



## DOD PRINCIPLES OF AI ETHICS

### NSCAI Recommended Practices:

		Responsible	Equitable	Traceable	Reliable	Governable
<b>Core Values</b>	A1 - Employ technologies and operational policies for privacy, fairness, inclusion, human rights, and LOAC	X	X			X
	B1 - Consider and document value considerations based on how tradeoffs with accuracy are handled	X	X	X	X	
	B2 - Consider and document value considerations in systems that rely on representations of objective or utility functions	X	X	X	X	
	B3 - Conduct documentation, reviews, and set limits on disallowed outcomes	X	X	X	X	X
<b>Engineering</b>	1 - Concept of operations development, and design and requirements definition and analysis	X	X	X	X	X
	2 - Documentation of the AI lifecycle			X		
	3 - Infrastructure to support traceability, including auditability and forensics		X	X		
	4 - Security and robustness: addressing intentional and unintentional failures				X	X
	5 - Conduct red-teaming				X	
<b>System Performance</b>	A1 - Standards for metrics & reporting	X	X	X	X	
	A2 - Representativeness of data and model for the specific context at hand	X	X	X	X	
	A3 - Evaluating an AI system's performance relative to current benchmarks			X		X
	A4 - Evaluating aggregate performance of human-machine teams	X				
	A5 - Reliability and robustness	X		X	X	
	A6 - For systems of systems, testing machine-machine/multi-agent interaction	X			X	
	B1 - Specifying maintenance requirements	X	X	X	X	
	B2 - Continuously monitoring and evaluating AI system performance	X	X	X	X	X
	B3 - Iterative and sustained testing and validation	X			X	X
	B4 - Monitoring and mitigating emergent behavior	X			X	X
<b>Human-AI Interaction</b>	A1 - Define functions and responsibilities of human operators and assign them to specific individuals	X		X		
	A2 - Policies should define the tasks of humans across the AI lifecycle	X				
	A3 - Enable feedback and oversight to ensure that systems operate as they should	X			X	
	B1 - Human-AI design guidelines	X	X	X		X
	B2 - Algorithms and functions in support of interpretability and explanation	X		X		X
	B3 - Designs that provide cues to human operator(s) about the confidence a system has in its results or behaviors	X		X		X
	B4 - Policies for machine-human handoff	X		X		X
	B5 - Leveraging traceability to assist with system development and understanding	X		X	X	X
B6 - Training	X	X	X	X	X	
<b>Accountability/Governance</b>	1 - Identify responsible actors	X		X		X
	2 - Adopt technology to strengthen accountability processes and goals	X		X		X
	3 - Adopt policies to strengthen accountability	X		X		
	4 - External oversight support			X		



## NSCAI Recommended Practices:

## IC PRINCIPLES OF AI ETHICS

		Respect the Law; Act with Integrity	Transparent & Accountable	Objective & Equitable	Human-Centered Development & Use	Secure & Resilient	Informed by Science & Technology
<b>Core Values</b>	A1 - Employ technologies and operational policies for privacy, fairness, inclusion, human rights, and LOAC	X		X	X		X
	B1 - Consider and document value considerations based on how tradeoffs with accuracy are handled	X	X	X	X	X	X
	B2 - Consider and document value considerations in systems that rely on representations of objective or utility functions	X	X	X	X	X	X
	B3 - Conduct documentation, reviews, and set limits on disallowed outcomes	X	X	X	X	X	X
<b>Engineering</b>	1 - Concept of operations development, and design and requirements definition and analysis	X	X	X	X	X	X
	2 - Documentation of the AI lifecycle	X	X			X	X
	3 - Infrastructure to support traceability, including auditability and forensics		X	X		X	X
	4 - Security and robustness: addressing intentional and unintentional failures			X		X	X
	5 - Conduct red-teaming			X		X	X
<b>System Performance</b>	A1 - Standards for metrics & reporting	X	X	X		X	X
	A2 - Representativeness of data and model for the specific context at hand	X	X	X	X	X	X
	A3 - Evaluating an AI system's performance relative to current benchmarks		X		X		X
	A4 - Evaluating aggregate performance of human-machine teams		X		X		
	A5 - Reliability and robustness		X			X	
	A6 - For systems of systems, testing machine-machine/multi-agent interaction					X	X
	B1 - Specifying maintenance requirements		X			X	X
	B2 - Continuously monitoring and evaluating AI system performance	X	X	X		X	X
	B3 - Iterative and sustained testing and validation		X			X	X
	B4 - Monitoring and mitigating emergent behavior					X	
<b>Human-AI Interaction</b>	A1 - Define functions and responsibilities of human operators and assign them to specific individuals		X		X		X
	A2 - Policies should define the tasks of humans across the AI lifecycle		X		X		X
	A3 - Enable feedback and oversight to ensure that systems operate as they should	X	X		X		X
	B1 - Human-AI design guidelines		X		X		X
	B2 - Algorithms and functions in support of interpretability and explanation		X		X		X
	B3 - Designs that provide cues to human operator(s) about the confidence a system has in its results or behaviors		X		X		X
	B4 - Policies for machine-human handoff		X		X		X
	B5 - Leveraging traceability to assist with system development and understanding		X		X		X
	B6 - Training	X	X	X	X	X	X
<b>Accountability/ Governance</b>	1 - Identify responsible actors	X	X				X
	2 - Adopt technology to strengthen accountability processes and goals	X	X				X
	3 - Adopt policies to strengthen accountability	X	X				X
	4 - External oversight support	X	X				X