Assessing the Impact of Microsoft's Generative AI Copilots on Enterprise Application Strategy

Published 26 September 2023 - ID G00795333 - 20 min read

By Analyst(s): Jason Wong, Adam Preset, Matt Cain

Initiatives: Digital Workplace Applications; Artificial Intelligence; Enterprise Applications Evaluation and Selection; Enterprise Applications Strategy; Software Engineering Technologies

> Microsoft's Copilot products impact not only digital workplace technologies, but also multiple enterprise application areas. Digital workplace application leaders must be the focal point to evaluate, coordinate and prioritize all Copilot investments as part of a broader generative AI strategy.

Overview

Impacts

- Microsoft Copilots are built on a complex interdependence of new and existing Microsoft technologies that, if not evaluated properly, can lead to risks such as poor compliance and data governance.
- Trying to deploy multiple Copilots and other generative AI tools without coordination leads to employee confusion and hinders their adoption of the technologies.
- Rapid, mass rollout of Copilots across the enterprise — without appropriate readiness — will lead to poor enablement of employees and long-term value realization.

Recommendations

- Establish new generative AI skills and policies by evaluating Microsoft Copilot as a new technology stack rather than merely a productivity tool.
- Maximize adoption and reduce features overlap by coordinating with business unit leaders on use of Copilots and other generative AI tools from enterprise applications.
- Improve digital employee experience by prioritizing deployments to high-digital-dexterity employee segments and personas to maximize value of Copilots.

Strategic Planning Assumption

By 2027, 25% of high maturity digital workplaces will have leaders with a background in AI, data and analytics, versus infrastructure and operations.

Introduction

Microsoft has announced numerous generative AI products and features across its portfolio that may offer compelling value for enhancing everyday AI benefits. Microsoft Copilot is the brand umbrella, but the products impact enterprise application strategy across multiple domains. Most organizations are not fully prepared for the uncertainties that deploying Microsoft's breadth of generative AI capabilities at scale bring. If not implemented and managed properly, the impact will be more adverse (in terms of costs and corporate risk) than beneficial (in terms of tactical, personal and productivity gains) to digital employee experience (DEX).

Digital workplace application leaders are at the forefront of assessing the critical impact areas, starting with Microsoft 365 (M365) Copilot and extending to other applications' generative AI tools. These leaders can use this research to evaluate, coordinate and prioritize Microsoft's Copilot products to mitigate risk while focusing on high-value DEX and digital workplace outcomes (see Table 1).

Table 1: Impacts and Recommendations

(Enlarged table in Appendix)

Impacts	Recommendations
Microsoft Copilots are built on a complex interdependence of new and existing Microsoft technologies that, if not evaluated properly, can lead to risks such as poor compliance and data governance.	■ Establish new generative AI skills and policies by evaluating Microsoft Copilot as a new technology stack rather than merely a productivity tool. ■ Establish a M365 product team with direct oversight of governance of generative AI services that interact with the Copilot stack. ■ Review and communicate to stakeholders key Microsoft online service terms and data protection and privacy commitments, all of which apply to M365 Copilot. ■ Reinforce the need for information governance and access controls in M365 with stakeholders to ensure users don't overshare information that could be exposed through the Copilot stack.
Trying to deploy multiple Copilots and other generative AI tools without coordination leads to employee confusion and hinders their adoption of the technologies.	■ Maximize adoption and reduce features overlap by coordinating with business unit leaders on use of Copilots and other generative AI tools from enterprise applications. ■ Lead a coalition with your stakeholders to make your initial Copilot investments immediately valuable, and pave the way for an impactful and successful long-term integration of multiple generative AI technologies. ■ Plan for a multivendor generative AI portfolio that includes Microsoft alongside other vendors, each likely with different approaches.
Rapid, mass rollout of Copilots across the enterprise — without appropriate readiness — will lead to poor enablement of employees and long-term value realization.	■ Improve DEX by prioritizing deployments to high-digital-dexterity employee segments and personas to maximize value of Copilots. ■ Pave the way for broader rollouts by initiating an influencer network for generative AI and connect with communities of practices to help bridge activities with your internal M365 product team. ■ Identify and prioritize high-value Copilot use cases by competitive impact, business value, urgency, cost and risk. ■ Mitigate security and compliance risks by working with cybersecurity teams to promote cyber judgment practices for generative AI.

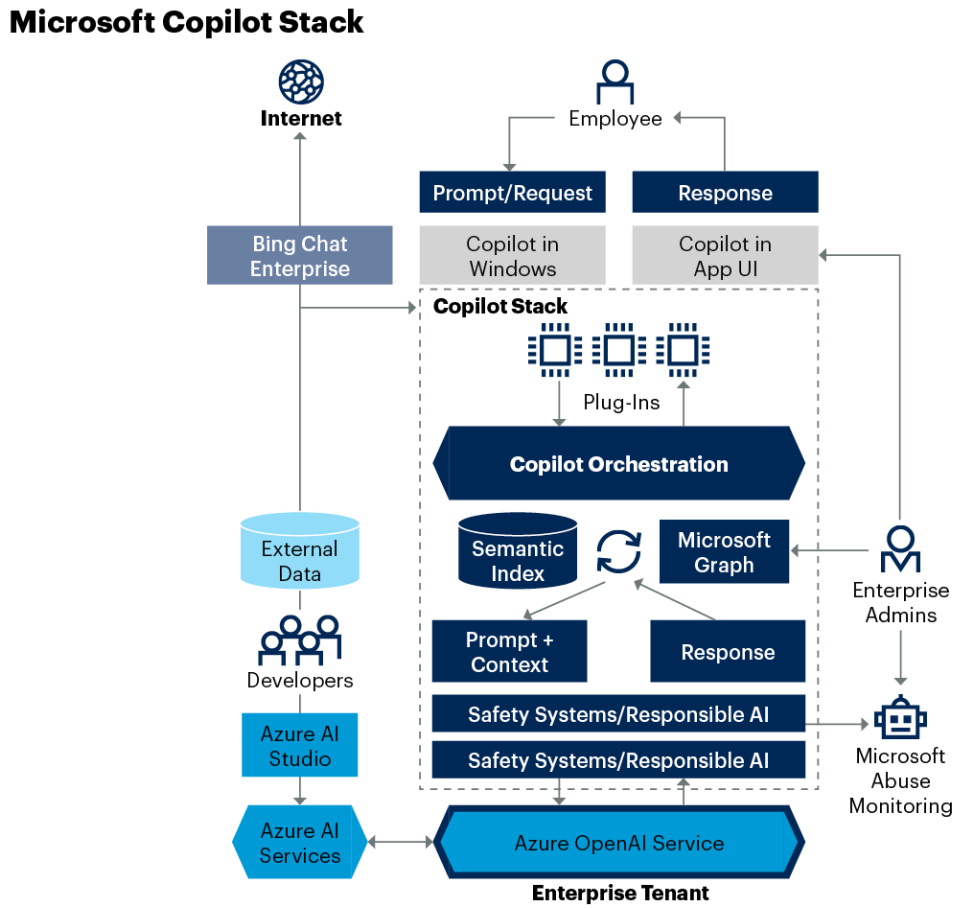
Source: Gartner (September 2023)

Impacts and Recommendations

Microsoft Copilots Have Complex Interdependencies

Copilot is not a single product or even single generative AI model. To evaluate Microsoft's Copilot offerings, you need to understand and explain the Copilot stack to executive stakeholders and your IT and business partners. The Copilot stack is a still evolving set of technologies that Microsoft has built as a layer over the OpenAI foundation models that it hosts as an Azure OpenAI Service. Gartner has analyzed publicly available information from Microsoft to create this high-level representation of the Copilot stack (see Figure 1).

Figure 1: Microsoft Copilot Stack



Source: Gartner
795333_C

Gartner

Functionally, when an employee uses M365 Copilot, such as in Outlook, Teams or Word, the Copilot function uses the prompt (a free-form text entry or prebuilt request) to initiate a prompt process using the Copilot stack. This process “grounds” the request and the response on the most relevant information accessible to the user from enterprise data sources. The approach is critical for providing organizational context and reducing “hallucinations” (the tendency of large language models [LLMs] to return false responses), although hallucinations still do occur. Grounding is part of the retrieval augmented generation (RAG) approach to find and filter the most relevant information that then gets passed to the base model in Azure OpenAI for processing. Azure OpenAI Service uses multiple models, including GPT-4, GPT-3.5-Turbo and DALL-E depending on the workload, for optimum performance and accuracy.

The Copilot stack enables Microsoft to combine and orchestrate fine-tuned models used by its various Copilots that then interoperate with the foundation models. The Copilot stack includes:

- **Plug-Ins:** These tools augment the capabilities of Copilots by interacting with APIs from other applications and services to retrieve information, incorporate company and other business data, and perform new types of computations. Plug-ins come in three types: ChatGPT plug-ins, Teams message extensions and Microsoft Power Platform connectors. The latter two types have backward compatibility for existing extensions and connectors. If developers want to create a plug-in for M365 Copilot from an existing API that follows the OpenAPI specification, they can use Teams Toolkit for Visual Studio to build it (see Quick Answer: How Will the Generative AI Plug-In Market Evolve?).

- **Copilot Orchestrator:** Every M365 tenant has a Copilot orchestrator instance that uses Microsoft Search to collate information. The orchestrator instance ensures data doesn't leave an organization's compliance boundary. It also acts as an AI orchestration layer that allows Microsoft to combine its own fine-tuned AI models and plug-ins to create connected Copilot experiences.

- **Microsoft Graph:** A fundamental technology in M365, it includes information about the relationships, permissions and policies between users, activities and your organization's data. It works in conjunction with the Semantic Index for Copilot, and it orchestrates information retrieval steps using Search. The Microsoft Graph API brings additional context from authorized data sources into the prompt. The prompt is the primary input into the model, which includes text plus files or content surfaced through the Copilot orchestrator, such as emails, chats, documents, calendars and meetings.

- **Semantic Index for Copilot:** Based on Microsoft Search technology, Semantic Index for Copilot uses multiple LLMs that sit on top of Microsoft Graph to interpret user queries and produce sophisticated, meaningful and multilingual responses that help users to be more productive. It allows users to search quickly through billions of vectors (mathematical representations of features or attributes) to connect with the most relevant and actionable information in the organization.

- **Safety Systems/Responsible AI:** To reduce the risk of harmful use of generative AI, Microsoft includes both content filtering and abuse monitoring features in the AI prompt pipeline. Azure OpenAI securely stores all prompts and responses for up to thirty days. Customers may request approval to modify abuse monitoring or turn it off. If approved, Microsoft does not store any prompts or completions associated with the Azure subscription that has abuse monitoring configured off. Once approved to turn off abuse monitoring, admins can verify the status in two ways in their Azure subscription: (1) Using the Azure portal (2) Azure CLI or any management API (see Microsoft Azure OpenAI vs. OpenAI: Comparing GenAI Trust, Risk and Security). [1]

Microsoft says customer data (including prompts, responses and data accessed through the Microsoft Graph) isn't used to train the foundation models that Copilot uses. Microsoft also announced that its Copilot Copyright Commitment extends Microsoft's existing IP indemnification coverage to copyright claims relating to the use of AI-powered Copilots, including the output they generate — specifically for generally available, paid versions of Microsoft commercial Copilot services. [2] These services include GitHub Copilot and Bing Chat Enterprise.

The Copilot stack combined with Azure AI Services enables the injection of RAG patterns and other prompt engineering techniques, such as Low-Rank Adaptation (LoRA), and model orchestration frameworks, like LangChain, to provide more enterprise-specific context. Developers can apply additional fine-tuning technologies, including Azure Machine Learning, Azure Cognitive Search and the Azure AI Studio tool (see Prompt Engineering With Enterprise Information for LLMs and GenAI).

Now is the time to establish a product team for M365, if your organization has not already done so. The complexity of managing the Copilot stack, especially with plug-ins and fine-tuned models, can be daunting without dedicated roles and new skills. The product owner leading this team must upskill to understand the architectural impact of the Copilot stack and plug-in integrations. Ensure there are sufficient staffing and training team members for new skills in data science, machine learning and prompt engineering needs (see Establish a Product Team to Maximize Value From Microsoft 365).

Recommendations

- Establish new generative AI skills and policies by evaluating Microsoft Copilot as a new technology stack rather than merely a productivity tool.

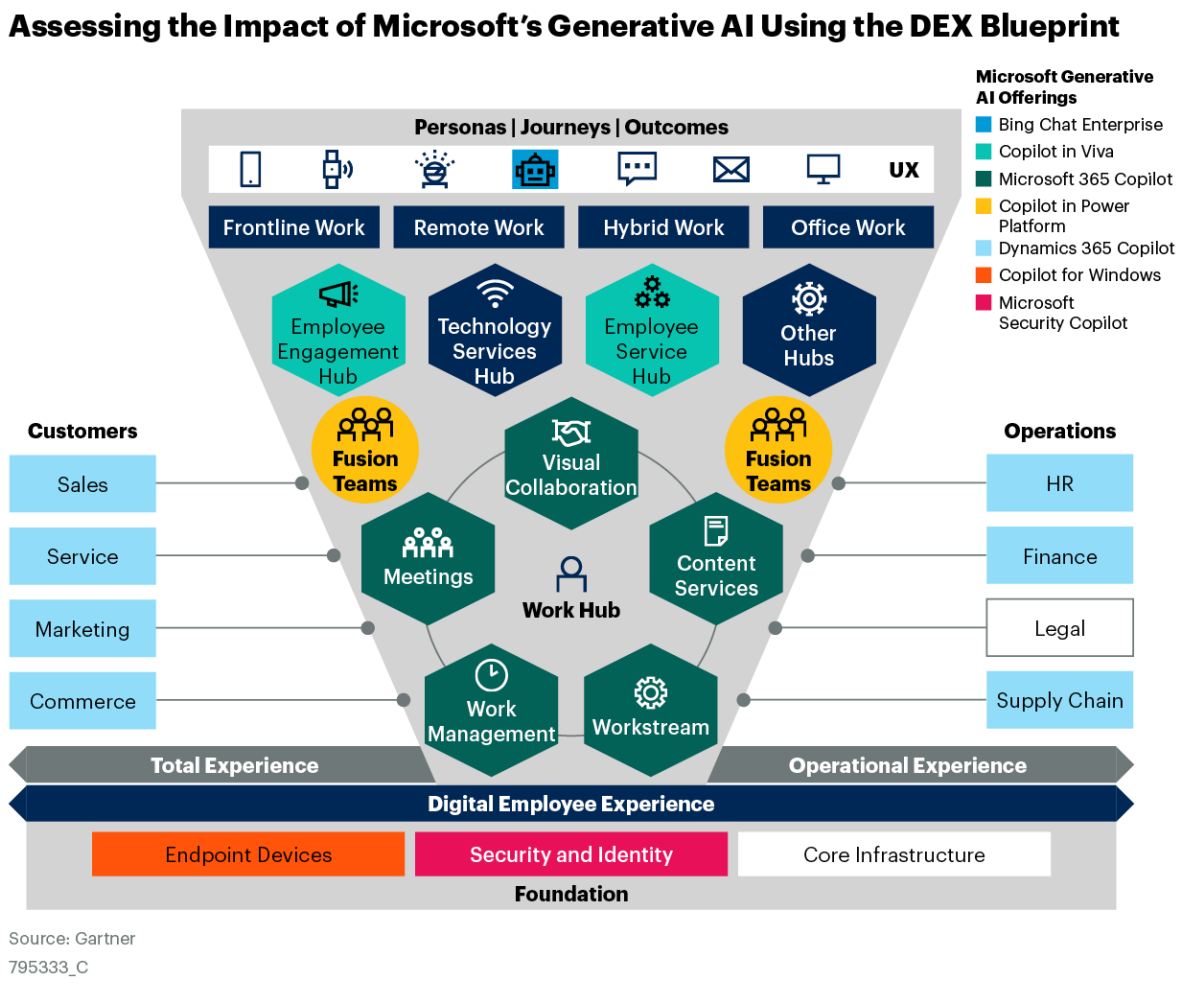
- Establish a M365 product team with direct oversight of governance of generative AI services that interact with the Copilot stack.
- Review and communicate to stakeholders key Microsoft online service terms and data protection and privacy commitments, all of which apply to M365 Copilot.
- Reinforce the need for information governance and access controls in M365 with stakeholders to ensure users don't overshare information that could be exposed through the Copilot stack.

Employee Confusion and Poor Adoption Stem From Poor Coordination

Generative AI capabilities will come from many enterprise application providers, such as Adobe, Salesforce, and ServiceNow, and new technology vendors. In addition to Copilot, there will be many digital workplace vendors, such as Box, Monday and Mural, that will have their own generative AI features, with some also offering integration with M365 Copilot via plug-ins. Vendors will push for their own embedded generative AI tools to be the entry point for users. Employees will be confused if they have too many options and too much to keep track of, leading to poor user adoption. Digital workplace application leaders must take charge of coordination across lines of business to implement a holistic approach to apply generative AI for better DEX outcomes.

The selection of generative AI tools must be coordinated and focus on serving the needs of specific employee personas and their journeys. A coordinated approach must center on improving DEX to drive sufficient value to justify the investment. The impact of Microsoft's generative AI offerings is not limited to the core work hub technologies of M365 — it affects the majority of technologies driving DEX as well (see Figure 2). Digital workplace application leaders can use Gartner's DEX blueprint to plan and communicate a shared vision toward using generative AI from Microsoft and other vendors (see Use Gartner's DEX Blueprint to Mature Your Digital Workplace Strategy).

Figure 2: Assessing the Impact of Microsoft Generative AI Using the DEX Blueprint



Source: Gartner
795333_C

Gartner

There are seven major product areas where Microsoft has generative AI capabilities impacting digital workplace and enterprise business applications:

- **Bing Chat Enterprise:** A generative-AI-powered internet search assistant with enterprise guardrails included in M365 E3, E5, Business Standard and Business Premium licenses. Alternatives include OpenAI's ChatGPT Business, You.com's YouPro and Anthropic's Claude.
- **Copilot in Viva:** Generative AI features embedded into Viva Glint, Viva Goals, Viva Engage, Viva Learning and Viva Topics for managing the employee experience. There are not really like-for-like alternatives, especially if your organization uses the Viva Suite in an integrated way.

- **Microsoft 365 Copilot:** Generative AI features offered as an add-on to M365 E3 and E5 licenses. Other stand-alone generative AI tools can be used instead, such as Jasper.ai, Grammarly or Otter.ai, depending on the use case. The main challenge will be integration of M365 data for use in third-party tools.
- **Copilot in Power Platform:** A mix of embedded and add-on generative AI capabilities to Power Apps, Power Automate, Power BI, Power Pages and Power Virtual Assistant. Low-code and no-code tools with generative AI for app and chatbot development, workflow and process auction, and data and analytics would be alternatives.
- **Dynamics 365 Copilot:** Embedded generative AI features in CRM and ERP applications at the enterprise licensing tier. There are stand-alone options for non-Dynamics customers, such as Sales Copilot. Alternatives would be to use native generative AI tools from the business application provider or use general purpose tools.
- **Copilot for Windows:** Generative AI assistant based on Bing Chat for Enterprise built into Windows 11 for assistance within the operating system. There is no alternative for OS-integrated generative AI.
- **Security Copilot:** Generative AI capabilities that support security operations center use cases. The product provides integrations across Microsoft Sentinel, Microsoft Defender for Endpoint and Microsoft Intune to improve threat identification, hunting, remediation and management.

Note that M365 Copilot will be generally available on 1 November 2023, but many Copilot features are in Preview release through the end of 2023. It's important to review terms of service for products and features in Preview, and their timeline for general availability production release, to determine fit for your organization's needs. Digital workplace leaders also need to know the alternatives to various Copilots and work with business stakeholders to determine the most appropriate technologies to maximize DEX outcomes.

Recommendations

- Maximize adoption and reduce features overlap by coordinating with business unit leaders on use of Copilots and other generative AI tools from enterprise applications.
- Lead a coalition with your stakeholders to make your initial Copilot investments immediately valuable, and pave the way for an impactful and successful long-term integration of multiple generative AI technologies.

- Plan for a multivendor generative AI portfolio that includes Microsoft alongside other vendors, each likely with different approaches.

Rapid, Mass Rollout Could Lead to Poor Enablement and Value Realization

Once you have evaluated the Copilot stack's technology requirements and coordinated use of Copilots with other generative AI technologies, digital workplace application leaders need to help prioritize the rollouts to employees in a controlled way. A "big bang" rollout will cause confusion among employees, leading to a surge of support issues. To ensure ROI is achieved, plan and execute a meticulous rollout strategy that includes a series of communications, multichannel training and support, and a holistic change management strategy with buy-in from executives and business leaders. Early preview customers for M365 Copilot indicated that selecting the appropriate employees for the pilot was a major challenge, as well as how to measure value. To set priorities with stakeholders that increase chances of successful adoption and value realization, digital workplace application leaders should prioritize digital dexterity, value creation and cyber judgment.

Prioritize Digital Dexterity

Results from the 2022 Gartner Digital Worker Survey show that 89% of respondents consider "improvement in digital technology skills" as highly important for career advancement. [3] In the same survey, 95% of respondents cite "improvement in digital technology skills" as being critical for work effectiveness. [3] Applying generative AI techniques should be at the top of the list of new digital skills to introduce into the workforce. Digital workplace application leaders must work with HR and business leaders to identify employees with high digital dexterity and ambition for using generative AI (see Create an Enablement Continuum to Advance Digital Skills Outside of IT).

High potential digital citizens — employees who use technology regularly to execute work — should be the first segment of employees to be provided access to and trained on use of Copilot tools. Such training may include integration with other role-based and specialized generative AI technologies, such as those in CRM, human capital management (HCM) and ERP. Identify these digital citizens within each business unit by asking who already has success using generative AI for personal use. Enlist them to apply generative AI for quick-win use cases for individual productivity gains, as well as for team-based outcomes, such as for meetings and collaboration workflow. Early anecdotal data suggests that curiosity and willingness to share Copilot experiences are the two leading traits correlated with an employee's ability to gain value from Copilot tools (see Bridge Technology Talent Gaps With "Untapped Technologists").

Eventually, these high-digital-dexterity employees can be offered “gigs” — short-term opportunities to apply their newly learned and applied skills outside of their business unit. The M365 product team can marshal these employees to form an influencer network to pave the way for broader rollouts by cultivating digital side hustles. This is a tactic to allow employees to develop technical expertise at their own pace and through diverse peer and project interactions by amplifying their generative AI skills (see How to Build and Manage a Digital Workplace Influencers Network).

Those employees that master Copilot prompting engineering skills and want to contribute more to the creation of new capabilities using generative AI could become business technologists. They remain in their business role but more actively participate in delivery teams with developers, engineers and data scientists. They form fusion teams that combine technical and business subject matter expertise to attack business problems. They can participate in communities of practice with employees who are already citizen developers and citizen data scientists. Business technologists — such as those building robotic process automation (RPA) bots, digital forms and workflows, and analytical models — would also be good candidates for early adopter programs for generative AI technologies (see Quick Answer: How Can Digital Workplace Leaders Support Business Technologists?).

Prioritize Value Creation

Quick personal productivity wins are expected from Copilot, but these gains may deliver diminishing returns over time. Generative AI capabilities must be focused on improving employee or customer journeys and their business outcomes to provide competitive differentiation. Value realization for generative AI can be quantified as financial benefits from improvements, such as task automation and greater work quantity and quality, with clear cost and revenue alignment. However, digital workplace application leaders must expand value by capturing improvements in employee experience, service levels, employee empowerment and upskilling, and customer engagement and value enhancement (see Assess the Value and Cost of Generative AI With New Investment Criteria).

Digital workplace application leaders must work with business leaders to identify and prioritize high-value Copilot use cases by competitive impact, business value, urgency, cost and risk. Prioritize use cases by collectively calculating incremental upfront and ongoing costs versus the business outcomes generated over time. Then build a portfolio of quick wins as well as differentiating and transformational generative AI use cases from the influencer network and communities of practice. Concurrently, optimize benefits realization by working with HR, finance and corporate strategy to create a change management strategy. This strategy should include an impact assessment, stakeholder communication, skills development plan, career path mapping and emotional support resources to navigate the change and redeployment of resources impacted by productivity gains.

Prioritize Cyber Judgment

Given the application and data interdependencies of generative AI, digital workplace application leaders will need to work even more closely with the IT security team or CISO organization counterparts to mitigate risks and ensure policy and regulatory compliance. Cybersecurity leaders recognize that their teams cannot scale to meet these needs directly. Digital workplace application leaders must work with these colleagues to ensure that appropriate tooling is in place for central monitoring and governance. Importantly, all leaders need to help their employees develop cyber judgment — the ability to make cyber-risk-informed decisions autonomously. For example, good generative AI cyber judgment would be for employees to trust but verify answers by examining the source data from prompt responses.

Cyber judgment is becoming a key part of cybersecurity strategy to empower employees. Employees with high cyber judgment are 2.2 times more likely to help their enterprise achieve digitization speed and obtain value from digital initiatives. [4] They are also 2 times more likely to avoid introducing additional risk. [4] Together, digital workplace and security leaders must prioritize these actions to build cyber judgment in those using generative AI (see Infographic: Building Cyber Judgment to Improve Risk Decision Making).

Recommendations

- Improve DEX by prioritizing deployments to high-digital-dexterity employee segments and personas to maximize value.
- Pave the way for broader rollouts by initiating an influencer network for generative AI and connect with communities of practices to help bridge activities with your internal Microsoft 365 product team.

- Identify and prioritize high-value Copilot use cases by competitive impact, business value, urgency, cost and risk.
- Mitigate security and compliance risks by working with cybersecurity teams to promote cyber judgment practices for generative AI.

Acronym Key and Glossary Terms

Fine-tuned model	A model focused on a specific context or category of information, such as a topic, industry or problem set.
Foundational model	A baseline model used for a solution set, typically pretrained on large amounts of data using self-supervised learning. Applications or other models are used on top of foundational models — or in fine-tuned contextualized versions.
Grounding	The ability of generative applications to map the factual information contained in a generative output or completion. It links generative applications to available factual sources — for example, documents or knowledge bases — as a direct citation, or it searches for new links.
Prompt engineering	The craft of designing and optimizing user requests to an LLM or LLM-based chatbot to get the most effective result, often achieved through significant experimentation.
Plug-in	A software component or module that extends the functionality of an LLM system into a wide range of areas, including travel reservations, e-commerce, web browsing and mathematical calculations.

Evidence

Data, Privacy, and Security for Microsoft 365 Copilot, Microsoft.

[1] Data, Privacy, and Security for Azure OpenAI Service, Microsoft.

[2] Microsoft Announces New Copilot Copyright Commitment for Customers, Microsoft.

[3] **2022 Gartner Digital Worker Survey:** This survey sought to understand workers' technological and workplace experience and sentiments. The research was conducted online from September through November 2022 among 4,861 respondents from the U.S. (n = 1,564), China (n = 1,167), the U.K. (n = 1,072) and India (n = 1,058). Participants were screened for full-time employment in organizations with 100 or more employees and were required to use digital technology for work purposes. Ages ranged from 18 through 74 years old, with quotas and weighting applied for age, gender, region and income, so that results are representative of working country populations. We defined "digital technology" as including any combination of technological devices (such as laptops, smartphones and tablets), applications and web services that people use for communication, information or productivity.Disclaimer: The results of this survey do not represent global findings or the market as a whole, but reflect the sentiments of the respondents and companies surveyed.

[4] **2022 Gartner Drivers of Secure Behavior Survey:** This survey was conducted via an online platform from May through June 2022 among 1,310 employees across functions, levels, industries and geographies. The survey examined the extent to which employees behave securely in their day-to-day work, root causes of insecure behavior, and the types of support and training that they received from their organizations to drive desirable secure behaviors. We used descriptive statistics and regression analysis to determine the key factors that drive or impede employees' secure behaviors and develop cyber judgment.

Recommended by the Authors

Some documents may not be available as part of your current Gartner subscription.

Quick Answer: How Will AI in Microsoft 365 Copilot Impact the Workplace?

Quick Answer: Evaluating Microsoft 365 Copilot Pricing & Bing Chat Enterprise

Quick Answer: How to Make Microsoft 365 Copilot Enterprise-Ready From a Security and Risk Perspective

Quick Answer: How Should Organizations Prepare for the Addition of Generative AI to the Microsoft Stack?

Quick Answer: What Should Application Leaders Do Now About GenAI?

Glossary of Terms for Generative AI and Large Language Models

Table 1: Impacts and Recommendations

Impacts	Recommendations
Microsoft Copilots are built on a complex interdependence of new and existing Microsoft technologies that, if not evaluated properly, can lead to risks such as poor compliance and data governance.	■ Establish new generative AI skills and policies by evaluating Microsoft Copilot as a new technology stack rather than merely a productivity tool. ■ Establish a M365 product team with direct oversight of governance of generative AI services that interact with the Copilot stack. ■ Review and communicate to stakeholders key Microsoft online service terms and data protection and privacy commitments, all of which apply to M365 Copilot. ■ Reinforce the need for information governance and access controls in M365 with stakeholders to ensure users don't overshare information that could be exposed through the Copilot stack.
Trying to deploy multiple Copilots and other generative AI tools without coordination leads to employee confusion and hinders their adoption of the technologies.	■ Maximize adoption and reduce features overlap by coordinating with business unit leaders on use of Copilots and other generative AI tools from enterprise applications. ■ Lead a coalition with your stakeholders to make your initial Copilot investments immediately valuable, and pave the way for an impactful and successful long-term integration of multiple generative AI technologies. ■ Plan for a multivendor generative AI portfolio that includes Microsoft alongside other vendors, each likely with different approaches.

Rapid, mass rollout of Copilots across the enterprise — without appropriate readiness — will lead to poor enablement of employees and long-term value realization.	■ Improve DEX by prioritizing deployments to high-digital-dexterity employee segments and personas to maximize value of Copilots. ■ Pave the way for broader rollouts by initiating an influencer network for generative AI and connect with communities of practices to help bridge activities with your internal M365 product team. ■ Identify and prioritize high-value Copilot use cases by competitive impact, business value, urgency, cost and risk. ■ Mitigate security and compliance risks by working with cybersecurity teams to promote cyber judgment practices for generative AI.

Source: Gartner (September 2023)