Assessing the use of CoPilot to enhance Ministerial processes

Logan Templer and Andrew Jackson

Policy Hub

The team

VUW

Professor Alex Richter - expert in the practical application for AI to transform work

Professor Markus Luczak Roesch – expert in the development of computational tools that securely and meaningfully augment human intelligence

Jingyu Zheng - expert in digital transformation to improve business performance

Hassan Mustafa – expert in digital knowledge production; algorithmic manipulation and AI governance

Andrew Jackson – expert in government and process design

MBIE

Logan Templer – Ministerial correspondence subject matter lead

Maddy Hirst – Manager, LSE Ministerial Servicing

Sarah Gibb – Principal Advisor, Private Secretaries

Goals and objectives of the project

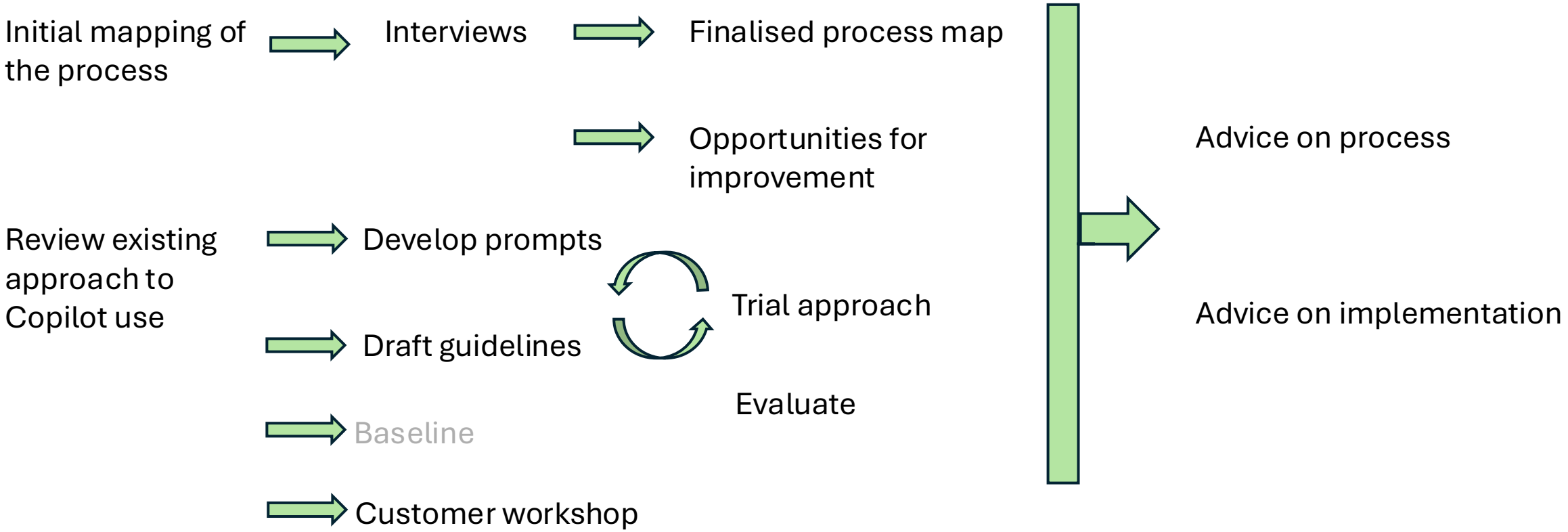
The goals of the project were to improve:

- the quality, and drafting efficiency of Ministerial Correspondence processes
- correspondents, Ministers' and officials experience
- AI literacy and capability within MBIE

The objectives of the project were to:

- review and assess how digital tools could enhance the overall processes
- trial the use of copilot drafting replies
- inform decisions in MBIE on implementation

The project assessed how the process to draft replies could be improved with effective adoption of Copilot



Policy Hub

Structured interviews were held to identify opportunities for process improvement and willingness to use Copilot

Key Findings

The current process is sound, but targeted enhancements could improve efficiency and consistency.

Copilot is already widely used, though the approach to use varies by group. Different levels of Copilot license affect the ability to get value from use.

Finding relevant input material for Copilot is hard – with no search on Nexus based on subject matter; no access to the Marco system; and final versions of signed off ministers replies not being available.

Concern that using Copilot to draft replies would remove a key learning opportunity for analysts.

Expectation of communication to Ministers and members of the public of when and how Copilot has been used is unclear.

There is a risk of negative public reaction from the use of Copilot without clear communication of when and how it is being used.

Copilot can help with sympathetic tone, but is less good at saying “no”.

Recommendations

Adopt a rigorous and consistent approach to the use of Copilot

The approach should build from the learnings of the prompt trial where possible reducing the number of layers involved in the drafting process

The approach should ensure it provides for staff development in relation to policy and supporting Ministers.

To maintain benefits and continue improvements MBIE should
- provide activity focused training in the use of Copilot
- establish mechanisms to support sharing of best practice

To manage risks associated with this approach MBIE should
- develop a governance and transparency framework
- front foot communication with the public on Copilot use

Recognised that prompt strategy should vary by letter complexity and user experience

	New to AI, New to Topic	New to AI, Knows Topic	Knows AI, New to Topic	Knows AI, Knows Topic
Simple				
Moderate				
Difficult				

A Copilot prompt strategy was developed and trialed building from existing MBIE experience

Overall process

Existing best practice reviewed

⇩

3 prompt strategies developed building from best practice

⇩

3 prompt strategies tested on 30 case to identify the best approach

⇩

Workshop where people with varying levels of skill in AI use and handling ministerials used the best prompt approach to draft replies to letters of differing complexity. Some with resources some without

⇩

Draft letters assessed by experts and managers

Best prompt approach included

Context, who is writing and why
Tone
Structure guidelines
Length and formatting
Limitations

Letter types

Type 1 Simple *easy to reply to, whether because it asks a simple routine and previously answered question or there has been a Ministerial direction to just acknowledge the letter or to speak with officials*

Type 2 **Moderate** Bespoke letters based on existing policies and information.

Type 3 Complex *Requires new thinking or the crafting of a careful reply. This may be the formulation of a new policy, dealing with an operational complaint where careful management is needed, or is politically sensitive area*

+

Capability types

New to AI and subject

New to AI but know the subject

AI experience but new to subject

AI experience and knows the subject

Policy Hub

The prompt strategy developed offers significant opportunities to deliver efficiencies if deployed across MBIE

Good results for simple or moderate letters when supporting resources were provided.

Needed human steer for difficult letters.

Use of Copilot with this prompt strategy significantly reduced the time taken to draft replies. Anecdotally between 4 fold and 10 fold

Access to supporting resources (prior letters, policy documents, legislative references) was key to quality AI drafts

Follow-up prompting was underutilised across all groups, suggesting experience with the use of AI overtime could see improvement in the results.

Replies were scored out of 5 against 6 criteria with a maximum score of 30

Dimension	What Was Assessed
Quality	Overall drafting standard, structure, and readability.
Relevance	Whether the response directly addressed the letter's specific concerns.
Accuracy	Factual correctness and absence of hallucination or unsupported claims.
Tone	Appropriateness of ministerial voice (formal, empathetic, and measured).
Policy Alignment	Consistency with stated positions and portfolio boundaries.
Completeness	Whether all key points were addressed without unnecessary additions.

All drafts for low and medium letters by those resourced were ready for sign off

Candidate	Group	Low (/30)	Medium (/30)	High (/30)
				Group A — Resourced
Candidate 1	Resourced	30	28	24
Candidate 2	Resourced	27	22	19
Candidate 3	Resourced	26	29	30
Candidate 4	Resourced	30	29	N/A*
				Group B — Non-Resourced
Candidate 5	Non-Resourced	15	23	20
Candidate 6	Non-Resourced	22	26	26
Candidate 7	Non-Resourced	17	24	24

A standardised tiered prompt approach should be adopted

Short-Term Recommendations | Next Six Months
1. Standardise a tiered prompting framework. Based on what worked in the trial, develop and publish a prompting playbook for ministerial correspondence, differentiating by complexity (single-shot for low, chain-of-thought for medium and high). Include example prompts drawn from the trial's highest-scoring outputs.
2. Develop a resource pack for Copilot-assisted drafting. Create a set of policy context cards, tone guides, and portfolio boundary references that officials can use to "resource" Copilot before drafting, replicating the advantage observed in Group A for all users.
3. Prioritise AI literacy training over writing skills training. The trial demonstrated that AI experience is the stronger predictor of output quality. Invest in practical AI prompting skills for correspondence staff, including hands-on practice with iterative prompting.
4. Establish a mandatory human review protocol. Introduce a clear sign-off requirement for all AI-drafted correspondence before it leaves MBIE, with an explicit checklist covering factual accuracy, policy alignment, tone, and portfolio scope.

Use of the prompt strategy should vary by letter complexity and user experience

	New to Copilot New to topic	New to Copilot Knows topic	Knows Copilot New to topic	Knows Copilot Knows topic
Simple	**Human drafts** Author drafts, potentially using summary prompt. SME to peer review and polish.	**Copilot drafts** Copilot draft. Author peer review and polish.	**Copilot drafts** Copilot draft and polish. SME peer review.	**Copilot drafts** Copilot drafts. Author peer review.
Moderate	**Human drafts** Author drafts, potentially using summary prompt. SME to peer review and polish.	**Copilot drafts** Copilot draft. SME peer review and polish.	**Copilot drafts** Copilot draft and polish. SME peer review.	**Copilot drafts** Copilot drafts and author polishes. SME peer review if necessary.
Difficult	**Human drafts** SME dialogue to agree approach. Human drafts using SME lines. Copilot polish and proof. SME peer review.	**Copilot assists** SME dialogue to agree approach. Author to draft, potentially using summary prompt. SME peer review.	**Copilot assists** SME dialogue to agree approach. Copilot drafts, author to polish. SME peer review.	**Copilot assists** SME dialogue to agree approach. Copilot drafts, author to polish. SME peer review.

Copilot drafts **Copilot assists** **Human drafts**

Policy Hub

There are significant additional opportunities using AI including agentic AI

Mid-Term Recommendations | Next One to Two Years
1. Move toward an agentic AI drafting pipeline.
2. Integrate Copilot with MBIE's policy and knowledge repositories.
3. Build a multi-agent review system.
4. Establish an internal feedback and improvement loop.
5. Expand use cases beyond ministerial correspondence.

Long-Term Recommendations | Beyond Two Years
1. Investigate and pilot a locally hosted, fine-tuned open-source LLM.
2. House the model entirely within MBIE's infrastructure.
3. Build in-house AI capability through the fine-tuning process.
4. Design for tasks beyond correspondence.
5. Establish a governance and ownership framework.

From Pilot to Practice

MBIE is taking forward the learnings from the Pilot through two main streams:

1. **Rolling out Copilot-assisted drafting approach**
    - Expanding beyond the Pilot Portfolio to others across the Labour, Science and Enterprise and Buildings and Resource Markets Groups.
    - Developing a “Minimum Operating Kit” for each portfolio, including specific prompts, resource libraries of approved materials, and exploring the use of Microsoft 365 agents.
    - Developing training to support a shift in drafting model.

2. **Exploring an agentic model for the ministerial correspondence process**
    - Working with the Government Digital Delivery Agency (GDDA) to develop an end to end agentic model using a series of more advanced agents.
    - MBIE has shared information from the Pilot and additional materials/feedback to support early thinking on an agentic model.
    - The GDDA is holding inter-agency workshops to inform solution design and understand needs across the Government.
    - Testing on an agentic model is scheduled for November 2026 onwards.

Policy Hub

In Summary – slide may not be needed

Findings

There was already widespread use of Copilot by officials to help draft replies to Ministerials. The way it was being used varied widely between teams and individuals as did the level of competence of the users.

The potential benefits of Copilot are not being realized fully and while current assurance processes mean quality of replies remains high, wider risks from use of Copilot remain; notably potential for negative public perception of use and the impact on officials' capabilities to draft letters and learn about a new area of policy.

A sophisticated approach to prompt use was developed which reduces the time taken to draft replies. The sophisticated allows a new process to be introduced, with non experts drafting replies. This allows for a significant reduction in the number of people involved, with the role shifting to quality assurance rather than drafting.

Recommendations

To maximise opportunities from Copilot, MBIE should:

- train correspondence teams in the use of the sophisticated approach to Copilot
- introduce mechanisms to deliver those teams with core information on key policies
- change its process so that correspondence teams send both the incoming letter and a draft reply to an assigned SMEs for sign off
- invest in further research to further enhance the use of AI to enhance processes, including trials with AI agents

To manage the risks MBIE should:

- establish effective governance of use of AI
- front foot public engagement on use
- limit initial use of copilot for staff new to Ministerials
- support a small cadre of AI specialists for ongoing enhancement and dissemination of AI deployment including the use of agents