Measure what matters: GenAI effectiveness

Why effectiveness is the measurement that matters

"We are encouraging our employees to use GenAI apps but are we more productive?"

This is not the first tech cycle where the answer is elusive. As economist Robert Solow noted in 1987, *"You can see the computer age everywhere but in the productivity statistics".* These cycles tend to follow a pattern. First, there is a tech build (at an expense) but the payoff only happens when organizations have 'rewired' themselves.

The Internet exhibited that pattern. GenAI is clearly following suit.

Personal productivity GenAI apps make the biggest difference for knowledge workers. Unfortunately, knowledge work has no single, measurable outcome. That's why NROC Security measures factors you *can* observe and act on, the **effectiveness of GenAI usage**. For example, how often employees use GenAI, what tasks are they trying to accomplish and how skillfully they prompt. These behaviors are the leading indicators of the rewiring needed that lead to the payoff.

For this research, we sampled 4,800 business users and 139,000 GenAI interactions during 2Q26. The users are across industries and represent a typical enterprise mix of knowledge worker profiles.

The big picture: adoption is climbing and broadening – effectiveness remains low

From Q1 to Q2, active usage nearly doubled, governed tools were added but personal consumer app plans continued to be popular. Personally Identifiable Information (PII) exposure fell as use cases evolved (with attachments as the new concern). The effectiveness index climbed but from a very low base.

Metric type	1Q26	2Q26	Change
% of eligible employees active with GenAI	19%	**31%**	▲ +12 pts
% employees active weekly	14%	**21%**	▲ +7 pts
App families in use	19	**53**	▲ 34 apps
% active users showing expert+ skill	28%	**48%**	▲ +20 pts
% active users on free & personal apps	80%	**52%**	▼ -28 pts
PII matches* / 100 prompts	52	**37**	▼ 15 instances
% of attachments to free & personal plans	N/A	**23%**	N/A
GenAI effectiveness index	0.7%	**1.4%**	▲ +0.7 pts

**On comparable basis. Inclusive of banking and credit cards, contact information and customer specific keywords. Exclusive of names and IP addresses.*

Key findings: governance and enablement are not keeping up with usage

GenAI usage nearly doubled in a single quarter to 31% of employees

The share of active users increased to 31% of eligible employees (from 19% three months earlier), and the share weekly active users went up by 7 points to 21%. Using GenAI has clearly crossed from a curious minority into mainstream for daily work tasks. Usage is the first prerequisite of effectiveness, and it is accelerating fast.

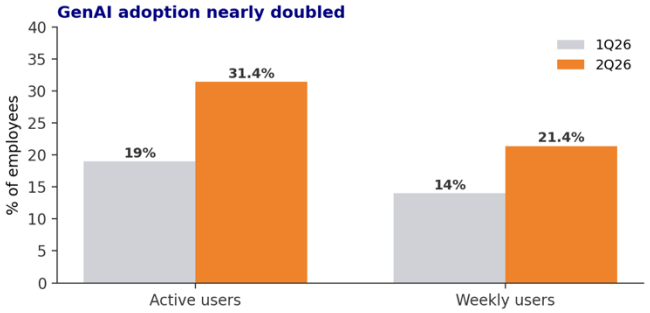

Figure 1 — Quarterly active and weekly GenAI users

Microsoft Copilot became ubiquitous, but 43% used ChatGPT on a personal plan

Governing access to apps has become more challenging. The active portfolio widened from 19 to 53 app families, of which the top 3 apps represented 84% of usage. Of the top three, Microsoft Copilot climbed to account for 40% of prompts, mainly due to its many appearances throughout the Microsoft 365 estate. Despite increased purchases of enterprise licenses, free/personal apps (primarily ChatGPT) maintained its appeal, attracting 53% active users (previously 80%). Claude took the 3rd spot, with a more expert use-case mix compared to Copilot and ChatGPT.

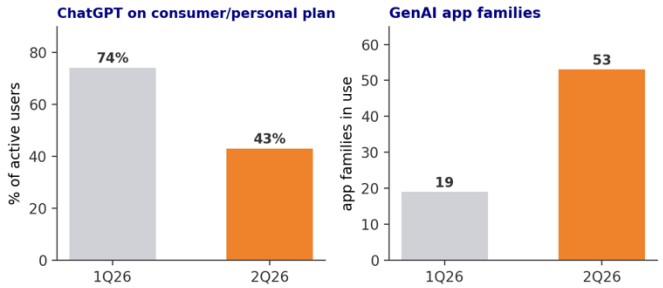

Figure 2 — ChatGPT consumer plan usage reduced due to increased purchases of enterprise subscriptions and Copilot availability. The number of app families nearly tripled.

89% of prompts resemble Google searches

On NROC Security's 6-point prompting skill scale, 89% of prompts are at most of intermediate skill (down from 96% in the previous quarter), but the share of users that reached expert+ skill level climbed from 28% to 48%. The ability for users to take more advantage of GenAI is increasing, but the skill-level delta between advanced users and beginners (or inactive) users is tremendous.

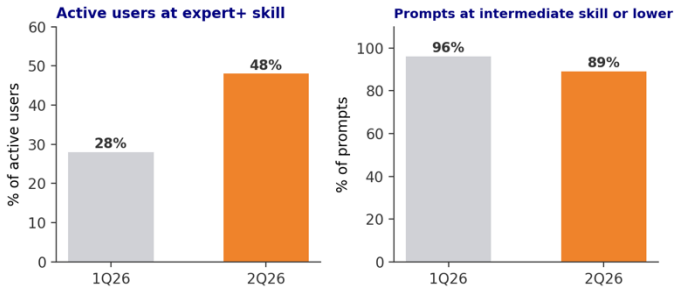

Figure 3 — Expert-level prompting increased heavily, but still nearly 89% of prompts where of the lowest two skill-levels out of six

23% of attachments are uploaded to consumer GenAI apps, increasing the risk of data leakage

Larger context windows and effective optical character recognition are encouraging users to upload documents to GenAI apps. About 23% of uploaded files reached free or consumer apps, and image files became the second-biggest file-type category (after Office documents). Among the image files, screenshots transcribed by AI are becoming a 'use case' for employees to extract data from locked-down IT systems.

Due to common use cases of email-writing and rewriting, there is still a lot of Personally Identifiable Information (PII) in prompts, but it decreased from 52 instances to 37 per 100 prompts.

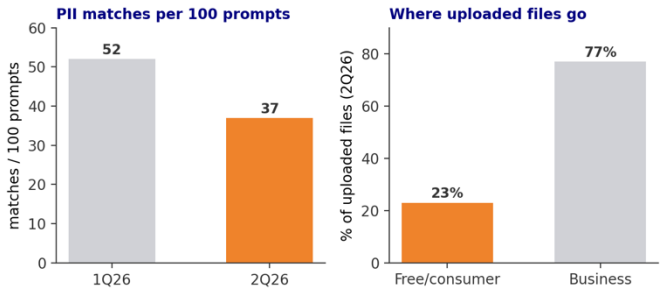

Figure 4 — PII per 100 prompts fell from 52 to 37, but attachments are a growing risk; about 23% of uploaded files end up in free and consumer apps.

GenAI Effectiveness Index is at 1.4% of working hours (up 0.7 pts from previous quarter)

To compute an illustrative GenAI Effectiveness Index, the NROC Security team divided end users up into 5 segments based on frequency and skill-level of prompts. The segments range from inactive users to those who used GenAI daily with high-skill prompts. Increased usage reduced in the inactive segment and grew the active ones. The most effective users - weekly and daily users with high-skill prompts - are the segments to watch but they represent still only about 5% of employees (up 10x in one quarter).

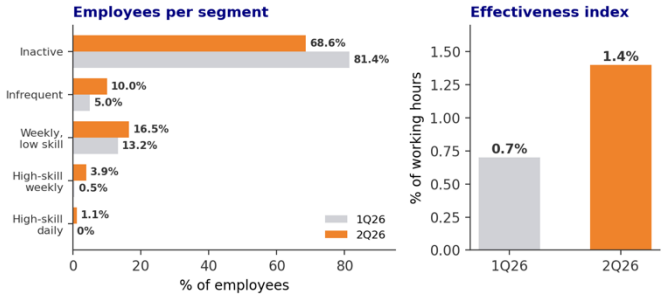

Figure 5 —Distribution of employees across the five effectiveness segments and index development

The segments were translated to an index by assuming a productivity potential for each segment, ranging from 0% to 20%. This results in a weighted average of 1.4% of GenAI effectiveness (or personal productivity potential), which went up 0.7 pts from the previous quarter.

Key take-away: drive GenAI enablement with productivity-first governance

AI challenges traditional IT governance, which relies on up-front analysis and approval of use cases. With GenAI apps, end users innovate without asking permission. In such a setting, observability and technical controls are key. Heavy-handed blocking tends to move usage to channels where the company has no visibility or control, like private laptops and accounts. Productivity-first governance balances exploration with safety in five steps:

1. **Allow, monitor, and react.** Let people use GenAI, keep full visibility, and intervene on real risks rather than blocking by default. You do not need to approve a multitude of apps, but 3-4 to satisfy use cases and preferences.
2. **Security as a safety net.** Use guardrails as real-time guidance, serving a dual purpose: protect data and act as a visible safety net for end users, thus inspiring confidence to use GenAI apps. Enforcing stricter guardrails for free and personal apps offer cost-efficient alternatives for lower-activity users.
3. **Identify and spread best practices.** Surface what your top 5% of users do well and turn it into training modules and inspiration for all employees.
4. **Train team by team.** Elevate prompting skills team by team, focusing first on repetitive common tasks where someone on the team already discovered useful GenAI prompts. These use cases are candidates for automation investments.
5. **Monitor, guide, and drive accountability.** Visibly monitor usage, reach out to the users bumping the guardrails most, and use real-time interventions, such as redactions or "are you sure?" nudges to drive accountability.

Collectively, these actions drive the factors of GenAI effectiveness: access to apps, confidence of usage, use case expansion, prompting skill development and safe use of company data.

About the research

How effectiveness is measured. Rather than chase an unmeasurable outcome, NROC Security scores the effectiveness of technology adoption across the whole population and breaks it into factors that can be inspected independently: usage frequency and apps, tasks, skill levels and data. The estimate is deliberately transparent rather than precise. It avoids a single abstract score, measures the behavior of the broader population rather than a handful of power users, and gives leaders a practical instrument for organizational change management – the approach NROC Security labels "productivity-first governance".

From behavior to an GenAI Effectiveness Index. Combining frequency and skill, every employee is placed in one of five effectiveness segments, each carrying a productivity assumption: inactive (0% of working hours), infrequent (0%), weekly but low-skill (5%), high-skill weekly (10%), and high-skill daily (20%). Only the top two segments are assumed to convert GenAI into real productivity. Prompting skill-level is scored against a six-level taxonomy from Level 1 Google search-like questions through structured, contextual, and few-shot prompting up to reusable agentic workflows and self-critique at Levels 5–6.