Transcript

Welcome to certified information securities training and certification program for certified NIST artificial intelligence risk Management Framework 1.0 architect. This is the second part of four for the introduction to this training and certification program. I'm Alan keel. Thanks for joining me. So now that we have an understanding of the intent of these risk management frameworks and how they are to be used to help organizations like ours to better, more effectively manage AI risk. Let's take a look at what we're up against when it comes to managing AI risk. We're going to look at setting the stage for AI risk management. This is where we start to look at the context, the business drivers and the business environment of AI risk. We need to recognize that AI could potentially lead to job displacement and significant economic and social disruption. That would be a risk. AI can also have. Unintended consequences, creating potential regulatory risks for privacy breach, and so on. AI systems can be biased and discriminatory since they learn from humans. This could also lead to ethical and legal issues as well. AI can also potentially increase cybersecurity risks, because AI can be used to make attacks smarter and more effective. So we also recognize, too, that AI itself could be compromised. And that could become a negative impact as well. So AI itself has vulnerabilities that need to be protected. And then finally, we recognize that AI powered autonomous weapons could pose a risk to human safety and security as well. As we continue to look at the business context of AI and how it will be adopted and used, we recognize it's going to change how we do business itself. In fact, AI can disrupt traditional industries as well as create entirely new ones. Businesses need to adapt and innovate in order to remain competitive in this age of artificial intelligence. Meaning that you can use AI to become more competitive or fail to use AI. And lose your position in the market. Either way, that is another business risk of artificial intelligence that is just inherent to our current society. So again, let's take a look at some of the potential concerns that AI introduces to our business. We have privacy regulations, such as the general Data Protection Regulation in Europe that requires us to protect the privacy and the personally identifiable information or. The I of EU citizens well towards that end, we have to recognize then that AI. Actually learns on its own, and it's researching from publicly available information, and this has the potential to potentially breach. Restrictions and requirements of privacy regulation like GDPR. So because data is usually processed on the provider servers, an AI provider can also be considered a responsible party along with the company that uses the AI system and this places responsibility on. That organization that they need to now understand the potential legal ramifications of of having assumed that responsibility. So that segues into legal compliance issues in the case of personal data processing, there is an obligation to obtain a declaration of consent from data subjects. Is AI doing that as it collects and aggregates information through its own learning? It's probably not stopping to get permission. We also have to be concerned then about other information disclosure requirements, because if personal data is involved, data subjects must be informed about the scope and intended outcome of using AI. And then this finally leaves us with data protection impact assessment. Before processing, companies have an obligation to really assess data and associated exposures and consequences as it

comes to using AI. In their business practice. So let's dive a little deeper into AI and privacy risk and business AI technologies can collect large amounts of data about individuals and organizations. It can then analyze that data to identify patterns and trends that actually may be protected as personally identifiable information. That means that the ability to identify patterns and trends may compromise privacy, which may in turn increase legal compliance risk along the way. AI can also be used to create targeted marketing campaigns that may be perceived as invasive. Another concern is how artificial intelligence actually increases the capabilities of potential malicious actors or hackers. Cyber attackers are now using AI to automate their attacks and to do a better job of evading detection, AI powered malware. Can adapt to new environments and defenses much more easily than when attacks were purely manually controlled. AI can be used to identify vulnerabilities and targets for attack. And it very. Effective and efficient and automated way AI can help cyber attackers then launch more sophisticated and more damaging attacks. So done well, risk assessment and risk treatment within risk management. Is actually not as intuitive as you might think. There's actually some well established clinical approaches to assessing and managing. And risk? Well, AI introduces a whole new level of ambiguity and complexity in assessing and managing risk. So we have challenges for managing AI risk and risk assessments. In understanding our risk tolerance and understanding how to prioritize our AI risk, and then finally how to integrate this AI risk management into our overall organizations, operational enterprise risk management. So I claim that AI introduces a new level of ambiguity and complexity into assessing and managing operational risk. How do I mean? Well, AI risk or failures that are not well defined or adequately understood are difficult to measure quantitatively or qualitatively. And the ability to appropriately measure AI risk. Does not automatically imply that an AI system necessarily poses a high or a low risk. We simply have a hard time determining what risk is there. Risk must be identified and analyzed to actually understand its overall level of risk. Typically, we can look to ISO 31,000 to provide a framework for identifying, analyzing, and evaluating risk that could then be applied to AI. Risk management as well. So we can actually look at categories of AI risk assessment challenges such as risk related to 3rd party software, hardware and data, third party data or systems can accelerate research and development and facilitate technology transition, which is a wonderful thing. However, they may also complicate risk measurement, since risk can emerge from third party data, software or the hardware itself. Risk metrics for methodologies used by the organization developing the AI system may not actually perfectly aligned to or with the risk metrics and methodologies used by the organization that is deploying or operating the AI system. Also, the organization developing the AI system may not actually be fully transparent about the risk metrics or methodologies that it use or failed to use for that matter. So we can see that we have a problem because the organization that develops the AI system isn't necessarily the organization that is using the AI system, and both of them have their own understanding of what tolerable risk is. Both of them have perhaps differing understandings of. Risk impacts to consider when they assess, measure and evaluate risk. And again, the dilemma is there is no current consensus on robust and verifiable measurement criteria or methods for AI risk and trustworthiness. Which means that our measurement approaches between the people who develop the AI and the people that use it. Are often inconsistent, oversimplified, gained, or just otherwise unreliable.

And yet another dilemma is the risk itself may change depending upon what part of the AI life cycle that you're looking at. AI risk and the development phase can be very different from the AI risk that is in the deployment phase or the usage phase. So measuring risk at an earlier stage in the AI life cycle may yield different results than measuring risk at a later stage, since some risk may be latent at a given point in time but then may increase as AI systems adapt and evolve. And we also have different AI actors across the a life cycle that could have different risk perspectives. For example, an AI developer who makes AI software available, such as with pre trained models. Can have a different risk perspective, risk tolerances than an AI actor who is responsible for deploying that pre trained model in a specific use case scenario. And such developers may not recognize that their particular use cases could entail risks which differ from those perceived by the initial developer. However, all involved AI actors share responsibilities for designing, developing and deploying trustworthy AI system that is fit for purpose. And yet another risk assessment challenge is that while measuring AI risk in a laboratory or controlled environment may yield important insights pre deployment, these measurements may differ from the risk that we actually see emerging in operational real world settings. Another challenge to AI risk assessment is known as inscrutability. Inscrutable AI refers to systems that are difficult to understand or interpret. Inscrutability can arise from opaque algorithms, lack of transparency, or inherent uncertainties in the AI. System itself. Interpreting and understanding these systems is essential for risk management and building trust in AI. Inscrutable AI systems can complicate AI risk measurement. Yet another risk assessment challenge is known as human baseline. Risk management of AI systems that are intended to augment or replace human activities, such as decision making, requires some form of baseline metrics for comparison, but this is difficult to systematize since AI systems carry out different tasks and perform tasks differently. Than humans do. So when we manage risk, we're not trying to eliminate all risk. That simply can't be done. Risk being uncertainty of outcome. There's always some uncertainty of outcome. So how far do we need to manage risk too? Well, typically the organization needs to establish risk tolerance levels for various impacts, whether it be reputational impact, financial impact, health and safety impact. And the dilemma is that how do we then incorporate the organizations perceive tolerances to risk and to our AI risk management. So again, risk tolerance refers to the organization or AI actors readiness to accept or tolerate or bear the risk. In order to achieve its objectives. And so risk tolerance is best determined in an organization's enterprise risk management framework, not necessarily at the AI risk management level. So we determine the risk tolerance that is applied to all flavors of risk at the enterprise level, but then we apply that understanding of risk tolerance to our AI risk and the AI risk management framework to then prioritize our risk accordingly.