

Navigating the Copilot Ecosystem

Information technology (IT), as an organizational function, has the primary purpose of enabling business counterparts to implement solutions that improve productivity at the task and business process levels. For the past decade, IT leaders have been under incredible pressure to deliver business value while also being charged with continuing to drive down costs. For every new wave of technological innovation, IT leaders must navigate the fine line between embracing the hype and delivering tangible ROI.

Generative AI (GenAI) has forced an almost “gold rush” mentality within the IT industry, with consultants and independent software developers (ISVs) alike working hard to bring forward the next wave of innovation. Microsoft has made significant investments in GenAI through its Azure AI Studio service, which enables organizations to safely and securely develop their own GenAI applications, as well as its Copilot brand of GenAI-as-a-Service offerings, which are being built-in to its entire portfolio of applications. Satya Nadella, Chairman and Chief Executive Officer (CEO) of Microsoft, has been quoted as saying:

We are the Copilot company. We believe in a future where there will be a Copilot for everyone and everything you do. Microsoft Copilot is that one experience that runs across all our surfaces, understanding your context on the Web, on your device. And when you're at work, bringing the right skills to you when you need them. Just like, say today you boot up an operating

system to access applications or a browser to navigate to a Web site, you can invoke a Copilot to do all these activities, and more—to shop, to call, to analyze, to learn, to create. We want the compiler to be everywhere you are.

`redmondmag.com/articles/2023/11/16/
nadella-ignite-2023-keynote.aspx`

This chapter will begin with a quick primer on GenAI and Microsoft's role in maturing this technology over the past few years. We will then step through Microsoft's Copilot brand of products to better understand its strategy of transforming how people work by introducing artificial intelligence (AI) into their workflow. Finally, we will end with a high-level overview of Copilot Studio and how it will enable citizen developers and IT professionals to drive even greater business value by combining a low-code application development platform with GenAI.

What Is GenAI?

If you are unfamiliar with the term, GenAI is a type of AI that focuses on creating content by analyzing and learning patterns from large datasets. Examples of content GenAI can create include text, images, code, and audio. Additionally, GenAI can analyze existing content and provide you with feedback based on questions you ask it. For example, you could write an email and then ask it for suggestions on how to rewrite it for tone, clarity, or brevity. It will then analyze both the content you provided and its dataset to provide you with recommendations in natural language, meaning text.

GenAI is on the same trajectory as other large disruptive technologies, such as the graphical user interface (GUI), the Internet, and the iPhone. Unlike other technologies, such as robotics process automation (RPA), monolithic enterprise resource planning platforms, or event-driven architectures, what makes GenAI appealing is how it blends creativity with computation using the most powerful interface that exists: language.

GenAI blends creativity with computational power, enabling people to draft compelling narratives, design complex visuals, write code, and even create business strategies in a fraction of the time it once took. The technology is intuitive and adaptable, making it accessible to a wide range of users—from seasoned professionals to those without technical expertise. Its ability to personalize interactions, learn from context, and continuously improve makes GenAI a powerful tool for enhancing productivity, boosting innovation, and driving meaningful engagement across industries.

How Does GenAI Work?

At a high level, GenAI works through a combination of machine learning and then through the development of neural networks, which are a type of AI modeled after the human brain. Machine learning is when you feed in large amounts of data into a computer application, and it begins to create patterns to organize the data. Neural networks are an architecture within AI that include a series of interconnected nodes organized into various layers. These nodes, often referred to as *neurons*, are computational units that process information by performing mathematical operations on inputs, applying a weight (to emphasize importance), and passing the result through an activation function to determine the output. Like humans, GenAI models are trained with enormous amounts of data. While humans are trained over decades, GenAI models are trained over months with large datasets and computing infrastructure.

Like the human brain, as information flows through these interconnected nodes within the neural networks, it is transformed—but by mathematical functions. What is often viewed as being “GenAI magic” is a combination of being trained on a very large amount of data and the organization of this data in a way that can identify patterns and structures. At the end of the day, GenAI does an amazing job of predicting your desired output to the question you have asked it because it has been trained on a tremendous amount of data.

GenAI Key Terms and Definitions

While this book isn’t meant to be a deep dive into GenAI, there are some key terms that are helpful to understand:

- **GenAI models:** Models are designed to create new data from machine learning patterns. The output could be text, images, music, etc.
- **Large language models (LLMs):** An LLM is a type of GenAI model that has been specifically trained on human language and can be used to create new content or summarize existing content.
- **Training data:** Data is the key to GenAI, as it is what feeds the GenAI models. Training data typically consists of large datasets, which include text, images, audio, etc. For LLMs specifically, the training data is human language in the form of text.
- **Pattern recognition:** This is where the model analyzes all the data you have trained it on and begins to identify patterns and structures that will be leveraged for content creation or content summarization.
- **Content creation:** This is where the model applies what it has learned and attempts to generate new content that mimics the style and structure of the training data it was provided when it responds to a user prompt.

The Risk of Bias in GenAI

When people talk about being careful about bias with GenAI, it is because the models are only as good as the data they are trained on. To provide a practical example, if you were building a model on baseball statistics and only provided statistics of the Boston Red Sox defeating the New York Yankees, the model would be biased toward the Boston Red Sox being the more dominant team. If you asked it who will win an upcoming game between the two teams, it would more than likely propose the Boston Red Sox based on its data. However, when you include all the matchups, the model might be biased toward the New York Yankees, since historically they have won more of the games between the two teams.

When you apply this same concept to people, the consequences can be even more drastic. For example, it has been proven that some ethnicities are statistically more prone to certain diseases. Therefore, it is important to have a wide set of training data for AI applications that are meant for health care use cases. If you train data only on a particular ethnicity, you may miss out on some of the potential nuances. Combined with a potential over-reliance on AI, this could lead to a situation where a clinician might miss a diagnosis even though they have access to this incredibly powerful AI application. The consequences of being over-reliant on technology that is wrong can have dire consequences, thus the need to plan for a comprehensive responsible AI strategy to help minimize these risks.

The technology industry and many data scientists are especially careful to point out the potential risks of bias that can happen with AI. Therefore, the need to obtain or create diverse datasets reflective of the various types of people, ethnicities, cultures, etc. is important to ensure that AI is for everyone, not just a subset of people. This risk is even more problematic given that the user experience of interacting with GenAI is very conversational in nature. In addition, the confidence with which LLMs provide responses is rather convincing, even when those answers are incorrect answers. The popularization of engaging with LLMs through conversation was catapulted into the spotlight through OpenAI's ChatGPT.

OpenAI Brings GenAI to the World

While OpenAI did not invent GenAI, it was able to bring it to the mainstream with its launch of ChatGPT in November 2022. ChatGPT was able to amass a user base of over 100 million users in less than a year, far quicker than any other technology to date. ChatGPT rapidly surpassed the adoption of commercial AI assistants such as Amazon's Alexa, Google's Echo, and Microsoft's Cortana by

providing what seems like infinite expertise. ChatGPT's user experience, both then and now, is simple, with a basic chat-based user interface that is forgiving of typos and grammatical errors.

ChatGPT Gains Excitement

People began asking ChatGPT everything from how to make certain food dishes to recommendations for fuel-efficient vehicles to providing programming code to build responsive websites. You can even ask ChatGPT philosophical questions, such as *"What is the meaning of life in three sentences,"* and it will provide an answer, such as *"The meaning of life often revolves around finding purpose, creating connections, and embracing the journey. It's about making an impact through love, growth, and kindness while learning from experiences. Ultimately, life's meaning is unique to each person, shaped by their values, passions, and aspirations."* The magic of this interaction between human and machine really lies in the simplicity of being able to have a conversation with technology and having it understand the intricate nuances of language and culture.

As people continued to interact with ChatGPT, businesses started to think through how they might be able to incorporate this powerful new capability into their products and services. The buzz was all around identifying GenAI use cases that would help to unlock new revenue streams; however, there also began to be a growing concern about what OpenAI was doing with both the prompts that people were sending to its popular ChatGPT service and its responses.

ChatGPT Data Leaks

As businesses started to leverage this new capability, they found themselves in a situation where people inadvertently exposed their organization's sensitive data to a capability that feeds on knowledge. One of the early instances of this risk becoming a reality was when Samsung workers inadvertently leaked company secrets by using ChatGPT to assist with fixing source code.¹

Additionally, they found instances where an employee used ChatGPT to summarize meeting minutes into a presentation. While today, the idea of having a meeting platform like Microsoft Teams or Zoom capturing meeting minutes and then summarizing them into action items, bullets, etc. might be considered common, this was incredibly impactful at the time. It combined the two things that most business professionals dislike: attending meetings and producing meeting summaries. This was perhaps one of the earliest examples of applying GenAI technology to streamline a business process, and it was a very relatable use case. As a reminder, this was still the early days of GenAI, when everyone was collectively learning how this technology worked in almost real time.

¹ techradar.com/news/samsung-workers-leaked-company-secrets-by-using-chatgpt.

This story became national news and set off a flurry of activity around how to enable a secure ChatGPT service that wouldn't expose company secrets to OpenAI. An additional challenge during the early years was OpenAI itself. It originally started as a nonprofit research organization when it organized in 2015, but then became a for-profit company structured as a "capped profit" organization. In many ways, with the almost overnight success of ChatGPT, OpenAI was thrust into the spotlight without having time to thoughtfully plan and scale to meet the needs of large enterprise customers. Just as we were all collectively learning about GenAI, OpenAI was learning about what it takes to become a company. Sam Altman, the Co-founder and CEO of OpenAI, has been very forthcoming of the potential societal impact of AI, which is reflective in the company's mission and incorporation. He also recognizes that the valuation of OpenAI is not in the models of today but in the continued research that will inevitably unlock potential we can barely even imagine in the future.

OpenAI's goal is to create general-purpose artificial intelligence that benefits all humanity.

Sam Altman, Co-founder and CEO of OpenAI

community.openai.com/t/the-possibilities-of-ai-interview-with-sam-altman-openai/733526

Microsoft's Strategic Investment in OpenAI

Microsoft has been an early investor in OpenAI, and its initial 1 billion US dollar investment in OpenAI may be one of the most strategic moves made under Satya Nadella's tenure as CEO. After over 13 years of lagging behind Amazon Web Services (AWS) in the pursuit of providing cloud services, Microsoft's investment in OpenAI helped it leapfrog over the other hyperscalers. This gave Microsoft early access to OpenAI's models and helped it to architect its own service to help securely deliver enterprise-ready GenAI services well before its competition.

NOTE The term *hyperscalers*, which has been around since the 1990s, simply refers to large-scale datacenters that are able to deliver massive amounts of computing resources typically delivered through cloud computing.

In early January 2023, Microsoft announced Azure OpenAI as a new service that would allow you to provision instances of OpenAI's LLMs, including GPT-3, ChatGPT, and Codex models, all within your own Azure subscription. This was pivotal in helping organizations looking to experiment early with GenAI build their own ChatGPT-like solution within an existing cloud Azure subscription, either in an existing or new "landing zone."

A *landing zone* includes best practice configurations for security, networking, access management, and provisioning within your cloud environment. It allows organizations to isolate cloud resources, manage multiple environments (e.g., development, testing, and production), and scale as needed. This architecture provides the same level of controls and separation—or greater—than is available in a typical on-premises datacenter. The benefit is that your landing zone is completely virtualized and can be provisioned and deprovisioned using Infrastructure as Code (IaC) templates instead of managing physical hardware.

Azure OpenAI was a key enabler in helping organizations avoid any murkiness around how OpenAI handled prompt data and whether they were training their LLMs with this data. Thus, it ushered in an era of CompanyGPT experiments to placate the C-Suite that IT teams were evaluating this new capability while also working to minimize any business risk associated with exposing proprietary business data to one of the fastest-adopted consumer-facing IT services in history.

As businesses continued to experiment, they found that while having access to an LLM trained on billions of data parameters was incredible, the real potential value comes from combining this new capability with their own business data. However, one of the challenges of developing an LLM trained on business data was the sheer compute, power, and technical knowledge required. Hyperscalers such as Microsoft have the advantage of near-infinite computing resources at their disposal, along with the capital needed to perform this task. Building LLMs was not feasible for most organizations, but luckily a new pattern of leveraging the power of LLMs with business enterprise data emerged: retrieval-augmented generation.

Retrieval-Augmented Generation

Retrieval-augmented generation (RAG) hit the mainstream in 2020 in a paper titled “Retrieval-Augmented Generation for Knowledge-Intensive NLP Task,”² authored by several members of Meta (formerly Facebook). The concept was developed to improve the accuracy of responses produced by LLMs by pulling in external data sources as part of the response, instead of them responding solely based on their training data.

As shown in Figure 1.1, the core elements of the RAG architecture include the following:

- **Retrieval:** This is where you are querying an external data source based on the user request.

² arxiv.org/pdf/2005.11401

- **Augmentation:** This is where you are combining the response from the external data query with the user prompt and sending both of those together to the LLM for processing.
- **Generation:** This is where the LLM analyzes both the user prompt and the additional context (the “augmentation”) and then provides a response back to the user.

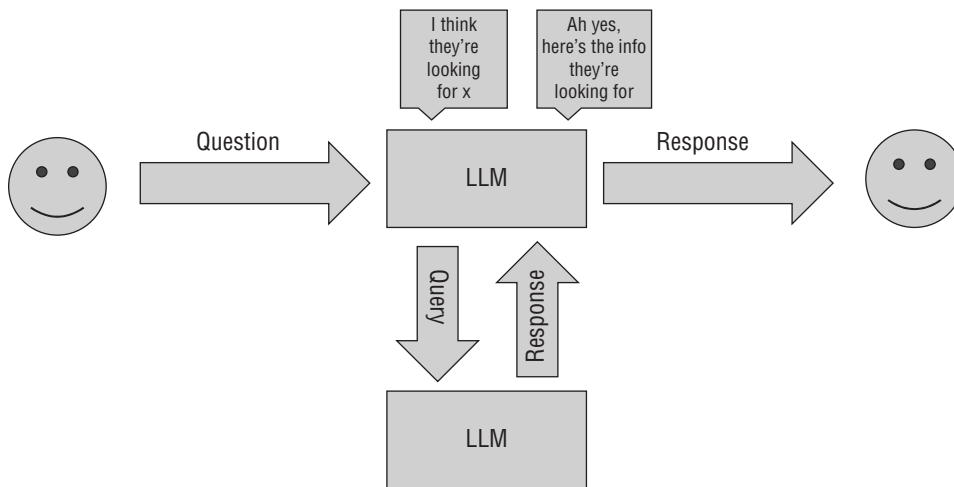


Figure 1.1: The core elements of RAG

The RAG approach enables you to gain the benefits of LLMs, which includes their reasoning and understanding skills. When you combine this capability with your business data, you are providing the LLM with additional context, which is also referred to as *grounding*. This grounding, in theory, should help the LLM predict a response that is closer to your anticipated outcome than if it were solely based on the training data alone. Additionally, this helps address one of the challenges with LLMs: they don’t necessarily provide you with the reasoning or source behind their answer unless prompted to do so.

Azure OpenAI RAG Pattern

The architecture was relatively straightforward and required minimal resources to build. First, you would configure Azure Cognitive Search (later renamed Azure AI Search) to crawl defined Azure Blob Storage, Azure SQL databases, Azure applications, and other data sources. Upon completion of crawling your data sources, it would create an index that you could query from your chat application.

Your chat application would take a user prompt and perform an Azure Cognitive Services keyword search against your business data, which would

return a subset of the data based on the additional parameters you may have specified. It would then pass that response, along with the original user prompt, to Azure OpenAI, reason over the data, and present the response back to the user.

An alternative to this architecture was to pass the user prompt to Azure OpenAI and then describe the type of data in your index, as well as potential keywords to look for. You would then take the modified prompt and perform a keyword search using Azure AI Search, passing that data to Azure OpenAI to reason over. Finally, you would return Azure OpenAI's response back to the user, along with the reference source file that it derived its response from. Typically, the developer would include an additional visual cue in the response to ensure that the user verifies the response provided by AI. Figure 1.2 illustrates an example of this RAG pattern in Azure OpenAI.

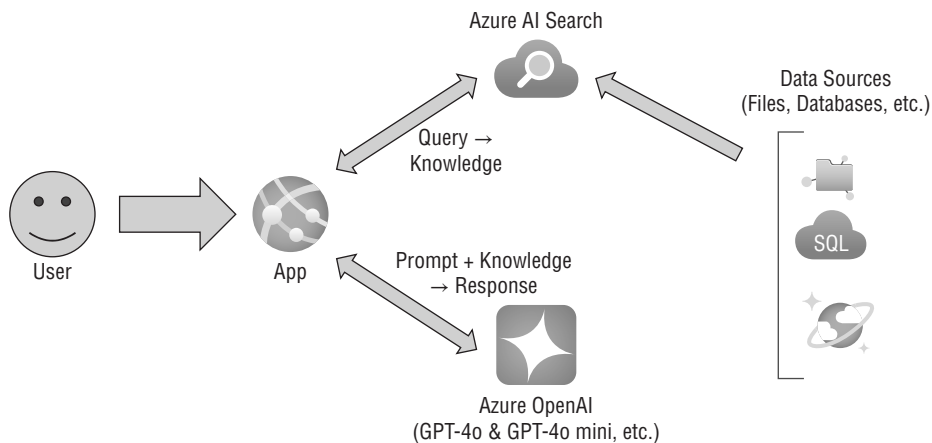


Figure 1.2: The Azure OpenAI RAG pattern

Source: Pablocastro (2023) / Microsoft

Enterprise Adoption of the RAG Pattern

This architecture became highly adopted across both custom GenAI applications developed in house and many off-the-shelf platforms being sold by software companies. Building custom LLMs is an incredibly expensive endeavor of scaling up large amounts of computing power and vast sets of training data to feed the model. Outlaying large amounts of capital without knowing precisely the effectiveness of the model is a risky endeavor.

However, the RAG pattern combined the power of GenAI and business data, while also checking the box of ensuring that your solution had some elements of responsible AI built in. This responsible AI check is satisfied by the LLM not providing a direct answer based on training data but rather citing the source

from where its response originated from. While it certainly wasn't foolproof, it at least provided the file reference for its answer, allowing the user to "check AI's work." This pattern's validity became cemented even further when Microsoft announced its upcoming new GenAI offering, Microsoft 365 Copilot (M365 Copilot).

Microsoft M365 Copilot

Microsoft's announcement of M365 Copilot in March 2023 represented yet another significant evolution in its service offering by becoming one of the first vendors to promise GenAI-as-a-service within its existing SaaS offering. Microsoft Office has been the market leader in the productivity suite category, and the idea of being able to leverage the power of ChatGPT-like functionality within its core application suite made the promises of GenAI even more relevant to business professionals. Microsoft named this new AI sidekick—which would help you with your tasks, answer questions, and engage in conversations—"Copilot." And the rest will someday be history.

Microsoft's investment in OpenAI was incredibly beneficial, as it gave the company a head start in incorporating GenAI into its SaaS offerings several months before its competition. Initially, M365 Copilot was released to select customers under an Early Access Preview (EAP) model, whereby you had to commit to a set 300 licenses for a year. This allowed Microsoft to control the rollout, minimizing risk to its reputation if users encountered shortcomings. The EAP program included assigning a Microsoft Cloud Architect to assist and oversee the deployment, as well as to help collect feedback to better inform the product team about what customers were experiencing through actual usage. This very risk-averse approach is actually something you can observe with the product itself.

As shown in Figure 1.3, M365 Copilot's architecture is an example of the RAG pattern discussed earlier. The user prompt is first sent to Copilot, which then performs a retrieval against the Microsoft Graph + Web sources. The Copilot service then sends that augmented payload to the LLM. The LLM then sends a response back to it, which it evaluates against the Microsoft Graph and Microsoft Purview for a compliance check. Finally, the response is sent back to the user. This additional compliance check before displaying the response back to the user is an example of Microsoft recognizing the needs of large enterprises to log activity and prevent malicious use of technology. This is a consistent theme with all of Microsoft's GenAI offerings: to design with security and compliance in mind.

Microsoft 365 Copilot

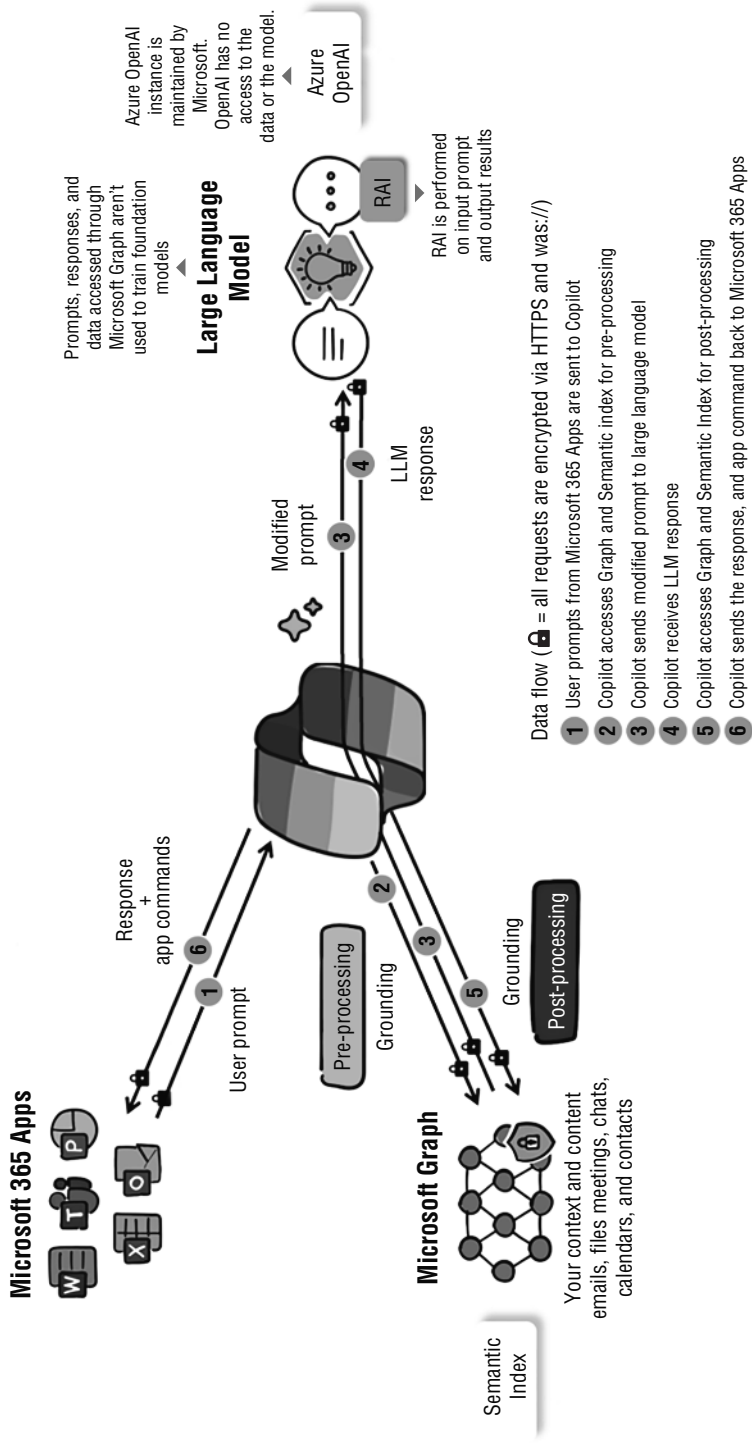


Figure 1.3: The M365 Copilot architecture

Source: Angela Byers (2024) / Microsoft

NOTE This book is a companion to *Microsoft 365 Copilot at Work: Using AI to Get the Most from Your Business Data and Favorite Apps* (ISBN: 978-1-394-25837-6), which goes into greater detail about how to leverage M365 Copilot in your company.

The Rise of the Copilots

Microsoft marketing took to heart Satya's comment that Microsoft is the "Copilot company" and began naming all products with integrated GenAI under the name "Copilot." Microsoft marketing has been a thorn in many enterprise IT professional's lives, as they continue to rename products mid-cycle, causing the need to continuously include both the current and previous name in conversations to ensure that both parties are talking about the same product. For example, "Bing Chat Enterprise" has been renamed to simply "Copilot." As more existing products have been renamed to use the Copilot branding, customers and Microsoft employees alike have been confused.

Copilot, aka Bing Chat Enterprise

Microsoft Bing was launched in June of 2009, when Steve Ballmer was the Chief Executive Officer. Google was by far the leader in terms of Internet search back then, with Yahoo and China's Baidu being the only other competition at the time. Google was the market leader in Internet search to the point that "Google something" was synonymous with performing an Internet search. Previous search engines like MSN, AOL Search, and even Ask Jeeves were unable to ever keep up with the speed, reliability, and relevance of a Google Internet search.

Additionally, Google's business was incredibly lucrative, as advertisers flocked to it and paid large sums of money to have their results promoted for their desired keywords. Google's Internet search business still brings in over 50% of Google's parent company Alphabet's revenue, even with the expansion into other products such as YouTube, Google Cloud, Google Pixel phones, and Google Home devices. Microsoft's foray into the Internet search business had a mix of altruism in "improving the Internet search experience," but in reality, there was a lot of money to be made in stealing away market share from Google's Internet search business.

Microsoft spent an enormous amount of money trying to market Bing, as well as forcing it upon their employees to help market it. After years of trailing Google, in July of 2023, it was announced that enterprise customers would receive AI-powered "Bing Chat Enterprise" as part of their core M365 service. This offering would allow users to perform "AI-powered searches," which were meant to combine the best of Internet search results with ChatGPT-like

AI reasoning. For example, when you asked a question, you might receive an answer directly from the LLM if the answer is within its training data, or alternatively, a RAG-based search would be performed, and you would be provided with a reference for where the answer was found. While Bing's relevance and ranking still lagged Google, the incorporation of GenAI within the search experience was a differentiator, at least for a period.

This new offering would be covered under the existing Microsoft commercial data protection agreement. This ensured that Microsoft would not have access to the prompts or responses, which removed the concern that it would be harvesting this data to further train its LLMs. Bing Chat Enterprise also had enterprise controls that you could manage through your M365 Administrator console. The requirement to be covered under the commercial data protection agreement is that a user would need to access bing.com/chat with a Microsoft Entra ID to ensure that the session was covered under the commercial data protection agreement.

However, the name Bing Chat Enterprise did not last long. On November 15, 2023, just two weeks after M365 Copilot became Generally Available, Microsoft decided to rename this new offering to "Copilot." This name change, while innocent enough, was a significant contributor to confusion around which Copilots bore an additional cost and which ones were included as part of current licensing agreements. Additionally, with this change, the homepage for Bing Chat Enterprise transitioned to copilot.microsoft.com and included a new toggle to flip between "Web" and "Work" for customers with an M365 Copilot license, as shown in Figure 1.4.

Interacting with the "Web" version of Copilot only returned Internet results, whereas interacting with "Work" meant you would have access to your enterprise data sources, such as documents in Microsoft Teams and OneDrive. To add even additional confusion, Microsoft's marketing renamed the "Work" mode of interacting with Copilot to "BizChat." However, just like anything else with Microsoft marketing, there's a good chance that name could be changed again.

Microsoft's Copilot Portfolio

Before we dive into the main topic of this book, Copilot Studio, it is important to have a baseline understanding of Microsoft's other offerings. As we will discuss later, the main value proposition for Copilot Studio is that it enables you to build additional functionality for an existing copilot product or to create a standalone copilot agent that can be deployed to a number of supported communication channels. So, for now, let's step through some of Microsoft's featured copilot products.

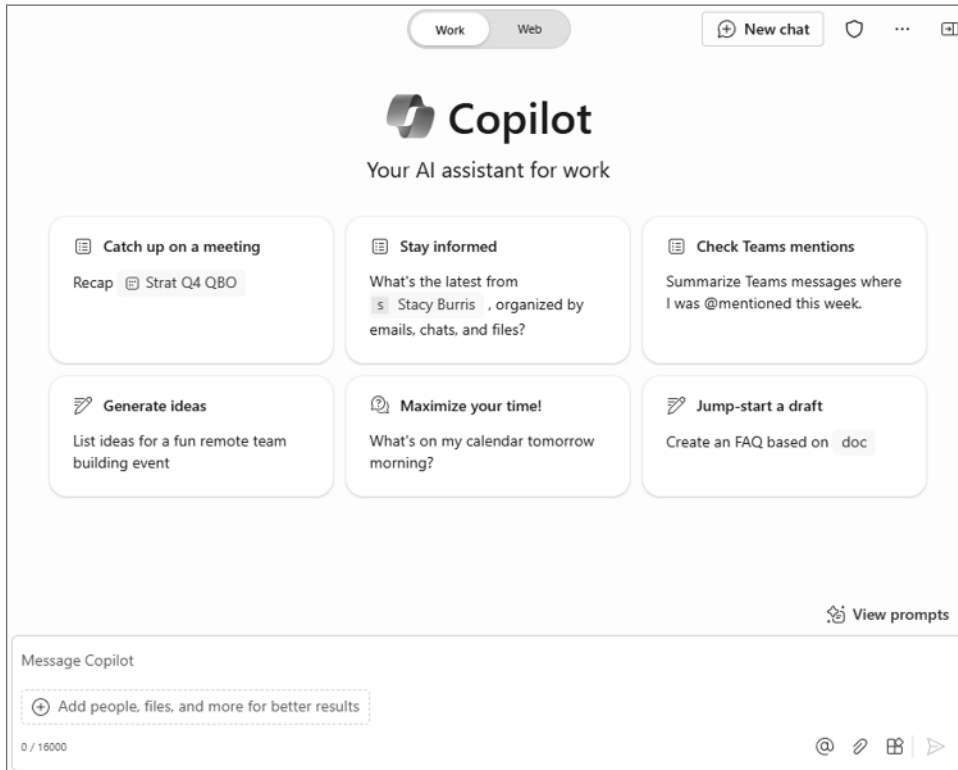


Figure 1.4: The Copilot homepage

Copilot for Sales

Copilot for Sales is an exciting AI product focused on individuals in sales-oriented roles. While M365 Copilot was focused on enabling knowledge workers to be more productive, the value proposition for Copilot for Sales is its integration with Customer Relationship Management (CRM) software, including both Dynamics 365 Sales & Marketing and Salesforce’s Sales Cloud product. There’s no question that Salesforce is the most dominant CRM in the marketplace today, with over 25% of the total market share, and their own investment in AI as part of its Einstein technology. Microsoft’s support of Salesforce’s CRM product is a prime example of Satya Nadella seeing an opportunity to serve a larger audience.

NOTE Copilot for Sales is another example of how a product can have more names than it has years of being on the market. What started out as “Viva Sales” became “Sales Copilot” and then finally “Copilot for Sales.”

The Copilot for Sales functionality is available in both Microsoft Teams and Outlook and includes the following features:

- **AI-assisted email generation:** Copilot can help draft an email and optionally reference information from a recent Teams meeting to help provide additional context.
- **Accessing Customer Relationship Management (CRM) objects from Outlook:** While within an email, Copilot will help pull up any associated CRM data, including the Account, Contact, and Opportunity objects based on the participants in the email conversation.
- **Updating CRM from Outlook:** In addition to having Copilot reference CRM data, you can also update key fields for an Opportunity or Contact record directly within Outlook, in addition to being able to save an email to that Opportunity record.
- **Opportunity summary:** Copilot will summarize the key updates for a given opportunity and provide what it believes is the next steps are to help close that opportunity.
- **In-Call insights:** Copilot will provide sales tips and pull up relevant information during Microsoft Teams calls to help coach your sellers toward winning an opportunity.
- **Call analysis:** Copilot will summarize the key points from a meeting that can be used in a recap email, suggest tasks based on the conversation flow, and highlight keywords mentions, along with sentiment during the call.
- **Sales automation:** Automatic provisioning of channels in Microsoft Teams enables the internal team to collaborate on proposals, contracts, etc., while also creating an external sharing channel to collaborate with your customers securely.
- **Sharing of adaptive cards in Teams messages:** To help reduce application context switching, you can share an opportunity, account, or contact record directly within Microsoft Teams and maintain any security trimming in the backend CRM.

Microsoft's vision for Copilot for Sales is enabling your sellers to stay in the applications they use daily to stay connected with your customers: Microsoft Teams and Outlook. AI helps pull relevant CRM data into the flow of work, thus removing the additional context switching. It also serves as a coach to your sellers, providing actionable insights and potential next steps to help continue driving opportunities forward.

There is a licensing consideration to be aware of when using Copilot for Sales. Copilot for Sales has a dependency on the user also being licensed to M365

Copilot. This would be applicable even if you have a Dynamics 365 Enterprise license, which includes a Copilot for Sales license as part of the bundle. It can be purchased as an add-on to an existing M365 Copilot license, or you can purchase them together as a bundle. At the time of writing, the list price of the combined M365 Copilot & Copilot for Sales bundle is the same as if you were to purchase them separately.

NOTE Always be sure to confirm licensing with an Authorized Microsoft Reseller, your Microsoft Account Team, or the licensing guide, which you can consult at microsoft.com/licensing/docs/view/Licensing-Guides.

Copilot for Service

Copilot for Service follows many of the same patterns as Copilot for Sales, including the same licensing model, presentation through Microsoft Teams and Outlook, and interoperability with the Dynamics and Salesforce CRM platforms. While both offerings are targeted at improving the customer experience, Copilot for Service focuses on after-sale service activities such as case resolution, collecting and processing customer feedback, identifying potential upsell opportunities, and handling warranty claims or return/exchange activities.

The target persona for Copilot for Service includes individuals who work in your Contact Center or help support Customer Service and even Customer Outreach activities. Some examples of where Copilot for Service can help improve the overall customer service experience include:

- Assisting with creating customer outreach emails based on their interaction history with your company
- Summarizing a customer case using AI to help a customer service representative get up to speed faster with the history and recommended next steps
- Quickly confirming policies and procedures associated with a given activity (e.g., warranty coverage, return policies)
- Assisting customer service representatives to understand their current workload through natural language—for example:
 - Show me all my active cases.
 - Show me all my escalated cases.
 - Show me all my cases due tomorrow.
- Enabling managers to quickly understand the health of their business through natural language prompts—for example:

- What cases missed their service-level agreement (SLA) in the past week?
- Show me active escalated support tickets for “Heidi Lamar” in the past 30 days.
- Show me the high-priority support tickets for Deepak’s Dairy in the past 15 days.

Customer service groups are constantly under pressure to improve the service quality across multiple KPIs, including net promoter score, customer satisfaction scores, first contact resolution, response time, ticket volume, resolution time, agent productivity, etc. The goal is to ensure that customers stay loyal to the company through positive experiences across the value chain. Copilot for Service’s goal is to help improve these scores and retain customers through AI-generated insights and personalization.

NOTE Additional information about Copilot for Service, including pricing and implementation details, can be found at microsoft.com/en-us/microsoft-365/copilot/copilot-for-service.

Copilot for Security

Copilot for Security is another role-based AI technology targeted at IT professionals responsible for cybersecurity tasks such as threat detection and analysis, incident response, threat hunting, and user activity monitoring. While M365 Copilot focuses heavily on the ability to create content, the real value of Copilot for Security is in its ability to parse large amounts of data and provide critical insights quickly. Microsoft has included the Copilot experience across its security portfolio of applications, including:

- **Microsoft Defender Extended Detection & Response (XDR)**
 - Assists with summarizing security incidents, including assessing the impact, providing recommendations for how to address, and then assembling the post-response activity report for tracking remediation efforts
 - Enables security professionals to assess risk through natural language to provide insights around potential security issues or vulnerabilities
- **Unified SOC Platform**
 - Assesses emerging threats and provides AI-driven insights
 - Streamlines investigation and response activities through AI-driven summarization and recommendations

- Enables security analysts to complete tasks through natural language, which Copilot translates to Kusto Query Language (KQL) for processing
- **Microsoft Purview**
 - Enables visibility across multiple solutions and layering in applicable compliance and regulatory requirements
 - Summarizes alerts across multiple channels to help prioritize where to focus based on business impact
- **Microsoft Entra**
 - Accelerates investigations of identity-related issues through natural language and by providing AI-driven recommendations to improve security posture
 - Finds gaps in current policies, generates workflows, and arrives at root causes faster
- **Microsoft Intune**
 - Remediates endpoint issues with what-if analysis and through AI-driven analysis of the potential problems and resolution actions
 - Assists with translating business intent into actionable configuration and policies using natural language

Microsoft's value proposition for Copilot for Security is that it enables IT security professionals to focus on the most important signals that are being generated through their enterprise security portfolio. It also removes some of the barriers that perhaps have prevented organizations from tapping into additional in-house talent to assist with incident response activities, including removing the need to know how to construct Kusto Query Language (KQL) for platform queries. Finally, given that Microsoft is one of the largest IT security companies, with a yearly \$1 billion investment, organizations can benefit from the learnings of these investments to better enhance their own security posture.

Copilot for Security differs from the other copilots Microsoft has brought to market because it does not have a per-user, per-month subscription model. Copilot for Security is licensed via a pay-as-you-go model. You are paying for the computing capacity to perform analysis of the various log files across Microsoft Security. This model is similar to a typical cloud computing model, where you provision resources and then pay for the time they execute transactions in the cloud. Microsoft recommends a minimum of three Security Compute Units (SCUs) to begin experimenting with Copilot for Security. These units carry a per-hour charge that is included in your monthly Azure bill along with any additional cloud infrastructure in your subscription.

While not a firm requirement, Microsoft also recommends that customers who are planning on using Copilot for Security also have Microsoft Sentinel SIEM service provisioned along with Microsoft Defender for Endpoint (MDE) Plan 2 (P2), which is an advanced endpoint security solution that includes features like automated investigation and remediation, advanced threat prevention, threat and vulnerability management, and threat hunting capabilities. So, when considering moving forward with Copilot for Security, you will want to factor in all the additional cloud computing costs and potential ancillary licensing costs associated with this cutting-edge service.

NOTE Additional Copilot for Security pricing details can be found at azure.microsoft.com/en-us/pricing/details/microsoft-copilot-for-security.

Copilot in Microsoft Viva

Microsoft Viva is not a single product but rather a digital employee experience platform focused on tools and technologies to improve the employee experience. While much of M365 focuses on productivity and collaboration, Microsoft Viva focuses on employee growth, learning, well-being, and connectedness to others. It is made up of several individual products that come together to help round out the experience. It is probably no surprise that certain modules have been prioritized with Copilot functionality to help provide additional value in Microsoft Viva.

NOTE Microsoft Viva licensing is complicated, as there are some features included in your base M365 license and others require a separate add-on license. It is best to review the licensing comparison table found at microsoft.com/en-us/microsoft-viva/pricing.

Viva Goals

Microsoft Viva Goals is a goal alignment solution that uses the popular objectives and key results (OKR) framework to help connect people to an organization's strategic priorities. It helps employees understand the impact of their work and how it aligns with the company's mission and drives business results. Setting these goals can be a bit challenging, which is where Copilot in Viva Goals comes in handy. You can collaborate with Copilot as part of your initial goal-setting process to brainstorm appropriate OKRs based on your company's strategic vision. For example, you could ask Copilot, "Please help me generate OKRs

for bringing a new GenAI product to market,” and it will assist with creating objectives and then key results that would roll up underneath each objective, as shown in Figure 1.5.

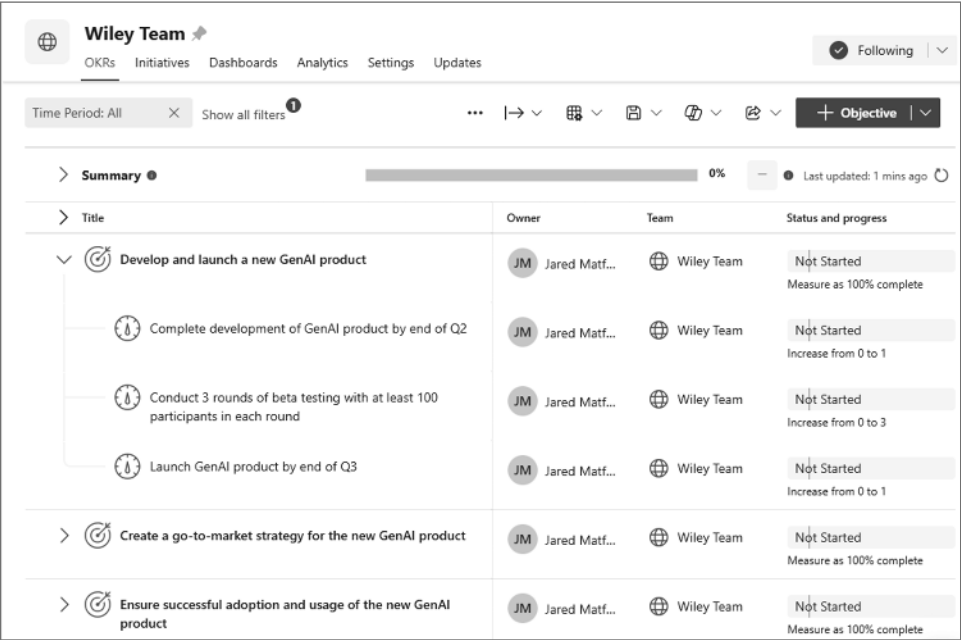


Figure 1.5: Copilot-generated OKRs

Viva Engage

Viva Engage is a social collaboration platform focused on employee engagement, knowledge sharing, corporate communications, and building communities within an organization. It also helps establish a channel for leaders to connect with employees at all levels of the organization through constructive dialogue. This two-way communication channel helps leaders keep their finger on the pulse of the organization while identifying high-potential individuals. For larger organizations with lengthy hierarchies, this also helps more junior employees feel part of something bigger than just their team or function.

NOTE In February of 2023, Yammer was rebranded as Viva Engage and given a new set of features to help drive interest and support the niche customers who had already invested in the platform. Users can leverage Copilot in Viva Engage through their home feed, storyline, community feed, and campaign pages. An example use case might be a communications manager or senior leader who wants to share an update with the company. For example, see Figure 1.6, where I asked Copilot to assist with creating a post announcing the release of this book. It created a draft for a brief message and included fun emoticons to enhance the post.

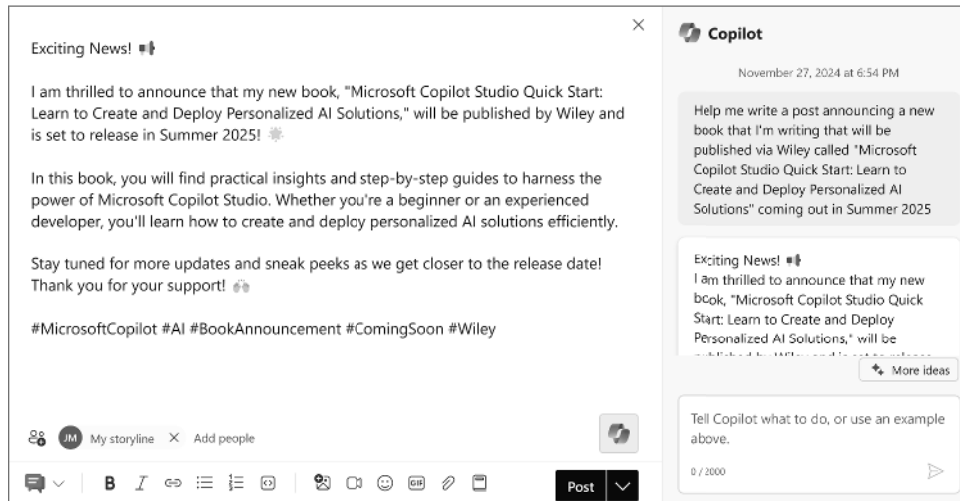


Figure 1.6: Copilot in Viva Engage

NOTE To leverage Copilot in Viva Engage, you will need both a Viva Engage license (purchased as part of Microsoft Viva Suite or Microsoft Viva Employee Communications and Communities) and an M365 Copilot license.

Viva Amplify

Viva Amplify is another solution in the Viva Suite designed to help communications professionals create, manage, and publish campaigns to engage employees across multiple channels. It centralizes communication processes, offering tools for campaign management, writing guidance, and analytics to measure the impact of messages. These communications are authored in modern SharePoint pages and then disseminated through email, SharePoint Sites (via SharePoint news), and also Microsoft Teams. Communications professionals can then measure which channel was most effective.

The Copilot experience in Viva Amplify is centered around helping communicators refine their messages. When drafting a new communication as part of a campaign, you can use Copilot to help “Auto rewrite” or adjust the tone of your communication, as shown in Figure 1.7. This functionality helps communicators refine their messages to make them more accessible to the target audience. A potential communications strategy might be to create multiple versions of the same communication for managers, individual contributors, and executives. They could start with the initial draft communications and then leverage Copilot to help refine the tone per target audience.

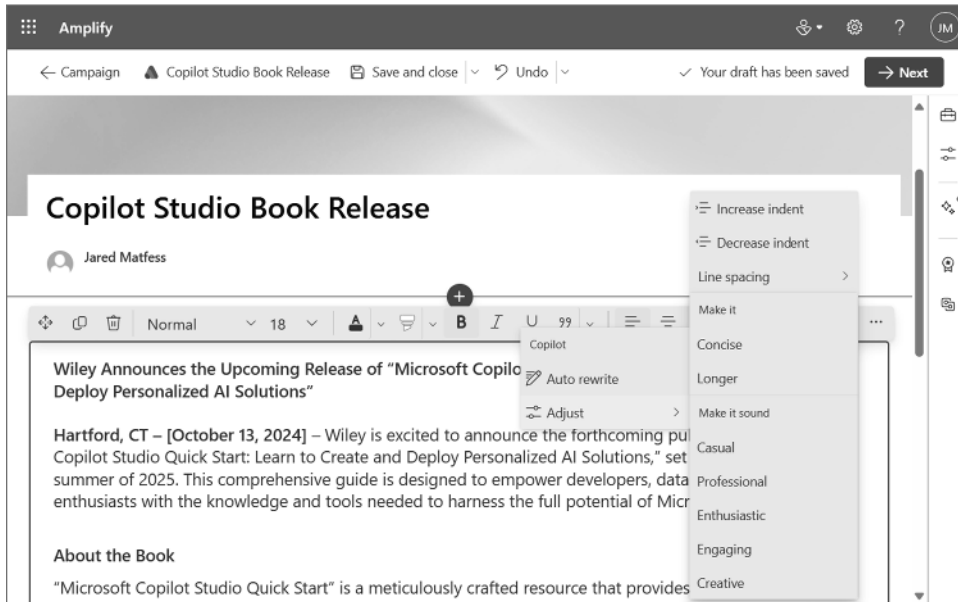


Figure 1.7: Copilot in Viva Amplify

Viva Roadmap

These are just a couple of examples of Microsoft integrating Copilot within the Microsoft Viva Suite, with many more features included in the roadmap. Known targets include Viva Glint, with planned functionality that enables Copilot to analyze employee feedback and survey data to unlock critical insights and identify outliers in the dataset. It will be very interesting to see how Microsoft addresses the licensing requirements across experiences and whether there are any concessions made to help boost the adoption of M365 Copilot.

Additional Copilots

The intent of this book isn't to go deep into all of Microsoft's Copilot offerings but rather to provide a high-level overview so you can start to see some of the emerging patterns for what is being brought to market. There are several additional copilots, including:

- **GitHub Copilot:** Microsoft acquired GitHub back in 2018, giving it access to one of the world's largest developer platforms, including all of its open-source software. This served as training data for GitHub Copilot, which is GenAI within popular integrated development environments (IDEs), such as Visual Studio, Visual Studio Code, Neovim, and JetBrains IDEs. GitHub Copilot can both suggest code and act as a pair-programmer,

along with being able to analyze existing code and, in natural language, provide an explanation for what it does.

- **Copilot for Finance:** Another role-specific offering, Copilot for Finance, is targeted at streamlining tasks like quote-to-cash processes, credit and collections, risk management, and compliance activities. In addition, it can help enable executives and managers who need real-time insights into business performance and strategic decision-making based on financial data.
- **Copilot in Azure:** Microsoft has developed a Copilot that's focused on your Cloud team and can assist with tasks such as summarizing bills, providing recommendations to optimize cost based on activity, receiving recommendations of services based on your business use case, or navigating the Azure portal using natural language.
- **Copilot in Microsoft Power Platform:** Similar to the Copilot in Azure, Microsoft has introduced GenAI into its key Power Platform applications, such as PowerApps, Power Automate, Power BI, and AI Builder. This enables makers to create compelling business solutions using the power of natural language. Example use cases include:
 - **Power Automate:** Describing a workflow to help solve a business problem, with Copilot creating the initial logic as a starting point for you to iterate on
 - **Power Automate:** Creating PowerApps, including the design, layout, fields, etc., to help create exciting low-code business applications
 - **Power BI:** Automatically creating reports, generating actionable insights, and designing dashboards using the power of natural language

With Microsoft focused on being the Copilot company, keeping up with the pace of its innovation can be incredibly challenging. For IT budget owners feeling extreme pressure from the business to reduce costs, navigating these various Copilot offerings and developing a strategy for enabling innovation while also controlling costs can be challenging. There is also likely pressure from your executives, the board, and even business stakeholders to understand where you are from an AI strategy perspective, which offerings might be coming, and when.

The consumerization of AI in solutions with similar capabilities to ChatGPT will continue to make it difficult for IT leaders to keep the demand at bay. Therefore, it is essential to plan and communicate your AI strategy to help inform your stakeholders and help you plan the financial investment required to enable these new capabilities. As part of that planning, you must factor in the people within your organization as part of a mapping exercise to determine who gets which copilots or other AI offerings.

Aligning Copilots with Company Personas

The approach you will want to take to organize this AI boom is first to take an inventory of the personas within your organization. The reality is that not everyone is going to need every copilot. There are some departments whose function and personnel will align to a particular body of work, and you will want to understand their primary business functions and begin aligning which copilot might deliver value. Through a persona mapping exercise, you will likely develop something similar to Table 1.1, which outlines the persona, potential products, and highlighted use cases or business outcomes you hope to drive.

Table 1.1: AI/Copilot persona mapping

PERSONA/DEPARTMENT	PRODUCT	USE CASES
Developers	GitHub Copilot	■ Increase developer productivity ■ Improve code quality ■ Help support junior developers through AI assistance
Knowledge workers (marketing, communications, legal)	M365 Copilot	■ Improve the quality of documents ■ Provide an intelligent recap of virtual meetings ■ Improve enterprise search
Customer service representatives	Copilot for Service	■ Reduce case resolution time ■ Provide enhanced customer service
Sales reps or alumni relations (university)	Copilot for Sales	■ Streamline customer sales activities ■ Provide AI coaching to improve presentation skills ■ Enable sellers to spend more time working with customers
Cybersecurity professionals	Copilot for Security	■ Help reduce the risk associated with security incidents ■ Improve the security posture of your organization ■ Focus on the most critical threats

Once you have finished the persona mapping exercise for your organization, you may find some personas are not represented, or some use cases may not be satisfied by current out-of-the-box capabilities. For those personas and use

cases, you can fall back on traditional IT analysis techniques to solve business problems and align technology solutions by:

- Understanding the business problem
- Gathering business requirements
- Assessing current system capabilities
- Researching and proposing solutions

As part of the researching and proposing solutions step, you may also consider build versus buy scenarios. Since you have already attempted to map use cases against Microsoft's existing portfolio, from an analysis of what you might be potentially "buying," you would be looking at other third-party products and the functionality they provide.

Assuming that you believe the business use case would benefit from a custom AI-related solution, there are two questions for you to consider:

- Is this solution something you would build as a standalone product, such as a fit-for-purpose copilot targeted at a particular persona?
- Would the solution have a better user experience if the functionality were made available through an existing Copilot product?

If you answered yes to either question, then the approach for building this solution might be a good use case for Microsoft's Copilot Studio.

Introducing Copilot Studio

Copilot Studio is not a standalone copilot like many of the other copilots we've reviewed thus far. It is a low-code application platform positioned to serve two purposes:

- **Extend first-party Copilots:** Similar to any enterprise software package, you can build specific customizations unique to your business for Microsoft's existing copilot products, such as M365 Copilot, Copilot for Sales, and Copilot for Service. For example, you could build a Copilot Agent that extends Microsoft 365 Copilot to pull your PTO balance or to return policies and procedures information from a preset list of SharePoint sites.
- **Build standalone Copilot agents:** There may be scenarios where the user might not be using a first-party copilot, such as external customers or your frontline workers. Copilot Studio allows you to develop standalone copilot agents that can help to assist them with tasks. Some examples include a copilot agent to assist customers with navigating your website or to assist your frontline workers with submitting safety incident reports from the factory floor. These agents can be published across several different channels to help deliver a custom experience for your target userbase.

Copilot Studio is part of Microsoft’s Business Applications Solution area and can integrate with core elements of Microsoft’s Power Platform portfolio of applications, as shown in Figure 1.8. The Power Platform also helps to ensure that the solutions developed can connect to other line-of-business applications through a mix of low-code and professional-code solutions.

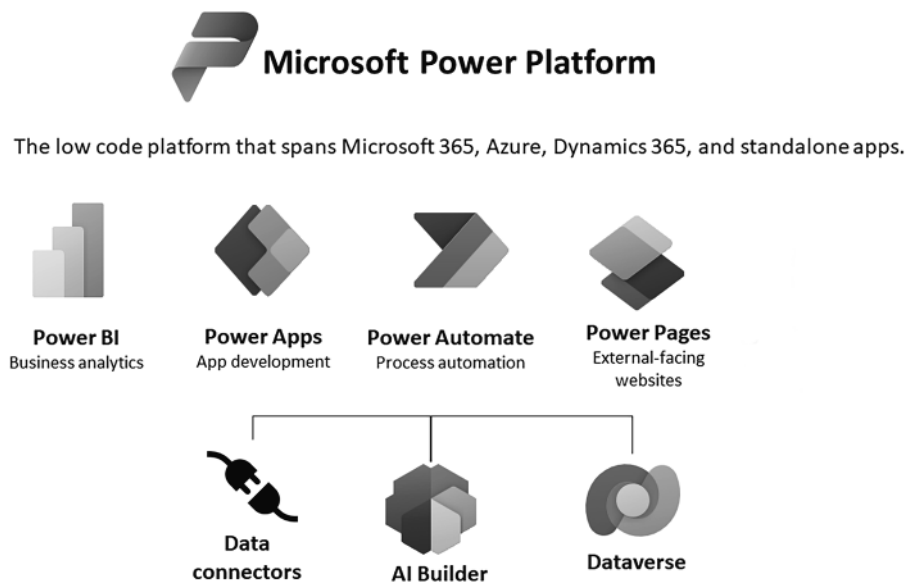


Figure 1.8: Microsoft Power Platform

In the next chapter, we will explore Copilot Studio in detail, including the licensing prerequisites, how to access it, and the key elements of building your own custom copilot. Later in the book, we will walk through extending M365 Copilot with additional functionality, also configured using Copilot Studio.

Conclusion

Microsoft’s copilot ecosystem is growing by the day, all powered by its strategic partnership with OpenAI. Satya’s direction of Microsoft is to be the copilot company and build a future with a copilot for everything you do. In this chapter, we went over a high-level overview of GenAI and how it catalyzed AI-driven innovation. We stepped through Microsoft’s investment in OpenAI and its impact on shaping the products and solutions that Microsoft is bringing to market. We then studied Microsoft’s numerous copilots, their target personas, and key use cases that drive business value.

We then discussed how to map Microsoft’s copilots to personas in your organization in a “when to use what” type manner. Finally, the key topic of this book, Copilot Studio, was introduced, and you learned where it fits in the current landscape. Now that you have been acclimated to Microsoft’s various copilots, let’s dive into Copilot Studio.

