Why should you care about agentic AI…

be the business

What is Agentic AI?

As artificial intelligence (AI) technologies evolve, one increasingly popular dimension has become the creation of AI agents. An AI agent is a piece of AI software designed to perform a series of tasks, either independently of human interactions, or as a response to an instruction from a human user.

Agentic AI Uses and Prompts

The potential applications of agentic AI seem boundless, but early use cases revolve around workflow automation and automated ordering and payment processes. In many cases, one or more instructions (“prompts”) are issued to the AI agents by a user to kick start each process.

Two scenarios illustrate this.

Scenario 1 – Workflow automation

An SMB implements a series of AI agents, working in tandem, to automatically reach out to existing customers via email, offering them special offers on selected services and products.

Customers who express interest are contacted by an AI voice agent for further information, and to confirm their desire to accept the offer, then a third AI payment system automatically bills the customer, taking payment card details or setting up invoicing, with no human involvement from the seller’s side.

Soon, the customer side of this process might involve an AI agent too; agents will chat to other agents and neither the human buyer nor the seller will ever interact directly.

Scenario 2 – Personal Assistants

Individuals will install AI personal assistants on their devices and give them access to their email, calendar, and payment information. Either in response to a user prompt, such as, “Order my grocery shopping tomorrow,” or on being triggered by a calendar appointment, the AI agent will search online, identify options, and ask the user to approve them, or act independently to complete the transactions.

What this means...

Prompt Injection Attacks

One category of attack that has already been demonstrated by experts is known as Prompt Injection. This method doesn't require attackers to directly access an SMB, as they would in a traditional data theft or ransomware attack. Instead, the attacker remotely triggers the victim's AI agent into executing malicious code hidden in a seemingly innocent prompt.

Agentic AI systems struggle to identify malicious instructions hidden within prompts. Examples include tricking AI agents into making payments or sharing sensitive data.

Addressing this category of risk is a major challenge. As Georg Zoeller, a leading AI security thinker stated recently, "To sanitize the inputs, you need to be able to either detect and remove undesirable content or transform the content completely into a secure structure. For most use cases, especially Chatbots, that's quite impossible."

Conclusion

Like any complex new technology, many of the risks resulting from the use of agentic AI are likely to be compounded by unrealistic expectations, poor systems design, poorly managed implementations, or some combination of these and other factors.

As with databases, experience suggests that giving AI systems access to payment information and personal data can increase the cyber-attack surface, and criminal hackers are likely to see this new technology wave as a bonanza.

It's therefore essential that SMBs use trusted suppliers of agentic AI tools, and that they carefully risk assess all planned rollouts of these systems and processes.

A Primer for Small and Medium Businesses

Written by Mark Johnson, Global Director of Training, The Risk Management Group.
In partnership with Be the Business.

July 2025