

Work IQ

Q: What is Work IQ?

Work IQ is a workplace intelligence layer that delivers a semantic understanding of everything happening across your business. Work IQ is built for an agent-first world, where AI agents dynamically reason, retrieve context, and act across systems. It unifies chat, context, tools, and workspaces into a single intelligence layer so agents can operate with the same depth, grounding, and reliability as Microsoft 365 Copilot. It is comprised of four components:

- Chat: Optimized for conversational intelligence. Work IQ supports A2A for agent-to-agent collaboration and REST for real-time interactions, enabling agents to delegate work, maintain continuity, and deliver Copilot-quality responses.
- Context: Optimized by eliminating orchestration overhead. Work IQ assembles and grounds context internally across organizational data, so agents don't need to stitch together raw signals or manage retrieval pipelines.
- Tools: Optimized for fast, composable actions. A compact set of generic tools and skills lets agents reason, retrieve data, and act across Microsoft 365 through a simplified, agent-friendly surface with centralized governance.
- Workspaces: Optimized for long-running agent workflows. Backed working storage gives agents a persistent space to manage intermediate data and outputs—supporting reliable task progression, reuse of results, and seamless handoff across agents and experiences.

Q: Where can you experience Work IQ?

Work IQ powers experiences across Microsoft 365 Copilot and agents, including Copilot Chat and in-app Copilot experiences (Word, Excel, PowerPoint, Outlook, Teams). You can also extend Work IQ to your own custom agents with the Work IQ API. It operates behind the scenes to ground responses and actions in live work context rather than surfacing as a standalone product.

Q: What is special about Work IQ APIs?

Work IQ APIs are unique because they deliver intelligence, speed, efficiency, scale, and governance in a single platform purpose-built for agentic systems. Work IQ provides intelligence instead of data, using a unified runtime to reason over organizational context, ground responses, and orchestrates actions automatically. ³ They are optimized for modern AI workloads, enabling faster task execution and reduced complexity by handling context assembly, reasoning, and tool selection within the platform, so developers don't have to build and maintain custom pipelines. At the same time, Work IQ APIs are built for enterprise scale with a single, unified surface that works across protocols and scenarios, while enforcing built-in security, compliance, and permission-aware access, ensuring every interaction is governed by the same policies as Microsoft 365 from day one.

Q: Who can use the Work IQ API?

The Work IQ API is for developers, partners, and organizations building custom agents and AI solutions outside Microsoft 365 Copilot. It is Generally Available (GA). There is no entitlement in the \$30 Microsoft 365 Copilot license; usage is billed via Copilot Credits, and an admin must enable usage-based billing in the Microsoft 365 Admin Center.

Q: How does billing work for Work IQ?

Usage-based billing requires an administrator to turn it on and set limits in the Microsoft Admin Center (MAC) before billing begins. Start here: <https://admin.microsoft.com> In MAC, tenant administrators can create named billing policies for credit-based access across the organization and set up access policies for all users, specific groups, or individual users based on need. This might include:

- Enabling specific services, such as the Work IQ API
- Choosing which agents and services selected users and groups can consume credits against and aligning credit access to approved services and agents only
- Sending alert emails to admins and stakeholders as usage approaches defined limits
- Applying optional user-level caps to protect shared credit capacity and reduce overspend risk

Q: How do admins buy and manage credits?

Admins can purchase Copilot Credits through the Cost Management dashboard in MAC and use MAC controls to manage how access is rolled out across the organization. There are two ways to acquire Copilot Credits:

1. Prepaid (Copilot Credits P3). This is an annual commitment that gives you the lowest per credit cost. It's ideal once you have a sense of your usage patterns and want to lock in predictable spend. If you're already running Copilot Studio agents or other usage-based workloads, you can pool those credits across services.
2. Pay-as-you-go at one cent per credit. There is no annual commitment and no usage term, you simply pay for what you use. This is the natural starting point if you don't have prior usage history yet or want to pilot Work IQ before committing to a larger license volume. It gives you full flexibility to ramp up or down without any lock-in.

Q: How do admins govern usage?

Admins monitor Work IQ usage through a single Cost Management dashboard that shows credit usage, remaining capacity, and spend across users and groups. They receive alerts as usage approaches limits or triggers pay-as-you-go, along with early warnings for over-utilization and a forecast of remaining capacity. Admins are notified when users request more capacity. These notifications are surfaced in the same dashboard, where they can review, approve, or bulk-manage the requests and make adjustments to limits as necessary. In subsequent releases, admins will also be able to drill down into usage by user or group to identify high-usage patterns and cost drivers, and spend trends. Get started at <https://admin.microsoft.com>.

