



AI copilots for knowledge workers

Introducing role-specific assistance that changes work without creating unmanaged shadow use.

WORKS WITH THE ASSISTANTS TEAMS ALREADY USE



EDITION 17 / EMBER & TEAL

EXECUTIVE WORKING EDITION - 08 OCT 2026

17

AT A GLANCE

Introducing role-specific assistance that changes work without creating unmanaged shadow use.



AI copilots draft, summarize, analyze and search alongside people in writing, meetings, spreadsheets, research and internal communication. The value is real but uneven: it depends on the role, the task, the quality of the surrounding data and how well people are trained. This playbook shows how to choose roles and tasks, roll out an approved copilot, set rules for acceptable use, measure time and quality gains honestly and avoid both shadow AI and shelfware.

WHO IT IS FOR Department heads, HR and learning leaders, IT and operations leaders and team managers responsible for rolling out AI assistants to office-based teams.

TIME TO USE 70 minutes to read; one workshop per role group to apply



What you will have at the end

- A shortlist of roles and tasks where a copilot is likely to pay off, and ones where it is not
- A task-level baseline method for time and quality, so gains can be shown rather than assumed
- Acceptable-use rules and review habits that staff can follow
- A staged rollout plan with training, champions and feedback loops
- A way to spot and respond to shadow AI without punishing curiosity



Contents

01	Copilots change tasks first and jobs later	03
02	Tasks and roles where a copilot helps, and where it does not	04
03	How copilots produce answers, and why that shapes good use	05
04	Ownership, champions and the rules of use	07
05	Risks to manage in everyday use	08
06	Baselines, metrics and honest results	10
07	Rolling out a copilot in stages	11
08	Choosing a copilot without being dazzled by a demo	13
09	Common failures and an illustrative scenario	14
10	A first quarter that produces evidence	16
11	Terms used in this publication	17



EXECUTIVE BRIEF

Copilots change tasks first and jobs later

01



A copilot does not replace a role; it changes how specific tasks within a role are done, and results depend on which tasks and who uses it.

A copilot is an AI assistant embedded in the tools people already use or offered as a general workspace assistant. It helps draft text, summarize long material, answer questions over documents, transform data, prepare meetings and suggest next steps. It works best as a fast first-draft partner for people who can judge the result. It works poorly when users cannot tell good from plausible, or when the task needs facts the tool has no access to.



Leaders often ask for a headline productivity figure. Be careful: published claims vary widely by task, role and study design, and your own mix of work will differ. The honest approach is to measure a few high-volume tasks in your own teams before and after, and to judge quality as well as speed. A faster email that must be rewritten is not a gain.

- Value concentrates in repeatable text and analysis tasks: first drafts, summaries, reformatting, searching, brainstorming options.
- Adoption is a change program. A license without training and manager support often means low use.
- If you do not provide an approved tool that is good enough, people will bring their own.
- Quality and confidentiality rules must be simple enough to remember under deadline.



POSITIONING

Treat the copilot as a capable but fallible junior colleague: quick, tireless, occasionally confidently wrong. The person who sends the work owns it.



WORKING QUESTION

Which three recurring tasks in your team consume the most hours of drafting, summarizing or searching?

WHERE IT FITS

Tasks and roles where a copilot helps, and where it does not

02

Choose tasks by volume, tolerance for error and ease of checking, not by how impressive the demo looks.

Evaluate at the task level. A single job may include tasks that suit a copilot well and others that do not. For each task ask three questions: how often does it occur, how costly is a mistake, and how easily can the user verify the output? High frequency, moderate cost of error and easy verification make the best candidates.

TASK	FIT	WHY
Summarizing a long meeting transcript into actions	Strong	Easy to check against the notes; high frequency
First draft of a routine customer or internal email	Strong	Low risk; user edits tone and facts
Reformatting or restructuring existing documents	Strong	Source is in front of the user
Spreadsheet formulas, cleanup and explanation of data	Good with checks	Verify results on sample rows
Research summaries from open sources	Good with checks	Verify claims and citations before use

TASK	FIT	WHY
Legal, medical or financial conclusions	Weak alone	Needs expert judgment and accountability
Anything needing facts the tool cannot see	Weak	Without access it guesses



Role examples

- Account managers: meeting prep, follow-up notes, proposal first drafts from approved templates.
- Finance analysts: commentary drafts on variances, formula help, explaining models to colleagues, with figures verified from source.
- HR and recruiters: job description drafts, policy question answers from approved documents, with care over any candidate evaluation, which is a high-risk area.
- Project managers: status reports, risk logs, meeting summaries, plan comparisons.
- Support leads: knowledge article drafts, reply suggestions that an agent reviews.



WORKING QUESTION

For your target role, which task would you be happiest to have a first draft of by Monday morning, and how would the person check it?

HOW IT WORKS

How copilots produce answers, and why that shapes good use

03



Understanding three facts about how copilots work prevents most disappointment.



First, a model generates likely text from the instructions and context it is given. It does not look up truth. When the needed facts are in the context, such as a pasted document or a connected file, output is far more reliable. When they are not, the model fills gaps with plausible text. Second, the output depends heavily on the quality of the request. Specific context, the

audience, the format and examples improve results. Third, results vary between runs, so a person must review each output they rely on.

A reliable way to ask

01 State the role and goal

Say what you need and who it is for, for example a summary for a finance director who has two minutes.

02 Supply the source

Provide or connect the actual documents, data or notes. Say to use only these sources when accuracy matters.

03 Set the format and limits

Specify length, structure, tone and what to avoid. Ask it to flag anything it could not find or is unsure about.

04 Review and iterate

Check facts, numbers and names against the source. Ask for revisions, and keep the version that you can defend.

NEVER COPY BLIND

Numbers, names, quotes, dates, citations and legal or policy statements must be checked against the source every time. Fabricated references and plausible but wrong figures are a known failure of these tools.

Connected copilots that can read your mailbox, files and calendar are powerful because they work from your real context. They also inherit your permissions, which makes document access hygiene important: if files are overshared, the copilot will surface them. Review sharing settings before enabling broad access, and check what the vendor stores and whether it uses your data for training under your contract.

WORKING QUESTION

Which of your team's current prompts include the source material, and which rely on the model to know?

OPERATING MODEL

Ownership, champions and the rules of use

04



A copilot rollout needs an owner, a small network of champions and rules short enough to remember.

Give the program a single accountable owner, often in operations, IT or a transformation function, with an executive sponsor who uses the tool visibly. Appoint champions in each team: respected practitioners who share what works, collect issues and help colleagues. Champions are more credible than central trainers because they know the real work.



ROLE	RESPONSIBILITY	TIME COMMITMENT (ILLUSTRATIVE)
Executive sponsor	Visible use, removes blockers, reviews results	Brief monthly review
Program owner	Plan, licenses, training, metrics, vendor contact	Part to full time depending on scale
Team champions	Peer coaching, prompt sharing, feedback collection	A few hours per week at start
Managers	Set expectations, allow time to learn, review quality	Embedded in normal management
IT and security	Access, permissions, data protection, monitoring	Setup plus ongoing review



Rules of use that people can follow

- Use only the approved copilot for company information. Do not paste confidential or personal data into other tools.
- You are accountable for anything you send, whatever helped you write it.
- Verify facts, figures and citations before use.
- Disclose AI assistance where customers, regulators or colleagues would reasonably expect to know, and where policy requires it.
- Do not use the copilot for decisions about individuals, such as hiring or performance, without the specific approved process.



- Report errors and near misses without blame.

GOVERNANCE LINK

Register the copilot and its connected data sources in your AI inventory, and apply your standard review for vendor terms, data handling and permissions. Check current legal and contractual requirements.

WORKING QUESTION

If someone asked your staff what they may and may not paste into the copilot, would they answer the same way?

RISKS AND CONTROLS

Risks to manage in everyday use

05

Most copilot risk is ordinary: wrong content sent confidently, data in the wrong place and skills that fade.

RISK	HOW IT SHOWS UP	CONTROL
Plausible errors	Wrong figure or invented citation in a client document	Mandatory source check; sampled quality review
Data leakage	Confidential text pasted into a personal chatbot	Approved tool, clear data rule, monitoring of unapproved use
Over-sharing through connected access	Copilot surfaces a restricted file to the wrong user	Permissions clean-up before enabling; sensitivity labels
Homogenized output	Everyone's writing sounds the same, losing voice	Style guidance; human editing; encourage original thinking
Skill erosion for juniors	New staff skip learning fundamentals	Pair with review; train judgment; rotate tasks

RISK	HOW IT SHOWS UP	CONTROL
Unclear accountability	Blame placed on the tool	Rule that the sender owns the output

Pay particular attention to meeting recording and transcription features. Participants may need to be told, and laws on consent differ by place. Decide when recording and summaries are allowed, who can access them and how long they are kept. Likewise, the use of copilots on employee data, performance or hiring raises fairness and legal questions that need approval before use, not afterward.

Before enabling a copilot team-wide

- Document sharing and permissions reviewed for the data it can read
- Vendor terms checked for training use, retention and data location
- Acceptable-use rules published and acknowledged
- Recording and transcription rules agreed with legal or privacy
- Training delivered, with real examples from the team's own work
- A way to report errors and a named owner to receive them

WORKING QUESTION

What is the first piece of restricted content a connected copilot might show to someone who should not see it?

MEASURING VALUE

Baselines, metrics and honest results

06

Measure a few tasks before and after, include quality, and be careful with self-reported savings.

Self-reported time savings are a start but tend to be generous and uneven. Combine three sources: system usage data, task-level timing on a small sample, and quality review. Choose three to five tasks per role and record the baseline for time, effort and quality over a couple of weeks before the rollout. Then repeat after training, with the same tasks and similar volumes.

METRIC	WHAT IT SHOWS	HOW TO MEASURE
Active use	Whether the tool is part of the workflow	Weekly active users divided by licensed users, by team
Task time	Speed on defined tasks	Timed sample of tasks, before and after
Quality score	Whether output is as good or better	Blind review of samples using a rubric
Rework rate	Hidden cost of fixing poor drafts	Share of AI-assisted outputs substantially rewritten
Reuse of good prompts	Maturity of practice	Shared library entries used by others
User confidence and trust	Calibration of reliance	Short survey on checking habits and perceived reliability
Cost per active user	Value for money	License and support cost divided by active users

ILLUSTRATIVE ARITHMETIC (NOT A BENCHMARK)

If a task takes 40 minutes now, and a measured sample after training takes 30 with equal quality, the gain is 10 minutes per task. Multiply by actual weekly volume, then discount for the share of people who really use it.



Decide up front what happens with saved time. If it vanishes into more email, you will not see value on the P&L. Name where freed capacity goes: more client time, faster turnaround, lower overtime or work not previously possible. Also watch the quality side: teams that are measured only on speed will trade quality for it.

WORKING QUESTION

For your pilot team, which three tasks will you time and score before anyone gets a license?



IMPLEMENTATION GUIDE

Rolling out a copilot in stages

07



Start with a motivated group, learn quickly, then widen with training and clear rules.

01 Pick roles and tasks

Select two or three roles with high-volume drafting, summarizing or analysis tasks. Record task baselines.

02 Prepare the ground

Review permissions on the content the copilot can read, finalize the rules of use and confirm vendor terms.





03 Run a pilot

Give licenses to a motivated group with champions. Hold weekly short sessions to share prompts, failures and findings.

04 Train by role

Teach with real tasks from the team, covering how to ask, how to check, what never to paste and when not to use it.

05 Decide and widen

Review the measured results against the baseline and decide who gets access next, what to change and what to stop.

06 Sustain

Keep a living prompt and example library, refresh training when features change and reclaim unused licenses.



Handling shadow AI

- Ask people what they already use and why. Their answers reveal the unmet need.
- Offer an approved option that is at least as convenient before restricting alternatives.
- Use network and spend data to see the extent, and treat it as information first, not a disciplinary list.
- Explain the reason for the rule in terms of client and employee data, not abstract policy.
- Keep the request path for new tools short, with a stated response time.



TRAINING TIP

Train on the team's own documents and tasks. Generic prompt tips fade quickly, while a worked example from this month's real report sticks.

WORKING QUESTION

Which team already has an informal AI habit you could bring into the open and support?

BUYING AND VENDOR QUESTIONS

Choosing a copilot without being dazzled by a demo

08

Compare tools on your own tasks, data and controls, and make the vendor show evidence on each.

Most copilots look excellent in a scripted demo. Run a structured trial instead. Give two or three candidate tools the same set of real tasks from your pilot roles, with realistic documents, and have the people who do the work score the results with a shared rubric. Include your hardest examples: messy notes, long threads, conflicting source files. Record not only quality but how much editing was needed and how easily users found the features.

AREA	WHAT TO ASK	WHY IT MATTERS
Data handling	Is our content used for training? Confidentiality and contract Where is it stored and for how long?	Compliance
Permissions	Does the tool respect existing document permissions per user?	Prevents surfacing restricted files
Fit with existing tools	Does it work inside the application our people already use?	Adoption depends on convenience
Admin controls	Can we manage users, features, logging and usage reports?	Governance and measurement
Quality on our tasks	How did it score on our trial tasks versus alternatives?	Evidence, not claims
Change management	How are feature and model changes announced?	Training and behavior can shift under you
Commercial terms	What are licensing terms, minimums and exit rights?	Avoid paying for unused seats and lock-in

Trial scorecard

- Same tasks and source material given to every candidate
- Scored blind by the people who do the work, using one rubric
- Editing time and error count recorded for each output

- Security and privacy review completed on the shortlisted tool
- Usage reporting confirmed so adoption can be tracked
- Exit terms and data export reviewed before signing


LICENSING CAUTION

Buy in stages tied to evidence. Contract for the pilot group first, with the option to expand, rather than committing to company-wide seats before you know who will use them.


WORKING QUESTION

Which five real tasks would you give each vendor in a head-to-head trial, and who would score them?

PITFALLS

Common failures and an illustrative scenario

09



Copilot programs usually stall from weak training and unclear rules, rarely from the technology.

- Buying licenses for everyone and measuring nothing.
- Training on features instead of tasks.
- No manager involvement, so use depends on individual enthusiasm.
- Enabling broad access to content before cleaning up permissions.
- Counting saved minutes without checking quality or where the time went.
- Banning outside tools without offering a good approved one.



ILLUSTRATIVE SCENARIO (HYPOTHETICAL)

A professional services firm buys 500 licenses. After three months, usage data shows a third of licensees active. Meanwhile, staff use personal chatbots for client summaries. The firm narrows to four role groups, adds champions, trains on real client tasks, sets a data rule and measures drafting time and rework. Active use and quality both improve, and personal tool use drops.



The case shows that distribution of a tool is not adoption of a practice. Adoption grows where managers expect use, champions help and the tool is tied to a task someone already wants to do faster. Treat the first 90 days as a learning period with a decision at the end, and do not expand licenses until the evidence supports it.

Also plan for the second-order effects. When drafting becomes cheap, the volume of documents and emails can rise, and so can the effort required to read them. Set norms for concise communication and for sharing the source rather than a long AI-generated summary of it. Without these norms, the organization risks producing more text with less meaning.

**WORKING QUESTION**

Six months from now, what evidence would make you expand this program, and what would make you stop?



30 / 60 / 90



A first quarter that produces evidence

10

The first quarter should produce a measured result for two or three roles and a clear decision about who goes next.



First 30 days: ground it

- Select two or three pilot roles and their top tasks
- Record task baselines for time and quality
- Review document permissions and finalize rules of use
- Name the program owner, sponsor and champions

Days 31 to 60: test it

- Launch the pilot with licenses and weekly champion sessions
- Train each role on its own tasks
- Start a shared prompt and example library
- Collect errors, near misses and unapproved tool use findings



Days 61 to 90: decide

- Measure time, quality and rework against baseline
- Decide on expansion by role, based on evidence
- Reclaim unused licenses and refresh training
- Report results and where freed capacity went to the sponsor

K

GLOSSARY

Terms used in this publication

11

TERM	MEANING
Copilot	An AI assistant that works alongside a person inside their tools, producing drafts, summaries and suggestions.
Prompt	The instruction and context given to an AI assistant.
Grounding	Giving the assistant real source material so its answer is based on it.
Hallucination	A confident output that is false or unsupported by sources.
Shadow AI	Use of AI tools for work outside approved channels.
AI champion	A respected team member who coaches colleagues and relays feedback.
Rework rate	The share of AI-assisted drafts that need substantial rewriting.
Sensitivity label	A tag marking a document's confidentiality so tools and people handle it correctly.
Connected access	Allowing the assistant to read the user's files, mail or calendar to give better answers.
Prompt library	A shared collection of tested prompts and examples for common tasks.
Automation bias	Over-reliance on machine output without enough checking.
Baseline	A measurement taken before change, used as the comparison point.

A note on use

This playbook is general guidance. Check current vendor terms and legal requirements, including on recording, privacy and employment, before rollout.



TAKE THE NEXT STEP

Take the free Halden assessment at halden.company/assessments, or talk to us at halden.company/contact about turning this into a plan for your organization.

