



Department for
Science, Innovation
& Technology



Government
Digital Service

Microsoft 365 Copilot Experiment: Cross-Government Findings Report

June 2025

Contents

1. Executive Summary	3
1.1. Overview	3
1.2. Key Findings	3
2. Introduction	4
2.1. M365 Copilot introduction	4
2.2. Experiment background	4
2.3. Experiment Objectives	5
3. Methodology	5
3.1. Experiment Design	5
4. Results and findings	7
4.1. Adoption	7
4.2. Time savings	10
4.3. Benefits	14
4.5. User feedback and insights	16
5. Additional considerations	19
6. Conclusions	21

1. Executive Summary

1.1. Overview

As part of the government's initiative to drive artificial intelligence (AI) adoption across the public sector, the Government Digital Service (GDS) is conducting trials into the capabilities of commercial AI tools. GDS organised and ran a cross government trial of M365 Copilot from 30 September 2024 to 31 December 2024, involving 20,000 government employees. A variety of government organisations participated and licenses were distributed in line with public sector employee distribution across departments.

Data collection was centralised through GDS, and quantitative data was collected on adoption rates and M365 Copilot usage. Qualitative data was collected via a user survey, feedback forms, and five focus groups representing different user groups. Focus groups aimed to capture in-depth experiences, while success stories provided additional insights.

This study focused on the impact of AI assistant tools on productivity. It aimed to answer research questions on whether:

- AI tools improve task quality and reduce time spent on routine tasks
- users are more satisfied with their work when using AI tools
- AI assistant tools impact the effort required to complete tasks

1.2. Key Findings

Trial participants saved an average of 26 minutes a day when using M365 Copilot. Results were consistent across grades and professions, with differences observed in how the tool was used and where benefits were realised.

Over 70% of users agreed that M365 Copilot reduced time spent searching for information, performing mundane tasks, and increased time spent on more strategic activities.

Perceived concerns with security and the handling of sensitive data led to reduced benefits in a minority of cases. Limitations were observed when dealing with complex, nuanced, or data-heavy aspects of work.

User sentiment was overwhelmingly positive, with 82% expressing they would not want to return to their pre-Copilot working conditions.

Satisfaction and recommendation scores were strong, with ratings of 7.7 and 8.2 out of 10 respectively. Users not only experienced time savings but also felt the tool

represented a positive step forward for both themselves and their organisations. An additional insight was the positive impact on users with accessibility needs, indicating untapped potential in this area that warrants further exploration.

2. Introduction

2.1. M365 Copilot introduction

M365 Copilot is an advanced AI-powered assistant integrated into Microsoft's suite of productivity tools, designed to enhance user productivity and streamline workflows. Built on large language models (LLMs), M365 Copilot assists users across widely used Microsoft applications, including Word, Excel, PowerPoint, Outlook, and Teams, offering context-aware suggestions and data-driven insights.

This integration allows users to leverage M365 Copilot's capabilities within the interfaces of these tools. Users can interact directly with M365 Copilot's suite of large language models through Copilot Chat, which is a new application that behaves as an AI chatbot, allowing users to ask queries and search information.

M365 Copilot integrates into existing online documents and data storage through Microsoft's OneDrive. This allows the underlying model to retrieve and query existing documentation. The tool follows all enterprise security policies when dealing with an organisation's data. The tool adopts the permissions of the end user and will only retrieve documents that a user could normally access.

The tool uses a method called retrieval augmented generation (RAG) to search and extract relevant information to a user's query. This allows the tools to generate the most relevant and best response.

However, organisations should ensure that current information and knowledge management practices are up to date and followed. M365 Copilot's ability to search and query using a user's access may highlight issues when users have access to files they shouldn't. This feature can be limited by removing search options and access to the internet (which is off by default) when generating responses. Most organisations in the trial did the latter to rely solely on internal data sources.

2.2. Experiment background

The M365 Copilot cross government experiment was conducted over a 3 month period from 30 September 2024 to 31 December 2024. This involved 20,000 government employees being provided a M365 Copilot license, with each organisation committing to at least 1000 licences.

The following organisations agreed to be a part of the M365 Copilot experiment:

- Department for Work and Pensions (DWP)
- His Majesty's Revenue & Customs (HMRC)
- Home Office (HO)
- Ministry of Justice (MoJ)
- Department for Energy Security and Net Zero (DESNZ)
- Department for Science, Innovation and Technology (DSIT)
- Foreign Commonwealth and Development Office (FCDO)
- Department for Environment, Food and Rural Affairs (DEFRA)
- Department for Business and Trade (DBT)
- Welsh Government
- Office for National Statistics (ONS)
- Companies House

This was the largest deployment of M365 Copilot in all organisations, with previous trials and pilots limited to 300 licences at most.

2.3. Experiment Objectives

The objective of this experiment was to understand the value that an AI tool such as M365 Copilot would bring when deployed across a large portion of the UK government. Value was defined as improvements in efficiency, task completion rates, and overall user satisfaction.

3. Methodology

The experiment's methodology was designed to gather comprehensive insights into user experiences, challenges, and the overall effectiveness of these tools.

3.1. Experiment Design

The experiment consisted of multiple trials across the public sector. Organisations committed to data collection and sharing with the Government Digital Service (GDS). In addition, some departments took part in some further data collection and analysis. Usage and adoption metrics were collected from the M365 Copilot

dashboard (which is now called Viva dashboard).

Data collection was handled centrally through GDS, with various form tools being used to collect information. Data is being stored in a central database for further analysis and reporting.

Quantitative data was collected by departments throughout the experiment, including adoption rates and M365 Copilot usage data of each application via the Microsoft inbuilt dashboard. This data was provided to GDS through a forms collection mechanism.

Qualitative data was collected from a user survey towards the conclusion of the experiment, and open ended feedback forms with questions around customer satisfaction.

In addition, five focus groups were conducted during the final phase of the M365 Copilot Experiment, representing five distinct user groups: Finance, HR, Operational Delivery, Digital and Data, and Policy. These sessions were designed to capture in-depth qualitative feedback to uncover real-world stories about users' interactions and experiences with M365 Copilot. Additionally, success stories were collected from participating departments which offered valuable anecdotal insights from users within central government.

Participants were asked to respond to a series of statements based on their general work experience and specific experience of M365 Copilot. Responses were measured using a Likert scale rating from 'Strongly Disagree' to 'Strongly Agree'. The survey covered the following areas:

- familiarity with AI tools and confidence in using new technologies
- overall usage of Microsoft 365 M365 Copilot
- usage across Microsoft 365 applications
- overall effectiveness of Microsoft 365 M365 Copilot
- task-specific effectiveness
- time saved
- impact and feelings on the potential loss of M365 Copilot
- demographic

Due to the complexity of estimating time savings, participants were asked to estimate the average amount of time saved daily by using M365 Copilot, with options ranging from 'none or less than 5 minutes' to 'more than an hour'.

The study primarily focused on impacts on productivity for the most common civil service professions, which include:

- operational delivery, who are defined as front line staff that ensure a department's public operations. These may include public interaction officers and case workers
- data, digital and technology staff who regularly interact with digital tools as part of their job description. These include software engineers, data analysts and product managers
- finance, which are individuals or teams engaged in the financial operations of a department, including accountants and commercial managers
- human resources
- policy staff who engage in the policy making operations of a department, including research and advising on future and current policies
- project delivery, which are individuals with responsibilities that cover project management and delivery of initiatives.

4. Results and findings

The survey was sent out towards the end of the experiment, with 7,115 responses by M365 Copilot users. Respondents covered 14 professions and were equally distributed across age and gender. Anonymised user data for M365 Copilot of 14,500 users was collected. Data was not available, and hence not supplied, for the remaining users.

Active users of M365 Copilot are defined as users with at least one interaction with M365 Copilot in the previous 30 days. The adoption rate is defined as the ratio between active users and number of users with a M365 Copilot licence. Similarly, adoption rate per application is the ratio between active users in that application and total number of users with a M365 Copilot licence.

4.1. Adoption

Change management activities and roll out of licences within the month of October saw an increase to an adoption rate of 83%. Adoption of around 80% was maintained for the remainder of the experiment.

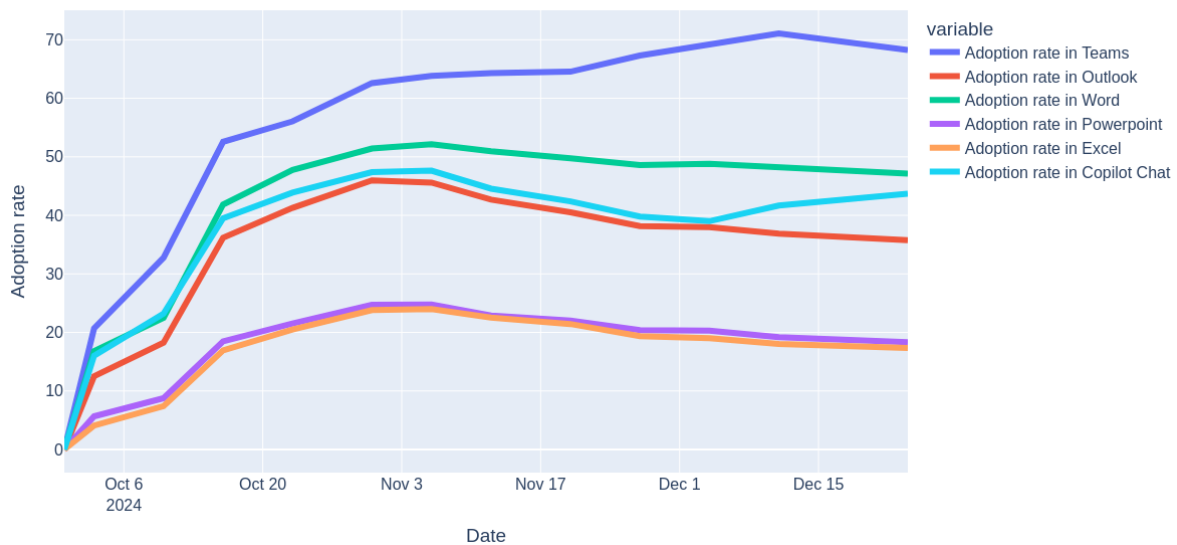
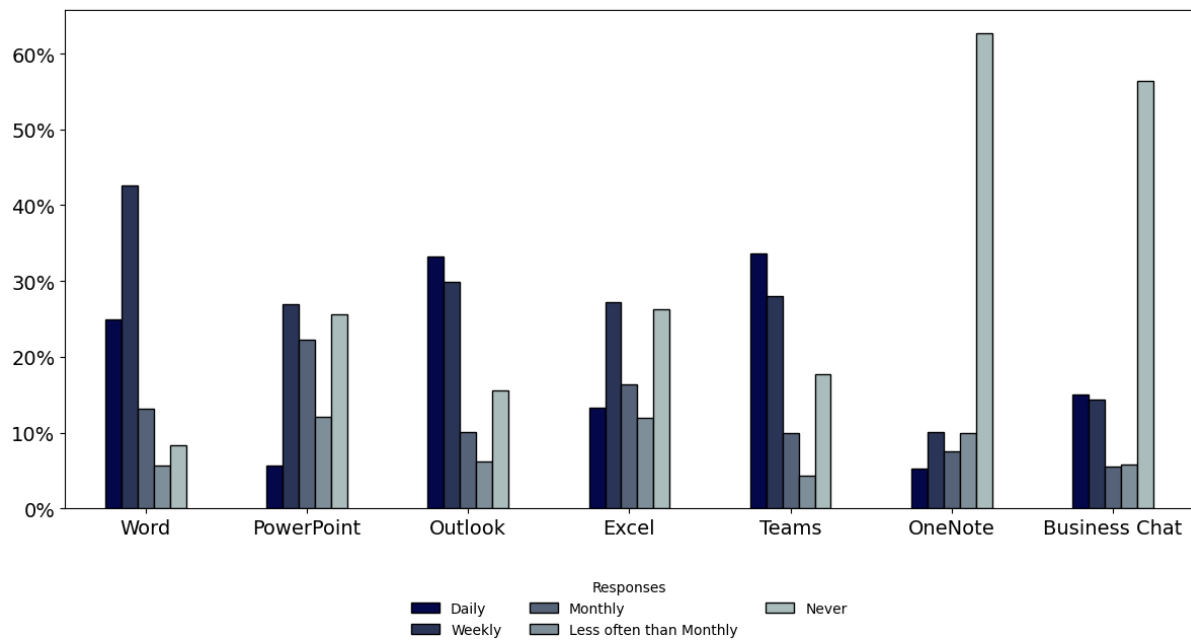


Figure 1. Active usage of M365 Copilot in Microsoft Tools over time represented by adoption rate from October 2024 to December 2024

Teams was the most popular tool for M365 Copilot and remained dominant throughout the experiment with a maximum adoption of 71%. On the other hand, adoption within Excel and Powerpoint remained low, with highest achieved adoption of 23% and 24% respectively. Despite the adoption rate remaining stable, a decrease was observed in the take up of M365 Copilot within Word and Outlook. Adoption of Word dropped by 5% from its peak to the end of the experiment and Outlook dropped by 10% from its peak. Adoption within M365 Copilot Chat recovered from a decline in usage by the experiment's end, with only a 3% dip from its peak.

Survey respondents used M365 Copilot heavily throughout the trial, with 39% of users using it multiple times a day, and a further 43% using it at least multiple times a week. Engagement was high throughout and only 2% of survey respondents did not use the tool at all.



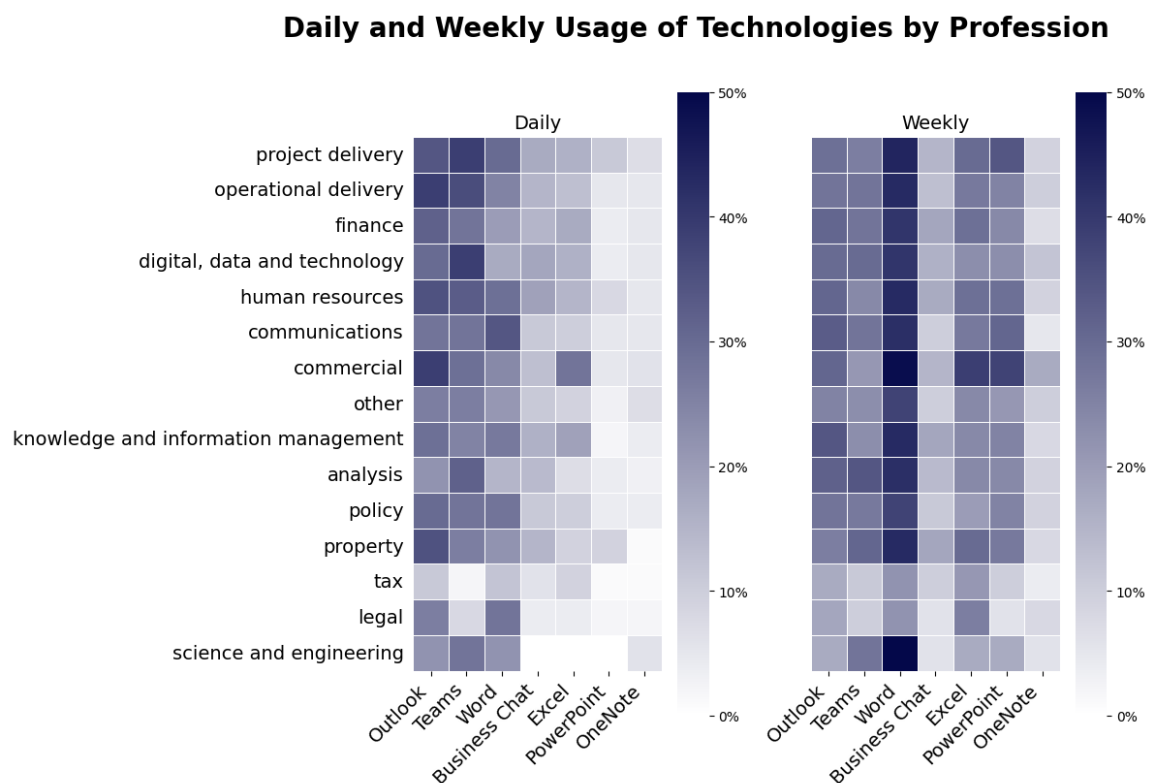
Users were asked 'Approximately how much time, on average, does Copilot save you...' for each activity

Figure 2. Usage of M365 Copilot across Microsoft Tools represented by software type and user-reported time

Survey respondents were asked how often they used Copilot within each Microsoft application. Daily usage was centred around using M365 Copilot for communications, compared to more infrequent weekly use for content creation.

Results showed that:

- 34% of people used M365 Copilot daily in Teams, and 33% used it daily in Outlook
- 25% of people used M365 Copilot daily in Word, with a further 43% using it weekly
- OneNote and Business chat saw limited use, with more than 55% of people not using it at all for these applications



Notes:

- Professions are ordered by average satisfaction scores.
- Percentages are calculated by the number of users that responded 'Daily' or 'Weekly' to 'How often do you use Copilot in Word?'
- Each technology had an identical question.

Figure 3. Daily and weekly usage of technologies by profession represented by software type, type of profession and percentage of usage

The survey responses align with the adoption data obtained from professions. M365 Copilot usage across professions follows similar trends with the majority seeing high levels of daily use for communications through Outlook and Teams, and weekly usage centred around content creation.

13 of the 15 professions in the trial saw more than 25% of respondents use M365 Copilot daily in Teams.

Meanwhile, weekly usage focussed on Word and document drafting. 11 of the 15 professions saw more than 40% of respondents use M365 Copilot weekly in Word.

On the other hand, OneNote and Business Chat saw very little use across all professions.

4.2. Time savings

The time savings presented throughout are self reported through the quantitative portion of the data capture. Time savings of this nature should be considered

alongside additional literature on the true value of these metrics to understand how that numerical value translates into real impact.

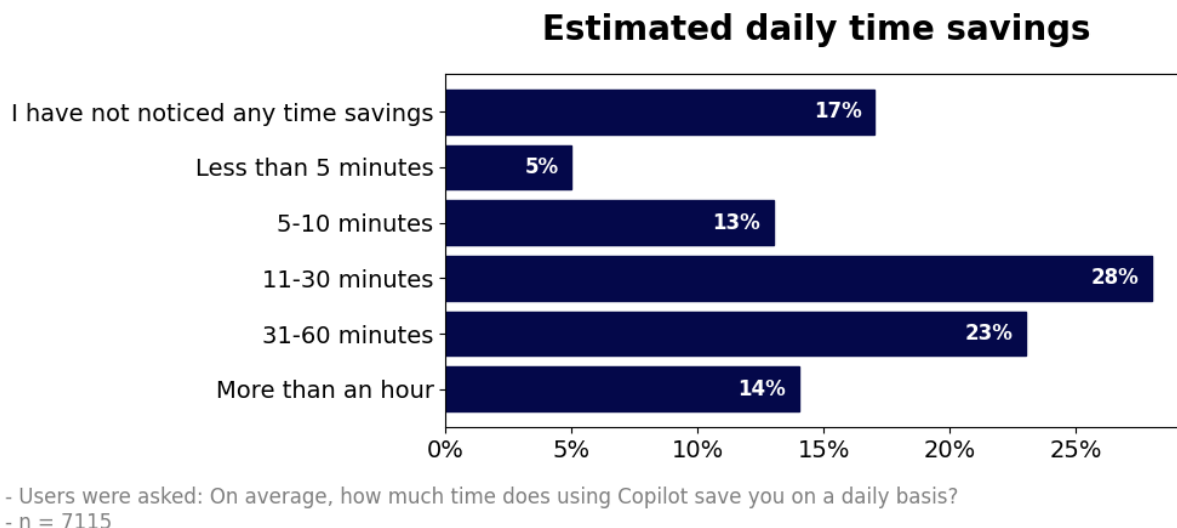
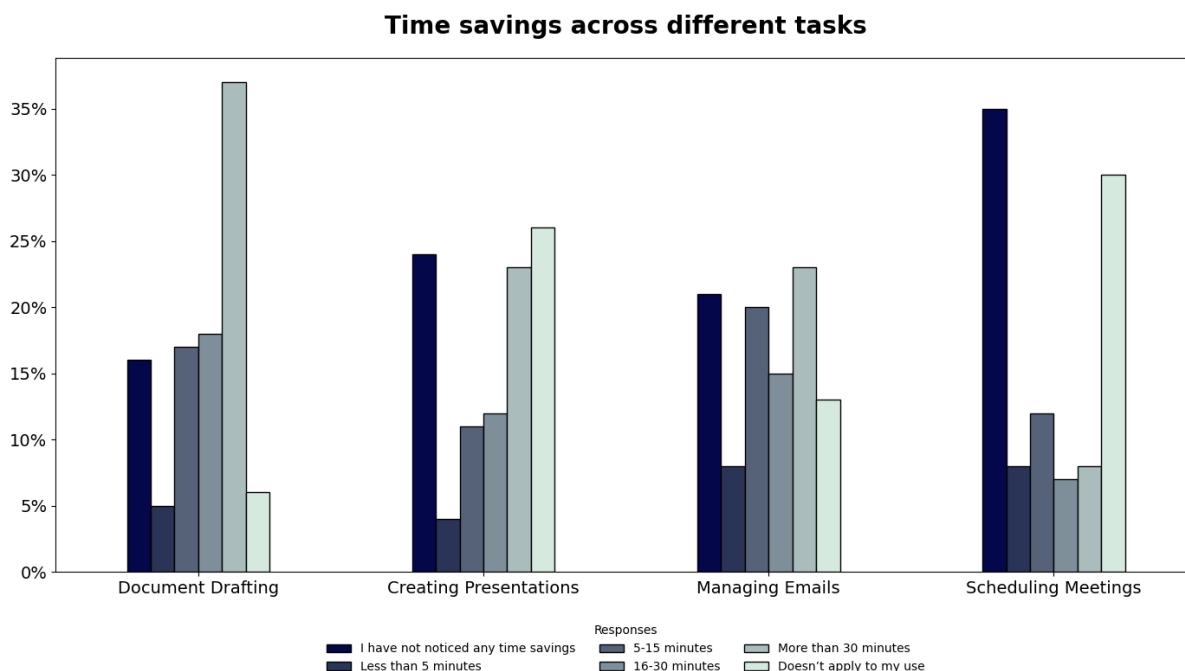


Figure 4. Daily time savings with M365 Copilot represented by user-reported time savings categories and percentage of users reporting each category

Users across the trial experienced significant daily time savings, with only 17% not noticing any clear time savings while using M365 Copilot. More than a third of users saved more than half an hour a day.

On average, users observed a time saving of 26 minutes per day. If this was to be replicated across a full working year, users could save 13 days. This estimate was calculated by using the middle point of each range, with the largest savings being estimated at 60 minutes.



Users were asked 'Approximately how much time, on average, does Copilot save you...' for each activity.
 Document Drafting: n = 5048; Creating Presentations: n = 4864; Managing Emails: n = 4921; Scheduling Meetings: n = 4749

Figure 5. Time savings with M365 Copilot across tasks represented by type of task and user-reported time savings category

For specific tasks, M365 Copilot achieved the most significant time savings with content creation, and most often with drafting documents with savings of 24 minutes. 19 minutes of time savings were observed for creating presentations and as such, there is potential for the aggregate time savings to be greater. The low adoption of M365 Copilot within presentations (see Section 4.1) indicates that with greater effort and upskilling, there is potential that more savings could be made.

Whilst much lower in terms of time saved, a high proportion of users experienced regular small time savings in communication tasks, such as scheduling meetings, with savings of 9 minutes. From Section 4.1, these tasks were done much more regularly.

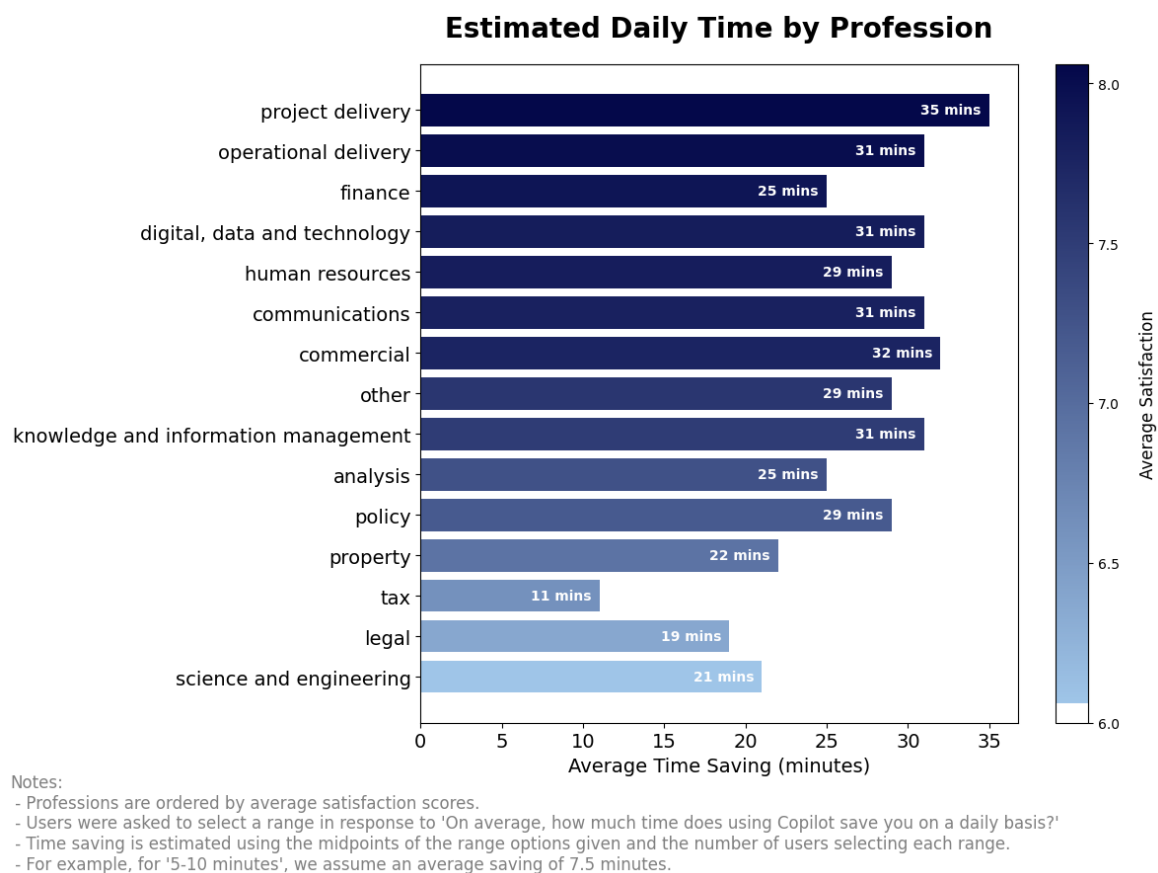


Figure 6. Estimated daily time by profession represented by time saved in minutes across professions and user-reported satisfaction score

Average time savings across the 15 professions in the trial were broadly consistent, with the majority of professions saving at least 2 minutes of the 26 minute average observed. 8 of the 15 professions in the trial saved at least 26 minutes per day on average.

Professions with the lowest satisfaction scores were also those that saved the least amount of time. Use of the tool within these roles were impacted by a range of factors. For example, policy-focused teams struggled with nuance and when data sources contained multiple contrasting opinions. Concerns were raised on the appropriateness of using M365 Copilot and therefore some users were cautious in its usage.

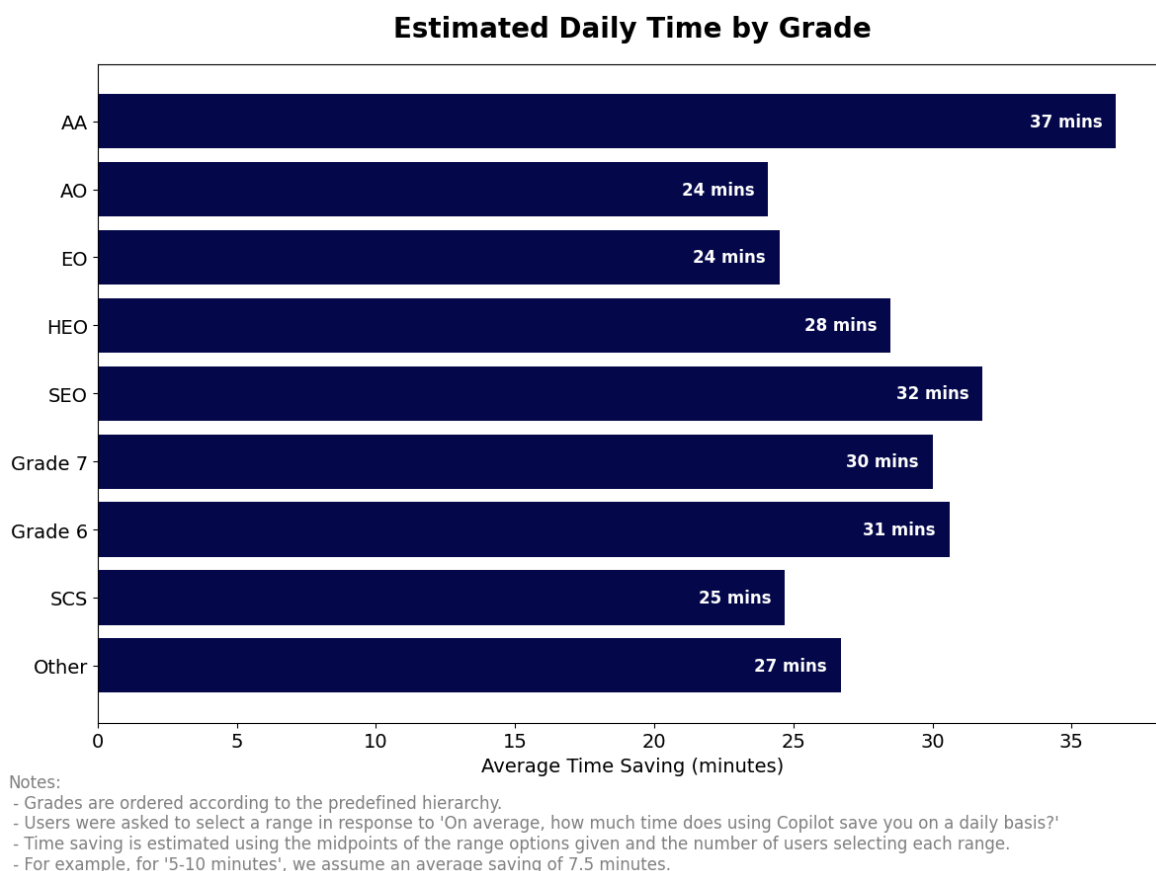
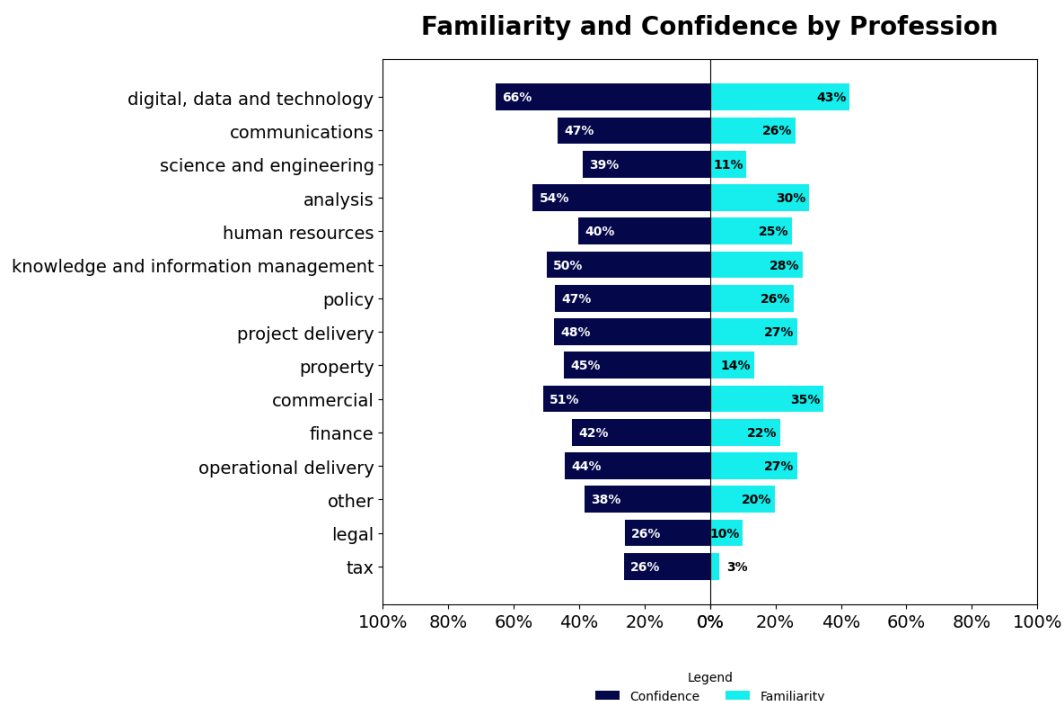


Figure 7. Estimated daily time savings by grade represented by average time saved in minutes against grade categories

Daily time savings were relatively consistent across grades, with the exception of staff at AA grades (with only 14 responses, or less than 1%). SEO to Grade 6 all saved more than 29 minutes on average. Despite the consistency in time saved, the trial observed differences in how grades perceived M365 Copilot to be most effective.

For the majority of grades (AO to Grade 7), M365 Copilot was most effective at saving time when drafting documents. For Grade 6 and SCS, responses indicated the highest effective savings were in improving collaboration in Teams. This is representative of the differing types of tasks undertaken by the more senior grades.



Notes:

- Bars are ordered by % of respondents that have never used AI tools before (high to low).
- Confidence is the count 'Extremely confident' or 'Very confident' responses to 'How would you rate your confidence in learning and using new technologies?'
- Familiarity is the count 'Extremely familiar' or 'Very familiar' responses to 'How familiar are you with AI-powered tools like Microsoft 365 Copilot?'

Figure 8. Familiarity and confidence by profession and department represented by user-reported confidence and familiarity percentages against professions

Professions and departments with the lowest familiarity and confidence in AI tools saw lower benefits and time savings from M365 Copilot. 11 professions observed that at least 20% of their users were very or extremely familiar with AI tools.

4.3. Benefits

The significant time savings with M365 Copilot meant it was scored highly across the board and users reported high satisfaction scores. Additionally:

- 85% of users agreed that M365 Copilot provided good value to the organisation
- 63% of users agreed that having access to M365 Copilot influenced satisfaction with their employer
- 63% of users believed their productivity would decline without access to M365 Copilot.

The most common benefits users highlighted from working with M365 Copilot were around increasing productivity and reducing time searching for information.

Many users also signalled benefits from changes in how they spend their time, with productivity gains allowing users to spend time on more strategic and satisfying

tasks. Overall, users had an average satisfaction score of 7.7 out of 10, and an average recommendation score of 8.2 out of 10.

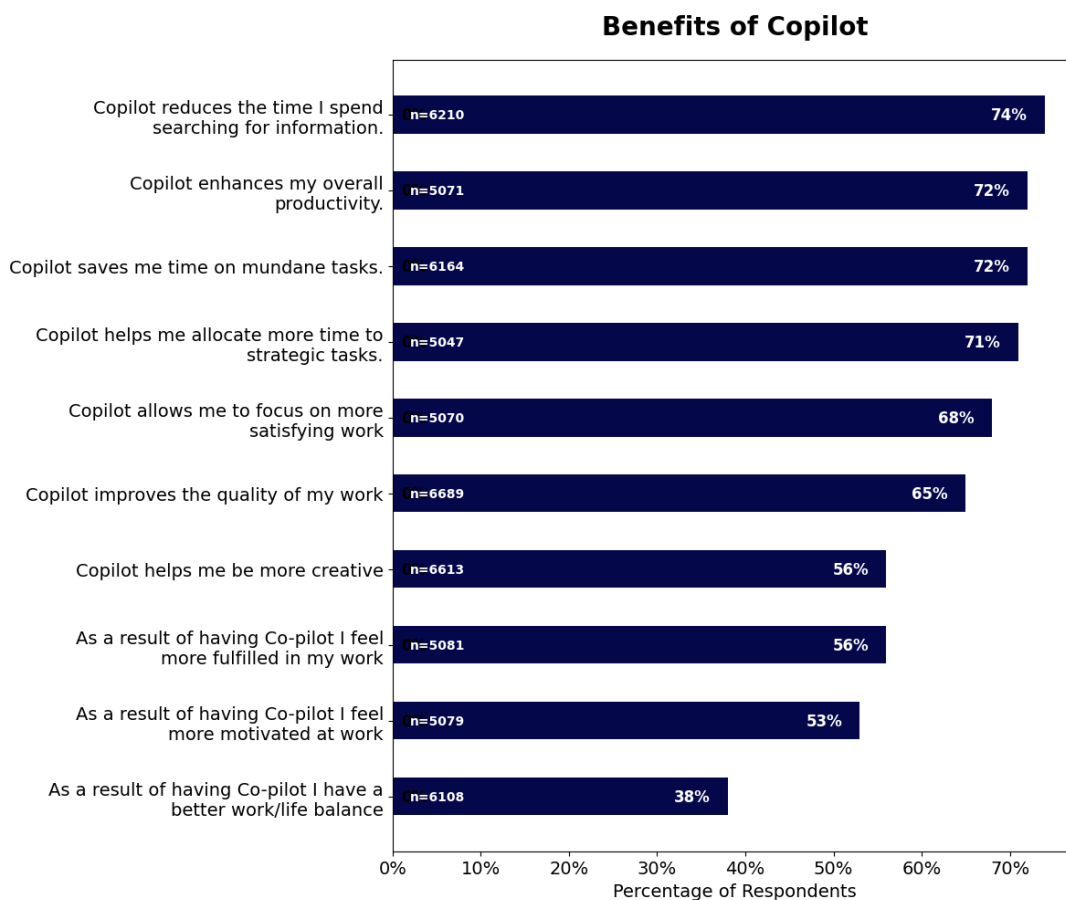
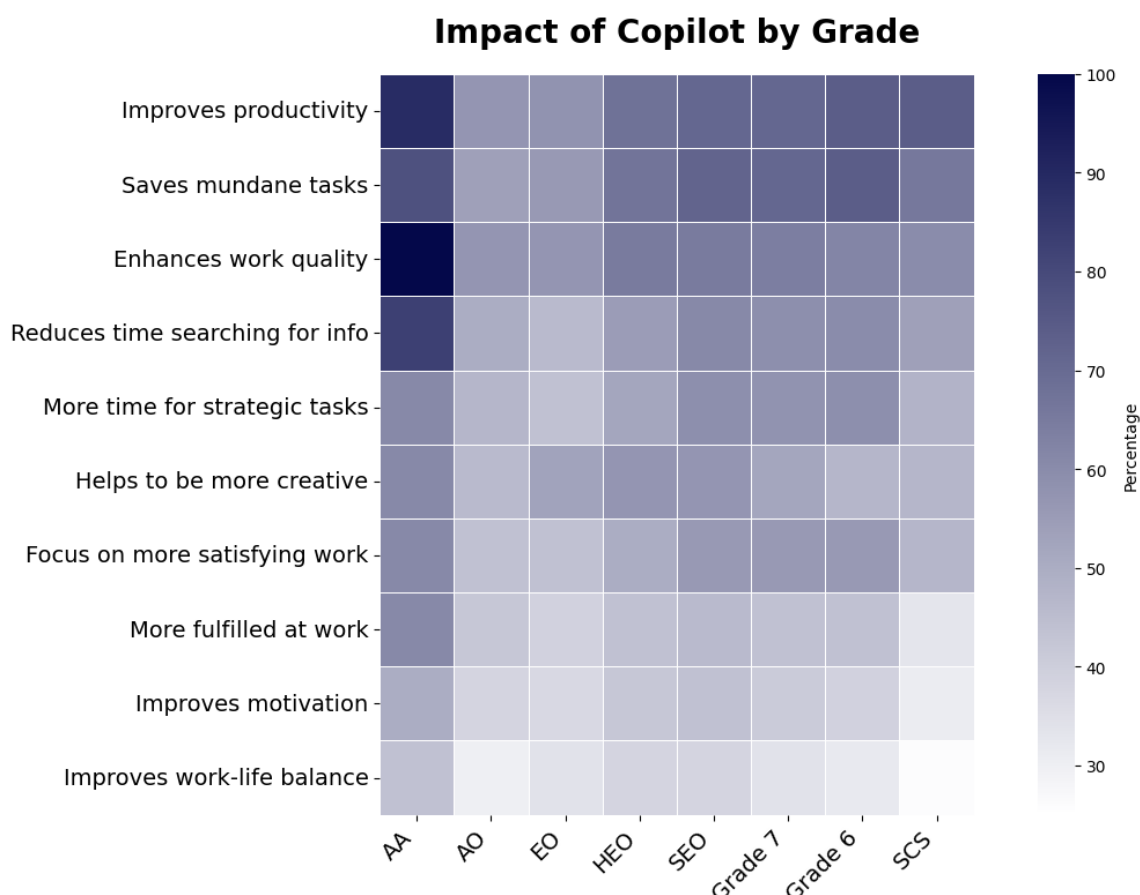


Figure 9. Benefits of M365 Copilot represented by user-reported benefit against percentage of users recording this benefit

The most common benefits across all professions were similar. For 10 of the 15 professions, 'improves productivity' was the highest scored benefit. For the remaining professions, 'saves mundane tasks' was the highest scored benefit.

In professions with the highest satisfaction scores, which were project delivery and operational delivery roles, around 75% of users agreed that M365 Copilot led to improved productivity and a reduction in time spent on mundane tasks.



Notes:

- Percentages are the proportion of users that responded 'Agree' or 'Strongly Agree' for each question.
- Questions are listed in order of average percentage that 'Agree' or 'Strongly Agree' across all grades

Figure 10. Benefits of M365 Copilot by grade represented by user-reported impact against percentage of users strongly agreeing with each impact against their grade

Across all grades, 'improves productivity' was either the first or second most agreed with benefit, with all grades agreeing that M365 Copilot allowed less time to be spent on mundane tasks.

M365 Copilot benefited the middle grades most. The average 'agree' responses across the listed benefits was around 55% in HEO to Grade 6. For AO, EO and SCS, this fell to about 45%.

Personal impacts, such as fulfillment, motivation and work-life balance, were significantly higher for lower grades. These benefits are also considerably fewer for SCS compared to all other grades.

4.5. User feedback and insights

Survey respondents were given a chance to highlight their own feedback to support findings. This user feedback suggested that productivity gains could be higher as

familiarity with the tool increased. The top 5 themes raised through this feedback were on:

- accessibility and inclusivity, where users agreed that M365 Copilot significantly aided individuals with disabilities. This has been observed from the early access programme and similar trials across the public sector
- efficiency and time saving, where users noted its ability to reduce effort on mundane and repetitive tasks
- quality of work, where users observed improvement to their work quality
- learning curves and usability where, outside of the benefits, many users mentioned the learning curve involved. Feedback indicated a need to learn how to effectively prompt and change usage patterns to get the most out of M365 Copilot
- reliability and accuracy, where some users expressed concerns on M365 Copilot's potential for generating incorrect information. There were also worries about dependency on AI for tasks that involve critical thinking and creativity, as well as the environmental impact of using AI technologies

Benefits:

There has been a clear trend of M365 Copilot being hugely beneficial for reducing effort on routine administrative tasks, such as scheduling, reporting, and documentation drafting. There was a significant indication for this benefit in both HR and policy teams.

Some examples of additional feedback from experiment participants in these roles include:

"I work in IDD/CoE and I've been using M365 Copilot to do most things from communication writing, proofreading, creating presentations, drafting documents, creating images"

"Whether I'm drafting communications, summarising meeting notes, or creating PowerPoint presentations to showcase the benefits of MS Teams and SharePoint, M365 Copilot has consistently proven to be incredibly helpful"

M365 Copilot allowed enhanced data analysis and summarisation by providing actionable insights and trends. There was an especially significant trend within the Finance profession. Some testimonials to support this include:

"I have found M365 Copilot to be an incredibly versatile tool, especially given my role.... M365 Copilot has been instrumental in ensuring that I pick up all the action

points in email chains and in analysing documents to double-check my understanding. Additionally, it provides summary overviews of unfamiliar technical subjects that arise during my work."

"M365 Copilot has added value through its ability to analyse data quickly, provide referenced viewpoints, and enhance my work visually. Overall, it has made my work more enjoyable, productive, efficient, and visually powerful."

Accessibility has consistently been raised by participating departments as a key area of benefit both before and during the experiment, which helped to navigate difficulties in various tools. Many departments have focused on these areas for previous and further M365 Copilot exploration following testimonials such as:

"M365 Copilot has added value through supporting me as a dyslexic and making my tone not only more refined but more sophisticated. I have always struggled with adapting my tone on paper and have felt more comfortable in conversations".

"I have dyspraxia and M365 Copilot has been brilliant. Emails are so much easier when using M365 Copilot. It has saved me so much time and effort in creating work"

Concerns:

One issue discussed across the focus groups was M365 Copilot's handling of highly nuanced or context-specific data. A policy participant expressed the following frustration:

"M365 Copilot's ability to extract key themes and insights from documents is strong, but it struggles with nuanced or context-heavy data requiring human judgement."

Another prevalent concern was M365 Copilot's reliance on its own training and external sources to process and compute requests. A finance participant shared their perspective:

"While M365 Copilot efficiently generates initial summaries and reports, it struggles with complex data requiring contextual input and often relies on external data sources without built-in verification."

These points underscore the ongoing challenges in fully optimising M365 Copilot for tasks requiring deep contextual understanding and verification. Furthermore, exploration of M365 Copilot agents by participants showed that these agents struggled to identify the exact documents used to generate responses when using a OneDrive folder as the source.

Concerns were raised on the use of M365 Copilot in tasks requiring deep domain knowledge, leading to questions about the accuracy of its outputs. An HR participant shared the following:

"In sensitive areas like grievance handling or performance evaluations, there is significant concern about the accuracy of M365 Copilot's outputs, as any errors could lead to reputational risks."

It was also emphasised that M365 Copilot, as an assistive tool, still requires human input and decision-making, creating an ongoing need for human oversight. This was reflected by a Project Delivery participant when discussing M365 Copilot's role in critical decision-making:

"For risk mitigation, M365 Copilot can identify trends, but decisions about which risks to escalate or how to approach them still require deep understanding and human judgment."

5. Additional considerations

There are certain considerations to keep in mind regarding this experiment's background and organisation. However, these factors have not changed the overall outcome and positive findings about M365 Copilot's value.

Inconsistent user experience

The experiment was marketed as a finite period for engagement with M365 Copilot, which may have limited adoption for some users hesitant to commit due to access being removed after the trial. Despite this, the clear messaging led to significant engagement and sufficient data.

Driving adoption required substantial effort and, given the scale, centralised management was not feasible, leaving the responsibility for user support to individual departments. Resources were supplied centrally from GDS and organisations were free to explore their own priorities and make changes as required. As a result, user experience was not consistent across the experiment group.

User capacity

Some users noted that their work capacity hindered full participation in the upskilling and exposure required by the experiment, as they were full-time civil servants with existing responsibilities. To address this, measures such as bite sized training videos were implemented to streamline engagement.

While this approach helped minimise disruption, some users may still have struggled to commit fully, which potentially affected adoption and success metrics.

However, this challenge reflects the real-world need to manage capacity during a live rollout.

Experiment timing

A limitation of the experiment was the timeframe, with the final month being disrupted by the festive period, leading to reduced availability of both users and departmental resources. For some, the original three-month trial effectively became shorter, which is reflected in the usage statistics. Internal priorities also delayed the full distribution of some licences until the end of the first month. Both factors potentially could have impacted adoption.

Growing functionality of M365 Copilot

Throughout the M365 Copilot experiment, several enhancements were made to the tool which added new capabilities to its existing functionality. One notable development that sparked significant interest among departments was M365 Copilot Agents.

These tailored AI assistants enhance the performance of Microsoft 365 M365 Copilot by connecting to your organisation's data and knowledge resources. However, as these features were introduced partway through the experiment, they were not included in the scope of research or analysis. However, there was strong positive feedback surrounding M365 Copilot Agents, with many departments eager to explore the tool further.

Changes to M365 Copilot's branding were also made, including the renaming of the M365 Copilot app to M365 Copilot Chat. While such changes are anticipated in emerging technologies, some participating departments reported challenges in managing the volume and dynamic nature of information from the tool and its provider. For organisations and users encountering AI for the first time, these ongoing adjustments can disrupt adoption processes. It is important to consider that these factors may have influenced participants' adoption rates and their overall experience with the tool.

6. Conclusions

M365 Copilot displays potential for significant time savings. Users reported an average of 26 minutes saved per day. However, due to experimental constraints it was not possible to identify how time saved was spent. However, users reported on saving time on mundane tasks and spending more time on strategic workflows.

Qualitative research and focus groups highlighted that M365 Copilot demonstrated value in reducing time spent on administrative tasks, and allowing users to focus on more meaningful work. On the other hand, Copilot showed limitations when dealing with complex data, especially for work requiring nuance and context. Human oversight was required at all times to ensure maximum benefits and reduce risks in its use.

Finally, resources were required to drive engagement and adoption of the tool. As a result, approximately 80% of users actively engaged with Copilot and training was essential to achieve benefits of the tool.

A correlation between a user's confidence and familiarity with AI and time savings was found, highlighting the need for well thought out change management programmes, including communications, engagement and training.