Microsoft's AI Security Warning: A Battle That Can't Be Won?

The legal industry has been all over the place regarding adopting artificial intelligence. Some attorneys have been on the AI bandwagon since the early days of ChatGPT, some have started to dip their toes in the water with Microsoft Copilot, and the rest have chosen to stay far away from anything AI-related or marketed.

The truth is that AI is here to stay. What we have been told about adopting it with caution and security in mind may not matter in the end.

Microsoft's internal testing of more than 100 of its AI generative products concluded that AI can never be made secure and suggested that securing AI systems will never be complete.

While this conclusion is not surprising, it does throw some cold water on the AI welcoming party. Here are several lessons to help with continuing to secure AI models.

To develop effective defenses, AI models must be understood thoroughly, including what the system can do and where it should be used. AI models behave differently depending on their design and application.

Defense in Depth Security Principals Apply Here Too

Just like with computer systems and information networks, defense-in-depth principles apply to AI models too. Layered security measures help reduce the inherent risks the models pose. The best security measures include controlling access through security permissions, restricting data the models have access to, auditing access, and requiring input validation. Policies and procedures are just as essential to prevent "Shadow AI," where users integrate their own selection of AI products at will into law firms without permission or thought. It is wise to have mechanisms in place to detect any "Shadow AI" attempts.

Having policies and security controls in place is vital, as is training. Training law firm staff on how to use implemented AI products is critical to maintaining the security of the application and law firm data. It is also important to reduce the risk of a user violating the firm's policies by implementing their preferred AI solution or prompting AI software with inappropriate data, such as client confidential data**.** End-user training is always the foundation of a good security program.

AI Risks are Unknown.

One problem cybersecurity professionals face is that the universe of risks posed by AI models is unknown. How AI models learn is unlike software vulnerabilities, which can be reproduced, verified, and patched. Yes, restrictions on input can be put in place to prevent user interface manipulation or validation of inputs to prevent hidden or unintended inputs. Still, the harms that AI can cause are more complex to quantify and reproduce, especially

since models can continue learning. Can you imagine what will happen to the volume of security patches for consumer products once AI gets involved – as if it weren't already bad enough with Microsoft?

Microsoft summarizes it well: If users can input private, confidential information, then it is safe to assume the models will output private, confidential information.

Because AI models may be able to learn, they will only continue to amplify existing risks and introduce new ones. That certainly provides job security for cybersecurity professionals moving forward. Still, it leaves users of AI products with difficult decisions about whether to accept the risk and how to become knowledgeable enough to decide whether a particular AI product application appears safe enough to use.

The decision becomes even more critical when you add the ethical requirements that attorneys must adhere to.

What happens when AI is embedded or added to a solution your firm has been using for years, whether you were informed or not? What information is shared with the product vendor and AI model, which will allow the model to continue to learn and grow? Microsoft is already incorporating AI into every piece of software that it can, including Microsoft 365.

Some examples include computers with AI prompts by default, such as Copilot for Windows or Apple Intelligence built into new iPhones, iPads, and Mac computers. This is another example of where reading the Terms of Services (TOS) is crucial, as well as understanding what settings for these AI prompts are turned on by default and how to access them to make any modifications.

If we know users (and trust us, we do), most will click and accept the TOS regardless of what it contains—that's human nature. Will lawyers voluntarily accept a possible ethics violation without knowing? The likelihood is very high.

By Michael C. Maschke, Sharon D. Nelson, Esq. and John W. Simek

Michael C. Maschke *is the President and Chief Executive Officer of Sensei Enterprises, Inc. Mr. Maschke is an EnCase Certified Examiner (EnCE), a Certified Computer Examiner (CCE #744), an AccessData Certified Examiner (ACE), a Certified Ethical Hacker (CEH), and a Certified Information Systems Security Professional (CISSP). He is a frequent speaker on IT, cybersecurity, and digital forensics, and he has co-authored 14 books published by the American Bar Association. He can be reached at mmaschke@senseient.com.*

Sharon D. Nelson *is the co-founder of and consultant to Sensei Enterprises, Inc. She is a past president of the Virginia State Bar, the Fairfax Bar Association, and the Fairfax Law Foundation. She is a co-author of 18 books published by the ABA. snelson@senseient.com*

John W. Simek *is the co-founder of and consultant to Sensei Enterprises, Inc. He holds multiple technical certifications and is a nationally known digital forensics expert. He is a co-author of 18 books published by the American Bar Association. jsimek@senseient.com*