Department for Business and Trade Copilot

IMPLEMENTED 2024–2025 · GOVERNMENT AND PUBLIC SERVICES · UK · BAND D · PARTLY PROVED

VERSION 1.0 · PUBLISHED 15 NOVEMBER 2026 · UPDATED 15 NOVEMBER 2026 · FACTS CHECKED 15 NOVEMBER 2026

The department's own evaluation of a 1,000-licence pilot found small time savings but no evidence of improved productivity.

ABOUT THIS SYNOPSIS

This is MarLuc's summary and grading of publicly available sources, for information only. It is not investment, legal, financial or professional advice, and should not be relied on for any decision. Readers must validate it against the original sources and, where appropriate, challenge its conclusions.

Facts and figures are as reported by the sources listed below, which may be incomplete, self-reported or out of date. Check the original source before relying on any figure.

The impact band, proof marker and SEAM view are MarLuc's opinion. The organisation is not a MarLuc client and has not reviewed or endorsed this synopsis. Names and trademarks belong to their owners.

NOTICE WORDING LAST UPDATED 16 NOVEMBER 2026

AT A GLANCE

AI used: Generative AI assistant: Microsoft 365 Copilot.

Implemented: 2024–2025. Grade: band D, Task support, Partly proved.

Headline result, as reported: Small time savings and no evidence of improved productivity, by the department's own evaluation

1 · The case

The facts, as reported by the 1 sources listed at the end. Where a source is silent, the field says "Not reported".

Organisation

Department for Business and Trade. UK central government department. One thousand Microsoft 365 Copilot licences were allocated to UK-based staff for the pilot. Canvas block: 5. People and skills. Also 8. Governance and risk, 6. Data and technology.

Business challenge

The challenge. The department wanted to understand the risks and benefits of deploying Microsoft 365 Copilot across the department, and which user groups might benefit most. Its evaluation asked what staff used it for, whether it saved time, whether users kept to acceptable use policies, and how satisfied they were.

How it was addressed. From October to December 2024 the department ran a three-month pilot

Who implemented it

Microsoft (Software vendor): Supplied the assistant.
The pilot was run and evaluated by the department's own teams.

Technology (SEAM pillar E)

SEAM types: T1 Chat assistants, T2 Embedded AI in existing software. Deployment: Vendor SaaS. Vendors and products: Microsoft, Microsoft 365 Copilot, Microsoft Copilot Chat.

Where it was built in: Embedded in Outlook, Teams, Word and Excel, plus a browser chat, for 1,000 pilot users; staff also had Copilot Chat.

What changed

Before: staff drafted, summarised and analysed without an assistant. After: pilot users saved small amounts of time on most tasks, most on drafting documents; scheduling and image generation took longer, and in observed tasks Excel data analysis was slower and of worse quality. [1]

What it cost

Not reported. The department said a value for money assessment was under way and not included in the report.

What it saved

Small time savings on most tasks, by the department's evaluation: the largest was 1.3 hours a task on average for drafting written documents, after adjusting for unused outputs and new tasks. The evaluation did not find evidence that time savings led to improved productivity. Achieved cash saving: Not reported. [1]

Governance and control

The evaluation tested use against acceptable use policies and departmental guidance. It found inconsistent quality assurance of outputs, and 22% of diary respondents said they had identified hallucinations. Colleagues outside the pilot had little knowledge of acceptable use policies. Accountable owner and controls after the pilot: Not reported.

S2 · Accountable owner · Not reported
Not reported.

A1 · Safe use · Claimed
Use was governed by acceptable use policies and departmental guidance; the evaluation found knowledge of them varied, some users had not read the security guidance, and staff outside the pilot knew little of them, per the department's own published evaluation. [1]

A2 · Human in the loop · Claimed
Users were expected to quality assure outputs, but the evaluation found this inconsistent; for scheduling, 36% of reported uses had no quality assurance or none comparable, per the department's own published evaluation. [1]

A3 · Monitoring and assurance · Claimed
The Digital, Data and Technology monitoring and evaluation team used usage dashboards, a diary study, interviews and observed tasks; 22% of diary respondents reported finding hallucinations, per the department's own published evaluation. [1]

A4 · Vendor AI · Claimed
The evaluation states that M365 Copilot is not trained on users' input data and so judged it more suitable for government information than other AI tools; contract terms are not reported. [1]

A5 · Agent controls · Not reported
No agents described.

A6 · Standards and regulation · Not reported
Not reported.

What worked, what did not, what is unproven

WORKED

– 72% of diary respondents were satisfied or very satisfied; the Net Promoter Score was 31.

– Written tasks such as drafting and summarising saved the most time.

– Neurodiverse colleagues were significantly more satisfied than other users.

– At least 2 hours of self-led training was linked to higher satisfaction.

DID NOT

– No evidence that time savings led to improved productivity; colleagues outside the pilot saw no productivity improvement.

– Scheduling and image generation took longer with Copilot.

– In observed tasks, Excel data analysis was slower and of worse quality, and slides were faster but of worse quality.

UNPROVEN

– Value for money, which was still being assessed.

– Environmental impact, for which the department was still developing methods.

– Whether the results hold, given a diary response rate of 32% and a very small observed-task sample.

Timeline

Oct 2024 Pilot starts with 1,000 licences for UK staff [1]

Dec 2024 Pilot licence period ends [1]

Mar 2025 Evaluation period ends [1]

Aug 2025 Evaluation report published [1]

2 · Grade and SEAM assessment

MarLuc's opinion, made on the evidence above. How cases are graded: marluc.co.uk/case-studies/how-we-rate

Level of change and impact

Claimed: Not applicable. Grade: D, Task support. Proof: Partly proved. Level of change: Task support.

Task support (Level 1, band D). Individuals used an assistant within existing tasks; no process or

The SEAM view and lessons learnt

The evaluation looked at people as well as the tool: training, colleagues' attitudes and who benefits most. That makes it more useful than a single time figure. Small savings scattered across tasks did not add up to measurable productivity, which is the usual pattern for task support. For a larger organisation, the transferable steps are a control group, an adjustment for unused outputs and observed tasks to test self-reports.

1. Adjust time savings for unused outputs and for new work the tool creates.
2. Test self-reported savings with observed tasks against a control group.
3. Put self-led training and clear acceptable use guidance in place before scaling.

Evidence

Score 3 of 4. Self-reported: a published evaluation by the department's own analysts; not independent of the department.

Evaluation method. Comparison group, small: a diary study with usage data, interviews and observed tasks, including 11 control-group colleagues; the report describes the design as evidence without a counterfactual.

Status. Completed: the pilot ended in December 2024 and the evaluation was published in August 2025.

3 · Original sources

Read these in full. To challenge a fact or a grade, use the contact form at marluc.co.uk or the link on https://marluc.co.uk/case-studies/department-for-business-and-trade-copilot

[1] Department for Business and Trade evaluation of Microsoft 365 Copilot
https://assets.publishing.service.gov.uk/media/68adbe409e1cebdd2c96a19d/dbt-microsoft-365-copilot-evaluation.pdf

Department for Business and Trade website: https://www.gov.uk/government/organisations/department-for-business-and-trade

Version 1.0 · Published 15 November 2026 · Updated 15 November 2026 · Facts checked against sources 15 November 2026 · Downloaded 26 November 2026. This PDF may be out of date: the latest version, and a way to report an error or a better source, is at https://marluc.co.uk/case-studies/department-for-business-and-trade-copilot

VERSION HISTORY

v1.0 15 November 2026 First published