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 first of four segments of the introduction to that course. I'm Alan keel. Thanks for joining me. The certified NIST AI, RMF 1.0 architect credential, certifies your ability to implement the formal structure, governance, and policy of a robust AI risk management framework conforming to internationally recognized and respected NIST best practices and standards. Upon completion of this training and certificate program, you'll be equipped with the knowledge and skills required to manage, monitor and improve a NIST AI risk management program. You'll expand your risk management competency as well as increase your credibility through gaining international recognition as well as improving your resume and helping to increase your earning potential. Becoming a certified NIST AI Risk Management Framework, 1.0 architect is fairly straightforward. You'll begin by becoming a member of the CIS body of certified professionals and completing your NIST AI risk management framework. Training this program, you'll follow that by taking your certification. Exam RM-102. And with your training and certification exams successfully completed, you will be eligible to apply for your certification again as a certified NIST AI Risk Management Framework 1.0 architect. So let's talk. About the training itself, the training begins with this introduction to AI risk management and related frameworks. It sets the context and the stage for understanding the benefits and the implementation of the AI risk management framework by NIST. We will then go on to begin our NIST AI risk Management Framework, 1.0 training with understanding the first core function governing. AI risk management. We'll follow that with core function #2 where we map AI risk. We'll then move on to core function #3, assessing and measuring AI risk, and then we will continue to the core function #4 where we learn to manage those AI risks. So the general approach to the learning objectives and content delivery of this course begins with a very comprehensive over 90 slides overview of the business context of AI, which is what this module is did. 2:00 to 2:00 we will cover the desired outcomes or opportunities for AI risk management as well as potential concerns considerations and AI system principles guiding the development and usage of AI in business. So the frameworks architecture lays out with an introduction this module which addresses again how organizations can frame the risks related to a. I and this introduction also describes the intended audience for AI risk management. In other words, the people who should be participating in AI risk management. AI risk and trustworthiness are then analyzed, outlining the characteristics of trustworthy AI systems. The core of the framework consists of four specific functions to help organizations address the risk of AI systems in practice. Then we also have certain profiles and use cases provided by nists to help us better understand how the core is implemented and audited in real life. So again, this all begins with governing moving on to mapping, then measuring and then managing AI risk management. So the frameworks architecture again can be broken down macro to micro by looking at the four functions. Those are the core of the AI risk management framework. However, within these functions each function has certain categories or subject matter areas. Of objectives that we need to accomplish. So in all, we have 19 categories spread across 4 functions. And within those 19 categories, we have a total of 76 subcategories, which are the desired objectives of the program that we actually need to set tasks to that we need to accomplish. So when we put all of this together, we find out that we have 460 recommended implementation actions to help fulfill those 76 subcategories across 19 subject matter categories of the four functions. So our pragmatic approach to teaching these four core functions begins by breaking down each category into individual lessons, explaining particular desired outcomes and objectives that we need to accomplish each subcategory. Lesson then. Contains, well, the desired outcome that we need to achieve also supported by NIST, recommended implementation actions and then finally finishes for each lesson with relevant documentation considerations. The primary reference sources for this course, or the authoritative documents used for the content include of course the NIST artificial intelligence risk Management Framework, 1.0, as well as the NIST AI RMF playbook. So let's take a look at the agenda for this introduction. We'll begin by learning how to leverage AI risk management frameworks from NIST as well as ISO. We'll also look at setting the stage for AI risk management. Managing AI risk throughout the AI lifecycle. Looking at the challenges and risks for AI trustworthiness. So let's get started. We'll learn to leverage risk management frameworks such as the NIST AI Risk Management Framework, 1.0, as well as the ISO standard for AI risk management 23894. So we begin by well, looking at AI risk itself, risk management can enable AI developers and users to better understand impacts and account for the inherent limitations and uncertainties in their models and systems, which in turn can improve overall system performance. And trustworthiness, and the likelihood that AI technologies will always be used in ways that are beneficial. While some AI risk and benefits are well known, it can be challenging to assess negative impacts and the degree of impact of those negative consequences. And since we are integrating artificial intelligence into our normal business operations, we can't really segregate managing the risk of AI from our normal business operations. So AI risk should not be considered in isolation from other business and operational risk. Different AI actors, which is the term we use for people who work with AI systems throughout their life cycle, have different responsibilities and awareness depending on their roles in that life cycle. We'll be looking at the life cycle soon. We also recognize that organizations developing an AI system will often not have information how the system will eventually be used. Since we often change how we'll be using the system as we go. So again, AI risk management should be considered a subset of overall enterprise risk management and it should be integrated and incorporated into broader enterprise risk management strategies and processes. Treating AI risks, along with other critical risks such as cyber security and privacy risk will yield a more integrated outcome as well as better improved organizational efficiencies. So again, risk management begins with establishing enterprise risk management. This is where an organization establishes consistent management of risk of all risk throughout the enterprise. This helps guide and harmonize niche types of risk management such As for artificial intelligence. Cyber security, privacy or perhaps even compliance risk. An example of a framework used to establish enterprise risk management includes ISO 31,000 for enterprise risk management. Then within that, we can specialize risk management for. Artificial intelligence. Now ISO actually has a framework for establishing an AI management system, and part of that management system is managing AI risks. So we can look to ISO 42,001 just released in 2020. Three, it's entitled information technology, artificial Intelligence management system, and this can be used to establish a formal system strategy and policy for leveraging effective and responsible use of AI throughout the enterprise. And within that, we can actually have managing the risk of AI and its utilization throughout the enterprise. And we have risk management frameworks to assist in this from both NIST with its AI risk management Framework, 1.0 as well as the ISO 23894 frame. Work that helps us again to better address risk of potential undesired outcomes of artificial intelligence and its adoption and used throughout the organization. So let's begin with getting a basic understanding of ISO 42,001 so we can better understand how AI risk management. Actually integrates into an overall integrated approach to managing. AI itself. So ISO 42,001 is a management system standard for organizations working with artificial intelligence. The standard has requirements for understanding and managing AI risk, and ISO has released a guidance standard to give us better assistance. With understanding how to actually manage risk within the overall context of an AI management. Program the AI risk management standard from ISO is 23894 entitled Information Technology Artificial Intelligence Guidance on Risk management and this is the ISO AI risk management framework, especially tailored for identifying. Analyzing and evaluating and controlling AI risk. Now one of the things that already I'm a bit hesitant on is the fact that, well, they've kind of narrowed this risk management standard to looking at information technology, AI risk management, but we use AI in areas. Other than information technology, we use it within our operations technology. We use it within our Internet of Things as well. So I'm a little concerned that they tend to take more of a biased, isolated approach towards IT. AI risk management now then, having said that, this standard also has requirements for data use and protection. We know that ISO 42,001 supports organizations in developing responsible and trustworthy AI systems. So ISO 42,001 is the overall ISO standard four and AI management system. Which includes requirements for managing AI risk. Now it has requirements for managing AI risk, but it doesn't really tell us how to do that well. It only has the requirements to get it done, so ISO has released a guidance document which is the ISO standard 23894. It's entitled Information technology, artificial Intelligence, guidance on risk management. And this is the standard that provides guidance on how organizations that utilize artificial intelligence to develop, produce, deploy or use products, systems and services can better manage risk, specifically related to AI. The guidance also aims to assist. Organizations to integrate risk management into their AI related activities and functions, and it also describes processes for the effective implementation and integration of AI risk management. Into overall enterprise risk management, the application of the guidance can be then customized to the unique business context of the organization. So when it comes to managing AI risk, it isn't necessarily an either or proposition as to whether we use ISO 23894 or if we choose to use NIST AI risk management framework. Matter of fact they complement each other. So let's go ahead and take a look at again the NIST. Artificial Intelligence risk management framework 1.0. As directed by the United States National Artificial Intelligence Initiative Act of 2020, the goal of AI RMF is to offer a resource to organizations designing, developing, deploying, or using AI systems to manage the many risk of AI. And to promote trustworthy and responsible development and use of AI systems. So the NIST AI risk management framework or the AI RMF. Is intended for voluntary use, meaning it doesn't have an audit specification that goes to it. It's intended for voluntary use and to improve the ability to incorporate trustworthiness considerations into the design, development, use and evaluation of AI products. Services and systems. So how did the AI RMF come about with this? Well, in collaboration with the private and public sectors, NIST developed a framework to better manage risk to individuals, organizations and society associated with artificial intelligence. The NIST AI risk management. Framework is intended again for voluntary use. And it is there to improve the ability to incorporate trustworthiness considerations into the design, development, use and evaluation of AI products, services and systems. So you can see the overlap with ISO 23894. So just released in January of 2023, the framework was developed through a consensus driven, open, transparent and collaborative. That included requests for information several draft versions for public comments, multiple workshops, as well as other opportunities to provide input. So as it was designed, the NIST AI Risk management framework was designed to be voluntary. Again, that is intended for flexible use by any organization, and the idea here is that rather than imposing requirements that organizations must meet specifically. It allows organizations to adapt the framework and use it as a general guidance document for how they can really build a fit for purpose risk management program for their own unique adoption and implementation of artificial intelligence. The AI Risk management framework is also intended to be rights preserving, that it will respect individual rights and freedoms, meaning that we're trying to use AI for the benefit of mankind rather than using it to actually create problems with breach of privacy. And breach of confidentiality, we want to make sure that the AI risk management framework is non sector specific. So that way it can be applicable across different industries as well as sectors. Again being fit for purpose to an organization's unique needs. And finally, the AI risk management framework should be used case agnostic, meaning that again it can be adapted to various applications of AI within different contexts. So the AI risk management framework is structured to support organizations and increasing the trustworthiness of AI systems. And fostering responsible practices overtime. And overtime means that it's a living document. This is expected to evolve with feedback from the AI community and changes in AI technologies that are developing very, very quickly. And as I said earlier, organizations shouldn't think that they have to go with one framework or another. They can actually blend 2 frameworks or more to get their right solution for managing their AI. List so the AI and risk management framework can be utilized along with related guidance and frameworks for managing AI system risk or even broader enterprise risk. Some risks related to AI systems are common across other types of software development and deployment. Examples of these kinds of overlapping risk include privacy concerns related to the use of underlying data to train AI systems, the energy and environmental implications associated with this. Resource heavy computing demands. As well as security concerns related to the confidentiality, integrity and availability of the system and its training and output data. As well as the general security of underlying software and hardware for AI systems. While we often use risk management processes to try to manage downside risk and negative impacts. The AI Risk management framework actually offers approaches to minimize the anticipated negative impacts of AI systems, but also to identify opportunities to maximize positive benefits and impacts. So effectively managing risk of potential undesired outcomes can help lead us to more trustworthy AI systems. And along the way, unleash potential benefits to people, including individuals, communities, and society, as well as our organizations and overall systems and ecosystems. So considering the overall mission of the AI risk management framework, we see that it's designed to equip organizations and individuals. As we've learned earlier, known as AI actors with approaches that increase the trustworthiness of AI systems, and to help foster the responsible. Design, development, deployment and use of AI systems over time. AI actors are defined by the Organization for Economic Cooperation and Development. The OECD, as those who play an active role in the AI system life cycle, including organizations and individuals that deploy and operate AI. So in other words. Any organization or individual who is a party involved in AI development, deployment, implementation use. Is called an AI actor. This is referenced in the OECD artificial intelligence. A society OECD eye library in Appendix A. In developing the AI risk management framework, it is intended that the framework be practical, adapting to the AI land. Escape as technologies continue to develop. So as I said earlier, it's intended to be an evolving and living program that evolves and adapts as AI changes over time. So the framework is intended to be utilized by organization in varying degrees and capacities. Again, fit for purpose. In the end, so society can benefit from AI while also being protected from any potential downside risk or impacts.