Transcript

Welcome to certified information securities training and certification program for the certified NIST artificial Intelligence Risk Management Framework, 1.0 architect. This is the third part of four of the introduction to the training and certification program. I'm Alan keel. Thanks for joining me. So you've heard me talk about the AI risk management life cycle. Well, we need to go ahead and get a basic understanding of that. So that way we can see as we look through the core with our four core functions, where those play out in the AI. Life cycle. We'll begin with looking at roles and responsibilities within that life cycle as well. Identifying and managing AI risk and potential impacts both positive and negative, requires a broad set of perspectives and actors across the entire AI life cycle, which is why we need to look at roles, responsibilities and how they're allocated to the a life cycle. From all spectrum of the organization. Ideally, AI actors will represent a diversity of experience, expertise and background. As well as comprising demographically and disciplinary diverse teams, the AI risk management framework is intended to be used by AI actors across the AI life cycle and its dimensions. So NIST provides this graphic to illustrate the AI lifecycle. The two inner circles show AI systems key dimensions and the outer circle shows AI lifecycle stages. Ideally, risk management efforts start with the plan and design function in the application context and are then performed throughout the entire AI system lifecycle. There are four key dimensions in an AI life cycle system. The AI dimensions are application context data and input the AI model and task and output AI actors involved in these dimensions who perform or manage the design, development, deployment. Evaluation and use of AI systems drive the AI risk management efforts. So let's look at some of the context for AI actors involved in AI design. AI design tasks are performed during the application contexts, and data and input phases of the AI lifecycle, AI design actors create the concept and objectives of AI systems. And are then responsible for the planning, design and data collection and processing task of the AI system so that the AI system is lawful and fit for purpose. Task and AI design include articulating and documenting the systems, concepts, and objectives, underlying assumptions, context and requirements gathering and cleaning data, and documenting the metadata and characteristics of the data set. AI actors in this category include data scientists. Domain experts, sociocultural analysts, experts in the field of diversity, equity and inclusion and accessibility. Members of impacted communities, human factors experts, governance experts, data engineers, data providers, system funders, product managers, third party entities, evaluators, and finally legal and privacy governance actors as well. So again, when you're trying to consider. Who needs to get an understanding of how to? Play a part. In AI risk management, there's your first set of individuals that need to understand AI risk management, as this course is teaching it. So now we've moved on from AI design to AI development. AI development tasks are performed in the AI model phase of the life cycle. AI development actors provide the initial infrastructure of AI systems and are responsible for model building and interpretation tasks which involve the creation. Selection, calibration training and or testing of models or algorithms. AI actors in this category include machine learning experts, data scientists, developers, third party entities, legal and privacy governance experts, as well as experts in the sociocultural and contextual factors associated with the deployment setting. Following the design and development phases, we now look at the AI deployment category AI deployment tasks are performed during the task and output phase of the life cycle. AI deployment actors are responsible for contextual decisions relating to how the AI system is used to assure deployment of the system into production. AI deployment tasks include palliating, the system, checking compatibility with legacy systems, ensuring regulatory compliance, managing organizational change, and evaluating the actual end user experience. AI actors in this category include system integrators. Software developers and end users, operators and practitioners evaluate. And domain experts with expertise in human factors, sociocultural analysis, and governance. And once we've deployed the AI system, we naturally need to now operate and monitor it. Operation and monitoring tasks are performed in the application. Contexts operate in monitor phase of the life. Cycle and these tasks were carried out by AI actors who are responsible for operating the AI system and working with others to regularly assess system output and impacts. AI actors in this category include system operators, domain experts, AI designers, users who interpret. Or incorporate the output of AI systems product developers, evaluators and auditors, compliance experts, organizational management and members of the research community at large. Despite all of the care and caution that we exercise during the design, development, deployment and operation of AI, we can't really be sure that the AI system is trustworthy if it hasn't been thoroughly tested first. So this is where. Test evaluation, verification and validation comes in. These are tasks that are performed throughout the AI life cycle, and they're carried out by AI actors who examine the AI system or its components, or detect and remediate problems. Ideally, AI actors carrying out verification and validation tasks are distinct from those. Who actually performed the tests and evaluation actions? Tev tasks can be incorporated into a phase as early as design, where tests are planned in accordance with the design requirement. You can learn more about test evaluation, verification and validation at the link provided on this page, where NIST provides further detailed. Guidance on this process itself. And then we also have roles devoted to human factors, human factors, tasks and activities are found throughout the dimensions of the AI life cycle. They include human centered design and practice methodologies promoting the active involvement of end users and other interested parties and relevant. AI actors incorporating context specific norms and values in system design, evaluating and adapting end user experiences, and broad integration of humans and human dynamics in all phases of the AI life cycle. Human factors professionals provide multidisciplinary skills and perspectives to understand context of use and form interdisciplinary and demographic diversity, engage in consultative processes, design and evaluate user experience, perform human centered evaluation and testing. And inform impact assessments as well. And in every phase of the AI lifecycle we have domain experts who have specific experience specialized to the context at hand. Domain experts tasks involve input from multidisciplinary practitioners or scholars who provide knowledge or expertise. In and about an industry sector, economic sector contacts or application area where an AI system is being used. AI actors who are domain experts can provide essential guidance for AI system design and development, and interpret outputs in support of work performed by Tev and AI impact assessment teams. Since risk is measured in impact. We also have to be concerned with roles that go to the AI impact assessment category, AI impact assessment tasks include assessing and evaluating requirements for AI system accountability, combating harmful bias, examining impacts of AI systems, Product Safety, liability and security. Among others, AI actors such as impact assessors and evaluators provide technical human factor, sociocultural, and legal expertise. And we have other roles within the AI lifecycle concerned with procurement procurement tasks are conducted by AI actors with financial, legal or policy management authority for acquisition of AI models, products or services from third party developers, vendors or contractors. And as we'll soon find out in the coming module, after the introduction, we'll begin by looking at AI, governance, AI governance and oversight. Tasks are assumed by AI actors, with management, fiduciary, legal authority, and responsibility for the organization in which an AI system is designed, developed. And Oregon deployed. Key AI actors responsible for AI governance include organizational management, senior leadership, even the board of directors. These actors are parties that are concerned with the impact and sustainability of the organization as a whole. And since AI can either facilitate that good end. Or potentially to destroy. Yet this again goes to show that even at the board level, we need a certain understanding of AI and risk to be able to manage the potential good or bad impact of AI to the organization's mission and objectives. So NIST also provides in the AI risk management framework playbook a really handy. The road map of AI development and adoption. It actually also includes roles and responsibilities that need to be assigned and performed at each sequential phase of that road map. So as we look at this, we can see in the plan and design phase. We will have certain activities where we articulate and document the system's concept and objectives, underlying assumptions and context in light of legal, regulatory requirements and ethical considerations. That will then need to be performed by. Representative actors that I've talked about in previous slides are system operators and users, domain experts, AI designers, and so on. From planning and design, we progressed to collecting and processing data. This is where we gather, validate and clean data and document the metadata and characteristics of the data. Set in light of the organization's AI objectives, legal and ethical considerations, these activities are performed by our data scientists, data engineers, data providers, domain experts, socio cultural analysts, human factors experts, as well as our. Tev experts? We then progress to the dimension of AI modeling, and you may remember that AI modeling actually has two life cycle stages building and use model as well as verify and validate. So within the build and use model we have activities that include. Creating or selecting algorithms training. Models and again with verify and validate. We also have verify, validate, calibrate and interpret model output and this is all performed by modelers, model engineers, data scientists, developers. Domain experts with consultation again a sociocultural experts and analysts who are familiar with the application contacts, as well as with the TEV experts. We then progress to the deploy and use life cycle stage. This is where we pilot the new AI system. We check compatibility with legacy systems, we verify regulatory compliance, manage organizational change and evaluate the user experience and this is performed by our systems. Integrators, developers, systems engineers, domain experts, procurement experts, third party suppliers, our C-Suite executives. Along with consultation from our human factors experts, sociocultural analysts, governance experts, and of course, RT EV experts as well. And as I explained, once we have deployed the system, we are now ready to operate and monitor that. System and this is where we operate the AI system and continuously assess its recommendations and impacts, both intended and unintended, in light of objectives, legal and regulatory requirements, as well as ethical considerations. This is all performed by our systems, operators and users and practitioners. AI designers impact assessors Tev experts, product managers, compliance experts, our governance experts, organizational management impacted individuals and communities as well as our evaluators. We now move to our final life cycle stage where we have to consider well the impacts of AI on our people and planet. This is where we assess the use or impacted by life cycle stage. So this is where we actually perform. Use system and technology. We monitor and assess impacts. We seek mitigation of negative impacts and we advocate for human rights. This is performed by end users, operators and practitioners impacted communities. The general public at large. The policymakers standards organizations, trade associations, advocacy groups, environmental groups, civil society organizations and researchers.