Trustworthy AI Requirements	The characteristics of trustworthy AI are integrated and have been evaluated to be: 1. Fair and Impartial 2. Transparent and Explainable 3. Responsible and Accountable 4. Robust and Reliable 5. Private and Respectful of Privacy 6. Safe and Secure	
□ Establish Risk Tolerances	Risk TOLERANCES are DETERMINED and documented according to ISO 31000:2018.	
□ Define Metrics	METRICS for measurement of SENSITIVE DATASETS and AI-SYSTEMS are adopted and regularly updated according to TRUSTWORTHY CHARACTERISTICS. Including the use of third-party data or software as are risks of infringement of a third party’s intellectual property or other rights.	
□ Inventory & Discovery	SYSTEMS, PRACTICES and RESOURCES are in place to routinely INVENTORY SENSITIVE DATASETS and generative or predictive AI-SYSTEMS.	

UNIFIED SPECIFICATIONS

□ Initial Risk Management	The RISK MANAGEMENT process and its OUTCOMES are routinely PERFORMED and EVALUATED for in-scope SENSITIVE DATA and AI SYSTEMS.	
□ Map Controls	Risk controls, including TECHNOLOGIES, DATA, PROCESSES, TECHNIQUES and SOFTWARE are identified and documented.	
□ Track risks	APPROACHES, PERSONNEL, and SYSTEMS are in place to regularly IDENTIFY and TRACK existing, unanticipated, and emergent RISKS.	
□ Treat risks	TREATMENT of risks is PRIORITIZED based on impact, likelihood, and available resources. Responses to EXTREME and HIGH risks are developed, planned, EXECUTED, and documented.	
□ Third-party Risk Management	RISK MANAGEMENT practices are in place that address SENSITIVE DATA and AI-SYSTEM risks associated with THIRD-PARTY ENTITIES, including risks of infringement of a third-party's INTELLECTUAL PROPERTY or other rights.	

□ Third-party Risk Remediation	CONTINGENCY PROCESSES are in place to handle FAILURES or INCIDENTS in third-party SENSITIVE DATA or AI-SYSTEMS deemed to be high-risk.	
□ Secure Decommissioning	Secure DECOMMISSIONING and PHASING out AI SYSTEMS and SENSITIVE DATASETS follows standards defined in the DATA PRIVACY & GOVERNANCE policy.	
DATA INTEGRITY AND ETHICS GOVERNANCE		
□ Governance Accountability	Establish a DATA INTEGRITY AND ETHICS BODY according to the DATA INTEGRITY AND ETHICS STARTER KIT, chartered to QUARTERLY, review DATA INTEGRITY AND ETHICS ASSESSMENTS and direct DATA INTEGRITY AND ETHICS MANAGEMENT activities for SENSITIVE DATASETS AND AI-SYSTEMS.	2.1
□ Establish Goals	Relevant GOALS for the COLLECTION, GENERATION or PREDICTION of SENSITIVE DATASETS are understood and documented.	
□ Governance Training	PERSONNEL and PARTNERS shall receive SENSITIVE DATA and AI RISK MANAGEMENT TRAINING to enable them to perform their duties and responsibilities.	2.2

UNIFIED SPECIFICATIONS

☐ Executive Leadership	EXECUTIVE LEADERSHIP takes RESPONSIBILITY for decisions about RISKS associated with SENSITIVE DATA and AI SYSTEM development and/or deployment by reviewing and overseeing the activities of the DATA INTEGRITY AND ETHICS BODY.	2.3
☐ Control the quality of Integrity & Ethics Compliance	IMPLEMENT and OVERSEE practices appropriate to the use case and dataset to enable TESTING, IDENTIFICATION OF INCIDENTS, and INFORMATION SHARING.	4.3
☐ Review efficacy	Measurable IMPROVEMENTS or DECLINES based on consultations with relevant ACTORS and CUSTODIANS, and field data about TRUSTWORTHINESS CHARACTERISTICS are identified and documented.	
DATA INTEGRITY AND ETHICS WORKFORCE DEVELOPMENT		
☐ Workforce Design	DECISION-MAKING related to mapping, measuring, and managing SENSITIVE DATASETS and AI RISKS throughout the lifecycle is informed by a DIVERSE TEAM (e.g., diversity of demographics, disciplines, experience, expertise, and backgrounds).	

UNIFIED SPECIFICATIONS

☐ Workforce Development	PROCEDURES and DUTIES are in place to DEFINE roles and responsibilities for HUMAN-AI CONFIGURATIONS and oversight of AI SYSTEMS and SENSITIVE DATASETS.	
☐ Establish Actors and Custodians	Regularly INCORPORATE FEEDBACK from relevant actors and custodians into DESIGN, IMPLEMENTATION, SUPERVISION & GOVERNANCE, including: Internal: ● Design ● Development ● Deployment ● Operation ● Maintenance ● Testing ● Evaluation ● Verification ● Validation External: ● Human Factors ● Domain Expertise ● Impact Assessment ● Procurement ● Governance and Oversight ● Third parties ● End Users ● Affected individuals and Communities	
NARROW ARTIFICIAL INTELLIGENCE CAPABILITY DEVELOPMENT		
☐ Identify Value	The business value or context of business use has been clearly defined.	

UNIFIED SPECIFICATIONS

□ Examine Benefits	Potential BENEFITS of intended SENSITIVE DATASET or AI-SYSTEM FUNCTIONALITY and PERFORMANCE are examined and documented.	
□ Examine Total Cost of Ownership	Potential COSTS, including non-monetary costs, which result from expected or realized ERRORS or SYSTEM FUNCTIONALITY, INTEGRITY, and TRUSTWORTHINESS.	
□ Define impact	LIKELIHOOD and MAGNITUDE of each identified IMPACT based on expected use, past uses of AI systems in similar contexts, public incident reports, feedback from those external to the team that developed or deployed the AI system, or other data are IDENTIFIED and DOCUMENTED.	
□ Establish Requirements	SYSTEM REQUIREMENTS are elicited from and understood by relevant ACTORS & CUSTODIANS. Design decisions take SOCIO-TECHNICAL IMPLICATIONS into account to address SENSITIVE DATA & AI-SYSTEM risks.	

UNIFIED SPECIFICATIONS

□ Fitness and Basis	The specific INTENTS, USE CASES and SCOPE OF PRACTICE are evaluated for RISKS and potential impacts of SENSITIVE DATA and AI TECHNOLOGY which is EVALUATED and COMMUNICATED prior to and throughout the design, development, deployment, evaluation, and use.	
□ Define Scope	Targeted application SCOPE OF PRACTICE is SPECIFIED and DOCUMENTED based on the system’s CAPABILITY, ESTABLISHED CONTEXT, and CATEGORIZATION.	
□ Establish Limits	INFORMATION about the KNOWLEDGE LIMITS and how SYSTEM OUTPUT may be utilized and overseen by humans is DOCUMENTED. Documentation provides sufficient information to assist relevant actors & custodians.	
□ Verify Integrity	SCIENTIFIC INTEGRITY and TEVV CONSIDERATIONS are identified and documented, including those related to experimental DESIGN, DATA COLLECTION and SELECTION (e.g., availability, representativeness, suitability), SYSTEM TRUSTWORTHINESS, and CONSTRUCT VALIDATION.	

UNIFIED SPECIFICATIONS

□ Define technical standards	PERFORMANCE and TRUSTWORTHINESS – and relevant TECHNICAL STANDARDS and CERTIFICATIONS – are defined, assessed, and documented.	
□ Supervision Parties	Regularly INCORPORATE FEEDBACK from relevant actors, custodians and supervision parties into system DESIGN and IMPLEMENTATION.	
□ Evaluate Trustworthiness	The TESTING, EVALUATION, VERIFICATION and VALIDATION test sets, metrics, and details are documented including; 1. Evaluations involving human subjects meet applicable requirements (including human subject protection) and are representative of the relevant population. 2. Assurance criteria are measured qualitatively or quantitatively and demonstrated for conditions similar to deployment setting(s). 3. The functionality and behavior of the AI system and its components are monitored when in production. 4. The AI system to be deployed is demonstrated to be valid and reliable. 5. The AI system is evaluated regularly for safety risks.	

	6. The AI system to be deployed is demonstrated to be safe, its residual negative risk does not exceed the risk tolerance, and it can fail safely, particularly if made to operate beyond its knowledge limits. 7. Safety metrics reflect system reliability and robustness, real-time monitoring, and response times for AI system failures. 8. AI system security and resilience are evaluated and documented. 9. Risks associated with transparency and accountability are examined and documented. 10. The AI model is explained, validated, and documented, and AI system output is interpreted within its context to inform responsible use and governance. 11. Privacy risk of the AI system is examined and documented. 12. Fairness and bias are evaluated, and results are documented. 13. Environmental impact and sustainability of AI model training and management activities are assessed and documented.	

UNIFIED SPECIFICATIONS

□ Sustain Operations	MECHANISMS and RESOURCES are in place and applied to SUSTAIN THE VALUE of deployed AI systems including mechanisms to SUPERSEDE, DISENGAGE, or DEACTIVATE AI systems that demonstrate performance or outcomes inconsistent with intended use.	
□ Continuous Improvement	Post-deployment MONITORING PLANS ARE IMPLEMENTED, Including: 1. Mechanisms for capturing and evaluating input from users and other relevant AI actors, appeal and override, decommissioning, incident response, recovery, and change management. 2. Measurable activities for continual improvements are integrated into AI system updates and include regular engagement with interested parties, including relevant AI actors. 3. Incidents and errors are communicated to relevant AI actors, including affected communities. Processes for tracking, responding to, and recovering from incidents and errors are followed and documented.	